

A general spatially-explicit model to inform rapid response to new biological invasions: why do we need one and what should it look like?

M. Renton^{a,b,c}, **D. Savage**^{a,b} and **J. Chopard**^{a,b,c}

^a *School of Plant Biology and UWA Institute of Agriculture, The University of Western Australia*

^b *Cooperative Research Centre for National Plant Biosecurity, Australia*

^c *Centre of Excellence for Climate Change, Forest and Woodland Health, The University of Western Australia*

Email: michael.renton@uwa.edu.au

Abstract: New biological invasions can cause great damage to natural, agricultural and urban ecosystems, particularly if the invasive species is able to establish across an area that is too large and in numbers that are too great for eradication to be feasible. Important decisions need to be made about what management strategies should be carried out after the initial discovery and as the invasion unfolds. For example, should we attempt a costly eradication of this invasive species at all? If so, what level of resources should be allocated to this attempt, and how should they be allocated? These decisions need to be made quickly, decisively and transparently, despite the fact that only limited information about the invasive species is likely to be available. An efficient functional characterisation system can help extract and distil the most important information from all available expert knowledge, and a suitable model that captures the essential processes involved in the early stages of a new biological invasion can help synthesize this information to provide assistance in making quick, decisive and transparent decisions.

But what is the most important information that should be extracted and distilled from all available expert knowledge? And what are the most essential processes involved in the early stages of a new biological invasion that should be represented in a model of this invasion? In this paper we discuss a project with the CRC for National Plant Biosecurity that aims to develop an efficient system for functional characterisation of newly detected invasive organisms, for quick parameterisation of models of biological invasion. As part of the project, a general model of biological invasion (GMBI) has been developed. In this paper we discuss the motivation for developing such a model; an initial proposed structure for the model; and a number of potential uses for the model. We also discuss how the GMBI can be used to address the question of what really needs to be included in a spatially explicit model of new biological invasions.

Keywords: *biological invasion, rapid response, spatially-explicit model, management decisions*

1. INTRODUCTION

Rapid response to incursions of invasive organisms into new environments is essential. Invasive organisms, including pathogens, insects, weeds and vertebrates can all establish and spread quickly through natural, agricultural or urban landscapes, so delays in implementing management can prove costly. However, there is usually a range of possible management options, such as attempted eradication, attempted containment or ongoing management, and choosing the wrong management option can also be extremely costly. For example, attempting an expensive eradication of an invasive organism based on a strategy that proves to be unsuccessful represents an expensive error. On the other hand, neglecting a potential opportunity for eradicating an organism that then establishes and causes significant environmental and/or economic damage also represents a costly mistake.

Models provide a means of synthesising all our available knowledge to predict establishment, spread and damage of invading organisms, and can also be used to predict the results of different surveillance strategies and the effects of various management options (Kot *et al.* 1996, Hasting *et al.* 2005). They range from empirical, descriptive or 'phenomenological' models, such as those that attempt to predict the potential range of an invasive organism based on correlations between environmental factors and the probability of observed occurrences, to mechanistic, explanatory or 'ontological' models that attempt to simulate the dynamic ecological processes underlying invasion at various levels of detail and realism. Very specific models can be developed to represent the details of particular species of concern, or alternatively more abstract models can be developed that try to capture the essential aspects of biological invasion in a more general way. Similarly, the environment or landscape being invaded can be modelled at varying levels of detail and abstraction.

Models can be time-consuming and expensive to develop and analyse. This means that if a model is to be used to inform rapid response to a new incursion, usually the model must exist before the incursion is detected. When an incursion is first discovered, it is unlikely that there will be enough time to gather all required information about the organism and the newly invaded landscape, develop a completely new model to synthesise this information, and then carry out simulations and analyses to identify optimal response options. For high-risk organisms that are particularly likely to invade and to have large negative consequences, it may be worth developing organism-specific models. However, to develop organism-specific models for every single organisms that pose potential biosecurity threats is not possible, due to the number of such threats. On the other hand, a very abstract, empirical or simple model of biological spread may not be able to represent the important factors influencing the initial invasion, or may have parameters that are difficult to identify values for because they are not biologically meaningful (eg. Kot *et al.* 1996, Skellam *et al.*, 1951). So, in order to help inform rapid response to new incursions of a wide range of biological organisms, there is a need for a model of biological invasion that is relatively general but still includes the most important processes at a sufficient level of mechanistic realism.

