Multiple time–scale dynamics of stage structured populations and derivative–free optimization

Ute Alexandra Schaarschmidt

Thesis for the Degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2018

UNIVERSITY OF BERGEN

Multiple time–scale dynamics of stage structured populations and derivative–free optimization

Ute Alexandra Schaarschmidt

Thesis for the Degree of Philosophiae Doctor (PhD)
at the University of Bergen

2018

Date of defence: 02.11.2018

Year: 2018

Title: Multiple time–scale dynamics of stage structured populations and derivative–free optimization

Name: Ute Alexandra Schaarschmidt

Print: Skipnes Kommunikasjon / University of Bergen

Acknowledgements

My deepest and sincere thank you goes to Trond Steihaug, for his support and guidance throughout my research work. I appreciate that you were never tired of giving me advice for any problem, often in the form of some metaphor.

I would like to thank Sam Subbey for his support and for many fruitful discussions.

The atmosphere at the Institute of Informatics was delightful, and I would like to thank the administration for letting me stay at my working place until submission. Many of my colleagues have made the last years a great experience. To mention but a few, Joanna Bauer, Arne Klein and Marika Ivanova, I enjoyed your company.

Thank you to my friends for listening to me talking about my research and for making sure that I did not forget life outside of academia. Thank you for proofreading parts of my work.

Sindre Sæther has been very supportive. Danke, dass du mir den Rücken frei gehalten hast. Ohne dich wäre der Kühlschrank leer und die Kleidung in der Wäsche geblieben.

To my family: Takk til Sæther-familien for at jeg får lov til å være en del av familien. Marianne, danke dafür, dass du immer mein Bestes willst. Manche Leute haben den Grundstein für etwas gelegt, das damals noch mehr oder weniger unabsehbar in der Zukunft lag. Mein Dank gebührt auch dir, Hannelore, denn du warst wichtig für mich. Den wichtigsten Stein aber hast du gelegt, Bodo. Ich hätte gerne mit dir über mein Thema geredet und ich wünschte, du wüsstest, wie dankbar ich dir bin.

Abstract

The parent-progeny (adult fish–juvenile) relationship is central to understanding the dynamics of fish populations. Management and harvest decisions are based on the assumption of a stock-recruitment function that relates the number of adults to their progeny. Multi-stage population dynamic models provide a modelling framework for understanding this relationship, since they describe the dynamics of fish in several life stages (such as adults, eggs, larvae, and juveniles). Biological processes at various life stages usually evolve at distinct time scales.

This thesis contains three papers, which address challenges in modelling and parameter estimation for multiple time-scale dynamics of stage structured populations. A major question is, whether a multi-stage population dynamic model supports the assumption of a stock-recruitment function.

In the first paper, we address the parent-progeny relationship admitted by slow-fast systems of differential equations that model the dynamics of a fish population with two stages. We introduce a slow-fast population dynamic model which replicates several well-known stock-recruitment functions.

Traditionally, the dynamics of fish populations are described by difference equations. In the second paper, discrete time models for several life stages are formulated. We demonstrate that a multi-stage model may not admit a stock-recruitment function. Sufficient conditions for the validity of two hypotheses about the existence and structure of a parent-progeny function are established.

Parameters in population dynamic models can be estimated by minimizing a function of the solution of the ordinary differential equations and available data. Efficient and accurate methods for the solution of differential equations usually evaluate conditional statements. In this case, the objective function may be noisy, instead of continuously differentiable. Furthermore, an algorithm which is used to evaluate the objective function may unexpectedly fail to return a (plausible) value. Then, the optimization problem includes constraints which are only implicitly stated and hidden from the problem formulation.

We demonstrate that derivative-free optimization methods find sufficiently accurate solutions for the challenging optimization problems. In the third paper, we compare the performances of several derivative-free methods for a set of optimization problems. We find that a derivative-free trust-region method is most robust to the choice of the initial iterate, but is in general outperformed by direct search methods. Additional numerical simulations in the thesis reveal that direct search methods which approximate a gradient or Hessian find the most accurate solutions. We observe that the optimization problems considered in this thesis are more challenging than a set of noisy benchmark problems.

The thesis includes scientific contributions in addition to the results from the three papers.

Contents

Part I

Overview of Research

Chapter 1

Introduction

1.1 Stock-recruitment relationships

Parent-progeny (adult fish–juvenile) relationships are important for the management of a fishery. Since the seminal paper by Hjort [74] in 1914, understanding the parent-progeny relationship of fish populations has been an integral part of research on population dynamics of fish. The life history of fish usually consists of several stages such as eggs, larvae, juveniles, and adults, [122]. Following the transitions from one life stage to the next, survival may be restricted by stage-specific factors such as cannibalism, competition, and spread of disease. Time series of the number of fish often show large temporal variations. As observed in [74], e.g. Ch. 6, the number of fish in the youngest year-classes varies over time, and highly abundant year-classes may remain highly abundant through their lifetime. Hjort concluded that understanding the variations in the number of fish in the youngest year-class is central to understanding the fluctuations in the number of fish, a hypothesis which still has relevance, [129]. Functions that represent the parent-progeny relationship are commonly used to derive target or limit values for the fishing mortality and for medium-term forecasts (5 to 10 years), [106].

Traditionally, the relationship between adult fish and juveniles is described by stock-recruitment (SR) functions. The term recruitment refers to the introduction of young fish to the population, for example, by becoming catchable or mature, or reaching a certain age or length. The part of the stock that is mature enough to contribute to the reproduction process is called the spawning stock. The spawning stock size, i.e., number or biomass of the reproductive part of the population (e.g. mature females), is often obtained as a weighted sum of the number of adults in several adult age-classes, [70, Ch. 10; 106]. In the following, we refer to any link between the time series of the number of adults in several age-classes and recruitment as the SR relationship – analogous to a SR function.

The fisheries literature contains a wide spectrum of SR functions, [106; 116, Ch. 3]. Denote by S the spawning stock size. Let $\mathbb{R}_{>0}$ be the set of positive real numbers and $a, b, \gamma \in \mathbb{R}_{>0}$. Two popular SR functions introduced by Ricker [119] and Beverton and Holt [21] are described by (1.1)–(1.2), respectively. The Deriso (sometimes referred to as Deriso–Schnute) model (1.3) corresponds to the Ricker model for $\gamma \to \infty$ and to the Beverton–Holt model for $\gamma = 1$. The Deriso model was first published by Deriso [52]

1

and (1.3) corresponds to the version by Schnute [126].

$$\text{Beverton–Holt:} \quad r(S) = aS \cdot (1+bS)^{-1}, \tag{1.1}$$

$$\text{Ricker:} \quad r(S) = aS \cdot e^{-bS}, \tag{1.2}$$

$$\text{Deriso:} \quad r(S) = aS \cdot (1+bS/\gamma)^{-\gamma}. \tag{1.3}$$

As described by Schnute [126], the Deriso function (1.3) is strictly increasing in S for $\gamma \leq 1$. For $\gamma > 1$, the SR function is dome-shaped, i.e., the function has exactly one stationary point $S_m \geq 0$, is strictly increasing for $S < S_m$ and strictly decreasing for $S > S_m$. The distinct types of asymptotic behaviour of a SR function as $S \to \infty$ have in the fisheries literature been linked to food availability. For example, hybrid SR functions, which are dome-shaped in case of low food supply and saturating otherwise, have been proposed in [79, 109]. The Beverton–Holt, Ricker, and Deriso functions are illustrated in Figure 1.1. The four curves share the same values for a and b.

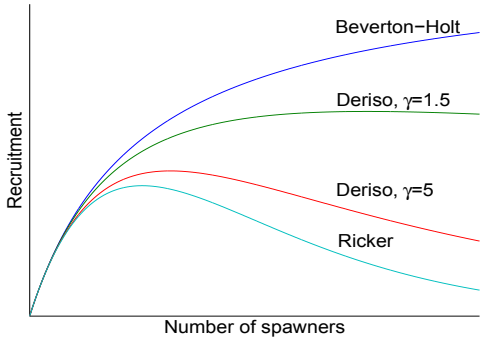


Figure 1.1: Traditional stock-recruitment functions.

Often, a SR function may be chosen based on a visual investigation of the available SR data, [134]. The data is used to estimate the uncertain parameters (a, b) of the chosen SR function. The estimation usually involves the task of minimizing a non-linear function, for example, the least squares error between the SR function and the data. Knowledge about the values of the parameters in the SR function is valuable for fishery management. For example, parameters in a SR function may inform about the resilience and productivity of a stock, [30] and references therein.

Hilborn and Walters [73, p. 241] note that the *"most important and generally most difficult problem in biological assessment of fisheries is the relationship between stock and recruitment"*. Data for recruitment and the stock often show high variability, as the estimates for recruitment and the spawning stock biomass of cod (*Gadus morhua*) in the North Sea, the Skagerrak, and the eastern Channel from [78, Tables 14.9 and 14.11a] and illustrated in Figure 1.2. Establishing a functional relationship between the mature stock and recruitment is nontrivial.

This is in part due to the high variability and uncertainty that characterize the data. Myers and Barrowman [103] investigated more than 350 stock and recruitment data sets to test the validity of the null hypothesis that recruitment is independent of the spawning stock. That their conclusion has been a subject of discussion in several papers

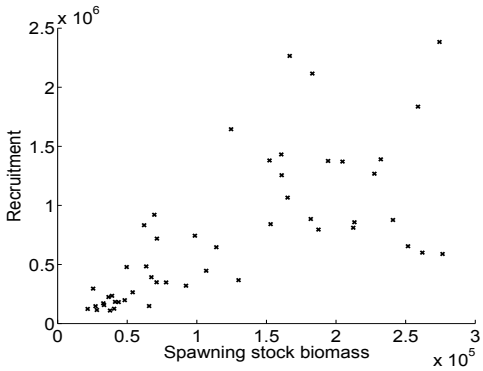


Figure 1.2: Data for recruitment and the spawning stock biomass for North Sea cod (*Gadus morhua*). Here, recruitment is defined as number of fish of age 1 and the spawning stock biomass is the total weight of mature fish in tons.

(references can be found in [70, Ch. 10]) shows that determining a SR function based solely on data is challenging. Furthermore, there is a drawback in specifying a SR function based on visual observations of trends in the data, [134]. A framework suitable for choosing a SR function is to define an infinite set of SR functions, which is adjusted to SR data in such a way that a plausible solution is guaranteed without unnecessary restrictions on the solution space, [134].

There are several other factors that can mask a SR function, or invalidate its existence. The link from the adult population to its progeny spans several life stages, and may be affected by temporal variations in the stage composition of a population. This motivates multi-stage population models which describe the evolution of fish from stock to recruitment, [30, 104, 110, 122]. In the literature, the evolution from spawning to recruitment including multiple early life stages has been considered for several fish populations and species, such as Northeast Arctic cod and European plaice, [25, 105].

Population dynamic models describe the changes in number (or biomass) of individuals in a population over time. A system of ordinary differential equations (ODE's) may represent our understanding of births, mortality, and transitions from one life-history stage to the next. For example, the average rate of deaths per individual and per time unit (called mortality rate) may be assumed to be constant with respect to time. Seen from a different angle, systems of difference equations may be considered for the variations in the number of individuals in several life stages, [34].

A multi-stage population dynamic model describes multiple time scale dynamics. The time scales and rates at which fish evolve and die are usually characteristic for the life stage of fish. For example, average mortality rates per year of prerecruits (eggs, larvae, and juveniles) may be several to several hundred times larger than for adults, [122, Ch. 3; 131].

Multi-stage population dynamic models may be used to investigate the type of SR relationship that evolves from a population consisting of several life stages. Recruitment may not be uniquely determined by the spawning stock size, and a SR function may not exist.

Touzeau and Gouzé [139] derive a system of differential equations with two time

scales which describes the fast evolution of a class of prerecruits and the concurrent dynamics of several adult age-classes. In general, the model describes a parent-progeny relationship which is not a SR function, [139]. Using singular perturbation theory, Touzeau and Gouzé prove that under suitable assumptions, trajectories of the slow-fast system approach and stay close to a graph of a Beverton–Holt function or another strictly increasing SR function. Two traditional sets of assumptions concerning the SR function are that the function is strictly increasing, or the function is dome-shaped. It is therefore surprising that all SR functions explained by the slow-fast population dynamic model are monotonic.

The evolution of trajectories of the differential equation in a neighbourhood of the graph of a SR function is no guarantee for the slow-fast population dynamic model to admit a SR function. The methodology used in [139] is not suitable to decide whether recruitment is uniquely determined given the spawning stock.

1.2 Parameter estimation for population dynamic models

The parameters in a population dynamic model for fish are generally highly uncertain and are estimated based on available data, for example, measurements of numbers of prerecruits and adults in several age-classes at several points in time. Following e.g. [6, 23, 29], we may consider an optimization problem that consists of finding a vector of parameters which minimizes the nonlinear least squares error between the solution of an initial value problem and the data points subject to bound constraints on the parameters.

Nonlinear least squares problems with Lipschitz continuously differentiable objective function are traditionally solved using the Gauss–Newton method, but also the Levenberg–Marquardt algorithm or sequential quadratic programming methods may be employed, [125, Sect. 2.3]. Weighted nonlinear least squares problems may, for example, be considered, when information about the variances of observational errors is available. Optimization methods as described in [68] are suitable for weighted nonlinear least squares problems with possibly large weights, and have been adapted for parameter estimation, [67], including for differential equations, [53]. Various approaches exist for addressing optimization problems with constraints which are explicitly stated in the problem formulation (e.g. [68, 108]).

A classical [18, Ch. 8; 121] and popular (e.g. [53, 125]) approach to formulating the task of estimating parameters in a differential equation as an optimization problem consists of reformulating it as a problem with only bound constraints. Every evaluation of the objective function requires a solution of the initial value problem. The approach has in the literature been referred to as initial value approach, [53], initial value problem approach, [23], or solution-based approach, [29]. If the initial value problem can be solved analytically, and the right-hand side of the differential equation is continuously differentiable with respect to the parameters and the states, then the analytical solution of the differential equation and the objective function are continuously differentiable functions of the parameters, as described e.g. in [71, Sect. I.14].

However, in many cases, including the stiff differential equations considered in this thesis, the initial value problems have to be solved numerically. The objective function is a function of the numerical solution of the differential equation. The accuracy and

smoothness of the objective function may vary with the algorithm used for the solution of the initial value problem. Most methods for the solution of differential equations adapt the step sizes dynamically, [92, Ch. 7]. The obtained approximation of the solution of an initial value problem may not be a continuously differentiable function of the parameters, [71, Sect. II.6]. In this case, the objective function is not Lipschitz continuously differentiable.

The objective function of an example of the optimization problem addressing parameters of a system of differential equations is illustrated in Figure 1.3. Due to the numerical solution of the differential equation by an algorithm with adaptive components and with finite precision arithmetic, we obtain a noisy objective function. The optimization problem and the stiff differential equation described in chapter 4 are approximated for 200 values of the parameter.

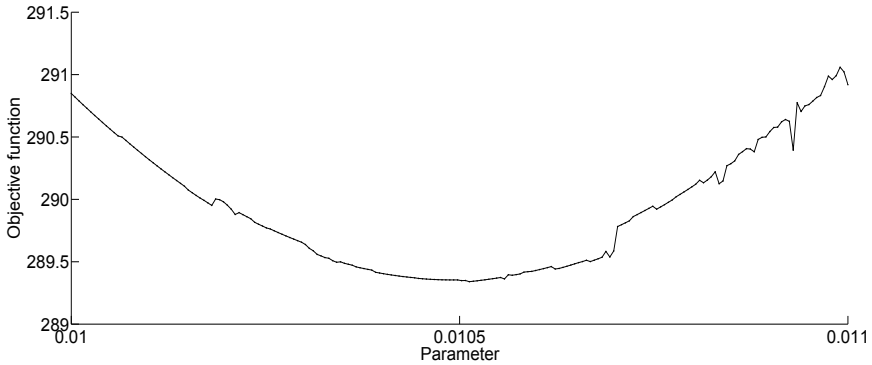


Figure 1.3: Example of a noisy objective function.

The solution of a differential equation may vary considerably with the values of its parameters. In some cases, the numerical (or analytical) solution of a differential equation is not defined for the entire time interval, [18, Ch. 8; 23]. Then, the objective function as a function of the numerical solution of the initial value problem is not defined for all parameters. The requirement that the objective function is defined is sometimes called 'hidden constraint', [38]. The latent constraint may only be detected when, for instance, the algorithm used to evaluate the objective function returns an error message. The feasible set of an optimization problem is unknown when constraints are only implicitly included in the problem formulation. The unpredictability of the success in evaluating the objective function is a challenge for optimization.

The properties of the differential equation and the method for the solution of initial value problems restrict the type of optimization method that can be employed for estimating the parameters in stiff systems of differential equations.

1.3 Scope of the research

This thesis considers challenges in quantifying and understanding the multiple time scale dynamics of multi-stage populations. We focus on issues related to fish population dynamics. However, the mathematical models and related questions are broader

in scope, and therefore applicable to other population systems, such as in terrestrial ecology.

1.3.1 Stock-recruitment relationships emergent from stage structured population models

We state conditions under which a SR function exists. Our approach consists of assuming the dynamics of a stage structured population to be described by a multi-stage population dynamic model and investigating the properties of the emergent SR relationship.

The first class of multi-stage population dynamic models considered in this thesis are systems of differential equations with two time scales such as the model by Touzeau and Gouzé [139]. We use a more generic approach and design a slow-fast population dynamic model that admits the Ricker, Beverton–Holt, and Deriso functions. We state the assumptions of the new slow-fast population dynamic model and relate them to the literature. To prove the existence of a SR function, we employ singular perturbation theory and results due to Fenichel [54], which will be shortly described in chapter 2.

The second class of multi-stage population dynamic models is a set of difference equations. We design a generic discrete-time population dynamic model that represents the dynamics of eggs, larvae, juveniles, and several adult age-classes. Our aim is to establish sufficient conditions for the existence of two types of parent-progeny functions, SR functions and recruitment as a function of the vector of the number of adults in several age-classes, referred to as adult-recruitment function.

In this thesis, we investigate the parent-progeny relationship, which is important from a management point of view. Recruitment may be affected by environmental and physical factors such as temperature, food availability or wind direction and intensity, as described in [73, Ch. 7; 110] and references therein. These mechanisms may act at several distinct temporal scales and indirectly through the spawning stock, [135]. Understanding the impact of environmental or physical factors on the SR relationship requires knowledge about the temporal variations in the parent-progeny relationship.

Deterministic models are suitable for addressing the SR relationship admitted by a multi-stage model. Furthermore, ordinary differential equation models may, for large populations such as fish, be justified from stochastic considerations, [101, Ch. 1]. The state of the art and challenges in predicting the SR relationship are reviewed in [135]. The article considers statistical questions and includes a discussion of factors that may influence the SR relationship.

As in [134], we avoid an a priori choice of the SR function. The use of a B-spline function in [134] may represent a broader range of SR functions than the SR function admitted by the new slow-fast population dynamic model. In comparison, we obtain a parent-progeny relationship, which may not be a SR function, from estimating parameters in a slow-fast population dynamic model.

1.3.2 Derivative-free optimization for population dynamic models

We also consider the task of estimating parameters for the new slow-fast population dynamic model, which is represented by a stiff system of ODEs. Our aim is to identify

optimization methods that are suitable for the nonlinear least squares problems with latent constraints and a noisy objective function.

Some methods for estimation of parameters for differential equations utilize a smooth (continuously differentiable) approximation of the objective function, as described in [27, 53]. However, this comes at the cost of obtaining a less accurate approximation of the objective function. We argue for the use of a class of optimization methods that is suitable for noisy problems with hidden constraints.

Derivative-free optimization methods have previously been suggested for noisy problems in general, [45], and some of the techniques have been designed with the aim of solving nonsmooth problems with hidden constraints, [10, 56]. We find that derivative-free methods in general are suitable for problems with implicitly stated constraints and for parameter estimation for stiff differential equations.

Two derivative-free optimization methods, the simplex-based Nelder–Mead algorithm [107] and the Hooke and Jeeves algorithm [76], have been employed for estimation of parameters in (discrete-time) population dynamic models, [114, 140]. These two optimization methods have also been used for parameter estimation in differential equations, [6, 27], and delay differential equations, [17, 31]. We consider optimization methods which have previously not been employed for parameter estimation for fish population models. The focus is on the performance of several classes of derivative-free optimization methods when estimating the parameters in a slow-fast population dynamic model.

Following a benchmark process introduced by Moré and Wild [99], we design a set of optimization problems and compare the performance of several derivative-free optimization methods based on traditional criteria for the choice of methods and based on criteria specific for parameter estimation for differential equations.

From the literature on derivative-free optimization, results of benchmark processes are available for analytical functions with bound constraints, [120]. A class of standard benchmark problems for derivative-free optimization contains a set of unconstrained problems with deterministically perturbed nonlinear least squares functions, [99]. The nonlinear least squares problems considered in this thesis have about the same dimension as the average dimension of the noisy benchmark problems. However, the evaluation of the objective functions considered in this thesis is based on numerical solution of differential equations. We evaluate the performance of a set of optimization methods for our time-dependent problems against results for the set of noisy benchmark problems. The aim is to test whether the choice of a derivative-free method for estimation of parameters in population dynamic models can rely on results from benchmark processes from the literature.

1.4 Outline of the thesis

The thesis includes two main parts. In the first part, we establish the background for the scientific research and the link between the interdisciplinary research problems. An aim of part I is to put the methods of choice and the key findings from three publications (A–C) into the context of the state of the art. The justification for the choice of methodology is strengthened, and the literature review is extended. Additional research questions that arise are answered in part I. The three publications, which contain new scientific

results, constitute part II of the thesis.

In chapter 2 of part I, terminology and methods used in the following chapters are introduced. The topics of the chapter are singular perturbation theory, models for age or stage structured populations, derivative-free optimization and numerical solution of stiff differential equations.

A framework for comparing general slow-fast population dynamic models to SR relationships is presented in paper A and described in chapter 3. In this chapter, the slow-fast population dynamic model introduced in paper A is derived from assumptions about the dynamics of prerecruits from the fisheries literature. We discuss the assumptions and applicability of the slow-fast population dynamic model in comparison to SR functions and other continuous-time multi-stage models.

Chapter 4 considers estimation of the parameters in the new slow-fast population dynamic model. We discuss several approaches to formulating the optimization problem addressing parameters of a system of differential equations. Formulation of the optimization problem involves the choice of the method for the solution of the stiff differential equations. We demonstrate that derivative-free optimization in combination with an initial value approach is suitable for solving the noisy optimization problems with latent constraints. A comparison of a set of derivative-free optimization methods for a set of time-dependent problems can be found in paper C. In chapter 4, we review the results from paper C, using a new set of numerical simulations. Furthermore, we consider a new topic, the importance of hidden constraints for the performance of the derivative-free optimization methods for parameter estimation for slow-fast population dynamic models.

In chapter 5, we address whether a general discrete-time multi-stage population model admits a SR function. Sufficient conditions for the existence of two parent-progeny functions are stated in manuscript B. In chapter 5, we demonstrate that results from manuscript B can be applied for a rich class of multi-stage models from the literature. Furthermore, we relate the results about emergent SR relationships from paper A to the ones from manuscript B.

The last chapter in part I includes the main conclusions from the scientific results and open questions for further research.

Chapter 2

Background

This chapter covers background material for the scientific contributions of this interdisciplinary thesis. Singular perturbation theory is introduced in section 2.1 as a tool for investigation of the slow-fast dynamics of populations. In section 2.2.1, we consider the assumptions underlying SR functions. Short introductions to discrete- and continuous-time models traditionally used for age and stage structured populations of fish can be found in sections 2.2.2 and 2.2.3, respectively. An example of a slow-fast population dynamic model is the model by Touzeau and Gouzé [139], which we consider in section 2.2.4.

The smoothness and accuracy of the objective functions of the optimization problems in this thesis depend on the numerical solution of the differential equation. Methods for the solution of stiff systems of ODEs are described in section 2.3. In section 2.4, we shortly describe three classes of algorithms which are interesting for noisy optimization problems, a simplex-based optimization method (in section 2.4.1), derivative-free trust-region methods, which build an interpolation model (in section 2.4.2), and directional direct-search methods (in section 2.4.3). The use of the term 'hidden constraint' in the derivative-free optimization literature is reviewed in section 2.4.4. In section 2.4.5, we describe why hidden constraints are a challenge for optimization, and consider derivative-free methods with convergence results for problems with hidden constraints.

In the following, all norms are Euclidean norms.

2.1 Singularly perturbed differential equations

The following description of singularly perturbed differential equations is based on [80, 81]. Let $k, l \in \mathbb{N}$, $f, g : \mathbb{R}^{k+l+1} \to \mathbb{R}$ and $0 < \varepsilon \ll 1$. For simplicity, we assume f, g to be C^∞ functions. The prime sign "′" denotes d/dT. Let $H_l = \{(x_1, \ldots, x_l) \in \mathbb{R}^l | x_l \geq 0\}$ and ∂H_l its boundary. Following [81], we say that $X \subset \mathbb{R}^{k+l}$ is a C^r, l-dimensional manifold with boundary if each point in X has a neighbourhood $U \cap X$ which is C^r diffeomorphic to $V \cap H_l$ for some open set V in $\mathbb{R}^l$. The subset of X that is C^r diffeomorphic to ∂H_q is the boundary ∂X of X.

We consider the singularly perturbed differential equation (S_ε), which describes slow and fast dynamics. Equation (S_ε) is formulated in terms of the slow time t. For $\varepsilon > 0$, (S_ε) is equivalent to (F_ε) for fast time $T = t/\varepsilon$. The vector $x \in \mathbb{R}^k$ consists of fast

variables and $y \in \mathbb{R}^l$ denotes the vector of variables that change at slow rates. Initial conditions for the two ODEs have the form $x(0) \in D \subset \mathbb{R}^k$ and $y(0) \in G \subset \mathbb{R}^l$.

In case of $\varepsilon \to 0$, (S_ε) becomes a differential algebraic equation (S_0) with algebraic constraint $f(x(t), y(t), 0) = 0$. Equation (S_0) with an initial condition of form $x(0) \in D$ is referred to as the reduced slow system. The limit of the system (F_ε) in terms of fast time T as $\varepsilon \to 0$ is given by (F_0). Here, y can be considered as a vector of constant parameters and $f(x(T), y(T), 0) = 0$ describes the set of fixed points (constant solutions) of (F_0). Assuming that the matrix $Df_x(x, y, 0)$ has k eigenvalues with negative real part, the fixed points are asymptotically stable fixed points of the differential equation $x'(T) = f(x(T), y(T), 0)$.

$$\begin{aligned} \varepsilon \dot{x}(t) &= f(x(t), y(t), \varepsilon) \qquad (S_\varepsilon) \\ \dot{y}(t) &= g(x(t), y(t), \varepsilon)\,. \end{aligned} \qquad \begin{aligned} x'(T) &= f(x(T), y(T), \varepsilon) \qquad (F_\varepsilon) \\ y'(T) &= \varepsilon g(x(T), y(T), \varepsilon)\,. \end{aligned}$$

$$\begin{aligned} 0 &= f(x(t), y(t), 0) \qquad (S_0) \\ \dot{y}(t) &= g(x(t), y(t), 0)\,. \end{aligned} \qquad \begin{aligned} x'(T) &= f(x(T), y(T), 0) \qquad (F_0) \\ y'(T) &= 0\,. \end{aligned}$$

Assume that there exists a continuous function $h : \mathbb{R}^l \to \mathbb{R}^k$ that defines isolated roots $h(y)$ of the equation $f(x, y, 0) = 0$, for all $y \in \mathbb{R}^l$. Then, the dynamics of the reduced slow system are described by ODE (2.1) with l, rather than $k + l$ variables.

$$\dot{y}(t) = g(h(y(t)), y(t), 0)\,. \tag{2.1}$$

We follow [142, p. 28-29] and define an invariant set for a solution $z(t, z_0, 0)$ of an initial value problem (F_0) with $z(0) = z_0$. A set S is called invariant under (F_0) if $z_0 = (x_0,\ y_0) \in S$ implies $z(t, z_0, 0) \in S$ for all $t \in \mathbb{R}$. The set of zeros of function f defines a l-dimensional manifold in $\mathbb{R}^{l+k}$ which is invariant under (F_0).

Tikhonov's theorem (see e.g. [141, Ch. 8.2]) states sufficient conditions for the convergence of the solution of an in general nonautonomous singularly perturbed differential equation to the solution of the reduced slow system as $\varepsilon \to 0$, in a bounded time interval. For $\varepsilon > 0$ and sufficiently small, the solution of (S_ε) may (possibly in a bounded time interval) be approximated by the solution of (S_0). Initially, the solution of (S_ε) may evolve very fast and the fast variables move considerably faster than the slow variables.

Results by Fenichel [54] include sufficient conditions for the existence of an analogue of the manifold invariant under (F_0), but for the case $\varepsilon > 0$ and sufficiently small. Solutions of (F_ε) starting on a manifold M_ε (possibly with boundary) that is locally invariant under the flow from (F_ε), remain on the manifold unless they leave the manifold via its boundary, [80]. M_ε is called the slow manifold. Under suitable assumptions, the slow manifold is the graph of a function h^ε, which gives the fast variables as a function of the slow variables. Furthermore, the results by Fenichel may be used to show that for $\varepsilon > 0$ and sufficiently small, solutions of the slow-fast system (F_ε) with suitable initial conditions, may converge exponentially towards the slow manifold.

Biological systems frequently evolve at two distinct time scales and singular perturbation theory may be applied for reduction of models in biology, [101]. As an example, slow-fast systems describing populations with several interacting organization levels (such as individual, population, ecosystem) have been approximated by the lower-dimensional ODE (2.1), as described in [16].

2.2 Models for the dynamics of stage structured populations

Characteristics of fish such as length, weight, mortality or fecundity may depend on age and the age structure of fish populations may vary considerably with time. In the following, the differences between fish of different age-classes and stages are assumed to be more important than the differences between individuals. Depending on the assumptions, the dynamics of age or stage structured populations may be described either by continuous-time or discrete-time models, [34, Ch. 8].

2.2.1 Assumptions of stock-recruitment functions

Several well-known SR functions may be derived from assumptions about egg production and mortality of prerecruits. Denote by $N_0(t)$ the number of prerecruits at time t. Mortality of prerecruits is a continuous-time process and egg production at time $t = 0$ is assumed to be a constant times the spawning stock size S. We obtain an initial value problem (2.2) with mortality rate m per individual.

$$\dot{N}_0(t) = -m(N_0(t), S) \cdot N_0(t), \quad N_0(0) = aS, \quad a > 0. \tag{2.2}$$

By solving an initial value problem of the form (2.2) and by substituting R for $N_0(T)$, $T > 0$, a SR function may be derived. Assuming the mortality rate to be linear in $N_0(t)$, the Beverton–Holt function is obtained, [21, Sect. II.6; 116, Ch. 3]. The Ricker function is based on the assumption that m is linear in S, where S denotes the spawning stock size at time of spawning, [116, Ch. 3; 119].

In general, a mortality rate that is a function of numbers of individuals is referred to as density-dependent and may be described by a nonlinear function. The term 'density-dependency' is explained by the fact that 'number' may stand for 'number per unit area', [22]. The Shepherd model (2.3) with $a, b, c > 0$ is a Beverton–Holt function in case of $c = 1$ and a dome-shaped SR function for $c > 1$. The SR function assumes mortality to be a power function of the number of eggs, [116, Ch. 3]. The generalized Ricker model (2.4) is obtained when assuming that the mortality rate is a power function of the spawning stock size, [116, Ch. 3].

$$r(S) = aS/(1 + bS^c). \tag{2.3}$$

$$r(S) = aS \cdot e^{-bS^\gamma}. \tag{2.4}$$

The generalized Ricker function has been introduced in [110] and [96].

2.2.2 Matrix models for stage structured populations

Matrix models are traditionally used to describe age structured and stage structured populations of fish, [116, p. 292 and Ch. 9]. The population is assumed to consist of discrete stages. Reproduction and transitions from a stage to the next are described as discrete-time processes.

We denote by $N_{i,t}$ the number of individuals in stage $i = 1, \ldots, n$ at time $t \in \mathbb{N}_0$ and by $\mathbf{N}_t$ the n-dimensional vector $(N_{1,t}, \ldots, N_{n,t})$. A population model, which describes

birth and death processes and stage transitions, is given by the matrix equation (2.5) with suitable initial values $\mathbf{N}_0 \in \mathbb{R}^n_{\geq 0}$.

$$\begin{pmatrix} N_{1,t} \\ \vdots \\ N_{n-1,t} \\ N_{n,t} \end{pmatrix} = \begin{pmatrix} \beta_1 & \dots & \beta_{n-1} & \beta_n \\ s_1 & & & 0 \\ & \ddots & & \\ 0 & & s_{n-1} & 0 \end{pmatrix} \begin{pmatrix} N_{1,t-1} \\ \vdots \\ N_{n-1,t-1} \\ N_{n,t-1} \end{pmatrix}. \tag{2.5}$$

The model description (2.5) may, for example, represent the dynamics of a population consisting of several age-classes projected from year $t-1$ to year t. The parameters $\beta_i \geq 0$ may denote the average number of eggs produced per spawner times the probability of surviving to reach age 1. The probability of surviving from age i to age $i+1$ is denoted by $s_i \in (0,1]$. The matrix in equation (2.5) with nonzero coefficients only in the first row and first subdiagonal is called a Leslie matrix and the model (2.5) is often referred to as Leslie matrix population model, after Leslie [91].

In general, all coefficients of the matrix in equation (2.5) may be non-zero and functions of $\mathbf{N}_t$ or of time. The linear difference equations are referred to as matrix population models. An introduction to matrix population models is given in [34].

As described in section 2.2.1, an assumption underlying several models for recruitment is that mortality of prerecruits is density-dependent, i.e., a function of numbers of individuals. In this case, survival s_0 of prerecruits to age 1 is a function of $\mathbf{N}_t$. Often, the coefficients of the population projection matrix are described as functions of a weighted sum of the number of individuals in several age-classes, [34, Ch. 16]. An example of a nonlinear matrix model is the population model described in [116, Ch. 7.4]. Survival of eggs to age-class 1 is assumed to be a function of egg production. Denote by $w_k \geq 0$ the number of eggs laid by individuals of age-class k. The population model described in [116, Ch. 7.4] is given by (2.6).

$$\begin{pmatrix} N_{1,t} \\ \vdots \\ N_{n-1,t} \\ N_{n,t} \end{pmatrix} = \begin{pmatrix} \beta_1(\mathbf{N}_t) & \dots & \beta_{n-1}(\mathbf{N}_t) & \beta_n(\mathbf{N}_t) \\ s_1 & & & 0 \\ & \ddots & & \\ 0 & & s_{n-1} & 0 \end{pmatrix} \begin{pmatrix} N_{1,t-1} \\ \vdots \\ N_{n-1,t-1} \\ N_{n,t-1} \end{pmatrix}. \tag{2.6}$$

Here, $\beta_k(\mathbf{N}_t) = w_k\, s_0(\sum_{k=1}^n w_k N_{k,t})$ is the number of eggs spawned and surviving to age 1, per individual of age k.

2.2.3 Ordinary differential equation models for stage structured populations

The dynamics of age structured populations may be represented by ODEs, as described e.g. in [33, 101]. We consider discrete age-classes $i \in \{1,\dots,n\}$. Denote by $N_i(t)$ the number of individuals in age-class i at time $t \in \mathbb{R}_{\geq 0}$. $\mathbf{N}(t)$ is the n-dimensional vector $(N_1(t),\dots,N_n(t))$. Mortality, recruitment and ageing are described in continuous time. The coefficient $\beta_i \geq 0$ represents the rate of recruitment to the population per individual in age-class i, such that $\int_{t_1}^{t_2} \sum_{i=1}^n \beta_i N_i(t)dt$ is the number of fish recruited to the population in time interval $[t_1,t_2] \subset \mathbb{R}_{\geq 0}$. The mortality rate per individual in class i is denoted by $m_i \geq 0$. The age progression rate per individual from class $i-1$ to class i

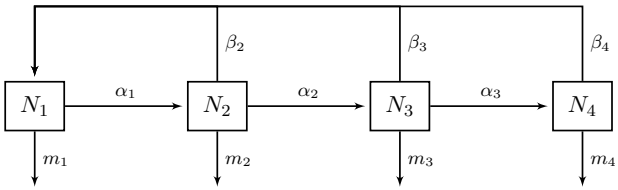


Figure 2.1: A schematic diagram for a linear ODE model for a population with four age-classes.

is $\alpha \geq 0$. A linear age structured population dynamic model is given by (2.7) and (2.8) for $i = 2, \ldots, n$ and an initial condition of form $\mathbf{N}(0) \in \mathbb{R}^n_{\geq 0}$.

$$\dot{N}_1(t) = \sum_{i=1}^{n} \beta_i N_i(t) - \alpha N_1(t) - m_1 N_1(t)\,, \tag{2.7}$$

$$\dot{N}_i(t) = \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t)\,. \tag{2.8}$$

The model (2.7)–(2.8) is illustrated in Figure 2.1.

The literature includes examples of nonlinear population dynamic models which assume at least one the rates m_i, α and β_i to be a function of $\mathbf{N}(t)$. In the fisheries literature, the parameters are usually assumed to be time-independent, [106]. However, the mortality rates and egg production may vary inter-annually, [73, Ch. 7].

2.2.4 The slow-fast population dynamic model by Touzeau and Gouzé

Touzeau and Gouzé [139] introduced a system of differential equations which describes the dynamics of numbers of prerecruits $N_0(t)$ and adults $N_i(t)$ of age-class $i = 1, \ldots, n$. In the following, we denote by $\mathbf{N}(t)$ the $(n+1)$-dimensional vector $(N_0(t), N_1(t), \ldots, N_n(t))$. The model by Touzeau and Gouzé is given by (2.9)–(2.10) and (2.8) for $i = 1, \ldots, n$. In comparison to the population dynamic model (2.7)–(2.8), the model by Touzeau and Gouzé includes the dynamics of prerecruits and describes the mortality rate of prerecruits as function (m) of $\mathbf{N}(t)$.

$$\dot{N}_0(t) = \frac{1}{\varepsilon} \sum_{i=1}^{n} f_i l_i N_i(t) - \alpha N_0(t) - \frac{1}{\varepsilon} m(\mathbf{N}(t)) N_0(t)\,, \tag{2.9}$$

$$m(\mathbf{N}(t)) = m_0 + \sum_{i=1}^{n} p_i N_i(t) + p_0 N_0(t)\,. \tag{2.10}$$

The rate of egg production at time t is $\sum_{i=1}^{n} f_i l_i N_i(t)$. Parameter $f_i \geq 0$ denotes the fraction of spawners in age-class i (for $i = 1, \ldots, n$). A spawner in age-class i is assumed to produce eggs at rate $l_i \geq 0$. Parameters $p_0, p_i \geq 0$ are coefficients for the degree of density-dependence attributed to fish of age i.

The parameter $0 < \varepsilon \ll 1$ represents the assumption that the rate of egg production and the mortality rate per prerecruit are larger than the rate of ageing and the mortality rate per adult. The model can be considered as an ODE which includes two distinct time-scales, a fast time T and a slow time $t = \varepsilon T$. The rates l_i, p_i, p_0 and m_0 are measured in the fast time-scale, while α and m_i, $i = 1, \ldots, n$, are measured in terms of

the slow time-scale $t = \varepsilon T$. For example, assume that the mortality rates per prerecruit and per day (m_0) and the mortality rate per adult and per year (m_i) are of order 10^{-1}. In this case, the mortality rate per prerecruit and per year (m_0/ε) is approximately 36.

The function $A^{TG} : \mathbb{R}^{n+1} \to \mathbb{R}$ defined by (2.11) represents the fast processes.

$$A^{TG}(\mathbf{N}(t)) = \sum_{i=1}^{n} f_i l_i N_i(t) - \left(m_0 + \sum_{i=1}^{n} p_i N_i(t) + p_0 N_0(t) \right) N_0(t) \,. \tag{2.11}$$

The superscript refers to the fact that the function is linked to the model by Touzeau and Gouzé.

As shown in [139], solutions of the initial value problem given by (2.8)–(2.10) and an initial condition $\mathbf{N}(0) \in \mathbb{R}^{n+1}_{\geq 0}$ are nonnegative and the ODE has a positive fixed point under assumptions I and II, which are defined as follows:

Assumption I: $\alpha\varepsilon + m_0 < \sum_{i=1}^{n} f_i l_i \cdot (\prod_{j=1}^{i} \alpha/(\alpha + m_j))$, i.e., the total rate of egg production in the average life time of an adult is for some $\mathbf{N} \in \mathbb{R}^{n+1}_{\geq 0}$ larger than the rate of decrease of the number of prerecruits.

Assumption II: $\exists i^* \in \{0, \ldots, n\}$ s.t. $p_{i^*} > 0$, i.e., the mortality rate of prerecruits is density-dependent.

By applying Tikhonov's theorem, the dynamics of the model by Touzeau and Gouzé are described as following in [139]. Solutions of the ODE are attracted to a neighbourhood of a surface in $\mathbb{R}^{n+1}$ which is described by $A^{TG}(\mathbf{N}) = 0$. The evolution towards the neighbourhood is dominated by (2.9). In a neighbourhood of the set of zeros of A^{TG}, the solutions of the population dynamic system can be approximated by the reduced slow system of (2.8)–(2.10).

The equation $A^{TG}(\mathbf{N}) = 0$ relates the number of prerecruits to the number of adults and stock to recruitment, as described by Touzeau and Gouzé [139]. Equation (2.12) defines the spawning stock size as the number of spawners, and recruitment as the transition rate to the youngest adult age-class.

$$R(t) = \alpha N_0(t) \quad \text{and} \quad S(t) = \sum_{i=1}^{n} f_i N_i(t) \,. \tag{2.12}$$

In general, recruitment is not uniquely defined given the spawning stock size. However, $A^{TG}(\mathbf{N}(t)) = 0$ is equivalent to,

$$0 = lS(t) - N_0(t)(m_0 + pS(t)/f + p_0 N_0(t)) \,,$$

if we assume:

Assumption 0: Let $p \geq 0$, f, $l > 0$. For all $i = 1, \ldots, n$, either $p_i = p$, $f_i = f$ and $l_i = l$ or $p_i = f_i = l_i = 0$. The class of adults consists of two homogeneous subclasses.

As shown in [139], solving the quadratic equation for $N_0(t)$ yields a strictly increasing function for $R(t) = \alpha N_0(t)$ given $S(t)$. In case of $p_0 = 0$, the SR function is given by (2.13).

$$R(t) = \frac{\alpha l S(t)}{m_0 + \frac{p}{f} S(t)} \,. \tag{2.13}$$

If we require $\alpha, m_0, p > 0$, we obtain a Beverton–Holt function as defined in section 1.1.

Tikhonov's theorem may explain that trajectories of the model by Touzeau and Gouzé approach, and remain in a neighbourhood of the graph of a function that links recruitment to numbers of adults. Recruitment may be approximated by a function of the number of adults if ε is sufficiently small and t sufficiently large. A dome-shaped SR function cannot be explained by the population dynamic model by Touzeau and Gouzé, [139].

2.3 Numerical solution of stiff differential equations

Consider the initial value problem (2.14) with solution $\mathbf{y} : \mathbb{R} \to \mathbb{R}^n$. Let $m, n, r \in \mathbb{N}$. For simplicity, we denote by $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ a function that is continuously differentiable with respect to $\mathbf{y}(t) \in \mathbb{R}^n$ and the parameters $\theta \in \mathbb{R}^m$.

$$\dot{\mathbf{y}}(t) = g(\mathbf{y}(t); \theta) \text{ for } t > 0\,, \quad \mathbf{y}(0) = \mathbf{d}_0\,. \tag{2.14}$$

We consider stiff systems of differential equations with the property that slight perturbations of a solution of a differential equation at any time cause variations which are considerably faster than the variations in the solution of the differential equation, [92, p. 167]. A slow-fast differential equation with solutions that exponentially converge towards a slow manifold is an example of a stiff differential equation. Stiffness of a differential equation may be detected when the stiffness ratio is large. Following [115], we define the stiffness ratio as the ratio between the smallest value of the real part of a negative eigenvalue of the Jacobian $\partial g(\mathbf{y}(t); \theta)/\partial \mathbf{y}$ and the largest value of the real part of a negative eigenvalue of the Jacobian.

The description of numerical methods for problem (2.14) is based on [92] and [115]. In general, a numerical solution of an initial value problem (2.14) on interval $[0, T]$ is approximated successively at times $t_j = t_{j-1} + h_j$ with step size $h_j > 0$ for $j = 1, \ldots, N$ and such that $t_0 = 0$ and $t_N = T$. For reasons for simplicity, we consider $h_j = h$, $j = 1, \ldots, N$. Denote by η_j an approximation of the solution of the initial value problem at t_j, $j = 1, \ldots, N$.

Some numerical methods for solution of (2.14) require small step sizes for the approximation of rapid transients and may involve large round-off errors and high computational costs. A numerical method for approximation of the solution of the test problem (2.15) is said to be absolutely stable for $h > 0$ and $\lambda \in \mathbb{C}$ if we have $\lim_{N\to\infty} |\eta_N| \to 0$, for the numerical solution $\{\eta_j\}_{0 \le j \le N}$ of problem (2.15) obtained by the numerical method using fixed step size $h > 0$.

$$\dot{y}(t) = \lambda y(t)\,, \;\; t > 0\,, \;\; y(0) = 1\,. \tag{2.15}$$

The region of absolute stability of a method is the set of $z = h\lambda \in \mathbb{C}$ for which the method is absolutely stable. The one-step error is the difference between η_{j+1} and the solution of the differential equation through η_j and possibly further points η_{j-1}, …, η_{j+1-r} used for computation of η_{j+1}. If a method is not absolutely stable for λ and h, a one-step error introduced by a numerical method for solving (2.15) at step size h may grow exponentially with the number of time steps, as explained e.g. in [92, Ch. 7]. If $\{z \in \mathbb{C} \setminus \{0\}$ with $|\arg(-z)| < \alpha\}$ is contained in the region of absolute stability of a method, the method is called A(α)-stable. In general, A(α)-stability is one characteristic of methods that are suitable to solve stiff problems. Only implicit methods,

which use an implicit formula to obtain the numerical solutions η_{j+1}, are A(α)-stable. For stiff differential equations, it is generally computationally more efficient to employ implicit methods.

Examples of classes of implicit methods are implicit Runge–Kutta methods, linearly implicit methods, such as Rosenbrock and W-methods, and backward differentiation formula (BDF) methods and the closely related family of numerical differentiation formula (NDF) methods.

Implicit Runge–Kutta methods use evaluations of the function g at intermediate steps between t_j and t_{j+1} to obtain an approximation of the solution of the ODE at time t_{j+1}. An r-stage Runge–Kutta method is characterized by equations (2.16)–(2.17).

$$\eta_{j+1} = \eta_j + h\sum_{i=1}^{r} b_i k_i, \tag{2.16}$$

$$\text{with } k_i = g(\eta_j + h\sum_{l=1}^{r} a_{il}k_l)\,, \quad i = 1,\dots,r. \tag{2.17}$$

A fully implicit Runge–Kutta method with $a_{il} \neq 0$ for all $i = 1,\dots,r$, $l = 1,\dots,r$, requires solution of a system of rn nonlinear equations in every step. Stability properties of implicit Runge–Kutta methods are e.g. topic of [72, Sect. IV.3].

