Dissertation for the degree of doctor of philosophy

Statistical models of marine multispecies ecosystems

Bjarki Þór Elvarsson

UNIVERSITY OF ICELAND

School of Engineering and Natural Sciences
Faculty of Physical Sciences
Reykjavík, May 2015

Dissertation submitted in partial fulfilment of a *Philosophiae Doctor* degree in Statistics

Doctoral committee
Prof. Gunnar Stefánsson, advisor
University of Iceland

Gísli A. Víkingsson
Head of division of Marine Mammal Research, Marine Research Institute

Dr. Eva Plaganyi–Loyd
Senior research scientist, CSIRO

Opponents

Dr. Carmen Fernandez
ACOM vice–chair, International Council for the Exploration of the Sea

Dr. Tore Schweder
Professor emeritus, University of Oslo

Abstract

Models of marine ecosystems are widely used to investigate possible structures of the oceans' vast and often unobservable ecosystems. Statistical models of marine ecosystems attempt to estimate substantial links and processes of a part of the ecosystem that is of interest through comparisons to data. In this thesis, properties of these statistical models are investigated within Gadget, a multi–area, multi–species, multi–fleet modeling framework designed for comparisons to minimally aggregated data.

The analyses presented in this thesis can broadly be split into two categories; 1) methodology for assessing goodness of fit, estimation of uncertainty and model selection, and 2) an application in the form of investigation into the link between minke whales (*Balaenoptera acutorostrata*) and cod (*Gadus morhua*) on the continental shelf around Iceland. These analyses are primarily reported in the five papers presented in Part II.

Paper I provides an analysis of the feeding ecology of minke whales in Icelandic waters. The main findings are that sandeels (*Ammodytidae sp*) appear to be a major component in the diet of minke whales in the southern areas, and to a lesser extent the northern. A greater part of the diet, than previously considered, are composed of gadoids (mainly cod and haddock (*Melanogrammus aeglefinus*)). Dramatic changes have been observed in the diet and abundance of minke whales, concurrent with an assumed collapse of the sandeel stock. The results are presented with uncertainty estimated using a bootstrap approach treating an individual whale as a sampling unit.

Paper II expands on the ideas underlying the uncertainty estimates in paper I. A bootstrap approach applied to highly disparate datasets, such as those used by Gadget, is developed and demonstrated on cod in Icelandic waters. The method is contrasted with a more commonly used normal approximation where the covariance matrix is derived from the inverse Hessian matrix of the negative log likelihood function at the optimum. This method appears to perform

favorably to the Hessian–based method for this class of models.

Paper III presents the development of Rgadget, an R-package containing a set of useful utitilities when developing models in Gadget. RGadget contains a number of methods which include tools to build model skeletons, iterative reweighting of likelihood components, assessment of fit and forward projections. In addition, an independent implementation of some of the key processes in Gadget is included. This package allows for the testing of various assumptions used by fisheries stock assessment models as well as to play with various parts of the model to allow for better understanding of age–length based assessment models.

Paper IV is an example application of Rgadget's simulation procedure to the question of stock structure of North–Atlantic fin whales (*Balaenoptera physalus*). The paper presents a power analysis of a proposed genetic tagging experiment contrasting two stock structure hypotheses based on information on genetically determined close relatives. The results indicate that in the case of small populations, such as fin whales in the North Atlantic (around 50 000 individuals), a tagging experiment can be improved considerably by augmenting it with information on genetically determined close relatives.

Paper V combines the data from the scientific survey of feeding ecology of minke whales described in paper I and a model of cod in Icelandic waters studied in paper II. The paper illustrates how different groupings in the iterative reweighting can affect the fit, and studies the utility and effects of adding information on consumption of minke whales to the fit of the cod model. The results suggest that fit is not altered considerably by linking the two species in a single model. However when predicting the stock status forward, substantial differences are observed between the class of models that model minke whale predations and those that ignore ignore multispecies interactions. These differences are driven by substantial changes in the abundance of minke whales.

Overall, this thesis has advanced the development of statistical multispecies models by developing rigorous approaches and routines for assessing goodness of fit, estimation of uncertainty and informing model selection. Moreover, the case study application to minke whales and their prey has advanced understanding of the complex interactions, as well as the effect of including or ignoring these interactions on the outcomes of model predictions.

Ágrip

Líkön af vistkerfum sjávar eru víða notuð til þess að rannsaka hugsanleg vensl milli einstakra vistkerfisþátta sem erfitt getur reynst að mæla. Með tölfræðilegum líkönum er reynt að rannsaka og bera saman við gögn, helstu vensl og ferla þess hluta vistkerfisins sem rannsókn beinist að. Í þessari ritgerð eru eiginleikar þessara tölfræðilegu líkana rannsakaðir innan sérsmíðaðs líkanaumhverfis, Gadget. Gadget auðveldar smíði fjölsvæða-, fjölstofna-, fjölflotalíkana og hefur Gadget sérstaklega verið hannaður með það að markmiði að bera saman úttak líkana við lítið unnin gögn.

Efni ritgerðarinnar má í stórum dráttum skipta í tvennt: 1) Þróun á aðferðafræði til þess að meta mátgæði, fá óvissumat og aðstoða við val á líkönum og 2) beitingu aðferðanna við rannsókn á sambandi hrefnu (*Balaenoptera acutorostrata*) og þorsks (*Gadus morhua*) á hafsvæðinu í kringum Ísland. Þessum rannsóknum er lýst í fimm greinum sem eru uppistaðan í hluta II.

Í grein I er gerð grein fyrir niðurstöðum greiningar á fæðuvistfræði hrefna á Íslandsmiðum. Helstu niðurstöður gefa það til kynna að sandsíli (*Ammodytidae sp*) virðast vera mikilvægur þáttur í fæðu hrefna suður af landinu, og einnig fyrir norðan það en þar þó í minna mæli. Þorskfiskar (að stærstum hluta ýsa (*Melanogrammus aeglefinus*) og þorskur) voru stærra hlutfall fæðu hrefnunnar en áður var talið. Töluverðar breytingar, sem verða samtímis hruni í sandsílastofninum, má merkja í fæðuvali og fjölda hrefna. Niðurstöður rannsóknanna eru settar fram með óvissumati byggt á endurvalsaðferð þar sem sérhver hvalur er meðhöndlaður sem úrtaksstærð.

Í grein II eru víkkaðar út þær hugmyndir sem grundvalla óvissumatið í grein I. Endurvalsaðferð fyrir eðlisólík gagnasett, eins og þeim er nýtt eru af líkönum smíðuð með Gadget, er þar þróuð og beitt á líkan fyrir viðgang þorsks á Íslandsmiðum. Aðferðin er þvínæst borin saman við hefðbundnar normal-nálganir á óvissu þar sem samdreifnifylkið er nálgað með andhverfu Hessian fylkisins af neikvæðum logra sennileikafallsins fengið við lággildi. Niðurstöðurnar gefa það

til kynna að endurvalsaðferðin henti betur en Hessian–nálganir við mat á óvissu fyrir þennan flokk líkana.

Grein III lýsir þróun á RGadget, R pakka sem inniheldur safn tóla sem nota má við þróun líkana með Gadget. RGadget býður upp á fjölda aðferða m.a. föll til að smíða stoðgrind fyrir líkön, endurvigtanir á þáttum sennileikafalls Gadget–líkans, mat á mátgæðum og framreikningum. Að auki inniheldur pakkinn óháða útfærslu á helstu ferlum Gadget í sérstökum hermi. RGadget auðveldar því prófanir á ýmsum forsendum að baki stofnmatsaðferða í fiskifræði auk þess að veita innsýn inn í helstu aðtriði aldurs-lengdarháðra stofnmatsaðferða.

Í grein IV er því lýst hvernig herminum í RGadget er beitt til stuðnings rannsókna á stofnsamsetningu langreyða (*Balaenoptera physalus*) í Norður – Atlantshafi. Greinin lýsir prófstyrksreikningum fyrir hugsanlegar erfðamerkingatilraunir sem ætlaðar eru til samanburðar tveggja tilgátna um stofnsamsetningu á grundvelli erfðafræðilegrar sifjagreiningar. Það að rannsaka sifjar, jafnframt því að beita hefðbundnum merkingaraðferðum, styrkir umtalsvert niðurstöður rannsókna á litlum stofneiningum eins og langreyðum í Norður-Atlantshafi, sem telja í kringum 50.000 dýr.

Í grein V eru teknar saman helstu niðurstöður vísindarannsókna á fæðusvistfræði hrefnu á Íslandsmiðum, sem lýst er í grein I, og líkans af þorski við Ísland, sem lýst er í grein II. Í greininni er því lýst hvaða áhrif það hefur á mátgæði þegar þáttum sennileikafallsins er hópað saman á mismunandi hátt við beitingu endurvigtunarreikniritsins. Á svipaðan máta er viðbótaráhrifum afráns hrefna á þorsk lýst í samhengi við mátgæði. Niðurstöðurnar gefa til kynna að mátgæði breytast ekki umtalsvert þótt afrán hrefna sé tekið með í reikninginn. En þegar kemur að því að spá fyrir um viðgang þorsks sýnir líkanið umtalsverðan mun eftir því hvort gert er ráð fyrir afráni eða það hunsað. Þessi munur skýrist einna helst af umtalsverðum mun á fjölda hrefna á Íslandsmiðum milli ára.

Í ritgerðinni er sýnt fram á mikilvægi tölfræðilegra fjölstofnalíkana í þróun aðferða og aðgerða til þess að meta og rannsaka samband mismunandi þátta innan vistkerfa, s.s. til að meta mátgæði og óvissu og aðstoða við val líkana til rannsókna. Þær rannsóknir sem lýst er í ritgerðinni varpa svo skýrara ljósi á flókið samband hrefna og þeirra fiskitegunda sem þær lifa á, auk þess sem þær lýsa þeim áhrifum sem það hefur að bæta við ellegar að hunsa þessi fæðutengsl þegar kemur að því að skoða spágetu fjölstofnalíkans.

Contents

List of Figures

List of Tables

List of Publications

This thesis is based on the following papers, which will be referred to in the text by their Roman numbers.

I. GA Vikingsson, BÞ Elvarsson, D Olafsdottir, V Chosson, J Sigurjonsson and Anton Galan. Recent changes in the diet composition of common minke whales (Balaenoptera acutorostrata) in Icelandic waters A consequence of climate change? *Marine Biology Research*, 10(2):138152, 2014.

II. BÞ Elvarsson, L Taylor, V Trenkel, V Kupca and G Stefansson. A bootstrap method for estimating bias and variance in statistical fisheries modelling frameworks using highly disparate data sets. *African Journal of Marine Science*, 36(1):99110, 2014.

III. BÞ Elvarsson. RGadget: a R–package for development, testing and analysis of Gadget models.

IV. BÞ Elvarsson. Evaluating stock structure hypotheses using genetically determined close relatives: A simulation study on North Atlantic fin whales. *ICES Journal of Marine Science: Journal du Conseil*, 72(2):661669, 2015.

V. BÞ Elvarsson, GA Víkingsson and G Stefánsson. Adding data on species interactions: A case study on cod (*Gadus morhua*) and minke whales (*Balaenoptera acutorostrata*) interaction in Icelandic waters.

Acknowledgements

First and foremost, I would like to thank my principal supervisor, professor Gunnar Stefánsson, for his guidance and support throughout this study. I would also like to thank the members of my doctoral committee, Gísli Víkingsson and Éva Plagányi-Lloyd for their professional support and willingness to assist me in any way. I would further like to thank my other co-authors for their essential contribution.

The Marine Research Institute has supported this study in many ways which I gratefully acknowledge. I would like to thank Jóhann Sigurjósson, the director of the Marine Research Institute, and Björn Ævar Steinarsson, head of the fishery advisory section, for providing excellent working facilities. I am also truly indebted to my coworkers at the Marine Research Institute. Especially I would like to thank Muggur for all his friendship and support, Höski and Einar for the monkey business, and Þorvaldur, Chris and Valerie for all the whaling.

I would also like to thank my fellow graduate students. This goes in particular to Sigrún and Anna Helga for all the R shenanigans and Helga, Óli, Warsha, Erla, Chrispine and Chris for for all the discussions, advice and comradery.

Last but certainly not least I would like to thank my family for all their support, love and kindness throughout the years. My parents and in-laws for helping me with the kids whenever I was busy. My wife, Linda María, for her incredible patience and endurance for which this study would have been impossible without.

Ég þakka börnunum mínum þremur, Ríkarði Flóka, Guðbjörgu Elísu og Þorsteini Hilmari, fyrir alla þolinmæðina, allt knúsið og endalausa uppsprettu af brosum.

So long and thanks for all the fish!

Part I

Thesis

1

Background

Fishing has been a part of human activities for thousands of years, and in some cultures fish and other marine animals are a vital part of the diet. The consumption of fish has an effect on many facets of human interactions, including commerce. One of the longest lasting trade histories is the trade of dry cod from the Lofoten area to southern parts of Europe, which started during the Viking period and has lasted to this day. Fishing techniques have, during the course of time, evolved to be more efficient. In the late 1800s the effects of overfishing had become apparent in various regions. In the North Sea this spawned a debate as to whether this was caused by human activities, which subsequently lead to an analysis that showed that indeed these effects were due to overfishing (Smith 2002). One of the strongest cases for effects of fishing came during the 1st world war, often referred to as the *Great fishing experiment*, when supporting the hypothesis of a nontrivial effect of fishing came the near cessation of fishing allowed fish populations to recuperate. The Great fishing experiment proved that fishing did have an effect and that fish populations could be rebuilt.

These considerations led to the development of models of marine ecosystems, which have been commonly used to determine satisfactory utilisation of a particular resource. The earliest examples of models trace their conceptual roots to simultaneous developements in the first few decades of the 1900s when many of the fundamental attributes of stock assessment were described. These include the works of Lotka and Volterra (Volterra 1928, Lotka 1925) on interac-

tions between species and Baranov's work linking fishing mortality, growth rates and fishing yields (Baranov 1918). Later Beverton and Holt (1957) provided a rigorous analysis combining growth rates, fishing intensity and fleet selection.

Since then significant improvements have been made on various fronts (Smith 2002). With reliable age readings it became possible to distinguish between fast and slow growing species in a rigorous manner. This paved the way for age–based stock assessment such as virtual population analysis (VPA, introduced by Gulland 1965). VPA as originally conceived was not a statistical modelling technique but underlies many popular stock assessment tools such as the popular eXtended Survivor Analysis (Shepherd 1999) and the statistical framework ADAPT (Gavaris 1988a). Another category of a stock assessment model are the variants of the surplus production model (e.g. Graham 1935, Pella and Tomlinson 1969). Their popularity stems from their relatively simple dynamics, which are therefore straightforward to interpret, and modest requirements on available data.

In the more recent years considerable improvements and investment in various data collections schemes have been made. These include long time series of standardised surveys (e.g. Pálsson et al. 1997, Borchers et al. 2009). Technological advances such as in tagging equipment and procedures (e.g. Walsh and Morgan 2004) and genetic structure (Bérubé et al. 2002, Pampoulie et al. 2006, Skaug 2001) have given further insights into population dynamics of exploited marine resources. With these longer dataseries containing detailed information on the marine resource comes the demand for more sophisticated models and increased computing power.

