Analytical methods for stochastic discrete-time metapopulation models

Fionnuala Mary Buckley

A thesis submitted for the degree of Doctor of Philosophy at
The University of Queensland in August 2010

School of Mathematics and Physics

Typeset in LaTeX.

Declaration by Author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the Copyright Act 1968.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material.

List of Publications

Citation numbers are the same as those given in the references section.

Statement of Contributions to Jointly Authored Works Contained in the Thesis:

[16] F. M. Buckley and P. K. Pollett. Analytical methods for a stochastic mainland-island metapopulation model. In R. S. Anderssen, R. D. Braddock, and L. T. H. Newham, editors, 18th IMACS World Congress - MODSIM09 International Congress on Modelling and Simulation, volume July 2009, pages 1767–1773, Canberra, Australia, 2009. Modelling and Simulation Society of Australia and New Zealand.
— P. K. Pollett proof-read the paper, made some editorial changes and wrote MATLAB code for the figures. F. M. Buckley was responsible for the rest of the work.

[15] F. M. Buckley and P. K. Pollett. Limit theorems for discrete-time metapopulation models. *Probability Surveys*, 7:53–83, 2010.
— P. K. Pollett proof-read the paper, made more significant editorial changes, wrote MATLAB code for the figures and is responsible for Theorem 2 and section 5.2. F. M. Buckley was responsible for the rest of the work.

[14] F. M. Buckley and P. K. Pollett, Analytical methods for a stochastic mainland-island model. *Ecological modelling*, in press (doi:101016/j.ecolmodel.2010.02.017).
— The contributions by each author are the same as for [16].

Statement of Contributions by Others to the Thesis as a Whole:

There has been no contribution by others to the writing of the thesis, however P. K. Pollett and R. S. McVinish have proof-read all of the thesis and provided suggestions. The principal advisor, P. K. Pollett, has also guided the path of the work while the associate advisor, R. S. McVinish, has made suggestions on the work as it progressed.

Statement of Parts of the Thesis Submitted to Qualify for the Award of Another Degree:

None.

Published Works by the Author Incorporated into the Thesis:

[16] F. M. Buckley and P. K. Pollett. Analytical methods for a stochastic mainland-island metapopulation model. In R. S. Anderssen, R. D. Braddock, and L. T. H. Newham, editors, 18th IMACS World Congress - MODSIM09 International Congress on Modelling and Simulation, volume July 2009, pages 1767–1773, Canberra, Australia, 2009. Modelling and Simulation Society of Australia and New Zealand.

— The material presented in [16] forms the basis of Chapters 3 and 4 where, specifically, the state-independent case is examined in Chapter 3 and state-dependent case is examined in Chapter 4.

[15] F. M. Buckley and P. K. Pollett. Limit theorems for discrete-time metapopulation models. *Probability Surveys*, 7:53–83, 2010.

— The infinite-patch models presented in section 5.1 of [15] form the basis of Chapter 5, while the majority of [15] is incorporated in Chapter 6.

[14] F. M. Buckley and P. K. Pollett, Analytical methods for a stochastic mainland-island model. *Ecological modelling*, in press (doi:101016/j.ecolmodel.2010.02.017).

— As above for [16].

Additional Published Works by the Author Relevant to the Thesis but not Forming Part of it:

None.

Acknowledgements

First and foremost, I would like to thank my principal advisor Prof. Phil Pollett. Thank you for your guidance and support throughout my PhD candidature at the University of Queensland (UQ) and for the opportunity to be part of the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems (MASCOS). And second, I would like to thank my associate advisor, Dr Ross S. McVinish. Thank you for your support also.

There are two sources of financial support I wish to acknowledge: the Australian Government and MASCOS. From the first I received a stipend under an Australian Postgraduate Award, and from the second I received a top-up stipend. I would also like to acknowledge further financial support from MASCOS, which allowed me to make use of a Macintosh Powerbook and attend various meetings and conferences within Australia.

A number of personnel within the Mathematics Department at UQ deserve to be mentioned also. I wish to thank Prof. Dirk Kroese and Prof. Hugh Possingham for their comments and suggestions whilst acting as independent assessors at my second-year (D.K.) and third-year (H.P.) postgraduate milestones. To Kathy Lyons and Murray Kane, thank you for your assistance with regard to administrative matters. To the Mathematics postgraduate cohort, in particular those who shared the corridor with me on Level 6, Building 69, thank you for your friendship and support.

Finally, I would like to thank my husband, Graham, and my mother, Christine, as well as the immediate members of the Buckley and Fox families, for their constant love and encouragement throughout my postgraduate years.

Abstract

The term ‘metapopulation’ is used to describe a population of individuals that live as a group of local populations in geographically separate, but connected, habitat patches. The patches are situated within an otherwise uninhabitable landscape which dispersing individuals traverse in search of suitable habitats. The central concepts associated with metapopulation dynamics are that of *local extinction*, the extinction of a local population, and *recolonisation*, where migrants establish new populations in unoccupied (empty) patches. The relationship between the processes of local extinction and colonisation is therefore an important consideration when formulating mathematical metapopulation models.

We shall consider a particular type of metapopulation dynamic where extinction events and colonisation events are assumed to occur during separate time periods, or *phases*, that alternate over time. This seasonal dynamic may be thought of as an annual cycle, say, where local populations are prone to extinction during winter and new populations establish during spring. In particular, we model the number n_t of occupied patches at time t as a discrete-time Markov chain $(n_t : t = 0, 1, 2, \dots)$ with transition probabilities that alternate according to the seasonal phases. The models are naturally constructed in a discrete-time setting due to the assumed dynamic, however it will be made clear that whether the Markov chains are time-homogeneous (where the population census is taken after every cycle) or time-inhomogeneous (where the census is taken after each seasonal phase) depends on the monitoring scheme under investigation.

We present a number of metapopulation models with the assumed seasonal dynamic where, in particular, the local extinction process is modelled in the same way in each case whilst the colonisation process is modelled according to various means of propagation. We assume that each local population goes extinct with the same, constant, probability, and that all events are independent. Hence, the number of extinction events that occur during the extinction phase is binomial. For metapopulation networks with a finite number N of patches, we also assume that the number of colonisation events that occur during the colonisation phase is binomial. We investigate both state-independent and state-dependent colonisation probabilities, where the former is defined with a constant probability and the latter depends on the current number of occupied patches. Metapopulation models defined with a state-independent colonisation process are referred to as *mainland models* because empty patches are thought to be colonised by migrants from an outside source population (the 'mainland') in this case. For models defined with state-dependent colonisation processes, we refer to these as *island models* when colonists originate from occupied patches (islands) or *mainland-island models* when both types of colonising behaviour are assumed. The overall two-phase model is called a *chain binomial metapopulation model* since the extinction and colonisation processes together define a sequence of binomial random variables. We also investigate similar models but for networks with infinitely-many patches ($N = \infty$); the number of colonisation events is modelled as a Poisson random variable in such cases.

The discrete-time Markov chain approach is well established in the applied metapopulation literature, however models of this type are usually examined via numerical methods and simulation. The models presented here are accompanied with extensive analytical treatments. For most of our finite-patch models, we evaluate conditional state distributions explicitly and use these distributions to establish convergence results (in the sense of convergence in distribution). These results include a law of large numbers, which identifies an approximating deterministic trajectory, and a central limit law, which establishes that the scaled fluctuations about the deterministic trajectory have an approximate normal (Gaussian) distribution. We show that the infinite-patch models are equivalent to branching processes.

This body of work culminates by presenting limit theorems for discrete-time metapopulation models. First, we prove limit theorems for a general class of inhomogeneous Markov chains that exhibit the particular property of density dependence. These theorems include a law of large numbers and a central limit law (in the sense of convergence in finite-dimensional distribution), which establishes that the scaled fluctuations about this deterministic trajectory have an approximating autoregressive structure. Second, we apply these results to our Markov chain metapopulation models (both finite-patch and infinite-patch) where we demonstrate that the limiting behaviour of any of our metapopulation models with density-dependent phases can be evaluated explicitly, even in situations where a conditional state distribution could not be determined.

Keywords:

Metapopulation; discrete-time Markov chain; chain binomial model; seasonal phases; mainland; mainland-island; law of large numbers; central limit law; branching process; limit theorems; autoregressive.

Australian and New Zealand Standard Research Classifications (ANZSRC):

010406 Stochastic Analysis and Modelling

Contents

List of Figures

List of Abbreviations

a.s.	almost surely
AR	autoregressive
CE	colonisation-extinction
cf	characteristic function
CLL	central limit law
EC	extinction-colonisation
iid	independent and identically distributed
LLN	law of large numbers
pgf	probability generating function
SIS	susceptible-infective-susceptible

List of Symbols

C	colonisation transition matrix, $C = (c_{ij})$
$\mathrm{Cov}\,(X, Y)$	covariance of X and Y
c	colonisation parameter
c_0	colonisation parameter attributed to a mainland
c_{ij}	transition probability for the colonisation process
$\mathbb{E}$	expectation
E	local extinction transition matrix, $E = (e_{ij})$
E_{N}	$\{1, 2, \ldots, N\}$
E_t	$\mathrm{N}(0, v(x_t))$ random variable (v is some continuous function)
e	extinction parameter
e_{ij}	transition probability for the local extinction process
exp	the exponential function
η	probability of extinction of a Galton-Watson branching process
F_{ij}	the generating function defined by $F_{ij}(z) = \sum_{t=0}^{\infty} z^t f_{ij}^{(t)}$
$f_{ij}^{(t)}$	starting in state i, $f_{ij}^{(t)}$ is the probability the first visit to state j occurs at time t
G	pgf of n_{t+1} conditional on n_t
G_t	pgf of n_t conditional on n_0 ($G_1 \equiv G$)
$\mathbf{i}$	$\sqrt{-1}$
i, j	states of a random process
K	carrying capacity
μ	$\mathbb{E}(\xi)$

μ_t	$\mathbb{E}(Z_t)$
N	total number of patches in a metapopulation network
n_t, n_t^N	number of occupied patches at time t (the symbol n_t^N is used in the context of time-inhomogeneous random processes, where the superscript N is an index)
P	1-step probability transition matrix, $P = (p_{ij})$
P_{ij}	the generating function defined by $P_{ij}(z) = \sum_{t=0}^{\infty} z^t p_{ij}^{(t)}$
p_{ij}	1-step transition probability
$p_{ij}^{(t)}$	t-step transition probability ($p_{ij}^{(1)} \equiv p_{ij}$)
$\mathbb{R}$	the set of real numbers
r	intrinsic growth rate
S	$\{0, 1, 2, \ldots\}$
S_N	$\{0, 1, 2, \ldots, N\}$
σ^2	$\mathrm{Var}\,(\xi)$
T	index set of a random process or first passage time
t	time
$\mathrm{Var}\,(X)$	variance of X
V_t	$\mathrm{Var}\,(Z_t)$
ξ	number of offspring produced by one particle
x^*	fixed point of a continuous function f defined by $x_{t+1} = f(x_t)$
X_t^N	proportion of occupied patches at time t (the superscript N is an index)
x_t	deterministic proportion of occupied patches at time t
ζ	number of immigrant particles
$\mathbb{Z}_+$	$\{0, 1, 2, 3, \ldots\}$
Z_t	$\mathrm{N}(\mu_t, V_t)$ random variable
Z_t^N	quantity defined as $\sqrt{N}(X_t^N - x_t)$ (the superscript N is an index)

Chapter 1

Introduction

In this thesis, population models are constructed as time-inhomogeneous Markov chains with transition matrices that alternate in discrete time. This alternating behaviour effectively models seasonal variation in population dynamics; a method that has been applied extensively in ecology (colonisation and extinction of local populations in a metapopulation network) and to a smaller extent in epidemiology (infection and recovery of individuals during the spread of a disease). This chapter reviews different types of population models that account for seasonal variation and motivates the use of inhomogeneous Markov chains for this particular purpose. The chapter closes by outlining the material presented in this thesis.

1.1 Population modelling

In 1798, Thomas R. Malthus published his philosophical views on the growth of the human population of the world [57]. He argued that if the population grew without restraint, its growth would forever be controlled (kept in 'check') by the means of subsistence. Malthus went so far as to propose a crude mathematical model where the (unchecked) population grew according to a geometric progression whilst subsistence grew according to an arithmetic progression; the former forever outstripping the latter.

Models introduced by Gompertz in 1825 [32] and Verhulst in 1838 [89] instead accounted for Malthus' idea of 'checked' growth. Their models defined the change in population size n at time t as a differential equation, where the quantity n retards over time and slows down to a fixed point in the model dynamics. Verhulst's model, as well as its variations, are perhaps the most well known classic population models. Using Pearl and Reed's rationale [71] (see also [69, 70]), we write Verhulst's model as

$$\frac{dn}{dt} = \frac{\lambda}{N} n(N-n) - \mu n \qquad (n \in [0, N]), \tag{1.1}$$

where λ is the (pair-wise) birth parameter, μ the death parameter and N the maximum population size that can be supported by limited resources; the parameter N is implicitly assumed to be large so that n is effectively considered as a continuous variable. Indeed, the model given by (1.1) is well known in two particular guises. In the epidemiology context, (1.1) is the *SIS (Susceptible-Infective-Susceptible) logistic model* where, for a closed population of N individuals, λ is the rate at which a disease spreads in infective-susceptible pairs and μ is the recovery rate. In the ecological context, this same model is known as the *Levins model* due to Levins' interpretation [51, 53] of (1.1) for a population occupying a network of N habitat patches; λ being the rate at which new populations establish within pairs of unoccupied-occupied patches and μ the extinction rate for established populations.

As Verhulst's model demonstrates, differential equation models are formulated in continuous time. If a population was considered to grow within discrete time intervals, say, where generations or breeding cycles do not overlap in time, then the population growth rate dn/dt may be discretized as $n_{t+1} - n_t$ (using an obvious notation) where $t \in \{0, 1, 2, \dots\}$. The discrete-time analogue of Verhulst's model,

$$n_{t+1} - n_t = \frac{\lambda}{N} n_t(N - n_t) - \mu n_t \qquad (n_t \in [0, N]), \tag{1.2}$$

is an example of a difference equation (for more examples, see [60] or [61]). Upon solving a differential equation or a difference equation explicitly for n, the solution is an exact (deterministic) trajectory for the population size as a function of the time variable.

Due to their deterministic nature, the classic differential and difference equation models do not account for random variation in demographic or environmental factors. Stochastic models, on the other hand, account for random variation and allow practitioners to calculate probability distributions for many quantities of interest. Markov processes formulated with discrete state spaces (commonly referred to as *Markov chains*) are particularly favoured for modelling population dynamics due to the advantage of both the Markov Property (that future predictions only rely on present observations) and being able to model population size with discrete values. Markov chains may be formulated in continuous time or discrete time depending on the application of interest.

Birth-death processes, a particular class of continuous-time Markov chains, are natural stochastic analogues of classic continuous-time population models. For example, the stochastic analogue of (1.1) is a birth-death process where the population size n is defined to increase by one unit at rate $(\lambda/N)n(N-n)$ and decrease by one unit at rate μn, where the parameters λ and μ are defined as before. Indeed (1.1) approximates its stochastic counterpart as the population ceiling N grows large [72]. Of course, birth-death processes can accommodate more general birth rates and death rates.

Population models that are based on discrete-time Markov chains are natural stochastic analogues for classic discrete-time population models. It will be demonstrated shortly that discrete-time Markov chains are particularly well suited to modelling seasonal population dynamics.

1.2 Accounting for seasonal variation

An assumption common to many population models is that ‘birth’ and ‘death’ events occur in random order over time, thus implying that the population being modelled does not exhibit any seasonal dynamics. Even when the population census occurs at discrete time points according to non-overlapping generations, discrete-time models do not necessarily account for seasonal behaviour because birth and death events may still be assumed to occur in random order between those time points. Equation (1.2), for example, does not assume a seasonal dynamic.

Population dynamics that are seasonal in behaviour clearly exhibit a time-dependent structure and population models that account for this kind of behaviour must necessarily be time-dependent in some way. Levins [51, 52, 53] was one of the first to account for seasonal variation and did so by considering Verhulst’s classic differential equation with time-dependent parameters. Similar studies were performed on the classic exponential growth model, $dn/dt = \lambda n$, where Lewontin and Cohen [54] focused on a discrete-time version of the model while the differential equation was studied by May [58]. In particular, all three studies are stochastic because the time-dependent parameters are defined as random variables. The population is subsequently said to be in a ‘randomly fluctuating environment’, however note that if the random variables are independent and identically distributed, then this type of time-dependence does not strictly account for 'seasonal' fluctuations since there is no serial correlation between successive time points. Explicit solutions were obtained in all three studies. However the method of using classic population models with time-dependent parameters did not gain traction in the ecological literature. A more complicated discrete-time method introduced by Roff [77, 78] gained traction instead. Roff's method utilised the concept of un-correlated random fluctuations, but it is the second type of time-dependence incorporated in his model that successfully accounts for seasonal variation.

The model developed by Roff [77, 78] considers a population of individuals as a set of local populations living in a spatially heterogeneous ('patchy') environment. This kind of population structure was recognised by Wright [94], Andrewartha and Birch [4], MacArthur and Wilson [56], and it is now commonly known as a *metapopulation* due to Levins' turn of phrase for a 'population of populations' introduced in 1970 [53]. Roff was particularly interested in investigating the effect dispersal had on the persistence of the population as a whole. Modelling the metapopulation on the grand scale (that is, at the patch-occupancy level where the size of each local population is ignored) does not take dispersal explicitly into account and so Roff developed a model that tracked the number of individuals in each local population and the dispersal of individuals around the network. More specifically, Roff's model assumed a discrete-time census schedule whereby, during the course of one time interval, the system underwent a period of dispersal followed by a period of local growth. Dispersal was either holistic (every individual disperses) or local (a proportion was exchanged between nearest neighbours) while growth in each local population was modelled using a classic discrete-time dynamic with time-dependent parameters (which were random though un-correlated). Not only did Roff account for random fluctuations in the environment, these fluctuations were defined to occur at a specific stage in the assumed dispersal-growth dynamic. Roff concluded that population networks connected by dispersal persisted for a significantly longer time than identical systems without dispersal.

A number of models based on Roff's discrete seasonal dynamic were later developed. Chesson [18] presented a more formal stochastic model while others tended to adopt a deterministic framework. Deterministic models focused heavily on studying the stability effects of different local dynamics and these include coupled logistic equation models [45, 36, 33], models with general density-dependent local dynamics [23, 96], and models developed in the framework of adaptive dynamics [24, 67, 68] where the evolution of dispersal rates is studied in populations comprising a number of genotypes. (For an introduction to adaptive dynamics, I direct the reader to [91].) One particular study by Parvinen [68] extended Roff's seasonal dynamic to define four different time periods corresponding to the following events: reproduction, emigration, immigration and catastrophes.

In the metapopulation literature, models that account for local population dynamics are classified as *structured population models* [37]. Although the models mentioned above are obviously structured models, the characteristic of importance here is that dispersal from each patch and internal growth within each patch are assumed to be *synchronous* events [24, 96]. In this way, these models account for seasonal variation and do so regardless of whether model parameters are defined to be time-dependent or not.

1.2.1 The discrete-time Markov chain approach

Roff demonstrated that the discrete-time setting is a natural framework for modelling population dynamics that exhibit seasonal variation because the time interval need only be divided in such a way that each sub-interval represents a particular seasonal period. However, a significant drawback associated with structured models is the complicated nature of their model analysis. For example, the deterministic structured models mentioned above exhibit bifurcation phenomena. A much simpler approach that avoids chaotic dynamics is a stochastic framework based on discrete-time Markov chains.

The discrete-time Markov chain approach was introduced by Akçakaya and Ginzburg [2] and their model is described as follows. Suppose that the state of a discrete-time Markov chain is a vector of size N and that this vector describes the presence/absence of occupants in an N-patch metapopulation, where the k-th component is 1 or 0 according to whether the k-th patch is occupied or empty (that is, whether the k-th local population is extant or extinct), respectively. Local patch dynamics are not taken into account here and so the model is *unstructured* (see [37]) as a result. Hence, there are 2^N possible states and the zero vector is the sole absorbing state, corresponding to metapopulation extinction. Suppose that each patch is associated with a patch-specific local extinction probability, that each patch-pair is associated with a patch-specific colonisation probability (which can be uni-directional or bi-directional, depending on the application), and that local extinction events and colonisation events occur independently of one another. Given this set-up, two probability transition matrices of size $2^N \times 2^N$ can be constructed: (i) the matrix E contains positive entries for transitions that only involve extinction events (all other transitions set with zero probability) and, likewise, (ii) the matrix C contains positive entries for transitions that only involve colonisation events (all other transitions set with zero probability). One can then choose to set $P = EC$ or $P = CE$ as the 1-step transition matrix of a time-homogeneous Markov chain.

Indeed, $P = EC$ and $P = CE$ correspond to the transition matrices of two different time-homogeneous chains that arise by observing a time-inhomogeneous chain with, for example (see [2]), transition matrix

$$P(t, t+1) = \begin{cases} C & \text{if } t \in \{0, 2, 4, \dots\} \\ E & \text{if } t \in \{1, 3, 5, \dots\} \end{cases}$$

at times $t = 1, 3, 5, \dots$ and $t = 0, 2, 4, \dots$, respectively. The two observation schemes are referred to as the EC and CE models [16] for obvious reasons, where the former is used if the population census is taken after each successive period of colonisation and the latter if census follows each period of local extinction. Whist the timing of the census does not affect the dynamics of the system, it is certainly important from an empirical perspective. For example, the model developed by Klok and De Roos [50] (explored in more detail in the following section) takes the census after colonisation because the real population under investigation was known to be more stable at this time.

The model developed by Akçakaya and Ginzburg assumes that local extinction events and colonisation events occur during separate time intervals in a repeating cycle. These assumptions effectively imply that the events are synchronous, just like that described for local growth and dispersal in Roff's model. It is unclear, though, as to whether Akçakaya and Ginzburg fully appreciated the discrete model dynamics because there was no evidence to suggest that the application of interest—a 3-patch metapopulation of mountain gorillas (*Gorilla gorilla beringei*)—exhibited this kind of seasonal behaviour. Indeed, they presented results in for CE model only, stating that "the order of multiplication did not make an important difference in the results". Their results consisted of numerical calculations of extinction probabilities, and so the ‘difference’ they are referring to is the difference between extinction probabilities that correspond to the EC and CE models defined with the same model parameters. The size of the difference *is* important, since it is an indication of the underlying dynamic, and the results presented in this thesis demonstrate that the difference between the same quantities of the respective models can be given explicitly. Akçakaya and Ginzburg’s model certainly simplifies the way in which combinations of events are counted between discrete time points (that is, by considering downward jumps and upward jumps in the state space separately), which may or may not have dictated the model formulation. Nonetheless, the Markov chain approach quickly took hold in the applied metapopulation literature.

Day and Possingham [22] soon modified the Markov chain approach by allowing local extinction probabilities to be inversely proportional to patch size and colonisation probabilities to be dependent on the distance between patch pairs. Their spatially explicit model was applied to an 8-patch metapopulation of malleefowl (*Leipoa ocellata*). The authors note a lack of empirical data with which to estimate extinction and colonisation probabilities, however the annual breeding behaviour of malleefowl [19] might be considered to fit within the assumed seasonal dynamic.

The greatest limitation with the Akçakaya-Ginzburg and Day-Possingham models is the size of the state space. Remembering that these models have 2^N states, the numerical analysis of these models quickly becomes computationally expensive as N increases. Instead of being concerned with which patches are occupied and which are empty, the most significant change adopted in later models was to consider the total number of occupied patches instead and this reduced the size of the state space to $N+1$ for an N-patch metapopulation[1]. These models are computationally inexpensive to analyse even for large ($N = 50$) networks and can incorporate spatial information implicitly by allowing colonisation probabilities to depend on the number of occupied patches.

Metapopulations models described by inhomogeneous Markov chains with scalar states commonly assume that local extinction events are independent and occur with the same probability, e. If i patches are occupied at the beginning of the extinction period or *phase*, then the number of events that occur during this phase follows the binomial $\mathrm{Bin}(i, e)$ law.

---

[1] Stochastic patch-occupancy models (SPOMs) defined with scalar states are known as homogeneous SPOMs in the ecological literature, which distinguish them from heterogeneous SPOMs (vector-state models); ‘heterogeneous’ signifies that patches are not assumed to be identical. See [44], for example.

Colonisation events, on the other hand, are modelled in various ways according to specific breeding habits or dispersal patterns, or both if necessary. The colonisation process described by Klok and De Roos' model [50], for example, considers breeding habits of individuals and dispersal of offspring separately whereby the reproduction and settlement transition matrices (non-square) used to model these two processes multiply in this order to produce the overall (square) colonisation matrix. The breeding process is described by a recursive formula and the settlement process is governed by a binomial law. In particular, Klok and De Roos designed their model to track the number of female individuals in a population of common shrews (*Sorex araneus*). The model considers each female as occupying one patch in a metapopulation network where this species is known to be susceptible to death during winter (local extinction) and later breed then disperse from spring through to autumn (colonisation).

The colonisation process described by Hill and Caswell [46] is based on the binomial law that was designed by Klok and De Roos for the settlement of shrews. This law is described as follows. Migrating individuals are implicitly assumed to arrive at each patch according to a time-homogeneous Poisson process with rate $\beta i/N$, where i is the number of patches currently occupied and β is the expected number of propagules produced by each occupied patch. Thus, the probability that one or more propagules arrive at any given patch is $1 - \exp(-\beta i/N)$. Given that i patches are occupied, the total number of colonisation events follows a binomial $\text{Bin}(N - i, 1 - \exp(-\beta i/N))$ law, however Hill and Caswell go one step further by allowing only a fixed subset of the N patches to be suitable for habitation. Since the extinction and colonisation processes together define a sequence of binomial random variables, the authors called their model a *chain binomial metapopulation* model.

The method by which seasonal phases are characterised with binomial distributions is a method reminiscent of the well known Greenwood [34] and Reed-Frost [93, 1] chain binomial epidemic models. I now review how the inhomogeneous Markov chain method has been applied to modelling epidemics.

1.2.2 Chain binomial epidemic models

Chain binomial models date back to stochastic formulations by Reed and Frost in 1928 [93, 1] and Greenwood in 1931 [34]. These models characterise the spread of disease in a discrete-time setting by assuming that the infectious period is very short in comparison with the latent period, the length of which represents the time interval. The infectious period is assumed to be concentrated at the time of the population census and, during this period, each susceptible individual is equally likely to contact infected individuals. The Reed-Frost model specifies that this probability is dependent on the number of infectives while the Greenwood model specifies a constant. The initial number of infectives are removed and the remaining individuals, a binomial number of which contracted the infection, effectively live out the latent period in isolation. The number of susceptible individuals at each census is described by a sequence of binomial random variables which, until 1971 [28], was not fully appreciated as a sequence that characterises a discrete-time Markov chain. Although events are defined to occur at a concentrated point in time, the classic epidemic models account for a unique type of seasonal variation.

Daley and Gani [20, see Section 4.4] effectively introduced the inhomogeneous Markov chain concept to epidemic modelling by considering the classic chain binomial epidemic models with 'replacement'. Each time interval begins by assuming that all infected individuals emigrate from the system and are *replaced* with an equal number of individuals, a binomial number of which are infected. A period of time then follows where the remaining susceptible population are infected according to another binomial law. The time between one population census and the next effectively consists of two separate periods (conceptually similar to the local extinction and colonisation phases in metapopulation models) so that the number of susceptible individuals at every time step is modelled by an inhomogeneous Markov chain. This model was later applied by Gani and Stals [29] to a viral plant epidemic in a garden nursery.

Clearly, the Daley-Gani and Hill-Caswell models are extensions of the Reed-Frost and Greenwood chain binomial models to a time-inhomogeneous framework. The chain binomial concept has been extended in a slightly different way by others, whereby the three variables of interest in a population of intravenous drug users—the number of susceptible users willing to share needles, the number of infective-susceptible contacts and the number of successful transmissions—are calculated in that order as though following a seasonal dynamic and with each being characterised by binomial distributions (see [31] and modifications by [30]). The underlying discrete-time Markov chain is not inhomogeneous however the model formulation still accounts for season variation.

1.2.3 Further developments

Inhomogeneous Markov chain population models continued to develop by modelling the seasonal processes in different ways. For example, the metapopulation paradigm for modelling a single population of female individuals with a fixed population ceiling has been adapted by both Tenhumberg et. al. [88] and Rout et. al. [84] for populations of Arabian Oryx (*Oryx leucoryx*) and bridled nailtail wallaby (*Onychogalea fraenata*), respectively. Females are assumed to produce either a maximum of one offspring [88] or multiple offspring [84], and recursive formulas are used to calculate the total number of female offspring produced in each breeding (colonisation) season.

Butterflies have also received some attention. Zhang, Liu and Xu [97] applied Day and Possingham's vector-state model to two species of butterfly (the Marsh Fritillary, *Euphydryas aurinia*, and the Knapweed Fritillary, *Melitaea phoebe*) where empirical studies supported the assumed seasonal dynamic. In particular, the incidence of *E. aurinia* was known to depend on the height of host plants in habitat patches while *M. phoebe* was sensitive to the density of host plants. These results were used to modify expressions for patch-specific event probabilities.

More recently, an analytical study of a vector-state metapopulation model with patch-specific extinction probabilities and a general colonisation probability that depends on the current proportion of occupied patches was presented by McVinish and Pollett [64]. The limiting behaviour of the model was investigated by letting the total number of habitat patches grow to infinity. The metapopulation models presented in this thesis are defined with scalar states and the limiting behaviour of these models are investigated in the same limit.

Although some inhomogeneous Markov chain models have been developed for epidemics, the approach clearly dominates in the applied metapopulation literature. In general, discrete-time Markov chain models are perhaps more popular in the ecological context due to efficient parameter estimation techniques with patch occupancy data [65, 46] and their applicability to optimal decision making techniques such as stochastic dynamic programming [88, 84, 76], although it may also be due to a misconception that discrete-time models are needed if populations are observed at discrete time points [79, 80]. Nonetheless, inhomogeneous Markov chains are particularly relevant for studying real metapopulations that exhibit seasonal behaviour.

1.3 Motivation

In order to develop population models that account for seasonal variation, researchers must necessarily model each part of the seasonal dynamic separately. For example, local population growth and dispersal were modelled separately within the structured population approach while the processes of local extinction and colonisation were modelled separately by the Markov chain approach. The seasonal periods were modelled in various ways depending on a particular local dynamic or colonising behaviour under investigation. Despite the simplicity and the growing popularity of the Markov chain approach, numerical methods and simulation are generally used for model analysis and typically in the EC case only [22, 46, 88, 84, 97]; the only exception mentioned so far being [64]. The analysis of structured population models is generally more explicit.

The chain binomial structure exhibited by the Daley-Gani epidemic model and the Hill-Caswell metapopulation model is perhaps the most likely characteristic that will accommodate analytical studies of scalar-state Markov chain models. Indeed this structure instigated the research presented in this thesis, where the research began by developing a chain binomial metapopulation model that described the colonisation process in a general way. The simplest case proved to be tractable and the model analysis for this case was published in 2009 [16] (later to appear in *Ecological Modelling* [14]). Of particular importance is that the analysis completely characterised both EC *and* CE monitoring schemes. In this thesis, I expand on the results presented in [16] and extend the framework to include infinite-patch systems. The following section gives an outline of the thesis, where I detail the types of models to be presented.

1.4 Outline of the thesis

This thesis concerns population models that account for a particular type of dynamic: seasonal variation. Chapter 1 has introduced the concept of seasonal variation and reviewed mathematical models that account for such behaviour. The inhomogeneous Markov chain approach is particularly popular with the applied metapopulation community, however these studies lack extensive analytical treatments. The research presented in later chapters adapts the inhomogeneous Markov chain approach to a range of metapopulation structures but first, Chapter 2 presents basic theory on stochastic discrete-time processes and difference equations necessary to the research. The material presented in this thesis is joint work with P.K. Pollett and is presented in the remaining chapters as follows.

The metapopulation models presented in Chapters 3 and 4 are based on the chain binomial structure developed by Daley and Gani [20] and by Hill and Caswell [46]. In turn, each chapter investigates a special case of the model introduced in [16]. First, Chapter 3 examines the state-independent case called the 'mainland model', the simplest and most tractable chain binomial metapopulation model. A number of explicit expressions for quantities of interest are presented, including conditional state distributions, limiting distributions and expected first hitting times. Second, Chapter 4 examines the more general state-dependent case which accommodates island-type and mainland-island-type metapopulation structures. The difference between the structures simply corresponds to the nature of the colonising process within the patch (island) network and an outside source population or 'mainland'. The state-dependent case is less tractable however a number of analytical results are presented.

In Chapter 5, three different metapopulation networks with infinitely-many habitat patches are considered. The models that describe each network account for a local extinction-colonisation dynamic, as before, where the local extinction process is governed by a binomial law and the colonisation process is governed by a Poisson law. The Poisson laws are derived as the asymptotic equivalent of binomial distributions used previously, and so the resulting infinite-patch models may be thought of as the natural infinite-patch analogues of our finite-patch chain binomial metapopulation models. Indeed, the conditional state distribution is shown to be that which describes a (discrete-time) branching process in each case and standard branching theory is then used to describe the long-term behaviour of the models.

The investigations carried out in Chapters 3, 4 and 5 examine the finite-patch and infinite-patch models in terms of the time-homogeneous EC and CE monitoring schemes, where some of these models are shown to be highly tractable. In Chapter 6, these models are re-examined in terms of the underlying time-inhomogeneous Markov chain. In particular, the chapter presents a theoretical framework for Markov chains that share the salient features of our metapopulation models, where limit theorems such as a law of large numbers and a central limit law are established in a general setting. Finite-state and infinite-state Markov chains are treated separately and it is these results that are used to examine the finite-patch and infinite-patch models, respectively. I demonstrate that the limiting behaviour of the metapopulation models can be evaluated explicitly and even in situations where the conditional state distribution could not be evaluated.

Finally, in Chapter 7, I conclude by summarising the results of the research and discussing possibilities for future work.

Chapter 2

Theory

This chapter provides the necessary background on stochastic processes, including discrete-time Markov chains, branching processes, autoregressive processes and Gaussian processes. Prior knowledge of commonly used random variables, their distribution functions and their probability generating functions, is assumed. Also included is theory on difference equations and a list of notational conventions.

2.1 Introduction

A stochastic process $(X_t : t \in T, X_t \in S)$ is a collection of random variables that are indexed by a set T and take values in a set S. The variable t usually represents time and so it may be that $T = \{0, 1, 2, \ldots\}$ for a discrete-time setting or $T = [0, \infty)$ for a continuous-time setting. The set S is called the *state space* and this set may be countable (discrete) or non-countable (continuous). We will be completely concerned with stochastic processes that evolve in discrete-time and so the time variable t shall henceforth take values in $\{0, 1, 2, \ldots\}$.

We are particularly interested in stochastic processes that satisfy the Markov Property (so named after the Russian mathematician A. A. Markov [10] for his famous work on these processes).

Definition 2.1 *A stochastic discrete-time process* $(X_t : t \geq 0)$ *is said to satisfy the Markov Property if*

$$\Pr(X_{t+1} = j | X_t = i_t, \ldots, X_1 = i_1, X_0 = i_0) \;=\; \Pr(X_{t+1} = j | X_t = i_t) \tag{2.1}$$

for all $t \geq 0$ *and all* $i_0, i_1, \ldots, i_t, j \in S$.

Thus, conditioned on the present state, the future state of a process that satisfies the Markov Property is independent of past values. For the case where S is discrete, the process is called a *discrete-time Markov chain*. Later, we will discuss a special class of these chains called *branching processes*. For the case where S is continuous, we are interested in one special class of processes called *Gaussian Markov processes*. We further classify these processes according to the time-dependence in equation (2.1).

Definition 2.2 *Suppose that the stochastic discrete-time process* $(X_t : t \geq 0)$ *satisfies the Markov Property. This process is then called time-homogeneous if it satisfies*

$$\Pr(X_{t+1} = j | X_t = i) \ = \ \Pr(X_1 = j | X_0 = i)$$

for all $t \geq 0$ *and all* $i, j \in S$*, otherwise the process is called time-inhomogeneous.*

Although the Markov Property simplifies the mathematical structure of the process, we note that it is a biologically sensible property for the population models that will be presented in later chapters.

In addition, we will also be concerned with deterministic processes that are described by difference equations, since these equations frequently arise in later chapters as deterministic approximations of discrete-time Markov chains.

The following sections present theory as it relates to the stochastic and deterministic processes mentioned above. Results are stated without proof, however the reader can find proofs in [35, 66] (discrete-time Markov chains), [8] (branching processes), [26] (Gaussian Markov processes) and [47, 85] (difference equations), unless otherwise stated. For those results where proof is given in only one of two quoted sources, the result is referenced accordingly. The chapter concludes by listing the notational conventions used in this thesis.

2.2 Discrete-time Markov chains

As the name suggests, a discrete-time Markov chain is a process that evolves in discrete time, satisfies the Markov Property and has a discrete state space. The state space S may be finite or countably infinite in the following discussion, and the distinction will be made clear where necessary.

