

What are the key drivers of spread in invasive plants: dispersal, demography or landscape: and how can we use this knowledge to aid management?

Shaun R. Coutts · Riëks D. van Klinken ·
Hiroyuki Yokomizo · Yvonne M. Buckley

Received: 11 February 2010 / Accepted: 27 November 2010 / Published online: 15 December 2010
© Springer Science+Business Media B.V. 2010

Abstract Invasive plants disrupt ecosystems from local to landscape scales. Reduction or reversal of spread is an important goal of many invasive plant management strategies, but few general guidelines exist on how to achieve this aim. We identified the main drivers of spread, and thus potential targets for management, using a spatially explicit simulation model tested on different life history categories in different spread and landscape scenarios. We used boosted regression trees to determine the parameters that most affected spread. Additionally, we analysed how spread reacted to changes in those parameters

over a broad realistic range. From our results we deduce four simple management guidelines: (1) Manage dispersal if possible, as mean dispersal distance was an important driver of spread for all life history categories; (2) short bursts of rapid spread or more usual year on year spread can have different drivers, therefore managers need to decide what type of spread they want to slow; (3) efforts to manage spread will have variable outcomes due to interactions between, and non-linear responses to, key drivers of spread; and (4) the most useful demographic rates to target depend on dispersal ability, life history and how spread is measured. Fecundity was found to be important for driving spread only when reduced to low levels and particularly when the species was short lived. For longer lived species management should target survival, or age of maturity, especially when dispersal ability is limited.

Electronic supplementary material The online version of this article (doi:[10.1007/s10530-010-9922-5](https://doi.org/10.1007/s10530-010-9922-5)) contains supplementary material, which is available to authorized users.

S. R. Coutts (✉) · Y. M. Buckley
School of Biological Sciences, University of Queensland,
St Lucia, QLD 4072, Australia
e-mail: s.coutts@uq.edu.au

R. D. van Klinken
CSIRO Entomology, 120 Meiers Road, Indooroopilly,
QLD 4068, Australia

H. Yokomizo
Research Center for Environmental Risk,
National Institute for Environmental Studies,
Onogawa 16-2, Tsukuba 305-8506, Japan

Y. M. Buckley
CSIRO Ecosystem Sciences, 306 Carmody Rd,
St Lucia, QLD 4067, Australia

Keywords Boosted regression trees ·
Heterogeneous landscapes · Plant dispersal · Plant
invasions · Sensitivity analysis · Simulation model

Introduction

Preventing invasive alien plants (IAPs) from reaching new areas can be the most cost effective means of control (Davies and Sheley 2007), especially when it is expensive to reduce densities to the point where the

damage caused is acceptable (Yokomizo et al. 2009). Management targeting demographic rates such as establishment, fecundity or survival, using biological control and herbicides (Shea and Kelly 1998; Buckley et al. 2003) is often undertaken to reduce the densities of IAP populations. As a positive side effect such management actions may also limit the spread of the species.

Spread can also be targeted by managing dispersal vectors directly, for example modifying the behaviour of birds by manipulating the landscape, canopy and fruits (Gosper et al. 2005; Buckley et al. 2006). However, this approach requires greater knowledge about species specific dispersal mechanisms, which in many cases involve multiple vectors (Davies and Sheley 2007). Alternatively control could be targeted at individuals in areas with a high frequency of dispersal opportunities, for example wind dispersed pine in windy sites (Buckley et al. 2005). Spread may also be managed by targeting satellite populations for surveillance and plant removal (Moody and Mack 1988; Fox et al. 2009; Pokorny and Krueger-Mangold 2007).

It should be noted that there are two main ways of defining spread, with previous models, including those discussed below, having typically defined spread as the velocity of the range front in one dimension (Hastings et al. 2005). In common use, however, spread is often used as a synonym for increase in area. In our model plants dispersed in two dimensions, and simulated invasions often had no definable range front. Therefore, here we define spread as an increase in area occupied.

Modelling studies suggest that targeting demographic rates for management may be of limited value in reducing spread (Clark 1998; Higgins and Richardson 1999; Neubert and Caswell 2000) because the frequency of long distance dispersal (LDD) is consistently the most important driver of spread. Models incorporating a higher probability of LDD show an order of magnitude increase in spread speed (Higgins and Richardson 1999). LDD also tends to result in many small satellite foci ahead of the invasion front, which allows the invasion to fill space more quickly (Moody and Mack 1988) if the dispersal kernel is not exponentially bounded.

The relationship between spread and demographic rates is not straight-forward (Clark 1998; Hastings et al. 2005). No matter how well a species disperses

its seeds, it will not spread unless those seeds reach suitable habitat and germinate and establish. If suitable habitat is fragmented then seeds will need to cross gaps in order to spread across the landscape (With and King 1999). How large the gaps are, seed production and dispersal distances will determine how many seeds cross gaps. Seed bank dynamics, germination and survival to age of maturity then determine what proportion of seeds arriving at a site give rise to seed producing plants.

Previous models of spread have used net reproductive rate (R_0) to summarise all these demographic processes in one number (Clark 1998; Clark et al. 2001; Kot et al. 1996). Recent integro-difference equation models tease out which components of population growth rate most affect spread (Neubert and Caswell 2000), but these use homogeneous landscapes. To date few models have combined realistic dispersal kernels with demography and spatially structured landscapes. Malanson and Cairns (1997) and Higgins and Richardson (1998) are exceptions but their models focus on trees, and there is little indication how species with other life history strategies might behave.

General guidelines about which demographic rates most affect local population growth rate are available (Ramula et al. 2008), but thus far we have no management relevant guidelines for which variables affect spread. Our aim is to identify the most important drivers of spread and use these to provide guidelines for spread management.

We categorised invasive species into one of four broad life history categories: annuals, and short, medium and long lived perennials. We use a landscape scale simulation model, run in a two dimensional landscape, where dispersal ability, demographic rates and landscape structure were all varied. As landscape structure and the dispersal process often cannot be quickly altered for management we subsequently varied just the demographic parameters while holding the other parameters constant at both high and low levels. Boosted regression trees (BRTs) (Friedman 2001) were used to determine the parameters most important for driving spread.