These developments aside, consumers and legislative bodies have put emphasis on sustainable utilisation of marine resources (UN 2002, Parliament 2008, Anon 2002). This has in turn been translated to further demands on reliability and robustness of harvesting strategies. Models of marine ecosystems aim to provide the tools for the inferences required in this regard. A number of approaches have been developed, from single species stock assessment models, with possible extension to allow for species interactions (eg. Magnússon 1995) to more all-encompassing models that aim to fully describe an ecosystem, such as described by Link et al. (2010). Needless to say, these approaches vary in scope and inferences possible.

1.1 Approaches to modelling marine multi–species ecosystems

Ecosystem considerations appear to be an obvious, and necessary part, of the management of marine resourcea. Several international organizations have stressed the importance of studying ecosystems interactions in terms of management (eg. Sinclair et al. 2002, NAMMCO 1997, IWC 2002). Optimum fisheries management based on ecosystem consideration is however an elusive target. Even if the model sufficiently describes the target ecosystem, Voss et al. (2014) illustrated that the definition of the this optimum is dependent on the manager's view of the resource. Although the rationale for ecosystem based approach to fisheries management is strong its adoption has been slow, which Möllmann et al. (2013) suggests is due to lack of strategy. However the effects of species interactions on the management of a commercially exploited stock is a fairly complicated and data intensive excercise, as it requires data not directly observed by fisheries. Dickey-Collas et al. (2013) note therefore that the state of knowledge is not sufficient to embark on ecosystem engineering, although in theory this would have desirable effects on the status of key resources, as such an undertaking may have undesirable and unforseen consequences.

The choice of a modelling approach used to assimilate biological information such as those obtained from scientific surveys also needs some consideration. There are a number of approaches to ecosystems models. According to Plagányi (2007) they broadly fall broadly into four categories:

1. Whole ecosystem approaches, where all trophic levels of the ecosystem are modelled (e.g. Christensen and Walters 2004).

2. Dynamic systems models, where where the whole ecosystem and possibly its effects to coastal communities is modelled (e.g. Link et al. 2010).

3. Extensions of single–species assesment models (e.g. Livingston and Methot 1998).

4. Minimum realistic models, where only a few but significant interactions are included. Examples of cases are models built using the Gadget framework (Begley and Howell 2004).

Obviously these approaches differ substantiantly both in scope and objectives. Minimimum realistic models and singles species extensions are, as the names suggest, more focussed on answering key questions and their parameters

are estimated statistically. Modelling approaches, like those using Atlantis (Fulton et al. 2005) or Ecopath with Ecosim (Christensen and Walters 2004), often try to include a larger part of the ecosystem where it becomes harder to estimate the model parameters through comparisons with real data. These grand models tend to be more useful when answering strategic questions, such as discussed by Fulton et al. (2008), rather than more tactical issues such as direct advice on next years quota.

1.1.1   Gadget – a toolbox for fisheries stock assessment

Modelling approaches such as those allowed by Atlantis have, although interesting, tended to have quite obvious drawbacks in terms of comparisons to real data. Estimates of process parameters are far from trivial in this setting, both in terms of available data and more importantly computation time. A fullblown Atlantis model will generally take a substantial time for a single simulation and as such not obvious how one would obtain an "optimal" fit in a statistical sense. That said, these types of model can provide valuable insight into how an ecosystems would behave. In particular one could study certain performance metrics, such as ecosystem indicators (as described in Fulton et al. 2005), and various stock assessment models.

It has been observed that in various settings single species models of a marine resource fail to explain model deviations. To answer these questions extensions to the normal single species have been developed (as discussed in Magnússon 1995, and references therein). These extensions attempt to ecompass significant processes that affect the species of interest. A special type of these approaches are the mimimally realistic models mentioned above (also referred to as MICE models as discussed in Plagányi et al. 2014).

A special modelling framework, Gadget (**G**lobally **a**pplicable, area–**d**isaggregated **e**cosystem **t**oolbox, described in Begley 2005), has been developed to aid in the study of various statistical processes involved in multi–species assessments. The development of Gadget conceptually dates back to MULTSPEC (Bogstad et al. 1997), which described a multi–species model for fish and marine mammals in the Barents sea. Many of the underlying concepts and most of the codebase that forms the Gadget framework was date back to Bormicon (Stefánsson and Pálsson 1997a, Stefansson and Palsson 1998), which was developed as a continuation of a multi-species sampling program (described in Gislason and Ástþórsson 1997), which has been under constant developement and iterations. Notable iterations include the Fleksibest framework (Frøysa et al. 2002), a fork of Bormicon which was merged back with Bormicon to form Gadget as

part of the dst^2 project (dst^2 2004).

The dst^2 project was multi-objective project funded by the European Union and began in 1999 (dst^2 2004). The basic idea was to "develop fisheries science as far as information will support" the development of statistically sound models of marine ecological processes, "but strictly no further". Four main objectives were the following:

1. Collect relevant data and provide objective means to analyse them. This included the development of Gadget and a specialised data warehouse.

2. Validate present stock assessment.

3. Evaluate the appropriateness of complicated models when providing management advice.

4. Evaluate and validate reference points used in management.

A substantial amount of analysis has subsequently been made on Cod in Icelandic waters using the Gadget modeling framework. Notably Taylor et al. (2007) which describes a thorough analysis of the available data and potential weaknesses. In it a single species single area model is considered. A continuation of this work in a multi-stock multi-species multi-area setting are described in Taylor (2011). There it is seen that substantial improvements in the model fit are possible by considering a north – south division in the cod stock.

These developments have not been restricted to single–species models. A number of species have been analysed in a multi–species setting using Gadget. Notable implementaions include a model of the Barents sea of cod, capelin, herring and minke whales (Lindstrøm et al. 2009) and subsquently Howell and Bogstad (2010) where FLR (Kell et al. 2007) is used in connection with Gadget in a harvest control simulation.

1.2 Sources of variation

Building a statistical stock assessment model is in general a multi–step procedure. The process, which is illustrated in Figure 3.1, involves collecting data from various sources, building a model or models that emulate the relevant parts of the ecosystem, and projecting the state of the ecosystem. In all of these steps considerable and unique uncertainties can be assigned. According to Francis and Shotton (1997) sources of uncertainty can be split into six categories:

• Observation uncertainty, error attributed to data sampling.

- Process uncertainty, natural variation in processes governing the ecosystem.

- Model uncertainty, that is uncertainty related to the correct form of the model.

- Error structure uncertainty, a special type of model error, which relates to the error arising when attributing sources of error to a wrong process.

- Estimation uncertainty is the uncertainty regarding the parameter estimation process.

- Implementation uncertainty, i.e. how well management advice is followed, which is relevant in the forecasting of the resources' status.

Naturally these are not independent dimensions of variation e.g. estimation is affected by variability in the steps leading to it. When managing a resource it is often beneficial to know the potential variation and uncertainty of the model's prediction of desirable yield. Variance estimates for fisheries stock assessments are typically derived either using a normal approximation with the hessian matrix interverted at the minimum as the variance–covariance matrix or using MCMC sampling within the Bayesian setting. When using Hessian based approaches several conditions need to be satisfied for statistical inference, e.g. confidence statements to hold in the finite-sample case. First, the model needs to be correct. Second, variance assumptions i.e. homoscedasticity and knowledge of the ratios of variances in individual data sets, need to be appropriate.

Methods of estimating variances in fish stock assessment models have been discussed and evaluated by many authors including Gavaris et al. (2000), Gavaris and Ianelli (2001), Magnusson et al. (2012) and Patterson et al. (2001). When the distributional properties of the data are not well understood or the models are incorrect, Hessian–based approaches have been seen to fail in several examples in fishery science (Patterson et al. 2001). Although this may seem to contradict the theoretical statements, the assumptions e.g. in Jennrich (1969) include independence of observations, a unique minimum, identically distributed errors and of course the results are only asymptotic. Any of these assumptions may fail. It follows that for problems in fishery science one cannot assume a priori that a Hessian–based method will give reasonable results. For example, disregarding correlation structure when present has been found to potentially lead to incorrect conclusions in single–species assessments, sometimes with serious consequences (Myers and Cadigan 1995). Similarly, multimodal likelihood

functions have been seen in real applications (Richards 1991) and typically correspond to incorrect model assumptions that are not detected with traditional analysis (Stefansson 2003) but may potentially be detected if histograms of bootstrap parameter estimates also become multimodal (see example in Hannesson et al. 2009).

1.2.1 Identifying the resource

The definition of a managment unit is a form of model uncertainty. Stocks have often been defined by management boundaries, set according to the distribution of the key species of commercial interest (Halliday and Pinhorn 1990). This could lead to separate advice given to the same stock depending on the management area, for obvious reasons. Therefore, when dealing with fish stocks, numerous methods have been employed to test the accuracy of stock definition and borders. Tagging, both mark recapture (Rayner 1940, Chenuil et al. 2000, Laurenson et al. 2005, Peakall et al. 2006, Hannesson et al. 2008) or satellite tracking (Mate et al. 2007, Víkingsson and Heide-Jörgensen 2014, Horton et al. 2011, Matthews et al. 2011), is commonly used to identify individual and stock movement between (and within) areas.

Mark–recapture experiments are however not always suited to track individual movements. For instance small fish may not survive the marking, marks do not survive moulting, are shed for other reasons or larvae may drift between populations. Other factors, not directly connected to the biology of the species can make traditional mark–recapture analysis impossible or ill suited to answer questions regarding stock boundaries. An example of such a case are tag – recapture experiments on baleen whales where a moratorium has been in place since 1986 preventing recapture. Satellite tracking methods are widely used in studies of animal population structure (eg. in Block et al. 2005, Claridge et al. 2009, Combreau et al. 2011, Kennedy et al. 2013, Mauritzen et al. 2002, Seegar et al. 1996). However, this method is difficult to apply to large cetaceans and despite decades of considerable effort in technical development and trials satellite tracking of cetaceans rarely last longer than a few months (eg. Mate et al. 2007, Silva et al. 2013, Víkingsson and Heide-Jörgensen 2014). Therefore satellite monitoring is currently not suitable to monitor annual migration patterns.

Even when tagging is possible it is not always sufficient to detect separate breeding populations. Breeding stocks can overlap on feeding grounds, as suggested by an analysis of otolith classification or genetic structure (Reynolds and Templin 2004, Wennevik et al. 2008, Jónsdóttir et al. 2007). These results indi-

cate that, if neglected, managing two (or more) separate breeding populations could have adverse effects, such as an overexploitation of one breeding population, without being detected by conventional assessment methods (Punt 2003). Genetic differences have been used to suggest separate breeding populations of marine mammals (Andersen et al. 1997, Bérubé et al. 2002, Parsons et al. 2006, Fontaine et al. 2007, Pampoulie et al. 2008).

Despite considerable efforts through decades, traditional population genetic studies have in many cases failed to give unequivocal answers to important questions concerning cetacean stock structure (Donovan 1991). A major obstacle for interpretations of these studies has been the fact that for most baleen whales the breeding grounds are unknown and sampling has thus been restricted to the summer feeding grounds. Furthermore, large baleen whales, such as the North Atlantic Fin whales, do not often exhibit sufficient genetic variability to detect separate breeding populations (Bérubé et al. 1998, Pampoulie et al. 2008). The possibility of two or more breeding populations can, however, not be ruled out as the hypothesized split occurred relatively recently and the stocks have not had time to differentiate to an extent that is detectable (Pampoulie et al. 2008).

In the absense of detectable genetic structures genetic tagging (as described by Palsbøll 1999) or other genetic methods such as information on close relatives could be used to answer questions related to stock structure (Skaug 2001, Palsbøll et al. 2010, Nielsen et al. 2001), such as the effect of larval drift between different breeding stocks (Planes et al. 2009). A simulation experiment of its potential application in management of marine mammals can be found in Økland et al. (2010), where managment units are defined for geographically segregated stocks using genetically determined close relatives. The authors, however, note that the methods presented therein would hardly be applicable to stocks that overlap on feeding grounds whilst separate on breeding grounds, as is common for baleen whales, due to low discriminatory power.

1.3   The Icelandic continental shelf area

The Icelandic continental shelf is positioned at a dynamic frontier of the Atlantic current and the artic circle, where an influx of warmer water meets the colder Artic waters (Astthorsson et al. 2007). As such it features a complicated ecosystem and, as in many other areas in the North Atlantic, gadoid species, especially cod, are the main focus of the fishery. Species interaction in Icelandic waters has received considerable attention (e.g. Gislason and Ástþórsson 1997, Stefánsson and Pálsson 1997a).

Historically the location of this dynamic front has been seen to vary considerably. On a longer time scale, during the small ice age, the frontier shifted further south which in turn shifted the distribution of several fish species further south. At that time contemporary descriptions indicated that little or no fish were caught by local fishermen (Jónsson 1994).

On a smaller timescale small scale variations in the spatial distribution of commercially important species have been observed. The great herring collapse of the 1960s is an extreme case (potentially due to overharvesting or changes in the ecosystem). Capelin migrations have been seen to vary according to temperature (Pálsson et al. 2014) and a north south migration has been observed in many species. Concomitant with these changes, fluctuations have been observed in the average sea temperatures.

The fluctuations in temperature in the North Sea have been linked to recruitment of sandeel (Arnott and Ruxton 2002). Similarly, albeit indirect, observations have been made in the Icelandic sandeel stock in the southern part of the continental shelf (Víkingsson et al. 2015). Studying the abundance of Vestmann Island puffin populations, through tagging series and direct export figures, further strengthens this relationship (Hansen pers. comm.).

1.3.1 Minke whales

The common minke whale (*Balaenoptera acutorostrata*) is the most abundant whale species in Icelandic continental shelf waters (Borchers et al. 2009, Pike et al. 2011a). Considerable uncertainty as to the role of minke whales in the Icelandic continental shelf ecosystem. Previous studies have indicated that cetaceans, and in particular minke whales, play an important role in the marine ecosystem by consuming several times the total Icelandic fishery landings (Sigurjónsson and Víkingsson 1997). There the annual consumption by minke whales alone was estimated as around 2 million tonns in the 1990's. Initial attempts to include three species of cetaceans, namely minke, fin and humpback whales, in a multispecies model indicated that their effect on the development of the stocks of cod and capelin may be considerable (Stefánsson et al. 1997). There was, however considerable uncertainty associated with this estimate. One of the greatest sources of uncertainty regarding the effects on the cod stock was associated with the very limited knowledge of the diet composition of minke whales in Icelandic waters. It was therefore of prime importance for further development of multispecies modelling in Icelandic waters to obtain data on the diet of minke whales and investigate multi-species interactions in more detail, in particular those between minke whales and the cod stock. The main objective

of the research programme on common minke whales, conducted in the years
between 2003 – 2007, was to address these questions as a pilot study using
various methods (see MRI 2003, for further details).

   Major changes in the abundance of minke whales have been observed on
the Icelandic continental shelf area. Areal surveys around Iceland indicate a
dramatic downward shift in the abundance since 2001. Several hypothesis have
been discussed as to what may have caused this shift. Although mass die offs can
not be ruled out, the NAMMCO (2009) concluded that the most likely scenarios
were shifts in spatial distribution of the common minke whale. Unfortunately
little or no data has been collected on the abundance after this hypothesised
shift was confirmed in 2007, and the few data available found no conclusive
evidence pertaining to the question at hand.

2

Aim

The major objective of this PhD project was to study the properties of statistical models of marine multi–species ecosystems, in particular models built using the Gadget framework. The main emphasis of the thesis is on estimating uncertainty and to investigate the types of difficulties when building multi–species models of the marine ecosystem in Icelandic waters. More specifically the aims were to:

- Analyse available data and methods for statistical stock assessments. In particular investigate the link between cod and minke whales.

