

ABSTRACT

THOMSON, NIKKALA ALISON. Pair-edge Approximation for Heterogeneous Lattice Population Models. (Under the direction of Stephen P. Ellner.)

To increase the analytical tractability of lattice stochastic spatial population models, several approximations have been developed. The pair-edge approximation is a moment-closure method that is effective in predicting persistence criteria and invasion speeds on a homogeneous lattice. Here the effectiveness of the pair-edge approximation is evaluated on a spatially heterogeneous lattice in which some sites are unoccupiable, or ‘dead’. This model has several possible interpretations, including a spatial SIS epidemic model, in which immobile host-species individuals occupy some sites while others are empty. As in the homogeneous model, the pair-edge approximation is found to be significantly more accurate than the ordinary pair approximation in determining conditions for persistence. However, habitat heterogeneity decreases invasion speed more than is predicted by the pair-edge approximation, and the discrepancy increases with greater clustering of dead sites. The accuracy of the approximation validates the underlying heuristic picture of population spread and therefore provides qualitative insight into the dynamics of lattice models. Conversely, the situations where the approximation is less accurate reveal limitations of pair approximation in the presence of spatial heterogeneity.

**Pair-edge Approximation
for Heterogeneous Lattice Population Models**

by

Nikkala Alison Thomson

A thesis submitted to the Graduate Faculty of

North Carolina State University

in partial fulfillment of the

requirements for the Degree of

Master of Science

BIOMATHEMATICS

Raleigh, NC

2002

Approved by:

Stephen P. Ellner, Chair of Advisory Committee

Marcia L. Gumpertz

John E. Franke

DEDICATION

To my husband,

Darren,

who always believed in me...

to my father,

Deloy,

who inspired me to pursue a higher education...

and to my children,

Alison and Brent,

who make every moment precious.

BIOGRAPHY

Nikkala Alison Thomson was born in Philadelphia, Pennsylvania, to M. Deloy and Alison Pack. She has four younger siblings: Benjamin, Jonathan, Roger, and Carolyn. In high school, she enjoyed science, mathematics, English, and music. There she also began doing independent research involving numerical analysis of density-dependent natural selection, which she presented at the 1990 Intermountain Junior Science and Humanities Symposium. When she entered Brigham Young University, Provo, Utah, in 1990, she continued her research under the guidance of Dr. Brian Maurer. From 1993 to 1994, Nikkala took a hiatus from her studies and spent a year and a half in Hong Kong, as a missionary for the Church of Jesus Christ of Latter-day Saints. In August 1996, she graduated magna cum laude from BYU with a Bachelor of Science degree in Honors Zoology. Her honors thesis was titled “Scale-sensitive Linear Analysis of Granivorous Rodent Population Dynamics.” The same month, Nikkala entered the Biomathematics Graduate Program at North Carolina State University in Raleigh, North Carolina, with a three-year National Science Foundation graduate research fellowship. She has been working under the direction of Dr. Stephen P. Ellner, with a focus on mathematical modeling in ecology. Nikkala married Darren Thomson in 1998. They have two children, Alison and Brent.

ACKNOWLEDGMENTS

I wish to acknowledge the unfailing support and encouragement of my graduate research advisor, Dr. Stephen P. Ellner. His guidance and patience have been truly outstanding. I would also like to thank Dr. Marcia Gumpertz and Dr. John Franke for their contributions as members of my advisory committee. My graduate experience was further enriched by the many other friendships and associations I made at North Carolina State University. Ann Ethridge, the Biomathematics Program administrative assistant, was a constant source of assistance and cheer.

Special thanks belong to Dr. Brian Maurer, my undergraduate research advisor. His enthusiasm for his work was contagious, and I appreciated the freedom he gave me to develop my own creative ideas. Many of my instructors from Brigham Young University, and several from high school, are responsible for igniting my interest in ecology and my desire to learn.

My father also contributed greatly to my decision to go to graduate school, by his example and his continuing search for knowledge. Both of my parents have been a constant source of inspiration to me, and I am deeply indebted to them for the values and lessons they have taught me through the years. Finally, I am grateful for the love and support of my husband Darren, who has made many sacrifices to help me achieve my goals.

TABLE OF CONTENTS

LIST OF FIGURES	vi
1. INTRODUCTION.....	1
2. METHODS.....	3
2.1 Simulation.....	3
2.2 Critical Birth Rate.....	5
2.3 Invasion Speed.....	6
2.4 Ordinary Pair Approximation (OPA)	7
2.5 Pair-edge Approximation	8
3. RESULTS.....	9
3.1 Critical Birth Rate.....	9
3.2 Invasion Speed.....	10
4. DISCUSSION	11
5. REFERENCES.....	14
6. APPENDICES.....	17
6.1 Derivation of ordinary pair approximation equations	18
6.2 Derivation of pair-edge approximation for invasion speed	20
6.2.1 Quasi-equilibria.....	20
6.2.2 Wave or invasion speed	20
6.3 Approximating $E[\delta]$	22

LIST OF FIGURES

Figure 1.	Heterogeneous lattices with different degrees of clustering.....	24
Figure 2.	An example of the invasion process on a heterogeneous lattice of height three.....	25
Figure 3.	Comparison of critical birth rates from simulation and approximation as they vary with the degree of clustering of the dead sites in the lattice....	26
Figure 4.	Comparison of invasion speeds derived from simulation and the pair-edge approximation.....	27
Figure 5.	Comparison of lattice composition, measured by the fraction of occupable sites that are occupied.....	28

1. INTRODUCTION

Lattice models, a form of discrete stochastic spatial population models, have been popular among theoretical biologists because they provide a simple, flexible framework for modeling spatial dynamics that arise from events at the individual level. Also known as stochastic cellular automata (Hogeweg 1988, Caswell and Etter 1993), interacting particle systems (Durrett and Levin 1994b) or artificial ecologies (Rand 1999), these models keep track of the presence or absence of individuals (of one or more types) in each cell of a regular spatial lattice.

The individuals in a lattice model could be plants (e.g., Lavorel et al. 1994; Harada and Iwasa 1994), the presence of a stage of an infectious disease (e.g., Levin and Durrett 1996; Rhodes and Anderson 1996abcd, 1997; Rhodes et al. 1997), a pair of territorial birds, or any other spatially localized biological entity (e.g., predator/prey in McCauley et al. 1993; host/parasitoid in Pacala et al. 1990 and Hassell et al. 1991). The primary assumption is that a maximum of one entity can occupy each cell.

Lattice models are individual-based, as births and deaths occur at the individual level; they thus provide a way to study the connection between individual interactions and larger-scale population dynamics such as sustainability and spatial spread. Reaction-diffusion models and other models that neglect individual interactions may be less effective in predicting the behavior of a spatial system (Durrett and Levin 1994a).

In a lattice model, two quantities often of interest are the conditions for population persistence the rate of population spread into unoccupied habitat ("invasion speed"). These, and any other quantities of interest, can be directly calculated by simulation. However, simulation has several limitations. First, it may require a detailed knowledge of the initial

state of the lattice. Second, simulation on a large lattice can be computationally intensive and time-consuming. Third, simulation results are difficult to analyze rigorously. For example, it would be difficult to construct bifurcation or stability diagrams. For these reasons, the utility of lattice models is greatly increased by the availability of various analytic approximations.

