The general continuum model for structured populations, with two case studies in plant ecology

Henri Laurie

Thesis presented for the degree of
DOCTOR OF PHILOSOPHY
in the Department of Applied Mathematics
UNIVERSITY OF CAPE TOWN

December 1993

The general continuum model for structured populations, with two case studies in plant ecology

Henri Laurie

Thesis presented for the degree of
DOCTOR OF PHILOSOPHY
in the Department of Applied Mathematics
UNIVERSITY OF CAPE TOWN

December 1993

Abstract

The broad aim of this thesis is to investigate the formulation and usefulness of a very general model for plant population dynamics.

In chapter 1, the goal of generality is discussed, particularly in the light of the lack of interaction between field and experimental population studies on the one hand and theoretical population dynamics on the other hand. A distinction is made between descriptive and axiomatic theories, and it is suggested that they serve different purposes. The advantages of a rigorous framework are pointed out and the basic elements of the continuum approach are introduced.

In chapter 2, the model is proposed, the existence and uniqueness of solutions to its equations is proved, and an algorithm for numerically approximating transient solutions is discussed. The question of generality is addressed in two places, and it is argued that the basic framework presented here is in principle adequate to model the processes of plant population dynamics in full detail, though the existence proof cannot to accomodate all possible models. In particular, models with time lags are excluded. Further limitations of the existence proof in terms of constitutive relations are pointed out. In consequence, the theory here presented does not fully exploit the possibilities for generality inherent in the basic equations.

In chapter 3, the question of what data would allow identification of factors determining somatic growth and mortality is investigated computationally. It is shown that using only the average size is insufficient. A class of models which includes all possible combinations of three types of size dependence in somatic growth and mortality is formulated. Qualitative parameter estimation for the various models yields size distributions that can be classified into the following biologically meaningful groups: group (i) has no models that use dependence on relative size; group (ii) has all the models in which somatic growth depends on relative size; group (iii) has the models where only mortality depends on relative size. Thus it appears that size distribution may be used to distinguish various forms of size dependence in somatic growth and mortality.

In chapter 4, a lottery model criterion for coexistence of plants with disjoint generations is developed, which is shown to require relative density dependence. Computer simulations aiming to initiate the use of exploratory calculations in studies of coexisting serotinous proteoids in fynbos indicate that the aspect of plant population dynamics most sensitive to density dependence is seed production, then somatic growth, while mortality is least sensitive to density dependence.

Declaration

I, Henri de Guise Laurie, declare that the contents of this thesis is my own work and that none of it has been presented for a degree at any other university.

H. Laurie

Dedication

To the hope that knowledge of the truth will make us (all) free.

Acknowledgements

To my supervisors Daya Reddy and Richard Cowling, without whom this project would not have been conceived, and without whose guidance and good humour it would not have been completed.

To my wife Lesley Miles, who had great faith in me and also cooked some crucial meals.

To the botanists at University of Cape Town, especially William Bond, Penny Mustart, Karen Esler and all the other post-graduate students, whose appetite and enthusiasm for every aspect of botany never ceases.

To Doug Butterworth, Penny Mustart and Ralf Wittenberg for careful readings at short notice.

And to my fellow denizens of room 401 (you know who you are), who provide living proof that mathematics is, after all, a fully human activity.

Thank you all very much.

Contents

3 Plant mortality and somatic growth: on model discrimination for even-aged monospecific stands 50

Chapter 1

Introduction: the role of continuum modelling in plant population dynamics

This work, which is a thesis in the modern sense of the word, should be read as a written defense of a thesis in the (more or less) archaic sense of the word, namely:

> Continuum population dynamics, which is based on the concept of population density as a function of arbitrarily many continuous variables, allows a mathematical treatment sufficiently simple and general to provide a rigorous framework for theories of plant population dynamics.

This claim can be seen to have the following component claims:

(1) That plant population dynamics admits of a general framework, and by implication that such a framework is desirable.

(2) That the general continuum model of population dynamics provides such a framework.

(3) That this framework is rigorous, in the sense that arguments can be fully analysed.

(4) That continuum population dynamics is adequately characterised by population density as a function of arbitrarily many continuous variables.

My strategy for defending these claims is first to argue for the desirability of a rigorous framework in plant population dynamics, then to show by construction that such a framework is possible, and then to show by two examples that such a framework is indeed useful. Thus this thesis has five chapters. In this introductory chapter I outline the advantages of a rigorous framework of plant population dynamics for ecological theory, and I establish that the model of a population as a continuously varying density is general enough to cover all cases of interest in plant population dynamics. Then in chapter 2, I discuss the general continuum model for population dynamics, formulating it as the solution to an equation in two forms, the one entirely local and the other in a certain sense non-local. I then prove the existence of unique solutions for a class of models general enough to embrace all initial conditions, but regrettably requiring all constitutive relations to satisfy Lipschitz conditions. However, the theorem does cover all the models used later, with one exception in chapter 3 for which I supply *ad hoc* arguments. The chapter ends with a discusion of the algorithm I used for finding the solution over a finite time interval. The next two chapters are devoted to examples of how a rigorous framework can illuminate issues in theoretical debate and serve to guide future research: in chapter 3 I report on simulations of even-aged monocultures, which suggest that the shape of the size distribution may characterise the type of size dependence in mortality and somatic growth, while in chapter 4 I discuss the relative importance of recruitment, mortality, and somatic growth for the coexistence of shrubs with disjoint generations, as suggested by simulations. Finally, in chapter 5 I review the approach taken in this thesis, and consider how it may usefully influence future research.

1.1 Background to plant population dynamics

As will become clear below, mathematical theory is not particularly well integrated into the study of plant populations, nor indeed into ecology as a whole. It is useful to review briefly how ecology as a separate science arose, and how mathematics came to play its role.

The earliest scientists called themselves natural philosphers, and until the nineteenth century people like Linnaeus and Hooke would not have thought of themselves as engaged in two different careers. However, specialities began to separate out, and by 1830 one could reasonably talk about chemists, such as Davy and Dalton; physicists, such as Ohm and Oersted, and natural historians such as Darwin. In the course of the nineteenth century, natural history itself specialised, so that by the middle of the century zoology and botany appear as quite distinct sciences (see for instance the account of the German precursors of plant ecology by Cittadino [33]), each of which separately developed taxonomy, physiology and palaeontology, with differences in style and approach as well as content. I mention this pattern of repeated division to stress that when ecology arose, it was as a specialisation within the existing fields. Despite claims that the scope of ecology includes all organisms, it is perfectly reasonable to talk about "plant ecology" and "animal ecology", since they arose separately and to a large extent remain separate. Ecological schools of thought have by force of descent been concerned with either animals or plants.

As far as mathematics is concerned, it is very important to realise that during the formation of ecology as a science in the 1890's, mathematicians had no role at all, nor did mathematics of any kind arise (see for instance Kingsland [90, chapter 1] and Cittadino [33, chapter 10]). In fact, quantitative issues, when mentioned at all, were not pursued. Harper [79] mentions an interesting case in point: Nägeli's 1874 paper on mathematical modelling of plant invasions, was (with one minor exception) "ignored in the development of a science[1] that gained no significant momentum until 50 years later". Mathematics were not part of the way one studied ecology, and innovations using mathmatics only

---

[1] By "science" Harper means population biology.

entered ecology through unusual circumstances. The first example was when Pearl introduced the logistic equation to ecology in 1920. He had not been an ecologist, but after the personal tragedy of losing by fire the data and other documents accumulated over several decades, he turned to population studies. The resistance which such mathematical arguments encountered has been well described by Kingsland [90], and I do not repeat it here. The second example of the external source of mathematics in ecology is the work of Lotka and Volterra (whose first publications were in 1925 and 1926, respectively, see Scudo and Ziegler [133]). Volterra had a high reputation as a physicist and Lotka had no real standing as a scientist. Their work was entirely abstract mathematics, and hence foreign to the ecology of their time. It was through the experimental work of Gause [64], published in 1934, that the possibility of such a simplified look at species interactions became known. Although this led within a decade to the widespread adoption of the "principle of competitive exclusion", the mathematical inspiration behind these experiments was indifferently received—see for example the 1935 review of Gause's work by Park [121]. Nowadays, introductory texts do use a mathematical formalism to teach this concept, but always with a liberal dose of illustrative data, and it is my impression from the informal talk of ecologists that concepts like "competitive exclusion" and "niche" have their basic meaning not in mathematical formalism but as generalisations of the real situations (necessarily few) that are familiar to the ecologist concerned.

One should not think that mathematical studies of population dynamics started with Pearl, Lotka and Volterra in the 1920's. On the contrary, at the time a very well elaborated body of mathematics existed, which was in daily use and provided extensive employment for mathematical talent. I am speaking of actuarial science, as used in the insurance industry. According to Gross [68], its roots stretch back to Stevin's publication of compound interest tables in 1585, while the first mathematically based insurance company (the Equitable) was founded in the middle of the eighteenth century. Remarkably enough, the intense mathematical activity concerning human populations has found very little audience among those concerned with other populations[2], so that it comes as a surprise that Gompertz, whose work of 1825 [65] is now honoured by an

---

[2] Kingsland mentions a few exceptions [90, page 94].

equation bearing his name, was an actuary, as does the fact that Lotka did his later research in human demography, employed by an insurance company.

It may be argued that the quality and quantity of data available to actuaries could not have been gathered by ecologists, and that therefore the mathematics of insurance should be irrelevant to them. This would not be true of all organisms, even in the 1920's. The basic data are ages at death, and for trees these could have been gathered. More to the point is that life expectancies would not have interested ecologists then, nor does it interest them now. One might plausibly suggest that it should, and that for example a 60-year record of life expectancies of a tree species in a particular boreal forest might reveal much in terms of climate, pollution, and other factors in mortality. However, the concerns of ecologists are not those of actuaries, and the kind of unusual event that brought about the interventions of Pearl, Lotka and Volterra simply did not happen to someone versed in actuarial science.

This history explains firstly why mathematical ecology has remained a speciality within ecology, strongly biased towards animals, and secondly why it exports to the mainstream its results as synthesising concepts rather than as forms of argument or methods of research. Examples of such concepts are "competitive exclusion" and "niche" (as mentioned above); "r- and K-selection"; "island biogeography" and "evolutionarily stable strategy". This has some unfortunate effects: a mathematical theorem only applies when its assumptions are true, but in a synthesis these assumptions are never included in full detail, so that the synthesising concepts are usually applied far beyond the scope of the axioms in terms of which the theorems were proved. On the other hand, the generality of the mathematics may pass unnoticed when the synthesising concept strongly suggests particular biological meanings.

At this point it is helpful to distinguish between what may be termed population studies and theories of population dynamics. In order to do so, I now elaborate a distinction between descriptive and axiomatic theories. This is based on Peters' [124] use of the term "tautology" in discussing ecological theory. He contrasts tautology with theory, claiming that only the latter involves assertions about what is the case in reality.

I prefer to emphasize the difference between theories that proceed from assumptions about reality and theories that rely on data.

A descriptive theory seeks to condense a large body of information, often numerical data, into a small set of statements, often equations. An axiomatic theory, on the other hand, is a set of axioms together with deductions from these axioms and specifications as to how to decide when the theory applies. To claim that a descriptive theory is good is simply to claim that it summarises a lot of information. Many theories may be used to describe the same or closely similar bodies of information; the criteria for choosing between them will change as scientific values change. The claim that an axiomatic theory is good is much stronger: it is to claim that the axioms are widely applicable and that many useful deductions can be made. Moreover, axiomatic theories furnish ways of reaching conclusions in the absence of data. While many biologists may distrust it, using axioms to extrapolate beyond available data has been the basis of much successful innovation in engineering (though of course also a fair amount of failure). As for choosing between alternative axiomatic theories, this can be done more decisively than in the case of descriptive theories: if they do differ in a significant way, they must differ in some of their conclusions. By experiment, one can decide which theory to favour.

As an application of this distinction, consider the critique by Hall [72] of traditional mathematical ecology on the basis that the classical equations lack the support of a significant body of data. Now, the classical equations were axiomatically derived, and despite the common practice of many ecologists they should not be expected to fit data unless the assumptions underlying the equations can be expected to hold in the case of a particular data set. In terms of the distinction advanced above, the critique is a plea for axiomatic theory to derive the descriptive theory, and is therefore misplaced: Hall should either argue that deductions from these assumptions cannot provide any useful insight into ecological phenomena, or that s/he prefers descriptive theory and can provide better equations for such purposes.

The attraction of axiomatising is strong: the security of rigorous deductions (see below); the continuity of reason in the face of novelty; the possibility of using mathematics to make applications routine. However, biologists have generally found axiomatics unsatisfactory (see Peters [124] and chapter 3 below). In large measure, this is because acceptably detailed axiom sets seem to be too complicated to yield useful deductions easily—this is usually presented as a trade-off between realism and tractability, following Levins [103]. In chapter 3 I refer to the so-called self-thinning law—a good descriptive theory which attracted much effort at axiomatisation. I think in general this is not a good idea. While good descriptive theories do hint at regularities, which axiomatic theorists would be foolish to ignore, it is equally foolish simply to aim at deducing a fitted equation. In fact, the best equation for describing a given set of data may sometimes be known to be wrong (see Linhart and Zucchini [106, chapter 1], for a good example). This is because simplistic descriptions often do better at summarising data than more complicated ones, due to the so-called bias-variance trade-off, which deserves to be better known among biologists. Unfortunately, one reason it is not well known is that a simple exposition is hard to find—I could not obtain a suitable reference. The bias-variance trade-off is one of the informal rules of thumb current among statisticians, and roughly it occurs because as the number of parameters in an estimator increases, the gain in reduced bias is eventually offset by the increase in variance.

On the other hand, descriptive theories, particularly when qualitative, may attempt to unify causally disparate phenomena, and axiomatisation can succeed only once the separate explananda are identified. For example, while the descriptions of genetic variation and heritability are complex and difficult (e.g. gene flow in plants), the recognition that sometimes one has two alleles at a one locus in a diploid organism was a fruitful starting point for mathematical genetics. Similarly, working axiomatically may lead one to very good descriptive theories; the work of Ford and Diggle [63] and Lloyd and Harms [107]) on self-thinning are cases in point.

I can now return to the distinction between population studies and theories of population dynamics. It is meant to identify two largely distinct bodies of work. I claim that theoretical population dynamics is axiomatic theory, and has remained an isolated speciality in ecology. Isolated, but not disconnected: theoretical population dynamics, like other mathematical theories, communicates via synthesising concepts with mainstream ecology, where these concepts are given meaning not within the theoretical context but within the field of population studies, which is marked by a plethora of empirical experiments, often designed in terms of some theoretical construct. These concepts may have become very influential, but on an empirical rather than an axiomatic basis. Descriptive theories are less frequent, though in the field of managing environmental resources such as fisheries they are much used. The division between population studies and theoretical population dynamics is also reflected in the literature. Of the first fifty names in the author index of Begon and Mortimer's introductory "population studies" text [9], only eight appear in the author index of Levin*et al.*'s introductory "theoretical population dynamics" text [102, note that it is based on lectures delivered in 1986].

As noted above, mathematical ecology arose within animal ecology, and has remained heavily biased towards zoology. Plant ecology may boast "the earliest scientifically significant paper on populations", but it was slow to develop population studies, and even slower to adopt mathematical arguments (Harper [80]). Plant ecology developed largely along the natural history paradigm, and for the most part its theoretical component was descriptive; this is true even for quantitative work such as Braun-Blanquet tables and other ways of statistically identifying plant communities, and pattern analyses. Thus in contrast to animal ecology, plant population studies historically preceded theoretical population dynamics. The fact that animal-oriented theories have by default been applied to plants has led to repeated calls for the development of plant-oriented theories, for example from Harper [80], van der Meijden [146] and Crawley [43], usually with clear statements as to why axioms appropriate for animals are inappropriate for plants. For a notable attempt to extend animal-based theory to plants, see Pacala [119].

1.2 Plant population studies

From the above it should be clear that in this thesis I hope to contribute to axiomatic theory. Accordingly, there is no need for extensive discussion of population studies beyond an overview of its history and how it relates to theoretical population dynamics for plants.

Recording the biomass and/or number of various plant species, presumably at first with a view to advising farmers on grazing, fertilisation and the like, has been going on since the middle of the last century in England and the beginning of this century in the USA (Crawley [43]). However, the definitive work on the subject was published only in 1977, when Harper's *Population Biology of Plants* appeared [80], and from the bibliography of this book it would appear that not much work specifically concerned with the numbers or biomass of plant populations was published prior to 1950. Since then, of course, there has been an abundance of studies, which has steadily increased in rate of supply, so that it is doubtful that a similarly comprehensive survey would now be possible, even covering only the 15 years since 1977. However, broad overviews such as those by Crawley [43] and Czaran and Bartha [46] have appeared.

Crawley's [43] approach has much in common with Harper's book: in both cases the concern is firstly to classify the patterns of plant population dynamics as observed in the field and laboratory and secondly to call for theoretical development, stressing those aspects of plants that render traditional animal-oriented theory inapplicable. Both works are organised by what they call models: a diagram in the case of Harper [80, p 29] and an equation for Crawley [43, p 127]. These "models" are not used to any quantitative purpose, but merely serve to classify phenomena, and in both cases there is an implicit assumption that in any given situation one category or other will dominate the dynamics, so categories are never considered in combination. Harper's categories are stages in the life history of plants, while those of Crawley are stabilising processes such as limiting and refuge effects. Reflecting on this similarity, one might conclude that very little population dynamic theory specifically for plants was developed in the years 1977–1990. As Czaran and Bartha's review make clear [46], this would be mistaken. They stress one class of model, namely that in which the limited distances over which plants interact are explicitly included, and classify these into three types. Thus Czaran and Bartha [46] share with Crawley [43] a focus on population dynamic processes. However, the former admit to being concerned with only one problem, namely the duration of plant populations, which can be seen as modelling the outcome of dynamical processes, while the latter is concerned much more broadly with patterns as they unfold in time. (In mathematical terms, they are concerned with asymptotic and transient phenomena respectively). Hence there is very little overlap between the two papers—indeed, only two of the references among the 36 cited in Czaran and Bartha [46] also occur in among the 73 cited in Crawley [43]. To this one should add Harper's observation [80] that extensive theory and experiment by agronomists are being ignored by ecologists; this still seems to be true, at least in population studies. Hence it appears that the field has become deeply fragmented, and the call by van der Meijden [146] for bridging the divisions seems more apt than Crawley's [43] suggestion that more theorising is needed. Indeed, theorists may feel that the models already exist (see below) and that experimenters have been tardy in taking them up.

1.3 Theoretical plant population dynamics

Theories specifically addressing the population dynamics of plants are often seen as a recent development (Southwood [141]), and there is no extended general treatment of theoretical plant population dynamics in the literature. This is not to say that no general models have been published, only that such models have been presented briefly, as suggested points of departure, rather than as fully developed theories.

Pacala [119] formulated a general model that specifically allowed for the limited interaction distance of plants; however in his analysis he considered a version simplified by additional assumptions to the point of becoming the Lotka-Volterra equation for organisms with limited interaction distance. The partial differential equation to be presented in chapter 2 has been used for detailed studies (Feldman and Curry [61], Kohyama [94]) and for general study of some diffusion models (Gurtin [69], Busenberg and Ianelli [19]). These diffusion models seem hard to apply to plants, since the peculiarities of seed dispersal are not modelled—indeed, the classical diffusion equation is often taken as the starting point (Bertsch *et al.* [11]), with the implied Gaussian distribution of velocities (and hence, at any fixed time, of dispersal distances) whereas seed rain is more appropriately modelled as a Poisson distribution (Pacala [119]). Most models of structured plant populations used in analysing data (i.e. most descriptive models) are matrix models (Caswell [23] gives many examples); such models are much less popular for theoretical work (but see Law [100], Solbrig *et al.* [140], Caswell [23] and Eriksson [55]). A number of models for populations with continuous size structure were proposed as derivations of a law of self-thinning—some of these are considered briefly in chapter 3.

The most popular strategy for proposing general population models that explicitly take into account plant size and position is to let the computer do the work. That is, some model for the effect of neighbours on each plant is proposed and then an initial population of plants (or patches of plants) is followed individually, for example Weiner and Conte [153], Pacala and Silander [120], Bonan [13], Coffin and Lauenroth [37] and Kareiva and Anderson [88]; for a review see Weiner [152]. This reliance on computers is unsatisfactory for theoretical development, since the results of such simulations have the character of data rather than deductions in that results are merely tabulation of a set of trials. One cannot develop axiomatic theory in this way, and the extrapolation of results to parameter situations not covered by the simulation seems rather uncertain. Rees and Long [128] do go some way towards a basis for generalisation by comparing the outcomes of two sets of models, the one a set of stochastic models and the other a set of deterministic models that correspond to modelling the mean of the stochastic case.

The work of Tuzinkevich [145] is unusual in that it starts with an integro-differential equation and ends as a time-discrete model involving integrals of the population density at the previous time; it fails to give much insight into plants because it uses classical growth and species interaction constants. I would also like to mention the work of Tilman [142, 143], which represents an ambitious project to unify plant ecology through population dynamics. So far, the only aspect of this work specific to plants is the root-stem-leaf trade-off discussed in Tilman [143], where again the results are data from computational experiments rather than deductions.

Thus, there is a multitude of largely incompatible models of varying degrees of generality, incompatible in the sense that they do not share variables and/or parameters, so that one cannot see the models as extensions, generalisations or specialisations of each other. Despite this profusion, they address remarkably few questions, Czaran and Bartha [46] admitting only the problem of population persistence, and if one includes the literature on crops, the further problem of yield appears (e.g, literature on constant final yield reviewed in Harper [80], Feldman and Curry [61]). In this sense, the impression that theory has served plant population biology badly (Crawley [43], van der Meijden [146]) seems to be justified.

The problem is not the absence of general models, but that it is hard to relate them to the detailed questions that interest experimenters. As an example, consider Pacala's [119] extension of the Lotka-Volterra equations to sessile individuals with limited range of interaction, which yields the result that this promotes coexistence of species. How is one to decide whether a boundary between two species in the field can be understood in terms of this theory? It is conceivable that the neighbourhood distance of each species can be measured, but the species interaction parameters seem beyond experimental grasp.

Other models for particular species and communities include transition models concerned with patterns of species richness (Yeaton and Bond [164], Hilbert [82], Bishir and Namkoong [12]), and other coexistence models (Kohyama [95], Iwasa *et al.* [86]). Not all of these are presented as axiomatic theories, but they could each be so presented, with parameters that are to be estimated in any particular case. The difficulty here is the inverse of above: these specific models lack a framework in which they can be compared, and so each is an *ad hoc* construct and general insights remain scarce and in any case cannot be put in quantitative terms.

Thus the weakness of theories of plant population dynamics is not so much a mismatch between theorists and experimenters as a jumble of theories, in which the divide between general and particular theories is particularly problematic, forming a divide across which little transfer takes place. Perhaps what is lacking is a general synthesis. The continuum approach elaborated in this thesis may be the way to achieve such a synthesis. It has already been used for a case study (Kohyama [94, 95]) and in chapter 2 I show that one can express the full variety of models in its formalism. I argue in chapter 2 that the axioms employed are general for plants; this would mean that any model whether general or particular may be expressed in this formalism. Even descriptive models suffer from a lack of a common framework, as can be seen from the recent debate between Rees and Crawley [126, 127] and Silvertown [137] on the effect of a possible lack of minimum size for initiating reproduction. As for discrete models, the approach of Bishir and Namkoong [12] may well prove generalisable, in the sense that as the discretisation parameters approach zero, their model should become equivalent to the general continuum model. If so, all of theoretical plant population dynamics may be amenable to a unified framework, but I do not pursue this interesting possibility further in this thesis.

1.4 The role of a rigorous framework

Above, I hinted that axiomatic theory is attractive because of rigour. The particular components of rigour are:

—Clear definitions and unambiguous statements.

—Explicit assumptions.

—Formal, or formalisable, deductions. (Note that validity is not assumed: formal proofs still need to be checked!)

The advantage of explicit assumptions is of course that, by scrutinising each assumption in turn, one can assess the applicability of a theory in detail. Moreover, modification of the theory can be systematic, in that any modification can be reduced to changes in one or more of the assumptions. Now, in some theories the axioms are so many as to overwhelm the comprehension of the reader. By contrast, the theory offered in this thesis aims to be as general as possible by using only those axioms that surely hold for all plant populations. In chapters 3 and 4 the usefulness of such a general rigorous theory is illustrated.

The advantages of formal deductions are manifold: more reliable and more precise arguments, flushing out fallacies, and clarity about the necessity and sufficiency of assumptions.

Thus it can be seen that a rigorous framework may greatly benefit a science; as an example I cite the effect of theoretical thermodynamics on what had been known until then as "theory of heat" (Ihde [84]). However, it is not always possible to achieve, or wise to attempt, a rigorous framework. In any field of human endeavour, there are things one does right for the wrong reason, and in a rigorous framework by noticing these wrong reasons one might be deterred from the right actions. To put it another way, a field of science might embody considerable knowledge in somewhat vague concepts, and clarification in itself would simply result in loss of knowledge, unless clarity about concepts brought with it clarity on gaps in knowledge and on how to fill those gaps effectively. Thus I do not quite agree with Loehle [108] that theory tenacity and confirmation bias are at fault in population dynamics; until a good synthesis is available there is considerable intellectual loss in converting to a purist line. In this thesis, I attempt to show by example that the time is ripe for synthesis.

1.5 The continuum assumption

Any theory of population dynamics is based on two elements: the representation of the population by a mathematical symbol, and an evolution equation[3] for this symbol. Hence, if the framework for population dynamics advanced in this thesis is to be general there are two conditions to be met:

—represention of any population as a density which is a function of continuous variables should be general for plants

—there should exist an evolution equation for this density which is gen-

---

[3] This could be a differential equation, in which case the initial condition and boundary conditions are given separately, or else an integral equation, or some combination, or the discrete versions of these i.e. difference and sum equations.

eral for the three population processes in plants (births, deaths and dispersal), and which accomodates the relevant features peculiar to plants (such as size plasticity and limited interaction distance; see Crawley [43] and van der Meijden [146]).

The appropriate place for arguing the second condition is chapter 2. The questions of existence and uniqueness of solutions to the evolution equation are vital—without positive answers to these questions the equation is meaningless—but can only be addressed after the mathematical theory has been elaborated. Here I show that the first condition is met, namely generality for plants of representing the population as a density over continuous variables. Let me start with attributes which can be represented by continuous variables, such as age, mass, height and position. Since in describing a population, one can consider only a finite number of attributes, finite-dimensional Euclidean space is sufficient to represent the character space of populations (for characters representable by continuous variables). It is convenient to represent the age a separately, and a vector $\mathbf{x} \in \mathbf{R}^k$ then specifies the other attributes; I shall refer to $\mathbf{x}$ as the structure vector. Thus an individual is completely characterised by the pair $(a, \mathbf{x})$, the *i*-state of Metz and Diekman [113]. I shall use the symbol $\rho(a, \mathbf{x}, t)$ to mean the density of those individuals of age a at time t with attributes $\mathbf{x}$. Thus ρ is the *p*-state of Metz and Diekman, and specifies the population as a density over the possible *i*-states. In this thesis $\mathbf{x}$ is a member of a bounded subset Ω of $\mathbf{R}^k$, and on the boundary of Ω the rate of change of $\mathbf{x}$ is zero, so that $S = [0, \infty) \times \Omega$ represents the set of all possible states. The boundedness of Ω means that no attribute can ever go to infinity; this does not limit generality, since unbounded attributes can always be represented by the reciprocal of a variable that goes to zero, provided the attribute itself is never zero (which can easily be arranged by the addition of an arbitrary constant—to recover the biologically meaningful quantity, this is subtracted out). However, there do not seem to be any ordinarily used attributes that require an unbounded interval (other than age, which is separately represented).

So much for attributes that can be modelled by continuous variables. There are others, such as species and sex (for dioecious plants), that cannot be so modelled as they do not vary continuously; moreover, they are often constant for the duration of an individual's life. In such cases, it may be convenient to define subpopulations, and consider

$$\rho(a, \mathrm{x}, t) = [\rho^1(a, \mathrm{x}, t), \rho^2(a, \mathrm{x}, t), \ldots, \rho^n(a, \mathrm{x}, t)]^{\mathrm{T}},$$

to be a vector, where every $\rho_i(a, \mathrm{x}, t)$ represents a subpopulation.

The total population P at time t is then represented by

$$P(t) = \int_0^\infty \int_\Omega \rho(a, \mathrm{x}, t) d\mathrm{x} da,$$

or

$$P(t) = \sum_{i=1}^{n} \int_0^\infty \int_\Omega \rho^i(a, \mathrm{x}, t) d\mathrm{x} da,$$

as the case may be; frequently the total in each subpopulation is what is required, in which case P is an n-dimensional vector.

Since every attribute can be modelled by either a discrete or a continuous variable, it is clear that every plant community can be described by a vector of population densities, each element being a function of a common set of continuous variables. This establishes the claim of generality for the representation of plant populations used in this thesis.

Chapter 2

General theory of continuum population dynamics

2.1 Introduction

The basis for non-linear theory of population models structured by continuous variables was laid by Webb [151], in which the nonlinearities are completely general and the population is considered as a vector of subpopulations, and by Tucker and Zimmerman [144], in which a scalar population has a completely general structure and some non-linearity. Separately, these are not sufficient to prove unique existence of the models I investigate in the case studies below, so in this chapter I combine them into one model (and even this model is not adequate to deal with all the models considered in chapter 3, where an *ad hoc* argument is necessary). Since my interest is in transient rather than asymptotic behaviour (which is discussed in many places, e.g. Cushing and Saleem [45], Di Blasio and Ianelli [49], Elderkin [52], Webb [151], Weinstock and Rorres [155]), I provide proof only of the existence and uniqueness of solutions, and discuss an algorithm for approximating the solution over a finite time interval. It turns out that the approach to analysing the problem in Webb [151] can be generalised to the case of arbitrarily many structure variables. On the other hand, the proof of existence of solutions in Tucker and Zimmerman [144] is flawed (see the appendix to the thesis), so this adaptation involves more than mere routine.

As mentioned in chapter 1, any theory of population dynamics can be seen as an evolution equation for the symbol representing the population. In the case of the symbol $\rho(a, \mathbf{x}, t)$ used in this thesis, the equation can be derived as a partial differential equation as follows: the population density is regarded as a vector depending on the continuous variables a, $\mathbf{x}$ and t. This may be regarded as the density of a generalised fluid [167], and well-known arguments from fluid mechanics [104] say that the rate of change of the mass (in a generalised sense) contained in an infinitesimal volume is

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial a} + \operatorname{div}(\mathbf{v}\rho),$$

where $\mathbf{v} = \dot{\mathbf{x}}$ is the vector of the rate of change in the state.

Of course, if mass is conserved this quantity is zero, giving the continuity equation of fluid dynamics. However, generally in population dynamics it is not conserved due to migration, births, changes of life stage (for organisms with complex life histories), and deaths. Changes of state are covered by $\mathbf{v}$ and need to be specified by a further set of models, and since births concern $\rho(0, \mathbf{x}, t)$ they are properly treated by a boundary condition concerning $a = 0$. The terminology introduced by Metz and Diekmann [113] is convenient here: $(a, \mathbf{x})$ is the i-state (for individual) and $\rho(\cdot, \cdot, t)$ is the p-state (for population). The change in an infinitesimal volume over an infinitesimal time interval depends on the i-state $(a, \mathbf{x})$ of that volume, and also on the current density $\rho(\cdot, \cdot, t)$ of all states. This can be symbolised by $G(\rho(\cdot, \cdot, t))(a, \mathbf{x})$, which I shall call the transition function: for a given p-state $\rho(\cdot, \cdot, t)$, the rate of change in the mass at every i-state $(a, \mathbf{x})$ is given by G. Then

$$\text{(LB)} \qquad \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial a} + \operatorname{div}(\mathbf{v}\rho) = G.$$

As it is derived from a balance condition, and applies locally, (LB) may be termed the balance law for continuum population dynamics in local form.

The boundary condition can likewise be as be written as

$$\rho(0,\mathrm{x},t) = F(\rho(\cdot,\cdot,t))(\mathrm{x}), \qquad t \in (0,T].$$

I shall refer to F as the birth function. Note that because F is undefined at $t = 0$, compatibility of F with ρ_0 is not required, and that hence the solution may have discontinuities arising from $\lim_{t\downarrow 0} F(\rho(\cdot,\cdot,t))(\mathrm{x}) \neq \rho_0(0,\mathrm{x})$.

Finally, the initial condition is $\rho(a,\mathrm{x},0) = \rho_0(a,\mathrm{x})$.

The history of (LB) and its associated conditions is somewhat obscure. In 1926, M'Kendrick [117] derived a linear version (with only age structure) as a minor part of a paper mostly devoted to statistical issues. This paper seems to have passed into obscurity, and in 1959 von Foerster derived the same equation independently [148], and the linear version with additional structural variables was derived apparently independently by Sinko and Streifer [139] and Bell and Anderson [10]. Subsequent authors usually fail to mention some or most of these early papers (for example, Rotenberg [130] gives essentially the same equation as Sinko and Streifer [139] and Bell and Anderson [10] but does not cite either paper). As a result, when referred to by the name of an author, (LB) may be called the M'Kendrick, the von Foerster, the M'Kendrick-von Foerster, the Sinko and Streifer or the Bell and Anderson equation; this does not exhaust the names it has been given!

In 1974, Gurtin and MacCamy [71] took the major step of introducing non-linearities into the equation, and this has resulted in a large number of papers (for example, the Scince Citation Index Cumulation 1975–1979 has 21 papers that cite it). As mentioned above, the degree of structure was increased by Tucker and Zimmerman [144] to x of arbitrary dimension, and by Webb [151] to vector interpretation of ρ. The degree of non-linearity was generalised by Webb [151] to dependence on $\rho(\cdot,\cdot,t)$ directly. A more complete history and comprehensive bibliography of the earlier work may be found in Metz and Diekman [113].

These equations constitute a framework for constructing models of population dynamics. Indeed, by the generality of the derivation, any continuum model of population dynamics is of this form. To construct a model, it is necessary to specify the constitutive relations F, G and $\mathbf{v}$— i.e. to develop formulae describing births, deaths, and chages in attributes. No generally valid formulae are known, though important applications such as actuarial science and fisheries management have developed simple approximations that have so far proved practicable over limited parameter ranges and time spans.

The first question is to prove that these equations can indeed be used to model real populations in the sense that solutions to the equations exist uniquely for given constitutive relations (i.e. F, G and $\mathbf{v}$), and initial condition ρ_0. In all the published work I've seen, the question is first reformulated in terms of integral equations, and I shall do the same. I follow Webb's procedure [151], modified in a way similar to the approach of Tucker and Zimmerman [144] to allow for a population characterised by the structure vector $\mathbf{x}$ as well as age. For the structure variable $\mathbf{x}$, the rate of change $\mathbf{v}$ specifies a constitutive relation that must be satisfied by $\mathbf{x}$ as a function of time, so the i-state is found by solving the following set of ordinary differential equations:

$$\text{(Chars)} \qquad \begin{cases} \dfrac{da}{dt}(a,\mathbf{x},t) &= 1 \\[2ex] \dfrac{dx_i}{dt}(a,\mathbf{x},t) &= v_i(\rho(\cdot,\cdot,t))(a,\mathbf{x}), \end{cases}$$

Note that Tucker and Zimmerman give [144] almost the same set of equations, the difference being that above the v_i depend directly on the current p-state $\rho(\cdot,\cdot,t)$, rather than on some functional $Q(t)$ of same. The only salient property of $Q(t)$ used in their argument is that it is continuous, but in their proof of this fact they assume that ρ is itself continuous in t. I prefer the more general statement, and appeal as they do to the continuity of $\rho(\cdot,\cdot,t)$ with respect to t and of v_i with respect to all its arguments and to the boundedness of v_i and all its derivatives and the assumption that v_i are zero on the boundary of Ω for the claim that, given $\rho(\cdot,\cdot,t)$ for $t \in [0,T]$, a and $\mathbf{x}$, there is a unique solution to (Chars) on the interval $[0,T]$; at this point T is just some (finite) positive number, and constraints on its size will emerge later. These assumptions are formalised in (Hyp 4) below. I shall denote this solution by $\phi_{\mathbf{x}}(s)$ and note that $s \in [-\tau, T-t]$, where $\tau = \min(a,t)$. These solutions constitute a flow $\Phi = \{\phi\}$ on S, (where as in chapter 1, $S = [0,\infty) \times \Omega$), and later it will be seen that they are the characteristic curves of the partial differential equation for structured population dynamics. The proofs below also require Φ not to be degenerate, that is, they require $\phi_{\mathbf{x}}(s) \neq \phi_{\mathbf{y}}(s)$ for any $s \in [-\tau, T-t]$ whenever $\mathbf{x} \neq \mathbf{y}$. This is achieved in the theory below as follows: v is required to be positive for all $T < \infty$ and for any ℓ in L_1 to be strictly monotone with respect to all components x_i of $\mathbf{x}$ in Ω. In chapter 3, some of the models violate this assumption, and are dealt with in an *ad hoc* fashion there.

Before proceeding I define some function spaces. Set $L_1 \overset{\text{def}}{=} L_1(S, \mathbf{R}^n)$. The density ρ is a vector of dimension n, so $\rho \in L_T$, defined by $L_T = C([0,T]; L_1)$, where the interval $[0, \infty)$ is the set of all possible ages, while $\Omega \subset \mathbf{R}^k$, the set of all possible structural attributes, is a bounded set in k-dimensional space, and as in chapter 1 $S = [0, \infty) \times \Omega$ is the set of possible i-states. L_1 itself is the set of equivalence classes of almost everywhere equal functions in the Banach space of Lebesgue integrable functions $\ell : S \rightarrow \mathbf{R}^n$ with norm

$$||\ell||_{L_1} = \int_0^\infty \int_\Omega ||\ell(a, \mathbf{x})||_1 \, d\mathbf{x} \, da,$$

where the norm $|| \cdot ||_1$ on $\mathbf{R}^n$ is defined by $||\mathbf{x}||_1 = \sum_i |x_i|$.