We developed an alternative general model of biological invasion (GMBI) that is simple and abstract enough to represent a broad range of organisms and environments, but includes enough important processes with enough detail to enable a range of information about these processes to be synthesised. Our aim was to ensure that this model could be easily adapted to represent a very wide range of viral, bacterial or fungal pathogens, insects and other invertebrates, weeds, or vertebrates that threaten agricultural, natural or urban landscapes through simply changing parameter values, rather than having to develop and implement new model algorithms. In this paper we explain the management context in which the GMBI would be used, and argue that there are a number of requirements for the GMBI in order that it can successfully fulfil its role within this context. We then present the GMBI framework and discuss a number of ways in which its usefulness could be extended in future.

2. RAPID RESPONSE MANAGEMENT CONTEXT

The GMBI has been developed as a component of a proposed system for rapid response to new biological incursions (Figure 1). The system is brought into play when a new incursion of a high risk invasive organism is discovered and a rapid response is critical. We assume that it is not possible to conduct a full detailed survey to determine precisely the current spatial extent of the organism's incursion within the limited time period required for rapid response decision making, due to the difficulty of detecting the organism, the small number of expert surveyors available or other such constraints. Nor can we conduct the experiments needed to accurately quantify functional traits of the organism, such as fecundity, dispersal, or duration of life stages, which would be needed to accurately predict its spread. However, the decision of whether to attempt eradication or containment depends on an estimate of the spatial extent of the incursion at the present moment, and projected into the near future.

All available experts on the organism and the invaded landscape are gathered together or consulted remotely. A structured series of questions is used to characterise those aspects of the organism that are the most important determinants of its colonisation and early spread, including population dynamics and dispersal. This process results in identifying a set of GMBI parameter values that effectively characterise the organism for our purpose. This set may consist of a single value for each parameter, or a distribution of possible values if experts are unable to agree or unwilling to propose a single most likely value. A spatially-explicit description of the suitability of the landscape for this invading organism also needs to be obtained or generated at this point, as discussed below in Section 3.3.

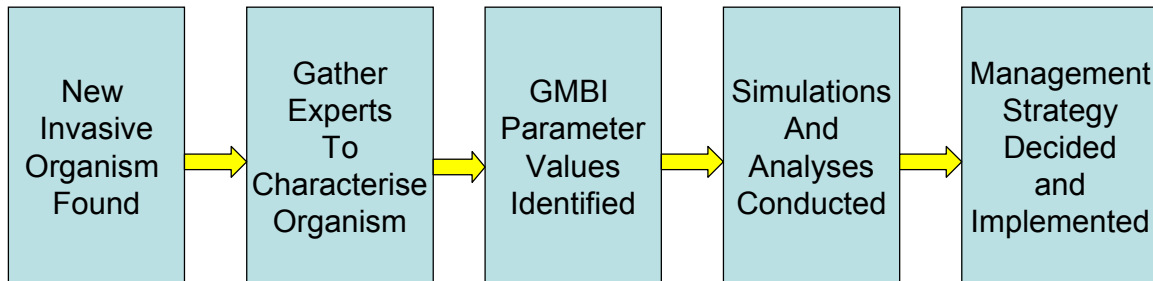


Figure 1. Proposed system for rapid response to new biological incursion within which the general model of biological invasion (GMBI) would be used.

A number of simulations are then conducted using the parameterised GMBI. Since the model is stochastic, the number must be great enough to account for the variability between individual model runs. It must also be great enough to adequately sample the range or distribution of parameter values specified, whenever single values were not obtained from the experts. While the simulations are being conducted, a number of management options are also formulated based on available expertise, with general guidelines for the criteria for deciding between them. These criteria will be largely based on an economic cost-benefit analysis. For example, it may be decided that if the organism is still contained within a certain distance, say 3 km of its original starting point, then the cost of an attempted eradication would be worth the long-term benefit of remaining free of the organism, but if it has spread beyond this distance, then it is not worth attempting eradication and containment is the best option. Or it may be decided that if the incursion has spread into a certain critical region, then eradication is no longer economically viable.