It is important to note that in order to keep the simulation model general enough to be applied across different life history categories, detail, and thus applicability to real life invasions was limited. As such the goal of this paper is not to predict the

outcome of particular invasions, but rather to find general trends.

Methods

We describe the simulation using Grimm et al.'s (2006) "Overview, Design concepts and Details" format. We first give a general outline, then a detailed description of each sub-model. Following this, the BRT analysis and simulated scenarios are explained.

Simulation model

Simulation model overview

Purpose This simulation describes how plants in different life history categories spread across landscapes, using a range of values for parameters controlling landscape fragmentation, dispersal and demography. BRTs were used to analyse simulation model outputs to determine the parameters most important in driving spread for each life history category and how the effect of important parameters on spread changed over the range of values tested.

State variables and scales The simulation base unit was the individual plant, each individual being either an adult (age $\geq o^*$) or juvenile (age $< o^*$), where o^* is the age at which individuals start to produce seed. Individuals were aggregated in square cells, w in m^2 , which were either suitable or unsuitable with no plants establishing in unsuitable cells. Cells contained the number of seeds that entered in the current time step and the number of seeds in the seed bank (for the dynamics of seeds and seed bank see Eqs. 4, 5). Each cell had a ceiling population, m , above which the population could not increase. w and m were constant for every cell within a landscape but differed between life history categories (see *simulation scenarios*).

Cells were aggregated into a landscape of 257×100 cells. Reflecting boundaries were used so that clumping of suitable habitat was consistent between the edges and the centre of the landscape. As described in With (1997) the clumping of suitable cells within a landscape was determined using the fractal generating mid-point displacement algorithm (described in Appendix 1—Electronic Supplementary Material). Landscapes were varied in two ways.

Firstly, the degree of clumping, controlled by the roughness constant (h), was varied from $h = 0$ (least clumped) to $h = 1$ (most clumped). Secondly, the proportion of suitable habitat (r) was varied between 0.1 and 1 (see With (1997) and Fig. A1 in Appendix 1—Electronic Supplementary Material, for examples of fractal landscapes).

The simulation used a yearly time step and was run for 50 time steps or until 95% of suitable cells had at least one individual present. A time horizon of 50 years was computationally tractable and suitable from a mid- to long-term management perspective. Initial simulations showed that once a population occupied roughly 95% of the landscape its growth rate slowed because finding suitable, empty habitat became harder.

Processes and scheduling At each time step (1) mortality was applied to all juveniles and adults (Eq. 1); (2) juveniles that survived were aged one year (Eq. 1), juveniles whose age increased to o^* became adults; (3) adults (including those created in the current time step) produced and dispersed seeds (Eq. 2); (4) seeds dispersed into suitable habitat and seeds in the seed bank, could germinate and become seedlings (Eq. 4); (5) new seeds that failed to germinate entered the seed bank or died (Eq. 5); (6) seed mortality was applied to the seed bank, including those seeds which entered in the current time step (Eq. 5); (7) seedling mortality was applied and those that survived became juveniles (Eq. 4). The relationship between states and mortality from each state is shown in Fig. 1.

Concepts Stochasticity in mortality, germination and seedling survival was included in the model by rounding, for example, 1.9 plants, up to two 90% of the time, and down to one, 10% of the time. Thus, even very low seed densities, far from the invasion front, could result in a new satellite infestations, as partial plants (e.g., 0.0001 of a plant) were occasionally rounded up to 1. We made the simplifying assumption that there was no environmental stochasticity through time, resulting in no bad years to slow population growth, or good years to accelerate it. A ceiling density was imposed to avoid unrealistically high numbers of plants in individual cells.

Initialisation At the start of each run the fractal landscape was generated (see subsection *State*

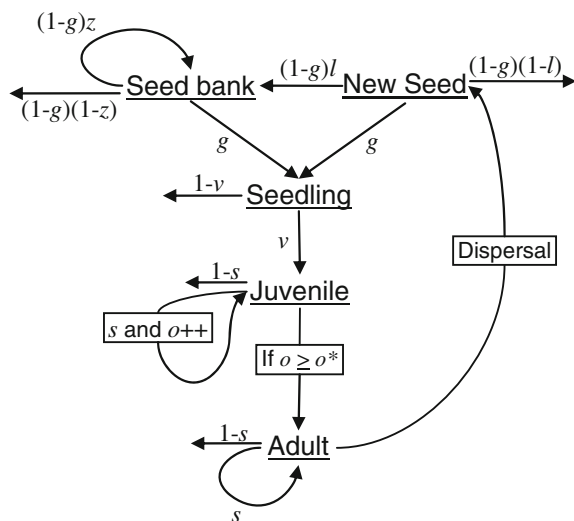


Fig. 1 Flow chart of simulation model. States are underlined and transitions out of states are represented by *arrows* and labelled with the probability that an individual will make that transition. Those *arrows* which do not point to a state indicate mortality (either seed or plant). *Boxes* show processes that are too complex to be captured by a single probability. *o++* means: increment age by one. Note that New Seed and Seedling only exist within a time step. The order that each process occurs is given in *Processes and scheduling*

variables and scales) and the appropriate cells for the scenario were seeded with m adults (for the ranges of m used see Table 2). The starting configuration for each scenario is out lined in the *simulation scenarios* section.

Input Demographic parameters were expressed as uniform distributions with an upper and lower limit. The upper and lower limits were chosen based on what we considered biologically reasonable, guided by reported rates in the literature for invasive species (Appendix 2—Electronic Supplementary Material) and unpublished data (see Table 2 for limits used). We could not simply use reported rates from the literature because the list of invasive species examined (Appendix 2—Electronic Supplementary Material) was not exhaustive and many species did not have all parameters reported for them.

Sub models

Adult and juvenile mortality and growth Mortality and growth determined how many adults and

juveniles survived to contribute to dispersal. The number of individuals in cell i at time step t of age o is

$$a_{i,t,o} = \begin{cases} sa_{i,t-1,o-1} : o > 1 \\ q_{i,t-1} : o = 1 \end{cases} \quad (1)$$

Because seedling survivorship, v , is generally much lower than subsequent survivorship, s , (see v and s in Table 2) the number of plants in age class 1 is $q_{i,t}$ (Eq. 4), the number of seeds that germinate and become juveniles in cell i time step t .