2.2.1 Preliminary notation and classifications

Transition probabilities

Let $(X_t : t \geq 0)$ be a discrete-time Markov chain on the set S, and let this chain be homogeneous in time. We may think of this chain of random variables as marking the state of a notional particle at times $t = 0, 1, 2, \ldots,$ where the particle is said to make a transition, or 'jump', at these times. The *transition probabilities* that govern these jumps are the elements of the *transition matrix* $P = (p_{ij} : i, j \in S)$, where $p_{ij} = \Pr(X_{t+1} = j | X_t = i)$. Similarly, the *t-step transition probabilities* are the elements of the *t-step transition matrix* $P^{(t)} = \big(p_{ij}^{(t)} : i, j \in S\big)$, where $p_{ij}^{(t)} = \Pr(X_t = j | X_0 = i)$. These matrices are *stochastic matrices*, which is to say that their entries are non-negative and their row sums equal 1. For the $t = 1$ case, we have that $P^{(1)} = P$, and for the $t = 0$ case, we use the convention that $p_{ij}^{(0)} = \delta_{ij}$ (the Kronecker delta) so that $P^{(0)} = I = P^0$ (the identity matrix, I). Indeed, we have the following simple result.

Proposition 2.1 *The t-step transition matrix is equal to t-th power of P. That is,* $P^{(t)} = P^t$ $(t \geq 0)$, *and thus* $p_{ij}^{(t)}$ *is equal to the* (i, j)*-th entry of* P^t.

Letting $p^{(t)} = \big(p_j^{(t)} : j \in S\big)$ denote the distribution of X_t, that is, $p_j^{(t)} = \Pr(X_t = j)$, it follows from Proposition 2.1 that $p^{(t)} = p^{(0)} P^t$ $(t \geq 0)$.

Recurrent states and transient states

The number of times the particle visits a particular state depends on whether the state is *recurrent* or *transient*.

Definition 2.3 *State i is called recurrent if* $\Pr(X_t = i \text{ for some } t \geq 1 | X_0 = i) = 1$, *otherwise it is called transient.*

Having started in a recurrent state i, the particle will return to this state with probability 1 and return to this state infinitely often. A transient state is visited only a finite number of times. Before we can describe recurrence and transience in terms of the transition probabilities, we define the generating functions

$$P_{ij}(z) = \sum_{t=0}^{\infty} z^t p_{ij}^{(t)} \quad \text{and} \quad F_{ij}(z) = \sum_{t=0}^{\infty} z^t f_{ij}^{(t)}, \tag{2.2}$$

where $p_{ij}^{(t)}$ is the t-step transition probability, as before, and

$$f_{ij}^{(t)} = \Pr(X_t = j, X_{t-1} \neq j, \dots, X_2 \neq j, X_1 \neq j, |X_0 = i)$$

is the probability that the first visit to state j, having started in state i, occurs at time t. By convention, we set $f_{ij}^{(0)} = 0$ for all i and j.

Theorem 2.1 ([35])

(a) $P_{ii}(z) = 1 + F_{ii}(z)P_{ii}(z)$,

(b) $P_{ij}(z) = F_{ij}(z)P_{jj}(z)$ *if* $i \neq j$.

Using this result, we can describe recurrence and transience in terms of $\left(p_{ij}^{(t)} : t \geq 0\right)$ as follows. Note that the quantity $f_{ij} = \sum_{t=0}^{\infty} f_{ij}^{(t)}$ is the probability that, having started in state i, the particle ever visits state j.

Corollary 2.1

(a) If $\sum_{t=0}^{\infty} p_{jj}^{(t)} = \infty$, *then state* j *is recurrent and* $\sum_{t=0}^{\infty} p_{ij}^{(t)} = \infty$ *for all* i *such that* $f_{ij} > 0$.

(b) If $\sum_{t=0}^{\infty} p_{jj}^{(t)} < \infty$, *then state* j *is transient and* $\sum_{t=0}^{\infty} p_{ij}^{(t)} < \infty$ *for all* i.

Corollary 2.2 *For a transient state* j, *we have that* $p_{ij}^{(t)} \to 0$ *as* $t \to \infty$ *for all* i.

Of course, in terms of the quantity $f_{jj} = \sum_{t=0}^{\infty} f_{jj}^{(t)}$, state j is recurrent if and only if $f_{jj} = 1$ and state j transient if and only if $f_{jj} < 1$.

Mean recurrence time

Let $T_i = \inf\{t \geq 1 : X_t = i\}$ denote the *first passage time* to state i, noting that the infimum of the empty set is equal to ∞. We are particularly interested in the special case where $T_i = \infty$ since this case implies that the process never reaches state i. For example, $\Pr(T_i = \infty | X_0 = i) > 0$ if and only if i is transient, and therefore $\mathbb{E}(T_i | X_0 = i) = \infty$ in this case. Noting that $f_{ii}^{(t)} = \Pr(T_i = t | X_0 = i)$, we present the following definition.

Definition 2.4 *The quantity* $\mu_i = \mathbb{E}(T_i | X_0 = i)$ *is called the mean recurrence time of state* i. *We have that* $\mu_i = \sum_{t=0}^{\infty} t f_{ii}^{(t)}$ *if* i *is recurrent and* $\mu_i = \infty$ *if* i *is transient.*

It may be that $\mu_i = \infty$ for a recurrent state i and so we present another definition.

Definition 2.5 *A recurrent state i is called*

(a) null-recurrent if $\mu_i = \infty$, or

(b) positive-recurrent if $\mu_i < \infty$.

The following result concerns the transition probabilities.

Theorem 2.2 *For a null-recurrent state i, we have that $p_{ii}^{(t)} \to 0$ as $t \to \infty$ and, as a result, $p_{ji}^{(t)} \to 0$ for all j.*

The period of a state

Another property associated with a state is its *period*.

Definition 2.6 *The period $d(i)$ of state i is the greatest common divisor of the set of possible return times to state i. State i is said to be periodic if $d(i) > 1$ or aperiodic if $d(i) = 1$.*

Therefore state i is aperiodic if the chain can return in $1, 2, 3, \ldots$ time-steps, since the greatest common divisor of these epochs is 1. We need only concern ourselves with aperiodic states.

Class structure

We begin with the following definition.

Definition 2.7 *State j is said to be accessible from state i (written as $i \to j$) if $p_{ij}^{(t)} > 0$ for some $t \geq 0$. If both $i \to j$ and $j \to i$, then i and j are said to communicate (written as $i \leftrightarrow j$).*

Recalling that $p_{ii}^{(0)} = 1$, then the $t = 0$ case implies that $i \leftrightarrow i$ for all $i \in S$. If $i \leftrightarrow j$ and $j \leftrightarrow k$, it is a simple matter to show that $i \leftrightarrow k$. Indeed, the relation $\leftrightarrow$ is an equivalence relation on the set S and S therefore partitions into one or more disjoint equivalence classes (sets) of communicating states.

Theorem 2.3 *If $i \leftrightarrow j$, then*

(a) i is recurrent (transient) if and only if j is recurrent (transient),

(b) i and j have the same period.

We can therefore talk of a recurrent (or transient) class and the period of a class.

Definition 2.8 *If the state space S forms a single communicating class, then the chain itself is called irreducible.*

A class may have the additional property that it is closed.

Definition 2.9 *A class C is called closed if $p_{ij} = 0$ for all $i \in C$ and $j \notin C$. If a closed class contains a single state i, then i is called an absorbing state.*

That is, if the particle jumps into a closed class, it remains in that class for ever.

Corollary 2.3 *Every recurrent class is closed.*

Corollary 2.4 *Every finite closed class is recurrent.*

A countably infinite closed class is not necessarily recurrent; it may be transient.

Corollary 2.5 *For the special case where S is a finite set, at least one state is recurrent and all recurrent states are positive-recurrent.*

Therefore, if S is finite and irreducible, the chain is positive-recurrent. Finally, we summarise the difference between finite and countably infinite state spaces by using the following result.

Theorem 2.4 ([35]) *The representation $S = Q \cup C_1 \cup C_2 \cup \ldots$ is a unique partition of the state space S, where Q is the set of transient states and $C_1, C_2, \ldots$ are closed classes of recurrent states.*

If the chain begins in one of the recurrent classes $C_1, C_2, \ldots$, then the chain will remain in that particular class for ever. If the chain begins in the set of transient states Q, then the size of Q is important: if Q is finite, the chain will eventually transition into one of the recurrent classes, while if Q is countably infinite, the chain may either transition into a recurrent class or remain in Q forever.

2.2.2 Stationary distributions

The previous section defined many of the important concepts associated with discrete-time Markov chains. Here, and in the following section, we consider some important distributions of these processes.

We begin by defining the term *stationary distribution*.

Definition 2.10 *Let* $\pi = (\pi_j : j \in S)$ *be a vector such that* $\pi_j \geq 0$ *for all* j *and* $\sum_{j \in S} \pi_j = 1$. *If this vector satisfies* $\pi = \pi P$, *then* π *is called a stationary distribution of the chain.*

The term 'stationary' is used since, if X_0 has distribution π, then X_t has distribution π for all $t \geq 1$. For irreducible chains in particular, we have the following result.

Theorem 2.5 *An irreducible chain has a stationary distribution* π *if and only if all the states are positive-recurrent. Furthermore, the stationary distribution* π *of an irreducible positive-recurrent chain is unique and its entries* $(\pi_j : j \in S)$ *are given by* $\pi_j = 1/\mu_j$ *for all* j.

Thus transient chains and null-recurrent chains do not have stationary distributions while a (unique) stationary distribution π exists for positive-recurrent chains. For the case where an irreducible chain is aperiodic, the following result describes the limiting ($t \to \infty$) behaviour of the transition probabilities and provides an important connection with stationary distributions.

Theorem 2.6 *For an irreducible aperiodic chain,* $\lim_{t\to\infty} p_{ij}^{(t)} = 1/\mu_j$ *for all* i *and* j.

If state j is transient or null-recurrent, then $\mu_j = \infty$ and it follows that $p_{ij}^{(t)} \to 0$; Theorem 2.6 clearly agrees with Corollary 2.2 and Theorem 2.2. For positive-recurrent and aperiodic chains, Theorem 2.5 together with Theorem 2.6 tell us that $p_{ij}^{(t)} \to \pi_j = 1/\mu_j$, where π_j is that given by the unique stationary distribution. Thus a proper limiting distribution exists in the positive-recurrent case and it is equal to the unique stationary distribution of the chain.

Remark. A number of irreducible aperiodic chains will be constructed in later chapters, where some are defined on finite state spaces while others are defined on countably infinite state spaces. The finite chains are positive-recurrent (Corollary 2.5) and so have a unique stationary (and limiting) distribution. The countably infinite chains cannot not be classified so easily because it is not immediately obvious whether the chains are transient, null-recurrent or positive-recurrent. In chapter 5, we draw upon branching theory (section 2.3) to determine conditions under which a proper limiting distribution exists for countably infinite chains.

2.2.3 Quasi-stationary distributions

Consider a discrete-time Markov chain $(X_t : t \geq 0)$ that takes values in $S = \{0\} \cup Q$, where state 0 is an absorbing state and Q is a communicating class of transient states from which 0 is accessible. Of course, if the process begins in state 0, it will remain in that state forever. It is not surprising, then, that there exists a unique stationary distribution π given by $\pi_0 = 1$ and $\pi_j = 0$ $(j \in Q)$. However, stationary distributions are not particularly useful in the context of absorbing chains since they do not say anything about the behaviour of the process before absorption occurs. Instead, it is more useful to study the process *conditional* on non-absorption and talk of stationary conditional distributions, or, as we call them, *quasi-stationary distributions.*

Suppose that our Markov chain reaches the absorbing state with probability 1. Recalling that $p^{(t)} = (p_j^{(t)} : j \in S)$ denotes the (unconditional) distribution of the chain at time t, we let $d^{(t)} = (d_j^{(t)} : j \in Q)$ denote the distribution at time t *conditional* on non-absorption, that is,

$$d_j^{(t)} = \Pr(X_t = j | X_t \neq 0) = \frac{\Pr(X_t = j, X_t \neq 0)}{\Pr(X_t \neq 0)} = \frac{p_j^{(t)}}{1 - p_0^{(t)}} .$$

We now state the following definition.

Definition 2.11 *Let $v = (v_j : j \in Q)$ be a distribution such that $v_j \geq 0$ and $\sum_{j \in Q} v_j = 1$. If this vector satisfies $vP_Q = \rho v$, where P_Q is the transition matrix P restricted to Q and ρ is a real constant, then v is called a quasi-stationary distribution of the chain.*

The term ‘quasi-stationary’ is appropriate since, if X_0 has conditional distribution v, then X_t has conditional distribution v for all $t \geq 1$. Darroch and Seneta [21] proved the following result for finite chains.

Theorem 2.7 *For the case where Q is finite and aperiodic, the chain has a unique quasi-stationary distribution v. There also exists a proper limiting conditional distribution in this case, where $\lim_{t \to \infty} d_j^{(t)} = v_j$ and v_j is that given by the quasi-stationary distribution.*

Thus the limiting conditional distribution is equal to the (unique) quasi-stationary distribution in the finite case. Indeed, it follows from the theory of non-negative matrices (see for example Seneta [86]) that v is the normalised left-eigenvector of P_Q corresponding to the maximal real eigenvalue of P_Q (the 'Perron-Frobenius' eigenvalue). The corresponding result for countably infinite and aperiodic chains was proved by Seneta and Vere-Jones (see Theorem 3.1 of [87]): assuming that absorption occurs with probability 1, then, if Q is R-positive and $R > 1$ (the concept of R-positivity will be defined shortly), the chain has a unique quasi-stationary distribution and there exists a proper limiting conditional distribution which is equal to the quasi-stationary distribution. Seneta and Vere-Jones also proved a similar result that takes absorption from a particular initial state into account (see Theorem 3.2 of [87]). For the matrix Q to be R-positive, one must show that $\lim_{t\to\infty} p_{ij}^{(t)} R^t > 0$, where R is the common radius of convergence of the generating functions $P_{ij}(z) = \sum_{t\to\infty} z^t p_{ij}^{(t)}$ $(i, j \in Q)$. These conditions hold automatically in the finite case but are, in general, difficult to establish for infinite matrices. We do not attempt to perform such calculations in this thesis and instead draw upon branching theory (section 2.3) to determine conditions under which a proper limiting conditional distribution exists.

Remark. In later chapters, we shall examine a number of chains for which $S = \{0\} \cup Q$. State 0 will always represent total extinction of a population, however this state will be absorbing when we assume that immigration from an external source is not possible.

2.2.4 Time-inhomogeneous chains

The discussion so far concerns discrete-time Markov chains that satisfy the property of time-homogeneity. Here we consider a particular type of inhomogeneous chain and demonstrate how it can be interpreted as a set of time-homogeneous chains.

Let $(X_t : t \geq 0)$ be a discrete-time Markov chain on the set S, where S may be either finite or countably infinite. For each time t, let $P(t, t+1) = (p_{ij}(t, t+1) : i, j \in S)$ now denote the 1-step transition matrix and $p_{ij}(t, t+1) = \Pr(X_{t+1} = j | X_t = i)$ the 1-step transition probability. If there exists a integer $M > 1$ such that

$$p_{ij}(t, t+1) = p_{ij}(t+M, t+M+1) \quad \text{for all } i, j \in S \text{ and for all } t \geq 0,$$

in which case

$$P(t, t+1) = P(t+M, t+M+1) \quad \text{for all } t \geq 0,$$

then the chain effectively cycles through M transition matrices over time. If $M = 2$, for example, then $P(0,1) = P(2,3) = P(4,5) = \ldots$ and $P(1,2) = P(3,4) = P(5,6) = \ldots$, and so there are two distinct transition matrices that govern the evolution of the chain. Furthermore, the sequences $(X_t : t = 0, 2, 4, \dots)$ and $(X_t : t = 1, 3, 5, \dots)$ define two time-homogeneous Markov chains whose transition matrices are given by $P = P(0,1) \times P(1,2)$ and $P = P(1,2) \times P(0,1)$, respectively. An inhomogeneous chain with 'period' M may therefore be observed in a homogeneous manner in any one of M ways, and we can therefore describe the behaviour of the homogeneous processes by using theory that was presented earlier.

Remark. We shall be using inhomogeneous Markov chains with $M = 2$ to model metapopulation dynamics with two distinct phases, namely local extinction and colonisation. The transition matrix of the time-homogeneous chain is either $P = EC$ or $P = CE$, depending on the timing of a the population census; E is the transition matrix that governs the local extinction phase and C is the transition matrix that governs the colonisation phase.

2.2.5 Simulating discrete-time Markov chains

The simulations and figures included in this thesis were produced using MATLAB® R2008a (version 7.6.0324). Included here is the algorithm used for simulating discrete-time Markov chains on a finite state space.

1. Initialise the transition matrix $P = (p_{ij} : i, j \in S)$, where S is some finite set. Suppose the states are labelled $1, 2, \dots, |S|$.

2. Evaluate the matrix $U = (u_{ij} : i, j \in S)$, where $u_{ij} = \sum_{k=1}^{j} p_{ik}$. Let u_i be the vector that denotes the i-th row of U.

3. Set $t = 0$ and generate the initial value X_0.

4. Increment t by 1. Generate the random variable X_t from the $u_{X_{t-1}}$ distribution.

5. Repeat step 4 up to some time horizon T. The vector $(X_0, X_1, \dots, X_T)$ is a simulated realisation of the chain.

For the case where S is countably infinite, steps 1 and 2 of the algorithm were ignored. At step 4, random variables were generated from binomial and Poisson distributions according to the structure defined by equation (6.26).

2.3 Branching processes

A *branching process* is a stochastic process that evolves in terms of sums of discrete random variables. These processes may be formulated in continuous time or discrete time, however we shall be concerned with discrete-time formulations. In particular, we examine the simplest type of discrete-time branching process, called a *Galton-Watson process* (so named after F. Galton's and H. W. Watson's famous efforts to calculate the probability of the extinction of family names [90]), and one of its variants, the *Galton-Watson-Immigration process*. In this context, we no longer visualise a notional particle jumping around on a set of states, but an evolving population of particles whose generations follow a discrete-time structure.

2.3.1 The Galton-Watson process

Imagine a population of particles that evolve between discrete time intervals. During any given time interval, suppose that the particles do not interact but that each particle splits into (or is otherwise replaced by) a random number ξ of offspring particles. Each particle in each generation produces offspring according to the same probability law, called the *offspring distribution*, and this distribution is denoted by $p = (p_j : j = 0, 1, 2, \ldots)$, where $p_j = \Pr(\xi = j)$. We shall henceforth assume that $p_0 > 0$ and $0 < p_0 + p_1 < 1$ to avoid trivial cases. If $\xi = 0$, then the particle has simply died without producing any offspring. Letting X_t denote the population size at time t, we have that

$$X_{t+1} = \xi_1^{(t)} + \xi_2^{(t)} + \cdots + \xi_{X_t}^{(t)} \qquad (t \geq 0), \tag{2.3}$$

where $\xi_1^{(t)}, \xi_2^{(t)}, \ldots$ are independent and identically distributed (iid) random variables with common law p; the term $\xi_j^{(t)}$ being the number of offspring produced by the j-th particle of the t-th generation. The sequence $(X_t : t \geq 0)$ defined by (2.3) is called a *Galton-Watson process*. This process can be modified in various ways to account for interactions between particles, overlapping generations and even multi-type populations (for examples see [38]), however we shall examine one particular modification in section 2.3.2 that accounts for immigrating particles.

The discussion above shows that a Galton-Watson process takes values in $S = \{0, 1, 2, \ldots\}$, satisfies the Markov Property and is homogeneous in time, and is therefore a particular type of discrete-time Markov chain defined on a countably infinite state space. We thus interpret the transition probability p_{ij} as the probability that j offspring particles are produced from i particles in one time interval, where state 0 represents population extinction. Indeed, state 0 is absorbing, the remaining states are transient and 0 is accessible from the transient class. However, instead of using the transition matrix $P = (p_{ij} : i, j \in S)$ to analyse the behaviour of the process, it is much easier to use probability generating functions. Let

$$g(z) = \sum_j z^j p_j \qquad \big(|z| \leq 1\big)$$

be the probability generating function (pgf) of the offspring distribution. Since the numbers of offspring on the right-hand side of (2.3) are iid, we can write $\mathbb{E}\big(z^{X_{t+1}}|X_t = i\big) = \sum_j z^j p_{ij} = g(z)^i$ for all $i \geq 1$ and $t \geq 0$. Letting $g_t(z)$ denote the t-th functional iterate of g, it is a simple matter to show that

$$\mathbb{E}\big(z^{X_t}|X_0 = 1\big) = \sum_{j=0}^{\infty} z^j p_{1j}^{(t)} = g_t(z)$$

and

$$\mathbb{E}\big(z^{X_t}|X_0 = i\big) = \sum_{j=0}^{\infty} z^j p_{ij}^{(t)} = [g_t(z)]^i,$$

where $p_{ij}^{(t)}$ denotes the t-step transition probability, as before. The last expression shows that the process is effectively made up of i independent branches, where one branch stems from one of the i individuals in the 0-th generation. Let μ and σ^2 denote the mean and variance of the offspring distribution, respectively. Again, it is a simple matter to show that $\mathbb{E}(X_t|X_0) = X_0\mu^t$ and

$$\operatorname{Var}(X_t|X_0) = \begin{cases} X_0\sigma^2 t & \text{if } \mu = 1 \\ X_0\sigma^2(\mu^t - 1)\mu^{t-1}/(\mu - 1) & \text{if } \mu \neq 1. \end{cases} \tag{2.4}$$

It follows that $\mathbb{E}(X_t) = \mathbb{E}(X_0)\mu^t$ and so, as $t \to \infty$, the average population size decreases to zero if $\mu < 1$, remains constant if $\mu = 1$ or grows to infinity if $\mu > 1$. The process is called *subcritical*, *critical* or *supercritical* according to whether $\mu < 1$, $\mu = 1$ or $\mu > 1$, respectively.

An important quantity associated with Galton-Watson processes is η, the probability that a population starting with one particle ever goes extinct. If the process began with i particles, then the extinction probability is η^i. We have the following result.

Theorem 2.8 *The extinction probability* η *of a Galton-Watson process is the smallest non-negative root of the equation* $z = g(z)$. *We have that* $\eta = 1$ *if* $\mu \leq 1$ *and* $\eta < 1$ *if* $\mu > 1$.

Thus the population is certain to go extinct if the mean number of offspring produced by each particle is less than or equal to 1. If the mean of the offspring distribution is greater than 1, then there is a chance the population will not go extinct. Indeed, the trivial case where $\eta = 0$ does not apply here; our assumption that $p_0 > 0$ implies that $\eta > 0$. In the subcritical case, we have Yaglom's [95] famous result.

Theorem 2.9 (Yaglom's Theorem) *If* $\mu < 1$, *then* $\lim_{t\to\infty} \Pr(X_t = j | X_t > 0) = b_j$ *exists and* $b = (b_j : j = 1, 2, \dots)$ *defines a proper distribution (a limiting conditional distribution); the generating function* $B(z) = \sum_j z^j b_j$ *satisfies* $B(g(z)) = \mu B(z) + (1 - \mu)$.

The next theorem provides more general results. Note that $g'(z)$ denotes the derivative of g with respect to z.

Theorem 2.10 *Let* T *represent the extinction time of a Galton-Watson process.*

(a) If $\mu = 1$, *then* $\lim_{t\to\infty} \Pr(X_t = j | t < T < \infty) = 0$.

(b) If $\mu \neq 1$, *then* $\lim_{t\to\infty} \Pr(X_t = j | t < T < \infty) = b_j$ *exists and* $b = (b_j : j = 1, 2, \dots)$ *defines a proper distribution (a limiting conditional distribution). The generating function* $B(z) = \sum_j z^j b_j$ *is the unique solution of*

$$B\left(\frac{g(z\eta)}{\eta}\right) = g'(\eta)B(z) + 1 - g'(\eta). \tag{2.5}$$

Notice that Theorem 2.10 conditions on the event that absorption ultimately occurs ($T < \infty$) and that part (b) applies to both subcritical and supercritical processes. Yaglom's Theorem only accounts for the subcritical case, where $\eta = 1$ and $\Pr(X_t = j | X_t > 0) \equiv \Pr(X_t = j | t < T < \infty)$. Although the limiting conditional distributions exist in the subcritical and supercritical cases, it is generally a difficult task to solve (2.5) explicitly for this distribution.

2.3.2 The Galton-Watson process with immigration

Consider the population of particles described above and suppose that, for each t, a random number Y_t of immigrant particles join the population during the t-th interval. Further suppose that the immigrant population size is independent of the current population size and that the immigrating particles produce offspring according to the same law p as the resident population. Let $h(z)$ be the pgf of the immigrant population size and suppose that $h(0) < 1$, in order to avoid trivial cases. We have that

$$X_{t+1} = \xi_1^{(t)} + \xi_2^{(t)} + \cdots + \xi_{X_t}^{(t)} + Y_{t+1} \qquad (t \geq 1, X_0 = Y_0), \tag{2.6}$$

where $\xi_1^{(t)}, \xi_2^{(t)}, \ldots$ are iid with common pgf $g(z)$, as before, and $Y_0, Y_1, \ldots$ are iid with pgf $h(z)$. The sequence $(X_t : t \geq 0)$ defined by (2.6) is thus called a *Galton-Watson-Immigration process.* Again, we can describe the process in terms of generating functions. Equation (2.6) tells us that $\mathbb{E}\big(z^{X_{t+1}} | X_t = i\big) = g(z)^i h(z)$ for all $i \geq 0$ and $t \geq 0$ and, upon letting $G_t(z) = \mathbb{E}\big(z^{X_t} | X_0 = 1\big)$ denote the pgf of X_t conditioned on there being one particle initially, we have that

$$G_{t+1}(z) = G_t(g(z))h(z) \qquad (t \geq 0)$$

and, hence, $G_t(z) = g_t(z) \prod_{k=0}^{t-1} h(g_k(z))$. The following result provides conditions under which a proper limiting distribution exists.

Theorem 2.11

(a) If $\mu > 1$, or $\mu = 1$ and $\sigma^2 < \infty$, then $\lim_{t\to\infty} \Pr(X_t = j) = 0$ $(0 < j < \infty)$.

(b) If $0 < h'(1) < \infty$ and $\mu < 1$, then $\lim_{t\to\infty} \Pr(X_t = j) = d_j$ exists and $d = (d_j : j \in S)$ defines a proper distribution (limiting distribution).

Notice how part (b) compares with Yaglom's Theorem (Theorem 2.9): if $\mu < 1$, then a proper limiting conditional distribution exists for a Galton-Watson process and a proper limiting distribution exists for a Galton-Watson-Immigration process. The following more general result was proved by Foster and Williamson [27].

Theorem 2.12 *Consider a Galton-Watson-Immigration process $(X_t : t \geq 0)$ with $\mu \leq 1$. Then a proper limiting distribution exists if and only if*

$$\int_0^1 \frac{1 - h(s)}{g(s) - s} ds < \infty \, .$$

2.4 Gaussian Markov processes

This section concerns a special type of *Gaussian process* called a *Gaussian Markov process*. Gaussian processes take values in continuous state spaces and are formulated in either discrete-time or continuous-time, however we shall, once again, be completely concerned with discrete-time formulations. It will soon become clear that Gaussian processes are described in terms of multivariate normal (Gaussian) distributions and that Gaussian Markov processes have an autoregressive structure. We therefore begin by defining the term *autoregressive process* and then follow with more formal definitions concerning Gaussian processes.

2.4.1 Autoregressive processes

Autoregressive processes are simple stochastic processes widely used to model and analyse time series data (see for example Chatfield [17]).

Definition 2.12 *A stochastic discrete-time process* $(X_t : t \geq 0)$ *is called an autoregressive process of order* p*, or an AR-*p *process, if*

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \cdots + a_p X_{t-p} + E_t \qquad (t \geq 1),$$

where $a_1, \ldots, a_p$ *are constants and* (E_t) *are iid random variables with mean* 0 *and variances* σ^2.

The definition clearly shows that the current value of an AR-p process is a weighted linear sum of the previous p values, and we may therefore think of the random variable X_t as being 'regressed' on $X_{t-1}, \ldots, X_{t-p}$. The random variable E_t is commonly referred to as the "error" and σ^2 as the "error variance", and we shall adopt these conventions.

The simplest autoregressive process is of course the AR-1 process. The following definition outlines an important extension in this case.

Definition 2.13 *Let* $\boldsymbol{X}_t = (X_{t_1}, \ldots, X_{t_d})^\top$ *be a vector of* d *random variables indexed by* t*. The* *discrete-time stochastic process* $(\boldsymbol{X}_t : t \geq 0)$ *defined by*

$$\begin{aligned}
X_{t_1} &= a_{11}X_{t-1_1} + \cdots + a_{1d}X_{t-1_d} + E_{t_1} \\
X_{t_2} &= a_{21}X_{t-1_1} + \cdots + a_{2d}X_{t-1_d} + E_{t_2} \\
&\vdots \\
X_{t_d} &= a_{d1}X_{t-1_1} + \cdots + a_{dd}X_{t-1_d} + E_{t_d}\,,
\end{aligned}$$

where (a_{ij}) *are constants and* $E_{t_1}, \ldots, E_{t_d}$ *are iid random variables with zero mean and variance* σ^2*, is called a d-variate autoregressive process of order* 1*. In matrix form we write* $\boldsymbol{X}_{t+1} = A\boldsymbol{X}_t + \boldsymbol{E}_t$*, where* $A = (a_{ij})$ *is the* $d \times d$ *matrix of constants and* $\boldsymbol{E}_t = (E_{t_1}, \ldots, E_{t_d})^\top$ *is the vector of errors at time* t*.*

The d-variate AR-1 process is an example of a *multivariate autoregressive process*. Multivariate autoregressive processes are commonly used to model several time series which, for example, may result by measuring several components of a system simultaneously or by measuring a system with a feedback loop several times. In chapter 6, we will show that multivariate AR-1 processes arise due to an underlying deterministic process which has a stable limit cycle (section 2.5).

2.4.2 Gaussian processes

We are now ready to discuss Gaussian processes more formally.

Definition 2.14 *A discrete-time stochastic process* $(X_t : t \geq 0)$ *is called a Gaussian process if the finite-dimensional distributions of the process are Gaussian.*

That is, $(X_t : t \geq 0)$ is a Gaussian process if $(X_{t_1} \ldots X_{t_k}) \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma_k)$ for any finite sequence of times $t_1, \ldots, t_k$, where $\mathrm{N}(\boldsymbol{\mu}, \Sigma_k)$ denotes the multivariate Gaussian distribution with mean vector $\boldsymbol{\mu} = (\mathbb{E}X_{t_1}, \ldots, \mathbb{E}X_{t_k})$ and covariance matrix $\Sigma_k = (\sigma_{ij})$, $\sigma_{ij} = \mathrm{Cov}(X_{t_i}, X_{t_j})$. A Gaussian process is thus characterised by its mean $\mu_t = \mathbb{E}(X_t)$ and covariance structure.

The following result outlines the conditions for which a Gaussian process satisfies the Markov Property. Note that $\rho(X_s, X_t) = \mathrm{Cov}(X_s, X_t)\big(\mathrm{Var}\, X_s\ \mathrm{Var}\, X_t\big)^{-1/2}$ denotes the correlation of X_s and X_t.

Theorem 2.13 *Let $(X_t : t \geq 0)$ be a Gaussian process. Then the following two conditions are necessary and sufficient for this process to satisfy the Markov Property.*

(i) For any finite sequence of times $t_1, \ldots, t_k$, $\mathbb{E}(X_{t_k} | X_{t_1} \ldots X_{t_{k-1}}) = \mathbb{E}(X_{t_k} | X_{t_{k-1}})$.

(ii) For $s \leq u \leq t$, the correlations satisfy $\rho(X_s, X_t) = \rho(X_s, X_u)\rho(X_u, X_t)$.

If the two conditions hold, then the process is called a *Gaussian Markov process*. The following more general result states that a Gaussian process is a Gaussian Markov Process if and only if the process is autoregressive.

Theorem 2.14 *A Gaussian process $(X_t : t \geq 0)$ is a Gaussian Markov process if and only if relations of the form*

$$X_0 \sim \mathrm{N}(\mu_0, v_0) \quad \text{and} \quad X_{t+1} = a_t + b_t X_t + E_t \qquad (t \geq 1) \tag{2.7}$$

hold, where (a_t) and (b_t) are both sequences of constants and (E_t) are independent $\mathrm{N}(0, \sigma_t^2)$ random variables with (σ_t^2) being a sequence of positive constants.

Letting $v_t = \operatorname{Var} X_t$, a Gaussian Markov process thus satisfies $a_t = \mu_{t+1} - b_t \mu_t$, $v_{t+1} = b_t^2 v_t + \sigma_t^2$ and, given $X_t = x$, $X_{t+1} \sim \mathrm{N}(a_t + b_t z, \sigma_t^2)$. If $X_0 = x$ is given, then $\mu_0 = x$ and $v_0 = 0$. The covariance structure is evaluated by iterating (2.7).

Remark. The Gaussian Markov processes encountered in chapter 6 have $X_0 = x$ given and $a_t = 0$ for all $t \geq 0$. In this case, we have that

$$\mu_t = x \prod_{u=0}^{t-1} b_u, \quad v_t = \sum_{u=0}^{t-1} \sigma_u^2 \prod_{v=u+1}^{t-1} b_v^2, \quad \operatorname{Cov}(X_t, X_{t+s}) = v_t \prod_{u=t}^{t+s-1} b_u, (s \geq 0),$$

thus explicitly showing the mean and covariance structure of the process. In some cases the sequences (b_t) and (σ_t^2) are independent of the time variable t, and the Gaussian Markov process is then simply referred to as an AR-1 process (Definition 2.12).

2.5 Difference equations

In the previous chapter, we explained that the difference equation model (1.2) is the natural discrete-time analogue of Verhulst's continuous-time population model (1.1). In later chapters, we will encounter several more difference equation models because they arise as deterministic approximations of stochastic discrete-time (Markov chain) population models. Each of these deterministic models exhibit the form

$$x_{t+1} = f(x_t) \qquad (t = 0, 1, 2, \dots), \tag{2.8}$$

where x_t represents the state of the system at time t and f is some function of interest. It is therefore the purpose of this section to present theory as it relates to studying systems described by first-order difference equations.

First we define the fundamental concepts associated with first-order difference equations. In the following discussion, f refers to the function that defines $(x_t : t \geq 0)$ through (2.8).

Definition 2.15 *A number u is called a fixed point of f if $f(u) = u$.*

Definition 2.16 *A number u is called a periodic fixed point of f if there exists an integer $d > 1$ such that $f_d(u) = u$, where f_d is the the d-th functional iterate of f. If such an integer exists, then u is said to have a period of d and the sequence $(u, f(u), f_2(u), \dots, f_{d-1}(u))$ is called a d-cycle or, more simply, a limit cycle.*

Thus fixed points are constant solutions of (2.8) and periodic solutions of (2.8) are determined by periodic fixed points. The stability of a fixed point is described as follows.

Definition 2.17 *Let u be a fixed point of f.*

(a) If there exists a number ϵ such that $\lim_{t\to\infty} x_t = u$ when $|x_0 - u| < \epsilon$, then u is called stable.

(b) If there exists a number ϵ such that, when $0 < |x_0 - u| < \epsilon$, $|x_t - u| > \epsilon$ for some $t > 0$, then u is called unstable.

Part (a) defines stability in a local sense since convergence to a stable fixed point u is dependent on the system (2.8) starting near u (if convergence occurs for every possible initial value then stability is defined in a global sense). A periodic fixed point of f with period d can also be described as stable or unstable provided part (a) or (b), respectively, is satisfied under iteration of f_d. The system (2.8) is therefore 'attracted' to stable fixed points (stable periodic fixed points) and 'repelled' by unstable fixed points (unstable periodic fixed points).

Once the fixed points of f have been identified, it is then a simpler matter of using the following theorem to establish their stability.

Theorem 2.15 *Suppose that u is a fixed point of f and that f has a continuous first derivative at u. Then u is stable if $|f'(u)| < 1$ and unstable if $|f'(u)| > 1$.*

Similarly, a periodic fixed point of f with period d is stable if $|(f_d)'(u)| < 1$ or unstable if $|(f_d)'(u)| > 1$. In addition, a limit cycle is said to be stable (unstable) if each point in the limit cycle is stable (unstable).

For the case where $f'(u) = 1$, we need to look at higher derivatives of f. We present the following definition.

Definition 2.18 *A fixed point u is said to be semi-stable if either (a) or (b) below applies.*

(a) If u is stable for $x_t > u$ and unstable for $x_t < u$, then u is semi-stable 'from above'.

(b) If u is stable for $x_t < u$ and unstable for $x_t > u$, then u is semi-stable 'from below'.

We can now state the following result.

Theorem 2.16 ([85]) *Suppose that u is a fixed point of f and that $f'(u) = 1$.*

(a) If $f''(u) \neq 0$, then u is semi-stable: it is semi-stable from above if $f''(u) < 0$ and semi-stable from below if $f''(u) > 0$.