Strictly speaking, $\rho(t)$ is an element of L_1, that is, a function from S to $\mathbf{R}^n$; the image of $(a, \mathbf{x}) \in S$ should therefore be written $\rho(t)(a, \mathbf{x})$. By lemma 1 below (page 34), one may regard $\rho \in L_T$ as equivalent to some $\rho' \in L_1([0,\infty) \times \Omega \times (0,T); \mathbf{R}^n)$. Abusing notation for the sake of avoiding an avalanche of symbols, I shall use $\rho(a, \mathbf{x}, t)$ to refer to the images of both ρ and ρ', and $|| \cdot ||_{L_1}$ for the norm of all the L_1 spaces that are used in this thesis. Finally, the norm on L_T is a supremum: for $\rho \in L_T$, $||\rho||_{L_T} = \sup\limits_{t \in [0,T]} ||\rho(\cdot, \cdot, t)||_{L_1}$. I shall sometimes denote $||\rho||_{L_T}$ simply by $||\rho||$.

The birth function F and the transition function G are from L_1 spaces, and $F : L_1 \times \Omega \rightarrow \mathbf{R}^k$ while $G : L_1 \times [0, \infty) \times \Omega \rightarrow \mathbf{R}^k$. I shall usually need to work with a fixed $\ell \in L_1$, and hence need the norms on $F(\ell)$ and $G(\ell)$. Again abusing notation, these are both denoted by $|| \cdot ||_\Omega$, and defined as follows:

$$||F(\ell)||_\Omega = \int_\Omega ||F(\ell)(\mathbf{x})||_1 \, d\mathbf{x}$$

$$||G(\ell)||_\Omega = \int_0^\infty \int_\Omega ||G(\ell)(a, \mathbf{x})||_1 \, d\mathbf{x} \, da$$

In what follows, I will assume that (Chars) is satisfied, that is, I shall use Φ as defined above. The problem associated with (LB) can now be specified.

Problem P′

For given F, G and ρ_0, find $\rho \in L_T$ which satisfies

$$\text{(Births)} \qquad \rho(0, \mathbf{x}, t) = F\big(\rho(\cdot, \cdot, t)\big)(\mathbf{x})$$

$$\text{(LB)} \quad \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial a} + \operatorname{div}(\mathbf{v}\rho) = G$$

$$\text{(I.C.)} \qquad \rho(\cdot, \cdot, 0) = \rho_0(\cdot, \cdot).$$

This is the local form of the general continuum model for population dynamics.

I now turn to Webb's derivation [151] of the general evolution equation. It avoids partial derivatives, which may seem a little odd when one is accustomed to thinking in terms of infinitesimals, but will be seen to be both more general and easier to use in proofs than the balance law (LB) given above. In Theorem 1 I prove that, in a weak sense, the two formulations are equivalent. One starts from the rate of change of the whole population, in terms of the Newton quotient:

$$\begin{aligned}
\dot{P}(t) &= \lim_{h \to 0} h^{-1}\big\{P(t+h) - P(t)\big\} \\
&= \lim_{h \to 0} h^{-1} \left\{ \int_0^\infty \int_\Omega \rho(a', \mathbf{y}, t+h)\, d\mathbf{y}\, da' - \int_0^\infty \int_\Omega \rho(a, \mathbf{x}, t)\, d\mathbf{x}\, da \right\} \\
&= \lim_{h \to 0} h^{-1} \left\{ \int_0^h \int_\Omega \rho(a', \mathbf{y}, t+h)\, d\mathbf{y}\, da' \right. \\
&\qquad \left. + \int_h^\infty \int_\Omega \rho(a', \mathbf{y}, t+h)\, d\mathbf{y}\, da' - \int_0^\infty \int_\Omega \rho(a, \mathbf{x}, t)\, d\mathbf{x}\, da \right\}
\end{aligned}$$

Change variables in the second integral via $a' = a + h$ and $\mathbf{y} = \phi_{\mathbf{x}}(h)$, and in the first integral via $a' = a$ and $\mathbf{y} = \mathbf{x}$ to obtain

$$\begin{aligned}\dot{P}(t) &= \lim_{h\to 0} h^{-1}\left\{\int_0^h \int_\Omega \rho(a,\mathbf{x},t+h)\,d\mathbf{x}\,da\right\} \\ &\quad + \lim_{h\to 0} h^{-1}\left\{\int_0^\infty \int_\Omega \rho(a+h,\phi_{\mathbf{x}}(h),t+h)J_\phi(h) - \rho(a,\mathbf{x},t)\,d\mathbf{x}\,da\right\}\end{aligned}$$

Here, $J_\phi(h)$ is the Jacobian of $\phi_{\mathbf{x}}(h)$, that is, the determinant of any matrix that maps $\mathbf{x}$ to $\phi_{\mathbf{x}}(h)$. It should be reasonably clear that the first integral refers to processes at the boundary where $a = 0$, i.e. to births, and that the second integral refers the other processes i.e. deaths, migration and stage transitions. On this basis, the following equations then define the model mathematically:

Problem P

For given F, G and ρ_0, find $\rho \in L_T$ which satisfies

(Births) $$\lim_{h\to 0} h^{-1} \int_0^h \int_\Omega \|\rho(a,\mathbf{x},t+h) - F(\rho(\cdot,\cdot,t))(\mathbf{x})\|_1\,d\mathbf{x}\,da = 0$$

(IntBal) $$\lim_{h\to 0} \int_0^\infty \int_\Omega \|h^{-1}\left[(\rho(a+h,\phi_{\mathbf{x}}(h),t+h)J_\phi(h) - \rho(a,\mathbf{x},t))\right] - G(\rho(\cdot,\cdot,t))(a,\mathbf{x})\|_1\,d\mathbf{x}\,da = 0$$

(I.C.) $$\rho(\cdot,\cdot,0) = \rho_0(\cdot,\cdot).$$

These three equations will be referred to as (P). This is the non-local form of the general continuum model for population dynamics. As one may expect, (P) is in a weak sense equivalent to (P′), as made precise by the following theorem.

Theorem 1 *(a) A solution ρ to* (P′) *is a solution to* (P).

(b) If the solution ρ to (P) *is sufficiently regular to allow the existence*

of the limit $\lim_{h\to 0} h^{-1}[\rho(a+h), \phi_{\mathbf{x}}(h), t+h)]$ *to exist for all* $t+h \in [0,T]$, *then* ρ *is a solution to* (P′) *in the sense that*

$$\begin{aligned} \rho(0,\mathbf{x},t) &\overset{a.e.}{=} F(\rho(\cdot,\cdot,t))(\mathbf{x}) \\ \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial a} + \operatorname{div}(\mathbf{v}\rho) &\overset{a.e.}{=} G. \end{aligned}$$

Proof. The initial condition in both problems is the same. To prove (a), suppose ρ is a solution to of (P′). Note that by the problem statement, the birth condition trivially satisfies (Births). Morever, since the bounds on the integral in (IntBal) do not depend on h, one can move the limit under the integral sign, and by direct calculation (the assumption $\dot{\mathbf{x}} = \mathbf{v}$ remains in force)

$$\begin{aligned} \lim_{h\to 0} h^{-1}[\rho(a+h), \phi_{\mathbf{x}}(h), t+h)] &= \rho_a + \rho_t + \dot{\mathbf{x}} \cdot \nabla\rho + \rho \operatorname{div}(\dot{\mathbf{x}}) \\ &= \rho_t + \operatorname{div}(\mathbf{v}\rho). \end{aligned}$$

Hence such a ρ satisfies (IntBal).

To prove (b), suppose ρ is a solution to of (P). Since (Births) hold, and its integrand is positive, it must be zero except on a set of measure zero. Hence the first almost everywhere equality. Moreover, since (IntBal) holds (again moving the limit under the integral sign), one may apply the limit as in (a) above, and again the integrand must be zero except on a set of measure zero, so that the second almost everywhere equality holds. □

The theorem enables one to work with whichever formulation is convenient, since solutions to the non-local problem would fail to solve the local problem only via the biologically irrelevant lack of regularity on sets of measure zero.

2.2 The problem (P) is general for all deterministic population dynamics

By the claim that (P) is general is meant the following: by appropriate choice of the n components of ρ and the k components of $\mathbf{x}$, and by specifying the relevant $\mathbf{v}$, F and G, the dynamical processes of any population may be modelled in full detail. What is not claimed is that all possible models fit this framework—indeed some useful classes of models are excluded. In particular the existence theory below does not cover models with time lags or with differential operators.

In chapter one I have already argued that $\rho(a, \mathbf{x}, t)$ can represent any population. What remains is to show that (P) is general enough to accommodate the dynamics of any population. Firstly, (P) is deterministic and, insofar as one regards biological data as random variables, is suitable only for modelling the mean of such variables. Second, although there is no explicit reference to environmental variation, it makes no difference to the formulation or the proofs to include in the constitutive laws G, F and $\mathbf{v}$ dependence on a function $E(t)$ symbolising the state of the environment. Therefore environmental variation can be represented in the framework of (P) and (LB). However, doing so clutters up the formulae and environmental variation was not considered in the applications given in chapters 3 and 4, so this level of generality is not analysed in this thesis. I now consider the fundamental processes of population dynamics (births, changes of state, migration, deaths), and show that (P) in conjunction with (Chars) is sufficiently general to model them all.

The four population dynamic processes are variously accommodated in the proposed framework. All four processes affect the density. The derivation of both (P′) and (P) make clear that, in the absence of time lags, the time course of $\rho(a, \mathbf{x}, t)$ is governed by (LB) and more generally by (IntBal). In fact, (IntBal) allows jump discontinuities, and no biological phenomenon needs to be modelled by any function more pathological than that. Thus (IntBal) is a balance law which can apply to any population. It may seem more general for dependence on ρ at some or all other times to be included, but I argue that this is not so. The absence of time lags does not detract from generality: models with time lags were developed to cope with aspects of life history while remaining in the tractable mathematical context of scalar models and ordinary differential equations (Aiello *et al.* give some references in [2]). Because of the possibility of a vector of subpopulations, life history stages are covered by the above framework; any influences that seem important can be modelled without lags. On the other hand, by specifying that component ρ^j at present equals component ρ^i at some earlier time, time lags can explicitly be added to the framework. However, this would represent an additional and rather artificial constraint on the solution and the theory below does not cover existence proofs for this case.

A note on using both $\mathbf{x}$ and ρ as vectors of dimension k and n respectively is in order here. It is of course possible to consider membership of a subpopulation as an attribute, and to include this in $\mathbf{x}$, so that ρ is a map to the real numbers. In that case, at least one of the components of $\mathbf{x}$ would not be a real number, but instead an integer—or even merely an ordinal number—so such a component would have to be explicitly excluded from operations that rely on the continuum concept, for instance integration and differentiation. It is also convenient, as shown above by the discussion of time lags, to be able to refer to subpopulations in a direct way through an index, rather than through the value of a variable. For these reasons I prefer to follow Webb ([151]) in regarding ρ as a vector.

For both (P) and (P′), when combined with (Chars), there are three sets of functions which need to be specified for any given model. They are birth law F, the transition law G and the state evolution law $\mathbf{v}$. These are the so-called constitutive relations. There is also the initial condition. I discuss these below, showing that they are general for plants in the sense that any plant population dynamics can be modelled by specifying the appropriate constitutive laws and initial condition.

The transition law concerns changes in density due to mortality, migration and changes of subpopulation. Since in an infinitesimal volume, mortality of the i-th subpopulation can be taken as constant, it can adequately be modelled by a modulus, giving $\lambda(a,\mathbf{x},\rho(\cdot,\cdot,t))\rho(a,\mathbf{x},t)$ as a term in G. Migration is usually seen as two distinct processes: diffusion and immigration/emigration. In either case, this can simply be included in G, as a net balance between source and sink terms. For example, in the case of dispersal of seeds, this could be modelled by a function integrating, for every position, the sources over the whole area (this could easily include limited distances of dispersal and preferred directions). Diffusion is more usually modelled in terms of local gradients, but as mentioned in chapter 1 the classic diffusion equation is not especially suitable for seed dispersal; nevertheless, diffusion is just local migration, and the Laplacian operator can be considered as a source/sink expression. Changes of stage are modelled by the density of one subpopulation ρ^i increasing at the expense of another ρ^j; because mortality is modelled separately the gain equals the loss and the instantaneous transitions are modelled by a matrix with ergodic properties. Thus one can formally represent G by a sum of three terms:

$$G = (\mathrm{diag}\lambda)\rho + \mathrm{A}\rho + \mathrm{D}(\rho)$$

where λ is the vector of death moduli, A is the instantaneous transition matrix, and D is the diffusion/dispersal model. In general, D is an operator incorporating derivatives of ρ, but for plants I argue below that this is not needed.

By the same argument that yielded the death modulus, the birth law defined as a birth modulus by $F(\rho(\cdot,\cdot,t)(\mathbf{x}) \overset{\mathrm{def}}{=} \beta(a,\mathbf{x},\rho(\cdot,\cdot,t))\rho(a,\mathbf{x},t)$ is quite general. Incidentally, the birth law is the reason that population dynamics is mathematically difficult: it forms a boundary condition that is inherently non-local and nonlinear. This difficulty extends to computations as well; in fact, in some of the earliest analyses of algorithms, the births were assumed to be known in advance (Douglas and Milner [50], Kostova [96]).

From now on, I shall refer to λ as the mortality (or mortality modulus or the *per capita* mortality), and to β as the birth rate (or birth modulus or the *per capita* birth rate).

The state evolution law specifies how the attributes of an individual changes with time. In the approach adopted here, each of these attributes are quantified separately, and the state of an individual is, besides age, a vector in the set Ω of dimension k. The range of plausible k is quite large. At one extreme, one could attempt to specify the number of various cell types in an organism and their relations, so that transitions in cell type could then be specified in this law, for example modelling the metamorphosis of a single insect. This would require k of order 100. At the other extreme are the classical models with no structure at all. For plant populations one would usually want to classify the population into types by considering ρ as a vector of subpopulations (for example seeds and adults of each species), each with a fairly small set of attributes (such as size and position), so that the overall set of attributes would not be overly large. Transitions from one type to another are specified by the transition law; the state evolution law consists of specifying constitutive relations for the attributes. For example, the somatic growth law would specify how size changes, and one can imagine a similar law for microsite changes. Position on the other hand will usually be considered constant for adults, and subject to change for seeds.

Note the difference between diffusive and directed changes: the former affect the density rather than the attributes and so go into the transition law because one cannot follow individuals and the effects are randomly distributed across the population; while the latter specify the behaviour of individuals with specific attributes. As an example, consider dispersal of seed by wind. The distribution of air currents in space and time can be specified in detail only up to a certain level of accuracy. This defines an average around which there is bound to be some random variation. This process would be conveniently modelled by dividing the model into two parts. The structure variable $\mathbf{x}$ would include spatial coordinates, and the solution to $\dot{\mathbf{x}} = \mathbf{v}$ would specify the average air flow for each time and place, and hence a velocity field that prescribing the movement of seeds there, while the random component would be modelled as a diffusive process through the transition law.

Finally, the initial condition. If $\rho(a, \mathbf{x}, t)$ is sufficiently general to represent any population, then $\rho(a, \mathbf{x}, 0)$ suffices for the initial condition.

It is therefore possible to formulate a model for plant population dynamics that is completely general in the sense that it could represent all possible states and processes. However, it is another matter to ensure that the solutions to such a model exist and can be interpreted in a biologically meaningful way. For instance, even though exotic situations may exist in which meaning could be contrived, one would usually like to avoid negative densities and multiple solutions. On the other hand, one would like to include constitutive relations that are mathematically difficult, such as ones with jump discontinuities. Below, I impose Lipschitz conditions on all constitutive relations, and in this way ensure the existence of unique solutions. This is clearly not sufficient for modelling all situations of biological interest—see for instance chapter 3 below—but does serve to show that the fully general model can be tractable.

2.3 Reformulation and proof of unique existence of solutions

2.3.1 Preliminaries

Consider the cohort function on $[-\tau, T-t_0] \to \mathbf{R}^n$ (where τ is as before) defined by

$$w_{(a_0,\mathbf{x}_0,t_0)}(t) = \rho(a_0+t, \phi_{\mathbf{x}_0}(t), t_0+t),$$

which specifies the evolution of the density of those individuals of age a_0 and state $\mathbf{x}_0$ at time t_0. Now suppose that either $t_0 = 0$ or $a_0 = 0$. Then $a_0 \geq t_0 \Rightarrow t_0 = 0$, while $a_0 < t_0 \Rightarrow a_0 = 0$, and hence one can write ρ in terms of its values along characteristics, where each characteristic is uniquely defined by the point where it meets the boundary:

$$\rho(a,\mathbf{x},t) = \begin{cases} w_{(0,\phi_{\mathbf{x}}(-a),t-a)}(a) & \text{for } a < t \\ w_{(a-t,\phi_{\mathbf{x}}(-t),0)}(t) & \text{for } a \geq t \end{cases}$$

By theorem 1 above, G can be identified with differentiation along characteristics (Gurtin and MacCamy [71], Webb [151], Tucker and Zimmerman [144]):

$$\begin{aligned} G(\rho(\cdot,\cdot,t))(a,\mathbf{x}) &= \mathrm{D}_{(1,\mathbf{v},1)}\rho(a,\mathbf{x},t) \\ &\overset{\text{def}}{=} \lim_{h\downarrow 0} h^{-1}(\rho(a+h,\phi_{\mathbf{x}}(h),t+h)J_{\phi}(h) - \rho(a,\mathbf{x},t)). \end{aligned}$$

Hence the along a characteristic curve through $(a_0, \mathbf{x}_0, t_0)$, in which we restrict the interpretation of a_0 and t_0 as indicated above, the cohort satisfies the (ordinary) differential equation

$$\begin{aligned} \frac{d}{ds}w_{(a_0,\mathbf{x}_0,t_0)}(s) &= G(\rho(\cdot,\cdot,t_0+s)))(a_0+s,\phi_{\mathbf{x}_0}(s)) \\ w_{(a_0,\mathbf{x}_0,t_0)}(0) &= \rho(a_0,\mathbf{x}_0,t_0), \end{aligned}$$

which upon integration yields an integral equation, as w appears on the left and (implicitly) on the right:

$$w_{(a_0,\mathbf{x}_0,t_0)}(t) = w_{(a_0,\mathbf{x}_0,t_0)}(0) + \int_0^t G(\rho(\cdot,\cdot,t_0+s)))(a_0+s,\phi_{\mathbf{x}_0}(s))ds.$$

Now, applying the relation between w and ρ and in the case of $t_0 > 0$ using $t_0 = t - a$ and changing variables, one gets to the equation

$$\text{(IntEq)} \qquad \rho(a,\mathbf{x},t) = \begin{cases} F(\rho(\cdot,\cdot,t-a))(\phi_{\mathbf{x}}(-a)) \\ \quad + \int_{t-a}^{t} G(\rho(\cdot,\cdot,s))(s-t+a,\phi_{\mathbf{x}}(s-t))\,ds \\ \qquad\qquad\qquad\qquad \text{a.a. } a < t \\ \rho_0(a-t,\phi_{\mathbf{x}}(-a)) \\ \quad + \int_0^t G(\rho(\cdot,\cdot,s))(s-t+a,\phi_{\mathbf{x}}(s-t))\,ds \\ \qquad\qquad\qquad\qquad \text{a.a. } a \geq t, \end{cases}$$

from which it can be clearly seen that integration of G gives the change in cohort density from the initial condition ρ_0 when $t \leq a$, and the births F when $t > a$.

In what follows, I first prove that solutions to (IntEq) exist uniquely, and then that (IntEq) is equivalent to (P) in that any solution of the one is also a solution of the other. First, by Lemma 1, $\rho \in L_T$ can be identified with an element (also denoted by ρ) in $L_1((0,\infty)\times\Omega\times(0,T);\mathbf{R}^n)$; second, by Lemma 2, the integrals in (IntEq) exist and hence solving it is a meaningful problem; by the remark after Lemma 3, translation in L_T is continuous. These technical lemmas enable the evaluation of integrals by changing variables and the order of integration and taking limits, which are used to prove Theorem 2: there is some non-zero T such that (IntEq) can be solved uniquely for $t \in [0,T]$. Then, using Lemma 4 on the semi-group property of solutions to (IntEq), I prove that (P) and (IntEq) pick out the same ρ in L_T as solutions (Theorem 3), and that these solutions either exist for all time or go to infinity at some finite time (Theorem 4).

This analysis is essentially a generalisation to $\rho(a,\mathbf{x},t)$ of Webb's [151] approach to $\rho(a,t)$, so that I shall often appeal to Webb's proof, with a brief argument of how it is to be generalised. For many of the proofs the argument for $\rho(a,t)$ extends to the case $\rho(a,\mathbf{x},t)$ in a very direct way, the main technical issues being the transformation of variables in Ω (when needed), and discontinuities in ρ_0.

The following hypotheses are assumed throughout; they are sufficient for the unique existence of solutions to the general population problem without time lags or differential operators.

(Hyp 1) For $F : L_1 \times \Omega \to \mathbf{R}^n$, there is an increasing function $c_1 : [0,\infty) \to [0,\infty)$ such that $\|F(\ell) - F(\hat\ell)\|_\Omega \le c_1(r)\|\ell - \hat\ell\|_{L_1}$ for all $\ell, \hat\ell \in L_1$ such that $\|\ell\|_{L_1}, \|\hat\ell\|_{L_1} \le r$, and that $\|F(\ell)(\mathbf{x}) - F(\ell)(\hat{\mathbf{x}})\|_1 \le c_1(r)\|\mathbf{x} - \hat{\mathbf{x}}\|_1$ for all $\mathbf{x}$, $\hat{\mathbf{x}}$ in Ω.

(Hyp 2) For $G : L_1 \times S \to [0,\infty) \times \mathbf{R}^n$, there is an increasing function $c_2 : [0,\infty) \to [0,\infty)$ such that $\|G(\ell) - G(\hat\ell)\|_\Omega \le c_2(r)\|\ell - \hat\ell\|_{L_1}$ for all $\ell, \hat\ell \in L_1$ such that $\|\ell\|_{L_1}, \|\hat\ell\|_{L_1} \le r$, and that $\|G(\ell)(a,\mathbf{x}) - G(\ell)(\hat a,\hat{\mathbf{x}})\|_1 \le c_1(r)(|a - \hat a| + \|\mathbf{x} - \hat{\mathbf{x}}\|_1)$ for all a, $\hat a$ in $[0,\infty)$ and all $\mathbf{x}$, $\hat{\mathbf{x}}$ in Ω.

(Hyp 3) The initial condition $\rho_0 \in S$ is non-negative, bounded by $\bar\rho_0$ and Lipschitz continuous with a common Lipschitz constant c_0 except perhaps at a finite number of $(k-1)$-dimensional manifolds Γ_i, $i = 1, \ldots, q$ in Ω, each of finite $(k-1)$-dimensional volume, where it may have a finite jump: if $\{(a,\mathbf{x}) + t((\hat a,\hat{\mathbf{x}}) - (a,\mathbf{x}))|t \in [0,1], a, \hat a \in [0,\infty), \mathbf{x}, \hat{\mathbf{x}} \in \Omega\} \cap \bigcup_{i=1}^{q} \Gamma_i = \emptyset$, then $\|\rho_0(a,\mathbf{x}) - \rho_0(\hat a,\hat{\mathbf{x}})\|_1 \le c_0(|a - \hat a| + \|\mathbf{x} - \hat{\mathbf{x}}\|_1)$, else $\|\rho_0(a,\mathbf{x}) - \rho_0(\hat a,\hat{\mathbf{x}})\|_1 \le n\bar\rho_0$.

(Hyp 4) For any $\rho \in L_T$, $\mathbf{v} : L_1 \times [0,\infty) \times \Omega \to \Omega$ is differentiable, all partial derivatives exist and are bounded by $b_{\mathbf{v}}$, $\mathbf{v}$ is zero on the boundary of Ω and all the v_i and their partial derivatives w.r.t. $\mathbf{x}$ are Lipschitz continuous with a common increasing function $c_v : [0,\infty) \to [0,\infty)$ as follows:

(a) $\|\mathbf{v}(\ell)(a,\mathbf{x}) - \mathbf{v}(\hat\ell)(\hat a,\hat{\mathbf{x}})\|_1 \le c_{\mathbf{v}}(\|\ell - \hat\ell\|_{L_1} + |a - \hat a| + \|\mathbf{x} - \hat{\mathbf{x}}\|_1)$

(b) $$|\frac{\partial v_i}{\partial x_j}(\rho(\cdot,\cdot,t))(a,\mathbf{x}) - \frac{\partial v_i}{\partial x_j}(\hat\rho(\cdot,\cdot,t))(\hat a,\hat{\mathbf{x}})| \le c_v(r)(\|\ell - \hat\ell\|_{L_1} + |a - \hat a| + \|\mathbf{x} - \hat{\mathbf{x}}\|_1$$

for all $\ell, \hat\ell \in \mathrm{L}_1$ such that $\|\ell\|_{L_1}, \|\hat\ell\|_{L_1} \le r$, all a, $\hat a \in [0,\infty)$ and all $\mathbf{x}$, $\hat{\mathbf{x}} \in \Omega$.

2.3.2 The framework with additional assumptions is (almost) general for plant populations

(1) The transition law can handle any mortality, subpopulation transitions and most seed dispersals. However, (Hyp 2) is restrictive for diffusion models that use derivatives: in general such models will not be Lipschitz continuous, or even well defined in the classical sense, because of the jump discontinuities in ρ. Moreover, quite natural models for population dispersal leads to difficult problems, for example, the diffusion model developed by Gurtin and MacCamy in the 1970's (see Gurtin [70] for a compact account). In order to model populations that disperse, these authors take the direction of diffusion as away from high total population densities, but allow the diffusivity to depend on age; thus different cohorts experience the same population pressure but respond with the varying degrees of movement. This situation is plausible for animals, who can sense densities directly and move in preferred directions, but I would argue that classical diffusion models (or a semi-classical one such as the one mentioned above) are inappropriate for plants: not only do the seed dispersal distances have a Poisson rather than a Gaussian distribution, but also the velocity (and hence distance) of seed movement is always bounded, so it will always be possible to use an explicit distribution in terms of sources, since any spatial domain will also be finite (see below). Hence the difficulties of modelling dispersal by means of differential operators can be avoided for plants. Moreover, preferred directions depend on external factors such as topography, animal movement and wind, so that the kind of density dependence implicit in diffusion models are often inappropriate for plants.

(2) The birth law can handle all birth processes, though in principle only, and where it is deemed necessary to use time lags in order to avoid detailed modelling of recruitment processes poorly understood and/or not observable, it fails.

(3) As stated above, the state evolution law $\mathbf{v}$ formally represents the constitutive relations for the attributes, and is assumed to allow the $\phi_{\mathbf{x}}$ to form a flow. This may be limiting on the kinds of changes possible, for example plants initially different are not allowed to become identical at some finite time, because in that case uniqueness of the characteristic

back to the initial or boundary condition will fail. Also, the assumption is that Ω is bounded and v is zero on the boundary. Boundedness is slightly limiting, in that models that rely on the convenience of an infinite spatial domain to ignore edge effects cannot be used, and the fact that v is zero on the boundary renders it absorbing, whereas for diffusion models one might prefer the looser condition that the exterior normal derivative is zero, as Webb does in [150]. However, this does not limit the generality for plant population models, where dispersal models are adequate and diffusion models are not needed (see above).

(4) The initial condition can handle any population with a finite number of jump discontinuities along hypersurfaces that are finite in extent. Since the domain is finite, it is inconceivable that any plant population would require anything more pathological. Anyway, no population data could ever violate this condition, because observed densities will always be finite.

The generality above is achieved at the expense of some mathematical inconvenience, in the sense that it excludes the mathematical conveniences of piecewise Lipschitz functions and time lags. As noted above, time lags are not covered by the theory developed here. They are attractive for instance as a way of simplifying the many steps from pollination to seedlling establishment. However, in principle these steps could be modelled explicitly, and hence in the sense that all population processes (as indeed all physical processes) are local in time and place, the framework in this thesis is sufficiently general. I accept that this is not adequate, since the point of science is a simplified understanding of reality, which is what time lagged models offer, so that the approach advocated here does fall short in this respect. The case of non-Lipschitz models is more favourable. For example, a model in which rapid changes in vital rates were handled by jumps or vertical tangents would need fewer parameters than a smoothed version of the same, and hence be preferred for the sake of fitting data, simplicity of biological interpretation or ease of computation. However, such a smoothed version can always be constructed, by replacing the condition at every point at which the Lipchitz condition fails by interpolations with finite tangents in a neighbourhood of the point. By making the neighbourhood smaller than biologically meaningful differences, the biological interpretation of the non-Lipschitz model may be retained. Thus for every non-Lipschitz model there is a biologically indistiguishable Lipschitz model. In this approximate sense, the theory above is general enough for to allow for all meaningful models without time lags.

2.3.3 Uniqueness and existence theorem

Lemma 1 *There is a unique* $\rho' \in L_1\big((0,\infty)\times\Omega\times(0,T);\mathbf{R}^n\big)$ *such that for any* $\rho \in L_T$

(a) $\rho'(a,\mathbf{x},t) = \rho(t)(a,\mathbf{x})$ *a.e. for all* $t \in [0,T]$.

(b) $$\int_0^T \|\rho(t)\|_{L_1} dt = \int_0^T \left\{\int_0^\infty \int_\Omega \|\rho'(a,\mathbf{x},t)\|_1 \, d\mathbf{x}\, da\right\} dt = \int_0^\infty \int_\Omega \int_0^T \|\rho'(a,\mathbf{x},t)\|_1 \, dt\, d\mathbf{x}\, da.$$

Note: Since the definition of ρ' is not used below, I do not give it here. The interested reader may consult Webb [151] on how ρ' is to be defined.

Proof: Identical to Webb's Lemma 2.1 (p 25) except that integration is here also over Ω. Since Ω is of finite volume, it cannot affect the validity of the proofs, which appeal only to products of measures and to Fubini's Theorem. □.

Lemma 2 *Define the set* $\Gamma_T = \big\{(c,s)\big|0 < s < T, -s < c < \infty\big\}$*, and assume (Hyp 2). Then for any* $\rho \in L_T$,

(a) *The function* $[0,T] \to L_1$ *that maps* t *to* $G\big(\rho(\cdot,\cdot,t)\big)$ *is in* L_T.

(b) *There exists a function* $h \in L_1\big((0,\infty)\times\Omega\times(0,T);\mathbf{R}^n\big)$ *s.t. for every* $t \in [0,T]$ *and* $a > 0$, $h(a,\mathbf{x},t) \overset{a.e.}{=} G\big(\rho(\cdot,\cdot,t)\big)(a,\mathbf{x})$.

(c) *Given a flow* $\Phi = \{\phi_{\mathbf{x}}\}$ *on* S*, there exists* $k \in L_1(\Gamma_T \times \Omega;\mathbf{R}^n)$ *s.t.* $k(c,s,\mathbf{x}) \overset{a.e.}{=} h(s+c,s,\phi_{\mathbf{x}}(s))$ *for* $(c,s) \in \Gamma_T$*, and*

$$\int_0^T \int_{-s}^{\infty} \int_{\Omega} \|k(c,s,\mathbf{x})\|_1 \, d\mathbf{x}\, dc\, ds$$
$$= \int_{-T}^{\infty} \left\{ \int_{\alpha}^{T} \int_{\Omega} \|k(c,s,\mathbf{x})\|_1 \, d\mathbf{x}\, ds \right\} dc, \text{ where } \alpha =$$
$\max[0,-c]$.

Proof: As for Lemma 1, the details are almost identical to the Lemma 2.2 in Webb. There is no change of variable in Ω. □

The following characterisation of compact subsets of L_1 is needed primarily to establish the continuity of translation in L_1. Webb appeals to Dunford and Schwartz [51, Theorem 20, p 298] for proof.

Lemma 3 *For a closed and bounded subset $M \in L_1$, compactness follows if and only if, for $\ell \in M$;*

(a) $\displaystyle\lim_{h\to 0} \int_0^{\infty} \int_{\Omega} \|\ell(a,\mathbf{x}) - \ell(a+h, \phi_{\mathbf{x}}(h))\|_1 \, d\mathbf{x}\, da = 0$ *uniformly, where one takes* $\ell(a+h,\phi_{\mathbf{x}}(h)) = 0$ *if* $a+h<0$,

(b) $\displaystyle\lim_{h\to 0} \int_h^{\infty} \int_{\Omega} \|\ell(a,\mathbf{x})\|_1 \, d\mathbf{x}\, da = 0$ *uniformly.*

□

To repeat Webb's proof of the continuity of translation, one needs to consider the Jacobian $J_\phi(\mathbf{x})$, which is defined for $\phi_{\mathbf{x}}(s)$ as the determinant of any matrix that maps $\mathbf{x}$ to $\phi_{\mathbf{x}}(s)$. Note the suppression of the flow in age, which is trivial, so that the Jacobian concerns a transformation on Ω only. Since $\phi_{\mathbf{x}}(s) = \mathbf{x}$ for all $\mathbf{x}$ on the boundary of Ω, using the flow at any time as a change of variable cannot affect the boundary, and hence $\int_\Omega \ell(\cdot,\phi_{\mathbf{x}}(\cdot))J_\phi(\mathbf{x})\,dx = \int_\Omega \ell(\cdot,\mathbf{y})\,dy$. Similar to the norms of $F(\ell)$ and $G(\ell)$, I use the norm $\|J_\phi\|_\Omega = \displaystyle\int_\Omega J_\phi(\mathbf{x})\,d\mathbf{x}$.

Define translation through h in the set M as $\mathrm{T}_h : L_1 \to L_1$ as follows:

$$(\mathrm{T}_h\ell)(a,\mathbf{x}) \overset{\text{a.e.}}{=} \begin{cases} \ell(a+h,\phi_{\mathbf{x}}(h)) & a > \max[0,-h] \\ 0 & a \le \max[0,-h]. \end{cases}$$

Now, following Webb, for a sequence $\{\ell_n\}$ in M s.t. $\ell_n \to \ell$,

$$\begin{aligned}||\mathrm{T}_h\ell_n - \mathrm{T}_h\ell||_{L_1} &= \int_{\max[0,-h]}^{\infty} \int_{\Omega} ||\ell_n(a+h, \phi_{\mathbf{x}}(h)) - \ell(a+h, \phi_{\mathbf{x}}(h))||_1 d\mathbf{x}da \\ &\leq \int_0^{\infty} \int_{\Omega} ||\ell_n(a, \mathbf{x}) - \ell(a, \mathbf{x})||_1 d\mathbf{x}da.\end{aligned}$$

Hence, by applying Lemma 3(a) to the compact set $\{\ell\}$, it follows that $\lim_{h\to 0} ||\mathrm{T}_h\ell - \ell||_1 = 0$.

Now follows the existence and uniqueness theorem.

Theorem 2 *Assuming (Hyp 1), (Hyp 2), (Hyp 3) and (Hyp 4), for any* $r > 0$ *there exists* $T > 0$ *such that, for* $||\rho_0||_{L_1} \leq r$, *there is a unique* $\rho \in L_T$ *solving (IntEq) on* $[0, T]$.

Proof: Consider the set $M \subset L_T$, defined by $M = \{\rho \in L_T : \rho(\cdot, \cdot, 0) = \rho_0$ and $||\rho||_{L_T} \leq 2r\}$. It is clear that M is closed. The aim is to show that the operator $K : M \to M$ yields a contraction on M when defined as follows:

$$K\rho(a, \mathbf{x}, t) = \begin{cases} F(\rho(\cdot, \cdot, t-a))(\phi_{\mathbf{x}}(-a)) \\ \quad + \displaystyle\int_{t-a}^{t} G(\rho(\cdot, \cdot, s))(s - t + a, \phi_{\mathbf{x}}(s-t))\, ds \\ \hfill \text{a.e. } a < t \\ \rho_0(a - t, \phi_{\mathbf{x}}(-a)) \\ \quad + \displaystyle\int_0^{t} G(\rho(\cdot, \cdot, s))(s - t + a, \phi_{\mathbf{x}}(s-t))\, ds \\ \hfill \text{a.e. } a \geq t. \end{cases}$$

Clearly, a fixed point of K is a solution to (IntEq). I need to show that K is continuous, that M is closed under K, and that K is contractive. Then, by the Banach fixed point principle, K has a unique fixed point in M. I follow Webb's proof, adding arguments concerning Ω and discontinuities in ρ_0; part (b) below requires only trivial changes (similar to the proof of continuity), but is included here because some of the calculations are used later.

(a) Continuity requires that for $\rho \in M$, the map $t \mapsto K\rho(\cdot,\cdot,t)$ is continuous. The proof in Webb goes through line by line (except of course that all the double integrals become triple integrals), because the argument concerns only one ρ, so there is only one flow on Ω, and therefore change of variables over Ω is never required, and so the position of the symbols $\int_\Omega (\cdot)\, d\mathbf{x}$ can be retained as the inner integral of iterated integrals without at all affecting the calculations.

(b) Closure further requires that $\|K\rho\|_{L_T} \leq 2r$. Consider the L_1 norm at some time t:

$$\begin{aligned}
&\int_0^\infty \int_\Omega \|K\rho(a,\mathbf{x},t)\|_1 \, d\mathbf{x}\, da \\
&\leq \int_0^t \int_\Omega \Big\{ \|F(\rho(\cdot,\cdot,t-a))(\phi_{\mathbf{x}}(-a))\|_1 \\
&\qquad + \int_{t-a}^t \|G(\rho(\cdot,\cdot,s))(s+a-t,\phi_{\mathbf{x}}(s-t))\|_1 ds \Big\} \, d\mathbf{x}\, da \\
&\qquad + \int_t^\infty \int_\Omega \Big\{ \|\rho_0(a-t,\phi_{\mathbf{x}}(-t))\|_1 \\
&\qquad + \int_0^t \|G(\rho(\cdot,\cdot,s))(s+a-t,\phi_{\mathbf{x}}(s-t))\|_1 ds \Big\} \, d\mathbf{x}\, da \\
&\leq \int_0^t \int_\Omega \|F(\rho(\cdot,\cdot,t-a))(\phi_{\mathbf{x}}(-a))\|_1 \, d\mathbf{x}\, da \\
&\qquad + \int_0^\infty \int_\Omega \|\rho_0(a,\phi_{\mathbf{x}}(-t))\|_1 \, d\mathbf{x}\, da \\
&\qquad + \int_0^t \int_\Omega \int_{t-s}^t \|G(\rho(\cdot,\cdot,s))(s+a-t,\phi_{\mathbf{x}}(s-t))\|_1 \, da\, d\mathbf{x}\, ds \\
&\qquad + \int_0^t \int_\Omega \int_t^\infty \|G(\rho(\cdot,\cdot,s))(s+a-t,\phi_{\mathbf{x}}(s-t))\|_1 \, da\, d\mathbf{x}\, ds \\
&\leq (c_1(2r)+c_2(2r)) \int_0^t \|\rho(\cdot,\cdot,s)\|_{L_1} ds \\
&\qquad + \int_0^t \|F(0)\|_\Omega \, ds \\
&\qquad + \int_0^t \|G(0)\|_\Omega \, ds + r \\
&\leq 2rt(c_1(2r)+c_2(2r)+\frac{1}{2r}(\|F(0)\|_\Omega + \|G(0)\|_\Omega)) + r.
\end{aligned}$$

Hence by taking $T \leq r/(2rc_1(2r) + 2rc_2(2r) + \|F(0)\| + \|G(0)\|)$ one ensures that the right hand side is less than or equal to $2r$, since the above holds for all $t \in [0,T]$.

