Spatially structured metapopulation models within static and dynamic environments

Andrew G. Smith

A thesis submitted for the degree of Doctor of Philosophy
at The University of Queensland in 2014
School of Mathematics and Physics
Abstract

Traditionally in population modelling, the mixing of individuals has been assumed to be homogeneous; that is, every individual can come into contact with every other individual. Within the last 40 years, however, a number of population models have been proposed that do not assume homogeneous mixing but rather assume populations are divided into disjoint habitable patches that are separated by uninhabitable space. Populations with this structure are known as metapopulations.

When metapopulation modelling was first proposed, the habitable patches would be classified as either colonised or extinct and the dynamics of colonisation and extinction would be the only dynamics accounted for in the model. For example, the logistic model adapted to a metapopulation would be

\[ \frac{dx}{dt} = \lambda x (1 - x) - \mu x, \]

where \( x(t) \) is the proportion of occupied patches at time \( t \), \( \lambda \) is the rate an unoccupied patch becomes colonised when all patches are unoccupied and \( \mu \) is the rate an occupied patch becomes extinct. Adding more detail to metapopulation models, the model examined in this thesis records the number of individuals in each location, thereby accounting for an individual’s dynamics, such as the rates of births, deaths and migrations, within the model.

The model is a continuous time Markov process that will be used to account for the demographic stochasticity within populations such as births, deaths and migration events. Results by Kurtz, and extended by Pollett, which can be applied to a family of Markov processes, termed asymptotically density dependent, will be used to determine an approximating system of
differential equations. These differential equations are then analysed to determine conditions for persistence and extinction. Furthermore, an Allee effect, where the initial conditions of the population determine whether it persists or goes extinct, is confirmed to exist in a two patch system that has a large difference between the migration rate for the two patches.

The model is extended in two ways. The first extension accounts for a deterministically changing environment. This is done by allowing the parameters of the system to depend on time. A new functional limit law is derived which can be applied to time inhomogeneous, asymptotically density dependent Markov processes. This functional limit law is used to derive a nonautonomous system of differential equations. This system is then analysed to provide conditions for persistence and extinction of the metapopulation.

The second extension to the model accounts for a stochastically changing environment. Again, the parameters of the system are allowed to vary in time. However, the parameters vary stochastically according to an underlying Markov process, designed to model a stochastically changing environment. A functional limit law provides a way to approximate the process with a random dynamical system. The random dynamical system is then analysed to determine a sufficient condition for extinction. While not a complete description of the long term behaviour, such an approach facilitates more research into models with a stochastically changing environment.
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. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.
Publications during candidature


- Research contained in (A1) was undertaken by Andrew Smith and guided by Prof. Pollett and Dr McVinish. Recommendations for mathematical corrections and improvements in writing were provided by Prof. Pollett and Dr McVinish.

Publications included in this thesis


- The majority of Chapter 3 is based on (B1).

<table>
<thead>
<tr>
<th>Contributor</th>
<th>Statement of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. G. Smith</td>
<td>Performed research (100%)</td>
</tr>
<tr>
<td></td>
<td>Wrote the paper (95%)</td>
</tr>
<tr>
<td>R. McVinish</td>
<td>Wrote the paper (5%)</td>
</tr>
<tr>
<td></td>
<td>Edited paper</td>
</tr>
<tr>
<td>P. K. Pollett</td>
<td>Edited paper</td>
</tr>
</tbody>
</table>

Contributions by others to the thesis

Research contained in this thesis has been guided by Prof. Pollett and Dr McVinish and suggestions on the work were provided by Prof. Pollett and Dr McVinish as the work progressed. Suggestions for improvements in writing and directions for mathematical problems were also provided by Prof. Pollett and Dr McVinish.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.
Acknowledgements

I would like to thank my principal supervisor, Phil Pollett. He has always been there to help further my research during my time at the University of Queensland. The encouragement I have received from Phil has been invaluable and motivates me to further my career in research. Furthermore, due to his role in MASCOS, I have been exposed to many interesting fields of research and made important contacts with people in, and because of, MASCOS, for which I am very grateful.

Next, I would like to thank my associate supervisor, Ross McVinish, who has been instrumental in channeling my (often unclear) ideas into tangible concepts. He has been very generous with his time, frequently meeting with me with no notice at all and replying to dozens of (possibly pointless) emails in short spaces of time. Thank you Ross for assisting my research and helping the development of my mathematical techniques.

I would also like to thank some other people at the University of Queensland. Murray Kane has been very helpful on the administration side of my candidature while other admin staff such as Kathryn Lyons, Clair Kettle and IT guru Karl Blakeney have always been able to answer my questions and direct me to where I need to be. Various conversations with, and seminars by, Leonardo Rojas-Nandayapa and Yoni Nazarathy have also been very helpful in the development of new ideas and paths of current and future research. Lastly, I’d like to thank my fellow postgraduate students, Robert Cope, Aminath Shausan, Jessica Chan and Dejan Jovanovic, for their input into my research as well as allowing my assistance with theirs. Especially, I would like to thank Robert, who I shared an office with for the majority of my candidature, for various mathematical, and not-so-mathematical, discussions.
My family has also been a great help in completing this degree. My parents assisted me greatly throughout my entire studies, through my undergraduate and honours degrees and continued their support through my PhD. My wife, Emily, has provided amazing support through times of stress and confusion and still agreed to marry me, knowing what it would be like beforehand. The emotional (and financial) support you have provided is unrivalled.

Lastly, I would like thank my saviour, Jesus Christ, for His unwavering hand in everything I do. He has been my rock when I have got nothing to stand on, and my strength when I have fallen.
Keywords:
density dependent Markov process, differential equation, metapopulation, random dynamical system

Australian and New Zealand Standard Research Classifications (ANZSRC):
ANZSRC code: 010204, Dynamical Systems in Applications, 50%
ANZSRC code: 010406, Stochastic Analysis and Modelling, 30%
ANZSRC code: 010202, Biological Mathematics, 20%

Fields of Research (FoR) Classification:
FoR code: 0102, Applied Mathematics, 80%
FoR code: 0602, Ecology, 20%
## Contents

### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>53</td>
</tr>
</tbody>
</table>

### 1 Introduction

1.1 Population Modelling ........................................ 1
1.2 Metapopulation Modelling .................................... 5
1.3 Environmental Influence ..................................... 12
1.4 Motivation .................................................... 14
1.5 Outline of the thesis ......................................... 15

### 2 Theory & Background

2.1 Introduction .................................................. 17
2.2 Continuous time Markov process .............................. 18
  2.2.1 Stationary Distributions ................................ 21
2.3 Functional Limit Laws ........................................ 23
  2.3.1 Density Dependent Processes ............................ 24
  2.3.2 Asymptotically Density Dependent Processes .......... 26
  2.3.3 Diffusion Approximation ............................... 27
  2.3.4 Functional Limit Laws for inhomogeneous time ...... 30
  2.3.5 Functional Limit Laws for different scalings ...... 31
2.4 Differential Equations ........................................ 34
2.5 Random Dynamical Systems ................................... 37
3 A Spatially Structured Metapopulation Model
3.1 Introduction ................................................................. 43
3.2 Model ................................................................. 45
3.3 Differential Equation Approximation ......................... 47
3.4 Equilibrium Behaviour .................................................. 50
  3.4.1 Extinction ............................................................... 53
  3.4.2 Persistence .............................................................. 58
  3.4.3 Persistence with Immigration ............................... 60
  3.4.4 Allee Effect ........................................................... 62
  3.4.5 Example ................................................................. 68
3.5 Diffusion ................................................................. 69
3.6 Discussion ................................................................. 72

4 Environmental Influence .............................................. 75
4.1 Introduction ................................................................. 75
4.2 Deterministic Influences .............................................. 76
  4.2.1 Model ................................................................. 77
  4.2.2 Differential Equation Approximation ..................... 77
  4.2.3 Equilibrium Behaviour ............................................... 84
  4.2.4 Discussion ............................................................. 89
4.3 Stochastic Influences ................................................... 90
  4.3.1 Model ................................................................. 90
  4.3.2 Piecewise-Deterministic Approximation .................. 91
  4.3.3 Equilibrium Behaviour ............................................... 94
  4.3.4 Discussion ............................................................. 102

5 Conclusion ................................................................. 104
5.1 Summary ................................................................. 104
5.2 Further Research ...................................................... 106
  5.2.1 A Model with a Static Environment ....................... 106
  5.2.2 A Model with a Dynamic Environment ................... 109

References ................................................................. 112
List of Figures

3.1 Illustration of the dynamics for patch $i$ and migration to and from patch $j$. \(47\)

3.2 Illustrating the different behaviour ODE (3.6) exhibits when $r = 0$ for a $J = 2$ system. The red dotted lines are the trajectories for a given initial condition and the black dots are fixed points. The birth rate function used was $b_i(x/M_l) = b_i(0)(1 - x/M_l)$. The parameters used are given in Table 3.1 \(54\)

3.3 The red solid line represents stable fixed points, the blue dashed line represents unstable fixed points and the black dash-dot line represents the first order approximation (3.20c). Parameters used for plots A and B were $b_2 = 0.1$, $d_1 = 0.1$, $d_2 = 0.7$, $\phi_1 = 0.4$, $\phi_2 = 0.9$, $M_1 = 0.99$ and $M_2 = 0.01$, while $\lambda_{10} = 0.4$ and $\lambda_{20} = 0.5$. \(65\)

3.4 The red dotted lines are the deterministic trajectories imposed by the flow (3.6), the squares are unstable fixed points, the dots are stable fixed points and the solid blue lines are the nullclines. Parameters used for plot A were $b_1 = 0.56$, $b_2 = 0.06$, $d_1 = 0.02$, $d_2 = 0.14$, $\phi_1 = 0.55$, $\phi_2 = 0.09$, $M_1 = 0.54$ and $M_2 = 0.46$, while $\lambda_{10} = 0.31$, and $\lambda_{20} = 0.99$. Parameters used for plot B were $b_1 = 5.24$, $b_2 = 7.5$, $d_1 = 1.14$, $d_2 = 5.59$, $\phi_1 = 7.03$, $\phi_2 = 0.29$, $M_1 = 0.78$ and $M_2 = 0.22$, while $\lambda_{10} = 0.06$ and $\lambda_{20} = 0.51$. \(65\)
3.5 An illustration of the effect the migration rates and the connectivity have on the persistence of a metapopulation with a population ceiling of 200 sea otters at Attu and Amchitka Islands. For the three cases mentioned in the text, A, D and G illustrate contours of the distribution of the two-dimensional approximating Gaussian diffusion (Theorem 3.12) at $t = 40$ years, together with one realisation of the metapopulation model; B, E and H show the same realisation (solid blue line) on a phase plane diagram together with the deterministic trajectory (3.6) (dotted red line); and C, F and I show this realisation and the deterministic trajectory along with $\pm 2$ standard deviations (dotted lines) determined from the distribution of the approximating Gaussian diffusion.

4.1 Illustration of the dynamics for patch $i$ and migration to and from patch $j$.

4.2 In A–C, the population density for three sets of parameters are plotted, where the thick lines are the deterministic trajectory given by (4.17), the thin lines are one realisation of the stochastic process detailed in Section 4.2.1. In D–F, the birth rate per-capita functions are plotted against time. In all plots, blue lines represent patch one and green lines represent patch two. The parameters used for A & D are $b = 2.1$, $b_1 = 1$, $d = 1$, $\phi = 1$, $M_1 = 0.75$, $M_2 = 0.25$, $\lambda_{12} = 0.2$ and $\lambda_{21} = 0.6$. The parameters used for B & E are $b = 1$, $b_1 = 1$, $d = 0.5$, $\phi = 1.5$, $M_1 = 0.75$, $M_2 = 0.25$, $\lambda_{12} = 1/3$ and $\lambda_{21} = 1$. And the parameters used for C & F are $b = 1$, $b_1 = 1$, $d = 1$, $\phi = 1$, $M_1 = 0.75$, $M_2 = 0.25$, $\lambda_{12} = 0.2$ and $\lambda_{21} = 0.6$.

5.1 A phase plane diagram for a $J = 3$ system, where trajectories are the red lines and fixed points are the circle (stable) and square (unstable) markers. The parameters used for this plot were $b_1(0) = 0.54$, $d_1 = 0.8$, $\phi_1 = 0.94$, $M_1 = 0.27$, $\lambda_{12} = 0.3$, $\lambda_{13} = 0.37$ and $b_2(0) = 4.42$, $d_2 = 1.1$, $\phi_2 = 4.82$, $M_2 = 0.67$, $\lambda_{21} = 0.61$, $\lambda_{23} = 0.2$ and $b_3(0) = 0.24$, $d_3 = 2.63$, $\phi_3 = 2.12$, $M_3 = 0.06$, $\lambda_{31} = 0.18$, $\lambda_{32} = 0.51$. 
Introduction

The evolution of the size of a population has long been of interest to both ecologists and mathematicians. A range of behaviours, from predetermined choices to random influences, have been singled out as important characteristics of populations and techniques that account for these behaviours have been developed. A select history of these developments is outlined in this chapter, providing the reader with motivation for the need of continued research in this area.

1.1 Population Modelling

Determining how the size of a population might change through time has been an interest of society for hundreds of years. One of the first insights into how the size of a population might evolve came in the late 18th century, where T. R. Malthus stated “by that law of our nature which makes food necessary to the life of man, the effect of these two unequal powers [population growth and subsistence growth] must be kept equal.” If left unchecked, the population will increase beyond a point that subsistence can provide for. He postulated that the amount subsistence will increase over twenty five years would be, at most, equal to the current subsistence and that this increase would remain unchanged over all twenty five year time periods following this. It is simple to see that if every individual in the population produces another individual in the same time period, there will not be enough subsistence to sustain the population, even after only 50 years. The population would increase exponentially as
Chapter 1: Introduction

the per-capita growth rate would remain constant. Malthus makes some assumptions about the subsistence and the population he was studying that may or may not be reasonable, depending on the species being modelled. Regardless of the type of species being researched, the question of population sustainability is as relevant today as it was in the 18th century.

Forty years later, in the 19th century, P. F. Verhulst [128] proposed a model that could account for the depleting subsistence as a population grew larger, and is commonly referred to as the logistic model. For his model, Verhulst altered the main assumption by Malthus of constant per-capita growth to allow a decreasing linear function of density for the per-capita growth rate. In its most common form, the logistic model introduces an ideal population size and is referred to today as the carrying capacity. As such, the logistic model permits a better understanding of the evolution of the population size when it becomes large. Before Verhulst, most populations would be modelled as if their growth could continue at an exponential rate forever, regardless of the space and resources available to them. For small populations with a substantial amount of resources, this assumption is reasonable. However, such an assumption is not reasonable for large populations. Verhulst modelled the size of a population $n(t)$ by means of the ordinary differential equation (ODE)

$$\frac{dn}{dt} = rn\left(1 - \frac{n}{K}\right),$$

where $r$ is the growth rate of the population and $K$ is the carrying capacity. The growth rate $r$ “is the exponential rate at which a population ... grows when no resource (for example food, space, shelter and water) is in short supply” [31, page 53]. When the population is small, the term $1 - n/K$ is approximately equal to 1 and the population grows almost uninhibited at the rate $rn$. However, as $n$ approaches $K$, the overall growth rate decreases and, if $n$ is ever larger than $K$, will become negative, causing a decrease in the population size. Pearl and Reed [100] proposed the logistic model, using a different parameterisation, for modelling the population of the United States of America.

The carrying capacity is affected by limited food, habitat and reproduction oppor-
tunities which results in competition between individuals. This competition leads to what is known in ecology as negative density dependence or simply, density dependence, where density relates to the population size per unit area. Broadly, it is when population dynamics, such as per-capita birth or death rates, depend on the population size only through the density of the population [63]. Conversely, density independence occurs when dynamics depend on the size of the population, rather than only the density. Therefore, a carrying capacity will always introduce density dependent dynamics into the population.

The majority of references to density dependence in the literature refer to negative density dependence. However, positive (or inverse) density dependence is also an important factor in population dynamics. Unlike negative density dependence, where population growth decreases as the population grows, positive density dependence causes the population growth to increase as the population grows. More commonly referred to as the Allee effect, named after its discoverer Warder Clyde Allee [4, 5], positive density dependence is more prominent in small populations. When the population is small, so is the population growth rate and therefore, the Allee effect can be a contributing factor to a population’s extinction. If the population growth rate decreases sufficiently that it becomes negative as the population size decreases, a threshold is introduced below which the population will tend to extinction. This is known as a strong Allee effect, while a weak Allee effect implies the population growth rate remains positive as the population size decreases. The mechanisms that cause such a phenomenon are usually one or more of the following; genetic inbreeding, demographic stochasticity, a reduction in cooperative interactions [39], mate limitation [73] and predator saturation [116]. When a population is small, the chance of genetic inbreeding is increased and the fitness of the population as a whole is decreased. It is then less likely a population will persist [39]. The second factor, demographic stochasticity, is the random fluctuations in the population size. A few deaths when there are a few hundred individuals will most likely be insignificant. However, when there are only a few individuals in the population, those few deaths become very important [39] and can cause the population to go extinct. Thirdly, a reduction in cooperative interactions refers to situations where the population relies on cooperative behaviour to survive and when the population size is
small, it is less likely this cooperative behaviour can occur. Cooperative behaviour exists in the reproduction and foraging for some species, as a certain population size is required for the activity to be successful. Mate limitation occurs when a certain type of individual that is required for reproduction (for example, the male or female species for sexual reproduction) is scarce and as such, reproduction declines, even though the entire population level might be reasonably high. Lastly, predator saturation is when species can defend themselves successfully as a group (higher population densities), but become more vulnerable as individuals (lower population densities).

Many other population models were produced in the 20th century around the same time as Pearl and Reed’s discovery. One of these models, which is still in use today, was by Lotka who introduced competitive population models for biological systems. Lotka allowed the population and the subsistence to effect each other, rather than only allow the subsistence to effect the population, as in the logistic model. Lotka’s model was also discovered independently by Volterra in 1926, which is why it is known today as the Lotka-Volterra model. When discussing population modelling, one cannot fail to mention the models of B. Gompertz and L. Von Bertalanffy, and the generalisation of these models by F. J. Richards.

Pearl and Reed’s formulation of the logistic model leads to many applications. Their model can be altered to accommodate growth which is attenuated by some form of per-capita removal from the system. Under this assumption, the logistic model can be written in the form

\[
\frac{dn}{dt} = \lambda n \left(1 - \frac{n}{J}\right) - \mu n, \tag{1.1}
\]

where \(\lambda\) is the intrinsic growth rate per capita, \(\mu\) is the removal rate per capita and \(J\) is the number of sites the population can occupy. The logistic model in this form has been applied to many different fields. In ecology, R. Levins in 1969 used this model to described the evolution of a population that is divided into disjoint locations. Rather than model the number of individuals in a system, \(n\) is the number of locations that are currently occupied by individuals which led to the naming of such models as patch occupancy models (POMs). A year later Levins coined the term
metapopulation to describe a “population of populations which go extinct locally and recolonise.” Since its introduction, the meaning of the term metapopulation has evolved to describe a collection of geographically distinct populations that allow migration between populations.

1.2 Metapopulation Modelling

Although Levins gave a precise formulation and named metapopulations in 1969–70 \cite{78,79}, the idea of subgroups of populations dates back to at least 1931, when S. Wright \cite{133,page100} proposed a model for the frequency of genes where “a large population \ldots is composed of subgroups each exchanging the proportion \(m\) of its population with a random sample of the whole population.” Around the same time, A. J. Nicholson and V. A. Bailey \cite{91,page590} also postulated that a population might break up into “numerous small widely separated groups which wax and wane and then disappear, to be replaced by new groups in previously unoccupied situations.” Furthermore, R. H. MacArthur and E. O. Wilson \cite{82,83} had presented a similar idea to metapopulations in the form of island biogeography.

There are two main differences between the field of metapopulations and the field of island biogeography. First, a metapopulation is a population of one species, whereas island biogeography accounts for many different species, both flora and fauna, on each island. The second difference is island biogeography devalues the smaller subgroups as immigration from a large mainland is the main (and sometimes, only) source of recolonisation of the smaller subgroups. This assumption is the main difference between island biogeography became metapopulation ecology. For a more in-depth discussion of island biogeography, see \cite{132}.

The habitable area of many species continues to be divided into smaller sections either by disasters such as bush fires or by humankind’s intervention which can be in the form of new roads and buildings, to list a few. Furthermore, many populations naturally exist in a metapopulation due to the segregation of resources or a naturally occurring limit on pack sizes, for example. As habitable area is segregated and more species are discovered existing in metapopulation structures, understanding the impact
Chapter 1: Introduction

such segregation has on the populations is vital. For example, Crooks et al. [40] found that that the level of mammalian carnivores’ habitat segregation and connectivity severely affected the species likelihood of eventual survival or extinction.

When a species’ habitat is fragmented it is necessary to distinguish between local extinction and extinction of the entire metapopulation. A species can become locally extinct, such that no individuals remain in a particular location. This type of extinction can be caused by demographic fluctuations, genetic factors, predation and unsuitable habitat. Only when all patches have become locally extinct is the metapopulation said to be extinct.

When a patch is locally extinct it may be colonised by another patch which is not locally extinct. Colonisation is linked to the migration of a species and the dispersal of offspring. The frequency of migration is also an important aspect when modelling a species’ population. Some species only migrate at specific times during the year or perhaps only once in their life while others have no predefined migration periods. For example, certain plant populations can be modelled as metapopulations. Once a seed has taken root, it can no longer migrate and hence, colonisation occurs if the seed is dispersed into a previously unoccupied location. Another example comes from the butterfly, Glanville fritillary Melitaea cinxia L., which does not have a set time of migration but instead continually migrates throughout its life [56].

There are times when these two dynamics, local extinction and colonisation, can combine in a way that decreases the rate of local extinction. If a small population is surrounded by large populations and becomes locally extinct, it is far more likely that this population, rather than populations in other configurations, will be rescued from extinction almost immediately from the neighbouring populations through colonisation. This is known as the rescue effect [27].

Extinction and colonisation are the two fundamental dynamics used for the majority of metapopulation models. If only these two dynamics are accounted for, each patch is considered either colonised (occupied) or extinct (unoccupied). This is known as the presence–absence assumption. It is often employed due to the dynamics within a patch occurring at much faster time scale and therefore could be modelled separately from the local extinction and colonisation dynamics. Furthermore presence/absence data is
more widely available for many populations than data regarding population levels is. If Levins’ model \((1.1)\) is used, \(\lambda\) and \(\mu\) are the rate an unoccupied patch is colonised and the rate that an occupied patch becomes extinct, respectively, while \(n\) and \(J\) are the number of occupied patches and the total number of patches available to be occupied, respectively. The inclusion of the factor \(1/J\) in the colonisation rate is a proxy for the probability that an occupied patch and unoccupied patch are close. In Levins’ 1969 \[78\] formulation, these rates can be time (and spatially) dependent; that is, \(\lambda = \lambda(t)\) and \(\mu = \mu(t)\), creating a more general model than the original logistic model.

One effort to introduce spatial dependency into deterministic models was presented by I. Hanski in 2001 \[57\] and has come to be known as the \textit{spatially realistic Levins model} (SRLM). Hanski’s model was originally designed to combine the dynamic theory of island biogeography (DTIB) \[83\] and the classic metapopulation theory (CMT) introduced by Levins \[78\]. MacArthur-Wilson’s model \[83\] in a simple form can be written as a system of \(J\) ordinary differential equations of the form

\[
\frac{ds_i}{dt} = \lambda (S - s_i) - \mu s_i, \tag{1.2}
\]

where \(J\) is the number of patches available to be occupied, \(S\) is the number of species, \(s_i\) is the number of species on patch \(i\) and \(\lambda\) and \(\mu\) are the colonisation and extinction rate parameters, respectively. Models \((1.1)\) and \((1.2)\) were combined by Hanski in 2001 \[57\] to give a system of \(J\) ODEs of the form

\[
\frac{dp_i}{dt} = c_i(p) (1 - p_i) - e_i(p) p_i, \tag{1.3}
\]

where \(c_i(p)\) and \(e_i(p)\) are the colonisation and extinction rate parameters for patch \(i\) and \(p_i\) is the probability of patch \(i\) being occupied. The form of equation \((1.3)\) has its origins in Gyllenberg and Hanski’s work in 1992 \[54\] (and extended five years later \[59\]), where the probability of a successful colonisation was assumed to be unaffected by the size of the population when it was large. Originally, Gyllenberg and Hanski \[54\] modelled the proportion of occupied patches and assumed that both the populations on the patches and the number of patches were very large. From their original model to \((1.3)\), the quantity being modelled changed to the probability a patch is occupied.
Chapter 1: Introduction

One might argue that for large populations the proportion of occupied patches closely models the probability of a patch being occupied. However, even if that statement is assumed to be true, the model only holds if local populations are considered large or zero, and therefore breaks down when the probability of occupancy is close to, but not equal to, zero. As ecologists are most concerned with the extinction of a species, the scenario that (1.3) excludes is the most important.

All the models presented above share one common feature; they are all deterministic. That is, starting with the same initial condition, the end result will always be the same. In nature this is very rarely the case. For example, the exact time an individual reproduces or dies is nearly always random and the same can be said for migration dynamics. When such factors are ignored in the modelling phase, the random fluctuations in the population that result from these factors are ignored too.

A stochastic patch occupancy model (SPOM) is one of the most common models for metapopulation modelling used today. Instead of modelling the probability, like SRLMs, SPOMs model the patches and allow them to alternate between occupied and unoccupied states stochastically, hence, ingrafting stochasticity into the model from the beginning. The closest stochastic analogue to (1.1), and simplest SPOM, is the stochastic logistic model, first proposed by W. Feller in 1939 [49], which will be briefly described now. Let $n(t)$ be the number of occupied patches at time $t$ in a metapopulation with $J$ patches and let $\lambda$ and $\mu$ be as above. Then $(n(t), t \geq 0)$ is a Markov chain, homogenous SPOM with a state space $S = \{1, \ldots, J\}$ and nonzero transitions

\begin{align*}
n \rightarrow n + 1 & \quad \text{at rate } \lambda n(J - n), \\
n \rightarrow n - 1 & \quad \text{at rate } \mu n.
\end{align*}

One type of SPOM, proposed by I. Hanski [56], has gained a large amount of traction in the metapopulation modelling community. This model, known as the incidence function model (IFM), is formulated in discrete time while also including stochasticity in the form of a Markov chain. The differentiating factor of the IFM compared to a general SPOM is that an IFM relates the physical landscape to the transition
probabilities of the Markov chain. Let $X_i(t)$ be the state on patch $i$ at time $t$, such that $X_i(t) = 0$ if patch $i$ is extinct and $X_i(t) = 1$ if patch $i$ is occupied. Patch $i$ has a probability $C_i$ of becoming recolonised in unit time and a probability of $E_i$ of becoming extinct in unit time. Hanski suggested that the extinction probability should be determined by $E_i = x/A_i^y$, where $x$ and $y$ are two constants and $A_i$ is the area of patch $i$. He also suggested the colonisation probability should be calculated by $C_i = (\beta S_i)^2 / \left( (\beta S_i)^2 + a \right)$, where $S_i = \sum_{j\neq i} X_j(t)e^{-\alpha d_{ij}}A_j$ and $a$, $\alpha$ and $\beta$ are constants and $d_{ij}$ is the distance between patches. Hanski’s original IFM has been extended to incorporate different formulations of the colonisation and extinction parameters. While the presence–absence assumption has simplified modelling, data collection and analysis for a number of metapopulation models, it is not always adequate, such as for stock dynamics where more detail is required and fails to provide information relating the individual behaviour to the population behaviour.

The presence–absence assumption, as previously stated, reduces the population dynamics to colonisation and extinction events. As the IFM demonstrates, characteristics of the population can be used to determine more appropriate values for the colonisation and extinction probabilities. Reproductivity, movement out of the system, predation, natural mortality, among others can be combined to give the rate of extinction for a given patch. Similarly, the rate of colonisation is determined from the rate of immigration from outside the metapopulation and the migration rate between patches, both of which are affected by the size of the patches and how the patches are connected but can also be affected by the reproductivity of the species. Therefore, when employing the presence–absence assumption, a trade off appears as the tractability of the model is increased while information that could be gained from the model is decreased. With the presence–absence assumption employed, there are less dynamics to model which results in a simpler model. However, results obtained from the model can only relate to the dynamics included, decreasing the information that can be gathered about the population.

If the presence-absence assumption is not employed, then the within patch population dynamics need to be incorporated into the modelling. These types of models are called structured metapopulation models (SMMs) and have been used by many
modellers \[54, 29, 30, 7, 18, 19, 51, 115, 48\]. SMMs model the births, deaths and migrations of individuals directly, and model the abundance on each patch in some capacity. Usually, either the number, proportion or probability of patches with \(i\) individuals is modelled \[29, 30, 7, 18, 19, 51\] or the number, proportion or probability of individuals on patch \(i\) is modelled \[54, 115, 48\]. The number of patches with \(i\) individuals can be modelled when the only characteristic differentiating the patches is the number of individuals currently occupying them; that is, there is no \textit{patch specific} dynamics. For example, Casagrandi and Gatto \[30\] model a metapopulation as such, using a Markov process where births, deaths and dispersals on a patch with \(i\) individuals occur at the per-capita rates \(\nu_i, \mu_i\) and \(D_i\), respectively. In Casagrandi and Gatto’s model, individuals migrate from a patch to a \textit{dispersers’ reservoir} and then migrate uniformly to all other patches. However, reaching the new patch, which required the individual to survive their time in the reservoir and reach an appropriate patch, only occurred with probability \(a\). While such an assumption had been used previously to study certain species \[98\], it removes any spatial structure the metapopulation might have as each patch is just as likely as any other patch to receive the migrating individuals. An analysis of the stationary distributions of their model showed that the while a stationary distribution corresponding to the extinction of the metapopulation always exists, it is possible for their model to exhibit other stationary distributions also. The existence of more than one stationary distribution was classified as the population persisting and they determined this is possible if

\[
\sum_{j=1}^{\infty} a \frac{D_j}{\mu_j + D_j} \rho_j > 1,
\]

where \(\rho_1 = 1\), \(\rho_2 = \nu_1/ (\mu_1 + D_1)\) and \(\rho_j = (\nu_1 \ldots \nu_j)/( (\mu_1 + D_1) \ldots (\mu_j + D_j))\) for \(j \geq 3\). Casagrandi and Gatto point out that this condition is equivalent to a persistence condition given by Chesson \[33\, page 105\], whereby the expected number of emigrants from a patch that is begun with one individual, and to which subsequent immigration is excluded, is greater than unity. While such a condition is informative and helpful in beginning to understand the contribution the dynamics within a patch might have, the inclusion of a dispersers’ pool causes this result to be less than ideal. The assumption
Chapter 1: Introduction

of the existence of a disperser’s pool limits the number of populations this result can be applied to but it is reasonable for some, perhaps limited number of, populations. However, spatial structure is an important aspect of a metapopulation as it almost exclusively determines if it is possible for one patch to colonise another patch. Therefore, spatial structure needs to be included in a model if more detailed results are required.