(b) If $f''(u) = 0$, then u is stable if $f'''(u) < 0$ or unstable if $f'''(u) > 0$.

It may be that a fixed point u is semi-stable from above with f only defined on the interval $[u, \infty)$, say, in which case u is then said to be stable.

2.6 Notation

We adopt the following conventions throughout this document:

- Random variables defined with the Bernoulli, binomial, Poisson and Gaussian distribution are denoted by $\mathrm{Ber}(p)$, $\mathrm{Bin}(n,p)$, $\mathrm{Poi}(\lambda)$ and $\mathrm{N}(\mu, \sigma^2)$, respectively. We will sometimes write "X follows the $\mathrm{Bin}(n,p)$ law", for example, where the notation instead refers to the *distribution* of X.

- The symbol $\xrightarrow{\mathrm{D}}$ denotes convergence in distribution and, similarly, $\xrightarrow{\mathrm{P}}$ and $\xrightarrow{r}$ denote convergence in probability and r-th mean, respectively. The symbol $\overset{\mathrm{D}}{=}$ reads as 'equal in distribution' and $:=$ reads as 'defined as'.

- We adopt the notation $f'(x)$ for the derivative of f with respect to x, and we shall sometimes write $f = g \circ h$ for the composition of g and h instead of the more formal $f(x) = g(h(x))$. We also adopt the conventions that $(\,\cdot\,)^\top$ denotes matrix or vector transpose and that empty products are to be interpreted as being equal to 1.

- The imaginary unit $\sqrt{-1}$ is denoted by $\mathbf{i}$ so as to distinguish it from i, which is used for other purposes. The exponential function is denoted by $\exp(\,\cdot\,)$ instead of, say, $\mathrm{e}^{(\,\cdot\,)}$, in order to prevent confusion with e, which is used for another purpose.

- Any variable accompanied with superscript N is indexed by the parameter N, not raised to the power N.

Chapter 3

A Mainland Model

This chapter introduces the simplest chain binomial metapopulation model. The local extinction and colonisation processes are both defined by binomial distributions, where the event probabilities are constant and independent of the current number of occupied patches. Note that this model was referred to as a 'mainland-island' model in [16, 14], however it is referred to as a 'mainland' model here.

3.1 Introduction

The concept of a *mainland* in metapopulation studies originates from MacArthur and Wilson's multi-species studies of a group of oceanic islands situated close to a mainland [55, 56]. A real island network is the obvious example of a metapopulation network and so the term 'mainland population' [12] was naturally adopted by modellers to describe populations connected to the network that were considered to be immune from extinction. Clearly, the movement of individuals is an important consideration in developing models for mainland-island networks. We divide such models into two groups: *mainland models*, those which account for migration from the mainland only, and *mainland-island models*, which account for migration from the mainland as well as the island network.

Considering a network of N patches connected to an outside source population (henceforth called the mainland), the classic mainland model is given by

$$\frac{dx}{dt} = c_0(1-x) - ex, \tag{3.1}$$

where x is the proportion of occupied patches ($x = n/N$ where n is the number of occupied patches), c_0 is the rate at which empty patches are colonised by migrants from the mainland and

e is the rate at which occupied patches go extinct [40, 43]. Indeed, equation (3.1) is the classic Ehrenfest urn model [25] in a continuous time setting: for N balls placed in two urns, x is the proportion of balls in the first urn, c_0 is the rate at which balls are added to this urn (from the second urn) and e is the rate that balls are taken from the first urn (and placed in the second urn). The differential equation has a unique positive equilibrium value at $x^* = c_0/(c_0 + e)$ which, in terms of metapopulations, implies that the metapopulation (island network) never goes extinct. If one is to account for migration from within the network of patches, the classic law of motion is given by

$$\frac{dx}{dt} = cx(1-x) + c_0(1-x) - ex, \tag{3.2}$$

where c is the rate that occupied patches colonise empty patches in unoccupied-occupied patch pairs [39]. For the case where $c_0 = 0$, (3.2) reduces to Levins' interpretation of Verhulst's model for a network of N patches without a mainland [53]. For models that do not consider a mainland, these are referred to as *island models*.

In this chapter, we shall concentrate on a stochastic mainland model. In particular, the model is formulated in discrete time and exhibits a chain binomial structure that accounts for seasonal variation. A number of analytical results are presented and I note a connection with the discrete-time version of the classic mainland model above. The following chapter concentrates on stochastic island and mainland-island models, which are based on the chain binomial structure introduced here.

3.2 The chain binomial mainland model

Consider a metapopulation residing in a network of finitely-many, N, habitat patches. The population is assumed to exhibit a seasonal dynamic whereby local extinction events occur during a particular seasonal period whilst colonisation events occur during another seasonal period. These periods, or *phases*, do not overlap and alternate over time. A census is assumed to take place either at the end of each colonisation phase (the EC model) or at the end of each extinction phase (the CE model).

Let n_t be the number of occupied patches at census time $t \in \{0, 1, 2, \ldots\}$ and further let $(n_t : t \geq 0)$ be a Markov chain taking values in the set $S_N = \{0, 1, \ldots, N\}$ with 1-step transition matrix $P = (p_{ij} : i, j \in S_N)$. The transition matrices $E = (e_{ij} : i, j \in S_N)$ and

$C = (c_{ij} : i, j \in S_N)$ govern the local extinction and colonisation phases, respectively, where the transition probabilities are defined below. Depending on the timing of the census, we have $P = EC$ or $P = CE$. Notice that the two observation schemes correspond to time-homogeneous observations of a time-inhomogenenous Markov chain (Section 2.2.4); we consider both schemes in the model analysis.

The network of patches is assumed to be connected to a mainland and the reason for this will be made clear below.

Extinction process. Suppose that the local population within each occupied patch goes extinct with probability e $(0 < e < 1)$ during the extinction phase and suppose that each event is independent. Thus, given i patches are occupied at the start of this phase, the number that survive local extinction follows a $\mathrm{Bin}(i, 1-e)$ law. Therefore,

$$e_{ij} = \begin{cases} \binom{i}{j}(1-e)^j e^{i-j} & \text{if } j = 0, \ldots, i \\ 0 & \text{if } j > i. \end{cases} \tag{3.3}$$

Observe that $e_{0j} = \delta_{0j}$ for all $j \in S_N$.

Colonisation process. Suppose that each empty patch is colonised with probability c_0 $(0 < c_0 \leq 1)$ during the colonisation phase and that all colonisation events are independent. Thus, given i patches are currently occupied (and hence $N - i$ unoccupied), the number of empty patches that are colonised during the period follows a $\mathrm{Bin}(N - i, c_0)$ law. Therefore,

$$c_{ij} = \begin{cases} \binom{N-i}{j-i}(1-c_0)^{N-j} c_0^{j-i} & \text{if } j = i, i+1, \ldots, N \\ 0 & \text{if } j < i. \end{cases} \tag{3.4}$$

Since c_0 does not depend on the current number of occupied patches, we may think of the empty patches as being colonised by migrants from a population outside the patch network. This source population is considered to be immune from extinction and will therefore be referred to as a mainland population.

Transition probabilities. Notice that the matrix E is lower triangular and the matrix C is upper triangular, however P is dense in both cases. The 1-step transition probabilities are therefore given by

$$\text{EC:} \qquad p_{ij} = \sum_{k=0}^{\min(i,j)} e_{ik}c_{kj} \tag{3.5a}$$

$$\text{CE:} \qquad p_{ij} = \sum_{\max(i,j)}^{N} c_{ik}e_{kj}. \tag{3.5b}$$

Since $e_{00} = 1$ and $c_{00} < 1$, we have that $p_{00} < 1$ and so there is a positive probability that the Markov chain jumps *out* of state 0 (the state corresponding to complete extinction of the patch network). This means that the metapopulation has a chance to recover and this formulation supports the assumption of a mainland population which, in this case, is thought to provide migrants that colonise each island with equal probability. Subsequently the set S_N forms an irreducible aperiodic class, and the Markov chain has a unique stationary (and hence limiting) distribution (Section 2.2.2).

Remark. The model outlined above accounts for migration from a mainland population only. In the ecological literature this model would be said to have a "source-sink" [74] structure, since the mainland population acts as a "source population" in this case while the island populations act as "sink populations".

3.3 Model analysis

The stochastic mainland model admits the following chain binomial structure:

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \text{Bin}(N - \tilde{n}_t, c_0) \qquad \tilde{n}_t = n_t - \text{Bin}(n_t, e) \tag{3.6a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \text{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \text{Bin}(N - n_t, c_0). \tag{3.6b}$$

I note that the CE case was interpreted by Daley and Gani [20, see equation 4.4.11] as a chain binomial SIS epidemic model that incorporates immigration-emmigration episodes. The results stated in this chapter apply equally to their model, with n_t interpreted as the number of susceptible individuals in a closed population with ceiling N.

The model analysis begins by evaluating the 1-step distribution of n_{t+1}. Let G define the probability generating function (pgf) of n_{t+1} conditional on n_t, that is $G(z) = \mathbb{E}(z^{n_{t+1}} | n_t = i)$.

For the EC model, $n_{t+1} = \tilde{n}_t + \tilde{B}$, where $\tilde{n}_t \sim \mathrm{Bin}(i, 1-e)$ and $\tilde{B} \sim \mathrm{Bin}(N - \tilde{n}_t, c_0)$, and so

$$G(z) = \mathbb{E}(z^{n_{t+1}}|n_t = i) = \mathbb{E}\left(\mathbb{E}(z^{\tilde{n}_t+\tilde{B}}|\tilde{n}_t, n_t = i)|n_t = i\right) = \mathbb{E}\left(z^{\tilde{n}_t}\mathbb{E}(z^{\tilde{B}}|\tilde{n}_t, n_t = i)|n_t = i\right)$$

by conditional expectation. By observing the pgf of $\tilde{B}$ and $\tilde{n}_t$ in turn, we simplify the last expression as follows:

$$\begin{aligned}
G(z) &= \mathbb{E}\left(z^{\tilde{n}_t}(1 - c_0 + c_0 z)^{N-\tilde{n}_t}|n_t = i\right) \\
&= (1 - c_0 + c_0 z)^N\, \mathbb{E}\left((z/(1 - c_0 + c_0 z))^{\tilde{n}_t}|n_t = i\right) \\
&= (1 - c_0 + c_0 z)^N (e + (1-e)z/(1 - c_0 + c_0 z))^i \\
&= (e(1-c_0) + [1 - e(1-c_0)]z)^i (1 - c_0 + c_0 z)^{N-i}.
\end{aligned}$$

For the CE model, $n_{t+1} \sim \mathrm{Bin}(i + \tilde{B}, 1-e)$, where $\tilde{B} \sim \mathrm{Bin}(N - i, c_0)$, and so

$$\begin{aligned}
G(z) &= \mathbb{E}\left(\mathbb{E}(z^{n_{t+1}}|\tilde{B}, n_t = i)|n_t = i\right) = \mathbb{E}\left((e + (1-e)z)^{i+\tilde{B}}|n_t = i\right) \\
&= (e + (1-e)z)^i\, \mathbb{E}\left((e + (1-e)z)^{\tilde{B}}|n_t = i\right) \\
&= (e + (1-e)z)^i (1 - c_0 + c_0(e + (1-e)z))^{N-i} \\
&= (e + (1-e)z)^i (1 - c_0(1-e) + c_0(1-e)z)^{N-i}.
\end{aligned}$$

Therefore $G(z) = (1 - p + pz)^i (1 - q + qz)^{N-i}$ in both cases where

$$\text{EC:} \qquad p = 1 - e(1 - c_0) \qquad q = c_0 \tag{3.7a}$$

$$\text{CE:} \qquad p = 1 - e \qquad q = c_0(1 - e). \tag{3.7b}$$

The pgf G is in the form of two binomial pgfs multiplied together and this establishes the following result.

Lemma 3.1 *Given $n_t = i$, $n_{t+1} \overset{\mathrm{D}}{=} B_1 + B_2$ where $B_1 \sim \mathrm{Bin}(i, p)$ and $B_2 \sim \mathrm{Bin}(N - i, q)$ are independent binomial random variables.*

Lemma 3.1 says that each of the currently occupied patches (i of them) behave as if they remain occupied with probability p and each of the $N - i$ unoccupied patches behave as though they become occupied with probability q, all patches being affected independently. The parameters p and q may therefore be interpreted as 'effective' survival and colonisation probabilities, respectively.

We now evaluate the distribution of n_t conditioned on the initial ($t = 0$) number of occupied patches. Suppose that $n_0 = i$ and let G_t define the pgf of n_t conditional on the initial value, that is $G_t(z) = \mathbb{E}(z^{n_t}|n_0 = i)$. Then, $G_0(z) = z^i$ and, from Lemma 3.1, we have

$$\begin{aligned} G_{t+1}(z) &= \mathbb{E}\left((1-p+pz)^{n_t}(1-q+qz)^{N-n_t}|n_0 = i\right) \\ &= (1-q+qz)^N\,\mathbb{E}\left(\left(\frac{1-p+pz}{1-q+qz}\right)^{n_t}|n_0 = i\right) \\ &= (1-q+qz)^N\,G_t\left(\frac{1-p+pz}{1-q+qz}\right), \end{aligned} \tag{3.8}$$

for all $t \geq 0$. The pgf G was shown to be in the form of two binomial pgfs multiplied together and we therefore intuit that

$$G_t(z) = (1-p_t+p_t z)^i(1-q_t+q_t z)^{N-i}, \tag{3.9}$$

for all $t \geq 0$, where p_t and q_t are time-dependent binomial parameters. Clearly (3.9) is true for $t = 0$ and $t = 1$ because $p_0 = 1$ and $q_0 = 0$ (since $G_0(z) = z^i$), while $p_1 = p$ and $q_1 = q$ (since $G_1 \equiv G$). Now suppose that (3.9) is true for some fixed $t \geq 0$. Substitution in to (3.8) then gives

$$\begin{aligned} G_{t+1}(z) &= (1-q+qz)^N\left(1-p_t+p_t\left(\frac{1-p+pz}{1-q+qz}\right)\right)^i\left(1-q_t+q_t\left(\frac{1-p+pz}{1-q+qz}\right)\right)^{N-i} \\ &= \left(1-(q+ap_t)+(q+ap_t)z\right)^i\left(1-(q+aq_t)+(q+aq_t)z\right)^{N-i}, \end{aligned}$$

where we have set $a = p - q = (1-e)(1-c_0)$, $0 < a < 1$, the same for both EC and CE cases. By the inductive hypothesis, $p_{t+1} = q + ap_t$ and $q_{t+1} = q + aq_t$, which solve to give

$$q_t = q^*(1-a^t) \quad \text{and} \quad p_t = q_t + a^t \qquad (t \geq 0), \tag{3.10}$$

where we have set $q^* = q/(1-a)$, $0 < q^* < 1$. By Mathematical Induction, equation (3.9) is proved where the sequences (p_t) and (q_t) are defined by (3.10). The result below follows.

Theorem 3.1 *Given* $n_0 = i$, $n_t \overset{\mathrm{D}}{=} B_t^{(1)} + B_t^{(2)}$ *where* $B_t^{(1)} \sim \mathrm{Bin}(i, p_t)$ *and* $B_t^{(2)} \sim \mathrm{Bin}(N - i, q_t)$ *are independent binomial random variables.*

Again, the behaviour of the Markov chain may be re-stated in terms of independent binomial random variables. It is as if the i initially occupied patches remain occupied with probability p_t and the $N - i$ initially unoccupied patches are occupied with probability q_t, all patches being affected independently. We may conclude that $\mathbb{E}(n_t | n_0 = i) = ip_t + (N - i)q_t$ and, since $B_t^{(1)}$ and $B_t^{(2)}$ are independent, $\mathrm{Var}\,(n_t | n_0 = i) = ip_t(1 - p_t) + (N - i)q_t(1 - q_t)$.

The equilibrium ($t \to \infty$) behaviour of n_t is simple to describe. Notice that $\lim_{t\to\infty} a^t = 0$ and therefore the sequences (p_t) and (q_t) have common limit q^*. Given $n_0 = i$, it is clear from Theorem 3.1 that $B_t^{(1)} \xrightarrow{\mathrm{D}} B^{(1)} \sim \mathrm{Bin}(i, q^*)$ and $B_t^{(2)} \xrightarrow{\mathrm{D}} B^{(2)} \sim \mathrm{Bin}(N - i, q^*)$, as $t \to \infty$, because the corresponding sequences of characteristic functions converge point-wise to the appropriate limits. The random variables $B_t^{(1)}$ and $B_t^{(2)}$ are independent (for each t) and hence $B^{(1)}$ and $B^{(2)}$ are independent also. Moreover, $B_t^{(1)} + B_t^{(2)} \xrightarrow{\mathrm{D}} B^{(1)} + B^{(2)}$ and, clearly, $B^{(1)} + B^{(2)} \sim \mathrm{Bin}(N, q^*)$. The following Corollary is thus established.

Corollary 3.1 *As* $t \to \infty$, $n_t \xrightarrow{\mathrm{D}} \mathrm{Bin}(N, q^*)$.

Thus the equilibrium (and stationary) distribution of the Markov chain is a binomial distribution with parameters N and q^*; it is as if each of the N patches are occupied with probability q^*. This result is demonstrated in Figure 3.1 for the EC model, along with a simulation of the random process (the algorithm for simulating a finite chain is provided in Section 2.2.5). Indeed, the significance of the parameters q^* and a are now clear: q^* is the equilibrium proportion of occupied patches and a (see Theorem 3.1) is the rate of (geometric) approach to equilibrium. Remembering that $q^* = q/(1 - a)$, where q is given by (3.7) and $a = (1 - e)(1 - c_0)$ for both EC and CE models, the equilibrium proportion of occupied patches is therefore smaller for the CE case by a factor of $1 - e$. The long-term proportions fluctuate between high and low values since the local extinction and colonisation processes are assumed to occur at distinct periods in time.

Remark. There are obvious connections between our chain binomial mainland model and the binomial AR-1 model proposed by McKenzie [62] (see also McKenzie [63] and Weiß [92]). Recall that, given $n_t = i$, n_t is equal in distribution to $B_1 + B_2$ where $B_1 \sim \mathrm{Bin}(i, p)$ and

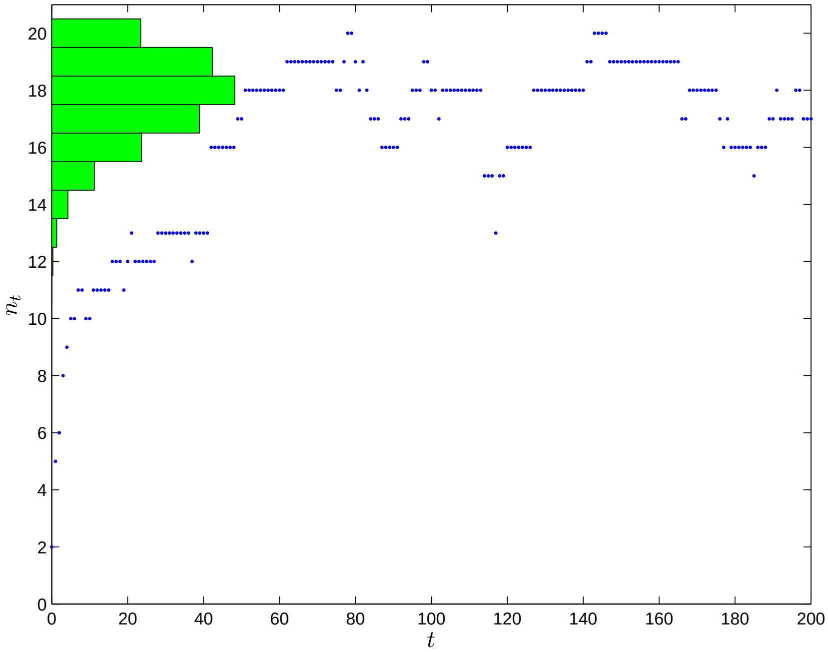

Figure 3.1: A single simulation of the EC model with $N = 20$, $e = 0.01$ and $c_0 = 0.05$, starting with $n_0 = 2$ patches occupied. The number n_t of occupied patches is plotted at times $t = 0, \ldots, 200$. The bar graph in green is the $\text{Bin}(N, q^*)$ stationary distribution ($q^* = 0.84034$).

$B_2 \sim \text{Bin}(N - i, q)$ are independent binomial random variables. Then, if we consider $n_{t+1} = \text{Bin}(n_t, p) + \text{Bin}(N - n_t, q)$ as a time-series model, the process $(n_t, t \geq 0)$ defined by this model is a binomial AR-1 process where p and q are known as the 'binomial thinning' parameters. In fact, Weiß proved that the stationary distribution of a binomial AR-1 process follows a single binomial law [92, Lemma 1.1]. Our result (Corollary 3.1) accords exactly with Weiß' result, though note that our proof is more detailed: we observed the limiting behaviour of the t-step conditional state distribution. Weiß did not evaluate the t-step distribution, though perhaps it did not occur to him to evaluate such a distribution since he was only concerned with studying *stationary* autoregressive processes.

3.3.1 Explaining the *rescue effect*

The term *rescue effect* [13, 40] is used to describe a decrease in local extinction rate when the immigration rate is high. I note a connection between the mainland model above and the incidence function model of Hanski [42] designed to account for the rescue effect.

Incidence function models [40, 41, 42] define the incidence (of occupancy) of a single patch as the stationary probability that the patch is occupied, whereby this probability is determined from a two-state Markov chain. If patch i is empty, it is defined to be recolonised with probability C_i and if the patch is occupied, it experiences local extinction with probability E_i. The incidence J_i of this patch is therefore

$$J_i = \frac{C_i}{C_i + E_i}. \tag{3.11}$$

In order to account for a rescue effect, Hanski [42] intuits that the local extinction probability should be replaced with $(1-C_i)E_i$ where E_i now accounts for environmental stochasticity only, that is, extinction as it occurs in the absence of migration. Hence, the incidence of patch i becomes

$$J_i = \frac{C_i}{C_i + (1 - C_i)E_i}. \tag{3.12}$$

If the colonisation and extinction probabilities are defined to be the same for all patches, a mainland-type model results with $J_i = c_0/(c_0 + (1-c_0)e)$ for all i (using our notation). Assuming seasonal extinction-colonisation dynamics, we established that $(1 - c_0)e$ *is* the 'effective' extinction probability in the EC case (Lemma 3.1) and the expected equilibrium proportion of occupied patches for this case is $q^* = q/(1 - a) = c_0/(c_0 + (1 - c_0)e)$ (Corollary 3.1), precisely that given by (3.12) with $C_i = c_0$ and $E_i = e$. More importantly, Theorem 3.1 establishes a time-dependent interpretation of the rescue effect. The 'effective' extinction probability at time t is $1 - p_t = (1 - q^*)(1 - a^t)$, which for the EC model is $1 - p_t = (1 - c_0)e(1 - a^t)/(1 - a)$, being $(1 - c_0)e$ when $t = 1$ and $(1 - e)c_0/(c_0 + (1 - c_0)e)$ in the long term.

A similar investigation of the CE case reveals that this model does not account for a rescue effect but rather an "inhibiting" effect that *decreases* the colonisation rate when the local extinction rate is high. To see this, recall that the 'effective' colonisation probability for the CE model is $c_0(1-e)$ and that the expected equilibrium proportion of occupied patches in this case is $c_0(1 - e)/(c_0(1 - e) + e)$. If C_i is replaced by $C_i(1 - E_i)$ in equation (3.11), then the incidence of patch i becomes

$$J_i = \frac{C_i(1 - E_i)}{C_i(1 - E_i) + E_i}.$$

Therefore Hanski's incidence function model also accounts for an inhibitive effect under the appropriate substitution. Furthermore, our CE model establishes a time-dependent interpretation of the "inhibiting" effect: the 'effective' colonisation probability at time t is $q_t = q^*(1-a^t) = c_0(1-e)(1-a^t)/(1-a)$, being $c_0(1-e)$ when $t=1$ and $c_0(1-e)/(c_0+(1-c_0)e)$ in the long term.

Remark. The calculations show, for the stochastic mainland model, the rescue effect and the inhibitive effect are artefacts of the timing of the population census. If the census is taken at the end of the colonisation phase, then the population has had a chance to recover from the preceding phase of local extinction and the result manifests as a rescue effect. Similarly, if the census is taken at the end of the local extinction phase, the population size may have declined during that phase and the result manifests as an inhibitive effect.

3.3.2 Large networks

We consider, now, the behaviour of the chain binomial mainland model as the parameter N gets large. Let $X_t^N = n_t/N$, the proportion of occupied patches at time t. The following law of large numbers gives the deterministic process (x_t) that approximates X_t^N for large metapopulation networks.

Theorem 3.2 *If $X_0^N \xrightarrow{\text{P}} x_0$ (a constant) as $N \to \infty$, then, for any $t \geq 1$, $X_t^N \xrightarrow{\text{P}} x_t$ as $N \to \infty$, where*

$$x_t = p_t x_0 + q_t(1-x_0) = q^* + a^t(x_0 - q^*). \tag{3.13}$$

Proof. From Theorem 3.1, we have $X_t^N \overset{\text{D}}{=} K_t^N + L_t^N$ where $K_t^N = B_t^{(1)}/N = X_0^N B_t^{(1)}/n_0$ and $L_t^N = B_t^{(2)}/N = (1-X_0^N)B_t^{(2)}/(N-n_0)$ are independent random variables. We are told that $X_0^N \xrightarrow{\text{P}} x_0$ as $N \to \infty$, and so $n_0 \to \infty$ and $N - n_0 = N(1-X_0^N) \to \infty$. It follows from the standard Weak Law of Large Numbers that $B_t^{(1)}/n_0 \xrightarrow{\text{P}} p_t$ and $B_t^{(2)}/(N-n_0) \xrightarrow{\text{P}} q_t$. Hence, $K_t^N \xrightarrow{\text{P}} x_0 p_t$ and $L_t^N \xrightarrow{\text{P}} (1-x_0)q_t$, and so $K_t^N + L_t^N \xrightarrow{\text{P}} p_t x_0 + q_t(1-x_0)$ $(= x_t)$. From (3.10), notice that $p_t x_0 + q_t(1-x_0) = q^* + a^t(x_0 - q^*)$ and the result follows. □

From (3.13), it is easy to see that the equilibrium $(t \to \infty)$ proportion of occupied patches for the deterministic process is q^*. From Corollary 3.1, recall that the equilibrium *expected* proportion is q^*. Thus, for both EC and CE models, the proportion of occupied patches converges to the same limit no matter the order in which the limits are taken. However, remember that q^* is different for each model and that this quantity is uniformly larger in the EC case. Indeed, the deterministic model is uniformly larger in the EC case since

$$x_t^{(EC)} - x_t^{(CE)} = \frac{c_0(1 - ea^t)}{1 - a} > 0$$

for all $t \geq 1$, using an obvious notation. This is not surprising since, even in equilibrium, the process is still being observed after each period of colonisation.

From Lemma 3.1, we deduce that $\mathbb{E}(n_{t+1}|n_t = i) = ip + (N - i)q$. Provided n_t/N converges to x_t in accordance with Theorem 3.2, we have that $x_{t+1} = x_t p + (1 - x_t)q$, for any $t \geq 1$. After some adjustments we have

$$x_{t+1} - x_t = q(1 - x_t) - (1 - p)x_t \tag{3.14}$$

which is immediately identified as a discrete-time version of the classic mainland model (3.1) with q and $1 - p$ as the colonisation and extinction rate parameters, respectively. The unique positive equilibrium of (3.14) was established to be $q^*(\equiv x^*)$ which has obvious connections with Hanski's incidence of occupancy given by (3.12).

Remark. The stochastic mainland model defined by equation (3.6) explicitly assumes a seasonal dynamic. The distribution stated in Lemma 3.1, which is *equivalent* to the distribution given by (3.6), is expressed in terms of independent random variables and it is this independence that effectively allows us to 'ignore' the assumed seasonal dynamic in further analysis. It is not surprising then that the difference equation (3.14) does not assume a seasonal structure. But, although the difference equation does not strictly exhibit seasonal variation, the seasonal behaviour inherent in the stochastic formulation is still retained: the equilibrium proportion of occupied patches is larger for the EC case.

Next, we define $Z_t^N = \sqrt{N}(X_t^N - x_t)$ and examine this quantity in the limit as $N \to \infty$. The following central limit law establishes that the fluctuation Z_t^N about the deterministic trajectory has a normal (Gaussian) distribution in this limit.

Theorem 3.3 *In addition to the conditions of Theorem 3.2, suppose that* $Z_0^N \xrightarrow{\mathrm{D}} z_0$ *(a constant) as* $N \to \infty$*. Then, for any* $t \geq 1$, $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(z_0 a^t, V_t)$ *as* $N \to \infty$*, where*

$$V_t = p_t(1 - p_t)x_0 + q_t(1 - q_t)(1 - x_0).$$

Proof. From Theorem 3.1 and equation (3.13), observe that

$$\begin{aligned}
X_t^N - x_t &= B_t^{(1)}/N + B_t^{(2)}/N - x_0 p_t - (1-x_0)q_t \\
&= X_0^N B_t^{(1)}/n_0 - x_0 p_t + (1-X_0^N)B_t^{(2)}/(N-n_0) - (1-x_0)q_t \\
&= X_0^N (B_t^{(1)}/n_0 - p_t) + (1-X_0^N)\big(B_t^{(2)}/(N-n_0) - q_t\big) + (X_0^N - x_0)(p_t - q_t),
\end{aligned}$$

and so, on multiplying by $\sqrt{N}$ and noting that $p_t - q_t = a^t$, we have

$$Z_t^N = \sqrt{X_0^N}\sqrt{n_0}\big(B_t^{(1)}/n_0 - p_t\big) + \sqrt{1-X_0^N}\sqrt{N-n_0}\big(B_t^{(2)}/(N-n_0) - q_t\big) + Z_0^N a^t, \qquad (3.15)$$

where $Z_0^N = \sqrt{N}(X_0^N - x_0)$. By the standard Central Limit Theorem $\sqrt{n_0}(B_t^{(1)}/n_0 - p_t) \xrightarrow{\mathrm{D}} \mathrm{N}(0, p_t(1-p_t))$ and $\sqrt{N-n_0}\big(B_t^{(2)}/(N-n_0) - q_t\big) \xrightarrow{\mathrm{D}} \mathrm{N}(0, q_t(1-q_t))$. Therefore, the first and second terms of (3.15) define independent sequences that converge in distribution to $\mathrm{N}(0, x_0 p_t(1-p_t))$ and $\mathrm{N}(0, (1-x_0)q_t(1-q_t))$ random variables, respectively. But, we are told that $Z_0^N \xrightarrow{\mathrm{D}} z_0$ (a constant) as $N \to \infty$. Therefore,

$$\begin{aligned}
Z_t^N &\xrightarrow{\mathrm{D}} \mathrm{N}\left(0, p_t(1-p_t)x_0\right) + \mathrm{N}\left(0, q_t(1-q_t)(1-x_0)\right) + z_0 a^t \\
&= \mathrm{N}\left(z_0 a^t, p_t(1-p_t)x_0 + q_t(1-q_t)(1-x_0)\right)
\end{aligned}$$

and the proof is complete. □

Notice that the approximating variance given in Theorem 3.3 is consistent with the exact variance given by Theorem 3.1, that is,

$$\mathrm{Var}(Z_t^N) = N\mathrm{Var}(X_t^N) = (1/N)\mathrm{Var}(n_t) = p_t(1-p_t)X_0^N + q_t(1-q_t)(1-X_0^N) \to V_t$$

as $N \to \infty$. They will be identical if $X_0^N = x_0$ and this is illustrated in Figure 3.2. Furthermore, the quality of the deterministic approximation may be assessed: the central limit law stated above implies that, for N sufficiently large, $\Pr(|X_t^N - x_t| > \epsilon_t) \simeq 2(1 - \Phi(\epsilon_t\sqrt{N/V_t}))$, where Φ is the standard normal distribution function.

A simpler distributional approximation for Z_t^N results when the random process is in equilibrium. On setting $x_0 = q^*$ in Theorems 3.2 and 3.3, the following corollary is obtained.

Corollary 3.2 *If* $X_0^N \xrightarrow{\mathrm{P}} q^*$ *as* $N \to \infty$*, then* $X_t^N \xrightarrow{\mathrm{P}} q^*$ *for all* $t \geq 1$*. Let* $Z_t^N = \sqrt{N}(X_t^N - q^*)$*. If, in addition,* $Z_0^N \xrightarrow{\mathrm{D}} z_0$ *(a constant) as* $N \to \infty$*, then* $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(z_0 a^t, q^*(1-q^*))$*.*

I note that stronger convergence results have been proved and that these are given later in Chapter 6. For example, there we show that the scaled process $(Z_t^N : t \geq 0)$ converges in finite dimensional distribution to the random process $(Z_t : t \geq 0)$ which is a Gaussian Markov process. Indeed, this type of convergence behaviour is the focus of Chapter 6 where limit theorems are presented for chain binomial metapopulations models in general.

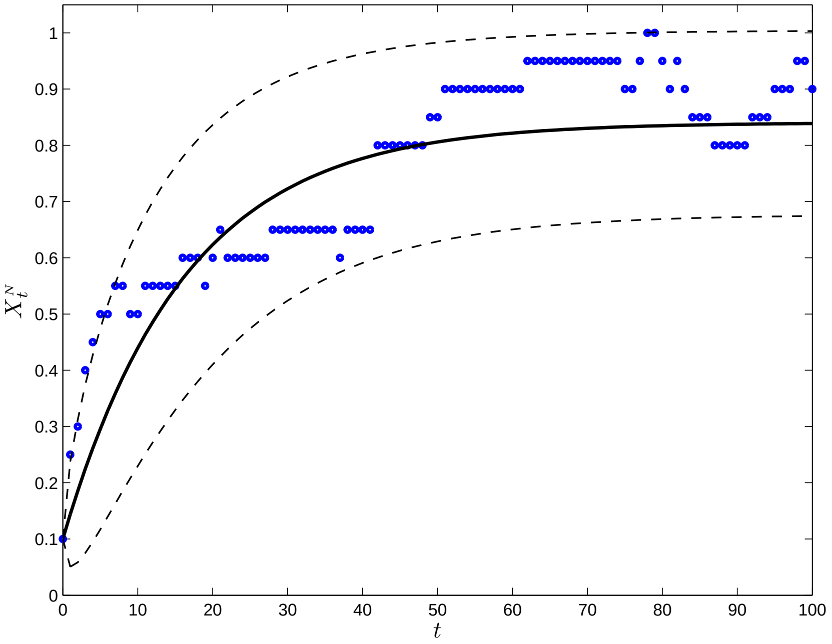

Figure 3.2: A single simulation of the EC model with $N = 20$, $e = 0.01$ and $c_0 = 0.05$, starting with $n_0 = 2$ patches occupied ($X_t^N = x_0 = 0.1$). The proportion X_t^N of occupied patches is plotted at times $t = 0, \ldots, 100$. The solid curve joins points on the limiting deterministic trajectory. The dashed curve is ± 2 standard deviations as predicted by the normal approximation.

3.3.3 Expected first passage time

Given that we have the t-step conditional state distribution for both the EC and CE mainland models, we can evaluate the mean first passage time to state 0 using Theorem 2.1 in both cases. The expected time at which the metapopulation first goes extinct is an important quantity because it can be used as a measure of population viability.

