

Received Date : 16-May-2016

Revised Date : 25-Aug-2016

Accepted Date : 21-Oct-2016

Article type : Standard Papers

Handling Editor: Tien Ming Lee

Using network connectivity to prioritise sites for the control of invasive species

George L. W. Perry¹, Kirk A. Moloney² and Thomas R. Etherington^{1,3,4}

1 - School of Environment, University of Auckland, Private Bag 92019, Auckland, New Zealand

2 – Department of E.E.O.B., Iowa State University, Ames, IA 50011-4009, USA

3 – Royal Botanic Gardens Kew, Wakehurst Place, West Sussex, RH17 6TN, UK

4 – Department of Zoology, South Parks Road, University of Oxford, Oxford OX1 3PS, UK

Corresponding author: george.perry@auckland.ac.nz

Running title: Using connectivity to limit invasive species spread

Summary

1. Habitat connectivity is a crucial determinant of population dynamics in fragmented landscapes. The corollary of the emphasis on maintaining connectivity to enhance the

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1365-2664.12827

This article is protected by copyright. All rights reserved.

Accepted Article

movement of organisms is that disrupting connectivity should minimise it. Here we evaluate the efficiency of an invasive species control strategy that targets the most connected habitats in a landscape.

2. A network (spatial graph) provides an intuitive representation of a landscape, and the topology of this network can be used to identify the most connected patches. We implemented a simulation model of the spread of an invasive species on a network and used it to evaluate whether targeting the better-connected components of the landscape enhances control effectiveness.
3. Control strategies based on network topology consistently outperformed both a null strategy of random habitat selection and one based on separation distance alone. The advantages of the connectivity-based strategy were strongest in the early phases of the invasion process, when a small number of habitats are occupied at low population density. However, if long distance dispersal events were common the advantages of the connectivity approach weakened.
4. The performance of the connectivity-based strategy is robust to habitat-level demographic stochasticity. In fact, connectivity-based targeting outperforms a strategy focussing on source habitats, with the additional benefit that it requires less information to be implemented.
5. *Synthesis and applications.* Our simulation model outcomes demonstrate that deliberately targeting the best-connected components of a landscape is an efficient control strategy for invasive species when long-distance dispersal is infrequent, and it is likely to be cheaper than other alternatives such as targeting population sources. Network scientists have developed a range of methods designed to identify the minimal set of nodes on a graph that will disrupt the network as a whole; these tools have potential to aid in the design of more effective control strategies.

Key-words: habitat connectivity, control strategy, invasion, network, population dynamics, simulation model, graph theory, invasive species,

Introduction

How biological entities, whether genes or organisms, spread through space is a fundamental and long-standing ecological question (Fisher 1937; Skellam 1951). In spatially heterogeneous landscapes the dynamics of spread are difficult to predict and emerge from interactions between an organism's life-history traits and the spatial structure of the landscape itself (Andow *et al.* 1990; With 2002; Hastings *et al.* 2005; Caplat, Coutts & Buckley 2012). Understanding how organisms disperse through such landscapes requires understanding how the elements in the landscape are connected. Conservation biologists have emphasised maintaining and promoting connectivity in order to facilitate the movement of plants and animals through fragmented landscapes (Fahrig 2003). This view, rooted in the equilibrium theory of island biogeography (MacArthur & Wilson 1967), underpins the often proposed use of habitat corridors and stepping stone habitats to enhance connectivity. However, if enhancing connectivity increases the movement of taxa through the landscape then reducing it should inhibit the movement of less desirable organisms such as invasive species or pathogens (Vilà & Ibáñez 2011; Rudnick *et al.* 2012; Etherington 2015). We will examine the latter approach in this paper.

Spatial networks (or graphs), composed of discrete habitat patches (nodes or vertices) connected by corridors (edges or links) between those habitats, provide an intuitive representation of how the components of a landscape are connected (Minor & Urban 2008; Urban *et al.* 2009; Galpern, Manseau & Fall 2011; Caplat *et al.* 2012). Network-based

approaches have been widely used to assess the ability of organisms to move between spatially isolated patches of habitat (Lookingbill *et al.* 2010). Despite concern that increased landscape connectivity might also favour the spread of undesirable taxa (Haddad *et al.* 2014), few studies have directly addressed this link in the context of invasive species, especially at the landscape level (Wilkerson 2013). However, Banks *et al.* (2015) demonstrate the importance of network topology in global trade networks in mediating human-assisted invasions and suggest that this information could be used to limit pre-border incursions. Likewise, considerable attention has been given to understanding network connectivity in the context of the spread of disease (Keeling & Eames 2005; Danon *et al.* 2011). For example, using simulations of disease spread through a social (rather than spatial) network (drawing on the Italian cattle industry), Natale *et al.* (2009) showed that well-connected ('central') nodes enable and facilitate spread and could be targeted to effectively control epidemic expansion.

A fundamental question in network science is identifying the minimal number of nodes (e.g. habitat elements in a landscape) that need to be removed or controlled to limit spread of some phenomena of interest on a graph. This question is directly relevant to the control of invasive species because survey and control activity is usually limited to a subset of all habitats (Keith & Spring 2013; Glen, Pech & Byrom 2013). It is well known that as nodes are removed from some network structures there are thresholds at which connectivity abruptly collapses (Callaway *et al.* 2000). However, in some networks (e.g. scale-free graphs) even high-intensity random control is insufficient to stop the spread of a disease or an invasive species (Pastor-Satorras & Vespignani 2002). Many metrics have been developed to quantify network topology and especially its connectivity (Borgatti 2005; Laita, Kotiaho & Mönkkönen 2011). One commonly used topological metric is betweenness-centrality (Freeman 1978), which seeks to identify the network components through which flows will

Accepted Article

be concentrated. In an ecological context, topological measures have been shown to be better predictors of the movement of organisms through the landscape than distanced-based measures alone (Lookingbill *et al.* 2010). Minor & Gardner (2011) demonstrated that, so long as long-distance dispersal rates were sufficiently low, targeting large, well-connected habitats inhibited the spread of invasive plants (hypothetical species) in a real landscape (Maryland, U.S.A). Likewise, Ruzenfeld *et al.* (2008) found that centrality measures successfully identified key stepping stone habitats (those where gene flow was highest) in a network describing the movement of the seagrass *Posidonia oceanica* in the Mediterranean. An important advantage of the use of connectivity metrics to inform landscape-level management is that, assuming the network structure is known, they are much cheaper to estimate than other possible factors influencing movement such as habitat quality or demographic parameters (see also Calabrese & Fagan 2004). Given that topological metrics may be able to identify the network components most important for spread, they could inform the prioritisation of control sites for invasive species.

Here we address the potential benefits of prioritising control sites based on their connectivity. Targeting control on the basis of connectivity should increase the isolation of local populations (Glen *et al.* 2013) and thus slow the rate of spread and potentially limit population growth via allee effects (Taylor & Hastings 2005). To this end, we develop a simple spatially explicit population model to evaluate the potential efficiency gains of control strategies targeted at reducing landscape-level connectivity. We hypothesise (as per Albert, Jeong & Barabási 2000; Pastor-Satorras & Vespignani 2002) that approaches targeting highly connected components will out-perform strategies where control locations are either randomly selected or selected on the basis of distance alone, especially in the early phases of an invasion when an organism is still spreading into unoccupied habitats. The strategy based

on distance alone represents situations where control is concentrated on habitats that are close to other habitats; such a strategy uses spatial, but not topological, information. Under this strategy control might be concentrated in locally aggregated areas under the assumption that limiting movement between these will inhibit spread to other parts of the landscape. We use our model to explore three overarching questions:

1. Is the efficacy of control improved by selecting control locations based on their connectivity relative to approaches based on either a null model of random selection or distance to nearest habitats?
2. Are efficiency gains from selecting control sites based on their connectivity consistent across all phases of an invasion and under different dispersal conditions?
3. Does the level of demographic stochasticity in a spatial population affect this decision?

Materials and methods

A graph-based landscape representation

We simulate the landscape as a graph with nodes representing the habitat patches the organisms can occupy in a simulation arena of an arbitrary 100×100 units. We construct the graph from a point pattern – in all cases a homogeneous Poisson process with 100 events – and then connect the nodes using a relative neighbourhood graph (RNG) algorithm (Toussaint 1980; Dale & Fortin 2014). We fix the number of nodes to minimise the effects of network density on connectivity metric performance (Laita *et al.* 2011). The RNG joins two nodes so long as there is not a third node closer to either of these nodes than they are to each other, and produces graphs with connectedness intermediate between minimum spanning trees and Delaunay tessellations (Toussaint 1980). As Galpern *et al.* (2011) note, since the

graph's links denote the possible exchange of individuals, rather than habitat *per se*, the graph provides a map of potential rather than structural connectivity (see Tischendorf & Fahrig 2000; Calabrese & Fagan 2004). The RNG network creates a flexible landscape model that phenomenologically represents a variety of invasion systems such as woodland connected by hedgerows, towns connected by roads, reserves connected by dispersal across urban areas, lakes connected by human transportation processes (Figure 1).