For Rosenbrock methods, computation of η_{j+1} involves only solution of linear systems. Denote by $\mathbb{I}$ the identity matrix and by $J_j = \partial g(\eta_j)/\partial y$ a Jacobian matrix. An r-stage Rosenbrock method as suggested in [145] is given by (2.16) with k_i, $i = 1,\dots,r$, such that (2.18) holds.

$$(\mathbb{I} - hd_{ii}J_j)k_i = g(\eta_j + h\sum_{l=1}^{i-1} a_{il}k_l + hJ_j\sum_{l=1}^{i-1} d_{il}k_l)\,. \tag{2.18}$$

Rosenbrock methods require evaluations of the Jacobian at each step. Instead of the Jacobian, W-methods [132] employ an arbitrary real square matrix A with the property that $W(h,d_{ii},A) = (\mathbb{I} - hd_{ii}A)$ is invertible.

Backward differentiation formulas (BDF) methods are linear multi-step methods. As described e.g. in [71, Sect. III.1], using backward differences,

$$\nabla\eta_j = \eta_j - \eta_{j-1}\,, \quad \nabla^{i+1}\eta_j = \nabla^i\eta_j - \nabla^i\eta_{j-1}\,, \quad i = 1,\dots,r\,,$$

the BDF formulas are given by (2.19) with $\kappa = 0$.

$$\sum_{i=1}^{r} \frac{1}{i}\nabla^i\eta_{j+1} - \kappa\nabla^{r+1}\eta_{j+1} = hg(\eta_{j+1})\,. \tag{2.19}$$

Numerical differentiation formulas (NDF) methods have the form (2.19), where $\frac{1}{r+1}\nabla^{r+1}\eta_{j+1}$ is an approximation for the leading term in the one-step error of a BDF method. The value of κ can be chosen such that the accuracy of a NDF method is improved in comparison to a BDF method, as shown by Shampine and Reichelt [128].

A linear r-step method computes η_{j+1} using $\eta_j, \dots, \eta_{j+1-r}$. Runge–Kutta methods, Rosenbrock methods, and W-methods are one-step methods, i.e., the integration scheme for η_{j+1} only involves η_j and no further previous values. One-step methods

are self-starting, while a multi-step method needs to obtain the approximations of the solution at times $t_1, \ldots, t_{r-1}$ by some other numerical method. However, if the values of function g at $t_j, \ldots, t_{j-r}$ are stored, they can be reused when approximating η_{j+1} by a multi-step method. This means fewer function evaluations per step than for one-step methods. In each step, implicit Runge–Kutta methods, BDF and NDF methods require solutions of nonlinear equations.

Implementations of methods for the solution of ODEs may adapt the step size and integration scheme to find a numerical solution for which an estimate $\mathbf{e}$ for the one-step error is bounded from above by a given tolerance. As an example, several ODE solvers implemented in MATLAB use inequality (2.20) for the i-th element of the vector $\mathbf{e}$ and the i-th element of $\mathbf{y}(t_{j+1})$, $i = 1, \ldots, n$, [128].

$$|e_i| \leq tol|y_i(t_{j+1})| + Tol_i\,. \tag{2.20}$$

Here, $tol > 0$ and $Tol_i > 0$ denote a relative and an absolute error tolerance, respectively. If inequality (2.20) fails to hold, a new approximation of $\mathbf{y}(t_{j+1})$ may be computed using a reduced step size. Numerical methods such as BDF and NDF methods may also vary r based on an estimate for the one-step error, [128]. For a method of order r, the one-step error is $\mathcal{O}(h^{r+1})$.

Due to the error control, the step size and the order of a method vary nonsmoothly with a parameter θ in the differential equation. In this case, the numerical solution of the differential equation is a nonsmooth or discontinuous function of the parameter, [18, Ch. 8; 63; 71, Sect. II.6].

2.4 Derivative-free optimization

We consider the optimization problem

$$\min_{x \in \mathbb{R}^n} f(x) \tag{2.21}$$

with f bounded from below in $\mathbb{R}^n$ and continuously differentiable with Lipschitz continuous gradient in an open set, which contains the level set $\mathcal{L} = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$. The aim is to find a local solution of a constrained optimization problem, a point $x^* \in \mathbb{R}^n$ with a neighbourhood $\mathcal{N}$ such that $f(x) \geq f(x^*)$ for all $x \in \mathcal{N}$.

In many cases, optimization problems are solved iteratively. Given an initial iterate x_0, an optimization method iteratively tries to replace the point x_k in iteration k by a new iterate with a smaller objective function value. The iterative search of derivative-free optimization methods (as described e.g. in [45]) is guided by evaluations of the objective function at sets of sample points or by trust-region models obtained by evaluating the objective function at sample points. Derivative-free methods control the geometry of the trial points and require no knowledge about derivatives or estimate them explicitly, for example, by finite differences. Avoiding the use of derivatives comes at the cost of often slower convergence compared to gradient-based optimization methods, [45]. An introduction to derivative-free methods is given in [45].

We refer to an optimization algorithm as globally convergent if it computes, for an arbitrary starting point, a sequence of iterates that has a limiting point which fulfils first-order necessary conditions for optimality. For the unconstrained problem, this

means that the algorithm builds a sequence of iterates x_k, $k = 0,1,2,\ldots$ such that a subsequence $\mathcal{K}$ of iterations exists with $\lim_{k\in\mathcal{K},k\to\infty} \|\nabla f(x_k)\| = 0$.

For simplicity, derivative-free methods are described for the unconstrained optimization problem (2.21). For some methods, these assumptions are stronger than the minimum requirements on the objective function needed to prove global convergence.

2

2.4.1 A simplex-based optimization method

The Nelder–Mead algorithm [107] iteratively improves a simplex. Let $\lambda^{(i)} \in \mathbb{R}$, $i = 0,\ldots,n$. A set of points $y^0,\ldots,y^n \in \mathbb{R}^n$ is affinely independent if the system of equations $\sum_{i=0}^n \lambda^{(i)} y^i = 0$, $\sum_{i=0}^n \lambda^{(i)} = 0$ has a unique solution $\lambda^{(i)} = 0$, $i = 0,\ldots,n$. The smallest convex set containing $n+1$ affinely independent vectors $y^0,\ldots,y^n \in \mathbb{R}^n$ is a simplex of dimension n.

The vertices $y^0,\ldots,y^n$ of the simplex are points in $\mathbb{R}^n$ at which the objective function has been evaluated. In each iteration, the vertices $y^0,\ldots,y^n$ of the simplex are ordered such that $f(y^0) \le f(y^1) \le \cdots \le f(y^n)$. The algorithm tries to replace the point with the highest objective function value by a point $y = y^n + \delta(y^c - y^n)$ on the line through y^n and the centroid $y^c = \sum_{i=1}^n y^i/n$ of $y^1,\ldots,y^n$. The parameter δ takes one of the values $-1 < \delta^{ic} < 0 < \delta^{oc} < \delta^r < \delta^e$, which correspond to an inside or outside contraction, reflection or expansion of the simplex. If neither a reflection nor a contraction of the simplex gives a point with lower objective function value than y^n, a shrink of the simplex is performed. In this case, the shape of the simplex remains unchanged, but its volume is decreased. All vertices y^i of the simplex except for y^0 are replaced by $y^0 + \gamma^s(y^i - y^0)$, where $\gamma^s \in (0,1)$. A simplex and the simplex operations are illustrated in Figure 2.2.

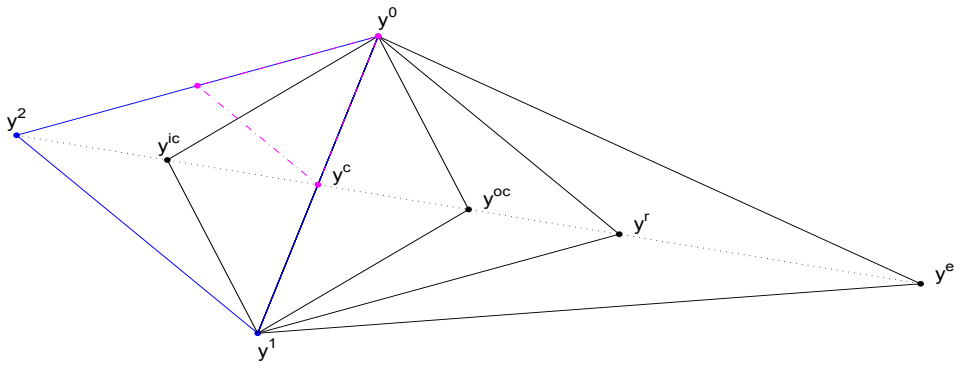


Figure 2.2: A simplex with vertices y^0, y^1, y^2 and its reflection (with vertex y^r), expansion (with vertex y^e), inside contraction (with vertex y^{ic}) and outside contraction (with vertex y^{oc}). The pink simplex is a shrink of the original simplex. Here, we follow the standard choice $\delta^{ic} = -\frac{1}{2}, \delta^{oc} = \frac{1}{2}, \delta^r = 1$, $\delta^e = 2$ and $\gamma = \frac{1}{2}$.

The simplices generated by a Nelder–Mead algorithm can capture local curvature information of the objective function, e.g. [45, Ch. 8]. Characteristic for the Nelder–Mead simplex method is its ability to often reduce objective functions well within few function evaluations, albeit global convergence is not guaranteed in general, since the simplices may become arbitrarily flat, [146]. The Nelder–Mead algorithm is the

most popular simplicial direct search method and references to further simplicial direct search methods can, for example, be found in [45, Ch. 8; 146].

2.4.2 Model-based derivative-free methods

A further class of derivative-free methods are trust-region methods that build a model function by evaluating the objective function at several sufficiently scattered sample points. The model function $m_k : \mathbb{R}^n \to \mathbb{R}$ approximates the objective function in the trust-region

$$B(x_k, \Delta_k) = \{x \in \mathbb{R}^n | \|x - x_k\| \leq \Delta_k\},$$

with the trust-region radius $\Delta_k > 0$. The new iterate $x_k + s$ is defined by a (possibly approximate) solution of the subproblem

$$\min_{s \in B(0,\Delta_k)} m_k(x_k + s).$$

Global convergence to points that fulfil first- or second-order necessary conditions for optimality can be ensured by the frameworks introduced by Conn, Scheinberg, and Vicente [44; 45, Ch. 10]. Lipschitz continuously differentiable model functions that satisfy (2.22) for positive constants κ_1, κ_2 and for all $s \in B(0, \Delta)$, are called fully linear [42] and approximate the objective function sufficiently well for first-order global convergence, [40].

$$\|f(x+s) - m(x+s)\| \leq \kappa_1 \Delta^2, \quad \|\nabla f(x+s) - \nabla m(x+s)\| \leq \kappa_2 \Delta. \tag{2.22}$$

The quality of the model function is ensured by control of the geometry (in the sense of well-poisedness, [42, 43]) of the sample points used to build a model.

An example of model functions are quadratic models of the form (2.23), where $c_k \in \mathbb{R}$, $g_k \in \mathbb{R}^n$ and $H_k \in \mathbb{R}^{n \times n}$.

$$m_k(x) = c_k + g_k^T x + \frac{1}{2} x^T H_k x. \tag{2.23}$$

In order to obtain a full quadratic model with symmetric matrix H_k by interpolation, the objective function is evaluated at $(n+1)(n+2)/2$ points. The quality of the approximation of the objective function by the model function depends on the Lipschitz constant of ∇f, the number and the geometrical properties of the interpolation points and the largest distance of two points in the interpolation set, [42, 43].

Powell's methods [111–113] are examples of derivative-free trust-region methods based on quadratic interpolation models. Per default, $2n+1$ interpolation points are utilized to obtain c_k, g_k and the diagonal elements of H_k. Models are built by minimization of the Frobenius norm of the difference between the Hessian matrix H_k in the new model and the Hessian matrix in the previous model. If f is quadratic, then the norm of the difference between H_k and the Hessian of f is nonincreasing with updates of the Hessian, [112]. Powell's method NEWUOA has been designed to be efficient in the number of function evaluations and in routine work necessary for updating the model function, [113].

By controlling the geometry of the sample points, polynomial models can be guaranteed to be fully linear and globally convergent derivative-free trust-region methods may be designed, [45]. For global convergence, the model is not required to be fully linear in every iteration, but when the gradient of the model function is small, [124]. Depending on the number p of sample points used, quadratic models may be obtained from a sample set of trial points by interpolation (for $p = q = (n+1)(n+2)/2$), regression (for $p > q$) or minimum norm interpolation (for $p < q$). For suitably chosen interpolation sets, an upper bound for the error between the gradient of f and the gradient g_k of an underdetermined quadratic model function is linear in the norm of the Hessian H_k, [45, Theorem 5.4]. For minimum Frobenius norm models [41], the Frobenius norm of matrix H_k is as small as possible.

A further example of model functions are radial basis functions, possibly with polynomial tail $p : \mathbb{R}^n \to \mathbb{R}$, given by (2.24), where $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}$, $\lambda_i \in \mathbb{R}$, $i = 0, \ldots, p$. The model function interpolates $y^0, \ldots, y^p \in \mathbb{R}^n$.

$$m_k(x_k + s) = \sum_{i=0}^{p} \lambda_i \phi(\|s - y^i\|) + p(s)\,. \tag{2.24}$$

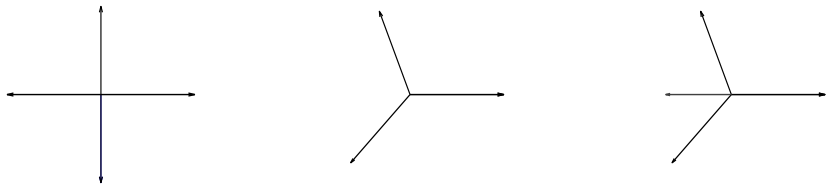
An example of a radial basis function is the cubic function $\phi(r) = r^3$. Radial basis functions can capture curvature information of the objective function and are suitable for the approximation of nonconvex objective functions, since they may be multimodal, [143, 144]. Radial basis functions can be built using at least $n+1$ sample points compared to $(n+2)(n+1)/2$ function evaluations necessary to interpolate a full quadratic model, [143]. The number of sample points interpolated is flexible and up to $3n$ sample points may be utilized, [143]. The global convergence of optimization by radial basis functions in trust-regions has been shown by Wild and Shoemaker [144].

2.4.3 Directional direct search methods

Frameworks for globally convergent directional direct search methods have been developed by Conn, Scheinberg and Vicente [45, Ch. 7] and Kolda, Lewis and Torczon [86]. Section 2.4.3 is based on [45].

Directional direct search methods are based on search along descent directions for the objective function at the current iterate. Following [45, p. 15], we say that a descent direction for f at x is a direction $d \in \mathbb{R}^n$ such that there exists an $\bar{\alpha} > 0$ with $f(x + \alpha d) < f(x)$ for all $\alpha \in (0, \bar{\alpha}]$. A vector d with $-\nabla f(x)^T d > 0$ is a descent direction of f at x, e.g. [108, Ch. 2]. In comparison to traditional line search methods, directional direct search methods employ sets of directions, which are positive spanning sets.

Let $\lambda^{(1)}, \ldots, \lambda^{(p)} \geq 0$. A set $D = \{d^{(1)}, \ldots, d^{(p)}\}$ of $p \geq n+1$ vectors in $\mathbb{R}^n$ is called a positive spanning set or generating set in $\mathbb{R}^n$ if every $v \in \mathbb{R}^n$ can be written as positive combination $v = \sum_{i=1}^{p} \lambda^{(i)} d^{(i)}$ of vectors in D. A positive basis is a positive spanning set for which no proper subset positively spans $\mathbb{R}^n$. The sets $D_{\oplus} = \{(1,0), (-1,0), (0,1), (0,-1)\}$ and $D_1 = \{(1,0), (-1,0), (-1,-1)\}$ are examples of a positive basis in $\mathbb{R}^2$. An approach for the generation of positive bases relies on the fact that given a positive basis $D = \{d^{(1)}, \ldots, d^{(p)}\}$ in $\mathbb{R}^n$ and a nonsingular matrix $W \in \mathbb{R}^{n \times n}$, the set $\{Wd^{(1)}, \ldots, Wd^{(p)}\}$ is a positive basis (a proof can e.g. be found in [45, Ch. 2.1]). Three examples of positive spanning sets are illustrated in Figure 2.3.



(a) Positive basis $D_{\oplus}$.

(b) A positive basis obtained from multiplication of the vectors in D_1 with a nonsingular matrix.

(c) A positive spanning set, which is not a positive basis.

Figure 2.3: Three examples of a positive spanning set.

The following property of positive spanning sets is central to convergence theory of directional direct search methods. Let $D = \{d^{(1)}, \ldots, d^{(p)}\}$, with $d^{(i)} \neq 0$ for $i = 1, \ldots, p$, be a positive spanning set in $\mathbb{R}^n$. Then, for every nonzero vector $w \in \mathbb{R}^n$, there exists $d \in D$ s.t. $w^T d > 0$. A proof can be found in [45, Ch. 2.2]. Consequently, for the continuously differentiable function f and $x \in \mathbb{R}^n$ with $\nabla f(x) \neq 0$, every positive spanning set D contains $d \in \mathbb{R}^n$ such that $-\nabla f(x)^T d > 0$.

Directional direct search methods iteratively try to find a new incumbent x_{k+1} with smaller objective function value than the current iterate x_k by evaluation of the objective function at the set of poll points (2.25) defined by a positive spanning set D_k in $\mathbb{R}^n$.

$$P_k = \{x_k + \alpha_k d_k | d_k \in D_k\}. \tag{2.25}$$

The step length is controlled by step size parameter $\alpha_k > 0$. Unless $\nabla f(x_k) = 0$, the set D_k contains at least one descent direction, and for a sufficiently small step size, the algorithm finds a point with lower objective function value than the current iterate. A poll point with (sufficiently) smaller objective function value than $f(x_k)$ may be accepted. In this case, $x_{k+1} = x_k + \alpha_k d_k$ and the step size parameter might be increased. If the decrease condition is not satisfied at any of the poll points, then α_k is decreased.

Every iteration of a directional direct–search method may consist of two phases, a search step and a poll step, [26]. The search step is optional and consists of evaluations of the objective function at a finite number of sample points. Clever choices of the directions in this step may speed up convergence. If the search step is unsuccessful, the poll step is performed. It ensures global convergence and is the main subject of the following paragraphs.

Generalized pattern search methods [9, 138] accept new steps if a simple decrease condition $f(x_{k+1}) < f(x_k)$ is satisfied. These methods build trial points on integer lattices translated by x_0. Denote by $D \in \mathbb{R}^{n \times n_D}$ a matrix whose column vectors are the vectors in a finite set of n_D poll directions. Poll and search points in iteration k are in the conceptual mesh M_k given by (2.26).

$$M_k = \{x_k + \alpha_k D z | z \in \mathbb{N}^{n_D}\}. \tag{2.26}$$

For generalized pattern search methods, every positive spanning set utilized is a subset of the finite set of column vectors of D.

Generating set search [86] is a class of directional direct search methods. It includes generalized pattern search methods and algorithms introduced in [62, 95], which require that a sufficient decrease condition holds for new iterates. Generating set search methods can utilize an infinite set of positive spanning sets if the positive bases are such that the cosine measure $\kappa(D_k)$, the angle between any $v \in \mathbb{R}^n$ and the vector in D_k with the smallest angle to v, is bounded away from zero and $\beta_{max} \geq \|d_k\|$ for $\beta_{max} > 0$, for all $d_k \in D_k$ and $k \geq 0$, [86].

We made two simplifications in the discussion. Stronger convergence results are available for subsets of the directional direct search methods, [1, 4, 9, 10, 138] and the assumptions about function f may be weakened for some of the directional direct search methods, [9; 10; 45, Ch. 7; 86].

Mesh adaptive direct search

Mesh adaptive direct search methods, introduced by Audet and Dennis [10], are direct search methods which evaluate the objective function on a mesh (2.26) defined by the finite set D and mesh size parameter α_k^m. In comparison to generalized pattern search methods, the set of normalized poll directions is allowed to become asymptotically dense in the unit sphere. To achieve this, the sets of poll directions and the lengths of poll steps may be chosen more freely than for generalized pattern search methods. The length of a step is bounded from above by the poll size parameter $\alpha_k^p \geq \alpha_k^m$ such that $\alpha_k^m \|d_k\| \leq \alpha_k^p \max_{d \in D} \|d\|$ for all poll directions d_k and iterations k.

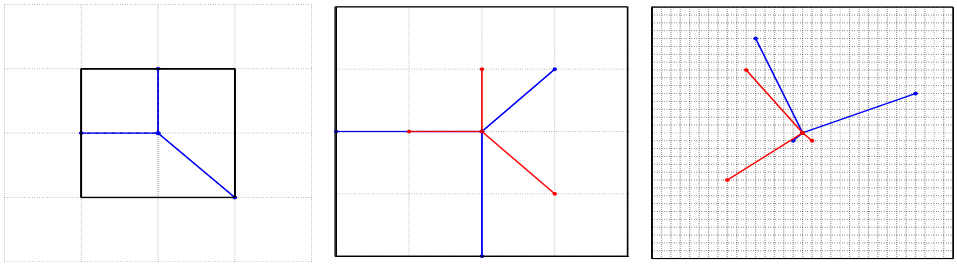
For mesh adaptive direct search methods, convergence results for nonsmooth functions exist, [10]. Furthermore, convergence of a subsequence of iterates to points that fulfil second-order necessary conditions for optimality, under suitable assumptions, has been shown, [1].

Figure 2.4 presents three examples of a mesh generated by $D_{\oplus}$ and sets of poll directions used by the mesh adaptive direct search algorithm described in [13]. In Figure 2.4 (a), the poll size parameter is equal to the mesh size parameter and the poll points could also have been generated by a pattern search method. The poll directions in Figures 2.4 (b) and (c) are generated by mesh adaptive direct search. The number of possible poll directions increases with increasing ratio between poll size parameter and mesh size parameter.

Efficiency

In the following, we consider several strategies for choices of the search step, the choice of the positive bases and the ordering of the poll steps, which may speed up convergence of directional direct search methods.

In the search step, model functions, which approximate the objective function, may be minimized in a trust-region, [46]. The approach described by Custódio, Rocha and Vicente [46] consists of computing a quadratic (minimum Frobenius norm or regression) model when the objective function value at more than $n+1$ sample points is available from previous function evaluations. In comparison to derivative-free trust-region methods, the convergence of the directional direct search method does not rely on the accuracy of the model function, and it is not necessary to control the geometry of the sample points, [46]. However, the trial points obtained from minimization of the



(a) For $\alpha_k^m = \alpha_k^p$. (b) For $\alpha_k^m = \frac{1}{2}\alpha_k^p$. (c) For $\alpha_k^m = \frac{1}{16}\alpha_k^p$.

Figure 2.4: Three examples of a mesh (the intersection of black, dashed lines) and poll directions (in colour) as used by mesh adaptive direct search. Poll points have to lie on the mesh and within the square indicated by black solid lines.

subproblem need to be projected to the mesh, [39]. A search step based on minimization of quadratic model functions and ordering of the poll points of a directional direct search methods by the value of the model function may speed up convergence, [39, 46].

Custódio and Vicente [48] suggested an efficient generalized pattern search algorithm that utilizes simplex derivatives, which are defined as follows. Let $Y = \{y^0, \ldots, y^n\}$ be a set of $n+1$ affinely independent vectors. The simplex gradient is the vector $g \in \mathbb{R}^n$ for which the linear model $m(x) = f(y^0) + g^T x$ interpolates f at the points in Y. Denote by $H \in \mathbb{R}^{n\times n}$ a symmetric matrix. If the set $Y = \{y^0, \ldots, y^p\}$ for $p = (n+1)(n+2)/2 - 1$ is such that the quadratic model

$$m(x) = f(y^0) + g^T x + \frac{1}{2} x^T H x$$

which interpolates Y is uniquely defined, the simplex Hessian of f at y^0 is the matrix H. The simplex gradient and Hessian may also be defined based on a set of $p \neq n$ and $p \neq (n+1)(n+2)/2 - 1$ sample points, [48]. At least after every unsuccessful iteration, a generalized pattern search method is guaranteed to have a suitable sample set for computing a simplex gradient without additional function evaluations, [48]. Considering the poll directions in the order of their angle to the negative of the simplex gradient, starting with the smallest angle, the number of function evaluations used by a generalized pattern search method may be considerably reduced, [48].

A generalized set search method introduced in [60] collects average curvature information in form of finite differences. Let $h, k > 0$. The matrix $R \in \mathbb{R}^{n\times n}$ with elements $R_{i,j}$, $i, j = 1, \ldots, n$, given by,

$$R_{i,j} = \frac{f(x + hq_i + kq_j) - f(x + hq_i) - f(x + kq_j) + f(x)}{kh},$$

is an approximation of a Hessian with respect to the coordinate system rotated by an orthogonal matrix $Q = [q_1, \ldots, q_n] \in \mathbb{R}^{n\times n}$. The algorithms by Frimannslund and Steihaug [60, 61] compute the matrix R using information from several iterations. The poll directions are ordered with the aim of calculating the curvature information based on few functional evaluations additional to the previously evaluated points. The curvature information is utilized to define sets of poll directions of the form $[V - V]$, for a

nonsingular and orthogonal matrix $V \in \mathbb{R}^{n\times n}$ that contains the eigenvectors of the approximation QRQ^T of the Hessian matrix. The use of average curvature information is suitable for noisy objective functions, [60]. Due to the choice of poll directions, the efficiency of a direct search method may be improved, [60], and, under suitable assumptions, convergence to points which satisfy second-order necessary optimality conditions may be guaranteed, [4].

2.4.4 Hidden constraints

The use of the term hidden constraint in the literature of derivative-free optimization is sometimes attributed to Choi and Kelley [38], who call *"the requirement that the objective be defined a 'hidden constraint', because there is no a priori way to tell if a point is feasible without attempting to evaluate the function"*. Such a constraint may also be called virtual constraint, [41].

Hidden constraints may be considered as examples of unrelaxable constraints (which have to be satisfied by all iterates, for example, since the user requires it), [11, 45]. Conn, Scheinberg and Vicente [45, p. 242] define hidden constraints as unrelaxable constraints, which *"are not part of the problem specification/formulation, and their manifestation comes in the form of some indication that the objective function could not be evaluated"*.

The term hidden constraint often refers to the case that the computer code used to evaluate the objective function fails to return a value, e.g. [36, 37, 55, 56, 77; 82, Ch. 4]. An internal iteration may fail to converge, as, for example, for the problems considered in [7, 8, 26, 41]. An iterative method applied to numerically solve PDEs fails to converge for a problem studied in [26, 41], and hidden constraints are violated for up to 60% of the function evaluations, [26]. Furthermore, an internal computation may fail, as e.g. in [32], evaluation of the objective function may be impossible with the given computational budget or available storage, [56], or the optimization algorithm may ask to evaluate the objective function at points for which the function is not defined (such as the logarithm of a negative number), [90].

An objective function may also be designed to return no value, when it detects violation of a constraint which is not explicitly computable, [36]. For a problem on water resource policies, constraints were imposed on functions of random variables sampled as a part of the evaluation of the objective function, [35; 82, Ch. 10; 85]. Constraints without explicit representation in the problem formulation, which were tested within the objective function evaluation, have also been reported for a hydraulic capture problem, [11; 82, Ch. 9].

The requirement that the algorithm used to evaluate the objective function returns a value is sometimes considered as an example of hidden constraints, [32, 83]. The feasibility of an implicitly stated constraint, which is either satisfied or not, may also be tested by a separate computation, [83]. For a constrained optimization problem, the term hidden constraint may refer to the case that an evaluation of a constraint fails, [11].

As remarked in [36], in some cases, constraints which could be represented by inequalities may be handled as if they were hidden constraints, i.e., the explicit representation of the constraint is not used. As an example, the formulation as a hidden constraint may be preferred when using an optimization method which has convergence theory for hidden constraints, as for the well problem considered in [11].

In this thesis, the objective functions are functions of a numerical solution of a differential equation, and we find in general two types of hidden constraints. Firstly, the (analytic) solution of the differential equation may not be defined. Secondly, an integrator may return an error message instead of a numerical solution for a differential equation. In comparison to the definition of hidden constraints in the taxonomy for constraints in simulation-based optimization [90], this includes cases for which the reason for the infeasibility is known.

2.4.5 Optimization with hidden constraints

In this section, we consider the constrained optimization problem

$$\min_{x\in\Omega} f(x) \tag{2.27}$$

with hidden constraints. The feasible set $\Omega \subset \mathbb{R}^n$ is a proper subset of the known set of constraints $\Theta \subset \mathbb{R}^n$. A local solution of the problem (2.27) is a feasible point $x^* \in \Omega$ with a neighbourhood $\mathcal{N}$ such that $f(x) \geq f(x^*)$ for all $x \in \mathcal{N} \cap \Omega$.

For a moment, let $\Omega = \mathbb{R}^n$. For continuously differentiable functions, Taylor's theorem can be used to approximate the objective function in a neighbourhood of a point and to find new iterates. As an example, due to Taylor's theorem, we know that by searching for a new iterate $x_k + \alpha_k d_k$ in a descent direction $d_k \in \mathbb{R}^n$, an improved estimate for the local minimizer of an unconstrained optimization problem can be found for sufficiently small $\alpha_k > 0$, unless the current iterate is a stationary point, [108, Ch. 2]. For a twice continuously differentiable function f, Taylor's theorem tells us that the quadratic model function m_k defined by (2.28) with an arbitrary symmetric matrix H_k has approximation error $\mathcal{O}(\Delta_k^2)$, as shown e.g. in [108, Ch. 4].

$$m_k(s) = f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T H_k s, \text{ with a symmetric matrix } H_k \in \mathbb{R}^{n\times n}. \tag{2.28}$$

The model function m_k may be employed in a trust-region framework. With the knowledge about the objective function and the gradient at the current iterate, the behaviour of the continuously differentiable function can be predicted in a neighbourhood of the current point. Globally convergent line search methods and trust-region methods exploit the local information about the objective function.

We consider now a problem with hidden constraints. An optimization algorithm that aims at solving the problem with feasible set $\Omega \subset \Theta$ may try to evaluate the objective function at points for which the function is not defined. Heuristics for handling this challenge exist and are, for example, described in [18, Ch. 8; 58]. However, if the evaluation of the objective function fails, the behaviour of the objective function can no longer be predicted in a neighbourhood of the current iterate based on knowledge about the objective function and the gradient at the current iterate. As an example, a descent direction may not be feasible with respect to the hidden constraints. To ensure that an optimization algorithm computes a subsequence of iterates that converges to a stationary point becomes a challenge. If the feasible set could be defined by inequalities and equalities, traditional approaches for handling constrained problems such as penalty, barrier, augmented Lagrangian methods and sequential quadratic programming

(described e.g. in [108]) could be employed. However, these optimization techniques require an explicit formulation of the feasible set.

A strategy to ensure global convergence for optimization problems with hidden constraints consists of producing trial points which form an asymptotically dense subset of Θ. This implies that for all $x \in \Theta$ and for all $\delta > 0$, the algorithm tries to evaluate the objective function at a point x_k with $\|x - x_k\| \leq \delta$. This strategy is, for example, utilized by the optimization methods SNOBFIT [77] and DIRECT [55].

Directional direct search methods with global convergence results for problems with hidden constraints optimize the barrier function f_Ω defined by (2.29), or assign NaN or an arbitrary value larger than the objective function value at the current iterate to the objective function if a hidden constraint is violated, [56].

$$f_\Omega(x) = \begin{cases} f(x), & x \in \Omega \\ \infty \quad , & \text{else.} \end{cases} \tag{2.29}$$

Global convergence results for problems with hidden constraints exist for mesh adaptive direct search, as shown in [10], and in general for directional direct search methods which evaluate the objective function value at a set of poll points (2.25), reduce the step length if a step is not successful, and build a set of poll directions that is asymptotically dense in the unit sphere, possibly after normalization, [56].

Denote by V_k the set of normalized poll directions in iteration k. Following [82, Ch. 5], a sequence $\mathcal{V} = \{V_k\}_{k=1}^{\infty}$ is called rich if for all unit vectors $v \in \mathbb{R}^m$ and any subsequence $\{W_{k_j}\}_{j=1}^{\infty}$ of $\mathcal{V}$, we have that $\liminf_{j\to\infty} \min_{w \in W_{k_j}} \|v - w\| = 0$. That the sequence of sets of search directions is rich and any of its subsets becomes arbitrarily close to every vector in the unknown tangent cone to Ω at a limit point of unsuccessful iterations, is important for global convergence results, [10, 56]. For mesh adaptive direct search, strategies for the generation of sequences of sets of poll directions have been designed, [3, 13]. Furthermore, by including a finite number of random directions to the set of poll directions, rich sequences of sets of directions can be obtained almost surely, [82].

Chapter 3

Slow-fast population dynamic models for fish

3.1 Introduction

In the following, we consider a continuous-time model for numbers of prerecruits $N_0(t)$ and adults $N_i(t)$ of age $i = 1,\dots,n$. Denote by $\mathbf{N}(t)$ the $(n+1)$-dimensional vector $(N_0(t),N_1(t),\dots,N_n(t))$. We use bold symbols to represent vectors. The general slow-fast population dynamic (SFPD) model is defined by (3.1)–(3.2) with initial condition $\mathbf{N}(0) \in \mathbb{R}^{n+1}_{\geq 0}$.

$$\dot{N}_0(t) = -\alpha N_0(t) + \frac{1}{\varepsilon} A(N_0(t),N_1(t),\dots,N_n(t)), \tag{3.1}$$

$$\dot{N}_i(t) = \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t), \quad i = 1,\dots,n. \tag{3.2}$$

The model describes the processes of egg production, mortality and ageing. Equation (3.2) is identical to (2.8) from section 2.2.3. Parameter $m_i \geq 0$ is the mortality rate of fish at age $i = 0,\dots,n$ and $\alpha > 0$ is the progression rate from stage $i-1$ to i. Following [139], we assume that the mortality rate per prerecruit and the rate of egg production are considerably larger than the progression and mortality rates per adult. The function A describes the fast processes. The differential equation includes two distinct time-scales, a fast time T and a slow time $t = \varepsilon T$. In the remainder of chapter 3, we assume the ratio ε between slow and fast time to be positive and small. Touzeau and Gouzé [139] assumed A to be quadratic. We are interested in a more generic approach.

We use the definitions for recruitment and the spawning stock size from [139], which we restate by (3.3).

$$R(t) = \alpha N_0(t) \quad \text{and} \quad S(t) = \sum_{i=1}^{n} f_i N_i(t)\,. \tag{3.3}$$

Recruitment is the progression rate of prerecruits to the youngest adult age-class and the spawning stock size is the sum of fecund fish. Parameter $f_i \geq 0$ denotes the fraction of spawners in age-class i.

The aim of this chapter is two-fold. Firstly, we give conditions under which a general SFPD model describes recruitment as a function of the number of adults in several age-classes. Next, we investigate which A would admit a class of SR functions that includes the Ricker and the Beverton–Holt functions.

Section 3.2 is based on paper A and describes a framework for the SR relationship derived from a general SFPD model. We apply the results by Fenichel [54] and prove the existence of a function that links recruitment to numbers of adults. Instead of the results by Fenichel, the reduction theorem from [16] could be applied if A had a root for all $(N_1,\dots,N_n) \in \mathbb{R}^n$. However, this assumption may not always be satisfied. A framework for the investigation of the parent-progeny relationship of a general slow-fast population dynamic model is important for fisheries science, since the literature includes several examples of assumptions about the dynamics of prerecruits.

In section 3.3.1, we derive an SFPD model from ecological considerations about the dynamics of prerecruits. This is the parametrized SFPD model introduced in paper A. A rich class of SR functions, including the Ricker and Beverton–Holt functions, defines locally invariant manifolds of the new model. A second SFPD model is shown to be linked to the Shepherd SR function. Based on paper A, we describe fixed points, the slow-fast dynamics and nonnegativity of the variables of the parametrized SFPD model in section 3.3.2. In section 3.3.3, we re-formulate the parametrized SFPD model in terms of a smaller set of parameters. Concluding remarks regarding the assumptions underlying the general SFPD model are offered in section 3.4. The approach of using model (3.1)–(3.2) is compared to the traditional SR functions. Furthermore, the parametrized SFPD model is compared to the model in [139].

3.2 Geometric singular perturbation theory for slow-fast population dynamic models

The application of geometric singular perturbation theory in this chapter is based on terminology introduced in section 2.1. We consider the general SFPD model and obtain the system of differential equations (3.4) by multiplication with $\varepsilon > 0$. In case of small ε, we have a singularly perturbed differential equation expressed in terms of the slow time. The variable N_0 changes at fast rates, while $N_1,\dots,N_n$ are slow variables.

$$\begin{aligned} \varepsilon \dot{N}_0(t) &= -\varepsilon\alpha N_0(t) + A(N_0(t),N_1(t),\dots,N_n(t)), \\ \dot{N}_i(t) &= \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t), \quad i = 1,\dots,n. \end{aligned} \tag{3.4}$$

$$\begin{aligned} N_0'(T) &= -\varepsilon\alpha N_0(T) + A(N_0(T),N_1(T),\dots,N_n(T)), \\ N_i'(T) &= \varepsilon\left(\alpha N_{i-1}(T) - \alpha N_i(T) - m_i N_i(T)\right), \quad i = 1,\dots,n. \end{aligned} \tag{3.5}$$

The algebraic constraint of the reduced slow system (the case $\varepsilon \to 0$ of (3.4)) is $A(\mathbf{N}(t)) = 0$. The set of zeros of A consists of fixed points of the reduced slow system, which are invariant under the reduced slow system.

Consider now the general SFPD model (3.5) expressed in terms of fast time T. We consider the solutions of the ODE in a domain of interest $\mathbb{R} \times [0,\bar{S}]^n$ with $\bar{S} > 0$ and state the following hypotheses about the general SFPD model.

(h1) The set $\hat{K} \subset \mathbb{R}^n$ is open and $A : \mathbb{R} \times \hat{K} \to \mathbb{R}$ is a C^∞ function.

(h2) The set $\hat{K}$ contains $[0,\bar{S}]^n$.

(h3) $M_0 = \{\mathbf{N} \in \mathbb{R} \times \hat{K} | A(N_0, N_1, \ldots, N_n) = 0\}$ is the graph of a C^∞ function $h^0 : \hat{K} \to \mathbb{R}$.

(h4) It holds that $\partial A(N_0, N_1, \ldots, N_n)/\partial N_0|_{M_0} < 0$, the derivative of A with respect to N_0 on M_0 is negative.

In paper A, the results by Fenichel (see [80]) for compact manifolds (possibly with boundary) are applied to the general SFPD model under assumptions **(h1)**–**(h4)** and the following is proven. If h^0 is defined on a suitable compact set K with $[0, \bar{S}]^n \subset K \subset \hat{K}$ (as defined in paper A), then the graph of h^0 on K is a compact manifold with boundary. Hypotheses **(h1)**–**(h3)** imply that h^0 can be defined on such a suitable set K, as shown in Section 2.4. of paper A. For $\varepsilon > 0$ sufficiently small and under hypotheses **(h1)**–**(h4)**, there exists a function $h^\varepsilon : K \to \mathbb{R}$, which is arbitrarily many times continuously differentiable, jointly in $N_1, \ldots, N_n$ and ε, such that the manifold with boundary given by (3.6) and denoted by M_ε is locally invariant under (3.5).

$$M_\varepsilon = \{(N_0, N_1, \ldots, N_n) \in \mathbb{R}^{n+1} | N_0 = h^\varepsilon(N_1, \ldots, N_n),\ (N_1, \ldots, N_n) \in K\}. \tag{3.6}$$

Furthermore, there exists a manifold $\mathcal{W}^s(M_\varepsilon)$ locally invariant under (3.5) with the property that solutions of the differential equation (3.5) which start in $\mathcal{W}^s(M_\varepsilon)$ and remain in a neighbourhood of M_ε converge exponentially to M_ε. The manifold $\mathcal{W}^s(M_\varepsilon)$ is an $\mathcal{O}(\varepsilon)$-perturbation of, and diffeomorphic to $\mathbb{R} \times K = \mathcal{W}^s(M_0)$.

Denote by $\mathbf{N}(t)$ a solution of (3.5). Local invariance of a manifold with boundary under (3.5) implies the following. If $N_0(\bar{t}) = h^\varepsilon(N_1(\bar{t}), \ldots, N_n(\bar{t}))$ for some $(N_1(\bar{t}), \ldots, N_n(\bar{t})) \in [0, \bar{S}]^n$ and $\bar{t} \geq 0$, then for all $t \geq 0$, $(N_1(t), \ldots, N_n(t)) \in [0, \bar{S}]$ implies $N_0(t) = h^\varepsilon(N_1(t), \ldots, N_n(t))$. The result is obtained from the fact that local invariance of a manifold with boundary means that $\mathbf{N}(t)$ can only leave the manifold with boundary through its boundary in the slow directions, from $K \supset [0, \bar{S}]^n$, and (3.7).

$$\partial M_\varepsilon = \{(N_0, N_1, \ldots, N_n) \in \mathbb{R}^{n+1} | N_0 = h^\varepsilon(N_1, \ldots, N_n),\ (N_1, \ldots, N_n) \in \partial K\}. \tag{3.7}$$

Here, ∂K denotes the boundary of set K.

Function h^ε describes the number of prerecruits as a function of numbers of adults. A SR relationship may be derived from h^ε by using (3.3).

3.3 A parametrized slow-fast population dynamic model

In this section, we consider examples of the general SFPD model (3.1)–(3.2). For the remainder of chapter 3, let l_i, p_i, $m_i \geq 0$ and for reasons of simplicity f_i, m_0, $\alpha > 0$ for $i = 1, \ldots, n$. The parameters represent properties of the population dynamics, as described in chapter 2 and in Table 3.1 at the end of this chapter. Based on [139], we use the following assumptions.

Assumption I: $\alpha\varepsilon + m_0 < \sum_{i=1}^n f_i l_i \cdot (\prod_{j=1}^i \alpha/(\alpha + m_j))$.

Assumption II: $\exists i^* \in \{1, \ldots, n\}$ s.t. $p_{i^*} > 0$.

Assumptions I and II are re-stated from section 2.2.4. Additionally, we use the following assumption from paper A.

Assumption III: The parameters $l_i = l > 0$, $f_i = f > 0$, $p_i = p > 0$ are age-independent and positive for all $i = 1, \ldots, n$.

3.3.1 Deriving a slow-fast population dynamic model

We now derive a new SFPD model from the following assumptions. The population dynamic model is a general SFPD model. We follow the standard assumption in the fisheries literature that the rate of egg production at time t is the sum of rates of egg production of spawners in age-classes $i = 1,\dots,n$, [73, Ch. 3]. Let m_0, c, $\gamma > 0$. We denote by $p_i \geq 0$ the coefficient for the degree of density-dependence attributed to fish of age i, for $i = 1,\dots,n$, and define $\hat{K}_c = \{(N_1,\dots,N_n) \in \mathbb{R}^n | \sum_{i=1}^n p_i N_i(t) \geq 0\}$ and $\hat{K}_\gamma = \{(N_1,\dots,N_n) \in \mathbb{R}^n | 1 + (\sum_{i=1}^n p_i N_i)/(\gamma m_0) > 0\}$. The mortality rate of prerecruits is assumed to be $m_c : \hat{K}_c \to \mathbb{R}_{>0}$ defined by (3.8) or $m_\gamma : \hat{K}_\gamma \to \mathbb{R}_{>0}$ defined by (3.9).

$$m_c(N_1(t),\dots,N_n(t)) = m_0 + \left(\sum_{i=1}^n p_i N_i(t)\right)^c, \tag{3.8}$$

$$m_\gamma(N_1(t),\dots,N_n(t)) = m_0 \left(1 + \sum_{i=1}^n \frac{p_i}{\gamma m_0} N_i(t)\right)^\gamma. \tag{3.9}$$

The definition of the mortality rate of prerecruits is motivated by assumptions of the generalized Ricker model, as described in the following.

The generalized Ricker model is derived from the assumption that the mortality rate per prerecruit is $m(S) = m_0 + bS^c$, with $b > 0$, as described in [116, Ch. 3]. Assume now that the coefficient b for the degree of density-dependence is a function of age. With (3.3), we obtain function m_c for the mortality rate of prerecruits. In case of $c = 1$, the function m_c is equal to m_γ with $\gamma = 1$. The functions m_c and m_γ are positive and represent the following assumptions. For $i = 1,\dots,n$, the rate of mortality of prerecruits increases strictly with $N_i(t) > 0$ if $p_i > 0$. The density-independent mortality rate is $m(0) = m_0 > 0$.

From the above assumptions, we obtain two functions for the fast processes egg production and mortality of prerecruits, function $A_c : \mathbb{R} \times \hat{K}_c \to \mathbb{R}$ given by (3.10) and $A_\gamma : \mathbb{R} \times \hat{K}_\gamma \to \mathbb{R}$ given by (3.11).

$$A_c(\mathbf{N}(t)) = \sum_{i=1}^n l_i f_i N_i(t) - \left(m_0 + (\sum_{i=1}^n p_i N_i(t))^c\right) N_0(t), \tag{3.10}$$

$$A_\gamma(\mathbf{N}(t)) = \sum_{i=1}^n l_i f_i N_i(t) - m_0 \left(1 + \sum_{i=1}^n \frac{p_i}{\gamma m_0} N_i(t)\right)^\gamma N_0(t). \tag{3.11}$$

Summarizing, we obtain two SFPD models defined by (3.1)–(3.2) and either (3.10) or (3.11).

Consider the SFPD model with the fast dynamics described by A_c. The function m_c is positive on $\hat{K}_c$, and the set of zeros of A_c is the graph of $h_c^0 : \hat{K}_c \to \mathbb{R}_{\geq 0}$ defined by (3.12).

$$N_0(t) = \frac{\sum_{i=1}^n l_i f_i N_i(t)}{m_0 + (\sum_{i=1}^n p_i N_i(t))^c}. \tag{3.12}$$

Under Assumption III, the function h_c^0 is a Shepherd function (2.3). To see this, multiply (3.12) with α. With Assumption III, we have p, l, f, α, $m_0 > 0$ and we can define

$a = \alpha l/m_0 > 0$ and $b = p^c/(f^c m_0) > 0$. With the definition of $R(t)$ and $S(t)$ by (3.3), we obtain a Shepherd function (3.13).

$$R(t) = \frac{(\alpha l/m_0)S(t)}{1 + p^c/(f^c m_0)S(t)^c}. \tag{3.13}$$

The Shepherd function is the Beverton–Holt function in case of $c = 1$, and a dome-shaped function in case of $c > 1$. In comparison to the Deriso function, the Shepherd function can only approximate the Ricker function.