Several moment-closure methods of approximation have been developed and applied to estimate the dynamics of lattice models. The simplest of these is the mean field approximation (Dickman 1986, 1988; Durrett and Levin 1994a). Based on the assumption that adjacent sites are independent, the mean field approximation incorporates only site-type frequencies (e.g., the fraction of occupied sites in the lattice). Because the mean field approximation ignores spatial structure, it often fails to accurately capture the dynamics of a lattice model (e.g., Kubo et al. 1996).

The ordinary pair approximation (OPA, Matsuda et al. 1992) goes a step further by also taking into account the pair-type frequencies (e.g., the fraction of neighboring site pairs in which both sites are occupied). It ignores correlations other than those between nearest neighbors. The OPA has been shown to be effective when applied to the estimation of the general dynamics of various lattice models (e.g. Harada et al. 1995; Harada and Iwasa 1994; Hiebeler 1997; Nakamura et al. 1997; Satō et al. 1994; van Baalen and Rand 1998; Kubo et al. 1996). Higher-order moment-closure approximations (e.g., Hiebeler 1997) take into account more detailed spatial structure than the OPA, but they require more detailed knowledge of the lattice and are less analytically tractable.

To estimate invasion speeds, the OPA has been extended to the pair-edge approximation (Ellner et al. 1998). The most accurate form of this approximation is based on the assumption that the area behind the invasion front is at OPA equilibrium, while the

leading edge of the front is at a quasi-steady state between the steady state behind the front and the unoccupied state beyond the front (Ellner et al. 1998). On homogeneous lattices, invasion speeds and critical birth rates for population persistence derived from pair-edge approximation agree well with simulation results and are considerably more accurate than OPA (Ellner et al. 1998).

However, the pair-edge approximation has been applied only to spatially homogeneous lattice models. Spatial heterogeneity can be incorporated into lattice models in several different ways, the simplest of which is to include ‘dead’ sites, which are lattice cells in which the organism cannot exist (e.g. Bascompte and Solé 1996). When such landscape heterogeneity is incorporated into a lattice model, both the quantity and the spatial arrangement of the dead sites is important. For a given fraction of the dead sites, there can be great variation in the degree of clustering (Figure 1). This clustering has the potential to dramatically affect the dynamics of the organisms on the lattice, depending on the rules governing the system (e.g. Lavorel et al. 1994). Hiebeler (2000) extended the OPA to heterogeneous landscapes. The purpose of this paper is to extend the pair-edge approximation to heterogeneous lattice models, and evaluate its effectiveness.

2. METHODS

2.1 Simulation

The basic contact process (Durrett 1999) is the foundation of the lattice model used in this paper. In the lattice model, each cell can be unoccupied (0), occupied (1), or dead (2). Dead sites cannot be occupied, and they remain fixed in location. Occupied and unoccupied sites, however, can change in state, as a result of birth and death processes. Each occupied

cell ‘gives birth’ into neighboring cells at rate b , and each cell becomes unoccupied (‘dies’) at rate d . Time was scaled so that the death rate d was standardized to 1. When a death occurs in an occupied cell, it becomes unoccupied. When a birth occurs from an occupied cell, the selection of which neighboring cell receives the new organism is random. If the selected cell is dead or occupied, no change occurs. If the selected cell is unoccupied, it becomes occupied.

There are several ways to define a ‘neighbor’ in a lattice model. Two of the most common are the four-neighbor and eight-neighbor definitions. In the four-neighbor case, cells to the right, left, up, and down of a given cell are considered to be its neighbors. In the eight neighbor case, cells diagonal to the given cell (e.g., to the right and up) are also considered neighbors. Although the four-neighbor case is simpler in concept, and easier to calculate, it is less spatially realistic, and sometimes leads to complications on a heterogeneous lattice. For example, if the dead sites on a heterogeneous lattice were arranged in a checkerboard pattern, with the four-neighbor definition, no births could occur, and any population would eventually die out, no matter how high the birth rate. For this reason, the eight-neighbor definition was used.

For simulations the proportion of dead sites was chosen to be $\rho_2 = 0.3$, indicating that three out of ten sites are dead. The degree of clustering was varied from $\rho_{22} = 0.09$ to $\rho_{22} = 0.25$, where ρ_{22} is the proportion of dead sites next to dead sites out of all neighboring site pairs. $\rho_{22} = 0.09$ results from random dead site placement, while $\rho_{22} = 0.25$ has a high degree of clustering (0.3 is the maximum possible value of ρ_{22} , as ρ_{22} cannot exceed ρ_2). Lattices with different degrees of clustering were created by first generating a lattice with random placement of dead sites, then randomly choosing two sites in the lattice and switching their

contents if that would bring the degree of clustering closer to the desired value. See Figure 1 for examples of lattices with differing degrees of clustering. In measuring the degree of clustering during lattice generation, pair probabilities were the only thing taken into account; higher order structure, such as triplet probabilities, was not incorporated.

Once the appropriate lattices were generated for each simulation, the simulation was performed in continuous time by implementing "exponential alarm clock" construction of the process. Each site was viewed as "ringing" independently (signaling a possible event) at stochastic rate $b+d$ (i.e. the time between rings is exponentially distributed with mean $(b+d)^{-1}$). The lattice as a whole thus rings at rate (*number of cells*) $\cdot(b+d)$, and the average time between rings is the inverse of this rate. However, when a very large number of rings is being simulated, the average time between rings is all that is necessary to calculate long-term average rates. Consequently, it is not necessary to actually simulate the time between each ring.

When a ring occurs, a uniform random variable from 0 to 1 is used to determine whether the event was a birth or a death. If the random variable is greater than $b/(b+d)$, the event is a death and the site (if occupied) becomes unoccupied. Otherwise the event is a possible birth, and a birth ($0 \rightarrow 1$ transition) occurs if the random variate is below the actual birth rate at the particular site (b times the fraction of occupied neighbors).

2.2 Critical Birth Rate

The critical birth rate is the minimum birth rate required for population persistence. Its value was determined by simulating 200 x 200-cell lattices with varying degrees of clustering. Wrap-around edges were assumed in order to approximate an infinite lattice. Each lattice was initially set so that each occupiable site was occupied. This reduced the

probability that the population would go extinct by chance, even though the birth rate was high enough to ensure persistence on a truly infinite lattice.

For each lattice, an initial birth rate was chosen, and one billion events were simulated with that birth rate. If the population was not extinct at that time, the birth rate was assumed to be greater than the critical birth rate. A new birth rate, presumably closer to the critical birth rate, was then chosen and simulated, and so on, until the critical birth rate was estimated to two decimal places.

2.3 Invasion Speed

Simulations to determine invasion speed required lattices with slightly different characteristics than those used in determining the critical birth rate. Assuming the invasion was to move from left to right on the lattice, only the top and bottom edges were wrapped around. The presence of dead sites, especially when highly clustered, could by chance totally obstruct invasion on a finite lattice or otherwise artificially reduce the invasion speed compared to a truly infinite lattice, so very tall lattices (2000 rows x 500 columns) were used to minimize the probability of this occurrence.

All occupiable sites in the leftmost 125 columns of the lattice were initialized as occupied; the rest of the occupiable sites were left unoccupied. Two hundred million events were then simulated, to give time for equilibrium to become established behind the invasion front and thus create a more realistic initial condition. The new initial location of the forerunner(s) was determined, and 1.8 billion more events were simulated. The large number of simulated events was necessary because of the large size of the lattices. In each simulation, the motion of the wave forerunner was tracked, and the wave speed was calculated by the equation:

$$\frac{\text{net distance traveled * rate of events}}{\text{number of events}} \quad (2.3.1)$$

where *net distance traveled* is the distance traveled by the wave, and *rate of events* is as mentioned above, $(\text{number of cells}) * (b+d)$.