(c) Contraction. The argument is similar to (b). Consider the L_1 norm at some time t:
$\|K\rho(\cdot,\cdot,t) - K\hat{\rho}(\cdot,\cdot,t)\|_{L_1}$

$$
\begin{aligned}
&= \| \int_0^\infty \int_\Omega K\rho(a,\mathbf{x},t) - K\hat{\rho}(a,\mathbf{x},t)\, d\mathbf{x}\, da \|_1 \\
&= \| \int_0^t \int_\Omega \Big\{ F(\rho(\cdot,\cdot,t-a))(\phi_{\mathbf{x}}(-a)) - F(\hat{\rho}(\cdot,\cdot,t-a))(\hat{\phi}_{\mathbf{x}}(-a)) \\
&\quad + \int_{t-a}^t G\big(\rho(\cdot,\cdot,s)\big)(s+a-t, \phi_{\mathbf{x}}(s-t)) \\
&\quad - G\big(\hat{\rho}(\cdot,\cdot,s)\big)(s+a-t, \hat{\phi}_{\mathbf{x}}(s-t))\, ds \Big\}\, d\mathbf{x}\, da \\
&\quad + \int_t^\infty \int_\Omega \Big\{ \rho_0(a-t, \phi_{\mathbf{x}}(-t)) - \rho_0(a-t, \hat{\phi}_{\mathbf{x}}(-t)) \\
&\quad + \int_0^t G\big(\rho(\cdot,\cdot,s)\big)(s+a-t, \phi_{\mathbf{x}}(s-t)) \\
&\quad - G\big(\hat{\rho}(\cdot,\cdot,s)\big)(s+a-t, \hat{\phi}_{\mathbf{x}}(s-t))\, ds \Big\}\, d\mathbf{x}\, da \|_1 \\
&\le \| \int_0^t \int_\Omega F(\rho(\cdot,\cdot,t-a))(\phi_{\mathbf{x}}(-a)) - F(\hat{\rho}(\cdot,\cdot,t-a))(\hat{\phi}_{\mathbf{x}}(-a))\, d\mathbf{x}\, da \|_1 \\
&\quad + \| \int_0^t \int_\Omega \int_{t-s}^\infty G\big(\rho(\cdot,\cdot,s)\big)(s+a-t, \phi_{\mathbf{x}}(s-t)) \\
&\quad - G\big(\hat{\rho}(\cdot,\cdot,s)\big)(s+a-t, \hat{\phi}_{\mathbf{x}}(s-t))\, da\, d\mathbf{x}\, ds \|_1 \\
&\quad + \| \int_t^\infty \int_\Omega \rho_0(a-t, \phi_{\mathbf{x}}(-t)) - \rho_0(a-t, \hat{\phi}_{\mathbf{x}}(-t))\, d\mathbf{x}\, da \|_1 \\
&= \mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3
\end{aligned}
$$

where

$$
\begin{aligned}
\mathrm{I}_1 &\overset{\text{def}}{=} \| \int_0^t \int_\Omega F(\rho(\cdot,\cdot,t-a))(\phi_{\mathbf{x}}(-a)) - F(\hat{\rho}(\cdot,\cdot,t-a))(\hat{\phi}_{\mathbf{x}}(-a))\, d\mathbf{x}\, da \|_1 \\
\mathrm{I}_2 &\overset{\text{def}}{=} \| \int_0^t \int_\Omega \int_0^\infty G\big(\rho(\cdot,\cdot,s)\big)(a, \phi_{\mathbf{x}}(s-t)) - G\big(\hat{\rho}(\cdot,\cdot,s)\big)(a, \hat{\phi}_{\mathbf{x}}(s-t))\, da\, d\mathbf{x} \\
\mathrm{I}_3 &\overset{\text{def}}{=} \| \int_0^\infty \int_\Omega \rho_0(a, \phi_{\mathbf{x}}(-t)) - \rho_0(a, \hat{\phi}_{\mathbf{x}}(-t))\, d\mathbf{x}\, da \|_1 .
\end{aligned}
$$

To complete the proof, I need to show that the each of the integrals I_1, I_2 and I_3 has a bound of the form $tA_i(t)\|\rho - \hat{\rho}\|_{L_t}$ for $i = 1, 2, 3$, where the $A_i(t)$ are positive non-decreasing functions of t. Now,

$$\begin{aligned}
\mathrm{I}_1 &\leq \int_0^t \int_\Omega \|F(\rho(\cdot,\cdot,t-a))(\mathbf{x})J_\phi(\mathbf{x}) - F(\hat{\rho}(\cdot,\cdot,t-a))(\mathbf{x})J_{\hat{\phi}}(\mathbf{x})\|_1 \, d\mathbf{x}\, da \\
&\leq \int_0^t \int_\Omega \|F(\rho(\cdot,\cdot,t-a))(\mathbf{x})J_\phi(\mathbf{x}) - F(\rho(\cdot,\cdot,t-a))(\mathbf{x})J_{\hat{\phi}}(\mathbf{x})\|_1 \, d\mathbf{x}\, da \\
&\quad + \int_0^t \int_\Omega \|F(\rho(\cdot,\cdot,t-a))(\mathbf{x})J_{\hat{\phi}}(\mathbf{x}) - F(\hat{\rho}(\cdot,\cdot,t-a))(\mathbf{x})J_{\hat{\phi}}(\mathbf{x})\|_1 \, d\mathbf{x}\, da \\
&\leq \int_0^t \|F(\rho)\|_\Omega \|J_\phi - J_{\hat{\phi}}\|_\Omega \, d\mathbf{x}\, da \\
&\quad + \int_0^t \|J_{\hat{\phi}}\|_\Omega \|F(\rho) - F(\hat{\rho})\|_\Omega \, da \\
&\leq 2rt\|J_\phi - J_{\hat{\phi}}\|_\Omega + t\|J_{\hat{\phi}}\|_\Omega c_1(2r)\|\rho - \hat{\rho}\|_{L_t},
\end{aligned}$$

where $\|J_\phi - J_{\hat{\phi}}\|_\Omega \overset{\text{def}}{=} \sup_{\mathbf{x}\in\Omega} \|J_\phi - J_{\hat{\phi}}\|_1$. By a similar argument,

$$\mathrm{I}_2 \quad = \quad 2rt\|J_\phi - J_{\hat{\phi}}\|_\Omega + t\|J_{\hat{\phi}}\|_\Omega c_1(2r)\|\rho - \hat{\rho}\|_{L_t}$$

The case of I_3 is a bit more complicated. By (Hyp 3), ρ_0 is Lipschitz except at most on q surfaces of dimension $(k-1)$ and of finite volume. Taking γ as a supremum for these volumes, one may conclude from the positivity and boundedness of ρ_0 that the difference $\|\rho_0(a, \phi_{\mathbf{x}}(-t)) - \rho_0(a, \hat{\phi}_{\mathbf{x}}(-t))\|_1$ is less than $n\bar{\rho}_0$, so that the contribution of the discontinuities to the integral is at most $nq\bar{\rho}_0\gamma B$, where B is a bound on the maximum norm of the difference between the points on the boundary picked out by the two flows. Now B clearly less than $\|\phi - \hat{\phi}\|_\Omega \overset{\text{def}}{=} \sup_{\mathbf{x}\in\Omega} \|\phi_{\mathbf{x}} - \hat{\phi}_{\mathbf{x}}\|_1$, on which a bound is obtained in the required form below. Now, consider the set $R \subset \Omega$ contained between the jumps in ρ_0 as picked out by $\phi_{\mathbf{x}}$ and $\hat{\phi}_{\mathbf{x}}$ respectively for all $\mathbf{x}$: the line joining $\phi_{\mathbf{x}}(s)$ to $\hat{\phi}_{\mathbf{x}}(s)$ is given by the points $\phi_{\mathbf{x}}(s) + \theta(\hat{\phi}_{\mathbf{x}}(s) - \phi_{\mathbf{x}}(s))$ for $\theta \in [0, 1]$. Along this line, let θ_1 (resp. θ_2) be the smallest (resp. largest) value of θ at which ρ_0 is discontinuous. Then $R \overset{\text{def}}{=} \left\{\mathbf{x}' \in \Omega | \mathbf{x}' = \phi_{\mathbf{x}}(s) + \theta'(\hat{\phi}_{\mathbf{x}}(s) - \phi_{\mathbf{x}}(s)), \quad \theta \in [\theta_1, \theta_2], \quad \forall \mathbf{x} \in \Omega\right\}$. Now on $\Omega' \overset{\text{def}}{=} \Omega \setminus R$, change of variable through $\phi_{\mathbf{x}}$ is well defined, ρ_0 is

Lipschitz continuous, so (remembering that $\|\rho_0\|_{L_1} \leq r$)

$$\begin{aligned} \mathrm{I}_3 &\leq \int_0^t \|\rho_0\|_{L_1}\|J_\phi - J_{\hat{\phi}}\|_\Omega + nq\bar{\rho}_0\gamma\|\phi - \hat{\phi}\|_\Omega \, da \\ &= t\big(r\|J_\phi - J_{\hat{\phi}}\|_\Omega + nq\bar{\rho}_0\gamma\|\phi - \hat{\phi}\|_\Omega\big). \end{aligned}$$

Thus bounds on $\|J_\phi\|_\Omega$, $\|J_\phi - J_{\hat{\phi}}\|_\Omega$ and $\|\phi - \hat{\phi}\|_\Omega$ are needed. Now, J_ϕ is the determinant of a $k \times k$ matrix, so J_ϕ is a sum of $k!$ terms of the form $(-1)^p \prod_{i=1}^{k} \frac{\partial \phi_{\mathbf{x}}^{\alpha_i}}{\partial x_{\beta_i}}$, where $\alpha_i = 1, \ldots, k$ and β_i is a permutation of α_i. Since all these partial derivatives are bounded by $b_{\mathbf{v}}$, the first of these bounds is just $\|J_\phi\|_\Omega \leq k! b_{\mathbf{v}}^k$. Furthermore, $\|J_\phi - J_{\hat{\phi}}\|_1$ is a sum of $k!$ terms of the form $\left|\prod_{i=1}^{k} \frac{\partial \phi_{\mathbf{x}}^{\alpha_i}}{\partial x_{\beta_i}} - \prod_{i=1}^{k} \frac{\partial \hat{\phi}_{\mathbf{x}}^{\alpha_i}}{\partial x_{\beta_i}}\right|$. Each of these terms is bounded by a sum of k terms like $|\frac{\partial \phi_{\mathbf{x}}^{\alpha_j}}{\partial x_{\beta_j}} - \frac{\partial \hat{\phi}_{\mathbf{x}}^{\alpha_j}}{\partial x_{\beta_j}}| \left|\prod_{i \neq j}^{k} \frac{\partial \phi_{\mathbf{x}}^{\alpha_i}}{\partial x_{\beta_i}}\right|$. Hence the last two bounds reduce to bounds on the absolute value of the difference between $\phi_{\mathbf{x}}(s)$ and $\hat{\phi}_{\mathbf{x}}(s)$ and on the difference between partial derivatives of $\phi_{\mathbf{x}}(s)$ and $\hat{\phi}_{\mathbf{x}}(s)$ for all $s \in [0, t]$. The approach taken Tucker and Zimmerman [144, pp 559– 560] can be adapted to give

$$\begin{aligned} \|\phi_{\mathbf{x}} - \hat{\phi}_{\mathbf{x}}\|_1 &= \sum_{i=1}^{k} |\phi_{\mathbf{x}}^i(s) - \hat{\phi}_{\mathbf{x}}^i(s)| \\ &\leq \int_0^s \sum_{i=1}^{k} |v_i(a, \phi_{\mathbf{x}}(\tau), \rho(\cdot,\cdot,t+\tau)) - v_i(a, \hat{\phi}_{\mathbf{x}}(\tau), \hat{\rho}(\cdot,\cdot,t+\tau)) \\ &\leq skc_v(2r)\|\rho - \hat{\rho}\|_{L_s} + kc_v(2r)\int_0^s \sum_{i=1}^{k} |\phi_{\mathbf{x}}^i(\tau) - \hat{\phi}_{\mathbf{x}}^i(\tau)| d\tau \\ &\leq skc_v(2r)e^{skc_v(2r)}\|\rho - \hat{\rho}\|_{L_s}, \end{aligned}$$

where the (integral equation) definition $\phi_{\mathbf{x}}^i(s) = x_i + \int_0^s v_i(a, \phi_{\mathbf{x}}(\tau), \rho(\cdot,\cdot,t+\tau))d\tau$ and (Hyp 4) was used in the first two steps, and Gronwall's inequality in the last step. The bound on the difference between the partial derivatives is then a straightforward imitation and application of this result and (Hyp 4), giving

$$\sum_{i=1}^{k} |\frac{\partial}{\partial x_j}\phi_{\mathbf{x}}^i(s) - \frac{\partial}{\partial x_j}\hat{\phi}_{\mathbf{x}}^i(s)| \leq skc_v(2r)\big(skc_v(2r)e^{kc_v(2r)s} + 1\big)\|\rho - \hat{\rho}\|_{L_s}.$$

These results then combine to give $\|K\rho(\cdot,\cdot,t) - K\hat{\rho}(\cdot,\cdot,t)\|_{L_1} \leq t(A_1(t) + A_2(t) + A_3(t))\|\rho - \hat{\rho}\|_{L_t}$, where

$$\begin{aligned}
A_1(t) &\overset{\text{def}}{=} t\big[2rk!b_{\mathbf{v}}^{k-1}tkc_v(2r)(tkc_v(2r)e^{tkc_v(2r)} + 1) + tk!b_{\mathbf{v}}^{k}c_1(2r)\big]\\
A_2(t) &\overset{\text{def}}{=} t\big[2rk!b_{\mathbf{v}}^{k-1}tkc_v(2r)(tkc_v(2r)e^{tkc_v(2r)} + 1) + tk!b_{\mathbf{v}}^{k}c_2(2r)\big]\\
A_3(t) &\overset{\text{def}}{=} t\big[nq\bar{\rho}_0\gamma tkc_v(2r)e^{tkc_v(2r)} + tkc_v(2r)(tkc_v(2r)e^{tkc_v(2r)} + 1)k!b_{\mathbf{v}}^{k-1}P_0\big]
\end{aligned}$$

are positive increasing functions of t.

Hence if one considers the supremum over $[0,T]$

$$\|K\rho - K\hat{\rho}\|_{L_T} \leq T(A_1(T) + A_2(T) + A_3(T))\|\rho - \hat{\rho}\|_{L_T}.$$

Since the A_i are increasing functions, the reciprocal of their sum gets larger as $T \to 0$ (not necessarily without bound, but necessarily larger than zero), so their certainly are T such that $T < (A_1(T) + A_2(T) + A_3(T))^{-1}$. By selecting such a T one ensures that $\|K\rho - K\hat{\rho}\|_{L_T} < \|\rho - \hat{\rho}\|_{L_T}$. ☐

I now outline Webb's technique for showing that solutions either remain bounded for all time, or else go to infinity at some finite time.

Lemma 4 *Select T such that ρ is a solution to* (P) *on* $[0,T]$. *Then there exists an extension of ρ on $[T, T+\hat{T}]$ such that ρ is a solution on $[0, T+\hat{T}]$, by $\rho(\cdot,\cdot,t) \overset{\text{def}}{=} \hat{\rho}(\cdot,\cdot,t-T)$ for $\hat{T} > 0$, where*

$$\hat{\rho}(a,\mathbf{x},t) = \begin{cases} F\big(\hat{\rho}(\cdot,\cdot,t-a)\big)(\hat{\phi}_{\mathbf{x}}(-a) \\ \quad + \int_0^a G\big(\hat{\rho}(\cdot,\cdot,s)\big)(s-t+a, \hat{\phi}_{\mathbf{x}}(s-t))\,ds \\ \qquad\qquad a.a.\ a \in (0,t) \\ \rho(a-t, \hat{\phi}_{\mathbf{x}}(-a), T) \\ \quad + \int_{a-t}^a G\big(\hat{\rho}(\cdot,\cdot,s)\big(s-t+a, \hat{\phi}_{\mathbf{x}}(s-t)\big)\,ds \\ \qquad\qquad a.a.\ a \in (t,\infty) \end{cases}$$

Proof: Note that the existence of some such $\hat{\rho}$ is guaranteed by the theorem above. This lemma is basically the same as Webb's Proposition 2.4 (p 42), and the proof goes through without any need for added comment, since no change of variable in Ω is required. ☐

Theorem 3 *ρ solves (P) if and only if it satisfies (IntEq) on $[0,T]$.*

Proof: That a solution to (IntEq) implies one to (P) is Webb's proposition 2.1 (p 30), while the converse implication is his theorem 2.3 (p 43). The proofs are again so similar as not to need repetition. □

Theorem 4 *Suppose ρ solves (P) on $[0,T_\phi)$. Then either $T_\phi = \infty$ or $\limsup_{t \to T_\phi} \|\rho(t)\| = \infty$.*

Proof: This is Webb's theorem 2.4; again the proof need not be repeated. □

2.4 Speculations on the regularity of solutions

Webb [151, chapter 3] gives several theorems are given concerning the regularity of solutions. I do not expand them for the general case, but I do note that they make use of semi-groups. Since Lemma 4 establishes the basic semi-group property of the finite solutions to (P), I expect that these theorems can be generalised in a similar way.

In his book [151, chapter 4], Webb follows tradition by emphasising the analysis of equilibria, and gives asymptotic results. Nowadays non-equilibrium phenomena are receiving increased attention (Dean [47], Chesson and Case [28]), and as the case studies in chapters 3 and 4 below concern transient behaviour, asymptotic results are not required.

2.5 An algorithm for solving P

In numerical work on (P), the aim has usually been to specialise the problem and then to focus on the features of the specialised problem. Thus Lopez and Trigiante [112] consider diffusion in $\mathbf{R}^2$ with finite maximum age w; the problem is in their terminology "stiff" because, as

$a \to w$, one has $\lambda \to \infty$. They propose using finite differences in both space and age, and set the time step Δt equal to the age step, and use a generalised scheme such that it is explicit for small a (thus allowing fast computation), and implicit for a near w (thus allowing a relatively large Δt, though at the price of having to solve a set of linear equations at each time step). They achieve, under a mild assumption on the birth rate relative to Δt, the following rate of convergence:

$$||\xi^{n+1}||_\infty < ||\xi^0||_\infty + \sigma n r \Delta t,$$

where σ is constant, r is a composite measure of mesh size, namely $r = h+k_1+k_2$ (k_1, k_2 discretise the spatial mesh), and ξ^n is the difference at t_n between the true and approximate solutions.

Milner and co-workers produced several papers on this topic. In Douglas and Milner [50] the population is structured by age only and the nonlinearity is restricted to letting mortality depend on the total population. Using uniform age-time mesh, the fact that characteristics are known in advance, and the rectangle rule for quadrature (i.e. the simplest possible), together with births given *a priori*, they obtain the convergence estimate

$$||\xi|| \leq C\Delta t.$$

Kostova [96] reports a similar convergence results, plus interesting examples of failure to converge. In [115], Milner uses finite elements in age, assumes initially finite age, and integrates over age to get a variational version of the problem. Application of a Crank-Nicholson step to the time integration then yields globally second order accuracy in time (on the assumption that the finite element mesh is refinable to give any desired accuracy). This requires, for the two-sex model he gives as an example, the solution of three finite element matrix equations at each time step. In Arbogast and Milner [6] a finite difference algorithm for this same problem was preferred. There results for an application are shown: taking US census data for 1970 and 1980, they find, by interpolation, the rates for various population processes. Then taking 1970 census data as initial values and using the estimated rates, they arrive at a good fit with the 1980 census data (except that the number of females over 75 is radically underestimated, and that the predicted number of couples is fairly good only in the age range where this number is high). This is purely descriptive theory, in that it uses 1980 data to predict 1980 data. Predictive power would only accrue through parameters that can plausibly be assumed to hold beyond the data, which would require a kind of constitutive relation absent in their work.

More recently, Milner [114] tackled the diffusion of an age-structured population in $\mathbf{R}^2$. This supersedes Lopez and Trigiante's paper [112], in that he considers nonlinear diffusion that depends on both ρ and $\mathbf{x}$—to be specific, the direction is of "least crowding", so that this diffusion is entirely analogous to that of classical physics. Numerically, there is now the problem that the characteristics are not known in advance. His basic strategy is to separate the approximate calculation of characteristics from that of the density, which he achieves by reformulating the problem through several changes of variable. As in Douglas and Milner [50], the quadrature is by the rectangle rule, and the time integration is of first order accuracy, so that the overall error in the age profile—i.e on the new variables—is first order. For the total population p and the characteristics ϕ, the error is somewhat poorer: $O\left((\Delta t)^{1/2} + h^{1/2} + h/(\Delta t)^{1/2}\right)$, where $h = O(\Delta t)$ is a measure of the space discretisation.

Milner's success with a simple finite element method holds the promise that the sophisticated methods developed for evolution equations, principally reaction-diffusion and advection-diffusion equations (for a good review, see Cockburn and Shu [36]), should be very efficient if any problem of sufficient economic importance to justify the investment in coding were to appear. Other interesting approaches that may profitably be adapted include the long time steps aimed at in Dey and Dey [48] and the boundary element approach of Ingber and Phan-Thien [85].

My approach to use the method of lines (cf. Douglas and Milner [50] and Kostova [96]), in which the population densities of individual cohorts are tracked along characteristics. I use fourth-order Runge-Kutta integration in time, and Simpson's two-thirds rule in space, so that the expected accuracy is $O\left((\Delta t)^4 + \|\mathbf{h}\|^3\right)$. However, on the basis of very limited evidence it appears that the accuracy achieved is somewhat better (see Figure 2.1). The computational results in chapters 3 and 4 were obtained using this algorithm, so that the results given here constitute an example of the rate of convergence, a good test of the code, and a practical guide to suitable discretisation. Much more sophisticated are the approaches to Runge-Kutta are algorithms explored in Vermiglio [147] and Cooper [39].

Consider $\rho(a_0, \mathbf{x}_0, t_0)$, where either $t_0 = 0$ or $a_0 = 0$, and the characteristics are $\phi_{\mathbf{x}_0}(s)$. Then $\rho(a_0 + s, \phi_{\mathbf{x}_0}(s), t_0 + s)$ gives the cohort density. The computation (and description) can be much simplified by using comoving coordinates in Ω:

$$\xi \overset{\mathrm{def}}{=} \phi_{\mathbf{x}_0}(s).$$

Then, with respect to ξ, the velocities are zero and, since I do not consider dispersal or transitions, the balance law reduces to a simple decay eqaution:

$$\frac{d\rho}{dt} - \lambda\rho = 0.$$

Put differently, there is only one characteristic through any $(a, \mathbf{x})$, so one can ignore the actual values of $(a, \mathbf{x})$ and merely keep track of the time. Whenever the current values of coordinates are required, they are referred back to values on the boundary and one uses $(a_0 + t, \phi_{\mathbf{x}_0}(t), t)$ to track the evolution of $\rho_0(a_0, \mathbf{x}_0)$ and $(t - t_0, \phi_{\mathbf{x}_0}(t - t_0), t)$ to track $F(\rho(\cdot, \cdot, t_0)(\phi_{\mathbf{x}}(t - t_0))$. This is computationally convenient since one need store only an array for ρ_0, which is filled during the initialisation stage of the calculation, and another array for $F\big(\rho(\cdot, \cdot, t)\big)(\mathbf{x})$, which is gradually filled as t advances.

On this basis, I propose the following algorithm:

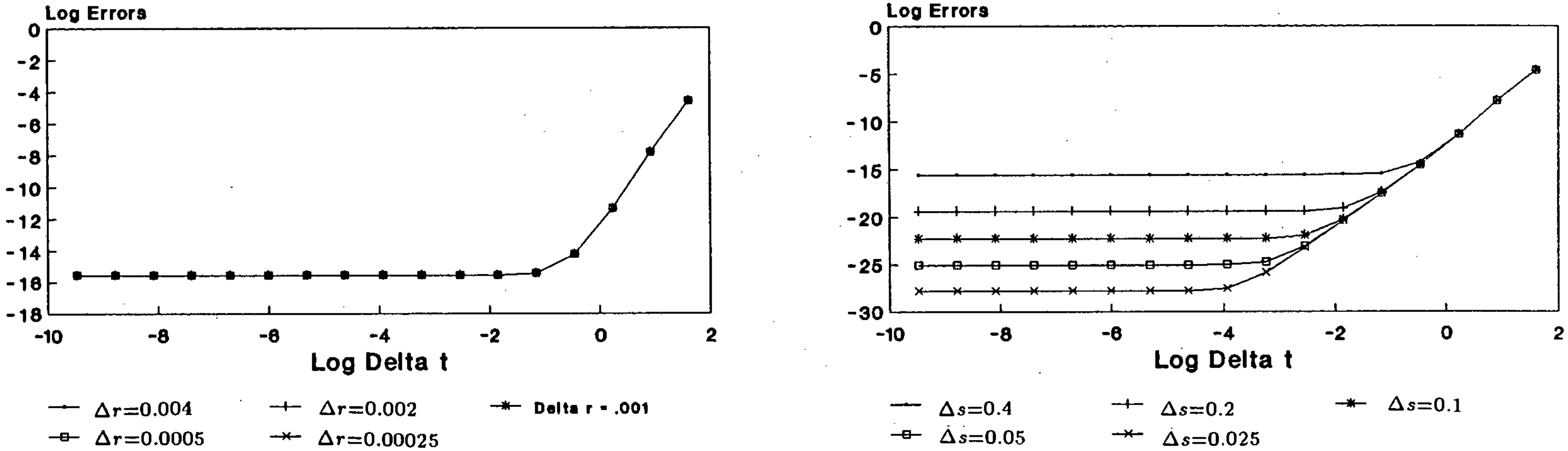
Algorithm A

Start at $\rho_0(a_j, \mathbf{x}_\alpha)$, where α is a multi-index to the discretisation of Ω and j is an index to the age discretisation; $j = 0$ corresponds to births.

To find $\rho_{i+1}\big(a_j, \phi_{\mathbf{x}_\alpha}(\tau)\big)$ for ≥ 0 and $j > 0$, solve the characteristic and cohort ODE's with the initial condition $\rho_i\big(a_{j-1}, \phi_{\mathbf{x}_\alpha}(\tau - h)\big)$ approximately by an n-th order Runge-Kutta step. To find $\rho(0, \mathbf{x})$, apply F to $\rho_{i+1}\big(a_j, \phi_{\mathbf{x}_\alpha}(\tau)\big)$.

The error in one step of the above algorithm is $O(h^{n+1})$ when the char-

Convergence behaviour of algorithm



All Δs=0.4

All Δr=0.004

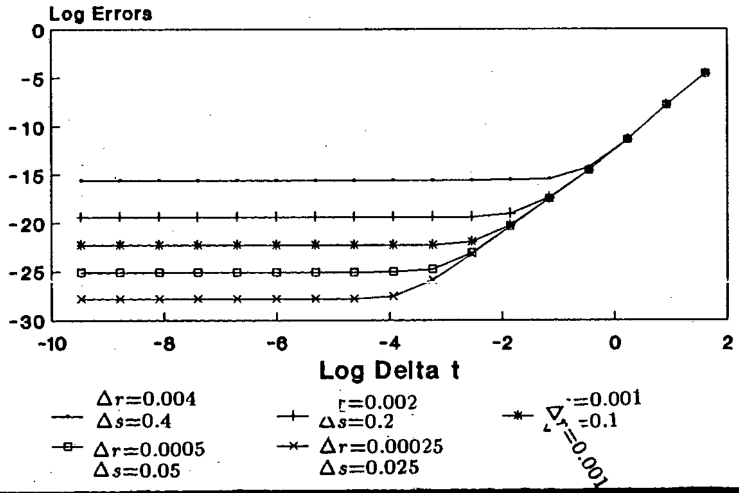


Figure 2.1: The model has mortality as the reciprocal of size, while size growth is linear. The coarsest mesh sizes are $\Delta r = .004$, $\Delta s = .4$ and $\Delta t = 2.5$. Errors represent the maxnorm at $t = 5$. It is clear that at such small Δr, further refinements are pointless. The slope of the straight section of all the curves is approximately 4.4, so the rate of convergence is approximately $O\big((\Delta t)^{4.4}\big)$ until $\Delta t \approx \Delta s$.

acteristics are known. However, even if the characteristics need to be approximated, one can ensure global $O(h^n)$ error by making all local approximations $O(h^n)$, as is shown for the case $n = 4$ by the following theorem.

Theorem 5 *For the initial value problem* $\dot{x} = f(g(x), t)$, $x(0) = x_0$, *if both* f *and* g *are analytical and* g *is approximated by* $\tilde{g}$ *such that* $\tilde{g}(x; h) = g(x) + O(h^m)$, *then*

$$x(t+h) = \mathrm{RK}(x(t), h; f, g) + O(h^{n+1}),$$

where $n = \min(4, m)$ *and RK is a fourth-order Runge-Kutta algorithm.*

<u>Proof</u>: I give the proof for a particular fourth-order Runge-Kutta algorithm, namely the standard

$$x(t+1) = x(t) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

where

$$\begin{aligned} k_1 &= f(g(x), t) \\ k_2 &= f(g(x + hk_1/2), t + h/2) \\ k_3 &= f(g(x + hk_2/2), t + h/2) \\ k_4 &= f(g(x + hk_3), t + h). \end{aligned}$$

Other explicit fourth-order Runge-Kutta algorithms differ merely in the weights given to the different factors, and would have entirely similar proofs.

Expanding f by Taylor series one has

$$f\big(\tilde{g}(x; h), t\big) = f\big(g(x), t\big) + O(h^m).$$

Using $\tilde{k}_i$ to denote the approximation to k_i, the error in each $\tilde{k}_i$ is $O(h^m)$, as follows:

$$\tilde{k}_1 = f\big(\tilde{g}(x; h), t\big) = f\big(g(x), t\big) + O(h^m) = k_1 + O(h^m).$$

The calculation for k_2 takes a bit longer:

$$\begin{aligned}
\tilde{k}_2 &= f(\tilde{g}(x + \frac{h}{2}(k_1 + O(h^m)); h), t + h/2) \\
&= f(\tilde{g}(x + hk_1/2; h) + O(h^{m+1}), t + h/2) \\
&= f(\tilde{g}(x + hk_1/2; h), t + h/2) + O(h^{m+1}) \\
&= f(g(x + hk_1/2; h), t + h/2) + O(h^m) \\
&= k_2 + O(h^m).
\end{aligned}$$

This calculation is repeated for k_3 and k_4.

In summary, one has $\tilde{k}_i = k_i + O(h^m)$. It is now immediate that

$$\mathrm{RK}(x(t), h; f, \tilde{g}) = \mathrm{RK}(x(t), h; f, g) + O(h^{m+1})$$

and the statement of the theorem follows. □

Thus, in order to achieve $O(h^4)$ global rate of convergence, one needs to ensure that the argument of f can be evaluated to fourth order accuracy.

2.5.1 Implementation of Algorithm A

In all the models in this thesis, $F \equiv 0$, at least for $t < t_{end}$, so that all individuals have the age $a_0 + t$. I assume for convenience that they are all in fact the same age and are born at $t = 0$, i.e. that $a_0 = 0$. Furthermore, the i-state reduces to $\mathbf{x} = (r, s)$, where s represents individual size, and r represents intrinsic somatic growth rate, which is constant throughout an individual's life. Thus $\mathbf{v}$ reduces to an expression for $\dot{s}$, and as already mentioned, G reduces to an expression giving λ.

The structure space Ω is discretised by an initially rectangular and regular mesh (r_i, s_j), where $r_{i+1} - r_i = \Delta r$ and $s_{j+1} - s_j = \Delta s$ are constants. The initial values of r_i and s_j are stored in the vectors $r0(i)$ and $s0(j)$ respectively. Because of the comoving coordinates, one needs the vector $s\_curr(j)$ storing the current (approximate) value of s_j. The mesh deformations can very large with respect to initial size, with step size becoming quite large, and of course not constant (see Figure 3.18 in chapter 3 for an example). One may estimate the Jacobian $J_\phi(\mathbf{x})$ by comparing the arrays $(r0(i), s0(j))$ and $(r0(i), s\_curr(j))$.

The density is stored as an array $\rho\_curr(i,j)$, and at each time step ρ_curr is overwritten by values obtained from the standard fourth order Runge-Kutta step. In order to take a time step of length $\triangle t$, one must simultaneously update ρ_curr and s_curr. This means that $\dot{s}$ and λ, which may depend on ρ, must be evaluated not only at t, but also at $t + \triangle t/2$ and $t + \triangle t$. Thus intermediate arrays giving the increments k_1, k_2 and k_3 for the Runge-Kutta steps of all the elements in ρ_curr and s_curr are needed. Note that sometimes the dependence on ρ requires integration over Ω (e.g. to obtain P or $\bar{s}$). For this I used Simpson's 1/3 rule, which gives accuracy compatible with the fourth order Runge-Kutta method, as long as ρ_0 is continuous and $F \equiv 0$, as was the case in all the models used in this thesis. Else, the integration is no better than first order, though tracking discontinuities along characteristics does help, as a brief and somewhat superficial investigation showed.

To validate the code written for this implementation, I used $\lambda = 1/s$ and $\dot{s} = r$ and constant $\triangle t$. Then solutions for characteristics and for ρ can be found analytically, so that global error can be calculated. Fig 2.1 shows the result: convergence was at the rate $O(h^{4.4})$, which is compatible with local error of $O(h^5)$ from one step of the Runge-Kutta method, where $h = \max(\triangle r, \triangle s, \triangle t)$. The program was further checked with a logistic model, namely $\lambda = rP(1 - P/K)$, to confirm that it gave correct results for a model requiring ρ to be used when evaluating λ. Some of the models in chapter 3 have analytical solutions, which further verified that the code gives correct results.

The above results were obtained on a VAX mainframe by means of rather lengthy code, which consists of a set of DCL batch files for controlling the compilation and execution of a set of FORTRAN 77 programmes. In this way a number of different models for mortality and somatic growth could be coded into independent modules, and combined with very little effort.

Chapter 3

Plant mortality and somatic growth: on model discrimination for even-aged monospecific stands

Plant mortality and somatic growth are central to any attempt to understand the dynamics of plant populations. There is an extensive literature of empirical studies, going back as far as 1928 (the work of Suskatchew, as cited in White and Harper [163]); the many recent studies include Cannel *et al.* [21] (1984), West and Borough [159], Peet and Christensen [122], Mithen *et al* [116], Penridge and Walker [123], Hara [74, 75, 76], Hara *et al.* [78], Shaw and Antonovics [134], Solbrig *et al.* [140]. Population simulations are of course based on these processes; recent studies includes Bonan [13], Bishir and Namkoong [12], Aikman and Watkinson [3], Ford and Diggle [63], Clark [35]. However, one often finds that either mortality or somatic growth is relatively neglected in a particular experiment, theoretical discussion or practical application. My purpose in this chapter is to demonstrate how the framework from chapter 2 enables one to pay attention to both processes at the same time. I illustrate the power of this approach by simulations, using a particular set of models for even-aged monospecific stands. Such populations are attractive experimental systems because of their relative simplicity, in that the processes of birth and migration (as occur in for example Czaran and Bartha [46]) are excluded. The results below suggest that by using the size distribution, three types of size dependence in mortality and somatic growth can be distinguished. Such a procedure is in a sense converse to that of Westoby [160], who uses properties of $\dot{s}$ (the distribution modifying function or DMF in his terminology) to infer properties of the size distribution.

One obtains a model from the framework expounded in chapter 2 by specifying the constitutive relations for births, migrations and other individual changes, and deaths. The use of such a model for description or prediction then involves estimating the relevant parameters (as was done by Kohyama [94, 95] and Hara [77]). The steps of developing a model (or a family of models) and estimating the best fit parameters for given data and model (family) are logically distinct, so that for any particular application it makes sense to ask whether some models can be excluded *a priori*. This may for instance help to reduce task of parameter estimation to manageable size.

Even-aged monospecific stands have another advantage: they were intensively studied both experimentally and theoretically during the the attempt to establish a self-thinning law for plant population dynamics (ably reviewed Weller [156]). I review the forms in which data are available from some of this literature in section 3.1, and go on to show in section 3.2 that by itself, a log biomass *vs.* log population plot, often called a B–N plot (Westoby [161]), does not allow for model discrimination. This motivates the use of structured models, and in section 3.3 I report on computational experiments which suggest that the combination of a B–N plot with a final size distribution may well furnish enough data for distinguishing between biologically important categories of models of somatic growth and mortality. Section 3.4 is devoted to discussion of the results of the two previous sections, and section 3.5 to how the work initiated here may be continued.

3.1 Brief account of data in the literature on self-thinning

The search for a self-thinning law has left a problematic legacy (Weller [156] Zeide [166], Lonsdale [111]). This is not the place to consider the extent to which the attempt to establish this law was ill-conceived, or to consider the validity of the statistical and other arguments used. However, much valuable data are contained in this literature. I agree with Peters [124] that "the data can be taken to support a very general, but imprecise, self-thinning rule"; and would also point out that in Zeide [166], along with a powerful critique of the claim that a general law has been discovered, one may find a plea for retaining awareness that the data are valuable and potentially very useful. I shall not refer to discussions of mortality and somatic growth from this literature, such as occur in Clark [34], Kenkel [89], Ellison [53], Lloyd and Harms [107] or Long and Smith [109]—below I point out that the data usually gathered in these studies could never distinguish between the effects of population density and individual size on the mortality and the somatic growth. My concern here is simply to discuss how information on mortality and somatic growth may be contained in the self-thinning data. Therefore, it is necessary to discuss the forms in which data have been presented, and their qualitative features.