We consider now A_γ. The limit of the right-hand side of equation (3.11) as $\gamma \to \infty$ defines $A_\infty : \mathbb{R}^{n+1} \to \mathbb{R}$. The SFPD model with A_γ and A_∞ is given by (3.14)–(3.16), where (3.16) is identical to (3.2).

$$\begin{aligned}\dot{N}_0(t) = &-\alpha N_0(t)\\ &+\frac{1}{\varepsilon}\underbrace{\left[-m_0 N_0(t)\left(1+\frac{1}{\gamma m_0}\sum_{i=1}^{n} p_i N_i(t)\right)^{\gamma} + \sum_{i=1}^{n} l_i f_i N_i(t)\right]}_{=A_\gamma(\mathbf{N}(t))}, \quad \text{case } \gamma > 0,\end{aligned} \tag{3.14}$$

$$\begin{aligned}\dot{N}_0(t) = &-\alpha N_0(t)\\ &+\frac{1}{\varepsilon}\underbrace{\left[-m_0 N_0(t)\cdot\exp\left(\frac{1}{m_0}\sum_{i=1}^{n} p_i N_i(t)\right) + \sum_{i=1}^{n} l_i f_i N_i(t)\right]}_{=A_\infty(\mathbf{N}(t))}, \quad \text{case } \gamma \to \infty,\end{aligned} \tag{3.15}$$

$$\dot{N}_i(t) = \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t), \quad i = 1,\dots,n. \tag{3.16}$$

This model is the parametrized SFPD model introduced in paper A.

As described in paper A, the model describes the following parent-progeny relationships. The set of zeros of $A_\gamma : \mathbb{R} \times \hat{K}_\gamma \to \mathbb{R}$ is equivalent to the graph of $h^0_\gamma : \hat{K}_\gamma \to \mathbb{R}_{\geq 0}$ with $h^0_\gamma(N_1(t),\dots,N_n(t)) = N_0(t)$ given by (3.17). The graph of the parent-progeny function $h^0_\infty : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ defined by (3.18) is the set of zeros of A_∞.

$$R(t) = \alpha N_0(t) = \alpha\left(\sum_{i=1}^{n} \frac{l_i f_i}{m_0} N_i(t)\right)\left(1 + \sum_{i=1}^{n} \frac{p_i}{\gamma m_0} N_i(t)\right)^{-\gamma}, \tag{3.17}$$

$$R(t) = \frac{\alpha}{m_0}\left(\sum_{i=1}^{n} l_i f_i N_i(t)\right)\cdot\exp\left(-\frac{1}{m_0}\sum_{i=1}^{n} p_i N_i(t)\right). \tag{3.18}$$

Under Assumption III, we obtain the SR function (3.19) for the case $\gamma > 0$. The SR function is the Deriso model (1.3) with $a = \alpha l/m_0$ and $b = p/(m_0 f)$. For the case $\gamma \to \infty$, the parent-progeny function (3.18) is a Ricker function (3.20) if Assumption III holds.

$$R(t) = \frac{\alpha l}{m_0} S(t)\left(1 + \frac{p}{\gamma m_0 f} S(t)\right)^{-\gamma}, \quad \text{case } \gamma > 0, \tag{3.19}$$

$$R(t) = \frac{\alpha l}{m_0} S(t)\cdot\exp\left(-\frac{p}{m_0 f} S(t)\right), \quad \text{case } \gamma \to \infty. \tag{3.20}$$

For $\gamma = 1$, the SR function (3.19) is the Beverton–Holt function (2.13) explained by the model in [139].

3.3.2 Properties of the parametrized population dynamic model

The solution for an initial value problem (3.14)–(3.16) with initial condition $\mathbf{N}(0) \in \mathbb{R}^{n+1}_{\geq 0}$ is nonnegative assuming $\alpha, m_0 > 0$ and $f_i > 0$, for $i = 1, \ldots, n$, as shown in paper A. Nonnegative solutions of the parametrized SFPD model are important, since the differential equation represents changes in the number of individuals.

In paper A, it is proven that 0 is a fixed point of (3.14)–(3.16), and the parametrized SFPD model has a positive fixed point if assumptions I and II hold and l_i, p_i, $m_i \geq 0$ and f_i, m_0, $\alpha > 0$ for $i = 1, \ldots, n$. A population may under suitable assumptions exist for an infinite time.

The function A_γ is defined on a proper subset of $\mathbb{R}^n$. The results by Fenichel (see [80]) are suitable to describe the slow-fast dynamics. As shown in paper A, hypotheses **(h1)**–**(h4)** hold for the parametrized SFPD model and the geometric approach to singular perturbation theory as described in section 3.2 can be utilized. Thus, for $\varepsilon > 0$ and sufficiently small, there exists a slow manifold M_ε and a manifold $\mathcal{W}^s(M_\varepsilon)$ for the parametrized SFPD model (3.14)–(3.16) with the following properties. Solutions of (3.14)–(3.16) which start in $\mathcal{W}^s(M_\varepsilon)$ and remain in a neighbourhood of M_ε, converge exponentially to M_ε. The slow manifold is locally invariant under (3.14)–(3.16) and the graph of a function h^ε defined on a suitable set $K \subset \hat{K}_\gamma$, which can be chosen such that $K \supset [0, \bar{S}]^n$ for any $\bar{S} > 0$. The function h^ε is an $O(\varepsilon)$-perturbation of h^0_γ or h^0_∞. Under Assumption III, the function h^0_γ defines a Deriso function (3.19) and h^0_∞ defines a Ricker function (3.20). The slow manifold M_ε is given by (3.6), and we have $\mathcal{W}^s(M_0) = \mathbb{R} \times K$ (with the same set K).

We extend the results as follows, using the description of local invariance of a manifold with boundary from section 3.2 and nonnegativity of the solution of (3.14)–(3.16). For all $t \geq 0$, assume that $\mathbf{N}(t)$ satisfies (3.14)–(3.16) with $\mathbf{N}(0) \in \mathbb{R}^{n+1}_{\geq 0}$ and $N_i(t) \leq \bar{S}$ for some $\bar{S} > 0$ and for all $i = 1, \ldots, n$. Then, for ε sufficiently small, there exists a function h^ε defined on a suitable $K \supset [0, \bar{S}]^n$ such that $N_0(0) = h^\varepsilon(N_1(0), \ldots, N_n(0))$ for some $(N_1(0), \ldots, N_n(0)) \in [0, \bar{S}]^n$ implies $N_0(t) = h^\varepsilon(N_1(t), \ldots, N_n(t))$ for all $t \geq 0$.

We consider an example of the parametrized SFPD model. The parameter values are such that Assumptions I, II and III hold and the ODE has a positive fixed point. Recruitment and spawning stock size are obtained from (3.3) and numerical solution of (3.14)–(3.16) at $t \in \{0, 0.005, 0.01, \ldots, 10\}$. We use the same parameter values and method for numerical solution of the ODE as in paper A.

Figure 3.1 shows recruitment and the spawning stock size as functions of time for the parametrized SFPD model with $\gamma = 1.5$. A point illustrates a numerical solution of the ODE at $t \in \{0, 0.005, 0.01, \ldots, 10\}$. The trajectories $k = 1, \ldots, 30$ represent the solution of an initial value problem with initial condition $\mathbf{N}(0) \in [0, 60] \times [0.9k, 2 + 0.9k]$ drawn from a uniform distribution. In a short first phase, recruitment may change considerably faster than the spawning stock size. All solutions of the initial value problems considered here are attracted to the graph of a SR function, and approach a positive fixed point.

The SR relationships of the parametrized SFPD model and the fast dynamics of recruitment in the four cases $\gamma \in \{1, 1.5, 5\}$ and $\gamma \to \infty$ are illustrated in Figure 3.2. For our choice of initial conditions, all trajectories converge to the graph of a SR function. The parent-progeny relationships approximate a Deriso function and a Beverton–Holt ($\gamma = 1$) or Ricker function ($\gamma \to \infty$) shown in Figure 1.1.

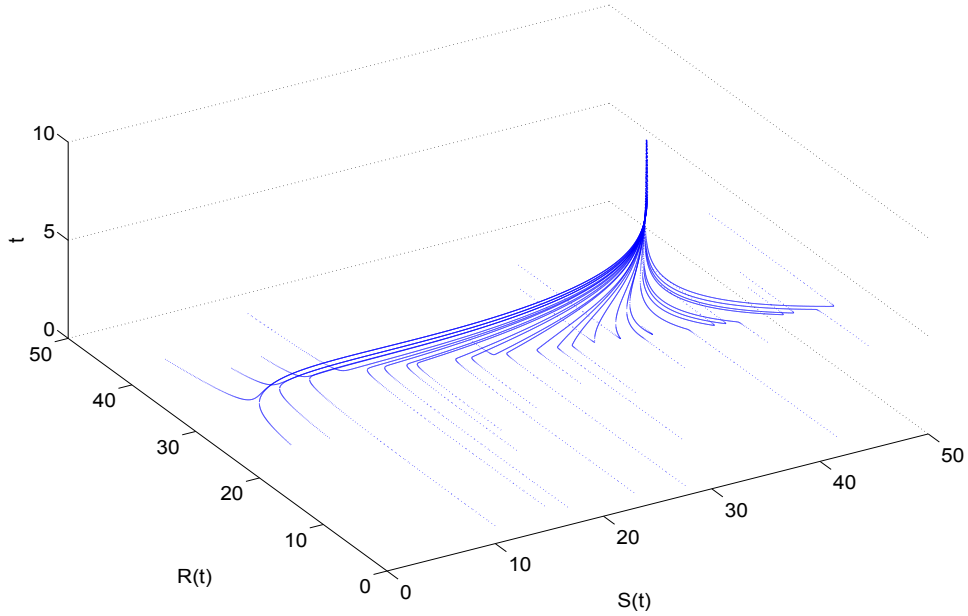


Figure 3.1: Simulation of recruitment and the spawning stock size as described by the SFPD model (3.14) and (3.16) with $\gamma = 1.5$. The solutions of the 30 initial value problems approach the graph of a SR function that approximates a Deriso function with $\gamma = 1.5$.

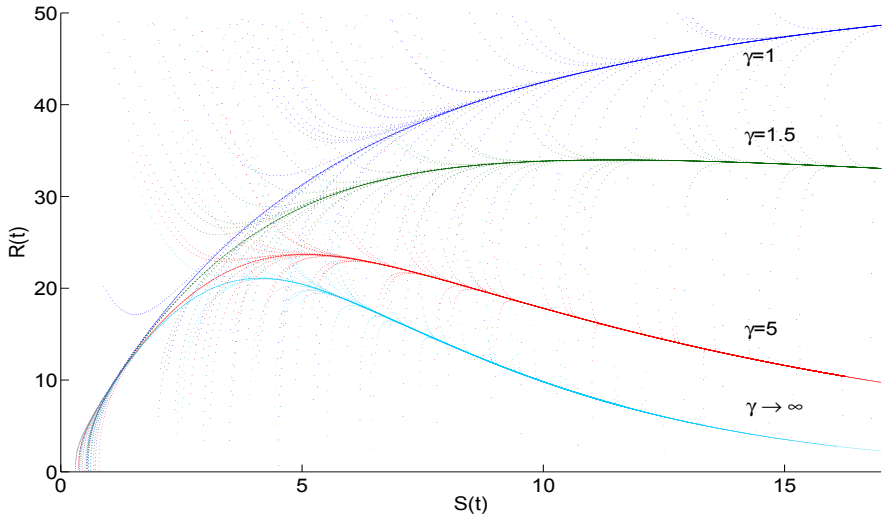


Figure 3.2: The SR relationships admitted by numerical solutions of the parametrized SFPD model (3.14)–(3.16). All trajectories approach the graph of a SR function, which can be approximated by the Deriso function. (From paper A.)

3.3.3 Dimensional analysis for the slow-fast population dynamic model

The aim of this section is to redefine the parametrized SFPD model in terms of dimensionless variables and parameters for simulation and parameter estimation, i.e., nondimensionalization of the ODE as described in [101, Sect. 2.1.2.8]. An introduction to dimensional analysis can be found in [94, Ch. 6]. The dimensions of the parameters of the parametrized SFPD model are given in Table 3.1.

The product $t\alpha$ is dimensionless and describes the time in which one individual progresses from age-class i to $i+1$, for $i = 0,\dots,n-1$. Denote by $c > 0$ a natural scale of $\mathbf{N}$. We substitute the time $\tilde{t} = t\alpha$ and $\tilde{\mathbf{N}} = \mathbf{N}/c$ into (3.14)–(3.16) and obtain the ODEs (3.21)–(3.23).

$$\frac{d}{d\tilde{t}}\tilde{N}_0(\tilde{t}) = -\tilde{N}_0(\tilde{t}) + \frac{1}{\varepsilon c\alpha} A_\gamma(c\tilde{\mathbf{N}}(\tilde{t})), \text{ case } \gamma > 0, \tag{3.21}$$

$$\frac{d}{d\tilde{t}}\tilde{N}_0(\tilde{t}) = -\tilde{N}_0(\tilde{t}) + \frac{1}{\varepsilon c\alpha} A_\infty(c\tilde{\mathbf{N}}(\tilde{t})), \text{ case } \gamma \to \infty, \tag{3.22}$$

$$\frac{d}{d\tilde{t}}\tilde{N}_i(\tilde{t}) = \tilde{N}_{i-1}(\tilde{t}) - \tilde{N}_i(\tilde{t}) - \frac{m_i}{\alpha}\tilde{N}_i(\tilde{t}), \quad i = 1,2,\dots,n. \tag{3.23}$$

From Eq. (3.17), we see that the following products of the parameters in the parametrized SFPD model fully describe the relationship between the numbers of adults and recruitment: α, γ, $l_i f_i/m_0$ and p_i/m_0, for $i = 1,\dots,n$. This motivates the definition of the dimensionless parameters θ_i, θ_{n+i} and $\theta_{2n+i} \geq 0$, with $i = 1,\dots,n$, and $\varepsilon^* > 0$ by (3.24).

$$\theta_i = \frac{l_i f_i}{m_0}, \quad \theta_{n+i} = c\frac{p_i}{m_0}, \quad \theta_{2n+i} = 1 + \frac{m_i}{\alpha}, \quad \varepsilon^* = \frac{\varepsilon\alpha}{m_0}. \tag{3.24}$$

We substitute the dimensionless parameters given by (3.24) into the parametrized SFPD model and obtain (3.25)–(3.26). The number of parameters has been reduced from $4n+4$ to $3n+2$.

$$\begin{aligned}\frac{d}{d\tilde{t}}\tilde{N}_0(\tilde{t}) = &-\tilde{N}_0(\tilde{t}) + \frac{1}{\varepsilon^*}\sum_{i=1}^{n}\theta_i\tilde{N}_i(\tilde{t}) \\ &- \begin{cases} \frac{1}{\varepsilon^*}\tilde{N}_0(\tilde{t})\left(1 + \frac{1}{\gamma}\sum_{i=1}^{n}\theta_{n+i}\tilde{N}_i(\tilde{t})\right)^\gamma, & \text{case } \gamma > 0, \\ \frac{1}{\varepsilon^*}\tilde{N}_0(\tilde{t}) \cdot \exp\left(\sum_{i=1}^{n}\theta_{n+i}\tilde{N}_i(\tilde{t})\right), & \text{case } \gamma \to \infty, \end{cases}\end{aligned} \tag{3.25}$$

$$\frac{d}{d\tilde{t}}\tilde{N}_i(\tilde{t}) = \tilde{N}_{i-1}(\tilde{t}) - \theta_{2n+i}\tilde{N}_i(\tilde{t}), i = 1,\dots,n. \tag{3.26}$$

The new formulation does not change the fact that the model describes slow-fast dynamics. As an example, assuming as in [139] that $\alpha = 0.8$, $m_0 = 0.5$ and $\varepsilon = 0.01$, we have $\varepsilon^* = 1.6\varepsilon = 0.016$.

The dimensionless parameters represent quantities of interest from a biological point of view. Parameters θ_i, with $i = 1,\dots,n$, describe the largest possible contribution $f_i l_i/m_0$ of age-class i to recruitment (the case $p_i = 0$). Parameter θ_{n+i} is a relative degree of density-dependence attributed to fish of age i. If $\mathbf{N}^*$ is a fixed point of (3.25)–(3.26), then $\theta_{2n+i}N_i^* = N_{i-1}^*$ for $i = 1,\dots,n$, and the age distribution is characterized by $\theta_{2n+1}, \dots, \theta_{3n}$.

Assumptions I, II and III in terms of the new parameters are, respectively:
Assumption I*: $(\varepsilon^* + 1) \cdot \prod_{j=1}^{n} \theta_{2n+j} < \sum_{i=1}^{n} \theta_i \cdot \prod_{j=i+1}^{n} \theta_{2n+j}$.
Assumption II*: $\exists i^* \in \{1,\dots,n\}$ s.t. $\theta_{n+i^*} > 0$.
Assumption III*: $\theta_i = \theta_1 > 0$ and $\theta_{n+i} = \theta_{n+1} > 0$ for all $i = 1,\dots,n$.
If Assumption III is satisfied, we obtain from (3.17) recruitment as a function of $\tilde{S}(t) = \sum_{i=1}^{n} \tilde{N}_i(t)$. The parent-progeny function is given by (3.27) and a Deriso function with $a = \alpha\theta_1$ and $b = \theta_{n+1}$.

$$R(t) = \alpha\theta_1 \tilde{S}(t) \left(1 + \frac{1}{\gamma}\theta_{n+1}\tilde{S}(t)\right)^{-\gamma}. \tag{3.27}$$

Instead of $\alpha > 0$, $m_i \geq 0$, for $i = 1,\dots,n$, we could assume $\theta_{2n+i} \geq 0$, $i = 1,\dots,n$. The existence of a locally invariant manifold and its representation as a graph rely on assumptions about A_γ and not on assumptions about θ_{2n+i}, $i = 1,\dots,n$. The proof of nonnegativity of the solution of the initial value problem (3.25)–(3.26) with $\tilde{\mathbf{N}}(0) \in \mathbb{R}^n_{\geq 0}$ and $\theta_k \geq 0$, $k = 1,\dots,3n$, can be done analogously to that for model (3.14)–(3.16) in paper A.

Furthermore, if $\theta \in \mathbb{R}^{3n}_{\geq 0}$, then the right-hand side of equations (3.25)–(3.26) is continuously differentiable with respect to $\tilde{\mathbf{N}}$ and θ in a neighbourhood of $\tilde{\mathbf{N}} \in \mathbb{R}^n_{\geq 0}$. This follows from $1 + \sum_{i=1}^{n} (\theta_{n+i}/\gamma)\tilde{N}_i > 0$ for $\tilde{\mathbf{N}} \in \mathbb{R}^n_{\geq 0}$.

3.4 Discussion

In the following, we compare the assumptions of the parametrized slow-fast population dynamic model to further models for recruitment and numbers of adults, including the model by Touzeau and Gouzé.

In case of $\gamma = 1$, the parametrized SFPD model (3.14)–(3.16) is the model by Touzeau and Gouzé with $p_0 = 0$. The parametrized SFPD model with $0 < \gamma \neq 1$ describes the processes egg production, ageing and mortality of adults in the same way as the model by Touzeau and Gouzé, but the mortality rate of prerecruits as a nonlinear function of numbers of adults. The model in [139] and the new SFPD model have nonnegative variables if all parameters are nonnegative, and have positive fixed points if Assumptions I and II hold, as shown in [139] and paper A, respectively.

In [139], it is shown that if ε is sufficiently small, then solutions of the parametrized SFPD model with $\gamma = 1$ can for sufficiently large t be approximated by a function of the number of adults, which is the Beverton–Holt function if Assumption 0 holds. Assumption 0 is restated from section 2.2.4.

Assumption 0: Let $p \geq 0$, f, $l > 0$. For all $i = 1,\dots,n$, either $p_i = p$, $f_i = f$ and $l_i = l$ or $p_i = f_i = l_i = 0$.

Assumption 0 with $f = l = 0$ would imply that h^0_γ is the constant zero function and $R(t) = 0$ for all $t \geq 0$. Assumption III considered in paper A implies Assumption 0. Assumption 0 is sufficient, but not necessary, for the set of zeros of A_γ and A_∞ to be the graph of a SR function. Assume more generally that Assumption 0b holds.

Assumption 0b: For all $i = 1,\dots,n$, either $p_i = cf_i$ and $l_i = l$, for $c \geq 0$, $l > 0$, or $p_i = f_i = l_i = 0$.

Under Assumption 0b, we have $\sum_{i=1}^{n} l_i f_i N_i(t) = lS(t)$ and $\sum_{i=1}^{n} p_i N_i(t) = cS(t)$. In this case, equations (3.17) and (3.18) define $R(t)$ as a function of $S(t)$. Assumption 0b could, for example, represent the case that fecundity and the degree of density-dependence attributed to fish of age i increase with age.

The general SFPD models are examples of continuous-time models with two stages, prerecruits and adults. The dynamics of stage structured populations may also be described by delay differential equations, [69; 102, Ch. 5]. An example of a delay differential equation is the two-stage model (3.28) for the number of juveniles $N_0(t)$ and the number of adults $N_1(t)$.

$$\begin{aligned}\dot{N}_0(t) &= \beta N_1(t) - m_0 N_0(t) - \underbrace{\beta N_1(t-\tau_0)\cdot \exp(-m_0\tau_0)}_{M_J(t)},\\ \dot{N}_1(t) &= \beta N_1(t-\tau_0)\exp(-m_0\tau_0) - m_1 N_1(t)\,.\end{aligned} \tag{3.28}$$

We use the notation for the parameters introduced in section 2.2.3. The duration of the juvenile stage is denoted by $\tau_0 \geq 0$. The rate $M_J(t)$ of maturation of juveniles to the adult stage at time t is assumed to be the product of the rate $\beta N_1(t-\tau_0)$ of transition to the juvenile stage at time $t-\tau_0$ and survival $\exp(-m_0\tau_0)$ from time $t-\tau_0$ to t.

In comparison to this assumption, the linear ODE model given by (2.7) (or (3.14) and (3.16) with $p_i = 0$, for all $i = 1,\ldots,n$, and $\varepsilon = 1$) describes a constant stage progression rate per individual ($\alpha \geq 0$). The delay differential equation (3.28) can be derived from a partial differential equation that models the dynamics of a population with continuous age structure, the McKendrick–Von Foerster equation. To this aim, the mortality rates and the rate of recruitment to the juvenile stage are assumed to be step functions of age, [69]. The linear ODE model (2.7) can also be derived from the McKendrick–Von Foerster equation, assuming that mortality, recruitment, and maturation rates are constant with respect to time and within age-classes, [33].

The parametrized SFPD model may explain stock and recruitment data which follows a Deriso, Beverton–Holt function or Ricker function. However, the SR functions and the SFPD models represent distinct assumptions about the mortality of prerecruits. Traditionally, SR functions are derived from the assumption that the mortality rate of prerecruits at time t is a function of the spawning stock size at the time of egg production (as described in section 2.2.1), but not at time t. Ricker [119] assumes that the mortality rate due to cannibalism is a function of the spawning stock size and that the spawning stock size can be measured by the number of eggs produced. The parametrized SFPD model admits a Ricker SR function and is based on the assumption that the mortality rate at time t is a function of the spawning stock size at time t. The assumption is suitable for populations with cannibalism in the time interval from spawning to recruitment. As an example, for Northeast Arctic cod, mortality of prerecruits of ages 0 to 3 may be due to cannibalism, [24, 25]. The general SFPD model describes spawning and recruitment as continuous processes. The timing of spawning (throughout all year, or in regular time intervals) may depend on the species (e.g. [139] and references therein). When recruitment can be assigned to a single spawning season, it might be suitable to describe recruitment as a discrete-time process. In many other cases, a continuous-time model might be preferred.

The SFPD models describe the dynamics of an age structured class of adults as a function of the number of prerecruits and vice versa. This two-sided relationship may

also be described by the reduced slow system of the parametrized SFPD model obtained from (3.17) and (3.2). The model is given by (3.29)–(3.31).

$$R(t) = \alpha\Big(\sum_{i=1}^{n} \frac{l_i f_i}{m_0} N_i(t)\Big) \left(1 + \sum_{i=1}^{n} \frac{p_i}{\gamma m_0} N_i(t)\right)^{-\gamma}, \tag{3.29}$$

$$\dot{N}_1(t) = R(t) - \alpha N_1(t) - m_1 N_1(t), \tag{3.30}$$

$$\dot{N}_i(t) = \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t), \quad i = 2, \ldots, n. \tag{3.31}$$

For ε sufficiently small, the solutions of the reduced slow system approximate the evolution of trajectories of the parametrized SFPD model on a slow manifold. The reduced slow system is based on the assumption that the ratio between the slow and the fast time scale ε is zero. The parametrized SFPD model describes, in addition to the reduced slow system, the fast initial evolution of prerecruits. The fast processes are important for the aim of studying the dynamics of prerecruits. Furthermore, the model (3.14)–(3.16) would also be applicable if ε was not (sufficiently) small.

$\alpha > 0$	transition rate; measured in numbers per individual and time unit t
$f_i > 0$	fecundity of fish of age $i \in \{1, \ldots, n\}$; dimensionless
$l_i \geq 0$	rate of eggs produced per fish of age i; measured in numbers per adult of age i and per time unit t
$p_i \geq 0$	degree of density-dependence attributed to fish of age i; measured in numbers per prerecruit, per adult of age i and per time unit t
$m_i \geq 0$	rate of density-independent mortality of age-class i; measured in numbers per individual and time unit t
$m_0 > 0$	rate of density-independent mortality of prerecruits; measured in numbers per prerecruit and time unit t
$\gamma > 0$	parameter that determines the asymptotic behaviour of the SRR; dimensionless

Table 3.1: Nomenclature for the parameters in the parametrized SFPD model.

Chapter 4

Derivative-free estimation of parameters in slow-fast population dynamic models

4.1 Introduction

We consider the problem of estimating parameters for a slow-fast differential equation, which describes the dynamics of a fish population. Let $n, r \in \mathbb{N}$, $m = 3n$ and $0 < \varepsilon < 1$. The vector $\mathbf{N}(t) = (N_0(t), \ldots, N_n(t)) \in \mathbb{R}^{n+1}$ represents the numbers of fish in $n+1$ classes at time $t \in \mathbb{R}$. The aim is to find the vector $\theta \in \Theta \subset \mathbb{R}^m_{\geq 0}$ that minimizes a nonlinear least squares error (4.1) between a numerical solution of an initial value problem (4.2) and a set of r data points $\mathbf{d}(t_j) \in \mathbb{R}^{n+1}$, $j = 1, \ldots, r$.

$$\min_{\theta \in \Theta} f(\theta) = \sum_{j=1}^{r} \|\mathbf{N}(t_j; \theta) - \mathbf{d}(t_j)\|_2^2 \tag{4.1}$$

$$\begin{aligned}
\text{s.t. } \dot{N}_0(t) &= -N_0(t) + \frac{1}{\varepsilon} \sum_{i=1}^{n} \theta_i N_i(t) \\
&\quad - \begin{cases} \frac{1}{\varepsilon} N_0(t) \left(1 + \frac{1}{\gamma} \sum_{i=1}^{n} \theta_{n+i} N_i(t)\right)^{\gamma}, & \text{case } \gamma > 0, \\ \frac{1}{\varepsilon} N_0(t) \cdot \exp\left(\sum_{i=1}^{n} \theta_{n+i} N_i(t)\right), & \text{case } \gamma \to \infty, \end{cases} \\
\dot{N}_i(t) &= N_{i-1}(t) - \theta_{2n+i} N_i(t), \quad i = 1, \ldots, n, \\
\mathbf{N}(t_1) &= \mathbf{d}_0 \ .
\end{aligned} \tag{4.2}$$

The data consists of measurements of fish at time $t_j \geq 0$, $j = 1, \ldots, r$. The slow-fast population dynamic model (4.2) is derived and described in section 3.3.3. When we wish to highlight the fact that the solution of initial value problem (4.2) is a function of θ, we denote it by $\mathbf{N}(t; \theta)$. The numerical solution of a differential equation may not be defined, and the objective function (4.1) may be noisy due to the adaptive strategies of methods for the solution of differential equations.

The problem of estimating parameters in differential equations can be formulated in several ways, and overviews of the approaches are e.g. given in [18, Ch. 8; 29, 53]. We employ the initial value approach and reformulate the problem (4.1)–(4.2) as a problem with constraints $\theta \in \Theta$, for which every evaluation of the objective function requires numerical solution of the initial value problem (4.2) on $[t_0, t_r]$. This choice of methodology is motivated as follows.

By choosing a different approach for the formulation of the optimization problem than the initial value approach, we might obtain a problem without hidden constraints, [23], or a smooth objective function, [137]. As described in [27], there are generally two strategies for optimization problems with noisy objective functions. The first strategy is to use a smooth approximation of the objective function or to try to avoid discontinuities. The sequence of step sizes and order of methods for the solution of differential equations may be fixed or the step size may vary smoothly, [63]. The same sequence of step sizes may be employed for several initial value problems, [53]. However, the adaptively chosen sequence of step sizes varies, in general, even with small perturbations of the parameters in the differential equation, [18, Ch. 8]. Accurate solution of the differential equation may require an adaptive choice of step sizes, particularly when the initial iterate is far from a local minimum, [53].

The second strategy consists of employing optimization methods which are suitable for noisy objective functions. Problems formulated using the initial value approach can be addressed by means of optimization methods which can handle noisy functions and hidden constraints. In this case, efficient methods for the accurate solution of differential equations with adaptive strategies can be employed. Furthermore, the initial value approach uses few variables in comparison to other approaches, and a solution of a differential equation, which is continuous in t, is obtained in every iteration.

In this chapter, we employ derivative-free optimization for the noisy problems with hidden constraints. A mesh adaptive direct search method, two of the efficient directional direct search methods described in section 2.4.3, a model-based derivative-free method, and the simplex-based Nelder–Mead algorithm are considered. We are interested in the ability of the derivative-free optimization methods to reduce the value of the objective function for a set of problems of form (4.1)–(4.2) that has been introduced in paper C. Our aim is to test whether the set of solvers finds sufficiently accurate solutions for the noisy optimization problems with hidden constraints, which are formulated using the initial value approach. Furthermore, we investigate the importance of hidden constraints for the performance of the derivative-free methods.

We employ a benchmark procedure due to Moré and Wild [99], which compares the ability of several derivative-free methods to reduce the objective function value as a function of the computational budget. In this chapter, the robustness of the algorithms to the initial iterate and the nonlinearity of the differential equation are examined.

Furthermore, we evaluate the performance of the derivative-free optimization methods for the problems from paper C against the performance for a set of unconstrained noisy benchmark problems introduced in [99]. The objective functions (4.3) are perturbations of nonlinear least squares functions from the CUTEr collection [65]. Function $\phi : \mathbb{R}^m \to [-1,1]$ represents deterministic noise and $\varepsilon_g > 0$ a level of noise.

$$g(\theta) = (1 + \varepsilon_g \phi(\theta)) \sum_{k=1}^{K} g_k(\theta)^2 . \tag{4.3}$$

In [46, 99], the performance of several derivative-free optimization methods has been tested based on the noisy benchmark problems. The benchmark set contains 53 problems with median dimension 7, and all functions $g_k : \mathbb{R}^m \to \mathbb{R}$ appear at most six times in the set. The time-dependent problems of form (4.1)–(4.2) from paper C are similar to the standard benchmark problems in the sense that they have dimension 6, each

objective function appears four times and is a function distorted by numerical noise arising from numerical solution of the differential equation with finite precision arithmetic. The question thus arises, whether we can expect derivative-free optimization methods to behave similarly for the set of noisy benchmark problems and the set of time-dependent problems.

A comparison of the relative performances of derivative-free optimization methods for the time-dependent problems can also be found in paper C. However, a comparison of several derivative-free methods may be implemented in many different ways. The numerical simulations in this chapter aim at verifying that the results from paper C remain valid when we consider a new set of optimization methods for a modified set of time-dependent problems.

4.1.1 Outline

Parameter estimation for stiff differential equations is topic of section 4.2. We demonstrate that objective functions which are functions of the numerical solution of an arbitrary differential equation may, for several reasons, be noisy and have hidden constraints. This leads us to argue for derivative-free estimation of parameters in stiff differential equations. Our choice of methodology is based on the review of algorithms with convergence theory for problems with hidden constraints in section 2.4.5. In sections 4.2.4–4.2.6, we consider the time-dependent problems from paper C as an example of the optimization problem addressing parameters of stiff differential equations. We motivate our choice of numerical method for solution of the ODEs. An example of problem (4.1)–(4.2) is used to illustrate numerical noise and hidden constraints of the optimization problems.

The numerical simulations are described in section 4.3 and the results in section 4.4. The results from paper C and from the new numerical simulations in this chapter are compared, and the performance of the derivative-free optimization methods is discussed in the context of the derivative-free optimization literature in section 4.5. Concluding remarks about the numerical simulations are given in section 4.6. Here, we mention further open problems in the field of parameter estimation for population dynamic models, which might be addressed using derivative-free optimization.

4.1.2 Related topics

Derivative-free optimization has been applied for several problems that involve systems of ODEs, [37, 50, 51, 57]. However, the fisheries literature includes few examples of the application of derivative-free optimization methods. In [114], a nondifferentiable objective function has been considered, and the pattern search method of Hooke and Jeeves [76] and the Nelder–Mead algorithm [107] are applied. The software Gadget (Globally applicable Area Disaggregated General Ecosystem Toolbox) [19, 20] is a tool for the design of stock assessment models. The toolbox includes a parallel implementation of the Hooke and Jeeves algorithm and has been applied for several stocks, [140]. However, only a hybrid approach using simulated annealing in a first phase and the pattern search method in a second phase has been tested, [140]. It would be interesting to test the performance of several derivative-free optimization methods, including the efficient and robust methods described in section 2.4.

The derivative-free optimization methods considered in this chapter aim at finding a local solution of the optimization problem (4.1)–(4.2). In general, strategies for identification of a good local solution of the optimization problem can be incorporated in the search step of directional direct search methods. As an example, a hybrid strategy for optimization, which is based on mesh adaptive direct search, has been suggested in [8]. A two-phase approach that combines an intelligent Monte-Carlo algorithm with a globally convergent local search and its application for estimation of parameters in population dynamic models has been presented in [136].

4.2 Estimation of parameters in stiff differential equations

We consider the optimization problem (4.1)–(4.2). From section 3.3.3, we know that $\mathbf{N}(t) \geq 0$ for $\theta \in \mathbb{R}^m_{\geq 0}$ and $\mathbf{N}(t_1) \in \mathbb{R}^n_{\geq 0}$. The function g defined by the right-hand side of (4.2) is continuously differentiable with respect to $\theta \in \mathbb{R}^m_{\geq 0}$ and $\mathbf{N}$ in a neighbourhood of $\mathbf{N} \in \mathbb{R}^n_{\geq 0}$. If the initial value problem (4.2) could be solved analytically, the objective function (4.1) would be continuously differentiable. Let $I = 0, \ldots, n$ and $J = 1, \ldots, m$. By [66], the derivatives $D = (\partial N_i(t;\theta)/\partial \theta_k)_{k\in J, i\in I}$ of the solution $\mathbf{N}(t;\theta)$ of (4.2) with respect to $\theta \in \Theta$ exist, are continuous and solutions of the sensitivity equations (4.4). In this case, the objective function (4.1), which is a quadratic function of the solution of (4.2), is a continuously differentiable function of $\theta \in \Theta$.

$$\dot{D}(t) = D(t) \left(\frac{\partial}{\partial N_i} g_j(\mathbf{N}(t;\theta),\theta) \right)_{j\in I, i\in I} + \left(\frac{\partial}{\partial \theta_k} g_j(\mathbf{N}(t;\theta),\theta) \right)_{k\in J, j\in I}, \qquad (4.4)$$

$$D(t_1) = 0 \ .$$

The objective function of problem (4.1)–(4.2) is a function of the numerical solution of (4.2). Depending on the approach for numerical solution of the ODE, the function f could be distorted by numerical noise, but it has an underlying smooth function.

More generally, we formulate an optimization problem with objective function (4.1) and constraint (4.5), where $\tilde{g} : \mathbb{R}^{n+1} \times \mathbb{R}^m \to \mathbb{R}^{n+1}$ is continuously differentiable with respect to $\mathbf{N}(t) \in \mathbb{R}^{n+1}$ and $\theta \in \mathbb{R}^m$. Equation (4.5) represents an arbitrary stiff differential equation that is solved numerically.

$$\dot{\mathbf{N}}(t) = \tilde{g}(\mathbf{N}(t);\theta) \ \text{ for } t > t_1 \ , \qquad \mathbf{N}(t_1) = \mathbf{d}_0 \ . \qquad (4.5)$$

In sections 4.2.1–4.2.3, we consider the generic problem. The specific ODE (4.2) is topic of sections 4.2.4–4.2.6.

4.2.1 Hidden constraints

An example of hidden constraints, which is relevant for parameter estimation for ODEs, is the case that the algorithm used to evaluate the objective function fails unexpectedly. As an example, a method for the solution of differential equations with adaptive step size control returns an error message if a required tolerance for the one-step error cannot be achieved without decreasing the step size below its lower bound, [71, Sect. II.4; 115]. This type of hidden constraint has been observed for a parameter estimation problem described in [37, 50, 51]. For some values of the parameters, the ODE considered is stiff and its numerical solution using an explicit method fails, [50, 83].

As an example in [23] shows, small variations of a parameter in a differential equation may considerably affect the asymptotic behaviour of the solution of a differential equation and may imply that it is impossible to obtain a numerical solution of the differential equation for all $t \in [t_1, t_f]$. In this case, the objective function is not defined, [23]. Defining explicit constraints which ensure stability of the solution of a differential equation is in general not a straightforward task, [18, Ch. 8].

From section 2.4.4, we know that hidden constraints may arise when the algorithm used to evaluate the objective function tests feasibility of constraints that cannot be explicitly computed. A numerical method for the solution of initial value problems may not always display an error message when the solution is not sufficiently accurate. As an example, in paper A, numerical solutions of the ODE (3.14)–(3.16) are reported to become negative for some $t \in [0, 10]$. The analytical solutions of the initial value problems are nonnegative, as shown in paper A. Negative numerical solutions of the ODE can only be explained by an inaccurate numerical solution of the ODE. The algorithm used to evaluate the objective function could be designed to return an error message when the numerical solution of the differential equation is negative. We would obtain an optimization problem with hidden constraints.

Additionally, the optimization problems described by (4.1) and (4.5) have hidden constraints if the analytical solution of the differential equation is not defined for all $t \in [t_1, t_r]$. An example of this type of hidden constraints has been described in [23].

4.2.2 Computational noise

Efficient methods for the solution of differential equations adapt the step size, often based on whether an approximation of the one-step error is smaller than a given tolerance, [71, Sect. II.4; 115]. Some algorithms, as an example the ones described in [127, 128], select the order of the method based on evaluation of conditional statements. In these cases, the numerical solution of a differential equation is a nonsmooth function of θ, [63; 71, Sect. II.6], and the objective function is nonsmooth, [27, 53].

Methods for the solution of stiff differential equations typically employ adaptive strategies additional to the choice of the step size or the order of the method. In each step, implicit methods require the solution of a system of (non)linear equations. The matrix defining the linear system may be updated dynamically. The number of iterations utilized by an iterative method for the approximation of a solution of the (non)linear equation may vary with the differential equation. The adaptive strategies explain discontinuities of the numerical solution of the differential equation, as described for BDF methods in [5]. If the algorithm used to obtain a numerical solution of a differential equation varies nonsmoothly with θ, computational noise (as defined in [100]) is introduced in the objective function.

Gear and Vu [63] found that the numerical solution of a differential equation obtained using a fixed step size and order varies smoothly with a parameter if the method is absolutely stable. A method which is not absolutely stable for a step size and ODE may amplify the one-step error with each step, [92, Ch. 7]. If a small variation in θ coincides with a loss of absolute stability of an integrator, we might observe a considerable jump in the numerical solution of the differential equation and the objective function. Computational noise may arise or be amplified.

4.2.3 Choice of the optimization methods

The optimization problems given by (4.1) and (4.5) may, for several reasons, have implicitly stated constraints or noisy objective functions. Optimization problems with hidden constraints may be challenging for some optimization methods, as described in section 2.4.5.

One of the challenges of noisy objective functions is that finite differences approximations of the derivatives of the objective function may be inaccurate. Hairer, Nørsett and Wanner [71, Sect. II.6.] give the approximation $\mathcal{O}(\sqrt{Tol})$ for the error of a finite difference as approximation of the derivative of a numerical solution of a differential equation with error tolerance Tol. As illustrated in [27], finite differences may have the opposite sign as the gradients when the absolute tolerance for the one-step error is 10^{-6}. For the problems considered in [27], an accurate approximation of the derivatives by finite differences is only obtained when a very small tolerance (10^{-9}) is used. The numerical solution of the differential equation is costly and round-off errors may become important if a very small tolerance is required, e.g. [125, Sect. 2.4].

For the general problem (4.1) with constraint (4.5), which is solved numerically, the derivatives of the solution of the ODE with respect to the parameters may be approximated by finite differences, by numerical solution of the sensitivity equations or by differentiation of the scheme used in numerical solution of the differential equation, while keeping all adaptive components fixed. Differentiation of the scheme may be combined with automatic differentiation. No matter which of the approaches described e.g. in [18, Ch. 8; 125, Sect. 2.4; 71, Sect. II.6] is used, exact derivatives are in general not easily obtained.

The noisy objective function (4.1) can be described as the sum of a smooth function $f_s : \mathbb{R}^m \to \mathbb{R}$ and small-amplitude noise $\phi : \mathbb{R}^m \to \mathbb{R}$, [37]. In this case, we have,

$$f(\theta) = f_s(\theta) + \phi(\theta).$$

For problems given by (4.1) and (4.5), the level of noise of the objective function and the underlying smooth objective function are in general unknown. The noisy objective functions may have local minima, which are nonstationary points of the underlying smooth function, and optimization methods may converge to local minima of the noisy objective function, [38, 133].

Gradient-based methods may in many cases be preferred to derivative-free optimization methods. Bard [18, Ch. 4] stated his view of direct search methods in 1974 as follows. *"Our own experience, however, has been disappointing; gradient methods, even using finite difference approximations, have outperformed direct search methods on all but the most trivial parameter estimation problems, both in reliability and speed of convergence."* In general, gradient-based methods are still more efficient than derivative-free optimization methods, e.g. [45, Ch. 1].

However, a strength of derivative-free optimization methods lies in the fact that they are suitable for problems with noisy or nonsmooth objective function and in cases when the computation of the derivatives is either not possible or considered to be too costly, [45, Ch. 1]. Derivative-free optimization methods that use average gradient information in form of simplex gradients may avoid being trapped in the local minima that are due to computational noise, [38, 47, 133]. Approximations of the Hessian utilized by some derivative-free methods represent average curvature information, [60].

Further results for derivative-free optimization in case of noisy problems with an upper bound for the level of noise or stochastic noise exist, and an overview can be found in [47].

Today, several classes of derivative-free optimization methods, for example, mesh adaptive direct search methods, have convergence theory for optimization problems with hidden constraints or nonsmooth objective functions, [10, 55, 56, 77]. These methods have been successfully employed for several problems with hidden constraints, as described in [83]. Global convergence for problems with hidden constraints can, as we saw in chapter 2, be achieved by building either a set of trial points that is asymptotically dense in the feasible set or an asymptotically dense set of poll directions. Derivative-free methods have been applied to problems with hidden constraints even before convergence results had been published, e.g. [26].

Model-based derivative-free methods might terminate at nonstationary points, but are for the following reasons suitable for problems with hidden constraints, as described in [41]. The optimization methods aim at evaluating the objective function at a set of trial points which are at least sufficiently affinely independent. Function evaluations at a finite set of several well-scattered points might be a good strategy to finding an improved iterate. Model-based derivative-free methods have been successfully employed for problems with hidden constraints, [41, 118].

For directional direct search methods, there is no guarantee that a limit point of a subsequence of iterates is a local solution of a problem with hidden constraints unless the sequence of poll directions is rich, [83]. However, the chance of finding a feasible descent direction might increase with the number of directions tested. In an unsuccessful iteration, a gradient-based line search method tests one direction in the half-space of descent directions, while a directional direct search method tests on average at least $(m+1)/2$ directions per iteration in the same half-space. An additional advantage of directional direct search methods is that variations of the set of poll directions (though with the requirements described in section 2.4.3) are possible. The use of multiple directions can be useful in case of hidden constraints. A similar argument is used in [86] to describe why multiple directions are useful in the case of nonsmooth objective functions.

Heuristic approaches such as simulated annealing [84] or genetic algorithms [117], have no (deterministic) convergence results (see e.g. [12, Ch. 5]). Model-based derivative-free methods and directional direct search methods are globally convergent and the trust-region radius and the step size parameter are reliable indicators of stationarity, [12, 45].

We have now shown that there are several reasons for derivative-free optimization, including methods without convergence theory in case of hidden constraints, to be suitable for the noisy optimization problem with hidden constraints given by (4.1) and (4.5).

4.2.4 A set of optimization problems

The following set of time-dependent problems of form (4.1)–(4.2) is introduced in paper C with the aim of testing the performance of several derivative-free optimization methods for estimation of parameters in a population dynamic model.

Thirty-six distinct data sets, which are designed as described in paper C, are given in Appendix A. Each data set consists of 10 values for $\mathbf{d}(t_j)$ at time $t_j = (4/9)(j-1)$, for $j = 1,\ldots,10$. Thirty-six optimization problems of form (4.1)–(4.2) are based on these data sets. Variations of the optimization problems are obtained by variations of the data and the three cases $\gamma = 1$, $\gamma = 2$ and $\gamma \to \infty$. For each of the 36 optimization problems, we consider the four distinct values for the initial iterate $\theta^{(0)}$ given in Table 4.1. For θ, the upper index refers to the iteration number.

The problems with $m = 6$ unknown parameters have bound constraints of the form $\theta \in \Theta = [\mathbf{0},\ \mathbf{b}] \subset \mathbb{R}^m$ for $\mathbf{b} = (100, 100, 2, 2, 5, 5)$. For all problems, parameters are linearly transformed by a mapping which assigns any value in $[\mathbf{0},\ \mathbf{b}]$ to a value in $[0,1]^m$.

In the following, we refer to the problems from Tables A.1, A.7, A.13, A.19, A.25 and A.31, which represent undistorted measurements, as Problem 1 to 6. Let $\tilde{\theta} = (6, 8, c, c, 2, 1.5)$ with $c = 0.05$ in case of $\gamma \to \infty$ and $c = 0.2$ otherwise. The parameter $\tilde{\theta} \in \Theta$ with $f(\tilde{\theta}) \in [10^{-5}, 0.75]$ is an a priori guess for a local solution of Problems 1 to 6. The exact values for $f(\tilde{\theta})$ for the undistorted problems are given in Table 4.2 in section 4.4.1.

In the remainder of this chapter, we denote by $g : \mathbb{R}^{n+1}_{\geq 0} \times \Theta \to \mathbb{R}^{n+1}_{\geq 0}$ the function defined by the right-hand side of the differential equation (4.2). The value of γ indicates the degree of nonlinearity of the differential equation.

4.2.5 Numerical solution of the slow-fast differential equations

In the following, we motivate our choice of method for numerical solution of (4.2). This is important, since the smoothness and accuracy of the objective function may depend on the numerical solution of the initial value problem. Let $Re(z)$ be the real part of $z \in \mathbb{C}$. We denote by $\lambda_{i,j}$ eigenvalue $i = 0, 1, 2$ of the Jacobian $\partial g(\mathbf{d}(t_j); \theta)/\partial \mathbf{d}$ for $t_j = (4/9)(j-1)$ and $j = 1,\ldots,10$. In this section, we consider the set of 24 differential equations from Problems 1 to 6 with the four values for θ from Table 4.1.