2.4 Ordinary Pair Approximation (OPA)

The ordinary pair approximation incorporates both global densities and nearest-neighbor correlations, but it ignores correlations at further distances. This does not mean that non-neighboring sites are independent; rather, it assumes that the relationship between those sites can be approximated by the product of nearest-neighbor correlations on a pathway between them (Nakamura et al. 1997). The general pair approximation assumption is

$$q_{j/k} \approx q_{i/j} \quad (2.4.1)$$

where $q_{j/k}$ is the probability that a randomly selected neighbor of a state- j site with state k neighbor is in state i ; similarly, $q_{i/j}$ is the probability that a randomly selected neighbor of a state- j site is in state i .

See Appendix 6.1 for the derivation of the three differential equations governing the OPA on a heterogeneous lattice, for the eight-neighbor case. These equations are as follows:

$$\frac{d\rho_1}{dt} = b(\rho_1 - \rho_{11} - \rho_{12}) - \rho_1 \quad (2.4.2)$$

$$\frac{d\rho_{11}}{dt} = \frac{b}{4}(\rho_1 - \rho_{11} - \rho_{12}) \left(1 + 7 \frac{\rho_1 - \rho_{11} - \rho_{12}}{1 - \rho_1 - \rho_2} \right) - 2\rho_1 \quad (2.4.3)$$

$$\frac{d\rho_{12}}{dt} = \frac{7b}{8}(\rho_1 - \rho_{11} - \rho_{12}) \left(\frac{\rho_2 - \rho_{12} - \rho_{22}}{1 - \rho_1 - \rho_2} \right) - \rho_{12} \quad (2.4.4)$$

Here, ρ_i is the fraction of sites in state i , and ρ_{ij} is the fraction of neighboring site pairs in which the first member of the pair is in state i and the second member is in state j .

OPA steady state values for ρ_1 , ρ_{11} , and ρ_{12} as a function of (b, ρ_2, ρ_{22}) can be obtained by setting the right-hand side of equations (2.4.2), (2.4.3), and (2.4.4) equal to zero and solving them simultaneously. Recall that ρ_2 and ρ_{22} are parameters describing the frequency and pattern of dead sites in the lattice, which are constant over time. Given the steady state values of ρ_1 , ρ_{11} , and ρ_{12} , the OPA of the critical birth rate can be calculated from equation (2.4.2) by solving $b(\rho_1 - \rho_{11} - \rho_{12}) - \rho_1 = 0$ for b .

2.5 Pair-edge Approximation

The pair-edge approximation is similar to the ordinary pair approximation in that it incorporates both local densities and nearest-neighbor correlations, and ignores correlations at further distances. See Figure 2 for a depiction of the invasion process on a heterogeneous lattice. The wave front advances when a front runner ‘gives birth’ into a cell in an unoccupied column in the direction the wave is travelling. The wave front retreats when a lone front runner ‘dies,’ and the occupied cells in the nearest occupied column then become the new front runners.

The pair-edge approximation assumes that in all columns behind the front runner, the lattice is at equilibrium, which can be approximated by the OPA. The column that includes the front runner is assumed to be at a quasi-equilibrium that is an intermediate state between the OPA equilibrium behind the front and the unoccupied state ahead of the front. This quasi-equilibrium at the invasion boundary is approximated by fixing ρ_1 at half of its OPA steady-state value, and calculating the other quasi-equilibrium values as you would equilibrium values for the OPA. See Appendix 6.2 for the full derivation of these quasi-equilibrium equations on a heterogeneous lattice, for the eight-neighbor case.

The pair-edge approximation also assumes that the wave speed in an infinite two-dimensional lattice can be well approximated by the wave speed in a lattice with only three rows which wrap around from top to bottom; simulations (Ellner et al. 1998) support this approximation. The rate of forward and backward moves by the wave front is calculated from the approximated quasi-equilibrium values at the wave front. After the backward steps are adjusted for the distribution of backward move distances, the net invasion speed can be estimated (see Appendix 6.2). The process of adjusting for the backward move distance in a heterogeneous lattice is described in Appendix 6.3. The critical birth rate is estimated by finding the value of b for which the pair-edge approximation of the invasion speed is equal to zero.

3. RESULTS

3.1 Critical Birth Rate

The critical birth rates from the simulations, the OPA, and the pair-edge approximation are compared in Figure 3. Recall that the critical birth rate in the pair-edge approximation was calculated as the birth rate at which the wave speed would equal zero. The pair-edge approximation was considerably better at predicting the critical birth rate than either the OPA or the mean field approximation, for all tested degrees of dead site clustering.

It is likely that the pair-edge approximation tends to under-predict the critical birth rate because it is based on OPA. OPA is an optimistic approximation that tends to overestimate ρ_1 . The assumption that the lattice is at OPA equilibrium behind the front results in an overestimation of the proportion of occupied sites behind the front, and thus an overestimation of the proportion of occupied sites at the front. This produces an

overestimation of the rate of forward moves (which result from births by front-row individuals), and thus underestimates the critical birth rate. Further factors that may affect the accuracy of the pair-edge approximation are discussed in Section 3.2, Invasion Speed.

Despite its slightly optimistic bias, the pair-edge approximation predicts well the general effect of clustering on the critical birth rate. The mean field approximation predicts no change in critical birth rate with degree of clustering, which is obviously in error. The OPA gets the sign and general magnitude of the change correct, but it is considerably more optimistic than the pair-edge approximation, and it fails to capture the slight curve in the simulation data. The pair-edge approximation is not only the most accurate of the compared approximations, but it is also the most consistent and does capture at least a portion of the simulation curve.

3.2 Invasion Speed

The invasion speeds from the simulations and the pair-edge approximation are compared in Figure 4. The heterogeneous pair-edge approximation predicts the wave speed very well for low degrees of clustering such as $\rho_{22} = 0.09$. However, as the degree of clustering increases, so does the level of inaccuracy, especially at high birth rates. The homogeneous (no dead sites) values are given for comparison.

As with the critical birth rate (Section 3.1), the primary suspect for the overprediction of the invasion speed is the optimism of OPA. As is evident in Figure 5, no matter whether the birth rate or degree of clustering are large or small, in simulations the fraction of occupiable sites that are occupied in the rows behind the front is less than 0.5 on average. However, for high clustering and high birth rate, the OPA estimate of this quantity can be as

high as 0.81. When this optimism about the state of the lattice behind the wave front is combined with high birth rates, the estimated wave speed also becomes overly optimistic.

In a homogeneous lattice, similar comparisons indicated that the OPA was a very accurate approximation for lattice sites behind the wave front (Ellner et al. 1998). One possible reason this is not the case with for a heterogeneous lattice is that the heterogeneity, especially when highly clustered, creates a more irregular, fingered wave front. The cells just behind the front runner are thus not typical of the steady-state that holds well behind the front.

Interestingly, the state of the lattice ahead of the front runner also proved to be unusual. One assumption of the pair-edge approximation is that the front runner is in a ‘typical’ site in the lattice, in relation to the dead sites. In reality, no matter what the degree of clustering, in the simulations performed the proportion of dead sites ahead of the front runner was consistently higher than that of a ‘typical’ site in the lattice. This may be because the front runner often gets ‘stuck’ behind one or more dead sites and the wave front must spread around the obstacle before continuing to progress.