Theorem 3.4 *Let* $T = \inf\{t \geq 1 : n_t = 0\}$ *be the first time at which all the patches in the metapopulation network are empty. Then*

$$\mathbb{E}(T|n_0 = i) = \sum_{k=1}^{N} \binom{N}{k} \left(\frac{q^*}{1-q^*}\right)^k \frac{1}{(1-a^k)} - \sum_{(k,m)\in A} \binom{i}{k}\binom{N-i}{m} \left(\frac{q^*}{1-q^*}\right)^m \frac{(-1)^k}{(1-a^{k+m})}$$

for both EC and CE chain binomial mainland models, where the set A *is defined as*

$$A = \{(k,m) : k = 0,1,\ldots,i;\ m = 0,1,\ldots,N-i;\ (k,m) \neq (0,0)\}.$$

Proof. Remembering that equation (2.2) defined $P_{ij}(z) = \sum_{t=0}^{\infty} z^t p_{ij}^{(t)}$ as the generating function of the t-step transition probabilities $(p_{ij}^{(t)}, t \geq 0)$, we see that $P_{i0}(z) = \sum_{t=0}^{\infty} z^t p_{i0}^{(t)}$ and $P_{00}(z) = \sum_{t=0}^{\infty} z^t p_{00}^{(t)}$. From equation (3.9), we have $p_{i0}^{(t)} = G_t(0) = (1-p_t)^i(1-q_t)^{N-i}$ and therefore $p_{00}^{(t)} = (1-q_t)^N$. Hence

$$\begin{aligned} P_{i0}(z) &= \sum_{t=0}^{\infty} z^t (1-p_t)^i\ (1-q_t)^{N-i} \\ &= \sum_{t=0}^{\infty} z^t \left(1-q^*-(1-q^*)a^t\right)^i\ \left(1-q^*+q^*a^t\right)^{N-i} \\ &= (1-q^*)^i \sum_{t=0}^{\infty} z^t\ (1-a^t)^i\ (1-q^*+q^*a^t)^{N-i} \end{aligned}$$

where we have used $q_t = q^*(1-a^t)$ and $p_t = q_t + a^t = q^* + q^*(1-a^t)$. Using the binomial theorem we may write $(1-a^t)^i = \sum_{k=0}^{i} \binom{i}{k}(-1)^k(a^t)^k$ and, similarly, $(1-q^*+q^*a^t)^{N-i} = \sum_{m=0}^{N-i} \binom{N-i}{m}(q^*a^t)^m(1-q^*)^{N-i-m}$. After grouping terms involving t, we have

$$P_{i0}(z) = (1-q^*)^N \sum_{k=0}^{i}\sum_{m=0}^{N-i} \binom{i}{k}\binom{N-i}{m}(-1)^k \left(\frac{q^*}{1-q^*}\right)^m \frac{1}{(1-za^{k+m})}. \qquad (3.16)$$

We know that $P_{00}(z) = \sum_{t=0}^{\infty} z^t(1-q_t)^N$ and, after similar simplifications, we have

$$P_{00}(z) = (1-q^*)^N \sum_{k=0}^{N} \binom{N}{k} \left(\frac{q^*}{1-q^*}\right)^k \frac{1}{(1-za^k)}.$$

Also remember that equation (2.2) defined $F_{ij}(z) = \sum_{t=0}^{\infty} z^t f_{ij}^{(t)}$ as the generating function of the first passage time probabilities $(f_{ij}^{(t)}, t \geq 0)$. From Theorem 2.1(b), we can write

$$F_{i0}(z) = \sum_{t=0}^{\infty} z^t f_{i0}^{(t)} = P_{i0}(z)/P_{00}(z).$$

The sequence $\left(f_{i0}^{(t)}, t \geq 0\right)$ defines the probability mass function of T and so, by evaluating the derivative of $F_{i0}(z)$ with respect to z in the limit as $z \uparrow 1$, we obtain the expected value of T conditional on the event $\{n_0 = i\}$. Clearly,

$$F'_{i0}(z) = \frac{P'_{i0}(z)P_{00}(z) - P_{i0}(z)P'_{00}(z)}{P_{00}(z)^2}. \tag{3.17}$$

After some simplifications,

$$P'_{i0}(z)P_{00}(z) = \frac{(1-q^*)^{2N}}{(1-z)^2}\left(\frac{1}{1-z} + \sum_{k=1}^{N}\binom{N}{k}\left(\frac{q^*}{1-q^*}\right)^k \frac{1}{(1-za^k)} + o(1)\right),$$

where $o(1) \to 0$ as $z \uparrow 1$, and

$$P_{i0}(z)P'_{00}(z) = \frac{(1-q^*)^{2N}}{(1-z)^2}\left(\frac{1}{1-z} + \sum_{A}\binom{i}{k}\binom{N-i}{m}\left(\frac{q^*}{1-q^*}\right)^m \frac{(-1)^k}{(1-za^{k+m})} + o(1)\right),$$

where the set A defines the sum in equation (3.16) without the first term (the term corresponding to $k = 0$ and $m = 0$). Also,

$$P_{00}(z)^2 = \frac{(1-q^*)^{2N}}{(1-z)^2}\left(1 + o(1)\right).$$

Notice that the factor $(1-q^*)^{2N}/(1-z)^2$ cancels upon substituting the last three expression into equation (3.17), and hence

$$\lim_{z\uparrow 1} F'_{i0}(z) = \sum_{k=1}^{N}\binom{N}{k}\left(\frac{q^*}{1-q^*}\right)^k \frac{1}{(1-a^k)} - \sum_{A}\binom{i}{k}\binom{N-i}{m}\left(\frac{q^*}{1-q^*}\right)^m \frac{(-1)^k}{(1-a^{k+m})},$$

which equals $\mathbb{E}(T|n_0 = i)$, and the proof is complete. $\square$

3.4 Discussion

The simplicity of the chain binomial mainland model is exemplified by Lemma 3.1 where the census at time $t + 1$, conditioned on the previous census at time t, was found to be equal in distribution to the sum of two *independent* binomial random variables. The parameters (p, q) were interpreted as the effective local extinction and colonisation probabilities and furthermore, by Theorem 3.1, the parameters (p_t, q_t) were interpreted as the effective time-dependent event probabilities. The equilibrium proportion of occupied patches was determined to be q^* for both time-homogeneous (EC and CE) models, however this quantity is slightly larger for the EC case since, even in equilibrium, the process is observed after the colonisation period.

The independence exhibited in Lemma 3.1 and Theorem 3.1 allowed a number of quantities to be evaluated explicitly. These included a law of large numbers (Theorem 3.2) and a central limit law (Theorem 3.3) for large networks, as well as the mean first passage time to state 0 (Theorem 3.4).

Since the chain binomial mainland model accounts for migration from a mainland population only, it is the simplest stochastic population model with a chain binomial structure. The model may be too simplistic for applied studies, however Daley and Gani's [20] epidemiological interpretation of the chain binomial mainland model has seen application in Gani and Stals' [29] viral plant epidemic. As the following chapters will show, the chain binomial mainland model can be adapted to account for a wide range of interesting population dynamics.

CHAPTER 4

Island Models and Mainland-Island Models

The chain binomial structure introduced in Chapter 3 is modified here to account for migration between local populations only (island models) or for migration between these populations as well as from a mainland (mainland-island models). This is done by allowing the colonisation probability to be state dependent.

4.1 Introduction

The previous chapter introduced the classic mainland model (3.1) and the classic mainland-island model (3.2). It is not surprising then that the classic island model is Verhulst's model (1.1) with n interpreted as the number of extant local populations (or more simply, the number of occupied patches) in a network of N habitat patches. This interpretation was made by Levins in 1969 [51] and upon rewriting the model in terms of the proportion $x = n/N$ of occupied patches [53] we have

$$\frac{dx}{dt} = cx(1-x) - ex, \tag{4.1}$$

where c is the (pairwise) rate at which empty patches are colonised and e is the local extinction rate. Alternatively, we may express Levins' model in a slightly different way:

$$\frac{dx}{dt} = (c-e)x\left(1 - \frac{x}{(1-e/c)}\right).$$

The second form clearly shows that the model has two equilibrium points, 0 and $1 - e/c$, and the factor $c - e$ is identified as the growth rate at low proportions. If $c > e$, the point $1 - e/c$ is stable and the metapopulation is said to persist with equilibrium proportion $1 - e/c$. If $c \leq e$, then 0 is stable and the metapopulation goes extinct. More generally, Levins' model satisfies the form

$$\frac{dx}{dt} = rx\left(1 - \frac{x}{x^*}\right), \tag{4.2}$$

where r is called the *intrinsic growth rate* and $x^*(\neq 0)$ is called the *carrying capacity* (expressed here as a proportion of the *population ceiling*, N). Thus $r = c - e$ and $x^* = 1 - e/c$ in Levins' case.

The classic island model tells us that persistence occurs when a particular stability condition is satisfied whilst extinction occurs when another condition is satisfied. Thus the term *persistence* has been introduced in the context of equilibrium stability criteria. For a metapopulation described by the classic mainland model (3.1), the population persists regardless due to a unique positive equilibrium point. The classic mainland-island model (3.2), given here again by

$$\frac{dx}{dt} = cx(1-x) + c_0(1-x) - ex,$$

has a unique positive and stable equilibrium point x^* provided $c \geq 0$ and $c_0 > 0$, this being the unique positive solution to $cx^2 - (c - c_0 - e)x - ex = 0$. The metapopulation persists in this case with equilibrium proportion x^*.

In a stochastic setting, population persistence is defined in reference to an absorbing or non-absorbing extinction state. For example, metapopulations described by our chain binomial mainland model persist because state 0 is non-absorbing; the mainland effectively rescues the metapopulation from extinction in this case. For population models defined with an absorbing extinction state, the population eventually goes extinct although it may persist for some time before this occurs.

In this chapter, I shall modify the chain binomial structure introduced in the previous chapter and demonstrate how the modifications can be used to construct stochastic island models and stochastic mainland-island models. The new structure proves to be less tractable than the stochastic mainland case, however we are still able to determine a number of explicit results. I also note connections with discrete-time versions of classic metapopulation models.

4.2 The model

Let $(n_t : t \geq 0)$ be a discrete-time Markov chain taking values in the set $S_{\scriptscriptstyle N} = \{0, 1, 2, \ldots, N\}$ where n_t models the number of occupied patches in a metapopulation network with N patches. We assume the same seasonal dynamic as that outlined in the previous chapter and so $P = EC$ or $P = CE$, as before. The local extinction process and the colonisation process are defined as follows.

Local extinction process. The assumptions about this process remain unchanged: the parameter $e \in (0, 1)$ is the local extinction probability and the transition probabilities for this phase are given by (3.3).

Colonisation process. Given that i patches are occupied at the beginning of the colonisation phase, we assume that each of the $N - i$ empty patches are colonised independently with probability $c(i)$ during that phase. Hence, the number of empty patches colonised during this phase follows the $\mathrm{Bin}(N - i, c(i))$ law and so

$$c_{ij} = \begin{cases} \binom{N-i}{j-i} \left(1 - c(i)\right)^{N-j} c(i)^{j-i} & \text{if } j = i, i+1, \ldots, N \\ 0 & \text{if } j < i. \end{cases} \tag{4.3}$$

We assume that the function c is continuous, increasing and concave, with $c(0) \geq 0$ and $c(N) \leq 1$. This means that c increases as i increases and does so as a 'law of decreasing returns' (where the colonisation probability increases by less and less with each additional occupied patch). Note that if $c(0) = 0$, then $c_{0j} = \delta_{0j}$.

Transition probabilities. The transition probabilities that define the matrices $P = EC$ and $P = CE$ are given by (3.5a) and (3.5b), respectively, but with c_{ij} as defined by (4.3).

Stationary and quasi-stationary distributions. If $c(0) > 0$, then $c_{00} < 1$ and, hence, $p_{00} < 1$ for both models. In this case, $S_{\scriptscriptstyle N}$ forms an irreducible aperiodic class and the Markov chain has a unique stationary, and limiting, distribution (Section 2.2.2). If $c(0) = 0$ instead, then $e_{0j} = c_{0j} = \delta_{0j}$ and, hence, $p_{0j} = \delta_{0j}$ for both models. State 0 is absorbing and $S_N = \{0\} \cup E_N$, where $E_N = \{1, 2, \ldots, N\}$ is an irreducible transient class from which state 0 is accessible. In this case, the Markov chain has a unique quasi-stationary distribution (Section 2.2.3).

Chain binomial structure. The Markov chain $(n_t, t \geq 0)$ admits the following structure:

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \text{Bin}(N - \tilde{n}_t, c(\tilde{n}_t)) \qquad \tilde{n}_t = n_t - \text{Bin}(n_t, e) \tag{4.4a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \text{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \text{Bin}(N - n_t, c(n_t)). \tag{4.4b}$$

Thus, the colonisation process is clearly state dependent. The state-independent case corresponds to $c(i) = c_0 \in (0, 1]$, which was investigated separately in the previous chapter.

4.3 The colonisation probability

It was defined above that colonisation events occur with probability $c(i)$ where i is the current number of occupied patches. What form, then, should this probability take? Here I explore a few examples.

If the colonisation probability is assumed to increase linearly with i then

$$c(i) = \frac{ci}{N} \qquad (0 < c \leq 1), \tag{4.5}$$

for example, where the constant c is set to lie in the interval $(0, 1]$ so that $c(i)$ obeys the conditions stated in the previous section. Notice that $c(1) = c/N$ and $c(N) = c$. Hence each occupied patch contributes c/N to the colonisation probability and the constant c may be thought of as the (hypothetical) probability that a fully occupied network colonises a single empty patch.

If propagules are assumed to arrive at each empty patch according to a homogeneous Poisson process with rate $\beta i/N$, where i is the current number of occupied patches, then

$$c(i) = 1 - \exp\left(-\frac{\beta i}{N}\right) \qquad (\beta > 0), \tag{4.6}$$

where $c(i)$ is equal to the probability that at least one propagule arrives at an empty patch (see [46, 50]). In other words, colonising individuals are assumed to propagate from each occupied patch at rate β.

Another 'law of decreasing returns' is given by

$$c(i) = c^* \left(1 - \left(1 - \frac{c_1}{c^*}\right)^i\right) \qquad (0 < c_1 \leq c^* \leq 1), \tag{4.7}$$

similar to that used in [73]. Notice that c_1 is the colonisation probability given that only one patch is currently occupied and c^* is the limiting colonisation probability since $c(N) \to c^*$ as $N \to \infty$.

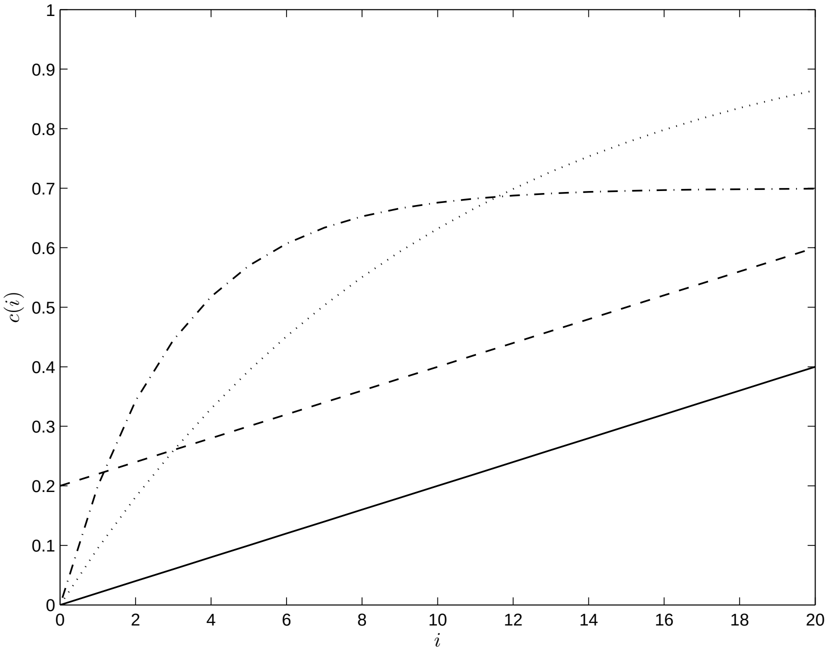

Figure 4.1: A plot of the colonisation functions given by (4.5), (4.6), (4.7) and (4.8) for $N = 20$: the linear case (–) with $c = 0.4$; the Poisson case ($\cdots$) with $\beta = 2$; the third law of decreasing returns ($\cdot$- - -) with $c_1 = 0.2$, $c^* = 0.7$; and the mainland-island case (- -) with $c = 0.4$, $c_0 = 0.2$.

The examples given so far satisfy $c(0) = 0$. Given this property, the chain binomial models defined by (4.4) have a sole absorbing state (state 0) and are examples of *island models*. In this case, there is no mainland to rescue the patch network from total extinction and so migrants originate from established populations (occupied patches) only. If one wishes to account for a mainland and therefore construct a *mainland-island model*, then the colonisation probability $c(i)$ can simply be a combination of a state-dependent case with the state-independent case. For example,

$$c(i) = \frac{ci}{N} + c_0 \qquad (c \geq 0,\ c_0 > 0,\ c + c_0 \leq 1), \tag{4.8}$$

is a probability that describes colonisation in a mainland-island network where ci/N is the component due to the islands (care of equation (4.5)) and c_0 is the component due to the mainland. Hence, mainland-island models satisfy the condition $c(0) > 0$. Figure 4.1 illustrates the four different colonisation probabilities outlined above.

4.4 Model analysis

First, we determine the 1-step distribution of n_{t+1} as done in Section 3.3. Let G be the probability generating function of n_{t+1} conditional on $n_t = i$, as before. For the EC model, $n_{t+1} = \tilde{n}_t + \tilde{B}$, where $\tilde{n}_t \sim \mathrm{Bin}(i, 1-e)$ and $\tilde{B} \sim \mathrm{Bin}(N - \tilde{n}_t, c(\tilde{n}_t))$, and so

$$\begin{aligned} G(z) &= \mathbb{E}(z^{n_{t+1}} | n_t = i) = \mathbb{E}\left(z^{\tilde{n}_t}\ \mathbb{E}(z^{\tilde{B}} | \tilde{n}_t, n_t = i) | n_t = i\right) \\ &= \mathbb{E}\left(z^{\tilde{n}_t}\left(1 - c(\tilde{n}_t) + c(\tilde{n}_t)z\right)^{N-\tilde{n}_t} | n_t = i\right). \end{aligned} \tag{4.9}$$

For the CE model, $n_{t+1} \sim \mathrm{Bin}(i + \tilde{B}, 1-e)$ where $\tilde{B} \sim \mathrm{Bin}(N - i, c(i))$, and so

$$\begin{aligned} G(z) &= \mathbb{E}\left(\mathbb{E}(z^{n_{t+1}} | \tilde{B}, n_t = i) | n_t = i\right) = \mathbb{E}\left((e + (1-e)z)^{i+\tilde{B}} | n_t = i\right) \\ &= \left(e + (1-e)z\right)^i\ \mathbb{E}\left((e + (1-e)z)^{\tilde{B}} | n_t = i\right) \\ &= \left(e + (1-e)z\right)^i \left(1 - c(i)(1-e) + c(i)(1-e)z\right)^{N-i}. \end{aligned} \tag{4.10}$$

The expectation inside expression (4.9) cannot be evaluated any further and so the conditional distribution of n_{t+1} remains unknown in the EC case. Fortunately, the expectation inside (4.10) could be evaluated and the end result is in the form of two binomial pgfs multiplied together. This gives rise to the following proposition.

Proposition 4.1 *Given $n_t = i$, then, for the CE case only, the distribution of n_{t+1} is equal to the sum of two independent binomial random variables B_1 and B_2 where $B_1 \sim \mathrm{Bin}(i, 1-e)$ and $B_2 \sim \mathrm{Bin}(N - i, c(i)(1-e))$.*

We may describe the CE model in the following way: each of the i occupied patches behave as though they survive independently of one another while the $N - i$ empty patches behave as though they are colonised independently and each with probability $c(i)(1-e)$. We may think of $c(i)(1-e)$ as the 'effective' state-dependent colonisation probability. Of course, Proposition 4.1 remains true for CE models with a state-independent colonisation probability (see Lemma 3.1).

Remark. Thus the 1-step conditional state distribution for both the EC and CE mainland models (Lemma 3.1) as well as the state-dependent CE model (Proposition 4.1) are equal to the sum of two independent binomial random variables. Indeed, in all three cases, the independent binomial random variables are expressed in such a way that we may view extinction events and colonisation events as though they occur independently of one another and without abiding by a seasonal dynamic (hence the terms *effective local extinction probability* and *effective colonisation probability*). It is unfortunate that we could not obtain the 1-step conditional state distribution for the state-dependent EC model as well since this prevents us from making a thorough analytical comparison between state-independent and state-dependent metapopulation models.

Numerical calculations show that the 1-step conditional state distribution for the EC model is similar in shape to that for the CE model (see Figure 4.2 for example), however these calculations do not tell us whether or not the distribution in the EC case is equal to a sum of independent random variables. We can draw other conclusions though, such as we should expect to observe more occupied patches when the census is taken after each colonisation phase because the mode is generally larger in the EC case.

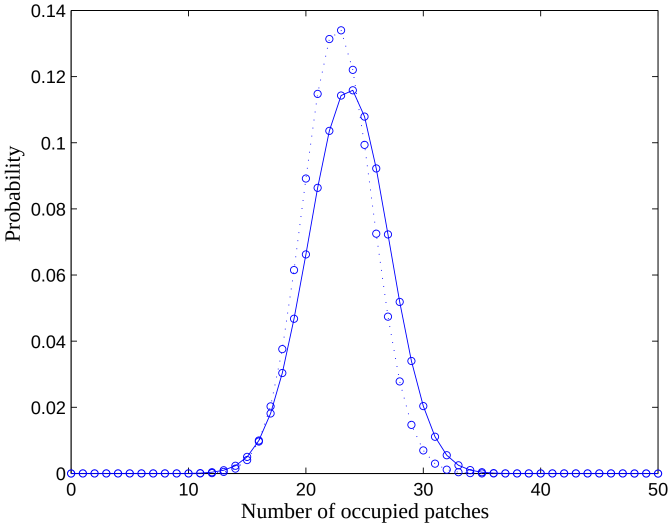

Figure 4.2: The distribution at time $t+1$ conditional on $n_t = 20$ for the chain binomial EC model (solid line) and CE model (dotted line) with $c(i) = ci/N$; $e = 0.2$, $c = 0.8$ and $N = 50$.

Since we have evaluated the 1-step conditional state distribution for the CE model, we can write

$$G_{t+1}(z) = \mathbb{E}\big((e + (1-e)z)^{n_t}(1 - c(n_t)(1-e)(1-z))^{N-n_t} | n_0 = i\big) \tag{4.11}$$

in a bid to solve for G_t, the pgf of n_t conditional on $n_0 = i$, explicitly. The state-independent case was solved successfully in the previous chapter (see equation (3.8)), however the same calculations cannot be performed here since the state-dependent case requires the factor $c(n_t)$ to be subject to the expectation on the right-hand side of equation (4.11). This means that we cannot determine the t-step conditional state distribution or prove limit theorems by way of observing the limiting behaviour of the t-step conditional state distribution (like that done for Theorems 3.2 and 3.3). Instead, we shall use the 1-step conditional state distribution and Mathematical Induction to prove limit theorems for the state-dependent CE model.

4.4.1 Limit theorems for the CE model

From Proposition 4.1 we deduce that $\mathbb{E}(n_{t+1}|n_t) = (1-e)n_t + c(n_t)(1-e)(N-n_t)$, and hence

$$\mathbb{E}(X_{t+1}^N|X_t^N) = (1-e)X_t^N + c(n_t)(1-e)(1-X_t^N),$$

where $X_t^N = n_t/N$ is the proportion of occupied patches at time t, as before. If the function c was such that

$$c(n_t) = \hat{c}(n_t/N),$$

where the function $\hat{c}$ is assumed to be continuous, increasing and concave, with $\hat{c}(0) \geq 0$ and $\hat{c}(x) \leq 1$, then one might expect $\mathbb{E}(X_{t+1}^N|X_t^N)$ to converge to $(1-e)x_t + \hat{c}(x_t)(1-e)(1-x_t)$ as $N \to \infty$, where x_t represents the value of X_t^N in this limit. The following theorem establishes a law of large numbers for the CE model.

Theorem 4.1 *Consider the CE model with parameters e and $c(n_t)$ and further suppose that $c(n_t) = \hat{c}(n_t/N)$, with $\hat{c}$ as defined above. If $X_0^N \xrightarrow{\text{P}} x_0$ (a constant) as $N \to \infty$ then, for any $t \geq 1$, $X_t^N \xrightarrow{\text{P}} x_t$ in this limit where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ with f given by*

$$f(x) = (1-e)(x + (1-x)\hat{c}(x)). \tag{4.12}$$

Proof. We will use Mathematical Induction. By assumption, $X_0^N \xrightarrow{\text{P}} x_0$. Now, suppose that $X_t^N \xrightarrow{\text{P}} x_t$ as $N \to \infty$ for some $t \geq 0$. From Proposition 4.1, we know that $X_{t+1}^N \overset{\text{D}}{=} K^N + L^N$, where $K^N = B_1/N = X_t^N B_1/n_t$ and $L^N = B_2/N = (1-X_t^N)B_2/(N-n_t)$ are independent random variables. Since $X_t^N \xrightarrow{\text{P}} x_t$, in which case $n_t \to \infty$, it follows from the standard Weak

Law of Large Numbers that $B_1/n_t \xrightarrow{\mathrm{P}} 1-e$, and hence $K^N \xrightarrow{\mathrm{P}} x_t(1-e)$. We shall prove a similar result for L^N but with the use of characteristic functions. Let α^N be the characteristic function (cf) of L^N and let β^N be the cf of B_2, that is, $\alpha^N(\omega) = \mathbb{E}\exp(\mathbf{i}\omega L^N)$ and $\beta^N(\omega) = \mathbb{E}\exp(\mathbf{i}\omega B_2)$. Then

$$\begin{aligned}\alpha^N(\omega) &= \mathbb{E}\exp(\mathbf{i}\omega B_2/N) = \beta^N(\omega/N)\\ &= \big(1-\hat{c}(X_t^N)(1-e)+\hat{c}(X_t^N)(1-e)\exp(\mathbf{i}\omega/N)\big)^{N(1-X_t^N)}\end{aligned}$$

and, since $\exp(\mathbf{i}\omega/N) = 1+\mathbf{i}\omega/N+o(1/N)$ as $N\to\infty$ by Taylor's Theorem, we may thus write

$$\alpha^N(\omega) = \left(1+\hat{c}(X_t^N)(1-e)\frac{\mathbf{i}\omega}{N}+o(1/N)\right)^{N(1-X_t^N)}.$$

We are told that $\hat{c}$ is continuous, but we also know that $\hat{c}(X_t^N)$ is almost surely uniformly bounded because $0 \le X_t^N \le 1$ and $\hat{c}(x) \le 1$. Therefore $\hat{c}(X_t^N) \xrightarrow{\mathrm{P}} \hat{c}(x_t)$ (a constant) as $N\to\infty$, in which case $\alpha^N(\omega) \to \exp\big((1-x_t)\hat{c}(x_t)(1-e)\mathbf{i}\omega\big)$ and it follows that $L^N \xrightarrow{\mathrm{D}} (1-x_t)\hat{c}(x_t)(1-e)$. Since the limiting value is a constant, we may then write $L^N \xrightarrow{\mathrm{P}} (1-x_t)\hat{c}(x_t)(1-e)$. Hence $K^N+L^N \xrightarrow{\mathrm{P}} x_t(1-e)+(1-x_t)\hat{c}(x_t)(1-e)$, implying that $X_{t+1}^N \xrightarrow{\mathrm{P}} x_{t+1}$ where $x_{t+1} = f(x_t)$ with f as given by (4.12). This completes the proof. □

Notice that Theorem 4.1 applies to CE models with colonisation probabilities given by (4.5), (4.6) or (4.8), for example, because all three can be expressed in terms of the proportion of occupied patches and satisfy the condition $c(\,\cdot\,) = \hat{c}(\,\cdot\,/N)$. The probability defined by (4.7) cannot be expressed in this way and so the limiting behaviour of the resulting CE model cannot be determined in the format used here. In Section 4.5, we examine a number of chain binomial models with colonisation probabilities of the form $\hat{c}(x)$.

Now that we have identified the deterministic process (x_t), we wish to examine the scaled fluctuation Z_t^N about this trajectory: $Z_t^N = \sqrt{N}(X_t^N - x_t)$. From Proposition 4.1, we write

$$\begin{aligned}X_{t+1}^N - x_{t+1} &= (B_1+B_2)/N - (1-e)(x_t-(1-x_t)\hat{c}(x_t))\\ &= X_t^N\frac{B_1}{n_t} + (1-X_t^N)\frac{B_2}{N-n_t} - (1-e)(x_t-(1-x_t)\hat{c}(x_t))\\ &= X_t^N\left(\frac{B_1}{n_t}-(1-e)\right) + (1-X_t^N)\left(\frac{B_2}{N-n_t}-\hat{c}(X_t^N)(1-e)\right) + f(X_t^N) - f(x_t),\end{aligned}$$

where $B_1 \sim \mathrm{Bin}(n_t, 1-e)$ and $B_2 \sim \mathrm{Bin}(N-n_t, \hat{c}(n_t/N)(1-e))$. On multiplying the last expression by $\sqrt{N}$, we have

$$Z_{t+1}^N = W_t^N + \sqrt{N}\big(f(X_t^N) - f(x_t)\big) \tag{4.13}$$

where

$$W_t^N = \sqrt{X_t^N}\sqrt{n_t}\left(\frac{B_1}{n_t} - (1-e)\right) + \sqrt{1-X_t^N}\sqrt{N-n_t}\left(\frac{B_2}{N-n_t} - \hat{c}(X_t^N)(1-e)\right). \tag{4.14}$$

In preparation for the next central limit law, we present the following result.

Lemma 4.1 *If $X_t^N \xrightarrow{\mathrm{P}} x$ (a constant) as $N \to \infty$, then $W_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(0, v(x))$ where*

$$v(x) = (1-e)\big[ex + \hat{c}(x)\big(1 - \hat{c}(x)(1-e)\big)(1-x)\big]. \tag{4.15}$$

Proof. We are told that, for fixed t and x, $X_t^N \xrightarrow{\mathrm{P}} x$ as $N \to \infty$, and so $n_t \to \infty$. Consider the first term in equation (4.14). It follows from the standard Central Limit Theorem that $\sqrt{n_t}(B_1/n_t - (1-e)) \xrightarrow{\mathrm{D}} \mathrm{N}(0, e(1-e))$ and hence $\sqrt{X_t^N}\sqrt{n_t}(B_1/n_t - (1-e)) \xrightarrow{\mathrm{D}} \mathrm{N}(0, xe(1-e))$. For the second term in equation (4.14), we shall use characteristic functions to show that this term converges in distribution to a normal random variable with mean 0 and variance $(1-x)\hat{c}(x)(1-e)(1-\hat{c}(x)(1-e))$. Letting α^N be the cf of the second term, that is

$$\alpha^N(\omega) = \mathbb{E}\exp\left(\mathbf{i}\omega\sqrt{1-X_t^N}\sqrt{N-n_t}\big(B_2/(N-n_t) - \hat{c}(X_t^N)(1-e)\big)\right),$$

and letting β^N be the cf of B_2, we have

$$\begin{aligned}
\alpha^N(\omega) &= \exp\left(-\hat{c}(X_t^N)(1-e)\mathbf{i}\omega\sqrt{N}(1-X_t^N)\right)\beta^N\left(\omega/\sqrt{N}\right)\\
&= \exp\left(-\hat{c}(X_t^N)(1-e)\mathbf{i}\omega\sqrt{N}(1-X_t^N)\right)\left(1-\hat{c}(X_t^N)(1-e) + \hat{c}(X_t^N)(1-e)\exp(\mathbf{i}\omega/\sqrt{N})\right)^{N-n_t}\\
&= \left(\exp\big(-\hat{c}(X_t^N)(1-e)\mathbf{i}\omega/\sqrt{N}\big)\left(1-\hat{c}(X_t^N)(1-e) + \hat{c}(X_t^N)(1-e)\exp\left(\mathbf{i}\omega/\sqrt{N}\right)\right)\right)^{N(1-X_t^N)}.
\end{aligned}$$

By Taylor's Theorem, $\exp(\mathbf{i}\omega/\sqrt{N}) = 1 + \mathbf{i}\omega/\sqrt{N} - \omega^2/N + o(1/N)$ and

$$\exp\left(-\hat{c}(X_t^N)(1-e)\mathbf{i}\omega/\sqrt{N}\right) = 1 - \hat{c}(X_t^N)(1-e)\mathbf{i}\omega/\sqrt{N} + \hat{c}(X_t^N)^2(1-e)^2\omega^2/N + o(1/N),$$

both as $N \to \infty$, and we may thus write

$$\alpha^N(\omega) = \Big(1 - \hat{c}(X_t^N)(1-e)\big(1 - \hat{c}(X_t^N)(1-e)\big)\frac{\omega^2}{2N} + o(1/N)\Big)^{N(1-X_t^N)}.$$

From the proof for Theorem 4.1, we know that $\hat{c}(X_t^N) \xrightarrow{\text{P}} \hat{c}(x)$ as $N \to \infty$ and so

$$\alpha^N(\omega) \to \exp\left(-(1-x)\hat{c}(x)(1-e)(1-\hat{c}(x)(1-e))\omega^2/2\right),$$

from which we deduce that

$$\sqrt{1-X_t^N}\sqrt{N-n_t}\,(B_2/(N-n_t) - (1-e)) \xrightarrow{\text{D}} \text{N}\left(0, \hat{c}(x)(1-e)(1-\hat{c}(x)(1-e))(1-x)\right),$$

as desired. Therefore, the first and second terms in equation (4.14) define independent random variables that converge in distribution to normal random variables, both with zero mean but with variance $xe(1-e)$ in the former case and variance $c(x)(1-e)(1-c(x)(1-e))(1-x)$ in the latter case. Therefore, for fixed t and x, it has been shown that $W_t^N \xrightarrow{\text{D}} \text{N}(0, v(x))$, where v is that given by (4.15). □

We now present a central limit law for the CE model. Recall that empty products are to be interpreted as being equal to 1.

Theorem 4.2 *For the CE model with parameters e and $\hat{c}(x)$, suppose that $\hat{c}$ is twice continuously differentiable and that, for $t \geq 0$, $X_t^N \xrightarrow{\text{P}} x_t$ as $N \to \infty$ where $x_\bullet$ satisfies $x_{t+1} = f(x_t)$ with f as given by* (4.12)*. If $Z_0^N \xrightarrow{\text{D}} z_0$ (a constant) as $N \to \infty$, then $Z_t^N \xrightarrow{\text{D}} \text{N}(\mu_t, V_t)$ for any $t \geq 1$, where*

$$\mu_t = z_0 \prod_{s=0}^{t-1} f'(x_s) \tag{4.16}$$

and

$$V_t = \sum_{s=0}^{t-1} v(x_s) \prod_{u=s+1}^{t-1} f'(x_u)^2, \tag{4.17}$$

with v as given by equation (4.15).

Proof. Recall that Z_{t+1}^N admits the representation $Z_{t+1}^N = \sqrt{N}\big(f(X_t^N) - f(x_t)\big) + W_t^N$. We are told that $\hat{c}(x)$ is twice continuously differentiable in x and so $f(x)$ is also twice continuously differentiable in x. By Taylor's Theorem, we have that

$$f(X_t^N) = f(x_t) + f'(x_t)(X_t^N - x_t) + f''(\theta_t^N)(X_t^N - x_t)^2/2$$

for some θ_t^N between X_t^N and x_t. On multiplying by $\sqrt{N}$, the last expression becomes

$$\sqrt{N}\big(f(X_t^N) - f(x_t)\big) = f'(x_t)Z_t^N + \frac{f''(\theta_t^N)(Z_t^N)^2}{2\sqrt{N}}.$$

But $f''(x)$ is bounded in x (since $0 \le X_t^N \le 1$ and $\hat{c}(x) \le 1$) and we may thus write

$$\sqrt{N}\big(f(X_t^N) - f(x_t)\big) = f'(x_t)Z_t^N + o_t^N(1),$$

where $o_t^N(1) \xrightarrow{\mathrm{P}} 0$ $(t \ge 0)$ as $N \to \infty$. Hence

$$Z_{t+1}^N = f'(x_t)Z_t^N + W_t^N + o_t^N(1). \tag{4.18}$$

We shall now use Mathematical Induction to show that $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(\mu_t, V_t)$ for all $t \ge 1$, where the parameters μ_t and V_t are given by (4.16) and (4.17), respectively. We are told that $X_0^N \xrightarrow{\mathrm{P}} x_0$ and $Z_0^N \xrightarrow{\mathrm{D}} z_0$ (a constant) and so, for the $t = 1$ case, we have $Z_1^N = f'(x_0)Z_0^N + W_0^N + o_0^N(1)$ from equation (4.18). Therefore, as $N \to \infty$,

$$Z_1^N \xrightarrow{\mathrm{D}} f'(x_0)z_0 + \mathrm{N}(0, v(x_0)) = \mu_1 + \mathrm{N}(0, V_1) = \mathrm{N}(\mu_1, V_1),$$

since $W_0^N \xrightarrow{\mathrm{D}} \mathrm{N}(0, v(x_0))$ by Lemma 4.1 and $o_0^N(1) \xrightarrow{\mathrm{P}} 0$. Now suppose that $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(\mu_t, V_t)$ for some $t \ge 1$ (noting that $X_t^N \xrightarrow{\mathrm{P}} x_t$ by assumption). It then follows from (4.18) that

$$\begin{aligned} Z_{t+1}^N &\xrightarrow{\mathrm{D}} f'(x_t)\mathrm{N}(\mu_t, V_t) + \mathrm{N}(0, v(x_t)) \\ &= \mathrm{N}(f'(x_t)\mu_t, f'(x_t)^2 V_t) + \mathrm{N}(0, v(x_t)) \\ &= \mathrm{N}(\mu_{t+1}, V_{t+1}), \end{aligned}$$

since, for fixed t and x_t, $W_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(0, v(x_t))$ by Lemma 4.1 and $o_t^N(1) \xrightarrow{\mathrm{P}} 0$. This completes the proof. □

Once again, we may assess the quality of the deterministic approximation x_t from a central limit law. That is, for N sufficiently large, the central limit law for the CE model implies that $\Pr(|X_t^N - x_t| > \epsilon_t) \simeq 2(1 - \Phi(\epsilon_t \sqrt{N/V_t}))$, where Φ is the standard normal distribution function and V_t is that given by (4.17).

4.4.2 Long-term behaviour of the CE model

The long-term behaviour of the stochastic process $(n_t, t \geq 0)$ is examined by way of the long-term behaviour of the deterministic CE model. First, we present stability criteria for the fixed points of the deterministic process and then use the criteria to describe the stationary or quasi-stationary behaviour of the stochastic process.