Analyses of the results of the simulations are then conducted to obtain a relevant probability on the model prediction of most relevance to the management decision or decisions to be chosen from. For the examples above, we would determine the probability that the organism is still contained within the threshold 3 km distance, or the probability that the organism has spread into the critical region. The probability of success of more cost-effective alternatives could also be calculated, so that all possibilities can be weighed. Based on these probabilities, the ‘best-guesses’ that we have in the situation, a decision is then made whether or not to pursue eradication or to instead implement a less expensive but less effective management strategy.

3. GMBI MODEL

3.1. Model requirements

As described above, the purpose of the GMBI is to help integrate and synthesise all available expert knowledge of a newly discovered invasive organism and the invaded landscape in order to make predictions of the likely spread of the organism that are as accurate as possible within the short time available to transparently inform critical management decisions. We argue that an empirical model would not be suitable, because an empirical model needs to be constructed from a large amount of directly relevant data, and within the management context described, this data would not be available. The spread of this particular organism in this particular landscape has never been measured before, because it is a new invasion. Moreover, sufficient data cannot be collected within the time available, because the organism is too difficult to detect, there is not sufficient resources, or there is not sufficient time. (If sufficient data could be quickly collected, the exact extent of current spread would be known, and decisions could be made based directly on this knowledge without need for simulation. However, a GMBI model would still be useful for predicting future spread under different management scenarios).

Experts on an invasive organism are likely to have some knowledge about the ecological and biological processes that affect its capacity to spread and invade, such as its fecundity, maturation rate, or dispersal

mechanism. They should thus be able to provide estimates for parameters concerning these kinds of processes. However, we also argue that a detailed and very realistic mechanistic or process-based model would not be suitable for our purpose. By definition, realistic models include many aspects of biological reality, and are thus tailored only to a particular species or class of organisms. A large number of detailed realistic models would thus be required to cover all possible types of invading organisms. These would take a long time to develop and validate, with high overheads for ongoing maintenance and organisation. Moreover, selecting the most relevant model for a newly discovered incursion would be time consuming, at a time when a rapid response is critical. Such a model would also contain algorithms representing a large number of interacting processes, making it computationally intensive, and thus the time required to run the model would be relatively long. Another possibility is to develop an extensive library of detailed and computationally optimised modules, each of which represents processes relevant to different classes of organisms and/or environments (Adeva *et al.*, 2011), and then select the relevant modules from the library when incursion happens, assemble them, and run simulations. But developing such a library to ensure coverage of all potential organisms and environments and processes would require a substantial initial investment in model building. Furthermore, any realistic mechanistic model would contain a large number of parameters, many of which it would be difficult and time consuming to obtain values for. The large number of parameters means that experts could disagree on many parameters, or find it difficult to estimate them all confidently, and so many could end up being specified with ranges or distributions of values rather than single estimates. Thus the number of model runs required to properly explore the high dimensional parameter space defining all possible invasion outcomes would be relatively high. Combined with the relatively long time for each run, this would make the total time required to conduct all required simulations very long. Since our primary goal is to inform rapid response to new incursions, this time requirement is not satisfactory.

We therefore argue that the ideal model is intermediate between these two extremes: a relatively abstract and simple model that includes mechanistic representation of only the key processes driving the early stages of a new biological invasion. Such a model is faster to run than a more detailed and complex simulation model. It has relatively few parameters, so that experts can focus their attention and work efficiently to provide the best possible estimates based on their knowledge of the organism. Only the most important parameters determining invasion are included, and these parameters are specified in a biologically-meaningful way that makes most sense to the experts likely to be parameterising the model. Fewer parameters with clear meaningful specification should help focus the attention of the experts and mean fewer parameters badly or imprecisely specified due to disagreement among experts or lack of confidence or understanding of what is required. Thus the amount of time required to specify parameter values and conduct required simulation runs would be relatively low compared to a more complex model, while still being tailored to the novel organism and landscape at hand. Ideally, despite the fact that the parameters of our ideal model are biologically meaningful, they also need to be general, so that for efficiency we can have just one GMBI, rather than having to work with a suite of models for different organisms.