4. DISCUSSION

The model is simple enough that critical birth rates and invasion speeds could be obtained by simulation. The ‘added value’ from this analysis is that the underlying heuristic picture of population growth and spread in a sharply defined, linearly advancing wave front is validated for situations where the approximation remains accurate despite the addition of spatial heterogeneity: unclustered heterogeneity, or low to moderate clustering with lower birth rates. Conversely, the circumstances and ways in which the approximation loses

accuracy indicate the limitations of pair approximation in the presence of spatial heterogeneity with clustering. The optimism of the approximation appears to become more severe, and larger-scale structures that pair approximation does not consider increase in importance – in particular, the large-scale irregularity of the wave front, and the fact that the leading edge of the invasion wave tends to occur in atypically unfavorable portions of the lattice.

It can be concluded, then, that habitat heterogeneity has a considerable effect on invasion dynamics, even if highly clustered. Theoretically, if heterogeneity is completely (100%) clustered, it should have little effect on invasion dynamics in a truly infinite spatial domain. Therefore, it may be tempting to neglect heterogeneity that is highly clustered relative to the scale of the organism of interest. However, this study has shown that that would be unwise. The factors identified as limiting the accuracy of the pair-edge approximation are in fact increased when habitat heterogeneity is highly clustered, and would need to be taken into account in any successful model.

As a possible interpretation of the heterogeneous lattice model, consider a pathogen spreading rapidly through an area where some sites are occupied by sessile host-species individuals, while other sites are vacant. For example, the newly-emerged fungal pathogen *Aspergillus sydowii* has recently spread through sea fan corals in the Caribbean, causing significant mortality in many local populations (Kim and Harvell 2001). The model in this paper is only applicable when the time scale of the disease is faster than that of population change—in particular, if it is assumed that infection does not cause any added mortality and that recovered hosts do not acquire any immunity. In this context the model in this paper is

just one very special case, but it is a useful step towards studying more general models for rapid spread of new pathogens attacking sessile organisms.

The corresponding homogeneous lattice models (in which all sites are occupied by a sessile host individual) have recently been studied using pair approximation (Satō et al. 1994, Levin and Durrett 1996, Filipe and Gibson 1998). Lattice models for the spread of plant disease within a homogeneous field (e.g. Kleczkowski et al. 1997, Bailey and Gilligan 1997, 1999), parameterized from individual-level data, have successfully predicted the mean time-course of infection and the level of variation between replicates. Rhodes and Anderson (1996a,b,c,d, 1997) have considered more biologically detailed lattice models for measles outbreaks in small communities which were able to reproduce features of the data that could not be explained by non-spatial models.

There are, of course, many other possible applications of the heterogeneous lattice model. It can be applied to the invasion not only of disease, but also of exotic species including plants and territorial organisms such as birds. If critical birth rate is the quantity of interest, the model and associated approximations can be applied to organisms that are not necessarily invading, but are being subject to varying degrees or forms of habitat heterogeneity.

In summary, the pair-edge approximation can give considerable insight into a heterogeneous lattice model without requiring simulation or exact knowledge of the initial state of the lattice. The approximation is strongest when the degree of habitat clustering and birth rates are low to moderate, and more limited if clustering or birth rates are high, but it does provide analytic tractability, which simulation does not. The fact that the approximation varies in effectiveness with degree of clustering and birth rate is itself

revealing; the nature of the actual invasion dynamics must also be changing as these parameter are modified.

This research could be extended by incorporating higher-order moments into the approximation, although that would require a more extensive knowledge of the initial state of the lattice than the pair-edge approximation. The limitations of the pair-edge approximation could also be further addressed, and possibly remedied, if a more complex model were developed. However, an increase in complexity may decrease analytic tractability. The pair-edge approximation, as it is based on the simple elegance of the pair approximation, provides a great deal of insight into invasion dynamics for the limited information it requires.

5. REFERENCES

- Bailey, D.J. and C. A. Gilligan. (1999). Dynamics of primary and secondary infection in take-all epidemics. *Phytopathology* 89: 84-91.
- Bailey, D.J. and C. A. Gilligan. (1997). Biological control of pathozone behaviour and disease dynamics of *Rhizoctonia solani* by *Trichoderma viride*. *New Phytologist* 136: 359-367.
- Bascompte, J., and R.V. Solé. (1996). Habitat fragmentation and extinction thresholds in spatially explicit models. *Journal of Animal Ecology* 65: 465-473.
- Caswell, H., and R.J. Etter. (1993). Ecological interactions in patchy environments: From patch-occupancy models to cellular automata. In: *Lecture Notes in Biomathematics* (Levin, S.A., T.M. Powell, and J.H. Steele, eds.) pp. 93-109. New York: Springer-Verlag.
- Dickman, R. (1986). Kinetic phase transitions in a surface reaction model: mean field theory. *Physics Reviews A* 34:4246-4250.
- Dickman, R. (1988). Mean field theory of the driven diffusive gas. *Physics Reviews A* 38:2588-2593.
- Durrett, R. (1999). Stochastic spatial models. *SIAM Review* 41: 677-718.

- Durrett, R., and S. Levin. (1994a). The importance of being discrete (and spatial). *Theoretical Population Biology* 46: 363-394.
- Durrett, R., and S.A. Levin. (1994b). Stochastic spatial models: a user's guide to ecological applications. *Philosophical Transactions of the Royal Society of London B* 343: 329-350.
- Ellner, S.P., A. Sasaki, Y. Haraguchi, and H. Matsuda. (1998). Speed of invasion in lattice population models: pair-edge approximation. *Journal of Mathematical Biology* 36: 469-484
- Filipe, J. A. N. and G. J. Gibson. (1998). Studying and approximating spatio-temporal models for epidemic spread and control. *Philosophical Transactions of the Royal Society of London B* 353: 2153-2162.
- Harada, Y., H. Ezoe, Y. Iwasa, H. Matsuda, and K. Satō. (1995). Population persistence and spatially limited social interaction. *Theoretical Population Biology* 48:65-91.
- Harada, Y., and Y. Iwasa. (1994). Lattice population dynamics for plants with dispersing seeds and vegetative propagation. *Researches on Population Ecology* 36: 237-249.
- Hassell, M.P., Comins, H.N., and May, R.M. (1991). Spatial structure and chaos in insect population dynamics. *Nature* 353:255-258.
- Hiebeler, D. (1997). Stochastic spatial models: from simulations to mean field and local structure approximations. *Journal of Theoretical Biology* 187: 307-319.
- Hiebeler, D. (2000). Populations on fragmented landscapes with spatially structured heterogeneities: landscape generation and local dispersal. *Ecology* 81:1629-1641.
- Hogeweg, P. (1988). Cellular automata as a paradigm for ecological modeling. *Applied Mathematics and Computation* 27:81-100.
- Kim, K. and C.D. Harvell. (2001). Aspergillosis of sea fan corals: dynamics in the Florida Keys. In: *Connections Between Ecosystems in the South Florida Hydroscapes: The River of Grass Continues* (Porter, J. W. and K. Porter, eds.) CRC Press.
- Kleczkowski, A., C.A. Gilligan and D.J. Bailey. (1997). Scaling and spatial dynamics in plant pathogen systems: from individuals to populations. *Proceedings of the Royal Society of London B* 264: 979-984.
- Kubo, T., Y. Iwasa, and N. Furumoto. (1996). Forest spatial dynamics with gap expansion: total gap area and gap size distribution. *Journal of Theoretical Biology* 180: 229-246.
- Lavorel, S., R.V. O'Neill, and R.H. Gardner. (1994). Spatio-temporal dispersal strategies and annual plant species coexistence in a structured landscape. *Oikos* 71: 75-88.