Each node on our spatial graphs represents a discrete patch of occupiable habitat and is characterised by a local population growth rate (r_i) and carrying capacity (K_i), with r_i and K_i drawn from a log-normal distribution of specified mean and standard deviation without spatial autocorrelation (setting both standard deviations to zero yields a system where all habitat is identical). When the standard deviation of r is non-zero, all habitats have a positive local growth rate, but some will be sinks when dispersal (emigration) is considered (as per Runge, Runge & Nichols 2006). We quantify each node's connectivity with the betweenness-centrality metric (Freeman 1978); locations with high centrality have been shown to facilitate spread during the early phases of invasion (Minor & Gardner 2011) and disease spread (Ortiz-Pelaez *et al.* 2006; Natale *et al.* 2009). We assume that the graph is static over the period simulated (i.e. habitats are neither destroyed nor created during any simulation).

Population dynamics

At the start of each simulation the invading population occupies some fraction of the habitats (p_0) with a population size that is some fraction of the habitat's carrying capacity (f_0); varying these two parameters (p_0 and f_0) allows us to represent different phases of an invasion. For example, small local population sizes (low f_0) in a small proportion of habitats (low p_0)

represents the initial phases of an invasion. The initially occupied habitats are selected at random. Local populations grow following a discrete-time logistic model (Pastor 2008) and so are unstructured with respect to either sex or age. Once the local population size (n_i) reaches some fraction of K_i a fraction of the individuals in excess of that value attempt to disperse to neighbouring patches that are connected by edges (both the fraction at which dispersal starts and the fraction dispersing are free parameters, but we set them both at 0.3 in all simulations). For example, if $K_i = 40$ and $n_i = 22$ then three individuals will disperse ($40 \times 0.3 = 12$, $22 - 12 = 10$, $10 \times 0.3 = 3$). The model scheduling is organised so that an individual cannot make more than one dispersal attempt per time-step. Dispersing individuals choose an edge at random (initial trials showed the results are robust to preferential selection of shortest edges) and suffer mortality at a rate (m) of:

$$m = 1 - \left(\frac{1}{d_{ij}}\right)^\alpha \quad \text{Eq. 1}$$

where: m is the probability of mortality of an organism dispersing between two patches separated by distance d_{ij} and α is a scalar (set to 0.5 in all cases).

We represented long-distance dispersal (LDD) via a habitat-level rate, with each individual habitat in each year having a probability equal to the number of occupied habitats multiplied by the LDD rate of receiving a long distance disperser (similar to Minor & Gardner 2011).

We do not seek to represent a specific taxa in great detail; rather, we are concerned more generally with how invasive species spread through fragmented landscapes (as represented by the RNG) and how this might best be inhibited. Nevertheless, the model's parameterisation

reflects the dynamics of a species with largely local dispersal such as exotic mammalian pests (e.g. *Trichosurus vulpecula* [brush-tail possum] in New Zealand) or plants with low rates of long distance dispersal (e.g. *Rhododendron ponticum* in the UK; Figure 1). The baseline population growth rate that we use ($r = 0.2$) is based on estimates for brush-tail possums (Ramsey & Efford 2010).

Prioritisation of control sites

At the start of each simulation n_C control habitats are selected and do not change subsequently (unless otherwise stated $n_C = 20$ in the experiments presented below). The control habitats (nodes) can be selected: (i) at random, (ii) on the basis of the mean distance to the other habitats to which it is linked (including reflexive neighbours; Cox 1981), (iii) on the basis of their connectivity, or (iv) on the basis of their demographic characteristics. If the control habitats are selected based on distance, the n_C nodes with the lowest mean distance to the patches they are linked to are selected; this is directly proportional to the closeness centrality metric (Freeman 1978). This distance-based approach represents a strategy that uses spatial information but not network topology, and is based on the idea that the rate of successful dispersal should be highest between patches that are close to each other (as per Eq. 1 and also empirical observation). If the control habitats are selected based on their connectivity we use an iterative approach where the most connected habitat is selected and then removed from the network and connectivity is recalculated, with this process continuing until n_C habitats are designated (Figure 2). The removal of habitats is simply for calculating node-level connectivity, and all habitats are included in the simulations. If demographic characteristics are used to select the habitats then the n_C habitats with the highest local growth rate (r_i) or the highest local carrying capacity (K_i) or the highest summed z -score for r_i and K_i are controlled irrespective of their location (where r_i and K_i are, respectively, the population

growth rate and carrying capacity in the i -th node). Once control starts, at each time-step some fraction of individuals (fixed at 0.9) are removed from each control habitat. In all simulations control started at the 25th time-step. There is a difference in the implied cost of the selection approaches; the random method carries effectively zero cost, the spatial approaches require assessment of the spatial structure of the landscape, which is achievable from static snapshot data, while the demographic approach requires more detailed population knowledge, which is costly and may require multiple years to collect.

We implemented the model in NetLogo 5.3.1 (Wilensky 1999) using the R-NetLogo library (Thiele 2014) and the spatstat (Baddeley & Turner 2005), spatgraphs (Rajala 2012) and igraph (Csardi & Nepusz 2006) R libraries for point pattern and network analyses.

Experiments

To evaluate the relative efficiency of connectivity-based prioritisation of control sites we conducted a series of *in silico* experiments. In each experiment we measured the final size of the population, the proportion of occupied habitats (habitat occupancy), the proportion of habitats that were never invaded and the time at which 50% of all habitats were invaded (t_{50}).

In all cases, the simulations ran for 250 time-steps and with one exception r and K were the same in every habitat. The experiments we conducted were (Table 1):

1. Multiple ($n = 1000$) realisations of population dynamics under the random, distance-based and connectivity-based control strategies on the *same* graph, varying the sites initially occupied at the start of each simulation.

2. Multiple realisations of population dynamics under the random and two spatial control strategies on 1000 *different* graphs but with the same macroscopic statistical structure (underlying point process).
3. A systematic analysis of the efficiency of each control strategy at different points in the invasion, from the initial phases (few habitats occupied with small initial population sizes) to established invasions (many habitats occupied with populations close to K_i). We conducted a broad sweep of the parameter space (Experiment 3a in Table 1) and a higher resolution sweep of the part of the parameter space relating to the early phases of an invasion (Expt. 3b in Table 1). The number of local populations controlled was varied as either 5, 20, or 50 sites.
4. The relative effectiveness of the different selection strategies as a function of the number of local populations controlled (n_c), which was varied from 0 to 100.
5. We evaluated how long distance-dispersal effects the relative effectiveness of the three strategies.
6. A repeat of (2) but with demographic stochasticity present in the form of inter-patch variation in r_i and K_i , and with control based on local demography or spatial structure.

In experiments 2-6 we used the same set of spatial graphs across each control strategy, in order to avoid any risk of confounding differences in landscape structure with the scenarios we evaluated. We use graphical and correlative analysis to assess the outcomes of these experiments; we do not use formal inferential statistical analyses for the reasons laid out in White et al. (2014).

Results

Control strategies under different incursion points on the same graph

In experiments where the three control target selection strategies were repeatedly ($n = 1000$) applied to the same graph but the incursion points varied, selection on the basis of connectivity consistently out-performed the other two approaches (proportion occupied habitat: 0.62-**0.73**-0.79, 0.78-**0.78**-0.80 and 0.37-**0.495**-0.62 for the random, distance and connectivity strategies [5th percentile-**median**-95th percentile]; Figure 3). In fact, selection based on mean distance to other habitats performed worse than the null model of random selection, presumably because it results in localised clusters being targeted. If the t_{50} measure is used as an index then all three strategies perform approximately equally well.

Control strategies under different incursion points on different graphs

Evaluating the performance of the three strategies on different graphs (each selection method on the same 1000 graphs) shows a similar trend, with connectivity-based selection again out-performing the others (proportion occupied habitat: 0.61-**0.73**-0.79, 0.61-**0.73**-0.79 and 0.39-**0.51**-0.63 for the random, distance and connectivity strategies; Figure 3). In both experiments the proportion of empty habitats at the end of the simulation is positively correlated with t_{50} ($r = 0.363$) and almost perfectly negatively correlated with proportion of habitats ever invaded and population size ($r = -0.991$ and -0.999 , respectively); thus in the subsequent analyses we use the proportion of occupied habitats as a measure of each strategy's performance.

Effects of invasion phase on spatial efficiency

The connectivity-based approach consistently out-performed the null model of randomly located control sites during the early phases of invasion across the three different levels of control we assessed (5, 20 and 50 sites; Figure 4). The distance-based approach performed somewhat worse than the null model across all the conditions we assessed (Figure 5), but no strategy performed well when less than 20% of sites were controlled (i.e. $n_C < 20$).

Effectiveness under different control intensities

We systematically varied the number of control sites to assess the relative efficacy of the different measures. In this context efficacy can be considered in terms of benefit maximisation (if we can only control n sites what is the best we can achieve?) or cost minimisation (what is the fewest sites that need to be controlled to achieve a pre-determined target?). Figure 6 shows that above some minimum effort (c. 5% of habitat) up to a point where c. 65% of the available habitats (graph nodes) are controlled the connectivity-based strategy outperforms the other methods, and beyond that threshold they perform approximately equally. If the control target were that 65% of habitats were unoccupied then this could be achieved with approximately half as few control sites with a connectivity-based approach (cost minimisation) as compared to the other two approaches; if only 20 sites could be controlled then around 40% fewer habitats are occupied with a connectivity-based strategy (benefit maximisation). The random placement approach out-performs a distance-based approach up to the point where all three strategies perform approximately the same (c. 65%; Figure 6).