Dispersal The number of propagules received by cell i at time step t ($p_{i,t}$), where the landscape contains K cells, is:

$$p_{i,t} = \sum_{k=1}^{k=K} a_{k,t,o \geq o^*} f x_{k,i} \quad (2)$$

f is the number of seeds produced per adult plant and $x_{k,i}$ is the propagule density sent from cell k to cell i , and is given by Eq. 3, which is Clark et al.'s (1998) dispersal kernel (p. 218).

$$x_{k,i} = \frac{c}{2\pi(\mu\Gamma(2/c)\Gamma(3/c))^2\Gamma(2/c)} \times \exp[-(d_{k,i}/[\mu\Gamma(2/c)/\Gamma(3/c)])^c] \quad (3)$$

$\Gamma(\cdot)$ is the gamma function, $d_{k,i}$ is the Euclidean distance between the centres of cells k and i , in cells, μ is the mean dispersal distance, also in cells, and c is a shape parameter. Lower c puts more of the distribution in the tail, making the tail fatter and lowering the y-intercept.

Establishment The number of seeds that establish and become juveniles at suitable sites is calculated as:

$$q_{i,t} = \begin{cases} (p_{i,t} + b_{i,t})gv : u_{i,t}gv \leq m - a_{i,t} \\ m - a_{i,t} : \text{otherwise} \end{cases} \quad (4)$$

$$b_{i,t} = z(1-g)(b_{i,t-1} + p_{i,t-1}l) \quad (5)$$

where g is the density independent germination rate and v is the density independent seedling survival rate. We assume that seedlings do not displace juveniles or adults, thus, we capped the number of seedlings that establish at $m - a_{i,t}$, where m is the ceiling population and $a_{i,t}$ is the number of individuals of all ages in cell i . $p_{i,t}$ is defined in Eq. 2 and $b_{i,t}$ is the number of seeds that are available for

establishment from the seed bank in cell i at time t . z is the proportion of seeds in the seed bank that remain viable into the next time step and l is the proportion of new seeds that enter the seed bank (see Table 2 for values of z and l and Fig. 1 for how the seed bank, new seed and seedlings relate to each other). To keep seed bank dynamics simple we assumed all seed decay processes were incorporated into z . This is reasonable for species without seed dormancy mechanisms; in species with seed dormancy this assumption overestimated germination rate from the seed bank and underestimated seed bank longevity.

Simulation scenarios

All simulation scenarios were conducted on four broad life history categories: annuals, and short lived, medium lived and long lived perennials. Because species in the annual and short lived categories tended to be smaller and grow more densely than those in the longer lived categories (see Appendix 2—Electronic Supplementary Material for examples of species in each category) many more of them could fit onto a landscape of the same size. Thus landscapes for the shorter lived categories either needed to contain many more, smaller, cells or allow more individuals per cell, or be of smaller size. It was not computationally tractable to make landscapes containing many more cells. Because dispersal is between cell centres, allowing too many plants per cell makes all dispersal between points that are, relative to individual size, far from each other.

We made cell area (w) smaller for shorter lived life history categories but kept the total number of cells the same (thus shrinking total landscape size), and increased maximum individuals per cell (m) for these categories. We scaled dispersal distance to landscape size by measuring all distances in cells. Therefore, while the drivers of spread can be compared between life history categories the actual spread rate cannot. We also assumed that longer lived life history categories had higher upper limits for age of maturity (o^*) and higher survivorship.

Consequently, there were four differences between the different life history categories: (1) the upper limit for maximum individuals per cell (m) was larger for shorter lived categories; (2) Cell area (w) was larger for longer lived categories; (3) survivorship

(s) was lower for shorter lived categories; and (4) the upper limit of age of maturity (o^*) was lower for shorter lived categories. All annuals died after they dispersed seed. Annuals were allowed to survive the mortality process in the first time step so that the initial population could disperse seed before dying. The upper and lower limit for each parameter is shown in Table 2, and a brief description of each parameter is given in Table 1.

Finding parameters that most influenced spread

Each combination of parameters was simulated for a single run of 50 time steps, and two metrics of spread, maximum and median rate of spread, were recorded. Spread was defined as an increase in area, which in the simulation was measured in cells. Maximum rate of spread was the largest net increase in the proportion of suitable cells occupied by at least one individual in a single time step and gives the fastest observed rate of spread in a run. Median rate of spread can be thought of as how much the invasion increased in a typical time step, and was the median per time step change in proportion of suitable cells occupied by at least one individual.

A data set of parameter values and the consequent maximum and median rates of spread was generated by repeating the following steps 1,500 times. (1) Parameter values were randomly drawn from a uniform distribution defined by upper and lower limits. The exceptions to this were the dispersal kernel parameters, which were varied in pre-determined levels to improve computation time (see Table 2). We did not incorporate potential co-variances between parameters. (2) A random landscape was generated according to fractal and suitability parameters and the bottom-most suitable cell (where the long axis is orientated vertically) was seeded with m adults. (3) The simulation was run for 50 time steps, the parameter values used and estimates of spread were recorded. If the population went extinct within 10 time steps the run was discarded. This removed the effect of harsh parameter sets under which small populations could not grow; an unrealistic situation for invasive species. Less than one percent of runs were rejected for any scenario.

Boosted regression trees (BRTs) were used to find the parameters that most influenced spread because they deal well with non-linear responses and higher

Table 1 All simulation model parameters, and their interpretations

Parameter	Interpretation
<i>Demographic</i>	
f	The number of seeds produced per adult plant per time step
v	Density independent probability that a seedling will survive to 1 time step old
g	Density independent probability that a single seed will germinate
s	The probability that a plant will survive 1 time step. Rates used (Table 2) correspond to 95% dead by 1–7 (short lived), 7–27 (medium lived) and 97–300 (long lived) time steps. Annuals die after dispersing seeds
o^*	The number of time steps it takes for a plant to start producing seeds also, referred to as “age of maturity”
l	The proportion of seeds from the current time step which enter the seed bank
z	Seed bank survivorship. Rates used are equivalent to 95% of the seeds in the seed bank disappearing from the population within 0–49 time steps
<i>Landscape</i>	
w	Cell area in m^2
m	The maximum number of plants that can exist in a cell at one time
h	The roughness constant. Controls how clumped the landscape is 1 = most clumped, 0 = least clumped
r	Proportion of habitat that is suitable
<i>Dispersal</i>	
c	Determines how fat the tail of the dispersal kernel is. $c < 1$ is fat tailed, $c = 1$ is a negative exponential, $c = 2$ is a Gaussian distribution
μ	Mean dispersal distance in cells

level interactions among explanatory variables (Elith et al. 2008). Using a series of small regression trees BRTs find the relationship between a set of inputs, the simulation parameter values, and an output, the maximum or median rate of spread (Elith et al. 2008). For examples of BRT use see Friedman (2001) and Elith et al. (2008).