- Levin, S.A., and R. Durrett. (1996). From individuals to epidemics. *Philosophical Transactions of the Royal Society of London Series B* 351: 1615-1621.
- Matsuda, H., N. Ogita, A. Sasaki, and K. Satō. (1992). Statistical mechanics of population: the lattice Lotka-Volterra model. *Progress of Theoretical Physics* 88: 1035-1049.
- McCauley, E., W.G. Wilson, and A.M. deRoos. (1993). Dynamics of age-structured and spatially structured predator-prey interactions: individual-based models and population-level formulations. *American Naturalist* 142: 412-442.
- Nakamaru, M., H. Matsuda, and Y. Iwasa. (1997). The evolution of cooperation in a lattice-structured population. *Journal of Theoretical Biology* 184:65-81.
- Pacala, S.W., Hassell, M.P., and May, R.M. (1990). Host-parasitoid associations in patchy environments. *Nature* 344:150-153.
- Rand, D. (1999). Correlation equations and pair approximations for spatial ecologies. In: *Advanced Ecological Theory : Principles and Applications* (J. McGlade, ed.) pp. 100-142. Oxford: Blackwell Science.
- Rhodes, C.J., and R.M. Anderson. (1996a). A scaling analysis of measles epidemics in a small population. *Philosophical Transactions of the Royal Society of London B* 351: 1679-1688.
- Rhodes, C.J., and R.M. Anderson. (1996b). Dynamics in a lattice epidemic model. *Physics Letters A* 210: 183-188.
- Rhodes C.J., and R.M. Anderson (1996c). Persistence and dynamics in lattice models of epidemic spread. *Journal of Theoretical Biology* 180: 125-133.
- Rhodes, C.J., and R.M. Anderson. (1996d). Power laws governing epidemics in isolated populations. *Nature* 381: 600-602.
- Rhodes, C.J., and R.M. Anderson. (1997). Epidemic thresholds and vaccination in a lattice model of disease spread. *Theoretical Population Biology* 52: 101-118.
- Rhodes C.J., H.J. Jensen, and R.M. Anderson (1997). On the critical behaviour of simple epidemics. *Proceedings of the Royal Society of London B* 264: 1639-1646.
- Satō, K., H. Matsuda, and A. Sasaki. (1994). Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology* 32:251-268.
- van Baalen, M., and D.A. Rand. (1998). The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology* 193:631-648.

6. APPENDICES

6.1 Derivation of ordinary pair approximation equations

Look first at the birth rate of ρ_1 , the rate at which single cells in the lattice change from 0 (unoccupied) to 1 (occupied). Incorporating the global birth rate b , the expected number of occupied nearest neighbors for an unoccupied cell $q_{1/0}$, and the proportion of cells that are unoccupied ρ_0 , the birth rate is the product $b q_{1/0} \rho_0$.

Similarly, the death rate for ρ_1 , the rate at which cells change from 1 to 0, is $d\rho_1$. This incorporates the intrinsic death rate d and the proportion of cells that are occupied, ρ_1 . Since it is assumed that $d = 1$, the death rate for ρ_1 is simply ρ_1 .

The net rate of change in ρ_1 is birth rate – death rate:

$$\frac{d\rho_1}{dt} = b q_{1/0} \rho_0 - \rho_1 = b(\rho_1 - \rho_{11} - \rho_{12}) - \rho_1 \quad (6.1.1)$$

since $q_{1/0} = \rho_{10}/\rho_0$ and $\rho_{10} = \rho_1 - \rho_{11} - \rho_{12}$.

Next, examine the birth and death rates for ρ_{11} . Since there are eight neighbors the rate at which 01's switch to 11's is $\frac{b}{8}(1 + 7q_{1/01})\rho_{10}$. This incorporates the knowledge that one neighbor of the occupied site is occupied, the probability that the other neighboring sites are occupied, and the proportion of sites that are in state 10. Because of symmetry, the rate at which 10's switch to 11's is exactly the same, so the total rate at which 01's and 10's switch to 11's, the birth rate for ρ_{11} , is $\frac{b}{4}(1 + 7q_{1/01})\rho_{10}$. From the basic OPA assumption that $q_{a/bc} \approx q_{a/b}$, the birth rate is approximated by $\frac{b}{4}(1 + 7q_{1/0})\rho_{10}$.

The rate at which pairs of neighboring sites switch from 11 (both occupied) to 01 or 10 (one occupied), the death rate for ρ_{11} , is $2d\rho_{11}$, or $2\rho_{11}$. This incorporates the proportion of 11 pairs and the possibility of either site experiencing death.

From these two rates, it is possible to calculate the rate of change in ρ_{11} (birth – death):

$$\frac{d\rho_{11}}{dt} = \frac{b}{4}\rho_{10}(1+7q_{1/0}) - 2\rho_{11} = \frac{b}{4}(\rho_1 - \rho_{11} - \rho_{12})\left(1 + 7\frac{\rho_1 - \rho_{11} - \rho_{12}}{1 - \rho_1 - \rho_2}\right) - 2\rho_{11} \quad (6.1.2)$$

Finally, look at the rates of change for ρ_{12} . The birth rate, the rate at which pairs switch from 02 to 12, is $\frac{7b}{8}q_{1/02}\rho_{02}$, or $\frac{7b}{8}q_{1/0}\rho_{02}$, from the OPA assumption. The death rate is simply $d\rho_{12}$ or ρ_{12} , since there is only one occupied cell in which death could occur in a 12 pair.

The net rate of change in ρ_{12} , then, is (birth – death):

$$\frac{d\rho_{12}}{dt} = \frac{7b}{8}\rho_{02}q_{1/0} - \rho_{12} = \frac{7b}{8}(\rho_1 - \rho_{11} - \rho_{12})\left(\frac{\rho_2 - \rho_{12} - \rho_{22}}{1 - \rho_1 - \rho_2}\right) - \rho_{12} \quad (6.1.3)$$

Since ρ_2 and ρ_{22} are constants, equations (6.1.1), (6.1.2), and (6.1.3) above form a closed set of equations that can be solved analytically.

6.2 Derivation of pair-edge approximation for invasion speed

6.2.1 Quasi-equilibria

The quasi-equilibria for the state of cells on the same row as the front runner will be denoted by tildes. The quasi-equilibria are derived by setting $\tilde{\rho}_1 = \frac{1}{2}\rho_1$ where ρ_1 is found by solving the OPA equations (see Appendix 6.1). $\tilde{\rho}_{11}$ and $\tilde{\rho}_{12}$ are then found by solving the set of equations:

$$\begin{aligned} \frac{d\tilde{\rho}_{11}}{dt} &= \frac{b}{4}(\tilde{\rho}_1 - \tilde{\rho}_{11} - \tilde{\rho}_{12}) \left(1 + 7 \frac{\tilde{\rho}_1 - \tilde{\rho}_{11} - \tilde{\rho}_{12}}{1 - \tilde{\rho}_1 - \tilde{\rho}_2} \right) - 2\tilde{\rho}_1 \text{ and} \\ \frac{d\tilde{\rho}_{12}}{dt} &= \frac{b}{4}(\tilde{\rho}_1 - \tilde{\rho}_{11} - \tilde{\rho}_{12}) \left(\frac{\tilde{\rho}_2 - \tilde{\rho}_{12} - \tilde{\rho}_{22}}{1 - \tilde{\rho}_1 - \tilde{\rho}_2} \right) - \tilde{\rho}_{21}. \end{aligned}$$

6.2.2 Wave or invasion speed

In estimating the invasion speed, it is assumed that the invasion dynamics on a lattice with only three rows will approximate the dynamics on a lattice that is infinite in two dimensions. Three components are needed for the computation: rate of forward steps (births into the row in front of the front runner), rate of backward steps (death of a lone front-runner), and the mean fallback distance.