From Theorem 4.1, recall that the deterministic model (x_t) is defined by $x_{t+1} = f(x_t)$ with x_0 given and the map f defined by $f(x) = (1-e)(x + (1-x)\hat{c}(x))$. Notice that x^* is a fixed point of f if and only if

$$\hat{c}(x^*) = \left(\frac{e}{1-e}\right)\frac{x^*}{1-x^*}$$

but we shall write $\hat{c}(x^*) = r(x^*)$ for convenience where $r(x) = \rho x/(1-x)$ and $\rho = e/(1-e)$. Notice, also, that the function r has slope ρ at $x = 0$ and increases strictly from 0 to ∞ whilst the function $\hat{c}$ is strictly increasing from $\hat{c}(0) \geq 0$ and concave with $\hat{c}(1) \leq 1$. Therefore, we always have precisely one stable fixed point and x_t approaches this point monotonically. The value of $\hat{c}(x)$ at $x = 0$, as well as the slope at this point, determines the number of fixed points in the interval $[0, 1]$. Using theorems from Section 2.5, we describe the stability of these fixed points as follows.

- If $c(0) = 0$ and $\hat{c}'(0) > \rho$, then there are two fixed points in the interval $[0, 1]$, namely 0 and $x^* \in (0, 1)$. We have $f'(0) = (1-e)(1 + \hat{c}'(0)) > 1$ and so 0 is unstable. The slope evaluated at x^* clearly satisfies $\hat{c}'(x^*) < r'(x^*)$ and we find that $f'(x^*) < 1$. Hence x^* is stable.

- If $c(0) = 0$ and $\hat{c}'(0) \leq \rho$, then 0 is the unique fixed point in the interval $[0, 1]$. For the case where $\hat{c}'(0) < \rho$, we have $f'(0) < 1$. For the case where $\hat{c}'(0) = \rho$, we have $f'(0) = 1$ but $f''(0) = (1-e)(\hat{c}''(0) - 2\hat{c}'(0)) < 0$. Hence 0 is stable in both cases.

- Lastly, if $c(0) > 0$, there is a unique fixed point x^* in the interval $[0, 1]$ and this satisfies $x^* \in (0, 1)$. We know that x^* is stable because $\hat{c}'(x^*) < r'(x^*)$ and so $f'(x^*) < 1$.

We can now describe the long-term behaviour of the stochastic process $(n_t, t \geq 0)$. Recall that if $\hat{c}(0) = 0$, then the stochastic process has an absorbing state (state 0) and there exists a unique quasi-stationary distribution. If, in addition, $\hat{c}'(0) \leq \rho$, we expect the process to be absorbed quickly (even for N quite large), while, if $\hat{c}'(0) > \rho$, then the process may reach a 'quasi equilibrium' before absorption occurs, where we expect the unique quasi-stationary distribution to be centered near Nx^*. Hence, the metapopulation dies out quickly in the former case or persists for some time before eventually going extinct in the latter case. The two sets of conditions define what we shall call *evanescence* and *quasi stationarity*, respectively. If, instead, $\hat{c}(0) > 0$, then the stochastic process has a unique stationary distribution and we expect this distribution to be centred near Nx^*. The metapopulation persists in this case and so the third condition outlines what we shall refer to as *stationarity*.

Indeed, we may describe the asymptotic ($N \to \infty$) behaviour of stochastic processes that are in equilibrium. On setting $x_0 = x^*$ in Theorems 4.1 and 4.2, the following corollary establishes a law of large numbers as well as a central limit law for CE models that have a positive stable deterministic equilibrium x^*.

Corollary 4.1 *Consider the CE model with parameters e and $\hat{c}(x)$. Suppose that $\hat{c}$ is twice continuously differentiable and that either (i) $\hat{c}(0) > 0$ or (ii) $\hat{c}(0) = 0$ and $\hat{c}'(0) > e/(1-e)$. Let x^* be a positive stable fixed point of f where f is as given in Theorem 4.1. If $X_0^N \xrightarrow{\mathrm{P}} x^*$ as $N \to \infty$, then $X_t^N \xrightarrow{\mathrm{P}} x^*$ for all $t \geq 1$. Let $Z_t^N = \sqrt{N}(X_t^N - x^*)$. If $Z_0^N \xrightarrow{\mathrm{D}} Z_0$ as $N \to \infty$, in addition to $X_t^N \xrightarrow{\mathrm{P}} x^*$ $(t \geq 1)$, then $Z_t^N \xrightarrow{\mathrm{D}} f'(x^*)^t Z_0 + E_t$ for any $t \geq 1$, where $E_t \sim \mathrm{N}(0, V_t)$ with $V_t = v(x^*)(1 - f'(x^*)^{2t})/(1 - f'(x^*)^2)$ and v as given in Lemma 4.1.*

4.5 Examples

Here we apply our limit theorems to a number of chain binomial CE models. The models differ in terms of the function $\hat{c}$ used to describe the colonisation process and their long-term behaviour is described by using the stability criteria outlined in the previous section.

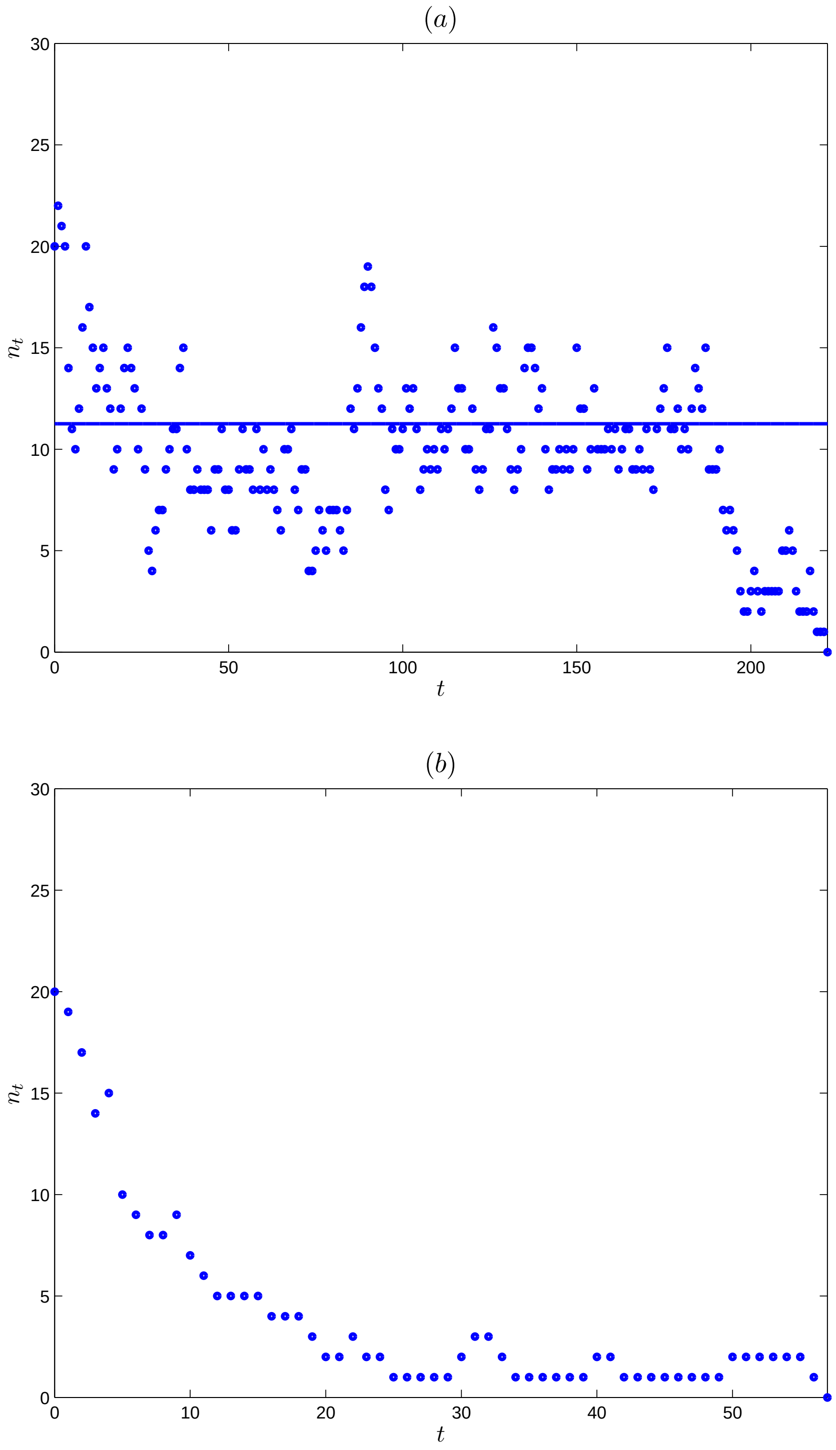

Figure 4.3: Simulations of the CE model with $N = 30$, $n_0 = 20$, $e = 0.2$ and $\hat{c}(x) = cx$: examples of (a) quasi stationarity, where $c = 0.4$ and $Nx^* = 30 \times 0.375 = 11.25$, with Nx^* marked by the solid line, and (b) evanescence, where $c = 0.2$.

Example 4.1 Consider the chain binomial CE model with $\hat{c}(x) = cx$ $(0 < c \leq 1)$. Theorem 4.1 tells us that the deterministic trajectory is defined by $x_{t+1} = (1-e)(1+c-cx_t)x_t$ but before we examine the fixed points of this difference equation, notice that the equation may be written as

$$x_{t+1} - x_t = c(1-e)x_t(1-x_t) - ex_t, \tag{4.19}$$

which is immediately identified as a discrete-time version of Levins' model (4.1) with $c(1-e)$ as the 'effective' colonisation rate parameter and e as the extinction rate parameter. Indeed, (4.19) satisfies the discrete-time logistic growth model, namely

$$x_{t+1} - x_t = rx_t\left(1 - \frac{x_t}{x^*}\right)$$

(see Section 3.2 of [75], for example), with $r = c(1-e) - e$ and $x^* = r/\big(c(1-e)\big)$. Hence, there are two fixed points, 0 and x^*, which are subject to the conditions $\hat{c}(0) = 0$ and $\hat{c}'(0) = c$. If $c > \rho$ (that is, $r > 0$), then 0 is unstable and x^* is stable. If $c \leq \rho$ (that is, $r \leq 0$), 0 is the unique stable fixed point in $[0, 1]$. Although the discrete-time logistic growth model is known to exhibit a wide range of dynamic behaviour, note that the deterministic model given by (4.19) does not exhibit any chaotic or periodic behaviour since $0 < 1+r = (1-e)(1+c) < 2$ (for example, see [59]). Hence, we have quasi stationarity if $c > e/(1-e)$ (see Figure 4.3(a)) and evanescence if $c \leq e/1-e$ (see Figure 4.3(b)). From Theorem 4.2, the function v is given by $v(x) = (1-e)x[e + c(1-x)(1-c(1-e)x)]$ and from Corollary 4.1, if $Z_0^N \xrightarrow{\mathrm{D}} Z_0$, then $Z_t^N \xrightarrow{\mathrm{D}} (1-r)^t Z_0 + \mathrm{N}(0, er(2-e-r)/(e+r))$ for the quasi-equilibrium case $(r > 0)$.

Example 4.2 For the case where $\hat{c}(x) = cx + c_0$ $(c \geq 0, c_0 > 0, c + c_0 \leq 0)$, Theorem 4.1 tells us that the deterministic trajectory is defined by $x_{t+1} = (1-e)x_t + (1-e)(cx_t + c_0)(1-x_t)$. We have that

$$x_{t+1} - x_t = c(1-e)x_t(1-x_t) + c_0(1-e)(1-x_t) - ex_t$$

which, unsurprisingly, is a discrete-time version of the classic mainland-island model (3.2) with $c(1-e)$ and $c_0(1-e)$ as the 'effective' colonisation rate parameters due to the island network and the mainland, respectively, and e as the extinction rate parameter. Since $\hat{c}(0) = c_0 > 0$, the stochastic process has a unique stationary distribution and we expect this distribution to be centred near Nx^* where x^* is the unique positive solution to

$$c(1-e)x^2 - ((c-c_0)(1-e)-e)x - c_0(1-e) = 0.$$

Similarly, the function v is determined from Theorem 4.2 and is given by

$$v(x) = (1-e)[ex + (1-x)(c_0+cx)(1-(1-e)(c_0+cx))].$$

Example 4.3 Consider the chain binomial CE model with $\hat{c}(x) = 1-\exp(-\beta x)$ $(\beta > 0)$. Since $\hat{c}(0) = 0$ and $\hat{c}'(0) = \beta > 0$, the stochastic process is absorbed quickly if $\beta \leq e/1-e$, or it reaches quasi equilibirium if $\beta > e/1-e$. In the latter case, we expect the quasi-stationary distribution to be centred near Nx^* where x^* is found by iterating the map $f(x) = (1-e)(1-(1-x)\exp(-\beta x))$ numerically. The function v is given by

$$v(x) = (1-e)[ex + (1-x)(1-\exp(-\beta x)(e+(1-e)\exp(-\beta x))].$$

Example 4.4 Of course, the limit theorems remain true for CE models with $\hat{c}(x) = c_0 > 0$. We have that $x_{t+1} = (1-e)(x_t + c_0(1-x_t))$ and $x^* = c_0(1-e)/(1-(1-e)(1-c_0))$ in this case, which is in complete agreement with the CE case of the chain binomial mainland model examined in Chapter 3. From Lemma 4.1, we have that

$$\begin{aligned} v(x) &= (1-e)\big(ex + c_0(1-c_0(1-e))(1-x)\big) \\ &= p(1-p)x + q(1-q)(1-x) \end{aligned}$$

where $p = 1-e$ and $q = c_0(1-e)$ from equation (3.7b). Recall that we evaluated the t-step conditional state distribution for the both the EC and CE chain binomial mainland models and that we used this distribution to establish $x_t = q^* + (x_0 - q^*)a^t$, where $a = p-q$ and $q^* = q/(1-a)$, and that $Z_t^N \overset{\mathrm{D}}{\to} a^t Z_0 + \mathrm{N}(0, V_t)$ whenever $Z_0^N \overset{\mathrm{D}}{\to} Z_0$, where $V_t = p_t(1-p_t)x_0 + q_t(1-q_t)(1-x_0))$ with $p_t = q_t + a^t$ and $q_t = q^*(1-a^t)$. These results agree with Theorem 4.2 because, for the CE case, $f'(x_t) = (1-e)(1-c_0) = a$ for all $t \geq 1$ and hence $\mu_t = a^t$. From equation (4.17), we see that

$$\begin{aligned} V_t &= \sum_{s=0}^{t-1} \big[p(1-p)x_s + q(1-q)(1-x_s)\big] a^{2(t-s-1)} \\ &= a^{2(t-1)} \sum_{s=0}^{t-1} \big[(p(1-p)-q(1-q))(q^* + (x_0-q^*)a^s) + q(1-q)\big] a^{-2s} \end{aligned}$$

from which we group terms involving s and solve the resulting geometric progressions. Observing that $p(1-p)-q(1-q) = a(1-a)(1-2q^*)$ and using the identity $q^* = q/(1-a)$, we then have

$$V_t = \frac{q(1-a^{2t})}{(1-a^2)}\big(a(1-2q^*) + (1-q)\big) + a^t(1-a^t)(1-2q^*)(x_0-q^*)$$

and so

$$V_t = \frac{q^*(1-a^t)}{(1+a)}\big[q(1+a^t)(a(1-2q^*)+1-q) - a^t(1+a)(1-2q^*)\big] + a^t(1-a^t)(1-2q^*)x_0.$$

The term in the square brackets reduces to $(1+a)(1-q^*(1-a^t))$. Upon using the identity $q_t = q^*(1-a^t)$ and observing that the coefficient of x_0 is equal to $p_t(1-p_t)-q_t(1-q_t)$, we have

$$\begin{aligned} V_t &= q_t(1-q_t) + \big(p_t(1-p_t) - q_t(1-q_t)\big)x_0 \\ &= p_t(1-p_t)x_0 + q_t(1-q_t)(1-x_0), \end{aligned}$$

which agrees with the expression for V_t as stated in Theorem 3.3.

4.6 Discussion

This chapter was concerned with chain binomial metapopulation models defined with state-dependent colonisation processes. We attempted to evaluate the 1-step conditional state distribution for both EC and CE models but found the EC case to be intractable. For the CE model, the 1-step conditional state distribution was found to be equal to the sum of two independent binomial random variables and so, by using Mathematical Induction, we proved a law of large numbers and a central limit law for models that depend on the current state of the system through the proportion of occupied patches. The law of large numbers identified a deterministic trajectory and, by analysing this trajectory, we determined explicit conditions that describe the stationarity of chain binomial mainland-island models as well as the quasi stationarity and evanescence of chain binomial island models.

A number of examples were investigated in this chapter, including several island models and one mainland-island model. Recall that the island model with $\hat{c}(x) = 1 - \exp(-\beta x)$ $(\beta > 0)$ is the metapopulation model proposed by Hill and Caswell [46]. Thus, in the CE case, all of the results presented in this chapter apply to Hill and Caswell's model. The model with $c(i) = c^*(1-(1-(c_1/c^*))^i)$ $(0 < c_1 \leq c^* \leq 1)$, similar to that used by Possingham [73], is another example of an island model though, since it does not exhibit a form in terms of the proportion $x(= i/N)$ of occupied patches, the CE limit theorems do not apply in this particular instance. Mainland-island models appear to be less popular in the applied metapopulation literature however future applications might lie in modelling endangered species that are supplemented with individuals from a captive breeding population. In such situations, a captive population would effectively act as a mainland population.

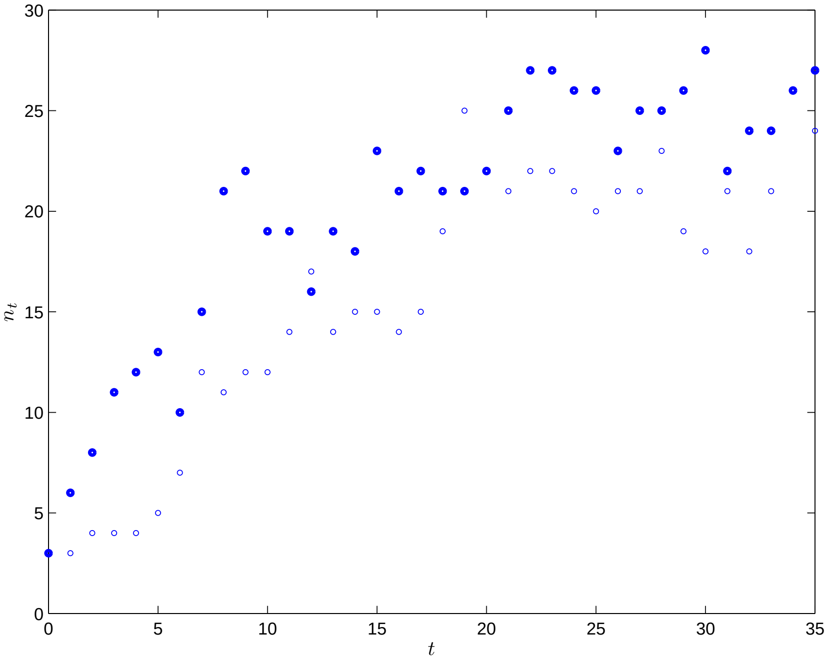

Figure 4.4: A simulation of the EC model (solid circles) and CE model (open circles) with $\hat{c}(x) = cx$; $N = 30$, $e = 0.2$ and $c = 0.6$.

Although we could not determine similar results for the EC model, simulation studies showed that this model appears to respond to the same stability criteria as that outlined above for the

CE model. For example, Figure 4.4 provides simulations for the quasi-stationary case with $\hat{c}(x) = cx$. However, the similarity is not surprising since the seasonal dynamic is not altered by the timing of the population census. Notice that the EC simulation appears to exhibit a slightly larger equilibrium value than the CE model, which is not surprising either since the EC process is observed immediately after the colonisation phase. This suggests that if one could prove limit theorems for EC models, then these theorems must be similar to those for the corresponding CE models and that the resulting limiting deterministic EC model exhibits similar stability criteria. Indeed, we revisit the chain binomial metapopulation structure in Chapter 6 where we prove limit theorems for both EC and CE models in a general framework and without the need for conditional state distributions.

CHAPTER 5

Infinite-patch Models

This chapter presents stochastic models for metapopulation networks with infinitely-many patches. These models, which may be viewed as infinite-patch analogues of our finite-patch chain binomial models, are shown to be equivalent to branching processes.

5.1 Introduction

The metapopulation models examined in Chapters 3 and 4 account for a seasonal dynamic where occupied patches experience local extinction during a particular seasonal phase whilst empty patches are colonised during another seasonal phase, the number of events that occur during each phase being governed by a binomial law. For these chain binomial models, the parameter N defines the total number of patches in the metapopulation network. A number of conditional state distributions were evaluated and then used to examine the models in the limit as $N \to \infty$, where, in particular, a deterministic trajectory was identified and used to obtain persistence criteria that describe the long-term behaviour of the stochastic process. How, then, would these results compare with stochastic models that describe metapopulation networks with infinitely-many habitat patches? In this chapter, we develop three different infinite-patch models (an island model, a mainland-island model and a mainland model, in that order) and compare their behaviour with that of their finite-patch counterparts.

5.2 An infinite-patch island model

Let n_t be the number of occupied patches at time t but now suppose that $(n_t : t \geq 0)$ is a discrete-time Markov chain taking values in $S = \{0, 1, 2, \dots\}$. We assume that local extinction events and colonisation events occur in separate seasonal phases and we assume that the population census takes place either at the end of each colonisation phase (the EC model) or at the end of each extinction phase (the CE model). The 1-step transition matrix $P = (p_{ij} : i, j \in S)$ is given by $P = EC$ or $P = CE$, depending on the timing of the census, where the transition matrices $E = (e_{ij} : i, j \in S)$ and $C = (c_{ij} : i, j \in S)$ that govern the local extinction and colonisation processes, respectively, are defined as follows.

Extinction process. We continue to assume that the extinction process is governed by the $\mathrm{Bin}(i, 1-e)$ law, where i is the number of patches occupied at the beginning of the local extinction phase and $e \in (0, 1)$. The transition probabilities are given by (3.3) although E is now an infinite matrix.

Colonisation process. Given that there are i patches occupied at the beginning of the colonisation phase, we assume that the number of empty patches colonised during that phase is governed by the $\mathrm{Poi}(mi)$ law where $m > 0$. Therefore,

$$c_{ij} = \begin{cases} \exp(mi)\ (mi)^{j-i}/(j-i)! & \text{if } j = i, i+1, i+2, \dots \\ 0 & \text{if } j < i, \end{cases} \tag{5.1}$$

where we note that $c_{0j} = \delta_{0j}$. Hence m may be interpreted as the expected number of empty patches colonised by *any one* occupied patch.

Transition probabilities. For the EC model, the 1-step transition probabilities are given by $p_{0j} = \delta_{0j}$ and $p_{i0} = e^i$ and

$$p_{ij} = \sum_{k=1}^{\min\{i,j\}} \binom{i}{k} (1-e)^k e^{i-k} \exp(-mk) \frac{(mk)^{j-k}}{(j-k)!} \qquad (i, j \geq 1).$$

For the CE model, $p_{0j} = \delta_{0j}$ and

$$p_{ij} = \sum_{k=\max\{i,j\}}^{\infty} \exp(-mi) \frac{(mi)^{k-i}}{(k-i)!} \binom{k}{j} (1-e)^j e^{k-j} \qquad (i \geq 1,\ j \geq 0).$$

In both cases, state 0 is absorbing and so $S = \{0\} \cup \{1, 2, \dots\}$ where $\{1, 2, \dots\}$ is an irreducible aperiodic class from which 0 is accessible. We think of these infinite-patch models as *island models* because there is no 'mainland' to rescue the patch network from total extinction.

Remark. Note that the class $\{1, 2, \dots\}$ is countably infinite and, hence, the chain may or may not reach the absorption state (Section 2.2.1). It will soon become apparent that branching theory can be used to determine conditions under which absorption (i.e. total extinction) takes place and conditions under which a proper limiting conditional distribution exists.

Seasonal dynamic. The infinite-patch island models admit the following seasonal dynamic:

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \text{Poi}(m\tilde{n}_t) \qquad \tilde{n}_t = n_t - \text{Bin}(n_t, e) \tag{5.2a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \text{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \text{Poi}(mn_t). \tag{5.2b}$$

Remark. The infinite-patch models derived above are natural analogues of the finite-patch (N-patch) chain binomial island models examined in Chapter 4 where, in particular, the colonisation process is defined by the $\text{Bin}(N-i, \hat{c}(i/N))$ law with $\hat{c}(0) = 0$. To see this, first consider the binomial random variable $X_n \sim \text{Bin}(n, p)$. If $p \to 0$ as $n \to \infty$ in such a way that $\mathbb{E}X_n = np$ approaches a constant $\lambda (> 0)$, then we have the well-known Poisson limit of the binomial distribution: as $n \to \infty$, $X_n \xrightarrow{\text{D}} X$ where $X \sim \text{Poi}(\lambda)$. Now, if $\hat{c}(0) = 0$ and $\hat{c}$ has a continuous second derivative near 0, then, for fixed i, $\text{Bin}(N-i, \hat{c}(i/N)) \xrightarrow{\text{D}} \text{Poi}(mi)$ as $N \to \infty$, where $m = \hat{c}'(0)$. Note that one need only assume the function $\hat{c}$ has a continuous second derivative near 0 and satisfy $\hat{c}(0) = 0$ in order that the desired asymptotic ($N \to \infty$) behaviour occurs; it is not necessary to assume that $\hat{c}$ also be continuous, increasing and concave (as done in Chapter 4). Our infinite-patch island models are therefore natural analogues of a wider class of N-patch island models than that previously investigated.

We remind the reader that the parameter m is interpreted as the expected number of empty patches colonised by any one occupied patch, and so we may take m to be any positive number ($m > 0$) in the following analysis. Of course, if one specifically wishes to study the infinite-patch limit of, say, the N-patch model with $\hat{c}(x) = cx$ ($0 < c \leq 1$), then $m = \hat{c}'(0) = c$ which is therefore restricted to $(0, 1]$.

5.2.1 Model analysis

From equations (5.2a) and (5.2b), notice that $n_{t+1} = \tilde{n}_t + \tilde{P}$ for the EC model, where $\tilde{n}_t \sim \text{Bin}(i, 1-e)$ and $\tilde{P} \sim \text{Poi}(m\tilde{n}_t)$, while $n_{t+1} \sim \text{Bin}(i + \tilde{P}, 1-e)$ for the CE model, where

$\tilde{P} \sim \mathrm{Poi}(mi)$. Therefore, the pgf $G(z) = \mathbb{E}(z^{n_{t+1}}|n_t = i)$ is given by

$$\begin{aligned} G(z) &= \mathbb{E}\big(z^{\tilde{n}_t}\mathbb{E}(z^{\tilde{P}}|\tilde{n}_t, n_t = i)|n_t = i\big) = \mathbb{E}\big(z^{\tilde{n}_t}\exp(-m\tilde{n}_t(1-z))|n_t = i\big) \\ &= \big(e + (1-e)z\exp(-m(1-z))\big)^i \end{aligned}$$

in the EC case and

$$\begin{aligned} G(z) &= \mathbb{E}\big(\mathbb{E}(z^{n_{t+1}}|\tilde{P}, n_t = i)|n_t = i\big) = \mathbb{E}\big((e + (1-e)z)^{i+\tilde{P}}|n_t = i\big) \\ &= (e + (1-e)z)^i \exp(-mi(1-e)(1-z)) \end{aligned}$$

in the CE case. We find that, for each $i \geq 1$, $G(z) = (g(z))^i$ where the pgf $g(z)$ is given by

$$\text{EC:} \qquad g(z) = e + (1-e)z\exp(-m(1-z)) \tag{5.3a}$$

$$\text{CE:} \qquad g(z) = (e + (1-e)z)\exp(-m(1-e)(1-z)). \tag{5.3b}$$

We may thus write

$$n_{t+1} \overset{\mathrm{D}}{=} \xi_1^{(t)} + \xi_2^{(t)} + \cdots + \xi_{n_t}^{(t)} \qquad (t \geq 0), \tag{5.4}$$

for both models, where $\xi_1^{(t)}, \xi_2^{(t)}, \ldots$ are iid random variables with common pgf $g(z)$. Furthermore, $g(z)$ is a mixture of Bernoulli and Poisson pgfs and so, for a random variable ξ with pgf $g(z)$, we have that $\xi = B(1 + P_1)$ for the EC model and $\xi = B + P_2$ for the CE model where $B \sim \mathrm{Ber}(1-e)$, $P_1 \sim \mathrm{Poi}(m)$ and $P_2 \sim \mathrm{Poi}(m(1-e))$ are independent Bernoulli and Poisson random variables, respectively. Therefore, our infinite-patch network behaves *as if* each occupied patch becomes extinct with probability e, or each occupied patch survives with probability $1-e$ *and* colonises $\mathrm{Poi}(m)$ patches (the EC case), or, instead, each occupied patch behaves as if it goes extinct with probability e *independently* of colonising $\mathrm{Poi}(m(1-e))$ patches (the CE case).

Indeed, equation (5.4) shows that both infinite-patch island models are equivalent to branching processes (Galton-Watson processes) with g being the pgf of the offspring distribution. The "particles" are the occupied patches and so we may think of each occupied patch at time t as being *replaced by* a random number of occupied patches (their "offspring") at time $t + 1$. The census times mark the generations of the branching process and, since the analysis above tells us that the number of "offspring" ξ produced by each occupied patch can be written in terms of Bernoulli and Poisson random variables, we can write down the offspring distribution $(p_j : j \geq 0)$, where $p_j = \Pr(\xi = j)$. We have that $p_0 = e$ and

$$p_j = (1-e)\exp(-m)\frac{m^{j-1}}{(j-1)!} \qquad (j \geq 1),$$

for the EC case, showing that the particle dies with probability e or survives and gives birth to a number of offspring that follows a Poisson distribution with mean m. For the CE model, $p_0 = e\exp(-m(1-e))$ and

$$p_j = (1-e)\exp(-m(1-e))\frac{(m(1-e))^{j-1}}{(j-1)!} + e\exp(-m(1-e))\frac{(m(1-e))^j}{j!} \qquad (j \geq 1),$$

and so the particle gives birth to a number of offspring that follows a Poisson distribution with mean $m(1-e)$ and dies with probability e.

We now invoke the theory of branching processes (Section 2.3.1) to answer questions about the long-term behaviour of the infinite-patch island models. Let μ be the mean of the offspring distribution. We have that $\mu = (1-e)(1+m)$, the same for both models, and so $\mathbb{E}(n_t|n_0) = n_0\mu^t$ $(t \geq 1)$. Both branching processes are therefore *subcritical*, *critical* or *supercritical* according to whether μ is *less than*, *equal to* or *greater than* 1 or, rather, according to whether the parameter m is *less than*, *equal to* or *greater than* the critical value $\rho = e/(1-e)$. We state the following result and then remark on an obvious connection with our N-patch chain binomial island models.

Theorem 5.1 *For both infinite-patch island models, total extinction occurs with probability 1 if and only if* $m \leq \rho$*; otherwise total extinction occurs with probability* η^{n_0} *where* η *is the unique solution of* $z = g(z)$ *on the interval* $(0,1)$*, with* g *as given by* (5.3)*, and* n_0 *is the initial number of occupied patches.*

The extinction probability η cannot be exhibited explicitly, but can of course be obtained numerically by iterating the map g.

Remark. The extinction criteria for the infinite-patch island models accord immaculately with the criteria for evanescence ($\hat{c}'(0) \leq e/(1-e)$) and quasi stationarity ($\hat{c}'(0) > e/(1-e)$) derived in the previous chapter for N-patch CE models with $\hat{c}(0) = 0$ (Section 4.4.2). We have that

(i) if $m = \hat{c}'(0) \leq e/(1-e)$, then extinction is certain for both infinite-patch island models while extinction is expected to occur quickly for the N-patch CE model with N large, and

(ii) if $m = \hat{c}'(0) > e/(1-e)$, then extinction is no longer certain for both infinite-patch island models while quasi equilibrium may ensue for the N-patch CE model with N large.

In Section 4.6, we mentioned that simulations of the N-patch EC model respond to the same stability criteria as that outlined above for the N-patch CE model. Since both infinite-patch models respond to the same criteria, then this is further evidence to support our comment that both finite-patch models do so as well.

We can also invoke the theory of branching processes to evaluate the variance of n_t conditional on n_0. Letting σ^2 be the variance of the offspring distribution, then $\mathrm{Var}(n_t|n_0)$ is that given by equation (2.4) with $\sigma^2 = (1-e)(c+e(1+c^2))$ in the EC case, $\sigma^2 = (1-e)(c+e)$ in the CE case and $\mu = (1-e)(1+m)$ in both cases. Also, a proper limiting conditional distribution exists in the case where $m < e/(1-e)$ (Yaglom's Theorem, Theorem 2.9) and, conditional on absorption occurring with probability 1, in the case where $m > e/(1-e)$ (Theorem 2.10).

5.3 An infinite-patch mainland-island model

A simple extension of the infinite-patch island model is obtained by making the following modification to the colonisation process. Given i patches are occupied at the beginning of the colonisation phase, suppose that the number of empty patches colonised during that phase follows the $\mathrm{Poi}(mi + m_0)$ law where $m \geq 0$ and $m_0 > 0$. The parameter m may once again be interpreted as the expected number of newly colonised patches by any one occupied patch, while the parameter m_0 may be interpreted as the expected number of empty patches colonised by an external source (i.e. a mainland). The local extinction process remains unchanged and so the natural infinite-patch mainland-island analogues are given by

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \mathrm{Poi}(m\tilde{n}_t + m_0) \qquad \tilde{n}_t = n_t - \mathrm{Bin}(n_t, e) \tag{5.5a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \mathrm{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \mathrm{Poi}(mn_t + m_0). \tag{5.5b}$$

The colonisation transition probabilities are therefore given by

$$c_{ij} = \begin{cases} \exp(mi + m_0)\ (mi + m_0)^{j-i}/(j-i)! & \text{if } j = i, i+1, i+2, \dots \\ 0 & \text{if } j < i, \end{cases}$$

where $c_{00} < 1$ and, subsequently, $p_{00} < 1$ for both the EC and CE cases. Hence, state 0 is non-absorbing and the set S forms a single irreducible and aperiodic class, thus supporting our interpretation of an infinite-patch network connected to a mainland population.

Remarks. (i) Although the Markov chain $(n_t : t \geq 0)$ is now irreducible and aperiodic, the state space remains countably infinite and so the chain does not necessarily have a unique stationary distribution (Section 2.2.2). We shall, however, provide conditions under which this distribution exists.

(ii) The infinite-patch mainland-island models may be derived from N-patch mainland-island models by replacing the colonisation probability $\hat{c}(i/N)$ with $\hat{c}(i/N) + m_0/N$; the second component (attributed to the mainland) is apportioned equally among all N patches. Hence, if $\hat{c}(0) = 0$ and $\hat{c}$ has a continuous second derivative near 0, then, for fixed i, we have that $\text{Bin}(N - i, \hat{c}(i/N) + c(0)/N) \xrightarrow{\text{D}} \text{Poi}(mi + m_0)$ as $N \to \infty$, where $m = \hat{c}'(0)$. The parameters m and m_0 need not be restricted to the interval $[0, 1]$ for the infinite-patch models and so we take $m \geq 0$ and $m_0 > 0$ in the following analysis.

5.3.1 Model analysis

Recall that the function G could be written in the form $G(z) = (g(z))^i$ $(i \geq 1)$ for both infinite-patch island models. For both infinite-patch mainland-island models, we find that G exhibits the form

$$G(z) = (g(z))^i h(z) \tag{5.6}$$

for each $i \geq 0$, where g is that given in Theorem 5.1 and h is defined by

$$\text{EC:} \qquad h(z) = \exp(-m_0(1 - z)) \tag{5.7a}$$

$$\text{CE:} \qquad h(z) = \exp(-m_0(1 - e)(1 - z)). \tag{5.7b}$$

We may thus write

$$n_{t+1} \overset{\mathrm{D}}{=} \xi_1^{(t)} + \xi_2^{(t)} + \cdots + \xi_{n_t}^{(t)} + \zeta_{t+1} \qquad (t \geq 1, n_0 = \zeta_0), \tag{5.8}$$

in both cases, where $\xi_1^{(t)}, \xi_2^{(t)}, \ldots$ are iid with pgf $g(z)$ and $\zeta_0, \zeta_1, \ldots$ are iid with pgf $h(z)$. Equation (5.7) tells us that $\zeta_0, \zeta_1, \ldots$ are iid $\mathrm{Poi}(d)$ random variables where $d = m_0$ in the EC case and $d = m_0(1-e)$ in the CE case. We conclude that the infinite-patch mainland-island models are equivalent to the Galton-Watson processes identified in Section 5.2.1 but modified so that there are $\mathrm{Poi}(d)$ immigrant particles in each generation (indeed, these processes are Galton-Watson-Immigration processes where $\mu = (1-e)(1+m)$ is the mean number of offspring produced by one particle and d is the mean number of immigrant particles). We may think of the mainland population as the source of "immigrant particles".