Effectiveness under long-distance dispersal

As the rate of long-distance dispersal increased, the benefits of the connectivity-based approach weakened (Figure 7). When the per-habitat probability of LDD was 1×10^{-3} per year all three strategies performed similarly and none stopped the organism spreading through the entire network.

Demographic stochasticity

The experiments presented above do not consider stochasticity in population dynamics; however, spatial variation in demographic rates may be important in population expansion if it drives source-sink type dynamics (Pulliam 1988). To assess the robustness of our findings we introduced stochasticity to the population dynamics, with population growth rate and carrying capacity varying between habitats (in a spatially uncorrelated way). We then tested the connectivity-based approach against one in which control sites were selected based on the habitats with the highest population growth rates, the highest carrying capacity or a combination of the two. The connectivity-based strategy consistently out-performed those based on demography, especially where either inter-habitat variation in carrying capacity or growth rate was low (Figure 8).

Discussion

Using network topology to inform control strategies

In our simulations, using topological information to target network components based on their connectivity consistently out-performed strategies based on either distance or a random

placement null model. The connectivity approach out-performed distance-based methods on the same graph (varying incursion point) and on randomised graphs with the same macroscopic structure (varying network topology). Previous studies have shown that distance-based measures may not adequately describe the interactions between habitat connectivity and organism movement in fragmented landscapes (Lookingbill *et al.* 2010). It is, however, a little surprising that there is no apparent increase in the performance of demographically informed targeting as variability in r and K increase. Travis and Parks (2004) analysed culling strategies for invasive species in a two-patch source-sink model and noted that focussing exclusively on the source habitat was sub-optimal when sinks were weak or dispersal from source to sink was frequent and only weakly density-dependent. In our case, so long as there are some uncontrolled patches acting as source populations then conditions for invasion may be sufficient, and at the levels of control we simulate (20% of all habitats) this is likely to be the case.

The early phases of invasion are particularly challenging to predict because contingency effects and demographic and environmental stochasticity can be overwhelmingly important. Ferrari and Lookingbill (2009) found neither topological distance- nor network-based connectivity measures consistently predict temporal dynamics during the early phases of an invasion. However, in our experiments connectivity-based measures *most* out-performed the other approaches during the early expansion phase. This outcome occurs because the network can be most effectively partitioned into small isolated sub-networks by strategically placing controls before the population has spread to some critical fraction of habitats. In this case, some of the subnetworks are never invaded even if not directly controlled. In summary, under all the conditions we considered where dispersal was local, irrespective of demographic stochasticity, a connectivity-based approach was the best performing strategy.

Under even moderate levels of long-distance dispersal none of the three strategies performed well, confirming the outcomes of Minor & Gardner (2011). In other words, the connectivity-based strategy performs best where most dispersal is localised and between neighbouring habitats. It is no surprise that LDD significantly reduces the performance of the connectivity-based approach because in effect it destroys the network topology by allowing the instantaneous invasion of unconnected patches. That LDD is so important suggests that, for taxa where this is sufficiently frequent, a connectivity-based approach will have only marginal benefits; however, no prioritisation scheme is likely to perform well where random long-distance dispersal is frequent.

As Calabrese & Fagan (2004) emphasise, a fundamental consideration in implementing any control strategy is assessing the information and monetary costs involved in developing the prioritisation (e.g., a site's connectivity, local demographic rates). Clearly, the null model (random site selection) is cheap to implement, but it is inefficient relative to the connectivity-based method. Characterising network topology based on either distance or connectivity is relatively cheap and simply requires that the structure of the network be known, meaning that the prioritisation can occur before any invasion occurs. Our approach assumes that the structure of the network including, critically, the connections between habitats are known in advance. Using a distance-based approach to define the links allows this, but, operationally, this definition should be informed by the biology of the species and the structure of the landscape of interest. Directly assessing connectivity from biological information via least-cost distances (Etherington 2016) or resistance distances (McRae *et al.* 2008) is not easy and so an approach that relies on potential linkages, as ours does, is easier to implement (Zeller, McGarigal & Whiteley 2012). Finally, assessing the local population dynamics at each site in order to derive local demographic rates is costly and requires long-term monitoring *after* a

Accepted Article
site has been invaded. Distance and connectivity-based metrics can be directly estimated from a network's topology, but local demographic rates can only be indirectly assessed (e.g. via habitat quality models) in advance of a site being invaded.

Although many centrality-type measures of network topology are somewhat correlated (Estrada 2007; Minor & Gardner 2011; Li *et al.* 2015), it is an open question as to whether the specific connectivity metric used influences the effectiveness of site prioritisation.

Different connectivity measures capture different facets of landscape structure (Kindlmann & Burel 2008) and so the appropriate metric may be taxon-specific. Although betweenness centrality was effective in informing targeted control, Ferrari & Lookingbill (2009)

concluded that a patch's degree (number of other patches it is connected to) was not a good predictor of spread dynamics and argued that a network-based approach must also consider patch-level processes. Quantitative methods have been developed that seek to identify the minimal subset of a graph that if controlled would limit spread (Ruths & Ruths 2014; Morone & Makse 2015) – these approaches have the potential to inform control strategies for invasive species (see also Hock & Mumby 2015). Our approach is strictly local in that we identify

nodes with the highest centrality, independent of the centrality of neighbouring nodes;

Morone and Makse (2015) outline an approach that takes a node's neighbourhood context into account when assessing its contribution to connectivity and flow. They demonstrate that their 'collective influence' method out-performs approaches that identify only locally well-connected components. It is clear that approaches such as those described by Ruths and Ruths (2014) and Morone and Makse (2015) are applicable to understanding and controlling the spread of invasive species through spatial networks and warrant more attention in this context, especially where movement is not limited to nearest neighbours.

Another extension to our prioritisation approach would be to make it more adaptive by changing the control locations over time, conditioned on the location of occupied sites. In other words, as the invasion proceeds different sites could be targeted as their relative importance changes. Targeting based on connectivity alone will be sub-optimal if the most connected patches are distal from the incursion points and hence unoccupied for long periods of time, and the dynamic selection of control locations would address this issue. Fundamentally, this type of prioritisation involves a trade-off between inhibiting spread to new habitats and eradicating/controlling in sites known to be invaded, and would need to carefully consider the risk of reinvasion (Lookingbill *et al.* 2014; Etherington 2015) as well as the efficiency of the control strategy itself (i.e. the kill rate). An adaptive design could be supported by representing the landscape as a dynamic network graph whose structure changes and emerges as the invasion event unfolds (Ferrari, Preisser & Fitzpatrick 2014); such a design would likely be superior to static approaches in combatting new satellite populations establishing via LDD.

The landscape as a graph

Underlying our approach is an assumption that a spatial graph adequately represents the landscape. The first assumption of representing the landscape as a network is that its components can be dichotomised as either habitat or matrix. There has been considerable discussion as to how reasonable this dichotomy is (Ricketts 2001,) and Wilkerson (2013) argues that for invasive plants the idea of a ‘matrix’ is tenuous as such habitat may favour them. In its simplest form a graph-based representation assumes that nodes are identical in terms of the other abiotic and biotic characteristics that drive the spread of invasive taxa (González-Moreno *et al.* 2013). To some extent this assumption can be relaxed by changing demographic rates in each node, as in our analyses of demographic stochasticity (i.e. node-

level variability in K could reflect patch size or habitat quality), or by assigning nodes other attributes (Galpern *et al.* 2011). However, as the detail with which the node attributes are represented increases, the benefits of the simplicity of a graph-based approach are diluted. While there are more computationally expensive, but ‘realistic’, ways to represent landscapes and the movement of organisms through them (Zeller *et al.* 2012), few studies directly compare them with graph-based models. A notable exception is provided by Minor and Urban (2007) who demonstrated that a graph-based approach and a more detailed spatially explicit individual-based model identified similar landscape components as crucial for the persistence of a declining songbird. Nevertheless, the conditions under which a graph-based representation of a landscape is adequate remains an important question.

Future directions for enquiry

The analyses we present are preliminary and other important issues need to be considered before connectivity measures are used to prioritise field-based control strategies. Perhaps most importantly, our approach is ‘network-centric’ in that biological information does not inform the network topologies we consider. While the RNG structure is phenomenologically similar to the types of networks seen in fragmented landscapes, the extent to which our results are robust to network structure is uncertain. Much of the theoretical understanding that informs our approach is based on analysis of random graphs (Callaway *et al.* 2000) or networks where node degrees are power-law distributed (i.e. scale-free networks; Pastor-Satorras & Vespignani 2001, 2002). Studies on other types of networks derived from landscape maps have not found topological measures (e.g., node degree) to be useful in determining spread dynamics (Ferrari & Lookingbill 2009). The extent to which such networks relate to ‘real’ landscapes is unclear. Likewise, the underlying pattern of the habitats (i.e. the point process model used) may also be important. The influence of the point

process model is, however, somewhat predictable. In graphs where habitat is regularly distributed through the landscape the distribution of node connectivity values will be more homogeneous and so connectivity is less informative than in aggregated landscapes where a few nodes link local clusters together.