- Develop methods to estimate uncertainty in fisheries stock assessment models.

- Develop (if required) additional functionality and/or methodologies to account for marine mammals in Gadget.

- Develop a multi–species model for minke whales and cod interactions in Icelandic waters.

3

Material and methods

This chapter describes the general setting in which the analyses of this thesis were conducted. It provides a overview of the methods used and the relevant sections and chapters of this thesis referenced. This chapter begins by introducing the main tool used in this thesis, the statistical toolbox Gadget designed to build models of marine multi–species ecosystems. Subsequently methods and tools dealing with disparate data used in stock assessment are described. The chapter then closes with a brief description of the analyses performed in this thesis. This description is ordered by the project aims, although most papers address more than one point simultaneously.

3.1 General context of Gadget models

3.1.1 Description of Gadget

Gadget is essentially three things, an ecosystem simulator, a likelihood function that takes the output from the ecosystem simulator and compares to data, and a function mininimizer. Gadget's simulation module allows for a fairly configurable ecosystem simulation. Its fundamental unit, a **stock** (or more accurately substock), represents a group of individuals that is homogenous with respect to various processes. These processes include growth, predation (including commercial fisheries) and migration. In this setup different stages of the life history
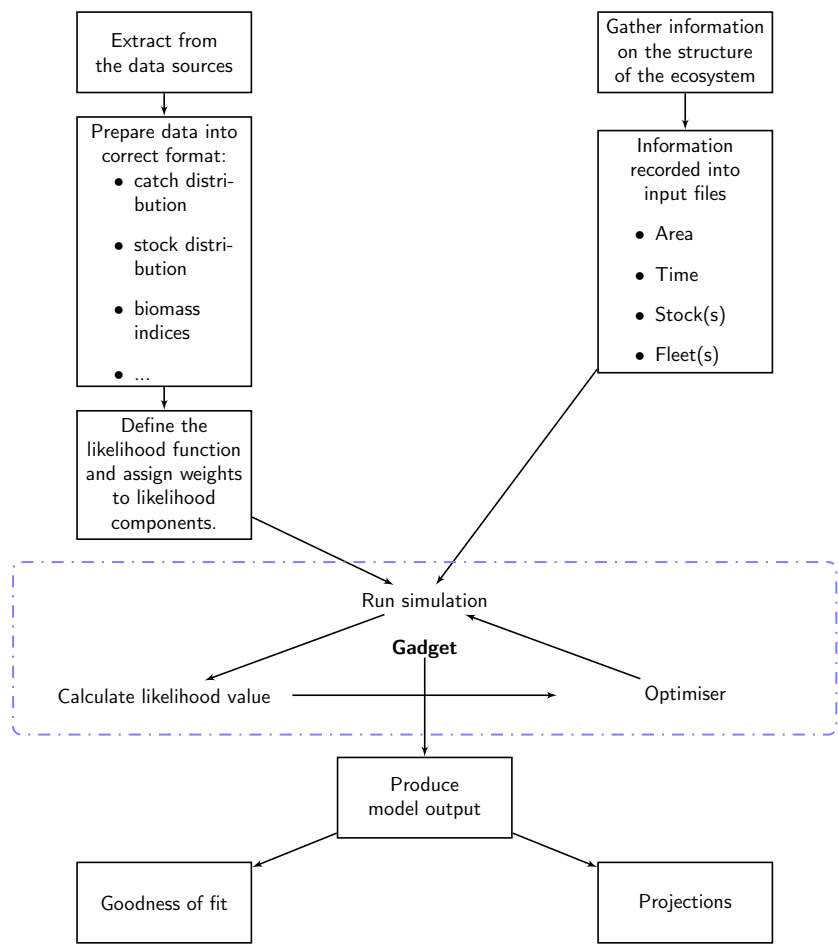
Figure 3.1: An illustration of a typical stock assessment process and the types of uncertainty associated with each step within Gadget.

of a particular species would be represented as separate stocks and individuals "moved" between stocks when required. The simulation takes place in a set number of **years** or **observations** and **time-steps** within an observation. The time-steps within the observations allow for the emulation of the annual cycles of the ecosystem, such as recruitment and stock migrations.

The stock unit within Gadget is simply a representation of the total number of individuals in a certain age and length group within a certain area. The stocks live in an area, or areas, where they optionally migrate to and from. In this setup processes such as fleet harvest or recruitment can be restricted to take place only in certain areas. Harvesting of the substocks is defined through fleets that fish according to effort and (length–based) selection functions.

Gadget's likelihood module processes the output from the ecosystem simulation based on aggregate dimensions. Within the likelihood module a number of datasets can be compared to the model output. In addition to a suite of functions designed to work with different types of survey indices, length distributions, tagging data, age and length distribution and maturity data, to name a few, can be contrasted to the model output. Each data set is included at its own aggregation level, with missing data handled in a robust manner.

In contrast with Gadget, age based or stock production type stock assessments require data in a fairly processed form. For instance when using VPA one requires the total catch in numbers of individuals by age. However, apart from catches of fin whales in the North Atlantic (IWC 2015), one rarely has all catches by numbers at age. Therefore the age distribution of catches needs to be approximated using some combination of age readings, length distributions, total catches in tons and weight at age (as noted in Hirst et al. 2005). Essentially using VPA requires a two step modelling process.

Gadget's function minimizer, based on the negative log–likelihood, varies the model parameters, runs a full simulation, and calculates a new output. This process is repeated until a minimum is obtained. The model has three alternative optimising algorithms linked to it, a wide area search `simulated annealing` (Corana et al. 1987), a local search `Hooke and Jeeves algorithm` (Hooke and Jeeves 1961) and finally one based on the Broyden-Fletcher-Goldfarb-Shanno algorithm hereafter termed `BFGS`.

The simulated annealing and Hooke-Jeeves algorithms are not gradient based, and there is therefore no requirement fir the likelihood surface to be smooth. Consequently neither of the two algorithms returns estimates of the Hessian matrix. Simulated annealing is more robust than Hooke and Jeeves and can find a global optima where there are multiple optima but needs about 2-3 times the order of magnitude number of iterations than the Hooke and Jeeves algorithm.

BFGS is a quasi-Newton optimisation method that uses information about the gradient of the function at the current point to calculate the best direction to look for a better point. Using this information the BFGS algorithm can iteratively calculate a better approximation to the inverse Hessian matrix. When compared to the two other algorithms implemented in Gadget, BFGS is a very local search compared to simulated annealing and is more computationally intensive than the Hooke and Jeeves. However the gradient search in BFGS is more accurate than the step-wise search of Hooke and Jeeves and may therefore give a more accurate estimation of the optimum. The BFGS algorithm used in Gadget is derived from that presented by Bertsekas (1999).

The total objective function to be minimised is a weighted sum of the different components. The estimation could be difficult due to groups of correlated parameters or multiple local optima. Therefore the optimisation procedure often involves a combination of the more robust simulated annealing, to make the results less sensitive to the initial (starting) values, and to the local search algorithms (Hooke and Jeeves and BFGS) in the neighborhood of the global optima.

3.1.2   Data handling

To facilitate flexible development of statistical fisheries models proper handling of data and related extraction processes is beneficial. For instance the bulk of ICES stock assessment reports deal primarily with available data and similarly most of the work goes into finding data (eg. ICES 2014, and other similar reports). Data availability is often subject to inter– and intra – institutional politics and excarcebated by differences in data storage.

The dst^2–project included the creation of a dedicated data-base specially designed to accomodate the needs of stock assessments. It contains minimally aggregated data on the species of interest, and aggregation based on geographical units specially chosen to reduce intra-division correlations with a minimal time step of 1 month. In addition the database is designed to be portable and widely applicable in the sense that the ETL[1] process is based on strict ascii files. The design and implementation is fully open source. One of the key attractions of an approach like this is that it is designed to create input files for modelling frameworks such as Gadget. The desired aggregation dimensions can be defined and altered with relative ease, allowing for testing of the effects of various aggregation schemes.

---

[1]Extract, Transform, Load

The spatial scheme for the data base is a hierarchical structure that represents spatial structure that is suitable for the types of models that can be implemented in Gadget. They should reflect hydrography, bathymetry, species distribution and to a lesser degree fishing controls. In this scheme all data is aggregated within a spatial subdivisions, but sufficiently disaggregated to allow for detailed modelling of relevant processes. For modelling, the spatial structure needs to allow for adequate definitions of spatial stock structure, migration patterns and predator-prey overlap. Each area should also be considered relatively homogeneous in terms of bathymetry and hydrography

The process of defining these spatial subdivisions is described in Taylor (2003). It follows previous work on Bormicon areas, introduced in Stefánsson and Pálsson (1997a). The main spatial features accounted for by this scheme are:

- Divisions at depth 500m, which represents the border of the Icelandic continental shelf area.

- Hydrography separates the shelf area to northern and southern areas.

- Larger divisions are split into subdivisions by 200m depth contour lines. These subdivisions are illustrated in Figure 3.2.

The motivation for this definition is mainly ecological, where the area definition follows the main features of the Icelandic continental shelf area. These subdivisons were validated by minimising their intra-correlation through a study of a combination of spatial rectangles and the output from a clustering analysis of samples obtained in annual spring and autumn surveys within the Icelandic EEZ.

3.2 Review of data

3.2.1 Diet data

The first paper of this thesis, Paper I, provides a comprehensive analysis of the feeding ecology of minke whales in Icelandic waters. The minke whale material used in this study was collected during 2003-2007 as a part of a comprehensive research programme (MRI 2003). To ensure representative sampling, searching effort was distributed temporally and spatially in proportion to densities of minke whales as known from previous surveys (Borchers et al. 2009). The sampling was designed according to an established delineation of the Icelandic
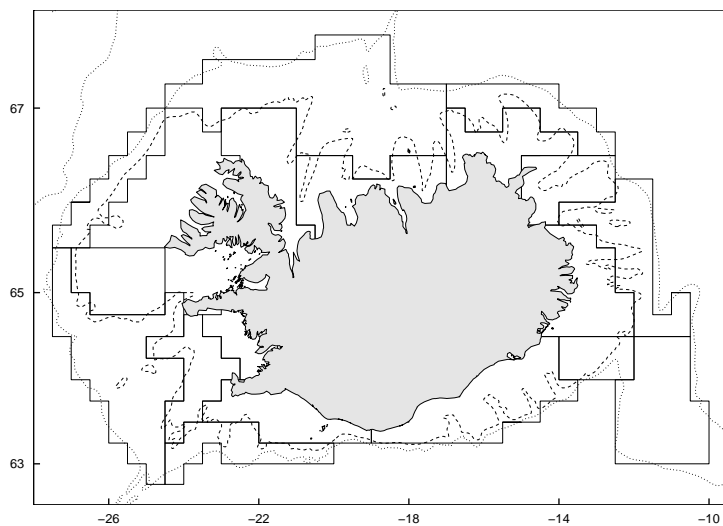
Figure 3.2: A graphical representation of the dst^2 subdivisions on the Icelandic continental shelf area used for the data warehouse.

continental shelf area into nine sub-areas (Figure I.1), as defined by Stefánsson
and Pálsson (1997a). To avoid selective sampling exhaustive attempts were
made to catch the first whale sighted within a given subarea and period. Whales
were taken onboard immediately after death for detailed dissection and sampling
for the various sub-projects (see MRI 2003, for further details on the sampling
and aims on the study).

The stomach contents of minke whales obtained from the scientific study on
the ecological status of minke whales in Icelandic waters were analysed based
on food remains in the stomach. The details of this analysis are described in
Paper I and discussed in section 4.1. Uncertainty estimates are obtained using
a bootstrap treating individual whales as sampling units, in effect treating the
stomach data as a disparate array of prey data.

3.3   Uncertainty methods

3.3.1   A bootstrap method for highly disparate datasets

The original definition of a subdivision in the specialised database was conceived
to allow for the rigorous development of multi–area models. As noted in the
introduction all data within those subdivisions are aggregated by relevant di-
mensions such as species, age, length and maturity, where relevant and possible.
A single subdivisison therefore represents a ragged array of measurements of a
particular stock within the spatial subdivision at a certain time. The input files
for the various stock assessment models are then created by simply adding the
ragged arrays together to form the input. This effectively treats the subdivisions
as independently and identically distributed samples. The independence arises
from the way the subdivisions are created, while the identical distribution arises
from observations made in Taylor (2002).

The second paper, Paper II, of this thesis presents a novel use of the boot-
strap. Similar to the approach used for the minke whales, where individual
whales where treated as disparate arrays, a more rigorous definition of sampling
units, the spatial subdivisions illustrated in Figure 3.2, is proposed. This ap-
proach is generic, in the sense that it can be applied to substantially different
models of marine populations, as it only involves combinations of elementary
sampling units. Furthermore the methodology was contrasted with a more tra-
ditional Hessian–based inference.

The methodology was illustrated on the model for Cod in Icelandic waters,
described in Taylor et al. (2007). Comparisons were made to variants of the

bootstrap approach. Initially a sample of 1000 bootstrap samples were drawn and a Gadget model fitted to each of them using the iterative reweighting procedure. This was compared to three other inference approaches:

- Bootstrap with 100 samples

- Bootstrap with 1000 samples but likelihood weights condititoned to those obtained using the original sample.

- Hessian–based approximation.

The results are discussed in section 4.2 and Paper II.

3.3.2   Testing stock structure hypotheses

As alluded to in section 1.2.1, proper assignment to management units is potentially crucial to sustainable managment of a marine resource. Paper IV performs simulation study on the utility of a genetic tagging experiment on North Atlantic fin whales, augmented with information genetically determined close kin, when contrasting two competing stock structure hypothesis.

In the analyses which follows comparisons will be made on the basis of two possible stock structures, mixing or dispersal type, as shown in Figure 3.3. When mixing dynamics are assumed, separate breeding stocks overlap (to some fixed degree) on the feeding grounds, while dispersion denotes the permanent migration between breeding stocks.

In this analysis two competing stock structure hypotheses, that consider substantially different stock dynamics that cannot be distinguished using conventional means, are simulated. The main difference between the two hypotheses lies in the type of stock overlap on the feeding grounds and how this relates to breeding grounds. One hypothesis suggests that in the waters around Iceland there are three separate breeding stocks that only marginally overlap, or mix, on feeding grounds. The other suggests that there are three breeding stocks while the whales disperse between them so effectively these breeding groups are not separate. The constrast between the two hypotheses is illustrated in Figure 3.3.