We compute the stiffness ratio for the set of 24 differential equations at $\mathbf{d}(t_j)$, $j = 1,\ldots,10$. All stiffness ratios are larger than 80 (rounded to two significant numbers) and smaller than $5.6 \cdot 10^{38}$. The median value of the 240 stiffness ratios is 1200 and indicates stiffness of the differential equations. For all differential equations considered, for all $i = 0, 1, 2$, and all $j = 1,\ldots,10$, we have $|\lambda_{i,j}| \geq 0.02$.

For small ε, the differential equation (4.2) is singularly perturbed. As described in chapter 3, for ε sufficiently small, there exists an n-dimensional manifold with boundary, such that trajectories with initial value not on the manifold with boundary and solutions that are slightly perturbed at any time, may converge exponentially to the manifold with boundary. The differential equation has the properties of a stiff differential equation as described e.g. in [92, p. 167].

For solving the initial value problem (4.2), we consider using one of the following two implementations of implicit methods for solution of ODEs described in [128]. NDF methods form the basis of the MATLAB program ode15s. An implementation of a one-step method for MATLAB, which is called ode23s, is based on the ideas of the Rosenbrock and the W-method. We employ ode23s and ode15s with the aim of approximating the solutions of 24 initial value problems for $t \in [0, 4]$. An algorithm for computation of the Jacobian is supplied. The default value for the absolute error

tolerance is 10^{-6}. Asymptotic values of the solutions of 24 differential equations as $t \to \infty$ are in interval $[17,42] \times [7,18] \times [4,10]$ and we set the relative error tolerance to 10^{-7}. We use MATLAB version R2012a [97] and 64 bit floating-point precision.

	Problem	Runtime of ode15s	Runtime of ode23s
$\theta = 1.2\tilde{\theta}$	1 (with $\gamma \to \infty$)	1.36e-01	2.63e+00
	2 (with $\gamma \to \infty$)	1.09e-01	1.63e+00
	3 (with $\gamma = 1$)	8.74e-02	1.29e+00
	4 (with $\gamma = 1$)	9.21e-02	1.49e+00
	5 (with $\gamma = 2$)	1.13e-01	2.39e+00
	6 (with $\gamma = 2$)	9.52e-02	1.35e+00
$\theta = 1.5\tilde{\theta}$	1 (with $\gamma \to \infty$)	1.52e-01	3.45e+00
	2 (with $\gamma \to \infty$)	1.13e-01	1.45e+00
	3 (with $\gamma = 1$)	1.04e-01	1.45e+00
	4 (with $\gamma = 1$)	9.80e-02	1.35e+00
	5 (with $\gamma = 2$)	1.31e-01	2.77e+00
	6 (with $\gamma = 2$)	9.51e-02	1.21e+00
$\theta = 0.5\mathbf{b}$	1 (with $\gamma \to \infty$)	1.86e-01	3.14e+00
	2 (with $\gamma \to \infty$)	1.40e-01	1.08e+00
	3 (with $\gamma = 1$)	1.20e-01	6.79e-01
	4 (with $\gamma = 1$)	1.26e-01	1.44e+00
	5 (with $\gamma = 2$)	1.36e-01	3.85e+00
	6 (with $\gamma = 2$)	1.32e-01	1.24e+00
$\theta = 0.8\tilde{\theta}$	1 (with $\gamma \to \infty$)	9.80e-02	1.22e+00
	2 (with $\gamma \to \infty$)	1.06e-01	1.83e+00
	3 (with $\gamma = 1$)	7.48e-02	8.90e-01
	4 (with $\gamma = 1$)	9.45e-02	1.79e+00
	5 (with $\gamma = 2$)	1.04e-01	1.71e+00
	6 (with $\gamma = 2$)	9.74e-02	1.66e+00

Table 4.1: Time (in seconds and rounded to three significant numbers) required to solve an initial value problem of form (4.2) when employing ode15s or ode23s. Each of the problems 1 to 6 is defined by an initial condition and a value for γ.

The execution time of ode15s and ode23s for each of the 24 initial value problems is given in Table 4.1. We observe that ode23s is slower than ode15s for all initial value problems considered here. On average, ode15s is 16 times faster than ode23s. This may partly be explained by the fact that ode23s uses on average 10 times more steps than ode15s for approximating a solution of (4.2). The algorithm ode23s is based on a second order numerical method, while ode15s is based on a class of NDF methods of orders 1 to 5, [128]. Our results are in accordance with the observation by Shampine and Reichelt [128] that ode15s is more efficient than ode23s when a high accuracy is required.

We now test whether ode15s has suitable stability properties for the solution of our initial value problems. The NDF methods implemented in ode15s have order $r \leq 5$ and are A(α)-stable with $\alpha \in [51°, 90°]$ and $\alpha = 90°$ for $r = 1, 2$, [128]. Our aim is to compare the region of absolute stability of the NDF method of order 1 with eigenvalues of the 240 Jacobians $\partial g(\mathbf{d}(t_j);\theta)/\partial \mathbf{d}$ described above. In [128], the region of absolute stability of the NDF method of order 1 is described as the subset of the points $z \in \mathbb{C}$

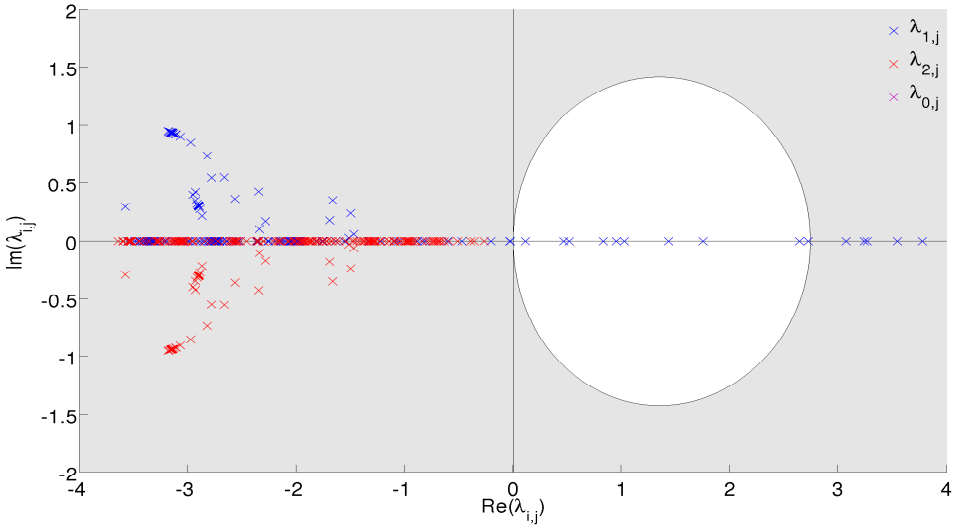


Figure 4.1: Region of absolute stability of the NDF method of order 1 described in [128] and eigenvalues $\lambda_{i,j}$, $i = 0,1,2$ of the Jacobian $\partial g(\mathbf{d}(t_j;\theta))/\partial \mathbf{d}$. The region of absolute stability is the shaded area.

with the property that,

$$z = (1-\kappa) - (1-2\kappa)\zeta - \kappa\zeta^2\,,\ \ \zeta \in \mathbb{C}\,,\ \ |\zeta| = 1\,.$$

The region of absolute stability of the NDF method of order 1 with κ as in [128] and the subset of the eigenvalues of the Jacobians with $|Re(\lambda_{i,j})| \leq 4$ are shown in Figure 4.1. Eigenvalues not indicated in Figure 4.1 are on the real axis.

We observe considerable variations in the eigendecomposition of the Jacobian with $\mathbf{d}(t)$. Some eigenvalues are real and positive, but we also observe negative eigenvalues. The NDF methods are not absolutely stable for real and positive eigenvalues of the Jacobian, unless sufficiently large step sizes are used. In this case, the one-step error may be amplified with the number of steps of ode15s. However, 97% of the eigenvalues considered here have negative real parts and we have $|\arg(-\lambda)| < 17^\circ$ for all λ with $Re(\lambda) < 0$. The NDF methods are A(α)-stable for $\alpha = 17^\circ$. Summarizing, our results show that the NDF methods are absolutely stable for the 24 initial value problems considered in this chapter and for all positive step sizes, unless the Jacobian has positive real eigenvalues.

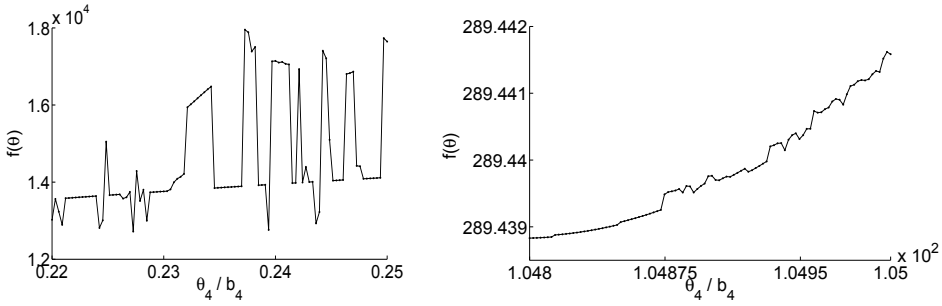
In the following, numerical solutions of initial value problems are obtained by ode15s. An approximation of $\mathbf{N}(t)$ at t_j, $j = 1,\ldots,r$, is calculated by interpolation. We caution that the results from employing ode23s and ode15s for the 24 initial value problems considered in this chapter can only give an indication for which numerical method might be suitable for numerical solution of the ODE (4.2).

The algorithm ode15s selects the number of iterations and the approximation of the Jacobian of an iterative method utilized to approximate a solution of a nonlinear equation, the order of the NDF methods, and the step size by evaluating conditional statements, [127, 128]. The objective function of problem (4.1)–(4.2) is a nonsmooth function with computational noise.

4.2.6 An example of the optimization problem

In this section, we consider an example of problem (4.1)–(4.2), the case $\gamma \to \infty$ of differential equation (4.2) with the data set from Table A.1. This is Problem 1. Let $\theta = (6, 8, 0.08, \theta_4, 2, 1.5)$. The values for θ_1, θ_2, θ_5 and θ_6 correspond to the ones in $\tilde{\theta} = (6, 8, 0.05, 0.05, 2, 1.5)$ with $f(\tilde{\theta}) = 3.25 \cdot 10^{-3}$. Here and in the following, objective function values are rounded to three significant figures.

Figure 4.2 illustrates the computational noise in $f(\theta)$ as a function of θ_4. The label on the x-axis of Figure 4.2 corresponds to $\theta_4/b_4 \in [0,1]$. Figure 1.3 from chapter 1 shows the objective function for the same problem, but the absolute and relative tolerance for the one-step error are required to be smaller or equal to 10^{-3} and 10^{-4}, respectively.



(a) For $\theta_4 \in [0.44, 0.5]$.

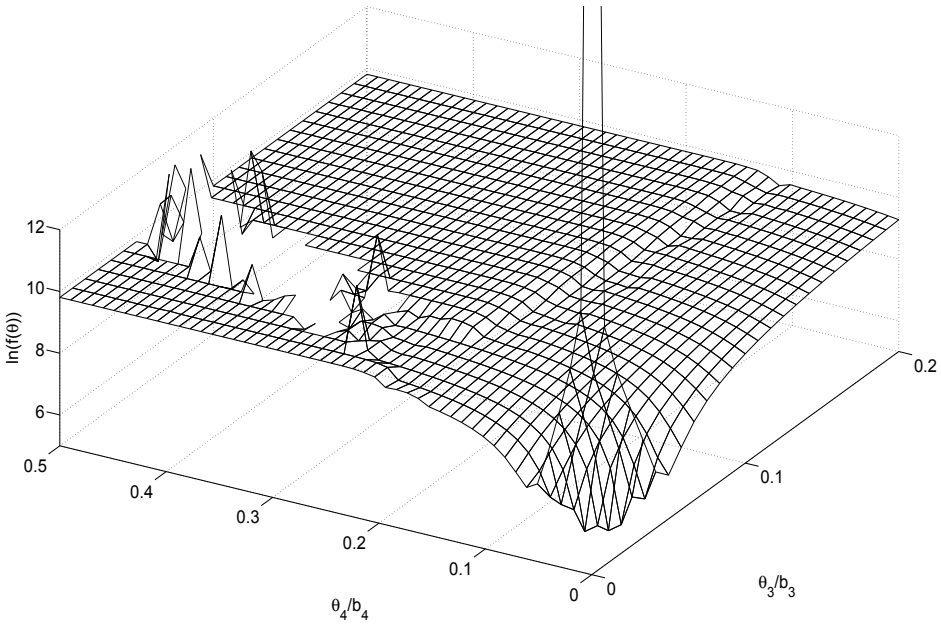
(b) For $\theta_4 \in [0.02096, 0.0210]$.

Figure 4.2: Graph of the objective function of problem (4.1)–(4.2) as a function of θ_4, evaluated at 100 points.

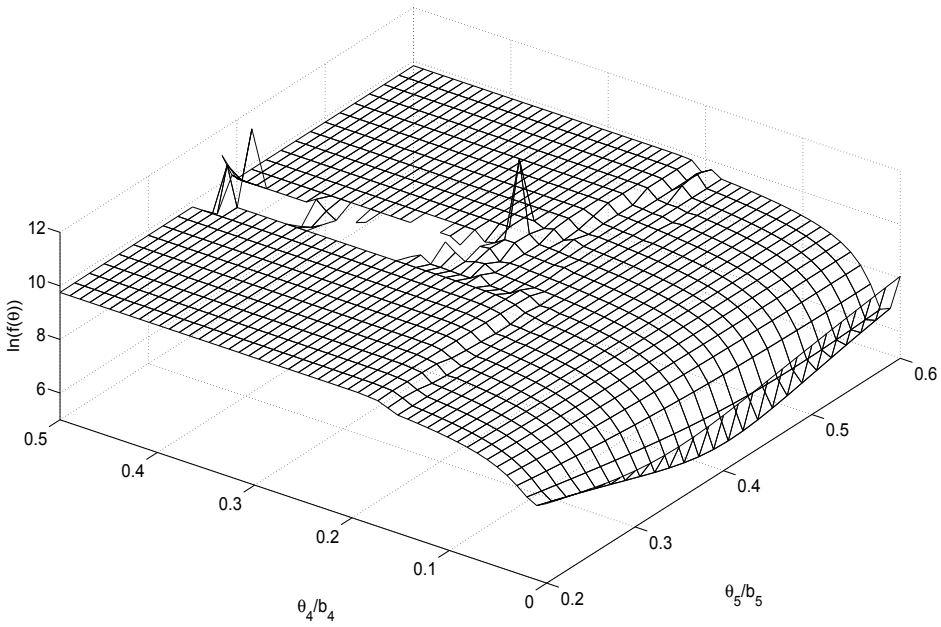
We observe that the level of relative noise varies with the tolerance for the one-step error and θ_4. Variations in the absolute stability of the NDF methods for the eigenvalues of the Jacobian of function g coincide with the 16 largest jumps in the value of the objective function that can be observed in Figure 4.2 (a). Here, the level of noise is considerably larger than the square root of the absolute tolerance for the one-step error 10^{-6}. We observe high- and low-amplitude noise in the objective function.

Figure 4.3 (a) presents the graph of the objective function as function of $\theta_3 \in [0,1]$ and $\theta_4 \in [0, 0.4]$. The graph of $f(\theta)$ as function of $\theta_4 \in [0, 0.4]$ and $\theta_5 \in [1,3]$ is given in Figure 4.3 (b).

The objective function is evaluated on a grid. Missing values in the graph correspond to parameters, for which the objective function could not be evaluated and indicate hidden constraints. The value 'NaN' (not a valid number) is assigned to the objective function in the following two cases. The algorithm ode15s returns an error message when the tolerances for the one-step error cannot be achieved without decreasing the step size below its lower limit. In case of $\theta_3 = \theta_4 = 0$, the differential equation (4.2) is linear, and the solution of the initial value problem increases exponentially, resulting in objective function value $1.15 \cdot 10^{23}$. When the numerical solution of the differential equation specified by some $\theta \in \Theta$ is too large to be represented by a computer, an error message is displayed. For the example illustrated in Figure 4.3 (a), the objective function is not defined for around 4% of the parameter values.



(a) $f(\theta)$ as a function of θ_3 and θ_4.



(b) $f(\theta)$ as a function of θ_5 and θ_4.

Figure 4.3: Graph of the objective function of problem (4.1)–(4.2) on a logarithmic scale.

The graphs of the objective function as a function of θ_3 and θ_4 and as a function of θ_5 and θ_4 show high frequency oscillations and have multiple saddle points for $\theta_4/b_4 \in [0.1, 0.25]$. For $\theta_4/b_4 \geq 0.25$, hidden constraints are observed and the graph of the objective function shows high frequency and high amplitude oscillations. Again, we observe considerable variations in the level of noise.

Figure 4.3 illustrates that the objective function may be highly sensitive to the values of θ_3 and θ_4 in comparison to the value of θ_5. The function f is poorly scaled (as defined in [108, Ch. 2]) for $\theta = (6, 8, 0.08, \theta_4, \theta_5, 1.5)$.

4.3 Numerical simulations

In paper C, the following derivative-free optimization methods are employed for problem (4.1)–(4.2). FMINSEARCH is an implementation of the Nelder–Mead algorithm as described in [88]. The simplex-based optimization method is known for a good practical performance, and has been employed as a point of reference in many cases, including for parameter estimation for population dynamics models, [114].

SID–PSM [46, 48] is an implementation of a generalized pattern search method which uses simplex gradients to order the poll directions and quadratic model functions in a search step. Both strategies can considerably improve the efficiency of generalized pattern search methods, [46, 48]. As described in [49], SID–PSM handles bound constraints by following an approach [86, 93] that may guarantee global convergence of generalized pattern search methods for bound constraint problems. SID–PSM has been tested for the set of standard benchmark problems, [46]. For noisy problems and computational budgets between $70(m+1)$ to $120(m+1)$ objective function evaluations, SID–PSM solves more problems than other derivative-free optimization methods, and it performs particularly well when high accuracy solutions are required, [46].

Mesh adaptive direct search has convergence theory for problems with bound and hidden constraints, [10], and to second-order stationary points, [1]. The implementation of a mesh adaptive direct search method NOMAD [2, 14, 89] is enhanced by a search step that consists of minimization of a quadratic model function. NOMAD orders the poll directions based on evaluations of a quadratic model function. A second search step is defined based on previously successful directions, [39]. For a set of benchmark problems and computational budgets of up to $1500(m+1)$ objective function evaluations, the mesh adaptive direct search method solves about as many problems as SID–PSM, [39].

The two directional direct search methods employ quadratic model functions, and a trust-region method based on radial basis functions is considered. The implementation of a derivative-free trust-region method using radial basis functions ORBIT has been observed to be efficient in case of small computational budgets, [143].

FMINSEARCH is not implemented to handle bound constraints. In paper C, the Nelder–Mead algorithm solves an unconstrained problem and returns negative values for some parameters and problems. In addition to being biologically implausible, negative values for one or several parameters may explain unexpected behaviour of the objective function. This is illustrated by the graph of the objective function as a function of θ_1 and θ_4 in Figure 4.4. The objective function (4.1) in case of $\gamma \to \infty$ and for the data set from Table A.1 is evaluated at $\theta = (\theta_1, 8, 0.08, \theta_4, 2, 1.5)$. Negativity

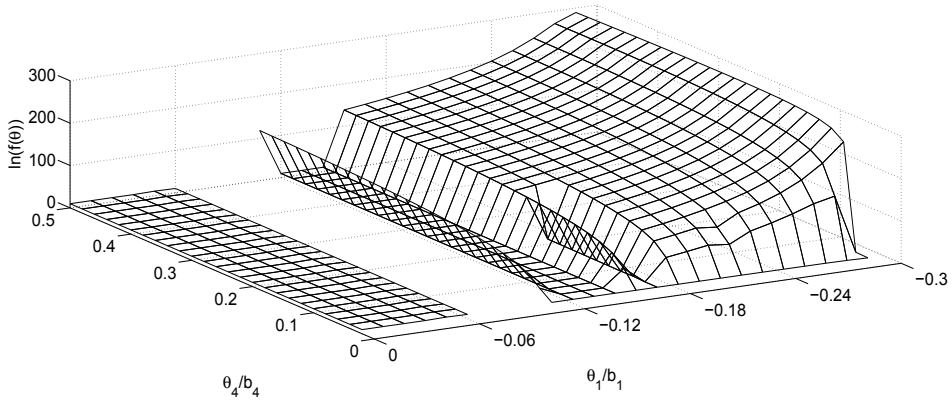


Figure 4.4: Graph of the objective function of problem (4.1)–(4.2) as a function of θ_1 and θ_4 on a logarithmic scale and for negative θ_1.

of one or several of $\theta_1, \ldots, \theta_n$ may imply that the solution of the differential equation (4.2) is exponentially decreasing. In case of $0 < \gamma$, the function g is only defined if $-\gamma \leq \sum_{l=1}^{n} \theta_{n+l} N_l(t)$. If one or several of $\theta_{n+1}, \ldots, \theta_{2n}$ are negative, the function g is not be defined for all $\mathbf{N}(t) \in \mathbb{R}^{n+1}_{\geq 0}$.

However, it is possible that FMINSEARCH finds an infeasible point with small objective function value. The results from paper C show that FMINSEARCH finds a good solution for more of the time-dependent problems than any other solver. That the other implementations of derivative-free optimization methods consider problems with bound constraints could be a reason for the good performance of FMINSEARCH. We might suspect that the performances of the derivative-free optimization methods will change if FMINSEARCH is required to return a feasible point.

Figure 4.3 suggests that some of the optimization problems are poorly scaled. In general, trust-region methods employing spherical trust-regions are more sensitive to poor scaling of the objective function than line search methods, [108, Ch. 2]. We test whether a generating set search method that is enhanced by a strategy different to minimization of a trust-region subproblem performs well for the set of time-dependent problems.

New numerical simulations

In this section, we suggest new numerical simulations for the set of time-dependent problems from paper C, as described in section 4.2.4. FMINSEARCH now handles bound constraints and we consider a generating set search method which we hypothesize to perform well for poorly scaled problems.

We follow the terminology in [61] and refer to the generating set search method introduced in [60, 61] as GSS–CI. We note however that in [4], a more general class of methods has been referred to as GSS–CI. The generating set search method described in [60, 61] utilizes average curvature information to define a Newton-like step. In comparison to the two directional direct search methods NOMAD and SID–PSM, GSS–CI updates a vector of m step size parameters individually with the aim of addressing poor

scaling, [60]. GSS–CI converges under suitable assumptions to second-order stationary points, [4]. This may be advantageous in case of problems with saddle points. GSS–CI has been observed to be robust and able to solve a large number of problems in comparison to a well-known trust-region method, [61].

Following [107], we employ the Nelder–Mead algorithm for the barrier function (4.6).

$$f_{\Theta}(\theta) = \begin{cases} f(\theta), & \theta \in [\mathbf{0}, \mathbf{b}], \\ \infty \quad , & \text{else.} \end{cases} \tag{4.6}$$

When a Nelder–Mead simplex method optimizes a barrier function, the simplex may become arbitrarily close to a subspace of the boundary of the feasible set, [28]. The algorithm GSS–CI handles bound constraints by minimizing the barrier function (4.6). This means that for two of the five solvers, the bound constraints are treated as hidden constraints.

The numerical simulations are implemented as follows. The value 'NaN' is assigned to the objective function when the algorithm used to numerically solve the differential equation returns an error message. Function evaluations at points $\theta \notin \Theta$ are not counted in the computation of the data profiles. An algorithm obtains information from a step to an infeasible point, as it learns that such a point is not a local solution to the optimization problem. However, the solvers that have information about the bound constraints can efficiently handle them.

For all solvers, termination criteria are chosen such that the number of function evaluations corresponds to 85 simplex gradient evaluations. Following [99], we denote by one simplex gradient evaluation $(m+1)$ function evaluations. The use of simplex gradient evaluations instead of function evaluations allows to compare the performances of solvers for problems with distinct numbers of variables, [99]. The computational budget of at most $85(m+1)$ function evaluations is rather small compared to the literature (e.g. [15, 99, 120]) and motivated by the fact that the computational budget is often a limiting factor for estimation of parameters in population dynamic models, [140]. We use the same versions of the implementations of derivative-free optimization methods and the same options for algorithm parameter values as in paper C.

4.4 Results

In this section, we consider the performance of the set S of five derivative–free optimization methods for the noisy problems (4.1)–(4.2) with hidden constraints. Percentages are rounded to the next integer and objective function values are represented with three significant numbers. We denote by $f_{L,s}$ the lowest objective function value obtained by a derivative-free optimization method $s \in S$ for the time-dependent problems and for a computational budget of $85(m+1)$ function evaluations.

4.4.1 Performance of a set of derivative-free optimization methods

This section starts with a study of the performance of all five solvers as one set of methods, as in [120]. We investigate Problems 1 to 6 with the four values for the initial iterate $\theta^{(0)}$ from Table 4.2. For Problem 1 to 6 with undistorted data, an upper bound

and approximation for the minimum is $f(\tilde{\theta})$. Table 4.2 presents the average over the smallest function values found by the five solvers and the smallest function value f_L obtained by any solver. We compare the objective function values obtained by the set of derivative-free optimization methods with $f(\tilde{\theta})$ and $f(\theta^{(0)})$.

Initial iterate	Problem	$\frac{1}{5}\sum_{s\in S} f_{L,s}$	f_L	$f(\tilde{\theta})$	$f(\theta^{(0)})$
$\theta^{(0)} = 1.2\tilde{\theta}$	1 (with $\gamma \to \infty$)	5.37e+00	2.49e-03	3.25e-03	2.73e+02
	2 (with $\gamma \to \infty$)	5.71e+00	1.70e-03	7.41e-01	1.81e+02
	3 (with $\gamma = 1$)	2.59e-03	3.98e-09	3.68e-03	2.61e+02
	4 (with $\gamma = 1$)	6.12e-02	4.56e-06	5.90e-02	6.16e+01
	5 (with $\gamma = 2$)	4.51e-03	8.90e-05	7.04e-04	1.23e+02
	6 (with $\gamma = 2$)	1.18e-02	3.99e-06	1.39e-02	2.49e+01
$\theta^{(0)} = 1.5\tilde{\theta}$	1 (with $\gamma \to \infty$)	3.10e+01	6.69e-03	3.25e-03	1.06e+03
	2 (with $\gamma \to \infty$)	5.22e+01	1.70e-03	7.41e-01	7.11e+02
	3 (with $\gamma = 1$)	2.98e-03	6.94e-05	3.68e-03	1.02e+03
	4 (with $\gamma = 1$)	1.27e+00	4.56e-06	5.90e-02	2.48e+02
	5 (with $\gamma = 2$)	6.00e-02	9.50e-04	7.04e-04	4.93e+02
	6 (with $\gamma = 2$)	4.85e-01	3.99e-06	1.39e-02	9.69e+01
$\theta^{(0)} = 0.5\mathbf{b}$	1 (with $\gamma \to \infty$)	8.68e+03	8.70e+00	3.25e-03	1.92e+04
	2 (with $\gamma \to \infty$)	5.38e+02	3.28e+02	7.41e-01	1.53e+04
	3 (with $\gamma = 1$)	5.71e-02	1.10e-03	3.68e-03	3.65e+03
	4 (with $\gamma = 1$)	6.21e+00	4.19e+00	5.90e-02	5.43e+03
	5 (with $\gamma = 2$)	6.39e-01	9.62e-02	7.04e-04	6.48e+02
	6 (with $\gamma = 2$)	4.99e+00	3.50e+00	1.39e-02	9.80e+01
$\theta^{(0)} = 0.8\tilde{\theta}$	1 (with $\gamma \to \infty$)	9.16e+00	1.23e-03	3.25e-03	6.50e+02
	2 (with $\gamma \to \infty$)	7.58e+00	1.70e-03	7.41e-01	4.16e+02
	3 (with $\gamma = 1$)	1.60e-02	1.77e-10	3.68e-03	6.09e+02
	4 (with $\gamma = 1$)	8.74e-02	4.56e-06	5.90e-02	1.24e+02
	5 (with $\gamma = 2$)	2.63e-02	1.21e-08	7.04e-04	2.76e+02
	6 (with $\gamma = 2$)	1.70e-02	3.99e-06	1.39e-02	5.59e+01

Table 4.2: Results from numerical simulations. The index $i \in \{1,\ldots,6\}$ refers to data set i with $\sigma = 0$ in Appendix A. We denote by $\frac{1}{5}\sum_{s\in S} f_{L,s}$ the average value of the lowest objective function values found by five derivative-free optimization methods within 595 objective function evaluations when starting at $\theta^{(0)}$. The lowest objective function value obtained by any solver is denoted by f_L, and $\tilde{\theta}$ is an a priori guess for a solution of the optimization problem.

From Table 4.2, we observe that for about 71% of the problems, the five derivative-free optimization methods together found an objective function value smaller than $f(\tilde{\theta})$. More precisely, f_L is smaller than $f(\tilde{\theta})$ for all problems with $\theta^{(0)} = 1.2\tilde{\theta}$ or $\theta^{(0)} = 0.8\tilde{\theta}$ and for two of six problems with $\theta^{(0)} = 1.5\tilde{\theta}$. When starting close enough to the a priori guess for a solution of the optimization problem, the five derivative-free methods together find a solution with objective function value smaller than $3 \cdot 10^{-3}$. For 25% of the problems overall, the smallest objective function value obtained by any solver within 85 simplex gradient evaluations is smaller than the absolute tolerance for the one-step error. The smallest relative reduction of the objective function by all five solvers together $(f(\theta^{(0)}) - f_L)/f(\theta^{(0)})$ is by 97%. The minimum value of the objective function is unknown, and may not be zero. Overall, we find that the set of five

solvers finds sufficiently accurate solutions for the least squares problem.

The point $\theta^{(0)} = 0.5\mathbf{b}$ is the initial iterate with largest distance (in the 2-norm) to $\tilde{\theta}$. The results in Table 4.2 show that the problems with $\theta^{(0)} = 0.5\mathbf{b}$ are most challenging. In general, we observe in Table 4.2 that for every initial iterate, the values of f_L and $1/5\sum_{s\in S} f_{L,s}$ are larger for the problems with $\gamma \to \infty$ than for the problems with $\gamma = 1$ or $\gamma = 2$. The problems with highly nonlinear differential equation appear to be challenging.

Comparing the third and fourth column in Table 4.2, we observe that the average solution found by the five solvers is on average 27 times larger (rounded to the next integer) than the best solution found by any of the five solvers. A good choice of the optimization method is important.

4.4.2 Comparison of the derivative-free optimization methods

In this section, we consider the results from a comparison of the five derivative-free optimization methods for the set of time-dependent problems from section 4.2.4. Following paper C, the optimization problems are distinguished by the value of γ and the initial iterate.

The comparison of the performances of derivative-free optimization methods relies on tools described in [99] and we utilize one of them, data profiles. Moré and Wild [99] suggest a test of convergence for derivative-free optimization methods, which compares objective function values with f_L. Given a tolerance $\tau > 0$, the inequality (4.7) holds if the decrease in the objective function from the starting point $\theta^{(0)}$ to θ is at least $(1-\tau)$ times the best possible reduction.

$$f(\theta^{(0)}) - f(\theta) \geq (1-\tau)(f(\theta^{(0)}) - f_L)\,. \tag{4.7}$$

The data profile $d_s(\alpha)$ is defined as the fraction of problems that an algorithm s solves with a computational budget corresponding to $\alpha > 0$ simplex gradient evaluations, [99].

The data profiles for FMINSEARCH, ORBIT, SID–PSM, NOMAD and GSS–CI for the time-dependent problems and for $\tau \in \{10^{-1}, 10^{-3}, 10^{-5}\}$ are given in Figure 4.5. In general, the required accuracy τ has a strong impact on the percentages of problems solved for any budget. Another difference in the data profiles for distinct values of τ is that the graphs of the data profiles flatten out, and the differences in the data profiles decrease with an increasing number of function evaluations for $\tau = 10^{-1}$, but not for $\tau = 10^{-3}$ and $\tau = 10^{-5}$. With the computational budget of $85\cdot(m+1)$ function evaluations, about 76%-99% of the problems are solved with accuracy $\tau = 10^{-1}$, about 16%-74% with accuracy $\tau = 10^{-3}$ and 2%-50% with accuracy $\tau = 10^{-5}$. All solvers solve most problems at least with accuracy 10^{-1}. It appears to be challenging to achieve accuracy 10^{-5} with the given computational budget.

Figure 4.5 (a) shows the fraction of problems, for which a solver reduces the objective function by 90% compared with the best reduction obtained by any solver. This corresponds to convergence criterion (4.7) for $\tau = 10^{-1}$. With a budget of 85 simplex gradient evaluations, SID–PSM solves about 99% of the problems and NOMAD, GSS–CI, FMINSEARCH and ORBIT solve approximately 97%, 95%, 93% and 76% of the problems, respectively. If a problem is solved, then this takes on average 11 simplex

gradient evaluations for FMINSEARCH, 21 for ORBIT, 16 for SID–PSM, 22 for NOMAD and 13 for GSS–CI. For $\tau = 10^{-1}$, NOMAD and ORBIT perform best for small computational budgets ($\alpha \leq 4$), FMINSEARCH, GSS–CI and SID–PSM perform well for $\alpha \in [11,61]$ and SID–PSM and NOMAD solve the largest percentage of the problems overall. Summarizing, the direct search methods are more successful than the trust-region method, and FMINSEARCH and GSS–CI are, overall, fastest.

Considering the data profiles for accuracy $\tau = 10^{-5}$ in Figure 4.5 (c) and the percentages of problems solved with accuracy $\tau = 10^{-3}$ in Figure 4.5 (b), we observe that the directional direct search methods GSS–CI and SID–PSM solve the highest number of problems overall. With a budget of at most 85 simplex gradient evaluations, average gradient and curvature information, which is utilized in a direct search framework, seems to allow for more accurate solutions for the time-dependent problems.

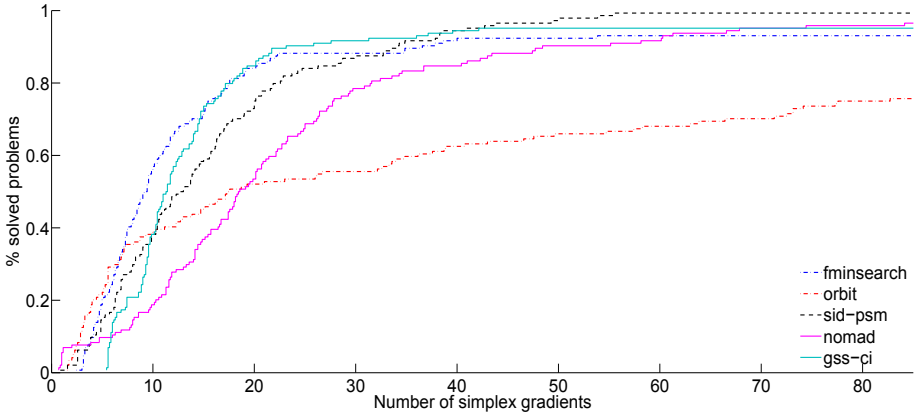
Comparison with the Noisy Benchmark Problems

In this section, FMINSEARCH, GSS–CI, ORBIT, NOMAD and SID–PSM are employed for the noisy benchmark problems (4.3) with deterministic noise and relative noise level $\varepsilon_f = 10^{-5}$. As for the time-dependent problems, the solvers terminate after $85 \cdot (m+1)$ function evaluations. We compare the data profiles for the noisy benchmark problems in Figure 4.6 with the data profiles for the time-dependent problems from Figure 4.5.

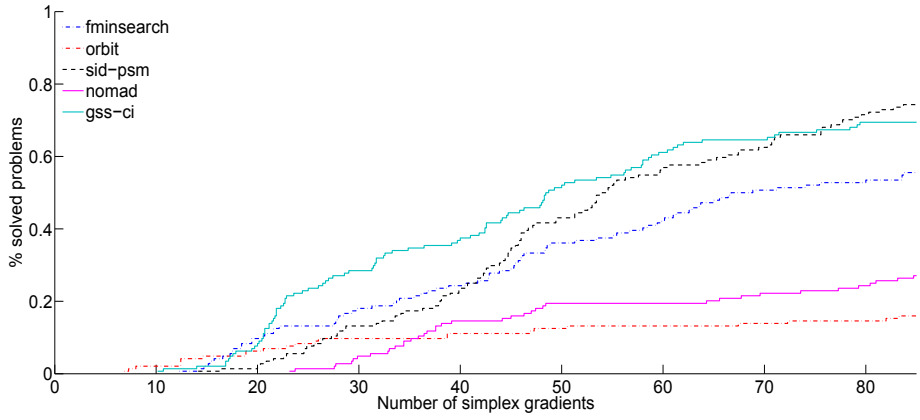
The data profiles for $\tau = 10^{-1}$ in Figure 4.6 (a), for $\tau = 10^{-3}$ in Figure 4.6 (b) and for $\tau = 10^{-5}$ in Figure 4.6 (c) illustrate that for all values of α considered here, the noisy benchmark problems are on average solved more frequently than the time-dependent problems. Especially for the two high accuracies $\tau = 10^{-3}$ and 10^{-5}, the time-dependent problems are considerably more challenging than the noisy benchmark problems. For the two values of τ and all values of α, no derivative-free optimization method solves a higher percentage of the time-dependent problems than of the noisy benchmark problems.

For all $0 \leq \alpha \leq 85$, we consider the difference between the largest percentage of problems solved by any solver and the smallest percentage of problems solved by any solver. The variations in performances for the time-dependent problems illustrated in Figures 4.5 are compared with the differences in the data profiles for the noisy benchmark problems presented in Figure 4.6. For $\tau = 10^{-1}$, the differences in percentages of solved problems are larger for the time-dependent problem than for the noisy benchmark problems if $\alpha \geq 20$. The variations between the data profiles for $\tau = 10^{-3}$ and for the time-dependent problems exceed the ones for the same accuracy and the benchmark problems if $\alpha \geq 50$. Summarizing, for many computational budgets, and for $\tau = 10^{-1}$ and $\tau = 10^{-3}$, we observe larger variations in the performances for the time-dependent problems than for the noisy benchmark problems.

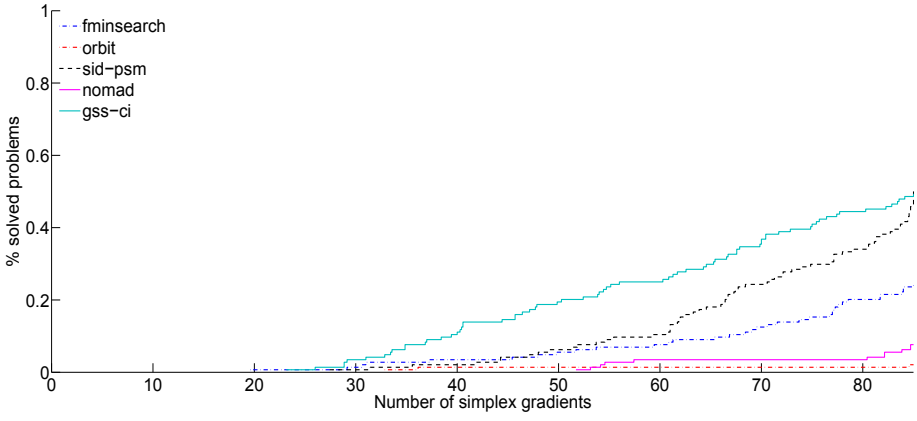
The ranking of the solvers regarding the highest fraction of problems solved for a given computational budget differs for the two types of problems. For example, for $\tau = 10^{-5}$, GSS–CI solves the highest number of time-dependent problems for almost all computational budgets considered, while SID–PSM solves most of the benchmark problems for all $\alpha \in [0,85]$. For $\tau = 10^{-5}$, NOMAD solves more of the noisy benchmark problems than FMINSEARCH, while the opposite is observed for the problems with ODE constraints. Overall and in comparison to the other derivative-free optimiza-



(a) Data profile for $\tau = 10^{-1}$.

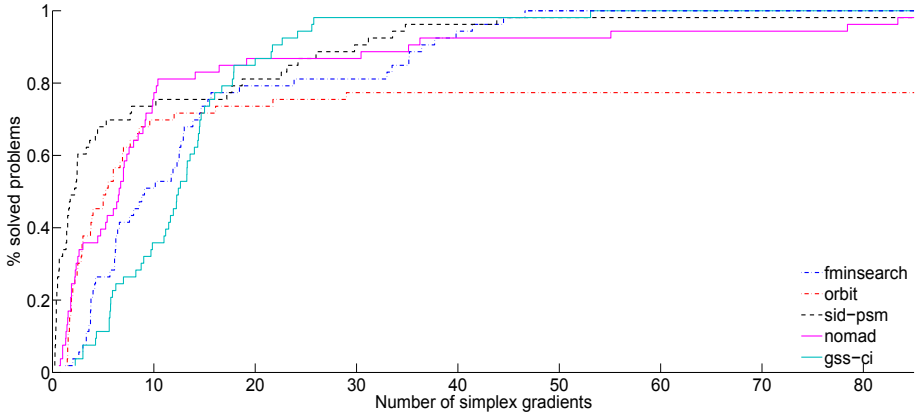


(b) Data profile for $\tau = 10^{-3}$.

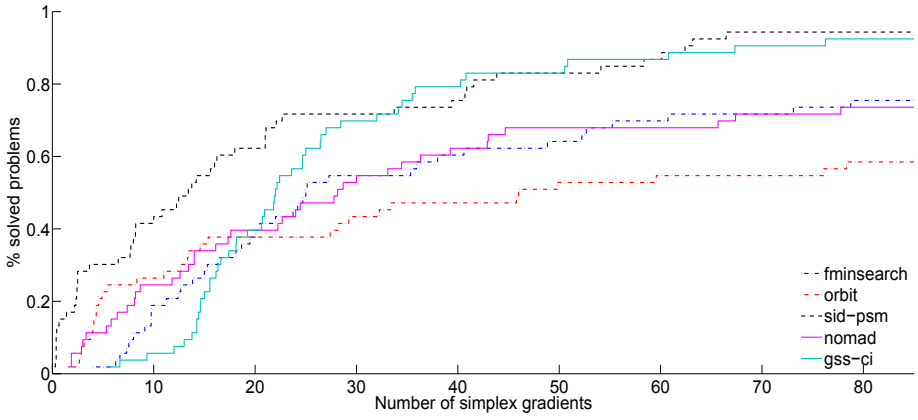


(c) Data profile for $\tau = 10^{-5}$.

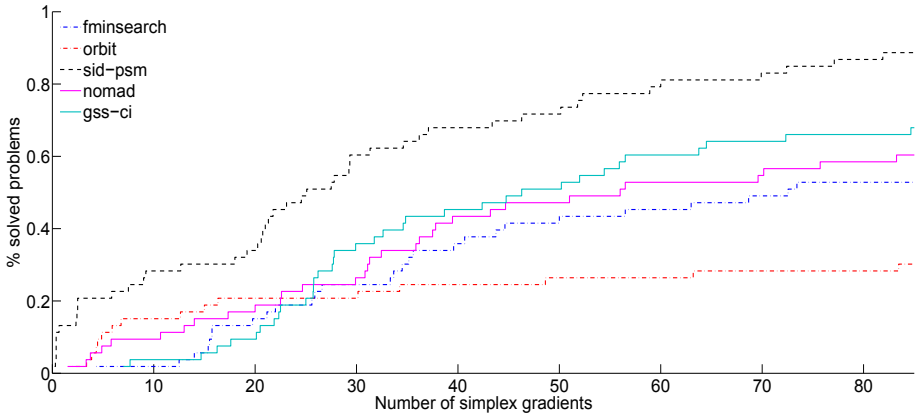
Figure 4.5: Percentage of problems solved by a solver for the time-dependent problems.



(a) Data profile for $\tau = 10^{-1}$.



(b) Data profile for $\tau = 10^{-3}$.



(c) Data profile for $\tau = 10^{-5}$.

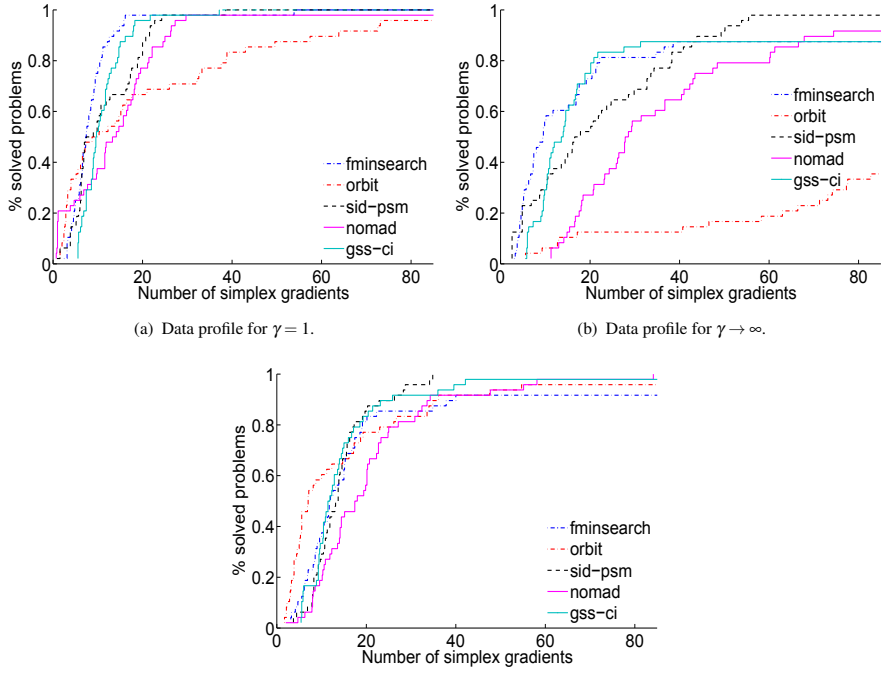
Figure 4.6: Percentage of problems solved by a solver for the noisy benchmark problems.

tion methods, the Nelder–Mead algorithm and GSS–CI perform better for the parameter estimation problems than for the noisy benchmark problems and the time-dependent problems are more challenging than the noisy benchmark problems, especially for NOMAD and ORBIT.

Our results are in accordance with previous results from employing derivative-free optimization for the set of benchmark problems. An earlier version of NOMAD than the one employed in this chapter has been observed to be slower than SID–PSM for a set of benchmark problems by Moré and Wild [99] (but not the noisy benchmark problems) and for level of accuracy 10^{-3}, [39]. As in [143], ORBIT is observed to be effective when the computational budget is small. Furthermore, SID–PSM has also previously been observed to be able to find accurate solutions for the noisy benchmark problems, [46].

Nonlinearity of the differential equation

The parameter γ indicates the nonlinearity of the differential equation. In case of $\gamma \to \infty$, the objective function may be highly sensitive to the values of θ_3 and θ_4, as illustrated in Figure 4.3.



(a) Data profile for $\gamma = 1$.

(b) Data profile for $\gamma \to \infty$.

(c) Data profile for $\gamma = 2$.