Conclusions

Fundamental results in network science suggest that controlling a few well-connected sites can efficiently limit spread through a network. Here we demonstrate that, in the context of the spread of invasive taxa, connectivity-based prioritisation using topological information when deciding *where* to control out-performs approaches using random placement or distance, especially in the early phases of invasion when control is critical. This result is robust to moderate demographic stochasticity, and potentially cheaper in terms of data to implement than demographically targeted approaches. Recently methods have been developed to identify the minimal network sub-components that most inhibit flow. While the prioritisation strategy we assess, and related methods, may be effective for limiting the spread of invasive species, the extent to which they are robust to network connectivity structure is uncertain and remains an important avenue for future research.

Acknowledgments

GP acknowledges the support of a University of Tasmania Visiting Scholar fellowship while this manuscript was being drafted. KAM thanks the University of Auckland for its logistical support and Iowa State University for its financial support, while this research was being developed during an extended stay in the beautiful country of New Zealand.

Data archiving

NetLogo and R scripts: uploaded to Figshare <https://dx.doi.org/10.17608/k6.auckland.4036131>

(Perry *et al.* 2016).

References

- Albert, R., Jeong, H. & Barabási, A.-L. (2000) Error and attack tolerance of complex networks. *Nature*, **406**, 378–382.
- Andow, D.A., Kareiva, P.M., Levin, S.A. & Okubo, A. (1990) Spread of invading organisms. *Landscape Ecology*, **4**, 177–188.
- Baddeley, A. & Turner, R. (2005) spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software*, **12**, Online: <http://www.jstatsoft.org/>.
- Banks, N.C., Paini, D.R., Bayliss, K.L. & Hodda, M. (2015) The role of global trade and transport network topology in the human-mediated dispersal of alien species (ed M Vila). *Ecology Letters*, **18**, 188–199.
- Borgatti, S.P. (2005) Centrality and network flow. *Social Networks*, **27**, 55–71.
- Calabrese, J.M. & Fagan, W.F. (2004) A comparison-shopper's guide to connectivity metrics. *Frontiers in Ecology and the Environment*, **2**, 529–536.
- Callaway, D.S., Newman, M.E.J., Strogatz, S.H. & Watts, D.J. (2000) Network robustness and fragility: percolation on random graphs. *Physical Review Letters*, **85**, 5468–5471.
- Caplat, P., Coutts, S. & Buckley, Y.M. (2012) Modeling population dynamics, landscape structure, and management decisions for controlling the spread of invasive plants: Management of spread of invasive plants. *Annals of the New York Academy of Sciences*, **1249**, 72–83.
- Cox, T.F. (1981) Reflexive nearest neighbours. *Biometrics*, **37**, 367–369.
- Csardi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal*, **Complex Systems**, 1695.
- Dale, M.R.T. & Fortin, M.-J. (2014) *Spatial Analysis: A Guide for Ecologists*, Second Edition. Cambridge University Press, Cambridge ; New York.
- Danon, L., Ford, A.P., House, T., Jewell, C.P., Keeling, M.J., Roberts, G.O., Ross, J.V. & Vernon, M.C. (2011) Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases*, **2011**, 1–28.
- Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns (2nd Ed.)*. Arnold, London, U.K.
- Estrada, E. (2007) Characterization of topological keystone species: local, global and “meso-scale” centralities in food webs. *Ecological Complexity*, **4**, 48–57.

- Etherington, T.R. (2015) Geographical isolation and invasion ecology. *Progress in Physical Geography*, **39**, 697–710.
- Etherington, T.R. (2016) Least-cost modelling and landscape ecology: concepts, applications, and opportunities. *Current Landscape Ecology Reports*, **1**, 40–53.
- Fahrig, L. (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology and Systematics*, **34**, 487–515.
- Ferrari, J.R. & Lookingbill, T.R. (2009) Initial conditions and their effect on invasion velocity across heterogeneous landscapes. *Biological Invasions*, **11**, 1247–1258.
- Ferrari, J.R., Preisser, E.L. & Fitzpatrick, M.C. (2014) Modeling the spread of invasive species using dynamic network models. *Biological Invasions*, **16**, 949–960.
- Fisher, R.A. (1937) The wave of advance of advantageous genes. *Annals of Eugenics*, **7**, 353–369.
- Freeman, L.C. (1978) Centrality in social networks conceptual clarification. *Social Networks*, **1**, 215–239.
- Galpern, P., Manseau, M. & Fall, A. (2011) Patch-based graphs of landscape connectivity: A guide to construction, analysis and application for conservation. *Biological Conservation*, **144**, 44–55.
- Glen, A.S., Pech, R.P. & Byrom, A.E. (2013) Connectivity and invasive species management: towards an integrated landscape approach. *Biological Invasions*, **15**, 2127–2138.
- González-Moreno, P., Pino, J., Carreras, D., Basnou, C., Fernández-Rebollar, I. & Vilà, M. (2013) Quantifying the landscape influence on plant invasions in Mediterranean coastal habitats. *Landscape Ecology*, **28**, 891–903.
- Haddad, N.M., Brudvig, L.A., Damschen, E.I., Evans, D.M., Johnson, B.L., Levey, D.J., Orrock, J.L., Resasco, J., Sullivan, L.L., Tewksbury, J.J., Wagner, S.A. & Weldon, A.J. (2014) Potential negative ecological effects of corridors. *Conservation Biology*, **28**, 1178–1187.
- Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvdkar, U., Melbourne, B.A., Moore, K., Taylor, C. & Thomson, D. (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, **8**, 91–101.
- Hock, K. & Mumby, P.J. (2015) Quantifying the reliability of dispersal paths in connectivity networks. *Journal of The Royal Society Interface*, **12**, 20150013.
- Keeling, M.J. & Eames, K.T.. (2005) Networks and epidemic models. *Journal of The Royal Society Interface*, **2**, 295–307.
- Keith, J.M. & Spring, D. (2013) Agent-based Bayesian approach to monitoring the progress of invasive species eradication programs. *Proceedings of the National Academy of Sciences*, **110**, 13428–13433.
- Kindlmann, P. & Burel, F. (2008) Connectivity measures: a review. *Landscape Ecology*.
- Laita, A., Kotiaho, J.S. & Mönkkönen, M. (2011) Graph-theoretic connectivity measures: what do they tell us about connectivity? *Landscape Ecology*, **26**, 951–967.
- Li, C., Li, Q., Van Mieghem, P., Stanley, H.E. & Wang, H. (2015) Correlation between centrality metrics and their application to the opinion model. *The European Physical Journal B*, **88**, 65.

- Lookingbill, T.R., Gardner, R.H., Ferrari, J.R. & Keller, C.E. (2010) Combining a dispersal model with network theory to assess habitat connectivity. *Ecological Applications*, **20**, 427–441.
- Lookingbill, T.R., Minor, E.S., Bukach, N., Ferrari, J.R. & Wainger, L.A. (2014) Incorporating risk of reinvasion to prioritize sites for invasive species management. *Natural Areas Journal*, **34**, 268–281.
- MacArthur, R.H. & Wilson, E.O. (1967) *The Theory of Island Biogeography*. Princeton University Press, Princeton, N.J.
- McRae, B.H., Dickson, B.G., Keitt, T.H. & Shah, V.B. (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, **89**, 2712–2724.
- Minor, E.S. & Gardner, R.H. (2011) Landscape connectivity and seed dispersal characteristics inform the best management strategy for exotic plants. *Ecological Applications*, **21**, 739–749.
- Minor, E.S. & Urban, D.L. (2007) Graph theory as a proxy for spatially explicit population models in conservation planning. *Ecological Applications*, **17**, 1771–1782.
- Minor, E.S. & Urban, D.L. (2008) A graph-theory framework for evaluating landscape connectivity and conservation planning. *Conservation Biology*, **22**, 297–307.
- Morone, F. & Makse, H.A. (2015) Influence maximization in complex networks through optimal percolation. *Nature*, **524**, 65–68.
- Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G. & Calistri, P. (2009) Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. *Preventive Veterinary Medicine*, **92**, 341–350.
- Ortiz-Pelaez, A., Pfeiffer, D.U., Soares-Magalhães, R.J. & Guitian, F.J. (2006) Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Preventive Veterinary Medicine*, **76**, 40–55.
- Pastor, J. (2008) *Mathematical Ecology of Populations and Ecosystems*. Wiley-Blackwell Pub, Oxford ; Hoboken, NJ.
- Pastor-Satorras, R. & Vespignani, A. (2001) Epidemic spreading in scale-free networks. *Physical Review Letters*, **86**, 3200–3203.
- Pastor-Satorras, R. & Vespignani, A. (2002) Immunization of complex networks. *Physical Review E*, **65**.
- Perry, G.L.W., Etherington, T.R. & Moloney, K.A. (2016) Source code (R and NetLogo) for invasive species targeting by connectivity model. Figshare. <https://dx.doi.org/10.17608/k6.auckland.4036131>
- Pulliam, H.R. (1988) Sources, sinks and population regulation. *American Naturalist*, **132**, 652–661.
- Rajala, T. (2012) *Spatgraphs: Graphs for Spatial Point Patterns*. <http://CRAN.R-project.org/package=spatgraphs>.
- Ramsey, D.S.L. & Efford, M.G. (2010) Management of bovine tuberculosis in brushtail possums in New Zealand: predictions from a spatially explicit, individual-based model: Spatially explicit model of Tb in possums. *Journal of Applied Ecology*, **47**, 911–919.