3.1.1 B–N plots

These plots were so named by Westoby [161], and are simply plots of the biomass (density) versus population (density), using a logarithmic scale for both axes, so that a simple power law, such as was first proposed for self-thinning by Yoda *et al.* [165], would appear as a straight line on the B–N plot. Since the biomass is simply the product of the average size and the population, plotting size versus population would be analytically equivalent. However, as pointed out by Weller [156], this leads to spurious correlation between the two variables whenever the actual measurement is of the biomass, from which the average size is then calculated, because the population data then appears in both population and (calculated) average size.

It is important to distinguish between the B–N plots that consist of several data points taken from a single site at a sequence of times, i.e. trajectories, and those that consist of single points from many sites. I shall not treat the latter, which are an intersting topic in their own right (Reineke [129], Gorham [66], Weller [157]). Only trajectories can be intepreted in terms of causal processes, such as mortality and somatic growth. Trajectories reported in the literature are usually concave[1] (Westoby [161], Aikman and Watkinson [3], Hara [73], Lonsdale and Watkinson [110], Peet and Christensen [122], Pitelka [125], White and Harper [163], Zeide [166]), though in several of these cases the concave curve is fitted (or assumed to be fitted) to data that are not concave (in the sense that linear collocation of the data yields a non-concave curve). One notable exception to concavity was reported by Carleton and Wannamaker [22], who found for one species at several sites that the B–N plot was sigmoid i.e. that it started concave and then became convex. In the experiments reported in section 3, the aim was to achieve concave trajectories; other shapes were observed in the course of experimentation and hence it appears that trajectories such as reported by Carleton and Wannamker could also be achieved by appropriate parameter combinations.

3.1.2 Size distributions

These are histograms, frequency distributions or probability density functions for individual size. Many measures of individual size have been considered, among them dry mass (Hara [75], Westoby and Howell [162], Weiner [152], Mithen *et al.* [116]), canopy volume (Aikman and Watkinson [3], Lonsdale and Watkinson [110]), height (Hara [75], Ford and Diggle [63]). Unfortunately, studies reporting both B–N plot trajectories and size distributions seem to be relatively rare–I found only three: Hara [76], Westoby and Howell [162] (for an example of their data, see Figure 3.17) and Aikman and Watkinson [3]; the last of these reports on a computational experiment. Occasionally, the size axis is given in a logarithmic scale; I have decided not to do the same. In

---

[1] In this thesis, a concave curve is such that a line segment joining two points on the curve never lies below the curve.

so doing fine structure at the lower end of the scale may be obscured (as shown by Mohler *et al.* [118]), but my purpose here is to highlight qualitative features that are relatively coarse.

3.2 Unstructured models

In this section, I show that in an unstructured model (in the sense that individuals are not distinguished from one another by any structure variable), it is not in general possible to distinguish among the effects of age, size, population density or biomass density on mortality or somatic growth. Of course, in the case of an unstructured population model, only a single size can be modelled at any one time, so that in order to use data from real populations it is necessary to calculate some sort of average size. The symbol s will denote size in the case of structured as well as unstructured population models; in the case of the former it denotes individual size, and in the case of the latter it denotes an average size.

I proceed in two ways: first by constructing an example, and then by considering the partial derivatives of the constitutive relations (in a way similar to Butterworth and Punt [20]). The example yields a set of models that have different biological interpretations and statistical properties (in a sense to made precise later), but are analytically equivalent. Although this does not rule out situations where a data set could be fitted by only one from a class of unstructured models, it certainly rules out the claim that such a selection is generally possible. Moreover, I argue below that data for which similar examples could be constructed occur frequently, so that where data are limited to B–N plots, one would as a rule be unable to distinguish between causal effects due to size or due to density.

Let $P(t)$ and $s(t)$ denote the population and the mean size respectively at time t. The slope of the B–N plot is $d(\log sP)/d(\log P)$; for the sake of this example I prefer to work with the relationship of s and P directly; to this end, define c as the slope of the plot of $\log s$ versus $\log P$. That is, $c = d(\log s)/d(\log P)$. Let $\dot{s}/s = r(1 - s/s_\infty)$, and fix $c = -3/2$, i.e.

logistic size growth and constant slope of $\log s$ versus $\log P$. Note that in the absence of structure, $\lambda = \dot{P}/P$ (where $\dot{P}$ denotes dP/dt), and that the assumption of constant slope is the exact analogy of fitting a straight line to data in the B–N plane. Now if one defines mortality by

$$\lambda(t,s,P,B) = a_1\lambda_s(s) + a_2\lambda_t(t) + a_3\lambda_P(P) + a_4\lambda_B(B)$$

that is, as the sum of size, age, population density and biomass density components, then four possible models for λ, in each of which only one component is operative, may be defined by taking only one $a_i \neq 0$ for $i = 1$ to 4. In these four cases, λ_s, λ_t, λ_P and λ_B can be obtained explicitly, as I now show.

By solving the differential equation for $\dot{s}$, one has $s(t) = s_0 s_\infty/(s_0 + e^{-rt}(s_\infty - s_0)$, where s_0 and s_∞ represent the initial mean size and the stable mean size respectively. Since by the definition of c, $\frac{3}{2}\lambda = \dot{s}/s$, one has

Case i=1: $$\lambda \equiv \lambda_s(s) = \alpha\left[1 - \frac{s}{s_\infty}\right],$$

where $\alpha = 2r/3$. Into this, one may substitute the explicit formula for $s(t)$ to obtain

Case i=2: $$\lambda \equiv \lambda_t(t) = \frac{\alpha}{\beta e^{rt} + 1},$$

where $\beta = s_0/(s_\infty - s_0)$. With a bit of work, one may now solve $\dot{P}/P = \lambda_t$, express t as function of P and substitute in the above to get

Case i=3: $$\lambda \equiv \lambda_P(P) = \alpha\left[1 - \frac{s_0}{s_\infty}\left(\frac{P_0}{P}\right)^{\gamma}\right],$$

where $\gamma = r/(\alpha\beta)$. Similarly, since $P(t)$ is now known, one may form $B(t) = s(t)P(t)$, invert this to get t as a function of B, and substitute as before to get

Case i=4: $$\lambda \equiv \lambda_B(B) = \alpha\left[1 - \frac{s_0}{s_\infty}\left(\frac{B_0}{B}\right)^{\gamma/(1-\gamma)}\right].$$

The four formulae for mortality are of course analytically equivalent, since they were derived from a single source by analytical procedures. However, they clearly have different biological interpretations. Moreover, they have different statistical properties in the following sense: if the parameters α, s_∞, β and γ (as appropriate) are estimated from data, then for many common methods of estimation, one would obtain slightly different fits with the different models. This means that a set of data relating size to population density can be interpreted in terms of any or all of age dependence, size dependence, population density dependence or biomass density-dependence.

The conclusions from this example rely only on knowing $s(t)$ and the relationship between s and P. In this case, $s(t)$ and c were given analytically, but they could equally well be specified in terms of curves fitted to data. All that is required for similar constructions is that $P(t)$ and $s(t)$ be invertible, because then s considered as a function of P will be invertible too. A similar procedure would succeed if $P(t)$ were known instead of $s(t)$. Monotonic population decline and somatic growth are quite general in the self-thinning literature, in which case the required relationships would certainly be invertible. Indeed, this is only biologically reasonable—the only exceptions would seem to be loss of biomass as a way of coping with drought or winter, through herbivory, or because of reproductive effort. Certainly in even-aged stands, the population density can only decline. It therefore seems likely that the confounding of influences on mortality and somatic growth on B–N plot data only for even-aged monospecific stands is to be expected generally.

Alternatively, following Butterworth and Punt [20]), one may indicate dependence on age, size, population density and biomass density in mortality and somatic growth by the symbols $\lambda(a, s, P, B)$ and $\dot{s}(a, s, P, B)$. This is perfectly compatible with the framework of chapter 2, since P and B are functionals of ρ. Now consider a plot of λ versus time. To explain changes in mortality with time is to explain the slope of this plot. However, this slope has four components:

$$\frac{d\lambda(a,s,P,B)}{dt} = \frac{\partial\lambda}{\partial a} + \frac{\partial\lambda}{\partial s}\dot{s} + \frac{\partial\lambda}{\partial P}\frac{dP}{dt} + \frac{\partial\lambda}{\partial B}\frac{dB}{dt}.$$

It follows that knowledge of mortality at various times cannot suffice to specify whether the mortality depends on age, size, population density or biomass density. The same argument applies to somatic growth. Thus, the effects of age, size, etc. are confounded; if one wants to disentangle these effects one has to consider additional data and more complicated models.

3.3 Computer experiments for a class of structured population models.

3.3.1 Model formulation

In general, a structured continuum model is based on the density $\rho(a, \mathbf{x}, t)$, where a is age, t is time, and $\mathbf{x}$ is a vector that specifies the value of the structure variables other than age. The individual or *i*-state (*sensu* Metz and Diekman [113], see chapter 2) is then the vector $(a, \mathbf{x})$. The choice of variables represented in $\mathbf{x}$ determines the structural detail, besides age, that the model could simulate. Kohyama [94, 95] and Hara [77] use one-dimensional $\mathbf{x}$, with size as the only component. The *i*-state in these cases consists of age and size, which together suffice to specify the attributes of an individual. I have chosen a second component for $\mathbf{x}$, denoted by r, which is constant for an individual plant and may be interpreted as the intrinsic growth rate of that plant as determined by genotype and micro-environment. There are two good reasons for this: firstly it allows one to model a non-homogeneous environment, and secondly it renders the model more dynamic, in that the initial size hierarchy may be upset by initially smaller but faster growing plants. The latter property seemed especially important, since I wanted to use the same initial size distribution for all models. Thus the initial sizes were not available as parameters for fitting the data. On the other hand, it is clear from the literature that small differences in initial size rapidly leads to large differences between the sizes of the smallest and largest plants (for example, Westoby [161], Hara [75]). This would be difficult for models where somatic growth is independent of relative size (i.e. types $(\dot{s}/s)_i$ and $(\dot{s}/s)_a$ described below), unless a range of growth rates were used. There is some cost to this, of course: r is like a parametric function, so that it increases the amount of data needed for reasonably good parameter estimates. In principle this requires an infinite number of parameters to be estimated; in practice one limits oneself to families of functions defined by only a few parameters—I chose the two-parameter family of constant population density between r_{lo} and r_{hi}.

Thus the basic model element in this chapter is the density $\rho(r, s, t)$, where $\mathbf{x} = (r, s)$, $r, s \in \mathbf{R}^+$, and $\mathbf{R}^+$ denotes the positive real numbers. Here, s is to be interpreted as individual size (the simplest would be always to think of it as dry mass). I assumed that r is constant throughout a plant's life; r should be interpreted as intrinsic growth rate as determined by the genotype and the microenvironment of an individual. Since r is constant, the only constitutive relations required are for mortality and somatic growth, which I denote generically by λ and $(\dot{s}/s)$ respectively. By the use of subscripts, three different types of dependence on size are indicated:

λ_i and $(\dot{s}/s)_i$	independent of size;
λ_a and $(\dot{s}/s)_a$	absolute size dependence, i.e. depends only on the size of the individual concerned;
λ_r and $(\dot{s}/s)_r$	relative size dependence, i.e. depends on how large the individual concerned is relative to other measures of size—in the models considered here, these other measures are current smallest size, mean size and largest size .

I used four different constitutive relations for λ and five for $(\dot{s}/s)$. The names and formulae of these relations are given in Table 3.1, and in Table 3.2 the interpretation of the various parameters may be found. These models form a family in the sense that all λ are of the form $\lambda_0 f(s) g(P)$ and all $(\dot{s}/s)$ are of the form $r f(s) g(P)$. The generic factor $g(P) = 1 - (P_\infty/P)^\alpha$ is common to all constitutive relations—see Figure 3.1, and the genericity of α means that α_λ and α_s are used in place of α to indicate parameters in equations for λ and in $(\dot{s}/s)$ respectively. Similar genericity applies to other parameters whenever confusion seems possible.

The names of these models were chosen to reflect some of their features. A name starting with "D..." refers to a model for $(\dot{s}/s)$, reflecting the dot on the s; otherwise the name refers to a model for λ. Since all models share $g(P)$, all names contain either "REC" or "RE" for the reciprocal that occurs in the formula for $g(s)$. Models differ in the actual $f(s)$ they use. In the cases of λ_i and $(\dot{s}/s)_i$, which are the simplest, $f(s) = 1$

Mortality	Type	$f(s)$
REC	λ_i	1
RECHP	λ_r	$\left(\frac{s+b}{s_2+b}\right)^{\frac{\ln a}{\ln(s_1+b)-\ln(s_2+b)}}$
RECPO	λ_r	$\left(\frac{s_2-s_1+b}{s-s_1+b}\right)^{\beta}$
RECLG	λ_a	$1-(s/s_\infty)^\beta$
Somatic growth	Type	$f(s)$
DREC	$(\dot{s}/s)_i$	1
DREPO	$(\dot{s}/s)_r$	$\left(\frac{s-s_1+b}{s_2-s_1+b}\right)^{\beta}$
DRESG	$(\dot{s}/s)_r$	$\begin{cases} a & \text{for } s < s_1 \\ a+(1-a)\exp\left(-d\left(\frac{s_2-s}{s-s_1}\right)^{\beta}\right) & \text{else} \end{cases}$
DRESC	$(\dot{s}/s)_r$	$a+(c-a)\frac{s-s_1}{s_2-s_1}$
DRELG	$(\dot{s}/s)_a$	$1-(s/s_\infty)^\beta$

Table 3.1: The names and formulae for $f(s)$ for the various constitutive relations used. These $f(s)$ are multiplied with $g(P)$ to give λ and $(\dot{s}/s)$—see Table 3.3 for the various combinations that make up the models. Subscripts denote type of dependence on size as follows: i: independent of size; r: dependent on size relative to size of others; a: dependent on own size only. See Figures 3.1–3.3 for graphic depiction of these relations.

and is not reflected in these two names—REC and DREC respectively. Else, the last two letters of the name refer to properties of the $f(s)$ that was used in that model. So RECLG and DRELG are in a generalised sense logistic in s; RECPO and DREPO are powers of a ratio; RECHP reminded me of an hyperbola; DRESG has a sigmoid shape; and DRESC is an affine scaling with respect to size. See Figures 3.2–3.3 for graphic depictions of the $f(s)$ used in these models.

These constitutive relations were chosen to achieve the following prop-

α	Shape parameter for population dependence in λ and $(\dot{s}/s)$.
β	Shape parameter for size dependence in λ and $(\dot{s}/s)$.
λ_0	The value of λ when $f(s)g(P) = 1$.
$\bar{s}, s_{max}$	Current mean and largest size.
a	Value of $f(s)$ at s_1.
b	Translation parameter in λ_r or $(\dot{s}/s)_r$.
c	Value of $f(s)$ at s_2.
d	Defined by $f(\frac{s_1+s_2}{2}) = e^{-\mathrm{d}}$.
P_∞	Stable population size.
r_{lo}, r_{hi}	The lowest and highest values of r.
s_1, s_2	The smaller and larger of the two reference sizes used in relative size dependence of λ and $\dot{s}/s$. Note that s_1 is either s_{min} or $\bar{s}$, and s_2 is either $\bar{s}$ or s_{max}.
s_∞	Stable individual size.
t_{end}	The final time.

Table 3.2: Referents of symbols used in the constitutive relations and discussion of parameter estimation. Note the genericity: for instance the α occurring in the $f(s)$ of a mortality (respectively somatic growth) model may be denoted as a_λ (respectively a_s) if confusion seems possible.

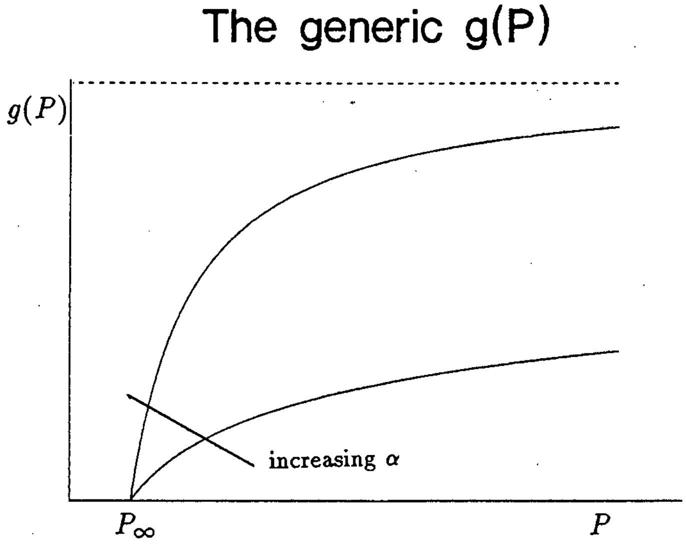


Figure 3.1: Graph of $g(P)$, the generic dependence on P used in all the constitutive relations, where P denotes the size of the total population.

erties:

(i) as $t \to \infty$, both λ and $(\dot{s}/s) \to 0$

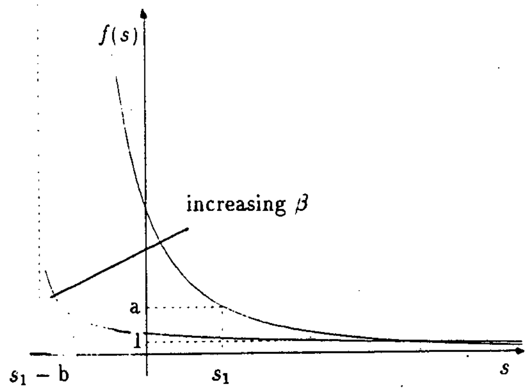
(ii) as $s \to s_\infty$, both λ_a and $s_a \to 0$

(iii) λ_r is larger for smaller s, and $\dot{s}_r$ is smaller for smaller s, and generically $f(s_1) = a$ and $f(s_2) = 1$.

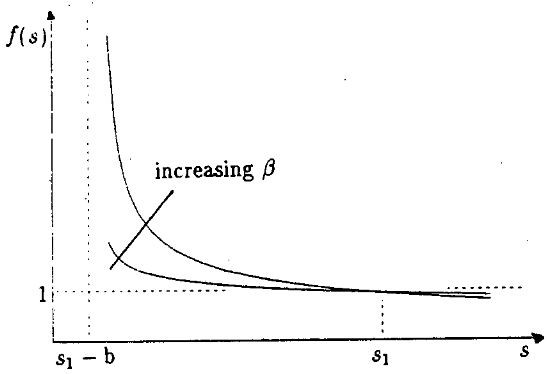
Property (i) is achieved by the common factor defining the family, $g(P)$. However, property (ii) is not always compatible with property (i), for in the cases of combining λ_a with either $(\dot{s}/s)_i$ or $(\dot{s}/s)_r$, there may be a size beyond which somatic growth continues while mortality is zero. In practice this difficulty was avoided by appropriate parameter choices.

Because each of λ and $(\dot{s}/s)$ may be of three different types, there are nine possible combinations with respect to the type of size dependence, and nine classes of model to consider. In some of these classes more than one combination was used—see Table 3.3 for details of which constitutive relations.

(i) RECHP



(ii) RECPO



(iii) RECLG

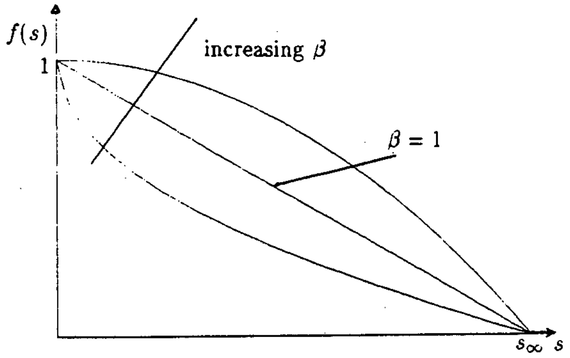


Figure 3.2: Graphs of $f(s)$ used in λ, the constitutive relation giving mortality. (i) $f(s)$ used in RECHP. (ii) $f(s)$ used in RECPO. (iii) $f(s)$ used in RECLG.

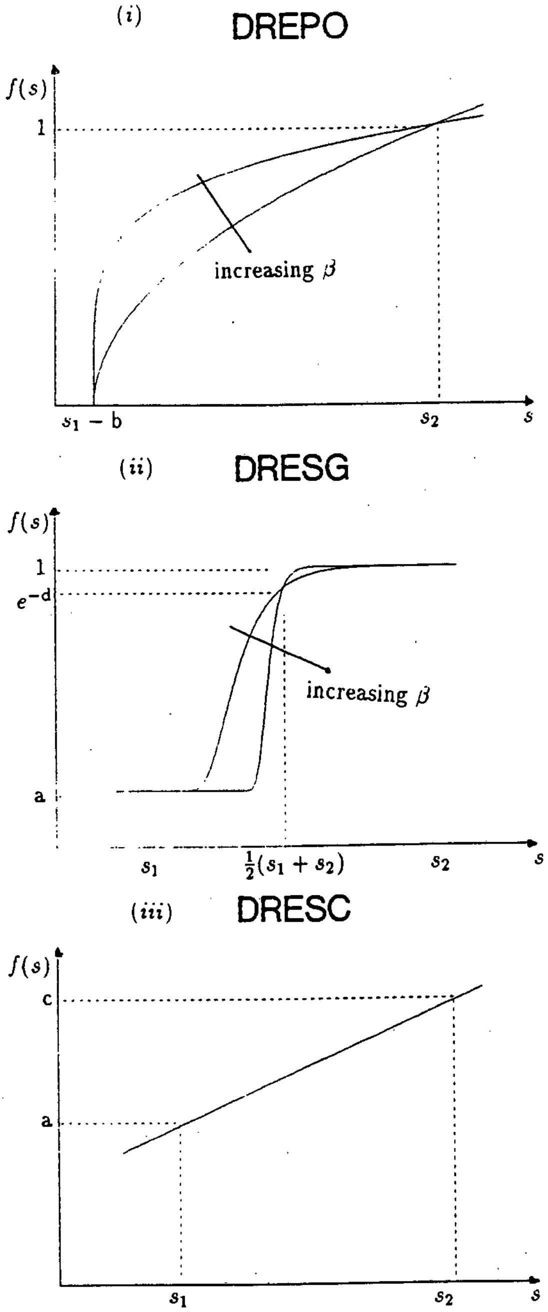


Figure 3.3: Graphs of $f(s)$ used in $(\dot{s}/s)$, the constitutive relation giving somatic growth. (i) $f(s)$ used in DREPO. (ii) $f(s)$ used in DRESG. (iii) $f(s)$ used in DRESC. Note that DRELG is not depicted here; its graph is similar to RECLG.

Model number	λ	$(\dot{s}/s)$	$(s_1)_\lambda$	$(s_2)_\lambda$	$(s_1)_s$	$(s_2)_s$
1	REC	DREC	-	-	-	-
2	RECLG	DREC	-	-	-	-
3	RECLG	DRELG	-	-	-	-
4	REC	DRELG	-	-	-	-
5a	RECLG	DREPO	-	-	$\bar{s}$	$s_{\max}$
5b	RECLG	DRESG	-	-	$\bar{s}$	$s_{\max}$
6a	REC	DRESC	-	-	$s_{\min}$	$s_{\max}$
6b	REC	DREPO	-	-	$\bar{s}$	$s_{\max}$
7a	RECHP	DREPO	$\bar{s}$	$s_{\max}$	$\bar{s}$	$s_{\max}$
7b	RECHP	DRESG	$s_{\min}$	$s_{\max}$	$\bar{s}$	$s_{\max}$
8	RECHP	DRELG	$s_{\min}$	$\bar{s}$	-	-
9a	RECHP	DREC	$\bar{s}$	$s_{\max}$	-	-
9b	RECPO	DREC	$s_{\min}$	$\bar{s}$	-	-

Table 3.3: Combinations of constitutive relations and arguments of relative size dependence used in the different models.

3.3.2 Existence and uniqueness of solutions

The theory presented in chapter 2 covers all of the models used in this chapter, with the exception of DREPO, which occurs in models 5a, 6b and 7a. This is because r and s are bounded, ρ_0 is finite, and, as is easily verified, λ and $\dot{s}$ are Lipschitz, because $g(P)$ and all the $f(s)$ other than the one used in DREPO are Lipschitz. The only cases where this claim is not obvious are the vertical asymptotes in RECHP and RECPO. In the first case, the asymptote is for a negative value of s, which cannot occur because $\dot{s}$ is positive. For RECPO, it is true that $s = s_1 - \mathrm{b}$ is possible, but in model 9b, which is where RECPO is used, s_1 is the current minimum size, and since $\mathrm{b} > 0$, there is never an s within a distance b from the asymptote.

The models containing DREPO have a vertical tangent at $s_1 - \mathrm{b}$. Moreover, in all these cases, $s_1 = \bar{s}$, so it is certainly possible that $s = s_1 - \mathrm{b}$. Hence in those three models, the theory of chapter 2 fails. However, the theory can be salvaged, at least in a neighbourhood of the parameters used here, and I now sketch a proof of the claim.

Given some $\rho \in L_T$, the characteristics form a flow in the sense that $\phi_{\mathbf{x}}(s) = \phi_{\mathbf{y}}(s) \iff \mathbf{x} = \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \Omega$, where, as in chapter 2, $s \in [t - \tau, T - t]$ and $\tau = \min(a, t)$. Existence of solutions follows because $\dot{s}$ is bounded and continuous. Furthermore, uniqueness follows from the fact that all initial sizes exceed the critical size for parameters used here, so that initially all plants grow at a positive rate. However, the larger plants grow faster. It is true that s_1 increases as well, and eventually growth rates drop to zero for some plants, but the smallest plants stop growing first. Hence if two plants with identical intrinsic growth rate start at different sizes, they will always remain at different sizes.

Because the characteristics form a flow, the changes of variable used in proving theorem 2 remain well defined. However, the estimates using Lipschitz properties of $\mathbf{v}$ are no longer valid. So the proof breaks down in the treatment of the integrals called I_1, I_2 and I_3. The last of these is zero, since ρ_0 is constant. Since there is no integration over age, changes of variable involving J_ϕ are not necessary. Then using the Lipschitz properties of F and G reduces the problem to showing the $\|\phi_{\mathbf{x}}(s) - \hat{\phi}_{\mathbf{x}}(s)\|_1 \leq sA(s)\|\rho - \hat{\rho}\|L_t$ for any $s \in [0, t]$ and all $\mathbf{x} \in \Omega$, where $A(s)$ is positive and non-decreasing. Using the integral definition of $\phi_{\mathbf{x}}$, this then follows from the fact that a bounded function which is non-Lipschitz on a finite set of points, when integrated over a finite domain, yields a function that is Lipschitz with respect to the positions of the non-Lipschitz points.

Therefore a unique solution exists for the equations corresponding to all the models in this section. These can be calculated in any convenient manner; I used Algorithm A as described in chapter 2.

3.3.3 Computational method and parameter search

The models specified above were solved by the technique described in chapter 2. That is, the initial density $\rho(r, s, 0)$ was discretised in terms of its value at nodal points in the subset of Ω where ρ is non-zero. This reduces the problem to the solution of a system of ordinary differential equations for ρ and s, which were solved by a standard fourth-order

Runge-Kutta algorithm. With this method, the non-linearities in the models posed not particular problem as long as the length of the time step was carefully managed. The only technical difficulty was to extract a satisfactory size distribution from the values of ρ and s at t_{end}. Let σ denote the final size density as a function of s only, so that

$$\sigma(s) = \int_{r_{lo}}^{r_{hi}} \rho(r, s, t_{end})dr.$$

In order to get a reasonable representation of the very sharp peaks in some of the σ that were observed, I found it useful to select the boundary of the mesh for the s-values at which to report σ. For a given s value, the interval of r-values at which ρ is not zero was obtained, by quadratic interpolation where necessary. Then r was given a regular mesh on this interval, and at the resultant points, the value of ρ was approximated by linear interpolation from its values at the corners of the quadrilateral in which the point occurred. Finally, the integral giving $\sigma(s)$ was evaluated by Simpson's rule.

For the models in which mortality is of type λ_i, it is possible to solve explicitly for $P(t)$, and hence for some of these models it is also possible to find $s(t)$ for given r and s_0. These equations were integrated with the help of Maple V, thus giving an objective test of the computer code and the error control. Since the discretisation in r and s was fixed, only the time step was adjusted, with doubling attempted at every instance, up to a maximum step length. Error tolerance was 10^{-5} in the maxnorm for both ρ and s. The times at which output in the form of $P(t)$ and $B(t)$ was generated formed part of a geometric series in order to have many more points near $t = 0$ than near t_{end}, since the dynamics were rapid at the start and slowed down consistently and considerably.

The program ran reasonably fast, and used less than half a minute of CPU time on a MicroVAX 3100–90 for the coarsest mesh. This enabled me to experiment extensively with a variety of models and many parameter combinations—in fact, to carry out a rudimentary estimation procedure by trial and error search. The results given below were obtained as follows: I fixed $P_\infty = 60$, $t_{end} = 100$ (a fixed final time is suitable because it models the fixed times for which size distribution data are available), and the initial sizes as the interval $[0.5, 1.5]$. The initial density $\rho(r, s, 0)$ was taken to be constant, with a value chosen to obtain $P(0) = 2000$. Then, for a given model, the parameters to be estimated were λ_0, α_λ, $\alpha_{\dot{s}}$, r_{lo}, r_{hi}, and also the parameters in the two factors $f(s)$ in λ and $(\dot{s}/s)$. The model with the fewest parameters to be estimated is model 1, with 5 parameters, and the one with the most is model 7b, with 10 parameters. The criterion for a successful parameter estimate was the following qualitative B–N plot: concave, with slopes approximately -1 near $t = 0$ and -0.5 near $t = t_{end}$. Such curves correspond to realistic B–N plots (Zeide [166]). The parameters were searched for by trial and error, guided by heuristic understanding of the effect of changing various parameters. This was first done roughly, using a 5×5 rectangular mesh for (r, s), and then with a 41×11 mesh for refinement, thus providing a check on convergence as well as a more acceptable plot of $\sigma(s)$.

3.3.4 Results

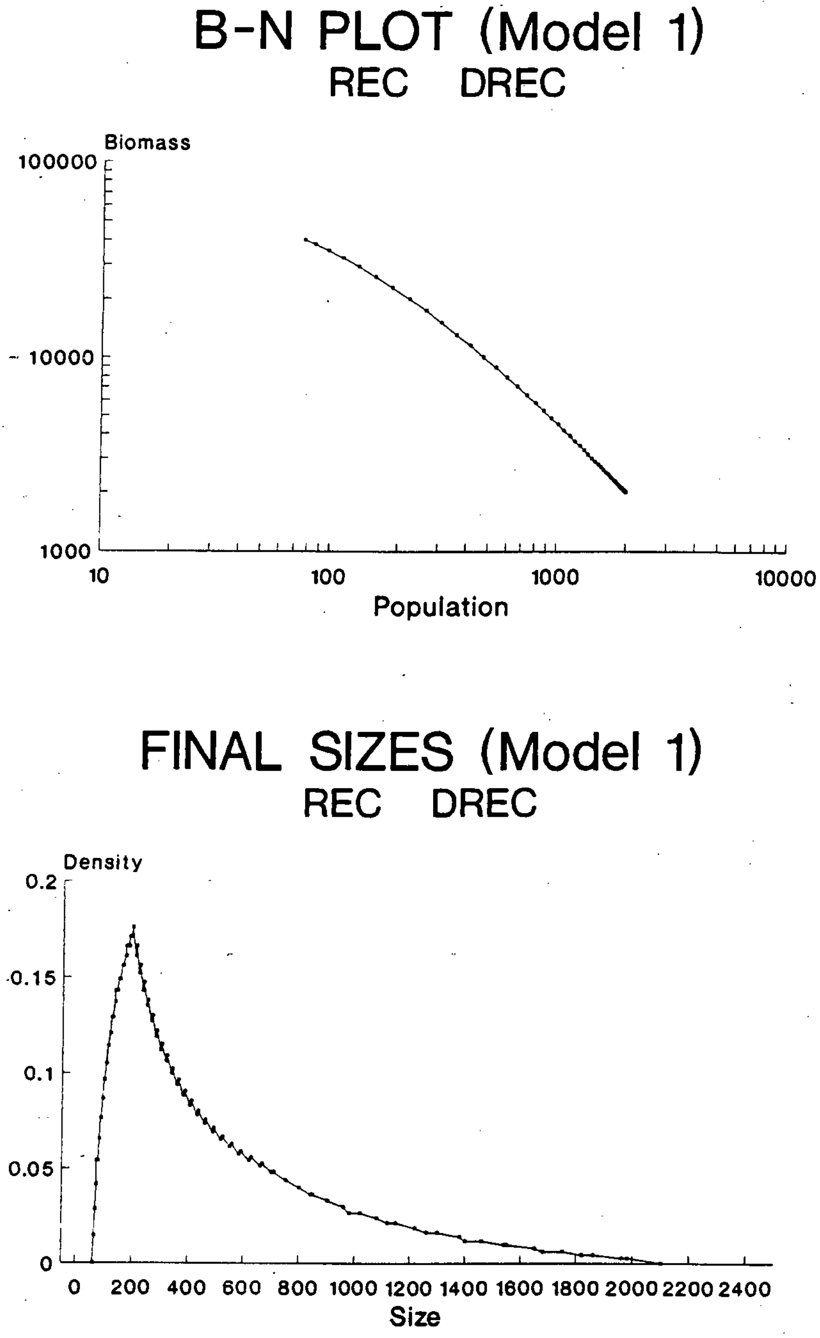
Table 3.4 gives the values of the parameters estimated by the technique above, and the B–N plots with their associated final size distributions are displayed in Figures 3.4–3.16. The initial and final slopes are recorded in the captions, and it will be seen that only Figure 3.13 seriously violates the criterion, through its lack of concavity. There is substantial variation in the values of $P(t_{end})$, but this is acceptable because only qualitative features in the B–N plot were sought. The jaggedness of some parts of the size distributions is an artefact of discretisation: they can be made smaller by refining the mesh, at considerable increase in computational expense. It should be noted that the size axis is not the same in the various figures.

3.4 Discussion

In section 3.2, I showed that one may often be able to construct a variety of analytically equivalent models to fit a B–N plot. This means that biological variables such as size and population density may be completely confounded in unstructured models. That is, models with different biological interpretations may be mathematically equivalent. However, this does not imply that they would yield the same predictions. For example, if the models in section 3.2 were fitted by least squares regression to data in the B–N plane, the resulting straight lines would not be

Model	Parameters									
	λ_0	α_λ	α_s	r_{lo}	r_{hi}					
1	0.048	1.0	0.5	0.1	0.15					
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	s_∞	β_λ			
2	0.055	0.5	0.25	0.13	0.17	20000	1.0			
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	s_∞	β_λ	β_s		
3	0.05	0.5	0.21	0.14	0.18	10000	1.0	1.0		
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	s_∞	β_s			
4	0.05	0.5	0.2	0.14	0.19	20000	1.0			
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	s_∞	β_λ	β_s	b	
5a	0.05	1.0	0.6	0.1	0.14	3000	3.0	0.04	10.0	
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	s_∞	β_λ	β_s	a	d
5b	0.045	0.8	0.2	0.22	0.26	3000	3.0	20.0	0.4	0.05
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	a	c			
6a	0.05	1.0	1.0	0.12	0.14	0.5	1.0			
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	β	b			
6b	0.055	1.0	0.7	0.1	0.14	0.05	10.0			
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	a	b_λ	β	b_s	
7a	0.025	1.0	0.8	0.07	0.12	1.8	10.0	0.04	10.0	
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	a_λ	b	β	a_s	d
7b	0.02	1.0	0.5	0.13	0.15	5.0	1.0	5.0	0.2	0.1
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	a	b	s_∞	β_s	
8	0.04	1.0	1.5	0.08	0.12	3.3	10.0	2500	0.25	
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	a	b			
9a	0.04	0.5	0.25	0.14	0.18	1.5	10.0			
	λ_0	α_λ	α_s	r_{lo}	r_{hi}	β	b			
9b	0.05	0.5	0.3	0.1	0.14	0.5	1.0			

Table 3.4: The parameter values as estimated by trial and error constrained by criteria on the resultant B–N plot.



B-N PLOT (Model 1)
REC DREC

FINAL SIZES (Model 1)
REC DREC

Figure 3.4: B–N plot and size distribution at final time for model 1. Initial slope is -1.2 and final slope is -0.5.

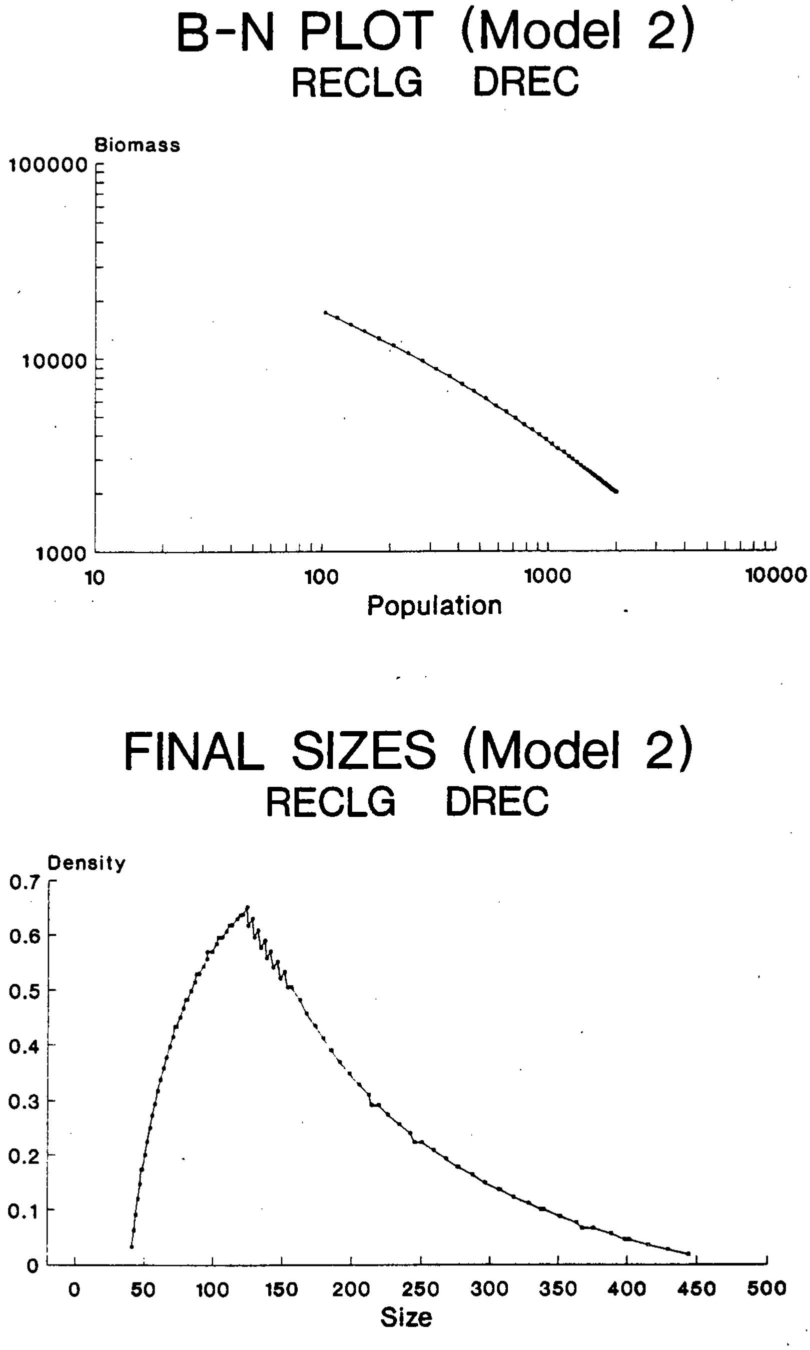


Figure 3.5: B-N plot and size distribution at final time for model 2. Initial slope is −0.93 and final slope is −0.52.