BRTs were fit to the data sets generated using the above process, closely following the approach and rules of thumb employed by Elith et al. (2008). BRTs have three parameters to specify, the number of regression trees used, the amount by which the prediction of each tree is scaled down (learning rate) and maximum interaction depth (tree complexity). We used hold out cross validation to determine the learning rate and number of trees that minimised the squared difference between predicted and observed values, summed across all predictions (cross validation error estimate). We split the data into 10 groups, one group was held out and all the other groups were combined and used to fit a BRT, which was then tested on the hold out group. This process was repeated until all groups had acted as the hold out

group. Using a tree complexity of 5 we varied the learning rate in levels between 0.05 and 0.0005, choosing the learning rate that produced a BRT that minimised the cross validation error estimate and contained between 1,000 and 35,000 trees. Testing showed that increasing tree complexity past 5 did not improve cross validation error. BRTs were created and interrogated using R (R Development Core Team 2006) and the “gbm” package (Ridgway 2010).

Exploring the effect of a parameter on the output

The relative influence of each variable was calculated using the index developed by Friedman (2001). This index is the number of times a parameter appears in each tree, multiplied by the squared improvement each time it appears in that tree, averaged over all trees in the BRT (Elith et al. 2008; Friedman 2001). The index is scaled so that it sums to 100 across all parameters. Higher values indicate a parameter had relatively more influence explaining variation in the data.

We generated 95% confidence intervals on relative influence values for each parameter using bootstrap

Table 2 The ranges used in the general (gen.) and management (man.) scenarios for each simulation parameter (rows) for each life history category (columns)

Parameter	Scenario	Annual	Life history category		
			Short lived	Medium lived	Long lived
f	Gen.	100–50,000	100–50,000	100–50,000	100–50,000
	Man.	1–50,000	1–50,000	1–50,000	1–50,000
g	Gen.	0.01–0.8	0.01–0.8	0.01–0.8	0.01–0.8
	Man.	0.00001–0.8	0.00001–0.8	0.00001–0.8	0.00001–0.8
v	Gen.	0.01–0.125	0.01–0.125	0.01–0.125	0.01–0.125
	Man.	0.00001–0.125	0.00001–0.125	0.00001–0.125	0.00001–0.125
s	Gen.	Included in v	0.2–0.7	0.7–0.9	0.97–0.99
	Man.	Included in v	0.1–0.7	0.1–0.9	0.1–0.99
o^*	Gen.	1	1	1–5	1–12
	Man.	1	1–2	1–8	1–18
l	Gen.	0–0.9	0–0.9	0–0.9	0–0.9
	Man.	0–0.9	0–0.9	0–0.9	0–0.9
z	Gen.	0–0.94	0–0.94	0–0.94	0–0.94
	Man.	0–0.94	0–0.94	0–0.94	0–0.94
w	Gen.	1 m ²	1 m ²	5 m ²	5 m ²
	Man.	1 m ²	1 m ²	5 m ²	5 m ²
m	Gen.	1–200	1–200	1–40	1–5
	Man.	1–200	1–200	1–40	1–5
h	Gen.	0–1	0–1	0–1	0–1
	Man.	0 or 1	0 or 1	0 or 1	0 or 1
r	Gen.	0.1–1	0.1–1	0.1–1	0.1–1
	Man.	0.25	0.25	0.25	0.25
c	Gen.	0.5, 1, 2	0.5, 1, 2	0.5, 1, 2	0.5, 1, 2
	Man.	0.5 or 2	0.5 or 2	0.5 or 2	0.5 or 2
μ	Gen.	1, 11, 21, 31, 41	1, 11, 21, 31, 41	1, 11, 21, 31, 41	1, 11, 21, 31, 41
	Man.	1 or 3	1 or 3	1 or 3	1 or 3

Where no range is given only one value was used. The comma separated values for c and μ indicate that these parameters were tested in levels, not randomly selected

re-sampling. 2,000 re-sampled data sets 1,000 or 1,500 rows in size (depending on the scenario) were generated by randomly selecting rows, with replacement, from the simulated data sets generated using the process in the previous section. A BRT was fit to each re-sampled data set using the same learning rate and number of trees as the BRT fit to the original simulated data set, the relative influence of each variable was recorded. BCa bootstrap confidence intervals (Davison and Hinkley 1997) were calculated using “boot.ci” (Canty and Ripley 2010) in R (R Development Core Team 2006).

Visualising the response

BRT's use partial dependence plots to visualise the relationship between a parameter and the output, integrating out the effect of the other parameters using Friedman's (2001) weighted tree traversal method. These plots do not show the exact relationship, particularly if there are strong interactions (Elith et al. 2008).

Confidence intervals (95%) for partial dependence plots were generated using a similar process to that described in *Exploring the effect of a parameter on the output*, except that only 1,000 re-sampled data

sets were generated, the partial dependence of the parameter of interest was recorded from each BRT instead of relative influence and percentile confidence intervals were calculated rather than BCa.

The Gini coefficient, G , measures the amount of area between a curve and a hypothetical straight line running between the end points of that curve (Cowell 1995). We used the Gini coefficient to quantify how concave (or convex) the partial dependence curves were using the trapezoid approximation

$$G = 1 - \sum_{n=1}^N (q_n - q_{n-1})(y_n + y_{n-1}) \quad (6)$$

where y_n is the predicted marginal effect of a parameter evaluated at value q_n , q_n is ordered so that $q_n > q_{n-1}$, and N is the number of evaluations made to create the partial dependence plots ($N = 100$). The area is scaled so that when $G = 0$ the curve is a straight line and when $G = 1$ the curve is a “L” (or inverted “L”) shape.