Figure 4.7: Percentage of problems solved for $\tau = 10^{-1}$ and for the time-dependent problems, classified by the nonlinearity of the differential equation.

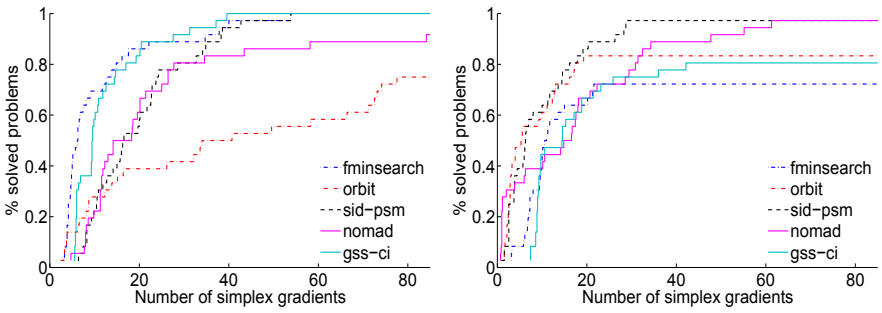
Figure 4.7 presents the percentage of problems solved by any of the five derivative-free optimization methods with accuracy $\tau = 10^{-1}$ for the subsets of the problems

defined by $\gamma = 1$, $\gamma = 2$ or the case $\gamma \to \infty$. Figures 4.7 (a) and (c) illustrate that all derivative-free optimization methods are efficient for the cases $\gamma = 1$ and $\gamma = 2$. The convergence test (4.7) is on average satisfied less frequently for the subset of problems described as the case $\gamma \to \infty$ than for the two cases $\gamma = 1$ and $\gamma = 2$. This is in accordance with the observation from section 4.4.1 that the problems with highly nonlinear differential equation are more challenging than the cases $\gamma = 1$ and $\gamma = 2$.

For all $\alpha \in [0, 85]$, ORBIT solves fewer of the problems for the case $\gamma \to \infty$ than for the other two cases. The nonlinearity of the differential equation has a considerable impact on the performance of the model-based derivative-free optimization method. Interestingly, ORBIT performs better for the problems with $\gamma = 2$ than with $\gamma = 1$. With a budget of 85 simplex gradient evaluations, NOMAD and SID–PSM frequently find high quality solutions for the three subsets of problems. However, the two directional direct search methods which are enhanced by a model-based search steps are less efficient than FMINSEARCH and GSS–CI for the problems with highly nonlinear ODE.

Robustness to the initial iterate

Figure 4.8 presents the data profiles for the problems with $\theta^{(0)} = 0.8\tilde{\theta}$ and the problems with $\theta^{(0)} = 0.5\mathbf{b}$ for $\tau = 10^{-1}$.



(a) Data profile for $\theta^{(0)} = 0.8\tilde{\theta}$.

(b) Data profile for $\theta^{(0)} = 0.5\mathbf{b}$.

Figure 4.8: Percentage of problems solved for $\tau = 10^{-1}$ and for the time-dependent problems, classified by the initial iterate.

We observe considerable variations in the performances of the five solvers with the initial iterate. Figure 4.8 illustrates that the problems with $\theta^{(0)} = 0.5\mathbf{b}$ are, on average, solved less frequently than the problems with $\theta^{(0)} = 0.8\tilde{\theta}$. NOMAD and SID–PSM are the only solvers to find a solution that satisfies the convergence test for $\tau = 10^{-1}$ for more than 90% of the problems with the largest distance between $\tilde{\theta}$ and the initial iterate.

ORBIT performs better for the problems defined by $\theta^{(0)} = 0.5\mathbf{b}$ than for the problems with $\theta^{(0)} = 0.8\tilde{\theta}$. As illustrated in Figure 4.8, FMINSEARCH and GSS–CI are less efficient and solve fewer of the problems with $\theta^{(0)} = 0.5\mathbf{b}$ than the other solvers. ORBIT is observed to be most robust and FMINSEARCH and GSS–CI are least robust to the choice of the initial iterates.

4.4.3 Hidden constraints of the parameter estimation problems

In the following, we examine the effect of hidden constraints on the data profiles from Figure 4.5. In paper C, hidden constraints are observed for about 8% of the optimization problems, more precisely the ones with $\gamma \to \infty$ and $\theta^{(0)} = 0.5\mathbf{b}$.

During optimization with NOMAD, hidden constraints arise for five of six of the problems defined by the data from Tables A.1–A.6, initial iterate $\theta^{(0)} = 0.5\mathbf{b}$ and the ODE in case of $\gamma \to \infty$. In total, 15 evaluations of the objective function failed, and in 22 cases, the value of the objective function exceeds the largest finite number represented in 64 bit floating-point precision, causing an overflow. For all time-dependent problems with hidden constraints, NOMAD obtains an objective function value that satisfies the convergence test (4.7) for $\tau = 10^{-1}$. For some of the problems with hidden constraints, the objective function value returned by NOMAD after 85 simplex gradient evaluations is the best solution found by any solver. The good performance of NOMAD is not surprising, since mesh adaptive direct search methods are designed to handle hidden constraints.

The objective function is defined for all $\theta \in \Theta$ considered by GSS–CI, FMINSEARCH, SID–PSM and ORBIT. That the number of unsuccessful attempts to evaluate the objective function is highest for NOMAD may to be due to the fact that NOMAD moves further away from the initial iterate $\theta^{(0)}$ than any of the other solvers.

FMINSEARCH and GSS–CI handle bound constraints as if they were hidden constraints. For 90 of the 144 time-dependent problems, FMINSEARCH evaluates the objective function at points that are not elements of Θ. This corresponds to about 7% of the function evaluations. For GSS–CI, hidden constraints occur for 130 problems and on average for about 8% of the function evaluations. For each problem, the objective function is not defined for at most 35% of the points considered by FMINSEARCH and at most 33% of the points tested by GSS–CI.

In general, minimization of a barrier function may be challenging for the simplex-based optimization method, [28], and GSS–CI is not designed to build a rich sequence of poll directions. However, for a large range of computational budgets considered in the numerical simulations, GSS–CI solves at least as many problems with accuracy $\tau = 10^{-5}$ as any other solver. GSS–CI obtains the lowest objective function value for more problems than any other algorithm and it performs well for the medium and higher values for α. For all values of τ and for most computational budgets, the simplex-based optimization method solves more problems than at least two other implementations of derivative-free optimization methods. FMINSEARCH and GSS–CI found iterates satisfying the convergence test (4.7) with $\tau = 10^{-1}$ for more than 90% of the time-dependent problems within 85 simplex gradient evaluations. Despite treating bound constraints as hidden constraints, the two solvers perform well for the time-dependent problems.

4.5 Discussion

4.5.1 Comparison of the results from two sets of numerical simulations

We now compare the data profiles from paper C for $\alpha \leq 85$ with the results from the numerical simulations from this chapter. The aim is to test, whether key results from paper C remain valid for the new simulations.

We employ the same versions of SID–PSM, ORBIT and NOMAD as in paper C, and the data profiles of the three algorithms in paper C resemble the ones in this chapter. The numerical simulations from paper C differ from the ones in this chapter by the number of solvers and by the approach used by FMINSEARCH to handle bound constraints.

In this chapter, we consider an additional solver (GSS–CI), which provides a better solution than the other solvers for 43% of the problems when FMINSEARCH minimizes a barrier function. This explains why the data profiles from paper C for SID–PSM, ORBIT and NOMAD are for some $\alpha \in [0, 85]$ larger than data profiles from this chapter for the same solvers and the same accuracy. However, for the three implementations of derivative-free optimization methods, the data profiles from paper C are smaller or equal to the data profiles from this chapter for most values of $0 \leq \alpha \leq 85$ and for all values of τ. Since the additional solver explains only smaller data profiles in this chapter than in paper C, this can only be due to the fact that in paper C, the Nelder–Mead algorithm solves an unconstrained problem, while we employ FMINSEARCH to minimize the barrier function (4.6).

The unconstrained problems considered in paper C are less challenging for FMINSEARCH than minimization of a barrier function for the time-dependent problems. For the numerical simulations from paper C, FMINSEARCH solves the highest number of problems for all three accuracies and for a computational budget of at most 85 simplex gradient evaluations. When minimizing a barrier function, FMINSEARCH ranks lower than SID–PSM for most computational budgets and accuracies. Overall, the differences in the data profiles of FMINSEARCH are small. The performances may be explained by the possibility to return $\theta \notin \Theta$ with small objective function value and the fact that optimization of barrier functions can be challenging for the simplex-based optimization method.

Several observations from paper C are confirmed by the numerical simulations in this chapter. This includes results concerning the behaviour of the derivative-free optimization methods with variations of the initial iterate and the value of γ. In paper C and in this chapter, it is observed that the Nelder–Mead algorithm performs better for the time-dependent problems than for the noisy benchmark problems.

As described in paper C, the variations in the performances of the solvers for the time-dependent problems are overall larger than for the noisy benchmark problems, and the time-dependent problems are more challenging than the benchmark problems. These observations remain valid for the new set of derivative-free optimization methods and the modified time-dependent problems.

4.5.2 Performance of the derivative-free optimization methods

The high variations in the data profiles for the time-dependent problems can be linked to the fact that the problems are challenging. This can be explained as follows. Moré and Wild [99] observe that the differences in data profiles may decrease with increasing number of function evaluations if many solvers find a sufficiently accurate solution for a sufficiently large computational budget. For all $\alpha \in [0, 85]$, the time-dependent problems are on average solved less often than the noisy benchmark problems. It is possible that more functions evaluations are required to solve the time-dependent problems than the noisy benchmark problems.

The following challenges may be encountered for optimization of the time-dependent problems. The problems (4.1)–(4.2) have hidden and bound constraints. Hidden constraints have a small effect on the performance of the derivative-free optimization methods for the problems considered in this chapter. Only two of the considered implementations, NOMAD and SID–PSM, have first-order global convergence properties for bound constraint problems.

Another difference between the time-dependent problems and the noisy benchmark problems is the type of noise. For the noisy benchmark problems, the level of noise is 10^{-5} and independent of θ. As illustrated in Figure 4.2, the level of computational noise of the objective function (4.1) may be considerably larger than 10^{-5} and varies with θ. Model-based derivative-free optimization may be less successful than direct search for accuracies that are about the level of noise than for smaller accuracies, [99]. A strong impact of computational noise on the performance of the solvers could explain why ORBIT is overall outperformed and NOMAD less efficient for the time-dependent problems. That SID–PSM and GSS–CI use average gradient and curvature information, respectively, may explain why the two methods obtain the most accurate solutions.

The problems with $\gamma \to \infty$ are more challenging than the problems with $\gamma = 1$ or $\gamma = 2$. The differential equation (4.2) is highly nonlinear in case of $\gamma \to \infty$, and the problems with $\gamma \to \infty$ may be poorly scaled, as observed in Figure 4.3. Spherical trust-regions or trust-regions of form $\{x \in \mathbb{R}^m | \|x - x_k\|_\infty \leq \Delta_k\}$, as employed by ORBIT, are not suitable for poorly scaled problems, [108, Ch. 4]. Newton's method is scale invariant, e.g. [108, Ch. 2], and individual step lengths are useful for poorly scaled functions, [60]. Poor scaling of the time-dependent problems with $\gamma \to \infty$ might explain why GSS–CI is more efficient for the problems with $\gamma \to \infty$ than NOMAD, ORBIT and SID–PSM.

The error bounds for the model functions utilized by ORBIT and in the search step of SID–PSM and NOMAD depend on the nonlinearity of the objective function, [44, 143]. While radial basis functions may perform better than quadratic models for multi-modal functions, [144], ORBIT is the only solver that relies on an accurate approximation of the objective function. High nonlinearity of the objective function could be a reason for ORBIT to be less efficient and for FMINSEARCH to rank high.

4.6 Concluding remarks

4.6.1 On the numerical simulations

The conclusions from this chapter are only valid for the specific implementations of the optimization methods and the specific optimization problems. The set of time-dependent problems is limited in the variety of the number of variables, the initial iterates, and the feasible set.

Further methods, such as NEWUOA [113], or methods designed to solve nonlinear least squares problems, such as POUNDERS [87] or DFLS [147], might be considered. However, a nonlinear least squares function is one of several types of objective functions commonly used for parameter estimation for population dynamic models, [73, Ch. 6]. The optimization methods employed in this chapter can be applied for a broad variety of objective functions. The set of solvers is chosen with the aim of testing the overall performance of derivative-free methods for parameter estimation in the singularly perturbed differential equations.

A remark is in place about the fact that a solver might fast approach a local solution with considerably higher objective function value than f_L. In general, data profiles consider the reduction of the objective function value and contain no information about whether a local solution of an optimization problem is approached.

The following software has been updated. SID–PSM now incorporates an option to build sequences of sets of poll directions that, after normalization, become asymptotically dense in the unit sphere, [49]. The implementation of mesh adaptive direct search utilizes now a vector of m mesh and poll size parameters, and the efficiency of NOMAD has been improved, [15].

The challenges of problem (4.1)–(4.2) are typical for parameter estimation for ODEs that are solved numerically. The objective functions are noisy due to the numerical solution of the differential equations and may, for several reasons, not be defined for all $\theta \in \Theta$. In general, the convergence may be slow if the objective function is more sensitive to a subset of the parameters than to others, e.g. [125, Sect. 4.2]. Derivatives or their approximations may inform about the local sensitivity of the objective function with respect to one point in the feasible set. However, unless the objective function is linear in the parameters, the local measure is only valid for a single point in the feasible set, and methods for investigation of the global sensitivity of the objective function with respect to the parameters require additional function evaluations, [123].

4.6.2 On estimation of parameters in population dynamic models

Implausible results are frequently obtained when estimating parameters in population dynamic models, [98], as, for example, in [64]. Heuristic approaches, which for example consist of keeping some parameters fixed and successively increasing the number of parameters to be addressed, [58], have been adopted in the fisheries literature, [98]. Derivative-free optimization methods such as mesh adaptive direct search provide a nonheuristic approach to handling hidden constraints.

Models for the dynamics of fish populations may incorporate several species and environmental factors with temporal and spatial variations at several distinct scales. Spatially and temporarily structured populations may be described by partial differen-

tial equations, [75]. Derivative-free optimization has been successfully applied for estimation of parameters in partial differential equations, e.g. [59, 87]. Individual-based models describe the characteristics of every individual in the population over its lifetime. Mechanisms such as growth may be described by piecewise-defined functions, thus resulting in nondifferentiable dynamical systems, [130]. We may anticipate that derivative-free optimization may in the future be applied more frequently for parameter estimation in population dynamic models.

Chapter 5

Emergent properties of a multi-stage population dynamic model

5.1 Introduction

A stage structured population may consist of several early life stages (such as eggs, larvae and juveniles) and several adult classes. For stages $i = 1,\dots,n$, denote by $\mathbf{z}_t = (z_{1,t}\dots z_{n,t}) \in \mathbb{R}^n_{\geq 0}$ the vector of numbers of individuals at time $t \in \mathbb{N}_0$. Based on standard assumptions from the fisheries literature, e.g. [116, Ch. 5], the population dynamics may be described by a difference equation (5.1) with nonnegative real-valued functions g_i and initial condition $\mathbf{z}_0 \in \mathbb{R}^n_{\geq 0}$.

$$z_{i,t} = g_i(\mathbf{z}_t, \mathbf{z}_{t-1}),\ i = 1,\dots,n\,. \tag{5.1}$$

The functions g_i represent birth and death processes and the transitions from one stage to the next.

Define recruitment as the transition of fish to stage $m \in \mathbb{N}$, with $1 \leq m < n$. A traditional assumption is that recruitment is given by a SR function

$$\begin{aligned} &r : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \ \text{ s.t. } \ z_{m,t} = \ r(S_{t-\tau})\,, \\ &\quad \text{for all } t \geq \tau, \text{ for } \tau \geq 1 \ \text{ and for all solutions } \mathbf{z}_t \text{ of } (5.1), \end{aligned}$$

where

$$S_t = \sum_{i=m}^{n} w_i\, z_{i,t}, \ \text{with } w_i \geq 0.$$

An example of w_i is the fraction of spawners in age class $i = m,\dots,n$. In this case, S_t denotes the total number of spawners at time t. If we assume $w_i = 1$ for $i = m,\dots,n$, then S_t is the total number of adults at time t. Parameter $\tau \geq 0$ may, as an example, denote the average time from spawning to recruitment.

From chapter 3, we know that in case no SR function exists, recruitment may be a function of the number of adults in several age classes. We distinguish between two parent-progeny relationships: a SR function, which describes recruitment as a function of a weighted sum of the number of adults in several age-classes and recruitment as a function of numbers of adults $z_{m,t}$, ... , $z_{n,t}$. In comparison to chapter 3, we consider a discrete-time model for several early life stages.

Our aim is to derive sufficient conditions for the existence of the two parent-progeny functions, and a framework that is applicable for a broad range of multi-stage models from the literature.

We start this chapter by considering an example of a life cycle of fish, which includes the egg, larval, juvenile and adult stages. The dynamics of the population are described by a discrete-time multi-stage model. In section 5.3, we state the general form of the emergent parent-progeny relationship. Sufficient conditions for the existence of two parent-progeny functions are given in section 5.4. Sections 5.2–5.4 present results from manuscript B.

In section 5.5, we give examples of discrete-time multi–stage models with SR function and compare them with multi–stage models from the fisheries literature. A SR function and recruitment as a function of the number of adults in several age-classes have been considered for a continuous-time two-stage population dynamic model in chapter 3. In section 5.6, we relate the sufficient conditions for existence of a SR function from chapter 3 to the ones from manuscript B.

Comparisons of models for populations with several early life stages to SR functions can be found in the literature. The question of existence of a SR function has not been considered for these multi-stage models, since the models are based on assumptions that guarantee their existence. We compare the assumptions underlying the multi-stage models from the literature with the results from manuscript B in section 5.7.

5.2 A general multi-stage model

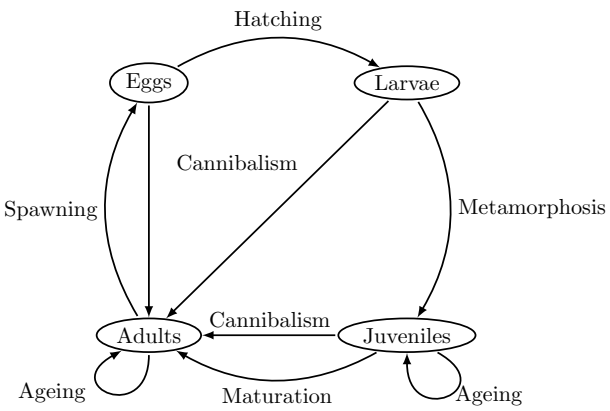


Figure 5.1: A schematic diagram of the life cycle of fish. Hatching, metamorphosis, maturation and spawning are the principal developmental processes that cause the transition from one life stage to the next. The linkages between the adults and the early life stages through cannibalism are indicated. (From manuscript B.)

An example of a schematic overview of the life history of fish, which involves several early life stages, is shown in Figure 5.1. The arrows indicate interactions between individuals of distinct life stages (such as cannibalism) and transitions from one stage to the next. As an example, egg production and survival may be influenced by the quantity

and structure of the parent population. We may consider several age-classes of juveniles and adults. The loop from the vertex that represents adults to itself indicates transitions from one age-class of adults to the next age-class of adults. Ageing of juveniles is represented by the loop from the vertex which represents juveniles. A SR relationship for the life cycle 5.1 subsumes the processes spawning, hatching, metamorphosis, maturation, cannibalism and ageing of juveniles. The life cycle 5.1 is considered in manuscript B, where a description of the life history of fish from an ecological point of view can be found.

$\mathbb{S} = \{e, l, j, a\}$	Set of life stages: eggs (e), larvae (l), juveniles (j), adults (a)
$t \in \mathbb{N}_0$	Simulation time
$m \in \mathbb{N}_0$	Number of juvenile age-classes
$n \in \mathbb{N}_0$	Index of the oldest age-class
$k \in \{0, .., m, .., n\}$	Indices of age-classes
$E_t \in \mathbb{R}_{\geq 0}$	Number of eggs at time t
$L_t \in \mathbb{R}_{\geq 0}$	Number of larvae at time t
$J_{k,t} \in \mathbb{R}_{\geq 0}$	Number of juveniles in class $k = 0, \ldots, m-1$ at time t
$N_{k,t} \in \mathbb{R}_{\geq 0}$	Number of adults in class $k = m, \ldots, n$ at time t
$0 < p_k < 1$	Proportion of adult female population in class $k = m, \ldots, n$
$f_k \geq 0$	Number of eggs produced per adult female of age k
$s_i : \mathbb{R}^d \to (0, 1]$	Probability of surviving stage $i \in \{e, l, (j, k_1), (a, k_2)\}$, for $k_1 = 0, \ldots, m-1$ and $k_2 = m, \ldots, n$, in each time step, with $d \in \mathbb{N}$
$\mathbf{J}_t = (J_{0,t}, \ldots, J_{m-1,t})$	Numbers of juveniles at time t
$\mathbf{N}_t = (N_{m,t}, \ldots, N_{n,t})$	Numbers of adults at time t

Table 5.1: Nomenclature for a general multi-stage model.

In manuscript B, a population dynamic model is introduced, which represents the dynamics of a stage structured population as indicated in Figure 5.1. The population consists of eggs, larvae, juveniles and adults. We distinguish between $m \in \mathbb{N}_0$ age-classes of juveniles and denote by $n \in \mathbb{N}$ the oldest age-class. The indices $m, \ldots, n$ refer to age-classes of adults. The population model describes egg production, mortality and transitions from one stage to the next.

The model is based on the following biological considerations. Total egg production is the sum of egg productions per age-class. Egg production per age-class $k = m, \ldots, n$ is proportional to fecundity $f_k \geq 0$ and the proportion $p_k \in (0, 1)$ of spawners in age-class k. Survival of eggs and early life stages might be limited by cannibalism and is described as a function of the number of adults. Food availability and competition may be limiting factors for survival of (feeding) larvae and juveniles. Survival of larvae and juveniles are described as functions of the number of larvae and juveniles, respectively. Fish reach the juvenile age-class 0 in the same time step as they are spawned (if at all). If surviving, juveniles and adults age by one in every simulation time step (with exception of fish of age n).

A nomenclature for the general discrete-time population model is given in Table 5.1. Numbers of individuals in a stage or age-class are denoted by uppercase letters, while functions and parameters are denoted by lowercase letters. The indices e, l, j and a refer to the life stages eggs (e), larvae (l), juveniles (j) and adults (a). The probabilities s_i of surviving a stage $i \in \{e, l, (j, k_1), (a, k_2)\}$ for $k_1 = 0, \ldots, m-1$ and $k_2 = m, \ldots, n$ are

described by positive functions. The notation in this chapter differs from the notation used in manuscript B, where survival of adults and fecundity are described as functions of age.

The general multi-stage model is defined by (5.2)–(5.8) with initial condition $(\mathbf{J}_0, \mathbf{N}_0) \in \mathcal{D} \subset \mathbb{R}^{n+1}_{\geq 0}$. Recruitment is denoted by R_t , and we assume $R_t = N_{m,t}$.

$$E_t = \sum_{k=m}^{n} f_k \, p_k \, N_{k,t-1} = e(\mathbf{N}_{t-1}), \tag{5.2}$$

$$L_t = E_t \cdot s_e(\mathbf{N}_{t-1}), \tag{5.3}$$

$$J_{0,t} = L_t \cdot s_l(L_t, \mathbf{N}_{t-1}), \tag{5.4}$$

$$J_{k,t} = J_{k-1,t-1} \cdot s_{j,k-1}(J_{k-1,t-1}, \mathbf{N}_{t-1}), \qquad k = 1, \ldots, m-1, \tag{5.5}$$

$$N_{m,t} = \begin{cases} J_{m-1,t-1} \cdot s_{j,m-1}(J_{m-1,t-1}, \mathbf{N}_{t-1}), & m \geq 1, \\ J_{0,t}, & m = 0, \end{cases} \tag{5.6}$$

$$N_{k,t} = s_{a,k-1} \cdot N_{k-1,t-1}, \qquad k = m+1, \ldots, n-1, \tag{5.7}$$

$$N_{n,t} = s_{a,n-1} \cdot N_{n-1,t-1} + s_{a,n} \cdot N_{n,t-1}. \tag{5.8}$$

5

We define a matrix model (introduced in section 2.2.2) to represent the dynamics of numbers of juveniles and adults as described by the general multi-stage model. To this aim, we define the probability $j_0 : \mathbb{R}^{n-m}_{\geq 0} \to (0, 1]$ of survival from the egg stage to the juvenile class 0 by (5.9). From the definition of the model by (5.2)–(5.4) and the definition of j_0, the number of juveniles in age-class 0 is given by (5.10).

$$j_0(\mathbf{N}_t) = s_e(\mathbf{N}_t) \cdot s_l\Big(e(\mathbf{N}_t) \cdot s_e(\mathbf{N}_t), \mathbf{N}_t\Big). \tag{5.9}$$

$$J_{0,t} = \Big(\sum_{k=m}^{n} f_k \, p_k \, N_{k,t-1}\Big) \cdot j_0(\mathbf{N}_{t-1}). \tag{5.10}$$

Equations (5.10) and (5.5)–(5.8) represent a matrix model (5.11) for the number of juveniles and adults.

$$\begin{pmatrix} \mathbf{J}_t \\ \mathbf{N}_t \end{pmatrix} = \begin{pmatrix} 0 & \ldots & 0 & \beta_{m+1} & \ldots & \beta_{n-1} & \beta_n \\ s_{j,1} & & & & & & 0 \\ & \ddots & & & & & \\ & & s_{j,m-1} & & & & \\ & & & s_{a,m} & & & \\ & & & & \ddots & & \\ 0 & & & & & s_{a,n-1} & s_{a,n} \end{pmatrix} \begin{pmatrix} \mathbf{J}_{t-1} \\ \mathbf{N}_{t-1} \end{pmatrix}. \tag{5.11}$$

The coefficients in the first row of the matrix are 0 and β_k as defined by (5.12).

$$\beta_k = f_k p_k j_0(\mathbf{N}_{t-1}), \quad k = m+1, \ldots, n. \tag{5.12}$$

The coefficients in the first subdiagonal are $s_{j,k}(J_{k-1,t-1}, \mathbf{N}_{t-1})$ for $k = 1, \ldots, m-1$ and $s_{a,k}$ for $k = m, \ldots, n-1$.

5.3 The parent-progeny relationship for the general multi-stage model

In the following, we consider the parent-progeny relationship that is admitted by the general multi-stage model.

The directed graph 5.2 illustrates the relationship between recruitment at time t and numbers of adults as given by the general multi-stage model. The functional relationships described by (5.2)–(5.6) are represented by edges. The directed graph includes an edge from a vertex denoted by v_1 to a vertex denoted by v_2 if the variable denoted by v_2 is a function of the variable denoted by v_1. As an example, equation (5.6) gives $N_{m,t}$ as a function of $\mathbf{N}_{t-1}$ and $J_{m-1,t-1}$. The graph includes vertices representing $\mathbf{N}_{t-1}$ and $J_{m-1,t-1}$ and a directed edge starting from each of the two nodes and ending at the node that represents $N_{m,t}$. The edges from nodes which represent numbers of adults are labelled by the functional relationship which they represent. From (5.2)–(5.6), we see that R_t is obtained from E_{t-m}, L_{t-m}, $J_{0,t-m}$, $J_{1,t-m+1}$, ... , $J_{m-1,t-1}$, $\mathbf{N}_{t-m-1}$, ... , $\mathbf{N}_{t-1}$. From every vertex denoted by $\mathbf{N}_u$, for $u = t-m-1,\ldots,t-1$, we have a path to the vertex denoted by $N_{m,t}$. The directed graph shows that $R_t = N_{m,t}$ is in general a function of $\mathbf{N}_{t-m-1},\ldots,\mathbf{N}_{t-1}$.

In case of $m = 0$, the directed graph includes only the vertices labelled E_t, L_t, $J_{0,t}$, $\mathbf{N}_{t-1}$ and the edges between these nodes.

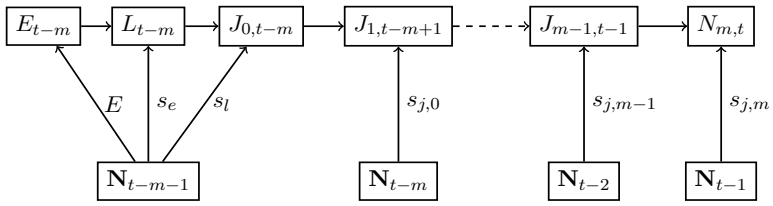


Figure 5.2: A schematic diagram of the relationship between recruitment at time t and numbers of adults at time $t-1,\ldots,t-m$, as given by (5.2)–(5.6).

An equation for the parent-progeny relationship described by the general multi-stage model can be derived from (5.2)–(5.6). The probability $j_k : \mathbb{R}_{\geq 0}^{k(n-m)} \to (0,1]$ of survival from the egg stage to the juvenile age-class k, for $k = 1,\ldots,m$, is defined recursively by (5.9) and (5.13).

$$\begin{aligned} j_k(\mathbf{N}_{t-1},\ldots,\mathbf{N}_{t-k-1}) = \; & j_{k-1}(\mathbf{N}_{t-2},\ldots,\mathbf{N}_{t-k-1}) \\ & \cdot s_{j,k-1}\Big(e(\mathbf{N}_{t-k-1})\cdot j_{k-1}(\mathbf{N}_{t-2},\ldots,\mathbf{N}_{t-k-1}),\, \mathbf{N}_{t-1}\Big), \quad k = 1,\ldots,m. \end{aligned} \tag{5.13}$$

With this definition, we obtain that recruitment is given by (5.14) for all $t \geq m+1$, as shown in manuscript B.

$$R_t = e(\mathbf{N}_{t-m-1}) \cdot j_m(\mathbf{N}_{t-1},\ldots,\mathbf{N}_{t-m-1}). \tag{5.14}$$

For the general multi-stage model, recruitment $R_t = N_{m,t}$ is in general a function of $\mathbf{N}_{t-m-1},\ldots,\mathbf{N}_{t-1}$.

5.4 Necessary and sufficient conditions for the existence of a SR function for the general multi-stage model

We state the following two hypotheses about the parent-progeny relationship described by model (5.2)–(5.8).

Hypothesis 1: There exists an adult-recruitment (AR) function

$$\tilde{r} : \mathbb{R}_{\geq 0}^{n-m} \to \mathbb{R}_{\geq 0} \text{ s.t. } N_{m,t} = \tilde{r}(N_{m,t-\tau}, \ldots, N_{n,t-\tau})$$
$$\text{for all } t \geq \tau, \text{ for } \tau \in [1, m+1] \text{ and for all solutions of (5.2)–(5.8).}$$

Hypothesis 2: Additionally, function $\tilde{r}$ is the composite of a function $r : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and the linear function $b : \mathbb{R}_{\geq 0}^{n-m} \to \mathbb{R}_{\geq 0}$ given by,

$$b(N_{m,t}, \ldots, N_{n,t}) = \sum_{k=m}^{n} w_k N_{k,t}, \text{ with } w_k \geq 0. \tag{5.15}$$

Hypothesis 1 entails that recruitment is a function of numbers of adults. The SR function for the general multi-stage model is a function

$$r : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \text{ s.t. } N_{m,t} = r(S_{t-\tau}) ,$$
$$\text{for all } t \geq \tau, \text{ for } \tau \in [1, m+1] \text{ and for all solutions of (5.2)–(5.8), where}$$
$$S_t = b(N_{m,t}, \ldots, N_{n,t}) = \sum_{k=m}^{n} w_k N_{k,t}, \text{ with } w_k \geq 0 .$$

Function r in Hypothesis 2 is the SR function.

If a SR function exists for a multi-stage model, we may define the AR function as $\tilde{r} = r \circ b$. Thus, a SR function exists if and only if Hypotheses 1 and 2 are true. A necessary condition for the existence of a SR function is the existence of an AR function.

Because R_t is in general given as a function of $\mathbf{N}_{t-m-1}$, …, $\mathbf{N}_{t-1}$, we require the delay τ in the definition of an AR and a SR function to be in interval $[1, m+1]$.

In the remainder of section 5.4 and in section 5.5, we assume $(E_t, L_t, J_{0,t}, \ldots, J_{m-1,t}, N_{m,t}, \ldots, N_{n,t})$ to be a solution of the initial value problem given by (5.2)–(5.8) with $(\mathbf{J}_0, \mathbf{N}_0) \in \mathcal{D}$, for nonempty $\mathcal{D} \subset \mathbb{R}_{\geq 0}^{n+1}$. The parameters p_k and f_k and the functions s_i are assumed to be as described in Table 5.1 and egg production is assumed to be nonzero, i.e., we have $\sum_{k=m}^{n} f_k p_k > 0$.

In the following, we state sufficient conditions for Hypotheses 1 and 2.

Existence of an AR function

By Theorem 2 in manuscript B, each of the following conditions **(C1)**, **(C2)** and **(C3)** is sufficient for Hypothesis 1.

(C1) $\mathbf{N}_t = \mathbf{N}_0$ for all $t \in [0, m+1]$,

(C2) $s_{j,k}(J_{k,t}, \mathbf{N}_t) = s_{j,k}(J_{k,t})$ for all $k = 0, \ldots m-1$ and for all $t \geq 0$,

(C3) $m = 0$ and $N_{m,t} = J_{0,t}$ is given by (5.4).

More specifically, it is shown in manuscript B that if at least one of **(C1)**, **(C2)** or **(C3)** is true, then recruitment is given by an AR function described by (5.16).

$$N_{m,t} = e(\mathbf{N}_{t-m-1}) \cdot j_m(\mathbf{N}_{t-m-1}), \text{ for } t \geq m+1. \tag{5.16}$$

Recruitment is total egg production multiplied with the probability of survival from the egg stage to recruitment. It is a function of the number of adults at the time of spawning. Hypothesis 1 is true for $\tau = m+1$.

As shown in manuscript B, **(C1)** implies that $\mathbf{N}_t = \mathbf{N}_0$ for all $t \geq 0$ and for all solutions of the initial value problem given by (5.2)–(5.8) with $(\mathbf{J}_0, \mathbf{N}_0) \in \mathcal{D}$. In this case, recruitment is constant with respect to time and Hypotheses 1 and 2 are true.

Assuming proposition **(C2)** to be true, we can remove the edges labelled $s_{j,k}$, with $k = 0, \ldots, m$ from the directed graph 5.2. In this case, there are no paths from the nodes denoted by $\mathbf{N}_{t-m}, \ldots, \mathbf{N}_{t-1}$ to the vertex denoted by $N_{m,t}$. The graph shows recruitment as a function of $\mathbf{N}_{t-m-1}$.

If proposition **(C3)** is true, then equation (5.10) from section 5.3 is identical to (5.16) and gives recruitment as a function of $\mathbf{N}_{t-1}$.

A multi-stage model for which none of the propositions **(C1)–(C3)** are true, may not admit an AR function. An example of the general multi-stage model for which Hypothesis 1 is false for all $\tau \in [1, m+1]$ is given in manuscript B.

Existence of a SR function

Assume that an AR function exists, i.e., recruitment is a function of $\mathbf{N}_{t-m-1}$ given by (5.16). **(C1)** implies that Hypothesis 2 is true, and we assume that either **(C2)** or **(C3)** are true. **(C3)** implies that the AR function is the product of e, s_e and s_l.

As shown in manuscript B, a sufficient condition for an AR function $\tilde{r}$ given by (5.16) to be a SR function is the logical conjunction of the following propositions **(C4)**(a), **(C4)**(b), and **(C4)**(c).

(C4)(a) $\exists\, \tilde{e} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ s.t. $e(\mathbf{x}) = \tilde{e}(b(\mathbf{x}))$, for all $\mathbf{x} \in \mathbb{R}_{\geq 0}^{n-m}$,

(C4)(b) $\exists\, \tilde{s}_e : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ s.t. $s_e(\mathbf{x}) = \tilde{s}_e(b(\mathbf{x}))$, for all $\mathbf{x} \in \mathbb{R}_{\geq 0}^{n-m}$,

(C4)(c) $\exists\, \tilde{s}_l : \mathbb{R}_{\geq 0}^2 \to \mathbb{R}_{\geq 0}$ s.t. $s_l(y, \mathbf{x}) = \tilde{s}_l(y, b(\mathbf{x}))$ for all $\mathbf{x} \in \mathbb{R}_{\geq 0}^{n-m}$, $y \in \mathbb{R}_{\geq 0}$.

If any of the conditions **(C4)**(a), **(C4)**(b), or **(C4)**(c) is false, then the AR function may not have the structure of a SR function, as shown in manuscript B.

5.5 Examples of discrete-time multi-stage models

In the following, we state the parent-progeny relationship for three examples of the general multi-stage model. In two cases, the matrix model (5.11) for numbers of juveniles and adults corresponds to one of the matrix models considered in section 2.2.2.

Example I

Example I of a general multi-stage model is given by (5.2)–(5.8) under the assumption that s_e, s_l and $s_{j,k}$, for $k = 0,\dots,m-1$, are constant with respect to L, $\mathbf{J}$ and $\mathbf{N}$.

In this case, **(C2)**, **(C4)**(b), and **(C4)**(c) are true. This implies Hypothesis 1. The AR function is given by (5.16). From the definition (5.9) and (5.13) of the probabilities j_k and the definition (5.2) of the function e, we obtain the AR function given by (5.17).

$$N_{m,t} = \Big(\sum_{k=m}^{n} f_k\, p_k\, N_{k,t-1}\Big) \cdot s_e \cdot s_l \cdot \prod_{k=0}^{m-1} s_{j,k}\,. \tag{5.17}$$

Whether **(C4)**(a) is true and the parent-progeny relationship is a SR function, depends on the definition of S_t. If S_t is defined as the total number of fecund fish, i.e., $S_t = \sum_{k=m}^{n} p_k N_{k,t}$, and $f_k = f$ for all $k = m,\dots,n$, then the AR function is a SR function. Under the additional assumption that $p_k = p$ for all $k = m,\dots,n$, recruitment could alternatively be formulated as a function of the total number of adults $S_t = \sum_{k=m}^{n} N_{k,t}$.

For Example I, the matrix model (5.11) for the dynamics of juveniles and adults is linear. If we assume $s_{a,n} = 0$, we obtain a Leslie matrix model with nonzero coefficients only in the first row and the first subdiagonal and with $\beta_k = 0$ for $k = 0,\dots,m-1$. For the Leslie matrix model, recruitment is a function of the total number of adults if $\beta_i = 0$ for $i = 1,\dots,m-1$ and $\beta_m = \beta_{m+1} = \cdots = \beta_n$, [116, Ch. 5]. From (5.12), we see that the assumption $\beta_m = \beta_{m+1} = \cdots = \beta_n$ implies $f_m p_m = \cdots = f_n p_n$ and **(C4)**(a), (b), and (c).

Example II

Example II of a general multi-stage model is derived from the assumptions that $s_{j,k}$, for $k = 0,\dots,m-1$, are constant with respect to $\mathbf{N}$. Furthermore, we assume that s_e and s_l are functions of egg production $\sum_{k=m}^{n} f_k p_k N_{k,t}$.

In this case, the probability of survival of eggs to juvenile class 0 is given by (5.18) and a function of egg production.

$$j_0(\mathbf{N}_t) = s_e\Big(\sum_{k=m}^{n} f_k\, p_k\, N_{k,t}\Big) \cdot s_l\Big(\sum_{k=m}^{n} f_k\, p_k\, N_{k,t}\Big)\,. \tag{5.18}$$

Since **(C2)** is true, the parent-progeny relationship is the AR function given by (5.16). From (5.18), we see that a SR function as described by Hypothesis 2 exists for $b(N_{m,t},\dots,N_{n,t}) = \sum_{k=m}^{n} f_k p_k N_{k,t}$.

In case of $m = 0$, the dynamics of the adults are given by the matrix model (5.19) with β_k as defined by (5.12) and $j_0(\mathbf{N}_t)$ given by (5.18).

$$\mathbf{N}_t = \begin{pmatrix} \beta_0(\mathbf{N}_t) & \beta_1(\mathbf{N}_t) & \dots & \beta_n(\mathbf{N}_t) \\ s_{j,1} & & & 0 \\ & \ddots & & \\ 0 & & s_{a,n-1} & s_{a,n} \end{pmatrix} \mathbf{N}_{t-1}\,. \tag{5.19}$$

For $s_{a,n} = 0$, we obtain the nonlinear matrix model considered in section 2.2.2. This model is derived from the assumption that survival of prerecruits is a function of a weighted sum of the number of adults, [116, Ch. 7.4].

Example III

As a third example of a general multi-stage model, we consider a general multi-stage model with $m > 0$ and s_e, s_l and $s_{j,k}$, for $k = 0,\dots,m-1$, constant with respect to $\mathbf{N}$. Furthermore, we assume that $S_t = \sum_{k=m}^{n} f_k p_k N_{k,t}$. In this case, the dynamics of the juveniles and recruitment are given by (5.20).

$$\begin{aligned} J_{0,t} &= e(\mathbf{N}_t)\cdot j_0(e(\mathbf{N}_t))\,, \\ J_{k,t} &= J_{k-1,t-1}\ \cdot\ s_{j,k-1}(J_{k-1,t-1})\,,\ k = 1,\dots,m-1\,, \\ N_{m,t} &= J_{m-1,t-1}\ \cdot\ s_{j,m-1}(J_{m-1,t-1})\,. \end{aligned} \tag{5.20}$$

The propositions **(C2)** and **(C4)**(a), (b) and (c) are true and recruitment is a function of S_{t-m-1}.

5.6 SR functions for slow-fast population dynamic models

Under a different set of assumptions than the one used to derive the general multi-stage model, we derived a continuous-time two-stage model in section 3.3.1. The population dynamic model is given by an initial condition $\mathbf{N}(0) \in \mathbb{R}^n_{\geq 0}$ and (3.14)–(3.16) in chapter 3. The equations are re-stated by (5.21).

$$\begin{aligned} \dot{N}_0(t) =\ & -\alpha N_0(t) + \frac{1}{\varepsilon}\sum_{i=1}^{n} l_i f_i N_i(t) \\ & - \begin{cases} \frac{1}{\varepsilon} m_0 N_0(t)\left(1 + \frac{1}{\gamma m_0}\sum_{i=1}^{n} p_i N_i(t)\right)^{\gamma}, & \text{case } \gamma > 0\,, \\ \frac{1}{\varepsilon} m_0 N_0(t)\cdot \exp\left(\frac{1}{m_0}\sum_{i=1}^{n} p_i N_i(t)\right), & \text{case } \gamma \to \infty\,, \end{cases} \\ \dot{N}_i(t) =\ & \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t), \quad i = 1,\dots,n\,. \end{aligned} \tag{5.21}$$

Here, $N_0(t)$ denotes the number of prerecruits and $N_i(t)$, $i = 1,\dots,n$ denotes numbers of adults in n age-classes at time $t \in \mathbb{R}_{\geq 0}$. For the slow-fast population dynamic model, recruitment is defined as $R(t) = \alpha N_0(t)$ and $S(t) = \sum_{i=1}^{n} f_i N_i(t)$ is the total number of fish contributing to egg production. Here, $f_i \geq 0$ denotes the fraction of spawners in age-class i.

Similarly to Hypotheses 1 and 2, we may state two hypotheses about the parent-progeny relationship described by the continuous-time model.

Hypothesis 3: There exists an adult-recruitment (AR) function

$$\begin{aligned} &\tilde{r} : \mathbb{R}^n_{\geq 0} \to \mathbb{R}_{\geq 0} \text{ s.t. } R(t) = \tilde{r}\,(N_1(t-\tau),\dots,N_n(t-\tau))\,, \\ &\quad \text{for all } t \geq \tau, \text{ for } \tau \geq 0 \text{ and for all solutions } \mathbf{N}_t \text{ of (5.21)}. \end{aligned}$$

Hypothesis 4: Additionally, function $\tilde{r}$ is the composite of a function $r : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and the linear function $b : \mathbb{R}^n_{\geq 0} \to \mathbb{R}_{\geq 0}$ given by,

$$b\,(N_1(t),\dots,N_n(t)) = \sum_{i=1}^{n} f_i\, N_i(t), \text{ with } f_i \geq 0.$$

Let $\mathbf{N}(t)$ be a solution of the slow-fast population dynamic model (5.21). From section 3.3.2, we know that for $\varepsilon > 0$ sufficiently small, there exists a function h defined on a suitable set $K \supset [0,\bar{S}]^n$ such that if $\mathbf{N}(0) \in h([0,\bar{S}]^n) \times [0,\bar{S}]^n$ and $N_i(t) \leq \bar{S}$ for all $i = 1,\ldots,n$ and all $t \geq 0$, then $N_0(t) = h(N_1(t),\ldots,N_n(t))$ for all $t \geq 0$. The function $\tilde{r} = \alpha h$ is an AR function with $\tau = 0$. The sufficient conditions for the existence of an AR function are requirements on the initial population $\mathbf{N}(0)$, the boundedness of the solution of the ODE and the ratio ε between the two time scales of the slow-fast system.

As shown in paper A, the function $\tilde{r} = \alpha h$ can be approximated by function $\tilde{r}^0 : \mathbb{R}^n_{\geq 0} \to \mathbb{R}_{\geq 0}$ defined by (5.22), which is identical to (3.17) from section 3.3.1. Here and in the following, we consider the case $\gamma > 0$. The case $\gamma \to \infty$ can be treated analogously.

$$R(t) = \frac{\alpha}{m_0}\Big(\sum_{i=1}^{n} l_i f_i N_i(t)\Big)\left(1 + \frac{1}{\gamma m_0}\Big(\sum_{i=1}^{n} p_i N_i(t)\Big)\right)^{-\gamma}. \tag{5.22}$$

In section 3.4, the following sufficient condition for function $\tilde{r}^0$ to be a SR function is suggested.

(C6) For all $i = 1,\ldots,n$, either $p_i = c f_i$ and $l_i = l$ with $c, l > 0$, or $p_i = f_i = l_i = 0$.

The sufficient condition **(C6)** for $\tilde{r}^0$ to be a SR function relates to the sufficient condition **(C4)** for Hypothesis 2 as follows. For the slow-fast population dynamic model, the rate of egg production $e(\mathbf{N}(t))$ is given by (5.23) and the mortality rate m_γ of prerecruits is defined by (5.24).

$$e(N_1(t),\ldots,N_n(t)) = \sum_{i=1}^{n} l_i f_i N_i(t)\,. \tag{5.23}$$

$$m_\gamma(N_1(t),\ldots,N_n(t)) = m_0\left(1 + \frac{1}{\gamma m_0}\sum_{i=1}^{n} p_i N_i(t)\right)^{\gamma}. \tag{5.24}$$

We define the following two propositions:

(C5)(a) $\exists\, \tilde{e} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ s.t. $e(\mathbf{x}) = \tilde{e}(b(\mathbf{x}))$, for all $\mathbf{x} \in \mathbb{R}^n_{\geq 0}$,

(C5)(b) $\exists\, \tilde{m}_\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ s.t. $m_\gamma(\mathbf{x}) = \tilde{m}_\gamma(b(\mathbf{x}))$, for all $\mathbf{x} \in \mathbb{R}^n_{\geq 0}$.