Again we can invoke the theory of branching processes. From Theorem 2.12 we see that the Markov chain $(n_t : t \geq 0)$ has a unique stationary distribution if and only if $m < e/(1-e)$.

We shall now examine a special case of the infinite-patch mainland-island model: the infinite-patch mainland model.

5.4 The infinite-patch mainland model

An important special case of the infinite-patch mainland-island model is obtained by setting m (the parameter attributed to the islands) equal to 0 in equation (5.5). Hence

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \mathrm{Poi}(m_0) \qquad \tilde{n}_t = n_t - \mathrm{Bin}(n_t, e) \tag{5.9a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \mathrm{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \mathrm{Poi}(m_0), \tag{5.9b}$$

where $m_0 > 0$, as before. State 0 is clearly non-absorbing and so the set S remains a single, irreducible and aperiodic class. The special case given by (5.9) is therefore the natural infinite-patch mainland analogue of our N-patch mainland model examined in Chapter 3.

The 1-step conditional distribution for the infinite-patch mainland model follows immediately from (5.8), however we will show that the mainland model is highly tractable and thus show that other quantities of interest can be evaluated explicitly.

5.4.1 Model analysis

By setting m equal to 0 in equation (5.3) and substituting the result into (5.6), we see that

$$G(z) = (e + (1-e)z)^i h(z) \qquad (t \geq 0),$$

for both infinite-patch mainland models, where $h(z)$ is given by (5.7). The distribution of n_{t+1} conditional on $n_t = i$ is therefore given by (5.8) but now $\xi_1^{(t)}, \xi_2^{(t)}, \ldots$ are independent $\text{Ber}(1-e)$ random variables and $\zeta_0, \zeta_1, \ldots$ are independent $\text{Poi}(d)$ random variables; $d = m_0$ in the EC case and $d = m_0(1-e)$ in the CE case. Hence, the infinite-patch mainland models are equivalent to Galton-Watson-Immigration processes where each particle (occupied patch) produces either 0 or 1 offspring according to the $\text{Ber}(1-e)$ distribution and each generation is joined by a random number of immigrant particles that follow a Poisson distribution with mean d. As before, the mainland may be interpreted as the source of "immigrant particles". Though instead of analysing the infinite-patch mainland models in terms of branching processes, we may re-express the 1-step conditional state distribution by

$$n_{t+1} \overset{\text{D}}{=} B + P, \tag{5.10}$$

where $B \sim \text{Bin}(i, 1-e)$ and $P \sim \text{Poi}(d)$. Due to the simple nature of (5.10), we can evaluate the distribution of n_t conditioned on the initial ($t = 0$) number of occupied patches using the same procedure as that outlined in Chapter 3 (see Theorem 3.1). The t-step conditional state distribution is as follows.

Theorem 5.2 *Define sequences* (r_t) *and* (s_t) *by* $r_t = (1-e)^t$ *and* $s_t = (d/e)(1-(1-e)^t)$ $(t \geq 0)$. *Then, given* $n_0 = i$, $n_t \overset{\text{D}}{=} B_t + P_t$ *for both infinite-patch mainland models where* $B_t \sim \text{Bin}(i, r_t)$ *and* $P_t \sim \text{Poi}(s_t)$ *are independent binomial and Poisson random variables, respectively.*

Proof. Suppose that $n_0 = i$ and recall that the function $G_t(z) = \mathbb{E}(z^{n_t} | n_0 = i)$ defines the pgf of n_t conditional on the initial observation. Then, $G_0(z) = z^i$ and, from (5.10), we have that

$$\begin{aligned} G_{t+1}(z) &= \mathbb{E}\big((e + (1-e)z)^{n_t} \exp(-d(1-z)) | n_0 = i\big) \\ &= \exp(-d(1-z))\, \mathbb{E}\big((e + (1-e)z)^{n_t} | n_0 = i\big) \\ &= \exp(-d(1-z))\, G_t(e + (1-e)z) \qquad (5.11) \end{aligned}$$

for all $t \geq 0$. The proof will be complete if we can show that

$$G_t(z) = (1 - r_t + r_t z)^i \exp(-s_t(1 - z)) \tag{5.12}$$

for all $t \geq 0$, where (r_t) and (s_t) are the given sequences. We will use Mathematical Induction. First, it is easy to see that (5.12) is true for $t = 0$ and $t = 1$ because $r_0 = 1$ and $s_0 = 0$ (since $G_0(z) = z^i$) while $r_1 = 1 - e$ and $s_1 = d$ (since $G_1 \equiv G$). Now, suppose that (5.12) is true for some fixed $t \geq 0$. Then, substitution into (5.11) gives

$$\begin{aligned} G_{t+1}(z) &= \exp(-d(1-z)) \big(1 - r_t + r_t(e + (1-e)z)\big)^i \exp\big(-s_t(1 - (e + (1-e)z))\big) \\ &= \big(1 - (1-e)r_t + (1-e)r_t z\big)^i \exp\big(-(d + (1-e)s_t)(1-z)\big). \end{aligned}$$

By the inductive hypothesis, we have $r_{t+1} = (1-e)r_t$ and $s_{t+1} = d + (1-e)s_t$. The difference equations solve to give $r_t = (1-e)^t$ and $s_t = (d/e)(1 - (1-e)^t)$ $(t \geq 0)$, which are the given sequences. Hence, equation (5.12) is proved by induction and the result follows. □

Therefore, for each $t \geq 1$, the metapopulation network behaves *as if* the i initially occupied patches remain occupied with probability r_t whilst the mainland colonises a $\mathrm{Poi}(s_t)$ number of empty patches, all patches being affected independently.

It is now a simple matter to establish the limiting $(t \to \infty)$ distribution of n_t. Since $\lim_{t\to\infty} r_t = 0$ and $\lim_{t\to\infty} s_t = d/e$, we have that, as $t \to \infty$, $\mathrm{Bin}(i, r_t) \xrightarrow{\mathrm{P}} 0$ and $\mathrm{Poi}(s_t) \xrightarrow{\mathrm{D}} \mathrm{Poi}(d/e)$. The following result is established.

Corollary 5.1 *As* $t \to \infty$, $n_t \xrightarrow{\mathrm{D}} \mathrm{Poi}(d/e)$.

Hence, in equilibrium, it is as if the mainland colonises a random number of patches that follows a Poisson distribution with mean d/e. The limiting mean number of occupied patches is larger for the EC case (since $d = m_0$ for the EC model and $d = m_0(1-e)$ for the CE model), though this is not surprising since, even in equilibrium, the EC model still corresponds to observing the metapopulation after successive colonisation phases. Branching theory establishes that the Markov chain always has a proper limiting distribution (see Theorem 2.12, noting that $\mu = 1 - e$ is always strictly less than 1), however Corollary 5.1 establishes a stronger result since it states this limiting distribution explicitly.

5.5 Discussion

This chapter was concerned with a number of metapopulation models that describe networks with infinitely-many patches ($N = \infty$). The models were designed to exhibit a seasonal dynamic and do so in such a way that they are viewed as natural infinite-patch analogues of our finite-patch chain binomial models examined in Chapters 3 and 4. The local extinction process was modelled in the same way for each infinite-patch model, however we considered three different ways in which empty patches are colonised: by migration from occupied patches (Section 5.2), by migration from occupied patches as well as an external source (Section 5.3) and by migration from an external source only (section 5.4).

Apart from the seasonal dynamic, the characteristic common to all three infinite-patch models is *branching*. We found that the 1-step conditional state distribution could be written as a sum of independent random variables and, due to the particular form of this sum, we deduced that occupied patches could be re-interpreted as "particles" of a branching process. We then invoked branching theory to prove results for the infinite-patch models, though, in one particular case (i.e. the mainland model), we obtained the t-step conditional state distribution and the limiting distribution explicitly without using branching theory. Indeed, our infinite-patch branching processes have an obvious epidemiological interpretation: given an assumed seasonal dynamic consisting of infectious periods and recovery (or death) periods, the "particles" may be thought of as infected individuals in a population with no ceiling.

One last remark should be made regarding the connection between our chain binomial models and the infinite-patch models above. We have already stated that the infinite-patch models may be derived as finite-patch limits of our chain binomial models, however observe that results pertaining to infinite-patch models turn out to be limits of results for finite-patch models as well. For example, consider the N-patch mainland model (equation (3.6)) but with c_0 replaced by m_0/N. Given $n_t = i$ for this N-patch model, Lemma 3.1 tells us that

$$\begin{aligned} n_{t+1} &\overset{\text{D}}{=} \text{Bin}(i, p) + \text{Bin}(N - i, q) \\ &= \text{Bin}(i, 1 - e(1 - m_0/N)) + \text{Bin}(N - i, m_0/N) \\ &\overset{\text{D}}{\to} \text{Bin}(i, r_1) + \text{Poi}(s_1) \end{aligned}$$

as $N \to \infty$. Given $n_0 = i$ now, Theorem 3.1 tells us that

$$n_t \overset{\text{D}}{=} \text{Bin}(i, p_t) + \text{Bin}(N - i, q_t) \overset{\text{D}}{\to} \text{Bin}(i, r_t) + \text{Poi}(s_t)$$

as $N \to \infty$. Hence, the conditional state distributions evaluated for the infinite-patch mainland models turn out to be equal to the asymptotic ($N \to \infty$) limits of the conditional state distributions of the corresponding N-patch mainland models. Of course, the same can be said for N-patch CE models by comparing Proposition 4.1 with the infinite-patch CE models above. Recall that $\xi \sim B + P_2$ for the infinite-patch island and mainland-island models in the CE case, where $B \sim Ber(1 - e)$ and $P_2 \sim \text{Poi}(m(1 - e))$. This allows us to write $n_{t+1} \overset{\text{D}}{=} \text{Bin}(i, 1 - e) + \text{Poi}(m(1 - e)i)$ for the infinite-patch island case and $n_{t+1} \overset{\text{D}}{=} \text{Bin}(i, 1 - e) + \text{Poi}(m(1 - e)i + m_0(1 - e))$ for the infinite-patch mainland-island case. Comparing these distributions to that in Proposition 4.1, it is easy to see that the 1-step conditional state distributions for the infinite-patch (CE) models are the asymptotic limits of the finite-patch (CE) counterparts.

Remark. Recall that $\xi \sim B(1 + P_1)$ for the infinite-patch island and mainland-island models in the EC case. The fact that ξ cannot be written as a simple sum of individual, independent random variables is perhaps one clue that explains why the 1-step conditional state distribution continues to remain elusive for the general finite-patch EC model.

Chapter 6

Limit Theorems for Discrete-time Metapopulation Models

Here we prove limit theorems for a general class of inhomogeneous Markov chains that exhibit the particular property of density dependence. We demonstrate how these theorems can be used to evaluate the limiting behaviour of any of our discrete-time metapopulation models (both finite-patch and infinite-patch) defined with density-dependent phases.

6.1 Introduction

The investigations carried out in Chapters 3, 4 and 5 established a number of analytical results for our stochastic discrete-time metapopulation models. In particular, recall that two types of limit theorems were proved for the N-patch mainland model (Section 3.3.2): the first, a law of large numbers, identified an approximating discrete-time deterministic trajectory and the second, a central limit law, showed that the fluctuations about this trajectory are approximately normally distributed. Indeed, we proved such laws for the N-patch CE model (Section 4.4.1) where the colonisation process depended on the current number i of occupied patches through the proportion, or *density*, i/N of occupied patches. In this chapter we revisit these types of limit laws. We show that it is possible to evaluate the limiting behaviour of both (EC and CE) N-patch models and that similar laws may also be found for infinite-patch models.

We begin this chapter by introducing a general class of time-inhomogeneous Markov chains that exhibit the particular property of density dependence. For this class of chains we prove a law of large numbers, which identifies a (time-inhomogeneous) discrete-time deterministic trajectory, and a central limit law, which shows that the random fluctuations about this trajectory have an approximating autoregressive structure. These results will then be used to determine limit theorems for our finite-patch (Section 6.3) and infinite-patch (Section 6.4) metapopulation models. The parameter N serves as an arbitrary index in the general setting, though it will soon become clear that N represents the total number of patches in the finite-patch setting and marks a particular threshold in the infinite-patch setting. The results are illustrated with examples.

6.2 General structure: density dependence

Let $(n_t^N : t \geq 0)$ be a family of discrete-time Markov chains indexed by N, each taking values in a set S_N which is a subset of $\mathbb{Z}_+$. Suppose that the family is *density dependent* in that there are sequences of non-negative functions (f_t) and (v_t) such that

$$\mathbb{E}(n_{t+1}^N | n_t^N) = N f_t(n_t^N / N) \quad \text{and} \quad \operatorname{Var}(n_{t+1}^N | n_t^N) = N v_t(n_t^N / N). \tag{6.1}$$

The "density process" $(X_t^N : t \geq 0)$ obtained by setting $X_t^N = n_t^N / N$ will have

$$\mathbb{E}(X_{t+1}^N | X_t^N) = f_t(X_t^N) \quad \text{and} \quad N \operatorname{Var}(X_{t+1}^N | X_t^N) = v_t(X_t^N).$$

The following result is a law of large numbers that establishes convergence of the density process to a deterministic trajectory. Here, and henceforth, convergence is established in the limit as $N \to \infty$.

Theorem 6.1 *Suppose that, for all $t \geq 0$, $f_t(x)$ and $v_t(x)$ are continuous in x and such that $f_t(X_t^N)$ and $v_t(X_t^N)$ are a.s. uniformly bounded. Then, if $X_0^N \xrightarrow{\text{P}} x_0$ (a constant), $X_t^N \xrightarrow{\text{P}} x_t$ for all $t \geq 1$, where $x_\bullet$ is determined by $x_{t+1} = f_t(x_t)$ $(t \geq 0)$.*

Proof. We will use Mathematical Induction. By assumption, $X_0^N \xrightarrow{\text{P}} x_0$. Now, suppose that $X_t^N \xrightarrow{\text{P}} x_t$ for some $t \geq 0$. Then, since f_t is continuous, $\mathbb{E}(X_{t+1}^N | X_t^N) = f_t(X_t^N) \xrightarrow{\text{P}} f_t(x_t)$. We are also told that $f_t(X_t^N)$ is a.s. uniformly bounded and so, by Theorem 7.4(b) of Grimmett and Stirzaker [35], $f_t(X_t^N) \xrightarrow{r} f_t(x_t)$ for all $r \geq 1$, which entails, in particular, that

$$\mathbb{E} X_{t+1}^N = \mathbb{E}\big(\mathbb{E}(X_{t+1}^N | X_t^N)\big) = \mathbb{E} f_t(X_t^N) \to f_t(x_t) \quad \text{and} \quad \operatorname{Var} f_t(X_t^N) \to 0.$$

Since v_t is continuous and $v_t(X_t^N)$ is a.s. uniformly bounded, we also have that $\mathbb{E} v_t(X_t^N) \to v_t(x_t)$. Therefore,

$$\begin{aligned}\operatorname{Var} X_{t+1}^N &= \mathbb{E}\operatorname{Var}(X_{t+1}^N|X_t^N) + \operatorname{Var}\mathbb{E}(X_{t+1}^N|X_t^N)\\ &= \frac{1}{N}\mathbb{E}\, v_t(X_t^N) + \operatorname{Var} f_t(X_t^N) \to 0.\end{aligned}$$

But, by Chebyshev's inequality (see for example Section 7.3 of [35]), we have

$$\begin{aligned}\Pr(|X_{t+1}^N - f_t(x_t)| \ge \epsilon) &\le \frac{1}{\epsilon^2}\mathbb{E}(X_{t+1}^N - f_t(x_t))^2\\ &= \frac{1}{\epsilon^2}\left(\operatorname{Var} X_{t+1}^N + (\mathbb{E}X_{t+1}^N - f_t(x_t))^2\right) \to 0\end{aligned}$$

for all $\epsilon > 0$. That is, $X_{t+1}^N \xrightarrow{\text{P}} x_{t+1}$, and the proof is complete. □

Theorem 6.1 will be applied in the context of our N-patch models (Section 6.3 to follow), with X_t^N being interpreted as a *proportion* and is itself bounded. To accommodate cases where X_t^N is unbounded (as in Section 6.4 where we treat our infinite-patch models), we relax uniform boundedness in favour of a Lipschitz condition, but at the expense of requiring a more stringent initial condition, that X_0^N converges to x_0 in mean square.

Theorem 6.2 *Suppose that, for all $t \ge 0$, $f_t(x)$ and $v_t(x)$ are Lipschitz continuous in x. If $X_0^N \xrightarrow{2} x_0$ (a constant), then $X_t^N \xrightarrow{2} x_t$ (and hence $X_t^N \xrightarrow{\text{P}} x_t$) for all $t \ge 1$, where $x_\bullet$ is determined by $x_{t+1} = f_t(x_t)$ $(t \ge 0)$.*

Proof. We will again use Mathematical Induction. Suppose $X_t^N \xrightarrow{2} x_t$ for some $t \ge 0$ (noting that $X_0^N \xrightarrow{\text{P}} x_0$ by assumption). Since $f_t(x)$ is Lipschitz continuous, then

$$|f_t(X_t^N) - f_t(x_t)| \le \kappa_t |X_t^N - x_t|,$$

and hence

$$(f_t(X_t^N) - f_t(x_t))^2 \le \kappa_t^2 (X_t^N - x_t)^2, \tag{6.2}$$

for some positive constant κ_t. On taking the expectation of (6.2), we see that $f_t(X_t^N) \xrightarrow{2} f_t(x_t)$. This implies, in particular, that (i) $\operatorname{Var} f_t(X_t^N) \to 0$ and (ii) $\mathbb{E} f_t(X_t^N) \to f_t(x_t)$, which means that $\mathbb{E}X_{t+1}^N \to x_{t+1}$. Similarly, since $v_t(x)$ is Lipschitz continuous, $\mathbb{E}v_t(X_t^N) \to v_t(x_t)$. We know that $\operatorname{Var} X_{t+1}^N = \mathbb{E}\, v_t(X_t^N)/N + \operatorname{Var} f_t(X_t^N)$ (from the previous proof) and so we see that $\operatorname{Var} X_{t+1}^N \to 0$. But, $\mathbb{E}(X_{t+1}^N - x_{t+1})^2 = \operatorname{Var} X_{t+1}^N + (\mathbb{E} X_{t+1}^N - x_{t+1})^2$, and so $X_{t+1}^N \xrightarrow{2} x_{t+1}$. This completes the proof. □

Having established convergence in probability to a limiting deterministic trajectory $x_{\bullet}$, we now consider the "fluctuations process" $(Z_t^N : t \geq 0)$ obtained by setting $Z_t^N = \sqrt{N}(X_t^N - x_t)$. The following additional structure dictates the conditions under which we determine a central limit law for Z_t^N. Suppose that

$$n_{t+1}^N = g_t^N + \textstyle\sum_{j=1}^{r_t^N} \xi_{jt}^N \qquad (t \geq 0), \tag{6.3}$$

where

$$r_t^N = N r_t(n_t^N/N) \quad \text{and} \quad g_t^N = N g_t(n_t^N/N)$$

with $r_t(x)$ and $g_t(x)$ being continuous in x, and ξ_{jt}^N $(j = 1, \ldots, r_t^N)$ are iid random variables having a distribution that depends only on t and on n_t^N/N, and which have bounded third moment. In particular, we assume that there are functions $m_t(x)$ and $\sigma_t^2(x)$ such that

$$\mathbb{E}\, \xi_{jt}^N = m_t(n_t^N/N) \quad \text{and} \quad \operatorname{Var}(\xi_{jt}^N) = \sigma_t^2(n_t^N/N),$$

and a function $b_t(x)$ such that

$$\mathbb{E}(\xi_{jt}^N - m_t(x))^3 = b_t(n_t^N/N),$$

which is bounded in x. Of course, these functions must be such that r_t^N and g_t^N are positive integers and, then, that $n_{t+1}^N \in S_N$. By the representation of (6.3), we see that $(n_t^N : t \geq 0)$ satisfies the properties of density dependence with sequences (f_t) and (v_t) given by

$$f_t(x) = g_t(x) + r_t(x) m_t(x) \quad \text{and} \quad v_t(x) = r_t(x)\sigma_t^2(x). \tag{6.4}$$

Remark. The structure outlined above is very similar to that of Klebaner and Nerman [49] (see also Klebaner [48]) who studied a generalisation of the Galton-Watson process where the offspring distribution was allowed to depend on the current population size measured as a proportion of some threshold N. They had $g_t \equiv 0$ and $r_t(x) = x$ and so their model is time-homogeneous with $n_{t+1} = \sum_{j=1}^{Nx} \xi_{jt}^N$ defining the size of a Galton-Watson process at time $t+1$. We note that many of their results (Section 3 of [49]) carry over to the present setting with only minor changes, however we content ourselves with proving two general results for our time-inhomogeneous setup (Lemma 6.1 and Theorem 6.3 below).

Observe that we may rewrite equation (6.3) as

$$X_{t+1}^N = f_t(X_t^N) + \frac{1}{\sqrt{N}}\,\eta_t^N(X_t^N), \tag{6.5}$$

where

$$\eta_t^N(x) = \frac{1}{\sqrt{N}} \sum_{j=1}^{Nr_t(x)} (\xi_{jt}^N - m_t(x)). \tag{6.6}$$

Note that, for fixed x, $\eta_t^N(x)$ is independent of X_t^N. We present the following result in anticipation of proving a central limit law for Z_t^N.

Lemma 6.1 *If $X_t^N \xrightarrow{\text{P}} x$ (a constant), then $\eta_t^N(X_t^N) \xrightarrow{\text{D}} \text{N}(0, v_t(x))$.*

Proof. Let $Y_{Nj} = \xi_{jt}^N - m_t(X_t^N)$ and consider the triangular array $(Y_{Nj} : j = 1, \ldots, r_N)$ where $r_N = [Nr_t(X_t^N)]$, noting that the sequence $Y_{N1}, \ldots, Y_{N,r_N}$ is independent for each N. It is clear that, for any j, $\mathbb{E}(Y_{Nj}) = 0$ and $\mathbb{E}(Y_{Nj}^2) = \text{Var}(\xi_{jt}^N) = \sigma_t^2(X_t^N)$. Hence $s_N^2 := \sum_{j=1}^{r_N} \mathbb{E}(Y_{Nj}^2) = r_N \sigma_t^2(X_t^N)$ and $Y_N := \sum_{j=1}^{r_N} Y_{Nj} = \sqrt{N}\eta_t^N(X_t^N)$. Provided $X_t^N \xrightarrow{\text{P}} x$ for fixed t, the Feller-Lindeberg Theorem (see for example Theorems 27.2 and 27.3 of Billingsley [11]) tells us that $Y_N/s_N \xrightarrow{\text{D}} \text{N}(0,1)$. But, first, we require *Lyapounov's condition*,

$$\lim_{N\to\infty} \sum_{j=1}^{r_N} \frac{1}{s_N^{2+\delta}} \mathbb{E}\left[|Y_{Nj}|^{2+\delta}\right] = 0 \qquad \text{(for some } \delta > 0\text{)},$$

to hold, since it is a sufficient condition for the Feller-Lindeberg Theorem. We will show that Lyapounov's condition is satisfied for $\delta = 1$. Now, for $j = 1$, we have

$$\frac{\mathbb{E}\left[|Y_{N1}|^3\right]}{s_N^3} = \frac{\mathbb{E}\left[|\xi_{1t}^N - m_t(X_t^N)|^3\right]}{\sigma_t^3(X_t^N)\left[Nr_t(X_t^N)\right]^{3/2}} \to 0$$

because $\mathbb{E}(|\xi_{1t}^N - m_t(x)|^3) \le b_t(x)$ and $b_t(x)$ is bounded in x. Indeed, we draw the same conclusion for any value of j and Lyapounov's condition is satisfied as a result. Therefore, if

$X_t^N \xrightarrow{\mathrm{P}} x$, then, by the Feller-Lindeberg Theorem, $Y_N/s_N \xrightarrow{\mathrm{D}} \mathrm{N}(0,1)$ and we deduce that

$$\frac{\sqrt{N}\eta_t^N(X_t^N)}{\sigma_t(X_t^N)\sqrt{Nr_t(X_t^N)}} \xrightarrow{\mathrm{D}} \mathrm{N}(0,1),$$

from which we see that $\eta_t^N(X_t^N) \xrightarrow{\mathrm{D}} \mathrm{N}(0, \sigma_t^2(x) r_t(x))$. This completes the proof. $\square$

We now present a central limit law for the fluctuations process. This law will be applied in both the N-patch and infinite-patch contexts, though the index N will take on a different meaning in each case.

Theorem 6.3 *Suppose that, for all $t \geq 0$, $f_t(x)$ is twice continuously differentiable in x with bounded second derivative and that $X_t^N \xrightarrow{\mathrm{P}} x_t$, where $x_\bullet$ satisfies $x_{t+1} = f_t(x_t)$ $(t \geq 0)$. If $Z_0^N \xrightarrow{\mathrm{D}} z_0$ (a constant), then $Z_\bullet^N$ converges weakly to the Gaussian Markov process $Z_\bullet$ defined by*

$$Z_{t+1} = f_t'(x_t) Z_t + E_t, \tag{6.7}$$

where (E_t) are independent with $E_t \sim \mathrm{N}(0, v_t(x_t))$.

Proof. Recall that $X_{t+1}^N = f_t(X_t^N) + \big(1/\sqrt{N}\big)\eta_t^N(X_t^N)$ where $\eta_t^N(X_t^N)$ is given by (6.6). We are told that $X_t^N \xrightarrow{\mathrm{P}} x_t$ for all $t \geq 1$, where $x_\bullet$ is determined by $x_{t+1} = f_t(x_t)$ $(t \geq 0)$, and so

$$X_{t+1}^N - x_{t+1} = f_t(X_t^N) - f_t(x_t) + \big(1/\sqrt{N}\big)\eta_t^N(X_t^N)$$

where, upon multiplying by $\sqrt{N}$, we have $Z_{t+1}^N = \sqrt{N}\left(f_t(X_t^N) - f_t(x_t)\right) + \eta_t^N(X_t^N)$. We are also told that $f_t(x)$ is twice continuously differentiable in x (for all $t \geq 0$) so that, by Taylor's Theorem, $f_t(X_t^N) = f_t(x_t) + f_t'(x_t)(X_t^N - x_t) + \frac{1}{2} f_t''(\theta_t^N)(X_t^N - x_t)^2$ for some θ_t^N between X_t^N and x_t, and therefore $Z_{t+1}^N = f_t'(x_t) Z_t^N + \eta_t^N(X_t^N) + \frac{1}{2\sqrt{N}} f_t''(\theta_t^N)(Z_t^N)^2$. But, since $f_t''(x)$ is bounded in x, we may thus write

$$Z_{t+1}^N = f_t'(x_t) Z_t^N + \eta_t^N(X_t^N) + o_t^N(1), \tag{6.8}$$

where $o_t^N(1) \xrightarrow{\mathrm{P}} 0$ as $N \to \infty$. Now, to establish weak convergence of $Z_\bullet^N$ to $Z_\bullet$ it is sufficient to establish convergence of the finite-dimensional distributions. To this end, consider the characteristic function

$$\phi_t^N(\omega_t, \ldots, \omega_0) = \mathbb{E} \exp\left(\mathbf{i}\left(\omega_t Z_t^N + \cdots + \omega_0 Z_0^N\right)\right)$$

of the sequence $(Z^N_0, \dots, Z^N_t)$. Then, from (6.8), the cf of $(Z^N_0, \dots, Z^N_{t+1})$ is given by

$$\begin{aligned}\phi^N_{t+1}(\omega_{t+1}, \omega_t, \dots, \omega_0) = \mathbb{E}\exp\Big(&\mathbf{i}\,\omega_{t+1}\eta^N_t(X^N_t) + \mathbf{i}\,(\omega_t + \omega_{t+1}f_t'(x_t))Z^N_t \\ &+ \mathbf{i}\big(\omega_{t-1}Z^N_{t-1} + \dots + \omega_0 Z^N_0\big) + o^N_t(1)\Big).\end{aligned}$$

We have assumed that, for fixed t and x_t, $X^N_t \xrightarrow{\text{P}} x_t$ and so $\eta^N_t(X^N_t) \xrightarrow{\text{D}} \text{N}(0, v_t(x_t))$ by Lemma 6.1. Then, from the Markov Property and our premise $Z^N_0 \xrightarrow{\text{D}} z_0$, it follows that $\lim_{N\to\infty}\phi^N_t = \phi_t$ exists for all $t \geq 1$ and satisfies

$$\phi_{t+1}(\omega_{t+1}, \omega_t, \dots \omega_0) = \exp(-\tfrac{1}{2}\omega^2_{t+1}v_t(x_t))\,\phi_t\big(\omega_t + \omega_{t+1}f_t'(x_t), \omega_{t-1}, \dots \omega_0\big),$$

with $\phi_0(\omega_0) = \exp(\mathbf{i}\,\omega_0 z_0)$ being the cf of $Z_0 = z_0$. But, this iteration defines the cf of $(Z_0, \dots, Z_t)$, where $Z_\bullet$ is the proposed limiting Gaussian process. This completes the proof. □

Hence, our central limit law establishes that the fluctuations converge in the sense of finite-dimensional distributions. This is to say that, for a finite sequence of times $t_1, \dots, t_k$, the sequence $(Z^N_{t_1}, \dots, Z^N_{t_k})$ converges in distribution to the random sequence $(Z_{t_1}, \dots, Z_{t_k})$, where $Z_\bullet$ is defined by (6.7). Indeed, the limiting process is 'Gaussian' because $Z_{t+1} \sim \text{N}\big(f_t'(x_t)z, v_t(x_t)\big)$ given $Z_t = z_t$ for all $t \geq 0$, and the representation of (6.7) clearly shows that this process is Markovian (Theorem 2.14). The mean and covariance function of $Z_\bullet$ are easy to evaluate by iterating (6.7):

$$\mu_t := \mathbb{E}Z_t = z_0\,\Pi_{0,\,t} \qquad (t \geq 1) \tag{6.9}$$

and

$$c_{t,\,s} := \text{Cov}\,(Z_t, Z_s) = V_t\,\Pi_{t,\,s} \qquad (s \geq t \geq 1), \tag{6.10}$$

where

$$\Pi_{u,\,v} := \prod_{w=u}^{v-1} f_w'(x_w) \qquad (v > u) \tag{6.11}$$

and

$$V_t := \text{Var}\,(Z_t) = \sum_{s=0}^{t-1} v_s(x_s)\Pi^2_{s+1,\,t} \qquad (t \geq 1). \tag{6.12}$$

Notice that the sequence of means satisfy $\mu_{t+1} = f_t'(x_t)\mu_t$ and the sequence of variances satisfy $V_{t+1} = (f_t'(x_t))^2 V_{t-1} + v_t(x_t)$, from which it is clear that the Gaussian Markov process is time-inhomogeneous. Also notice that, for any $t \geq 1$, $Z_t^N \xrightarrow{\text{D}} Z_t$, where $Z_t \sim \text{N}(\mu_t, V_t)$, and so we can use these formulae to approximate the mean and covariance function of $n_\bullet^N$. Indeed, for any finite set of times $t_1, \ldots, t_k$, the joint distribution of $n_{t_1}^N, \ldots, n_{t_k}^N$ can be approximated by a k-dimensional Gaussian distribution with $\mathbb{E} n_{t_i}^N \simeq N x_{t_i} + \sqrt{N} \mu_{t_i}$ and $\text{Cov}\,(n_{t_i}^N, n_{t_j}^N) \simeq N c_{t_i, t_j}$.

Remark. Each of our discrete-time metapopulation models were studied in terms of two time-homogeneous Markov chains, corresponding to the EC and CE observation schemes. It will soon become clear that some of our time-homogeneous Markov chains do not exhibit the required property of density dependence, however the level of generality incorporated in the limit theorems proved above allow us to establish the required limiting ($N \to \infty$) behaviour for time-inhomogeneous Markov chains with density-dependent seasonal phases.

We now demonstrate how the time-inhomogeneous limit theorems apply to our time-homogeneous metapopulation models.

6.3 N-patch metapopulation models

Here we explain how the results of Section 6.2 can be applied to our N-patch metapopulation models. First, we define $(n_t : t \geq 0)$ as a discrete-time Markov chain that takes values in the set $S_N = \{0, 1, 2, \ldots, N\}$ and evolves as follows:

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \text{Bin}(N - \tilde{n}_t, \hat{c}(\tilde{n}_t/N)) \qquad \tilde{n}_t = n_t - \text{Bin}(n_t, e) \tag{6.13a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \text{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \text{Bin}(N - n_t, \hat{c}(n_t/N)), \tag{6.13b}$$

where n_t is the number of occupied patches at time t in a group of N patches, e is the local extinction probability ($0 < e < 1$) and $\hat{c}(x)$ is the colonisation probability expressed as a function of the proportion x of occupied patches. The function $\hat{c} : [0, 1] \to [0, 1]$ is assumed to be continuous, increasing and concave, with $\hat{c}(0) \geq 0$ and $\hat{c}(x) \leq 1$. The set S_N is irreducible unless $\hat{c}(0) = 0$, in which case there is a single absorbing state (state 0, corresponding to total extinction of the population) with the remaining states forming an irreducible transient class $E_N = \{1, 2, \ldots, N\}$ from which 0 is accessible.

6.3.1 Limit theorems

Given the structure outlined in (6.13), let us now consider the time-inhomogeneous Markov chain $(n_t^N : t \geq 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$. It is important to note that the chain $(n_t^N : t \geq 0)$ corresponds to taking a metapopulation census after *each* seasonal phase, whilst the chain $(n_t : t \geq 0)$ corresponds to taking a census after each second phase. With $X_t^N = n_t^N/N$, the sequences (f_t) and (v_t) satisfy

$$f_{2t}(x) = f_0(x) = (1-e)x \qquad f_{2t+1}(x) = f_1(x) = x + (1-x)\hat{c}(x) \tag{6.14}$$

$$v_{2t}(x) = v_0(x) = e(1-e)x \qquad v_{2t+1}(x) = v_1(x) = (1-x)\hat{c}(x)(1-\hat{c}(x)) \tag{6.15}$$

for the EC model and

$$f_{2t}(x) = f_0(x) = x + (1-x)\hat{c}(x) \qquad f_{2t+1}(x) = f_1(x) = (1-e)x \tag{6.16}$$

$$v_{2t}(x) = v_0(x) = (1-x)\hat{c}(x)(1-\hat{c}(x)) \qquad v_{2t+1}(x) = v_1(x) = e(1-e)x \tag{6.17}$$

for the CE model. Notice that the roles of f_0 and f_1, as well as v_0 and v_1, are reversed. All of these functions are continuous, and $f_t(X_t^N)$ and $v_t(X_t^N)$ are uniformly bounded because $0 \leq X_t^N \leq 1$ and $\hat{c}(x) \leq 1$. Hence, by Theorem 6.1, the inhomogeneous density process $(X_t^N : t \geq 0)$ is approximated by a deterministic process $(x_t : t \geq 0)$ that satisfies, in particular,

$$x_{2(t+1)} = f(x_{2t}) \quad \text{where} \quad f = f_1 \circ f_0.$$

Thus we have the following law of large numbers (LLN) for the time-homogeneous density process.

Theorem 6.4 (LLN for N-patch models) *For the N-patch metapopulation models with parameters e and $\hat{c}(x)$, let $X_t^N = n_t/N$ be the proportion of occupied patches at census t. If $X_0^N \xrightarrow{\mathrm{P}} x_0$, then $X_t^N \xrightarrow{\mathrm{P}} x_t$ for all $t \geq 1$, where $x_{\bullet}$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with*

$$f(x) = (1-e)x + (1-(1-e)x)\hat{c}((1-e)x), \tag{EC model}$$

$$f(x) = (1-e)(x + (1-x)\hat{c}(x)). \tag{CE model}$$

Thus, Theorem 6.4 provides a new result for the EC model and reiterates the result for the CE model as given by Theorem 4.1.

Remark. In fact, Theorem 6.1 can be applied directly to the CE model (6.13b) because $\mathbb{E}(n_{t+1}|n_t)$ and $\operatorname{Var}(n_{t+1}|n_t)$ can be evaluated explicitly. That is,

$$\begin{aligned}\mathbb{E}(n_{t+1}|n_t) &= \mathbb{E}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t,\, n_t)|n_t\big) = \mathbb{E}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t)|n_t\big) = (1-e)\mathbb{E}\big(\tilde{n}_t|n_t\big)\\ &= (1-e)(n_t + (N-n_t)\hat{c}(n_t/N)),\end{aligned}$$

implying that $f_t(x) = f(x) = (1-e)(x + (1-x)\hat{c}(x))$, and

$$\begin{aligned}\operatorname{Var}(n_{t+1}|n_t) &= \mathbb{E}\big(\operatorname{Var}(n_{t+1}|\tilde{n}_t,\, n_t)|n_t\big) + \operatorname{Var}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t,\, n_t)|n_t\big)\\ &= \mathbb{E}\big(\operatorname{Var}(n_{t+1}|\tilde{n}_t)|n_t\big) + \operatorname{Var}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t)|n_t\big)\\ &= e(1-e)\mathbb{E}(\tilde{n}_t|n_t) + (1-e)^2 \operatorname{Var}(\tilde{n}_t|n_t)\\ &= e\mathbb{E}(n_{t+1}|n_t) + (1-e)^2(N-n_t)\hat{c}(n_t/N)(1-\hat{c}(n_t/N)),\end{aligned}$$

implying that

$$\begin{aligned}v_t(x) = v(x) &= ef(x) + (1-e)^2(1-x)\hat{c}(x)(1-\hat{c}(x))\\ &= (1-e)\left(ex + (1-x)\hat{c}(x)(1-(1-e)\hat{c}(x))\right).\end{aligned}$$

Both f and v are continuous, and $f(X_t^N)$ and $v(X_t^N)$ are bounded because $0 \le X_t^N\,(= n_t/N) \le 1$ and $\hat{c}(x) \le 1$. For the N-patch EC model (6.13a), similar calculations reveal that we cannot evaluate the conditional mean and variance explicitly. Unless we impose further restrictions on $\hat{c}$, we cannot show that the time-homogeneous EC model exhibits the required density dependence properties and so we are unable to apply Theorem 6.1 directly in this case.