Management scenarios

To explore the potential of current weed management options for controlling median and maximum rate of spread, BRT's were fitted and examined using the processes explained above, with a few key differences. The changes were informed by the constraints and goals of management. Initially 25 suitable cells were selected at random and seeded with the maximum number of adults, m . This represents a situation where management doesn't start until the invasive plant exists throughout the landscape, but in small scattered patches.

Because managers often have no means to influence dispersal or landscape structure these were held constant at two levels for each. The two types of dispersal were long distance dispersers, where $c = 0.5$ and $\mu = 3$, and short distance dispersers, where $c = 2$ and $\mu = 1$. Mean dispersal distances for both short and long dispersers are similar to reported ranges for several species with different life history strategies (Bullock and Clarke 2000; Dauer et al. 2006; Martinez and Gonzalez-Taboada 2009; Skarpaas and Shea 2007). The values of c for short and long dispersers are those used by Clark (1998) to model, respectively, a Gaussian and a fat tailed kernel. It should be noted that the aim was not to

mimic dispersal for different types of plant, but to see if changes in dispersal ability influenced which demographic rates were important in driving spread. The two types of landscape were continuous, $h = 1$, and fragmented, $h = 0$, $r = 0.25$ for both. Since the same value of r was used for all simulations there was no need to express maximum or median rate of spread as a proportion.

Because many management programs take demographic rates outside their normal range, the ranges of the demographic parameters used in these scenarios, (man. in Table 2), are far outside those reported in Appendix 2—Electronic Supplementary Material. Demographic rates were varied in ways that are consistent with common management goals; the lower limits of survivorship, seedling survival and germination probability were decreased, and the upper limit of age of maturity was increased. Finally, because one of the main goals of managers of IAPs is to cause populations to go locally extinct, all runs were counted, no matter how quickly the population died out.

A separate analysis was carried out for both long and short distance dispersers, in fragmented and continuous landscapes for each life history category for both maximum and median rate of spread, giving a total of 32 different scenarios. For each analysis a data set of 1,000 entries was created in the manner outlined above.

Results

General analysis

For both annual and short lived life history categories, for both spread measures, mean dispersal distance (μ) was overwhelmingly the most important parameter in explaining spread (Fig. 2a, b, striped bars).

For the two longer lived life history categories μ was important, but less dominant. The relative importance of μ in predicting maximum rate of spread was lower for the long lived category than all other categories. Also, the medium lived category was lower than the annuals (Fig. 2a). Both μ and o^* (age of maturity) had high relative influence on median rate of spread for long lived and medium lived categories (Fig. 2b).

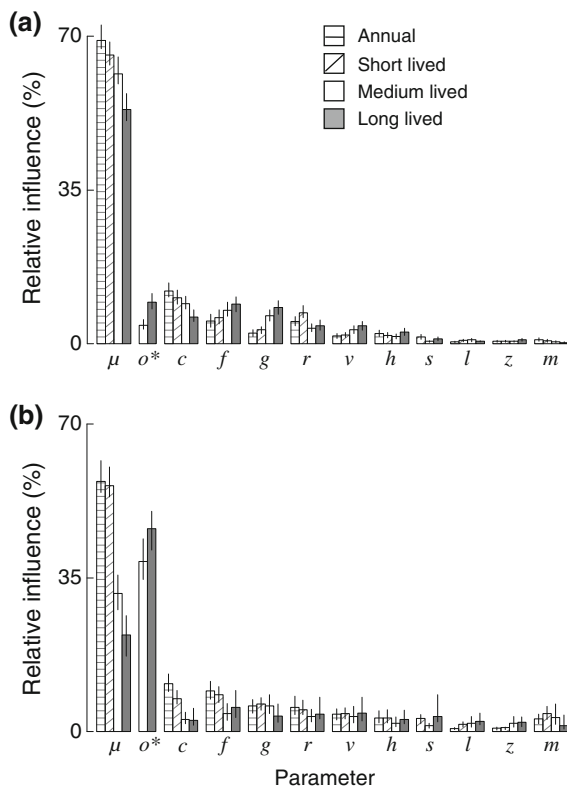


Fig. 2 The relative influence of each parameter for the four life history categories, on **a** maximum rate of spread, and **b** median rate of spread. Lines show upper and lower 95% confidence intervals calculated using a BCa bootstrap

Management scenario

In the interests of clarity only selected results are shown, the full results are available in Appendix 3—Electronic Supplementary Material.

Fecundity was one of the most important parameters under the majority of the management scenarios tested (Fig. 3a, b). Germination rate and seedling survival were also important under a wide range of management scenarios, but tended to be less important than fecundity (Appendix 3, Fig. A2 c, d, e and f).

Both life history category and dispersal ability had an effect on which demographic parameters drove spread. Fecundity, for example, was significantly more important for annuals and short lived categories than the long lived category, and tended to be more important for maximum rate of spread (Fig. 3a, b). Age of maturity offers another example, only being important when median rate of spread was the metric

used, and generally was more important for the two longer lived life history categories (Fig. 3c, d). The latter result is not surprising given that age of maturity could only increase to two time steps for the short lived life history category but to 8 and 18 time steps for the medium and long lived categories.

The relative influence of survival was conditional on life history category and dispersal ability, tending to be more important for the two longer lived categories and when dispersal was short (Fig. 3e, f).

Drivers of spread were not necessarily equally influential across their whole range. The partial dependence curves for fecundity were an inverse “L” shape (Fig. 4), where decreasing fecundity at the low end of its range (1 to ~5,000 seeds per plant) caused steep declines in maximum rate of spread for all life history categories. The shorter lived the life history category the more pronounced this pattern became (in Fig. 4 G gets larger as life history category gets shorter lived). Similar patterns are shown in several other management scenarios (see Appendix 3—Electronic Supplementary Material, Fig. A3 a–g).

When age of maturity was important both maximum and median rates of spread showed a negative, highly convex, “L” shaped relationship with age of maturity (for example $G = 0.84$ in Fig. 5, continuous landscape). In exception to this, on fragmented landscapes with short dispersal the decline in median rate of spread was more linear with increasing age of maturity ($G = 0.53$). In Fig. 5 we show an example of the difference between this latter scenario and others where age of maturity is important. Partial dependence plots produced for other scenarios where age of maturity was important did not differ much from the continuous landscape curve shown in Fig. 5 (see Appendix 3—Electronic Supplementary Material Fig A4 medium and long lived columns).