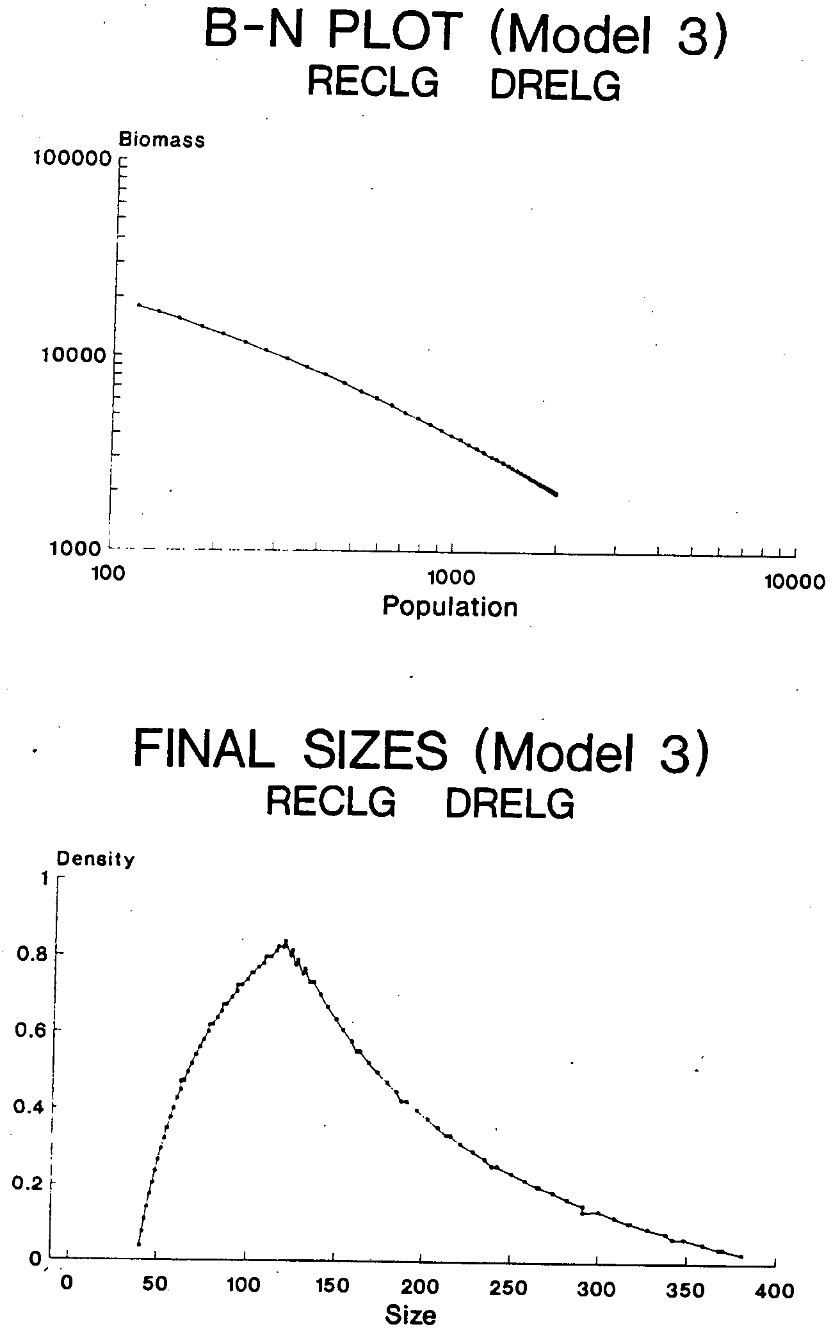


Figure 3.6: B–N plot and size distribution at final time for model 3. Initial slope is −1.0 and final slope is −0.52.

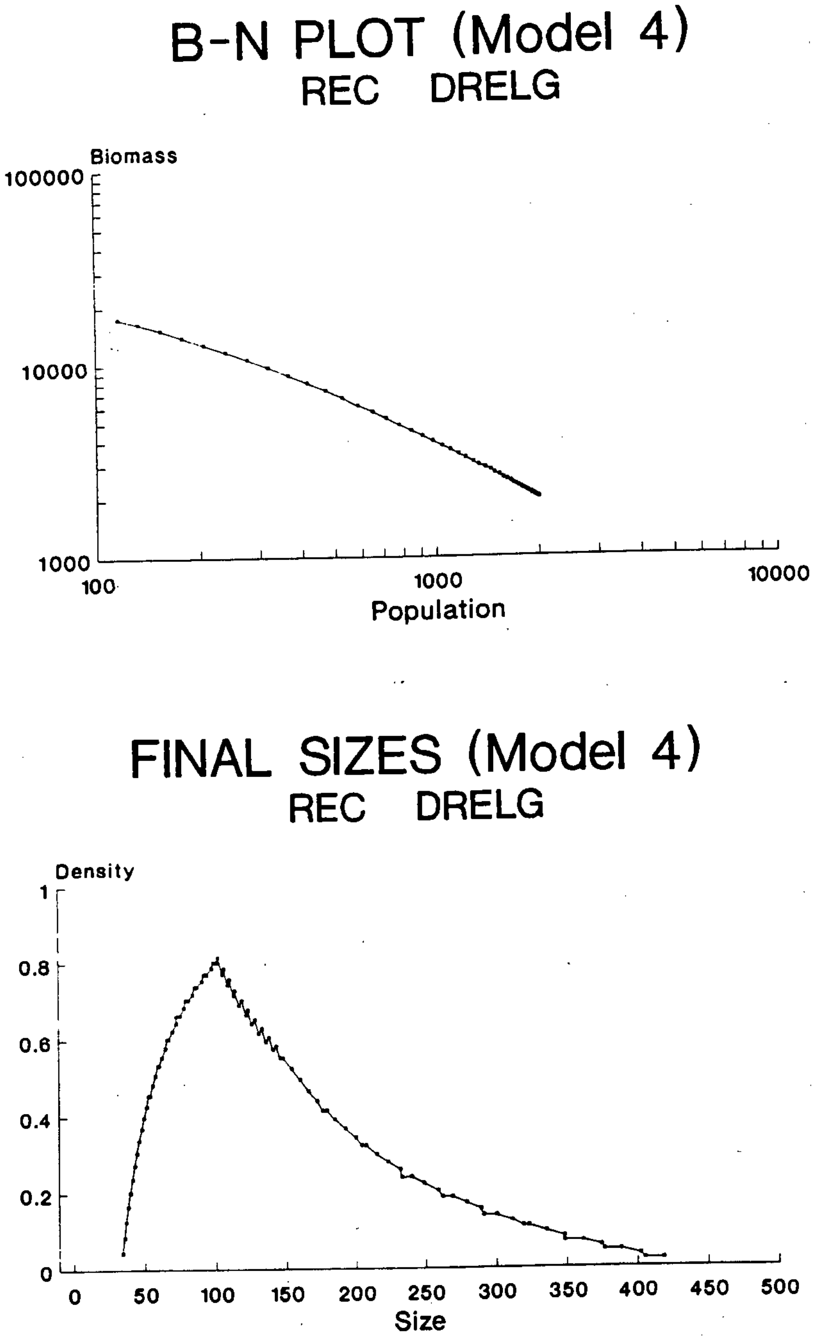


Figure 3.7: B–N plot and size distribution at final time for model 4. Initial slope is -1.0 and final slope is -0.50.

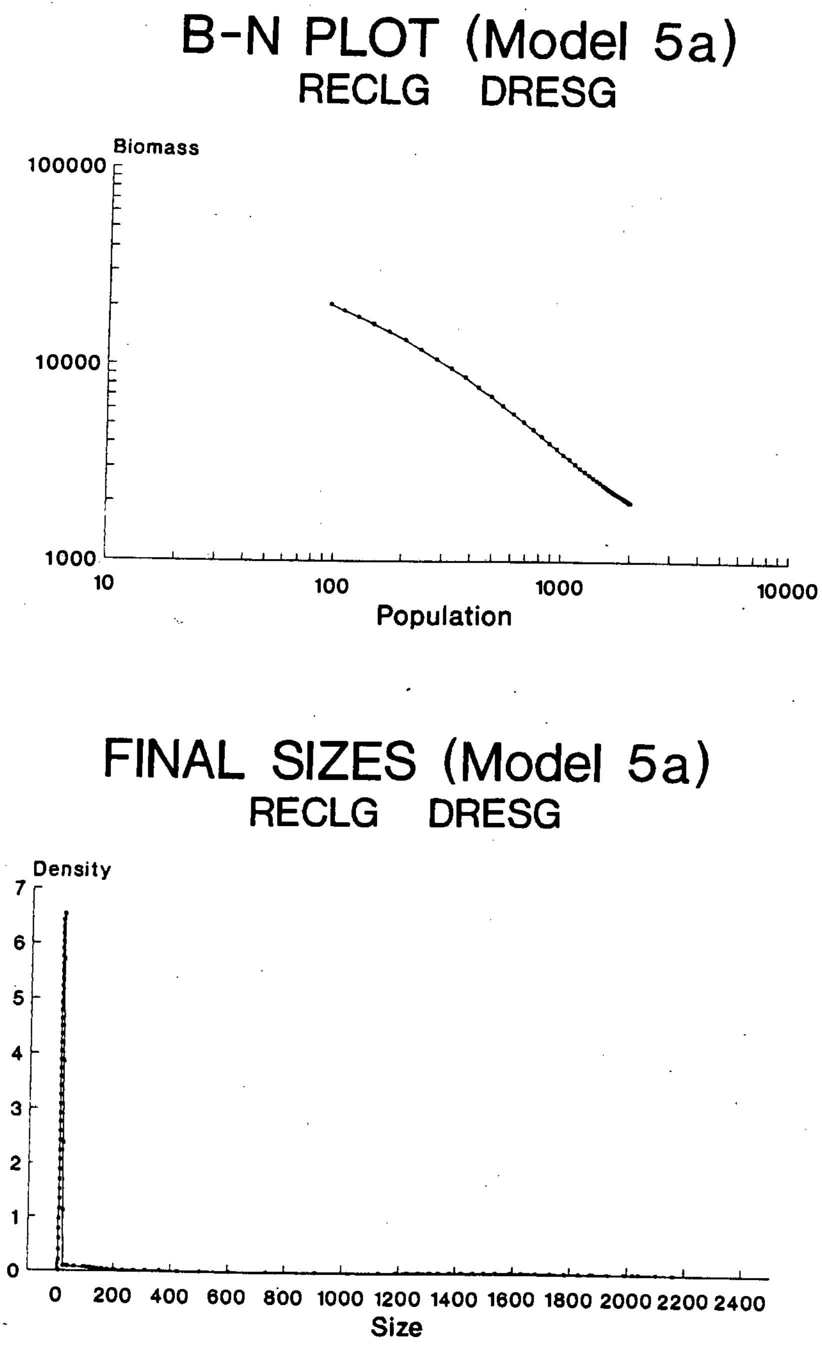


Figure 3.8: B–N plot and size distribution at final time for model 5a. Initial slope is −0.7, rising to −0.96 and final slope is −0.52.

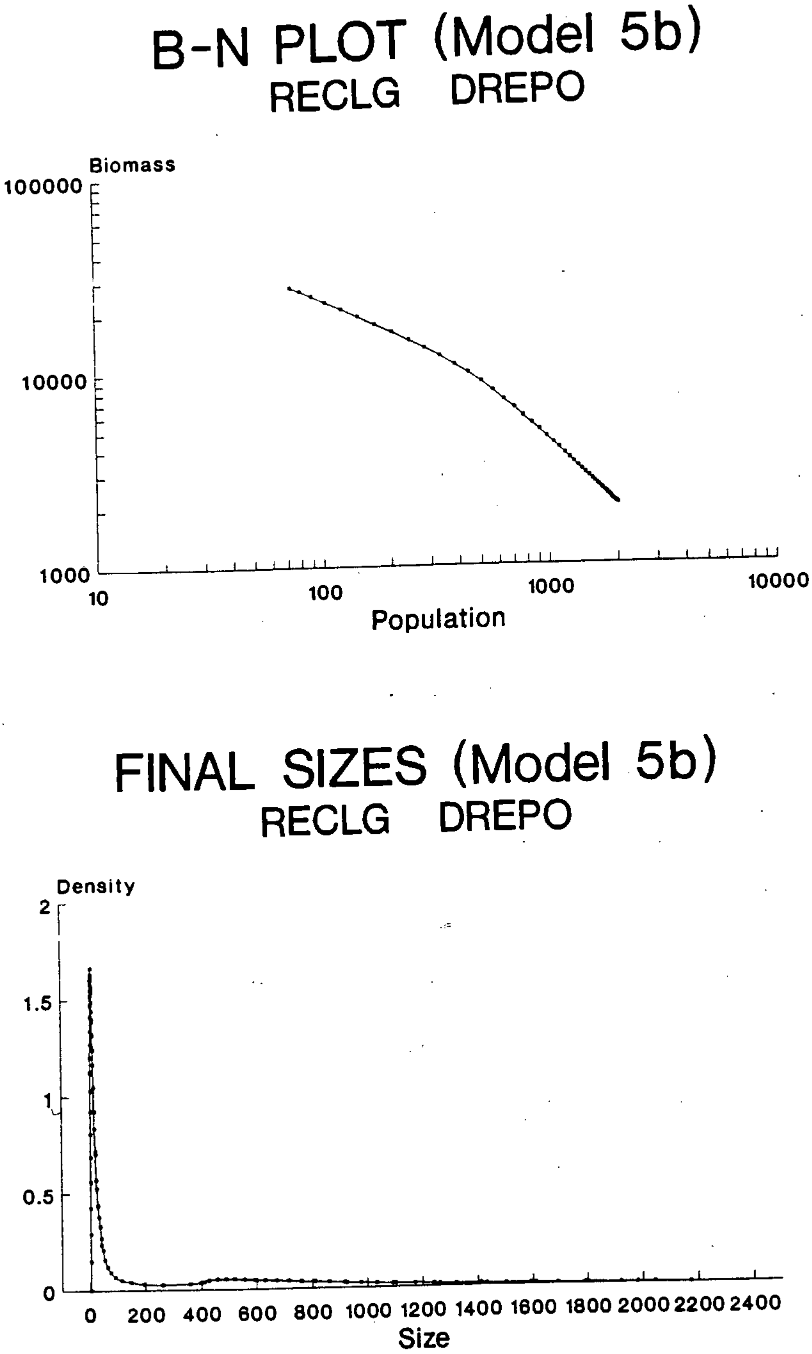


Figure 3.9: B–N plot and size distribution at final time for model 5b. Initial slope is -1.2 and final slope is -0.50.

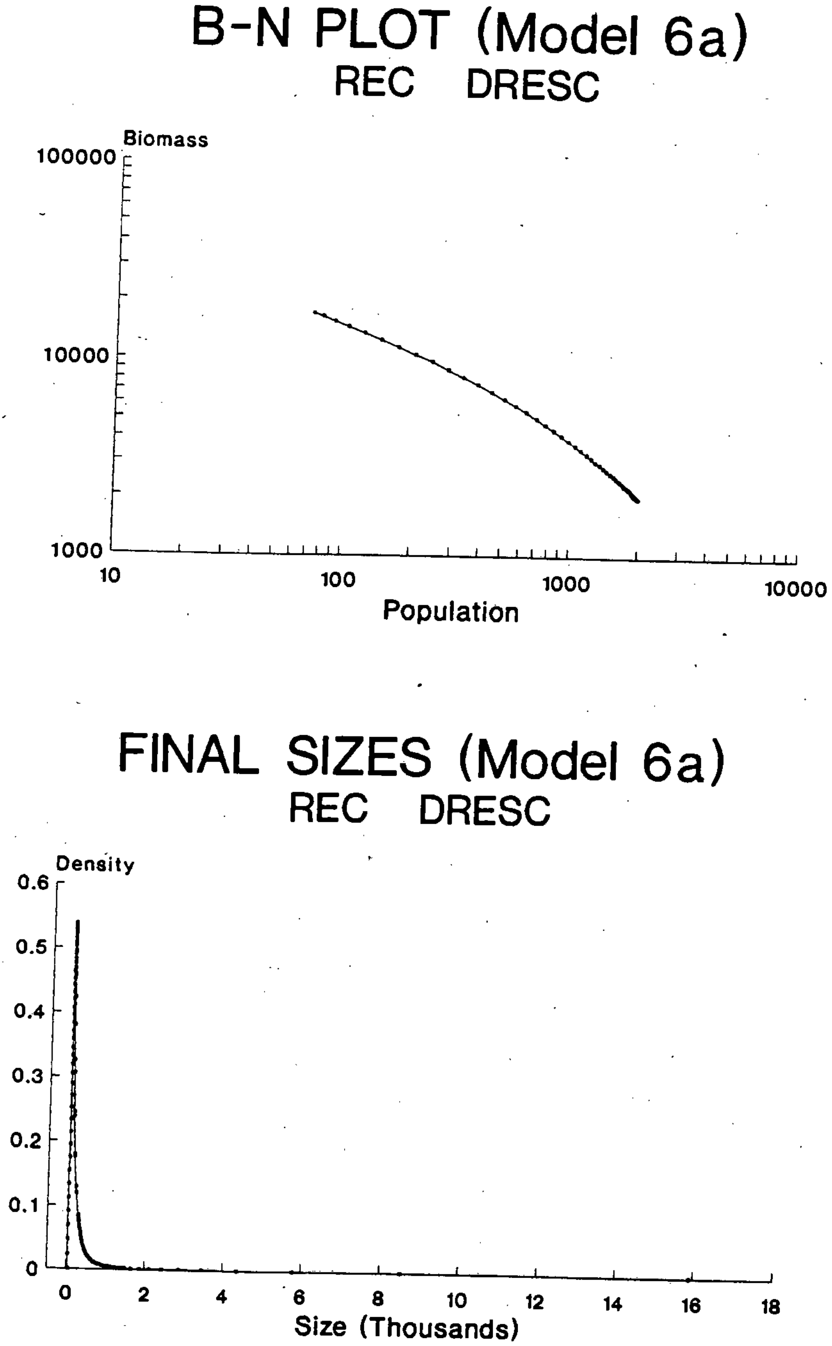


Figure 3.10: B–N plot and size distribution at final time for model 6a. Initial slope is -1.1 and final slope is $-.44$.

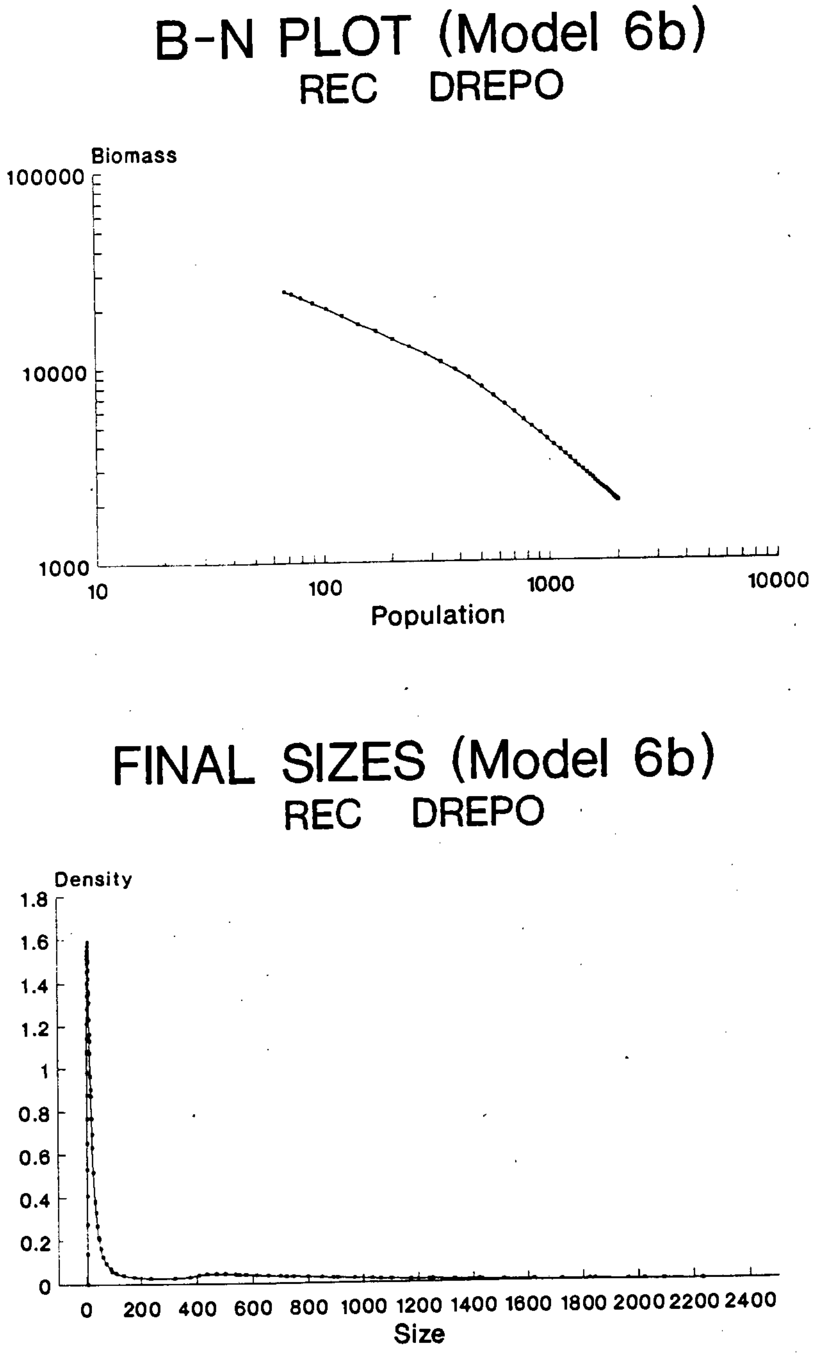


Figure 3.11: B–N plot and size distribution at final time for model 6b. Initial slope is −1.1 and final slope is −0.49.

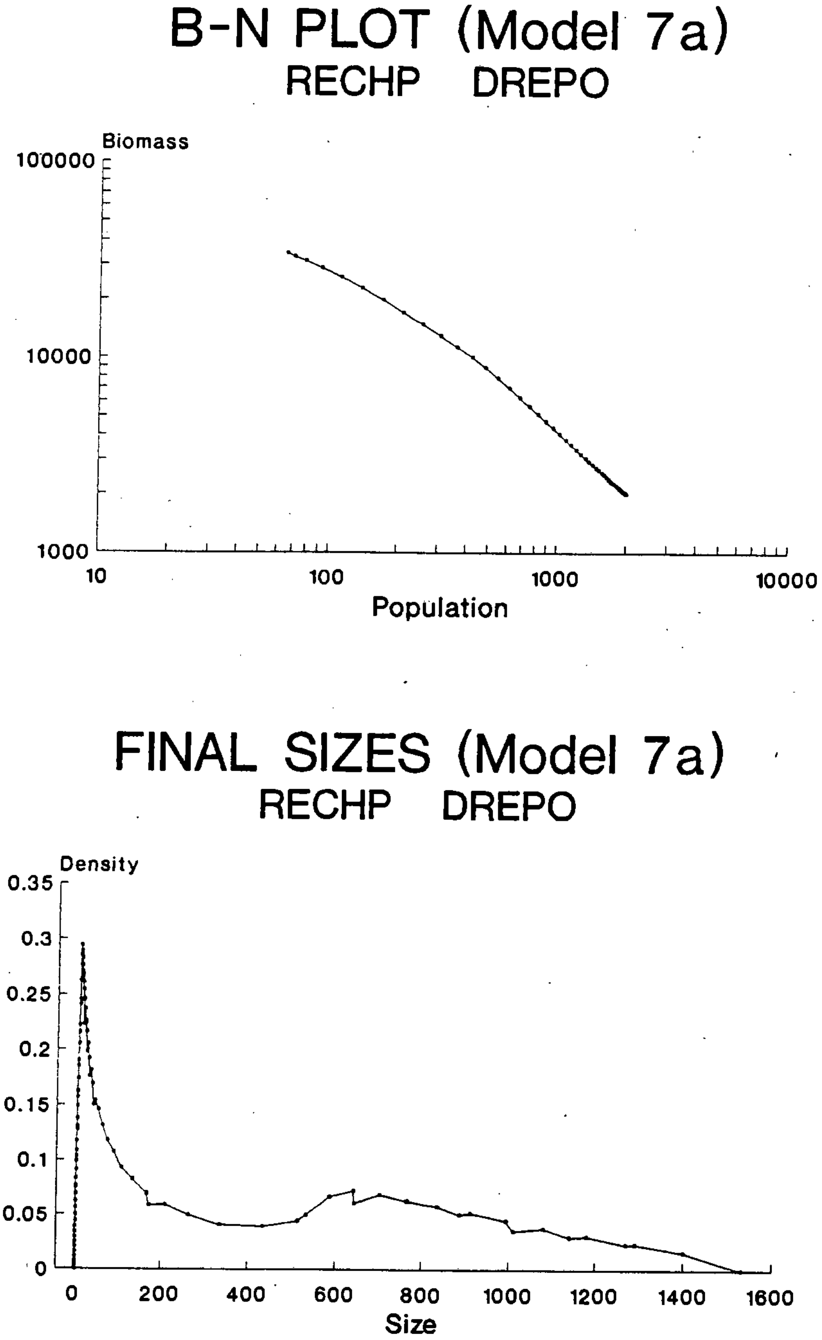
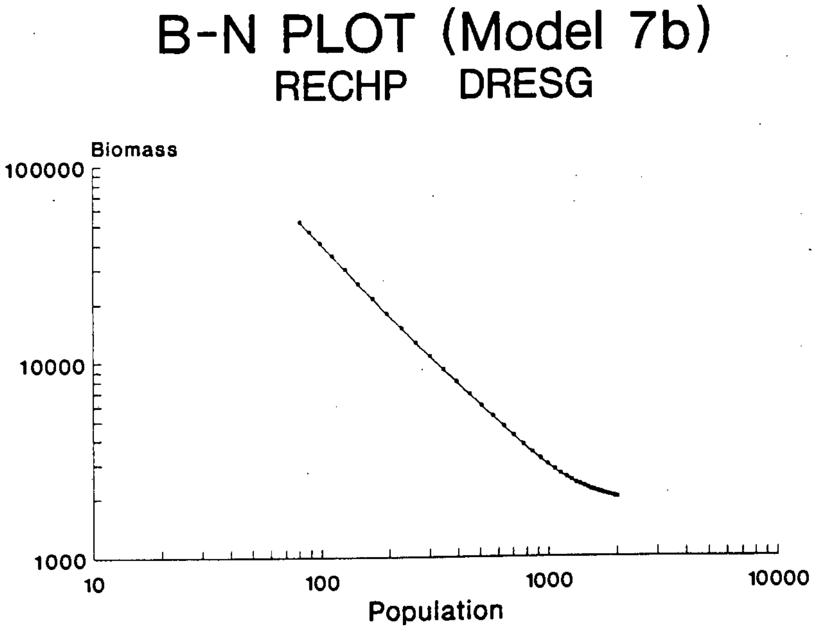


Figure 3.12: B–N plot and size distribution at final time for model 7a. Initial slope is -1.0 and final slope is -0.48.



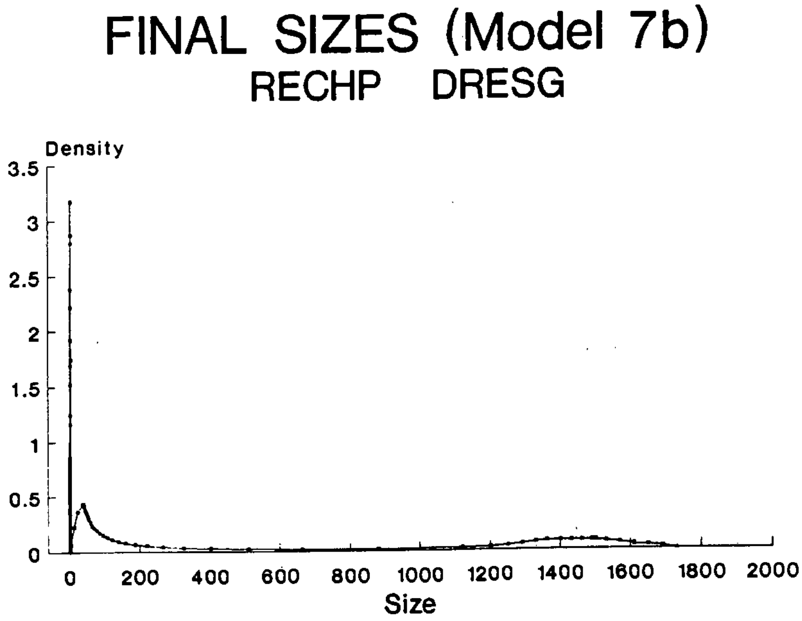


Figure 3.13: B–N plot and size distribution at final time for model 7b. Initial slope is −0.46 and final slope is −1.2.

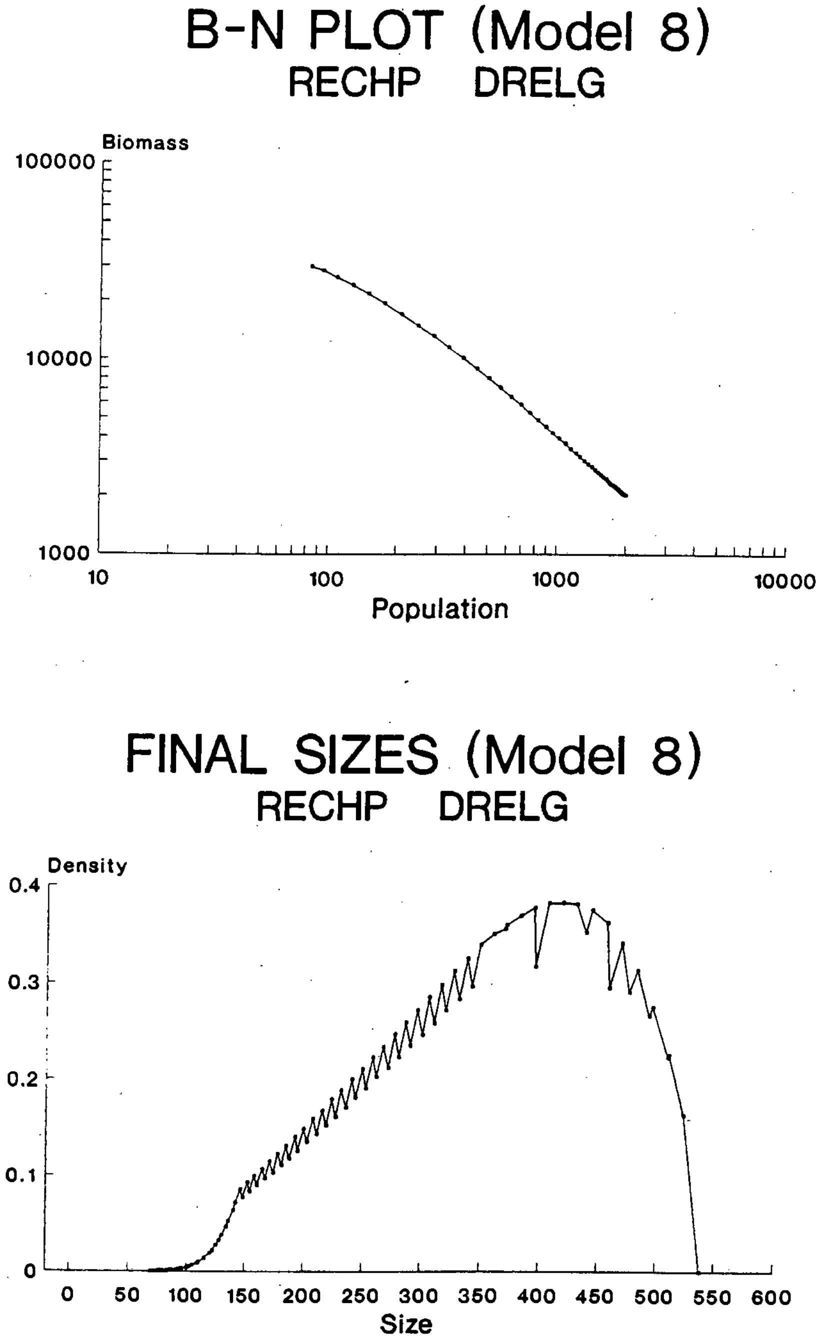


Figure 3.14: B–N plot and size distribution at final time for model 8. Initial slope is −0.93 and final slope is −0.45.

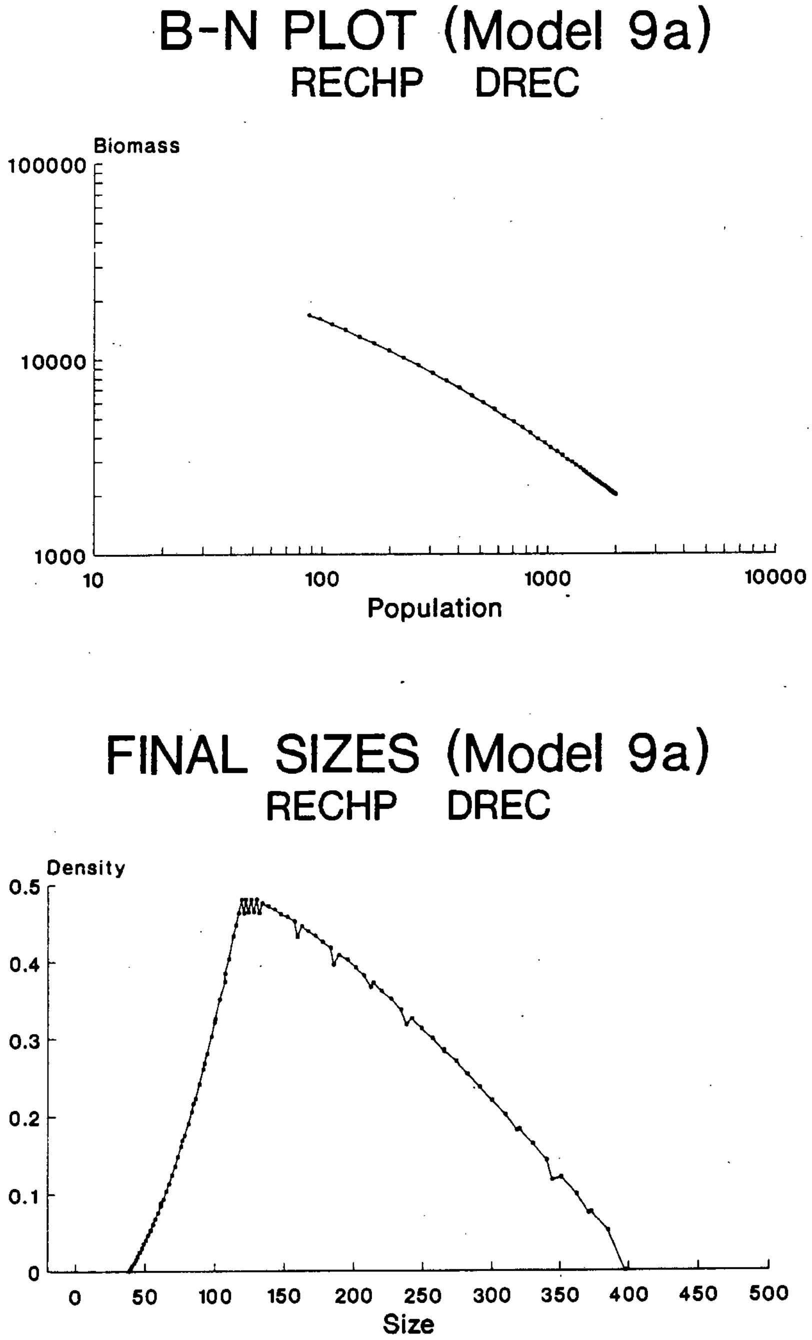


Figure 3.15: B–N plot and size distribution at final time for model 9a. Initial slope is −0.89 and final slope is −0.46.