For $b : \mathbb{R}^n_{\geq 0} \to \mathbb{R}$ given by $b(\mathbf{N}(t)) = \sum_{i=1}^n f_i N_i$, **(C6)** implies **(C5)**(a)–(b). This can be verified by defining $\tilde{e}(x) = lx$ and $\tilde{m}_\gamma(x) = m_0(1 + (c/\gamma m_0)x)^\gamma$.

We observe that the sufficient conditions for an adult-recruitment relationship to be a SR function given in manuscript B, reformulated for the continuous-time model, imply the conditions given in chapter 3. The sufficient conditions from chapter 3 include sufficient conditions given in [139] and paper A.

5.7 Discussion

A link between a SR function and a multi-stage population model for fish has been established by Paulik [110]. The model in [110] describes the dynamics of a species

with three spatially separated life stages. It is derived from the assumptions that the number of individuals in the first stage is a function of the spawning stock size and that the number of individuals in stages 2 and 3 are functions of the number of individuals in stages 1 and 2, respectively. Using the notation from the introduction of this chapter, the model by Paulik may be described by (5.25).

$$z_1 = g_1(S)\,,\ \ z_2 = g_2(z_1)\,,\ \ R = g_3(z_2)\,. \tag{5.25}$$

For the multi-stage model (5.25), the SR function is the composite of g_1, g_2 and g_3.

Brooks and Powers [30] described the SR functions emergent from a multi-stage population model. Under the additional assumption that $S_t = e(\mathbf{N}_t)$, and denoting by $J_{k,t}$ the number of individuals in an arbitrary prerecruit stage, equation (5.20) is satisfied for their model. We obtain Example III of a general multi-stage model. For Example III, conditions **(C2)** and **(C4)**(a), (b), and (c) are true and guarantee the existence of a SR function.

Here and in section 5.5, we have demonstrated that the framework from manuscript B can be applied for several discrete-time models from the fisheries literature.

The general multi-stage model (5.2)–(5.8) includes two important extensions in comparison to a model of form (5.25) or (5.20). Firstly, we consider an age structured class of adults. Secondly, the mortality of prerecruits may be a function of numbers of adults. This assumption is, for example, suitable when mortality is correlated with the spawning stock biomass as, e.g., for North East Arctic cod, [25].

For a multi-stage model, which includes several classes of juveniles and which assumes survival to be a nonlinear function of numbers of adults, a SR function might not exist. This is illustrated in paper B, where the R_t-S_{t-m-1} relationship is not a function, but the trajectories of the general multi-stage model spiral inwards. The SR relationship is affected by the temporal variations of the structure and state of the adult population.

Chapter 6

Summary of papers and future outlook

In this thesis, we have addressed two challenges that are prerequisites for the adoption of multiple time-scale multi-stage population dynamic models in fisheries science. Firstly, a description of the parent-progeny relationship is required. An important question is whether the assumption of the existence of a (specific) SR function, or an adult-recruitment function can be justified. Secondly, one needs to identify a suitable approach for estimating the parameters in a slow-fast population dynamic model. This task includes the formulation of the optimization problem, choice of a numerical method to solve the differential equation, and choice of an optimization method.

The scientific contributions of the thesis are summarized, and open problems for future research are proposed in the following.

6.1 Paper A. A parametrized stock-recruitment relationship derived from a slow-fast population dynamic model

A singularly perturbed differential equation may describe the dynamics of a two-stage (prerecruits and adults) population of fish. The evolution of prerecruits is assumed to be considerably faster than ageing and mortality of adult fish.

Paper A has addressed the existence of a SR function, or an adult-recruitment function, assuming that population dynamics are given by a singularly perturbed differential equation. Sufficient conditions for trajectories of a generic slow-fast population dynamic model to approach or remain on the graph of a parent-progeny function have been stated. The assumptions concern the set of fixed points of the differential equation that describes the fast dynamics of prerecruits and the smoothness of the differential equation. The results require the use of geometric singular perturbation theory.

In paper A, we have provided an answer to whether a dome-shaped SR function can be explained by slow-fast population dynamic models. The parametrized slow-fast population dynamic model introduced in paper A admits a rich class of SR functions, including the Beverton–Holt, the Ricker, and the Deriso SR functions. If the mortality rate of prerecruits increases more than linearly with the number of adults, then a SR function is dome-shaped. Strictly increasing SR functions result when the mortality rate is a linear function of the numbers of adults and have been previously observed for slow-fast population dynamic models.

The solutions of the new slow-fast population dynamic model are nonnegative for

nonnegative initial values. For suitable values of the parameters, the differential equation has a positive fixed point.

6.2 Paper B. Emergent properties of a multi-stage population dynamic model

In manuscript B, a difference equation model has been used to describe the dynamics of a fish population that includes several life stages (egg, larvae, juveniles, and adults). The model represents egg production, mortality, and transitions from one stage to the next. Two hypotheses about the parent–progeny relationship admitted by the general discrete time multi–stage model have been defined.

It has been shown that the general discrete-time multi-stage model may not admit an adult-recruitment function. Sufficient conditions for the existence of this type of parent-progeny relationship have been given. An adult-recruitment function exists if the time between spawning and recruitment is sufficiently short, or mortality of juveniles is constant with respect to the adult population. An adult-recruitment function may not have the structure of a SR function unless egg production and mortality of eggs and larvae are functions of the spawning stock size.

The manuscript investigated the population dynamics by incorporating functional relationships from the fisheries literature, which describe egg production and mortality in the four life stages. Our results show that temporal variations of the structure of the adult population may influence the parent–progeny relationship and prevent the existence of a SR function.

6.3 Paper C. Derivative-free optimization for population dynamic models

When estimating parameters in the parametrized slow-fast population dynamic model, the objective function is not defined for all parameters that satisfy some bound constraints. We deal with the additional challenge that the objective function is noisy due to the numerical solution of the differential equations and nondifferentiability of the functions involved.

The performances of the following four well-known derivative-free optimization methods, three direct search methods and one trust-region method, have been compared in paper C for the time-dependent problems. A derivative-free trust-region method utilizes nonconvex model functions, which are obtained from interpolation. A simplex-based optimization method adapts a simplex to the local curvature of the objective function. The following two direct search methods employ quadratic model functions, but their convergence does not rely on the quality of the models. Mesh adaptive direct search methods are designed to handle problems with hidden constraints and nonsmooth objective functions. A generalized pattern search method utilizes average gradient information.

In paper C, a set of time-dependent problems has been defined based on variations of the data sets with and without observational errors and the nonlinearity of the differen-

tial equation. For each problem, several initial iterates have been considered. Solutions of the singularly perturbed and stiff differential equations have been approximated by numerical differentiation formulas (NDF's). Criteria for a good performance of the derivative-free optimization methods are the ability to improve the objective function as a function of the computational budget, the robustness to the initial iterate, and the robustness to the nonlinearity of the differential equation.

It has been shown that the derivative-free trust-region method is more successful in solving the time-dependent problems than the direct search methods when the computational budget is small, but the problems with highly nonlinear differential equation are challenging for the model-based method. For medium and large computational budgets (exceeding 150 function evaluations), the direct search methods outperform the derivative-free trust-region method. The Nelder–Mead simplex method is the most successful and fastest solver overall for the time-dependent problems. However, occasionally, the simplex-based optimization method, which is the only algorithm to consider an unconstrained problem, has returned infeasible iterates.

Higher variations in performances of the solvers for the time-dependent problems than for a set of noisy benchmark problems demonstrate that estimation of parameters in singularly perturbed differential equations is, at least for some of the optimization methods, more challenging than the standard benchmark problems.

6.4 Additional scientific contributions and concluding remarks

At several places, additional questions have arisen from the discussion of the three papers. Results concerning the topics of paper A and manuscript B are summarized in section 6.4.1. Scientific contributions additional to the ones from paper C can be found in section 6.4.2. In section 6.4.3, we sum up how the main findings of the thesis as a whole contribute to the research field.

6.4.1 On stock-recruitment relationships emergent from stage structured population models

As shown in chapter 3, the parametrized slow-fast population dynamic model introduced in paper A may be derived from ecological assumptions. We have considered slow-fast population models additional to the parametrized slow-fast population dynamic model and the model by Touzeau and Gouzé. For the aim of parameter estimation, the slow-fast population dynamic model has been reformulated in terms of a smaller number of parameters than used in paper A.

In chapter 5, we have established that several well-known multi-stage models are examples of the model from manuscript B. It has been shown that the framework from manuscript B can be employed for a rich class of discrete-time models. The sufficient conditions for an adult-recruitment function to be a SR function also apply to the parametrized slow-fast population dynamic model.

6.4.2 On derivative-free optimization for population dynamic models

In chapter 4, we have described several types of hidden constraints of optimization problems addressing parameters in differential equations. Because of their efficiency, we have employed NDF methods of orders 1 to 5 to solve the stiff differential equations. Numerical simulations show that the stability properties of the algorithms are overall suitable for numerical solution of the slow-fast population dynamic models. However, the differential equations represent several classes of dynamic behaviour. We have observed high- and low-amplitude noise in the objective function, and the level of noise varies with the parameters in the differential equation.

It has been demonstrated that the derivative-free optimization methods find sufficiently accurate solutions for the optimization problems with hidden constraints and noisy objective functions. This includes several classes of derivative-free optimization methods without convergence theory for problems with hidden constraints. Existing algorithms with convergence theory in case of hidden constraints build sets of sample points that become asymptotically dense in the feasible set or an asymptotically dense set of poll directions. Model-based derivative-free methods generate sets of well scattered sample points in a neighbourhood of the current iterate. Direct search methods employ sequences of poll directions, which are positive spanning sets. We have found that several classes of derivative-free optimization methods are suitable for problems with unknown feasible sets and estimation of parameters in differential equations.

Chapter 4 includes numerical simulations additional to the ones from paper C. The new set of optimization methods contains a generating set search method, which utilizes average curvature information. A considerable modification in the implementation of the numerical simulations in comparison to paper C is that two algorithms, the Nelder–Mead algorithm and the generating set search method, now treat bound constraints as hidden constraints. We have observed that the simplex-based optimization method solves fewer of the bound constrained problems than of the unconstrained problems and is now outperformed by several direct search methods. However, many of the key results from paper C are confirmed by the new numerical simulations.

The time-dependent problems with numerical noise arising from numerical solution of the differential equations are, for several reasons, more challenging than a class of noisy benchmark problems. The problems considered in this thesis have bound and hidden constraints, and some of the objective functions are more sensitive to some parameters than to others. Furthermore, we have observed larger variations in the level of noise for the time-dependent problems than for the noisy benchmark problems. The properties of the optimization problems may explain why direct search methods overall outperform model-based derivative-free methods for estimation of parameters in the slow-fast population dynamic models.

The mesh adaptive direct search method with convergence theory for nonsmooth problems with hidden constraints solves second most problems overall, and the highest percentage of the problems with largest distance between an a priori guess for a solution of the optimization problem and the initial iterate. However, a method without convergence theory in case of hidden constraints, the generating set search method, treats bound constraints as hidden constraints and is one of the two most successful solvers for the time-dependent problems. We have found that the two methods which employ average gradient and curvature information in a direct search framework, the

generating set search method and the generalized pattern search method, achieve the most accurate solutions for the noisy optimization problems.

6.4.3 On derivative-free optimization and multiple time-scale dynamics of stage structured populations

Central to the choice of methodology in this thesis is the assumption that a model for a population of fish that includes the dynamics of prerecruits is a dynamical system with two time-scales. The assumption justifies the use of geometric singular perturbation theory for investigation of the parent-progeny function admitted by a continuous-time population model. Some of the adaptive strategies of methods for the solution of initial value problems are typical for stiff systems of differential equations.

We have suggested a new framework for quantification of the parent-progeny relationship in fish populations. Instead of assuming the existence of a SR function, a mathematical model for a stage structured population is derived from ecological assumptions. The models considered in this thesis describe the temporal variations of multi-stage demographic mechanisms, which act at several distinct temporal scales and may influence the SR relationship. The parameters in the population dynamic model may be estimated using the initial value approach, derivative-free optimization and fast and accurate methods for the solution of differential equations. With the frameworks proposed in paper A and B, the assumption of existence a SR function (or an adult-recruitment function) may be justified or questioned. Even if the assumptions that would support a SR function fail to be validated, we obtain parameters that describe the parent-progeny relationship.

6.5 Outlook

Open problems for future research can be seen in several areas.

A further challenge for modelling the SR relationship is to describe the relationship between environmental or physical factors and the dynamics of prerecruits and adults. The parametrized slow-fast population dynamic model could be extended to incorporate the multiple time-scale dynamics of a varying environment and ecosystem. The temporal variations of, e.g. temperature, food availability, or wind direction and intensity may be represented by additional variables or time-dependent parameters. The resulting differential equation might be nonautonomous. Using geometric singular perturbation theory or Tikhonov's theorem, the relationship between recruitment, spawning stock, and additional variables might be described. The emergent model for recruitment could be compared with the literature.

The following open problems can be identified for parameter estimation for population dynamic models. Hidden constraints and nondifferentiability of objective functions have been reported in the fisheries literature. To test the applicability of derivative-free optimization methods for parameter estimation for discrete-time models and data used in fisheries science remains a topic for further research. A set of benchmark problems could be designed for the specific purpose of application in fisheries science. As criteria for the performance of optimization methods, we could consider the ability to find sufficiently accurate solutions for a rich class of optimization problems. We might

observe that the selected benchmark problems have more parameters than the optimization problems considered in this thesis.

The following open problem for future research can be identified in the field of parameter estimation for differential equations. The level of noise of the objective function depends on the algorithm used for the solution of ODEs. Additionally, the algorithm employed for evaluation of an objective function may affect the number of occurrences of hidden constraints. As an example, some implicit Runge–Kutta methods have better stability properties than the NDF methods employed in this thesis and might result in fewer instances of hidden constraints. It could be tested, how the choices of the method for the numerical solution of ODEs and the tolerance for the one-step error affect the performance of optimization methods. The results of the numerical simulations might guide the choice and development of methods for parameter estimation for differential equations.

Appendix A

Data for the optimization problems

The following data is utilized in chapter 4. It is obtained as described in paper C.

Data set 1: $\mathbf{d}_0 = (8.98646736, 25, 60)$ and case $\gamma \to \infty$.

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	8.98646736	25.00000000	60.00000000
t_2	27.69436176	16.10227985	36.88118817
t_3	37.94783180	16.77225746	24.18020381
t_4	40.79748468	18.64289672	18.20469792
t_5	41.46307462	19.81182432	15.62469156
t_6	41.60914443	20.38333939	14.56338504
t_7	41.63561072	20.63703369	14.14053742
t_8	41.63678330	20.74389482	13.97614135
t_9	41.63430935	20.78735220	13.91354268
t_{10}	41.63238718	20.80455470	13.89015267

Table A.1: Data set 1 with $\sigma = 0$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	7.46854134	24.81846959	57.81713724
t_2	27.64581085	13.94542887	36.42847537
t_3	39.24581391	16.25875730	24.19284702
t_4	39.39028446	18.84947634	17.92929358
t_5	41.21717674	19.12502156	14.88162472
t_6	42.62890869	20.56848854	14.92309032
t_7	42.87948633	21.88169622	14.24031778
t_8	41.45694573	20.50287043	14.32913487
t_9	43.89054086	20.90293915	14.86927886
t_{10}	39.21394628	22.83228515	15.76334086

Table A.2: Data set 1 with $\sigma = 1$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	4.43268929	24.45540876	53.45141173
t_2	27.54870904	9.63172691	35.52304977
t_3	41.84177814	15.23175698	24.21813344
t_4	36.57588401	19.26263559	17.37848490
t_5	40.72538100	17.75141602	13.39549104
t_6	44.66843721	20.93878682	15.64250086
t_7	45.36723755	24.37102129	14.43987849
t_8	41.09727060	20.02082165	15.03512193
t_9	48.40300388	21.13411306	16.78075123
t_{10}	34.37706448	26.88774605	19.50971724

Table A.3: Data set 1 with $\sigma = 3$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	1.39683723	24.09234794	49.08568621
t_2	27.45160723	5.31802495	34.61762417
t_3	44.43774237	14.20475666	24.24341986
t_4	33.76148357	19.67579484	16.82767622
t_5	40.23358525	16.37781049	11.90935736
t_6	46.70796573	21.30908511	16.36191140
t_7	47.85498876	26.86034636	14.63943919
t_8	40.73759547	19.53877288	15.74110899
t_9	52.91546690	21.36528698	18.69222360
t_{10}	29.54018268	30.94320695	23.25609362

Table A.4: Data set 1 with $\sigma = 5$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	23.72928711	44.71996069
t_2	27.35450542	1.00432299	33.71219857
t_3	47.03370660	13.17775634	24.26870629
t_4	30.94708312	20.08895409	16.27686755
t_5	39.74178951	15.00420495	10.42322367
t_6	48.74749425	21.67938340	17.08132194
t_7	50.34273998	29.34967143	14.83899990
t_8	40.37792034	19.05672411	16.44709605
t_9	57.42792992	21.59646089	20.60369597
t_{10}	24.70330087	34.99866785	27.00247000

Table A.5: Data set 1 with $\sigma = 7$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	23.36622629	40.35423518
t_2	27.25740361	0.10000000	32.80677297
t_3	49.62967083	12.15075602	24.29399271
t_4	28.13268268	20.50211334	15.72605887
t_5	39.24999376	13.63059942	8.93708999
t_6	50.78702277	22.04968169	17.80073248
t_7	52.83049119	31.83899650	15.03856061
t_8	40.01824521	18.57467533	17.15308311
t_9	61.94039295	21.82763480	22.51516833
t_{10}	19.86641907	39.05412874	30.74884638

Table A.6: Data set 1 with $\sigma = 9$

Data set 2: $\mathbf{d}_0 = (7.38258415, 0.5, 0.6)$ and case $\gamma \to \infty$.

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	7.38258415	0.50000000	0.60000000
t_2	32.98613094	6.26457416	1.32903058
t_3	47.42159120	15.46143490	4.42924855
t_4	45.26782326	20.03692690	8.24919503
t_5	43.06329820	21.15697960	10.98806725
t_6	42.12142048	21.19703473	12.52325802
t_7	41.78473061	21.05138536	13.27985596
t_8	41.67489429	20.93622750	13.62472871
t_9	41.64173250	20.87150297	13.77384258
t_{10}	41.63262694	20.83986054	13.83586676

Table A.7: Data set 2 with $\sigma = 0$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	5.86465813	0.31846959	0.10000000
t_2	32.93758003	4.10772318	0.87631778
t_3	48.71957331	14.94793474	4.44189176
t_4	43.86062304	20.24350652	7.97379069
t_5	42.81740032	20.47017683	10.24500041
t_6	43.14118474	21.38218388	12.88296329
t_7	43.02860622	22.29604790	13.37963631
t_8	41.49505672	20.69520312	13.97772223
t_9	43.89796401	20.98708992	14.72957876
t_{10}	39.21418604	22.86759099	15.70905495

Table A.8: Data set 2 with $\sigma = 1$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	2.82880608	0.10000000	0.10000000
t_2	32.84047822	0.10000000	0.10000000
t_3	51.31553754	13.92093442	4.46717819
t_4	41.04622260	20.65666577	7.42298201
t_5	42.32560458	19.09657130	8.75886672
t_6	45.18071326	21.75248217	13.60237384
t_7	45.51635743	24.78537297	13.57919702
t_8	41.13538159	20.21315434	14.68370929
t_9	48.41042703	21.21826384	16.64105113
t_{10}	34.37730424	26.92305189	19.45543133

Table A.9: Data set 2 with $\sigma = 3$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	32.74337641	0.10000000	0.10000000
t_3	53.91150177	12.89393410	4.49246461
t_4	38.23182215	21.06982502	6.87217333
t_5	41.83380883	17.72296576	7.27273304
t_6	47.22024178	22.12278045	14.32178438
t_7	48.00410865	27.27469804	13.77875773
t_8	40.77570646	19.73110557	15.38969635
t_9	52.92289005	21.44943775	18.55252350
t_{10}	29.54042244	30.97851279	23.20180771

Table A.10: Data set 2 with $\sigma = 5$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	32.64627460	0.10000000	0.10000000
t_3	56.50746600	11.86693378	4.51775103
t_4	35.41742170	21.48298427	6.32136466
t_5	41.34201309	16.34936023	5.78659936
t_6	49.25977030	22.49307874	15.04119492
t_7	50.49185986	29.76402311	13.97831844
t_8	40.41603133	19.24905679	16.09568341
t_9	57.43535308	21.68061166	20.46399587
t_{10}	24.70354063	35.03397369	26.94818409

Table A.11: Data set 2 with $\sigma = 7$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	32.54917279	0.10000000	0.10000000
t_3	59.10343023	10.83993346	4.54303745
t_4	32.60302126	21.89614352	5.77055598
t_5	40.85021734	14.97575470	4.30046568
t_6	51.29929882	22.86337703	15.76060546
t_7	52.97961108	32.25334818	14.17787914
t_8	40.05635620	18.76700802	16.80167047
t_9	61.94781610	21.91178557	22.37546823
t_{10}	19.86665883	39.08943458	30.69456047

Table A.12: Data set 2 with $\sigma = 9$

Data set 3: $\mathbf{d}_0 = (35, 25, 60)$ and $\gamma = 1$.

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	35.00000000	25.00000000	60.00000000
t_2	33.62885306	20.35724876	37.98652456
t_3	32.18781796	18.02578111	25.64919974
t_4	30.92885136	16.66920968	18.75478858
t_5	29.96313360	15.78939548	14.86854202
t_6	29.28188726	15.19248686	12.64210605
t_7	28.82467757	14.78596206	11.34151752
t_8	28.52628544	14.51246151	10.56712722
t_9	28.33445058	14.33121238	10.09835937
t_{10}	28.21208613	14.21266625	9.81084598

Table A.13: Data set 3 with $\sigma = 0$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	33.48207397	24.81846959	57.81713724
t_2	33.58030215	18.20039778	37.53381176
t_3	33.48580008	17.51228095	25.66184295
t_4	29.52165114	16.87578931	18.47938424
t_5	29.71723573	15.10259272	14.12547518
t_6	30.30165152	15.37763601	13.00181132
t_7	30.06855318	16.03062459	11.44129787
t_8	28.34644788	14.27143712	10.92012075
t_9	30.59068209	14.44679933	11.05409555
t_{10}	25.79364523	16.24039670	11.68403417

Table A.14: Data set 3 with $\sigma = 1$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	30.44622192	24.45540876	53.45141173
t_2	33.48320034	13.88669582	36.62838616
t_3	36.08176431	16.48528063	25.68712937
t_4	26.70725069	17.28894856	17.92857556
t_5	29.22543998	13.72898718	12.63934150
t_6	32.34118004	15.74793429	13.72122186
t_7	32.55630440	18.51994966	11.64085858
t_8	27.98677275	13.78938835	11.62610780
t_9	35.10314511	14.67797325	12.96556792
t_{10}	20.95676342	20.29585760	15.43041055

Table A.15: Data set 3 with $\sigma = 3$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	27.41036987	24.09234794	49.08568621
t_2	33.38609853	9.57299386	35.72296056
t_3	38.67772853	15.45828031	25.71241579
t_4	23.89285025	17.70210780	17.37776689
t_5	28.73364424	12.35538165	11.15320782
t_6	34.38070856	16.11823258	14.44063241
t_7	35.04405561	21.00927473	11.84041929
t_8	27.62709762	13.30733957	12.33209486
t_9	39.61560813	14.90914716	14.87704029
t_{10}	16.11988162	24.35131850	19.17678693

Table A.16: Data set 3 with $\sigma = 5$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	24.37451782	23.72928711	44.71996069
t_2	33.28899672	5.25929190	34.81753496
t_3	41.27369276	14.43127999	25.73770221
t_4	21.07844980	18.11526705	16.82695821
t_5	28.24184849	10.98177611	9.66707413
t_6	36.42023708	16.48853087	15.16004295
t_7	37.53180683	23.49859980	12.03998000
t_8	27.26742249	12.82529080	13.03808192
t_9	44.12807115	15.14032107	16.78851266
t_{10}	11.28299982	28.40677939	22.92316331

Table A.17: Data set 3 with $\sigma = 7$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	21.33866577	23.36622629	40.35423518
t_2	33.19189491	0.94558994	33.91210936
t_3	43.86965699	13.40427967	25.76298863
t_4	18.26404935	18.52842630	16.27614953
t_5	27.75005275	9.60817058	8.18094045
t_6	38.45976560	16.85882916	15.87945349
t_7	40.01955804	25.98792488	12.23954071
t_8	26.90774736	12.34324202	13.74406898
t_9	48.64053417	15.37149498	18.69998503
t_{10}	6.44611802	32.46224029	26.66953969

Table A.18: Data set 3 with $\sigma = 9$

Data set 4: $\mathbf{d}_0 = (6.39344262, 0.5, 0.6)$ and $\gamma = 1$.

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	6.39344262	0.50000000	0.60000000
t_2	15.48705454	3.64386476	1.00159575
t_3	21.24682122	7.12560640	2.33893143
t_4	24.12866065	9.71704348	4.00667153
t_5	25.67833115	11.38217391	5.53019834
t_6	26.57238015	12.40223924	6.72879783
t_7	27.11026020	13.01990637	7.59729528
t_8	27.44147177	13.39476746	8.19653914
t_9	27.64803024	13.62389799	8.59758866
t_{10}	27.77774083	13.76507189	8.86081251

Table A.19: Data set 4 with $\sigma = 0$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	4.87551660	0.31846959	0.10000000
t_2	15.43850363	1.48701378	0.54888295
t_3	22.54480333	6.61210624	2.35157464
t_4	22.72146043	9.92362310	3.73126719
t_5	25.43243328	10.69537114	4.78713150
t_6	27.59214441	12.58738839	7.08850311
t_7	28.35413581	14.26456890	7.69707563
t_8	27.26163421	13.15374307	8.54953267
t_9	29.90426175	13.73948495	9.55332484
t_{10}	25.35929993	15.79280234	10.73400070

Table A.20: Data set 4 with $\sigma = 1$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	1.83966455	0.10000000	0.10000000
t_2	15.34140182	0.10000000	0.10000000
t_3	25.14076756	5.58510592	2.37686106
t_4	19.90705998	10.33678235	3.18045852
t_5	24.94063753	9.32176561	3.30099782
t_6	29.63167293	12.95768668	7.80791365
t_7	30.84188702	16.75389397	7.89663634
t_8	26.90195908	12.67169430	9.25551973
t_9	34.41672477	13.97065886	11.46479721
t_{10}	20.52241813	19.84826323	14.48037708

Table A.21: Data set 4 with $\sigma = 3$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	15.24430001	0.10000000	0.10000000
t_3	27.73673179	4.55810560	2.40214748
t_4	17.09265954	10.74994160	2.62964984
t_5	24.44884179	7.94816007	1.81486413
t_6	31.67120145	13.32798496	8.52732419
t_7	33.32963824	19.24321904	8.09619705
t_8	26.54228395	12.18964553	9.96150679
t_9	38.92918779	14.20183278	13.37626958
t_{10}	15.68553633	23.90372413	18.22675346

Table A.22: Data set 4 with $\sigma = 5$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	15.14719819	0.10000000	0.10000000
t_3	30.33269602	3.53110528	2.42743391
t_4	14.27825909	11.16310085	2.07884116
t_5	23.95704604	6.57455454	0.32873045
t_6	33.71072997	13.69828325	9.24673473
t_7	35.81738945	21.73254411	8.29575776
t_8	26.18260882	11.70759675	10.66749384
t_9	43.44165082	14.43300669	15.28774194
t_{10}	10.84865452	27.95918503	21.97312983

Table A.23: Data set 4 with $\sigma = 7$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	15.05009638	0.10000000	0.10000000
t_3	32.92866025	2.50410496	2.45272033
t_4	11.46385865	11.57626010	1.52803248
t_5	23.46525030	5.20094900	0.10000000
t_6	35.75025849	14.06858154	9.96614528
t_7	38.30514066	24.22186918	8.49531847
t_8	25.82293369	11.22554798	11.37348090
t_9	47.95411384	14.66418060	17.19921431
t_{10}	6.01177272	32.01464593	25.71950621

Table A.24: Data set 4 with $\sigma = 9$

Data set 5: $\mathbf{d}_0 = (6.98060942, 25, 60)$ and $\gamma = 2$.

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	6.98060942	25.00000000	60.00000000
t_2	10.47741075	12.91465354	36.41550691
t_3	13.68647107	8.95958456	22.04390622
t_4	15.59355317	8.06993366	14.02140913
t_5	16.33300738	8.05383789	9.80511983
t_6	16.53491512	8.16113221	7.66631431
t_7	16.57286006	8.23185234	6.59738016
t_8	16.57424370	8.26451613	6.06416397
t_9	16.57083279	8.27725250	5.79699239
t_{10}	16.56833580	8.28160638	5.66223642

Table A.25: Data set 5 with $\sigma = 0$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	5.46268339	24.81846959	57.81713724
t_2	10.42885984	10.75780256	35.96279411
t_3	14.98445319	8.44608440	22.05654943
t_4	14.18635294	8.27651329	13.74600479
t_5	16.08710951	7.36703512	9.06205299
t_6	17.55467938	8.34628135	8.02601959
t_7	17.81673567	9.47651487	6.69716051
t_8	16.39440614	8.02349174	6.41715750
t_9	18.82706430	8.39283946	6.75272858
t_{10}	14.14989490	10.30933683	7.53542460

Table A.26: Data set 5 with $\sigma = 1$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	2.42683134	24.45540876	53.45141173
t_2	10.33175803	6.44410060	35.05736851
t_3	17.58041742	7.41908408	22.08183585
t_4	11.37195250	8.68967253	13.19519611
t_5	15.59531376	5.99342958	7.57591930
t_6	19.59420790	8.71657964	8.74543013
t_7	20.30448688	11.96583994	6.89672122
t_8	16.03473101	7.54144296	7.12314456
t_9	23.33952732	8.62401337	8.66420094
t_{10}	9.31301309	14.36479772	11.28180098

Table A.27: Data set 5 with $\sigma = 3$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	24.09234794	49.08568621
t_2	10.23465622	2.13039863	34.15194291
t_3	20.17638165	6.39208376	22.10712228
t_4	8.55755205	9.10283178	12.64438743
t_5	15.10351802	4.61982405	6.08978562
t_6	21.63373642	9.08687793	9.46484067
t_7	22.79223810	14.45516501	7.09628193
t_8	15.67505588	7.05939419	7.82913162
t_9	27.85199034	8.85518728	10.57567331
t_{10}	4.47613129	18.42025862	15.02817736

Table A.28: Data set 5 with $\sigma = 5$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	23.72928711	44.71996069
t_2	10.13755441	0.10000000	33.24651731
t_3	22.77234588	5.36508344	22.13240870
t_4	5.74315161	9.51599103	12.09357876
t_5	14.61172227	3.24621852	4.60365194
t_6	23.67326494	9.45717622	10.18425121
t_7	25.27998931	16.94449008	7.29584264
t_8	15.31538075	6.57734541	8.53511868
t_9	32.36445336	9.08636119	12.48714568
t_{10}	0.10000000	22.47571952	18.77455374

Table A.29: Data set 5 with $\sigma = 7$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	23.36622629	40.35423518
t_2	10.04045259	0.10000000	32.34109171
t_3	25.36831010	4.33808312	22.15769512
t_4	2.92875116	9.92915028	11.54277008
t_5	14.11992653	1.87261298	3.11751826
t_6	25.71279347	9.82747450	10.90366176
t_7	27.76774052	19.43381515	7.49540335
t_8	14.95570562	6.09529664	9.24110573
t_9	36.87691638	9.31753510	14.39861805
t_{10}	0.10000000	26.53118042	22.52093012

Table A.30: Data set 5 with $\sigma = 9$

Data set 6: $\mathbf{d}_0 = (6.33065498, 0.5, 0.6)$ and $\gamma = 2$.

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	6.33065498	0.50000000	0.60000000
t_2	13.63765122	3.39515736	0.97279446
t_3	16.10336014	5.91690364	2.08704437
t_4	16.53789398	7.26595637	3.26071918
t_5	16.58656087	7.86779731	4.15303316
t_6	16.57968558	8.11745481	4.73533360
t_7	16.57161480	8.21750786	5.08485580
t_8	16.56760483	8.25696884	5.28438985
t_9	16.56608271	8.27247466	5.39473361
t_{10}	16.56563248	8.27860653	5.45449698

Table A.31: Data set 6 with $\sigma = 0$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	4.81272895	0.31846959	0.10000000
t_2	13.58910032	1.23830638	0.52008166
t_3	17.40134225	5.40340348	2.09968758
t_4	15.13069375	7.47253599	2.98531484
t_5	16.34066300	7.18099454	3.40996632
t_6	17.59944984	8.30260395	5.09503887
t_7	17.81549040	9.46217040	5.18463615
t_8	16.38776726	8.01594445	5.63738337
t_9	18.82231422	8.38806162	6.35046980
t_{10}	14.14719158	10.30633698	7.32768517

Table A.32: Data set 6 with $\sigma = 1$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	1.77687690	0.10000000	0.10000000
t_2	13.49199851	0.10000000	0.10000000
t_3	19.99730648	4.37640316	2.12497400
t_4	12.31629331	7.88569524	2.43450616
t_5	15.84886725	5.80738900	1.92383264
t_6	19.63897836	8.67290224	5.81444942
t_7	20.30324162	11.95149547	5.38419686
t_8	16.02809213	7.53389568	6.34337043
t_9	23.33477724	8.61923553	8.26194216
t_{10}	9.31030978	14.36179788	11.07406155

Table A.33: Data set 6 with $\sigma = 3$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	13.39489669	0.10000000	0.10000000
t_3	22.59327071	3.34940284	2.15026042
t_4	9.50189286	8.29885449	1.88369748
t_5	15.35707151	4.43378347	0.43769896
t_6	21.67850688	9.04320053	6.53385996
t_7	22.79099283	14.44082054	5.58375757
t_8	15.66841700	7.05184690	7.04935749
t_9	27.84724026	8.85040945	10.17341453
t_{10}	4.47342798	18.41725878	14.82043793

Table A.34: Data set 6 with $\sigma = 5$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	13.29779488	0.10000000	0.10000000
t_3	25.18923494	2.32240252	2.17554684
t_4	6.68749242	8.71201374	1.33288881
t_5	14.86527576	3.06017794	0.10000000
t_6	23.71803540	9.41349881	7.25327050
t_7	25.27874405	16.93014561	5.78331828
t_8	15.30874187	6.56979813	7.75534455
t_9	32.35970328	9.08158336	12.08488690
t_{10}	0.10000000	22.47271967	18.56681431

Table A.35: Data set 6 with $\sigma = 7$

	$d_0(t)$	$d_1(t)$	$d_2(t)$
t_1	0.10000000	0.10000000	0.10000000
t_2	13.20069307	0.10000000	0.10000000
t_3	27.78519917	1.29540219	2.20083326
t_4	3.87309197	9.12517299	0.78208013
t_5	14.37348001	1.68657240	0.10000000
t_6	25.75756392	9.78379710	7.97268104
t_7	27.76649526	19.41947068	5.98287898
t_8	14.94906674	6.08774935	8.46133161
t_9	36.87216630	9.31275727	13.99635927
t_{10}	0.10000000	26.52818057	22.31319069

Table A.36: Data set 6 with $\sigma = 9$

Bibliography

[1] ABRAMSON, M. A., AND AUDET, C. Convergence of mesh adaptive direct search to second-order stationary points. *SIAM J. Optim. 17*, 2 (2006), 606–619. 2.4.3, 2.4.3, 4.3

[2] ABRAMSON, M. A., AUDET, C., COUTURE, G., DENNIS, JR., J. E., LE DIGABEL, S., AND TRIBES, C. *The NOMAD project*. Software available at `https://www.gerad.ca/nomad/`. 4.3

[3] ABRAMSON, M. A., AUDET, C., DENNIS, JR., J. E., AND LE DIGABEL, S. OrthoMADS: a deterministic MADS instance with orthogonal directions. *SIAM J. Optim. 20*, 2 (2009), 948–966. 2.4.5

[4] ABRAMSON, M. A., FRIMANNSLUND, L., AND STEIHAUG, T. A subclass of generating set search with convergence to second-order stationary points. *Optim. Methods Softw. 29*, 5 (2014), 900–918. 2.4.3, 2.4.3, 4.3

[5] ALBERSMEYER, J., AND BOCK, H. G. Sensitivity generation in an adaptive BDF-method. In *Modeling, simulation and optimization of complex processes* (Berlin, Heidelberg, 2008), H. G. Bock, E. Kostina, H. X. Phu, and R. Rannacher, Eds., Springer, Berlin Heidelberg, pp. 15–24. 4.2.2

[6] ASHYRALIYEV, M., FOMEKONG-NANFACK, Y., KAANDORP, J. A., AND BLOM, J. G. Systems biology: parameter estimation for biochemical models. *FEBS J. 276* (2009), 886–902. 1.2, 1.3.2

[7] AUDET, C., BÉCHARD, V., AND CHAOUKI, J. Spent potliner treatment process optimization using a MADS algorithm. *Optim. Eng. 9* (2008), 143–160. 2.4.4

[8] AUDET, C., BÉCHARD, V., AND LE DIGABEL, S. Nonsmooth optimization through mesh adaptive direct search and variable neighborhood search. *J. Global Optim. 41*, 2 (Jun 2008), 299–318, doi: 10.1007/s10898-007-9234-1. 2.4.4, 4.1.2

[9] AUDET, C., AND DENNIS, JR., J. E. Analysis of generalized pattern searches. *SIAM J. Optim. 13*, 3 (2003), 889–903. 2.4.3, 2.4.3

[10] AUDET, C., AND DENNIS, JR., J. E. Mesh adaptive direct search algorithms for constrained optimization. *SIAM J. Optim. 17*, 1 (2006), 188–217, doi: 10.1137/040603371. 1.3.2, 2.4.3, 2.4.3, 2.4.5, 4.2.3, 4.3

[11] AUDET, C., DENNIS, JR., J. E., AND LE DIGABEL, S. Globalization strategies for mesh adaptive direct search. *Comput. Optim. Appl. 46*, 2 (2010), 193–215. 2.4.4

[12] AUDET, C., AND HARE, W. *Derivative-free and blackbox optimization*. Springer series in operations research and financial engineering book series (ORFE). Springer, Cham, 2017. 4.2.3

[13] AUDET, C., IANNI, A., LE DIGABEL, S., AND TRIBES, C. Reducing the number of function evaluations in mesh adaptive direct search algorithms. *SIAM J. Optim. 24*, 2 (2014), 621–642, doi: 10.1137/120895056. 2.4.3, 2.4.5

[14] AUDET, C., LE DIGABEL, S., AND TRIBES, C. NOMAD user guide. *Les cahiers du GERAD G-2009-37* (2009). 4.3

[15] AUDET, C., LE DIGABEL, S., AND TRIBES, C. Dynamic scaling in the mesh adaptive direct search algorithm for blackbox optimization. *Optim. Eng. 17*, 2 (2016), 333–358, doi: 10.1007/s11081-015-9283-0. 4.3, 4.6.1

[16] AUGER, P., BRAVO DE LA PARRA, R., POGGIALE, J. C., SÁNCHEZ, E., AND SANZ, L. Aggregation methods in dynamical systems and applications in population and community dynamics. *Phys. Life Rev. 5* (2008), 79–105. 2.1, 3.1

[17] BARBAROSSA, M. V., AND KUTTLER, C. Mathematical modeling of bacteria communication in continuous cultures. *Appl. Sci. 6(5), 149* (2016), doi: 10.3390/app6050149. 1.3.2

[18] BARD, Y. *Nonlinear parameter estimation*. Academic Press, New York, 1974. 1.2, 1.2, 2.3, 2.4.5, 4.1, 4.2.1, 4.2.3

[19] BEGLEY, J. *Gadget user guide*. Marine and Freshwater Research Institute, Reykjavík, Iceland, August 2012. Retrieved May, 2018 from `http://www.hafro.is/gadget/files/userguide.pdf`. 4.1.2

[20] BEGLEY, J., AND HOWELL, D. An overview of Gadget, the Globally applicable Area-Disaggregated General Ecosystem Toolbox. *ICES CM 2004/FF:13*. 4.1.2

[21] BEVERTON, R. J. H., AND HOLT, S. J. *On the dynamics of exploited fish populations*, vol. 19 of *Fishery Investigations. Series 2*. HMSO, UK, 1957. 1.1, 2.2.1

[22] BEVERTON, R. J. H., AND ILES, T. C. Mortality rates of 0-group plaice (*Platessa platessa* L.), dab (*Limanda limanda* L.) and turbot (*Scophthalmus maximus* L.) in European waters: III. Density dependence of mortality rates of 0-group plaice and some demographic implications. *Neth. J. Sea. Res. 29*, 1-3 (1992), 61–79, doi: 10.1016/0077-7579(92)90008-3. 2.2.1

[23] BOCK, H. G., KOSTINA, E., AND SCHLÖDER, J. P. Direct multiple shooting and generalized Gauss-Newton method for parameter estimation problems in ODE models. In *Multiple shooting and time domain decomposition methods* (2015), T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, Eds., vol. 9 of *Contributions in mathematical and computational sciences*, Springer, Cham, pp. 1–34. 1.2, 1.2, 4.1, 4.2.1, 4.2.1

[24] BOGSTAD, B., LILLY, G. R., MEHL, S., PÁLSSON, Ó. K., AND STEFÁNSSON, G. Cannibalism and year-class strength in Atlantic cod (*Gadus morhua* L.) in Arcto-boreal ecosystems (Barents Sea, Iceland, and eastern Newfoundland). *ICES Mar. Sci. Symp. 198* (1994), 576–599. 3.4

[25] BOGSTAD, B., YARAGINA, N. A., AND NASH, R. D. M. The early life-history dynamics of Northeast Arctic cod: levels of natural mortality and abundance during the first 3 years of life. *Can. J. Fish. Aquat. Sci. 73*, 2 (2016), 246–256. 1.1, 3.4, 5.7

[26] BOOKER, A. J., DENNIS, JR., J. E., FRANK, P. D., SERAFINI, D. B., TORCZON, V., AND TROSSET, M. W. A rigorous framework for optimization of expensive functions by surrogates. *Struct. Multidiscip. Optim. 17*, 1 (1999), 1–13. 2.4.3, 2.4.4, 4.2.3

[27] BORGGAARD, J., PELLETIER, D., AND VUGRIN., K. On sensitivity analysis for problems with numerical noise. AIAA Paper 2002–5553. In *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization Atlanta, Georgia* (2002). 1.3.2, 4.1, 4.2.2, 4.2.3

[28] BOX, M. J. A new method of constrained optimization and a comparison with other methods. *Comput. J. 8*, 1 (1965), 42–52. 4.3, 4.4.3

[29] BREWER, D., BARENCO, M., CALLARD, R., HUBANK, M., AND STARK, J. Fitting ordinary differential equations to short time course data. *Phil. Trans. R. Soc. A 366*, 1865 (2008), 519–544. 1.2, 4.1

[30] BROOKS, E. N., AND POWERS, J. E. Generalized compensation in stock-recruit functions: properties and implications for management. *ICES J. Mar. Sci. 64* (2007), 413–424. 1.1, 1.1, 5.7

[31] BUDDRUS-SCHIEMANN, K., RIEGER, M., MÜHLBAUER, M., BARBAROSSA, M. V., KUTTLER, C., HENSE, B. A., ROTHBALLER, M., UHL, J., FONSECA, J. R., SCHMITT-KOPPLIN, P., SCHMID, M., AND HARTMANN, A. Analysis of *N*-acylhomoserine lactone dynamics in continuous cultures of *Pseudomonas putida* IsoF by use of ELISA and UHPLC/qTOF-MS-derived measurements and mathematical models. *Anal. Bioanal. Chem. 406*, 25 (2014), 6373–6383, doi: 10.1007/s00216-014-8063-6. 1.3.2

[32] CARTER, R. G., GABLONSKY, J. M., PATRICK, A., KELLEY, C. T., AND ESLINGER, O. J. Algorithms for noisy problems in gas transmission pipeline optimization. *Optim. Eng. 2*, 2 (2001), 139–157, doi: 10.1023/A:1013123110266. 2.4.4

[33] CASTILLO-CHAVEZ, C., AND BRAUER, F. Models for populations with age structure. In *Mathematical models in population biology and epidemiology*, 2nd ed. Springer, New York, 2012. 2.2.3, 3.4

[34] CASWELL, H. *Matrix population models: Construction, analysis, and interpretation*, 2nd ed. Sinauer Associates, Inc., Sunderland, MA, 2001. 1.1, 2.2, 2.2.2

[35] CHARACKLIS, G. W., KIRSCH, B. R., RAMSEY, J., DILLARD, K. E. M., AND KELLEY, C. T. Developing portfolios of water supply transfers. *Water Resour. Res. 42*, 5 (2006), W05403–1–W05403–14, doi: 10.1029/2005WR004424. 2.4.4

[36] CHEN, X., AND KELLEY, C. T. Optimization with hidden constraints and embedded Monte Carlo computations. *Optim. Eng. 17*, 1 (2016), 157–175, doi: 10.1007/s11081-015-9302-1. 2.4.4

[37] CHOI, T. D., ESLINGER, O. J., KELLEY, C. T., DAVID, J. W., AND ETHERIDGE, M. Optimization of automotive valve train components with implicit filtering. *Optim. Eng. 1*, 1 (2000), 9–27, doi: 10.1023/A:1010071821464. 2.4.4, 4.1.2, 4.2.1, 4.2.3

[38] CHOI, T. D., AND KELLEY, C. T. Superlinear convergence and implicit filtering. *SIAM J. Optim. 10*, 4 (2000), 1149–1162. 1.2, 2.4.4, 4.2.3

[39] CONN, A. R., AND LE DIGABEL, S. Use of quadratic models with mesh-adaptive direct search for constrained black box optimization. *Optim. Methods Softw. 28*, 1 (2013), 139–158, doi: 10.1080/10556788.2011.623162. 2.4.3, 4.3, 4.4.2

[40] CONN, A. R., SCHEINBERG, K., AND TOINT, PH. L. On the convergence of derivative-free methods for unconstrained optimization. In *Approximation theory and optimization: Tributes to M. J. D. Powell*, M. D. Buhmann and A. Iserles, Eds. Cambridge University Press, Cambridge, UK, 1997, pp. 83–108. 2.4.2

[41] CONN, A. R., SCHEINBERG, K., AND TOINT, PH. L. A derivative free optimization algorithm in practice. In *Proceeedings of 7-th AIAA/USAF/-NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St Louis, MO* (1998). 2.4.2, 2.4.4, 4.2.3

[42] CONN, A. R., SCHEINBERG, K., AND VICENTE, L. N. Geometry of interpolation sets in derivative free optimization. *Math. Program. 111*, 1–2 (2008), 141–172, doi: 10.1007/s10107-006-0073-5. 2.4.2, 2.4.2, 2.4.2

[43] CONN, A. R., SCHEINBERG, K., AND VICENTE, L. N. Geometry of sample sets in derivative-free optimization: polynomial regression and underdetermined interpolation. *IMA J. Numer. Anal. 28*, 4 (2008), 721–748. 2.4.2, 2.4.2