To test the difference between the two stock structure hypothesis a tagging experiment east of Greenland, with recaptures by Icelandic whaling vessels west of Iceland, was simulated to calculate a satisfactory number of tags to reliably detect differnces between the two. Furthermore the effects of adding information on genetically determined close relatives was investigated. The simulation model
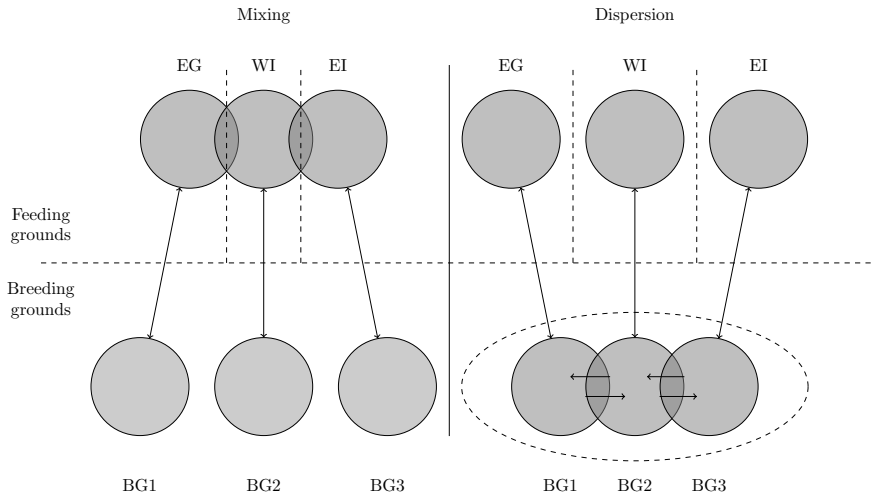
Figure 3.3: An overview of the competing stock structure hypotheses for the North Atlantic Fin whales in the Irminger sea. The figures illustrate three breeding components with unknown breeding grounds BG1, BG2 and BG3 that migrate to feeding grounds East of Greenland (EG), West (WI) and East of Iceland (EI). The figure on the left illustrates the mixing stock structure, where there is no gene sharing with some degree of overlap on the feeding grounds (shaded region). The figure on the right shows the dispersion hypothesis, i.e. genetic sharing between the breeding components, where arrows indicate movement between breeding components.

was an age–based model configured to mimic the functionality of Punt (1999). The model was implemented using simulation tools described in Paper III.

3.4 Additional functionality

3.4.1 Standard process for fitting models within the Gadget framework

One of Gadget's key strengths is the assimilation of disparate datasets into a weighted likelihood function at a level that requires minimal processing of data. For instance length distribution and age – length distribution are contrasted to the model output simultaneously. This obviously increases the number of

comparisons substantially as now the model compares to the length distributions, that are in general at finer scale than ages. This can, depending on the likelihood component weight and the starting value of the optimiser, lead to a biased result as one component may end up dominating the end result.

To account for these potential biases in the weighting scheme Stefansson (2003) proposed an objective heuristic for assigning these weights to individuals likelihood components. This heuristic, which is a generalisation of the normal weighted regression procedure, was first applied to a real data set in Taylor et al. (2007). This thesis provides further advances in the application of this heuristic, as described in chapter III in section III.4.1. The iterative reweighting heuristic was implemented in R (R Development Core Team 2011) along with as set of diagnositic tools to analyse the final fit. These tools where applied in Papers II and V.

3.4.2   Digesting output from Gadget

Gadget models can produce fairly detailed output. This includes model prediction of stock status by age and length, consumption of predators by the same dimensions for both predator and prey, and comparison to data used by the likelihood function. Apart from the iterative reweighting procedure, Paper III describes tools to digest this output from Gadget and standardise the results. Routines were developed to present the results from iterative reweighting, produce biomass results, do yield per recruit analysis and forward projections, and plot the fit to data (section III.5).

3.4.3   Simulation model

In addition to methods designed to interact with Gadget, Rgadget includes an independent implementation of the key processes of Gadget in R, described in section III.6 of Paper III. An application of this simulator can be found in Paper IV described above.

3.5   A model of cod and minke whales

Paper V details the developement of single area model of the interaction between minke whales and cod. It extends the model of cod developed in Taylor et al. (2007) and estimates the interaction between cod and minke whales using data obtained in Paper I. The paper explores what effects the inclusion of data

Variants	SI_i 4 - 17	SI_i 17 - 33	SI_i 33 - 140	SI_a 16 - 27	SI_a 27 - 39	SI_a 39 - 140
V1	G_1 F	G_1 F	G_1 F	-	-	-
V2	G_1 S	G_1 S	G_1 F	-	-	-
V3	G_1 F	G_1 F	G_1 F	G_2 F	G_2 F	G_2 F
V4	G_1 S	G_1 S	G_1 F	G_2 S	G_2 S	G_2 F
V5	G_1 F	G_2 F	G_3 F	G_1 F	G_2 F	G_3 F
V6	G_1 S	G_2 S	G_3 F	G_1 S	G_2 S	G_3 F

Table 3.1: Model variants tested in the case of cod in Icelandic waters. G. denotes the weighting group, F denotes a fixed slope in the log-linear relationship between survey index and model biomass and S the case where the slope is estimated.

on minke whale predation of cod has on the fit to data when compared to the respective single species models. In addition to linking the two models it tests model sensitivity by implementing six different model variants for cod, illustrated in Table 3.1, where different relationships and weighting schemes of survey indices are explored.

As noted in section 1.3.1 appreciable changes are observed in abundance of minke whales on the Icelandic continental shelf area. To follow these changes migrations in and out of Icelandic waters is allowed to vary to follow these trends. As suggested in Paper I sandeel has been a substantial part of minke whale diet. Therefore fluctuations in minke whale abundance in the model are explained by linking the migration proportion to indices of sandeel biomass as follows:

$$\begin{pmatrix} 1 & \delta_s m + m_c \\ 0 & 1 - (\delta_s m + m_c) \end{pmatrix} \tag{3.1}$$

where columns indicate area of origin ("Icelandic waters" and "other"), rows movement to area, δ_s is the sandeel abundance index and m and m_c are the relative proportion of minke whales that migrate to the Icelandic continental shelf during the summer. Autumn migrations are estimated similarly but in the opposite direction.

4

Results

The analyses in this thesis are roughly split between themes according to different aims of this thesis. A major theme of this thesis is the consideration and estimation of the types of uncertainty encountered when building models of marine multi–species ecosystems, as described by Butterworth and Punt (1999) and references therein. The principal focus is on observation uncertainty, that is uncertainty that arises from population sampling, e.g. those that arise from surveys, port sampling etc. Other sources of uncertainty, such as process error, model error and estimation error are also discussed. Additional functionality, developed for this thesis and discussed in Paper III is used throughout this thesis.

The aims of the thesis serve as substance for the five papers presented in Part II, of which three have been published in peer reviewed journals. This chapter provides an overview of the main results on which this thesis is based. For comprehension the sections are arranged by the aims of this thesis.

4.1 Feeding ecology of minke whales

Paper I details the study of the feeding ecology of minke whales. The main findings indicate a substantial shift in the diet of minke whale to previous studies as illustrated in Figure I.3. Furthermore, as evident from Figure 4.1, significant changes are observed in stomachs in the southern areas. For example, the

frequency of occurence of sandeel in stomachs goes from being present in all stomachs to being present in almost a quarter of stomachs. The size of prey eaten by the whales appeared to be fairly large, as illustrated in Figure 4.2.

4.2   Uncertainty estimates

Paper II outlines a novel application of the bootstrap to estimate uncertainty of fisheries stock assessment models. Figures 4.3 and 4.4 illustrate the relevant biomass estimates and confidence bounds based on the bootstrap and the Hessian–based approximation. The bootstrap approaches appear to behave similarly, while the Hessian–approach appears to be overly confident with the results. The bootstrap approach with conditioned weights does however appear to miss possible biases as shown in Figure 4.5.

An analysis of the convergence of this approach, with respect to the number of samples, is illustrated in Figure 4.6. The figure, in combination with that illustrated in Figure 4.3, shows that the estimated CV of the variables is sufficiently determined after 100 bootstrap samples.

The Hessian–based approximation appears to behave poorly for this particular type of model. The reason for this appears to be due to the combination of the type of likelihood function used and the number of data points, which in this case is ≈ 15000 points. If the number of points is artificially inflated by increasing the number of intra–year time steps from 4 to 6 the Hessian–based confidence bounds shrink accordingly while the bootstrap inferences are unaffected.

4.3   Rgadget

The third paper, Paper III, develops a set of tools to interact with and process output from Gadget written in R (R Development Core Team 2011) as a special package. Initially the key motivation behind this developement was the implementation of the iterative reweighting heuristic. This heuristic assigns weights iteratively by finding an optimal solution for each dataset separately, basing the weight of that dataset on the optimal fit and then optimising the weighted likelihood of all datasets. In contrast with other weighting schemes this approach provides an important diagnostic, as illustrated in Table 4.1, in the composition of the variance estimates after each of the two stages (Taylor et al. 2007). Substantial departures in the model fit from the secondary stage
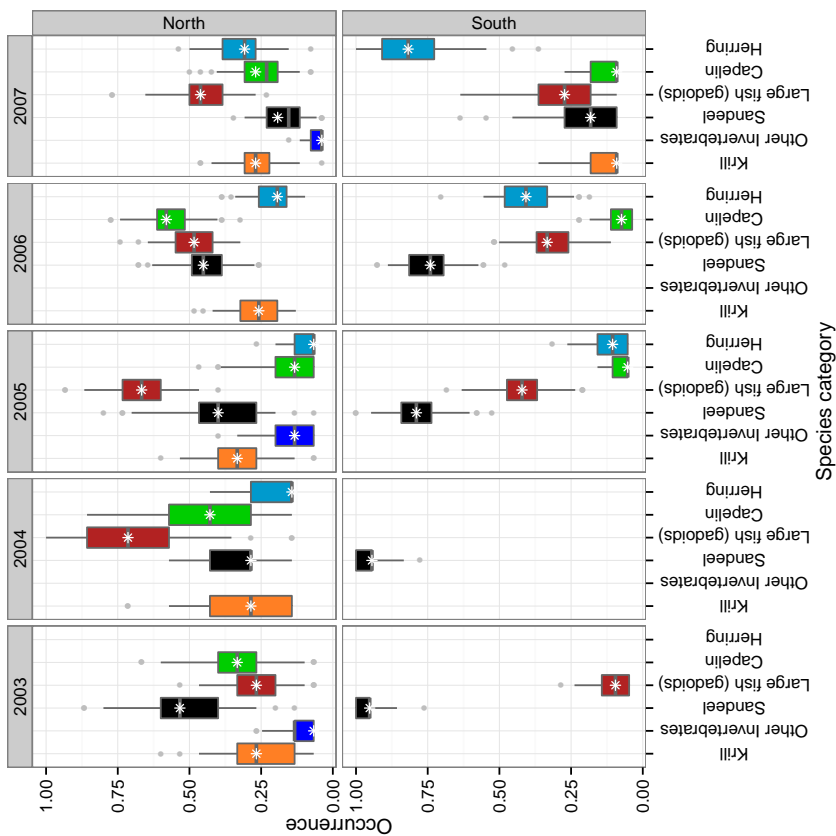
Figure 4.1: Boxplots of spatial and temporal variation in diet composition of minke whales expressed as frequency of occurrence (FO). The point estimate is indicated by a white star while the solid grey lines through the boxes represent the bootstrapped medians. The box indicates the interquartile range and the whiskers 95% confidence intervals. Any further outlying data points are indicated as points. Area division (NorthSouth) as in Figure I.1
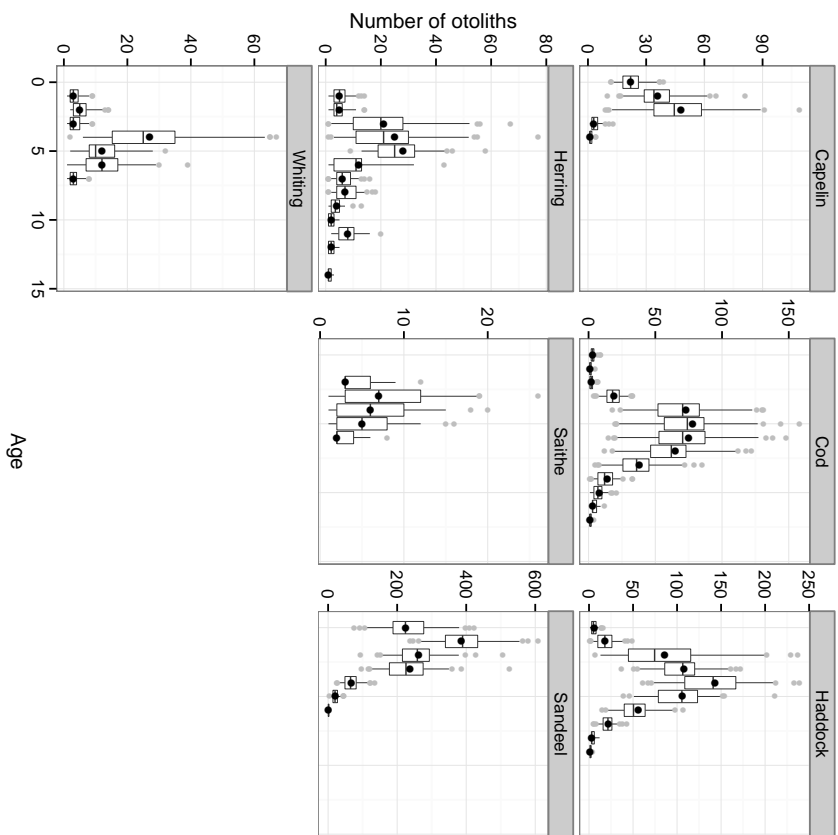
Figure 4.2: Boxplots of age distributions of the most common fish species identified in the fore-stomachs of minke whales sampled in 2003 2007. Boxplot parameters as in Figure I.6.

from the optimal fit obtained in it the primary stage would indicate model mis-specification. Furthermore Table 4.1 could provide further insights into where these conflicts lie. For example, if the estimated survey variance is much higher when maturity data are up-weighted, this indicates a model error, which needs to addressed, through model re-specification, not merely through an automated weighing scheme or estimation in phases.

	ALKc	ALKs	LDc	LDs	SI 20-39	SI 40-69	SI 70-110
Catch	7787	9158	26720	50630	15.91	36.19	83.25
Survey	8706	4944	154900	23030	14.46	36.22	99.91
Sind	11990	10780	113400	74260	1.797	4.335	17.75
Final	7676	4993	22480	15950	2.641	4.817	21.03

Table 4.1: An example fit diagnostics for tusk (*Brosme brosme*) in Icelandic waters from the iterative reweighting procedure. Rows indicate the weighting group, as described in Table III.3, and columns the final score from the data set when a particular weighting group is emphasized. Here ALK and LD denotes age–length and length distributions from surveys *s* and commercial *c* catches respectively, and SI survey based on the length intervals indicated in the subscript.

Other tools, described in Paper III, include functions that interact with, digest and standardize output from a typical Gadget model. This makes the comparisons of model estimates to data, as illustrated in Figure 4.7, and derivation of biomass estimates and forecasts, shown in Figure 4.8, fairly simple.

4.3.1   Fin whales in the North Atlantic

The results from the simulation experiment in the fourth paper, Paper IV, indicates that considerable improvements can be made to traditional tagging experiments by adding information on genetically determined close relatives. Table 4.2 shows the rejection probability as a function of tags/biopsy samples. This results illustrates that a traditional tagging experiment would have had substantial difficulty in detecting the difference between two hypothesess, i.e. the mixing hypothesis where separate breeding stocks only overlap on feeding grounds and the dispersion hypothesis where whales disperse annually between three breeding components. Additional power is sought simply by the increased effective

Table 4.2: The probability of rejecting the mixing hypothesis conditioned on the stock hypotheses using a direct tag-recapture experiment (time trend in tags), total recaptures and ρ, the ratio of total recaptures to related individuals within the catch. The rejection interval was chosen such that the probability of type I error was 5%

Number of tags	Rejection probability		
	Trend	Total recaptures	ρ
100	0.11	0.05	0.39
200	0.09	0.09	0.62
300	0.04	0.27	0.77
400	0.13	0.21	0.81
500	0.21	0.26	0.91
600	0.17	0.25	0.86
700	0.17	0.22	0.90
800	0.19	0.42	0.95
900	0.25	0.51	0.98
1000	0.25	0.35	0.97
1100	0.38	0.58	1.00
1200	0.34	0.41	1.00
1300	0.32	0.54	0.99
1400	0.58	0.59	1.00
1500	0.38	0.60	1.00

sample size when considering relatives of tagged animals as being tagged also. This is apparent from the difference in the total number of recaptures. This is further improved by comparing the ratio of the total number of recaptures to relatives within the catch, ρ, as a test of the alternative hypotheses.