The key decisions regarding management options are spatial because they depend on the probability that the invasive organism has or has not yet spread beyond threshold eradication or containment distances or into certain critical regions. The way that these probabilities depend on the organism's ecology and biology will also clearly depend on spatial characteristics of the particular landscape being invaded, such as the size, number and degree of connectivity of areas suitable for the organism to establish and reproduce. Various studies have shown that these spatial aspects are crucial (Hastings *et al.*, 2005, Lindström *et al.*, 2011, Nathan *et al.* 2000). Therefore, we also argue that the model needs to include a spatially explicit representation of the suitability of the landscape to invasion.

3.2. Overview of GMBI model structure

In the GMBI, space is divided into a grid or matrix of square cells, each of which has a specified suitability for the invading organism. The processes represented within the model can be divided into those that occur within a cell and the processes that involve movement between cells. The model can thus be conceptualised as shown in Figure 2. The within-cell processes include growth and maturation of individuals, reproduction and mortality. These processes may depend on the suitability of the cells, but not on the spatial relationship between locations. Dispersal or movement occurs between locations and depends directly on the spatial relationship between locations. Dispersal affects populations within locations by adding or removing individuals from locations, and population dynamics within locations affects dispersal, by affecting the number of individuals dispersing, or the timing of dispersal due to triggers dependent on maturation or population density. For the sake of simplicity, we can assume that the spatial pattern of suitability is not affected by the population dynamics or dispersal.

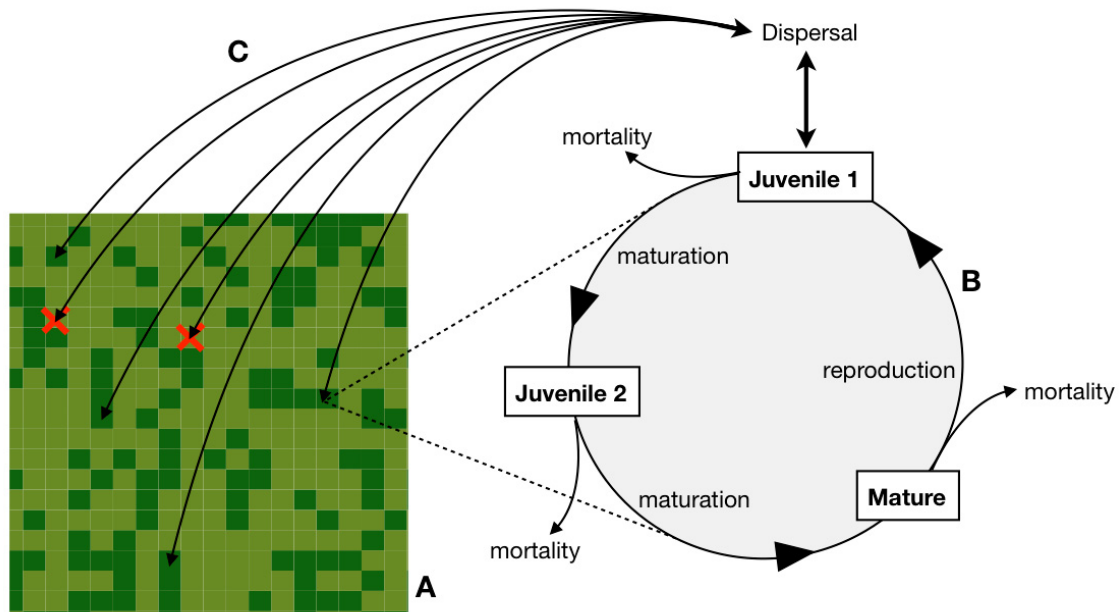


Figure 2. Overview of GMBI structure, showing three main components of the model, a spatially-explicit representation of landscape suitability (A), within-cell population dynamics (B), dispersal (C), and the relationship between them.

3.3. Landscape

The GMBI requires a spatially-explicit representation of landscape suitability as specific as possible to the invaded landscape and the invading organism. This may be obtained in various ways, all of which will need input from our collected experts on what kind of environments will suit the organism. If the organism is restricted to a particular well-characterised environment (a fungal pathogen that requires a particular host crop species, for example) and GIS data on this environment is available, then this can be used to generate the spatially-explicit description of suitability. If this is not available, then aerial or satellite images may be used to generate the spatially-explicit description of suitability, based on best available knowledge of where suitable environments, hosts and/or conditions occur within the image. Alternatively, a more abstract representation may be used, based on approximations of the percentage of the landscape likely to be suitable, and characteristics such as the mean and variability in size of suitable blocks, or the degree of aggregations versus evenness in the spatial distribution of suitable cells (Figure 3). The experts can choose from a library of these, or a new one can be generated based on their knowledge.