[44] CONN, A. R., SCHEINBERG, K., AND VICENTE, L. N. Global convergence of general derivative-free trust-region algorithms to first- and second-order critical points. *SIAM J. Optim. 20*, 1 (2009), 387–415, doi: 10.1137/060673424. 2.4.2, 4.5.2

[45] CONN, A. R., SCHEINBERG, K., AND VICENTE, L. N. *Introduction to Derivative-free Optimization*. MOS-SIAM series on optimization. SIAM, Philadelphia, 2009. 1.3.2, 2.4, 2.4.1, 2.4.2, 2.4.2, 2.4.3, 2.4.3, 2.4.3, 2.4.4, 4.2.3, 4.2.3

[46] CUSTÓDIO, A. L., ROCHA, H., AND VICENTE, L. N. Incorporating minimum Frobenius norm models in direct search. *Comput. Optim. Appl. 46*, 2 (2010), 265–278. 2.4.3, 4.1, 4.3, 4.4.2

[47] CUSTÓDIO, A. L., SCHEINBERG, K., AND VICENTE, L. N. Methodologies and software for derivative-free optimization. In *Advances and trends in optimization with engineering applications*, T. Terlaky, M. F. Anjos, and S. Ahmed, Eds., MOS-SIAM series on optimization. SIAM, Philadelphia, 2017, ch. 37, pp. 495–506. 4.2.3

[48] CUSTÓDIO, A. L., AND VICENTE, L. N. Using sampling and simplex derivatives in pattern search methods. *SIAM J. Optim. 18*, 2 (2007), 537–555. 2.4.3, 4.3

[49] CUSTÓDIO, A. L., AND VICENTE, L. N. *SID-PSM: A pattern search method guided by simplex derivatives for use in derivative-free optimization (version 1.3)*, December 2014. Retrieved May, 2018 from `http://www.mat.uc.pt/sid-psm`. 4.3, 4.6.1

[50] DAVID, J. W., CHENG, C. Y., CHOI, T. D., KELLEY, C. T., AND GABLONSKY, J. Optimal design of high speed mechanical systems. Tech. Rep. CRSC-TR97-18, Center for Research in Scientific Computation, North Carolina State University, 1997. 4.1.2, 4.2.1

[51] DAVID, J. W., KELLEY, C. T., AND CHENG, C. Y. Use of an implicit filtering algorithm for mechanical system parameter identification. SAE paper 960358. In *Modeling of CI and SI engines* (1996), SAE international congress and exposition conference proceedings, Society of Automotive Engineers, Washington, DC, pp. 189–194. 4.1.2, 4.2.1

[52] DERISO, R. B. Harvesting strategies and parameter estimation for an age-structured model. *Can. J. Fish. Aquat. Sci. 37* (1980), 268–282. 1.1

[53] EDSBERG, L., AND WEDIN, P.-Å. Numerical tools for parameter estimation in ODE-systems. *Optim. Methods Softw. 6*, 3 (1995), 193–217, doi: 10.1080/10556789508805633. 1.2, 1.3.2, 4.1, 4.2.2

[54] FENICHEL, N. Geometric singular perturbation theory for ordinary differential equations. *J. Diff. Eq. 31* (1979), 53–98. 1.3.1, 2.1, 3.1

[55] FINKEL, D. E., AND KELLEY, C. T. Convergence analysis of the DIRECT algorithm. Tech. Rep. CRSC-TR04-28, Center for Research in Scientific Computation, North Carolina State University, 2004. 2.4.4, 2.4.5, 4.2.3

[56] FINKEL, D. E., AND KELLEY, C. T. Convergence analysis of sampling methods for perturbed Lipschitz functions. *Pac. J. Optim. 5*, 2 (2009), 339–350. 1.3.2, 2.4.4, 2.4.5, 2.4.5, 4.2.3

[57] FOMEKONG-NANFACK, Y., KAANDORP, J. A., AND BLOM, J. Efficient parameter estimation for spatio-temporal models of pattern formation: case study of *Drosophila melanogaster*. *Bioinformatics 23*, 24 (2007), 3356–3363. 4.1.2

[58] FOURNIER, D. A., SKAUG, H. J., ANCHETA, J., IANELLI, J., MAGNUSSON, A., MAUNDER, M. N., NIELSEN, A., AND SIBERT, J. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw. 27*, 2 (2012), 233–249, doi: 10.1080/10556788.2011.597854. 2.4.5, 4.6.2

[59] FOWLER, K. R., REESE, J. P., KEES, C. E., DENNIS, JR., J. E., KELLEY, C. T., MILLER, C. T., AUDET, C., BOOKER, A. J., COUTURE, G., DARWIN, R. W., FARTHING, M. W., FINKEL, D. E., GABLONSKY, J. M., GRAY, G., AND KOLDA, T. G. A comparison of derivative-free optimization methods for groundwater supply and hydraulic capture community problems. *Adv. Water Resour. 31* (2008), 743–757. 4.6.2

[60] FRIMANNSLUND, L., AND STEIHAUG, T. A generating set search method using curvature information. *Comput. Optim. Appl. 38* (2007), 105–121. 2.4.3, 4.2.3, 4.3, 4.5.2

[61] FRIMANNSLUND, L., AND STEIHAUG, T. On a new method for derivative free optimization. *Int. J. Adv. Softw* (2011). 2.4.3, 4.3

[62] GARCÍA-PALOMARES, U. M., AND RODRÍGUEZ, J. F. New sequential and parallel derivative-free algorithms for unconstrained minimization. *SIAM J. Optim. 13*, 1 (2002), 79–96, doi: 10.1137/S1052623400370606. 2.4.3

[63] GEAR, C. W., AND VU, T. Smooth numerical solutions of ordinary differential equations. In *Numerical treatment of inverse problems in differential and integral equations*, P. Deuflhard and E. Hairer, Eds. Birkhäuser, Boston, 1983, pp. 2–12. 2.3, 4.1, 4.2.2

[64] GJØSÆTER, H., BOGSTAD, B., AND TJELMELAND, S. Assessment methodology for Barents Sea capelin, *Mallotus villosus* (Müller). *ICES J. Mar. Sci. 59*, 5 (2002), 1086–1095, doi: 10.1006/jmsc.2002.1238. 4.6.2

[65] GOULD, N. I. M., ORBAN, D., AND TOINT, PH. L. CUTEr and SifDec: A constrained and unconstrained testing environment, revisited. *ACM Trans. Math. Software 29* (2003), 373–394. 4.1

[66] GRONWALL, T. H. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Ann. Math. 20*, 4 (1919), 292–296, doi: 10.2307/1967124. 4.2

[67] GULLIKSSON, M., AND SÖDERKVIST, I. Surface fitting and parameter estimation with nonlinear least squares. *Optim. Methods Softw. 5*, 3 (1995), 247–269, doi: 10.1080/10556789508805614. 1.2

[68] GULLIKSSON, M., SÖDERKVIST, I., AND WEDIN, P.-Å. Algorithms for constrained and weighted nonlinear least squares. *SIAM J. Optim. 7*, 1 (1997), 208–224, doi: 10.1137/S1052623493248809. 1.2

[69] GURNEY, W. S. C., NISBET, R. M., AND LAWTON, J. H. The systematic formulation of tractable single-species population models incorporating age structure. *J. Anim. Ecol. 52*, 2 (1983), 479–495. 3.4, 3.4

[70] HADDON, M. *Modelling and quantitative methods in fisheries*, 2nd ed. Chapman and Hall, Boca Raton, 2011. 1.1, 1.1

[71] HAIRER, E., NØRSETT, S. P., AND WANNER, G. *Solving ordinary differential equations I: Nonstiff Problems*, 2nd ed., vol. 8 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1993. 1.2, 2.3, 2.3, 4.2.1, 4.2.2, 4.2.3

[72] HAIRER, E., AND WANNER, G. *Solving ordinary differential equations II: Stiff and differential–algebraic problems*, vol. 14 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1991. 2.3

[73] HILBORN, R., AND WALTERS, C. J. *Quantitative fisheries stock assessment: Choice, dynamics and uncertainty*. Chapman and Hall, London, 1992. 1.1, 1.3.1, 2.2.3, 3.3.1, 4.6.1

[74] HJORT, J. Fluctuations in the great fisheries of Northern Europe: Viewed in light of biological research. *Rapp. p.-v. réun. - Cons. int. explor. mer 20* (1914), 1–228. 1.1

[75] HOFMANN, E. E. Structured-population models: many methods, a few basic concepts. In *Models for marine ecosystems*, H. Caswell and S. Tuljapurkar, Eds., vol. 18 of *Population and community biology series*. Chapman and Hall, New York, 1997, pp. 409–432. 4.6.2

[76] HOOKE, R., AND JEEVES, T. A. "Direct search" solution of numerical and statistical problems. *J. ACM 8*, 2 (1961), 212–229, doi: 10.1145/321062.321069. 1.3.2, 4.1.2

[77] HUYER, W., AND NEUMAIER, A. Snobfit - stable noisy optimization by branch and fit. *ACM Trans. Math. Softw. 35*, 2 (2008), 1–25, doi: 10.1145/1377612.1377613. 2.4.4, 2.4.5, 4.2.3

[78] ICES. Report of the Working Group for the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK), 30 April–7 May 2014, ICES HQ, Copenhagen, Denmark. ICES CM 2014/ACOM:13, 30 April–7 May 2014, 2014. 1.1

[79] JOHANSEN, R. A model for the interaction between gadoid larvae and their nauplii prey. *Math. Biosci. 208*, 1 (2007), 177–192, doi: 10.1016/j.mbs.2006.10.005. 1.1

[80] JONES, C. K. R. T. Geometric singular perturbation theory. In *Dynamical systems*, R. Johnson, Ed., vol. 1609 of *Lecture notes in mathematics*. Springer, Berlin, 1995, pp. 44–118. 2.1, 2.1, 3.2, 3.3.2

[81] KAPER, T. J. An introduction to geometric methods and dynamical systems theory for singular perturbation problems. In *Analyzing multiscale phenomena using singular perturbation methods* (1999), J. Cronin and R. E. O'Malley, Jr., Eds., vol. 56 of *Proceedings of symposia in applied mathematics*, pp. 85–132. 2.1

[82] KELLEY, C. T. *Implicit filtering*. No. 23 in Software, environments, and tools. SIAM, Philadelphia, 2011. 2.4.4, 2.4.5

[83] KELLEY, C. T. Implicit filtering and hidden constraints. In *Advances and trends in optimization with engineering applications*, M. A. T. Terlaky and S. Ahmed, Eds. SIAM, Philadelphia, 2017, ch. 38, pp. 507–518. 2.4.4, 4.2.1, 4.2.3

[84] KIRKPATRICK, S., GELATT, JR., C. D., AND VECCHI, M. P. Optimization by simulated annealing. *Science 220*, 4598 (1983), 671–680, doi: 10.1126/science.220.4598.671. 4.2.3

[85] KIRSCH, B. R., CHARACKLIS, G. W., DILLARD, K. E. M., AND KELLEY, C. T. More efficient optimization of long-term water supply portfolios. *Water Resour. Res. 45*, W03414 (2009), doi: 10.1029/2008WR007018. W03414. 2.4.4

[86] KOLDA, T. G., LEWIS, R. M., AND TORCZON, V. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Rev. 45*, 3 (2003), 385–482. 2.4.3, 2.4.3, 2.4.3, 4.2.3, 4.3

[87] KORTELAINEN, M., LESINSKI, T., MORÉ, J., NAZAREWICZ, W., SARICH, J., SCHUNCK, N., STOITSOV, M. V., AND WILD, S. Nuclear energy density optimization. *Phys. Rev. C 82* (2010), 024313, doi: 10.1103/PhysRevC.82.024313. 4.6.1, 4.6.2

[88] LAGARIAS, J. C., REEDS, J. A., WRIGHT, M. H., AND WRIGHT, P. E. Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optim. 9*, 1 (1998), 112–147. 4.3

[89] LE DIGABEL, S. Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm. *ACM Trans. Math. Softw. 37*, 4 (2011), 44:1–44:15. 4.3

[90] LE DIGABEL, S., AND WILD, S. M. A taxonomy of constraints in simulation-based optimization. *Les cahiers du GERAD G-2015-57* (2015). 2.4.4

[91] LESLIE, P. H. On the use of matrices in certain population mathematics. *Biometrika 33*, 3 (1945), 183–212. 2.2.2

[92] LEVEQUE, R. J. *Finite difference methods for ordinary and partial differential equations*. SIAM, Philadelphia, 2007. 1.2, 2.3, 2.3, 4.2.2, 4.2.5

[93] LEWIS, R., AND TORCZON, V. Pattern search methods for linearly constrained minimization. *SIAM J. Optim. 10*, 3 (2000), 917–941. 4.3

[94] LIN, C. C., AND SEGEL, L. A. *Mathematics applied to deterministic problems in the natural sciences*. Classics in applied mathematics 1. SIAM, Philadelphia, 1988. With material on elasticity by G.H. Handelman. 3.3.3

[95] LUCIDI, S., AND SCIANDRONE, M. On the global convergence of derivative-free methods for unconstrained optimization. *SIAM J. Optim. 13*, 1 (2002), 97–116, doi: 10.1137/S1052623497330392. 2.4.3

[96] LUDWIG, D., AND WALTERS, C. J. A robust method for parameter estimation from catch and effort data. *Can. J. Fish. Aquat. Sci. 46*, 1 (1989), 137–144, doi: 10.1139/f89-018. 2.2.1

[97] *MATLAB Release 2012a*. The MathWorks, Inc., Natick, Massachusetts, United States, 2012. 4.2.5

[98] MAUNDER, M. N., AND PUNT, A. E. A review of integrated analysis in fisheries stock assessment. *Fish. Res. 142* (2013), 61–74, doi: 10.1016/j.fishres.2012.07.025. 4.6.2

[99] MORÉ, J. J., AND WILD, S. M. Benchmarking derivative-free optimization algorithms. *SIAM J. Optim. 20*, 1 (2009), 172–191. 1.3.2, 4.1, 4.1, 4.3, 4.4.2, 4.4.2, 4.4.2, 4.5.2, 4.5.2

[100] MORÉ, J. J., AND WILD, S. M. Estimating computational noise. *SIAM J. Sci. Comput. 33*, 3 (2011), 1292–1314, doi: 10.1137/100786125. 4.2.2

[101] MÜLLER, J., AND KUTTLER, C. *Methods and models in mathematical biology*. Springer, Berlin, 2015. 1.3.1, 2.1, 2.2.3, 3.3.3

[102] MURDOCH, W. W., BRIGGS, C. J., AND NISBET, R. M. *Consumer-resource dynamics*. MPB-36. Princeton University Press, 2003. 3.4

[103] MYERS, R. A., AND BARROWMAN, N. J. Is fish recruitment related to spawner abundance? *Fish. Bull. 94*, 4 (1996), 707–724. 1.1

[104] NASH, R. D. M. Exploring the population dynamics of Irish Sea plaice, *Pleuronectes platessa* L., through the use of Paulik diagrams. *J. Sea Res. 40*, 1–2 (1998), 1–18. 1.1

[105] NASH, R. D. M., AND GEFFEN, A. J. Mortality through the early life-history of fish: What can we learn from European plaice (*Pleuronectes platessa* L.)? *J. Mar. Syst. 93* (2012), 58–68. 1.1

[106] NEEDLE, C. L. Recruitment models: diagnosis and prognosis. *Rev. Fish Biol. Fish. 11* (2002), 95–111. 1.1, 2.2.3

[107] NELDER, J. A., AND MEAD, R. A simplex method for function minimization. *Comput. J. 7*, 4 (1965), 308–313. 1.3.2, 2.4.1, 4.1.2, 4.3

[108] NOCEDAL, J., AND WRIGHT, S. J. *Numerical Optimization*, 2 ed. Springer series in operations research and financial engineering. Springer, New York, 2006. 1.2, 2.4.3, 2.4.5, 2.4.5, 4.2.6, 4.3, 4.5.2

[109] OLSEN, E. M., OTTERSEN, G., LLOPE, M., CHAN, K.-S., BEAUGRAND, G., AND STENSETH, N. C. Spawning stock and recruitment in North Sea cod shaped by food and climate. *Proc. R. Soc. B 278*, 1705 (2011), 504–510, doi: 10.1098/rspb.2010.1465. 1.1

[110] PAULIK, G. J. Studies of the possible form of the stock-recruitment curve. *Rapp. p.-v. réun. - Cons. int. explor. mer 164* (1973), 302–315. 1.1, 1.3.1, 2.2.1, 5.7

[111] POWELL, M. J. D. Direct search algorithms for optimization calculations. *Acta Numer. 7* (1998), 287–336, doi: 10.1017/S0962492900002841. 2.4.2

[112] POWELL, M. J. D. On trust region methods for unconstrained minimization without derivatives. *Math. Program. 97*, 3 (2003), 605–623, doi: 10.1007/s10107-003-0430-6. 2.4.2

[113] POWELL, M. J. D. The NEWUOA software for unconstrained optimization without derivatives. In *Large-scale nonlinear optimization*, G. Di Pillo and M. Roma, Eds. Springer, New York, 2006, pp. 255–297. 2.4.2, 4.6.1

[114] PUNT, A. E., AND ELVARSSON, B. Þ. Improving the performance of the algorithm for conditioning Implementation Simulation Trials, with application to North Atlantic fin whales. Paper SC/D11/NPM1 presented to the First Intersessional Workshop for the Implementation Review of western North Pacific common minke whales, 12-16 December 2011, Tokyo, Japan (unpublished). 7pp., 2011. 1.3.2, 4.1.2, 4.3

[115] QUARTERONI, A., SACCO, R., AND SALERI, F. Numerical solution of ordinary differential equations. In *Numerical mathematics*. Springer, New York, 2007, pp. 469–530. 2.3, 4.2.1, 4.2.2

[116] QUINN, T. J., AND DERISO, R. B. *Quantitative fish dynamics*. Oxford University Press, New York, Oxford, 1999. 1.1, 2.2.1, 2.2.2, 2.2.2, 3.3.1, 5.1, 5.5, 5.5

[117] RECHENBERG, I. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. PhD thesis, Technische Universität Berlin, 1971. 4.2.3

[118] REGIS, R. G., AND WILD, S. M. CONORBIT: constrained optimization by radial basis function interpolation in trust regions. *Optim. Methods Softw. 32*, 3 (2017), 552–580, doi: 10.1080/10556788.2016.1226305. 4.2.3

[119] RICKER, W. E. Stock and recruitment. *J. Fish. Res. Board Can. 11* (1954), 559–623. 1.1, 2.2.1, 3.4

[120] RIOS, L. M., AND SAHINIDIS, N. V. Derivative-free optimization: a review of algorithms and comparison of software implementations. *J. Glob. Optim. 56*, 3 (2013), 1247–1293. 1.3.2, 4.3, 4.4.1

[121] ROSENBROCK, H. H., AND STOREY, C. Chapter 8. In *Computational techniques for chemical engineers*. Pergamon Press, Oxford, 1966. 1.2

[122] ROTHSCHILD, B. J. *Dynamics of Marine Fish Populations*. Harvard University Press, USA, 1986. 1.1, 1.1

[123] SALTELLI, A., TARANTOLA, S., CAMPOLONGO, F., AND RATTO, M. *Sensitivity Analysis in practice*. John Wiley and Son, Ltd., Chichester, 2004. 4.6.1

[124] SCHEINBERG, K., AND TOINT, PH. L. Self-correcting geometry in model-based algorithms for derivative-free unconstrained optimization. *SIAM J. Optim. 20*, 6 (2010), 3512–3532, doi: 10.1137/090748536. 2.4.2

[125] SCHITTKOWSKI, K. *Numerical data fitting in dynamical systems*. Springer, Dordrecht, 2002. 1.2, 4.2.3, 4.6.1

[126] SCHNUTE, J. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci. 42* (1985), 414–429. 1.1, 1.1

[127] SHAMPINE, L. F. Implementation of implicit formulas for the solution of ODEs. *SIAM J. Sci. and Stat. Comput. 1*, 1 (1980), 103–118, doi: 10.1137/0901005. 4.2.2, 4.2.5

[128] SHAMPINE, L. F., AND REICHELT, M. W. The MATLAB ODE Suite. *SIAM J. Sci. Comput. 18*, 1 (1997), 1–22. 2.3, 2.3, 2.3, 4.2.2, 4.2.5, 4.2.5, 4.1, 4.2.5, 4.2.5

[129] SHELTON, A. O., AND MANGEL, M. Fluctuations of fish populations and the magnifying effects of fishing. *Proc. Natl. Acad. Sci. 108*, 17 (2011), 7075–7080, doi: 10.1073/pnas.1100334108. 1.1

[130] SHIN, Y.-J., AND CURY, P. Exploring fish community dynamics through size-dependent trophic interactions using a spatialized individual-based model. *Aquat. Living Resour. 14*, 2 (2001), 65–80, doi: 10.1016/S0990-7440(01)01106-8. 4.6.2

[131] SMITH, P. E., HORNE, J. K., AND SCHNEIDER, D. C. Spatial dynamics of anchovy, sardine, and hake pre-recruit stages in the California Current. *ICES J. Mar. Sci. 58*, 5 (2001), 1063–1071, doi: 10.1006/jmsc.2001.1092. 1.1

[132] STEIHAUG, T., AND WOLFBRANDT, A. An attempt to avoid exact Jacobian and nonlinear equations in the numerical solution of stiff differential equations. *Math. Comput. 33*, 146 (1979), 521–534. 2.3

[133] STONEKING, D. E., BILBRO, G. L., GILMORE, P. A., TREW, R. J., AND KELLEY, C. T. Yield optimization using a GaAs process simulator coupled to a physical device model. *IEEE Trans. Microw. 40*, 7 (1992), 1353–1363, doi: 10.1109/22.146318. 4.2.3

[134] SUBBEY, S. Regularization of a parameter estimation problem using monotonicity and convexity constraints. *ESAIM: Proc. Surv. 57* (2017), 86–96, doi: 10.1051/proc. 1.1, 1.1, 1.3.1

[135] SUBBEY, S., DEVINE, J. A., SCHAARSCHMIDT, U., AND NASH, R. D. M. Modeling and forecasting stock recruitment: current and future perspectives. *ICES J. Mar. Sci. 71*, 8 (2014), 2307–2322. 1.3.1

[136] SUBBEY, S., AND SCHAARSCHMIDT, U. A two-phase strategy for effective parameter estimation. In *Abstracts of MMA2015* (Sigulda, Latvia, May 26–29 2015). 4.1.2

[137] TJOA, I.-B., AND BIEGLER, L. T. Simultaneous solution and optimization strategies for parameter estimation of differential-algebraic equation systems. *Ind. Eng. Chem. Res. 30*, 2 (1991), 376–385, doi: 10.1021/ie00050a015. 4.1

[138] TORCZON, V. On the convergence of pattern search algorithms. *SIAM J. Optim. 7*, 1 (1997), 1–25, doi: 10.1137/S1052623493250780. 2.4.3, 2.4.3

[139] TOUZEAU, S., AND GOUZÉ, J.-L. On the stock-recruitment relationships in fish population models. *Environ. Model. Assess. 3* (1998), 87–93. 1.1, 1.1, 1.3.1, 2, 2.2.4, 2.2.4, 2.2.4, 2.2.4, 3.1, 3.1, 3.3, 3.3.1, 3.3.3, 3.4, 3.4, 5.6

[140] VÁZQUEZ, S., MARTÍN, M. J., FRAGUELA, B. B., GÓMEZ, A., RODRÍGUEZ, A., AND ELVARSSON, B. Þ. Novel parallelization of simulated annealing and Hooke & Jeeves search algorithms for multicore systems with application to complex fisheries stock assessment models. *J. Comput. Sci. 17* (2016), 599–608, doi: 10.1016/j.jocs.2016.07.003. Recent Advances in Parallel Techniques for Scientific Computing. 1.3.2, 4.1.2, 4.3

[141] VERHULST, F. *Methods and applications of singular perturbations: Boundary layers and multiple timescale dynamics*. Springer, New York, 2005. 2.1

[142] WIGGINS, S. *Introduction to applied nonlinear dynamical systems and chaos*, 2nd ed. Springer, New York, 2003. 2.1

[143] WILD, S. M., REGIS, R. G., AND SHOEMAKER, C. A. ORBIT: optimization by radial basis function interpolation in trust-regions. *SIAM J. Sci. Comput. 30*, 6 (2008), 3197–3219. 2.4.2, 4.3, 4.4.2, 4.5.2

[144] WILD, S. M., AND SHOEMAKER, C. A. Global convergence of radial basis function trust region derivative-free algorithms. *SIAM J. Optim. 21*, 3 (2011), 761–781. 2.4.2, 4.5.2

[145] WOLFBRANDT, A. *A study of Rosenbrock processes with respect to order conditions and stiff stability*. PhD thesis, University of Gothenburg, 1977. 2.3

[146] WRIGHT, M. H. Direct search methods: once scorned, now respectable. In *Numerical analysis 1995 (Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis)*, D. F. Griffiths and G. A. Watson, Eds., vol. 344 of *Pitman research notes in mathematics*. Addison Wesley Longman, Harlow, 1996, pp. 191–208. 2.4.1

[147] ZHANG, H., CONN, A. R., AND SCHEINBERG, K. A derivative-free algorithm for least-squares minimization. *SIAM J. Optim. 20*, 6 (2010), 3555–3576, doi: 10.1137/09075531X. 4.6.1

Part II

Included Papers

Chapter 7

List of papers

Paper A Schaarschmidt, U., Steihaug, T., Subbey, S. A parametrized stock-recruitment relationship derived from a slow-fast population dynamic model. *Math. Comput. Simul. 145* (2018), 171–185. The 5th IMACS Conference on Mathematical Modelling and Computational Methods in Applied Sciences and Engineering, in honour of Professor Owe Axelsson's 80th birthday.

Paper B Schaarschmidt, U., Subbey, S., Nash, R.D.M., and Frank, A.S.J. Emergent properties of a multi-stage population dynamic model. Submitted to *J. Math. Biol.*

Paper C Schaarschmidt, U., Steihaug, T., Subbey, S. Derivative-free optimization for population dynamic models. In *Model. Comput. & Optim. in Inf. Syst. & Manage. Sci.* (2015), Le Thi, H.A., Pham Dinh, T., and Nguyen, N.T., Eds., vol. 359 of *Advances in Intelligent Systems and Computing*, Springer, Cham, 391–402.

The following publication is also related to the thesis:

Paper D Subbey, S., Devine, J.A., Schaarschmidt, U., Nash, R.D.M. Modeling and forecasting stock recruitment: Current and future perspectives, *ICES J. Mar. Sci. 71*, 8 (2014), 2307–2322.

Paper A

A parametrized stock-recruitment relationship derived from a slow-fast population dynamic model

Ute Schaarschmidt, Trond Steihaug, Sam Subbey

Math. Comput. Simul. 145 (2018), 171–185. The 5th IMACS Conference on Mathematical Modelling and Computational Methods in Applied Sciences and Engineering, in honour of Professor Owe Axelsson's 80th birthday.

Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 145 (2018) 171–185
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original articles

A parametrized stock-recruitment relationship derived from a slow-fast population dynamic model

Ute Schaarschmidt[a,*], Trond Steihaug[a], Sam Subbey[b]

[a] *Department of Informatics, University of Bergen, Pb 7803, N-5020 Bergen, Norway*
[b] *Institute of Marine Research, Pb 1870, N-5817 Bergen, Norway*

Received 15 August 2014; received in revised form 12 October 2017; accepted 12 October 2017
Available online 6 November 2017

Abstract

The Beverton–Holt, Ricker and Deriso functions are three distinct descriptions of the link between a parental population size and subsequent offspring that may survive to become part of the fish population.

This paper presents a model consisting of a system of ordinary differential equations, which couples a stage of young fish with several adult stages. The slow-fast dynamics captures the different time scales of the dynamics of the population and leads to a singular perturbation problem.

The novelty of the model presented here is its capability to replicate a rich class of the stock-recruitment relationship, including the Beverton–Holt, Ricker and Deriso dynamics. The results are explained using geometric singular perturbation theory and illustrated by numerical simulations.

Keywords: Slow-fast; Population dynamic; Stock-recruitment relationship; Geometric singular perturbation theory; Stage-structured model

A

1. Introduction

Fish populations decrease due to (natural and fishing) mortality and increase when new generations of fish resulting from egg production are added to the existing population. In the literature, the spawning stock refers to the part of the stock that is matured enough (called spawners) to contribute to the reproduction process, while stock-recruitment means the relationship between the spawning stock and number of young fish resulting from egg production.

Recruitment is traditionally described as a function of the number of spawners (spawning stock size). The two most popular representations of the relationship have been introduced by Beverton–Holt [2] and Ricker [15]. Both models assume a decrease in recruitment per spawner with stock size, but the degree of decline, often referred to as

* Corresponding author.
E-mail addresses: ute.schaarschmidt@ii.uib.no (U. Schaarschmidt), Trond.Steihaug@ii.uib.no (T. Steihaug), samuel.subbey@imr.no (S. Subbey).

https://doi.org/10.1016/j.matcom.2017.10.008
0378-4754/

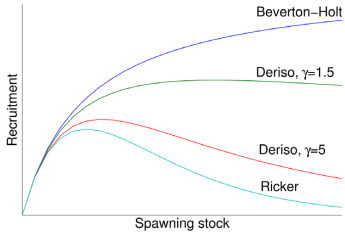


Fig. 1. Traditional stock-recruitment relationships.

Table 1
Processes described by the age-structured model (4)–(5).

$\sum_{i=1}^{n} l_i f_i X_i(t)$	Rate of egg production
$-m_i X_i(t)$	Rate of mortality of individuals in age-class $i = 0, \ldots, n$
$-\sum_{i=1}^{n} p_i X_i(t) X_0(t)$	Rate of mortality as function of numbers of adults
$-p_0 X_0(t)^2$	Rate of mortality as function of the number of prerecruits
$\pm\alpha X_i(t)$	Rate of ageing of individuals in age-class $i = 0, \ldots, n$

density-dependence, distinguishes the two models from each other. A Beverton–Holt function, as described by Eq. (1), is strictly increasing with an asymptotic maximum. A Ricker type of stock-recruitment relationship (SRR) given by Eq. (2) has a maximum and is dome-shaped, i.e., recruitment may decrease with increasing spawning stock size (as illustrated in Fig. 1). The Deriso model [4] corresponds to the Ricker model for $\gamma \to \infty$ and to the Beverton–Holt model for $\gamma = 1$, see Eq. (3). Here, R denotes recruitment and S the number of spawners. In the following, we let $a, b, \gamma \in \mathbb{R}_+ \setminus \{0\}$, with $\mathbb{R}_+$ the set of non-negative real numbers.

$$\text{Beverton–Holt:} \quad R = \frac{aS}{1+bS} \tag{1}$$

$$\text{Ricker:} \quad R = aSe^{-bS} \tag{2}$$

$$\text{Deriso:} \quad R = aS(1+\frac{b}{\gamma}S)^{-\gamma}. \tag{3}$$

Not only is recruitment a function of numbers of spawners, but the spawning stock consists of survivors of previous recruitments. This two-sided relationship has for example been investigated by Touzeau and Gouzé [21]. A prerecruit stage (including eggs, larvae and juveniles) is added to a continuous time population dynamic model with discrete age structures. For further discussion about age-structured models, see e.g. [13]. Dynamic processes at the prerecruit stage evolve on shorter time scales than at other stages and in [21], two distinct time scales are used.

More specifically, a system of differential equations describes the dynamic behaviour of numbers of prerecruits $X_0(t)$ and adults $X_i(t)$ of age-class $i = 1, \ldots, n$ at time $t \geq 0$, with positive natural number n. The processes described by Eqs. (4)–(5) are listed in Table 1. Initial conditions of the form $X_i(0) \geq 0$ for $i = 0, \ldots, n$ ensure that $X_i(t) \geq 0$ for all $t > 0$ and $i = 0, \ldots, n$, [21]. The rate (of change in numbers of individuals with time) of egg production is assumed to be proportional to the numbers of adults and the average rate $l_i > 0$ of eggs spawned per individual of age i. The rate of mortality per prerecruit is assumed to be linear in X_0 and X_i, for $i = 1, \ldots, n$. The parameters $p_0 \geq 0$ and $p_i \geq 0$ express the degree of the limiting effect of the numbers of prerecruits and adult age-class i, respectively. Rates of mortality per individual of age i, which are independent of X_j, $j = 0, \ldots, n$, are called density-independent and are denoted by m_i, for $i = 0, \ldots, n$. The rate of ageing per individual is $\alpha > 0$. A nomenclature including parameters and their units is given in Table A.2. In the following, we let parameters be chosen as described in Table A.2. We remark that the parameters may in practice vary with time, but the assumption of time-independent model parameters is standard practice in fisheries science, [14].

In order to cope with the fact that evolution of prerecruits happens at faster rates than ageing and mortality of adult fish, a fast time T and a slow time $t = \epsilon T$ are used. The parameter $0 < \epsilon \ll 1$ describes the ratio between the two

A

time scales. For reasons of simplicity, let $X(t)$ denote the $(n+1)$-dimensional vector $(X_0(t), X_1(t), \ldots, X_n(t))^t$. The function $A_{TG} : \mathbb{R}^{n+1} \to \mathbb{R}$ defined in Eq. (4) summarizes the fast events associated with spawning and mortality of prerecruits. Changes in numbers of prerecruits consist (partly) of rapidly varying processes, while sizes of adult age-classes are slow variables. The processes which drive the dynamics of adults in age-classes $i = 1, \ldots, n$ are ageing and natural mortality. Summarizing, Touzeau and Gouzé [21] derived the following model,

$$\dot{X}_0(t) = -\alpha X_0(t) + \frac{1}{\epsilon} \underbrace{\left[-m_0 X_0(t) + \sum_{i=1}^{n} f_i l_i X_i(t) - \sum_{i=1}^{n} p_i X_i(t) X_0(t) - p_0 X_0(t)^2 \right]}_{=A_{TG}(X(t))}, \tag{4}$$

$$\dot{X}_i(t) = \alpha X_{i-1}(t) - \alpha X_i(t) - m_i X_i(t), \quad i = 1, \ldots, n. \tag{5}$$

The age-structured model also characterizes the dynamics of spawning stock size and recruitment. At any point in time, recruitment is assumed to be a proportion α of the number of prerecruits and the spawning stock is the sum of fecund fish (see Eq. (6), where $f_i \geq 0$ denotes the fraction of spawners in age-class $i = 1, \ldots, n$).

$$R(t) = \alpha X_0(t) \quad \text{and} \quad S(t) = \sum_{i=1}^{n} f_i X_i(t). \tag{6}$$

Dynamical systems where a subset of variables is assumed to change at a faster rate than the rest of the variables are often referred to as slow-fast systems. Continuous slow-fast systems are singularly perturbed differential equations, since they behave singularly in the limiting case $\epsilon \to 0$. For a general literature about slow-fast systems, see [23], while specific fisheries applications include [1,3] and references therein. A method for reduction of the number of dependent variables (e.g. age-classes) of a (population) dynamical system, which bases on geometric singular perturbation theory, is described in [1].

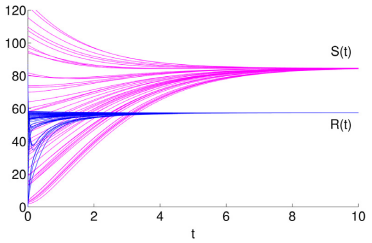
In [21], singular perturbation theory is used to show that for ϵ sufficiently small, the dynamic behaviour of the system (4)–(5) can in general be separated into the following two phases. In the first phase, fast processes are dominant. The number of prerecruits is subject to larger variations than numbers of adults in age-class i, for all $i = 1, \ldots, n$. This is illustrated in Fig. 2(a). In the second phase, solutions remain in a neighbourhood of a surface described by the set of zeros of function A_{TG} (given by equation $A_{TG}(X) = 0$). The equation $A_{TG}(X) = 0$ describes a relationship between the fast variable X_0 and the slow variables X_i. Due to the connection between prerecruits and recruitment, as well as adults and spawners (given by Eq. (6)), this yields a SRR (as illustrated in Fig. 2(b) and (c)). Under suitable conditions, the dynamical system has an unstable equilibrium point at 0 and an asymptotically stable fixed point, [21]. Here and in the following, we refer to constant solutions of an ordinary differential equation (ODE) as fixed or equilibrium points. For further discussion of fixed points and definitions of their stability, see e.g. [24, Ch. 1].

As shown in [21], a function assigning a unique value for recruitment to a spawning stock size may in general not exist. But, assuming that the parameters in Eq. (4) are independent on age and positive ($l_i = l > 0$, $f_i = f > 0$, $p_i = p > 0$, for all $i = 1, \ldots, n$), $A_{TG}(X(t)) = 0$ yields a function which links recruitment to the spawning stock size. For $p_0 = 0$, the approximation for the SRR (7) corresponds to a Beverton–Holt function.

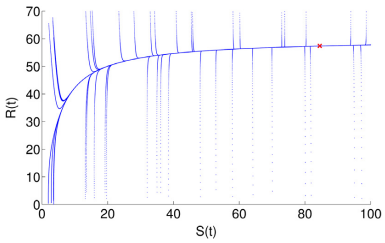
$$R(t) = \frac{\frac{\alpha l}{m_0} S(t)}{1 + \frac{p}{m_0 f} S(t)}. \tag{7}$$

The slow-fast population dynamic model by Touzeau and Gouzé with prerecruit stage explains a Beverton–Holt type of stock-recruitment function, but a Ricker type of SRR cannot be derived from the population dynamic model. Further, it has been shown that the SRR obtained under assumption of positive and age-independent parameters (with exception of $m_i \geq 0$) is strictly increasing, [21]. The system of differential equations in case of $p_0 \neq 0$ is described in [21] and we let $p_0 = 0$ for reasons of simplicity.

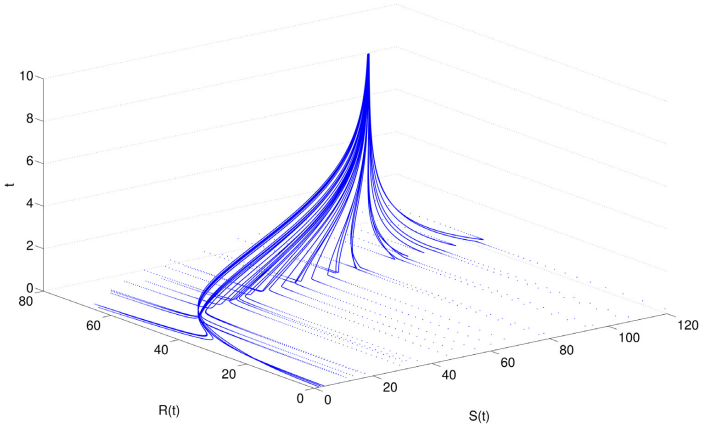
A system of slow-fast ordinary differential equations, which could replicate a rich class of SRRs, including the Ricker and Beverton–Holt functions would be of interest for several reasons. Heterogeneity of time scales is a well-known challenge in modelling fish populations, [16, Ch. 3]. The durations of egg and larval stage are of the order of a couple of days and weeks, respectively. Using geometric singular perturbation theory, both fast and slow processes can be examined and we can describe how micro- and macro-scale events influence each other. On the slow time scale, a relationship between the spawning stock and recruitment emerges. The relationship between the parental population and reproduction is important for management of fisheries ([6, Ch. 7]; [22]) and its description is often given by



(a) Spawning stock size $S(t)$ and recruitment $R(t)$ as described by the model by Touzeau and Gouzé as functions of time.



(b) After a short, first phase, all trajectories evolve similarly to a Beverton–Holt stock-recruitment curve. The cross marks the equilibrium states of recruitment and spawning stock size.



(c) Independent of the initial states, the SRR soon behaves similarly to a Beverton–Holt curve. The vertical axis is time t.

Fig. 2. Numerical solution of the SRR as described by the model (4)–(6) by Touzeau and Gouzé. The numerical experiments are described in Section 4.

or based on Ricker and Beverton–Holt functions, see e.g. ([6, Ch. 7]; [8, Ch. 8]). The model assumptions by Ricker explain a dome-shape of stock-recruitment curves, which may arise from important biological mechanisms. Examples are aggregation of cannibals or predators or density-dependence of growth rates coupled with size-dependent predation ([6, Ch. 7]; [15]). Recent reviews on stock-recruitment models include [14,20].

We keep the structure of a system of differential equations with two time scales, which describes the changes in numbers of adults and prerecruits. But we define a set of dynamical systems that is able to explain a set of SRRs including the Ricker model.

The next section starts with a short introduction to geometric singular perturbation theory, application of which allows us to derive a population dynamic system. This is able to replicate the typical behaviour of recruitment as a function of the stock size in the sense of both Ricker and Beverton–Holt. Additionally, the model presented here can be interpreted as a generalization of the model introduced in [21]. This and other properties of the model presented here are summarized in Section 3 and illustrated in Section 4. Conclusions are given in Section 5.

A

2. Deriving a slow-fast population model using geometric singular perturbation theory

2.1. Introduction to geometric singular perturbation theory

This subsection is based on [7,9]. Geometric singular perturbation theory provides tools for investigation of the geometric properties of slow-fast systems of the form (F_ϵ). Here, vector $x \in \mathbb{R}^k$ consists of variables which change at fast rates, vector $y \in \mathbb{R}^l$ summarizes the slow variables and the prime sign "′" denotes $\mathrm{d}/\mathrm{d}T$. Furthermore, f and g denote smooth functions, $f, g \in C^\infty$. The set of differential equations can also be expressed in terms of slow time $t = \epsilon T$. It is then referred to as the slow system (S_ϵ).

$$\begin{aligned} x'(T) &= f(x(T), y(T), \epsilon) \\ y'(T) &= \epsilon g(x(T), y(T), \epsilon) \end{aligned} \tag{F_ϵ}$$

$$\begin{aligned} x'(T) &= f(x(T), y(T), 0) \\ y'(T) &= 0 \end{aligned} \tag{F_0}$$

$$\begin{aligned} \epsilon \dot{x}(t) &= f(x(t), y(t), \epsilon) \\ \dot{y}(t) &= g(x(t), y(t), \epsilon) \end{aligned} \tag{S_ϵ}$$

$$\begin{aligned} 0 &= f(x(t), y(t), 0) \\ \dot{y}(t) &= g(x(t), y(t), 0). \end{aligned} \tag{S_0}$$

Geometric singular perturbation theory studies the persistence of characteristics of the reduced forms (the case $\epsilon \to 0$) of the two systems for the general case $0 < \epsilon \ll 1$. More specifically, the fast and the slow system are reduced to lower dimensional problems (F_0) and (S_0), respectively, for the limiting case $\epsilon \to 0$. Central to the geometric theory of singular perturbations is a compact manifold (possibly with boundary) contained in the set of zeros of f, the *critical manifold* $M_0 \subset \{(x, y)^t \in \mathbb{R}^{k+l} | f(x, y, 0) = 0\}$. Orbits of the reduced slow system (S_0) are restricted to the set of zeros of function f. From point of view of the reduced fast system (F_0), the critical manifold is a set of equilibrium points.

Let $z(t, z_0, 0)$ denote the solution of the initial value problem (F_ϵ) with $z(0) = z_0$. A set M is called *locally invariant under* (F_ϵ), if it has a neighbourhood V such that $z_0 = (x_0,\ y_0)^t \in M$ with $z(t, z_0, 0) \in V$ for all $t \in [0, \bar{t}]$ and for $\bar{t} \in \mathbb{R}_+$ implies $z(t, z_0, 0) \in M$ for all $t \in [0, \bar{t}]$, and if $z_0 = (x_0,\ y_0)^t \in M$ with $z(t, z_0, 0) \in V$ for all $t \in [-\bar{t}, 0]$, then $z(t, z_0, 0) \in M$ for all $t \in [-\bar{t}, 0]$. In the following, we refer to $\mathcal{W}^s(M_0)$ defined by Eq. (8) as the *stable manifold of* M_0. Here, V denotes a neighbourhood of z^* and $z(t, z_0, 0)$ the solution of the initial value problem (F_0) with $z(0) = z_0$. $\mathcal{W}^s(M_0)$ is the union of the stable manifolds of fixed points $z^* \in M_0$. The unstable manifold of M_0 is defined by Eq. (9). The definition can e.g. be found in [10, p. 20 and p. 55], but our notation follows [7].

$$\mathcal{W}^s(M_0) = \bigcup_{z^* \in M_0} \{z_0 \in V | z(t, z_0, 0) \in V\ \forall t \geq 0 \text{ and } z(t, z_0, 0) \to z^* \text{ exponentially as } t \to \infty\} \tag{8}$$

$$\mathcal{W}^u(M_0) = \bigcup_{z^* \in M_0} \{z_0 \in V | z(t, z_0, 0) \in V\ \forall t \leq 0 \text{ and } z(t, z_0, 0) \to z^* \text{ exponentially as } t \to -\infty\}. \tag{9}$$

For $\epsilon > 0$ and sufficiently small, results by Fenichel [5] include conditions sufficient for the existence of a manifold M_ϵ (possibly with boundary) which is locally invariant under (F_ϵ) and $\mathcal{O}(\epsilon)$ close and diffeomorphic to M_0 and for the existence of manifolds $\mathcal{W}^s(M_\epsilon)$ and $\mathcal{W}^u(M_\epsilon)$, which are locally invariant under (F_ϵ) and $\mathcal{O}(\epsilon)$ close and diffeomorphic to $\mathcal{W}^s(M_0)$ and $\mathcal{W}^u(M_0)$, respectively. Two manifolds V and W are called diffeomorphic, if there exists a continuously differentiable map $f : V \to W$ which is bijective and its inverse is continuously differentiable. The manifolds with boundary are locally invariant in the sense that trajectories of (F_ϵ) can leave a locally invariant manifold M with boundary only through its boundary.

2.2. Geometric singular perturbation theory for a population dynamic model

In this section, we use geometric singular perturbation techniques to derive a generic population dynamic model that explains a rich class of SRRs including (1)–(3), and for which the model by Touzeau and Gouzé [21] given by

A

Eqs. (4)–(5) with $p_0 = 0$ is a special case. The results by Fenichel which are relevant for the following are referred to as Fenichel's theorem and are restated in Appendix B, where we follow [7].

Consider a generalization of the model (4)–(5), a population dynamic model of form (10) with initial conditions $X(0) \in \mathbb{R}_+^{n+1}$. This is a slow-fast system expressed in terms of the slow time t as in Eq. (S_ϵ). The fast time scale equivalent of (10) is given by Eq. (11). X_0 is the fast variable and there are n slow variables $X_1, \ldots, X_n$.

$$\begin{aligned} \epsilon \dot{X}_0(t) =& -\epsilon\alpha X_0(t) + A(X_0(t), X_1(t), \ldots, X_n(t)) \\ \dot{X}_i(t) =& \alpha X_{i-1}(t) - \alpha X_i(t) - m_i X_i(t), \quad i = 1, \ldots, n \end{aligned} \tag{10}$$

$$\begin{aligned} X_0'(T) =& -\epsilon\alpha X_0(T) + A(X_0(T), X_1(T), \ldots, X_n(T)) \\ X_i'(T) =& \epsilon\left(\alpha X_{i-1}(T) - \alpha X_i(T) - m_i X_i(T)\right), \quad i = 1, \ldots, n. \end{aligned} \tag{11}$$

We restate Fenichel's theorem for the specific case of a slow-fast population dynamic model (11) to obtain a set of hypotheses for the function A.