- Ricketts, T.H. (2001) The matrix matters: effective isolation in fragmented landscapes. *American Naturalist*, **158**, 87–99.
- Rozenfeld, A.F., Arnaud-Haond, S., Hernandez-Garcia, E., Eguiluz, V.M., Serrao, E.A. & Duarte, C.M. (2008) Network analysis identifies weak and strong links in a metapopulation system. *Proceedings of the National Academy of Sciences*, **105**, 18824–18829.
- Rudnick, D.A., Ryan, S.J., Beier, P., Cushman, S.A., Dieffenbach, F., Epps, C.W., Gerber, L.R., Hartter, J., Jenness, J.S., Kintsch, J., Merenlender, A.M., Perkl, R.M., Preziosi, D.V. & Trombulak, S.C. (2012) The role of landscape connectivity in planning and implementing conservation and restoration priorities. *Issues in Ecology*, **16**, 1–20.
- Runge, J.P., Runge, M.C. & Nichols, J.D. (2006) The role of local populations within a landscape context: defining and classifying sources and sinks. *The American Naturalist*, **167**, 925–938.
- Ruths, J. & Ruths, D. (2014) Control profiles of complex networks. *Science*, **343**, 1373–1376.
- Skellam, J.G. (1951) Random dispersal in theoretical populations. *Biometrika*, **38**, 196–218.
- Taylor, C.M. & Hastings, A. (2005) Allee effects in biological invasions. *Ecology Letters*, **8**, 895–908.
- Thiele, J.C. (2014) R marries NetLogo: introduction to the RNetLogo package. *Journal of Statistical Software*, **58**.
- Tischendorf, L. & Fahrig, L. (2000) On the usage and measurement of landscape connectivity. *Oikos*, **90**, 7–19.
- Toussaint, G.T. (1980) The relative neighborhood graph of a finite planar set. *Pattern Recognition*, **12**, 261–268.
- Travis, J.M.J. & Park, K.J. (2004) Spatial structure and the control of invasive alien species. *Animal Conservation*, **7**, 321–330.
- Urban, D.L., Minor, E.S., Treml, E.A. & Schick, R.S. (2009) Graph models of habitat mosaics. *Ecology Letters*, **12**, 260–273.
- Vilà, M. & Ibáñez, I. (2011) Plant invasions in the landscape. *Landscape Ecology*, **26**, 461–472.
- White, J.W., Rassweiler, A., Samhouri, J.F., Stier, A.C. & White, C. (2014) Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos*, **123**, 385–388.
- Wilensky, U. (1999) *NetLogo*. Center for Connected Learning and Computer-Based Modeling, Northwestern University., Evanston, IL.
- Wilkerson, M.L. (2013) Invasive plants in conservation linkages: a conceptual model that addresses an underappreciated conservation issue. *Ecography*, **36**, 1319–1330.
- With, K.A. (2002) The landscape ecology of invasive spread. *Conservation Biology*, **16**, 1192–1203.
- Zeller, K.A., McGarigal, K. & Whiteley, A.R. (2012) Estimating landscape resistance to movement: a review. *Landscape Ecology*, **27**, 777–797.

Figures

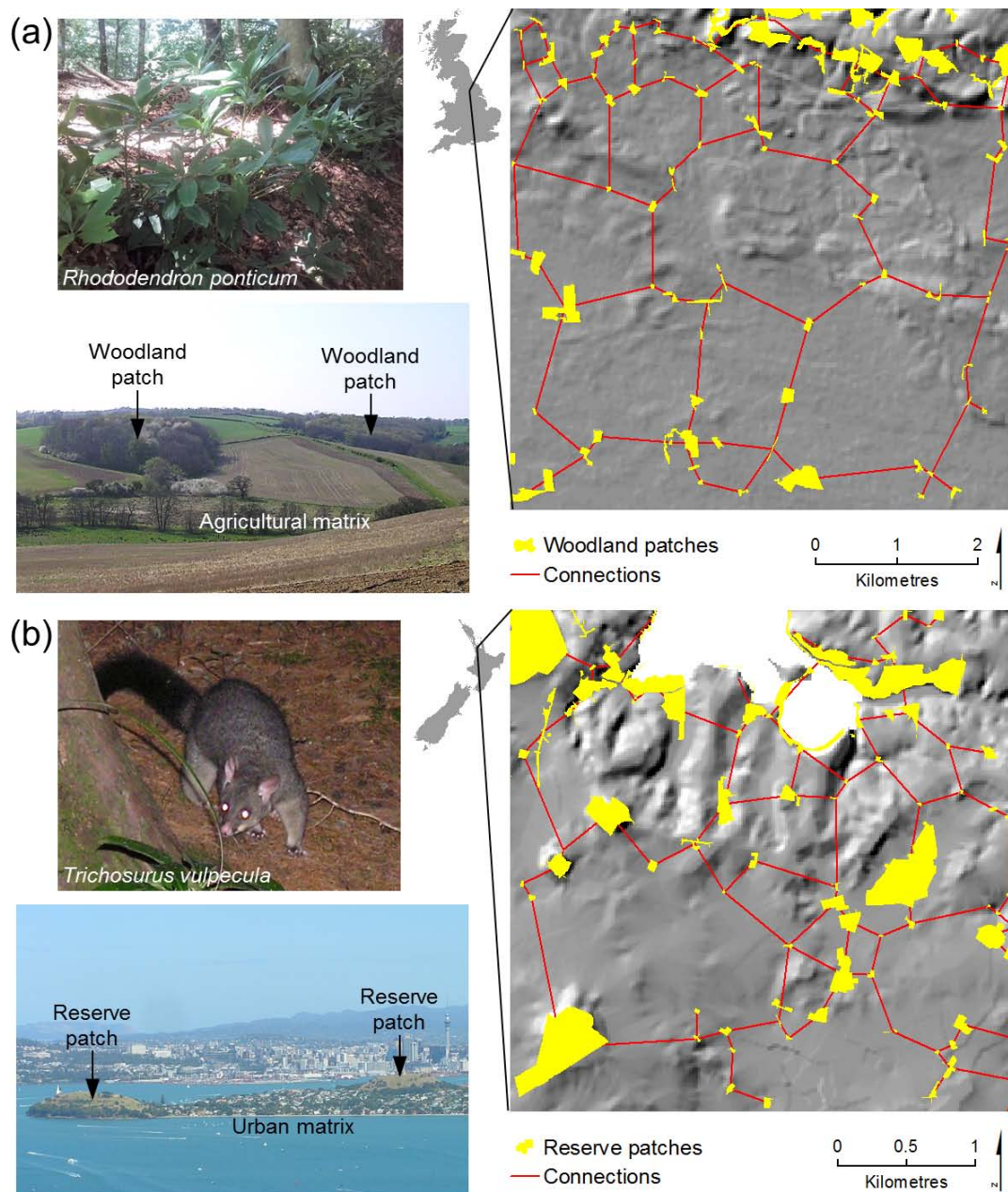


Figure 1. Examples of invasion systems where network connectivity methods could be applied. (a) In the UK *Rhododendron ponticum* invades woodland that often occurs, such as in North Yorkshire, as patches embedded in a broader agricultural landscape. Contains OS data and information supplied by the Forestry Commission. © Crown copyright and database right [2016] Ordnance Survey [100021242] licensed under the Open Government Licence v3.0. (b) In urban Auckland, New Zealand, the brushtail possum (*Trichosurus vulpecula*) occupies reserves that are patches among an urban matrix. Data reproduced with the permission of Landcare Research New Zealand Limited.