The rate of forward steps is approximated by

$$\frac{3}{8}b \left(1 + 2 \frac{\tilde{\rho}_{11}}{\tilde{\rho}_1} \right) \frac{(1 - 2\rho_2 + \rho_{22})}{1 - \rho_2}.$$

The $3/8b$ part of this expression reflects birth taking place in one of the 3 cells in front of the front runner, out of the 8 total neighbors of the front runner. The central part of the expression (in parenthesis) reflects the possibility that the birth could occur from the front

runner or one of its occupied neighbors. The final fraction in the expression accounts for the possibility that one or more of the 3 cells in front of the front runner are dead.

A backwards move occurs if there is a single front-runner and it dies. Since $d = 1$, the rate of backwards movement is equal to the probability that the front-runner has no occupied

neighbors on the front row, which is $\left(1 - \frac{\tilde{\rho}_{11}}{\tilde{\rho}_1}\right)^2$.

The estimated invasion speed is calculated by taking the rate of forward steps and subtracting the rate of backward steps, multiplied by the estimated mean fallback distance $E[\delta]$, which is derived in Appendix 6.3:

$$\text{invasion speed} = \frac{3}{8} b \left(1 + 2 \frac{\tilde{\rho}_{11}}{\tilde{\rho}_1}\right) \frac{(1 - 2\rho_2 + \rho_{22})}{1 - \rho_2} - \left(1 - \frac{\tilde{\rho}_{11}}{\tilde{\rho}_1}\right)^2 \left(1 + \frac{q_{0/1} q_{0/0}}{1 - q_{0/0}^3}\right).$$

6.3 Approximating $E[\delta]$

The pair-edge estimate of the mean fallback distance δ uses the 3-strip approximation and assumes that OPA holds in the occupied area behind the front row. The estimate therefore can only be computed when OPA predicts that the population persists. To compute $E[\delta]$, first find $\Pr\{\delta \geq k\}$ for $k=1,2,3,\dots$. A fallback occurs when the front runner dies, and the two sites immediately above and below the front runner are both vacant.

Let $g_k(i)$ be the probability that there are no occupied sites in rows $1,2,\dots,k$ back from the front, and that the center site in the k^{th} row back is in state i , $i=0,2$. Then $g_k(0) + g_k(2)$ is the probability that rows $1,2,\dots,k$ back are all vacant, which equals $\Pr\{\delta \geq k+1\}$. From the fact that the center site in the front row was a 1 until the death occurred, OPA gives us the (approximate) values

$$\begin{aligned} g_1(0) &= q_{0/1}(1 - q_{1/0})^2 \\ g_1(2) &= q_{2/1}(1 - q_{1/2})^2 \end{aligned} \quad (6.3.1)$$

for the first row back. In each of the expressions in (6.3.1) the $q_{i/1}$ term is the conditional probability for the center site (based on the just-deceased front runner), and the second term is the conditional probability (based on the center site) that the sites one up and one down are also not occupied.

The same reasoning allows us to move back iteratively through the lattice. Let \bar{g}_k be the 2×1 column vector $[g_k(0), g_k(2)]$. Using OPA and conditioning on the center site in row k we then have

$$\bar{g}_{k+1} = \begin{bmatrix} q_{0/0}(1 - q_{1/0})^2 & q_{0/2}(1 - q_{1/0})^2 \\ q_{2/0}(1 - q_{1/2})^2 & q_{2/2}(1 - q_{1/2})^2 \end{bmatrix} \bar{g}_k . \quad (6.3.2)$$

Letting \mathbf{A} denote the matrix in (6.3.2), we therefore have $\bar{\mathbf{g}}_k = \mathbf{A}^{k-1} \bar{\mathbf{g}}_1$, and

$$\begin{aligned}\mathbf{E}[\delta] &= \sum_{k>0} \Pr\{\delta \geq k\} = 1 + [1,1]^T \bar{\mathbf{g}}_1 + [1,1]^T \bar{\mathbf{g}}_2 + [1,1]^T \bar{\mathbf{g}}_3 + \dots \\ &= 1 + [1,1]^T (\bar{\mathbf{g}}_1 + \mathbf{A} \bar{\mathbf{g}}_1 + \mathbf{A}^2 \bar{\mathbf{g}}_1 + \dots) \\ &= 1 + [1,1]^T (\mathbf{I} - \mathbf{A})^{-1} \bar{\mathbf{g}}_1\end{aligned}\tag{6.3.3}$$

so long as the series in the second row of (6.3.3) converges as claimed. Note that the first and second column sums of \mathbf{A} are bounded above by $q_{0/0} + q_{2/0} = 1 - q_{1/0}$ and $1 - q_{1/2}$, respectively. So unless $q_{1/0} = 0$ or $q_{1/2} = 0$, \mathbf{A} therefore has spectral radius < 1 and (6.3.3) is valid. When OPA predicts that the population is persistent (which is a prerequisite for pair-edge approximation) we have $q_{1/0} > 0$. Moreover $q_{1/2} > 0$ unless the lattice is totally clustered ($q_{2/2} = 1$), meaning that no occupiable site has a dead site as a neighbor. In that case, however, the dead sites are irrelevant to the dynamics in the occupiable part of the lattice, so the expression for $\mathbf{E}[\delta]$ in Ellner et al. (1998) would still apply,

$$\mathbf{E}[\delta] = 1 + \frac{q_{0/1} q_{0/0}^2}{1 - q_{0/0}^3}. \tag{6.3.4}$$

In fact for perfect clustering \mathbf{A} reduces to a diagonal matrix with diagonal entries $(q_{0/0}, 1)$, so partial sums of the second line in (6.3.3) can be computed explicitly, and the infinite sum yields (6.3.4). The approximation (6.3.3) therefore predicts correctly that as $q_{2/2} \rightarrow 1$ the wave speed converges to the value for a homogeneous lattice regardless of the frequency of dead sites in the lattice.