As Casagrandi and Gatto’s model demonstrates, when using a SMM, the information about the population increases but the tractability of the model decreases. Including an individual’s dynamics means the parameters are easily interpreted as per–capita birth, death and migration rates, rather than abstract parameters such as local extinction and colonisation rates. This means information deduced from the model about the population can be related back to the individual, providing much more insight into the population. However, obtaining such results is difficult as the complexity of the model is substantially increased. A simple example of this new level of complexity is the change in the state space from \( \{0, 1\}^J \) to \( \{0, 1, \ldots, N\}^J \), where \( N \) relates to the size of a patch and is often much larger than \( J \), possibly even infinity. To overcome the problem of tractability, often models are analysed via simulations which, for SMMs, can be quite computationally expensive. And as the models become more heterogeneous, understanding the dynamics over a wide parameter range becomes intractable, even via simulations.

As SMMs model the local population level, density dependent effects can occur within the patches just like population models, such as the logistic model. However, with subsistence spread over multiple patches, density dependence can be considered not only at the patch level, but also at the metapopulation level. This can be explained by considering a patch with only a small number of individuals, relative to the patch size. These individuals initially have ample resources (or space) available to them but as the population on that patch increases, the amount of resources available to each individual decreases; this is the normal density dependence at the patch level. However, when migration to new patches is allowed, these new patches become colonised, decreasing the rate of new colonisations; this is density dependence at the metapopulation level. Furthermore, positive density dependence, or the Allee effect, can also occur
Chapter 1: Introduction

at the local population and metapopulation levels. If each individual patch exhibits the Allee effect individually, then the metapopulation can also exhibit the Allee effect [37, section 3.5.1]. But just as density dependence had an added dimension in metapopulations, so does the Allee effect. It is known as the Allee-like effect [37, section 3.5.2] and is due to the metapopulation requiring a minimum number of occupied patches to colonise unoccupied patches. An example of this is the African wild dog, *Lycaon pictus*, which requires a minimum number of packs for the population size to increase [37, section 3.5.2].

The deterministic models presented previously differ from SPOMs not only by their lack of stochasticity but also by the approximation that population size is a continuous function of time. This leads to an important modelling question that must be answered before any scenario can be modelled. Should time and population size be modelled as if they were continuous or discrete? The logistic model assumes that time is continuous but also assumes that the population size is continuous. Whilst this might seem a reasonable approximation for large populations, it is definitely not reasonable for small populations. Furthermore, discrete time models might perform better in scenarios where populations evolve according to generations; that is, over set time intervals. On the other hand, continuous time models allow the population to evolve at any time, not restricting dynamics to set periods. All four options have their advantages and disadvantages with respect to their tractability and application to real metapopulations. This is why models with all four combinations exist; discrete time and discrete population size [50], discrete time and continuous population size [71], continuous time and discrete population size [111] and continuous time and continuous population size [78].

1.3 Environmental Influence

The logistic model presented by Verhulst in 1838 is the basis of many modern formulations of population models. There is one aspect of all systems, however, that his model and nearly all models presented in Sections 1.1 and 1.2 omit from their formulations: environmental variation.
The environment a species inhabits will always influence dynamics of the population at some level. For example, reproduction in many species such as crustaceans (*Daphnia*) [122], beetles (*Chrysomela aeneicollis*) [41], frogs (*Rana lessonae*) [93] and birds (*Passer domesticus*) [110] is affected by the temperature of the environment of the habitat. The survival and migration of the population can also be altered when the environment changes, as has been found in penguin (*Pygoscelis adeliae*) metapopulations [45]. The cause of the temperature change might be due to seasonal variation [114] or an unpredictable heat wave or cold snap, to list a few. Accounting for all these sources in a mathematical model would be superfluous. However, a model that accounts for how the population might react to changes in the environment should predict the long term outcome of the population more accurately than one that cannot. There are two types of changes in environment that must be accounted for: deterministic and stochastic.

Deterministic environmental influence comes in the form of seasonal variations. The frequency and severity of changes in the environment are largely, if not completely, known in advance. A species’ breeding season is one example of seasonal variation. During some period of the year, the reproduction rate of the species is increased and this is usually repeated at the same time every year, although some species are known to have a nine month period between breeding seasons [125].

The second type of environmental influence is stochastic changes in the environment. The breeding season introduced previously could, for some species, be considered a stochastic influence as the breeding season occurs during periods of rain which are inherently random [125]. Weather in general is an example of a stochastic influence. During times of heavy rain which cause flooding, various patches might be made inaccessible or, if you consider water dwelling species, flooding might cause patches to be connected. Furthermore, flooding will likely increase the mortality rate as well as cause species to migrate to higher ground [26].

Environmental influences, both deterministic and stochastic, need to be included in metapopulation models. Levins’ original model (1.1) with time dependent parameters is an example of a deterministic model accounting for a deterministically changing environment. Furthermore, some discrete time SPOMs have assumed that metapopu-
lations go through two phases which then repeat over time \cite{3, 43, 28} which is another method to account for a deterministically changing environment. One of these phases is extinction and the other phase is colonisation, giving rise to naming such models “CE” or “EC” models depending on the order of the phases in one time period. Such an assumption does not allow extinction in the colonisation phase and colonisation cannot occur during the extinction phase. While this assumption may seem unreasonable over any extended period of time, it has been suggested \cite{43} that some species exhibit very short periods of migration (often aligned with the dispersal of juveniles), which gives credence to a phase type setup. A continuous time SPOM that accounts for similar periodic changes between two (or more) phases does exist \cite{112}, however not much analysis has been done on the model.

Regarding stochastic influence, it was shown by J. R. Watson et al. \cite{131} that introducing stochastic connectivity between patches decreased the metapopulation growth rates by as much as 30% when near extinction, and up to 40% when near equilibrium. These large differences illustrate the need to include environmental stochasticity if the extinction of a species is be to accurately predicted. Smith & Wilkinson \cite{121} and Athreya & Karlin \cite{9} began the modelling of environmental stochasticity around 1970 by proposing models that utilised branching processes. More recently, authors have continued developing techniques to account for environmental stochasticity \cite{20, 52, 61, 131}, indicating that this is currently an area of interest.

\section{Motivation}

The aim of this thesis is to improve our understanding of how individual behaviour affects the survival of the metapopulation as a whole. The majority of metapopulation models to date have been patch occupancy models which are developed by first observing a segregated population and then attempting to mimic the observed behaviour of colonisation and extinction in a model. While useful in the analysis phase, the presence–absence assumption that patch occupancy models utilise leaves out the behaviour of individuals in the local populations and as such, patch occupancy models cannot be used to fully understand how individual behaviour affects the survival of a
Chapter 1: Introduction

metapopulation.

The primary reason individual based models are not as widely utilised as SPOMs is the high level of complexity such models contain, thereby making the analysis difficult. Results via simulation are often the norm for individual based models and are applied on a case-by-case basis. As such, the effect individuals have on the survival of the metapopulation has only been understood for a limited number of cases, motivating the need for research into understanding how individuals affect the metapopulation as a whole.

Therefore, this thesis aims to determine how the fecundity, mortality and movement of individuals affect the survival of a metapopulation. The important factors of stochastic dynamics of the individual and a changing environment will be included in the modelling phase to provide a more complete understanding of a metapopulation’s evolution toward extinction or survival.

1.5 Outline of the thesis

Chapter 2 provides an overview of the tools that are used in the later chapters. It begins with a definition for a stochastic process generally, followed by a deeper description of a particular stochastic process; a continuous time Markov chain. Several functional limit laws are provided for different families of continuous time Markov chains which all demonstrate the convergence of the Markov chain to another, possibly stochastic, process. Definitions and notations are given for a family of differential equations and random dynamical systems.

Chapter 3 begins with the definition of the model that is used to describe the evolution of a metapopulation with a static environment. One of the functional limit laws presented in Chapter 2 is then used to determine a system of differential equations which approximates the evolution of the metapopulation as an index parameter goes to infinity. The system of differential equations is analysed to determine the long term behaviour of the metapopulation, providing conditions for extinction, persistence and evidence for the Allee effect in a two patch metapopulation. A diffusion approximation is also outlined which quantifies the difference between the stochastic model and the
differential equation.

The model described in Section 3.2 is extended in Chapter 4 to include a dynamic environment. This chapter begins with a model that incorporates an environment that changes deterministically and presents a new functional limit law that shows asymptotically density dependent and time dependent Markov chains converge to a system of a differential equations. This is followed by an analysis of the system of differential equations is undertaken, which provides sufficient conditions for persistence and extinction of the metapopulation. In the second half of Chapter 4 a model that incorporates a stochastically changing environment is presented. Applying a functional limit law, the process is shown to be approximated by a piecewise-deterministic Markov process. The approximating piecewise-deterministic Markov process is analysed using the theory of random dynamical systems and a sufficient condition for extinction is determined.

The thesis concludes with a summary of results and a brief outline of extensions to the models and paths of future research in Chapter 5.
Chapter 2

Theory & Background

Population modelling can be accomplished via many different approaches. This chapter aims to present the specific tools used throughout this thesis. It includes definitions, notations and theorems concerning continuous time Markov processes, density dependent processes, functional limit laws for such processes, ordinary differential equations and random dynamical systems.

2.1 Introduction

A stochastic process is a collection of random variables that are classified by two sets. The stochastic process, \((X(t), t \in T, X(t) \in S)\), is indexed by a time variable \(t\) that can either be discrete, in which case \(T\) is usually of the form \(T = \mathbb{N}\) or \(\mathbb{Z}\) or it can be continuous, in which case \(T = [0, \infty)\) or \(\mathbb{R}\) usually. These two categories are classified as discrete time processes and continuous time processes, respectively. The second set that classifies a stochastic process is the state space \(S\) which, like the time variable, can be discrete or continuous. Research described in this thesis concerns continuous time and discrete state space processes.

One of the most important classes of stochastic processes used today is the class Markov processes, named after Andrei Andreevich Markov who gave the Markov process its foundations in the early 1900’s. Markov processes are characterised by the following property.
Definition 2.1. A stochastic process is said to satisfy the Markov property if, for all \( n \geq 0 \), set of times \( 0 \leq t_0 \leq \cdots \leq t_n \) where \( t_j \in T \) and set of states \( i_0, i_1, \ldots, i_n \) where \( i_j \in S \)

\[
\begin{align*}
\Pr(X(t_n) = i_n | X(t_0) = i_0, X(t_1) = i_1, \ldots, X(t_{n-1}) = i_{n-1}) &= \Pr(X(t_n) = i_n | X(t_{n-1}) = i_{n-1}). \tag{2.1}
\end{align*}
\]

In other words, the Markov property states that the (conditional) transition probabilities depend only on the current state of the process, forgetting all previous states. This leads to the idea of Markov processes being memoryless. For many applications in biology, such as population processes and epidemiology, this property appears to hold.

A further categorisation of stochastic processes is whether it is a time homogeneous or inhomogeneous Markov process. A Markov process is time homogeneous if for any \( t, s > 0 \), \( \Pr(X(s + t) = j | X(s) = i) = \Pr(X(t) = j | X(0) = i) \), and time inhomogeneous otherwise. The remainder of this chapter is set out as follows. Continuous time Markov processes and some of their properties are described in Section 2.2. Then in Section 2.3, some functional limit laws for density dependent processes are discussed. Properties of differential equations are provided in Section 2.4 and random dynamical system conclude the chapter in Section 2.5.

2.2 Continuous time Markov process

A continuous time Markov process (CTMP), \( X(t) \), is a stochastic process that satisfies the Markov property, takes values in a state space \( S \) and is indexed by \( t \in T = [0, \infty) \). The state space \( S \) can be infinite, however only processes with a finite state space are considered in this thesis. Therefore, from now on, \( S \) is assumed to be finite. The transition rates are given in a matrix \( Q_t = (q_t(i, j)) \), whose element \( q_t(i, j) \) is the transition rate from state \( i \in S \) to state \( j \in S \) at time \( t \). In the time-homogeneous case, these rates are constant through time. The diagonal elements of \( Q_t \) represent the total rate out of each state and are given by \( q_t(i, i) = -q_t(i) \) where \( q_t(i) \geq \sum_{j \neq i} q_t(i, j) \). In this thesis, \( Q_t \) will always be conservative, that is \( q_t(i) = \sum_{j \neq i} q_t(i, j) \) for all \( i \). Let \( P_{ij}(t) \) denote the probability that the Markov chain starting in state \( i \) at time zero is
in state $j$ at time $t$. The matrix $P(t) = (P_{ij}(t))$ is the transition matrix. For CTMPs with finite state spaces, the transition matrix satisfies

$$\frac{dP(t)}{dt} = Q_t P(t), \quad P(0) = I,$$  \hspace{1cm} (2.2a)

$$\frac{dP(t)}{dt} = P(t) Q_t, \quad P(0) = I.$$  \hspace{1cm} (2.2b)

Equations (2.2) are known as the Kolmogorov backward and forward equations, respectively. In the case of a time homogenous process, equations (2.2) have a unique solution of the form

$$P(t) = \exp(Qt),$$

where $\exp$ is the matrix exponential.

A realisation of a Markov process can be constructed in the following manner. The process begins in a state $i$ (chosen depending on the distribution of $X(0)$) and remains there for a random amount of time, called the holding time. The holding time in state $i$ at time $t$ has a cumulative distribution function given by

$$F_{t,i}(x) = 1 - \exp(-\int_t^{t+x} q_u(i)du).$$

If the process is time homogeneous ($q_t(i) = q(i)$), this expression simplifies and the holding time has an exponential distribution with parameter $q(i)$. Once the process has reached the end of the holding time, a transition occurs. Given a transition occurs at time $t$ and the process is in state $i$, the probability that a transition is made to state $j$ is given by $q_t(i,j)/q_t(i)$.

Markov processes can exhibit very different long term behaviour, depending on how the states are connected. A state $j$ is said to be accessible from state $i$ if $P_{ij}(t) > 0$ for some $t > 0$. A sufficient condition for state $j$ to be accessible from state $i$ is that $(Q_t)_{ij} > 0$ for all $t$. If state $j$ is accessible from state $i$ and state $i$ is accessible from state $j$, then it is said that state $i$ and state $j$ communicate. A set of states where all pairs communicate forms a communicating class. A state is called absorbing if the probability of leaving that state is 0. For such states, $q_t(i) = 0$. The entire state space can then be broken up into these two classes. Usually, in population modelling,
the extinction state would be an absorbing state and all other states would form a communicating class. For a more detailed look into the classification of states, see [25, Section 3.2].

A Markov population process is a particular class of Markov process that was introduced by Kingman [69]. One of the simplest examples of a Markov population process is a single population, \( n(t) \), which accounts for births and deaths occurring one at a time, hence its name of a “birth–death process” (BD process). A BD process has only two types of transitions such that the process will either increase or decrease by one to a new state. Suppose the size of the population is \( n \) at time \( t \). A birth occurs at a rate \( q_l(n, n + 1) = \lambda_l(t, n) \) and a death occurs at a rate \( q(n, n - 1) = \mu(t, n) \), where \( n \in S \). If more than one population or type of individual is to be modelled, then a BD process will not suffice. A multidimensional form of a BD process is required when modelling multiple populations and such models are often termed a “birth–death–migration process” (BDM process). Such a process is represented by a vector \( n(t) = (n_1(t), n_2(t), \ldots, n_J(t)) \), where \( n_i(t) \) is the number of individuals in population \( i \) (or of type \( i \)) and \( J \) is the total number of populations (or types). Transitions of a Markov population process can be due to births or deaths in a population and migrations of individuals between populations. The following definition is a modification of one due to Kingman [69] to cover the time inhomogeneous case.

**Definition 2.2.** Define, for any \( J \in \mathbb{Z}_+ \), \( S^J \) as the set of \( J \)-vectors \( n = (n_1, \ldots, n_J) \) where the \( n_i \) are non-negative integers. A Markov population process is a Markov process on a subset \( S \) of \( S^J \) with the transition rates

\[
q_l(n, n + e_i) = \lambda_l(t, n),
\]

\[
q_l(n, n - e_i) = \mu_l(t, n),
\]

\[
q_l(n, n - e_i + e_j) = \gamma_{ij}(t, n_i, n_j) \quad \text{for all } j \neq i,
\]

where \( e_i \) is the unit vector with a one in the \( i \)th position and \( \lambda_l(t, n), \mu_l(t, n) \) and \( \gamma_{ij}(t, n_i) \) are positive functions.

Due to the generality of Markov population processes, they can be applied to a number of situations. In this thesis, they are applied to populations of one type of
individual at multiple locations. However, as stated previously, such processes could
model populations involving multiple types such as predator prey interactions, epi-
demics and others.

2.2.1 Stationary Distributions

The long term behaviour of models is often a sought after characteristic. In the Markov
process context, one characterisation of the long term behaviour is the stationary dis-
tribution. The concept of a stationary distribution only applies to time homogeneous
processes; the Markov processes discussed in this section are assumed to be time ho-
mogeneous, that is $Q_t \equiv Q$. A stationary distribution, $\pi$, is found by solving

$$\pi P(t) = \pi, \text{ for all } t \geq 0.$$  \hspace{1cm} (2.3)

For finite state spaces, differentiating $\pi P(t)$ and applying (2.2a), yields $\pi Q = 0$ results,
while $\pi Q = 0$ implies (2.3), which is why $\pi$ is often referred to as an equilibrium
distribution. For infinite state spaces $\pi Q \leq 0$ still holds, with equality when $\pi$ is the
minimal solution to (2.3) or, more generally, the process is nonexplosive (see [106] for
more details) but the proof is more complex as interchanging the differentiation and
summation in the matrix multiplication cannot be done; see [92, Theorem 3.5.5] for
the complete proof. When the state space is a finite communicating class and $Q$ is
conservative, a solution to (2.3) exists and is unique if the additional criteria that $\pi$
is a distribution (requiring the elements in $\pi$ to sum to one) is met. In the case of
an infinite state space, (2.3) can have, at most, one solution but there may be none.
For example, a BD process where $\lambda_i = \lambda$ and $\mu_i = \mu$ for all $i \in \mathbb{N}_+$, (2.3) has a solution
$\pi_i = (\lambda/\mu)^i$, but it is only a distribution ($\sum_{i=1}^{\infty} \pi_i = 1$) when $\lambda < \mu$. Otherwise, no
solution will exist. More generally, for a finite state space BD process, where $\lambda(0) > 0$
and that $\mu(i) > 0$ if $\lambda(i-1) > 0$ for all $i$, a stationary distribution exists. Furthermore,
for a multidimensional Markov population process with $J$ populations, a closed form
solution for (2.3) does not always exist. However, under the assumption of reversibility,
that is $\pi$ satisfies the detailed balance equations

$$\pi(n)q(n, m) = \pi(m)q(m, n), \quad (2.4)$$

a closed form solution for the stationary distribution does exist and is given by

$$\pi(n) = B\prod_{i=1}^{J} \prod_{r_i=0}^{n_i-1} \frac{\lambda_i(r_i)}{\mu_i(r_i + 1)},$$

where $B$ is chosen so that $\sum_{n \in S} \pi(n) = 1$. Although there are no explicit restrictions on the functions $\gamma_{ij}$, the assumption of reversibility is quite strong, limiting the forms that the functions $\lambda$, $\mu$ and $\gamma$ can take. In particular, the assumption of reversibility excludes absorbing states. For more information about the limiting behaviour of Markov chains, see [25, Section 3.4].

The impossibility of an absorbing state in a process is a crucial exclusion. It is typical for CTMPs modelling population processes to have a finite state space consisting of one absorbing state and one communicating class. Such processes always have a stationary distribution; one with a point mass concentrated at the absorbing state. This is due to the process reaching the absorbing state in finite time with positive probability. However, the time to absorption can be very large. These processes often exhibit some form of stationarity away from the absorbing state over any reasonable time period, despite inevitably ending in the absorbing state (with probability one). This behaviour is caused by the existence of a \textit{quasi-stationary} distribution, $m$, which is a stationary distribution of the process, conditioned on not being absorbed \[126\]. Barbour and Pollett \[16, 17\] investigated Markov processes, $X(t)$, that exhibited such behaviour to determine simple approximations for quasi-stationary distributions, if such a distribution should exist. To calculate an approximation to the quasi-stationary distribution, they introduce a returned process, $X^\mu(t)$, that evolves exactly like $X(t)$ until the absorbing state is reached, at which point $X^\mu(t)$ is \textit{returned} to the communicating class according to the probability measure $\mu$. Interestingly, they show, under appropriate conditions, that the quasi-stationary distribution $m$ and the stationary distribution of the returned process (that is, as $t \to \infty$) , $\pi^\mu$, will be close in total variation \[16, Theorem 2.1\] and as such, $\pi^\mu$ offers a reasonable approximation for $m$. This is ex-
tended further to show that a similar result holds even for finite $t$ \cite[Theorem 2.2]{16}. Furthermore, the results do not depend on the choice of $\mu$. Barbour and Pollett \cite[Section 4]{17} later apply these results to density dependent Markov population processes (see Definition \ref{def:2.3}) and show that an appropriate Gaussian approximation for such processes, described later in Section \ref{subsec:2.3.3} is close in total variation to a particular returned process, for some applicable time range \cite[Theorem 4.1]{17}. For a detailed up to date bibliography of quasi-stationary distributions, see \cite{104}.

Therefore, as the quasi-stationary distribution provides greater insight into the long term behaviour of the process and the appropriate, easier to calculate, Gaussian approximation is always close to the quasi-stationary distribution, the Gaussian approximation will be used as a proxy for the quasi-stationary distribution. Furthermore, even if a quasi-stationary distribution were calculated, manipulating it analytically to derive useful results can still be difficult. This provides motivation for studying approximations to Markov processes.

\section{2.3 Functional Limit Laws}

In practice, complicated stochastic processes are often approximated by differential equations as the analysis of differential equations is simpler. The stochastic processes are analysed so as to extract the mean behaviour of the process, which is then used as the deterministic approximation for the overall trajectory. The reason such an approximation works well is due to the law of large numbers (LLN) principle, which has had such distinguished names as Jacob Bernoulli \cite{22}, Pafnuty Chebyshev \cite{32} and Andrey Kolmogorov \cite{70}, among many others, contribute to its formulation. The LLN states that the sum of independent and identically distributed random variables divided by the number of summands, $n$, converges to the expected value of the random variables as $n \to \infty$. Such a concept can be extended from random variables to random processes. One such method was proposed by Thomas Kurtz.
2.3.1 Density Dependent Processes

Kurtz’ method is relied on substantially in this thesis and relates to the approximating methods of ordinary differential equations. It was first formalised by Kurtz in 1969 \[75\] as a functional law of large numbers but applied directly to certain types of Markov processes, termed density dependent, a year later \[76\]. The formal definition for density dependence is now given.

**Definition 2.3** (Kurtz, \[76\]). A family of Markov processes \(\{n^{(N)}(t)\}\) indexed by \(N > 0\) (with a state space \(S_N \subset \mathbb{Z}^J\)) is said to be “density dependent” if there exists a continuous function, \(f : E \times \mathbb{Z}^J \mapsto \mathbb{R}\), where \(E \subseteq \mathbb{R}^J\) is an open set, such that the transition rates of \(n^{(N)}(t)\) are given by

\[
q\left(n^{(N)}, n^{(N)} + l \right) = Nf\left(\frac{n^{(N)}}{N}, l\right), \quad l \neq 0.
\]

The fundamental idea behind this characterisation is that the transition rates of the density process, \(X^{(N)}(t) := n^{(N)}(t)/N\), depend on \(n^{(N)}\) only through the density of the process. The stochastic logistic model in the context of a metapopulation will be used as an example throughout this section to demonstrate how to apply the theorems therein. The stochastic logistic model is a BD process on the state space \(S_N = \{0, \ldots, N\}\) with the birth rate \(\lambda(n) = (\lambda/N)n(N - n)\) and death rate \(\mu(n) = \mu n\). In the context of metapopulation modelling, \(n(t)\) is the number of currently occupied patches at time \(t\), \(N > 0\) is the total number of patches and \(\lambda, \mu\) are the strictly positive per-capita colonisation and extinction rates, respectively. The colonisation rate is derived by first noting there are \(n(N - n)\) pairs consisting of an occupied patch and unoccupied patch and the factor \(1/N\) accounts for the likelihood that the pair of patches occupy the same area, resembling a proximity requirement. The factor of \(1/N\) is important in terms of density dependence. With \(N\) as the index parameters, define \(E := (0, 1)\) and note that the results to follow will only hold when the approximation is in \(E\). The transition rates may be written in the form given in Definition 2.3 as

\[
q\left(n^{(N)}, n^{(N)} + l \right) = Nf\left(\frac{n^{(N)}}{N}, l\right).
\]
where
\[
f(x, l) = \begin{cases} 
\lambda x (1 - x) & \text{if } l = 1, \\
\mu x & \text{if } l = -1, \\
0 & \text{otherwise}.
\end{cases}
\]

Note that if the per-capita colonisation was not inversely proportional to \(N\), then the process would not be density dependent. In general, the index parameter \(N\) relates to the size of the system such as volume in a chemical reaction [105] or the number of patches in a metapopulation [8, 111]. However, this is not required to apply the theory and could be related to the initial size of the system.

The function \(f(\cdot, l)\) is the transition rate of a jump of size \(l\). Therefore, if an approximation to the trajectory of the density process \(n^{(N)}(t)/N\) were to be formulated, a starting point would be a differential equation where \(F(x) := \sum_l l f(x, l)\) is the rate of change. Hence, with this definition for density dependence, it may seem obvious to associate a deterministic model governed by the differential equation
\[
\frac{dx}{dt} = F(x(t)), \quad \text{with } x(0) = x_0,
\]
with the stochastic counterpart. Kurtz [76] determined the formal relationship, deriving a functional law of large numbers to describe the convergence of density dependent stochastic processes to deterministic ones. Kurtz’s result is presented below with the implicit assumption that the state space is finite. For processes with infinite state spaces, the theorems hold with extra conditions on \(f\).

**Theorem 2.4** (Theorem 3.1 of Kurtz [76]). Suppose that \(f(x, l)\) is bounded for each \(l\), and \(F\) is Lipschitz on \(E\). Then, for every trajectory \(x(s, x_0) \in E, 0 \leq s \leq t\) satisfying
\[
\frac{\partial}{\partial s} x(s, x_0) = F(x(s, x_0)), \quad \text{where } x(0, x_0) = x_0,
\]
\[
l_{\lim_{N \to \infty}} X^{(N)}(0) = x_0 \text{ implies for every } \delta > 0
\]
\[
\lim_{N \to \infty} P \left( \sup_{s \leq t} |X^{(N)}(s) - x(s, x_0)| > \delta \right) = 0.
\]
Theorem 2.4 states that a density dependent process will converge in probability on finite time intervals to a differential equation (2.5), provided the initial condition $X^{(N)}(0)$ converges to $x_0$.

### 2.3.2 Asymptotically Density Dependent Processes

If patches in the metapopulation could be colonised from sources originating outside the metapopulation (from a mainland, for example), the colonisation rate could be of the form $\lambda(n) = (\lambda n + \nu)(1 - n/N)$, where $\nu$ is a constant immigration rate. Such a process would not be density dependent according to Definition 2.3. To account for such processes, Philip K. Pollett [101] extended the results of Kurtz and defined asymptotically density dependent processes.

**Definition 2.5** (Pollett, [101]). A family of Markov processes $\{n^{(N)}(t)\}$ indexed by $N > 0$ (with a state space $S_N \subseteq \mathbb{Z}^J$) is said to be "asymptotically density dependent" when there exists a continuous function, $f^{(N)} : E \times \mathbb{Z}^J \to \mathbb{R}^J$, where $E \subseteq \mathbb{R}^J$ is an open set, such that the transition rates of $n^{(N)}(t)$ are given by

$$q\left(n^{(N)}, n^{(N)} + l\right) = N f^{(N)}\left(\frac{n^{(N)}}{N}, l\right), \quad l \neq 0,$$

and $F^{(N)}(x) := \sum_l l f^{(N)}(x, l)$ converges uniformly to $F(x)$ on $E$.

In Pollett’s definition, the transition rates of $X^{(N)}(t)$ can depend on the index parameter $N$ as long as $F^{(N)}(x)$ converges uniformly. Essentially, Definition 2.5 includes processes that are only density dependent (in terms of Definition 2.3) in the limit as $N \to \infty$. The stochastic logistic model with immigration is then classified as asymptotically density dependent and in fact, $F^{(N)}(x)$ converges to the same $F(x)$ as the model without immigration. This agrees with intuition as when the number of patches increases, the affect constant immigration has decreases and the model with immigration closely resembles that of one without immigration.

Pollett [101] extended Theorem 2.4 for asymptotically density dependent processes. It follows Theorem 2.4 except that the definition for the function $F(x)$ changes.
Chapter 2: Theory & Background

**Theorem 2.6** (Theorem 3.1 of Pollett [101]). Suppose that $f^{(N)}(x,l)$ is bounded for each $l$ and $N$, and $F$ is Lipschitz on $E$ and satisfies

$$
\lim_{N \to \infty} \sup_{x \in E} |F^{(N)}(x) - F(x)| = 0.
$$

Then, for every trajectory $x(s,x_0) \in E$, $0 \leq s \leq t$ satisfying

$$
\frac{\partial}{\partial s} x(s,x_0) = F(x(s,x_0)), \quad \text{where } x(0,x_0) = x_0,
$$

$\lim_{N \to \infty} X^{(N)}(0) = x_0$ implies for every $\delta > 0$

$$
\lim_{N \to \infty} \mathbb{P} \left( \sup_{s \leq t} |X^{(N)}(s) - x(s,x_0)| > \delta \right) = 0.
$$

Therefore, an asymptotically density dependent Markov process can be approximated by the solution to the same ODE as the equivalent density dependent process. This approximation becomes more accurate as $N \to \infty$. The convergence in probability of the Markov process to a deterministic one gives reason to performing an analysis on the long term behaviour of the differential equation, which is discussed in Section 2.4.

**2.3.3 Diffusion Approximation**

When $N$ is finite, the differential equation is merely an approximation for the overall trajectory of the stochastic process and, due to random jumps, the stochastic process fluctuates about the deterministic trajectory. This leads to using diffusion approximations. Such processes are closely related to the continuous time Markov process presented in Section 2.2 such that they must exhibit the Markov property but also have continuous sample paths almost surely. A diffusion process $Y(t)$ can be characterised by its infinitesimal parameters for the mean (or drift), $\mu(t,x)$ and variance, $\sigma^2(t,x)$ where

$$
\mu(t,x) := \lim_{s \to 0} \mathbb{E} \left[ \frac{Y(t+s) - Y(t)}{s} | Y(t) = x \right]
$$

and

$$
\sigma^2(t,x) := \lim_{s \to 0} \mathbb{E} \left[ \frac{(Y(t+s) - Y(t))^2}{s} | Y(t) = x \right].
$$
Brownian motion is an example of a diffusion process. The drift and variance parameters for Brownian motion are $\mu(t, x) = 0$ and $\sigma^2(t, x) = \gamma^2 t$, for some $\gamma \geq 0$, respectively.

Kurtz [77, Theorem 3.5] accounts for these fluctuations with a functional central limit theorem which Pollett [101, Theorem 3.2] extends to asymptotically density dependent processes. It is to be noted that Andrew Barbour remarked [12], without a formal theorem, about such a result 15 years prior to Pollett. Let $D^J[0,t]$ be the space of càdlàg functions (right continuous functions from $[0,\infty)$ to $\mathbb{R}^J$ with left limits everywhere) on $[0,t]$.