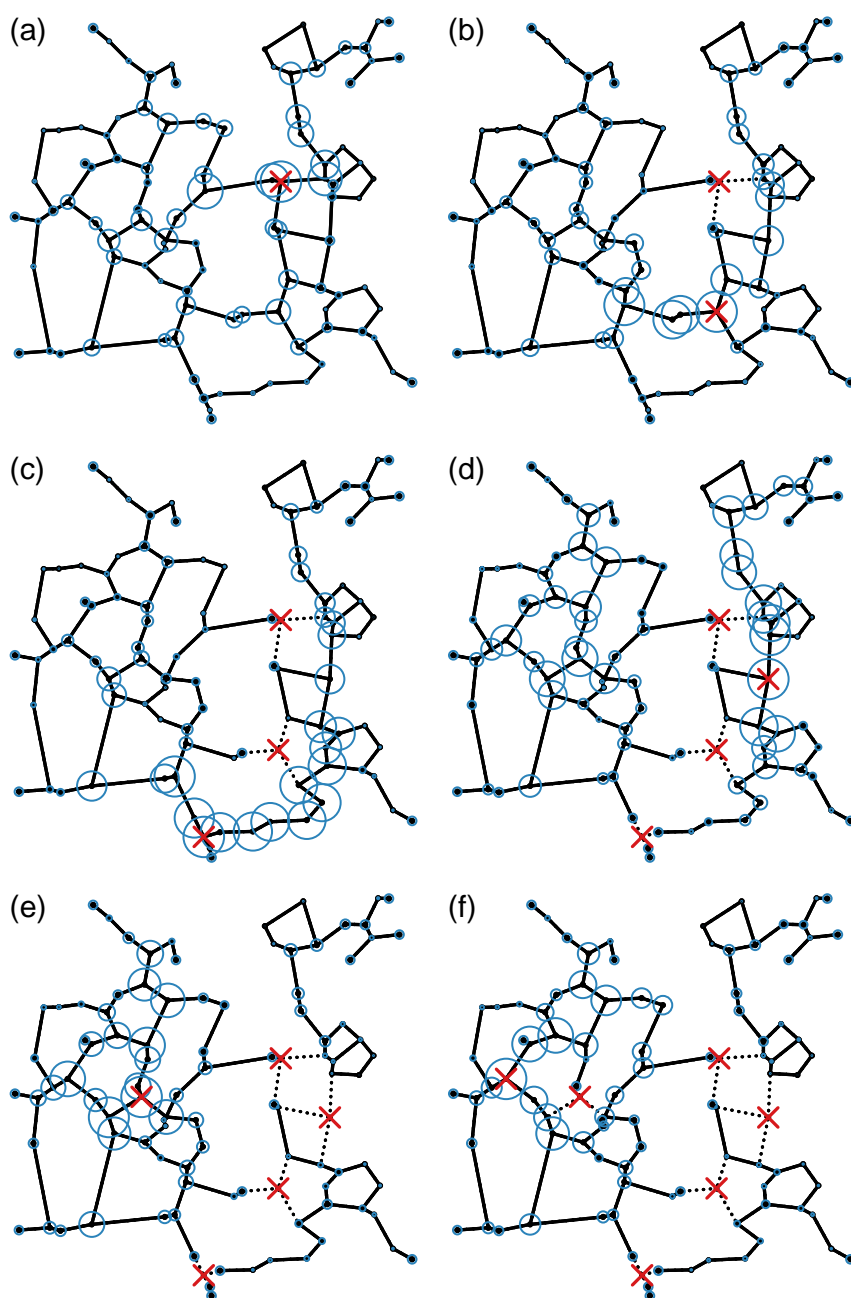


Figure 2 To select control locations based on connectivity, the betweenness centrality (indicated by size of blue circles) of a network was calculated, and the most central node was selected (red cross). The selected node was then removed, the betweenness centrality recalculated, and the most central node was selected and removed. This process continued iteratively (a→ f) until the desired number of control locations was selected – in this case six locations. Note that the nodes removed in the calculation are reinstated for the invasion simulation.

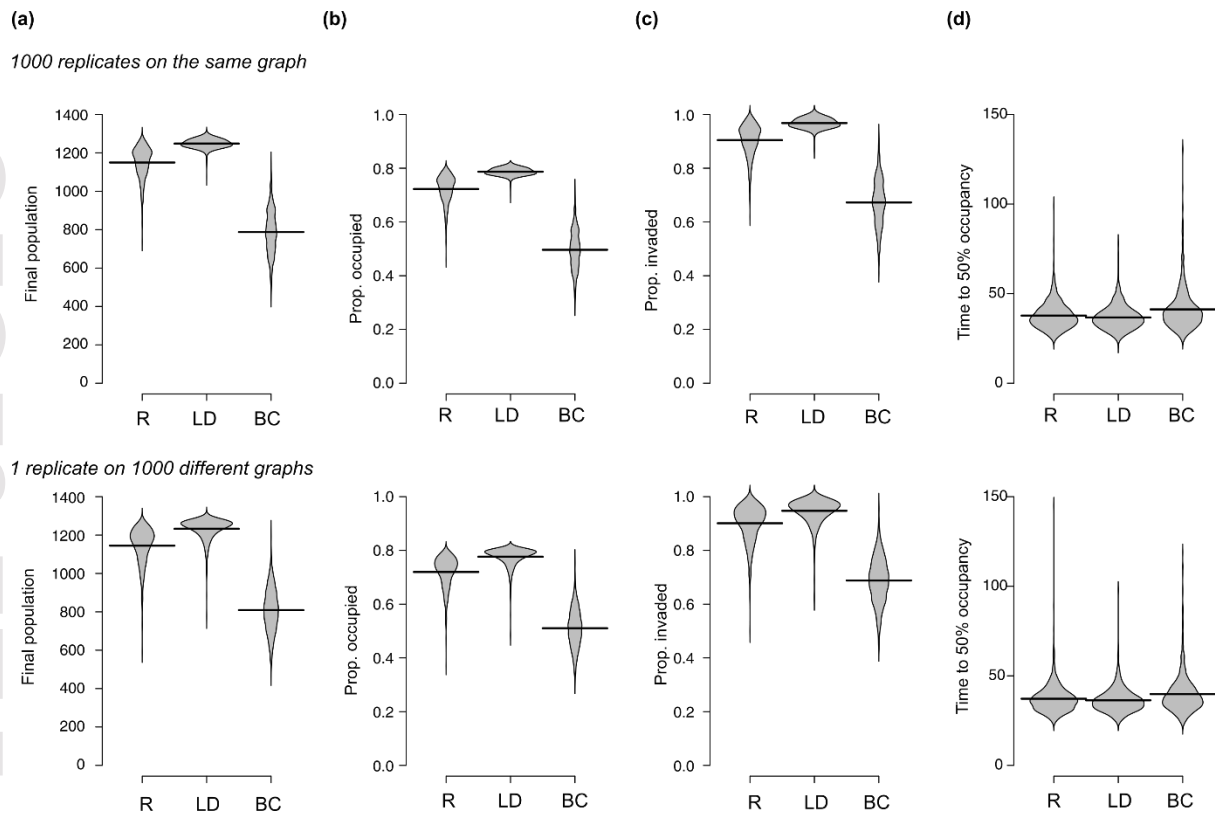


Figure 3 Effectiveness of three different prioritisation strategies – random (R), least distance (LD) and best connected (BC) – as measured by (a) final population size, (b) proportion of habitats occupied at end of simulation, (c) proportion of habitats ever invaded, and time until (d) 50% of habitats occupied (where that occurred); Expts 1 and 2 in Table 1. Top row repeated simulations on the same graph but with different population starting conditions; lower row repeated simulations on different graphs of identical macroscopic structure. Horizontal black bars are distribution means.

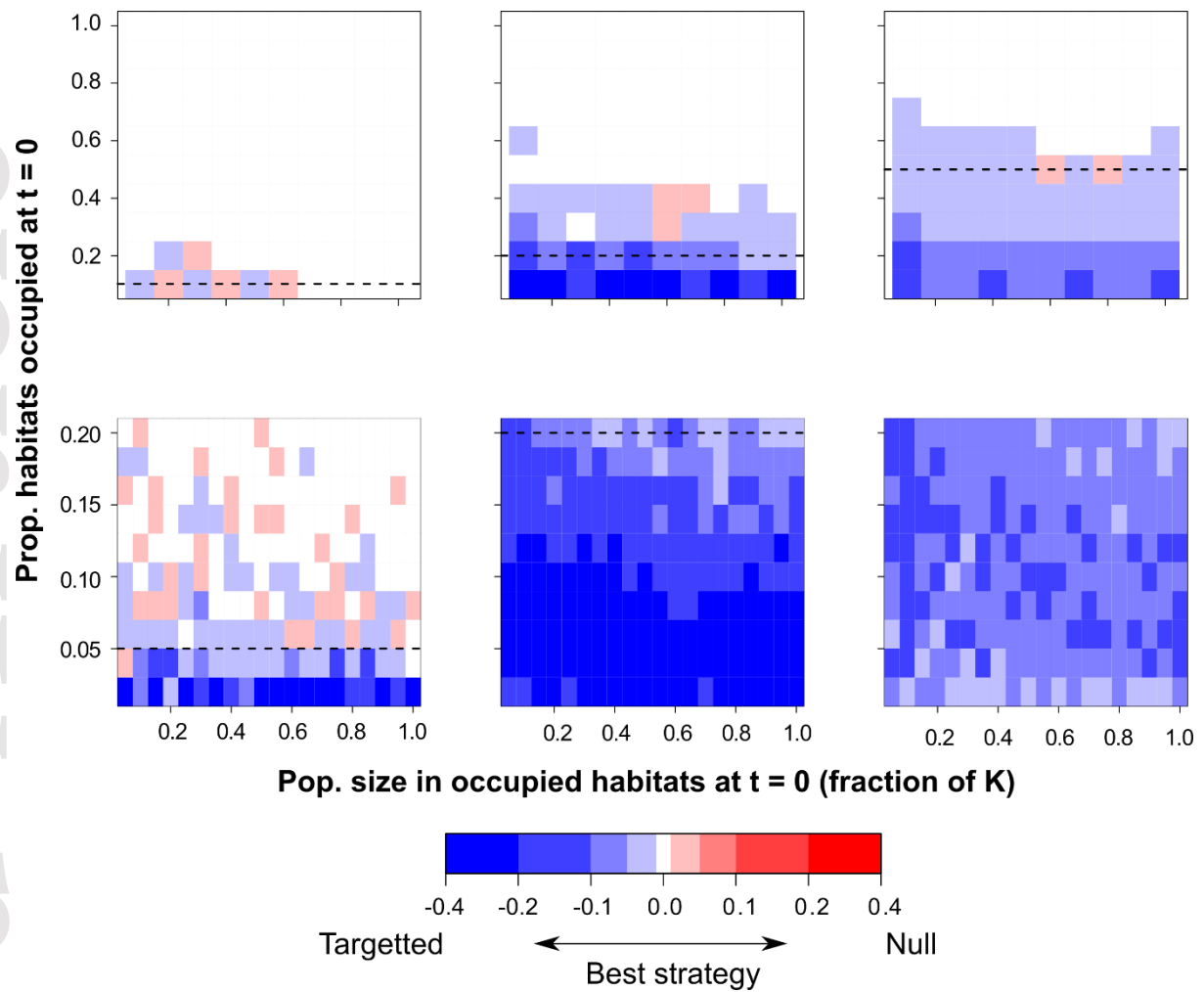


Figure 4 Effectiveness of the connectivity-based prioritisation strategy relative to a null model in different phases of invasion (Expt 3 in Table 1). Plot shows median difference in habitat occupancy (proportion of habitats occupied) after 250 time-steps between the null and connectivity-based strategies from blue (connectivity better) to red (null better) via white (no difference), with the lower row showing results from a higher resolution investigation of the area in the dashed box in the upper row. Negative values denote lower median occupancy and thus more effective control. Left to right are different control intensities: $n_c = 5, 20, 50$ sites (indicated by dashed horizontal lines, which are at the comparable habitat occupancy). The large white areas in the upper part of each plot in the top row are cases where the invasive species spreads through the entire landscape irrespective of control strategy.