4.4   A multi–species model of cod and minke whale interactions

The fifth paper, Paper V, implements a model of cod–minke whales interactions. Results indicate that all model variants appear to follow the trends in the data reasonably well (Figure 4.7). For variants, listed in Table 3.1, where species interactions are modelled, modelled cod indices appears to grow faster after 2000, when at the same time the abundance of minke whales decreases. These differences become more apparent when the stock status is projected for-

ward. Figure 4.8 shows mature biomass of cod projected forward 40 years given constant fishing mortality, based on the average fishing mortality for the last 5 years of the model. Model variants where interactions are ignored appear on average to predict substantially lower biomass than those that represent inter- actions. Figure 4.9 shows the total consumption of cod by minke whales as estimated in the model. There the consumption ranges between 4500 to 7500 tons in 1980 whereas in 2012 it was around 2500 to 3000 tons.
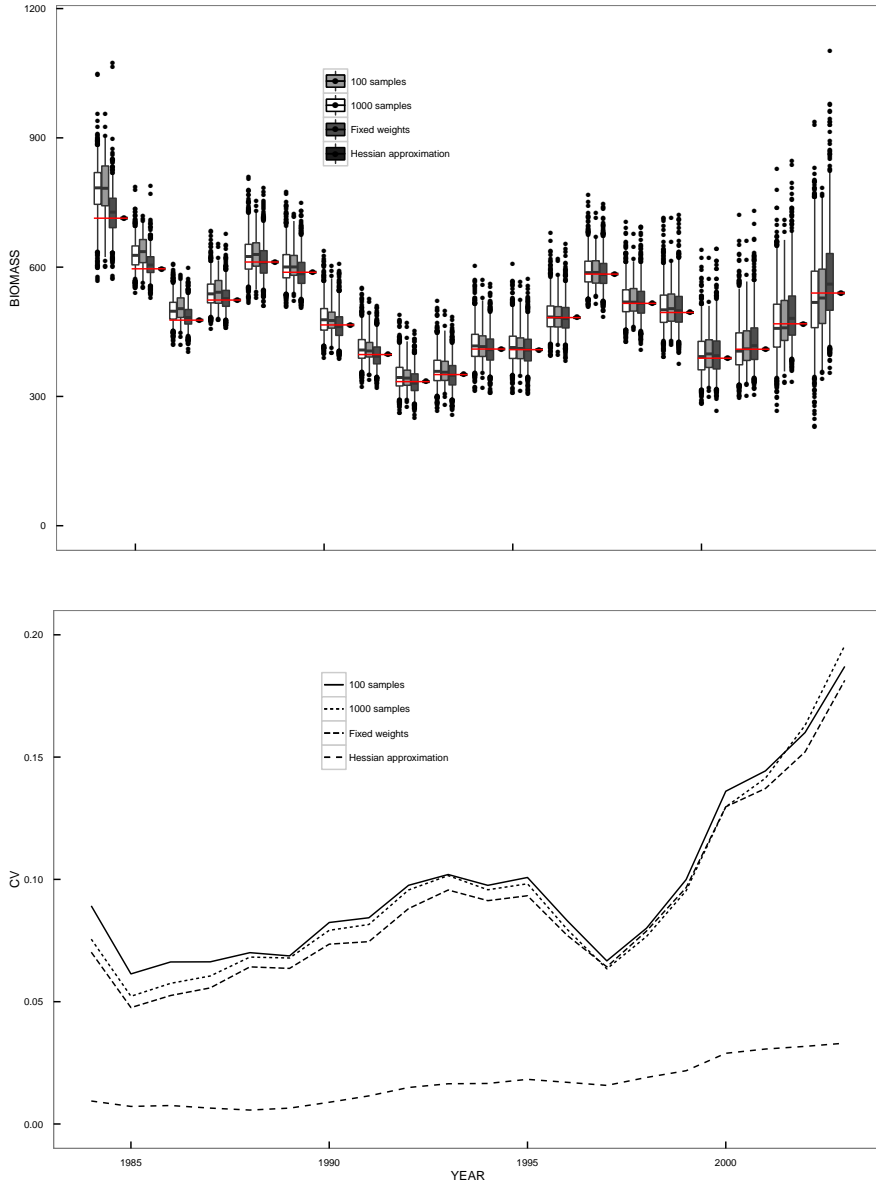
Figure 4.3: Boxplot (top panel) of the end of year biomass for cod of age 4 and older estimated from 1000 bootstrap samples, using both iterative weighting for each sample and using the fixed weights for all samples, compared to 100 bootstrap samples. The fixed weights were obtained using iterative weighting for the original dataset. The point estimate is indicated by the central red line through the boxes. The box indicates the interquartile range and the whiskers 95% confidence intervals. Any further outlying data points are indicated as points. Bottom panel shows the estimated CV for the age 4^+ biomass using the same methods as above.
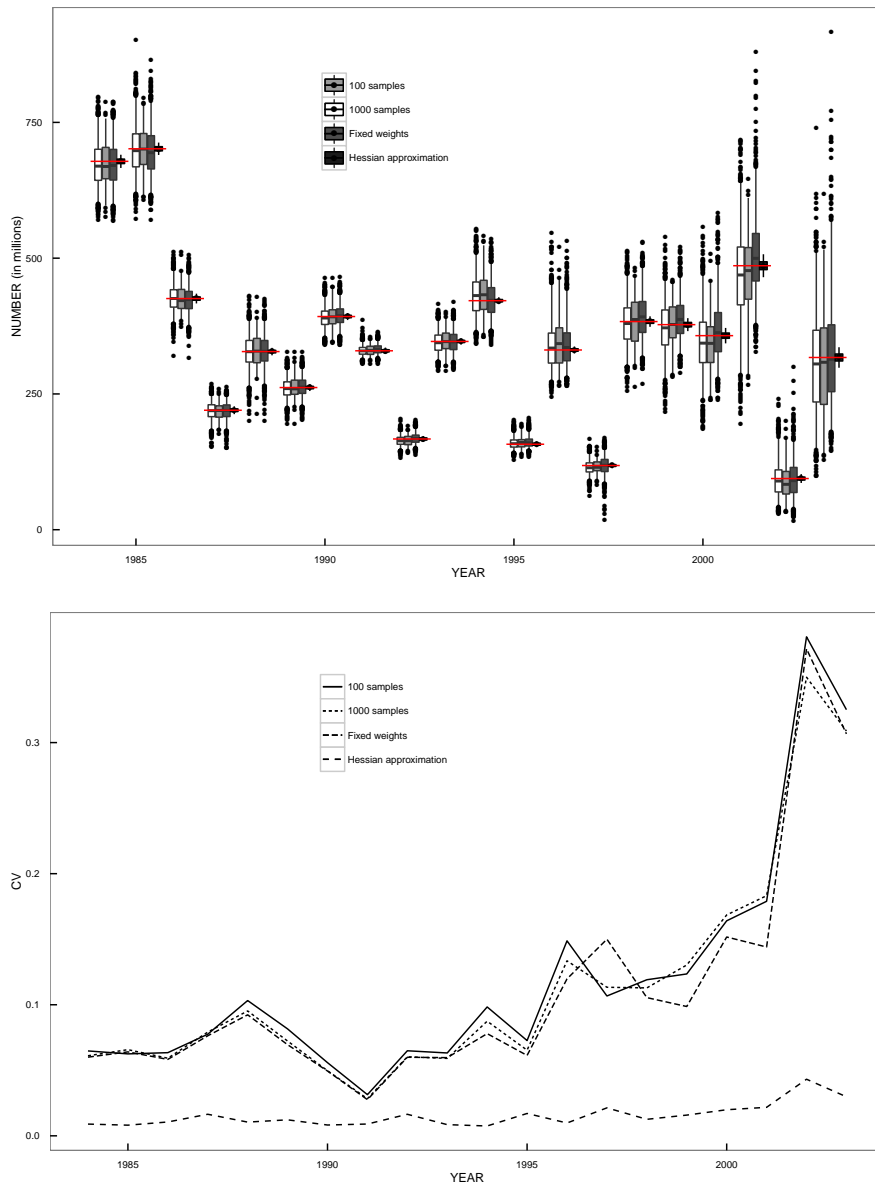
Figure 4.4: Boxplot (top panel) of the number of recruits (age 1) in each year estimated by 1000 and 100 bootstrap models compared to 1000 bootstraps with fixed weights and a Hessian–based approximation to the 95% confidence interval. The point estimate indicated by a central red line through the boxes. Bottom panel shows the estimated CV for the recruitment using the same methods as above.
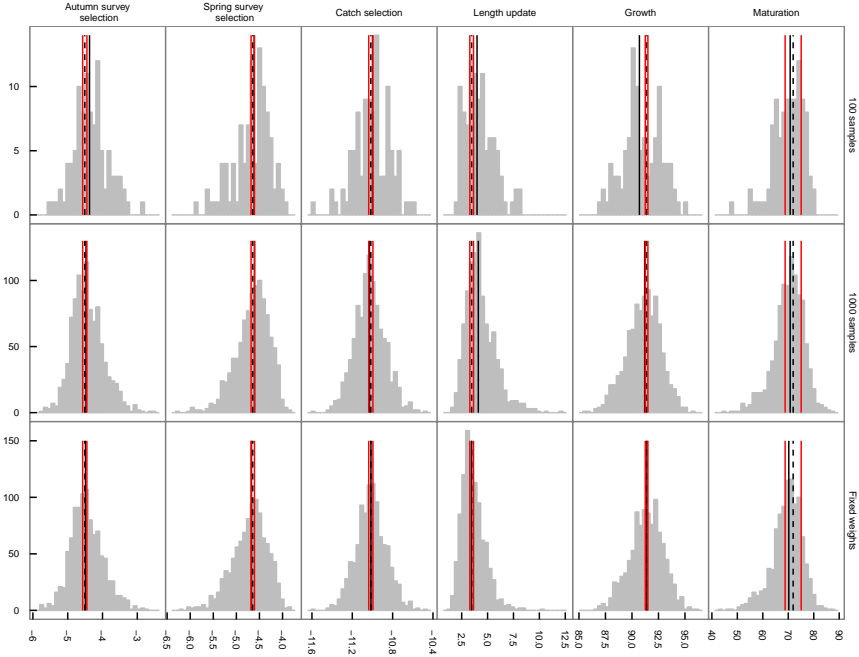
Figure 4.5: Histograms of the estimated fleet selection parameter a_f for the three fleets (October survey, March survey, Commercial catch), β the parameter defining the length update matrix, k the growth rate and the maturity l_{50}. The parameter estimates were obtained from 1000 bootstrap samples, compared to a smaller number of bootstrap samples, 100, and for each of these two samples, iterative weighting was applied to all bootstrap samples. This was then compared to 1000 bootstrap samples obtained through a process in which the parameter estimation the weighted likelihood function is conditioned on the original weights. The point estimate (black broken line) and bootstrap mean (black solid line) along with 95% confidence bound obtained from a Hessian–based approximation to the variance covariance matrix (red solid lines) are indicated.
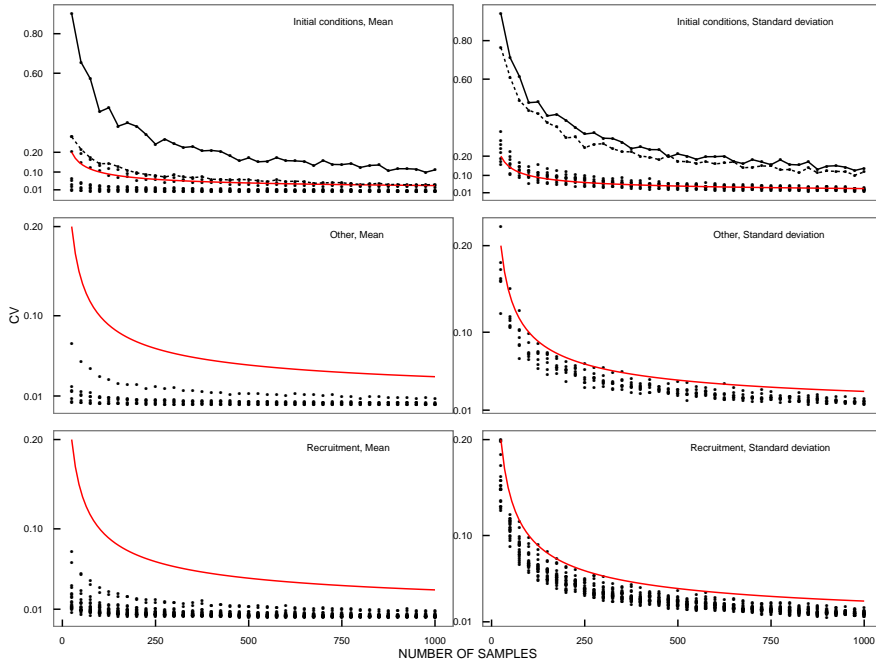
Figure 4.6: Results of a restrospective bootstrap sampling on the parameter estimates from the 1000 bootstrap samples, with iterative weighting applied to all samples. This retrospective bootstrap studies the variation of the mean and standard deviation of each parameter estimate by calculating the coefficient of variation (CV) as a function of the number of bootstrap samples, n, of both the mean and standard deviation (SD). A point on the graph shows the CV of the mean (panels on the left hand side) or SD (panels on the right hand side) for a particular parameter and number of samples, n. The different panels contain the CVs of the initial number at age (ν_a in eq. II.6), "Other" variables i.e. the variables which are shown in Figure II.3, and yearly recruitment shown in Figure II.5 (R_y in eq. II.5). The CVs of the initial number at ages 8 and 9 are illustrated with solid and broken lines respectively. For comparison $1/\sqrt{n}$ is shown (red solid line) on all panels.
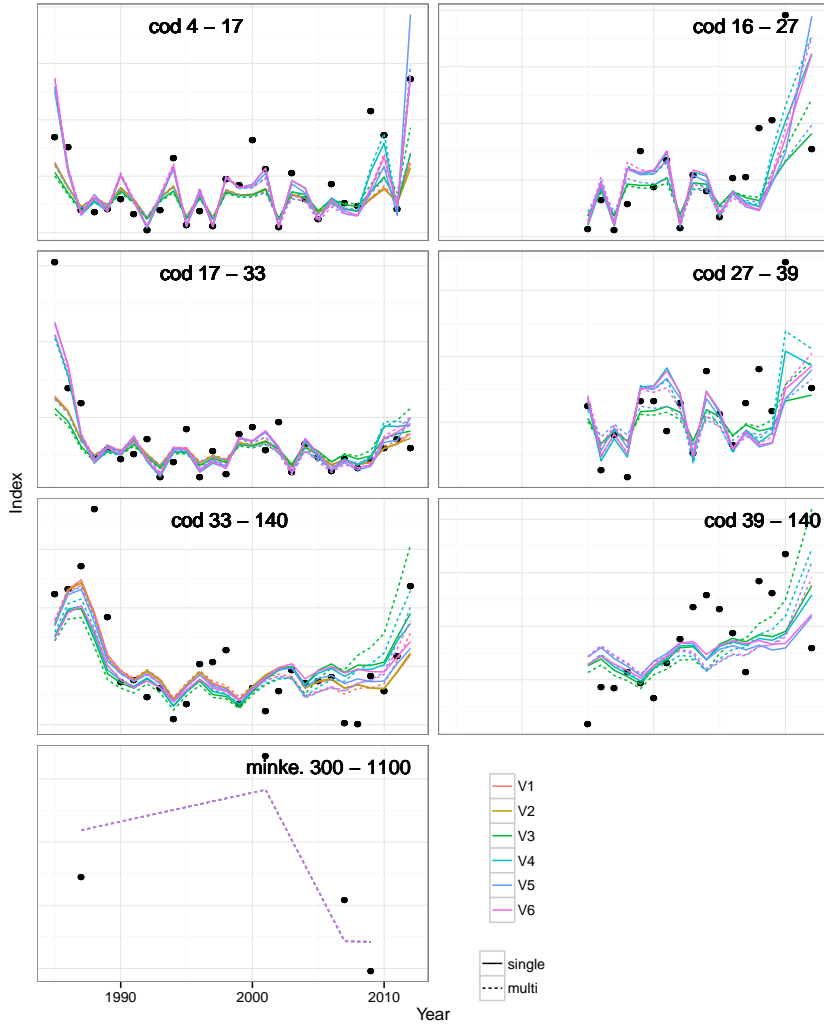
Figure 4.7: Results from the different model variants of cod and minke whales in Icelandic waters, described in Table 3.1. The figure shows the survey index by length group (cod) and abundance index (minke). Figures on the left indicate for cod indicate the spring survey indices and on the right hand side the autumn survey indices. Lines represent the fit to survey indices (points). Solid lines indicate a single species fit while the broken lines fit with interactions.
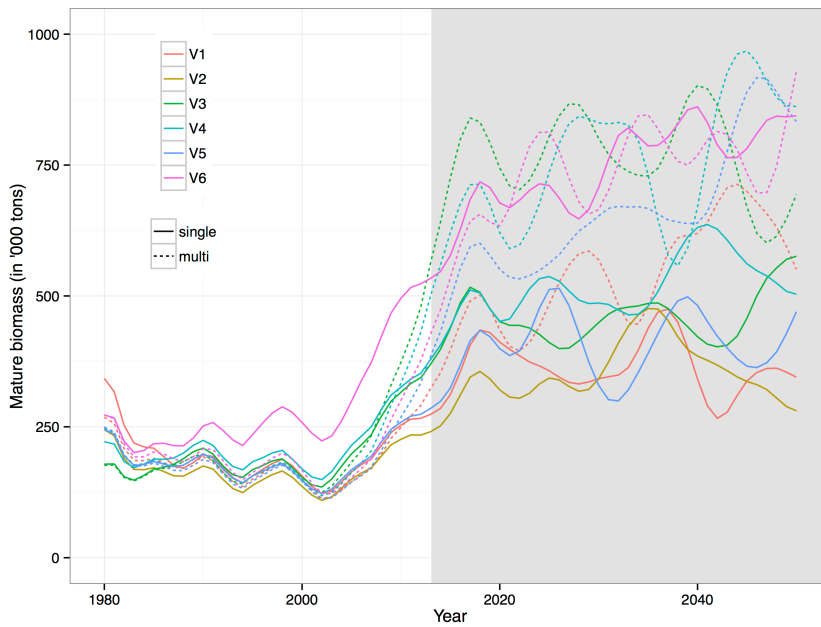
Figure 4.8: Projections of mature biomass of cod as a function of year by model variant and shown for each of the model variants as indicated in Table 3.1. The shaded area shows the projection period and olid lines indicate a single species fit while the broken lines fit with interactions.
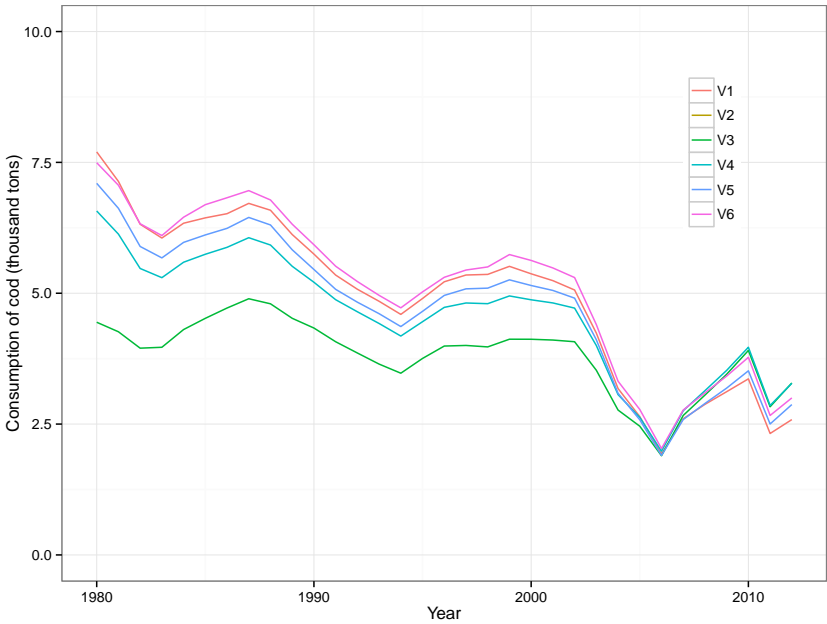
Figure 4.9: Total consumption of cod, both mature and immature, by minke whales as a function of year and by model variants as indicated in Table 3.1.