**Theorem 2.7** (Theorem 3.2 of [101]). Suppose that $F^{(N)}$ converges uniformly to $F$ and $F$ is Lipschitz continuous on $E$ and has uniformly continuous first partial derivatives. Suppose also that $G^{(N)}(x) = (g^{(N)}_{ij}(x))$, a $J \times J$ matrix with elements

$$g^{(N)}_{ij}(x) = \sum_l l_il_jf(x,l), \quad \text{where } l = (l_1, \ldots, l_J) \in \mathbb{Z}^J \text{ and } x \in E,$$

converges uniformly to $G$, where $G$ is bounded and uniformly continuous on $E$. Furthermore, suppose that $f^{(N)}(x,l)$ is bounded for each $l$ and $N$ and

$$\lim_{N \to \infty} \sup_{x \in E} \sqrt{N} |F^{(N)}(x) - F(x)| = 0.$$

Then, provided

$$\lim_{N \to \infty} \sqrt{N} (X^{(N)}(0) - x_0) = z,$$

the family of processes $\{Z^{(N)}(t)\}$, defined by

$$Z^{(N)}(s) = \sqrt{N} \left( X^{(N)}(s) - x(s, x_0) \right), \quad 0 \leq s \leq t$$

converges in distribution on the space $D^J[0,t]$ to a Gaussian diffusion $Z(t)$, with initial value $Z(0) = z$ and characteristic function $\psi = \psi(t, \theta)$ that satisfies

$$\frac{\partial \psi(t, \theta)}{\partial t} = -\frac{1}{2} \sum_{j,k} \partial_j g_{jk}(x(t,x_0)) \theta_k \psi(t, \theta) + \sum_{j,k} \theta_j \frac{\partial F_j(x(t,x_0))}{\partial x_k} \frac{\partial \psi(t, \theta)}{\partial \theta_k}.$$

Using (2.6), the mean, $\mu(t)$, and covariance matrix, $\Sigma(t)$, of $Z(t)$ may be expressed
by
\[ \mu(t) = M(t)z \quad \text{and} \quad \Sigma(t) = M(t) \left( \int_0^t M(u)^{-1} G(x(u, x_0)) (M(u)^{-1})^T du \right) M(t)^T, \]
where
\[ M(t) = \exp \left( \int_0^t H(u) du \right) \quad \text{and} \quad H(t) = \nabla F(x(t, x_0)). \]

While Theorem 2.7 provides a method to approximate the fluctuations about the deterministic trajectory, the expressions used to determine the mean and covariance matrix can often be troublesome. Barbour [13, Theorem 2] determined a special case of Theorem 3.5 of [77] (of which Theorem 2.7 is a generalisation) where the initial value of the deterministic system, \( x_0 \), is a fixed point of (2.5), that is, \( F(x_0) = 0 \). Under such conditions, the process \( Z^{(N)}(t) \) converges to a particular stochastic process called an Ornstein-Uhlenbeck (OU) process. An OU process is a particular type of diffusion process where \( \mu(t, x) = -\alpha x \) and \( \sigma(t, x) = \sigma^2 \), for some \( \alpha, \sigma \geq 0 \). Pollett provided the explicit formulation regarding asymptotically density dependent process in 2001 [103].

**Corollary 2.8** (Corollary 2.1 of [103]). If \( x^* \) satisfies \( F(x^*) = 0 \), then under conditions of Theorem 2.7, the family of processes \( \{Z^{(N)}(t)\} \), defined by
\[ Z^{(N)}(s) = \sqrt{N} \left( X^{(N)}(s) - x^* \right), \quad 0 \leq s \leq t \]
converges in distribution on the space \( D^J[0, t] \) to an OU process \( Z(t) \), with initial value \( Z(0) = z \) and local drift matrix \( H = \nabla F(x^*) \) and local covariance matrix \( G(x^*) \).

As in the case of Theorem 2.7 explicit expressions for \( \mu(t) \) and \( \Sigma(t) \) may be determined. Corollary 2.8 says that \( Z(t) \) follows a Gaussian diffusion, with mean \( \mu(t) = e^{Ht}z \) and covariance matrix
\[ \Sigma(t) = e^{-Ht} \left( \int_0^t e^{-Hs} G(x^*) e^{-HTs} ds \right) e^{-HTt}. \]

Then, for sufficiently large \( N \), the asymptotically density dependent process \( n^{(N)}(t) \) with an initial value of \( n^{(N)}(0) = \lfloor x_0 N \rfloor \) is approximately normally distributed with
mean
\[ \mathbb{E}(n^{(N)}(t)) \approx N \left( x^* + \exp(\lambda t)(x_0 - x^*) \right), \]
and covariance matrix
\[ \text{Cov}(n^{(N)}(t)) \approx N \Sigma(t). \]

2.3.4 Functional Limit Laws for inhomogeneous time

When a Markov process is time inhomogeneous, the results given in the previous sections do not hold. As such, Daniel Pagendam and Philip K. Pollett [96] developed results for cases where the \( Q \)-matrix is time dependent. Alongside his note about asymptotic density dependence, Barbour [12] also remarked that Kurtz’ results should hold in the time inhomogeneous case.

**Definition 2.9** (Pagendam & Pollett, [96]). A family of Markov processes \( \{n^{(N)}(t)\} \) indexed by \( N > 0 \) (with a state space \( S_N \subset \mathbb{Z}^J \)) is said to be “density dependent in time” when there exists a continuous function, \( f : [0, \infty) \times E \times \mathbb{Z}^J \mapsto \mathbb{R} \), where \( E \subset \mathbb{R}^J \) is an open set, such that the transition rates of \( n^{(N)}(t) \) are given by

\[ q_l \left( n^{(N)}(t), n^{(N)}(t) + l \right) = N f \left( t, \frac{n^{(N)}(t)}{N}, l \right), \quad l \neq 0. \]

This definition and Definition 2.3 only differ by the absence of \( t \) in Definition 2.3. Theorem 1 from [96] regarding the convergence of a density dependent in time process is also required. Similar to Section 2.3.1, define the function \( F(t, x) := \sum_l l f(t, x, l) \).

**Theorem 2.10** (Theorem 1 of Pagendam & Pollett [96]). Suppose that \( f(x, l) \) is bounded for each \( l \), \( F(t, \cdot) \) is Lipschitz on \( E \) and for each \( l \neq 0 \), \( f(t, x, l) \) is continuous in \( t \), uniformly in \( x \in E \). Then, for every trajectory \( x(s, x_0) \in E, 0 \leq s \leq t \) satisfying

\[ \frac{\partial}{\partial s} x(s, x_0) = F(s, x(s, x_0)), \quad \text{where } x(0, x_0) = x_0, \quad (2.7) \]

\[ \lim_{N \to \infty} X^{(N)}(0) = x_0 \text{ implies for every } \delta > 0 \]

\[ \lim_{N \to \infty} \mathbb{P} \left( \sup_{s \in [t]} \left| X^{(N)}(s) - x(s, x_0) \right| > \delta \right) = 0. \]
The density process \( X^{(N)}(t) \) converges in probability over finite time intervals to the process governed by \((2.7)\). Therefore, for large values of \( N \), the density process \( X^{(N)}(t) \) can be approximated by the solution to \((2.7)\). Pagendam and Pollett [96] also extended Kurtz’s diffusion approximation results but these will be omitted as they are not utilised.

### 2.3.5 Functional Limit Laws for different scalings

The limit laws presented in this section so far have relied on the processes being at least asymptotically density dependent in time (Definition 2.9). All the transitions of such a process need to possess the same density dependent scaling in relation to \( N \). However, this may not always be plausible as in some cases as \( N \) becomes large, some transition rates of \( X^{(N)}(t) \) do not change. Some chemical reactions such as simple crystallisation, enzyme kinetics and reversible isomerization are examples of processes where not all transition rates scale equally as \( N \) gets large [11]. In these cases, the methods of Kurtz, Pollett and Pagendam will fail and therefore other methods are required. One method was presented by Uwe Franz, Volkmar Liebscher and Stefan Zeiser [50] which allowed transition rates to scale unequally. Franz et al. have separated the transition rates into two groups; one group contains all the transition rates which have a form (defined later) similar to density dependence while the other group contains the remainder. Under some technical conditions, Franz et al. have shown that such Markov processes converge to piecewise deterministic Markov processes (PDMP) as \( N \to \infty \). The Markov process will be assumed to have \( K \) transitions, with rates of the form

\[
q\left(n^{(N)}, n^{(N)} + l_k^{(N)}\right) = N f_k^{(N)}\left(\frac{n^{(N)}}{N}\right), \quad k \in \{1, \ldots, K\}, \tag{2.8}
\]

where \( l_k^{(N)} \) is the size of the jump of the \( k \)th transition. Notice that the form of the rates in \((2.8)\) are identical to the ones defined in Definition 2.5, except for the dependence \( l_k^{(N)} \) can have on \( N \). The difference, however, is that \( F^{(N)} \) from Definition 2.5 does not converge in this case. This is because there exists a set, \( A \), such that \( \sum_{k \in A} l_k^{(N)} f_k^{(N)}(x) \) does not converge. However with the appropriate scaling, a limit result can still be obtained.
Chapter 2: Theory & Background

The results that follow relate to the density process $X^{(N)}(t) := n^{(N)}(t)/N$. Therefore, to simplify notation, it is assumed that the transition rates of the density process are of the form

$$ q \left( X^{(N)}, X^{(N)} + e^{(N)}_{k} \right) = r^{(N)}_{k} \left( X^{(N)} \right), \quad k \in \{1, \ldots, K \}. \quad (2.9) $$

To make this change, a simple transformation of $e^{(N)}_{k} = l^{(N)}_{k}/N$ and $r^{(N)}_{k}(x) = Nf^{(N)}_{k}(x)$ is applied. Under the assumption that the transition rates can be written in the form (2.9) and that $X^{(N)}(0) = x^{(N)}_{0}$, the trajectory of $X^{(N)}(t)$ may be written as

$$ X^{(N)}(t) = x^{(N)}_{0} + \sum_{k=1}^{K} e^{(N)}_{k} \Pi_{k} \left( \int_{0}^{t} r^{(N)}_{k} \left( X^{(N)}(s) \right) ds \right), \quad (2.10) $$

where $\Pi_{k}$ are independent Poisson processes with unit rate. Without loss of generality, the transitions are assumed to be ordered such that the first $K_{L}$ transitions do not scale with $N$ (termed large jumps) and that the remaining $K - K_{L}$ transitions do scale appropriately with $N$ (termed small jumps). Furthermore, the state space of $X^{(N)}$ is given by $S = M \times E$ where, again without loss of generality, $M \subset \mathbb{N}_{0}$ and $E \subset \mathbb{R}^{J}$, such that $S \subset \mathbb{R}^{J+1}$. Franz et al. impose the following conditions on the process.

**Condition 1** (PDMP conditions, Condition 4.1 of [50]). For the various parameters and functions below, the following statements are assumed to hold.

- (i) $\lim_{N \to \infty} x^{(N)}_{0} = x_{0}$.
- (ii) $e^{(N)}_{k} \to e_{k}$ as $N \to \infty$ for $k = 1, \ldots, K_{L}$.
- (iii) $N e^{(N)}_{k} \to e_{k}$ as $N \to \infty$ for $k = K_{L} + 1, \ldots, K$.
- (iv) $r^{(N)}_{k} \to r_{k}$ as $N \to \infty$ for $k = 1, \ldots, K_{L}$ uniformly on compacts.
- (v) $(1/N)r^{(N)}_{k} \to r_{k}$ as $N \to \infty$ for $k = K_{L} + 1, \ldots, K$ uniformly on compacts.
- (vi) $r_{k}$ is locally Lipschitz on $\mathbb{R}^{J+1}$ for all $k$.  

32
(vii) There is a locally bounded function $c : \mathbb{R}^{J+1} \to \mathbb{R}_+$ such that, for all $x \in S$, all $L, N \in \mathbb{N}$, and every sequence $(k_l)_{l \in \mathbb{N}} \subset \{K_L + 1, \ldots, K\}^\mathbb{N}$ with $x + \sum_{l=1}^L c^{(N)}_{k_l} \in S$,

$$
\frac{1}{N} f^{(N)}(x + \sum_{l=1}^L c^{(N)}_{k_l}) \leq c(x) \left( \frac{L}{N} + 1 \right).
$$

Condition [1] has components that are very similar to the conditions for Theorem 2.6. Parts (iii) and (v) together form conditions that are comparative to the condition in Definition 2.6 concerning the convergence of $F^{(N)}$ and describe the scaling of small jumps and their corresponding rates, respectively. It is parts (ii) and (iv) that allow the limit process to retain some stochasticity and describe the convergence of large jumps and their corresponding rates, respectively.

As Franz et al. [50] note, uniform convergence like the results of Kurtz [76] and Pollett [101] would be difficult to obtain due to the fact that the limit process still contains jumps. Instead, Franz et al. use the Skorokhod metric on the space $D^{J+1}[0, t]$. This metric is defined as

$$
d_S(x, y) = \inf_{\lambda \in \Lambda} \left( \max \left\{ \gamma(\lambda), \int_0^\infty \exp(-u) d(x, y, \lambda, u) du \right\} \right),
$$

where $\Lambda$ is the family of Lipschitz continuous functions $\lambda : [0, \infty) \to [0, \infty)$ which are strictly increasing and onto. Define

$$
\gamma(\lambda) = \sup_{0 \leq t < s} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right|,
$$

and for two càdlàg functions $x, y \in D$, $\lambda \in \Lambda$ and $u \geq 0$, set

$$
d(x, y, \lambda, u) = \sup_{t \geq 0} \frac{\|x(t \wedge u) - y(\lambda(t) \wedge u)\|}{1 + \|x(t \wedge u) - y(\lambda(t) \wedge u)\|}
$$

Convergence in Skorokhod topology is assured under Condition [1]. The following theorem presents this formally.

**Theorem 2.11** (Theorem 4.1 of [50]). Let $(X^{(N)}(t), t \geq 0)$, be a family of Markov jump processes, indexed by $N \geq 1$ and let $(X(t), t \geq 0)$ be a piecewise deterministic Markov
process given by

\[ X(t) = x_0 + \sum_{k=1}^{K_L} e_k \Pi_k \left( \int_0^t r_k(X(s))ds \right) + \int_0^t V(X(s))ds, \]

where \( V(x) = \sum_{k=K_L+1}^K e_k r_k(x) \) and which satisfies Condition \( \square \). Then, almost surely,

\[ X^{(N)}(\cdot) \to X(\cdot), \]

in the Skorokhod topology as \( N \to \infty \).

Theorem 2.11 allows the original process, \( X^{(N)}(t) \), to be approximated by a PDMP. Therefore, an analysis of the PDMP can help to understand the original process.

### 2.4 Differential Equations

As has been demonstrated in the previous sections, the stochastic component of some processes can have quite small effects on the trajectory of the process. To this end, it is then important to understand how processes without stochastic components are realised and analysed. Consider an ordinary differential equation of the form

\[ \frac{dx(t, x_0)}{dt} = f(t, x(t, x_0)), \quad x(0, x_0) = x_0. \] (2.11)

In this case, \( t \in \mathbb{R} \) or \( \mathbb{R}_+ \) and is identified with time. The function \( f : E \to \mathbb{R}^J \) is a vector function that is continuous in \( t \) and \( x \), and \( E \) is some subset of \( \mathbb{R}^J \), often termed the phase space. Just as stochastic processes were separated into time homogeneous and time inhomogeneous processes, such a classification also exists for differential equations. An ODE of the form (2.11) is said to be “autonomous” if the function \( f(t, x) \equiv f(x) \), and “nonautonomous” otherwise. Therefore, an autonomous differential equation is similar to a time homogeneous stochastic process. Section 3.4 includes an analysis of an autonomous system while a nonautonomous system is analysed in Section 4.2.3.

A solution to (2.11) only exists under certain conditions on \( x_0 \) and \( f \), the most important being that of Lipschitz. Under the assumption of \( f(t, x) \) being Lipschitz, a solution to (2.11) with an initial condition \( x(0, x_0) = x_0 \) exists and is unique. This is
commonly referred to as the Cauchy-Lipschitz Theorem or Picard-Lindelöf Theorem. The theorem presented below is for forward differential equations.

**Theorem 2.12** (Cauchy-Lipschitz Theorem). If \( f(t,x) \) is continuous in \( t \) on \([0,T]\) for some \( T \geq 0 \) and Lipschitz in \( x \) on \( E \subset \mathbb{R}^J \), the initial value problem (2.11) admits one and only one solution.

Therefore, as \( f \) being Lipschitz is a condition for Theorems 2.4, 2.6, 2.10 and 2.11 to hold, a deterministic solution to their corresponding ODE approximations always exists and is unique.

A further classification for ODEs is where \( f(t,x) = A(t)x \), where \( A(t) \) is a \( J \times J \) matrix and these ODEs are defined as *linear* ODEs and *nonlinear* otherwise. The nonlinear system of ODEs associated with population processes often have a special property. To define it precisely, new notation is required. For vectors \( a = (a_i) \) and \( b = (b_i) \), write \( a \leq b \) if \( a_i \leq b_i \) for all \( i \) and \( a < b \) if \( a \leq b \) and \( a_i < b_i \) for at least one \( i \). Finally, write \( a \ll b \) will be written if \( a_i < b_i \) for all \( i \). This also holds for the greater than relation, in that \( a \geq b \) if \( a_i \geq b_i \) for all \( i \), \( a > b \) if \( a \geq b \) and \( a_i > b_i \) for at least one \( i \) and \( a \gg b \) if \( a_i > b_i \) for all \( i \). Finally, the solution to the initial value problem (2.11) may also be written \( \phi_t(x_0) \).

**Definition 2.13.** A flow \( \phi_t(x) \) induced by \( f(t, x) \) is said to be partially ordered if \( x \leq y \) implies that \( \phi_t(x) \leq \phi_t(y) \) for all \( t \geq 0 \).

A partially ordered flow can also be referred to as an order preserving flow. Conditions exist to ensure a flow is partially ordered. The following provides one such condition.

**Theorem 2.14** (Lemma 2.1 of [120]). If \( E \in \mathbb{R}_+^J \) is open and convex and (2.11) is autonomous, then the flow \( \phi_t(x) \) will be partially ordered if the off-diagonal elements of

\[
\nabla_x f(x) := \left( \frac{\partial f_i}{\partial x_j} \right)_{ij},
\]

are nonnegative.
Note that a similar result holds for nonautonomous systems, however an extra condition on the concavity of $f$ is required. If the matrix (2.12) has nonnegative off-diagonal elements it is said to be a Metzler matrix or an $ML$ matrix. A Metzler matrix $M$ has many properties in common with some nonnegative matrix $T$ due to the relation $T = M + aI$, for some $a \geq 0$. These properties are listed in the following and are derived via a manipulation of the Perron-Frobenius Theorem applied to the matrix $T$.

**Theorem 2.15** (Theorem 2.6 of [117]). Suppose $M$ is a Metzler matrix. Then there exists an eigenvalue $r$ such that

(a) $r$ is real,

(b) with $r$ are associated strictly positive left and right eigenvectors,

(c) $r > 9\Re\lambda$ for any eigenvalue $\lambda$, $\lambda \not= r$, of $M$,

(d) $r$ is a simple root of the characteristic equation of $M$,

(e) $r \leq 0$ if and only if there exists $y \in \mathbb{R}_+ \setminus \{0\}$ such that $My \leq 0$ and $r < 0$ if and only if there is inequality in at least one position in $My \leq 0$.

The chapters that follow contain analyses deriving the long term behaviour of (2.11). The system (2.11) can exhibit various types of long term behaviour. A good reference for all types is Khalil [68, Chapter 1]. A simple characterisation of the long term behaviour of an ODE is through its stationary points. A stationary point $x^*$, often also called an equilibrium point, is a point in the phase space that does not change with time. As such, it satisfies $f(t,x^*) = 0$ for all $t$ and implies $\phi_t(x^*) = x^*$ for all $t$. Each stationary point for (2.11) is classified as Lyapunov stable (or simply stable) or unstable. Lyapunov stable points can then be further classified as asymptotic. These classifications are expressed in the following definition.

**Definition 2.16.** A stationary point $x^*$ of (2.11) is

- Lyapunov stable if, for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$\|x_0 - x^*\| < \delta \implies \|x(t,x_0) - x^*\| < \epsilon \text{ for all } t \geq 0;$$
• unstable if it is not stable;
• asymptotically stable if it is Lyapunov stable and $\delta$ can be chosen such that

$$
\|x_0 - x^*\| < \delta \implies \lim_{t \to \infty} x(t, x_0) = x^*.
$$

(2.13)

For a system defined on a set $E \subseteq \mathbb{R}^J$, a stationary point is globally asymptotically stable if it is asymptotically stable for all $x_0 \in E$. Otherwise, it is locally asymptotically stable. Theorems 7.1 and 7.3 of [127] relate the stability of a fixed point to the sign of the eigenvalue with the largest real part, $r$, of the Jacobian (2.12) evaluated at the fixed point. Under some conditions on the form of $f(x)$ in (2.11), which are discussed in Chapter 3, the fixed point will be asymptotically stable if $r < 0$ and unstable if $r > 0$.

2.5 Random Dynamical Systems

Introducing random dynamics into dynamical systems can be done a number of ways. One method is via a stochastic differential equation which adds a random noise term to the ODE. A random dynamical system (RDS) is a generalisation of such a process. An RDS is made up of two parts; a metric dynamical system (MDS) which models the random perturbations and a cocycle over this system.

**Definition 2.17** (Definition 1.1.1 of [34]). An MDS $\theta \equiv (\Omega, \mathcal{F}, \mathbb{P}, \{\theta_t, t \in \mathbb{T}\})$ with time set $\mathbb{T}$ is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a family of transformations $\{\theta_t, t \in \mathbb{T}\}$ such that

1. it is one-parameter group, that is

$$
\theta_0 = \text{id}, \quad \theta_t \circ \theta_s = \theta_{t+s} \quad \text{for all } t, s \in \mathbb{T};
$$

2. $(t, \omega) \mapsto \theta_t \omega$ is measurable;

3. $\theta_t \mathbb{P} = \mathbb{P}$ for all $t \in \mathbb{T}$, that is $\mathbb{P}(\theta_t B) = \mathbb{P}(B)$ for all $t \in \mathbb{T}$.

A set $B \in \mathcal{F}$ is called $\theta$-invariant if $\theta_t B = B$. A metric dynamical system $\theta$ is said to be ergodic under $\mathbb{P}$ if for any $\theta$-invariant set $B \in \mathcal{F}$ we have either $\mathbb{P}(B) = 0$ or
There are many examples of systems that are MDSs, such as a system of ODEs and certain Weiner Processes (see [34, page 10] for examples). Of particular interest are stationary processes. For a stationary random process $X(t)$ where $\Omega$ is the space of càdlàg functions and $\mathcal{F}$ is an appropriate $\sigma$-algebra, the shifts $X(t) \mapsto (\theta_s X)(t) = X(t + s)$ generate an MDS.

Let $E$ be a separable complete metric space, equipped with the Borel $\sigma$-algebra $\mathcal{B} = \mathcal{B}(E)$ generated by open sets of $E$. The definition of an RDS is given below.

**Definition 2.18** (Definition 1.2.1 of [34]). An RDS with time set $\mathbb{T}$ and state space $E$ is a pair $(\theta, \varphi)$ consisting of an MDS $\theta \equiv (\Omega, \mathcal{F}, \mathbb{P}, \{\theta_t, t \in \mathbb{T}\})$ and a cocycle $\varphi$ over $\theta$ of continuous mappings of $E$ with time set $\mathbb{T}_+$, that is a measurable mapping

$$
\varphi : \mathbb{T}_+ \times \Omega \times E \to E, \quad (t, \omega, x) \mapsto \varphi(t, \omega, x),
$$

such that

(i) the mapping $x \mapsto \varphi(t, \omega, x) \equiv \varphi(t, \omega)x$ is continuous for every $t \geq 0$ and $\omega \in \Omega$,

(ii) the mappings $\varphi(t, \omega) : = \varphi(t, \omega, \cdot)$ satisfy the cocycle property:

$$
\varphi(0, \omega) = \text{id}, \quad \varphi(t + s, \omega) = \varphi(t, \theta_s \omega) \circ \varphi(s, \omega)
$$

for all $t, s \in \mathbb{T}_+$ and $\omega \in \Omega$.

RDS enable the modelling of piecewise deterministic processes with random changes. Under some technical conditions, the limiting process of Theorem 2.11 belongs to this class of process. Therefore it is useful to have results describing the long term behaviour of an RDS. Cooperative random differential equations is one particular class of RDS that is amenable to analysis. The following definition from Igor Chueshov [34, Chapter 5] is used.

**Definition 2.19.** The system of equations

$$
\dot{x}(t) = f(\theta_t \omega, x(t)),
$$

38
is a cooperative random differential equation (CRDE) if \( f : \Omega \times \mathbb{R}^J_+ \to \mathbb{R}^J \) is a measurable function such that \( f(\omega, \cdot) \) possesses the following properties for all \( \omega \in \Omega \):

(I) \( f(\omega, \cdot) \) is continuously differentiable and \( f_i(\omega, \cdot) \) and \( \partial f_i(\omega, \cdot)/\partial x_j \), \( i, j = 1, \ldots, J \) are bounded on compact sets \( K \subset \mathbb{R}^J_+ \) by \( C_K(\omega) \) such that \( t \mapsto C_K(\theta_t \omega) \) is locally integrable;

(II) there exist random variables \( C_1 \) and \( C_2 \) such that \( t \mapsto C_j(\theta_t \omega) \) is locally integrable and \( \langle x, f(\omega, x) \rangle \leq C_1(\omega)|x|^2 + C_2(\omega) \), where \( \langle \cdot, \cdot \rangle \) is the standard inner product in \( \mathbb{R}^J \) and \( |x|^2 = \langle x, x \rangle \);

(III) \( f(\omega, \cdot) \) is weakly positive, that is

\[ f_i(\omega, x) \geq 0, \quad \text{for all} \quad x \in \Gamma_i, \omega \in \Omega, i = 1, \ldots, J, \]

where \( \Gamma_i = \{ x = (x_1, \ldots, x_J) \in \mathbb{R}^J : x_i = 0 \} \);

(IV) \( f(\omega, \cdot) \) is cooperative, that is

\[ f_i(\omega, x) \leq f_i(\omega, y), \quad i = 1, \ldots, J, \omega \in \Omega, \]

for all \( x, y \in \mathbb{R}^J_+ \) such that \( x_i = y_i \) and \( x_i \leq y_i \) for \( j \neq i \).

Proposition 5.2.1 of [34] shows that a CRDE generates an RDS \((\theta, \varphi)\) in \( \mathbb{R}^J_+ \) with the cocycle defined by

\[ \varphi(t, \omega)x_0 \equiv x(t) = x_0 + \int_0^t f(\theta_u \omega, x(u))du. \]

Furthermore, according to Theorem 5.2.1 of [34], if the matrix

\[ \nabla_x f(\omega, x) := \left( \frac{\partial f_i(\omega, x)}{\partial x_j} \right)_{i,j} \]

is irreducible for all \( x \in \text{int}\mathbb{R}^J_+ \) and \( \omega \in \Omega \), that is for every pair \( i, j \), there is an integer \( m \) such that \( (\nabla_x f(\omega, x)^m)_{ij} > 0 \), then the CRDE will generate a partially ordered RDS
$$(\theta, \varphi)$$ in $\mathbb{R}^d_+$; that is

$$\varphi(t, \omega)x \ll \varphi(t, \omega)y$$ if $$0 \ll x < y$$ for all $$\omega \in \Omega$$.

Various results regarding the long term behaviour of a CRDE exist but most limit the type of systems that can be modelled (see [34, Chapter 5]). A particularly useful class of CRDE is one where $f$ is $s$-concave.

**Definition 2.20.** A function $f(\omega, x)$ will be defined as $s$-concave if

$$\nabla_x f(\omega, x) < \nabla_x f(\omega, y)$$ for $$0 \ll y \ll x$$, and $\omega \in \Omega$.

As in Section 2.4, a linear RDS is defined as one where $\varphi(t, \omega)x = \Phi(t, \omega)x$, where $\Phi(t, \omega)$ is a cocycle over $\theta$ consisting of bounded linear operators of $E$. An ODE can have fixed points, or equilibrium points, that are invariant in time. The equivalent for an RDS will be referred to also as an equilibrium and is defined in the following, from Definition 1.7.1 of [34].

**Definition 2.21.** A random variable $u : \Omega \mapsto \mathbb{R}^d_+$ is said to be an equilibrium of the RDS $(\theta, \varphi)$ if it is invariant under $\varphi$, that is, if

$$\varphi(t, \omega)u(\omega) = u(\theta_t \omega)$$ for all $$t \geq 0$$ and $$\omega \in \Omega$$.

Despite the similar terminology, the concept of an equilibrium for an RDS varies significantly from an equilibrium for an ODE. For example, consider the differential equation $\dot{y}(t) = a - Ay$, where $a > 0$ and $A > 0$. It is straightforward to see that $u = a/A$ is an equilibrium, by setting $y = u$ and verifying $\dot{y}(t) = 0$. However, now consider the random differential equation $\dot{x}(t) = a - A(\theta_t \omega)x$, where $a > 0$ and $A(\omega)$ is a random variable with finite strictly positive mean. The value $u(\omega) = a/A(\theta_t \omega)$ is now not an equilibrium for $x(t)$ because, while $A$ can allow jumps, $u$ must transition smoothly. As such, the equilibrium for such a system is the more complex solution $u(\omega) = a \int_{-\infty}^{0} e^{-\int_{s}^{0} A(\theta_r \omega) dr} ds$, which one can use to verify that $\dot{x}(t) = 0$ with $x = u(\omega)$.

For an ODE, stability is determined by the eigenvalue of the Jacobian. For an RDS, stability is determined by the top Lyapunov exponent.
Definition 2.22. The top Lyapunov exponent for a linear RDS \((\theta, \varphi)\) in a separable Banach space \(E\) is the minimal real number \(\lambda\) with the following property: there exists a \(\theta\)-invariant set \(\Omega^* \subset \Omega\) of full measure such that

\[
\|\Phi(t, \omega)x\| \leq R_\epsilon(\omega) \exp((\lambda + \epsilon)t) \|x\|, \quad \omega \in \Omega^*, \quad t \geq 0,
\]

for every \(\epsilon > 0\) and all \(x \in E\), where the invariant set for which

\[
\lim_{t \to \infty} \frac{1}{t} \log R_\epsilon(\theta(t)\omega) = 0
\]

has full measure.

We can now determine the long term outcome of a CRDE. Within this thesis, and more generally also, this will involve calculating the quantity \(\lim_{t \to \infty} \varphi(t, \theta^{-t}\omega)x\). This is due to the fact that

\[
\mathbb{P}(\omega: \varphi(t, \omega)x \in D) = \mathbb{P}(\omega: \varphi(t, \theta^{-t}\omega)x \in D),
\]

for all \(x \in E\) and \(D \in B(E)\), where \(B(E)\) is the Borel-\(\sigma\) algebra of \(E\), due to \(\{\theta_t\}\) being measure preserving. As such,

\[
\lim_{t \to \infty} \mathbb{P}(\omega: \varphi(t, \omega)x \in D) = \lim_{t \to \infty} \mathbb{P}(\omega: \varphi(t, \theta^{-t}\omega)x \in D),
\]

if such a limit exists. Hence, the quantity \(\lim_{t \to \infty} \varphi(t, \theta^{-t}\omega)x\) allows one to determine the convergence of \(\varphi(t, \omega)x\) in probability, thus demonstrating its usefulness.