To obtain the corresponding central limit law for the time-homogeneous fluctuations process $(Z_t^N : t \ge 0)$, where $Z_t^N = \sqrt{N}(X_t^N - x_t)$ with $X_t^N = n_t/N$, first observe that our N-patch models can be represented as

$$n_{t+1} = \tilde{n}_t + \textstyle\sum_{j=1}^{N-\tilde{n}_t}\mathrm{Ber}_j(\hat{c}(\tilde{n}_t/N)) \qquad \tilde{n}_t = n_t - \textstyle\sum_{j=1}^{n_t}\mathrm{Ber}_j(e) \qquad \text{(EC model)}$$

$$n_{t+1} = \tilde{n}_t - \textstyle\sum_{j=1}^{\tilde{n}_t}\mathrm{Ber}_j(e) \qquad \tilde{n}_t = n_t + \textstyle\sum_{j=1}^{N-n_t}\mathrm{Ber}_j(\hat{c}(n_t/N)), \qquad \text{(CE model)}$$

where the ($\mathrm{Ber}_j(p)$) are collections of iid Bernoulli random variables with success probability p. Thus, we may apply Theorem 6.3 to the time-inhomogeneous Markov chain $(n_t^N : t \geq 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$ (as above) because this chain will have the form exhibited by equation (6.3) with the $(\pm\, \xi_{jt}^N)$ being appropriate sequences of iid Bernoulli random variables. For both inhomogeneous models, we have $g_{2t}(x) = g_{2t+1}(x) = x$. For the EC case,

$$\begin{aligned} r_{2t}(x) &= x \qquad r_{2t+1}(x) = 1 - x \\ m_{2t}(x) &= -e \qquad m_{2t+1}(x) = \hat{c}(x) \\ \sigma_{2t}^2(x) &= e(1-e) \qquad \sigma_{2t+1}^2(x) = \hat{c}(x)(1-\hat{c}(x)), \end{aligned}$$

which lead to the same sequences (f_t) and (v_t) as that given by (6.14) and (6.15), respectively. We also have $b_{2t}(x) = e(1-e)(1-2e)$ and $b_{2t+1}(x) = \hat{c}(x)(1-\hat{c}(x))(1-2\hat{c}(x))$ in this case. For the CE case,

$$\begin{aligned} r_{2t}(x) &= 1 - x \qquad r_{2t+1}(x) = x \\ m_{2t}(x) &= \hat{c}(x) \qquad m_{2t+1}(x) = -e \\ \sigma_{2t}^2(x) &= \hat{c}(x)(1-\hat{c}(x)) \qquad \sigma_{2t+1}^2(x) = e(1-e), \end{aligned}$$

which then lead to the same sequences (f_t) and (v_t) as that given by (6.16) and (6.17). Also, $b_{2t}(x) = \hat{c}(x)(1-\hat{c}(x))(1-2\hat{c}(x))$ and $b_{2t+1}(x) = e(1-e)(1-2e)$. If we assume that $\hat{c}$ is twice continuously differentiable (in addition to the conditions we have already imposed on $\hat{c}$), then $f_t(x)$ will be twice continuously differentiable in x with bounded second derivative, $v_t(x)$ will be continuous in x, and $b_t(x)$ will be bounded in x. We already know that the deterministic trajectory for both inhomogeneous Markov chains satisfy $x_{2(t+1)} = f(x_{2t})$ where $f = f_1 \circ f_0$, and so, by Theorem 6.3, the limiting Gaussian Markov process $Z_\bullet$ must satisfy

$$Z_{2(t+1)} = f'(x_{2t}) Z_{2t} + \hat{E}_{2t}, \quad \text{with} \quad \hat{E}_{2t} \sim \mathrm{N}(0, v(x_{2t})),$$

where $v = v_1 \circ f_0 + (f_1' \circ f_0)^2 v_0$, with $f_1'(x) = 1 - \hat{c}(x) + (1-x)\hat{c}'(x)$ in the EC case and $f_1'(x) = (1-e)$ in the CE case. Thus we arrive at the following central limit law (CLL) for our time-homogeneous metapopulation models.

Theorem 6.5 (CLL for N-patch models) *For the N-patch metapopulation models with parameters e and $\hat{c}(x)$, suppose that $\hat{c}$ is twice continuously differentiable. Let $Z_t^N = \sqrt{N}(X_t^N - x_t)$, where $X_t^N = n_t/N$ is the proportion of occupied patches at census t and where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with f given as in Theorem 6.4. Then, if $Z_0^N \xrightarrow{\mathrm{D}} z_0$, $Z_\bullet^N$ converges weakly to the Gaussian Markov process $Z_\bullet$ defined by*

$$Z_{t+1} = f'(x_t)Z_t + E_t,$$

with (E_t) independent and $E_t \sim \mathrm{N}(0, v(x_t))$, where

$$\begin{aligned} v(x) &= (1-(1-e)x)\hat{c}((1-e)x)(1-\hat{c}((1-e)x)) \\ &\quad + e(1-e)x\big[1-\hat{c}((1-e)x) + (1-(1-e)x)\hat{c}'((1-e)x)\big]^2, && \text{(EC model)} \\ v(x) &= (1-e)\big[ex + (1-x)\hat{c}(x)\big(1-(1-e)\hat{c}(x)\big)\big]. && \text{(CE model)} \end{aligned}$$

An immediate consequence of Theorem 6.5 is that the time-homogeneous fluctuations process satisfies $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(\mu_t, V_t)$ for any $t \geq 1$, where

$$\mu_t = z_0 \prod_{s=0}^{t-1} f'(x_s) \quad \text{and} \quad V_t = \sum_{s=0}^{t-1} v(x_s) \prod_{u=s+1}^{t-1} f'(x_u)^2. \tag{6.19}$$

The central limit law that was presented earlier for the N-patch CE model (Theorem 4.2), gives the same expression for the mean and variance of $Z_\bullet$, however notice that the central limit law as stated by Theorem 6.5 is more detailed: not only can we write down the mean and variance of the limiting Gaussian random variables, we can also write down the covariance function of the limiting Gaussian process. Hence, the joint distribution of $n_{t_1}, \ldots, n_{t_k}$, observed at census times $t_1, \ldots, t_k$, can be approximated by a k-dimensional Gaussian distribution with $\mathbb{E}n_{t_i} \simeq Nx_{t_i} + \sqrt{N}\mu_{t_i}$ and $\mathrm{Cov}\,(n_{t_i}, n_{t_j}) \simeq Nc_{t_i,t_j}$, where $c_{t,s} := \mathrm{Cov}\,(Z_t, Z_s) = V_t \prod_{u=t}^{s-1} f'(x_u)$ $(s \geq t)$; μ_t and V_t are given by (6.19). Of course, Theorem 6.5 demonstrates that these results apply to the EC case as well (a new result) where f and v are chosen appropriately.

Remark. These limit theorems were proved without using any of the conditional state distributions provided by Lemma 3.1, Theorem 3.1 or Proposition 4.1. Thus, instead of evaluating such distributions and observing their behaviour as $N \to \infty$, the limiting behaviour of *any* of our N-patch models with density-dependent seasonal phases can be determined by simply observing the conditional means and variances of those phases. It should be clear, then, that this programme extends to models with more than two density-dependent phases. For example, our methods can easily accommodate an extinction-reproduction-settlement dynamic like that described by Klok and De Roos [50] (although note that their particular model does not directly translate into our setting, since their reproduction and settlement transition matrices are non-square).

6.3.2 Long-term behaviour

We start by investigating the long-term ($t \to \infty$) behaviour of the deterministic process ($x_t : t \geq 0$), where, in an obvious notation, we use subscripts 'EC' and 'CE' to distinguish between the same quantities for the respective models. The LLN for N-patch models tells us that

$$f_{\mathrm{CE}}((1-e)x) = (1-e)f_{\mathrm{EC}}(x), \tag{6.20}$$

and so, letting x^* denote a fixed point, the fixed points of f_{CE} and f_{EC} are related by

$$x^*_{\mathrm{CE}} = (1-e)x^*_{\mathrm{EC}}. \tag{6.21}$$

From (6.20), it is easy to see that $f'_{\mathrm{CE}}((1-e)x) = f'_{\mathrm{EC}}(x)$ and $(1-e)f''_{\mathrm{CE}}((1-e)x) = f''_{\mathrm{EC}}(x)$, and so (6.21) tells us that $f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f'_{\mathrm{EC}}(x^*_{\mathrm{EC}})$ and $(1-e)f''_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f''_{\mathrm{EC}}(x^*_{\mathrm{EC}})$. Hence, x^*_{CE} and x^*_{EC} share the same stability properties. It is now a simple matter of recalling results presented in Section 4.4.2 where the stability properties were evaluated for the deterministic CE model in terms of the parameter $\rho = e/(1-e)$. Remember that x^*_{CE} is a fixed point of f_{EC} if and only if $\hat{c}(x^*_{\mathrm{CE}}) = r(x^*_{\mathrm{CE}})$, where $r(x) = \rho x/(1-x)$, and that, since $\hat{c}$ is strictly increasing from $\hat{c}(0) \geq 0$ and concave with $\hat{c}(x) \leq 1$, there is always precisely one unique stable fixed point in $[0,1]$. We may summarise the long-term behaviour of both the EC and CE models as follows.

(i) *Stationarity*: $\hat{c}(0) > 0$. There is a unique fixed point x^* in $[0,1]$, which is stable and satisfies $0 < x^* < 1$. The Markov chain has a unique stationary distribution in this case and we expect this distribution to be centred near Nx^*.

(ii) *Evanescence*: $\hat{c}(0) = 0$ and $\hat{c}'(0) \leq \rho$. The unique fixed point in $[0, 1]$ is 0 and it is stable. The Markov chain has an absorbing state (state 0) in this case, which it reaches in finite time, however, under the condition that $\hat{c}'(0) \leq \rho$, we expect the stochastic process to be absorbed quickly (even for large N).

(iii) *Quasi stationarity*: $\hat{c}(0) = 0$ and $\hat{c}'(0) > \rho$. There are two fixed points in $[0, 1]$, namely 0 and $x^* \in (0, 1)$; 0 is unstable and x^* is stable. The Markov chain reaches its absorbing state (state 0) in finite time, however, under the condition that $\hat{c}'(0) > \rho$, quasi-equilibrium will be reached and we expect the unique quasi-stationary distribution (being the limiting conditional state probabilities, conditional on non-extinction) of the chain to be centred near Nx^*.

These results are in complete agreement with our comments made in Section 4.6, where we suggested that the same stability criteria describes both deterministic models and that the equilibrium proportion of occupied patches is larger in the EC case. Indeed, we now know that the equilibrium value for the CE model is smaller by a factor of $1-e$ and that the deterministic trajectory is uniformly smaller for the CE model (since $f_{\text{CE}}(x) < f_{\text{EC}}(x)$). These remarks are supported by illustrations in plots (a), (b) and (c) of Figure 6.1, where simulations are depicted for both the EC and CE models with $\hat{c}(x) = cx$ (see Example 6.1 below), as well as the corresponding deterministic trajectories and quantities relating to the limiting Gaussian processes. Figure 6.1(d) depicts the quasi-stationary distribution, $p^N = (p_i^N, i \in E_N)$, of $n_\bullet$ in the EC case (where, from Section 2.2.3, this distribution was evaluated as the normalized left eigenvector of the transition matrix restricted to E_N corresponding to its Perron-Frobenius eigenvalue), and this was compared with the approximating Gaussian pdf with mean Nx^* and variance NV^*, where $V^* = v(x^*)/(1 - f'(x^*)^2)$ (see Corollary 6.1 below).

Our next result shows that in the stationary and quasi-stationary cases, where there is a positive stable deterministic equilibrium x^*, the fluctuations $Z_t^N = \sqrt{N}(X_t^N - x^*)$ of $X_\bullet^N$ about x^* can be approximated by an AR-1 process whose parameters can be exhibited explicitly. It follows from Theorems 6.4 and 6.5 by setting $x_0 = x^*$.

Corollary 6.1 *For the N-patch metapopulation models with parameters e and $\hat{c}(x)$, let $Z_t^N = \sqrt{N}(X_t^N - x^*)$ where $X_t^N = n_t/N$. In addition to $\hat{c}$ being twice continuously differentiable, suppose that (i) $\hat{c}(0) > 0$ or (ii) $\hat{c}(0) = 0$ and $\hat{c}'(0) > e/(1-e)$. Further let x^* be the stable fixed* *point of f where f is given as in Theorem 6.4. Then, $X_0^N \xrightarrow{\text{P}} x^*$ implies that $X_t^N \xrightarrow{\text{P}} x^*$ for all $t \geq 1$, in which case if $Z_0^N \xrightarrow{\text{D}} z_0$, then $Z_\bullet^N$ converges weakly to the AR-1 process $Z_\bullet$ defined by $Z_{t+1} = f'(x^*)Z_t + E_t$, where (E_t) are iid with $E_t \sim \mathrm{N}(0, v(x^*))$ and v as given in Theorem 6.5.*

One consequence of the corollary is that $Z_t^N \xrightarrow{\text{D}} \mathrm{N}(z_0 a^t, V_t)$ for all $t \geq 1$, where $a = f'(x^*)$ is the decay rate and now

$$V_t = v(x^*)(1 - a^{2t})/(1 - a^2). \tag{6.22}$$

Another is that there will be a sequence of times (t_N) such that $Z_{t_N}^N \xrightarrow{\text{D}} \mathrm{N}(0, V^*)$, where

$$V^* = v(x^*)/(1 - a^2), \tag{6.23}$$

which is to say that, as $N \to \infty$ and $t \to \infty$, we expect $Z_\bullet^N$ to converge weakly to a stationary AR-1 process with variance V^*. Also, we expect that if the process has reached equilibrium/quasi equilibrium then the joint distribution of the numbers of occupied patches, observed at census times $t_1, \dots, t_k$, can be approximated by an k-dimensional Gaussian distribution with means $N x_{t_i} + \sqrt{N}\mu_{t_i}$ and covariances $N c_{t_i, t_j}$, where $\mu_{t_i} = z_0 a^t$, $c_{t,s} := \mathrm{Cov}\,(Z_t, Z_s) = V_t\, a^{|s-t|}$ $(s \geq t)$ and V_t is given by (6.22).

We have already noted the simple relationship between the deterministic equilibria of our two models, $x^*_{\text{CE}} = (1-e)x^*_{\text{EC}}$, and that the decay rates are the same: $a = f'_{\text{CE}}(x^*_{\text{CE}}) = f'_{\text{EC}}(x^*_{\text{EC}})$. The stationary variances of the approximating AR-1 processes are also related. First, because $\hat{c}(x^*) = r(x^*)$, it is easy to prove that

$$v_{\text{CE}}(x^*_{\text{CE}}) = e x^*_{\text{CE}} \left(2 - e/(1 - x^*_{\text{CE}})\right). \tag{6.24}$$

And, since it can also be shown that

$$(1-e)^2 v_{\text{EC}}(x) = v_{\text{CE}}((1-e)x) + e\big((1-e)x(f'_{\text{CE}}((1-e)x))^2 - f_{\text{CE}}((1-e)x)\big),$$

we have

$$(1-e)^2 v_{\text{EC}}(x^*_{\text{EC}}) = v_{\text{CE}}(x^*_{\text{CE}}) - e x^*_{\text{CE}}(1 - a^2), \tag{6.25}$$

and therefore $(1-e)^2 V^*_{\text{EC}} = V^*_{\text{CE}} - e x^*_{\text{CE}}$.

6.3.3 Examples

We shall illustrate these results by revisiting the four particular instances of $\hat{c}(x)$ examined in Section 4.5, noting that we are now able to explore both the EC and CE models in each case.

Example 6.1 Suppose that $\hat{c}(x) = cx$ $(0 < c \leq 1)$. In this case we may write

$$f(x) = x(1 + r(1 - x/x^*)),$$

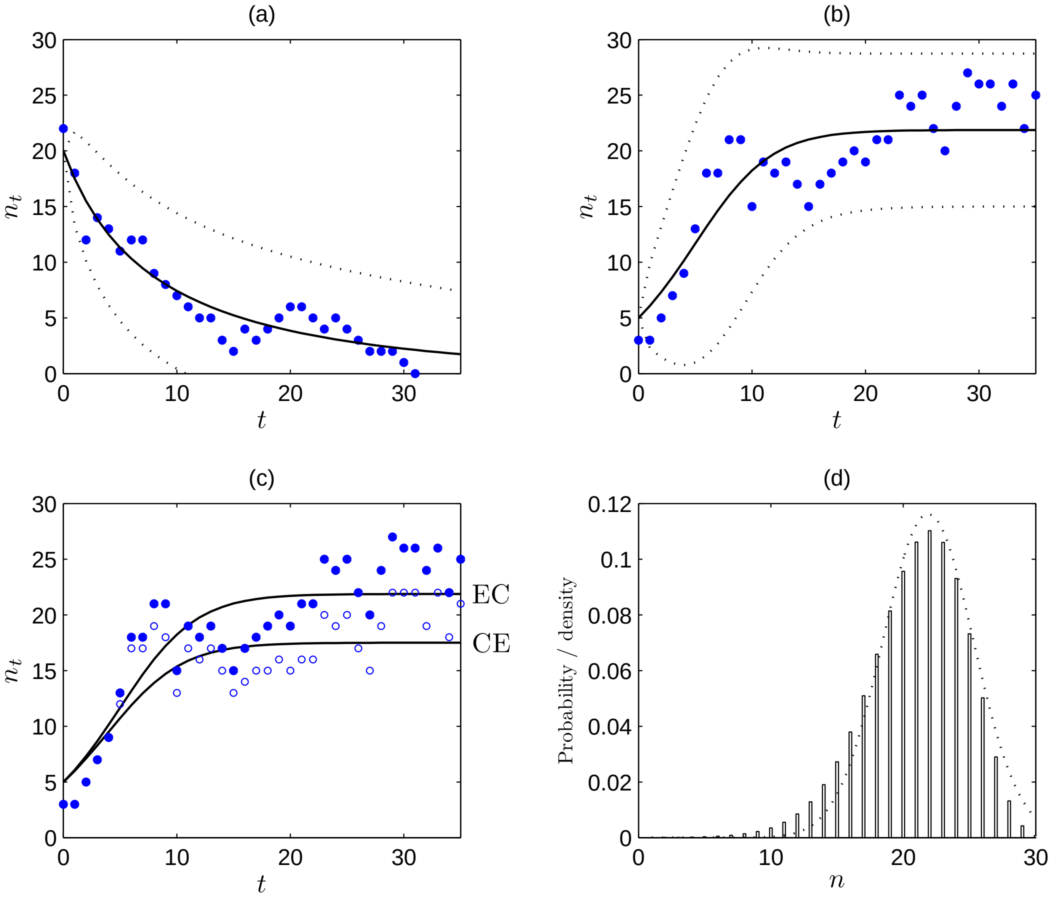

Figure 6.1: N-patch metapopulation model with $N = 30$, $e = 0.2$ and $\hat{c}(x) = cx$. Simulation (solid circles) of the EC model with (a) $c = 0.2$ (evanescence) and (b) $c = 0.6$ (quasi stationarity); deterministic trajectories are shown (solid), together with ±2 standard deviations of the Gaussian approximation (dotted). (c) Simulation of the EC (solid circles) and CE (open circles) models with $c = 0.6$; both deterministic trajectories shown (solid). (d) Quasi-stationary distribution (bars) of the EC model with $c = 0.6$ and the stationary Gaussian pdf (dotted).

where $r = c(1-e) - e$ for both models and x^* is the appropriate equilibrium value: $x^*_{\text{EC}} = r/(c(1-e)^2)$ or $x^*_{\text{CE}} = r/(c(1-e))$ (both being strictly positive, and then stable, if and only if $c > e/(1-e)$). Thus, in both cases, the limiting deterministic trajectory follows the discrete-time logistic growth model (again, see Section 3.2 of Renshaw [75]), where r is the intrinsic growth rate and x^* is the carrying capacity expressed as a proportion of the ceiling N, although note that neither of the deterministic models exhibit any chaotic or periodic behaviour since $0 < 1 + r = (1-e)(1+c) < 2$. Our limiting Gaussian Markov chain has error variance

$$v(x) = (1-e)x\big[c(1-(1-e)x)(1-c(1-e)x) + e\big(1+c-2c(1-e)x\big)^2\big] \qquad \text{(EC model)}$$

$$v(x) = (1-e)x\big[e + c(1-x)(1-c(1-e)x)\big]. \qquad \text{(CE model)}$$

In the quasi-equilibrium case ($r > 0$), the limiting AR-1 process is defined by $Z_{t+1} = aZ_t + E_t$, where $a = 1 - r$ ($0 < a < 1$), with $E_t \sim \mathrm{N}(0, v(x^*))$, where, from (6.24) and (6.25), $v_{\text{CE}}(x^*_{\text{CE}}) = er(1-e+a)/(e+r)$ and $(1-e)^2 v_{\text{EC}}(x^*_{\text{EC}}) = er(a(1+a)-e)/(e+r)$. Since $a \in (0,1)$, the stationary variance $V^* = v(x^*)/(1-a^2)$ of $Z_{\bullet}$ decreases with a and so the faster the decay in the mean, the smaller the stationary variance.

Example 6.2 Suppose that $\hat{c}(x) = c_0 + cx$ ($c_0 > 0$, $c > 0$ and $c_0 + c \le 1$). Now we may write

$$f(x) = \nu + x(1 + r(1 - x/K)),$$

where $\nu = c_0(1-e)$ and $r = (c-c_0)(1-e) - e$ for both models, and K depends on which model: $K_{\text{EC}} = r/(c(1-e)^2)$ or $K_{\text{CE}} = r/(c(1-e))$. Since $\hat{c}(0) = c_0 > 0$, we have unique stable equilibria $x^*_{\text{EC}} = x^*_{\text{CE}}/(1-e)$ with x^*_{CE} being the unique positive solution to $c(1-e)x^2 - rx - \nu = 0$. The error variance is given by

$$\begin{aligned} v(x) = {} & (1-(1-e)x)(c_0 + c(1-e)x)(1 - c + 0 - c(1-e)x) \\ & + e(1-e)x\big[1 - c_0 - 2c(1-e)x + c\big]^2, \qquad \text{(EC model)} \end{aligned}$$

$$v(x) = (1-e)\big[ex + (1-x)(c_0 + cx)\big(1 - (1-e)(c_0 + cx)\big)\big]. \qquad \text{(CE model)}$$

The limiting AR-1 process is again defined by $Z_{t+1} = aZ_t + E_t$, but now $a = 1 + r(1 - 2x^*/K)$, where $E_t \sim \mathrm{N}(0, v(x^*))$ with $v(x^*)$ given by (6.24) or (6.25). The stationary variance is given by (6.23).

Example 6.3 Suppose that $\hat{c}(x) = 1 - \exp(-\beta x)$ $(\beta > 0)$. Since $\hat{c}(0) = 0$ and $\hat{c}'(0) = \beta > 0$, we have evanescence if $\beta \leq e/(1-e)$ and quasi stationarity if $\beta > e/(1-e)$. The error variance is given by

$$\begin{aligned} v(x) &= \exp(-\beta(1-e)x)\Big((1-(1-e)x)(1-\exp(-\beta(1-e)x)) \\ &\qquad + e(1-e)x\big[1+(1-(1-e)x)\beta\big]^2\exp(-\beta(1-e)x)\Big), && \text{(EC model)} \\ v(x) &= (1-e)\big[ex + (1-x)(1-\exp(-\beta x))\big(e+(1-e)\exp(-\beta x)\big)\big]. && \text{(CE model)} \end{aligned}$$

In the quasi-stationary case the deterministic equilibria cannot be exhibited explicitly, but can be evaluated numerically by iterating the map $f_{\mathrm{CE}}(x) = (1-e)(1-(1-x)\exp(-\beta x))$, remembering that $x^*_{\mathrm{EC}} = x^*_{\mathrm{CE}}/(1-e)$. The limiting AR-1 process has stationary variance $v(x^*)/(1-a^2)$, where $v(x^*)$ is evaluated using (6.24) or (6.25). A simple calculation reveals that

$$a = \frac{(1+\beta(1-x^*_{\mathrm{CE}}))(1-e-x^*_{\mathrm{CE}})}{(1-e)(1-x^*_{\mathrm{CE}})}.$$

Example 6.4 Suppose that $\hat{c}(x) = c_0$ $(0 < c_0 \leq 1)$. From Chapter 3, we know that the metapopulation behaves as if, at every census, each occupied patch remains occupied with probability p, and, *independently*, each unoccupied patch is colonised with probability q, where

$$\begin{aligned} p &= 1-e(1-c_0) & q &= c_0 && \text{(EC model)} \\ p &= 1-e & q &= (1-e)c_0. && \text{(CE model)} \end{aligned}$$

We proved that, for all $t \geq 1$, n^N_t has the same distribution as the sum of two independent random variables, $\mathrm{Bin}(n^N_0, p_t)$ and $\mathrm{Bin}(N - n^N_0, q_t)$, with success probabilities $q_t = q^*(1-a^t)$ and $p_t = q_t + a^t$ $(t \geq 0)$, where $a = p - q = (1-e)(1-c_0)$ (the same for both EC and CE) and $q^* = q/(1-a)$. The proportion X^N_t of occupied patches at time t has mean and variance given by $\mathbb{E}X^N_t = x_t(X^N_0)$ and $N\mathrm{Var}(X^N_t) = V_t(X^N_0)$, where

$$x_t(x_0) = x_0 p_t + (1-x_0)q_t \quad \text{and} \quad V_t(x_0) = x_0 p_t(1-p_t) + (1-x_0)q_t(1-q_t).$$

So, on the one hand, as $t \to \infty$, $\mathbb{E}X_t^N$ converges to q^* at geometric rate a (note that $0 < a < 1$) and $N\mathrm{Var}(X_t^N) \to q^*(1-q^*)$ (indeed, $n_\bullet^N$ has a $\mathrm{Bin}(N, q^*)$ stationary distribution). On the other, letting $N \to \infty$ with t fixed, $\mathbb{E}X_t^N \to x_t(x_0)$ and $N\mathrm{Var}(X_t^N) \to V_t(x_0)$ whenever $X_0^N \xrightarrow{\mathrm{P}} x_0$. Furthermore, because n_t^N is the sum of two independent binomial random variables, it is clear that $Z_t^N := \sqrt{N}(X_t^N - x_t)$ and $Y_t^N := \sqrt{N}(X_t^N - q^*)$ will converge in distribution to Gaussian random variables if their initial values converge. Theorem 6.5 and Corollary 6.1 provide more detailed information. Since $f(x) = px + q(1-x)$, and hence $f'(x) = a$, and $v(x) = p(1-p)x + q(1-q)(1-x)$, we deduce that if $Z_0^N \xrightarrow{\mathrm{D}} z_0$, then $Z_\bullet^N$ converges weakly to a Gaussian Markov chain $Z_\bullet$ with $\mathbb{E}Z_t = a^t z_0$ and $\mathrm{Cov}(Z_t, Z_s) = V_t(x_0)\, a^{|s-t|}$, while if $Y_0^N \xrightarrow{\mathrm{D}} y_0$ (a constant), then $Y_\bullet^N$ converges weakly to an AR-1 process $Y_\bullet$ with $\mathbb{E}Y_t = a^t y_0$ and $\mathrm{Cov}(Y_t, Y_s) = q^*(1-q^*)a^{|s-t|}(1-a^{2t})$; the error variance here is $v(q^*) = q^*(1-q^*)(1-a^2)$.

Next we explain how the inhomogeneous Markov chain approach may be applied to our infinite-patch metapopulation models.

6.4 Infinite-patch metapopulation models

As before, let n_t be the number of occupied patches at time t, but now suppose that $(n_t : t \geq 0)$ is a discrete-time Markov chain taking values in $S = \{0, 1, 2, \dots\}$. We shall be concerned with chains that evolve as follows:

$$\text{EC:} \qquad n_{t+1} = \tilde{n}_t + \mathrm{Poi}(m(\tilde{n}_t)) \qquad \tilde{n}_t = n_t - \mathrm{Bin}(n_t, e) \tag{6.26a}$$

$$\text{CE:} \qquad n_{t+1} = \tilde{n}_t - \mathrm{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \mathrm{Poi}(m(n_t)), \tag{6.26b}$$

where $m(n) \geq 0$. The local extinction and colonisation phases alternate over time in the usual way, where each occupied patch goes extinct independently with probability e $(0 < e < 1)$, but now the colonisation phase follows a Poisson law where $m(n)$ is the mean number of empty patches colonised during that phase; the mean $m(n)$ is a function of the number n of occupied patches. Even though there is no ceiling on the number of occupied patches (indeed, we interpret the metapopulation network as having infinitely-many habitat patches in this setting), the dependence of m on n_t allows us to account for a range of colonising behaviours. For example, we have already seen three special cases of (6.26) that correspond to *island*, *mainland-island* and *mainland* models in Chapter 5 (where, of course, we showed that these special cases can be interpreted as Galton-Watson processes).

The following section considers what happens when the initial number of occupied patches becomes large. We will suppose that there is an index N such that $m(n) = N\mu(n/N)$, where the function μ is continuous with bounded first derivative. We may take N to be simply n_0 or, more generally, following Klebaner [48], we may interpret N as being a "threshold" with the property that $n_0/N \to x_0$ as $N \to \infty$. By choosing μ appropriately, we can still allow for a degree of regulation in the colonisation process. Apart from the three examples in Chapter 5 (where, clearly, we may take $\mu(x) = mx$, $\mu(x) = mx + m_0$ and $\mu(x) = m_0$, respectively) $\mu(x)$ might be of the form

$$\mu(x) = rx(a - x) \qquad (0 \le x \le a) \qquad \text{(logistic growth)}$$

$$\mu(x) = x\exp(r(1 - x)) \qquad (x \ge 0) \qquad \text{(Ricker growth)}$$

$$\mu(x) = \lambda x/(1 + ax)^b \qquad (x \ge 0) \qquad \text{(Hassell growth)}$$

(see Sections 3.2 and 4.5 of Renshaw [75], for example).

6.4.1 Limit theorems

We shall consider the time-inhomogeneous Markov chain $(n_t^N : t \ge 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$ with $n_\bullet$ and $\tilde{n}_\bullet$ as given by (6.26). With $X_t^N = n_t^N/N$, the number of occupied patches at census t measured *relative to* the threshold, the sequences (f_t) and (v_t) satisfy

$$f_{2t}(x) = f_0(x) = (1 - e)x \qquad f_{2t+1}(x) = f_1(x) = x + \mu(x) \tag{6.27}$$

$$v_{2t}(x) = v_0(x) = e(1 - e)x \qquad v_{2t+1}(x) = v_1(x) = \mu(x) \tag{6.28}$$

for the EC case and

$$f_{2t}(x) = f_0(x) = x + \mu(x) \qquad f_{2t+1}(x) = f_1(x) = (1 - e)x \tag{6.29}$$

$$v_{2t}(x) = v_0(x) = \mu(x) \qquad v_{2t+1}(x) = v_1(x) = e(1 - e)x \tag{6.30}$$

for the CE case. Again, the roles of f_0 and f_1 as well as v_0 and v_1 simply reverse depending on the model of interest. All of these functions are Lipschitz continuous and, given that X_t^N is unbounded, we can apply Theorem 6.2 to the inhomogeneous density process: given $X_0^N \xrightarrow{2} x_0$ as $N \to \infty$, then $X_\bullet^N$ converges in probability to a limiting deterministic trajectory $x_\bullet$ that satisfies, in particular, $x_{2(t+1)} = f(x_{2t})$ with $f = f_1 \circ f_0$. We thus have the following result for the time-homogeneous density process.

Theorem 6.6 (LLN for infinite-patch models) *For the infinite-patch metapopulation models with parameters e and $\mu(x)$, let $X_t^N = n_t/N$ be the number of occupied patches at census t relative to the threshold N. Suppose that μ is continuous with bounded first derivative. If $X_0^N \xrightarrow{2} x_0$ as $N \to \infty$, then $X_t^N \xrightarrow{2} x_t$ (and hence $X_t^N \xrightarrow{\mathrm{P}} x_t$) for all $t \geq 1$, where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with*

$$f(x) = (1-e)x + \mu((1-e)x) \qquad \text{(EC model)}$$

$$f(x) = (1-e)(x + \mu(x)). \qquad \text{(CE model)}$$

Remark. Indeed, the CE model is always density-dependent because

$$\mathbb{E}(n_{t+1}|n_t) = (1-e)(n_t + m(n_t)) = (1-e)(n_t + N\mu(n_t/N)),$$

implying that $f_t(x) = f(x) = (1-e)(x + \mu(x))$, and

$$\begin{aligned}\operatorname{Var}(n_{t+1}|n_t) &= e(1-e)\mathbb{E}(\tilde{n}_t|n_t) + (1-e)^2 \operatorname{Var}(\tilde{n}_t|n_t) \\ &= e(1-e)(n_t + m(n_t)) + (1-e)^2 m(n_t) \\ &= e(1-e)(n_t + N\mu(n_t/N)) + (1-e)^2 N\mu(n_t/N),\end{aligned}$$

implying that $v_t(x) = v(x) = (1-e)(ex + \mu(x))$. Both f and v are Lipschitz continuous in x, since μ is continuous with bounded first derivative, and so we could have applied Theorem 6.2 directly in this case. It was necessary to work with the phases separately in the EC case because this model is not always density-dependent.

Having established that $X_t^N \xrightarrow{\mathrm{P}} x_t$ for all $t \geq 0$, we can also prove a central limit law for the scaled fluctuations $Z_t^N = \sqrt{N}(X_t^N - x_t)$. First observe that our infinite-patch models can be represented as

$$n_{t+1} = \tilde{n}_t + \textstyle\sum_{j=1}^{N} \mathrm{Poi}_j(\mu(\tilde{n}_t/N)) \qquad \tilde{n}_t = n_t - \textstyle\sum_{j=1}^{n_t} \mathrm{Ber}_j(e) \qquad \text{(EC model)}$$

$$n_{t+1} = \tilde{n}_t - \textstyle\sum_{j=1}^{\tilde{n}_t} \mathrm{Ber}_j(e) \qquad \tilde{n}_t = n_t + \textstyle\sum_{j=1}^{N} \mathrm{Poi}_j(\mu(n_t/N)), \qquad \text{(CE model)}$$

where the $(\mathrm{Poi}_j(\,\cdot\,))$ are collections of iid Poisson random variables with mean $\mu(n_t/N)$. Thus, we may apply Theorem 6.3 to the time-inhomogeneous Markov chain $(n_t^N : t \geq 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$ with $n_\bullet$ and $\tilde{n}_\bullet$ as given directly above. This inhomogeneous chain exhibits the form (6.3), but now $(\pm\,\xi_{jt}^N)$ alternates between appropriate sequences of iid Poisson random variables and iid Bernoulli random variables. For both cases, $g_{2t}(x) = g_{2t+1}(x) = x$. For the EC case,

$$\begin{aligned} &r_{2t}(x) = x \qquad r_{2t+1}(x) = 1 \\ &m_{2t}(x) = -e \qquad m_{2t+1}(x) = \mu(x) \\ &\sigma_{2t}^2(x) = e(1-e) \qquad \sigma_{2t+1}^2(x) = \mu(x), \end{aligned}$$

which lead to the same sequences (f_t) and (v_t) as given by (6.27) and (6.28), and $b_{2t}(x) = e(1-e)(1-2e)$ and $b_{2t+1}(x) = \mu(x)$. For the CE case,

$$\begin{aligned} &r_{2t}(x) = 1 \qquad r_{2t+1}(x) = x \\ &m_{2t}(x) = \mu(x) \qquad m_{2t+1}(x) = -e, \\ &\sigma_{2t}^2(x) = \mu(x) \qquad \sigma_{2t+1}^2(x) = e(1-e), \end{aligned}$$

which lead to (f_t) and (v_t) as given by (6.29) and (6.30), and $b_{2t}(x) = \mu(x)$ and $b_{2t+1}(x) = e(1-e)(1-2e)$. Hence, the roles f_0 and f_1 as well as v_0 and v_1 are reversed (again). If the model satisfies the stronger condition that μ be twice continuously differentiable with bounded second derivative, then, in both cases, $f_t(x)$ will be twice continuously differentiable in x with bounded second derivative, $v_t(x)$ will be continuous in x, and $b_t(x)$ will be bounded in x. We already know that the limiting deterministic trajectory satisfies $x_{2(t+1)} = f(x_{2t})$, where $f = f_1 \circ f_0$, and so from Theorem 6.3 it is clear that our limiting Gaussian Markov process $Z_\bullet$ should take the form $Z_{2(t+1)} = f'(x_{2t})Z_{2t} + \hat{E}_{2t}$, with $\hat{E}_{2t} \sim \mathrm{N}(0, v(x_{2t}))$, where $v = v_1 \circ f_0 + (f_1' \circ f_0)^2 v_0$, with

$f_1'(x) = 1 + \mu'(x)$ in the EC case and $f_1'(x) = 1 - e$ in the CE case. Thus we arrive at the following result.