5

Conclusions and further perspectives

Various aspects of the use of statistical models of marine multi–species ecosystems have been discussed and analysed in this thesis. Section 5.1 discusses how the bootstrap approach developed in this thesis has allowed for a more general comparison of models based on the same dataset. Section 5.2 describes how information on the status of minke whales in the ecosystem has allowed for the modeling of the interaction between cod and minke whales. Section 5.3 discusses the potential utility of the tools developed in this thesis. Finally Section 5.4 discusses the overall contribution of this thesis.

5.1 Uncertainty methods

This thesis has presented a novel application of bootstrap methods suitable for models of population dynamics. Several modifications and alternatives to the original bootstrap methodology (Efron 1979, Efron and Tibshirani 1994) have been presented. For example, to account for correlations in simple non–replacement sampling schemes (as used for most questionnaires or "sample surveys"), without–replacement bootstraps and with–replacement bootstraps have been suggested along with somewhat more general resampling procedures for

complex survey data (McCarthy and Snowden 1985, Gross 1980, Rao and Wu 1988, Sitter 1992). Theoretical assumptions and derivations behind these approaches do not easily extend to the present situation with disparate data sets, composite likelihoods in the estimation phase and last but not least, the highly nonlinear population dynamics models used as a basis for obtaining predicted values and error sums of squares or likelihood functions. The "trick" in the current proposal is not a theoretical development but the methodology of having the bootstrap sampling unit y_i as a collection of all relevant datasets sufficiently aggregated such that they can be assumed to be independent.

When designing an aggregated data base to be used for modelling, several issues need to be taken into account. The most important statistical condition on the choice of the "data units" is that correlations between them should be minimal. On the other hand, there also needs to be a fair number of them within each model area if the bootstrap mechanism is to provide some variation in results. For a given measurement type one can in many cases investigate spatial correlation or variograms to determine the distances at which those become negligible (Petitgas 2001). This can not easily be done for many data types, however (age-length tables, tagging experiments etc). In fact, the original reasoning for the areas used in this paper was ecological (Stefánsson and Pálsson 1997a, Taylor 2003) rather than based on spatial correlation. It is likely that in most real situations data will be aggregated either according to such criteria or pragmatically into "statistical rectangles" of some form.

5.2   Multi–species models of ecosystems of the Icelandic continental shelf

The analysis presented in this thesis provides the first attempt to link cod in Icelandic waters to a large mammalian predator, minke whales, in a statistical age-length based model. It provides a basis to answer some of the questions set out by Stefánsson et al. (1997) and subsequently MRI (2003). Furthermore, it provides insight into the restraints on the model by adding interactions between species, particularly species with such disparate data available as cod and minke whales, and their impact on the overall fit to individual datasets.

The abundance of minke whales has been observed, both by whalers and whalewatchers, to have decreased substantially in the last decade. Here the major driver in the abundance is considered to be linked to sandeel biomass, based on observations from Paper I. Additional drivers for minke whale abun-

dance in Icelandic waters have not been excluded. For instance capelin has been identified as an important part of the diet of baleen whales (Sigurjónsson and Víkingsson 1997) and a shift in capelin summer distribution away from Icelandic waters has been observed in recent years (Pálsson et al. 2012). Preliminary work to include capelin in a multi–species model has however been unsuccessful due to the nature of the data (Taylor 2011) but work is on-going.

The effects of minke whales on cod are surprisingly small, judging by the fit to data. The difference between model variants with minke whales predating cod and those with no predation becomes apparent when projecting the stock status. When interpreting stock projections for cod, considerable care needs to be taken as the changes in ecosystem that cause the departure of minke whales, which are not included in the model, could also negatively affect cod. The consumption of minke whales estimated here may seem minor in comparison with fleet catches. However most of the consumption is focused on juvenile cod within the models, while the commercial fishery targets larger fish. The mortality related to whale predation is also smaller by an order of magnitude. This may contradict results in Paper I which indicated substantially bigger consumption than estimated here. Here the form of the suitability function may play a role, as more dome–shaped function may be more appropriate. It cannot, however, be ruled out that the data on consumption are inconsistent with other data used with respect to the model.

As with all statistical models, access to representative data is necessary to estimate the processes of interest. Here for the first time information on the feeding ecology of minke whales in Icelandic waters is analysed. Data available from whaling operations before the moratorium, although limited, suggest considerable changes in the diet. Therefore the effects suggested by the model may not be as pronounced.

Sensible next steps could include a wide range of analyses. Minor changes could include modifying minke whale selection to be changed to be more dome shaped, testing different levels of depletion and productivity, and altering the assumptions of their energy requirements. Major changes could include continued developement of the two area model of cod described in Taylor (2011), inclusion of other species such as sandeel and capelin, which are seen as a large part of minke whale diet. Fleet composition will need some scrutiny as well.

5.3   Rgadget

RGadget provides utilities for rapid developement of Gadget models and the subsequent analyses. It is hoped that this will encourage the application of Gadget in new areas. During package development, parts of RGadget have been applied in various settings. The R simulator has seen some published applications, when tagging was developed in Gadget (see Hannesson et al. 2008, for further details) and IV where a simulation study on fin whale stock structure was presented. The iterative reweighting procedure has seen many applications during its developement (e.g. Elvarsson et al. 2014a, Elvarsson and Þórðarson 2014, Þórðarson and Elvarsson 2013, to name a few).

Rapid model development is a fairly large issue for models of the scale and magnitude of a typical Gadget model. For instance scale changes can be fairly cumbersome and errorprone to implement, while potentially necessary to prevent unnecessary noise in the input data (Demyanov et al. 2006). dst^2 (2004) recommended the development of a specialised database system (Kupca 2006) which allows for exactly this kind of analysis and flexible extraction of ecosystem data. This database system is currently being revised as part of MareFrame (2013). RGadget compliments these developments by providing a scriptable generation of model settings for Gadget and digestion of results.

The development of RGadget echoes efforts of other similar projects. Notably r4ss (Taylor et al. 2014) and to a lesser extent FLR (Kell et al. 2007) as these packages aim to standardise model results in the general stock assements and management strategy evaluations process. Both RGadget and r4ss are designed around their respective modelling frameworks, i.e. Gadget and Stock synthesis. FLR appears to be aimed for a more general integration with age based stock assessment models such as XSA (Shepherd 1999) and SAM Nielsen and Berg (2014). A link between Gadget and FLR was developed in Howell and Bogstad (2010) for harvest control rule evaluation. Further integration may be considered at some stage, and RGadget could in some sense be considered a first step in that direction.

The utility of an independent Gadget-like simulation engine in R is considerable. Appart from the obvious consistency check, it can be used as a tools to test the feasibiltity of potential new features for Gadget. Additionally it could be used as a tool to study different assumptions of the interaction between the two substocks and harvesting strategies as done in Paper IV. As such the simulator could be used as a teaching aid in fisheries science. Data can also be generated for complicated modelling frameworks such as Gadget. For example in (Hannesson et al. 2008), an earlier version of the simulator was used to simulate a

tagging experiment designed to test Gadget's performance.

5.4   Final words

This thesis introduces novel tools that aid development of statistical models of marine ecosystems. Although this thesis focuses on models developed using the Gadget framework the work presented here is sufficiently general to be applicable in a more general setting. Notably the bootstrap method, described in Paper II, is applicable to a range of problems that stem from samples of marine ecosystem. An example could be that it provides a common ground between substantially different approaches to fisheries stock assessment, such as models developed using Ecopath with Ecosim and Gadget.

Methods described in this thesis have already seen applications outside of it. These include applications described by Þórðarson et al. (2011), where the iterative reweighting procedure from Rgadget is used for golden redfish and tusk in Icelandic waters. Elvarsson and Þórðarson (2014) used the whole suite of iterative reweighting, bootstrap and forward projections implemented in Rgadget are used to perform an evaluation of an advice rule for ling in Icelandic waters.

In addition to methodological developments this thesis provides new insights into the interactions of minke whales with commercial exploited fish species in Icelandic waters. Although many question remain unanswered the work described in Papers I and V advances current knowledge, and provides indications for the next steps.

Part II

Papers

I

Paper I

Recent changes in the diet composition of common minke whales (*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate change?

Gísli A. Víkingsson, Bjarki Þór Elvarsson, Droplaug Ólafsdóttir, Valerie Chosson, Jóhann Sigurjónsson and Anton Galan

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
change?

50

Abstract

The common minke whale (*Balaenoptera acutorostrata*) is the most abundant mammalian top predator in Icelandic continental shelf waters. Here, an account is given on the first systematic research programme into the feeding ecology of common minke whales in Icelandic waters based on an analysis of minke whale stomach contents data collected in Icelandic waters during 2003-2007. The results show pronounced spatial and temporal variation in the diet. The temporal changes include a decrease in the proportion of sandeel in the diet over the study period and a corresponding increase in herring (*Clupea haerengus*) and haddock (*Melanogrammus aeglefinus*) particularly in the southern area. The diet also differed markedly from the previously available, limited data from Icelandic waters with less krill and the cold water species capelin (*Mallotus villosus*) and more gadoids and herring in the more recent period. These changes in diet composition are consistent with recent changes in the Icelandic continental shelf ecosystem including increased sea surface and bottom temperatures and changes in distribution and abundance of several fish species including sandeel (*Ammodytidae*) and capelin. Although natural fluctuations cannot be ruled out at this stage, these dietary changes, together with decreased abundance in coastal waters, may reflect the responses of minke whales to a changed environment possibly driven by global warming.