Under the assumption of \(s\)-concavity and irreducibility of the Jacobian, the RDS generated by a CRDE can have one of three outcomes. This result is outlined in Theorem 5.5.3 in [34].

Theorem 2.23 (Theorem 5.5.3 of [34]). Assume that a CRDE possesses an \(f(\omega,x)\) that is \(s\)-concave and a Jacobian \(\nabla_x f(\omega,x)\) that is irreducible for all \(x \in \text{int}\mathbb{R}^d_+\) and \(\omega \in \Omega\). If \(f(\omega,0) \in \mathbb{R}^d_+\setminus\{0\}\) for all \(\omega \in \Omega\), then either

(a) the orbit \(\gamma_0\) emanating from \(v\) is unbounded for all \(v(\omega) \geq 0\), or
(b) there exists a unique equilibrium $u > 0$ such for every $v(\omega)$ possessing the property $0 \leq v(\omega) \leq \alpha \cdot u(\omega)$ with some $\alpha > 0$ the orbit emanating from $v$ converges to $u$, that is

$$\lim_{t \to \infty} \varphi(t, \theta_{\omega}^{-t})v(\theta_{\omega}^{-t}) = u(\omega) \quad \text{for all } \omega \in \Omega^*, \quad (2.14)$$

where $\Omega^*$ is a $\theta$-invariant set of full measure. Otherwise, if $f(\omega, 0) \equiv 0$, $\theta$ is an ergodic metric dynamical system and the top Lyapunov exponent of the linear RDS generated by

$$\dot{y} = \nabla_x f(\theta_{\omega} t, 0)y$$

is less than zero, then we have

$$\lim_{t \to \infty} \phi(t, \theta_{\omega}^{-t})x = 0 \quad \text{for all } x \in \mathbb{R}_+^J \quad (2.15)$$

on a $\theta$-invariant set of full measure.

Theorem 2.23 shows that when $0$ is not a stationary point of the RDS, the RDS (a) has unbounded paths or (b) possess one or more nonzero equilibrium points. And when $0$ is a stationary point of the RDS, then if the top Lyapunov exponent is less than the zero, the RDS converges to $0$. However, situations Theorem 2.23 fails to cover are when $0$ is a stationary point for only some $\omega$ and, secondly, when $0$ is a stationary point for all $\omega \in \Omega$ but the top Lyapunov exponent is greater than zero.
A Spatially Structured Metapopulation Model accounting for within patch dynamics

The habitats of many species are becoming more fragmented, making metapopulation modelling more important than ever. Within this chapter, a model is examined that utilises dynamics that are easily interpreted as birth, death and migration rates of a species rather than the abstract dynamics of colonisation and extinction that metapopulations generally use. An analysis of this metapopulation model is undertaken via a differential equation approximation and conditions for persistence and extinction of the metapopulation are determined. Furthermore, an Allee effect is discovered in a two patch metapopulation whereby the potential survival of the metapopulation is dependent on the initial size of the population.

3.1 Introduction

The field of metapopulation ecology concerns the study of populations with a specific spatial structure where the population is separated into geographically distinct patches or islands. There has been a high level of interest in the field since the late 60s [78, 58], and this has continued to the present [108, 56, 60, 28] and references therein. Of significant concern to ecologists is the survival of the population and under what conditions the population might become extinct. Mathematical models have proved useful in addressing these questions.
As stated in Section 1.2, many metapopulation models employ the presence–absence assumption, that is, they simply record whether or not each patch is occupied. This assumption is employed in the two most widely used metapopulation models: Levins’ model [78] and Hanski’s incidence function model [56]. Hanski’s model has proven extremely successful in incorporating landscape structure and quality into the metapopulation dynamics. More generally, the presence–absence assumption has simplified modelling, data collection and analysis for a number of metapopulations [124, 86, 87, 89, 90, 95, 94, 111]. However, this assumption is not always adequate, for example in stock dynamics where more detail is required [74].

An alternative approach is to use a structured metapopulation model (SMM) that models the births, deaths and migration of individuals directly, and the number of individuals present on each patch is recorded. The parameters of SMMs are easily interpreted as per–capita birth, death and migration rates, rather than abstract parameters such as patch level extinction and colonisation rates. Furthermore, SMMs give far more detail about the state of the metapopulation than is possible under the presence–absence assumption. Unfortunately, many SMMs [7, 18, 19, 51, 54] impose a number of unrealistic assumptions on the metapopulation; they fail to account for the spatial configuration of patches and assume that migration patterns are homogeneous across all patches.

We study a metapopulation model that is structured in respect of both spatial configuration and within patch dynamics. This model has the form of a Markov population process defined in Definition 2.2. Previous analyses of this class of models have focused on determining expressions for moments and stationary distributions [108]. However, the restrictions that these analyses require are not natural in the present context since, under the assumption of no immigration from an external source, this model has an absorbing state corresponding to extinction. In this case, the stationary distribution would necessarily assign all its probability mass to the extinction state, and thus would not provide useful information about any quasi-stationary regime (being a common feature of metapopulation models [102]). Instead, this model is analysed by determining a simpler approximating differential equation based on the work of Kurtz [76] and Pollett [101].
Using the differential equation, we are able to determine conditions under which the metapopulation will go extinct quickly or persist for an extended period of time. Also, more complex dynamics such as the presence of an Allee effect for some range of parameters are identified. An Allee effect refers to populations exhibiting an increasing per capita growth rate at low population density levels. When the per-capita growth rate is initially negative, a critical threshold emerges below which the population goes extinct. In populations displaying an Allee effect, conservation strategies need to be adapted to account for this, particularly if a critical threshold is present [37, section 5.1.4].

The remainder of this chapter is organised as follows. In Section 3.2, the model is detailed. The differential equation approximation is described in Section 3.3. In Section 3.4, an analysis is provided for the long-term behaviour of the approximating deterministic model, conditions for extinction or persistence are derived, and the possibility of an Allee effect is demonstrated. Some examples are given to illustrate our results. The results are summarised and discussed in Section 3.6.

3.2 Model

The model under consideration is an example of Kingman’s Markov population process (Definition 2.2). In the present context, \( J \) is the number of patches in the metapopulation and \( n_i(t) \) is the number of individuals occupying patch \( i \) at time \( t \). The process \( (n(t), t \geq 0) \) describing the state of the metapopulation takes values in \( S_N = \{0, \ldots, N_1\} \times \cdots \times \{0, \ldots, N_J\} \) and has nonzero time homogeneous transition rates

\[
\lambda_i(n_i) = n_i b_i \left( \frac{n_i}{N_i} \right) + \nu_i (N_i - n_i), \quad (3.1a)
\]

\[
\mu_i(n_i) = \phi_i(n_i) \lambda_{i0} + d_i n_i, \quad (3.1b)
\]

\[
\gamma_{ij}(n_i, n_j) = \phi_i(n_i) \lambda_{ij} \frac{N_j - n_j}{N_j} \quad \text{for all } j \neq i, \quad (3.1c)
\]

where \( \phi_i(0) = 0, \phi_i(n) > 0 \) for \( n \geq 1 \) and \( b_i : [0, 1] \mapsto \mathbb{R}_+ \) such that \( b_i(x) = 0 \) for \( x = 1 \). These rates correspond to: an increase on patch \( i \) due to a birth or immigration from
outside the system (3.1a), a decrease on patch \(i\) due to a death or removal from the system (3.1b) and a migration from patch \(i\) to patch \(j\) (3.1c). Figure 3.1 illustrates these transitions. The parameters \(\nu_i\), \(d_i\), \(\lambda_{ij}\) and \(N_i\) are the immigration rate, per-capita death rate, proportion of individuals migrating from patch \(i\) to patch \(j\) (or out of the system if \(j = 0\)) and the population ceiling for patch \(i\), respectively. The birth rate function \(b_i(\cdot)\) determines the per-capita birth rate given how densely populated patch \(i\) is. The function \(\phi_i(\cdot)\), henceforth referred to as the migration function, is the rate at which individuals leave patch \(i\). Therefore, the product \(\phi_i(n_i)\lambda_{ij}\) is the rate individuals migrate from patch \(i\) to patch \(j\) (or out of the system if \(j = 0\)). The migration rate is attenuated by the proportion of space available on the destination patch which is a proxy for the probability an individual will arrive at an available space on the new patch. This assumes the migrating individuals simply cannot migrate when the destination patch is too full and hence remain on the original patch. A modelling question arises as to what happens to individuals upon arrival at an overcrowded patch. Modelling migration with rates (3.1) assumes the individuals either are aware of the current capacity of other patches before migration and do not migrate or that, upon arrival at an overcrowded patch, they return to their original patch. While it might seem impossible for individuals to be aware of the current capacity of other patches, a delayed knowledge of the system can be expected to filter through the population from individuals moving around the patches. This is assumed for ideal free distributions, where individuals are assumed to distribute themselves to patches with higher resources. Haché et al. [55] have found evidence of such migration in the Ovenbird (Seiurus aurocapilla), while a particular ground beetle (Pterostichus vernalis) has also been found to exhibit similar behaviour [62]. Therefore, it can be assumed that individuals are aware of the remaining capacity on other patches, which is one way to interpret (3.1c). It this assumption is invalid, another interpretation of (3.1c), as said above, is that individuals return to their original patch, if they arrive at an overcrowded patch. The alternative to modelling migration this way is to assume individuals, upon arrival at an overcrowded patch, either die or migrate to another patch. While neither of these scenarios will be explored in this thesis, if migrating individuals die upon arrival at an overcrowded patch, a term of \(\sum_j \phi_i(n_i)\lambda_{ij} n_j / N_j\) should be added to rate (3.1b).
Figure 3.1: Illustration of the dynamics for patch $i$ and migration to and from patch $j$.

If, instead, individuals choose to migrate to another patch, the resulting model becomes significantly more complex, as the next patch must also have space for the individual and the problem continues. The only patch with certain space is the originating patch and the assumption that individuals return to this patch is already accounted for in rate (3.1c).

Note that the models of Renshaw [108] and Arrigoni [7] have a number of features in common with this model. The main difference with Renshaw’s model is in the linearity of the birth and migration rates. That linearity excludes the possibility of a carrying capacity at each patch. Arrigoni’s model included catastrophes, that is, the possibility of the instantaneous death of all individuals on a given patch. However, it assumed that the birth, death and migration rates were the same for all patches and, as in Renshaw’s model, it could not incorporate a carrying capacity at each patch.

### 3.3 Differential Equation Approximation

We now apply Theorem 2.6 which facilitates an approximation of the path of the process by the solution to a system of differential equations. To do this we first need to establish that the model is density dependent in the sense of Definition 2.3 or at least asymptotically density dependent by Definition 2.5.

Define the population ceiling as the sum of all patch ceilings $N := \sum_j N_j$. The population density at patch $i$ is the number in patch $i$ measured relative to $N$ and is given by $X_i^{(N)}(t) := n_i(t)/N$. We are interested in the convergence of the density process $X_N := (X_1^{(N)}, \ldots, X_J^{(N)})$ as $N \to \infty$. Define the relative ceiling for patch $i$ as
\[ M_i^{(N)} := N_i / N \text{ and assume that } M_i^{(N)} \to M_i > 0 \text{ as } N \to \infty. \]

The density process \( X_N \) is a Markov process on the state space \( E_N := S_N / N \).

Suppose that the functions \( \hat{\phi}_i^{(N)} : [0, M_i^{(N)}] \to \mathbb{R}_+ \) satisfy

\[
\hat{\phi}_i^{(N)} \left( \frac{n}{N} \right) = \frac{\phi_i(n)}{N},
\]

for all \( n \geq 1 \) and \( N \geq 1 \). Then, the rates (3.1) can be written as

\[
q(n, n + l) = N f_N \left( \frac{n}{N}, l \right),
\]

where

\[
f_N(x, l) = \begin{cases} 
  x; b_i \left( \frac{x_i}{M_i^{(N)}} \right) + \nu_i (M_i^{(N)} - x_i) & \text{if } l = e_i, \\
  \hat{\phi}_i^{(N)}(x_i) \lambda_{i0} + d_i x_i & \text{if } l = -e_i, \\
  \hat{\phi}_i^{(N)}(x_i) \lambda_{ij} \left( 1 - \frac{x_i}{M_i^{(N)}} \right) & \text{if } l = -e_i + e_j, \\
  0 & \text{otherwise.}
\end{cases}
\]

Let \( F_i^{(N)}(x) := \sum_l l f_N(x, l) \) and observe that

\[
F_i^{(N)}(x) = M_i^{(N)} \nu_i + \left( b_i \left( \frac{x_i}{M_i^{(N)}} \right) - d_i - \nu_i \right) x_i \\
+ \sum_{j \neq i} \hat{\phi}_j^{(N)}(x_j) \lambda_{ji} \left( 1 - \frac{x_i}{M_i^{(N)}} \right) - \hat{\phi}_i^{(N)}(x_i) \sum_{j \neq i} \lambda_{ij} \left( 1 - \frac{x_j}{M_j^{(N)}} \right).
\]

Define \( E := [0, M_1] \times \cdots \times [0, M_J] \). Assume there exists bounded Lipschitz continuous functions \( \hat{\phi}_i : [0, M_i] \to \mathbb{R}_+ \) satisfying

\[
\lim_{N \to \infty} \sup_{x \in [0, M_i]} \left| \hat{\phi}_i^{(N)}(x) - \hat{\phi}_i(x) \right| = 0, \quad \text{for all } i,
\]

and also

\[
\lim_{N \to \infty} \sup_{x \in [0, M_i]} \left| b_i \left( \frac{x_i}{M_i^{(N)}} \right) - b_i \left( \frac{x_i}{M_i} \right) \right| = 0.
\]
We may then conclude that $F^{(N)}(x) \to F(x)$ as $N \to \infty$, uniformly on $E$, where

$$F_i(x) = M_i \nu_i + \left( b_i \left( \frac{x_i}{M_i} \right) - d_i - \nu_i \right) x_i + \sum_{j \neq i} \phi_j(x_j) \lambda_{ji} \left( 1 - \frac{x_i}{M_i} \right) - \phi_i(x_i) \left( \lambda_{i0} + \sum_{j \neq i} \lambda_{ij} \left( 1 - \frac{x_j}{M_j} \right) \right),$$

for $i = 1, \ldots, J$. Therefore, the family of processes indexed by the population ceiling $N$ is asymptotically density dependent according to Definition 2.5. Next we apply Theorem 2.6, the analogue of Theorem 2.4 for asymptotically density dependent families of processes. The conditions of this theorem are fulfilled as $f_N(x, l)$ is bounded on $E$ for all $N$ and $l$ and is nonzero for only finitely many $l$. Recall that $\lambda_{ij}$ is the proportion of individuals emanating from patch $i$ who are destined for patch $j$. Thus,

$$\sum_{j \neq i} \lambda_{ij} + \lambda_{i0} = 1,$$

and so we may rewrite $F(x)$ as

$$F_i(x) = \nu_i M_i + \left( b_i \left( \frac{x_i}{M_i} \right) - d_i - \nu_i \right) x_i - \phi_i(x_i) + \sum_{j \neq i} \phi_j(x_j) \lambda_{ji} \left( 1 - \frac{x_i}{M_i} \right) - \phi_j(x_j) \lambda_{ij} \left( 1 - \frac{x_j}{M_j} \right),$$

for $i = 1, \ldots, J$. It can be seen that $F$ is Lipschitz continuous on $E$. Hence, the conditions of Theorem 3.1 of [101] are satisfied, and we conclude that the density process converges in probability over finite time intervals to the solution $x(t, x_0)$ of the deterministic model

$$\frac{dx(t, x_0)}{dt} = F(x(t, x_0)), \quad x(0, x_0) = x_0,$$

as $N$ increases, where the elements of $F$ are defined in (3.5). More precisely, we have the following result.

**Theorem 3.1.** Let $x(t, x_0)$ be the solution to (3.6). Suppose there exists bounded Lipschitz continuous functions $\hat{\phi}_i : [0, M_i] \to \mathbb{R}_+$ satisfying (3.2) and assume also that the functions $b_i$ satisfy (3.3). If $X_N(0) \to x_0 \in E \setminus \partial E$ as $N \to \infty$ and $x(s, x_0) \in E \setminus \partial E$
Chapter 3: A Spatially Structured Metapopulation Model

for $0 \leq s \leq t$. Then, for every $t > 0$ and $\delta > 0$,

$$\lim_{N \to \infty} \Pr\left(\sup_{s \leq t} |X_N(s) - x(s, x_0)| > \delta \right) = 0.$$

Although the above result holds only in the limit as $N \to \infty$, explicit bounds on
\Pr\left(\sup_{s \leq t} |X_N(s) - x(s, x_0)| > \delta \right) for finite $N$ could be determined using Theorems 4.1 and 4.2 of [42]. Furthermore, the fluctuations of the stochastic process $X_N$ about the deterministic trajectory $x(\cdot, x_0)$ can be scaled as $N \to \infty$ to yield a Gaussian diffusion in the limit, which is determined in Section 3.5. These results go some way to justifying the use of the deterministic model (3.6) to approximate the behaviour of our metapopulation model when the population ceiling $N$ is large. Note that, for a given population ceiling $N$, the deterministic model is expected to provide a better approximation to the stochastic model when the number of patches is small. In the extreme case where the number of patches is comparable to $N$, only a small number of individuals would occupy any given patch, a situation where a differential equation approximation would be ineffective.

In what follows, we identify the fixed points of our deterministic model, investigate their stability, and thus elucidate conditions for persistence and extinction of the metapopulation.

3.4 Equilibrium Behaviour

The basic problem we seek to address is to understand what happens to the metapopulation in the long term. Assuming there is no external immigration and the population ceiling $N$ is finite then it is known that the metapopulation will eventually go extinct. However, for $N$ sufficiently large, the time to extinction may be very large and the metapopulation may settle into some quasi-equilibrium state before going extinct. In this section, we examine the quasi-equilibrium state of the metapopulation described by the Markov population process with rates (3.1) by examining the fixed points of the limiting deterministic model given by (3.6). The following assumptions about the population dynamics are assumed to hold:
(A) There is no immigration from outside the system into any patch. That is $\nu_i = 0$ for all $i$.

(B) The migration functions are linear: $\phi_i(n) = \phi_i n$, where $\phi_i > 0$ for all $i$ (implying that $\hat{\phi}_i(x) = \phi_i x$ for all $i$).

(C) For all $i$ and $j$, $\lambda_{ij} = 0$ implies $\lambda_{ji} = 0$.

(D) For all $i$ and $j$, there is a finite sequence $(a_k)$ such that $\lambda_{ia_1}\lambda_{a_1a_2}\ldots\lambda_{a_{m}j} \neq 0$.

(E) For all $i$, $b_i(x)$ is a continuously differentiable, strictly decreasing function on $[0, 1]$ such that $b_i(1) = 0$ and $xb_i(x)$ is strictly concave on $[0, 1]$.

(F) The parameters $\phi_i$, $\lambda_{ij}$ and $M_i$ satisfy $\phi_i\lambda_{ij}M_i = \phi_j\lambda_{ji}M_j$ for all $i, j$.

Assumption (A) allows the metapopulation to go extinct as it permits the system to contained a fixed point at zero. The result presented in Theorem 3.11, however, assumes that (A) does not hold. Assumption (B) stipulates that the rate at which individuals leave a patch is proportional to the number of individuals in that patch; as the patch becomes more crowded, individuals leave at a greater rate. Whilst it is true that this assumption simplifies the analysis, there is considerable empirical evidence to support density-dependent migration [1, 54, 71, 86], a linear migration term often being used [54, 71, 10]. Assumption (C) implies that all migration paths must allow two-way movement; if an individual moves from patch $i$ to patch $j$, it must be possible to return without needing to go through other patches. Assumption (D) excludes the possibility that any patch or group of patches is isolated. If Assumption (D) did not hold, then the metapopulation could be divided into a number of smaller metapopulations that could each be analysed separately. Assumption (E) implies that as a patch becomes full and less space is available for new individuals, the birth rate decreases. When the patch is full, the birth rate is zero. In an empty network, Assumption (F) translates to every patch having the same maximum migration rate to any other patch. For example, in the symmetric case where $\lambda_{ij} = \lambda_{ji}$ for all $i, j$, individuals would migrate away faster from a smaller patch than a larger patch. Although this restriction is quite strict, it is not required for all of the results. Under these assumptions, an analysis of the behaviour of a population governed by (3.6) is provided.
In the analysis to follow, the concept of a partially ordered flow is employed. For notation regarding the relation of vectors, see Section 2.4. According to Theorem 2.14, any flow $F$ on an open set $U \subset \mathbb{R}^n$ is partially ordered if all off–diagonal elements of $\nabla F(x)$ are nonnegative for all $x \in U$, that is, if $\nabla F(x)$ is a Metzler matrix. For the ODE (3.6) the off–diagonal elements of $\nabla F(x)$ are given by

$$\nabla F_{ij}(x) = \hat{\phi}'_j(x_j)\lambda_{ji} \left(1 - \frac{x_i}{M_i}\right) + \hat{\phi}_i(x_i)\frac{\lambda_{ij}}{M_j}. $$

Define the open set $U_\epsilon = (-\epsilon, M_1 + \epsilon) \times \cdots \times (-\epsilon, M_J + \epsilon)$ for some $\epsilon > 0$. Under Assumptions (B) and (C), $\nabla F_{ij}(x) \geq 0$ for all $x \in U_\epsilon$ for $\epsilon$ sufficiently small. Thus, the flow of $F$ is partially ordered on $U_\epsilon$.

**Proposition 3.2.** Assume (B) and (C) hold. The set $E := [0, M_1] \times \cdots \times [0, M_J]$ is a positive invariant set for the flow defined by (3.6). That is, if $x_0 \in E$, then $x(t, x_0) \in E$ for all $t \geq 0$.

**Proof.** Note that $E \subset U_\epsilon$ and for any $x_0 \in E$, $0 \leq x_0 \leq M$. As the flow is partially ordered on $U_\epsilon$, if $x_0 \in E$, then

$$x(t, 0) \leq x(t, x_0) \leq x(t, M), \quad t \geq 0. \quad (3.7)$$

Therefore, to show that $E$ is a positive invariant set it is sufficient to show that $x(t, 0) \in E$ and $x(t, M) \in E$ for all $t \geq 0$. Consider the trajectory starting at $M$.

The elements of $F(M)$ are given by

$$F_i(M) = - \left( d_i M_i + \hat{\phi}_i(M_i) \left(1 - \sum_{j \neq i} \lambda_{ij}\right) \right) = - \left( d_i M_i + \hat{\phi}_i(M_i)\lambda_{i0}\right) < 0. $$

Therefore, for all $s > 0$ sufficiently small, $x(s, M) < M$, which implies that $x(ns, M) \leq M$, where $n \in \mathbb{Z}_+$, again by partial ordering. Therefore, for all $t \geq 0$,

$$x(t, M) \leq M. \quad (3.8)$$

At the lower boundary, $F(0) \geq 0$ so using a similar argument as was used to derive (3.8), $x(t, 0) \geq 0$ for all $t \geq 0$. This, combined with (3.7) and (3.8), implies that
$x(t, M) \in E$ for all $t \geq 0$. This completes the proof. 

### 3.4.1 Extinction

As noted earlier, although the eventual extinction of the Markov population process is certain under Assumption (A), the process may take a very long time to reach the extinction state. However, if the deterministic process (3.6) converges to the extinction state quickly, then, from Theorem 3.1, it would be reasonable to conclude that the Markov population process also goes extinct quickly. The following theorem describes the behaviour of the deterministic process in a neighbourhood of the extinction state.

**Theorem 3.3.** Assume (A) - (E) hold. If there exists a $y \in \mathbb{R}_+^J\setminus \{0\}$ such that

$$
(b_i(0) - d_i - \phi_i) y_i + \phi_i \sum_{j \neq i} \lambda_{ij} y_j \leq 0, \quad \text{for all } i,
$$

with strict inequality for at least one $i$, the fixed point 0 is asymptotically stable. If there is no $y \in \mathbb{R}_+^J\setminus \{0\}$ satisfying (3.9), then 0 is unstable.

**Proof.** Under Assumption (A) and (B), 0 is a fixed point of (3.6), that is $F_i(0) = 0$. The elements of the Jacobian of $F$ at 0 are given by

$$
\nabla F(0)_{ij} = \begin{cases} 
    b_i(0) - d_i - \phi_i & \text{if } j = i \\
    \phi_j \lambda_{ji} & \text{if } j \neq i.
\end{cases}
$$

As all parameters are non-negative, $J_0 := \nabla F(0)$ is a Metzler matrix. From Assumption (D), $J_0$ is irreducible, meaning for every pair $(i, j)$ there is an integer $m$ such that $(J_m^0)_{ij} > 0$. Henceforth properties of $J_T^0$ will be exploited, noting that its eigenvalues are the same as those of $J_0$. Since $J_T^0$ is also a Metzler matrix, part (c) of Theorem 2.15 implies that $J_T^0$ has a real eigenvalue $r$ which is greater than the real part of any other eigenvalue of $J_T^0$. Furthermore, from Part (e) of that theorem, $r \leq 0$ if and only if there is a vector $y > 0$ such that $J_T^0 y \leq 0$ and $r < 0$ if and only if $J_T^0 y < 0$. The condition $J_T^0 y \leq 0$ gives rise to (3.9). Now write $F(x)$ in (3.6) as

$$
F(x) = J_0 x + \tilde{F}(x),
$$

(3.10a)
Figure 3.2: Illustrating the different behaviour ODE (3.6) exhibits when \( r = 0 \) for a \( J = 2 \) system. The red dotted lines are the trajectories for a given initial condition and the black dots are fixed points. The birth rate function used was \( b_i(x/M_i) = b_i(0)(1 - x/M_i) \). The parameters used are given in Table 3.1.

where

\[
\tilde{F}_i(x) = x_i \left( b_i \left( \frac{x_i}{M_i} \right) - b_i(0) \right) + \sum_{j \neq i} x_i x_j \left( \frac{\phi_i \lambda_{ij}}{M_i} - \frac{\phi_j \lambda_{ji}}{M_j} \right). \tag{3.10b}
\]

It can be seen that \( \tilde{F}(x) \) is Lipschitz continuous on \( E \) and that

\[
\lim_{\|x\| \to 0} \frac{\|\tilde{F}(x)\|}{\|x\|} = 0. \tag{3.11}
\]

If \( r < 0 \), then \( 0 \) is an asymptotically stable fixed point by Theorem 7.1 of [127]. If \( r > 0 \), then \( 0 \) is unstable by Theorem 7.3 of [127].

It must be noted that Theorem 3.3 does not deal with the case of equality in (3.9) for all \( i \). This corresponds to the case of where the maximum eigenvalue of \( J_0 \) is 0. Numerical studies have shown that, in such cases, the stability of \( 0 \) is model specific. This is demonstrated in Figure 3.2. In plot A, the trajectory begins with both patches being full and the metapopulation tends towards the extinction state. As the system is partially ordered, every other trajectory in \( E \) will also tend towards the extinction state implying it is globally stable. However, in plot B, the trajectory begins with both patches near extinction and the metapopulation moves away to a nonzero fixed point. Hence, in the second case, the extinction state is unstable.

The conditions given in Theorem 3.3 are not easily interpreted, particularly for
metapopulations consisting of a large number of patches. The following corollaries provide simpler sufficient conditions for stability/instability of the extinction state.

**Corollary 3.4.** Assume (A) - (E) hold. If

$$\frac{b_i(0)}{d_i + \phi_i \lambda_0} \leq 1, \quad \text{for all } i,$$

(3.12a)

with a strict inequality for at least one $i$, 0 is asymptotically stable, while if

$$\frac{b_i(0)}{d_i + \phi_i \lambda_0} > 1, \quad \text{for all } i,$$

(3.12b)

0 is unstable.

**Proof.** Take $y = 1$. The condition for the asymptotic stability of 0 from Theorem 3.3 is satisfied if

$$b_i(0) - d_i - \phi_i + \phi_i \sum_{j \neq i} \lambda_{ij} \leq 0, \quad \text{for all } i,$$

(3.13)

with a strict inequality for at least one $i$. Using (3.4) we may express (3.13) as $b_i(0) - d_i - \phi_i \lambda_0 \leq 0$, for all $i$, which gives rise to (3.12a). To prove the second part of the corollary, apply Corollary 1 of Theorem 2.8 of [117] to $J_0$ to give the following lower bound on $r$, the largest real part of the eigenvalues of $J_0$:

$$\min_i \left( b_i(0) - d_i - \phi_i + \sum_{j \neq i} \phi_i \lambda_{ij} \right) \leq r.$$

(3.14)

Recalling (3.4) again, (3.14) becomes

$$\min_i (b_i(0) - d_i - \phi_i \lambda_0) \leq r.$$
If (3.12b) holds, then \( \min_i (b_i(0) - d_i - \phi_i \lambda_i) > 0 \) and \( r > 0 \). Hence, \( 0 \) is unstable from Theorem 3.3.

\[ \text{Remark: For homogeneous systems, where } b_i(0)/(d_i + \phi_i \lambda_i) = \alpha \neq 1 \text{ for all } i, \] conditions (3.12) are almost necessary and sufficient for the stability of \( 0 \). When \( \alpha = 1 \), the maximum eigenvalue of \( J_0 \) is 0 and the stability of \( 0 \) is not characterised by determining the maximum eigenvalue of \( J_0 \), as stated previously.

**Corollary 3.5.** Assume (A) - (E) hold. Then if
\[
\frac{b_i(0)}{d_i + \phi_i} > 1, \quad \text{for at least one } i, \quad (3.15)
\]
\( 0 \) is unstable.

**Proof.** If (3.15) holds for some \( i \), then, for any vector \( y \in \mathbb{R}^J \setminus \{0\} \),
\[
y_i (b_i(0) - d_i - \phi_i) + \phi_i \sum_{j \neq i} y_j \lambda_{ij} > 0.
\]
Thus there is no \( y \in \mathbb{R}^J \setminus \{0\} \) satisfying inequality (3.9) and, according to Theorem 3.3, \( 0 \) is unstable.

The above results only address the behaviour of the system when it starts in a neighbourhood of the extinction state. It may be that the metapopulation can persist if it is initially densely populated, such as for a metapopulation with Allee effect [6, 37]. Due to the partial ordering of the flow, an Allee effect would be observed if there exists a nonzero fixed point and \( 0 \) were asymptotically stable. The following result shows that if Assumption (F) is imposed and inequality (3.9) holds, then the metapopulation goes extinct regardless of the initial condition. Hence, the metapopulation does not display an Allee effect under Assumption (F).

**Theorem 3.6.** Assume (A) - (F) hold. If there exists a \( y \in \mathbb{R}^J \setminus \{0\} \) such that (3.9) holds, then \( x(t,x_0) \to 0 \) for all \( x_0 \in E \).