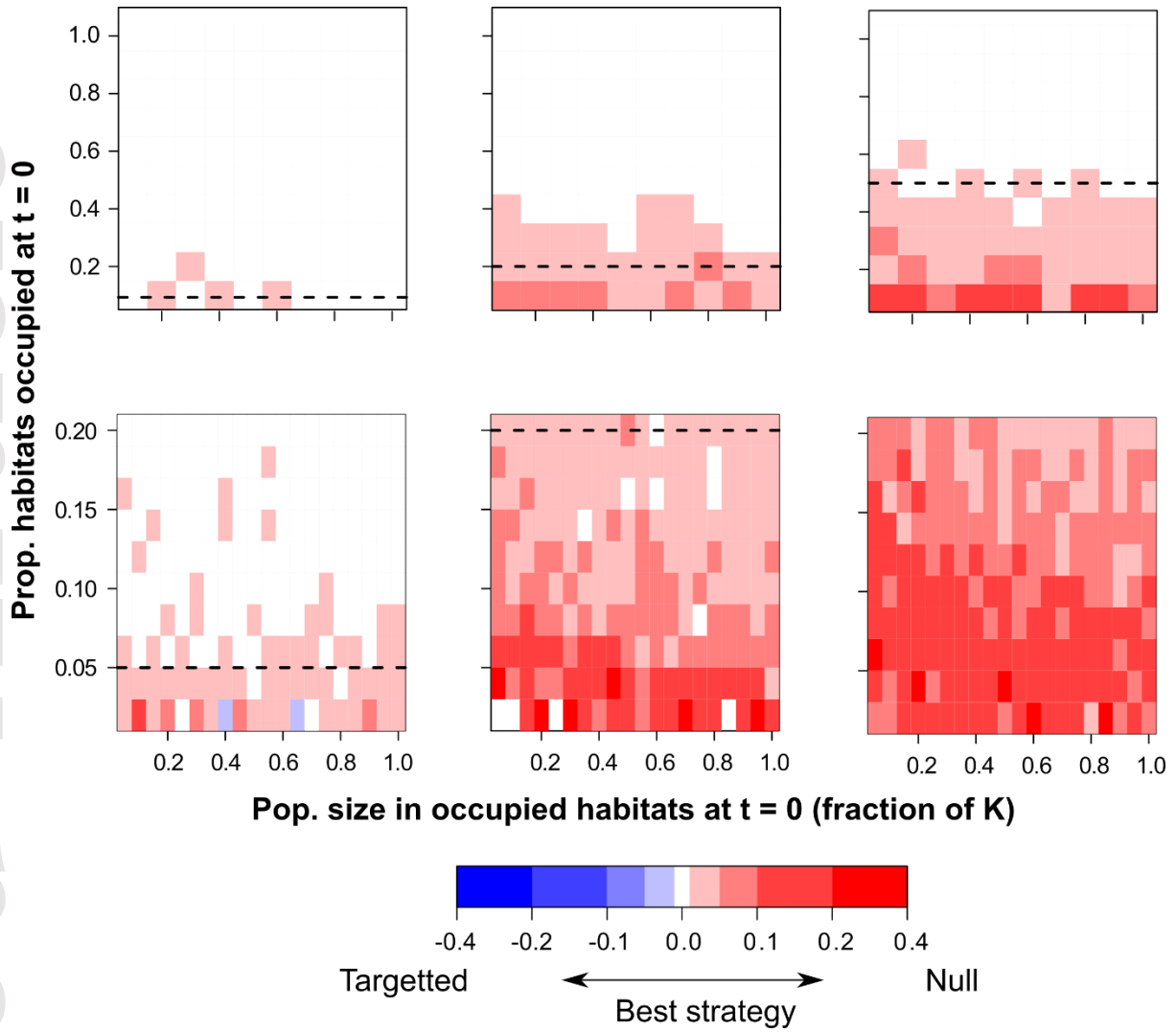
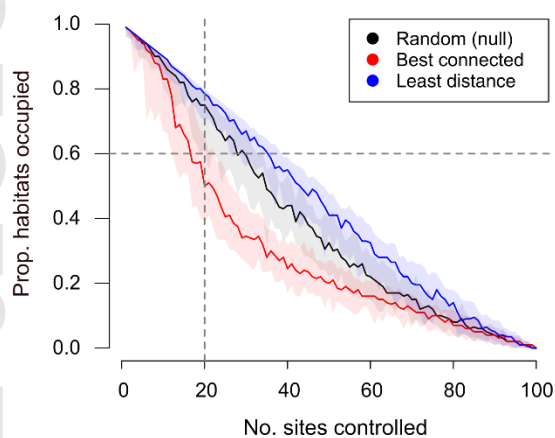


Figure 5 Effectiveness of the distance-based prioritisation strategy relative to a null model in different phases of invasion (Expt 3 in Table 1). Plot shows median difference in habitat occupancy (proportion of habitats occupied) after 250 time-steps between the null and distance-based strategies from blue (distance better) to red (null better) via white (no difference), with the lower row showing results from a higher resolution investigation of the area in the dashed box in the upper row. Negative values denote lower median occupancy and thus more effective control. Left to right are different control intensities: $n_C = 5, 20, 50$ sites (indicated by dashed horizontal lines, which are at the comparable habitat occupancy).

++

(a)



(b)

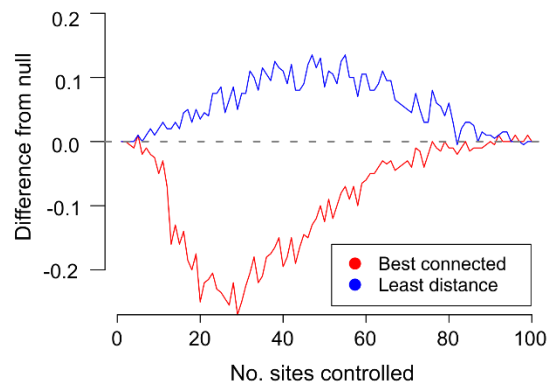


Figure 6 Absolute (a) and relative (b) effectiveness (as measured by proportion of occupied habitats at end of simulation) of the distance- and connectivity-based prioritisation strategies as a function of control intensity (Expt 4 in Table 1). The relative assessment is based on comparison with the null (random placement) strategy. Lines are medians from $n = 30$ replicates, shaded areas in (a) are the 5th-95th percentile limits.