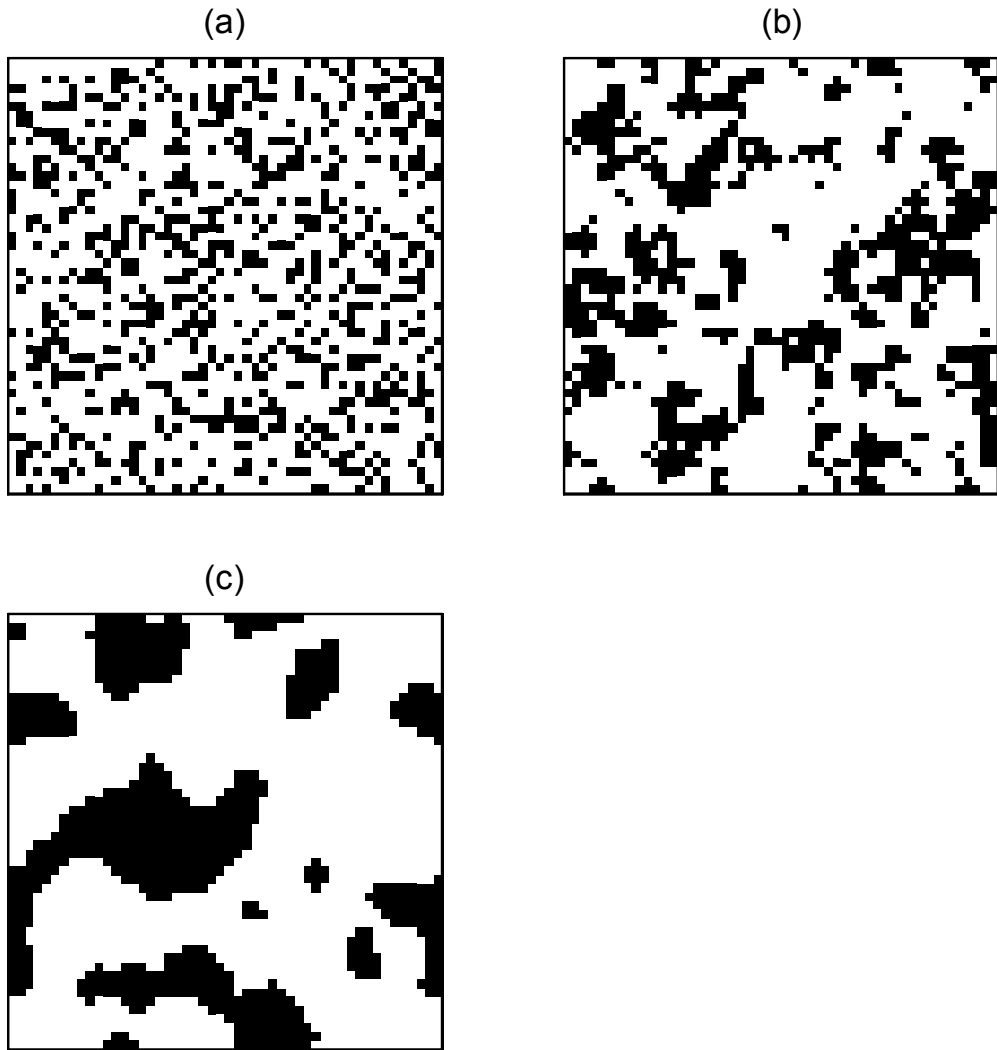


Figure 1. Heterogeneous lattices with different degrees of clustering. Occupiable cells are white and dead cells are black. Each lattice has 2500 cells, of which 30 percent (750) are dead. The degree of clustering is measured by ρ_{22} , which is the proportion of neighboring site-pairs that consist of two dead sites. A higher value of ρ_{22} indicates a higher degree of clustering. In (a), $\rho_{22} = 0.09$, which is equivalent to random placement of dead sites in the lattice. In (b), $\rho_{22} = 0.17$, and in (c), $\rho_{22} = 0.25$.

Invasion Process on a Heterogeneous Lattice

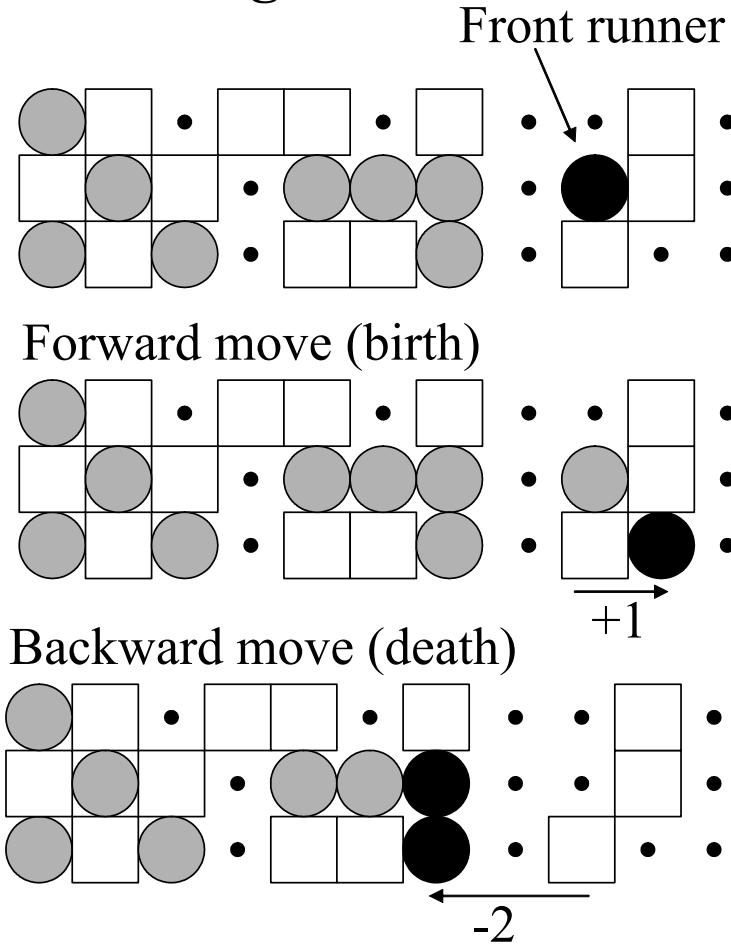


Figure 2. An example of the invasion process on a heterogeneous lattice of height three. Dead sites are represented by squares, unoccupied sites are represented by dots, and occupied sites are represented by circles. The front runners of the invasion are indicated by black circles. A forward move in the invasion front occurs when there is a birth into the column to the right of the front runner. A backward move occurs when the front runner dies, which in this case results in a backwards move of length two.

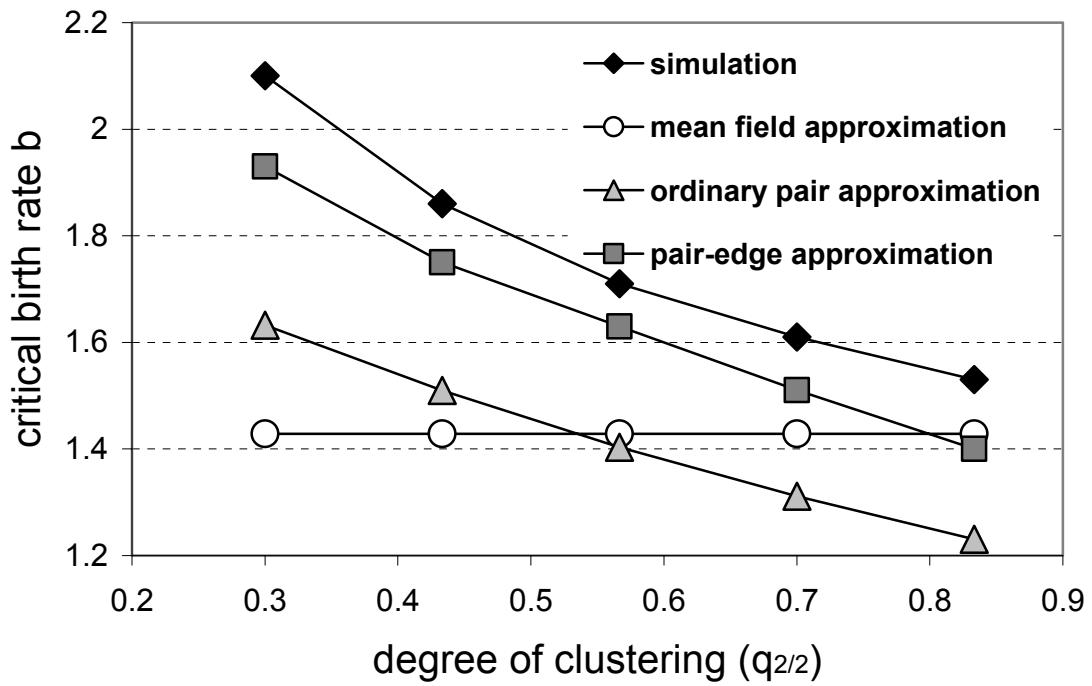


Figure 3. Comparison of critical birth rates from simulation and approximation as they vary with the degree of clustering of the dead sites in the lattice. Each simulated critical birth rate was determined to two decimal places by repeated trials.