Before giving the proof of this theorem we first derive an upper bound on \( x(t,x_0) \).
Lemma 3.7. Assume hold and define as the solution to
\[
\frac{dy}{dt} = J_0 y, \quad y(0) = y_0.
\]
(3.16)

Then \( x(t, x_0) \leq y(t, x_0) \) for all \( t \geq 0 \).

Proof. The solution to (3.16) is given by \( y(t, x_0) = e^{tJ_0} x_0 \) and the unique solution to (3.6) may be written as
\[
x(t, x_0) = y(t, x_0) + \int_0^t e^{(t-s)J_0} \tilde{F}(x(s))ds,
\]
where \( \tilde{F}(x) \) is given by (3.10b). Under Assumptions (E) and (F), \( \tilde{F}(x) \leq 0 \) for all \( x \in E \).

Furthermore, as \( J_0 \) is an irreducible Metzler matrix (from the proof of Theorem 3.3), we can apply Theorem 2.7 of [117] to conclude that \( e^{tJ_0} \) is positive for all \( t \geq 0 \). It follows that \( x(t, x_0) \leq y(t, x_0) \) for all \( t \geq 0 \).

We can now use Lemma 3.7 to show that 0 is globally stable under the conditions of Theorem 3.6.

Proof of Theorem 3.6. When there exists \( y \in \mathbb{R}^n \setminus \{0\} \) such that (3.9) holds, the eigenvalue of \( J_0 \) with largest real part, \( r \), satisfies \( r \leq 0 \) and has algebraic multiplicity one (parts (a), (c), (d) and (e) of Theorem 2.15). Applying Theorem 6.1(b) of [127], there exists a positive constant \( C \) such that \( \|y(t, x_0)\| \leq C\|x_0\| \). Therefore, \( y(t, x_0) \) is bounded uniformly in \( t \geq 0 \) and \( x_0 \in E \). Next it is known that \( x_0 \leq M \) for all \( x_0 \in E \). Therefore, if \( x(t, M) \to 0 \) then, due to partial ordering, \( x(t, x_0) \to 0 \) for all \( x_0 \in E \) also. So we will only consider \( x(t, M) \). As \( E \) is a positive invariant set, from Proposition 3.2 \( x(s, M) \leq M \) for any \( s \geq 0 \). Due then to partial ordering \( x(t, x(s, M)) = x(s + t, M) \leq x(s, M) \leq M \) for any \( s, t \geq 0 \). Hence, \( x(t, M) \) is monotone decreasing and bounded, implying that it has a limit as \( t \to \infty \) which we denote by \( \alpha \).

The proof will be complete if we can show that \( \alpha = 0 \).

Suppose \( r < 0 \). From Theorem 4.5 of [68], \( y(t, x_0) \to 0 \) for any \( x_0 \in E \). Since \( y(t, x_0) \) bounds \( x(t, x_0) \), \( x(t, x_0) \to 0 \) for any \( x_0 \in E \), showing \( \alpha = 0 \). Now suppose \( r = 0 \) and that \( 0 < \alpha \). With Assumptions (E) and (F), this implies that there exists a \( \beta > 0 \) such that \( -\tilde{F}(x(t, M)) \geq \beta \) for all \( t > 0 \). Since \( e^{tJ_0} \) is positive for all \( t \geq 0 \) [117] Theorem 2.7,
it follows that
\[
\int_0^t e^{(t-s)J_0} \tilde{F}(x(s,M))ds \leq \int_0^t e^{(t-s)J_0} \beta ds.
\]  
(3.18)

From Theorem 2.7 of [117] it is known that 
\[e^{sJ_0} = wv^T + O(e^{\lambda s}),\]
where \(\lambda < 0\), and \(w\) and \(v\) are the positive right and left eigenvectors of \(J_0\) corresponding to the eigenvalue \(r\) normed so that \(v^Tw = 1\). Therefore, the integral on the right-hand side of (3.18) tends to negative infinity as \(t \to \infty\). Hence, from Lemma 3.7 (equation (3.17)) and recalling that \(y\) is bounded, we can take \(t\) sufficiently large so that \(x(t,M) < 0\). This is a contradiction since, from Proposition 3.2, \(x(t,M) \in E\) for all \(t\). Therefore \(\alpha = 0\) and \(x(t,x_0) \to 0\) for all \(x_0 \in E\). 

3.4.2 Persistence

Theorem 3.3 shows how the metapopulation behaves when near extinction but does not provide any information concerning the behaviour of the metapopulation away from the extinction state. Theorem 3.6 provides a complete description of the metapopulation when the model satisfies Assumption (F) and inequality (3.9). This subsection studies the behaviour of the metapopulation when condition (3.9) does not hold.

**Theorem 3.8.** Assume \(A - E\) hold. If there is no \(y \in \mathbb{R}^d_+ \setminus \{0\}\) satisfying (3.9), \(E\) contains at least one nonzero fixed point \(x^{(1)*}\) and, for all \(x_0\) such that \(0 < x_0 \leq x^{(1)*}\), \(x(t,x_0) \to x^{(1)*}\).

**Proof.** Suppose that there is no \(y \in \mathbb{R}^d_+ \setminus \{0\}\) satisfying (3.9), then the eigenvalue of \(J_0\) with largest real part, \(r\), satisfies \(r > 0\) and has algebraic multiplicity one (parts (a), (c), (d) and (e) of Theorem 2.15). The corresponding eigenvector \(v\) of \(J_0\) satisfies \(v > 0\). Finally, as \(E\) is a positive invariant set we may apply Theorem 2.8 of [120]. Combining parts (1), (4) and (5) of Theorem 2.8 [120], it can be concluded that either \(x(t,x_0) \to x^{(1)*}\) or \(\|x(t,x_0)\| \to \infty\) for all \(x_0 \geq 0, x_0 \neq 0\) where \(x^{(1)*} > 0\). As \(E\) is a positive invariant set, \(\|x(t,x_0)\| \leq \|M\|\). Hence, \(x(t,x_0) \to x^{(1)*}\) for all \(x_0\) such that \(0 < x_0 \leq x^{(1)*}\). 

This theorem shows that for sufficiently small \(x_0 \neq 0\), trajectories will tend to a
nonzero fixed point. Furthermore, due to partial ordering, \( x^{(1)^*} \leq \liminf_{t \to \infty} x(t, x_0) \) for all \( x_0 \in E \setminus \{0\} \), implying that the metapopulation will persist. However, Theorem 3.8 does not preclude the possibility of another nonzero fixed point. This issue is addressed in the following theorem under Assumption (F).

**Theorem 3.9.** Assume (A) - (F) hold. If there is no \( y \in \mathbb{R}_+^J \setminus \{0\} \) satisfying (3.9), then there is a unique nonzero fixed point \( x^{(1)^*} \) and \( x(t, x_0) \to x^{(1)^*} \) for all \( x_0 \in E \setminus \{0\} \).

As in the proof of Theorem 3.6, we first derive an upper bound on \( x(t, x_0) \).

**Lemma 3.10.** Assume (A) - (F) hold and let \( y(t, y_0) \) be the solution to

\[
\frac{dy}{dt} = J_1(y - x^{(1)^*}), \quad y(0) = y_0,
\]

where \( J_1 = \nabla F(x^{(1)^*}) \). Then \( x(t, x_0) \leq y(t, x_0) \) for all \( t \geq 0 \).

**Proof.** It is known that \( y(t, x_0) = x^{(1)^*} + e^{J_1 t}(x_0 - x^{(1)^*}) \). With the simple change of coordinates \( z = x(t, x_0) - x^{(1)^*} \), together with Assumption (F), the solution to (3.6) may be written as

\[
x(t, x_0) = y(t, x_0) + \int_0^t e^{(t-s)J_1} \bar{F}(x(s, x_0)) \, ds, \tag{3.19}
\]

where \( \bar{F}(x) \) is given by

\[
\bar{F}_i(x) = x_i \left( b_i \left( \frac{x_i}{M_i} \right) - b_i \left( \frac{x_i^{(1)^*}}{M_i} \right) \right) + \left( x_i^{(1)^*} - x_i \right) b_i' \left( \frac{x_i^{(1)^*}}{M_i} \right) \frac{x_i^{(1)^*}}{M_i}.
\]

Under Assumption (F), \( \bar{F}(x(s, x_0)) \leq 0 \) for \( x(s, x_0) \in E \). As in Lemma 3.7, \( e^{J_1 t} \) is a positive matrix since \( J_1 \) is an irreducible Metzler matrix. We then conclude that \( x(t, x_0) \leq y(t, x_0) \) for all \( t \geq 0 \).

**Proof of Theorem 3.9.** From Theorem 3.8 it is known that for all \( x_0 \) that satisfy the condition \( 0 < x_0 \leq x^{(1)^*} \), \( x(t, x_0) \to x^{(1)^*} \). If we can show that \( x(t, M) \to x^{(1)^*} \), then we can conclude that \( x(t, x_0) \to x^{(1)^*} \) for any \( x_0 \in E \setminus \{0\} \) as the flow is partially ordered. Following the arguments in the proof of Theorem 3.6, \( \lim_{t \to \infty} x(t, M) \) exists. Denote this limit by \( \alpha \). The proof will be complete if we can show that \( \alpha = x^{(1)^*} \). As in
Theorem 3.3. $J_1$ has a real eigenvalue $r_1$ which is greater than the real part of any other eigenvalue of $J_1$.

Suppose $r_1 < 0$. From Theorem 4.5 of [68], $y(t, x_0) \to x^{(1)*}$ for any $x_0 \in E$. Since $y(t, x_0)$ bounds $x(t, x_0)$, $x(t, x_0) \to x^{(1)*}$ for any $x_0 \in E$, showing $\alpha = x^{(1)*}$. Now suppose that $r_1 = 0$ and $x^{(1)*} < \alpha$. With Assumptions (E) and (F) this implies that there exists a $\beta > 0$ such that $-\bar{F}(x(t, M)) \geq \beta$ for all $t > 0$. Note that $e^{tJ_1}$ is positive for all $t \geq 0$ as $J_1$ is a Metzler matrix [117, Theorem 2.7]. It now follows from Lemma 3.10 (equation (3.19)) and recalling that $y$ is bounded, we can take $t$ sufficiently large so that $x(t, M) < x^{(1)*}$. This is a contradiction as the flow is partially ordered and $x^{(1)*}$ is a fixed point. Therefore $\alpha = x^{(1)*}$ and $x(t, x_0) \to x^{(1)*}$ for all $x_0 \in E \setminus \{0\}$. 

As before, the conditions for persistence can be quite difficult to interpret. If it is assumed that Assumption (A) to (F) hold then, according to Theorem 3.9, if condition (3.9) does not hold, the population will tend towards a unique nonzero fixed point. We may write this fixed point explicitly if $b_i(0)/(d_i + \phi_i\lambda_{i0}) = \alpha > 1$ for all $i$. It is rather simple to show that $x^* \in E \setminus \{0\}$, with elements $x^*_i = M_i(1 - \alpha^{-1})$, satisfies $F(x^*) = 0$. Furthermore, all trajectories converge to $x^*$.

### 3.4.3 Persistence with Immigration

Sections 3.4.1 and 3.4.2 have shown that when there is no immigration into the system, the persistence of the metapopulation is unknown and it can depend on a number of factors. When immigration into the system from outside is nonzero, however, the population will persist in any case. Assumption (F) is not required for persistence, as a simple argument can be made that $0$ is not a fixed point of any system with immigration (that is, $F(0) \neq 0$). However, a stronger result is valid when Assumption (F) holds. The following theorem provides the details precisely.

**Theorem 3.11.** Assume (A) does not hold while (B) - (E) hold. Then $E$ contains at least one non-zero fixed point and

$$
\liminf_{t \to \infty} x(t, x_0) > 0, \quad \text{for all } x_0 \in E.
$$

If (F) holds also, then $E$ contains one unique asymptotically stable non-zero fixed point.
Chapter 3: A Spatially Structured Metapopulation Model

$x^*$, such that

$$\lim_{t \to \infty} x(t, x_0) = x^*, \quad \text{for all } x_0 \in E.$$  

Proof. If Assumption [A] does not hold, then $F(0) \neq 0$ and hence the vector $0$ cannot be a fixed point. From the proof of Proposition 3.2, it is known that $x(t, 0)$ is a strictly increasing function and $E$ is invariant, therefore $\lim_{t \to \infty} x(t, 0)$ must exist and it will be denoted by $L_0$. From Lemma 2.3 of [120], it follows that $L_0$ is an equilibrium point. Furthermore, $0 < x(t, 0)$ for sufficiently small $t$, which leads to $0 < L_0$ due to partial ordering. But as $0 < x_0$ for all $x_0 \in E \setminus \{0\}$ then, again due to partial ordering, $0 < L_0 \leq \lim \inf_{t \to \infty} x(t, x_0)$, which gives rise to the first statement in the theorem.

Again from the proof of Proposition 3.2, it is known that $x(t, M)$ is a strictly decreasing function and $E$ is invariant which together imply that $\lim_{t \to \infty} x(t, M)$ must exist and it will be denoted by $L_M$, noting that, due to the partial ordering property of the ODE, $L_0 \leq \lim \inf_{t \to \infty} x(t, x_0) \leq \lim \sup_{t \to \infty} x(t, x_0) \leq L_M$ for all $x_0 \in E$. It is also known that $0 < x(t, 0) \leq x(t, M) < M$ and $E$ is invariant so again Lemma 2.3 of [120] is applied which states that along with $L_0$, $L_M$ is also an equilibrium point. If we can show that $L_0 = L_M$, the proof is complete.

It is known there exists at least one value $y \in E \setminus \{0\}$, such that $F(y) = 0$. As $J_z := \nabla F(x)$ is a Metzler matrix for all $x \in E$, then from parts (a), (c) and (e) of Theorem 2.15 there exists a real eigenvalue of $J_y$, $r_y$, such that $r_y > \Re \lambda$, $r_y \neq \lambda$ for any eigenvalue $\lambda$ of $J_y$ and $r_y < 0$ if there exists $z \in \mathbb{R}^\lambda \setminus \{0\}$ such that $J_z z \leq 0$, with an inequality for at least one element. Let $y$ satisfy $F(y) = 0$. Under Assumption (F), we may then write the elements of $J_y y$ as

$$(J_y y)_i = \left[ -(d_i + \nu_i + \phi_i) + b_i \left( \frac{y_i}{M_i} \right) \right] y_i + \frac{y_i^2}{M_i} \frac{d}{dy_i} b_i (y_i) + \sum_{j \neq i} \phi_j \lambda_{ij} y_j,$$

$$= -M_i \nu_i + \frac{y_i^2}{M_i} \frac{d}{dy_i} b_i (y_i).$$

As $M_i \nu_i > 0$ for at least one $i$ due to Assumption [A] not holding and $\frac{d}{dx} b_i (x) < 0$ for all $i$ due to Assumption [E], then $(J_y y)_i < 0$ for any $y \in \mathbb{R}^\lambda \setminus \{0\}$ for at least one $i$. Therefore, if $y \in E \setminus \{0\}$ is chosen such that $F(y) = 0$ is satisfied, then $r_y < 0$, showing that any fixed point in $E$ must be stable.

As $L_0, L_M \in E$, then $r_{L_0} < 0$ and $r_{L_M} < 0$. Proposition 2.9 of [120] may now be
applied which states that there must exist at least one fixed point \( s \) (that is, \( F(s) = 0 \)) that satisfies \( L_0 \leq s \leq L_M \) and \( r_s \geq 0 \). However, \( L_0 \leq s \leq L_M \) implies that \( s \in E \) and it is known that for any \( s \in E \) that satisfies \( F(s) = 0, r_s < 0 \). This is a contradiction and therefore \( L_0 = L_M \), which completes the proof.

Theorem 3.11 shows if Assumption \( (A) \) does not hold, the metapopulation will not go extinct and if Assumption \( (F) \) holds also, the metapopulation will converge to a unique nonzero fixed point, regardless of its initial value.

3.4.4 Allee Effect

Under Assumption \( (F) \), we have shown that when there exists a \( y \in \mathbb{R}_+ \setminus \{0\} \) satisfying (3.9), the metapopulation will tend towards the extinction state \( 0 \). Otherwise, the metapopulation will tend towards a nonzero unique equilibrium level regardless of its initial value. However, it has been observed for some populations that whether it progresses towards extinction or a nonzero equilibrium depends on the initial population size. This is known as the Allee effect [5]. Courchamp et al. [38] have shown the existence of an Allee effect for metapopulations in their study of the African Wild dog, *Lycaon pictus*.

The possibility of our metapopulation model displaying an Allee effect when Assumption \( (F) \) does not hold is now investigated. To address this question, we focus on the two–patch metapopulation with the decreasing birth rate function of the form \( b_i(x/M_i) = b_i(0)(1 - x/M_i) \). To simplify notation, below we write \( b_i \) for \( b_i(0) \). Let \( M_1 = 1 - \varepsilon \) and \( M_2 = \varepsilon \) and \( y_i := x_i^* / M_i = y_i^{(0)} + \varepsilon y_i^{(1)} + \varepsilon^2 y_i^{(2)} + \ldots \) for \( i = 1, 2 \), where \( x_i^* \) satisfies \( F(x^*) = 0 \). Using perturbation theory, an expansion for the fixed points of the system is determined.

The fixed points are the solutions to a system of two quadratic equations in two variables, entailing four solutions. One of these is \( 0 \) and, for sufficiently small \( \varepsilon > 0 \),
we can approximate the remaining three solutions \( y_{(i)} := (y_1, y_2) \) to first order by

\[
\begin{align*}
Y_{(1)} &= \left(\frac{\phi_2 \lambda y_1 C_1}{(d_1 + \phi_1 - b_1 - C_1 \phi_1 \lambda_1)} \varepsilon + O(\varepsilon^2)}{C_1 + C_3(C_1) \varepsilon + O(\varepsilon^2)}
\end{align*}
\]

\( 3.20a \)

\[
\begin{align*}
Y_{(2)} &= \left(\frac{\phi_2 \lambda y_2 C_2}{(d_1 + \phi_1 - b_1 - C_2 \phi_1 \lambda_1)} \varepsilon + O(\varepsilon^2)}{C_2 + C_3(C_2) \varepsilon + O(\varepsilon^2)}
\end{align*}
\]

\( 3.20b \)

\[
\begin{align*}
Y_{(3)} &= \left(\frac{1 - \phi \lambda_0 + d_1}{b_1} - \frac{d_2 + \phi \lambda_2 y_2}{b_1 - d_1 - \phi_1 \lambda_0} \varepsilon + O(\varepsilon^2)}{1 - \phi_1 \lambda_2 (b_1 - d_1 - \phi_1 \lambda_0)} \varepsilon + O(\varepsilon^2)}
\end{align*}
\]

\( 3.20c \)

where

\[
C_1 = \frac{1}{2} \left(1 + \alpha \left(R_1^{(1) - 1} - R_0^{(2) - 1}\right) - B_0^{(0) - 1}\right) \right) + \gamma
\]

\[
C_2 = \frac{1}{2} \left(1 + \alpha \left(R_1^{(1) - 1} - R_0^{(2) - 1}\right) - B_0^{(0) - 1}\right) - \sqrt{\left(1 + \alpha \left(R_1^{(1) - 1} - R_0^{(2) - 1}\right) - B_0^{(0) - 1}\right)^2 + \gamma}
\]

\[
\gamma = \alpha \left(R_0^{(1) - 1} - 1\right) \left(R_1^{(2) - 1} - 1\right) + R_0^{(2) - 1} - 1,
\]

\( 3.21 \)

and

\[
\alpha = \frac{b_1}{\phi_1 \lambda_1}, \quad R_0^{(i)} = \frac{b_i}{d_i + \phi_1 \lambda_0}, \quad R_1^{(i)} = \frac{b_i}{d_i + \phi_i}, \quad i = 1, 2.
\]

The expression for \( C_3(x) \) is not given here owing to its length.

For the metapopulation to display an Allee effect, \( 0 \) must be stable. Using Theorem 3.3, the fixed point \( 0 \) is stable if there exists \( y \in \mathbb{R}_0^2 \setminus \{0\} \) such that

\[
y_1(b_1 - d_1 - \phi_1) + \phi_1 y_2 \lambda_{12} \leq 0,
\]

\( 3.22 \)

\[
y_2(b_2 - d_2 - \phi_2) + \phi_2 y_1 \lambda_{21} \leq 0,
\]

\( 3.23 \)

with strict inequality in one of (3.22) and (3.23). Such a \( y \) exists if \( d_i + \phi_i - b_i > 0 \) for \( i = 1, 2 \) and

\[
1 < \left(\frac{(d_1 + \phi_1 - b_1)(d_2 + \phi_2 - b_2)}{\phi_1 \lambda_{12} \phi_2 \lambda_{21}}\right)
\]

\( 3.24 \)

Inequality (3.24) implies \( \gamma > 0 \) which implies \( C_1 > 0 \). A nonzero fixed point is present
in $[0, 1]^2$ if $C_1 < 1$, which is satisfied if

$$R_0^{(1)} > \alpha \frac{1 + \frac{1}{3} R_1^{(2)} - 1}{R_0^{(2)} - 1}.$$  \hspace{1cm} (3.25)

Therefore, if the metapopulation parameters are such that $R_0^{(1)} > 1$, $\gamma > 0$ and (3.25) is met, then there exists an $\varepsilon > 0$ sufficiently small such that the three points $0$, $y_{(1)}$, and $y_{(3)}$ are contained in $[0, 1]^2$.

Thus we have one stable fixed point at $0$ and two nonzero fixed points whose stability is unknown. Let $x^{(i)*}$, $i = 1, 2$, denote the nonzero fixed points. As the flow is partially ordered, for any $x_0$ such that $x^{(i)*} \leq x_0$, we have $x^{(i)*} \leq \lim\inf_{t \to \infty} x(t, x_0)$. Therefore, the metapopulation will persist if it is initially sufficiently large. We can conclude that the metapopulation can display an Allee effect for a certain range of parameters. Figure 3.3 plots the scaled fixed points of a metapopulation as a function of $b_1$. The Allee effect is present when there is a dotted line, representing an unstable fixed point, between two solid lines, representing stable fixed points. In this example, the Allee effect is present when $b_1$ is between 0.415 and 0.428. This also illustrates the sensitivity of the metapopulation to disturbance, as only a very small change in the birth rate on the larger patch is needed to reduce the occupancy of the metapopulation from $(y_1, y_2) \approx (0.1, 0.3)$ to the extinction state.

Although we have only demonstrated the Allee effect for a two–patch metapopulation where the patch ceilings are significantly different, numerical results show that the Allee effect can be present in metapopulations where the ceilings are not significantly different. Figure 3.4A shows a phase plane diagram for such a system and illustrates the different long term behaviour the system exhibits with different initial points. In one case the system starts with patch 1 full and patch 2 is empty, and the system converges to the largest fixed point. However, in the other case, patch 2 is full and patch 1 is empty, and the system goes extinct.

In Figures 3.3 and 3.4A, the three fixed points are partially ordered with the largest fixed point stable and the other nonzero fixed point unstable. There is some theoretical evidence that this holds in general. If all the eigenvalues of $\nabla F(x^{(2)*})$ have negative real
Figure 3.3: The red solid line represents stable fixed points, the blue dashed line represents unstable fixed points and the black dash-dot line represents the first order approximation (3.20c). Parameters used for plots A and B were $b_2 = 0.1$, $d_1 = 0.1$, $d_2 = 0.7$, $\phi_1 = 0.4$, $\phi_2 = 0.9$, $M_1 = 0.99$ and $M_2 = 0.01$, while $\lambda_{10} = 0.4$ and $\lambda_{20} = 0.5$.

Figure 3.4: The red dotted lines are the deterministic trajectories imposed by the flow (3.6), the squares are unstable fixed points, the dots are stable fixed points and the solid blue lines are the nullclines. Parameters used for plot A were $b_1 = 0.56$, $b_2 = 0.06$, $d_1 = 0.02$, $d_2 = 0.14$, $\phi_1 = 0.55$, $\phi_2 = 0.09$, $M_1 = 0.54$ and $M_2 = 0.46$, while $\lambda_{10} = 0.31$, and $\lambda_{20} = 0.99$. Parameters used for plot B were $b_1 = 5.24$, $b_2 = 7.5$, $d_1 = 1.14$, $d_2 = 5.59$, $\phi_1 = 7.03$, $\phi_2 = 0.29$, $M_1 = 0.78$ and $M_2 = 0.22$, while $\lambda_{10} = 0.06$ and $\lambda_{20} = 0.51$. 

65
parts, then Proposition 2.9 of [120] implies that $x^{(1)*} \leq x^{(2)*}$ and at least one eigenvalue of $\nabla F(x^{(1)*})$ has nonnegative real part. On the other hand, if one of the eigenvalues of $\nabla F(x^{(1)*})$ has positive real part, then Theorem 2.8 of [120] implies that $x^{(1)*} \leq x^{(2)*}$ and all the eigenvalues of $\nabla F(x^{(2)*})$ have nonnegative real parts. Difficulties in dealing with the case where the eigenvalues of $\nabla F(x^{(i)*})$ have real parts equal to 0 prevent us from proving the general result.

Metapopulation models exhibiting an Allee effect are not new. Both Amarasekare [6] and Zhou & Wang [134] have proposed models displaying an Allee effect. However, in contrast to those models, the Allee effect observed here is not due to a manipulation of the birth rate function. Without migration between patches, the metapopulation model described in this chapter will not exhibit the Allee effect; it is induced by the migration of individuals which is now described.

For a metapopulation to display an Allee effect, it is necessary that the extinction state be stable and for there to be a non-zero equilibrium. The non-zero equilibrium derived through perturbation analysis arises as follows. When the small patch is near capacity, migration between the two patches is reduced to a very low level. This has different effects on the two patches. The population dynamics on the large patch are dominated by the birth and death events. As the per-capita birth rate is initially larger than the death rate, a stable population becomes established. On the small patch, the immigration rate is still considerable relative to the size of the patch, and is sufficient to maintain the population close to capacity.

The extinction state can be seen to be stable by considering what happens when the populations on both patches are small relative to capacity. When the population on the small patch is far from capacity, the per-capita emigration rate on the large patch is significant. The birth rate is not sufficiently large to balance the deaths and emigration events resulting in a decreasing population on the large patch. As previously noted, the small patch has a high per-capita death rate. At the non-zero equilibrium, this was balanced by immigration from the large patch. However, when the metapopulation is near extinction, the migration from the large patch is much smaller and is no longer sufficient to balance the high death rate on the small patch.

The way births, deaths and migrations are accounted for in the model significantly
Chapter 3: A Spatially Structured Metapopulation Model

affect whether or not the Allee effect occurs, as these dynamics encapsulate all the behaviour of individuals. However, as stated in Section 3.2, the way in which migration has been modelled can be altered to account for two additional types of behaviour relating to what happens to individuals if they arrive at an overcrowded patch and, therefore, it is important to discuss the Allee effect in relation to these two types of behaviour. If it is assumed individuals die upon arrival at an overcrowded patch then it is likely an Allee effect will not occur. This is because individuals will migrate to the smaller patch, regardless of its current occupancy, and therefore will not remain on the more habitable, larger patch. This difference of behaviour relating to the current occupancy of the smaller patch is vital in the existence of an Allee effect as, for the metapopulation to persist, individuals need to remain on the larger patch. However, in a two patch system, if it is assumed individuals, upon arrival at an overcrowded patch, choose to migrate again, then they must return to their original destination. Such a scenario is exactly what happens in the current model and therefore, an Allee effect is to be expected if individuals are assumed to move on to another patch.

We have so far focussed on the case where \( \theta \) is stable, that is, when the quantity \( \gamma \) given by equation (3.21) is greater than zero. However, interesting behaviour can also be observed if \( \gamma < 0 \). In that case, if \( \left(1 + \alpha \left(R_1^{(1)} - 1\right) - R_0^{(2)}\right)^2 > |\gamma| \) and \( 1 + \alpha \left(R_1^{(1)} - 1\right) - R_0^{(2)} > 0 \), there potentially exists an \( \varepsilon > 0 \) sufficiently small such that all four fixed points are in \([0,1]^2\). With \( \gamma < 0 \), \( \theta \) is unstable, and hence trajectories tend away from \( \theta \) (the system persists regardless of the initial values). This possibility is discussed by Courchamp et al. [37, section 6.5.3], where Allee effects occur at intermediate population sizes or densities, resulting in up to three interior steady states, two of which are locally stable. This type of behaviour is illustrated in Figure 3.4B. One trajectory starting with patch 2 empty and patch 1 at 8% capacity tends towards a nonzero fixed point. However, when patch 1 begins at 32% capacity the trajectory tends to a larger (by partial ordering) fixed point. These two fixed points are also separated by an unstable fixed point.
3.4.5 Example

We now consider the metapopulation of sea otter (*Enhydra lutris*) in the north-east Pacific Ocean, which has been studied by various groups [118, 46]. Our purpose here is to illustrate the results obtained in the previous section, rather than provide a complete description of the population. Estes [46] studied five populations in this region: Attu Island, south-east Alaska, British Columbia, Washington State and central California. The migration rate between these 5 locations is not large. However, Amchitka Island, a previously unoccupied island adjacent to Attu Island, became occupied with sea otters. We consider a two-patch model to describe the evolution of the sea otter population in these two islands. Siniff and Ralls [118] determined an age distribution for the lifespan of the sea otter from which we can determine the average age of a female to be 5.18 years. Therefore, we set \( d_i = 0.2 \approx 5.18^{-1} \) per year for both islands.

Since the reproductive rate of females is between 0.43 and 0.45 per year [46], we set \( b_i(x) = 0.44(1 - x/M_i) \). The per-capita migration rate \( \phi_i \) is set to be the inverse of the average time an individual spends on island \( i \) before migrating. Finally, since Attu Island is approximately three times the size of Amchitka Island, we set \( M_1 = 0.75 \) (Attu) and \( M_2 = 0.25 \) (Amchitka).

From Corollary 3.5 if the average time an otter spends on one island before migrating is greater than 4.17 years, the extinction fixed point 0 is unstable and trajectories will tend away from it. In that case, Theorem 3.8 implies the population will persist.

Next we consider the case where the average time an otter spends on one island before migrating is less than 4.17 years. By Theorem 3.3, extinction occurs if there exists a \( y \in \mathbb{R}_2^+ \setminus \{0\} \) such that inequalities (3.22) and (3.23) hold, equivalently, if inequality (3.24) holds. Rearranging inequality (3.24), it can be seen that if the average time spent on one island is greater than \( 4.17(1 - \lambda_{12}^i \lambda_{21}^j) \) years, the population will persist. However, if the average time spent on Attu Island and Amchitka Island before migrating are both less than \( 4.17(1 - \lambda_{12} \lambda_{21}) \) years and

\[
\phi_i^{-1} > \frac{4.17\left[4.17(1 - \lambda_{12} \lambda_{21}) - \phi_j^{-1}\right]}{4.17 - \phi_j^{-1}}, \quad i, j \in \{1, 2\}, \ j \neq i,
\]

then the population will persist. If none of these conditions is met, the population
Chapter 3: A Spatially Structured Metapopulation Model

will become extinct. Note that if \(4.17(1 - \lambda_{12}\lambda_{21}) < \phi_i^{-1} < 4.17\) for any \(i\), \((3.26)\) will automatically be satisfied.