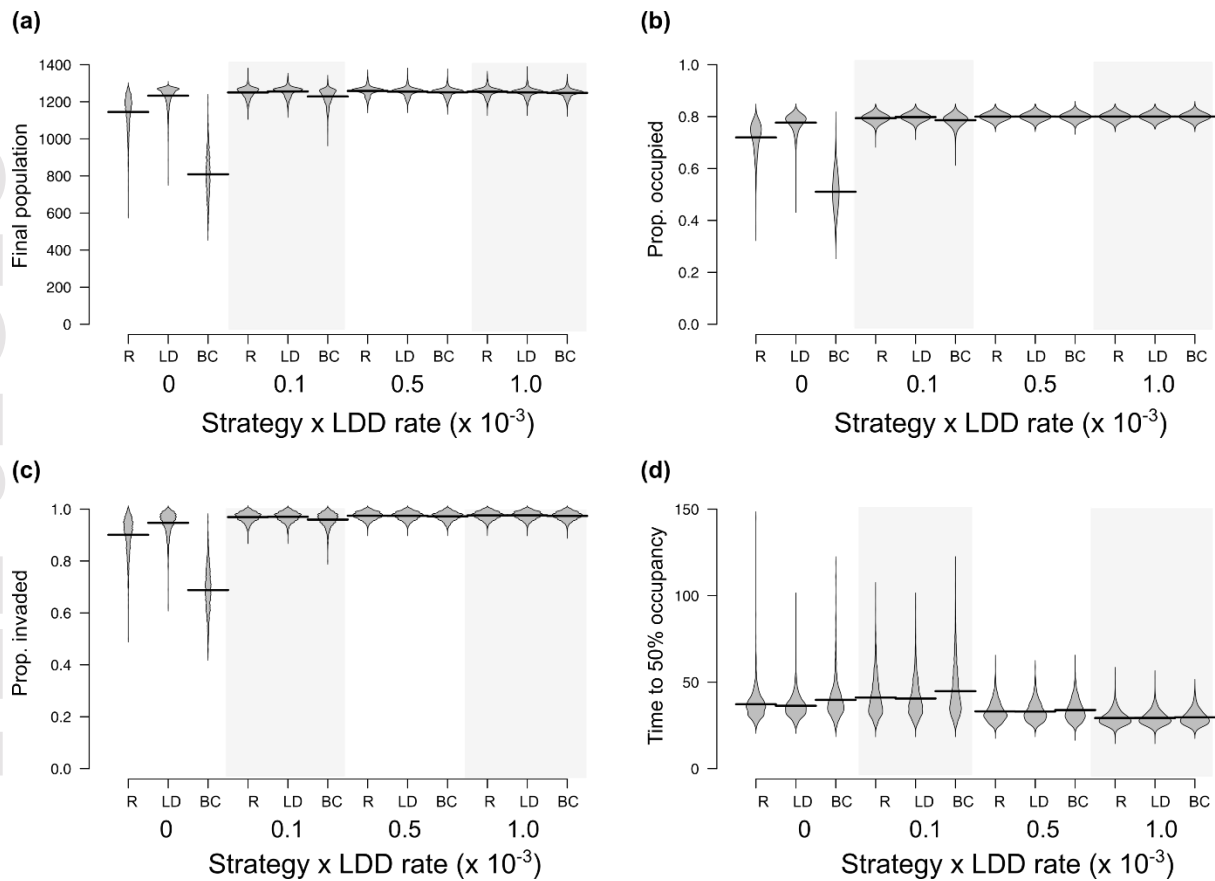


Figure 7 Effectiveness of three different prioritisation strategies – random (R), least distance (LD) and best connected (BC) – under different rates of long distance dispersal (LDD) as measured by (a) final population size, (b) proportion of habitats occupied at end of simulation, (c) proportion of habitats ever invaded, and time until (d) 50% of habitats occupied (where that occurred). This is Expt. 5 in Table 1. LDD rates are per time-step per habitat probability of a LDD event ($\times 10^{-3}$).

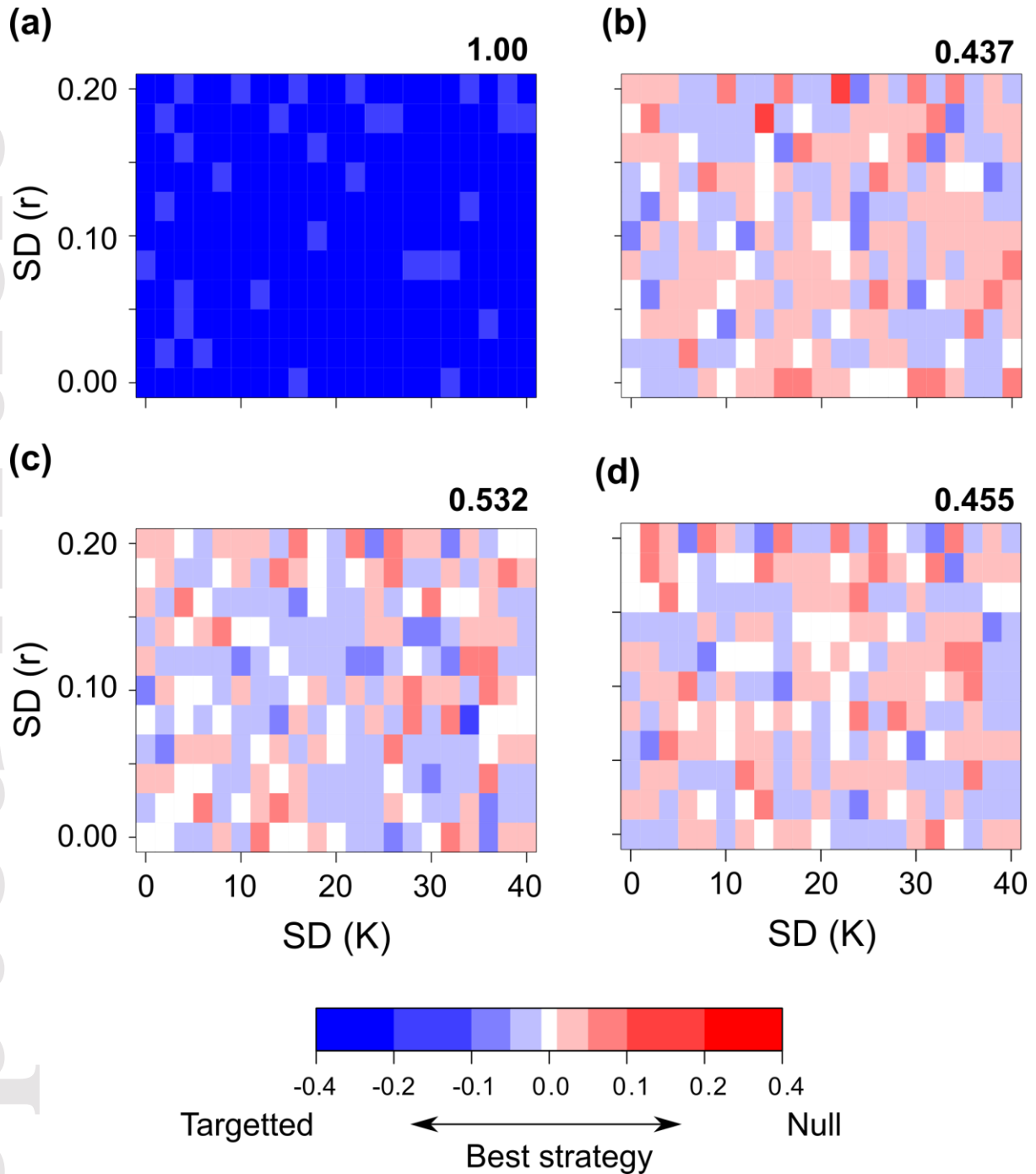


Figure 8 Effectiveness (median difference in proportion of habitats occupied between the given strategy and the null model) of prioritisation based on (a) connectivity, (b) highest local population growth (r_i), (c) highest local carrying capacity (K_i) and (d) highest r_i and K_i combined (summed z-score); Expt. 6 in Table 1. Negative values denote lower median habitat occupancy and so more effective control. Values above the plots are the proportion of the parameter combinations evaluated where each strategy out-performs the null strategy.

Table 1 Summary of the simulation experiments we conducted and the parameter ranges they encompassed. In all cases the graphs were constructed using the RNG algorithm (Toussaint 1980; Dale & Fortin 2014) based on an homogeneous Poisson point process (i.e., “complete spatial randomness” - Diggle 2003) comprising 100 nodes. The initially occupied habitats were randomly selected in all cases.

Graph		Initial Conditions			Incursion locations		Demography		Control
Experiment	Point pattern	No. of habitats (nodes)	Prop. habitats occupied	Initial pop. size (prop. K_i)		r	K		Start control
1. Same graph	CSR	100	0.1	0.25	Random	0.2	40		25
2. Multiple graphs with same macroscopic structure	CSR	100	0.1	0.25	Random	0.2	40		25
3a. Invasion phase	CSR	100	[0.05 ,1.0, 0.05]	[0.05 ,1.0, 0.05]	Random	0.2	40		25
3b. Invasion phase (early)	CSR	100	[0.02 ,0.2, 0.02]	[0.05 ,1.0, 0.05]	Random	0.2	40		25
4. Efficiency with no. of control sites	CSR	100	0.1	0.25	Random	0.2	40		25
5. Long-distance dispersal	CSR	100	0.1	0.25	Random	0.2	40		25
6. Demographic stochasticity	CSR	100	0.1	0.25	Random	N(0.2, [0.0, 0.2, 0.02])	N(40, [0.0, 40, 2.0])		25

Control		Dispersal			Replicates		No. of simulations	
Experiment	Strategy	Number of control sites (n_c)	Fraction controlled	Fraction of K for dispersal	Fraction of local pop that move	α	LDD	
1. Same graph	R, LD, BC	20	0.9	0.3	0.3	0.5	0	1000
2. Multiple graphs with same macroscopic structure	R, LD, BC	20	0.9	0.3	0.3	0.5	0	1
3a. Invasion phase	R, LD, BC	5, 20, 50	0.9	0.3	0.3	0.5	0	10
3b. Invasion phase (early)	R, LD, BC	5, 20, 50	0.9	0.3	0.3	0.5	0	10
4. Efficiency with no. of control sites	R, LD, BC	[0, 100, 1]	0.9	0.3	0.3	0.5	0	30
5. Long-distance dispersal	R, LD, BC	20	0.9	0.3	0.3	0.5	0, 0.0001, 0.0005, 0.001	10
6. Demographic stochasticity	R, LD, BC, DR, DK, DRK	20	0.9	0.3	0.3	0.5	0	10

Strategies: random (R), least distance (LD), best connected (BC), demographic by r (D), demographic by K (DK), demographic by r and K (DRK); [x, y, z] notation denotes parameter sweep from x to y by step z; N(m, s) denotes a draw from a log-normal deviate with mean m and standard deviation s