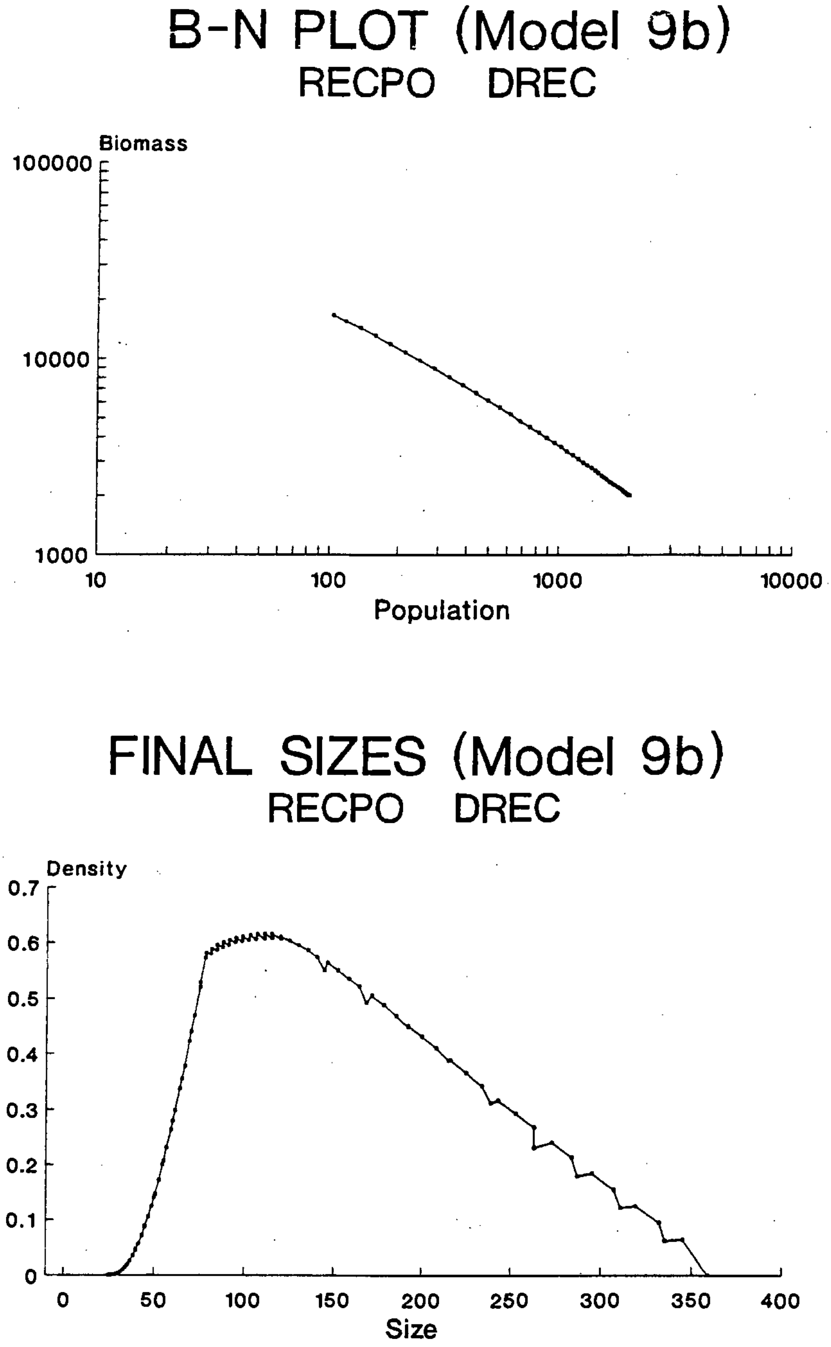


Figure 3.16: B–N plot and size distribution at final time for model 9b. Initial slope is −0.89 and final slope is −0.57.

collinear. Thus these models differ not only in terms of biological interpretation but also in the results they give when statistical procedures are applied. However, it is not to be expected that these differences will be large enough to lead one to prefer one formula, and hence one biological interpretation, above the rest.

The fact that biologically diverse models are mathematically identical means of course that the minimum mathematics implicit in a B–N plot is too crude distinguish between important biological ideas. This implies the need for caution in interpreting B–N plots: they do not contain enough information to support claims about density dependence or size dependence in mortality or in somatic growth. Given this indeterminacy, it is not surprising that one finds claims of both phenomena in the literature: reference is made to density-dependent mortality in Aikman and Watkinson [3], Antonovics and Levin [5] and to size-dependent mortality in Hara [76] and Peet and Christensen [122]. Turning now to the structured models of section 3.3, it can be seen that this indeterminacy persists to some extent—for example, consider models 1, 2, 3 and 4, which are qualititatively similar both with respect to the B–N plot and final size structure. On the other hand, these models have in common the biological property that neither mortality nor somatic growth depend on relative size. This offers the possibility that classes of biologically similar models may be found, where the final size distribution is similar within classes and varies between classes, so that the size distribution may serve to discriminate between different types of model.

In fact, one can discern three fairly homogeneous groups with respect to final size distributions that were obtained:

(i) Models 1, 2, 3 and 4 have a sharply defined peak on the left, tailing off fairly slowly to the right. In biological terms, one observes mostly small plants, but large individuals are by no means rare.

(ii) Models 5, 6 and 7 have a very narrow peak at the extreme left, with a long tail of near zero densities. In biological terms, one observes an abundance of small plants from a rather narrow range, and a very few plants that are large and occasionally very large.

(iii) Models 8 and 9 have a rather flat peak somewhere near the middle of the size range, with a nearly linear drop in density with size to the other side. In biological terms, plants are abundant across a relatively large part of their size range.

The interesting thing about the corresponding three groups of models is that they are readily interpreted in biological terms. Models in group (i) are characterised by the absence of relative size dependence. The other two groups have relative size dependence, but are distinguished by whether this is in mortality or in somatic growth: in group (ii), the somatic growth depends on relative size irrespective of the type of size dependence in mortality, while in group (iii), the mortality depends on relative size and the somatic growth does not. This is the pattern predicted under self-thinning in papers such as Mohler *et al.* [118], Peet and Christensen [122], and Weiner and Thomas [154].

The models used in this section have much in common with the model discussed in Bonan [13], except for the absence of spatial structure. There, an ordinary differential equation is developed for an individual plant, and then a population of individuals is followed; the resulting model is a system of ordinary differential equations, which is the same form as here (after discretisation). Furthermore, in both cases a range of initial sizes and growth rates is considered. Bonan found that for a size hierarchy to develop, relative growth rate must correlate positively with mass. It is not immediately clear that this is also true here, since here $(\dot{s}/s)$ decreases with time, even though s increases, so that positive correlation is not to be expected. Furthermore, Bonan claims that increased mortality on smaller plants—in other words, a mortality model of type λ_r—leads to "equality in size". However, I find that this holds only if the somatic growth model is not also of type $(\dot{s}/s)_r$. In other words, on Bonan's hypothesis, one would expect model 7 to give results in group (iii), which it did not.

A comment on causal inference from pattern is in order here. In ecology it is generally true that process cannot always be deduced from pattern, in that one is usually able to specify several possible causal mechanisms all of which could result in the observed pattern. However, (*pace* Peters [124, p. 267]), there are exceptions. If one can show that, among a class of causal mechanisms under consideration, only one can give the observed pattern, then that suffices to select a unique mechanism from that class. Of course, uniqueness may be lost if the class were to be enlarged. On the other hand, it is logically impossible to be sure that one has enumerated all possible causal models: the preferred causes are always selected from a relatively small set of possibilities. That is precisely the case here: only nine classes of model are considered, and on the basis of results they are classified in only three groups. If a pattern is unique to a group, then observing it in nature means that of these nine models, only the ones in that group should be considered. Within this class of models, one may infer cause from pattern. Whether such inference is valid for a class of models sufficiently large to be of practical and theoretical importance remains to be seen.

The patterns observed above can be recast in terms of asymmetric competition, which is defined as effects disproportional to the relative sizes of two interacting plants. (Weiner [152]). Usually, the plants are neighbours and the effect of the larger plant on the smaller is disproportionately large. Here, there is no spatial information, and so the asymmetry is in an average sense: smaller plants are disproportionately affected; it is modelled by $f(s)$ being other than constant. In group (i), competition, where it occurs, is symmetric. In group (ii), competition is asymmetric in somatic growth, and may be symmetric or asymmetric in mortality. In group (iii), competition is asymmetric in mortality and symmetric in somatic growth.

One can also interpret some of the results reported from earlier studies in this context. Weiner [152] reports that the skew in size distribution is most pronounced when both roots and leaves are crowded, and this is consistent with the idea that in that case asymmetric competition is most intense in size growth, corresponding to group (ii) above. Hara [73] suggests that skew in the size distribution indicates crowding—however, it should be clear that skew can develop (as in model 1 above) solely on the basis of a range of initrinsic growth rates. Some studies (e.g. Westoby and Howell [162]—see Figure 3.17, Knox *et al.* [93]) did not find much change in the size distribution, which can be seen as either as group (iii) above or as due to a very narrow range of intrinsic growth rates plus symmetric competition. Finally, the pattern in Mohler *et al.* [118], where the distribution flattens out after being very strongly skewed, is quite consistent with initial development dominated by a range of growth rates, and then the shape of the size distribution determined by a model from group (iii).

Bimodality appears here for several models (5, 6 and 7), and this is

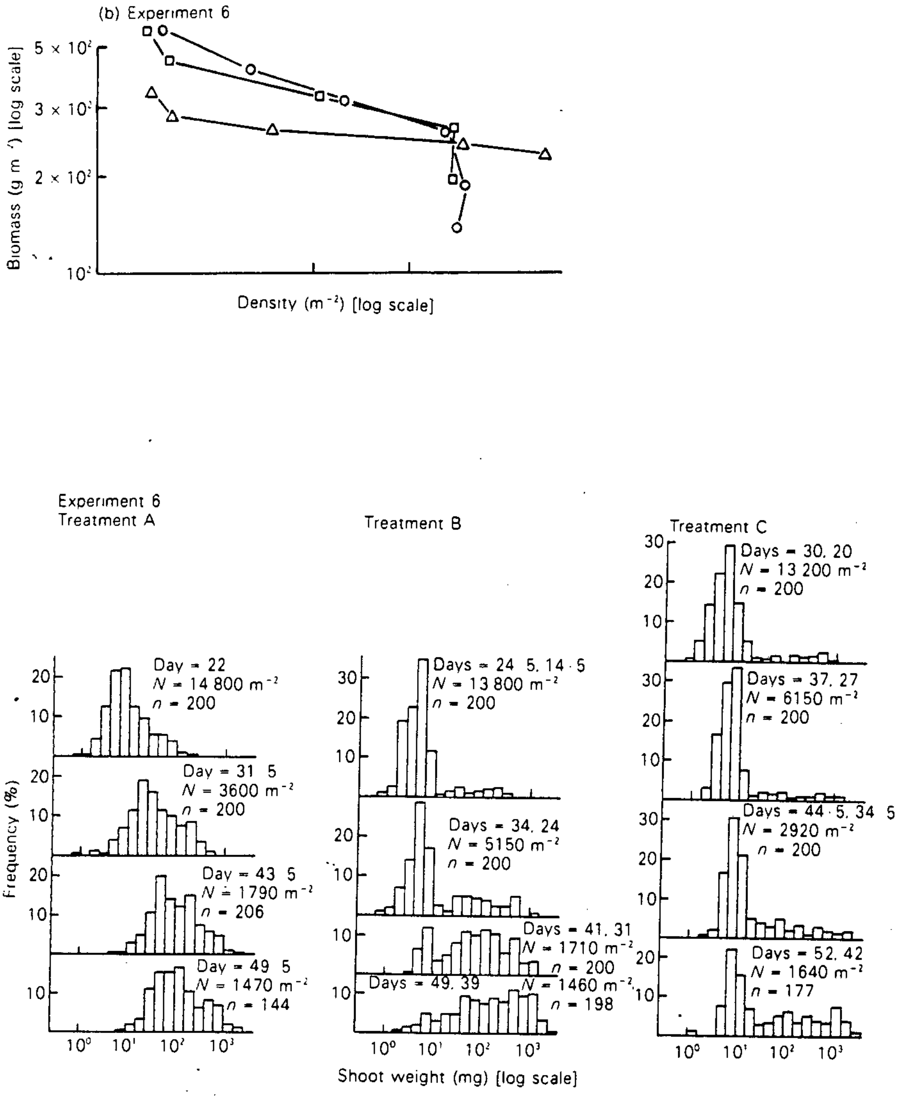


Figure 3.17: A sample of B–N plots plus size distribution data. Note the logarithmic scale on the size axis, which means these diagrams are not directly comparable to the size distributions displayed elsewhere in this thesis. The species is *Raphanus sativus;* reproduced from Westoby and Howell [162].

consistent with the conclusion of Huston and DeAngelis [83]. However, all these cases are from group (ii), which raises the intriguing possibility that bimodality is the result of asymmetric competition in somatic growth. This would be in contrast to the hyptheses in Westoby [160], where bimodality is also seen as determined by the somatic growth model, but where there is no relative size dependence.

In summary then, it appears possible that the shape of the size distribution conveys information on the kind of size dependence in mortality and size growth, at least for this class of models. There is some heuristic motivation that the pattern should be as observed here: for example, somatic growth dependent on relative size should give extreme divergence in size, and when the smaller plants do not die off very rapidly, this will yield a very sharp peak far to the left—as in group (ii) above. Similar heuristic arguments justify the expectation of the identification of group (i) with models where relative size dependence is absent, and of group (iii) with models where mortality depends on relative size and somatic growth does not. However, advancing this as an hypothesis is premature, and it would be even more premature to suggest a generalisation of such an hypothesis. There are several reasons.

(**1**) The parameter estimation here is crude, comprising nothing more than trial and error adjustments, no sensitivity analysis has been conducted all, nor any attempt made to map out the subset of parameter space that gives acceptable results.

(**2**) The B–N plots do not reflect fits to the same data, but have only qualitative features in common.

(**3**) If a successful fit for a model from group (i) were achieved with a greatly increase range of intrinsic growth rate, the size distribution would be far more skewed, making differences between groups (i) and (ii) much less conspicuous. My experience from the computations attempted is that it would be hard to maintain the convexity of the B–N plot in such a case, but even so a satisfactory fit might be achieved.

(4) I used a range of models for relative size dependence, and I have not looked at population density dependence (*cf*. Firbank and

Watkinson [62] and Silander and Pacala [136], where density but not size dependence is considered). That makes this family less coherent or complete than one would like for model selection.

Until the patterns observed here are repeated for best-fit parameter estimates for sets of independent data and using several families of models, they remain merely suggestive.

3.4.1 The role of growth rate

One point which does stand out is that, where the size differential between propagule and adult plant is very large, the major determinant of success is the intrinsic growth rate. This is illustrated by the fact that in model 9, the effect of dependence on relative size in somatic growth overshadows its effect on mortality. One can put this another way by looking at the change of the part of the (r, s) plane occupied at the start and end of the computation—see Fig 3.18, which shows initial and final mesh for the models that exhibited the highest and the lowest deformation. It is clear that initial size differences are far less important than intrinsic growth rate differences.

One may therefore say that, if size determines reproductive success (rather than timing of reproductive effort, or pattern of resource allocation to roots, stems, leaves and reproduction, or pollination rate, or some other factor such as chemical defence), then natural selection would lead to a very narrow range of genotypes with respect to growth rate. In fact, one would expect that from the genetic point of view, intrinsic growth rate would be constant. Furthermore, if intrinsic growth rate showed genetic variation, one would expect that some factor other than individual plant size played the decisive role in reproductive success, for example seed dispersal, or that the size difference between adult and propagule was not very large (which to my knowledge does not occur in plants that bear seeds). Therefore, in the absence of evidence to the contrary, differences in observed growth rate should be interpreted as due to environmental rather than genetic factors.

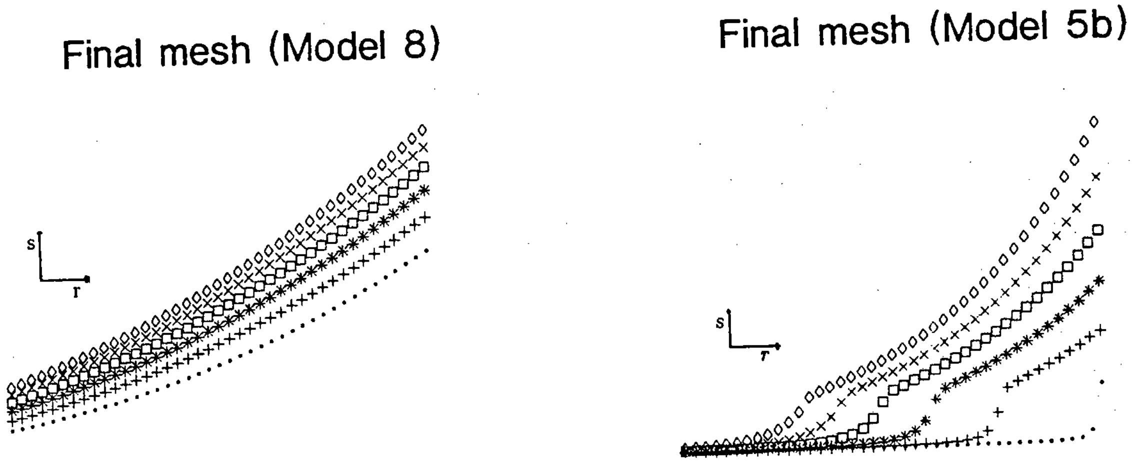


Figure 3.18: Mesh deformation for (a) Model 8 and (b) Model 5b. The rate axis is the same in both cases; for the size axis see Figures 3.14 and 3.9 respectively.

3.4.2 The interconnectedness of mortality and somatic growth

A major theme of this chapter is that a proper understanding of plant population dynamics is only possible if equal emphasis is placed on both mortality and somatic growth. This is easier in modelling than in experimental work, and most research in the former regard has been done on discrete models, using transition matrices and size classes (see for example Caswell [23], Solbrig *et al.* [140], Klinkhamer *et al.* [91]). A discrete analogue of the approach used here is Bishir and Namkoong [12] (because of its use of non-linearities), and serves to illustrate another advantage of a continuum model: it may well be possible to construct a detailed model with fewer parameters. Bishir and Namkoong, for instance, present an example in which size structure is limited to two size classes, and which nevertheless requires 21 parameters to be estimated. Bigger models such as used by Groendael and Slim [67] require of the order of a 100 parameters. By contrast, the model in this chapter with the most parameters uses 10, and if the parameters held constant were included, there would be 13 parameters.

3.5 Future possibilities

The possibility that there is a way to infer types of density dependence from size distribution is very attractive, and here I would like to suggest how one may proceed to investigate this further.

First, computational experiments similar to the one reported here should make it clear whether the patterns reported above are robust across all choices for parameters within plausible ranges. If so, it would be interesting to see how soon the pattern develops and whether it deteriorates at large times. Such work would also indicate if the patterns were very sensitive to changes in particular parameters. An experiment that would test many of these features indirectly could be constructed as follows: assemble a model in which all three types of size dependence are present both in mortality and in somatic growth, in such a way that the relative contribution of each type of size dependence could be varied by changing a few weights. Then fit this model to data (which for the sake of preliminary investigation could be computer generated, and thus not require the time and expense of field and laboratory), including the size distribution. If the pattern above has predictive value, this should be revealed by systematic variation in the relative contribution of each type of size dependence as evidenced by its weight.

Secondly, it would be very useful to fit these models to observational data such as those reported by Carleton and Wannamaker [22] and Hara [75] (where size data appear to be available), to see if the same pattern was observed. It may well be that in the massive data set assembled by Weller [156, 157, 158] there are several more cases where both the B–N data and the size distribution are available. The obverse of this approach would be to record the size distributions of monospecific even-aged stands of a species at various densities and site qualities. Based on knowledge of changes in population density and type of competition inferred using the pattern above, it should be possible to predict what changes, if any, in the size distribution would be observed (assuming of course that Hypotheses 1 to 4 held).

Even if these tests showed that the pattern did generalise, its practical value would at first be unknown. The simplest model above required the estimation of five parameters, and most required more, so that successful parameter estimation or characterisation of size distribution may be impossible for lack of data, and the pattern could not be tested or used. Moreover, the type of questions addressed in Kohyama [95], concerning several species and including births (i.e. recruitment) would need to be considered. However, where patterns like the above held, they would clearly be useful. For instance, in silviculture they could inform decisions about thinning and species composition, and in plant conservation they could yield predictions of reproductive potential based on size structure (as investigated by Samson and Werk [132] and Klinkhamer *et al.* [92]) well into the future.

Chapter 4

Density dependence in lottery models for plants with disjoint generations, with applications to serotinous proteoid shrubs

4.1 Introduction

Fynbos, the shrubland vegetation at the extreme south-western tip of Africa, has aroused much interest because of its extremely high levels of species diversity (Cowling [41]). The vegetation burns readily, and frequent fires dominate all aspects of its ecology, and are thought to have played a major role in its evolution (Cowling [40, 41], Linder *et al.* [105]). Fynbos often has a canopy of shrubs 2–4 metres tall. The canopy species are usually from the most speciose South African genera in the Proteaceae, and are usually referred to as proteoids. Many of them are killed by fire and store seeds in cones on the plant, i.e. they are serotinous (Lamont *et al.* [98]). I shall refer to this guild of shrubs as serotinous proteoids. Not only are coexisting taxa closely related, but they often have very similar growth form and life histories (Cowling and Holmes [42]). One might expect to find resource partitioning, but I have found it to be undetectable in at least one species pair (Laurie, unpublished data). As a result, conventional theories of coexistence such as niche separtation have been seen as less attractive than non-equilibirum theories such as lottery models (Bond *et al.* [16]). In this chapter I explore the issue in detail for one guild (serotinous proteoids as defined below), and in the conclusion consider to what extent similar arguments may hold for the rest of the flora. First I show that coexistence requires that the recruitment success of a species must depend on its relative density, and second I use simulations based on the formulation and algortithms developed in chapter 2 to show that such density dependence is most effective if it acts on seed production. Hence I suggest that one should look for aspects of fynbos biology which contribute to reduced seed production in relatively dense stands, such as increased seed predation.

4.2 Lottery models for plants with disjoint generations

The original lottery models found that coexistence of species was possible only if generations overlapped (Fagerström [58]). However, many plants in fynbos have disjoint generations. For example, many of the canopy shrubs are weakly serotinous in the sense that seeds are retained in cones on the plant for a few years, after which they are released and within a year lose viability (Bond [15]). It is therefore presumed that seeds in the canopy form the only source of viable seed at the time of fire, so that where adults cannot survive fire (which is the case for most fynbos shrubs) the pre- and post-fire generations are disjoint.

In verbal form, lottery models have existed for many years (e.g. Andrewartha and Birch [4], Sale [131]). Fagerström and Ågren [59, 60], who were interested in phenological spread as a mechanism for coexistence, formulated a simple mathematical model for plants. In their model, the environment consists of a set of sites that can be occupied by at most one adult of any species. Adults die at species-specific rates, and seedlings take their place on a random basis, with probabilities determined by the relative frequency of seedlings present for each species as well as by competitive ability. Fluctuations in recruitment success among species are determined entirely by the seedling stage: a species succeeds according to the proportion of its seeds among all the seeds present at a site together with their relative ability to exploit the environmental conditions they find at the time an adult dies. Fagerström and Ågren [59] found that, when generations do not overlap, the superior competitor will always completely displace the other, but that when several generations overlap, the weaker competitor could stay in the system due to a variety of mechanisms based on phenological spread, uneven dispersal, or temporal variation in the environment.

Chesson and Warner [31] produced a very similar model, but with a more general mathematical technique, and with animals rather than plants in mind. They considered only temporal variation in the environment, and only insofar as it determines propagule number and competitiveness. Their results confirmed those of Fagerström and Ågren: disjoint generations in competing species lead to extinction of all but the best competitor, but in the case of overlapping generations, poor competitors can stay in the system if they have episodes of good recruitment. This was termed the storage effect by Chesson [25].

The term "lottery model" was first used by Chesson and Warner [31]. The two models above, and others mentioned below, all use what is called the lottery assumption: namely, that recruitment success for a given species is proportional to its relative frequency in the pool of propagules (Lavorel and Lebreton [99]). Hence, one may define a lottery model as follows: it is a mathematical model of population dynamics for two or more species in which recruitment is based on the lottery assumption.

Several lottery models have since appeared: Shmida and Ellner [135] using the sophisticated mathematical analysis of Ellner [54], showed that lottery models could explain coexistence in the case of annual plants by considering overlap in the seed bank due to dormancy of the seeds. Comins and Noble [38], in order to overcome the lack of coexistence in disjoint generations, considered a system of ecologically differing patches with equally likely dispersal to all patches and coined the phrase "transient niche" to describe the mechanism whereby a species that competed poorly but dispersed extensively could avoid extinction. Thus in their model one may observe coexistence of plants with disjoint generations, but only because some patches favoured the one species and other patches the other one. Ågren and Fagerström [1] showed that the spatial variation in seed density due to random variation in dispersal was sufficient to allow coexistence of two plant species, even where the one was a vastly superior competitor, provided that their mean lifetime seed output per plant was nearly the same. Warner and Chesson [149] focussed on testable predictions for field experiments; Hatfield and Chesson [81] included dispersal of a general kind in a lottery model; Chesson [26, 24] and Chesson and Huntly [30] considered general effects of various types of environmental variation on coexistence. Recently, Chesson and collaborators (Chesson and Ellner [29], Chesson [27]) have proposed a general competition model which emphasises competition and environmental variation.

It should be noted that all these models are linear in the sense that population processes such as births and deaths have no density dependence. Below, I show how lottery models should be generalised so as to predict coexistence for species with disjoint generations. The central theme that emerges is that for coexistence to occur in a lottery model with disjoint generations, there must be relative density dependence in reproductuve output. That is, the *per capita* propagule production must depend on the relative abundance of the two species.

4.3 A density-dependent lottery model

Following Chesson and Warner [31], I write the density-independent lottery model for two species with disjoint generations as follows:

$$\frac{P_1(t+1)}{P_2(t+1)} = \frac{\beta_1^*(t)}{\beta_2^*(t)} \frac{P_1(t)}{P_2(t)}. \tag{4.1}$$

Here, $P_i(t)$ is the total population (it could also be the population density) of the i-th species at time t, and $\beta_i^*(t)$ is the effective per capita output of propagules of the i-th species at time t. In $\beta_i^*(t)$, the asterisk indicates that the actual number of propagules is weighted by a constant factor to model competitive differences at the seedling stage. In words, the model says that the ratio between the population sizes of the species at time $t+1$ is the ratio at time t multiplied by the ratio between their effective *per capita* propagule outputs.

For purpose of this and subsequent analysis, I rewrite equation 4.1 as follows:

$$u(t+1) = u(t) + w(t) \tag{4.2}$$

where $u(t) = \log \frac{P_1(t)}{P_2(t)}$ and $w(t) = \log \frac{\beta_1^*(t)}{\beta_2^*(t)}$.

One can interpret equation 4.2 as a time series model, where $w(t)$ plays the role of a random variable, and therefore $u(t)$ is also considered to be a random variable. Now, coexistence clearly implies that neither $P_1(t)$ nor $P_2(t)$ goes to zero as t goes to infinity. In terms of $u(t)$, the equivalent requirement is that all observations of $|u(t)|$ remains bounded. Since $u(t)$ is a random variable, it is more appropriate to think of this bound as meaning that the probability of any observation exceeding the bound is negligibly small. When considering two species in the field, a specific meaning for the bound above is available: for example in an environment with space for 10^6 plants, negligible probability of $|u(t)| > \log(10^6)$ would be required.

Using the concept of an autoregressive time series with lag one [18], denoted by AR(1), I now show that $|u(t)|$ cannot be bounded, in the sense that arbitrarily large values of $u(t)$ have non-zero probabilities with sufficiently large values of t. Indeed, if $w(t)$ has non-zero expectation, then the expectation $\mathrm{E}[u(t)]$ goes to $\pm\infty$, the sign depending on the sign of the sign of $\mathrm{E}[w(t)]$, whereas if $w(t)$ has zero expectation, the variance $\mathrm{Var}[u(t)]$ goes to infinity as time goes to infinity, so that the probability of $|u(t)|$ exceeding any finite bound approaches certainty as time goes to infinity. In either case, since $u(t)$ is not bounded for all t, coexistence under the model given by equation 4.2, and hence equation 4.1, is impossible. These arguments are often summarised by saying that the AR(1) process described above is a random walk.

The above merely reformulates, in the language of time series analysis, the well-known result that coexistence is not possible in density-independent lottery models with disjoint generations (Fagerström [58]). I now show that, if seed production depends in a suitable way on the densities of the competing species, coexistence is possible. To this end, modify equation 4.1 by writing:

$$\frac{P_1(t+1)}{P_2(t+1)} = \frac{\beta_1^*(t,u(t))}{\beta_2^*(t,u(t))}\frac{P_1(t)}{P_2(t)}, \tag{4.3}$$

where $u(t)$ is as defined before.

The difference is now that $\beta_1^*(t, u(t))$, the effective *per capita* propagule output, depends not only on time (i.e. environmental conditions) but also on the ratio of densities of the competing species via $u(t)$. Again, for purpose of analysis, it is convenient to rewrite the model as a time series:

$$u(t+1) = u(t) + w(t, u(t)), \tag{4.4}$$

where $u(t)$ is as before, and $w(t, u(t)) = \log \frac{\beta_1^*(t,u(t))}{\beta_1^*(t,u(t))}$. In this form, the model is awkward to analyse, since the random variable $w(t, u(t))$ does not have a stationary distribution. I thus simplify somewhat, by assuming that one can write $w(t, u(t)) = \bar{w}(t, u(t)) + e(t)$, the sum of the random variable $e(t)$ and the mean of the function w. Assuming that $e(t)$ has bounded variance then gives the following model equation for density-dependent lottery with non-overlapping generations:

$$\text{(DDL)} \qquad u(t+1) = g(u(t)) + e(t),$$

which I shall use henceforth.

As before, $u(t)$ is a random variable and $e(t)$ is a noise process. However, from now on I merely assume $\mathrm{Var}[e(t)] \leq \sigma^2 < \infty$, so that $e(t)$ need not be stationary. In (DDL), $g(u(t)) = u(t) + \bar{w}(t, u(t))$; it turns out that when g satisfies conditions 1 and 2 below, coexistence of the two species is assured.

Condition 1 There exists $M \in \mathbf{R}$, with $0 \leq M < \infty$, such that

$$|g(x)| < \begin{cases} M & |x| < M \\ |x| & |x| \geq M. \end{cases}$$

Condition 2 $\lim_{|x| \to \infty} \left| \frac{g(x)}{x} \right| < 1.$

Note that bounded mean and variance of $u(t)$ imply that large values of $|u(t)|$ have negligible probability. That means that $P_i(t) \neq 0$ with probability 1, i.e. that neither population goes extinct.

Biologically, condition 1 means that the species which is relatively more abundant recruits fewer propagules *per capita* into the next generation (when its predominance is so large that $|u| > M$). That is, the more abundant species is relatively less successful at recruitment. Condition 2 means that this disadvantage does not tend to disappear as the disparity in densities becomes very large. Together, they imply that coexistence is crucially dependent on what happens at very low densities (which is when $|u(t)|$ becomes large), and is not reallly sensitive to behaviour at medium densities. Similar emphasis can be found explicitly in Ellner [54], Chesson [26]. and Chesson and Ellner [29].

Theorem 6 *If $u(t)$ is defined by equation* DDL, *$u(0)$ is of finite mean and variance, and g satisfies conditions 1 and 2, then the non-stationary time series $u(t)$ has finite mean and variance.*

Proof: Below, I use the following definitions of the expectation E and the variance Var of a random variable x with p.d.f $f(x)$:

$$\mathrm{E}[x] = \int_{-\infty}^{\infty} x f(x) dx$$

$$\mathrm{Var}[x] = \int_{-\infty}^{\infty} x^2 f(x) dx = \mathrm{E}[x^2] - (\mathrm{E}[x])^2.$$

Let $0 < \phi < 1$ be a common bound on $|g(x)/x|$ and $\lim_{|x| \to \infty} |g(x)/x|$.

(a) First I prove that the mean is bounded. Note that for any random variable x, $|\mathrm{E}[x]| \leq \mathrm{E}[|x|]$; I prove that $\mathrm{E}[|u(t)|]$ is bounded.

From condition 1, $|g(x)| \leq M + \phi|x| = \phi(M' + |x|)$, where $M' = M/\phi$. Moreover, since the variances of the $e(t)$ are all bounded by σ^2, there exists a bound η on $\mathrm{E}[|e(t)|]$. For example, $\eta = \sigma^2 + 1/4$ will do, since for any random variable x with p.d.f. $f(x)$, one has $\mathrm{E}[|x|] \leq \int_{-\infty}^{\infty} x^2 f(x) dx + \int_{-1}^{1} (|x| - x^2) f(x) dx$.

Then a bound on $\mathrm{E}[|u(t)|]$ can be calculated as follows:

$$\begin{aligned}
\mathrm{E}[|u(t)|] &= \mathrm{E}[|g(u(t-1)) + e(t-1)|] \\
&\leq \mathrm{E}[|g(u(t-1))|] + \mathrm{E}[|e(t-1)|] \\
&\leq \mathrm{E}[|\phi(M' + |u(t-1)|)|] + \mathrm{E}[|e(t-1)|]
\end{aligned}$$

$$
\begin{aligned}
&\leq \phi(M' + \mathrm{E}\big[|u(t-1)|\big]) + \eta \\
&\leq \phi(M' + \phi(M' + \mathrm{E}\big[|u(t-2)|\big]) + \eta) + \eta \\
&= \phi^2\mathrm{E}\big[|u(t-2)|\big]) + \phi^2 M' + \phi M' + \phi\eta + \eta \\
&\vdots \\
&\leq \phi^t\mathrm{E}\big[|u(0)|\big]) + (\phi M' + \eta)\frac{1-\phi^t}{1-\phi} \\
&\leq \mathrm{E}\big[|u(0)|\big]) + (M + \eta)/(1-\phi).
\end{aligned}
$$

Hence there exists a $B \geq 0$ such that $\mathrm{E}\big[|u(t)|\big] \leq B$ for all t, so that the mean is bounded independently of time.

(b) The bound on the variance is arrived at similarly.

$$
\begin{aligned}
\mathrm{Var}\big[u(t)\big] &= \mathrm{Var}\big[g(u(t-1)) + e(t-1)\big] \\
&= \mathrm{E}\big[g^2(u(t-1))\big] - (\mathrm{E}\big[|g(u(t-1))|\big])^2 + \sigma^2 \\
&\leq \mathrm{E}\big[g^2(u(t-1))\big] + \sigma^2 \\
&\leq \mathrm{E}\big[(M + \phi|u(t-1)|)^2\big] + \sigma^2 \\
&= \mathrm{E}\big[(M^2 + 2M\phi|u(t-1)| + \phi^2|u(t-1)|^2\big] + \sigma^2 \\
&= \phi^2\mathrm{E}\big[u^2(t-1)\big] + M^2 + 2M\phi\mathrm{E}\big[|u(t-1)|\big] + \sigma^2 \\
&\leq \phi^2(\mathrm{Var}\big[u(t-1)\big] + (\mathrm{E}\big[u(t-1)\big])^2) + M^2 + 2M\phi\mathrm{E}\big[|u(t-1)|\big] \\
&\leq \phi^2\mathrm{Var}\big[u(t-1)\big] + \phi^2B^2 + M^2 + 2M\phi B + \sigma^2 \\
&\leq \phi^4\mathrm{Var}\big[u(t-2)\big] + \phi^2(\phi^2B^2 + M^2 + 2M\phi B + \sigma^2) \\
&\qquad + \phi^2B^2 + M^2 + 2M\phi B + \sigma^2 \\
&\vdots \\
&\leq \phi^{2t}\mathrm{Var}\big[u(0)\big] + (\phi^2B^2 + M^2 + 2M\phi B + \sigma^2)\frac{1-\phi^{2t}}{1-\phi^2} \\
&\leq \mathrm{Var}\big[u(0)\big] + (\phi^2B^2 + M^2 + 2M\phi B + \sigma^2)/(1-\phi^2)
\end{aligned}
$$

□

This shows that conditions 1 and 2 are sufficient. Neither are necessary (see below), but there is a biologically meaningful property which is necessary for g to satisfy, if coexistence is to be assured: in some sense g must shrink its argument in absolute value for large arguments. For any other g, there would exist an $N \in \mathbf{R}$ such that for $|x| > N \Rightarrow |g(x)| \geq |x|$. To guarantee coexistence, one should have bounded mean and variance of $u(t)$ for any $u(0)$ and $e(t)$ with finite variance. This does not happen, and one example suffices: for g and N as above, consider an initial condition $u(0) > N$ with probability 1, but finite variance (the p.d.f. $f(x) = e^{N-x}$ for $x > N$ and zero elsewhere will do), and similar $e(t)$. Then for all time, $u(t) > N$ with probability 1, whence $\mathrm{Var}\big[g(u(t))\big] \geq \mathrm{Var}\big[u(t)\big]$, whence $\mathrm{Var}\big[u(t+1)\big] \geq \mathrm{Var}\big[u(t)\big] + \mathrm{Var}\big[e(t)\big]$, whence the variance grows out of bounds.

However, neither condition is strictly necessary. Consider of the function

$$g(x) = \begin{cases} x + 1/(n+1) & x \in [n, n+1/(n+1)] \\ x - 1/(n+1) & x \in [-n-1/(n+1), -n] \\ 0 & \text{elsewhere,} \end{cases}$$

where $n \in \mathrm{Z}$. This violates both conditions, but clearly there is a $\phi < 1$ such that $\mathrm{Var}\big[g(u(t))\big] \leq \phi^2 \mathrm{Var}\big[u(t)\big]$, where ϕ^2 is strictly less than 1, unless it so happens that $g(u(t)) \neq 0$ with probability 1. Since the set of points x such that $g(x) \neq 0$ is not an invariant set for g, this is not the case for finite t nor approached as t gets large. Hence in this case $u(t)$ is of bounded mean and variance.

Finally, neither condition is sufficient by itself. Consider the model defined by $g(x) = ax + 1/x$, with $0 < a < 1$ and $u(0) = e(t) \sim \mathrm{N}(0, \sigma^2)$. Then g violates condition 1 but satisfies condition 2. It is easy to see, by direct integration, that $\mathrm{E}\big[u(1)\big]$ does not exist. On the other hand, consider the model defined by

$$g(x) = \begin{cases} x - 1/2 & \text{if } x \geq 0 \\ x + 1/2 & \text{else,} \end{cases}$$

with $u(0) = e(t) \sim \mathrm{N}(0, \sigma^2)$, which violates condition 2 but satisfies condition 1. Then $\mathrm{Var}\big[u(t+1)\big] = \mathrm{Var}\big[g(u(t))\big] + \sigma^2 = \mathrm{Var}\big[u(t)\big] + \sigma^2$, whence as time goes to infinity the variance grows without bound.

On the basis of these examples, one can say that a necessary condition is that g must shrink large arguments "on average" in some sense of average. Conditions 1 and 2 ensure that this is the case. They are somewhat too strong, but not excessively so, and give a good indication of the requirement that g must shrink large arguments, which is used in the body of the paper. In fact, if one lays down the biologically reasonable requirement that g must be monotonic, then condition 1 can be proved to be strictly necessary.