Figure 3. Examples of abstract landscapes, based on approximations of the percentage of the landscape likely to be suitable, and characteristics such as the mean and variability in size of suitable blocks, or the degree of aggregations versus evenness in the spatial distribution of suitable cells. Each of the landscapes shown has 30% suitable habitat (dark green), but using different methods of generation leads to differences in the distribution of these habitats.

3.4. Population dynamics

Population dynamics within the GMBI consist of reproduction, maturation and mortality (Figure 2). The life cycle of the organism being modelled is broken up into a series of discrete stages and for each cell in the landscape the model tracks the number of individuals within each stage. The model requires at least two stages to be defined, representing juvenile and mature individuals respectively. However, depending on the organism being modelled, multiple juvenile stages may be included, and in each time step, individuals move from one stage to the next. Once an individual reaches the final stage before maturity they remain within this stage until they mature, which may occur in each subsequent time step with a specified probability p_m . In each time step, a specified proportion of individuals in each stage is assumed to die and is removed from the simulation. Mature individuals are assumed to reproduce at a specified reproduction rate, which is defined according to expert knowledge, and the resulting cohort of new individuals is added to the first stage in the life cycle. Allee effects are included simply as a minimum population density for reproduction to occur (Keitt *et al.*, 2001) and carrying capacity as a maximum population density that cannot be exceeded.

3.5. Dispersal

Dispersal can be set to occur within any of the life stages used to represent population dynamics, and can occur multiple times over an individual's life time. For each life stage associated with dispersal, all individuals are assumed to undergo dispersal, and all individuals are assumed to be equally likely to travel over a particular distance or in a particular direction, with the distance travelled assumed to be randomly distributed according to a Weibull distribution. The shape and scale parameters for the Weibull distribution are calculated from the 50th and 99th percentiles of distances travelled for the organism in question, with these parameters estimated using expert knowledge. The direction of travel is drawn from either a uniform distribution or a von-Mises distribution if experts believe dispersal in particular directions is more likely. The destination cell is then calculated, and, if this cell represents an area of suitable habitat, the individual is added to the appropriate life stage in this cell. If the destination cell is unsuitable, the individual dies, and is removed from the simulation, or, to represent more active dispersal, it moves to the nearest suitable cell within a certain threshold distance and only dies if no such cell exists. Individuals that leave the area being simulated are assumed to have escaped the region under surveillance and would usually indicate a situation where eradication would no longer be considered feasible.

3.6. Example outputs

The GMBI produces various outputs, both visual and numerical. Visual outputs include spatially explicit representations of simulated invasion from either particular model runs, or to summarise results of a large number of simulations, accounting for model stochasticity (Figure 4). Numerical outputs include total population numbers within the whole landscape or a specified sub-section, maximum distance of invasion, or total area occupied, for each time step.

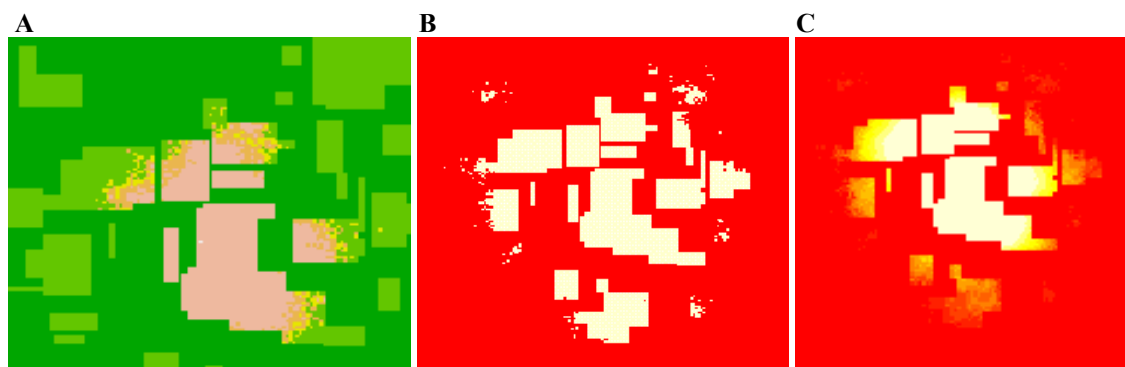


Figure 4. Example model outputs, including population densities from a particular model run after a certain time (A), the union of all cells invaded to any density after the same time in any of a large number of model runs (B), the expected density of each cell after the same time, based on the same set of model runs (C).