I.1   Introduction

The common minke (*Balaenoptera acutorostrata* Lacepede, 1804.) whale is the most abundant whale species in Icelandic coastal waters with mean densities up to 0.07 animals/nm^2 in the continental shelf area in recent decades (Pike et al. 2008a). The species is generally regarded as the most ichthyophagous of baleen whales and also the most opportunistic with regard to feeding habits (Tomilin 1957, Horwood 1990). Skaug et al. (1997) concluded that despite being quite flexible in feeding habits, minke whales off Norway show some preference for fish over krill. In the Northern Hemisphere, minke whale diet varies considerably among areas and large temporal variations have also been demonstrated for the species (Jonsgård 1982, Horwood 1990, Klinowska 1991, Haug et al. 1995 1996 1997, Neve 2000, Tamura and Fujise 2002). Although hunting of minke whales in Iceland dates back to 1914, knowledge on feeding habits of the species in Icelandic and adjacent waters, has been very limited. Sæmundsson (1932 1939) concluded from visual observations and accounts from fishermen that herring (*Clupea haerengus* Linnaeus, 1758) was the minke whales' most important prey species in Icelandic waters. During the 20th century, only 68 minke whale stomachs were examined, 56 from the commercial fishery during 1977-1978, one caught in 1984 and 11 stranded or by-caught animals between 1988 and 1997. Most of these animals were sampled off North Iceland in June and July. Thus, large parts of the distribution area and residence period of minke whales in Icelandic waters were virtually unknown. Based on the frequency of prey occurrence, this limited data indicated that approximately 65% of the diet consisted of fish while the remaining 35% were euphausiid crustaceans (Sigurjónsson et al. 2000). Among the identified fish species were capelin (*Mallotus villosus* Muller, 1776), sandeel (*Ammodytidae*), and cod (*Gadus morhua* Linnaeus, 1758), while the identified krill species were *Thysanoessa raschii* (M. Sars, 1864) and *Meganyctiphanes norvegica* (M. Sars, 1857). Preliminary investigations into the ecological role of minke whales in Icelandic waters have indicated that the species may have a significant effect on the long-term yield from fisheries, in particular cod and capelin (Sigurjónsson and Víkingsson 1997, Stefánsson et al. 1997). There was however, large uncertainty associated with these estimated effects, limited knowledge of the diet of minke whales in Icelandic waters being among the most important identified sources of uncertainty (Stefánsson et al. 1997). Considering the large spatial and temporal variations in the diet of minke whales observed in other areas, there is clearly a need for further studies into the feeding ecology of the species in Icelandic waters. Like most other baleen whales, common minke whales are migratory, spending the

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
52   change?

summer at high latitude feeding grounds and the winters breeding at lower latitudes (Horwood 1990). The Icelandic continental shelf constitutes the most important summer feeding ground of the Central North Atlantic population of minke whales judging from the NASS series of large scale standardized sightings surveys (Gunnlaugsson and Sigurjónsson 1990, Borchers et al. 2009, Pike et al. 2009). Abundance in this area has varied considerably since systematic monitoring of abundance began in 1986 with a maximum estimate of 43,633 (30,148 - 63,149) animals obtained in 2001 (Borchers et al. 2009). The most recent surveys (2007-2009) have shown considerably lower densities (Pike et al. 2011ab). While the reasons for the decline are unclear, the limited capture of minke whales during this period (a total of 207 whales taken during 2003-2007) cannot have caused a decline of this magnitude (NAMMCO 2009). This sudden change in minke whale density in coastal Icelandic waters seems more likely to represent a shift in distribution than a crash in population size. However, due to poor coverage of offshore areas, particularly to the north and west of Iceland, such a shift in distribution cannot be demonstrated from the survey data and hence a real population decrease cannot be ruled out. A potential explanation for such a shift in distribution of minke whales could be a shift in the distribution of important prey species that may be related to appreciable changes in the physical and biological environment observed in Icelandic waters in recent years (Astthorsson et al. 2007, NAMMCO 2009). Given the euryphagous nature of the minke whale, such a shift in prey species would likely be reflected in the diet of minke whales. This paper presents a study on the diet composition of minke whales in Icelandic coastal areas. The results are discussed in connection with the limited previously available data, spatial variation and in relation to recent changes in the marine environment of Icelandic coastal waters.

I.2   Material and methods

I.2.1   Sampling

The minke whale material used in this study was collected during 2003-2007 as a part of a comprehensive research programme (MRI 2003). The whales were caught using five minke whale catching boats hired by the Marine Research Institute (MRI). Cruise leaders from the MRI were in charge of the operation onboard each vessel and up to four scientist in total. In most cases, dissection and sampling took place onboard the vessels within three hours post mortem. To ensure representative sampling, searching effort was distributed temporally

and spatially in proportion to densities of minke whales as known from previous surveys (Borchers et al. 2009). The sampling was designed according to an established delineation of the Icelandic continental shelf area into nine sub-areas (Figure I.1), as defined by Stefánsson and Pálsson (1997a). This division is based on ecological and hydrographical properties of Icelandic waters which are characterized by the meeting of warm Atlantic water masses from the south and Polar waters (Subarctic water) from the north (Gislason et al. 2009). Off southern and western Iceland, the warm Atlantic water masses are dominant while the northern and eastern waters are characterized by a mixture of Atlantic and Polar water. In the analysis, the nine sub-areas were merged into South and North regions to reflect these hydrographical conditions (Figure I.1). To avoid selective sampling exhaustive attempts were made to catch the first whale sighted within a given subarea and period. Whales were taken onboard immediately after death for detailed dissection and sampling for the various sub-projects (MRI 2003).

I.2.2   Analysis of stomach contents

Based on comparison between contents of different stomach compartments, Lindstrøm et al. (1997) concluded that the fore-stomach contents are sufficient to describe the diet of minke whales. Therefore this study is confined to the analysis of the fore-stomach. The contents of the stomach were transferred to a special container for measuring the total volume. Preliminary identification of dominant prey species and identification of stomach fill was conducted at sea by visual inspection. Then the liquid was removed by a series of sieves (20mm, 5mm and 1mm) separating differently sized particles as described by Haug et al. (1995). Sub-samples were taken from stomachs with large contents of uniform prey and preserved frozen (see Haug et al. 1996) while stomachs with diverse contents were sampled completely. In the laboratory, food remains were identified to the lowest possible taxonomic level. For undigested fish prey, up to 50 measurements were made of length and weight for each species and the age determined from growth zones in sagittal otolith sections . Most of the stomach contents were, however, too digested for such measurements. In these cases, otoliths were used for species identification, counted and then measured under a stereo-microscope with a micrometer eyepiece.

   In addition to whole undigested fish, the number of fish in each stomach was calculated by dividing the total number of sagittal otoliths by two or, in the case of larger otoliths, by counting the number of otolith pairs. It was not possible to estimate the number of highly digested small crustaceans. For this

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
54   change?
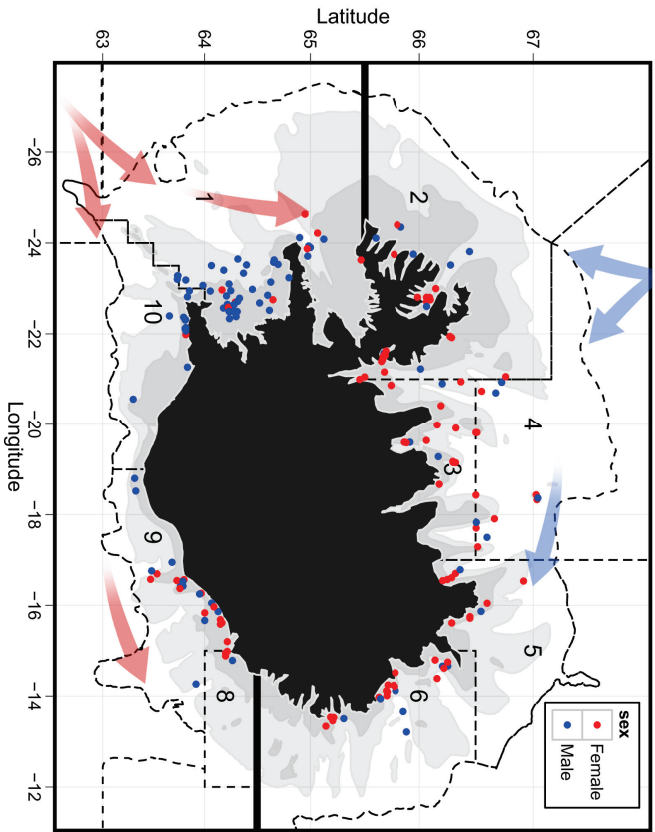
Figure I.1: Catch distribution of the minke whales caught in 2003 – 2007. In the analysis areas 2 – 6 were combined to form the Northern Area and the remaining subareas (1, 8, 9, 10) are referred to as the Southern Area. Depth is indicated as follows: dark grey: 0 – 100 m; light grey: 101 – 200 m; white: more than 200 m. Red and blue arrows indicate warm and cold water masses, respectively.

group, the measured volume was used to estimate it's contribution to the diet as done in previous studies (Tamura and Fujise 2002, Windsland et al. 2007). The stomach volume and weight was analysed using a generalized linear model with a Gamma link function as described by Stefánsson and Pálsson (1997b). The parameters considered were area, year, month, sex and maturity. The best model was selected on the basis of the Akaike Information Criterion (AIC), as discussed in Akaike (1974), by stepwise removal of the above parameters.

The results are expressed as frequency of occurrence (FO) and reconstructed weight (RW), feeding indices traditionally used in analyses of stomach content (Hyslop 1980, Pierce et al. 1991, Víkingsson et al. 2003). In addition a weighted frequency index (WFO) was used for comparison with a previous Icelandic study (Sigurjónsson et al. 2000). Comparisons between major prey groups with respect to digestion (undigested versus states of digestion) and comparisons of whether one or more prey species were found in the stomachs varied based on geography was made using a χ^2-test.

To estimate the uncertainty of prey contribution to the overall diet an adapted form of the traditional bootstrap procedure Efron and Tibshirani (1993) was used. When dealing with data originating from large animals, such as minke whales, considerable care must be taken when defining the proper sampling unit. For reasons stated in Pennington and Volstad (1994a), i.e. the "intra-haul correlation", and further discussed in Hrafnkelsson and Stefánsson (2004) individual prey items can not be treated as the sampling unit. To account for this correlation within a whale stomach individual whales are treated as sampling units, thus augmenting the traditional bootstrap Efron and Tibshirani (1993). Similar approaches have been implemented by Hannesson et al. (2008) where whole tagging experiments were treated as sampling units. Individual whales are considered sufficiently uncorrelated due to rules applied to spread the catch in time and space. To maintain geographical and temporal structure the resampling was stratified by year and the North-South division (Figure I.1).

For calculating fish size from the size of sagittal otoliths, formulae based on Icelandic published (Vilhjálmsson 1994, Víkingsson et al. 2003) and unpublished measurements were applied if available. In the absence of Icelandic data, formulae from other areas Härkönen (1986), were used. Otoliths that were noticeably eroded by digestion, corresponding to digestion state 3 and above (Recchia and Read 1989), were not used in calculations of fish size. Some stomachs containing large amounts (>50) of otoliths were subsampled for measurements and in few cases also for counting of prey items. Further details on laboratory techniques and methods are given in Víkingsson et al. (2003).

Length dependent prey selection was analysed using a linear regression model

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
56   change?

for prey length where prey species within individual whales were treated as a
random effect. The best model was selected, as in the case of stomach volume,
on the basis of AIC.

I.3   Results

The spatial and temporal distribution of the sampling is shown in Figure I.1
and Table I.1 respectively. A total of 190 minke whales (96 males and 94 fe-
males) sampled during April-September 2003-2007 were examined with respect
to stomach contents. Total volume and weight of stomach contents Most (97.4
%) of the stomachs contained some food remains. The median and mean vol-
ume of the unfiltered fore-stomach contents was 32 litres (l) (95% quantiles: 0
– 132.3) and 42 l respectively with a maximum of 200 l in an 8.1 m long female.
The median and mean weight of stomach contents after removal of fluid was
7.82 kg (95% quantiles: 0.03 – 79.5) and 14.83 kg respectively with a maximum
weight of 106.25 kg in a 7.85 m long male. The best models of stomach vol-
ume and weight, in terms of AIC, were both based on the whale length and
geographic area (Table I.2). Other variables such as sex, month, year and time
of the day did not offer a significant improvement to the model. Two sampling
areas, 2 and 10 (see Figure I.1) were significantly lower than the other seven in
terms of weight of stomach contents and whales caught in areas 2, 5, 9 and 10
had significantly lower stomach volumes than those in other areas (Table I.3).

I.3.1   Overall diet composition

In total 14 prey types were found in the stomachs including 10 species of fish and
2 species of euphausids (Table I.4). In addition two species of crabs (*Pagurus
bernhardus* and *Hyas coarctatus*) found in one stomach each were considered
secondary prey ("prey of prey") and thus not included in further analyses. Most
(56%) of the stomachs contained only one prey species while the maximum
number of species in a single stomach was six.

Sandeel was the most frequently encountered prey, occurring in 110 (58%)
of the stomachs. It was not possible to distinguish between the three species of
sandeel occurring in Icelandic waters. While it seems likely that the bulk of the
sandeel prey is *Ammodytes marinus* (Raitt, 1934), the overwhelmingly dominant
species in these waters (V. and K. 2009), the large size of some individuals
shows that Hyperoplus lanceolatus was also included to some extent. Other
frequently occurring prey types include haddock (*Melanogrammus aeglefinus*

Table I.1: Temporal distribution of the sampled common minke whales in 2003–2007. The areas (North and South) are shown in Figure I.1

	Area	Apr	May	Jun	Jul	Aug	Sep	Total
2003	North	-	-	-	-	8	7	15
2003	South	-	-	-	-	10	11	21
2003	Total	-	-	-	-	18	18	36
2004	North	-	-	6	1	-	-	7
2004	South	-	-	10	8	-	-	18
2004	Total	-	-	16	9	-	-	25
2005	North	-	-	-	8	7	-	15
2005	South	-	-	-	10	9	-	19
2005	Total	-	-	-	18	16	-	34
2006	North	-	-	10	15	6	-	31
2006	South	-	-	3	10	14	-	27
2006	Total	-	-	13	25	20	-	58
2007	North	3	3	16	-	4	-	26
2007	South	-	4	6	-	-	1	11
2007	Total	3	7	22	-	4	1	37
Grand total		3	7	51	52	58	19	190

Table I.2: Overview of AIC scores for the fitted GLM models for weights and volume of stomach contents. Full model includes the variables whale length, geographic area, month, year, sex and maturity. D.f.: Degrees of freedom.

Step	D.f.	Deviance	Resid. D.f.	Resid. Dev.	AIC
Weight of stomach contents					
Full model			169.00	329.70	1388.87
- month	5.00	0.80	174.00	329.50	1379.44
- year	4.00	5.32	178.00	334.82	1375.21
- sex	1.00	0.20	179.00	335.02	1373.35
Volume of stomach contents					
Full model			160.00	242.40	1815.17
- year:month	5.00	1.39	165.00	243.80	1806.46
- month	5.00	1.86	170.00	245.66	1798.17
- year	4.00	2.17	174.00	247.84	1792.16
- maturity	4.00	2.38	178.00	250.22	1786.32
- sex	1.00	0.33	179.00	250.55	1784.62

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
58   change?

Table I.3: Summary of the variables of the best model, in terms of AIC, for the
weight and volume of stomach contents. The intercept value represents the model fit
for length 0 whales in Area , the t-value and the corresponding p value represent the
results from a Wald's test for a single parameter difference from zero.