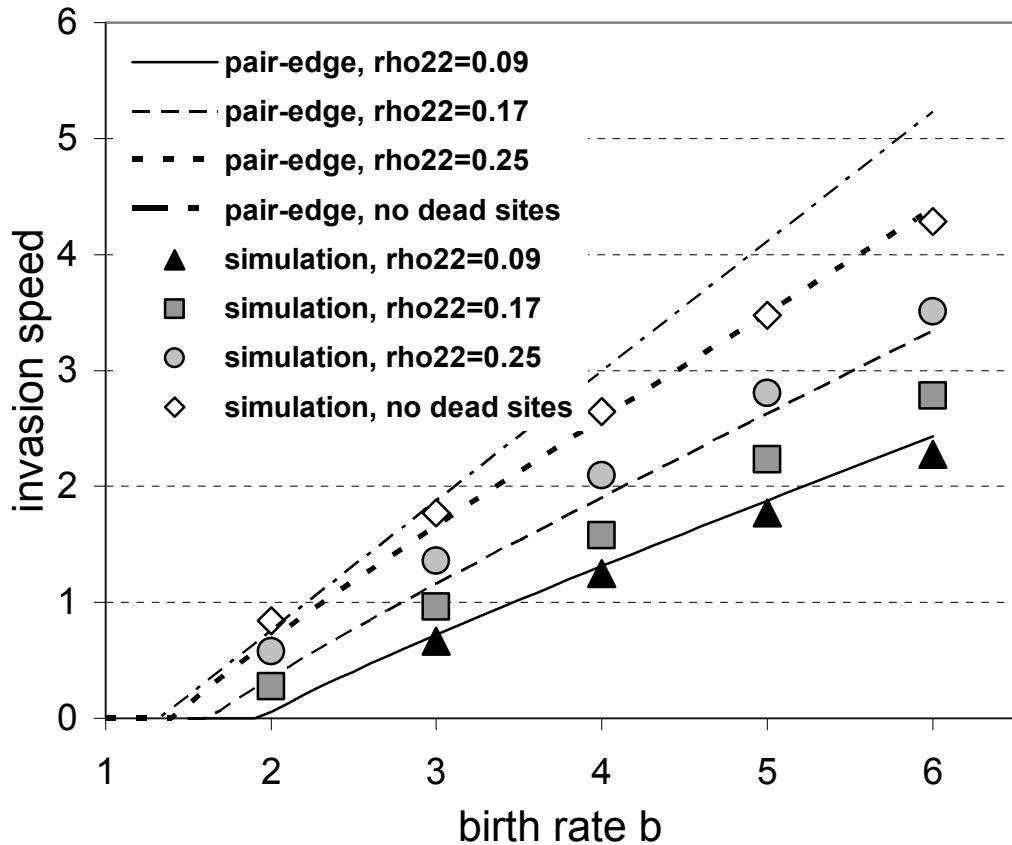


Figure 4. Comparison of invasion speeds derived from simulation and the pair-edge approximation. Invasion speed is measured in cells per unit time. With the exception of the line labeled ‘no dead sites,’ the fraction of dead sites was equal to 0.3. ρ_{22} represents the degree of clustering in the dead sites, with higher ρ_{22} indicating a higher degree of clustering. Each simulation data point represents the average of ten different simulations, on ten different randomly-generated lattices (in the heterogeneous cases).

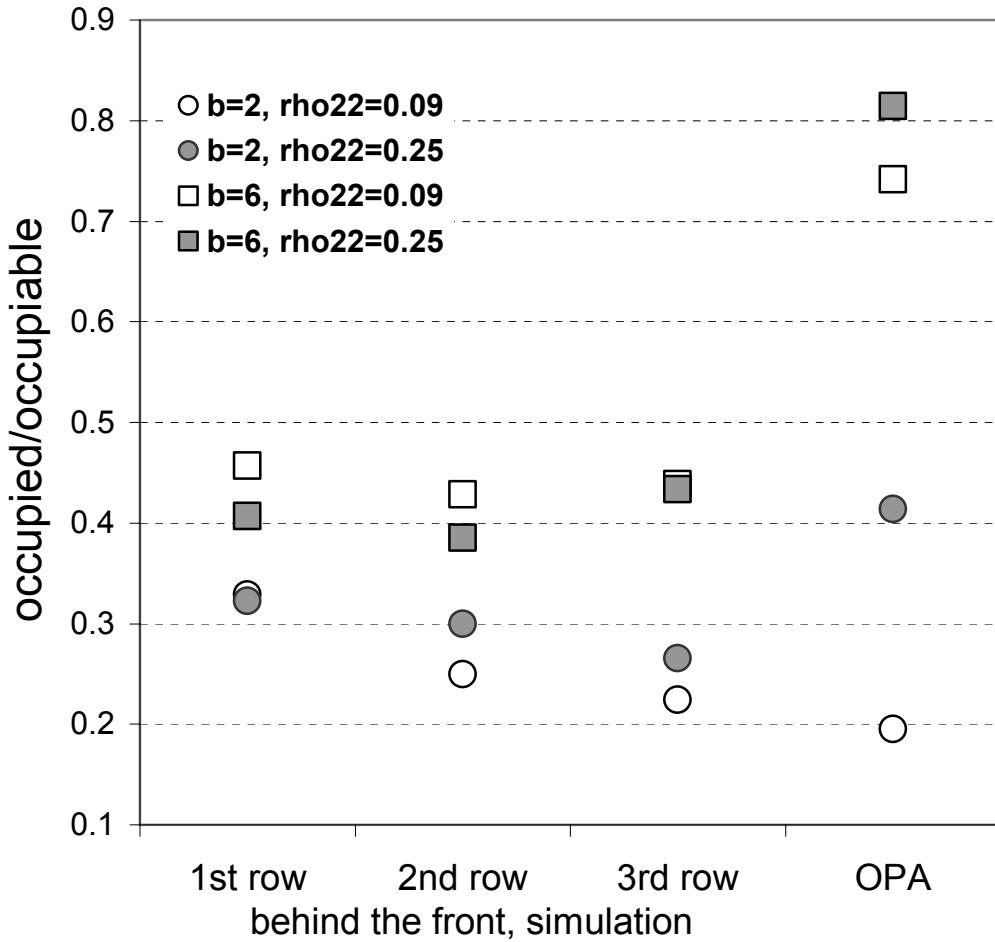


Figure 5. Comparison of lattice composition, measured by the fraction of occupiable (not dead) sites that are occupied. The average composition in the first three rows behind the invasion front in simulations is compared with that predicted by the ordinary pair approximation, which is assumed by the pair-edge approximation to apply for all rows behind the front. The simulation data points each represent the average of at least 188 different measurements taken at regular intervals during simulation. In the legend, ‘ b ’ represents the birth rate of an occupied cell, and ρ_{22} represents the degree of clustering in the dead sites, with higher ρ_{22} indicating a higher degree of clustering. The fraction of dead sites was equal to 0.3 in all cases.