Some comment on the generality of the theorem is in order. Although it was developed to deal with lottery models for organisms with disjoint generations, it applies to any discrete model for the coexistence of two species, as follows. Define f_i by

$$f_i(\cdot;t) = P_i(t+1)/P_i(t).$$

Then equation 4.4 becomes

$$u(t+1) = u(t) + \log(\frac{f_1(u,t)}{f_2(u,t)}),$$

(cf. Chesson [27, equation 3]) so that

$$\bar{w}(t) = \mathrm{E}\big[\log(\frac{f_1(u,t)}{f_2(u,t)}).\big]$$

Of course, in practice one usually fails to express $\bar{w}$ analytically, so the theorem is not particularly widely applicable.

I now present a series of examples of increasing complexity.

Example 1 The simplest example satisfying conditions 1 and 2 is the trivial case of $g(x) = 0$. In that case, the density from year to year is completely independent of previous densities, and fluctuates in a random manner modeled by the random variable $e(t)$. Clearly, if $\mathrm{Var}[e]$ is bounded, then so is $\mathrm{Var}[u(t)]$, so that extinction is of negligible probability. This model has the biologically unsatisfactory meaning that there is a density effect exactly opposite to the current population ratio. This "unrealism" is because the lottery assumption was built into the model, so that to achieve $g \equiv 0$ it must be subtracted out.

Example 2 Next, consider the case $g(x) = ax$. If $a = 1$, then we simply have equation 4.2 again, but if $|a| < 1$ then the conditions for coexistence are met. This means (ignoring the noise) that whenever one species outnumbers the other, the less abundant species will recruit better. Thus in the long run, they will be equally abundant. In terms of the population levels themselves, this model says

$$\frac{P_1(t+1)}{P_2(t+1)} = \left(\frac{P_1(t)}{P_2(t)}\right)^a * \text{noise}.$$

If a is small, one would observe nearly random fluctuation in $u(t)$, whereas if a was near 1, one would observe a trajectory similar to that of a random walk.

Example 3 In the case where a rare and an abundant species are to coexist, one needs $E[u(t)] = L \neq 0$. Change variables to $u'(t)$ as follows: $u'(t) = u(t) - L$. Then choose $u'(t)$ so that it satisfies the conditions 1 and 2. Hence $u'(t)$ remains bounded in mean and variance, and since $u(t)$ is just $u'(t)$ plus a constant, it too remains similarly bounded, as required.

As one considers examples 1, 2 and 3 in turn, they are seen to require the estimation of more and more parameters, with the attendant cost of gathering more and better data. Example 3 is the least parsimonious, and requires estimates of L, a and σ^2. The other two are in effect special cases of example 3: both have $L = 0$ and example 1 also has $a = 0$.

4.4 Models of density-dependent recruitment for serotinous proteoids

It is clear from the theory of lottery models for discrete generations given above that coexistence has to be mediated by density-dependent recruitment—to be exact, that in order to avoid extinction some apparent density-dependence in the data is necessary. However, the theory sheds no light on possible mechanisms of this apparent dependence. Knowledge of such mechanisms would enable various important predictions; for example, whether an indigenous species can coexist with an invasive alien species, whether a rare species in a reserve is in danger of local extinction, whether weeds will proliferate in various agricultural practices and so on. While it is true for short-term predictions that mechanistic explanations are not necessary and that in fact purely empirical relationships derived from data may be superior (as discussed in chapter 1), this is by no means the case for the kind of long-term predictions mentioned above.

My purpose in the rest of this chapter is to show how the structured population models of chapter 2 can be used to investigate various hypotheses concerning the coexistence of fynbos species. In particular, I look at the question of whether one can argue that a particular element in the population dynamics such as seed production is more important in mediating coexistence than other elements. Unfortunately, I can go no further than showing that such an investigation is possible, and how to go about it, because of the dearth of data. The results I give are for imaginary fynbos-like systems, rather than for actually existing fynbos. For a meaningful investigation of the relative importance of density-dependence in somatic growth, mortality and seed set in the coexistence of proteoids, one would need sufficient data for useful approximations to the constitutive relations. In the absence of such data, I supply *ad hoc* forms below. In the final section of this chapter I discuss various possible ways of getting data that would enable more realistic calculations than the ones I report here.

A note on terminology. The phrase "depending on relative density" is frequently used below; it is well to emphasize that it refers to a species the members of which are sensitive to the ratio between the two co-occurring species. Of course, the more usual type of density dependence is where individuals are sensitive either to an overall or local density composed of all individuals or to such a density composed of conspecifics only. This type is not precluded; however it may or may not lead to relative density dependence. If both species are sensitive only to densities composed of all species, then they cannot be sensitive to any ratio of species densities, since each individual contributes equally to the total irrespective of species. On the other hand, if both species are sensitive only to the density of conspecifics then one does have dependence on relative density as follows: suppose the total number of individuals is constant, then if one species is rare the other is abundant, so that the two species will be differently affected by the densities. However, constructing relative density dependence in this way is not entirely satisfactory, since all effects are intraspecific, and there is no interaction between the species. Because of its convenenient simplicity, relative density dependence is constructed simply by interspecific effects in the simulations below, but I do mention an example with both inter- and intraspecific effects.

4.4.1 Components of recruitment and model variables

Earlier, the general case of the coexistence of two species was modelled by:

$$u(t+1) = g(u(t)) + e(t),$$

where $u(t)$ is the logarithm of the ratio of the two species at time t, e is a random variable of zero mean and finite variance, and g is a function that shrinks the argument.

In the case of serotinous proteoids, the function g can be regarded as a composition of functions:

$$g(x) = g_3 \circ g_2 \circ g_1(x),$$

where g_1 is the lottery process whereby an initial log-ratio of seeds is converted into two seedling populations, structured by size and somatic growth rate; g_2 is the somatic growth process, which also involves mortality; and g_3 is a birth process, which returns the log-ratio of seeds for two given adult populations structured by size. Thus, one should be aware of constitutive relations on two levels: firstly the functions g_i, which combine to model post-fire seedling establishment, interfire evolution, and pre-fire seed production, and secondly the constitutive relations within g_2 (the interfire model), which give the mortality and somatic growth laws.

Returning to the components of g, it is clear that by taking any two of the constitutive relations g_i to be their simplest possible forms, that for the third relation one could propose many mathematical formulae that would allow g to achieve the required relative density dependence. That is, from a mathematical point of view, density dependence could occur in any of the constitutive relations. However, I will limit myself to only two of the three, namely the interfire model and the seed production model, for the following biological reasons.

Firstly, density dependence in the post-fire lottery is unlikely for these plants, because it occurs when they are very small, and so the interactions between them are at their weakest at the time of the lottery. One might want to argue that density dependence may occur in seed germination, early mortality, and so on. The point is that it may very well be true that each species is globally sensitive for example to the total amount of seeds soaking up water, but the density dependence required is that they should be more sensitive to the presence of conspecific seeds than contraspecific ones. This seems to me so unlikely as to exclude the possibility of relative density dependence in respect of resource use. (I suppose that ideopathic compounds released by germinating seed would result in dependence of relative density, but will not pursue this). Note that in a different system, Shaw and Antonovics [134] did not find that seedling mortality was affected by seed or seedling density.

Secondly, although one should certainly consider seed predation as a possible density-dependent effect subsequent to seed production (Botha and le Maitre [17]), I feel it can adequately be represented in the seed production model itself. The point is simply that I am ignoring density-dependence in the lottery process: after the pre-fire seed set and prior to the mortality process. Otherwise, there would be no need for investigating somatic growth, mortality and seed production, since one could already satisfy the conditions of theorem 6 in the lottery process itself. However, such a model would be unable to reveal any mechanism, and would of necessity be purely descriptive. Besides the fact that it would severely limit discussion of long-term issues, there is at present not enough data on any single population to justify the estimate of such an empirical relation. Unfortunately, it will take hundreds of years to accumulate data series of the length recommended by Crowley [44]. Lumping data from many different populations presents many statistical pitfalls, but may be possible for plants that are sufficiently widespread, for instance some of the more common proteas (Bond *et al.* [14]).

Thus I will examine only the effect of relative density dependence in g_2 and g_3. In case of g_3, there is some data on size dependence in seed production by serotinous proteoids (Esler and Cowling [56]), and in general an agreement that seed production should increase with size, though whether plants should have a minimum size below which seed production fails is controversial (Rees and Crawley [126, 127], Silvertown [137]). I have assumed that proteoids do in fact fail to set seed below a certain size. Whether the proportion of reproductive biomass declines, remains stationary, or increases with size is also unclear (Samson and Werk [132], Klinkhamer *et al.* [92]). I have assumed it to be a constant proportion of the biomass in excess of the minimum size for setting seeds.

For the interfire period, g_2 must take as input a population of propagules and give as output a population of adults. This should be done using the framework of chapter 2, to which may be added the assumption that recruitment is limited to the immediate post-fire period, so that $F \equiv 0$ except at t_{end}, since the post-fire recruitment is modelled as a lottery by means of g_1. Accordingly, I have chosen to represent the population in the interfire period with the symbol $\rho_j(r, s, t)$, representing the density ρ of the individuals of the j-th population with attributes of growth rate r and biomass s at time t. The age is suppressed as a variable, as in chapter 3.

The models investigated below do not, in my opinion, allow explicit solutions, and hence one cannot estimate the constant ϕ used in theorem 1. Instead, I present computational experiments, using a (necessarily) limited range of density ratios. The range of these ratios are fairly large, so it is adequate for the coexistence of plants that are in medium-sized communities, i.e. the total number of adult plants of both species is at most of the order of a million.

I cannot guarantee that simpler models will fail to provide the kind of insight demonstrated below. Indeed, one could dispense with both the variables r and s and still satisfy the hypotheses of Theorem 1 by appropriate density dependence in λ. This would yield the following testable prediction: initially relatively more dense stands should thin faster. I anticipate that this prediction will fail. Firstly because it is very strong: it says that a very sparse population of species A would die faster than a dense population of the same species in the case where the sparse population of species A shared space with an even sparser population of species B. Secondly, density-dependent mortality on the whole is hard to observe in fynbos, although the evidence is largely anecdotal and only one detailed study has been done (Kruger [97] as cited in le Maitre and Midgley [101]).

If one discards the mortality-only model of relative density dependence as too simple, it may seem wise to include only one further variable, which should obviously be size. It is possible to gather data giving at least some indication of size-dependent population processes, and indeed some data are available for the seed production (Esler and Cowling [56]). This would be much more difficult for the other variable, namely somatic growth rate, because to determine rates requires the more onerous process of long-term monitoring, and because of the greater statistical difficulty, since the data take the form of ratios. One can make the same point using the jargon of numerical analysis by saying the estimating growth amounts to approximating a derivative, which is a notoriously ill-conditioned problem. On the other hand, Esler and Phillips' work on Karoo plants [57] intriguingly suggests that once a plant is locked into fast growth, it becomes more susceptible to drought, so that in some cases mortality may be directly related to growth rate. Even in the case of this work one cannot unambiguously conclude that mortality was related to growth rate, since although size effect and differences in root/shoot ratios were not significant, they cannot be ruled out. Therefore this example supports the case of preferring size to somatic growth rate if only one of the two variables in the above framework is used.

Another consideration is age effects, which I do not consider in my examples. The first reason for this is simplicity: it is much easier to look for effect size in the absence of age dependence, since one does not need to worry about interaction between age and density in producing observed effects. The second reason is that in fynbos the populations are even-aged, so that for a given fire interval there is only one final age to consider, which is of course independent of density. If age were included, the function g would vary with interfire interval, so what I need to explain is why this is in fact immaterial as far as coexistence is concerned. There are at least two ways of putting it, both relying on the fact that for any population, all plants of both species have the same age. Following Chesson [32], one invokes the fact that the two populations are indistinguishable by age, and hence their log population ratio would perform a simple random walk, which must lead to extinction. Following Fagerström [58], one would emphasize that the crucial requirement of covariance less than 1 would not be met, since the species' ages are always equal.

It would be interesting to see how much of the variance in seed production would be accounted for by the use of two variables: size and density. If these are all that is really needed, then it becomes important to be able to predict the size distribution of a population, and this is why I choose to include growth rate in the examples below. It has been reported that plant populations may go from a unimodal distribution to a flat one or even a bimodal one (Hara gives a review in [75]). It is certainly true that initially similar populations can evolve in quite different ways, for instance due to site differences. If needed, this can be accomodated by considering r as a site index. Thus, in order to obtain a good prediction of size distribution, r would be a useful way to introduce effects that can change the shape of the size distribution (as was done for a fish population Banks and Fitzpatrick [8]).

4.5 A technique for investigating the effects of density-dependence

4.5.1 Preliminaries

As in chapter 3, and again using the notation of chapter 2, I choose $\mathbf{x} = (r, s)$, where $r, s \in \mathbf{R}^+$, with s denoting size and r denoting intrinsic growth rate. Note that serotinous proteoids reproduce only at the moment they die, so

$$F(\rho(\cdot,\cdot,t))(s) = \begin{cases} 0 & t < t_{end} \\ \beta(\rho(\cdot,\cdot,t))\rho(r,s,t) & \text{else.} \end{cases}$$

As before, G is defined via the mortality modulus λ. Since I wish to compare two species, it is convenient to represent them each as a subpopulation, so $n = 2$ and $\rho(r,s,t) = (\rho^1(r,s,t), \rho^2(r,s,t))^{\mathrm{T}}$. Then $\dot{s}_j$, λ_j and β_j suffice to specify a model up to initial conditions, where $j = 1$ (respectively $j = 2$) denotes model and parameters used for species 1 (respectively species 2).

In the simulations below, I am interested in comparing the effect sizes of density dependence in the three constitutive relations of mortality, somatic growth and seed production. In order to do this very simply, I use a carrying capacity K_j, which may refer either to the total population or the total biomass of a species. The density dependence is then a term of the form $(P_j/K_j)^{a_j}$; the larger the shape parameter a_j is, the weaker is the effect of density dependence at low densities—that is, the closer P_j must be to K_j in order to exert a substantial effect. Systematically varying the shape parameters allows one to detect trends in the strength of the density dependence. Note that species differ only in the value of the shape parameters and initial densities.

In the first example, the models for mortality and size growth are derived from logistic population growth. In the second example the mortality derived from an exponentially declining death rate in size (Hara [76]) and somatic growth is derived from the von Bertalanffy equation. In both cases, the seed production is derived from simple proportionality above a size minimum, which fits the reported data (Esler and Cowling [56]).

The calculations were done using Algorithm A, as discussed in chapter 2 and by means of the programs mentioned there.

4.5.2 Limitation via maximum population in mortality and somatic growth

The model is defined by the following constitutive relations:

$$\begin{aligned}
\lambda_j &= -(1-(\frac{P_j}{K})^a) \\
\dot{s}_j &= rs(1-(\frac{P_j}{K})^b) \\
\beta_j &= \begin{cases} k(s-s_{\min}) & \text{if } s > s_{\min} \\ 0 & \text{else.} \end{cases}
\end{aligned}$$

Here, P_j is the current total population of the j-th population. Note that the models and parameters are identical for both species.

The shape parameters a and b are positive; for the sake of comparison, The shape parameters a and b are positive; for the sake of comparison, $a + b = 0.9$. All other parameters remain constant as follows:

K	$=$	10 000	carrying capacity
k	$=$	5	proportionality constant of seeds per size unit
s_{min}	$=$	2	the minimum size for seed production
r	$\in$	$[0.1, 0.2]$	range of somatic growth rates
s_0	$\in$	$[1.0, 2.0]$	range of initial sizes
t_{end}	$=$	2.5	final time

The initial values of the density ρ are such that $\rho_0(r, s_0) = a\mathbf{e}$, where $\mathbf{e}$ is bivariate normal: $\mathbf{e} \sim (N(.15, .05^2), N(1.5, 0.5^2))$ with zero covariance, and a is a constant such that $\int_{.1}^{.2}\int_1^2 \rho_0 \, ds \, dr = P_i(0)$. The two initial populations add to a constant: $P_1(0) + P_2(0) = 1000$, and the ratio between them was varied so that it ran, in the logarithm, from -16 to 0. That is, $\log(P_1(0)/P_2(0))$ varied from -16 to 0 in order to simulate various initial relative densities. The above defines a family of models, parameterised by the values of a and b. I chose to let $a = 0.1, 0.2, \ldots, 0.9$. The density effect is strongest when the power is nearest to zero.

The results are displayed in Figure 4.1, where it can be seen that the density effect via somatic growth is stronger than via mortality.

There is a basic problem with this example, namely that logistic growth in population implies (for a population as here starting below carrying capacity) that recruitment is taking place all the time, which of course is not the case in fynbos. So in fact one should not read any significance into these results, which are presented here only as an example of a simple case, and to display the parameters.

4.5.3 Limitation via biomass in all three constitutive relations

The model is defined by the following constitutive relations:

$$\lambda_j = \lambda_{\text{max}} e^{-c_\lambda s} \left(\frac{B_j}{K}\right)^{r_\lambda}$$

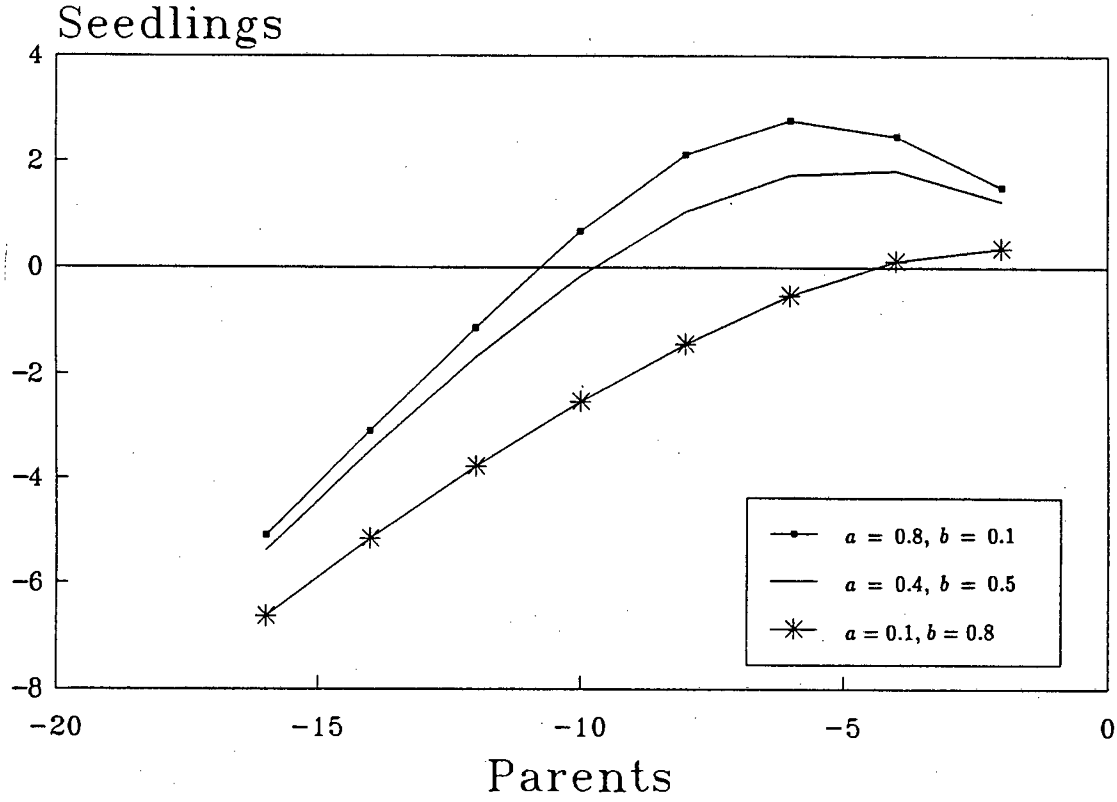


Figure 4.1: The density effect in somatic growth dominates the density effect in mortality for the model of section 4.5.2. The unit on both axes is the logarithm of the ratio between the species' population densities at $t = t_{end}$. Clearly the seedlings of the rare species are more abundant than their parents in all cases, thus demonstrating that the density dependence is effective in assuring recovery from low density. The shape parameters a and b are positive and control the intensity of the density effect in mortality and somatic growth respectively. As a parameter approaches zero, the relevant effect increases, especially at low densities.

$$
\begin{aligned}
\dot{s}_j &= r(s_\infty - s)\left(1-\left(\frac{B_j}{K}\right)^{r_s}\right) \\
\beta_j &= \begin{cases} k(s-s_0)\left(1-\left(\frac{B_i}{K}\right)^{r_\beta}\right) & \text{if } s > s_0 \\ 0 & \text{else.} \end{cases}
\end{aligned}
$$

Here, B_j is the current total biomass of the j-th population, and K is the maximum possible biomass. Again, both species have identical models and parameters.

There are now three density effect size parameters, and again $r_s + r_\lambda + r_\beta = 1$, each of them taking values from $0.1, 0.2, \ldots, 0.8$. The other parameters are as for example 1, together with the following new ones:

c_λ	$=$	0.005	decline rate constant for mortality with size
$\lambda_{\max}$	$=$	0.5	maximum mortality rate (occurs at zero size and at carrying capacity)

The initial values of the density ρ are as before.

With this set of constitutive relations, biomass grows to a constant level, while population monotonically declines. The rate of population decline is a trade-off between somatic growth, which increases individual size and hence reduces mortality, and biomass accumulation, which increases mortality.

The results are depicted in Figures 4.2–4.4, where each panel is a comparison of the relative effect of two parameters by varying them while keeping the third one constant.

The results suggest a ranking according to effect size: relative density dependence in seed production has more effect than in somatic growth where it has more effect than in mortality.

4.5.4 Other possible investigations

(a) In both the above examples, coexistence is really due to each species separately being regulated through its own carrying capacity. Using the

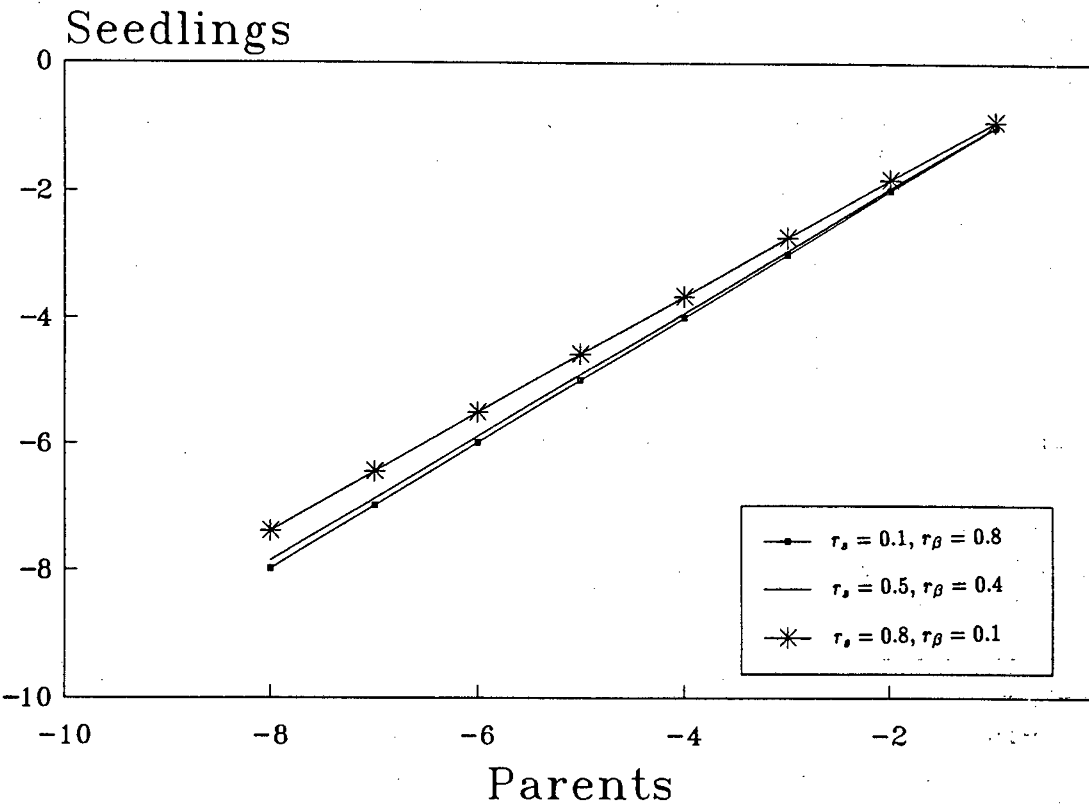


Figure 4.2: The density effect in the model of section 4.5.3 is stronger in seed production than in somatic growth. The parameters r_λ, r_s and r_β control the density effect in mortality, somatic growth, and seed production respectively. Units and parameter behaviour are as for Figure 4.1. In this diagram, $r_\lambda = .1$ is constant.

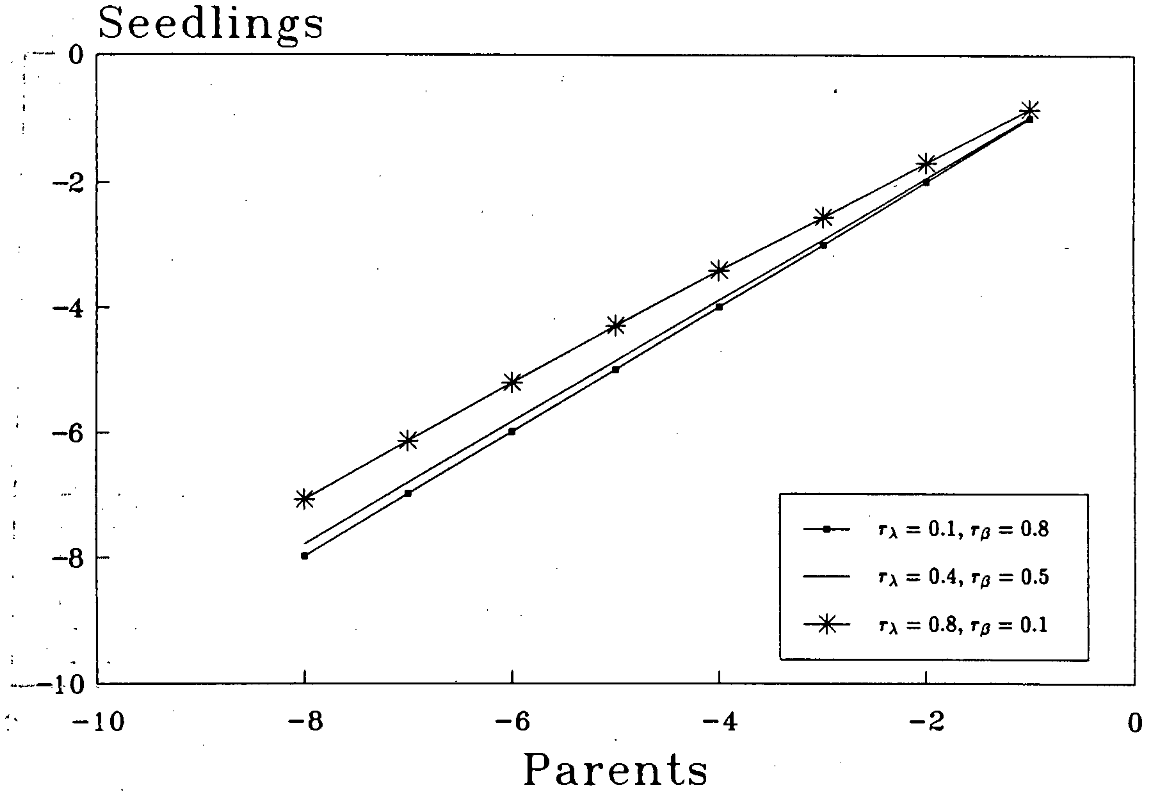


Figure 4.3: The density effect in the model of section 4.5.3 is stronger in seed production than in mortality. The parameters r_λ, r_s and r_β control the density effect in mortality, somatic growth, and seed production respectively. Units and parameter behaviour are as for Figure 4.1. In this diagram, $r_s = .1$ is constant.

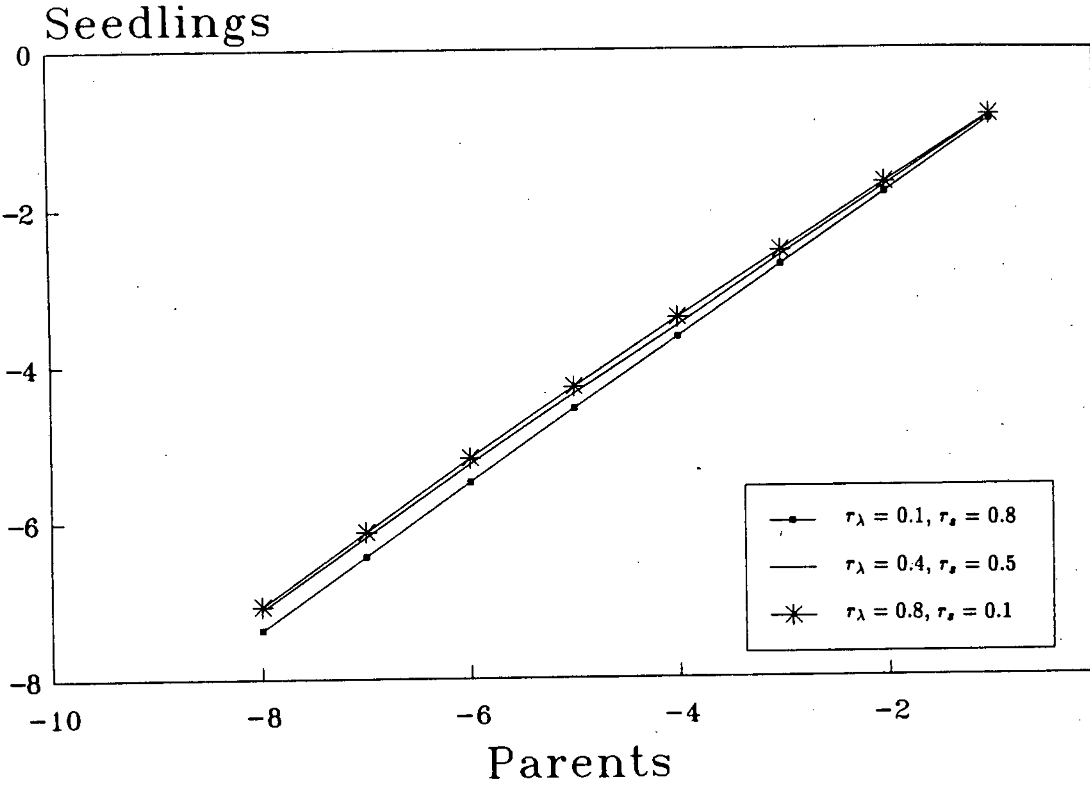


Figure 4.4: The density effect in the model of section 4.5.3 is stronger in somatic growth than in mortality. The parameters r_λ, r_s and r_β control the density effect in mortality, somatic growth, and seed production respectively. Units and parameter behaviour are as for Figure 4.1. In this diagram, $r_\beta = .1$ is constant.

language of competition one might say all competitive effect was intraspecific. One may extend this approach to a comparison of intra- vs. interspecific competition by replacing the factors $(B_j/K)^{r_j}$ for species j by the factors $((a_jB_1 + (1 - a_j))B_2/K)^{r_j}$, where $0 < a_j < 1$. The parameters a_1 and a_2 may be equal or different. The carrying capacity is shared, and the parameter a models the intra- vs. interspecific sensitivity to density. When $a = .5$, these two sensitivities are equal, and it is immaterial what a plant's neighbours are. In that case the model reduces to a random walk to extinction. When $a_j < 0.5$, the plants of that species are more sensitive to interspecific density, so that they perform worse the more of the companion species there is. If this is the case, using the approach above would still lead to two species equivalent in all respects, but $u = 0$ would be an unstable equilibrium, and any species that has sufficient dominance will increase that dominance until the other dies out. When $a_j = 1$ this example reduces to the previous one, and when $a_j > 0.5$, I expect coexistence, similar to the previous example but more mildly the closer a_j is to 0.5. Comparisons similar to the above on the strength of the various kinds of interspecific competition could then be made.

The perception may arise that the kind of species-specific density dependence required by lottery dynamics for disjoint generations is just another form of niche separation. Such a perception would be based on including in the concept of niche separation all the differences between two species that facilitate coexistence. "Niche" would then be a generic term descriptive of certain biotic interactions, and would have no explanatory value of its own not already contained in the account of the biotic interactions themselves. In terms of the distinction made in chapter 1, niche theory would then be seen as descriptive theory. On the other hand, the more narrow conception of niche as a region in resource space (Begon and Mortimer [9]) may go beyond being merely descriptive by assumptions on how physiology, morphology, life history and so forth limit the fundamental niche. However, the realised niche is again a pure descriptive concept, being determined by data. In the absence of a commonly accepted distinction between niche separation and mechanisms of coexistence, the debate about whether the one implies the other is premature (Silvertown and Law [138], Chesson [32]).

(b) The resource-mediated coexistence theories proposed by Tilman [142, 143] are more complex, but have the virtue of not contriving the coexistence of species through attributes as dubious as species-specific carrying capacity. Following Tilman's lead in the framework of this thesis would require explicit spatial modelling of resource levels and population density, as well as bringing dependence on resources into the population dynamics, and the investigation would be quite expensive computationally. It would be interesting to see whether lottery recruitment could change the outcome of any of Tilman's scenarios, for instance whether spatial variation in seed dispersal suffices to stabilise an otherwise unstable equilibrium in a homogeneous two species–two resource system.

(c) In recent work, Chesson and co-workers (Chesson and Huntly [30], Chesson and Ellner [29], Chesson [27]) have investigated how the effect of environmental variation on competition may influence coexistence. Simulating this would entail the use of $E(t)$ to model the state of the environment in the constitutive relations λ, β and $\dot{s}$ I have used above. As indicated in chapter 2, this should be possible. However, it is not clear that this would offer any advantages for plants with disjoint generations, since in this theory coexistence appears to rely on overlapping generations to take advantage of the competitive release occasionally favoured by environmental conditions.

4.5.5 Discussion

One problem with the above results is that they amount to an exploration of a miniscule part of parameter space. A substantial, but not huge, computation is required to cover more meaningful parameter ranges. One could then discover whether the shape of the rate distribution influences coexistence, and whether the pattern that greatest sensitivity to density dependence occurs in seed production is general. However, it seems to me desirable to postpone these calculations until enough data beecome available for one to be more confident about what meaningful parameter ranges might be.

Suppose the effect size hierarchy revealed here does extend to fynbos in general. It would still be very hard to untangle the direction of selection pressure. Further calculations concerning plants with differing attributes are needed to throw light on the effects of possible evolutionary changes, and at present any specultation about such effects seem empty to me. The one robust requirement of the theory is that to persist, a plant requires mechanisms of recovery from low density. I have argued that for serotinous proteoids it may well be reproductive weakness of the companion species when at high density. Conversely, one may consider Johnson's [87] suggestion that many fynbos geophytes and orchids are pollinator limited and his further suggestion that pollination efficiency drops at low densities. This would constitute a depsensatory mechanism so that zero would be a stable equilibrium and below a critical density the population would decline to zero. Thus one would expect numerous local extinctions, which may explain rapid speciation, and hence the high diversity and frequent endemism in these groups.

4.6 Future possibilities

The above suggests a number of important questions for fynbos. Firstly, that relative density dependence may be generated by all the various components of the population dynamics, but not to equal effect. It seems to me important that one finds out what happens in reality. The kind of simulation done above could usefully support, but never take precedence over field studies. I suggest the most useful way to use simulation is for helping to choose between experiments. Because the framework advocated in this thesis is comprehensive and rigorous, it allows one to simulate the effect of more accurate data—indeed, it leads one to the areas where data are lacking entirely and by investigating the effect of assumptions one can rank gaps in knowledge.

It may be that one then decides to investigate density dependence more broadly (Klinkhamer *et al.* [92] give brief suggestions on experimentation, but note that the tests developed by Crowley [44] are not directly useful in fynbos), or in seed production only, or in seed/seedling predation (Botha and le Maitre [17]). It may be that effects of relative density on fire are important. Whatever is chosen for data collection, it is clear the main obstacle to a population dynamics approach to species coexistence in fynbos is the difficulty of obtaining life table data. In this circumstance, the role of theory is in the first place to point out the most important data still needed. I should add that the results above merely indicate that this role could be played, but do not amount to such a ranking of research priorities.

The range of topics amenable to the above approach is vast. In this chapter, I have concentrated on within-guild coexistence of serotinous proteoids. There is also the question of between-guild coexistence in fynbos: reseeders vs. resprouters; canopy shrubs vs. understory shrubs. There are numerous questions of abundance: why is the restioid guild so important? Why are seedling-to-parent ratios so variable? There are economic issues: how to find optimal densities for flower production; development of harvesting strategies that safeguard recruitment. The difficult task will be to decide where the appoach will be most effective.

Chapter 5

Conclusion: strengths and weaknesses of taking the fully general point of view

Biologists revel in detail, while mathematicians strive for generality. Part of my intention is to show how they need each other, and that a strong interaction between detail and generality is highly desirable. Thus, in a certain sense, this thesis is an exercise in using general results to address particular questions, and it is appropriate to review this project as an approach to plant population dynamics. In this chapter I give an outline of generality as a point of view, and consider its strengths and weaknesses. I end the chapter, and the thesis, by discussing some aspects of possible future research; however, the specific suggestions at the ends of chapters 3 and 4 are not repeated here.

5.1 The fully general point of view: a research heuristic for plant population dynamics

In biology, generality as a point of view consists of a synthesis of widely different elements, where much important detail is suppressed. As such it is not particularly of interest to most biologists. I prefer the mathematical view of generality, which is that one should proceed rigorously from the smallest possible set of axioms. It is then possible to add further (and perhaps more restrictive) axioms, and in this way to build up the detail appropriate to particular problems.

This provides a useful heuristic: when facing a problem, start with the widest possible framework, i.e. the minimal assumptions. Very likely only a few conclusions can be drawn, but they do guide further investigation, in particular by suggesting further assumptions, but also by pointing to gaps where more data are needed, and by allowing assessment of which new data would most effectively assist progress. This interaction proceeds until the problem is solved or abandoned. (Of course, this is a somewhat idealised picture in that one is never in full control of all the assumptions involved.)