These scenarios are illustrated in Figure 3.5. In plots A – C the average time spent on Attu Island is 10 years (\(\phi_1 = 0.1\)), while the average time spent on Amckitka Island is 1.25 years (\(\phi_2 = 0.8\)). The connectivity between patches is very weak, with \(\lambda_{12} = 0.1\) and \(\lambda_{21} = 0.05\). However, from Corollary 3.5 the metapopulation persists. In D – F, \(\phi_1\) is increased to 0.5, but \(\lambda_{12} = 0.9\) and \(\lambda_{21} = 0.4\). As inequality \((3.26)\) is satisfied, the metapopulation persists. Finally, in G – I, \(\phi_1\) is increased to 0.7 per year. Now, inequality \((3.26)\) is not satisfied and the metapopulation goes extinct. With the chosen values for the birth and death rates, it is not possible for condition \((3.25)\) to be satisfied, and so this system will not exhibit the Allee effect.

3.5 Diffusion

The results above do not provide any information concerning the stochastic fluctuations about the deterministic trajectory. Therefore we appeal to Theorem 2.7. To this end define \(G(x) = (g_{ij}(x))\) by

\[
g_{ij}(x) = -\hat{\phi}_i(x_i)\lambda_{ij}\left(1 - \frac{x_j}{M_j}\right) - \hat{\phi}_j(x_j)\lambda_{ji}\left(1 - \frac{x_i}{M_i}\right), \quad \text{for } j \neq i,
\]

\[
g_{ii}(x) = \nu_i M_i + (b_i + d_i - \nu_i)x_i + \hat{\phi}_i(x_i) - \frac{b_i}{M_i}x_i^2 + \sum_{j \neq i} \left(\hat{\phi}_j(x_j)\lambda_{ji}\left(1 - \frac{x_i}{M_i}\right) - \hat{\phi}_i(x_i)\lambda_{ij}\frac{x_j}{M_j}\right), \quad \text{for all } i.
\]

It can be seen that \(G\) is bounded and uniformly continuous if \(\hat{\phi}_i\) has this property for all \(i\), while \(F\) has uniformly continuous first partial derivatives whenever \(\hat{\phi}_i\) has this property for all \(i\). Noting that the Jacobian \(H(x) = (h_{ij}(x))\) of \(F\) has entries given by

\[
h_{ij}(x) = \hat{\phi}_j(x_j)\lambda_{ji} + \hat{\phi}_i(x_i)\frac{\lambda_{ij}}{M_j} - \hat{\phi}_i(x_i)\frac{\lambda_{ij}x_j}{M_j}, \quad \text{for } j \neq i,
\]
Figure 3.5: An illustration of the effect the migration rates and the connectivity have on the persistence of a metapopulation with a population ceiling of 200 sea otters at Attu and Amchitka Islands. For the three cases mentioned in the text, A, D and G illustrate contours of the distribution of the two-dimensional approximating Gaussian diffusion (Theorem 3.12) at $t = 40$ years, together with one realisation of the metapopulation model; B, E and H show the same realisation (solid blue line) on a phase plane diagram together with the deterministic trajectory (3.6) (dotted red line); and C, F and I show this realisation and the deterministic trajectory along with $\pm 2$ standard deviations (dotted lines) determined from the distribution of the approximating Gaussian diffusion.
Chapter 3: A Spatially Structured Metapopulation Model

\[ h_i(x) = b_i - \nu_i - d_i - \frac{2b_i}{M_i} x_i - \frac{\hat{\phi}_i'(x_i)}{\hat{\phi}_i'(x_i)} \left( \lambda_{ij} + \sum_{j \neq i} \lambda_{ij} \left( 1 - \frac{x_j}{M_j} \right) \right) - \sum_{j \neq i} \frac{\hat{\phi}_j(x_j)}{\hat{\phi}_j(x_j)} \frac{\lambda_{ji}}{M_i}, \text{ for all } i, \]

We can apply Theorem 2.7 to show that the fluctuations of \( X_N(t) \) about the deterministric trajectory follow a Gaussian diffusion.

**Theorem 3.12.** Under the conditions of Theorem 3.1, assume also that \( \hat{\phi}_i \) has uniformly continuous first partial derivatives and

\[ \lim_{N \to \infty} \sqrt{N} (X_N(0) - x_0) = z. \]

Then the family of processes \( \{Z_N(t)\} \), defined by

\[ Z_N(s) = \sqrt{N} (X_N(s) - x(s, x_0)), \quad 0 \leq s \leq t, \]

converges weakly in \( D[0, t] \) to a Gaussian diffusion \( Z(t) \) with initial value \( Z(0) = z \) and characteristic function \( \psi = \psi(t, \theta) \) that satisfies

\[ \frac{\partial \psi(t, \theta)}{\partial t} = -\frac{1}{2} \sum_{j,k} \theta_j g_{jk}(x(t)) \theta_k \psi(t, \theta) + \sum_{j,k} \theta_j \frac{\partial F_j(x(t))}{\partial x_k} \frac{\partial \psi(t, \theta)}{\partial \theta_k}. \]

The scaled fluctuations \( Z_N(t) \) about the deterministic trajectory can thus be approximated by a Gaussian distribution with mean \( \mathbb{E}(Z(t)) = M(t)z \) and covariance matrix \( \text{Cov}(Z(t)) = \Sigma(t) \), where

\[ \Sigma(t) = M(t) \left( \int_0^t M(u)^{-1} G(x(u, x_0))(M(u)^{-1})^T du \right) M(t)^T \]

and

\[ M(t) = \exp \left( \int_0^t H(u) du \right). \]

Notice in particular that \( \text{Cov}(X_N(t)) \approx \Sigma(t)/N \) for large \( N \). The covariance matrix can only be evaluated explicitly in some simple cases. However, since populations are often observed in equilibrium, it makes sense to assume that the initial value \( x_0 \) is a fixed point \( x^* \) of \( F \), in which case we may appeal to Theorem 2.8 which can give information about the fluctuations around \( x^* \). The approximating diffusion is now an Ornstein-Uhlenbeck (OU) process.
Theorem 3.13. Assume the conditions of Theorem 3.12 hold and
\[
\lim_{N \to \infty} \sqrt{N} (X_N(0) - x^*) = z.
\]

The family of processes \( \{Z_N(t)\} \), defined by
\[
Z_N(s) = \sqrt{N} (X_N(s) - x^*), \quad 0 \leq s \leq t,
\]
converges weakly in \( D[0, t] \) to an OU process \( Z(t) \) with initial value \( Z(0) = z \), local drift matrix \( H = \nabla F(x^*) \) and local covariance matrix \( G(x^*) \). \( Z(t) \) follows a Gaussian diffusion with mean \( \mu(t) = e^{Ht} z \) and covariance matrix
\[
\Sigma(t) = e^{Ht} \left( \int_0^t e^{-Hs} G(x^*) e^{-HTs} ds \right) e^{HTt}.
\]

3.6 Discussion

We have proposed a structured metapopulation model that incorporates heterogeneous within patch dynamics and spatial structure, and identified conditions under which the metapopulation persists or goes extinct. As the extinction state is absorbing when no immigration from outside the system is assumed, we were not able to identify these conditions by identifying a stationary distribution for the model. Instead, we based the analysis on a dynamical system (3.6) that approximates the stochastic density process when the population ceiling is large.

The theorems presented in this chapter combine to give a detailed description of the long term behaviour of the model. Under Assumption (F), the long term behaviour of the system is completely described; inequality (3.9) determines the extinction or persistence of the metapopulation. When Assumption (F) does not hold, however, the metapopulation may display more complex behaviour.

The discovery of an Allee effect in the two–patch metapopulation model is unexpected as the birth rate functions for each patch are strictly decreasing. It appears that the Allee effect arises as a result of a large variation in the migration rates which depends on the population density of the patches. For a metapopulation where the two
patches are of greatly different sizes, a high population density on the smaller patch allows the population to become established on the larger patch as emigration from the larger patch is reduced to a very low level. It is to be noted that the presence of an Allee effect in a metapopulation has important implications for the design of conservation strategies. Conservation targets need to be set taking into account the critical threshold below which the metapopulation goes extinct [37, section 5.1.4]. Similarly, for the successful reintroduction of a species, the release size needs to be sufficiently large for the population density to exceed the critical threshold [44]. The Allee effect can also be exploited to create more efficient strategies for managing invasive pests [123].

We also discovered the possibility of three nonzero fixed points in the two patch model when the zero fixed point is unstable. For the $J$ patch system, there are $2^J - 1$ possible nonzero fixed points, which, if all are contained in $E$, would result in the Allee effect occurring at many intermediate population sizes. Although the perturbation analysis conducted for the two–patch model could be generalised to larger number of patches, the expressions involved quickly become cumbersome. Thus, a different approach will be required to improve our understanding of the Allee effect in this model.

The appropriateness of the conclusions presented is dependent on how well the stochastic model (3.1a) – (3.1c) is approximated by the dynamical system (3.6). It was mentioned briefly in Section 3.3 that the accuracy could be quantified using the results of [42]. However, it is important to note that, in general, the accuracy of the dynamical systems approximation deteriorates near the extinction state [14, 103]. Therefore, alternative methods might be needed to improve our understanding of the stochastic metapopulation model near extinction.

One aspect of population dynamics that our model excludes, but has been incorporated in other models [19, 99, 7], is the potential for catastrophes. A catastrophe occurs when every individual on a given patch dies instantly or when all individuals on a patch are subject to a higher death rate over some small period of time. An obvious extension then would be to include catastrophes. Doing so would increase the chance an individual dies, hence the birth rate required for the metapopulation to survive
would naturally have to be higher. One way to introduce catastrophes into the model would be to randomly switch between two sets of parameters where one set includes a much higher death rate. However, if catastrophes were introduced, a deterministic approximation could not be used to study the stochastic model, since catastrophes are inherently random events and affect a large number of individuals. These two aspects are not approximated well by a deterministic system, even in the limit as population ceiling $N$ gets large. However, the behaviour of the metapopulation between catastrophes would remain unchanged. A piecewise deterministic approximation to a model including catastrophes might be obtained using the functional limit laws of Franz et al. [50], which is addressed in Chapter 4.
Chapter 4

Environmental Influence

The evolution of a population is governed by how individuals interact with other individuals in the population and how they interact with the environment. Previously, the environment was assumed to be static. However, this excludes important changes such as breeding season, cold snaps and catastrophes. To better understand a population’s evolution, the model from Chapter 3 is enriched to account for dynamic environments.

4.1 Introduction

In Chapter 3, a metapopulation model that accounts for within patch dynamics was examined. However, the environment in that model was assumed static and did not allow any temporal variation in the parameters. Such environmental variation is very common in populations. For example, many species of birds exhibit breeding seasons. A breeding season is an interval of time where the rate of births is higher than at other times and usually repeats every year. The time the season starts and the length of the season depends on the species itself and its geographic location but the periodicity of the season remains unchanged. However, there are minor exceptions to this rule such as two breeding seasons in one year, no distinct breeding season, a breeding season that occurs during periods of rain and the Wideawake Sterna fuscata, which is reported to have a 9 month period between breeding season.

Another environmental influence is temperature. The temperature can affect the population through an increased death rate when it becomes too hot or cold for a
species. However, this is not the only type of influence temperature has. Species tend to migrate away from extreme temperatures, be it hot or cold, and therefore a sudden decrease in the population on one patch could be attributed to this migration, rather than deaths. However, a change in temperature may not always be the cause of this increased migration, with some species exhibiting migratory seasons regardless of the temperatures.

However, not all environmental influences follow a largely predicted pattern. Seasons can be predicted fairly accurately. Temperatures, whilst having a higher incidence of variation, can also be predicted with some accuracy. On the other hand, catastrophes are largely random and hence a deterministic approach to include such factors would be ineffective. Wild fires are an example of such a catastrophe. The chance a fire occurs increases with temperature, along with various other factors, however this process is still stochastic.

The examples given demonstrate how changes in the environment will affect the evolution of a population. The long term behaviour of the population will also be affected if a dynamic environment is added. Furthermore, questions regarding the changing environment, such as the influence breeding and migratory seasons have on the population and how a population handles catastrophes, can be answered. In this chapter, models accompanied with analyses are proposed, that can contribute to answering these questions more effectively. The models presented account for variation in the environment that is deterministic in Section 4.2 and stochastic in Section 4.3.

4.2 Deterministic Influences

The first type of environmental influence to be introduced into the model will be deterministic. As previously mentioned, breeding seasons are an example of such an influence and many species exhibit some form of breeding season. And there are many more examples of deterministic influences such as migratory seasons for birds, sharks, elephants, butterflies and some types of bears, snakes, turtles, bees and bats hibernate throughout certain parts of the year, which decreases their migration and breeding rates for a period. Lastly, harsh temperatures in winter can increase the death rate of
some insects. To account for these dynamics, the parameters of the model presented in Chapter 3 need to be allowed to vary in a certain predictable way during the year.

### 4.2.1 Model

The model presented in Chapter 3 is extended by allowing the immigration, migration, birth and death rates as well as the connectivity parameters to depend on time. Therefore, the new model will be a time inhomogeneous process. Let $J$ be the number of patches that can be occupied in the metapopulation and $n_i(t)$ be the number of individuals occupying patch $i$ at time $t$ and define $n(t) := \{n_1(t), \ldots, n_J(t)\}$. The Markov population process (see Definition 2.2) $(n(t), t \geq 0)$ takes values in the state space $S_N := \{0, \ldots, N_1\} \times \cdots \times \{0, \ldots, N_J\}$, where $N := \sum_i N_i$ and has nonzero transition rates

\begin{align*}
\lambda_i(t, n_i) &= \nu_i(t)(N_i - n_i) + n_i b_i \left( t, \frac{n_i}{N_i} \right), \quad (4.1a) \\
\mu_i(t, n_i) &= \phi_i(t, n_i) \lambda_{i0}(t) + d_i(t)n_i, \quad (4.1b) \\
\gamma_{ij}(t, n_i, n_j) &= \phi_i(t, n_i) \lambda_{ij}(t) \frac{N_j - n_j}{N_j} \quad \text{for all } j \neq i. \quad (4.1c)
\end{align*}

These rates correspond to: an increase on patch $i$ due to a birth (4.1a), a decrease on patch $i$ due to a death or removal from the system (4.1b) and a migration from patch $i$ to patch $j$ (4.1c), all at time $t$. The functions $d_i(t)$, $\lambda_{ij}(t)$ and parameters $N_i$ are the per–capita death rate function, proportion of individuals migrating from patch $i$ to patch $j$ (or out of the system if $j = 0$) at time $t$ and the population ceiling for patch $i$, respectively. The birth rate function $b_i(t, \cdot)$ determines the per–capita birth rate at time $t$ given how densely populated patch $i$ is. The migration function $\phi_i(t, \cdot)$ represents the rate at which individuals leave patch $i$ at time $t$. Figure 4.1 illustrates the transitions (4.1).

### 4.2.2 Differential Equation Approximation

As discussed in Section 2.3.4, Kurtz’ functional limit laws [76] extend to the time inhomogeneous case. To apply the results of Pagendam & Pollett [96] (Theorem 2.10), the rates need to be written in the form given in Definition 2.9. The transition rates
Figure 4.1: Illustration of the dynamics for patch $i$ and migration to and from patch $j$.

Equation (4.1) can be written in the form

$$q_i(n, n + l) = N f^{(N)} \left( t, \frac{n}{N}, l \right),$$

where

$$f^{(N)}(t, x, l) = \begin{cases} 
\nu_i(t)(M_{i}^{(N)} - x_i) + x_i b_i \left( t, \frac{x_i}{M_{i}^{(N)}} \right) & \text{if } l = e_i, \\
\phi_i^{(N)}(t, x_i) \lambda_{i0}(t) + d_i(t) x_i & \text{if } l = -e_i, \\
\phi_i^{(N)}(t, x_i) \lambda_{ij}(t) \left( 1 - \frac{x_j}{M_{j}^{(N)}} \right) & \text{if } l = -e_i + e_j, \\
0 & \text{otherwise},
\end{cases}$$

and the functions $\hat{\phi}_i^{(N)} : [0, \infty) \times [0, M_{i}^{(N)}] \to \mathbb{R}_+$ satisfy

$$\hat{\phi}_i^{(N)} \left( t, \frac{n}{N} \right) = \frac{\phi_i(t, n)}{N}.$$

Define $F^{(N)}(t, x) := \sum_l l f^{(N)}(t, x, l)$. Due to the dependence on $N$ in $f^{(N)}(t, x, l)$, the Markov process $(n(t), t \geq 0)$ is not density dependent in time according to Definition 2.9. Therefore, it is necessary to extend Definition 2.9 to include a new type of process; an asymptotically density dependence process.

**Definition 4.1.** A family of Markov processes \{n^{(N)}(t)\} indexed by $N > 0$ (with a state space $S_N \subset \mathbb{Z}^J$) is said to be “asymptotically density dependent in time” if there exists a continuous function, $f^{(N)} : [0, \infty) \times E \times \mathbb{Z}^J \to \mathbb{R}$, where $E \subseteq \mathbb{R}^J$, such that the


transition rates of $n^{(N)}(t)$ are given by

$$q_t(n^{(N)}, n^{(N)} + l) = N f^{(N)}(t, \frac{n^{(N)}}{N}, l), \quad l \neq 0,$$

and $F^{(N)}(t, x) := \sum_l l f^{(N)}(t, x, l)$ converges uniformly over $[0, \infty) \times E$ to $F(t, x)$ as $N \to \infty$.

The process with rates (4.1) satisfies Definition 4.1 under some additional mild assumptions. Assume that $M_i^{(N)} \to M_i$ and there exists functions $\hat{\phi}_i : S_i \mapsto \mathbb{R}_+$, where $S_i := [0, \infty) \times [0, M_i]$ satisfying

$$\lim_{N \to \infty} \sup_{(t, x) \in S_i} \left| \hat{\phi}_i^{(N)}(t, x) - \hat{\phi}_i(t, x) \right| = 0, \quad \text{for all } i,$$

and also functions $b_i : [0, \infty) \times [0, 1] \mapsto \mathbb{R}_+$ satisfying

$$\lim_{N \to \infty} \sup_{(t, x) \in S_i} \left| b_i\left(t, \frac{x_i}{M_i^{(N)}}\right) - b_i\left(t, \frac{x_i}{M_i}\right) \right| = 0, \quad \text{for all } i.$$

Then the process with rates (4.1) is asymptotically density dependent in time according to Definition 4.1 if the totally population ceiling, $N$, is the index parameter. Just as the definition of density dependence was extended to include asymptotic density dependence, Theorem 2.9 needs to be extended to also include asymptotic density dependent processes. Informally, as $N$ gets large, the density process, $X^{(N)}(t) := n^{(N)}(t)/N$ converges to the solution of a differential equation. Formally, this result is presented in Theorem 4.2.

**Theorem 4.2.** Let $(n^{(N)}(t), t \geq 0)$, indexed by $N$, be asymptotically density dependent in time, $F$ Lipschitz,

$$\sup_{x \in E} \sum_l f^{(N)}(t, x, l) < \infty, \quad \text{for all } t > 0, N \geq 1 \quad \text{and} \quad (4.4)$$

$$\sup_{x \in E} \sum_l \|l\|^2 f^{(N)}(t, x, l) < \infty \quad \text{for all } t > 0, N \geq 1. \quad (4.5)$$
Chapter 4: Environmental Influence

If \( n^{(N)}(0)/N \to x_0 \) in probability as \( N \to \infty \), then for any \( \varepsilon > 0 \),

\[
\lim_{N \to \infty} \mathbb{P} \left( \sup_{s \leq t} \left| \frac{n^{(N)}(s)}{N} - x(s) \right| > \varepsilon \right) = 0, \tag{4.6}
\]

where \( x(t, x_0) \) is the solution to

\[
\frac{dx}{dt} = F(t, x), \quad x(0) = x_0.
\]

**Proof.** Define \( X^{(N)}(t) := n^{(N)}(t)/N \). Under the conditions of Theorem 7.3 of [47], we may express \( X^{(N)}(t) \) as

\[
X^{(N)}(t) = X^{(N)}(0) + M^{(N)}(t) + \int_0^t F^{(N)}(s, X^{(N)}(s)) \, ds \tag{4.7}
\]

where \( M^{(N)}(t) \) is a martingale w.r.t. the filtration \( \mathcal{F}^{(N)}_t = \sigma \{ X^{(N)}(s), 0 \leq s \leq t \} \). To verify the conditions of Theorem 7.3 of [47] hold, set

\[
\mu(t, x, y) = \frac{N f^{(N)}(t, x, y - x)}{\lambda(t, x)} \quad \text{and} \quad \lambda(t, x) = N \sum_l f^{(N)}(t, x, l)
\]

and define \( \gamma(t) := N \sup_{x \in E} \sum_l f^{(N)}(t, x, l) \). It then can be seen that \( \lambda(t, x) \leq \gamma(t) \) for all \( t \geq 0 \). Therefore, the conditions of Theorem 7.3 of [47] are satisfied if \( \int_0^t \gamma(s) \, ds < \infty \) for all \( t > 0 \), which is true for all finite \( N \), due to (4.4). As such,

\[
X^{(N)}(t) - x(t) = X^{(N)}(0) - x_0 + \int_0^t F^{(N)}(s, X^{(N)}(s)) \, ds - \int_0^t F(s, x(s)) \, ds + M^{(N)}(t).
\]

Therefore, recalling the Lipschitz condition on \( F \),

\[
\left| X^{(N)}(t) - x(t) \right| \leq \left| X^{(N)}(0) - x_0 \right| + \int_0^t \left| F^{(N)}(s, X^{(N)}(s)) - F(s, X^{(N)}(s)) \right| \, ds + L \int_0^t \left| X^{(N)}(s) - x(s) \right| \, ds + \left| M^{(N)}(t) \right| \tag{4.8}
\]

From Definition [4.1] it is known that \( F^{(N)}(t, x) \) converges uniformly in \( t \) and \( x \) over \([0, \infty) \times E \) to \( F(t, x) \), thereby implying that for any \( \delta > 0 \) there exists an \( N \) sufficiently large such that \( \sup_{x \in E, s > 0} \left| F^{(N)}(s, x) - F(s, x) \right| < \delta \). On the event \( \{ \sup_{s \leq t} \left| M^{(N)}(s) \right| \leq \delta \} \),
equation (4.8) becomes
\[
\sup_{s \leq t} |X^{(N)}(s) - x(s)| \leq |X^{(N)}(0) - x_0| + \delta(1 + t) + L \int_0^t |X^{(N)}(s) - x(s)| \, ds.
\]

Therefore, applying Gronwall’s inequality and for sufficiently large \( N \),
\[
\sup_{s \leq t} |X^{(N)}(s) - x(s)| \leq \left( |X^{(N)}(0) - x_0| + \delta(1 + t) \right) e^{Lt}.
\]

Confirming (4.6) now amounts to finding a bound for \( \mathbb{P}( \sup_{s \leq t} |M^{(N)}(s)| > \delta ) \). Let \( M^*(t) := \sup_{0 \leq s \leq t} |M^{(N)}(t)| \). Markov’s inequality gives
\[
\mathbb{P}(M^*(t) > \varepsilon) \leq \frac{\mathbb{E}(M^*(t)^2)}{\varepsilon^2}.
\]

Furthermore, by Proposition 2.16 of [47] with Corollary 3 of Theorem 27 of Chapter 2 of [107] gives
\[
\mathbb{E}(M^*(t)^2) \leq 4\mathbb{E}(M^{(N)}(t)^2) = 4\mathbb{E}([M]_t),
\]
where \([M]_t\) is the quadratic variation of \( M^{(N)}(t) \). The quadratic variation of \( M^{(N)}(t) \) is equivalent to the quadratic variation of \( X^{(N)}(t) \) due to the piecewise constant paths of \( X^{(N)}(t) \). As such we first determine an expression for \( (X^{(N)}(t))^2 \) in the same fashion as (4.7) which gives
\[
(X^{(N)}(t))^2 = X^{(N)}(0)^2 + \hat{M}^{(N)}(t) + \frac{1}{N} \int_0^t \sum_i i^2 f^{(N)}(r, X^{(N)}(r), l) \, dr
\]
\[
+ 2 \int_0^t X^{(N)}(r) F^{(N)}(r, X^{(N)}(r)) \, dr,
\]
for some martingale \( \hat{M}^{(N)}(t) \) w.r.t. the filtration \( \mathcal{F}^{(N)}_t = \sigma\{(X^{(N)}(s))^2, 0 \leq s \leq t\} \). Then, following Theorem 22 of Chapter 2 of [107], for a sequence \( \sigma_n \) which is the sequence \( 0 = T_0^{(n)} \leq T_1^{(n)} \leq \cdots \leq T_k^{(n)} \) where \( T_i^{(n)} \) are stopping times, we have
\[
[M]_t = [X]_t = \lim_{n \to \infty} \sum_i \left(X^{(N)}\left(T_{i+1}^{(n)}\right) - X^{(N)}\left(T_{i}^{(n)}\right)\right)^2.
\]
Using (4.11), we have

\[
\left( X^{(N)}(t) - X^{(N)}(s) \right)^2 = \\
\hat{M}^{(N)}(t) - \hat{M}^{(N)}(s) + \frac{1}{N} \int_s^t \sum_i l^2 f^{(N)}(r, X^{(N)}(r), l) \, dr \\
- 2X(s) \left( M^{(N)}(t) - M^{(N)}(s) \right) + 2 \int_0^t X^{(N)}(r) F^{(N)}(r, X^{(N)}(r)) \, dr \\
- 2X^{(N)}(s) \int_s^t F^{(N)}(r, X^{(N)}(r)) \, dr.
\]

Then, by Assumption (4.5),

\[
\lim_{n \to \infty} \sum_i X^{(N)}(T^{(n)}_{i+1}) \left( M^{(N)}(T^{(n)}_{i+1}) - M^{(N)}(T^{(n)}_i) \right) = \int_0^t X^{-}(r) \, dM^{(N)}(r),
\]

and

\[
\lim_{n \to \infty} \sum_i X^{(N)}(T^{(n)}_{i+1}) \int_{T^{(n)}_i}^{T^{(n)}_{i+1}} F^{(N)}(r, X^{(N)}(r)) \, dr = \int_0^t X^{(N)}(r) F^{(N)}(r, X^{(N)}(r)) \, dr.
\]

Thus, combining (4.12), (4.13) and (4.14), the quadratic variation for \( M^{(N)}(t) \) is given by

\[
[M]_t = \hat{M}^{(N)}(t) - \hat{M}^{(N)}(0) + \frac{1}{N} \int_0^t \sum_i l^2 f^{(N)}(r, X^{(N)}(r), l) \, dr \\
- 2 \int_0^t X^{-}(r) \, dM^{(N)}(r) + 2 \int_0^t X^{(N)}(r) F^{(N)}(r, X^{(N)}(r)) \, dr \\
- 2 \int_0^t X^{-}(r) F^{(N)}(r, X^{(N)}(r)) \, dr.
\]

Furthermore, applying Theorem 15.12 of [66] gives

\[
\mathbb{E} \left( \int_0^t X^{-}(r) \, dM^{(N)}(r) \right) = \int_0^t X^{(N)}(r) \, dM^{(N)}(r) = 0.
\]

Then, by Assumption (4.5), \( \sum_i l^2 f^{(N)}(s, x, l) \leq B \) for all \( x \in E \), \( s \geq 0 \) for some \( B > 0 \).
Chapter 4: Environmental Influence

giving

\[ \mathbb{E}([M]_t) = \frac{1}{N} \mathbb{E} \left( \int_0^t \sum_i i^2 f^{(N)}(r, X^{(N)}(r)) \, dr \right) \leq \frac{Bt}{N}. \]  

(4.15)

Combining (4.9), (4.10) and (4.15) gives

\[ \mathbb{P} \left( \sup_{s \leq t} |M^{(N)}(s)| > \delta \right) \leq \frac{4Bt}{N\delta^2}. \]  

(4.16)

Using (4.16), it can be seen that for any \( \varepsilon > 0 \), there is a \( \delta(\varepsilon) > 0 \) such that

\[ \mathbb{P} \left( \sup_{s \leq t} |X^{(N)}(s) - x(s)| > \varepsilon \right) \leq \frac{4Bt}{N\delta^2} + \mathbb{P} \left( |X^{(N)}(0) - x_0| > \varepsilon \right), \]

which results in (4.6) if \( \mathbb{P} \left( |X^{(N)}(0) - x_0| > \varepsilon \right) \to 0 \) as \( N \to \infty \), completing the proof.

Therefore, when conditions (4.2) and (4.3) are satisfied, the density process defined by \( X^{(N)}(t) := n(t)/N \) converges in probability over finite time intervals, as the population ceiling increases, to the solution \( x(t, x_0) \) of the deterministic model

\[ \frac{dx(t, x_0)}{dt} = F(t, x(t, x_0)), \quad x(0, x_0) = x_0, \]  

(4.17)

where \( F(t, x) \) has elements

\[ F_i(t, x) = M_i \nu_i(t) - (\nu_i(t) + d_i(t) + \phi_i(t)) x_i + x_i b_i \left( t, \frac{x_i}{M_i} \right) \]

\[ + \sum_{j \neq i} \left( \phi_j(t) x_j \lambda_{ji}(t) + \left( \phi_i(t) x_i \lambda_{ij}(t) - \phi_j(t) x_j \lambda_{ji}(t) \frac{x_i}{M_i} \right) \right). \]

More precisely, we have the following result.

**Theorem 4.3.** Suppose there are bounded, continuous functions \( b_i, d_i, \nu_i \) and \( \phi_i \) for all \( i \). If \( X_N(0) \to x_0 \) as \( N \to \infty \), then, for every \( t > 0 \) and \( \delta > 0 \),

\[ \lim_{N \to \infty} \mathbb{P} \left( \sup_{s \leq t} |X_N(s) - x(s, x_0)| > \delta \right) = 0, \]

where \( x(t, x_0) \) satisfies (4.17).
4.2.3 Equilibrium Behaviour

Theorem 4.3 provides a way to approximate the behaviour of the density process by a differential equation and therefore, the analysis of the differential equation (4.17) will provide information about the long term behaviour of the density process. To determine this behaviour, the following assumptions about the dynamics of the system are imposed:

1. The migration functions are linear; that is \( \phi_i(t, x) = \phi_i(t)x \) for all \( i \), and \( \phi_i(t) > 0 \) for all \( i \) and \( t \geq 0 \) (implying that \( \hat{\phi}_i(t, x) = \phi_i(t)x \) for all \( i \)).

2. The functions \( b_i(t, \cdot), d_i(t), \phi_i(t) \) and \( \lambda_{ij}(t) \) are \( 2\pi \)-periodic in \( t \) for all \( i \).

3. For all \( i \) and \( j \) and \( t \in \mathbb{R}_+ \), there is a finite sequence \( (a_k) \) such that

\[
\lambda_{ia_1}(t)\lambda_{a_1a_2}(t)\ldots\lambda_{a_ma_j}(t) \neq 0.
\]

4. For all \( i \) and \( t \in \mathbb{R}_+ \), \( b_i(t, x) \) is a continuously differentiable, strictly decreasing function on \([0, 1]\) such that \( b_i(\cdot, 1) = 0 \) and \( xb_i(\cdot, x) \) is strictly concave on \([0, 1]\).