4. DISCUSSION

The next step for the GMBI is to test and then demonstrate its generality, or its ability to represent a diverse range of types of invasive organisms. For example, we aim to show that the model can be parameterised to

represent invasive fungal pathogens, insects, vertebrates and weeds without needing any further adaptation to underlying model algorithms. This will show that the model is ready to be deployed in the specified role within the rapid response management context described in Section 2.

We also argue that the GMBI has important roles to play beyond this direct management application. Because of its generality and relative simplicity, the GMBI as described can act as a base reference point for theoretical investigations into what should and should not be included within such a model, and what kinds of questions the experts should focus on during the model parameterisation procedure described in Section 2. We will first choose sets of parameter values to represent a range of the types of organisms that we would like to apply the model to, such as fungal pathogens, insects, vertebrates and weeds. We will then conduct an extensive sensitivity analysis on all parameters for each of these sets, in order to determine which parameters have the greatest effect on simulation results. We can then conclude that these are the parameters to focus on most intensely during discussions with experts. Parameters that are relatively less important will need less attention, and could even be dropped from the GMBI completely. Similarly, we can investigate whether processes that experts believe are missing from the GMBI should be included in future, based on whether including them makes much difference to simulation results. In the same way, we could test whether we may actually require two or three versions of the GMBI to represent the full diversity of possible invasive organisms, and thus move a little towards the idea of a library. The GMBI could also be used to help prioritise costly research effort, directing it towards first quantifying the most important parameters in the model for species of particular concern.

The usefulness of the GMBI will also be increased by extending it to include surveillance, detection and management options. This will allow these processes to be included in analyses like those discussed above, and also to directly test the usefulness of different surveillance and management strategies. This will further improve the reliability of decisions between eradication, containment and ongoing management.

The usefulness of the GMBI also has the potential to be further improved through meta-modelling. A very large number of runs of the model will be conducted, across many combinations of model parameter values. Important outputs of the model will be recorded across time for each run. This can then be used as data to construct empirical models relating GMBI parameters to outputs, using relatively simple functional forms rather than simulation. These empirical models can then act as meta-models or emulators of the GMBI, giving similar output more transparently and much more quickly. These will be particularly suited for the role within the rapid response management context, since they will provide estimates of expected values and likely variability for key predictions almost instantly, unlike the mechanistic GMBI itself, which would require many runs to give similar results. They will also be suited to adaptive management approaches such as testing whether parameter values actually describe the invasion and quickly updating these values if necessary, as real data on species occurrences across the landscape becomes available.

ACKNOWLEDGMENTS

The authors acknowledge the support of the Cooperative Research Centre for National Plant Biosecurity, established and supported under the Australian Government's Cooperative Research Centres Programme.

REFERENCES

- Adeva, J.J.G., Sousa-Majer, M.J.D., Botha, J.H., Hardie, D.C., Reynolds, M. (2011). A simulation technology approach to *Bactrocera spp.*: lessons from a past incursion for improving future responses. *Proceedings of the National Plant Biosecurity CRC Science Exchange 2011*.
- Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, 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.
- Keitt, T.H., Lewis, M.A., Holt, R.D. 2001. Allee effects, invasion pinning, and species' borders. *The American Naturalist* 157, 203-216.
- Kot, M., Lewis, M.A., van den Driessche, P. (1996). Dispersal data and the spread of invading organisms. *Ecology*, 77, 2027.
- Lindström, T., Håkansson, N., Wennergren, U. (2011). The shape of the spatial kernel and its implications for biological invasions in patchy environments. *Proceedings of the Royal Society B: Biological Sciences*, 278,1564 -1571.
- Nathan, R., Muller-Landau, H.C. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology & Evolution*, 15, 278-285.
- Skellam, J.G. (1951). Random dispersal in theoretical populations. *Biometrika*, 38, 196-218.