I claim no special virtue for this approach to research, or for the axioms of chapter 2. Research plans are an expression of creativity of the highest order, and it would never do to prescribe one form for all. Similarly, axiom sets proliferate notoriously fast. My contention is simply that this does provide a fully general approach, and that this approach is worthwhile.

5.2 Strengths

The strengths considered here arise from the combination of generality and rigour, and I shall illustrate them with examples from earlier chapters.

Firstly, one can compare models. Chapter 3 is a good example of this, where a family of models was rigorously formulated, and all nine possible types of model could be listed. The discrimination among groups of models suggested there depends crucially on the fact that all nine types of model were investigated. That suggestion was made tentatively because of the preliminary nature of the investigation, and not from any fear that some unknown type of model could confound the argument.

Secondly, the comprehensiveness of a general treatment allows one to combine models in a coherent way, and to spot gaps in models that consist of a network of submodels. In chapter 4, I used this opportunity to identify all possible sources of density dependence in serotinous proteoids. It would be interesting to know whether a similar analysis is possible and useful for more complex models, but that is a question I am not equipped to answer.

The third great strength of the general point of view is as a guide to thought. This is a subtle effect, which operates behind the scenes and while no concrete example from this thesis suggests itself, it pervaded the preparation. It is true that generality as a style of thought is not to be imposed on every researcher; nevertheless it can be very useful.

5.3 Weaknesses

There are at least four weaknesses in the approach to plant population dynamics taken here.

Firstly, the redundancy in ρ: as pointed out in chapter 2, it is possible in this formulation to model structure either through subpopulations or through attributes. This is in fact the dichotomy between discrete and continuum models in another guise. I believe a multiplicity of modelling possibilities does reduce our ability to order thought; however, a fairly large measure of redundancy is unavoidable, if only because a large variety of models serving a wide range of purposes are needed.

Secondly, the general model is not in itself applicable, but needs additional assumptions in the form of constitutive relations. However, the constitutive relations of plant population dynamics are not known. By this I mean that, while a myriad of relations can be and have been considered, there seems to be no consensus on exactly how to model things like mortality and somatic growth. This weakness appears in both chapter 3, where the actual constitutive relations used were not based on any demonstrable value they may have, and in chapter 4, where the constitutive relations have illustrative value only, and cannot with any confidence be said to hold for serotinous proteoids. A good general model where only poorly known constitutive relations exist amounts to a persistent temptation to draw easily available but unwarranted conclusions.

Thirdly, the model considered here is of course deterministic, and this is a problem when so many parameters and environmental variables are known only with a high degree of uncertainty. It has been suggested that the main existence theorem of this thesis can be generalised to the set-valued case (J.-P. Aubin, personal communication; for an account of set-valued anylysis see Aubin and Frankowska [7]). If so, it would certainly cover the full generality of error and uncertainty in population dynamics, but whether set-valued analysis can provide results of use in the theory and practice of ecology remains to be seen.

This leads to the final weakness, which is that the main problem in biology is that data are sparse, expensive and noisy. The general point of view taken here does not contribute directly to the optimal use of available data.

5.4 Future possibilities

A number of directions for mathematical exploration suggest themselves, as is to be expected from such a general project. For example: how much can one relax the assumptions of (Hyp 1), or indeed any other assumptions? Does Theorem 2 carry over with full generality to the set-valued case? Can environmental variation really be included without loss of generality? Does the general discrete model become equivalent to the general continuum model in the small-scale limit? And so on. However, it seems to me inappropriate to pursue the mathematical possibilities without consideration of the biological issues. Interesting biological questions arising from this thesis include the role of spatial structure in lottery dynamics; the question whether differences in intensity of density dependence in the various population processes have biological importance; and a more detailed understanding of why plants die. I also believe that analysis within a rigorous general framework can distinguish adequately between various possible meanings of "competition" and "population regulation", and so help to unclutter the ecological debate so deplored by Peters [124]. However, from the general point of view, the outstanding problem in population dynamics is the weakness of models of constitutive relations. The general framework proposed here is suitable for exploring a wide variety of models, particularly since they have to be combined, but that is hardly the point. Rather, it should be possible to use the general framework to specify what the properties of a good constitutive relation are, and this might help the search by clarifying what it is one is looking for. If one had good constitutive relations, the problem of optimal use of data might be solvable, and then the general approach to population dynamics will have been justified many times over.

Appendix A

A mistake in Tucker and Zimmerman [1]

Tucker and Zimmerman's [144] paper is important, in that it aims at the basic mathematical analysis of a potentially very important model in population dynamics. However, in spite of the admirable aim and the fact that the basic issues are thoroughly considered, there is a fundamental flaw in their paper, which is a mistake in the proof offered for their existence theorem.

Specifically, there is an invalid claim on page 561. The two relevant formulae are (3.15) and (3.16) in their numbering; they are given here exactly as in the paper. The reader is not expected to follow each individual symbol, as most of them play no role in what follows, and those that do are fully explained below (in particular, ignore the difference between $J\phi_{-\alpha}$ and J_ϕ used in chapter 2). The claim is that (3.15) implies (3.16).

$$\begin{aligned}(3.15):\quad & |B(\mathbf{x},t) - \hat{B}(\mathbf{y},t)| \\ \leq\ & \int_0^t\int_\Omega |\beta(\alpha,\mathbf{w},\mathbf{x},Q(t))B(\phi_{-\alpha})S(\alpha,\phi_{-\alpha})J\phi_{-\alpha} \\ & - \beta(\alpha,\mathbf{w},\mathbf{y},\hat{Q}(t))\hat{B}(\hat{\phi}_{-\alpha})\hat{S}(\alpha,\hat{\phi}_{-\alpha})J\hat{\phi}_{-\alpha}|(\alpha,\mathbf{w},t)\,d\mathbf{w}\,d\alpha\end{aligned}$$

---

[1] Reference [144]

$$+\int_0^\infty\int_\Omega \big|\beta(\alpha,\mathbf{w},\mathbf{x},Q(t))n_0(\phi_{-t})S(t,\phi_{-t})J\phi_{-t}$$
$$-\beta(\alpha,\mathbf{w},\mathbf{y},\hat{Q}(t))n_0(\hat{\phi}_{-t})\hat{S}(t,\hat{\phi}_{-t})J\hat{\phi}_{-t}\big|(\alpha,\mathbf{w},t)\,d\mathbf{w}\,d\alpha$$

(3.16): $$\big|B(\mathbf{x},t)-\hat{B}(\mathbf{y},t)\big|$$
$$\leq\ r_3(T)|Q-\hat{Q}|_T + r_4(T)\|\mathbf{x}-\mathbf{y}\|_k$$
$$+\,\bar{\beta}\int_0^t\int_\Omega \big|B(\phi_{-\alpha})-\hat{B}(\hat{\phi}_{-\alpha})\big|J\hat{\phi}_{-\alpha}(\alpha,\mathbf{w},t)\,d\mathbf{w}\,d\alpha$$

In their text, Tucker and Zimmerman explain the passage from (3.15) to (3.16) as follows: "By ... the insertion and deletion of like terms, the RHS of (3.15) can be bounded by the sum of eight terms to which the estimates derived above can be applied". I will first show that this is false, in that only six of these terms reduce to earlier work.

It is convenient to abbreviate (3.15) as follows:

(3.15abbr) $$\big|B(\mathbf{x},t)-\hat{B}(\mathbf{y},t)\big| \ \leq\ \int_0^t\int_\Omega |\beta BSJ-\hat{\beta}\hat{B}\hat{S}\hat{J}|\,d\mathbf{w}\,d\alpha + \int_0^\infty\int_\Omega |\beta n_0 SJ-\hat{\beta}\hat{n}_0\hat{S}\hat{J}|\,d\mathbf{w}\,d\alpha$$

where the symbols in (3.15abbr) relate to (3.15) in the obvious way. What Tucker and Zimmerman mean by "insertion and deletion of like terms" is adding and subtracting terms like $\hat{\beta}BSJ$, in order to form eight terms like the following:

(B1) $$\int_0^t\int_\Omega |\beta-\hat{\beta}|BSJ\,d\mathbf{w}\,d\alpha$$

(B2) $$\int_0^t\int_\Omega \hat{\beta}|B-\hat{B}|SJ\,d\mathbf{w}\,d\alpha$$

(B3) $$\int_0^t\int_\Omega \hat{\beta}\hat{B}|S-\hat{S}|J\,d\mathbf{w}\,d\alpha$$

(B4) $$\int_0^t\int_\Omega |\hat{\beta}\hat{B}\hat{S}|J-\hat{J}|\,d\mathbf{w}\,d\alpha$$

(B5) $$\int_0^\infty\int_\Omega |\beta-\hat{\beta}|n_0SJ\,d\mathbf{w}\,d\alpha$$

(B6) $$\int_0^\infty\int_\Omega \hat{\beta}|n_0-\hat{n}_0|SJ\,d\mathbf{w}\,d\alpha$$

$$\text{(B7)} \qquad \int_0^\infty \int_\Omega \hat{\beta}\hat{n}_0 |S - \hat{S}| J \, d\mathbf{w} \, d\alpha$$

$$\text{(B8)} \qquad \int_0^\infty \int_\Omega \hat{\beta}\hat{n}_0 \hat{S} |J - \hat{J}| \, d\mathbf{w} \, d\alpha$$

Following Tucker and Zimmerman's suggestion in this way does not lead to a unique set of eight terms. However, what the set above has in common with other sets derived a similar way are the factors like $|\beta - \hat{\beta}|$. For each of the two integrals, there are four such factors, so that each integral expands to four terms. Earlier in the paper, Tucker and Zimmerman used the same device on a similar pair of integrals, and came up with a set of six terms, which they labelled (A1) to (A6). Matching by factors like $|S - \hat{S}|$, one finds that (B3) is like (A2), (B4) is like (A3), (B6) is like (A4), (B7) is like (A5), and (B8) is like (A6). (B2) does not need to be matched (see below), but neither of the terms (B1) or (B5) can be matched with earlier estimates, for the simple reason that the factors β and $\hat{\beta}$ do not occur in them. Thus their claim is false.

However, this results merely in a gap in the proof. I now show that this gap cannot be filled, by showing that it is impossible to derive (3.16) from (3.15) in the suggested way i.e. via expanding (3.15) into the sum of terms (B1)–(B8). I start with a result from earlier in the paper: the five terms (A2)–(A6) have all been shown to be bounded by either $r(t)|Q - \hat{Q}|_T$ or by $r(t)||\mathbf{x} - \mathbf{y}||_k$ (or by a sum of these two terms), where $r(t)$ is a non-negative function that goes to zero with t. Thus there is no problem in relating (3.15) to (3.16) via these five terms, and this disposes of (B3), (B4) and (B6)–(B8). Furthermore, (B2) gives the integral term in (3.16). This leaves (B1) and (B5), and in discussing these some of the symbols will have to be explained. The function β is intended to model the birth process, and, by hypothesis (H5) stated on page 554, it is nonnegative and satisfies a Lipschitz condition with Lipschitz constant c_β. I interpret this condition in a way that proves that (B1) can be bounded in the desired way as follows:

$$\int_0^t \int_\Omega |\beta - \hat{\beta}| BSJ \, d\mathbf{w} \, d\alpha \leq \int_0^t \int_\Omega c_\beta \big(||\mathbf{x} - \mathbf{y}||_k + |Q(t) - \hat{Q}(t)| \big) BSJ \, d\mathbf{w} \, d\alpha.$$

Recalling the appropriate Lipschitz properties and that B, S and J are bounded, it is clear that using $r(t) = \int_0^t \int_\Omega BSJ \, d\mathbf{w} \, d\alpha$ yields the desired form.

However, (B5) cannot be so bounded, because the interval of integration does not get small as t gets small. To be specific, by using the Lipschitz property of β one obtains

$$c_\beta(\|\mathbf{x}-\mathbf{y}\|_k + |Q(t)-\hat{Q}(t)|) \int_0^\infty \int_\Omega n_0 S(t,\phi_{-t}) J\phi_{-t}\, d\mathbf{w}\, d\alpha.$$

Neither S nor J go to zero with t. It is true that n_0 has been assumed to have compact support (p. 553), but the interval over which $n_0(\phi_{-t})$ is non-zero is in fact constant with respect to α and t. Hence there is no way in which (B5) can be considered to contain a factor of the form $r(t)$, and thus Tucker and Zimmerman's claim that (3.16) can be derived from (3.15) is false.

Bibliography

[1] G. I. Ågren and T. Fagerstrom. Limiting dissimilarity in plants: randomness prevents exclusion of species with similar competitive abilities. *Oikos*, 43:369–375, 1984.

[2] W.G. Aiello, H.I. Freedman, and J. Wu. Analysis of a model representing stage-structured population growth with state-dependent time delay. *SIAM Journal of Applied Mathematics*, 52:855–869, 1992.

[3] D.P. Aikman and A.R. Watkinson. A model for growth and self-thinning in even-aged monocultures of plants. *The Annals of Botany*, 45:419–427, 1980.

[4] H.G. Andrewartha and L.C. Birch. *The Distribution and Abundance of Animals.* University of Chicago Press, Chicago, 1954.

[5] J. Antonovics and D.A. Levin. The ecological and genetic consequences of density-dependent regulation in plants. *Annual Review of Ecology and Systematics*, 11:411–452, 1980.

[6] T. Arbogast and F.A. Milner. A finite difference method for a two-sex model of population dynamics. *SIAM Journal of Numerical Analysis*, 26:1474–1486, 1989.

[7] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis.* Birkhäuser, Boston, 1990.

[8] H.T. Banks and B.G. Fitzpatrick. Estimation of growth rate distributions in structured populations models. *Quarterly of Applied Mathematics*, 49:215–235, 1991.

[9] M. Begon and M. Mortimer. *Population Ecology.* Blackwell, Oxford, second edition, 1986.

[10] G.I. Bell and E.C. Anderson. Cell growth and division. I. A mathematical model with applications to cell volume distributions in mammalian suspension cultures. *Biophysical Journal*, 7:329–351, 1967.

[11] M. Bertsch, M.E. Gurtin, and D. Hilhorst. On interacting populations that disperse to avoid crowding: the case of equal dispersal velocities. *Nonlinear Analysis, Theory, Methods & Applications*, 11:493–499, 1987.

[12] J. Bishir and G. Namkoong. Density-dependent dynamics in size-structured plant populations. *Journal of Theoretical Biology*, 154:163–188, 1992.

[13] G.B. Bonan. The size structure of theoretical populations : spatial populations and neighbourhood effects. *Ecology*, 69:1721–1730, 1988.

[14] W.J. Bond, J. Vlok, and M. Viviers. Variation in seedling recruitment of Cape Proteceae after fire. *Journal of Ecology*, 7:209–221, 1984.

[15] W.J. Bond. Canopy-stored seed reserves (serotiny) in Cape Proteaceae. *South African Journal of Botany*, 51:181–186, 1985.

[16] W.J. Bond, R.M. Cowling, and M.B. Richards. Competition and coexistence. In R.M. Cowling, editor, *The Ecology of Fynbos*, pages 206–225. Oxford University Press, 1992.

[17] S.A. Botha and D.C. le Maitre. Effects of seed and seedling predation by small mammals on seedling recruitment of *Protea neriifolia* in Swartboskloof, Cape Province. *South African Journal of Zoology*, 27:60–69, 1992.

[18] G.E.P. Box and G.M. Jenkins. *Time Series Analysis.* Holden-Day, San Francisco, 1970.

[19] S. Busenberg and M. Ianelli. A class of nonlinear diffusion problems in age-dependent population dynamics. *Nonlinear Analysis, Theory, Methods & Applications*, 7:501–529, 1983.

[20] D.S. Butterworth and A.E. Punt. Some preliminary examinations of the potential infromation content of age-structure data from Antarctic minke whale research catches. *Reports of the International Whaling Commission*, 60:301–315, 1990.

[21] M.G.R. Cannel, P. Rothery, and E.D. Ford. Competition within stands of *Picea sitcensis* and *Pinus contorta. The Annals of Botany*, 53:349–362, 84.

[22] T.J. Carleton and B. Wannamaker. Mortality and self-thinning in postfire black spruce. *Annals of Botany*, pages 621–628, 1987.

[23] H. Caswell. *Matrix Population Models.* Sinauer Associates, Sunderland, 1989.

[24] P. Chesson. Interactions between environment and competition: how fluctuations mediate coexistence and competitive exclusion. In S.A. Levin, editor, *Community Ecology. Lecture Notes in Biomathematics 77*, pages 51–71. Springer-Verlag, Heidelberg, 1988.

[25] P.L. Chesson. Coexistence of competitors in a stochastic environment: the storage effect. In H.I. Freedman and C. Strobeck, editors, *Population Biology. Lecture Notes in Biomathematics 52*, pages 188–198. Springer Verlag, Berlin, 1983.

[26] P.L. Chesson. Environmental variation and the coexistence of species. In J. Diamond and T.J. Case, editors, *Community Ecology*, pages 240–256. Harper and Row, New York, 1986.

[27] P.L. Chesson. Geometry, heterogeneity and competition in variable environments. *Philosophical Transactions of the Royal Society of London B*, 330:165–173, 1990.

[28] P.L. Chesson and T.J. Case. Overview: nonequilibrium community theories: chance, variability, history and coexistence. In J. Diamond and T.J. Case, editors, *Community Ecology*, pages 229–239. Harper and Row, New York, 1986.

[29] P.L. Chesson and S. Ellner. Invasibility and stochastic boundedness in monotonic competition models. *Journal of Mathematical Biology*, 27:117–138, 1989.

[30] P.L. Chesson and N. Huntly. Community consequences of life history traits in a variable environment. *Annales Zoologica Fennici*, 25:5–16, 1988.

[31] P.L. Chesson and R.R. Warner. Environmental variability promotes coexistence in lottery competitive systems. *The American Naturalist*, 117:923–943, 1981.

[32] P.L. Chesson. A need for niches? *Trends in Ecology and Evolution*, 6:26–28, 1991.

[33] E. Cittadino. *Nature as the laboratory.* Cambridge University Press, Cambridge, 1990.

[34] J.S. Clark. Integration of ecological levels: individual plant growth, population mortality and ecosystem process. *Journal of Ecology*, 78:275–299, 1990.

[35] J.S. Clark. Disturbance and population structure on the shifting mosaic landscape. *Ecology*, 72:1119–1137, 1991.

[36] B. Cockburn and C-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin FEM for conservation laws. II: General framework. *The Mathematics of Computation*, 52:411–435, 1989.

[37] D.P. Coffin and W.K. Lauenroth. A gap dynamics simulation model of succession in a semiarid grassland. *Ecological Modelling*, 49:229–266, 1990.

[38] H.N. Comins and I.R. Noble. Dispersal, variability and transient niches : species coexistence in a uniformly varying environment. *The American Naturalist*, 126:706–723, 1985.

[39] G.J. Cooper. On the implementation of singly implicit Runge-Kutta methods. *The Mathematics of Computation*, 57:663–672, 1991.

[40] R.M. Cowling. Fire and its role in coexistence and speciation in Gondwanan shrublands. *South African Journal of Science*, 83:106–111, 1987.

[41] R.M. Cowling, editor. *The Ecology of Fynbos.* Oxford University Press, Cape Town, 1992.

[42] R.M. Cowling and P.M. Holmes. Flora and vegetation. In R.M. Cowling, editor, *The Ecology of Fynbos*, pages 23–61. Oxford University Press, 1992.

[43] M.J. Crawley. The population dynamics of plants. *Philosophical Transactions of the Royal Society of London B*, 330:125–140, 1990.

[44] P.H. Crowley. Density dependence, boundedness, and attraction: detecting stability in stochastic systems. *Oecelogia*, 90:246–254, 1992.

[45] J.M. Cushing and M. Saleem. A predator prey model with age structure. *Journal of Mathematical Biology*, 14:231–250, 1982.

[46] T. Czaran and S. Bartha. Spatioemporal dynamic models of plant populations and communities. *Trends in Ecology and Evolution*, 7:38–42, 1992.

[47] D.L. DeAngelis. Equilibrium and non-equilibrium concepts in ecological models. *Ecological Monographs*, 57:1–21, 1987.

[48] C. Dey and S.K. Dey. Explicit finite difference predictor and convex corrector with applications to hyperbolic partial differential equations. *Computers and Mathematics with Applications*, 9:549–557, 1983.

[49] G. Di Blasio, M. Ianelli, and E. Sinestrari. Approach to equilibrium in age structured populations with an increasing recruitment process. *Journal of Mathematical Biology*, 13:371–382, 1982.

[50] J. Douglas Jr. and F.A. Milner. Numerical methods for a model of population dynamics. *Calcolo*, 24:247–254, 1987.

[51] N. Dunford and J. Schwartz. *Linear Operators, Part I: General Theory*. Interscience, 1958.

[52] R.H. Elderkin. Population models with globally age-dependent dynamics: on computing the steady state. *Computers and Mathematics with Applications*, 9:371–376, 1983.

[53] A.M. Ellison. Density-dependent dynamics of *Salicornia europaea* monocultures. *Ecol*, 68:737–741, 1987.

[54] S. Ellner. Asymptotic behaviour of some stochastic difference equation population models. *Journal of Mathematical Biology*, 19:169–200, 1984.

[55] O. Eriksson. Population structure and dynamics of the clonal dwarf-shrub *Linnaea borealis*. *Journal of Vegetation Science*, 3:61–68, 1992.

[56] K. Esler and R.M. Cowling. Effects of density on the reproductive output of *Protea lepidocarpodendron*. *South African Journal of Botany*, 56:29–33, 1990.

[57] K.J. Esler and N. Phillips. Experimental effects of water stress on semi-arid karoo seedlings: implications for field seedling survivorship. In press, *Journal of Arid Environments*.

[58] T. Fagerström. Lotteries in communities of sessile organisms. *Trends in Ecology and Evolution*, 3:303–306, 1988.

[59] T. Fagerström and G. Ågren. Theory for coexistence of species differing by regeneration porperties. *Oikos*, 33:1–10, 1979.

[60] T. Fagerström and G. Ågren. Phenological spread in plants: a result of adaptations to environmental stochasticity? *Vegetatio*, 43:83–86, 1980.

[61] R.M. Feldman and G.L. Curry. A PDE formulation and numerical solution for a boll weevil-cotton crop model. *Computers and Mathematics with Applications*, 9:393–402, 1983.

[62] L.G. Firbank and A.R. Watkinson. On the analysis of competition at the level of the individual plant. *Oecologia*, 71:308–317, 1987.

[63] E.D. Ford and P.J. Diggle. Competition for light in a plant monoculture modelled as a stochastic process. *The Annals of Botany*, 48:481-500, 1981.

[64] G.F. Gause. *The struggle for existence.* Hafner, New York, 1969, first published in 1934.

[65] B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value

of life contingencies. *Proceedings of the Royal Society of London*, 115:513–585, 1825.

[66] E. Gorham. Shoot height, weight and standing crop in relation to density of monospecific plant stands. *Nature*, 279:148–150, 1979.

[67] J.M. van Groenendael and P. Slim. The contrasting dynamics of two populations of *Plantago lanceolata* classified by age and size. *Journal of Ecology*, 76:585–599, 1988.

[68] H-E. Gross. The employment of mathematicians in insurance companies in the 19th century. In H. Mehrtens, H. Bos, and I. Schneider, editors, *Social history of nineteenth century mathematics*, pages 179–196. Birkhäuser, Boston, 1981.

[69] M.E. Gurtin. A system of equations for age-dependent population diffusion. *Journal of Theoretical Biology*, 40:389–392, 1973.

[70] M.E. Gurtin. Some questions and open problems in continuum mechanics and population dynamics. *Journal of Differential Equations*, 48:293–312, 1983.

[71] M.E. Gurtin and R.C. MacCamy. Non-linear age-dependent population dynamics. *Archives of Rational Mechanics*, 54:281–300, 1974.

[72] C.A.S. Hall. An assessment of several of the historically most influential theoretical models used in ecology and of the data provided in their support. *Ecological Modelling*, 104:5–31, 1988.

[73] T. Hara. Dynamics of stand structure in plant monocultures. *Journal of Theoretical Biology*, 110:223–239, 1984.

[74] T. Hara. Modelling the time course of self-thinning in crowded plant populations. *The Annals of Botany*, 53:181–188, 1984.

[75] T. Hara. A stochastic model and the moment dynamics of the growth and size distribution in plant populations. *Journal of Theoretical Biology*, 109:173–190, 1984.

[76] T. Hara. A model for mortality in a self-thinning plant population. *The Annals of Botany*, 55:667–674, 1985.

[77] T. Hara. Effects of mode of competition on stationary size distribution in plant populations. *The Annals of Botany*, 69:509–513, 1992.

[78] T. Hara, M. Kimura, and K. Kikuzawa. Growth patterns of tree height and stem diameter in populations of *Abies veitchii*, *A. mariesii* and *Betula ermanii*. *Journal of Ecology*, 79:1085–1098, 1991.

[79] J.L. Harper. A centenary in population biology .... *Nature*, 252:526–527, 1974.

[80] J.L. Harper. *Population Biology of Plants.* Academic Press, New York, 1977.

[81] J.S. Hatfield and P.L. Chesson. Diffusion analysis and stationary distribution of the two-species lottery competition model. *Theoretical Population Biology*, 36:251–266, 1989.

[82] D.W. Hilbert. A model of life history strategies of chaparral shrubs in relation to fire frequency. In J.D. Tenhunen, F.M. Catarino, O.L. Lange, and W.C. Oechel, editors, *Plant Response to Stress*, pages 597–606. Springer-Verlag, Berlin, 1987.

[83] M.A. Huston and D.L. DeAngelis. Size bimodality in monospecific populations: a critical review of potential mechanisms. *The American Naturalist*, 129:678–707, 1987.

[84] A.J. Ihde. *The Development of Modern Chemistry.* Harper and Row, New York, 1964.

[85] M.S. Ingber and N. Phan-Thien. A boundary element approach for parabolic differential equations using a class of particular solutions. *Applied Mathematical Modelling*, 16:124–132, 1992.

[86] Y. Iwasa, K. Sato, and T. Maud Kubo. Modelling biodiversity: latitudinal gradient of forest species diversity. In E.D. Schulze and H.A.L. Mooney, editors, *Biodiversity and Ecosystem Function.* Springer-Verlag, 1993.

[87] S.D. Johnson. Pollinator-driven speciation in the Cape fynbos. In submission.

[88] P. Kareiva and M. Andersen. Spatial aspects of species interactions: the wedding of models and experiments. In S.A. Levin, editor, *Community Ecology*, pages 35–50. Springer-Verlag, 1988.

[89] N.C. Kenkel. Pattern of self-thinning in jack pine: testing the random mortality hypothesis. *Ecology*, 69:1017–1024, 1988.

[90] S.E. Kingsland. *Modelling Nature.* University of Chicago Press, Chicago, 1985.

[91] P.G.L. Klinkhamer and T.J. de Jong. A deterministic model to study the importance of density-dependence for regulation and the outcome of intraspecific competition in populations of sparse plants. *Acta Botanica Neerlandia*, 38:57–65, 1989.

[92] P.G.L. Klinkhamer, E. Meelis, T.J. de Jong, and J. Weiner. On the analysis of size-dependent reproductive output in plants. *Functional Ecology*, 6:308–316, 1992.

[93] R.G. Knox, R.K. Peet, and N.L. Christensen. Population dynamics in loblolly pine stands: changes in skewness and size inequality. *Ecology*, 70:1153–1166, 1989.

[94] T. Kohyama. Simulating stationary size distribution of trees in rain forests. *The Annals of Botany*, 68:173–180, 1991.

[95] T. Kohyama. Size-structured multi-species model of rain forest trees. *Ecology*, 6:206–212, 1992.

[96] T.V. Kostova. Numerical solutions of a hyperbolic differential-integral equation. *Computers and Mathematics with Applications*, 15:427–436, 1988.

[97] F.J. Kruger. Succession after fire in selected fynbos communities of the south-western cape. PhD thesis, University of the Witwatersrand, 1987.

[98] B.B. Lamont, D.C. le Maitre, R.M. Cowling, and N.J. Enright. Canopy seed storage in woody plants. *Botanical Review*, 57:277–317, 1991.

[99] S. Lavorel and J.D. Lebreton. Evidence for lottery recruitment in mediterranean old fields. *Journal of Vegetation Science*, 3:91–100, 1992.

[100] R. Law. A model for the dynamics of a plant population containing individuals classified by age and size. *Ecology*, 64:224–230, 1983.

[101] D.C. le Maitre and J.J. Midgley. Plant reproductive ecology. In R.M. Cowling, editor, *The Ecology of Fynbos*, pages 135–174. Oxford University Press, 1992.

[102] S.A. Levin, T.G. Hallam, and L.J. Gross, editors. *Applied Mathematical Ecology. Biomathematics 18.* Springer-Verlag, Berlin, 1989.

[103] R. Levins. The strategy of model building in population biology. *American Scientist*, 54:421–431, 1966.

[104] J. Lighthill. *An Informal Introduction to Theoretical Fluid Mechanics.* Clarendon Press, Oxford, 1986.

[105] H.P. Linder, M.E. Meadows, and R.M. Cowling. History of the Cape flora. In R.M. Cowling, editor, *The Ecology of Fynbos*, pages 113–134. Oxford University Press, 1992.

[106] H. Linhart and W. Zucchini. *Model Selection.* John Wiley and sons, New York, 1986.

[107] F.T. Lloyd and W.R. Harms. An individual stand growth model for mean plant size based on the rule of self-thinning. *The Annals of Botany*, 57:681–688, 1986.

[108] C. Loehle. Hypothesis testing in ecology: psychological aspects and the importance of theory maturation. *Quarterly Review of Biology*, 62:397–409, 1987.

[109] J.N. Long and F.W. Smith. Relation between size and density in developing stands : a description and possible mechanisms. *Forest Ecology and Management*, 7:191–206, 1984.

[110] W.M. Lonsdale and A.R. Watkinson. Light and self-thinning. *New Phytologist*, 90:431–445, 1982.

[111] W.M. Lonsdale. The self-thinning rule: dead or alive? *Ecology*, 71:1371–1388, 1990.

[112] L. Lopez and D. Trigiante. A finite difference scheme for a stiff problem arising in the numerical solution of a population dynamical model with spatial diffusion. *Nonlinear Analysis, Theory, Methods & Applications*, 9:1–12, 1985.

[113] J.A.J. Metz and O. Diekmann, editors. *The dynamics of physiologically structured populations*, volume 68 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, 1986.

[114] F.A. Milner. A numerical method for a model of population dynamics with spatial diffusion. *Computers and Mathematics with Applications*, 19:31–41, 1990.

[115] F.A. Milner. A finite element method for a two-sex model of population dynamics. *Numerical Methods for Partial Differential Equations*, 4:329–345, 1988.

[116] R. Mithen, J.L. Harper, and J. Weiner. Growth and mortality of individual plants as a function of "available area". *Oecologia*, 62:57–60, 1984.

[117] A.G. M'Kendrick. Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–190, 1926.

[118] C.L. Mohler, P.L. Marks, and D.G. Sprugel. Stand structure and the allometry of trees during self-thinning of pure stands. *Journal of Ecology*, 66:599–614, 1978.

[119] S.W. Pacala. Competitive equivalence: the coevolutionary consequences of the sedentary habit. *The American Naturalist*, 132:576–593, 1988.

[120] S.W. Pacala and J.A. Silander. Neighbourhood models of plant population dynamics. I. Single-species models of annuals. *The American Naturalist*, 125:385–411, 1985.

[121] T. Park. Experimental evolution. *Quarterly Review of Biology*, 10:209–212, 1935.

[122] R.K. Peet and N.L. Christensen. Competition and tree death. *Bioscience*, 37:586–595, 1987.

[123] L.K. Penridge and J. Walker. Effect of neighbouring trees on eucalypt growth in a semi-arid woodland in Australia. *Journal of Ecology*, 74:925–936, 1986.

[124] R.H. Peters. *A Critique for Ecology.* Campbridge University Press, Cambridge, 1991.

[125] L.F. Pitelka. Application of the -3/2 power law to clonal herbs. *The American Naturalist*, 123:442–449, 1984.

[126] M. Rees and M.J. Crawley. Growth, reproduction and population dynamics. *Functional Ecology*, 3:645–653, 1989.

[127] M. Rees and M.J. Crawley. Do plant populations cycle? *Functional Ecology*, 5:580–583, 1991.

[128] M. Rees and M.J. Long. Germination biology and the ecology of annual plants. *The American Naturalist*, 139:484–508, 1992.

[129] L.H. Reineke. Perfecting a stand-density index for even-aged forests. *Journal of Agricultural Research*, 46:627–638, 1933.

[130] M. Rotenberg. Equilibrium and stability in populations whose interactions are age-specific. *Journal of Theoretical Biology*, 54:207–224, 1975.

[131] P.F. Sale. Maintenance of high diversity in coral reef fish communities. *The American Naturalist*, 111:337–359, 1977.

[132] D.A. Samson and K.S. Werk. Size-dependent effects in the analysis of reproductive effort in plants. *The American Naturalist*, 127:667–680, 1986.

[133] F.M. Scudo and J.R. Ziegler. *The Golden Age of Theoretical Ecology: 1923-1940.* Springer-Verlag, Berlin, 1978.

[134] R.G. Shaw and J. Antonovics. Density-dependence in *Salvia lyrata,* a herbaceous perennial: the effects of experimental alteration of seed densities. *Journal of Ecology*, 74:797–813, 1986.

[135] A. Shmida and S. Ellner. Coexistence of plants with similar niches. *Vegetatio*, 58:29–55, 1984.

[136] J.A. Silander and S.W. Pacala. Neighbourhood predictors of plant performance. *Oecologia*, 66:256–263, 1985.

[137] J. Silvertown. Modularity, reproduction thresholds and plant population dynamics. *Functional Ecology*, 5:577–582, 1991.

[138] J. Silvertown and R. Law. Do plants used niches? Some recent developments in plant community ecology. *Trends in Ecology and Evolution*, 2:24–26, 1987.

[139] J.W. Sinko and W. Streifer. A new model for age-structure of a population. *Ecology*, 48:910–918, 1967.

[140] O.T. Solbrig, R. Sarandón, and W. Bossert. Effect of varying density and life table parameters on the growth rate and population size of *Viola fimbriatula*. *Acta Oecologia*, 11:263–280, 1990.

[141] T.R.E. Southwood. Introductory remarks. *Philosophical Transactions of the Royal Society of London B*, 330:123–124, 1990.

[142] D. Tilman. *Resource Competition and Community Structure*. Princeton University Press, Princeton, 1982.

[143] D. Tilman. *Dynamics and Structure of Plant Communities*. Princeton University Press, Princeton, 1988.

[144] S. Tucker and S. Zimmerman. A model of non-linear population dynamics with an arbitrary number of structure variables. *SIAM Journal of Applied Mathematics*, 17:549–591, 1988.

[145] A.V. Tuzinkevich. Bifurcations and chaos in a time-discrete integral model of population dynamics. *Mathematical Biosciences*, 109:99–126, 1992.

[146] E. van der Meijden. Mechanisms in plant population control. In P.J. Grubb and P.J. Whittaker, editors, *Toward a more exact ecology*, pages 163–181. Blackwell, Oxford, 1989.

[147] R. Vermiglio. On the etability of Runge-Kutta methods for delay integral equations. *Numerische Mathematik*, 61:561–577, 1992.

[148] H. von Foerster. Some remarks on changing populations. In F. Stohlman, Jr., editor, *Kinetics of Cellular Proliferation*, pages 382–407. Grune and Stratton, 1959.

[149] R.R. Warner and P.L. Chesson. Coexistence mediated by recruitment fluctuations: a field guide to the storage effect. *The American Naturalist*, 125:769–787, 1985.

[150] G.F. Webb. Dynamics of populations structured by internal variables. *Mathematische Zeitschrift*, 189:319–335, 1985.

[151] G.F. Webb. *Theory of Non-linear Age-dependent Population Dynamics*. Academic Press, New York, 1985.

[152] J. Weiner. Assymetric competition in plant populations. *Trends in Ecology and Evolution*, 5:360–364, 1990.

[153] J. Weiner and P.T. Conte. Dispersal and neighbourhood effects in an annual competition model. *Ecological Modelling*, 13:131–147, 1981.

[154] J. Weiner and S.C. Thomas. Size variability and competition in plant monocultures. *Oikos*, 47:211–222, 1986.

[155] E Weinstock and C. Rorres. Local stability of an age-structured population with density-dependent fertility and mortality. *SIAM Journal of Applied Mathematics*, 47:589–604, 1987.

[156] D.E. Weller. A reevaluation of the -3/2 power law of plant self-thinning. *Ecological Monographs*, 57:23–43, 1987.

[157] D.E. Weller. Self-thinning exponent correlated with allometric measures of plant geometry. *Ecology*, 68:813–821, 1987.

[158] D.E. Weller. The interspecific size-density relationship among crowded plant stands and its implications for the -3/2 power rule of self-thinning. *The American Naturalist*, 133:20–41, 1989.

[159] P.W. West and C.J. Borough. Tree suppression and the self-thinning rule in a monoculture of *Pinus radiata* D. Don. *Annals of Botany*, 32:149–158, 1983.

[160] M. Westoby. Frequency distributions of plant size during competitive growth of stands: the operation of distribution-modifying functions. *Annals of Botany*, 50:733–735, 1982.

[161] M. Westoby. The self-thinning rule. *Advances in Ecological Research*, 14:167–225, 1984.

[162] M. Westoby and J. Howell. Influence of population structure on self-thinning plant populations. *Journal of Ecology*, 74:343–359, 1986.

[163] J. White and J.L. Harper. Correlated changes in plant size and number in plant populations. *Journal of Ecology*, 58:467–485, 1970.

[164] R.I. Yeaton and W. Bond. Competition between two shrub species: dispersal differences and fire promote coexistence. *The American Naturalist*, 138:328–341, 1991.

[165] K. Yoda, T. Kira, H. Ogawa, and K. Hozumi. Self-thinning in overcrowded pure stands under cultivated and natural conditions. *Journal of Biology of Osaka City University*, 14:107–129, 1963.

[166] B. Zeide. Analysis of the 3/2 power law of self-thinning. *Forest Science*, 33:517–537, 1987.

[167] S. Zimmerman and R.A. White. Generalizations of a fluid dynamic model for analyzing multiparameter flowcytometric data. In M. Rothenberg, editor, *Biomathematics and cell kinetics*, pages 403–409. Elsevier/North Holland Biomedical Press, New York, 1981.