Discussion

Our results show that spread was driven primarily by dispersal ability, with different demographic rates being of secondary importance depending on dispersal ability, life history category and whether maximum or median rate of spread was the measure. In particular we found mean dispersal distance (μ) was more important than the tail of the dispersal kernel (c) in predicting spread. In light of previous

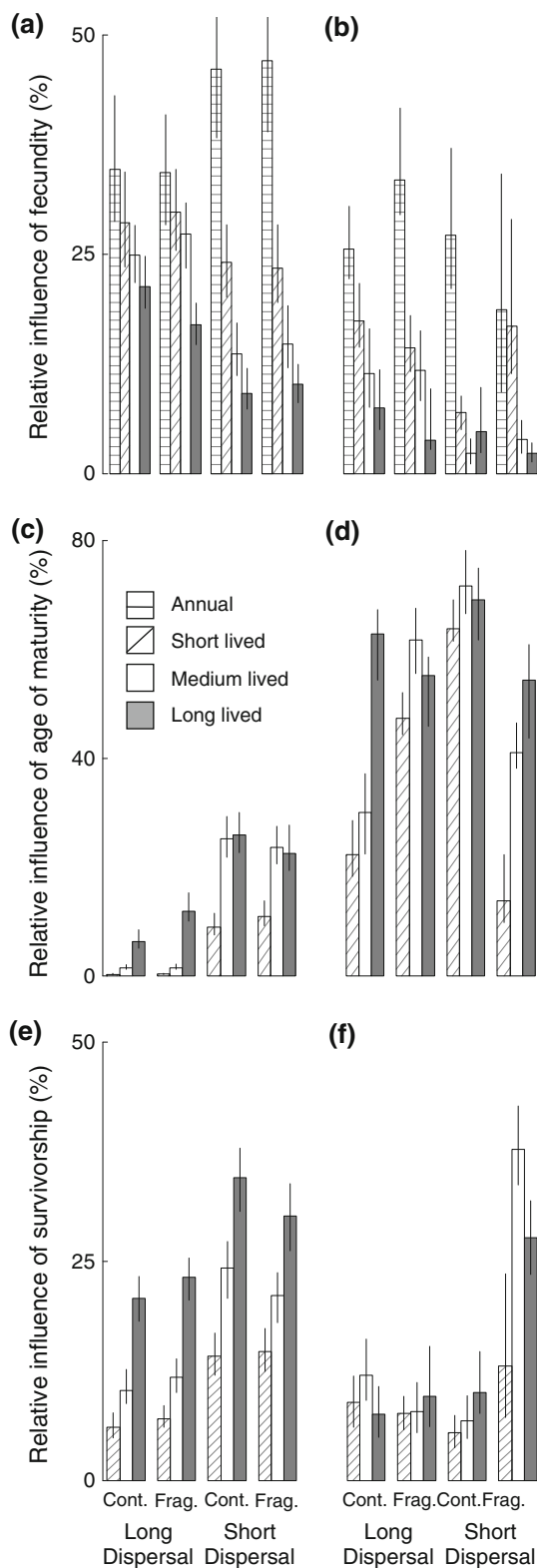


Fig. 3 The relative influence of fecundity (**a** and **b**), age of maturity (**c** and **d**) and survivorship (**e** and **f**) for both maximum (**a**, **c** and **e**) and median (**b**, **d**, **f**) rate of spread, for each life history category, for each scenario (scenario shown on x-axis). Note the different scale for age of maturity. On the x-axis *Cont.* Continuous landscape and *Frag.* Fragmented landscape. *Lines* show upper and lower 95% confidence intervals calculated using a BCa bootstrap

work (Clark 1998; Higgins and Richardson 1999; Neubert and Caswell 2000) it is likely that the changes caused by μ to mid range dispersal (10's to a few hundred cells) played a key role in driving changes in spread, and thus should be targeted when managing dispersal.

The two most important demographic parameters, fecundity and age of maturity, displayed “L” shaped relationships with spread, thus, how much effect these rates had on spread rate depended in part on their value, similar to results from local population models (e.g., Rees and Paynter 1997). Fecundity had a positive relationship with spread while age of maturity a negative relationship, this led to different potential dangers to managers of IAPs resulting from their “L” shape. If fecundity is not driven low enough then resources will be wasted because little result will be achieved. However, if fecundity is pushed low enough then substantial reductions in spread are possible. For example seed feeding biocontrol has been shown to reduce spread rates of woody weeds (Le Maitre et al. 2008). For age of maturity the danger is in the opposite direction, if it is pushed too high then resources are wasted because while spread is reduced, less could have been spent for the same result.

The relationship between spread and fecundity became less sharply curved (lower G) as life history category became longer lived. That between spread and age of maturity became less sharply curved on more fragmented landscapes when dispersal was short. Thus, depending on circumstance managers need to give more or less weight to the potential pitfalls and benefits caused by the shape of the relationship between these two demographic rates and spread.

The importance of many demographic drivers of spread was also context dependant, only being important under certain dispersal, landscape and/or life history conditions. Age of maturity was only important for median rate of spread. We hypothesise

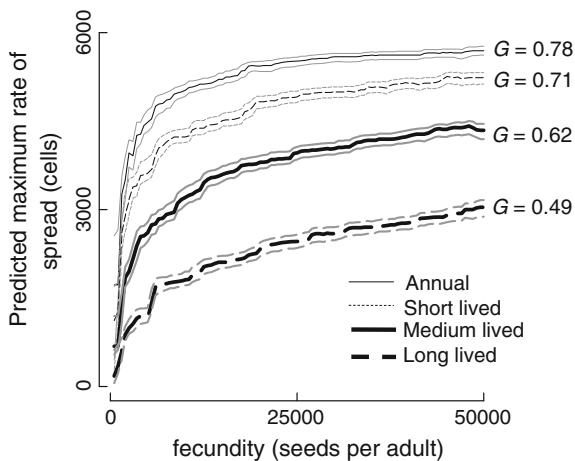


Fig. 4 Partial dependence plots showing the marginal effect of fecundity on maximum rate of spread for long dispersers on a continuous landscape for all life history categories. *Grey lines* show upper and lower 95% confidence intervals, calculated using a percentile bootstrap. Because landscape size was scaled with life history the two shorter lived categories cannot be compared with the two longer lived categories. However, annual and short lived categories (*thin lines*) can be compared to each other, as can medium and long lived categories (*thick lines*). G (Eq. 6 in text) is scale free, and so can be compared between all life history categories