5. The functions \( \phi_i(t), \lambda_{ij}(t) \) and \( M_i \) satisfy \( \phi_i(t)\lambda_{ij}(t)M_i = \phi_j(t)\lambda_{ji}(t)M_j \) for all \( i,j \).

6. There is no external immigration to any patch, that is, \( \nu_i = 0 \) for all \( i \).

With the exception of Assumption (2), these assumptions are natural generalisations of those employed in Section 3.4. Assumption (2) assumes all dynamics have the same period, which agrees with studies showing that species will often exhibit some sort of annual cycle for breeding and migration [24, 125]. As in Chapter 3, define \( E := [0, M_1] \times \cdots \times [0, M_J] \).

**Theorem 4.4.** Assume (1)-(6) hold. Let \( \Phi(t) \) be the transition matrix for the linear ODE \( \frac{dz}{dt} = \nabla_x F(t, 0) z \), and define \( \gamma_i \) as the eigenvalues of \( \Phi(2\pi) \). Then if

\[
\gamma_i > 1, \quad \text{for any } i,
\]

(4.18)
the ODE (4.17) possesses a unique nonzero $2\pi$-periodic solution $q(t) \in E$, which is such that $\lim_{t \to \infty} x(t, x_0) = q(t)$ for all $x_0 \in E \setminus \{0\}$. Otherwise, $\lim_{t \to \infty} x(t, x_0) = 0$ for all $x_0 \in E$.

If (6) does not hold, then there exists a unique nonzero $2\pi$-periodic solution $q(t) \in E$.

Proof. The proof begins with verifying the conditions Theorem 3.1 of [119]. Under Assumption (2), $F(t, x)$ is also $2\pi$-periodic. If $x \geq 0$ with $x_i = 0$, then

$$F_i(x, t) = M_i \nu_i(t) + \sum_{j \neq i} \phi_j(t) x_j \lambda_{ji}(t) \geq 0.$$ 

Under Assumption (1) and (3) it is known that $\nabla_x F(t, x)$ exists, is continuous in $\mathbb{R} \times E$ and is irreducible for each $(t, x) \in \mathbb{R} \times E$. Furthermore, under Assumption (5), $\nabla_x F(t, x)$ has elements given by

$$\nabla_x F(t, x) = \begin{cases} \phi_j(t) \lambda_{ji}(t) & \text{if } i \neq j, \\ b_i(t) \left(1 - \frac{2 x_i}{M_i}\right) - d_i(t) - \phi_i(t) & \text{if } i = j. \end{cases} \quad (4.19)$$

If $i \neq j$, then $\nabla_x F(t, x)_{ij} = \nabla_x F(t, y)_{ij}$ for any $x, y$. Also, if $i = j$ and $y_i > x_i$, then $\nabla_x F(t, x)_{ii} - \nabla_x F(t, y)_{ii} = \frac{2b_i(t)}{M_i} (y_i - x_i) > 0$ from Assumption (4). Therefore $\nabla_x F(t, x) \geq \nabla_x F(t, y)$ element-wise, with an inequality for at least element. Finally, under Assumption (6), $F(0, t) = 0$.

Therefore, Theorem 3.1 of [119] may be applied, which states that if (4.18) is satisfied, then provided $x_0 \in E \setminus \{0\}$, the solution $x(t, x_0)$ will do one of two things; (a) $x(t, x_0) \to \infty$ as $t \to \infty$ or (b) $x(t, x_0) \to q(t)$, where $q(t)$ is a unique nonzero $2\pi$-periodic solution to (4.17). If it can be shown that the process $x(t, x_0)$ is bounded, then (a) cannot be true and hence (b) must be true. Considering the trajectory starting at $M$, that is $x(t, M)$. The elements of $F(t, M)$ are given by

$$F_i(M, t) = - (d_i(t) + \phi_i(t) \lambda_{i0}(t)) M_i < 0.$$ 

Therefore, for sufficiently small $s$, $x(s, M) < M$. Then according to Lemma 3.2 of [119], $x(s, x(s, M)) := x(2s, M) < x(s, M)$, implying that $x(ns, M) < M$, where $n = 1, 2, \ldots$. Lemma 3.1 of [119] gives $0 \leq x(t, 0)$ for all $t > 0$. Therefore, $0 \leq x(t, x_0) \leq M$ for all
Chapter 4: Environmental Influence

$t > 0$ if $0 \leq x_0 \leq M$, showing $E$ is an invariant set, thereby implying all solutions to (4.17) remain bounded and $q(t) \in E$.

If (4.18) is not true, then according to Theorem 3.1 of [119], $x(t) \to 0$ as $t \to \infty$.

Conversely, if Assumption (6) does not hold then $F(0, t) > 0$ and $E$ remains invariant. As such, all solutions are bounded, implying $x(t, x_0) \to q(t)$, where $q(t) \in E$ which completes the proof.

Theorem 4.4 shows under what conditions the metapopulation will persist and otherwise go extinct. Condition (4.18) determines which outcome the metapopulation will have, however it is rather hard to interpret. The following examples will address this issue while still maintaining some generality.

Example 1. The first example is a two patch system where only the birth rates differ between patches ($d_1 = d_2 = d$, $\phi_1 = \phi_2 = \phi$ but $\lambda_{12} \neq \lambda_{21}$). Note that Assumption (5) is required to hold and, as such, the relation $M_1\lambda_{12} = M_2\lambda_{21}$ is assumed to hold. There will be a higher birth rate on patch one which is designed to model a species where individuals migrate to patch one to reproduce, while a smaller reproduction rate will occur on patch two. The birth rate functions will be $b_1(t, x/M_1) = b_1(1 - \sin(t))(1 - x/M_1)$ and $b_2(t, x/M_2) = b_2(1 - \sin(t))(1 - x/M_2)$, where $b > 0$ and $b_1 > 0$ are constants. These type of functions will model a breeding season in part of the year. Under these conditions, (1) - (6) are satisfied and as such, Theorem 4.4 can be applied. The linear ODE can then be written as $\frac{dz}{dt} = A(t)z$ where

$$A(t) := \nabla_x F(t, 0) = \begin{pmatrix} b(b_1 + 1 - \sin(t)) - d - \phi & \phi\lambda_{21} \\ \phi\lambda_{12} & b(1 - \sin(t)) - d - \phi \end{pmatrix}.$$ 

Due to the fact that $A(t) \left( \int_0^t A(s)ds \right) = \left( \int_0^t A(s)ds \right) A(t)$, the solution matrix $\Phi(t)$ can be expressed as $\Phi(t + 2\pi) = \Phi(t) \exp(2\pi B)$, where $B = (1/2\pi) \int_0^{2\pi} A(s)ds$ and $\exp$ is the matrix exponential (see the proof of Theorem 2.42 of [67]). Therefore the eigenvalues of $\Phi(2\pi) = \exp(2\pi B)$ are given by $\gamma_i = e^{2\pi \sigma_i}$ for $i = 1, 2$ where

$$\sigma = \begin{pmatrix} b - d - \phi + \frac{bb_1}{2} - \frac{\sqrt{b^2b_1^2 + 4\lambda_{12}\lambda_{21}\phi^2}}{2} \\ b - d - \phi + \frac{bb_1}{2} + \frac{\sqrt{b^2b_1^2 + 4\lambda_{12}\lambda_{21}\phi^2}}{2} \end{pmatrix}.$$
Figure 4.2: In A–C, the population density for three sets of parameters are plotted, where the thick lines are the deterministic trajectory given by (4.17), the thin lines are one realisation of the stochastic process detailed in Section 4.2.1. In D–F, the birth rate per-capita functions are plotted against time. In all plots, blue lines represent patch one and green lines represent patch two. The parameters used for A & D are $b = 2.1$, $b_1 = 1$, $d = 1$, $\phi = 1$, $M_1 = 0.75$, $M_2 = 0.25$, $\lambda_{12} = 0.2$ and $\lambda_{21} = 0.6$. The parameters used for B & E are $b = 1$, $b_1 = 1$, $d = 0.5$, $\phi = 1.5$, $M_1 = 0.75$, $M_2 = 0.25$, $\lambda_{12} = 1/3$ and $\lambda_{21} = 1$. And the parameters used for C & F are $b = 1$, $b_1 = 1$, $d = 1$, $\phi = 1$, $M_1 = 0.75$, $M_2 = 0.25$, $\lambda_{12} = 0.2$ and $\lambda_{21} = 0.6$.

Therefore, (4.18) is satisfied if either

$$b \left(1 + \frac{b_1}{2}\right) - d - \phi > 0,$$

or

$$\frac{(d + \phi - b)(d + \phi - b(1 + \frac{b_1}{2}))}{\phi^2 \lambda_{12} \lambda_{21}} < 1,$$

which are similar conditions to condition (3.24), given for the autonomous system. Therefore, if either of the conditions (4.20) is satisfied the metapopulation will persist, otherwise it will go extinct. Figure 4.2 illustrates these three cases. In A & D, condition (4.20a) holds while in B & E, condition (4.20b) holds. And finally, in C & F neither condition holds.
The method to find the eigenvalues for $\Phi(2\pi)$ used in Example 1 holds more generally, for heterogeneous patches and an arbitrarily sized system as long as $\nabla_x F(t,0)$ and $\int_0^t \nabla_x F(s,0) ds$ commute. In this context, these two matrices commute whenever all of the diagonal elements of $\nabla_x F(t,0)$ have constant differences for all $t$ and the off diagonal elements are constant and greater than zero. The difficulty arises in determining tractable analytical expressions for the eigenvalues of $\Phi(2\pi)$. However, this difficulty disappears if the eigenvalues are determined numerically. This is the aim of second example below.

**Example 2.** Assume there are five patches with the parameters given in Table 4.1, where the birth rate functions are given by $b_i(t,x) = (b_i + 1 - \sin(t))(1 - x)$, $b_i > 0$ is constant and all other functions are constant in time.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$i = 1$</th>
<th>$i = 2$</th>
<th>$i = 3$</th>
<th>$i = 4$</th>
<th>$i = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_i$</td>
<td>0.5</td>
<td>0.25</td>
<td>0.1</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>$b_i$</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$d_i$</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$\varphi_i$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\lambda_{ij}$</th>
<th>$j = 1$</th>
<th>$j = 2$</th>
<th>$j = 3$</th>
<th>$j = 4$</th>
<th>$j = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i = 1$</td>
<td>0</td>
<td>0.16</td>
<td>0.094</td>
<td>0.1042</td>
<td>0.156</td>
</tr>
<tr>
<td>$i = 2$</td>
<td>0.32</td>
<td>0</td>
<td>0.08</td>
<td>0.0760</td>
<td>0.32</td>
</tr>
<tr>
<td>$i = 3$</td>
<td>0.47</td>
<td>0.2</td>
<td>0</td>
<td>0.27</td>
<td>0.06</td>
</tr>
<tr>
<td>$i = 4$</td>
<td>0.21</td>
<td>0.19</td>
<td>0.27</td>
<td>0</td>
<td>0.15</td>
</tr>
<tr>
<td>$i = 5$</td>
<td>0.39</td>
<td>0.4</td>
<td>0.03</td>
<td>0.075</td>
<td>0</td>
</tr>
</tbody>
</table>

Then the linear ODE is given $\frac{dz}{dt} = A(t)z$ where

$$A(t) := \nabla_x F(t,0) = \begin{pmatrix}
-1 - \sin(t) & 0.32 & 0.47 & 0.21 & 1.56 \\
0.16 & -\sin(t) & 0.2 & 0.19 & 1.6 \\
0.094 & 0.08 & -1 - \sin(t) & 0.27 & 0.12 \\
0.042 & 0.076 & 0.27 & -\sin(t) & 0.3 \\
0.156 & 0.32 & 0.06 & 0.15 & -3 - \sin(t)
\end{pmatrix}.$$  

Again the equality $A(t) \left( \int_0^t A(s) ds \right) = \left( \int_0^t A(s) ds \right) A(t)$ holds, and therefore the relation $\Phi(2\pi) \equiv \exp(2\pi B)$ can be used where $B = (1/2\pi) \int_0^{2\pi} A(s) ds = A(0)$ in this case.
Therefore the eigenvalues of $\Phi(2\pi) \equiv \exp(2\pi B)$ are given by $\gamma_i = e^{2\pi \sigma_i}$ for $i \in \{1, \ldots, 5\}$ where $\sigma_i$ are the eigenvalues of $B$, of which the largest is approximately 0.42, implying that $\gamma_i > 1$ for at least one $i$ and then, by Theorem 4.4 implies the ODE (4.17) possesses a unique nonzero $2\pi$-periodic solution $q(t) \in E$ and, more importantly, $x(t) \not\rightarrow 0$.

4.2.4 Discussion

We have seen how the model from Chapter 3 can be extended to account for a deterministically changing environment. To do this, the parameters of the system have been allowed to vary temporally. To determine the long term behaviour of the process, a new functional limit law was derived which accounts for a larger family of processes that are not only time dependent but also asymptotically density dependent. Using this result, it was shown that the metapopulation model can be approximated by a nonautonomous ODE for large $N$. The analysis of the ODE resulted in a necessary and sufficient condition for extinction.

Utilising condition (4.18) can be quite difficult as it involves determining the eigenvalues of the transition matrix $\Phi(2\pi)$, which is not always easy to calculate. Some methods exist for calculating the entire transition matrix $\Phi(t)$ (see [23]), however such methods can be particularly burdensome. These calculations can be simplified when the matrices $\nabla_x F(t,0)$ and $\int_0^t \nabla_x F(s,0) ds$ commute. However, when these matrices do not commute, one of many other methods to compute the transition matrix can be used to determine the eigenvalues of $\Phi(2\pi)$. For example, Pandiyan & Sinha [97] showed that there always exists some matrix $C$ such that $\Phi(2\pi) = \exp(2\pi C)$, reducing condition (4.18) to finding the eigenvalues of $C$. Utilising this property, a clearer condition than (4.18) could be formulated.

In Chapter 3 it was shown that when an equivalent condition to Assumption (5) was not satisfied, the system could exhibit an Allee effect, whereby the initial conditions of the system determined the eventual survival or extinction of the metapopulation. Therefore, it might be expected that the nonautonomous system with parameters varying sufficiently small would also exhibit an Allee effect, motivating more research into this area. Furthermore, such a phenomenon has been characterised for nonautonomous systems only recently [81].
4.3 Stochastic Influences

In Section 4.2, issue of deterministic variation in the environment was addressed, which account for factors such as breeding and migratory seasons. However, some alterations in the environment cannot be predicted with a high level of accuracy. For example, the connectivity between patches can be abruptly reduced or even cut throughout various times of the year, or perhaps the suitability of a patch is drastically reduced when a disease infects the food source in the area. Occurrences like these are inherently hard to predict and as such, are classified as stochastic influences in the system. The model presented in Section 4.2 cannot account for stochastic influences on the environment.

In this section, a model similar to the one in Section 4.2 is presented but rather than allowing the parameters of the system to changing deterministically, they are allowed to change stochastically.

4.3.1 Model

Extending the model presented in Chapter 3, we define a family of Markov processes \((Y^{(N)}(t), t \geq 0)\), indexed by \(N \geq 1\) as follows. The parameter \(N\) measures the size of the ceiling of the total population over all patches. Let \(J\) be the number of patches in the metapopulation and \(n(t) = \{n_1(t), \ldots, n_J(t)\}\), where \(n_i(t)\) is the number of individuals on patch \(i\). The process \(n(t)\) has state space \(S_N = \{0, \ldots, N_1\} \times \cdots \times \{0, \ldots, N_J\}\).

The process \(C(t)\) is introduced to model the environment; it represents which configuration the environment is in at time \(t\). The process \((C(t), t \geq 0)\) has a state space \(S_C := \{1, \ldots, K\}\). The birth, death and migration rates depend on the configuration of the environment; \(C\) in superscript is used to denote the rates when the environment is in configuration \(C\). The process \((C(t), n(t))\) then has the following transition rates:

\[
q((C, n), (C, n) + (0, e_i)) = \nu_i^{(C)}(N_i - n_i) + n_i b_i^{(C)} \left( \frac{n_i}{N_i} \right), \quad (4.21a)
\]

\[
q((C, n), (C, n) - (0, e_i)) = \phi_i^{(C)}(n_i) \lambda_i^{(C)} + d_i^{(C)} n_i, \quad (4.21b)
\]

\[
q((C, n), (C, n) - (0, e_i + e_j)) = \phi_i^{(C)}(n_i) \lambda_j^{(C)} \frac{N_j - n_j}{N_j} \quad \text{for all } j \neq i, \quad (4.21c)
\]

\[
q\left( (C, n), (C, n) + (l_i^{(N)}, 0^T) \right) = g_i^{(N)}(C, n/N), \quad \text{for } i = 1, \ldots, k. \quad (4.21d)
\]
Chapter 4: Environmental Influence

The rates (4.21a) – (4.21c) are equivalent to the rates (3.1) if $C$ is constant. The parameters $\nu^{(c)}_i$, $d^{(c)}_i$, $\lambda^{(c)}_{ij}$ and $N_i$ are the immigration rate per unoccupied space, death rate per-capita, proportion of individuals migrating from patch $i$ to patch $j$ (or out of the system if $j = 0$) and the patch ceiling for patch $i$, respectively, when in configuration $c$. The birth rate function $b^{(c)}_i(\cdot)$ determines the per–capita birth rate given how densely populated patch $i$ is when in configuration $c$. The function $\phi^{(c)}_i(\cdot)$, henceforth referred to as the migration function, represents the rate at which individuals leave patch $i$. The migration function satisfies $\phi^{(c)}_i(0) = 0$ and $\phi^{(c)}_i(n) > 0$ for $n \geq 1$. The final transition rate (4.21d) describes transitions between configurations. To that end, it is assumed that there are $k$ transitions between configurations and $l^{(N)}_i$ is the jump size of the $i$th jump while $g^{(N)}_i(c, x)$ is the transition rate of the $i$th jump.

Collectively, rates (4.21) define all the possible non zero transitions for the Markov process $((C(t), n(t)), t \geq 0)$, with state space $S_C \times S_N$. The process $Y^{(N)}(t)$ is then defined as $Y^{(N)}(t) = (C(t), n(t)/N)$, with a state space of $S_C \times E_N$ where $E_N := S_N / N$.

4.3.2 Piecewise-Deterministic Approximation

We now employ Theorem 2.11 to show that the evolution of the CTMP $Y^{(N)}(t), t \geq 0$ can be approximated by a piecewise-deterministic Markov process (PDMP). To apply Theorem 2.11 we first verify that Condition 1 holds. Assume $M^{(N)}_i \rightarrow M_i$ for all $i$ and define $E := [0, M_1] \times \cdots \times [0, M_J]$. Suppose that there are functions $\hat{\phi}^{(N,c)}_i : [0, M^{(N)}_i] \rightarrow \mathbb{R}_+$ such that, for all $n \geq 1, N \geq 1$ and $c \in S_C$,

$$\hat{\phi}^{(N,c)}_i \left( \frac{n}{N} \right) = \frac{\phi^{(c)}_i(n)}{N},$$
and let the functions $r_{n,i}^{(N)} : S_C \times E \to \mathbb{R}$ be defined as

\begin{align*}
    r_{1,i}^{(N)}(c,x) &= g_i^{(N)}(c,x) \\
    r_{2,i}^{(N)}(c,x) &= N \left( \nu_i^{(c)}(M_i^{(N)} - x_i) + x_i b_i^{(c)} \left( \frac{x_i}{M_i^{(N)}} \right) \right), \\
    r_{3,i}^{(N)}(c,x) &= N \left( \tilde{\Phi}_i^{(N,c)}(x_i) \lambda_i^{(c)} + d_i^{(c)} x_i \right), \\
    r_{4,i,j}^{(N)}(c,x) &= \begin{cases} 
        N \tilde{\Phi}_i^{(N,c)}(x_i) \lambda_i^{(c)} \left( 1 - \frac{x_i}{M_j^{(N)}} \right) & \text{for all } j \neq i, \\
        0 & \text{for } j = i.
    \end{cases}
\end{align*}

Finally, define the jump sizes $e_{n,i}^{(N)}$ as

\begin{align*}
    e_{1,i}^{(N)} &= \left( t_i^{(N)}, 0^T \right), \\
    e_{2,i}^{(N)} &= (0, (1/N)e_i), \\
    e_{3,i}^{(N)} &= (0, -(1/N)e_i), \\
    e_{4,i,j}^{(N)} &= (0, (1/N)(-e_i + e_j)),
\end{align*}

where $e_i$ is a unit vector with a one in the $i$th position. Then, the transitions of $Y^{(N)}(t)$ can be written as

\begin{align*}
    q \left( Y^{(N)}, Y^{(N)} + e_{n,i}^{(N)} \right) &= r_{n,i}^{(N)} \left( Y^{(N)} \right), \quad \text{for } n \in \{1, 2, 3\} \text{ and } i \in \{1, \ldots, J\}, \\
    q \left( Y^{(N)}, Y^{(N)} + e_{4,i,j}^{(N)} \right) &= r_{4,i,j}^{(N)} \left( Y^{(N)} \right), \quad \text{for } (i, j) \in \{1, \ldots, J\}^2.
\end{align*}

To confirm Condition 1 holds, the transitions that constitute large and small jumps must be defined. There are $J$ transitions involving an increase due to a birth or immigration, $J$ transitions involving a decrease due to a death or removal and $J(J - 1)$ transitions accounting for a migration from one patch to another patch. There are $k$ transitions for $C(t)$ and therefore there are $k + J(J + 1)$ possible transitions for $Y^{(N)}(t)$. Order the transitions such that large jumps are the first $k$ transitions, that is transitions between configurations, ((4.23a) with corresponding rate (4.22a)), and small jumps are the remaining $J(J + 1)$ transitions ((4.23b)–(4.23d) with corresponding rates (4.22b)–(4.22d)).
Parts (ii) and (iii) of Condition 1 are satisfied if \( l_i^{(N)} \to l_i \) for some \( l_i < \infty \) as \( Ne_i^{(N)} \), \( Ne_i^{(N)} \) and \( Ne_i^{(N)} \) do not depend on \( N \). If there exists functions \( g_i : S_C \times E \to \mathbb{R} \) that are bounded and Lipschitz continuous and satisfy
\[
\lim_{N \to \infty} \sup_{x \in S_C \times E} \left| g_i^{(N)}(x) - g_i(x) \right| = 0, \quad \text{for all } i,
\]
then part (iv) of Condition 1 is satisfied. Next, assume there are some functions \( \hat{\phi}_i^{(c)} : [0, M_i] \to \mathbb{R}_+ \) that are bounded and Lipschitz continuous and satisfy
\[
\lim_{N \to \infty} \sup_{x \in [0, M_i]} \left| \hat{\phi}_i^{(N,c)}(x) - \hat{\phi}_i^{(c)}(x) \right| = 0, \quad \text{for all } i,
\]
and the birth functions \( b_i^{(c)} : [0, M_i] \to \mathbb{R}_+ \) are bounded and Lipschitz and satisfy
\[
\lim_{N \to \infty} \sup_{x \in [0, M_i]} \left| b_i^{(c)} \left( \frac{x}{M_i^{(N)}} \right) - b_i^{(c)} \left( \frac{x}{M_i} \right) \right| = 0, \quad \text{for all } i.
\]
Then parts (v) and (vi) of Condition 1 are satisfied. And as the functions \( r \) are bounded and all transitions are bounded, part (vii) of Condition 1 is also satisfied. Now define a new process \( Y(t) \) with space \( S_C \times E \) such that
\[
Y(t) = y_0 + \sum_{i=1}^{K} (l_i, 0^T) \Pi_i \left( \int_0^t g_i(Y(s)) \, ds \right) + \int_0^t V(Y(s)) \, ds,
\]
where \( \Pi_i(\cdot) \) are independent Poisson processes with unit rate and the elements of \( V \) are defined as
\[
V_i(c, x) = 0,
\]
\[
V_{1+i}(c, x) = M_i \nu_i^{(c)} + b_i^{(c)} - \nu_i^{(c)} - d_i^{(c)} x_i - 2 \frac{b_i^{(c)}}{M_i} x_i^2 + \sum_{j \neq i} \left( \hat{\phi}_j^{(c)}(x_j) \lambda_{ij}^{(c)} x_j - \hat{\phi}_j^{(c)}(x_j) \lambda_{ij}^{(c)} \frac{x_j}{M_j} - \hat{\phi}_j^{(c)}(x_j) \lambda_{ij}^{(c)} \frac{x_i}{M_i} \right),
\]
for all \( i \in \{1, \ldots, J\} \). Then Condition 1 is satisfied and this leads to the following theorem.

93
Chapter 4: Environmental Influence

Theorem 4.5. Suppose there are functions $g_i : S_C \times E \rightarrow \mathbb{R}$ which are bounded and Lipschitz continuous and satisfy (4.24). Further suppose there exist some functions $\hat{\phi}_i^{(c)} : [0, M_i] \rightarrow \mathbb{R}$ which are bounded and Lipschitz continuous and satisfy (4.25). Then if, a.s.,

$$\lim_{N \rightarrow \infty} Y^{(N)}(0) \rightarrow y_0,$$

then

$$Y^{(N)}(t) \rightarrow Y(t),$$

almost surely in Skorokhod topology, where $Y(t)$ is given by (4.27).

4.3.3 Equilibrium Behaviour

Theorem 4.5 suggests that the process $Y(t)$ approximates the original process $Y^{(N)}(t)$ well for large $N$. Therefore, the long term behaviour of the approximating process, $Y(t)$, will provide a good approximation of the long term behaviour of the original process. Recall that $Y(t) = (C(t), x(t))$ where $x(t)$ is the approximation ($N \rightarrow \infty$) for the state of the density of the metapopulation at time $t$, $(n(t)/N)$. Differentiating (4.27), we see that

$$\frac{dx(t, x_0)}{dt} = F(C(t), x), \quad x(0, x_0) = x_0,$$  \hspace{1cm} (4.28)

where

$$F_i(c, x) = M_i\nu_i^{(c)} + \left(b_i^{(c)} - \nu_i^{(c)} - d_i^{(c)}\right)x_i - \hat{\phi}_i^{(c)}(x_i) - \frac{b_i^{(c)}}{M_i}x_i^2$$

$$+ \sum_{j \neq i} \left(\hat{\phi}_j^{(c)}(x_j)\lambda_{ji} + \left(\hat{\phi}_i^{(c)}(x_i)\lambda_{ij}^{(c)} \frac{x_j}{M_j} - \hat{\phi}_j^{(c)}(x_j)\lambda_{ji}^{(c)} \frac{x_i}{M_i}\right)\right),$$

and $C(t)$ is the configuration process. To determine the long term behaviour of the system (4.28), the following conditions on the parameters and functions of the system are imposed:

94
Chapter 4: Environmental Influence

(i) The parameters \( \phi_i^{(c)} > 0 \) for all \( i \) and \( c \).

(ii) For all \( i, j \) and \( c \) and \( t \in \mathbb{R}_+ \), there is a finite sequence \((a_k)\) such that

\[
\lambda_{ia_1}^{(c)} \lambda_{a_1a_2}^{(c)} \cdots \lambda_{a_{m}j}^{(c)} \neq 0.
\]

(iii) There is no external immigration to any patch, that is, \( \nu_i^{(c)} = 0 \) for all \( i \).

(iv) For all \( i \) and \( c \), \( b_i^{(c)}(x) \) is a continuously differentiable, strictly decreasing function on \([0,1]\) such that \( b_i^{(c)}(x) = 0 \) for all \( x \geq 1 \) and \( xb_i^{(c)}(x) \) is strictly concave on \([0,1]\).

(v) The parameters \( \phi_i^{(c)}, \lambda_{ij}^{(c)} \) and \( M_i \) satisfy

\[
\phi_i^{(c)} \lambda_{ij}^{(c)} M_i = \phi_j^{(c)} \lambda_{ji}^{(c)} M_j \quad \text{for all } i, j, c.
\]

(vi) The transition rate functions \( g_i(c,x) \) do not depend on \( x \). That is \( g_i(c,x) \equiv g_i(c) \).

(vii) The process \( C(t) \) is irreducible and begins in the stationary distribution. That is \( C(0) \sim \pi \), where \( \pi \) is the solution to \( \pi \tilde{Q} = 0 \), and \( \tilde{Q} \) has elements given by \( \tilde{q}_{i,i+1} = g_j(i) \) and \( \sum_i g_i(c) > 0 \) for all \( c \).

Assumptions (i)–(v) are similar to Assumptions (A)–(F) in Section 3.4 for the time homogeneous model. The difference between these assumptions and the assumptions presented previously is that these assumptions must hold in every configuration that exists. Refer to Section 3.4 for a thorough explanation of assumptions (i) - (v). Assumption (vi) allows the transition rate of the configuration process to depend only on the current state of the configuration process, rather than the state of the metapopulation as well. Despite this limitation, the model can still adequately describe many stochastic influences such as fires, diseases in food sources and human influences since they are typically independent of the size of the metapopulation. Assumption (vii) is reasonable if the environment has been in existence for a long time. In the context of this model, this assumption relates to the stationary distribution which is ensured to exist and be unique due to the finite state space \( S_C \) and the irreducibility of \( C(t) \).

Under Assumptions (i)–(vii), we determine the long term behaviour of the system (4.28). Unfortunately, due to the random nature of the environment, a typical analysis
which determines the fixed points of the system and their stability cannot be done. Therefore, we first need to show that the system (4.28) is a random dynamical system. Tools from the field of random dynamical systems can then be used to determine the long term behaviour.

The first step is to show that the collection \( \theta \equiv (\Omega, \mathcal{F}, \mathbb{P}, (\theta_t, t \geq 0)) \), where \( \theta_s C(t) \equiv C(t + s) \), is a metric dynamical system. Hence this system (4.28) generates a random dynamical system (RDS).

**Proposition 4.6.** Let \( C(t) \) be on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \( \Omega \) is the space of càdlàg functions and \( \mathcal{F} \) is an appropriate \( \sigma \)-algebra and assume (v) and (vii) hold. Then the collection

\[
\theta \equiv (\Omega, \mathcal{F}, \mathbb{P}, (\theta_t, t \geq 0))
\]

where \( \theta_s C(t) \equiv C(t + s) \), is an ergodic metric dynamical system.

**Proof.** The proof proceeds by confirming the conditions of Definition 2.17 are met. The Markov process \((C(t), t \geq 0)\) is known to be an irreducible Markov process on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with values in \( S_C = \{1, \ldots, K\} \), where \( \mathcal{F} \) is the \( \sigma \)-algebra generated by \( C \). Furthermore, \( C \) posses the càdlàg property. Therefore

\[
\theta_0 = \text{id}, \quad \theta_t \circ \theta_s = \theta_{t+s}.
\]

Define \( \theta \equiv (\Omega, \mathcal{F}, \mathbb{P}, (\theta_t, t \geq 0)) \). For each \( t, \omega \rightarrow \theta_t(\omega) \) is a measurable mapping from \((\Omega, \mathcal{F}) \rightarrow S_C \). Under Assumption (vii), \( C(t) \) is a stationary process. Therefore \( \mathbb{P}(\theta_t B) = \mathbb{P}(B) \) for all \( B \in \mathcal{F} \) and \( t \geq 0 \). More so, \( \theta_t B = B \) for all \( t \geq 0 \) and \( \theta \)-invariant \( B \in \mathcal{F} \), which completes the proof. \( \square \)

It can now be shown that (4.28) is a CRDE by Definition 2.19 that generates an RDS of a particular form.