Theorem 6.7 (CLL for infinite-patch models) *For the infinite-patch metapopulation models with parameters e and $\mu(x)$, suppose that μ is twice continuously differentiable with bounded second derivative. Let $X_t^N = n_t/N$ be the proportion of occupied patches at census t, and suppose that $X_0^N \xrightarrow{2} x_0$, so that $X_t^N \xrightarrow{\text{P}} x_t$ for all $t \geq 0$, where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with f given as in Theorem 6.6. Let $Z_t^N = \sqrt{N}(X_t^N - x_t)$ and suppose that $Z_0^N \xrightarrow{\text{D}} z_0$. Then, $Z_\bullet^N$ converges weakly to the Gaussian Markov process $Z_\bullet$ defined by*

$$Z_{t+1} = f'(x_t)Z_t + E_t,$$

with (E_t) independent and $E_t \sim \text{N}(0, v(x_t))$, where

$$v(x) = \mu((1-e)x) + e(1-e)x\,(1 + \mu'((1-e)x)) \qquad \text{(EC model)}$$

$$v(x) = (1-e)(ex + \mu(x)). \qquad \text{(CE model)}$$

The central limit law for the infinite-patch models tell us that $Z_t^N \xrightarrow{\text{D}} Z_t \sim \text{N}(\mu_t, V_t)$ for any $t \geq 1$, where μ_t and V_t are given by (6.19) (same as for the N-patch models but with f as given in Theorem 6.6 and v as given in Theorem 6.7). The covariance function is given by $c_{t,s} := \text{Cov}\,(Z_t, Z_s) = V_t \prod_{u=t}^{s-1} f'(x_u)$ $(s \geq t)$.

6.4.2 Long-term behaviour

Using the notation adopted in Section 6.3.2, notice that the infinite-patch deterministic models are related by $f_{\text{CE}}((1-e)x) = (1-e)f_{\text{EC}}(x)$ and so the fixed points of f_{CE} and f_{EC} are related by $x^*_{\text{CE}} = (1-e)x^*_{\text{EC}}$. As before, $f'_{\text{CE}}(x^*_{\text{CE}}) = f'_{\text{EC}}(x^*_{\text{EC}})$ and $(1-e)f''_{\text{CE}}(x^*_{\text{CE}}) = f''_{\text{EC}}(x^*_{\text{EC}})$, implying that x^*_{CE} and x^*_{EC} have the same stability properties. Notice also that x^*_{CE} will be a fixed point of f_{CE} if and only if

$$\mu(x^*_{\text{CE}}) = \rho x^*_{\text{CE}}$$

where $\rho = e/(1-e)$. So, if $\mu(0) = 0$ then 0 is a fixed point; it is stable if $\mu'(0) < 1$ and unstable if $\mu'(0) > 1$ (if $\mu'(0) = 1$ its stability is determined by higher derivatives of μ near $x = 0$).

However, even when $\mu(0) = 0$, there might be other (conceivably many) fixed points since our conditions on μ do not preclude this.

We cannot be as precise in classifying the equilibrium behaviour as we were earlier for the N-patch models, however we can still present two results concerning the fluctuations of X_t^N about a positive stable equilibrium point (Corollary 6.2) or a stable limit cycle (Corollary 6.3). First, if there is a unique positive fixed point x^*, it will be stable if $\mu'(x^*) < 1$ and unstable if $\mu'(x^*) > 1$ (again we need to consider higher derivatives when $\mu'(x^*) = 1$). The next result follows by setting $x_0 = x^*$ in Theorem 6.7.

Corollary 6.2 *Suppose that f given in Theorem 6.6 admits a unique positive fixed point x^* satisfying $\mu'(x^*) < 1$. Then, if $X_0^N \xrightarrow{2} x^*$, $x_t = x^*$ for all t and, assuming $Z_0^N \xrightarrow{\mathrm{D}} z_0$, the limiting process $Z_{\bullet}$ determined by Theorem 6.7 is an AR-1 process $Z_{\bullet}$ defined by $Z_{t+1} = aZ_t + E_t$, where $a = (1-e)(1+\mu'(x^*_{\mathrm{CE}}))$ (being the same for both models), and (E_t) are idd where $E_t \sim \mathrm{N}(0, v)$, with $v = e(1+a)x^*$ (EC model) or $v = e(2-e)x^*$ (CE model).*

Second, notice that the d-th iterates of our maps are also related by

$$f_{\mathrm{CE}}^{(d)}((1-e)x) = (1-e)f_{\mathrm{EC}}^{(d)}(x),$$

which means that if $x_0^*, x_1^*, \ldots, x_{d-1}^*$ is a stable limit cycle for the deterministic EC model, then $(1-e)x_0^*, (1-e)x_1^*, \ldots, (1-e)x_{d-1}^*$ is a stable limit cycle for the deterministic CE model. The next corollary follows by setting $x_0 = x_0^*$ in Theorem 6.7, where we see that the deterministic trajectory $x_{\bullet}$ tracks the limit cycle, $x_{td+j} = x_j^*$ $(t \geq 0, j = 0, \ldots, d-1)$, and the limiting process $Z_{\bullet}$ is a d-variate AR-1 process $\boldsymbol{Y}_{\bullet}$ (Definition 2.13). The representation of $\boldsymbol{Y}_{\bullet}$, as well as the particular form of the coefficient matrix A and the error covariance matrix Σ_d, follow by iterating $Z_{t+1} = f'(x_t)Z_t + E_t$, with (E_t) independent $\mathrm{N}(0, v(x_t))$ random variables: using expressions (6.9) to (6.12) with $f_t = f$ and $v_t = v$, noting that

$$\Pi_{i,j} = \prod_{k=i}^{j-1} f'(x_k^*) = a_j/a_i \qquad (1 \leq i \leq j \leq d),$$

we obtain a representation of Z_{td+j} $(j = 1, \ldots, d)$ in terms of Z_{td} $(t \geq 0)$, as well as the stationary covariance matrix V. Note that here and henceforth the subscript td represents the time variable t as being multiplied by the index d.

Corollary 6.3 *Suppose that* f *given in Theorem 6.6 admits a stable limit cycle* $x_0^*, x_1^*, \ldots, x_{d-1}^*$ *with* $X_0^N \xrightarrow{2} x_0^*$. *Then,* $x_{td+j} = x_j^*$ $(t \geq 0,\ j = 0, \ldots, d-1)$ *and, assuming* $Z_0^N \xrightarrow{\mathrm{D}} z_0$, *the limiting process* $Z_{\bullet}$ *determined by Theorem 6.7 has the following representation:* $(\boldsymbol{Y}_t : t \geq 0)$, *where* $\boldsymbol{Y}_t = (Z_{td}, Z_{td+1}, \ldots, Z_{(t+1)d-1})^\top$ *is a* d*-variate AR-1 process of the form* $\boldsymbol{Y}_{t+1} = A\boldsymbol{Y}_t + \boldsymbol{E}_t$, *where* $(\boldsymbol{E}_t)$ *are independent and* $\boldsymbol{E}_t \sim \mathrm{N}(\mathbf{0}, \Sigma_d)$; *here* A *is the* $d \times d$ *matrix*

$$A = \begin{pmatrix} 0 & 0 & \cdots & a_1 \\ 0 & 0 & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_d \end{pmatrix},$$

where $a_j = \prod_{i=0}^{j-1} f'(x_i^*)$, $\Sigma_d = (\sigma_{ij})$ *is the* $d \times d$ *symmetric matrix with entries*

$$\sigma_{ij} = a_i a_j \sum_{k=0}^{i-1} v(x_k^*)/a_{k+1}^2 \qquad (1 \leq i \leq j \leq d),$$

where v *is given as in Theorem 6.7, and the random entries,* $(Z_1, \ldots, Z_{d-1})$, *of* $\boldsymbol{Y}_0$ *have a Gaussian* $\mathrm{N}(\boldsymbol{a}z_0, \Sigma_{d-1})$ *distribution, where* $\boldsymbol{a} = (a_1, \ldots, a_{d-1})$. *Furthermore,* $\boldsymbol{Y}_{\bullet}$ *has a Gaussian* $\mathrm{N}(\mathbf{0}, V)$ *stationary distribution, where* $V = (v_{ij})$ *has entries given by*

$$v_{ij} = \frac{a_i a_j}{1 - a_d^2} \sum_{k=0}^{d-1} v(x_k^*)/a_{k+1}^2 \qquad (1 \leq i \leq j \leq d).$$

Remark. We established that there is always precisely one stable fixed point for the deterministic N-patch models (Section 6.3.2) and so we did not have to concern ourselves with results like that outlined in Corollary 6.3. We note that such results would be necessary if, for example, $\hat{c}$ was no longer assumed to be concave, since then the deterministic N-patch model could possibly have more than one stable fixed point or have a stable limit cycle. If this were the case, then an N-patch analogue of Corollary 6.3 is simply determined by following the same method outlined above but applied to the autoregressive process in Theorem 6.5.

To illustrate the results for infinite-patch models, we look at the case where the colonisation probability obeys a Ricker law.

6.4.3 An example

Suppose that $\mu(x) = x\exp(r(1-x))$, where $r > 0$. The colonisation probability is greatest when the number of occupied patches is close to N/r and so we may interpret the parameter r as the intrinsic growth rate and N as the carrying capacity of the metapopulation *in the absence of extinction*. The fixed points of f_{CE} are 0 and $x^*_{\mathrm{CE}} = 1 - r_0/r$, where $r_0 = \log(\rho)$. Notice that $f'_{\mathrm{CE}}(x) = (1-e)(1+(1-rx)\exp(r(1-x)))$ and $f''_{\mathrm{CE}}(x) = -(1-e)(2-rx)r\exp(r(1-x))$, implying that $f'_{\mathrm{CE}}(0) = (1-e)(1+\exp(r))$ and $f''_{\mathrm{CE}}(0) = -2(1-e)r\exp(r)$. Therefore, if $r \le r_0$, 0 is the unique non-negative fixed point, and it is stable. If $r > r_0$, then x^*_{CE} is an additional positive fixed point; it is stable because $f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = 1 - e(r - r_0) < 1$ (and 0 is unstable). However, if r is sufficiently large, we get limiting cycles with period doubling towards chaos, as illustrated in Figure 6.2.

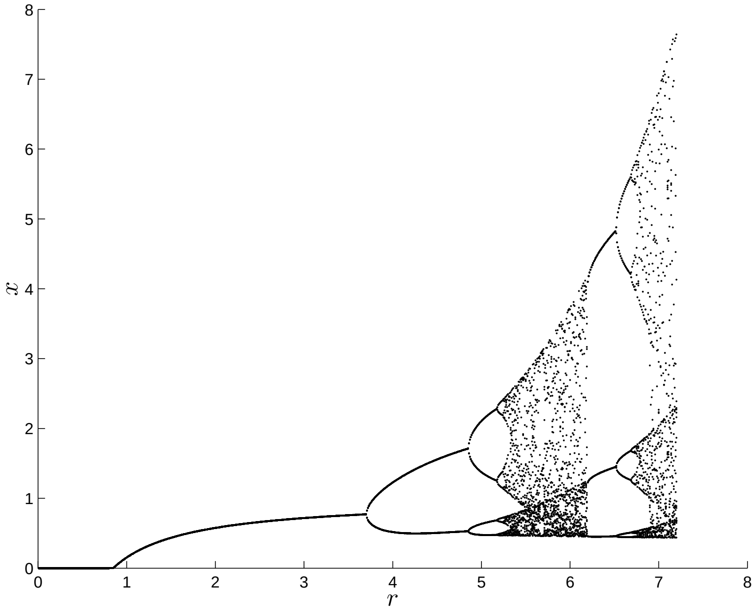

Figure 6.2: Bifurcation diagram for the infinite-patch deterministic CE model with Ricker growth dynamics: $x_{t+1} = (1-e)x_t(1+\exp(r(1-x_t)))$. Here $e = 0.7$ and r ranges from 0 to 7.2. The vertical axis plots the possible equilibrium values x of the difference equation.

Figure 6.3 illustrates some of the range of behaviour exhibited by the stochastic CE model with Ricker growth dynamics. The cases depicted in plots (a)-(d) are described below.

(a) *Evanescence*: $0 < r < r_0$ ($r_0 \simeq 0.8473$). The unique stable fixed point is 0, and the process dies out quickly.

(b) *Quasi stationarity*: $r_0 < r < r_1$ ($r_1 \simeq 3.7$). Here x^*_{CE} ($\simeq 1 - 0.8473/r$) is the unique stable fixed point, and the process exhibits quasi-equilibrium behaviour. The parameter r_1 is the point of first period-doubling, which we have not been able to determine analytically.

(c) *Oscillation*: $r > r_1$, with period 2. The process 'tracks' the limit cycles of the deterministic model.

(d) *Oscillation*: same as for (c) but with period 4.

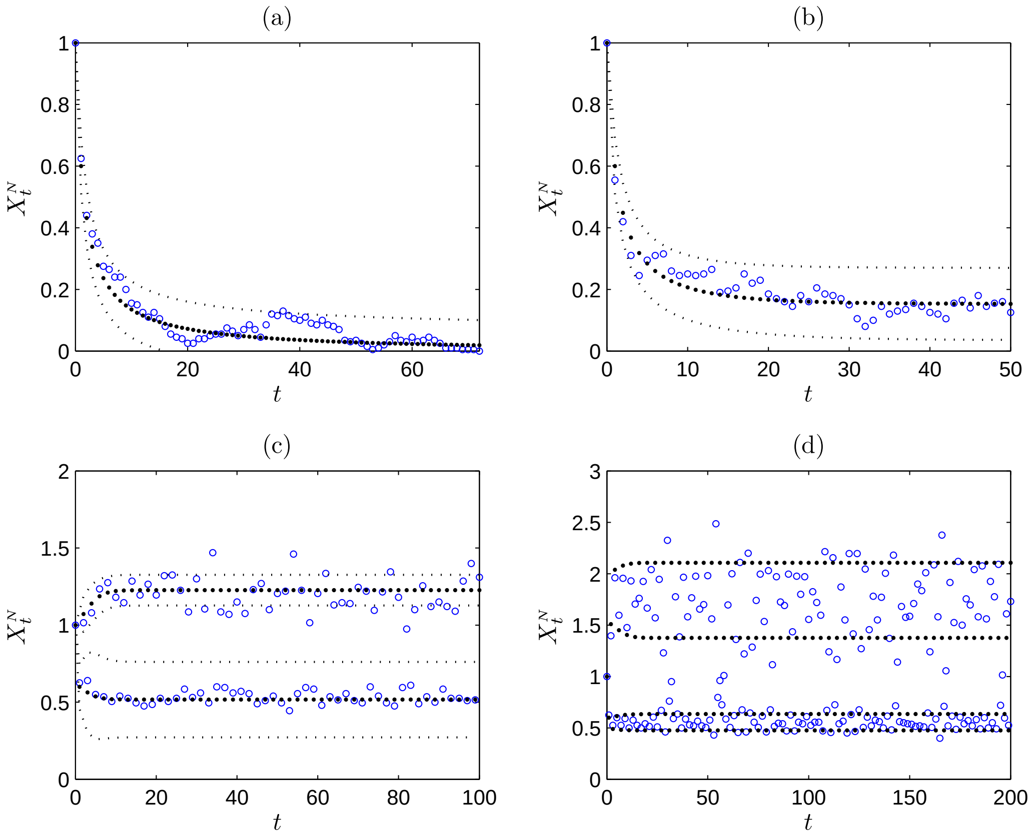

Figure 6.3: Simulation (open circles) of the infinite-patch CE model with Ricker growth dynamics, together with the corresponding limiting deterministic trajectories (small solid circles). Here $e = 0.7$ and $N = 200$, and, (a) $r = 0.84$, (b) $r = 1$ (c) $r = 4$ and (d) $r = 5$. In (a), (b) and (c), the dotted lines indicate ±2 standard deviations of the Gaussian approximation (in (c) every *second* point is joined to indicate variation about each of the two limit cycle values).

Plots (a)-(d) include the limiting deterministic trajectories while plots (a)-(c) also include ±2 standard deviations as determined by the appropriate Gaussian approximation.

6.5 Discussion

In this chapter, we proved limit theorems for a class of time-inhomogeneous Markov chains that exhibit the particular property of density dependence. From these results we established limit theorems, a law of large numbers and a central limit law, for our finite-patch and infinite-patch metapopulation models defined with density-dependent phases. Indeed, the central limit laws exhibit a stronger form of convergence than that displayed by central limit laws in Sections 3.3.2 and 4.4.1; here we proved convergence for a finite sequence of times instead of point-wise convergence at a particular time t. This type of convergence enables us to look at the limiting process over time and calculate such quantities as a covariance function. Also, in each of the finite-patch and infinite-patch settings, we successfully treated both of the EC and CE monitoring schemes.

The particular method we used to establish these limit theorems for our discrete-time metapopulation models was to work with the density-dependent phases separately. Although this is not necessary in all situations (such as the CE models, for example), the method does not require the use of time-homogeneous conditional state distributions and therefore easily accommodates seasonal population models with more than two density-dependent phases.

Chapter 7

Conclusion

I have presented several stochastic metapopulation models that describe the number of occupied patches in a discrete-time setting. In particular, these models are based on the discrete-time Markov chain approach first used by Akçakaya and Ginzburg [2], and so account for a dynamic where local extinction events and colonisation events occur in distinct seasonal phases. Extensive analytical treatments of these models were provided. For example, I showed that some models are highly tractable and that the limiting behaviour of any discrete-time Markov chain with density-dependent phases can be exhibited explicitly. This chapter summarises the results of these treatments, discusses some applications of the work and looks at possible future directions.

In Chapters 3 and 4, I investigated a class of *chain binomial metapopulation models*, so called because the number of events that occur during each of the local extinction and colonisation phases follow binomial laws. The simplest chain binomial model, the *mainland* model, was analysed in Chapter 3. Many quantities of interest were evaluated explicitly in this special case, including the 1-step and t-step conditional state distributions, the equilibrium distribution and the expected first passage time to state 0. Also, a time-dependent interpretation of the *rescue effect* was given and limit theorems that describe large networks were proved. In Chapter 4, the chain binomial model was examined in a more general setting where the colonisation probability was allowed to depend on the current number of occupied patches. There, I demonstrated how to construct an *island* model or a *mainland-island* model by choosing an appropriate form for the colonisation probability. In the general setting, only the 1-step conditional state distribution for the CE case was evaluated and so, for the particular case where the colonisation probability depends on the number of occupied patches through the *proportion* of occupied patches, limit theorems that describe large networks were proved.

In Chapter 5, three different models that describe metapopulation networks with infinitely-many habitat patches were presented. These models also accounted for a two-phase dynamic: the local extinction process followed a binomial law and the colonisation process followed a Poisson law, where the mean of the Poisson law depended on the current number of occupied patches in such a way that the three configurations—mainland, island and mainland-island—were accommodated. In this way, these models may be viewed as infinite-patch analogues of the chain binomial (finite-patch) models. The 1-step conditional state distribution was evaluated in each case and it was shown that all three infinite-patch models could be reinterpreted as branching processes. Standard branching theory was used to answer questions about the long-term behaviour of these models. The infinite-patch mainland model proved to be particularly tractable, since the t-step conditional state distribution and the equilibrium distribution were also evaluated.

Finally, in Chapter 6, limit theorems were proved for finite-patch and infinite-patch models defined with density-dependent phases. These theorems established convergence in the limit as $N \to \infty$, where N represented the total number of patches in the finite-patch setting and marked a certain threshold in the infinite-patch setting. I included a law of large numbers, that identified an approximating deterministic trajectory, and a central limit law, which showed that the scaled fluctuations about this trajectory have an approximating autoregressive structure, where both EC and CE monitoring schemes were treated. These results were made possible by using limit theorems that were proved for a particular class of time-inhomogeneous density-dependent Markov chains. The results for these chains allowed us to work with the density-dependent phases separately and effectively by-passed the need for evaluating time-homogeneous conditional state distributions. This method provides greater scope for model analysis because it easily accommodates models with any number of density-dependent phases.

My research thus provides (i) a range of tractable discrete-time Markov chain models that account for a two-phase dynamic and (ii) methods that allow practitioners to approximate any discrete-time Markov chain model defined with density-dependent phases. In particular, the approximating models are Gaussian Markov processes whose structures are exhibited explicitly. There is plenty of scope for practitioners to account for other colonisation behaviours of interest, to develop new models that account for dynamics with more than two density-dependent phases and, indeed, to adapt any of these models to a range of seasonal dynamics outside of the metapopulation context (including epidemics, stock market fluctuations, climate patterns, to name a few). Although these types of Markov chain models are not spatially explicit, I have demonstrated that spatial information can be incorporated implicitly by allowing the seasonal phases to depend on the current state of the system.

In terms of applying these results to real metapopulations, practitioners need to be able to estimate model parameters from observed data. Practitioners can apply standard maximum likelihood techniques to the tractable models, since the 1-step and t-step conditional state distributions that have been provided can be used to construct likelihood functions. For the intractable models, as well as the tractable models, practitioners can use time-series methods to estimate parameters of the approximating autoregressive process. Indeed, there is even a third option for the N-patch mainland model and that is to estimate parameters by way of McKenzie's binomial AR-1 process [62] (see Section 2 of [92] for further details). It may be that a general framework can be developed for estimating parameters of density-dependent Markov chain models via an approximating autoregressive model. A possible avenue for future work would therefore be to develop this framework.

In terms of broadening the range of stochastic models that account for seasonal dynamics, future projects might focus on developing models in a continuous-time Markov chain setting. In Section 6 of our most recent paper [15], we looked at a number of well known continuous-time Markov chain population models as analogues of our finite-patch and infinite-patch models, however these analogues assume that local extinction events and colonisation events occur in random order over time. This is an assumption that holds in many continuous-time Markov chain metapopulation models [3, 5, 6, 7, 9, 82, 83, 72]. Of course, in order to account for seasonal phases, the transition rates of the continuous-time Markov chain must be time-dependent. For example, the formulation outlined in a recent paper by Ross [81] achieves this by predefining the time-length of each phase and then assigning a different set of transition rates to each phase. In this way, direct comparisons could then be made with our discrete-time models.

References

[1] H. Abbey. An examination of the Reed-Frost theory of epidemics. *Human Biology*, 24: 201–233, 1952.

[2] H. R. Akçakaya and L. R. Ginzburg. Ecological risk analysis for single and multiple populations. In A. Seitz and V. Loeschcke, editors, *Species Conservation: A Population-Biological Approach*, pages 78–87. Birkhauser, Basel Switzerland, 1991.

[3] D. Alonso and A. McKane. Extinction dynamics in mainland-island metapopulations: an N-patch stochastic model. *Bulletin of Mathematical Biology*, 64:913–958, 2002.

[4] H. G. Andrewartha and L. C. Birch. *The Distribution and Abundance of Animals*. University of Chicago Press, Chicago USA, 1954.

[5] F. Arrigoni. Approssimazione deterministica di modelli stocastici di metapopulazione. *Bollettino dell'Unione Matematica Italiana, Serie VIII*, IV-A:387–390, 2001.

[6] F. Arrigoni. A deterministic approximation of a stochastic metapopulation model. *Advances in Applied Probability*, 3:691–720, 2003.

[7] F. Arrigoni and A. Pugliese. Limits of a multi-patch SIS epidemic model. *Journal of Mathematical Biology*, 45:419–440, 2002.

[8] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, New York, 1972.

[9] A.D. Barbour and A. Pugliese. Asymptotic behaviour of a metapopulation model. *Annals of Applied Probability*, 15:1306–1338, 2005.

[10] G. P. Basharin, A. N. Langville, and Naumov V. A. The life and work of A. A. Markov. *Linear Algebra and its Applications*, 386:3–26, 2004.

[11] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 3rd edition, 1995.

[12] S. A. Boorman and P. R. Levitt. Group selection on the boundary of a stable population. *Theoretical Population Biology*, 4:85–128, 1973.

[13] J. H. Brown and A. Kodric-Brown. Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology*, 58:445–449, 1977.

[14] F. M. Buckley and P. K. Pollett. Analytical methods for a stochastic mainland-island model. *Ecological modelling*, in press, 2010.

[15] F. M. Buckley and P. K. Pollett. Limit theorems for discrete-time metapopulation models. *Probability Surveys*, 7:53–83, 2010.

[16] F. M. Buckley and P. K. Pollett. Analytical methods for a stochastic mainland-island metapopulation model. In R. S. Anderssen, R. D. Braddock, and L. T. H. Newham, editors, *18th IMACS World Congress - MODSIM09 International Congress on Modelling and Simulation*, volume July 2009, pages 1767–1773, Canberra, Australia, 2009. Modelling and Simulation Society of Australia and New Zealand.

[17] C. Chatfield. *The Analysis of Time Series: An Introduction*. Chapman & Hall Texts in Statistical Science Series. Chapman & Hall, London, 5th edition, 1996.

[18] P. L. Chesson. Models for spatially distributed populations: the effect of within-patch variability. *Theoretical Population Biology*, 19:288–325, 1981.

[19] F. Chiera. *It's Gnow or Never*. Documentary by filmmaker Franco Chiera, aired on ABC1 Television in Queensland on 4th February, 2010.

[20] D. J. Daley and J. Gani. *Epidemic Modelling: An Introduction*. Cambridge Studies in Mathematical Biology. Cambridge University Press, Cambridge UK, 1999.

[21] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, 2:88–100, 1965.

[22] J. R. Day and H. P. Possingham. A stochastic metapopulation model with variability in patch size and position. *Theoretical Population Biology*, 48:333–360, 1995.

[23] M. Doebeli. Dispersal and dynamics. *Theoretical Population Biology*, 47:82–106, 1995.

[24] M. Doebeli and G. D. Ruxton. Evolution of dispersal rates in metapopulation models: branching and cyclic dynamics in phenotype space. *Evolution*, 51:1730–1741, 1997.

[25] P. Ehrenfest and T. Ehrenfest. Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem. *Physikalische Zeitschrift*, 8:311–314, 1907.

[26] W. Feller. *An Introduction to Probability Theory and its Applications: Volume II.* Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, Inc., New York, 1966.

[27] J. Foster and J. A. Williamson. Limit theorems for the galton-watson process with time-dependent immigration. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 20:227–235, 1971.

[28] J. Gani and D. Jerwood. Markov chain methods in chain binomial epidemic models. *Biometrics*, 27:591–603, 1971.

[29] J. Gani and L. Stals. The spread of a viral infection in a plantation. *Environmetrics*, 15: 555–560, 2004.

[30] J. Gani and S. Yakowitz. Modelling the spread of HIV among intravenous drug users. *IMA Journal of Mathematics Applied in Medicine and Biology*, 10:51–65, 1993.

[31] J. Gani, C. R. Heathcote, and D. F. Nicholls. A model for HIV infection through needle sharing. *Journal of Mathematical and Physical Sciences*, 25:665–675, 1991.

[32] B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical transactions of the Royal Society of London*, 115:513–585, 1825.

[33] J. L. Gonzalez-Andujar and J. N. Perry. Chaos, metapopulations and dispersal. *Ecological Modelling*, 65:255–263, 1993.

[34] M. Greenwood. On the statistical measure of infectiousness. *Journal of Hygiene*, 31: 336–351, 1931.

[35] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes*. Oxford University Press, Oxford, 3rd edition, 2001.

[36] M. Gyllenberg, G. Söderbacka, and S. Ericsson. Does migration stablize local population dynamics? Analysis of a discrete metapopulation model. *Mathematical Biosciences*, 118: 25–49, 1993.

[37] M. Gyllenberg, I. Hanski, and A. Hastings. Structured metapopulation models. In I. A. Hanski and Gilpin M. E., editors, *Metapopulation biology: ecology, genetics and evolution*, pages 93–122. Academic Pres, San Diego, 1997.

[38] P. Haccou, P. Jagers, and V. A. Vatutin. *Branching Processes: Variation, Growth and Extinction of Populations*. Cambridge Studies in Adaptive Dynamics. Cambridge University Press, Cambridge UK, 2005.

[39] I. Hanski. *Metapopulation Ecology*. Oxford Series in Ecology and Evolution. Oxford University Press, Oxford UK, 1999.

[40] I. Hanski. Single species metapopulation dynamics: concepts, models and observations. *Biological Journal of the Linnean Society*, 42:17–38, 1991.

[41] I. Hanski. Inferences from ecological incidence functions. *The American Naturalist*, 139: 657–662, 1992.

[42] I. Hanski. A practical model of metapopulation dynamics. *Journal of Animal Ecology*, 63: 151–162, 1994.

[43] I. Hanski and M. Gilpin. Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society*, 42:3–16, 1991.

[44] I. Hanski and O. Ovaskainen. Metapopulation theory for fragmented landscapes. *Theoretical Population Biology*, 64:119–127, 2003.

[45] A. Hastings. Complex interactions between dispersal and dynamics: lessons from coupled logistic equations. *Ecology*, 74:1362–1372, 1993.

[46] M. F. Hill and H. Caswell. The effects of habitat destruction in finite landscapes: a chain-binomial metapopulation model. *Oikos*, 93:321–331, 2001.

[47] W. G. Kelley and A. P. Peterson. *Difference Equations: An Introduction with Applications*. Academic Press, Boston, 1991.

[48] F. C. Klebaner. Population-dependent branching processes with a threshold. *Stochastic Processes and Their Applications*, 46:115–127, 1993.

[49] F. C. Klebaner and O. Nermann. Autoregressive approximation in branching processes with a threshold. *Stochastic Processes and Their Applications*, 51:1–7, 1994.

[50] C. Klok and A. M. De Roos. Effects of habitat size and quality on equilibrium density and extinction time of *Sorex araneus* populations. *Journal of Animal Ecology*, 67:195–209, 1998.

[51] R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15:237–240, 1969.

[52] R. Levins. The effect of random variations of different types of population growth. *Proceedings of the National Academy of Sciences of the United States of America*, 62:1061–1065, 1969.

[53] R. Levins. Extinction. In M. Gerstenhaber, editor, *Some Mathematical Questions in Biology*, pages 75–107. American Mathematical Society, Providence RI, 1970.

[54] R. C. Lewontin and D. Cohen. On population growth in a randomly varying environment. *Proceedings of the National Academy of Sciences of the United States of America*, 62: 1056–1060, 1969.

[55] R. H. MacArthur and E. O. Wilson. An equilibrium theory of insular zoogeography. *Evolution*, 17:373–387, 1963.

[56] R. H. MacArthur and E. O. Wilson. *The Theory of Island Biogeography*. Princeton University Press, New York, 1967.

[57] T. R. Malthus. *An Essay on the Principle of Population*. London, 1798.

[58] R. M. May. Stability in model ecosystems. *Proceedings of the Ecological Society of Australia*, 6:18–56, 1971.

[59] R. M. May. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. *Science*, 186:645–647, 1974.

[60] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261: 459–467, 1976.

[61] R. M. May and G. F. Oster. Bifurcations and dynamic complexity in simple ecological models. *The American Naturalist*, 110:573–599, 1976.

[62] E. McKenzie. Some simple models for discrete variate time series. *Water Resources Bulletin*, 21:645–650, 1985.

[63] E. McKenzie. Discrete variate time series. In C.R. Rao and D. N. Shanbhag, editors, *Handbook of Statistics*, pages 573–606. Elsevier Science, Amsterdam, 2003.

[64] R. S. McVinish and P. K. Pollett. Limits of large metapopulations with patch dependent extinction probabilities. *Advances in Applied Probability*, 42:1172–1186, 2010.

[65] A. Moilanen. Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. *Ecology*, 80:1031–1043, 1999.

[66] J. R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, 1997.

[67] K. Parvinen. Evolution of migration in a metapopulation. *Bulletin of Mathematical Biology*, 61:531–550, 1999.

[68] K. Parvinen. Evolution of dispersal in a structured metapopulation model in discrete time. *Bulletin of Mathematical Biology*, 68:655–678, 2006.

[69] R. Pearl. *The Biology of Population Growth*. Alfred A. Knopf, New York, 1925.

[70] R. Pearl. The growth of populations. *Quarterly Review of Biology*, 2:532–548, 1920.

[71] R. Pearl and L. J. Reed. On the rate of growth of the population of the united states since 1790 and its mathematical representation. *Proceedings of the National Academy of Sciences of the United States of America*, 6:275–288, 1920.

[72] P. K. Pollett. Diffusion approximations for ecological models. In F. Ghassemi, editor, *Proceedings of the International Congress on Modelling and Simulation*, volume 2, pages 843–848, Canberra, Australia, 2001. Modelling and Simulation Society of Australia and New Zealand.

[73] H. P. Possingham. Decision theory and biodiversity management: how to manage a metapopulation. In R. B. Floyd, A. W. Sheppard, and P. J. De Barro, editors, *Frontiers of Population Ecology*, Melbourne, Australia, 1996. CSIRO Publishing.

[74] H. R. Pulliam. Sources, sinks and population regulation. *American Naturalist*, 132:652–661, 1988.

[75] E. Renshaw. *Modelling Biological Populations in Space and Time*, volume 11 of *Cambridge Studies in Mathematical Biology*. Cambridge University Press, Cambridge, 1991.

[76] S. A. Richards, H. P. Possingham, and J. Tizard. Optimal fire management for maintaining community diversity. *Ecological Applications*, 9:880–892, 1999.

[77] D. A. Roff. Spatial heterogeneity and the persistence of populations. *Oecologica*, 15: 245–258, 1974.

[78] D. A. Roff. The analysis of a population model demonstrating the importance of dispersal in a heterogeneous environment. *Oecologica*, 15:259–275, 1974.

[79] J. Ross, T. Taimre, and P. Pollett. On parameter estimation in population models. *Theoretical Population Biology*, 70:498–510, 2006.

[80] J. Ross, D. Pagendam, and P. Pollett. On parameter estimation in population models II: multi-dimensional processes and transient dynamics. *Theoretical Population Biology*, 75: 123–132, 2009.

[81] J. V. Ross. Computationally exact methods for stochastic periodic dynamics: spatiotemperal dispersal and temporally forced transmission. *Journal of Theoretical Biology*, 262: 14–22, 2010.

[82] J. V. Ross. Stochastic models for mainland-island metapopulations in static and dynamic landscapes. *Bulletin of Mathematical Biology*, 68:417–449, 2006.

[83] J. V. Ross. A stochastic metapopulation model accounting for habitat dynamics. *Journal of Mathematical Biology*, 52:788–806, 2006.

[84] T. M. Rout, C.E. Hauser, and H. P. Possingham. Minimise long-term loss or maximise short-term gain? Optimal translocation strategies for threatened species. *Ecological Modelling*, 201:67–74, 2007.

[85] J. T. Sandefur. *Discrete Dynamical Systems: Theory and Applications*. Clarendon Press, Oxford, 1990.

[86] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, 1981.

[87] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *Journal of Applied Probability*, 3:403–434, 1966.

[88] B. Tenhumberg, A. J. Tyre, K. Shea, and H. P. Possingham. Linking wild and captive populations to maximise species persistence: optimal translocation strategies. *Conservation Biology*, 18:1304–1314, 2004.

[89] P. F. Verhulst. Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique*, X:113–121, 1838.

[90] H. W. Watson and F. Galton. On the probability of the extinction of families. *Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.

[91] D. Waxman and S. Gavrilets. 20 questions on adaptive dynamics. *Journal of Evolutionary Biology*, 18:1139–1154, 2005.

[92] C. H. Weiß. A new class of autoregressive models for time series of binomial counts. *Communication in Statistics: Theory and Methods*, 38:447–460, 2009.

[93] E. B. Wilson and M. H. Burke. The epidemic curve. *Proceedings of the National Academy of Sciences of the United States of America*, 28:361–367, 1942.

[94] S. Wright. Breeding structure of populations in relation to speciation. *American Naturalist*, 74:232–248, 1940.

[95] A. M. Yaglom. Certain limit theorems of the theory of branching processes. *Doklady Academii Nauk SSSR*, 56:795–798, 1947.

[96] A. Yakubu and C. Castillo-Chavez. Interplay between local dynamics and dispersal in discrete-time metapopulation models. *Journal of Theoretical Biology*, 218:273–288, 2002.

[97] Y. Zhang, L. Liu, and R. Xu. The effect of migration on the viability, dynamics and structure of two coexisting metapopulations. *Ecological modelling*, 220:272–282, 2009.