(h1) Let $A : \mathbb{R} \times \hat{K} \to \mathbb{R}$ be a C^∞ function, for an open set $\hat{K} \subset \mathbb{R}^n$.

Then, the functions $f(X, \epsilon) = -\epsilon\alpha X_0 + A(X_0, X_1, \ldots, X_n)$ and $g(X, \epsilon) = (g_1(X), \ldots, g_n(X))^t$, with $g_i(X) = \alpha X_{i-1} - \alpha X_i - m_i X_i$ for $i = 1, \ldots, n$, are C^∞ on $\mathbb{R} \times \hat{K} \times \mathbb{R}$.

(h2) Let $K \subset \hat{K} \subset \mathbb{R}^n$ be a compact and simply connected set, whose boundary is an $(n-1)$-dimensional C^∞ submanifold.

Define the critical manifold $M_0 = \{X \in \mathbb{R} \times K | A(X_0, X_1, \ldots, X_n) = 0\}$ and assume that

(h3) M_0 is the graph of a C^∞ function $h^0 : K \to \mathbb{R}$ which assigns to each $(X_1, \ldots, X_n)^t \in K$ the $X_0 \in \mathbb{R}$ s.t. $A(X_0, X_1, \ldots, X_n) = 0$.

Assumption (h3) implies that M_0 is given by a function associating the number of prerecruits to the numbers of adults. We consider finite numbers of fish ($X_i \leq \bar{S}$ for $i = 1, \ldots, n$ and $\bar{S} > 0$) and assume we may choose a compact set $K \supset [0, \bar{S}]^n$ which contains the domain of interest. Premises (h2) and (h3) correspond to hypothesis (H3) in [7] and Appendix B.

Since M_0 is the graph of a C^∞ function on compact set K described by (h2), it is a compact manifold with boundary. One of the hypotheses of Fenichel's theorem is that the critical manifold is such that $(\partial f(X_0, X_1, \ldots, X_n, 0)/\partial X_0)|_{M_0}$ has no zero real part. We assume that

(h4) the derivative of A with respect to X_0 on M_0 is negative,

$$\left.\frac{\partial A(X_0, X_1, \ldots, X_n)}{\partial X_0}\right|_{M_0} < 0\,. \tag{12}$$

For the slow-fast population dynamic model (4)–(5) by Touzeau and Gouzé, we have (13)

$$\frac{\partial A_{TG}(X_0, X_1, \ldots, X_n)}{\partial X_0} = -m_0 - \sum_{i=1}^{n} p_i X_i - 2p_0 X_0 < 0, \tag{13}$$

for $m_0 > 0$, $p_i \geq 0$ and $X_i \geq 0$ with $i = 0, \ldots, n$. This entails that the spawned population size less the proportion of prerecruits that dies due to natural mortality, decreases strictly with the number of prerecruits. Let $(X_1, \ldots, X_n)^t \in K$. The point $X^* = (h^0(X_1, \ldots, X_n), X_1, \ldots, X_n)^t \in \mathbb{R} \times K$ is a fixed point of the reduced fast system given by Eq. (14).

$$\begin{aligned} X_0'(T) =& A(X_0(T), X_1(T), \ldots, X_n(T)) \\ X_i'(T) =& 0, \quad i = 1, \ldots, n. \end{aligned} \tag{14}$$

Trajectories of the reduced fast system which start in $(X_0, X_1, \ldots, X_n)^t$, for $X_0 \in \mathbb{R}$, remain in manifold $\{(X_0, X_1, \ldots, X_n)^t | X_0 \in \mathbb{R}\}$ and converge under (h4) exponentially to the fixed point X^* as $t \to \infty$. The fixed point has a one-dimensional stable manifold $\{(X_0, X_1, \ldots, X_n)^t | X_0 \in \mathbb{R}\}$. Since we considered an arbitrary $(X_1, \ldots, X_n)^t \in K$, we obtain $\mathcal{W}^s(M_0) = \mathbb{R} \times K$ by Eq. (8).

Given hypotheses (h1)–(h4) and for $\epsilon > 0$ sufficiently small, we obtain the following conclusions from Fenichel's theorem as provided in [7] and restated in Appendix B. By Theorems 1 and 2 in [7], there exists a function h^ϵ defined

A

on K, such that the graph M_ϵ (15) is locally invariant under the fast system (11) and h^ϵ is C^r, $\forall r < \infty$, jointly in y and ϵ. The graph M_ϵ is $\mathcal{O}(\epsilon)$ close and diffeomorphic to M_0.

$$M_\epsilon = \{(X_0, X_1, \ldots, X_n)^t \in \mathbb{R}^{n+1} | X_0 = h^\epsilon(X_1, \ldots, X_n),\ (X_1, \ldots, X_n)^t \in K\}. \tag{15}$$

From Theorem 3 in [7], it follows that there exists a stable manifold $\mathcal{W}^s(M_\epsilon)$, which is locally invariant under the fast system (11) and an $\mathcal{O}(\epsilon)$-perturbation of and diffeomorphic to $\mathcal{W}^s(M_0) = \mathbb{R} \times K$. Since the stable manifold of M_ϵ is diffeomorphic to the non-empty manifold $\mathcal{W}^s(M_0)$ and $\mathcal{W}^s(M_0)$ has dimension $n + 1$, it follows that $\mathcal{W}^s(M_\epsilon)$ has dimension $n + 1$ (e.g. by Theorem 2.17 in [12]). Further, let V be a neighbourhood of M_ϵ. By Theorem 5 in [7], trajectories of the slow-fast system (11) which start in $\mathcal{W}^s(M_\epsilon)$ and remain in neighbourhood V converge exponentially to M_ϵ.

Thus, under (h1)–(h4) and for $0 < \epsilon$ sufficiently small, trajectories of the slow-fast population dynamic model (10) which start in $\mathcal{W}^s(M_\epsilon)$ and stay within a neighbourhood of M_ϵ converge exponentially to the graph of h^ϵ, which is locally invariant. The function h^ϵ can be approximated by solving $A(X) = 0$ for X_0 and assigns the number of prerecruits to numbers of adults. Since spawning stock size is a weighted sum of numbers of adults and the number of prerecruits is assumed to be proportional to recruitment, h^ϵ also describes a link between spawning stock size and recruitment (but not necessarily a stock-recruitment function).

2.3. A slow-fast population dynamic model

Specification of the function A and introduction of a parameter $\gamma \in \mathbb{R}_+ \setminus \{0\}$ will yield a population dynamic model, which explains the set of SRRs (1)–(3). We aim at generating a Deriso type of SRR which contains the Ricker model and the Beverton–Holt model as special cases. The Deriso function, with recruitment and spawning stock size as functions of time, is given by Eq. (16). We use Eq. (6) to express the Deriso model (16) in terms of numbers of adults and prerecruits and obtain Eq. (17). Here, we allow age-dependent and non-negative parameters $a_i, b_i \in \mathbb{R}_+$ with $i = 1, \ldots, n$.

$$R(t) = aS(t)(1 + \frac{1}{\gamma} bS(t))^{-\gamma}, \tag{16}$$

$$\alpha X_0(t) = (\sum_{i=1}^{n} a_i f_i X_i(t)) \left(1 + \frac{1}{\gamma} (\sum_{i=1}^{n} b_i f_i X_i(t))\right)^{-\gamma}. \tag{17}$$

Now, define a function $A_\gamma : \mathbb{R} \times \hat{K} \to \mathbb{R}$ with the following properties:

- Eq. (17) describes the solutions of $A_\gamma(X_0, X_1, \ldots, X_n) = 0$ for $(X_1, \ldots, X_n)^t \in K$ for a suitable set $K \supset [0, \bar{S}]^n$, which is introduced in Section 2.4
- and the function A_1 (with $\gamma = 1$) is the function A_{TG} defined in Eq. (4) in case of $p_0 = 0$.

This can be achieved by setting $a_i = \alpha l_i / m_0$, $b_i = p_i/(m_0 f_i)$ and defining A_γ on $\mathbb{R} \times \hat{K}$ with $\hat{K} = \{(X_1, \ldots, X_n)^t \in \mathbb{R}^n | 1 + (\sum_{i=1}^n p_i X_i(t))/(\gamma m_0) > 0\}$ by

$$A_\gamma(X(t)) = \left[-m_0 X_0(t) \left(1 + \frac{1}{\gamma m_0} \sum_{i=1}^{n} p_i X_i(t)\right)^{\gamma} + \sum_{i=1}^{n} l_i f_i X_i(t) \right]. \tag{18}$$

For parameters chosen as described in Table A.2, the set $\hat{K}$ contains the domain of interest $[0, \bar{S}]^n$, for $\bar{S} > 0$.

The slow-fast population dynamic model is described by Eqs. (19)–(20).

$$\dot{X}_0(t) = -\alpha X_0(t) + \frac{1}{\epsilon} \underbrace{\left[-m_0 X_0(t) \left(1 + \frac{1}{\gamma m_0} \sum_{i=1}^{n} p_i X_i(t)\right)^{\gamma} + \sum_{i=1}^{n} l_i f_i X_i(t) \right]}_{= A_\gamma(X(t))} \tag{19}$$

$$\dot{X}_i(t) = \alpha X_{i-1}(t) - \alpha X_i(t) - m_i X_i(t), \quad i = 1, \ldots, n. \tag{20}$$

The slow-fast population dynamic model describes ageing, mortality and egg production and may be justified by the same mechanisms as a Ricker type of SRR. The function A_γ describes the rate of mortality of prerecruits as function

of the numbers of adults, which is non-linear in the sizes of age-classes X_i, $i = 1, \ldots, n$ for $\gamma \in \mathbb{R}_+ \setminus \{0, 1\}$. Proportionality between the rate of mortality and spawning stock size is a biological assumption which may explain the Ricker type of SRR. More details may e.g., be found in [6, Ch. 7].

2.4. Geometric singular perturbation theory for the population dynamic model

In the following, we verify that assumptions (h1)–(h4) hold for A_γ (18) and that the population dynamic model (19)–(20) yields a representation of numbers of prerecruits in terms of numbers of adults, which can be approximated by a Deriso type of relationship (21).

$$X_0(t) = (\sum_{i=1}^{n} \frac{l_i f_i}{m_0} X_i(t)) \left(1 + \frac{1}{\gamma m_0} (\sum_{i=1}^{n} p_i X_i(t))\right)^{-\gamma}. \tag{21}$$

The function A_γ is C^∞ on $\mathbb{R} \times \hat{K}$ and $\hat{K}$ is an open set. The critical manifold $M_0 = \{X \in \mathbb{R} \times K | A_\gamma(X) = 0\}$ is the graph of function $h^0 : K \to \mathbb{R}$ defined by $h^0(X_1(t), \ldots, X_n(t)) = X_0(t)$, with $X_0(t)$ given by Eq. (21). Concerning the choice of a suitable domain K, note that the function h^0 is well defined on $\hat{K}$. Since $\hat{K}$ is open and $[0, \bar{S}]^n$ is compact and convex, we can enlarge $[0, \bar{S}]^n$ within $\hat{K}$ to a compact and convex set K (such that $[0, \bar{S}]^n \subset K \subset \hat{K} \subset \mathbb{R}^n$), whose boundary is an $(n-1)$-dimensional C^∞ submanifold. $K \subset \mathbb{R}^n$ convex implies K simply connected. A definition and properties of simply connected sets can e.g. be found in appendix A in [12]. Summarizing, there exists a domain K which fulfils assumption (h2) and contains $[0, \bar{S}]^n$.

The function h^0 is the product of function $q(X_1, \ldots, X_n) = [1 + (\sum_{i=1}^{n} p_i X_i(t))/(\gamma m_0)]^{-\gamma}$ and a linear function. The term $1 + (\sum_{i=1}^{n} p_i X_i(t))/(\gamma m_0)$ is positive for all $(X_1, \ldots, X_n)^t \in \hat{K}$ and the function $q(X_1, \ldots, X_n)$ is defined and C^∞ on $\hat{K}$. Thus, h^0 is C^∞ and assumption (h3) holds.

The image $h^0([0, \bar{S}]^n)$ is a subset of $\mathbb{R}_+$ and the numbers of prerecruits are non-negative. Concerning (h4), Eq. (22) shows that the derivative of A_γ with respect to X_0 on M_0 is negative for $m_0 > 0$, since $1 + (\sum_{i=1}^{n} p_i X_i(t))/(\gamma m_0) > 0$ on M_0.

$$\left.\frac{\partial A_\gamma(X_0, X_1, \ldots, X_n)}{\partial X_0}\right|_{M_0} = -m_0 \left(1 + \frac{1}{\gamma m_0} \sum_{i=1}^{n} p_i X_i(t)\right)^{\gamma}. \tag{22}$$

How realistic the assumption $\epsilon > 0$ sufficiently small is, depends on the definition of recruitment, and the duration time of the recruitment period in relationship to the total life span of a particular fish species; i.e., short or long lived species. For example, rates of mortality of eggs and larvae may be of order 0.1, of order 0.01 for juveniles and of order 0.001 for adults, [19]. If recruitment is defined as number of juveniles, one might assume $\epsilon \approx 0.1$.

3. Properties of the slow-fast population dynamic model

3.1. Stock-recruitment relationships

As the slow-fast population dynamic model (19)–(20) is a special case of the dynamical system (10) investigated in Section 2.2, the behaviour of its trajectories can be explained using geometric singular perturbation theory. In this section, we let $\epsilon > 0$ sufficiently small. In a first, short phase, trajectories are governed by the fast system (11). The spawning stock size undergoes small changes in comparison to the number of prerecruits. Trajectories which start in $\mathcal{W}^s(M_\epsilon)$ are attracted to the graph of a function h^ϵ, as long as they stay in a neighbourhood of the graph. The function h^ϵ is an $\mathcal{O}(\epsilon)$-perturbation of h^0 described by Eq. (21). Thus, in the second phase, reproduction can be approximated by a Deriso type of relationship (23). On the critical manifold, recruitment is given as a function of numbers of adults, but in general not of the spawning stock size. However, in case of positive age-independent parameters ($l_i = l > 0$, $f_i = f > 0$ and $p_i = p > 0$, for $i = 1, \ldots, n$), the critical manifold can be associated with a Deriso function (24), with $a = \alpha l / m_0$ and $b = p/(m_0 f)$.

$$R(t) = \alpha (\sum_{i=1}^{n} \frac{l_i f_i}{m_0} X_i(t)) \left(1 + \frac{1}{\gamma m_0} (\sum_{i=1}^{n} p_i X_i(t))\right)^{-\gamma} \tag{23}$$

$$R(t) = \frac{\alpha l}{m_0} S(t) \left(1 + \frac{p}{\gamma m_0 f} S(t)\right)^{-\gamma}. \tag{24}$$

By variations of parameter γ, the model (19)–(20) can replicate a rich class of the link between adult population and reproduction. The parameter γ allows to assign distinct degrees of density-dependence to the stock-recruitment function.

3.2. Limiting case

Define the limiting case $\gamma \to \infty$ of the population dynamic system (19)–(20) by Eqs. (25) and (20). The slow-fast population dynamic model is given by the two systems of differential equations. Analogously to Section 2.4, we can show that function A_∞ fulfils assumptions (h1)–(h4). The set of zeros of function $A_\infty : \mathbb{R}^{n+1} \to \mathbb{R}$ is the set of $(X_0(t), X_1(t), \ldots, X_n(t))^t \in \mathbb{R}^{n+1}$ for which Eq. (26) holds. The function h^0 defined by $h^0(X_1(t), \ldots, X_n(t)) = X_0(t)$, with $X_0(t)$ given by Eq. (26) is well defined and C^∞ on $\mathbb{R}^n$. We may define a manifold M_0 for any compact and convex set K such that $[0, \bar{S}]^n \subset K \subset \mathbb{R}^n$ and the boundary of K is an $(n-1)$-dimensional C^∞ submanifold, including the set K defined in Section 2.4.

$$\dot{X}_0(t) = -\alpha X_0(t) + \frac{1}{\epsilon} \underbrace{\left[-m_0 X_0(t) \cdot \exp\left(\frac{1}{m_0} \sum_{i=1}^{n} p_i X_i(t) \right) + \sum_{i=1}^{n} l_i f_i X_i(t) \right]}_{=A_\infty(X(t))} \tag{25}$$

$$R(t) = \alpha X_0(t) = \frac{\alpha}{m_0} \left(\sum_{i=1}^{n} l_i f_i X_i(t) \right) \cdot \exp\left(-\frac{1}{m_0} \sum_{i=1}^{n} p_i X_i(t) \right). \tag{26}$$

Trajectories of the system of differential equations, which start in $\mathcal{W}^s(M_\epsilon)$ and remain in a neighbourhood of M_ϵ, converge to the graph of a perturbation of the Ricker type of stock-recruitment function (26).

3.3. Fixed points

The population dynamic model (19)–(20) has an equilibrium point $X_0 = X_1 = \cdots = X_n = 0$. Further, if the parameters are chosen as described in Table A.2 and in addition

- Assumption I: $\alpha\epsilon + m_0 < \sum_{i=1}^{n} f_i l_i \pi_i$
- Assumption II: $\exists i^* \in \{1, \ldots, n\}$ s.t. $p_{i^*} > 0$,

then a second fixed point is X^{**} as given by

$$X_0^{**} = \frac{\gamma m_0}{\sum_{i=1}^{n} p_i \pi_i} \left[\left(\frac{\sum_{i=1}^{n} l_i f_i \pi_i}{m_0} - \frac{\alpha\epsilon}{m_0} \right)^{\frac{1}{\gamma}} - 1 \right], \tag{27}$$

$$X_i^{**} = \pi_i X_0^{**}, \text{ with } \pi_i = \prod_{j=1}^{i} \frac{\alpha}{\alpha + m_j}, \; i = 1, \ldots, n\,. \tag{28}$$

We prove the existence of $X^{**} > 0$ for $\gamma > 0$. The case $\gamma = 1$ has been considered in [21]. From $\alpha > 0$ and $m_i \geq 0$, it follows that $\pi_i > 0$ for $i = 1, \ldots, n$. By definition of the slow-fast population dynamic model (20), Eq. (28) implies constant numbers of adults. Thus, we can substitute the fractions of X_0^{**} for X_i^{**} into Eq. (19) and obtain that $\dot{X}_0^{**} = 0$ is equivalent with Eq. (29). Assuming $X_0^{**} \neq 0$ and with $m_0 > 0$, we obtain (30). Using assumption II and $\pi_i > 0$ for $i = 1, \ldots, n$, the fixed point is given by Eq. (27). $X_0^{**} > 0$ and $X_i^{**} > 0$ follow from assumptions I, II, non-negativity of $p_i \geq 0$ and positivity of $m_0, \gamma, \pi_i > 0$ for $i = 1, \ldots, n$.

$$\epsilon\alpha X_0^{**} = -m_0 X_0^{**} \left(1 + \frac{1}{\gamma m_0} \sum_{i=1}^{n} p_i \pi_i X_0^{**} \right)^{\gamma} + \sum_{i=1}^{n} l_i f_i \pi_i X_0^{**} \tag{29}$$

$$\left(1 + \frac{1}{\gamma m_0} \sum_{i=1}^{n} p_i \pi_i X_0^{**} \right)^{\gamma} = \frac{\sum_{i=1}^{n} l_i f_i \pi_i}{m_0} - \frac{\alpha\epsilon}{m_0}. \tag{30}$$

A

In terms of recruitment and number of spawners, the equilibrium point (27)–(28) is given by

$$R_\gamma^{**} = \frac{\gamma m_0 \alpha}{\sum_{i=1}^n p_i \pi_i} \left[\left(\frac{\sum_{i=1}^n l_i f_i \pi_i}{m_0} - \frac{\alpha\epsilon}{m_0} \right)^{\frac{1}{\gamma}} - 1 \right] , \tag{31}$$

$$S_\gamma^{**} = \sum_{i=1}^n f_i \pi_i \frac{1}{\alpha} R_\gamma^{**} . \tag{32}$$

It can be shown that if assumptions I and II hold, then R_γ^{**} (as a function of parameter γ) is strictly decreasing for $\gamma > 0$. To see this, substitute $\theta_1 = (m_0\alpha)/(\sum_{i=1}^n p_i \pi_i)$ and $\theta_2 = (\sum_{i=1}^n l_i f_i \pi_i - \alpha\epsilon)/m_0$ into Eq. (31) and obtain (33). Here, we use that assumptions I and II and the choice of the parameters described in Table A.2 imply that $\theta_1 > 0$ and $\theta_2 > 1$.

$$R_\gamma^{**} = \theta_1 \gamma (\theta_2^{1/\gamma} - 1) , \quad \text{with} \quad \partial R_\gamma^{**}/\partial\gamma = \theta_1 \left(\theta_2^{1/\gamma} - 1 - \theta_2^{1/\gamma} \ln(\theta_2^{1/\gamma}) \right) . \tag{33}$$

The function $p : (0, \infty) \to \mathbb{R}$ defined by $p(x) = x - 1 - x \ln(x)$ is differentiable with $p'(x) = -\ln(x)$. Since p is non-increasing for $x \in [1, \infty)$, strictly decreasing for $x \in (1, \infty)$ and has zero $p(1) = 0$, we conclude that $p(x) < 0$ for all $x > 1$. With $\theta_2^{1/\gamma} > 1$, it follows that the first derivative of R_γ^{**} with respect to γ is negative.

The slow-fast population dynamic model described by Eqs. (25) and (20) has also an equilibrium point $X_0 = X_1 = \cdots = X_n = 0$ and, under assumptions I and II and for the choices of parameters described in Table A.2, a positive fixed point. The existence of the positive fixed point given by Eqs. (34), (28) and (32) can be proven analogously to the case $\gamma > 0$.

$$R^{**} = \alpha X_0^{**} = \frac{m_0 \alpha}{\sum_{i=1}^n p_i \pi_i} \ln \left(\frac{\sum_{i=1}^n l_i f_i \pi_i}{m_0} - \frac{\alpha\epsilon}{m_0} \right) . \tag{34}$$

3.4. Non-negativity of the variables

The slow-fast population dynamics model represents changes in numbers of individuals. As mentioned in the introduction, for the model by Touzeau and Gouzé (with $p_0 \geq 0$), the solution $X(t)$ of the initial value problem (4)–(5) with non-negative $X(0) \in \mathbb{R}_+^{n+1}$ is non-negative, [21].

Concerning the slow-fast population dynamic model introduced in this paper, let parameters be chosen as described in Table A.2. Then, we have $X(t) \in \mathbb{R}_+^{n+1}$ for all $t > 0$ for all solutions of the initial value problem (19)–(20) with non-negative $X(0) \in \mathbb{R}_+^{n+1}$ and for all solutions $X(t)$ of the initial value problem (25) and (20) with non-negative $X(0) \in \mathbb{R}_+^n$.

To see this, let $t_1 > 0$ the smallest $t > 0$ such that there exists $i \in \{0, 1, \ldots, n\}$ with $X_i(t) \leq 0$. Since $X(t)$ is continuous in $t \geq 0$, we have $X_i(t_1) \geq 0$ for $i = 0, \ldots, n$. Assume $X_i(t_1) = 0$ for $0 < i \leq n$. Then, by Eq. (20) and $\alpha > 0$, we have $\dot{X}_i(t_1) \geq 0$ and X_i remains non-negative. Now assume $X_0(t_1) = 0$. By Eq. (19) (or Eq. (25)), the choice of parameters described in Table A.2 and $X_i(t_1) \geq 0$, we obtain $\dot{X}_0(t_1) = (1/\epsilon) \sum_{i=1}^n l_i f_i X_i(t_1) \geq 0$. Thus, the variable $X_0(t)$ remains non-negative.

4. Numerical simulations

The presence of two distinct time scales is characteristic for stiff systems. Eqs. (19)–(20) and Eqs. (25) and (20) are solved by numerical differentiation formulas for stiff problems as implemented in MATLAB and described in [18]. We use default options for algorithm parameter values except for absolute error tolerance 10^{-10}. For given $\gamma > 0$ and initial conditions $X(0) \in \mathbb{R}_+^{n+1}$, the ODE solver computes an approximation of the solution $X(t)$ of the differential equation at specific points in time $t \in \{0, 0.0005, 0.001, \ldots, 10\}$ by interpolation. Recruitment and spawning stock size are given by Eq. (6). The limit of the slow-fast system as $\gamma \to \infty$ refers to Eqs. (25) and (20). The following age-independent parameter values are used: $n = 4, \alpha = 0.8, \epsilon = 0.1, m_0 = 0.7, p_0 = 0, p = 0.1, f = 0.5, l = 15$ and $m_i = 0.2$ for $i = 1, \ldots, n$. In [21], a realistic example for values for the parameters has been suggested and we use the same parameter values with the following exceptions. Following Section 3.4, we assume $\epsilon = 0.1$. Further, we choose $m_0 = 0.7$ (instead of $m_0 = 0.5$), $p_0 = 0$ and age-independent parameters. Initial conditions are uniformly distributed with $X(0) \in ([0, 1] \cup [2, 7] \cup [85, 115]) \times [0, 57]^4$ (Figs. 2 and 4) and $X(0) \in ([0, 1] \cup [2, 7] \cup [45, 60]) \times [0, 57]^4$

A

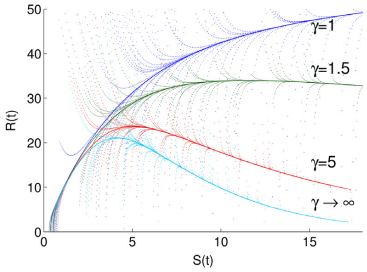


Fig. 3. Simulation of the stock-recruitment relationship corresponding to the slow-fast population dynamic model at $t \in \{0, 0.005, 0.01, \ldots, 10\}$ and for four distinct values for γ.

(Fig. 5). We use MATLAB version R2012a and 64 bit floating-point precision. For the sake of clarity, the figures present a subset of the results of the numerical experiments.

The SRRs corresponding to the slow-fast population dynamic model with positive age-independent parameters are illustrated in Fig. 3. The initial value problems given by Eqs. (19)–(20) or Eqs. (25) and (20) with initial conditions $(X_0(0), X_1(0), \ldots, X_4(0))^t \in [0, 340] \times [0, 12]^4$ are solved numerically and in the cases $\gamma \in \{1, 1.5, 5\}$ and $\gamma \to \infty$. We observe that for our choice of initial conditions, all trajectories approach graphs of stock-recruitment functions. These are approximations of the Beverton–Holt, Deriso and Ricker functions illustrated in Fig. 1. For Fig. 3, $(X_0(0), X_1(0), \ldots, X_4(0))^t \in [0, 340] \times [0, 12]^4$ are not uniformly distributed but chosen with the aim of illustrating the stock-recruitment functions.

For $\gamma = 1$, the system corresponds to the case $p_0 = 0$ of the model by Touzeau and Gouzé [21] described by Eqs. (4) and (5). Under assumptions I and II and further conditions, the equilibrium point (27)–(28), with $\gamma = 1$, is stable, [21]. In case of positive age-independent parameters (with exception of $m_i \geq 0$), trajectories of the system of differential equations, which start in $\mathcal{W}^s(M_\epsilon)$ and remain in a neighbourhood of M_ϵ, approach the graph of a function, which may be approximated by a Beverton–Holt type of SRR. Recruitment per spawner decreases with increasing spawning stock size, but the SRR is strictly increasing on the critical manifold. This is illustrated in Fig. 2. In our simulations, all numerical solutions starting in $([0, 1] \cup [2, 7] \cup [85, 115]) \times [0, 57]^4$ approach the graph of a perturbation of a Beverton–Holt function.

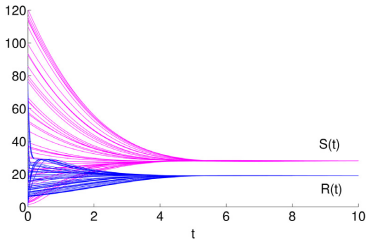
With increasing γ, density-dependence of the stock-recruitment relationship increases. In case of positive and age-independent parameters (with exception of $m_i \geq 0$), a Deriso function describes the relationship between spawning stock and recruitment. This function has a maximum and is dome-shaped for all $\gamma > 1$ (for details, see [17]). Fig. 4 presents the case $\gamma = 2$. For our choice of initial conditions, all trajectories approach the graph of a perturbation of the Deriso function. Here, it can also be observed that the equilibrium recruitment R_γ^{**} is lower in case of $\gamma = 2$ than for $\gamma = 1$.

The limiting case $\gamma \to \infty$ is illustrated in Fig. 5. Comparing Figs. 2(a), 4(a) and 5(a), it can be observed that the value of recruitment at the fixed point is lower for the case $\gamma \to \infty$ than for $\gamma = 2$ and $\gamma = 1$. For about 95% of the initial conditions, the solutions are attracted to the graph of a perturbation of the Ricker type of stock-recruitment function (26). The rest of the numerical solutions is observed to be negative. From Section 3.4, we know that the solution of the initial value problem (25) and (19) with non-negative initial values is non-negative. Negative numerical solutions indicate a failure of the numerical solution of the differential equation. Only biologically plausible solutions are illustrated in Fig. 5.

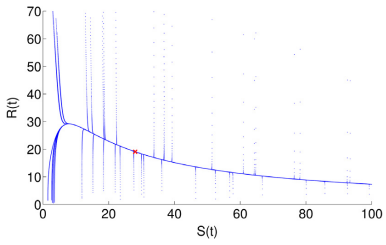
5. Conclusions

The solutions of the slow-fast population dynamic model introduced here, are under suitable assumptions attracted to the graph of a perturbation of a Deriso type of SRR. If the rate of mortality of prerecruits is a non-linear function of the numbers of adults, then a non-monotonic SRR may emerge from a slow-fast population dynamic model. The

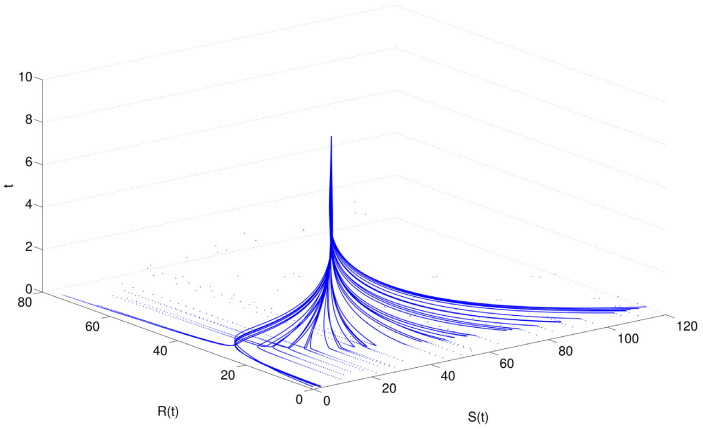
A



(a) Spawning stock size and recruitment as functions of time converge towards a fixed point.



(b) Here, the SRR is dome-shaped and can be approximated by a Deriso function.



(c) Independent of the initial states, the SRR soon behaves similarly to a Deriso function.

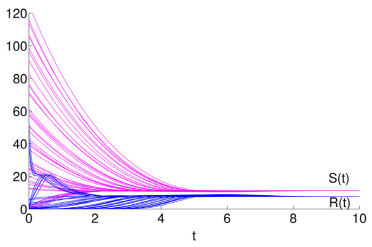
Fig. 4. Numerical solution of the slow-fast population dynamic model with $\gamma = 2$.

A

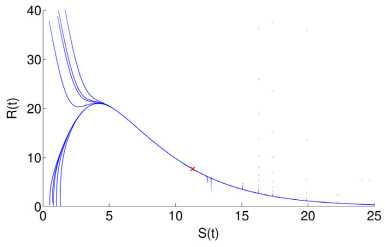
approach proposed in this paper is not restricted to the Deriso model. Given a specific stock-recruitment function and a description of slow dynamics, a set of candidates of descriptions of the fast dynamics can be identified. Sufficient conditions for a slow-fast population dynamic model to generate a specific SRR have been derived using geometric singular perturbation theory.

The assumption of age and time-independent model parameters is a simplification, but is standard practice in fisheries science as pointed out in the introduction. In the case of age-dependent parameters, the population dynamic model describes recruitment as a function of an age-structured spawning stock, but not of the total number of spawners. In this paper, we have assumed parameters such as fecundity and rates of mortality to be time-independent. Analyses such as determination of fixed points necessitate this assumption.

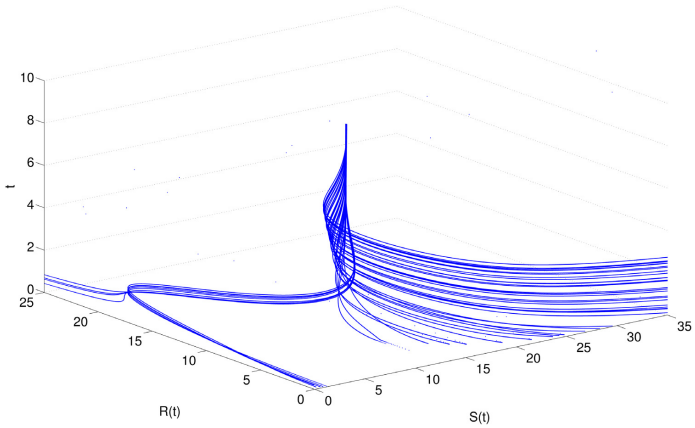
Environmental and physical factors may alter the stock-recruitment function, and affect the complete stock, including asymptotic behaviour and equilibrium states. A SRR describes the endpoint of the development from spawning to recruitment as a function of the number of spawners. Time-dependent parameters in a SRR may reflect long-term variations such as regime shifts. But also daily or seasonal effects, e.g., temperature in the first growing season, may cause high variations in mortality of larvae and recruitment, [11]. The model introduced here, and the one by Touzeau and Gouzé, describes the fast evolution of numbers of prerecruits. It can therefore reflect high frequency



(a) Spawning stock size and recruitment as functions of time converge towards a fixed point.



(b) The stock-recruitment relationship can be approximated by a Ricker function.



(c) Independent of the initial states, the SRR soon behaves similarly to a Ricker function.

Fig. 5. Numerical solution of the slow-fast population dynamic model with $\gamma \to \infty$. The axes are shorter than in Figs. 2 and 4.

A

variations in parameters such as natural mortality and describe the reaction of the dynamics of stock and recruitment to these variations. The strength of cannibalism or predation may also change, and thus affect density-dependence. This may be investigated using the slow-fast population dynamic model presented here, because of its ability to replicate a rich class of density-dependent mortalities.

Appendix A. Nomenclature

See Table A.2.

Appendix B. Fenichel's theorem

In the following, we restate a set of theorems described in [7] and used in Section 2. Consider a slow-fast system of the forms (F_ϵ) and (S_ϵ) with the lower dimensional problems (F_0) and (S_0). Assume
(H1) f, g are C^∞ on a set $U \times I$, where $U \subset \mathbb{R}^{k+l}$ is open and I is an open interval containing 0.
(H2) There exists a set $M_0 \subset \{(x, y)^t \in U | f(x, y, 0) = 0\}$, s.t.

Table A.2
Nomenclature.

$t \geq 0$	Slow time
$T \geq 0$	Fast time
$0 < \epsilon \ll 1$	Ratio between slow and fast time $t = \epsilon T$ (dimensionless)
$R(t)$	Numbers of fish recruited to the stock at time t
$S(t)$	Total number of fish in the spawning stock at time t
$i \in \{0, \ldots, n\}$	Index of age-class
$X_i(t)$	Number of fish in age-class $i \in \{1, \ldots, n\}$ at time t
$X_0(t)$	Number of prerecruits at time t
$\alpha > 0$	Rate of ageing; measured in numbers per individual and time unit t
$f_i > 0$	Fecundity of fish of age $i \in \{1, \ldots, n\}$ (dimensionless)
$l_i \geq 0$	Rate of eggs produced per fish of age i; measured in numbers per adult of age i and per time unit t
$p_i \geq 0$	Degree of density-dependence attributed to fish of age i; measured in numbers of deaths per prerecruit, per adult of age i and per time unit t
$p_0 \geq 0$	Degree of density-dependence attributed to juvenile competition; measured in numbers of deaths per prerecruit squared, and per time unit t
$m_i \geq 0$	Rate of density-independent mortality of age-class i; measured in numbers of deaths per individual and time unit t
$m_0 > 0$	Rate of density-independent mortality of prerecruits; measured in numbers of deaths per prerecruit and time unit t
$\gamma > 0$	Degree of density-dependence of the SRR (dimensionless)

1. M_0 is a compact manifold, possibly with boundary and
2. M_0 is normally hyperbolic relative to (F_0), i.e. for all $(x, y)^t \in M_0$, the matrix $Df_x(x, y, 0)$ has no eigenvalues with zero real part.

(H3) $M_0 = \{(x, y)^t \in \mathbb{R}^{k+l} | x = h^0(y),\ y \in K\}$ with a C^∞ function h^0 defined on K where

1. $K \subset \mathbb{R}^l$, compact and simply connected,
2. and the boundary of K is an $(l-1)$-dimensional C^∞ submanifold.

Let $\epsilon > 0$, but sufficiently small.

- By Theorems 1 and 2 in [7]: Then, there exists a function $h^\epsilon : K \to \mathbb{R}^k$ s.t. the graph $M_\epsilon = \{(x, y)^t \in \mathbb{R}^k | x = h^\epsilon(y),\ y \in K\}$ is locally invariant under (F_ϵ) and h^ϵ is C^r, $\forall r < \infty$, jointly in y and ϵ. The graph M_ϵ lies within $\mathcal{O}(\epsilon)$ of and is diffeomorphic to M_0.
- By Theorem 3 in [7]: There exist manifolds $\mathcal{W}^s(M_\epsilon)$ and $\mathcal{W}^u(M_\epsilon)$ of M_ϵ,
  1. which are $\mathcal{O}(\epsilon)$ close and diffeomorphic to their unperturbed counterparts $\mathcal{W}^s(M_0)$ and $\mathcal{W}^u(M_0)$, respectively.
  2. They are each locally invariant under (F_ϵ) and they are C^r, including in ϵ, $\forall r < \infty$.
- By Theorem 5 in [7]: Denote by V a neighbourhood of M_ϵ. There exist $\kappa_s > 0$ and $\alpha_s < 0$ so that $z_0 \in \mathcal{W}^s(M_\epsilon)$ and $z(t, z_0, 0) \in V$ for all $t \in [0, \bar{t}]$, with $\bar{t} > 0$ implies $d(z(\bar{t}, z_0, 0), M_\epsilon) \leq \kappa_s \exp(\alpha_s \bar{t})$ (with d the Euclidean distance).

The manifolds $\mathcal{W}^s(M_\epsilon)$ and $\mathcal{W}^u(M_\epsilon)$ of M_ϵ are called stable and unstable manifolds of M_ϵ, respectively.

References

[1] P. Auger, R. Bravo de la Parra, J.C. Poggiale, E. Sánchez, L. Sanz, Aggregation methods in dynamical systems and applications in population and community dynamics, Phys. Life Rev. 5 (2008) 79–105.
[2] R.J.H. Beverton, S.J. Holt, On the Dynamics of Exploited Fish Populations, Fishery Investigations Series II, Vol. XIX, HMSO, London, 1957.
[3] N. Charouki, N. Raïssi, P. Auger, R. Mchich, H. Atmani, A management oriented competitive model with two time scales: The case of sardine fishery along the Atlantic coast between Cantin Cape and Blanc Cape, Ecol. Model. 222 (2011) 1253–1261.

A

[4] R.B. Deriso, Harvesting strategies and parameter estimation for an age-structured model, Can. J. Fish. Aquat. Sci. 37 (1980) 268–282.
[5] N. Fenichel, Geometric singular perturbation theory for ordinary differential equations, J. Differential Equations 31 (1979) 53–98.
[6] R. Hilborn, C.J. Walters, Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty, Chapman and Hall, London, 1992.
[7] C.K.R.T. Jones, Geometric singular perturbation theory, in: R. Johnson (Ed.), Dynamical Systems, in: Lect. Notes in Math., vol. 1609, Springer, Berlin Heidelberg, 1995, pp. 44–118.
[8] B. Jonsson, N. Jonsson, Ecology of Atlantic Salmon and Brown Trout: Habitat As a Template for Life Histories, in: Fish & Fisheries Series, vol. 33, Springer, Netherlands, 2011. http://dx.doi.org/10.1007/978-94-007-1189-1.
[9] T.J. Kaper, An introduction to geometric methods and dynamical systems theory for singular perturbation problems, in: J. Cronin, R.E. O'Malley Jr (Eds.), Analyzing Multiscale Phenomena using Singular Perturbation Methods, in: Proc. Symp. Appl. Math., vol. 56, Am. Math. Soc., Providence, RI, 1999, pp. 85–131.
[10] C. Kuehn, Multiple Time Scale Dynamics, in: Applied Mathematical Sciences, vol. 191, Springer, Switzerland, 2015. http://dx.doi.org/10.1007/978-3-319-12316-5.
[11] J. Lappalainen, Effects of Environmental Factors, Especially Temperature, on the Population Dynamics of Pikeperch (Stizostedion Lucioperca L.), University of Helsinki, Finland, 2001 (Ph.D. thesis).
[12] J.M. Lee, Introduction to Smooth Manifolds, second ed., in: Graduate Texts in Mathematics, Springer, New York, 2013.
[13] J. Li, F. Brauer, Continuous-time age-structured models in population dynamics and epidemiology, in: F. Brauer, P. van den Driessche, J. Wu (Eds.), Mathematical Epidemiology, in: Mathematical Biosciences Subseries, vol. 1945, Springer, Berlin Heidelberg, 2008, pp. 205–227.
[14] C.L. Needle, Recruitment models: diagnosis and prognosis, Rev. Fish Biol. Fish. 11 (2002) 95–111. http://dx.doi.org/10.1023/A:1015208017674.
[15] W.E. Ricker, Stock and recruitment, J. Fish. Res. Board Can. 11 (1954) 559–623.
[16] B.J. Rothschild, Dynamics of Marine Fish Populations, Harvard University Press, USA, 1986.
[17] J. Schnute, A general theory for analysis of catch and effort data, Can. J. Fish. Aquat. Sci. 42 (1985) 414–429.
[18] L.F. Shampine, M.W. Reichelt, The MATLAB ODE suite, SIAM J. Sci. Comput. 18 (1) (1997) 1–22.
[19] P.E. Smith, J.K. Horne, D.C. Schneider, Spatial dynamics of anchovy, sardine, and hake pre-recruit stages in the California Current, ICES J. Mar. Sci. 58 (5) (2001) 1063–1071. http://dx.doi.org/10.1006/jmsc.2001.1092.
[20] S. Subbey, J.A. Devine, U. Schaarschmidt, R.D.M. Nash, Modeling and forecasting stock recruitment: current and future perspectives, ICES J. Mar. Sci. 71 (8) (2014) 2307–2322. http://dx.doi.org/10.1093/icesjms/fsu148.
[21] S. Touzeau, J.-L. Gouzé, On the stock-recruitment relationships in fish population models, Environ. Model. Assess. 3 (1998) 87–93.
[22] E.A. Trippel, Estimation of stock reproductive potential: history and challenges for Canadian Atlantic gadoid stock assessments, J. Northwest Atl. Fish. Sci. 25 (1999) 61–82.
[23] F. Verhulst, Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics, Springer, New York, 2005.
[24] S. Wiggins, Introduction to Applied Nonlinear Dynamical Systems and Chaos, second ed., Springer, New York, 2003.

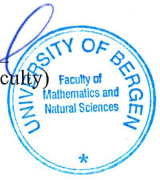
Errata for
Multiple time–scale dynamics of stage structured populations and derivative–free optimization

Ute Alexandra Schaarschmidt

Thesis for the degree philosophiae doctor (PhD)
at the University of Bergen

12.10.18 Schaarschmidt
(date and sign. of candidate)

Birte Godell
(date and sign. of faculty)



2

Page iii. Line break: Line 1 and 2 from below: the line break is now before "three"
Page vi and 73. Spacing in the title of "5.3 The parent-progeny relationship for the general multi-stage model" has been corrected.

Chapter 1.
Page 3. Misspelling: "matured enough" has been replaced by "mature enough".
Page 5. Punctuation: "For example, the average rate of deaths per individual and per time unit (called mortality rate), may ...." corrected to "For example, the average rate of deaths per individual and per time unit (called mortality rate) may ..."
Page 6. Punctuation: "[71, Sect. I.14.]" has been corrected to "[71, Sect. I.14]"
Page 7. Misspelling: "differentiable equation" has been replaced by "differential equation".

Chapter 2.
Page 15. Omission of "(t)": Equation (2.8) should read $N_i(t) = \alpha N_{i-1}(t) - \alpha N_i(t) - m_i N_i(t)$
Page 21. Line break: The line break in (2.22) has been removed.
Page 27. "consider a constrained optimization problem" has been replaced by "consider the constrained optimization problem".

Chapter 3.
Page 31. Misspelling: replaced "and under Hypotheses (h1)–(h4)" by "and under hypotheses (h1)–(h4)"
Page 33. Removed a superfluous parenthesis in Eq.s (3.17) and (3.19).
Page 34. "The function A_γ is defined on proper subsets" has been replaced by "The function A_γ is defined on a proper subset".
Page 36. Omission of "c": Equations (3.21) and (3.22) should read

$$\frac{d}{d\tilde{t}}\tilde{N}_0(\tilde{t}) = -\tilde{N}_0(\tilde{t}) + \frac{1}{\varepsilon c\alpha} A_\gamma(c\tilde{\mathbf{N}}(\tilde{t})), \text{ case } \gamma > 0. \tag{3.21}$$

$$\frac{d}{d\tilde{t}}\tilde{N}_0(\tilde{t}) = -\tilde{N}_0(\tilde{t}) + \frac{1}{\varepsilon c\alpha} A_\infty(c\tilde{\mathbf{N}}(\tilde{t})), \text{ case } \gamma \to \infty, \tag{3.22}$$

Page 39. Removed a superfluous parenthesis in Eq. (3.29).

Chapter 4.
Page 44. Layout: The numbering of (4.4) has been moved to the first line.
Page 46. Punctuation: ", and small-amplitude noise" has been replaced by "and small-amplitude noise".
Page 47. Line break: Line 30 and 31: the line break is now before "objective".
Page 49: "of the nonlinear equation" corrected to "of a nonlinear equation"
Page 51: Misspelling: "largest jumps in the value of the objective functions" has been corrected to "largest jumps in the value of the objective function".

Chapter 5.
Page 75. Omission: "$N_t = N$" has been corrected to "$N_t = N_0$"
Page 76. Line break: corrected to "(C4)(a), (b), and (c)."
Page 77. Line break: in "recruitment is a function of S_{t-m-1} is now before "of".
Page 78. Indention before "(C6)" has been removed.
Page 79. Misspelling: "stage 2 and 3" has been corrected to "stages 2 and 3".

3

Chapter 7.
Page 107. "H. A. Le Thi, T. Pham Dinh, and N. T. Nguyen" has been corrected to "Le Thi, H.A., Pham Dinh, T., Nguyen, N.T."
Page 107. "pp. 391–402" has been corrected to "391–402"
Page 107. "*ICES J. Mar. Sci.*, 71, 8 (2014), p. 2307–2322, 2014" has been corrected to "*ICES J. Mar. Sci. 71,* 8 (2014), 2307–2322"

Paper C.
The layout is changed.

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 978-82-308-3657-6