**Theorem 4.7.** Assume [i] - [vi] hold. The system (4.28) is a CRDE and generates an RDS \((\theta, \varphi)\) in \( \mathbb{R}_+^d \) with the cocycle defined by

\[
\varphi(t, \omega)x_0 \equiv x(t) = x_0 + \int_0^t F(C(u), x(u))du.
\]
Chapter 4: Environmental Influence

Proof. We need to verify that the system (4.28) satisfied the conditions of Definition 2.19 Part (I) holds for $F(C(t,\omega),x)$ as it is continuously differentiable in $x$ and bounded for all $x \in E$. Furthermore it can be seen that the partial derivatives $\partial F_i/\partial x_j$ defined by

$$
\nabla_x F(c,x)_{ij} = \begin{cases} \phi_j^{(c)} \lambda_j^{(c)} & \text{if } i \neq j, \\
\frac{b_i^{(c)}}{M_i} \left( \frac{x}{M_i} \right) + \frac{x}{M_i} b_i^{(c)^T} \left( \frac{x}{M_i} \right) - d_i^{(c)} - \phi_i^{(c)} & \text{if } i = j,
\end{cases}
$$

are also bounded for all $x \in E$. To confirm part (II), let $C_1 = \max_i \sum_j \phi_j^{(c)} \left( 1 - \lambda_j^{(c)} \right)$. Then $(x,F) \leq C_1|x|^2 + C_2$. Part (III) holds as $F_i(c,x) = x_i g_i(c,x) + \left( J_0^{(c)} x \right)_i \geq 0$ if $x_i = 0$, for some function $g$ as $\left( J_0^{(c)} x \right)_i \geq 0$ because $J_0^{(c)}$ is a Metzler matrix. To confirm part (IV) holds, define two vectors $x,y$, such that $x_i = y_i$ and $x_j \leq y_j$ for $j \neq i$. Then

$$
F_i(c,x) - F_i(c,y) = - \sum_{j \neq i} (y_j - x_j) \left( \frac{\phi_i \lambda_{ij}}{M_j} + \frac{\phi_j \lambda_{ji}}{M_i} (1 - M_i) \right) \leq 0,
$$

showing that $F_i(c,x) \leq F_i(c,y)$, confirming part (IV). Therefore (4.28) is a CRDE and by Proposition 5.2.1 of [34], it generates an RDS $(\theta,\varphi)$ in $\mathbb{R}_+^J$ with the cocycle defined as per the theorem.

We see that (4.28) generates an RDS. Hence, Theorem 2.23 can be used to determine the long term behaviour of the system under Assumptions (i) - (vii).

**Theorem 4.8.** Let $\Phi(t,\omega)$ be the cocycle of the linear RDS governed by the equation $\frac{dz}{dt} = \nabla_x F(C(t,\omega),0)z$ and define $\lambda$ as the top Lyapunov exponent of $\Phi(t,\omega)$ by Definition 2.22. If $\lambda$ is such that

$$
\lambda < 0,
$$

then for any $\varepsilon > 0$,

$$
\lim_{t \to \infty} \Pr(\{x(t) > \varepsilon\}) \to 0
$$

Proof. We aim to confirm the conditions of Theorem 2.23 are met. The function $\nabla_x F(c,x)$ is known to be irreducible for all $x \in E$ and $c \in \Omega$ due to assumptions (i) and
And to show $F(c, x)$ is s-concave under Assumption (v), recall the definition of $\nabla_x F(c, x)$ in (4.29). If $i \neq j$, $\nabla_x F(c, x)_{ij} = \nabla_x F(c, y)_{ij}$ for any $x, y$. For $i = j$, note that $\frac{d^2}{dx^2} F_i(c, x) = M^{-1} \frac{d^2}{dx^2} b_i^{(c)}(x) < 0$ from Assumption (iv), which shows that $\nabla_x F(c, x)_{ii}$ is a strictly decreasing function, implying that $\nabla_x F(c, y)_{ii} < \nabla_x F(c, x)_{ii}$ if $x < y$. Therefore $\nabla_x F(c, y) \leq \nabla_x F(c, x)$ if $x < y$, showing $F(c, x)$ is s-concave. Finally, under Assumption (iii) $F(c, 0) = 0$ and from Proposition 4.6, $\theta$ is an ergodic MDS. Therefore, if (4.30) is satisfied, according to Theorem 2.23 for any $\varepsilon > 0$, $\lim_{t \to \infty} \Pr(|x(t)| > \varepsilon) \to 0$, which completes the proof.

Condition (4.30) requires finding the top Lyapunov exponent of the linear system, which can often be quite difficult. In the following examples, Theorem 4.8 is applied to determine sufficient conditions for the metapopulation to go extinct.

**Example 1.** Consider a metapopulation where the migration rates between patches are symmetric for a particular configuration; that is

$$\phi_i^{(c)} \lambda_{ij}^{(c)} = \phi_j^{(c)} \lambda_{ji}^{(c)} \text{ for all pairs } i, j \text{ and for all } c. \quad (4.31)$$

The trajectory of the cocycle $\Phi(t, \omega)$ of the linear RDS given by $\frac{dx}{dt} = \nabla_x F(C(t, \omega), 0)x$ before the first jump is

$$\Phi(t, \omega)x_0 = \exp \left(A_{C(0)}t\right)x_0, \quad 0 \leq t < T_1,$$

where $\exp$ is the matrix exponential, $A_i := \nabla F(i, 0)$ and $T_i$ is the time of the $i$th switch of configurations (noting that $T_0 = 0$). To extend this concept, define $c_i := C(T_i)$ (that is, $c_i$ is the configuration after the $i$th switch of configurations) and $N(t) := \sup\{i : T_i < t\}$ (that is, the number of switches between configurations at time $t$). Then

$$\Phi(t, \omega)x_0 = g(t)h(N(t))x_0, \quad (4.32)$$

where $h : \mathbb{Z} \mapsto \mathbb{R}^{J \times J}$ and $g : \mathbb{R} \mapsto \mathbb{R}^{J \times J}$ are given by

$$h(n) = \prod_{i=0}^{n} \exp \left(A_{c_i}(T_{i+1} - T_i)\right), \quad \text{and } g(t) = \exp \left(A_{c_{N(t)}}(t - T_{N(t)})\right).$$
Therefore, finding the top Lyapunov exponent is equivalent to finding a particular bound for the product of $h$ and $g$. We first determine a bound for the norm of matrices of the form $\exp(Mt)$, where $M$ is assumed to be Metzler. It was shown in Theorem 2 of [72] that, for any Metzler matrix $M \in \mathbb{R}^{J \times J}$, there exists a $\beta \geq 1$ such that

$$\|\exp(Mt)\| \leq \beta e^{(r+\epsilon)t}, \quad \text{for all } t \geq 0,$$

(4.33)

for every $\epsilon \geq 0$, where $r$ is the maximum real eigenvalue of $M$ and $\epsilon$ can be chosen to be zero because $r$ is unique due to part (d) of Theorem 2.15. Furthermore, if $M$ is a normal matrix ($MM^T = M^TM$), the inequality turns to an equality and $\beta = 1$. While not the only requirements, $M$ will be a normal matrix when it is symmetric or orthogonal. Note that when $t = 0$, $\|\exp(Mt)\| = 1$ and $e^{(r+\epsilon)t} = 1$, which is why $\beta \geq 1$ is required. Finding a value for $\beta$ is done via the following. Assume $M$ has $p$ multiple eigenvalues with algebraic multiplicities $t_k$ and geometric multiplicities $m_k$ for $k = 1, \ldots, p$. Then $M$ can be expressed in Jordan normal form as $M = VJV^{-1}$ where

$$J = \text{diag}(J_{11}, J_{12}, \ldots, J_{1m_1}, J_{21}, \ldots, J_{p1}, \ldots, J_{pm_p})$$

and

$$J_{ki} = \begin{pmatrix} \lambda_k & 1 & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & \lambda_k \end{pmatrix},$$

for $k = 1, \ldots, p$ and $i = 1, \ldots, m_k$. Assume $V$ is chosen such that $K(V) := \|V\|_2\|V^{-1}\|_2$ is minimised. Furthermore define $D = \text{diag}(D_{ki})$, where $D_{ki} = \text{diag}(1, \delta, \delta^2, \ldots, \delta^{\gamma_{ki}-1})$ for $k = 1, \ldots, p$, $i = 1, \ldots, m_k$, and $\delta = \min\{1, \epsilon\|\|N\|_F\}$ and $N$ is the shift operator matrix with unit elements in the same positions $(i, i+1)$ as $J$ and $\epsilon > 0$. Then, $\beta = K(VD)$. Numerically, this can be accomplished in MATLAB by computing $[V, J] = \text{jordan}(M)$ and then $\text{beta} = \text{norm}(V)\times\text{norm}(\text{inv}(V))$.

We now turn to finding a bound for the product of $h$ and $g$. From (4.31), we see that $A_i$ is symmetric and hence, the inequality in (4.33) is an equality and $\beta_i = 1$. 

99
Therefore for every $\epsilon \geq 0$

$$
\|g(t)h(N(t))\| \leq \left\| \exp\left( A_{eN(t)} \left( t - T_{N(t)} \right) \right) \right\| \prod_{i=1}^{N(t)} \| \exp\left( A_{c_i} (T_i - T_{i-1}) \right) \|
\leq e^{(r_{N(t)} + \epsilon) (t - T_{N(t)}) + \sum_{i=1}^{N(t)} (r_{c_i} + \epsilon) (T_i - T_{i-1})}.
$$

(4.34)

The power of the exponential can be rewritten in the following form:

$$(r_{cN(t)} + \epsilon) (t - T_{N(t)}) + \sum_{i=1}^{N(t)} (r_{c_i} + \epsilon) (T_i - T_{i-1}) = t \sum_{i=1}^{K} (r_i + \epsilon) \frac{1}{t} \int_{0}^{t} 1\{C(s) = i\} ds.$$

By the ergodic theorem, $\frac{1}{t} \int_{0}^{t} 1\{C(s) = i\} ds \to \pi_i$ almost surely as $t \to \infty$ (recalling that $\pi$ is the stationary distribution described in Assumption (vii)). As such, (4.34) is reformulated as

$$
\|g(t)h(N(t))\| \leq e^{t \sum_{i=1}^{K} (r_i + \epsilon)} \left( \frac{1}{t} \int_{0}^{t} 1\{C(s) = i\} ds - \pi_i \right) e^{t \sum_{i=1}^{K} (r_i + \epsilon) \pi_i}.
$$

If a new random variable is defined as $R_{\epsilon}(\omega) := e^{t \sum_{i=1}^{K} (r_i + \epsilon)} \left( \frac{1}{t} \int_{0}^{t} 1\{C(s, \omega) = i\} ds - \pi_i \right)$, then

$$
\|g(t)h(N(t))\| \leq R_{\epsilon}(\omega) e^{t \sum_{i=1}^{K} (r_i + \epsilon) \pi_i} = R_{\epsilon}(\omega) e^{t \sum_{i=1}^{K} r_i \pi_i}.
$$

(4.35)

Lastly, it is required that the set where

$$
\lim_{t \to \infty} \frac{1}{t} \log R_{\epsilon}(\theta(t)\omega) = 0
$$

(4.36)

be of full measure. It can be seen that

$$
\frac{1}{t} \log R_{\epsilon}(\theta(t)\omega) = \sum_{i=1}^{K} (r_i + \epsilon) \left( \frac{1}{t} \int_{0}^{t} 1\{C(s, \theta(s)\omega) = i\} ds - \pi_i \right),
$$

but from the ergodic theorem, it is known that $\frac{1}{t} \int_{0}^{t} 1\{C(s, \theta(s)\omega) = i\} ds \to \pi_i$ a.s. for all $i$. This shows that (4.36) is satisfied on a set of full measure. Combining (4.32) and (4.35) gives the inequality $\Phi(t, \omega)x_0 \leq R_{\epsilon}(\omega) e^{t (\epsilon + \lambda_1)} \|x_0\|$, where $\lambda_1 = \sum_{i=1}^{K} r_i \pi_i$. Hence, $\lambda_1$ is larger than to equal to the top Lyapunov exponent. Therefore, if $\sum_{i=1}^{K} r_i \pi_i < 0$, then $x(t) \to 0$ in probability as $t \to \infty$. The quantity $\lambda_1$ has an intuitive interpretation;
Chapter 4: Environmental Influence

it is a sum of growth rates weighted according to the expected time spent in each configuration and as such, if that sum is less than 0, the population should decrease overall.

Example 2. The next example drops Assumption (4.31). In Example 1, finding the top Lyapunov exponent was achieved by finding a bound for the product \( g(t)h(N(t)) \). Without (4.31), the inequality sign in (4.33) remains and \( \beta > 1 \). Therefore, inequality (4.34) becomes

\[
\|g(t)h(N(t))\| \leq e^{\int \ln \beta_N(t) + \left(r_{c_N(t)} + e\right)(t-T_{\Delta(t)}) + \sum_{i=1}^{N(t)} \ln \beta_{c_i} + \left(r_{c_i} + e\right)(T_i - T_{i-1})}.
\]

where \( N_i(t) \) is the number of visits to configuration \( i \) at time \( t \). It is to be noted that \( \frac{1}{t} \int_0^t 1\{C(s) = i\}ds \rightarrow \pi_i \) a.s. due to the ergodic theorem and \( \frac{N_i(t)}{t} \rightarrow \pi_i q_i \) a.s. where \( q_i = \sum_n g_n(i) \), due to the ergodic theorem applied with Theorem 8.1 of [25]. As such, (4.37) is reformulated as

\[
\|g(t)h(N(t))\| \leq R_\epsilon(\omega) e^{t + \sum_{i=1}^{K} (q_i \ln \beta_i + r_i) \pi_i}.
\]

where

\[
R_\epsilon(\omega) = e^{t \sum_{i=1}^{K} \ln \beta_i \left(\frac{N_i(t) - \epsilon}{t} - \pi_i q_i\right) + (r_i + \epsilon)\left(\frac{1}{t} \int_0^t 1\{C(s) = i\}ds - \pi_i\right)}.
\]

Note that

\[
\frac{1}{t} \log R_\epsilon(\theta(t)\omega) = \sum_{i=1}^{K} \ln \beta_i \left(\frac{N_i(t, \theta(t)\omega)}{t} - \pi_i q_i\right)
+ \sum_{i=1}^{K} (r_i + \epsilon) \left(\frac{1}{t} \int_0^t 1\{C(s, \theta(s)\omega) = i\}ds - \pi_i\right),
\]

which converges to 0 a.s. as \( t \rightarrow \infty \) due to the ergodic theorem applied with Theorem 8.1 of [25]. Combining (4.32) and (4.38) gives the inequality \( \Phi(t, \omega)x_0 \leq R_\epsilon(\omega) e^{t(\epsilon + \lambda_2)} \|x_0\| \), where \( \lambda_2 = \sum_{i=1}^{K} (q_i \ln \beta_i + r_i) \pi_i \), showing that \( \lambda_2 \) is larger than or equal to the top Lyapunov exponent. One can see that, if \( \beta_i = 1 \) for all \( i \), then \( \lambda_2 = \lambda_1 \), reducing this problem to Example 1. So, applying Theorem 4.8 if \( \sum_{i=1}^{K} (q_i \ln \beta_i + r_i) \pi_i < 0 \), then \( x(t) \rightarrow 0 \) in probability as \( t \rightarrow \infty \).
4.3.4 Discussion

In this section, we have presented an extension of the model from Chapter 3 that accounts for a stochastically changing environment. This was done by introducing a new process that models the current configuration of the environment. Due to the complexity of the system and the possible inclusion of an absorbing state, an approximating process was used to determine the long term behaviour of the process. Precisely, the original process converges to a piecewise-deterministic Markov process as the population ceiling goes to infinity.

While the PDMP is a simpler process than the original, the analysis remains difficult. This fact is the reason that the only condition determined was a sufficient condition for extinction. While not as advantageous for population dynamics as a sufficient condition for persistence, condition (4.30) can still be useful. For example, it provides ecologists with a condition to ensure a population of pests goes extinct and epidemiologists with a condition to ensure a disease dies out. In the context of the spread of a disease, the model can be used to look at the spread of parasites either within a host where the patches are different locations within the host or between several hosts where each patch is a host.

As the two examples have demonstrated, finding the top Lyapunov exponent for a specific metapopulation structure can be rather difficult, due to the relation for the matrix exponential of $\exp(A) \exp(B) \neq \exp(A + B)$ for any matrices $A$ and $B$, unless $A$ and $B$ commute. In the context of finding the top Lyapunov exponent in (4.34), the matrices $A_i$ only commute if the system is homogeneous, that is $b_{ij}^{(c)}/\left(d_{ij}^{(c)} + \phi_{ij}^{(c)} \lambda_{ij}^{(c)}\right) = \alpha_c$ and $\phi_{ij}^{(c)} \lambda_{ij}^{(c)} = \rho_c$ for all $i, j$ and $c$ and some constants $\alpha_c$ and $\rho_c$. This is a stronger condition than only (4.31), therefore such a system is encapsulated in Example 1 with the additional information that $r_i = \rho_i (J-1) + \alpha_i$. Unsurprisingly, both examples showed that the growth rate in each configuration (represented by the eigenvalue $r_i$) plays a vital role in determining whether a population will definitely go extinct. However, when the migration rates were assumed asymmetric, the sufficient condition for extinction is altered slightly by the introduction of the factor $q_i \ln \beta_i$, where $\beta_i$ can be as low as one, when the migration rates are symmetric, and increases as the difference between migration rates increases. However, the method to obtain the value for $\beta_i$ outlined
in Example 1 nearly always produces a value greater than 1, even when the rates are symmetric. Therefore, a method that produces a tighter bound than (4.33) is required. Whether the factor $q_i \ln \beta_i$, or something similar, is required for the necessary condition is unknown. To address this, a lower bound and tighter upper bounded on (4.34) are required. Such bounds would require different relations between the matrix norm and maximal real eigenvalues, such as the ones suggested by Cohen et al. [35].

A phenomenon that has recently been of interest is the persistence of a metapopulation consisting only of sink, or unfit, patches [65]. In the notation used in this thesis, this would correspond to a metapopulation persisting where, for some $c$, $b_i^{(c)}/(d_i^{(c)} + \phi_i^{(c)} \lambda_i^{(c)}) < 1$ for all $i$. This has been extended to stochastic environments [113, 115, 48], where the same phenomenon has been shown to exist. A precise statement about whether such a phenomenon occurs in my model is not possible due to Theorem 4.8 relating to the extinction of a species, rather than the persistence. However, a similar result by Roy et al. [113] could be possible. Roy et al. [113] show that persistence is possible even when all patches are sinks, on average. This last component of averaging is required for the phenomenon to occur in the model presented in this section, as it allows some patches to become sources for some environments (configurations). Without this, this model would show the opposite; that is, persistence is not possible in a metapopulation consisting of only sinks. This is also true for the model presented in Chapter 3. The main reason for the discrepancy between this model and the models of others [113, 115, 48] is the inclusion of deterministic trajectories in my model. For example, Schreiber [115] shows that the stochastic metapopulation growth rate, $\chi$, closely follows $\lambda_1$ (referred to as $\langle \mu \rangle$ by Schreiber) but is attenuated by the long term infinitesimal variance. This shows that although $\lambda_1$ is an important part of determining whether or not the metapopulation will go extinct, it is not the entire picture.
Conclusion

5.1 Summary

In this thesis, three stochastic SMMs have been presented that model a metapopulation within a static environment, a deterministically varying environment and a stochastically varying environment. These models were based on continuous time Markov population processes for which at least some of the rates were density dependent. These processes can be used to model the population size of a wide range of species. These models could also be used to model parasitic infections where the patches are interpreted as hosts. Here I will summarise the results presented, as well as possible extensions and further research.

In Chapter 3, the question of how an individual’s dynamics affects a population’s long term behaviour was addressed. A continuous time Markov process on a finite state space was used to model the individuals in a metapopulation. A SMM was formulated that was spatially structured in terms of patch connectivity and size, as well as accounting for the within patch dynamics of births, deaths and migrations. An analysis of the stationary distribution of this model would have proved unhelpful as the extinction state is reached with probability one. However, it has been observed that the time taken for this event to occur can be quite large for other metapopulation models. Therefore, a differential equation approximation was used. It was determined that under some technical conditions, the SMM converges in probability over finite time intervals to a system of nonlinear ordinary differential equations as the population
ceiling increased to infinity. The analysis of the limiting ODE resulted in some very interesting behaviours. A necessary and sufficient condition was determined that would ensure a metapopulation’s long term persistence or extinction under a strict symmetry assumption on some parameters. It is also important to note that this condition did not depend on the patch ceilings in any way, illustrating that increasing or decreasing patch ceilings (for example, the land area of a patch) will have no impact on the persistence of a species satisfying the strict symmetry assumption.

When the strict symmetry assumption did not hold, it was shown that metapopulations consisting of two patch could exhibit an Allee effect, which is where the initial condition of the metapopulation determines whether or not the metapopulation persists. A necessary condition for the metapopulation to exhibit an Allee effect was that the metapopulation contained only one patch suitable for reproduction (that is, \( b_i/(d_i + \phi_i\lambda_{i0}) > 1 \)) while the other patch was unsuitable (that is, \( b_i/(d_i + \phi_i\lambda_{i0}) < 1 \)). Under this assumption, the source of the Allee effect was shown to be the migration between patches, which differs from the typical source of an initially increasing birth rate function. When the unsuitable patch was near extinction, the rate of migration to the unsuitable patch could be sufficiently large so as to draw individuals away from the suitable patch. This caused the metapopulation to tend towards extinction. However, when the unsuitable patch was full, the migration rate away from the suitable patch was small. Individuals then tended to remain on the suitable patch and hence, ensured the metapopulation’s survival.

In Chapter 3, the long term behaviour of a metapopulation in a static environment was addressed. The extent to which these results hold for dynamic environments was then investigated in Chapter 4. In Section 4.2 a SMM similar to that presented in Chapter 3 was formulated with the addition of time dependent parameters. The resulting model was only asymptotically density dependent in time so required a new functional limit law to determine the differential equation approximation. This result encapsulated all the functional limit laws by Kurtz, Pollett and Pagendam (listed in Section 2.3) and further extended their results to a new family of processes. Applying this result to the SMM provided an approximating nonautonomous ODE for large population ceilings. Under a strict symmetry assumption similar to the one presented
Chapter 5: Conclusion

in Chapter 3 a necessary and sufficient condition was established for the persistence of a metapopulation in a deterministically changing environment. When this condition is satisfied, the size of the metapopulation tends to a nonzero level which changed continuously in time, otherwise it quickly decreases to zero.

Following this, in Section 4.3, the model from Chapter 3 was extended to account for a stochastically changing environment. The environment was assumed to have a finite set of configurations, each of which was associated with a set of parameters for the system. A Markov process was used to describe the state of the environment, allowing the environment to transition to a new configuration stochastically. A functional limit law was applied which showed that as the population ceiling increases, the process converges almost surely in Skorokhod topology to a piecewise deterministic Markov process (PDMP). Under some technical conditions on the initial distribution and transition rates of the environment, the PDMP was shown to be a random dynamical system (RDS). An analysis was undertaken on the RDS, from which a sufficient condition for extinction resulted. To demonstrate how to apply this condition, two examples were provided. Both examples required the calculation of the stationary distribution of the environment process, \( \pi \), as well as the eigenvalue of \( \nabla_x F(i,0) \) with maximal real part, \( r_i \), for each configuration \( i \). It was shown that \( r_i \) and \( \pi_i \) play vital roles in determining whether the population could go extinct.

5.2 Further Research

5.2.1 A Model with a Static Environment

The most interesting, and surprising, discovery from the results of Chapter 3 is the possible occurrence of the Allee effect. As previously mentioned, this is because the Allee effect is usually introduced into a model by a manipulation of the birth rate function \([2, 123]\). Although sufficient conditions for a two patch metapopulation to exhibit an Allee effect were provided, a complete characterisation of the initial conditions that allow the metapopulation to persist were not determined. It was only possible to state that if the trajectory begins smaller (larger), by partial ordering, than the unstable fixed point, \( x_u^* \), it will tend to zero (a nonzero fixed point \( x_s^* \)). A complete character-
Conclusion

isation of the initial conditions involves determining the domain of attraction for the fixed points, that is determining a set $D_x$ for a fixed point $x$, such that if $x_0 \in D_x$, then $x(t, x_0) \to x$ and $x(t, x_0) \not\to x$ for all $x_0 \not\in D_x$.

Preliminary work for a three patch metapopulation has shown that there is an Allee effect for higher dimensions. Figure 5.1 illustrates an example of a three patch system exhibiting an Allee effect. It shows a phase plane diagram from four different perspectives; (A) shows the phase plane in three dimensions while (B) is the phase plane diagram taken at $x_3 = 0$, (C) is for when $x_2 = 0$ and (D) is for when $x_1 = 0$. Two trajectories are plotted and, just like in Figure 3.4A, the metapopulation persists when all patches are full, whereas it tends towards extinction when the metapopulation starts small. Therefore, the Allee effect is not a phenomena unique to a two patch system. Further research into the existence of the Allee effect is required to better understand when it occurs in higher dimensional systems. While the perturbation analysis performed on the two patch system could be extended to higher dimensional systems, this approach becomes cumbersome and intractable as the dimension increases. One approach to proving the existence of an Allee effect in metapopulations with an arbitrary number of patches would be to apply an extended version of Lemma 3.7 and rework inequality (3.18).

Without Assumption (F), there is no means to exclude the possibility of an Allee effect. Therefore, the persistence of a metapopulation is not purely dependent on the extinction condition (3.9). However, numerical results have shown that the conclusions of Theorems 3.6 and 3.9 may still hold (that is, there is no Allee effect) even when Assumption (F) does not hold. Therefore, the necessary conditions for an Allee effect remain to be determined and finding such conditions is a high priority. In summary, there are three main courses of research resulting from this work; first is finding the domain of attraction for a two patch metapopulation, with higher dimensional systems a longer term goal; second, a proof that determines sufficient conditions that a metapopulation with any number of patches can exhibit an Allee effect; and thirdly, a proof that determines necessary and sufficient conditions for a system to exhibit an Allee effect.

One aspect of research that has gone mostly untouched is the distribution of the
Figure 5.1: A phase plane diagram for a $J = 3$ system, where trajectories are the red lines and fixed points are the circle (stable) and square (unstable) markers. The parameters used for this plot were $b_1(0) = 0.54$, $d_1 = 0.8$, $\phi_1 = 0.94$, $M_1 = 0.27$, $\lambda_{12} = 0.3$, $\lambda_{13} = 0.37$ and $b_2(0) = 4.42$, $d_2 = 1.1$, $\phi_2 = 4.82$, $M_2 = 0.67$, $\lambda_{21} = 0.61$, $\lambda_{23} = 0.2$ and $b_3(0) = 0.24$, $d_3 = 2.63$, $\phi_3 = 2.12$, $M_3 = 0.06$, $\lambda_{31} = 0.18$, $\lambda_{32} = 0.51$. 
process, at time $t$ or in the limit as $t \to \infty$, derived from Theorems 3.12 or 3.13. As the deterministic approximation negates the possibility of extinction when certain criteria are met (unlike the original stochastic process), Barbour and Pollett [17] have shown that there exists a relationship between these distributions and quasi-stationary distributions. Exploring this further for our particular process, the distributions from Theorems 3.12 or 3.13 could be investigated to determine how they change when the number of fixed points in the deterministic system changes. Of particular interest would be a metapopulation of two patches, for which the occurrence of an Allee effect is somewhat understood. Investigating this distribution further should provide a better understanding about how the underlying stochastic dynamics are affected when an Allee effect does, or does not, exist.

5.2.2 A Model with a Dynamic Environment

In Chapter 4, models that accounted for either a deterministically or stochastically changing environment were proposed. However, many environments would be exposed to both stochastic and deterministic changes and therefore, a model that allows for such changes needs to be developed. It is unknown whether a result similar to Theorem 2.11, but with time dependent rates, holds. However, it is expected that such a result holds if the time dependent rates remain bounded, similar to the conditions given in Theorem 4.2.

**Deterministically Dynamic.** It is possible that an Allee effect of some form exists in the nonautonomous system (4.17) when Assumption (5) does not hold. A numerical analysis of the system has generated promising results, showing that when variation in parameters is small, the initial conditions in $x$ and $t$ play some role in determining the long term behaviour of the system. To confirm the existence of an Allee effect, a perturbation analysis on a two patch model should be undertaken. However, another approach will be required for higher dimensional systems. An initial step before a perturbation analysis is undertaken should be to determine a necessary condition for persistence when Assumption (5) does not hold, as this will be more applicable to populations for ecologists.
**Stochastically Dynamic.** It may be possible to strengthen the type of convergence in Theorem 4.8 similar to the results that Benaim and Schreiber [21] proved for a discrete time model. That is, if condition (4.30) is satisfied, then \( x(t) \to 0 \) almost surely. Theorem 4.4.2 from [34] could be used as a starting point to show such a result is true.

As was the case with a deterministically changing environment, a necessary condition for persistence, with or without Assumption (v), should be the initial aim of any further research. Persistence is guaranteed if it is assumed that Assumption (iii) does not hold such that \( \nu_i(c) > 0 \) for all \( i, c \) as then \( 0 \) will not be an equilibrium point for any configuration and by Theorem 2.23 the population will persist. However, a more useful result will be determining a necessary condition for persistence when \( F(c, 0) = 0 \) for all \( c \) and \( 0 \) is unstable only for some configurations. As such, it might be expected that if the top Lyapunov exponent is greater than one, then there exists a nonzero equilibrium \( u \) (see Definition 2.21) such that \( x(t) \to u \), at least in probability, for all initial conditions. However, as Example 1 and Example 2 in Section 4.3.3) demonstrates, calculating the top Lyapunov exponent is not always straightforward. As such, a more useful result could find sub-equilibria and super-equilibria (see [34, Section 3.4]) and one of the many results from [34, Chapter 5] could then be applied. In our context, a sub-equilibrium (super-equilibrium) is a set of points that remain smaller (larger) by partial ordering than the trajectory of the metapopulation density for all \( t \), taking into account that the environment switches configurations. The point \( 0 \) is an example of a sub-equilibrium, while \( M \) (the relative patch ceilings) would be classified as a super-equilibrium.

Even when applied to simple systems (see Example 1 in Section 4.3.3), condition (4.30) does not fully characterise the relationship between an individual’s dynamics and the population’s outcome. Therefore, an important step is to determine the relationship the eigenvalue with largest real part, \( r_i \), has with the parameters of configuration \( i \). Condition (3.9) provides a method to determine whether \( r_i \) is positive, negative or zero. A logical next step would be to determine the relationship more precisely, at least for simple systems. An example of this is the homogeneous system discussed in Section 4.3.4 where \( r_i = \rho_i (J - 1) + \alpha_i \).
Chapter 5: Conclusion

The presence of an Allee effect in a model with a stochastically changing environment could also be established. If every configuration exhibits an Allee effect, it might be expected that the RDS would exhibit an Allee effect of some form as well. This is likely to be true, however a better description for an Allee effect in an RDS will be required before confirmation. An initial step would be to numerically confirm the presence of an Allee effect for some appropriate range of parameters.
References


References


References


References


References