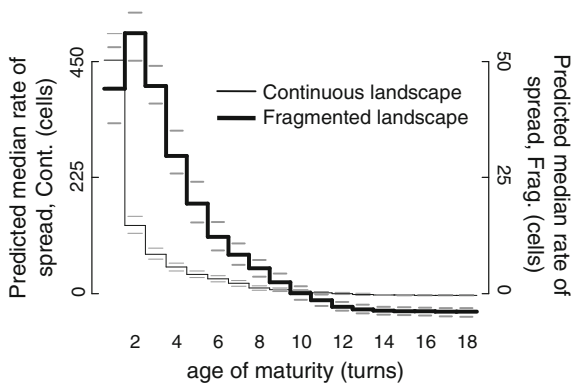


Fig. 5 Partial dependence plot showing the marginal effect of age of maturity on median rate of spread on continuous and fragmented landscapes for the long lived life history category with short dispersal. For ease of comparison between curves the scale for continuous landscapes is shown on the left y-axis and the scale for fragmented landscapes is shown on the right y-axis. *Grey lines* show upper and lower 95% confidence intervals, calculated using a percentile bootstrap. $G = 0.84$ for the continuous landscape, $G = 0.53$ for the fragmented landscape, see Eq. 6 in text for definition of G

this is because the edges of invaded patches supplied a lot of seed to unoccupied cells. Higher age of maturity introduced a longer pause before newly

occupied edge cells produced seed. This would not affect maximum rate of spread, which could be recorded in a time step after the edge of the largest patch had reached adulthood. Survivorship was only important for the two longer lived life history categories and was more important when dispersal was short. Because age of maturity was higher for longer lived life history categories lower survivorship meant fewer juveniles would have reached adulthood.

Two parameters were generally important in driving spread irrespective of context. Mean dispersal distance was important for all life history categories for both maximum and median rates of spread. Fecundity was important under a majority of scenarios tested, particularly for annual and short lived life history categories. This supports previous work on spread (Jongejans et al. 2008), and is in agreement with guidelines for slowing local population growth (Ramula et al. 2008).

The large number of interactions suggests there is not going to be a “one size fits all” solution to managing spread; however, we suggest four general guidelines for managers of invasive plants based on the results of this work.

Guideline 1 manage dispersal to manage spread

Mean dispersal distance was found to be the most important driver of spread for all life history categories. Thus, more effort should go into managing spread for those species for which containment and prevention of dispersal is feasible. Dispersal could be managed by controlling dispersal vectors directly, e.g., controlling seed feeding birds (Gosper et al. 2005; Buckley et al. 2006); or targeting surveillance and plant removal at satellite populations (Moody and Mack 1988; Fox et al. 2009; Pokorný and Krueger-Mangold 2007), or populations in areas with a high frequency of dispersal vectors (Buckley et al. 2005).

Guideline 2 what you measure changes how you view the drivers of spread

Maximum and median rates of spread were driven by different factors. For example fecundity tended to be more important in driving maximum rate of spread, while age of maturity was more important in driving median rate of spread. This suggests that those interested in the spread of plants, either as managers

or researchers, need to use a definition of spread that is appropriate to their needs; a similar message emerges from the management of mountain pine beetle (*Dendroctonus ponderosae*) (Trzcinski and Reid 2008). We used two measures of spread, but others will be relevant in different contexts; for example, defining spread as the furthest forward individual would be appropriate if the goal is to stop any individuals of a species reaching an important area that lies ahead of the invasion front.

Guideline 3 targeting demographic rates will have variable effectiveness in controlling spread

There are two mechanisms behind this guideline: key drivers of spread will vary between populations due to spatial and temporal variation in dispersal opportunities (Jongejans et al. 2008; Martinez and Gonzalez-Taboada 2009); and there is a strong “L” shaped response of spread to the most common demographic drivers, fecundity and age of maturity. For example, if one population’s fecundity is low and another population’s is very high, then control in the first population may have dramatic results, but very little effect in the second population.

Guideline 4 managing spread with context dependant drivers

While it is true that the effect of key demographic rates on spread changed with dispersal, landscape and/or life history conditions, common IAP control methods, such as herbicides and bio-control, can attack several important demographic rates simultaneously. Gall forming wasps, for example, can reduce both fecundity and survivorship (Hoffmann et al. 2002). Thus many control options will be beneficial in a of range contexts. When a control option that affects multiple demographic rates is not available context should be considered, but life history category, and in some cases landscape structure, are easy to deduce with limited biological data.

Conclusion

Dispersal is the main factor that drives spread, but targeting it specifically is difficult. Demographic rates also affect spread and if control strategies target

fecundity, survivorship and/or age of maturity we predict that spread rates will be reduced. A major caveat is that demographic rates (particularly fecundity) have to be moved to a point where they will have an effect. In many cases this point will be unknown. We also find that different types of spread may be driven by different factors, and thus spread needs to be narrowly defined in order to target management at, or determine, its key drivers. For researchers, when there is no obvious choice for a spread metric several should be used to see how results differ between them.

Acknowledgments This work was funded by the Condamine Alliance, Queensland Murray-Darling Committee, the Co-operative Research Centre for Australian Weed Management and the Australian Research Council (LP0667489). YMB is supported through an Australian Research Council Australian Research Fellowship (DP0771387). Thanks to Ric Colasanti for initial modelling ideas and discussion.

References

- Buckley YM, Brieser DT, Rees M (2003) Demography and management of the invasive plant species *Hypericum perforatum*. II. Construction and use of an individual-based model to predict population dynamics and the effects of management strategies. *J Appl Ecol* 40:494–507
- Buckley YM, Brockerhoff E, Langer L, Ledgard N, North H, Rees M (2005) Slowing down a pine invasion despite uncertainty in demography and dispersal. *J Appl Ecol* 42:1020–1030
- Buckley YM, Anderson S, Catterall CP, Corlett RT, Engel T, Gosper CR, Nathan R, Richardson DM, Setter M, Spiegel O, Vivian-Smith G, Voigt FA, Weir JES, Westcott DA (2006) Management of plant invasions mediated by frugivore interactions. *J Appl Ecol* 43:848–857
- Bullock JM, Clarke RT (2000) Long distance seed dispersal by wind: measuring and modelling the tail of the curve. *Oecologia* 124:506–521
- Canty A, Ripley B (2010) Boot: bootstrap R (S-Plus) functions. R package version 1.2, p 42
- Clark JS (1998) Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. *Am Nat* 152:204–224
- Clark JS, Macklin E, Wood L (1998) Stages and spatial scales of recruitment limitation in southern Appalachian forests. *Ecol Monogr* 68:213–235
- Clark JS, Lewis M, Horvath L (2001) Invasion by extremes: population spread with variation in dispersal and reproduction. *Am Nat* 157:537–554
- Cowell FA (1995) Measuring inequality, 3rd edn. Prentice Hall/Harvester Wheatsheaf, London
- Dauer JT, Mortensen DA, Humston R (2006) Controlled experiments to predict horsetweed (*Conyza canadensis*) dispersal distances. *Weed Sci* 54:484–489

- Davison AC, Hinkley D (1997) Bootstrap methods and their application. Cambridge University Press, Cambridge
- Davies KW, Sheley RL (2007) A conceptual framework for preventing the spatial dispersal of invasive plants. *Weed Sci* 55:178–184
- Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *J Anim Ecol* 77:802–813
- Fox JC, Buckley YM, Panetta FD, Bourgoin J, Pullar D (2009) Surveillance protocols for management of invasive plants: modelling Chilean needle grass (*Nassella neesiana*) in Australia. *Divers Distrib* 15:577–589
- Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29:1189–1232
- Gosper CR, Stansbury CD, Vivian-Smith G (2005) Seed dispersal of fleshy-fruited invasive plants by birds: contributing factors and management options. *Divers Distrib* 11:549–558
- Grimm V, Berger U, Bastiansen F et al (2006) A standard protocol for describing individual-based and agent-based models. *Ecol Model* 198:115–126
- Hastings A, Cuddington K, Davies KF, Dugaw CJ, Elmendorf S, Freestone A, Harrison S, Holland M, Lambrinos J, Malvadkar U, Melbourne BA, Moore K, Taylor C, Thomson D (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecol Lett* 8:91–101
- Higgins SI, Richardson DM (1998) Pine invasions in the southern hemisphere: modelling interactions between organism, environment and disturbance. *Plant Ecol* 135:79–93
- Higgins SI, Richardson DM (1999) Predicting plant migration rates in a changing world: the role of long-distance dispersal. *Am Nat* 153:464–475
- Hoffmann JH, Impson FAC, Moran VC, Donnelly D (2002) Biological control of invasive golden wattle trees (*Acacia pycnantha*) by a gall wasp, *Trichilogaster* sp. (Hymenoptera: Pteromalidae), in South Africa. *Biol Control* 25:64–73
- Jongejans E, Shea K, Skarpaas O, Kelly D, Sheppard AW, Woodburn TL (2008) Dispersal and demography contributions to population spread of *Carduus nutans* in its native and invaded ranges. *J Ecol* 96:687–697
- Kot M, Lewis MA, van den Driessche P (1996) Dispersal data and the spread of invading organisms. *Ecol* 77:2027–2042
- Le Maitre DC, Krug RM, Hoffmann JH, Goydon AJ, Mgidi TN (2008) *Hakea sericea*: development of a model of the impacts of biological control on population dynamics and rates of spread of an invasive species. *Ecol Model* 212:342–358
- Malanson GP, Cairns DM (1997) Effects of dispersal, population delays, and forest fragmentation on tree migration rates. *Plant Ecol* 131:67–79
- Martinez I, Gonzalez-Taboada F (2009) Seed dispersal patterns in a temperate forest during a mast event: performance of alternative dispersal kernels. *Oecologia* 159:389–400
- Moody ME, Mack RN (1988) Controlling the spread of plant invasions—the importance of nascent foci. *J Appl Ecol* 25:1009–1021
- Neubert MG, Caswell H (2000) Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecol* 81:1613–1628
- Pokorny ML, Krueger-Mangold JM (2007) Evaluating Montana's dyer's woad (*Isatis tinctoria*) cooperative eradication project. *Weed Technol* 21:262–269
- R Development Core Team (2006) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL:<http://www.R-project.org/>
- Ramula S, Knight TM, Burns JH, Buckley YM (2008) General guidelines for invasive plant management based on comparative demography of invasive and native plant populations. *J Appl Ecol* 45:1124–1133
- Rees M, Paynter Q (1997) Biological control of Scotch broom: modelling the determinants of abundance and the potential impact of introduced insect herbivores. *J Appl Ecol* 34:1203–1221
- Ridgway G (2010) gbm: generalized boosted regression models. R package version 1.6–3.1. <http://CRAN.R-project.org/package=gbm>
- Shea K, Kelly D (1998) Estimating biocontrol agent impact with matrix models: *Carduus nutans* in New Zealand. *Ecol Appl* 8:824–832
- Skarpaas O, Shea K (2007) Dispersal patterns, dispersal mechanisms, and invasion wave speeds for invasive thistles. *Am Nat* 170:421–430
- Trzcinski MK, Reid ML (2008) Effect of management on the spatial spread of mountain pine beetle (*Dendroctonus ponderosae*) in Banff National Park. *For Ecol Manage* 256:1418–1426
- With KA (1997) The application of neutral landscape models in conservation biology. *Conserv Biol* 11:1069–1080
- With KA, King AW (1999) Dispersal success on fractal landscapes: a consequence of lacunarity thresholds. *Landscape Ecol* 14:73–82
- Yokomizo H, Possingham HP, Thomas MB, Buckley YM (2009) Managing the impact of invasive species: the value of knowing the density-impact curve. *Ecol Appl* 19:376–386