	Estimate	Std. Error	t value	
Weight of stomach contents				
(Intercept)	0.1853	0.0741	2.50	0.0133
Length	-0.0001	0.0001	-1.38	0.1681
Location:				
Area 2	-0.0379	0.0189	-2.00	0.0466
Area 3	0.0177	0.0377	0.47	0.6389
Area 4	-0.0043	0.0314	-0.14	0.8916
Area 5	0.0099	0.0416	0.24	0.8122
Area 6	0.0404	0.0371	1.09	0.2771
Area 8	0.0566	0.0996	0.57	0.5710
Area 9	-0.0281	0.0209	-1.34	0.1813
Area 10	-0.0442	0.0200	-2.21	0.0285
Volume of stomach contents				
(Intercept)	0.1110	0.0176	6.30	0.0000
Length	-0.0001	0.0000	-4.51	0.0000
Location:				
Area 2	-0.0139	0.0045	-3.11	0.0022
Area 3	0.0002	0.0075	0.03	0.9779
Area 4	-0.0082	0.0063	-1.31	0.1928
Area 5	-0.0146	0.0061	-2.41	0.0172
Area 6	-0.0059	0.0058	-1.02	0.3113
Area 8	0.0102	0.0173	0.59	0.5554
Area 9	-0.0125	0.0048	-2.62	0.0097
Area 10	-0.0146	0.0050	-2.92	0.0039

Table I.4: Identified prey species of minke whales and the number of stomachs in which they occurred stratified by sampling year and area. N: Northern areas S: Southern areas (see text). n: sample size.

Year:	2003		2004		2005		2006		2007		All years	
Area:	N	S	N	S	N	S	N	S	N	S	N	S
Sample size:	15	21	7	18	15	19	31	27	26	11	94	96
Crustacea												
Euphausiacea												
Thysanoessa rasci	1		1								14	7
Meganychtiphanes norvegica	2										5	
Krill (unidentified)	3			1	1	4	3		3	1	13	2
Copepoda (unidentified)	1										1	
Decapoda												
Pandalus borealis					1						1	
Vertebrata												
Gadidae												
Gadus morhua	3	1	2	2	4	1	8	3	4		21	7
Melanogrammus aeglefinus				1	7	4	10	8	10	3	27	16
Pollachius virens					2			1	1		3	1
Merlangius merlangus	1		1		3	1	1	3	1		7	4
Micromesistius poutassou	1	1			1	1	1	1			3	3
Trisopterus esmarki	1										1	
Unidentified. large fish	1				1	2	1		4	1	7	2
Clupeidae												
Clupea harengus			1		1	2	6	11	8	9	16	22
Osmeridae												
Mallotus villosus	4		3		2	1	17	2	7	1	33	4
Ammodytidae												
Ammodytes marinus/ Hyperoplus lanceolatus	9	20	2	17	6	15	13	21	5	2	35	75

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
60   change?
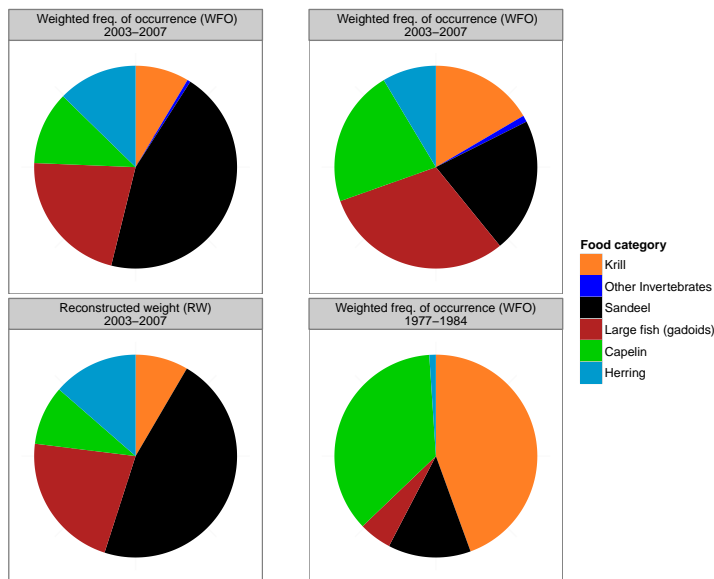
Figure I.2: (a) Diet composition of minke whales sampled in Icelandic waters 2003–2007 shown as weighted frequency of occurrence (WFO) and as reconstructed weight (RW) for all areas combined. (b) Diet composition of minke whales in terms of WFO for the northern area (see Figure I.1) and the period June–July only for comparability with previous data (1977–1984).

Linnaeus, 1758), herring, capelin, cod and krill found in 23%, 20%, 19%, 15% and 14% of the examined stomachs respectively (Table I.4).

Figure I.2a shows the diet composition in terms of weighted frequency of occurrence (WFO) and reconstructed weight (RW) of different prey species pooled for the research period 2003-2007 The two measures of diet composition (WFO and RW) show remarkably similar results (Figure I.2a). The diet was primarily composed of fish, with krill contributing only 8.6% in terms of WFO and 8.4% in terms of RW to the diet. Sandeel was the single most important prey type overall with 45% and 47% prevalence in terms of WFO and RW, respectively. Other common prey species were herring, capelin, haddock and cod. Together, large demersal fish (gadoids) constituted 22% of the diet respectively according to both these measures.

I.3.2   Geographical and temporal variation

The previously available information on the diet of minke whales in Icelandic
waters is very limited as regards total sample size as well as temporal and spatial
coverage. Thus, of the 68 animals reported by Sigurjónsson et al (2000) for the
period 1977-1997, 50 were sampled off North Iceland during June and July 1977-
1984. Comparisons between these two periods are therefore restricted to this
area and months. In Figure I.2b the results for June-July off North Iceland
are compared between the two sampling periods 1977-1984 and 2003-2007. In
the former period, krill and capelin were the dominant prey species, together
constituting 81% of the diet (Figure I.2b). These two prey types accounted for
38% in the latter period, where gadoids (30%), sandeel (22%) and herring (9%)
contributed appreciably more to the diet than during 1977-1984.

Figure I.3 shows the diet composition in terms of FO (Figure I.3a) and RW
(Figure I.3b) by year and area for the period 2003-2007. These two measures
show generally similar results although invertebrates have a somewhat higher
score in terms of FO. The diet composition varied considerably with geographic
location. Sandeel dominated the diet in the southern and western areas, while
the diet was more diverse off northern and eastern Iceland. The dominance of
sandeel was particularly striking in the beginning (2003) when its proportion
of the diet in terms of reconstructed weight amounted to 94 and 47 % in the
southern and northern areas respectively (Figure I.3b). This proportion de-
creased steadily throughout the research period in both areas and was 18 and
12 % respectively in 2007. Of the pelagic fish prey, capelin appeared to form a
steady part of the diet in the northern area throughout the observation period
while being hardly detected in the southern area (Figure I.3). The importance
of herring in the diet however increased from being non-existent in the stomachs
from 2003 to constituting nearly 80% of the diet in the southern area during
2007 (Figure I.3). Figure I.5 shows the contribution of gadoid species (and
other large bony fishes) in more detail. Overall, haddock and cod were the
most important gadoids, contributing 10% and 8% of the reconstructed weight
respectively. Cod consumption was largely confined to the northern area where
it constituted up to 29% of the diet. Haddock was only a minor component
of the diet in 2003-2004, but thereafter it constituted an important part of the
diet in both areas (Figure I.5). Other gadoid species such as saithe (*Pollachius
virens* (Linnaeus, 1758)), whiting (*Merlangius merlangus* Linnaeus, 1758) and
blue whiting (*Micromesistius poutassou* Risso, 1827) appeared in the diet to a
lesser extent.

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
change?

62

Figure I.3: Boxplots of spatial and temporal variation in diet composition of minke whales expressed as (a) frequency of occurrence (FO) and (b) reconstructed weight (RW). The point estimate is indicated by a white star while the solid grey lines through the boxes represent the bootstrapped medians. The box indicates the interquartile range and the whiskers 95% confidence intervals. Any further outlying data points are indicated as points. Area division (North–South) as in Figure I.1
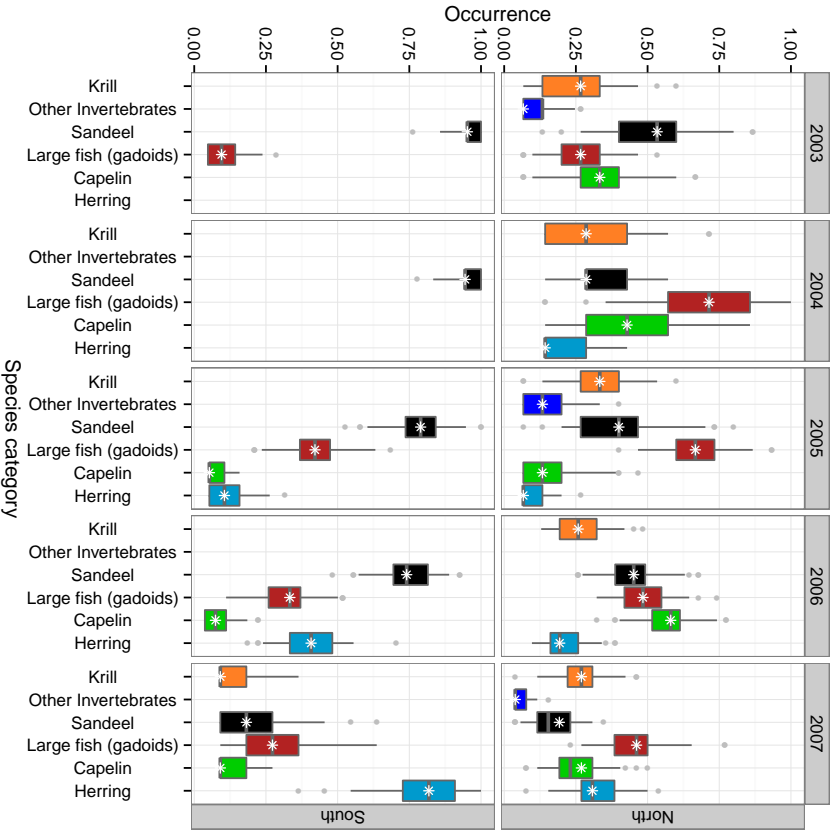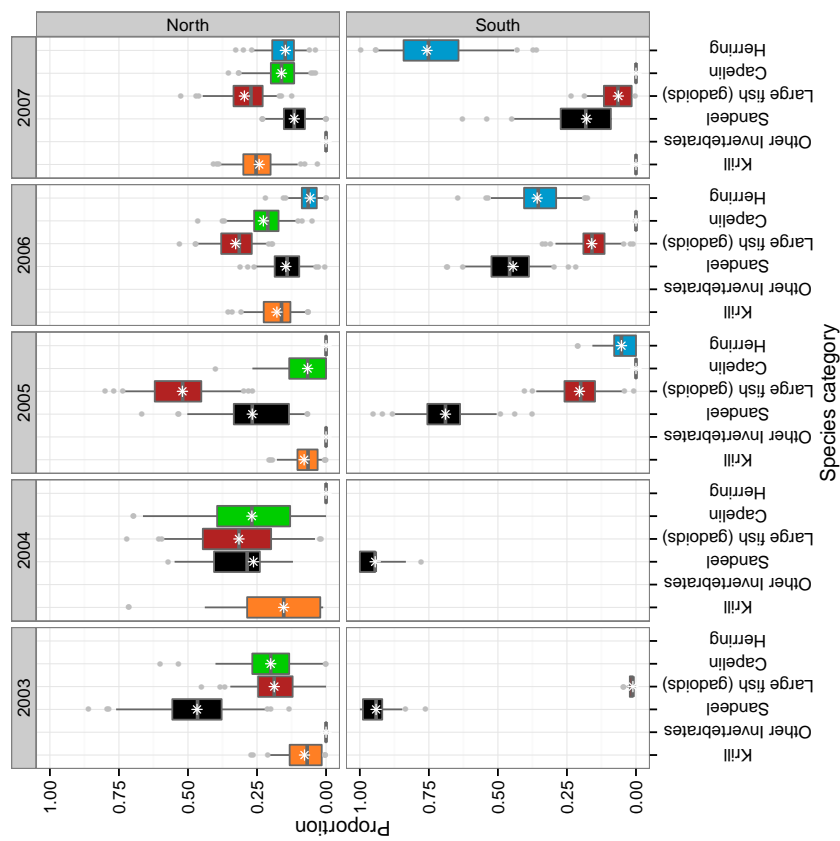
Figure I.4: Fig I.3 continues

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
64   change?
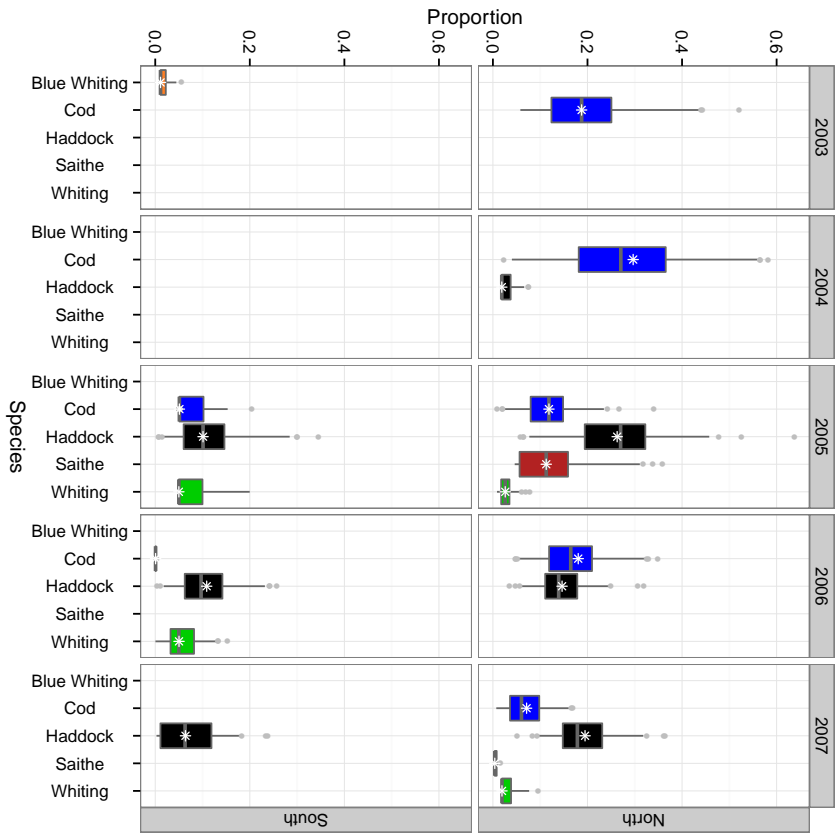
Figure I.5: Boxplot of spatial and temporal variation in the contribution of the most important gadoid fish species to the diet composition of minke whales (RW). Boxplot parameters as in Figure I.3 and area division (North–South) as in Figure I.1.

I.3.3 Prey length and age

The reconstructed size range of fish prey for the minke whale was 1 - 92 cm in total length (Figure I.6). Among the fish species taken in considerable numbers, sandeel and capelin had the lowest mean length of 12.8 and 9.3 cm respectively, although individuals up to 36.7 and 16.2 cm were found. All age classes (0 to 4 years) of capelin were consumed by the minke whales, although a large majority was less than three years old (Figure I.7). The age of sandeel ranged between 0 and 6 years with year classes 0-3 dominating in the diet. The herring consumed varied from 0 to 41 cm in length and 1-14 years in age. The size range for cod was 0 to 92.51 cm with a mean size of 61.99 cm. The age of cod prey ranged from 0 to 14, with a mean of 6 years. The mean length of haddock was 41.63 cm, ranging from 2.6 to 91.78 cm, and mean age was 3.9 years, ranging from 0 to 9 years.

I.4 Discussion

I.4.1 Potential sources of bias

Minor sampling constraints were imposed by exclusion from whale watching areas and generally worse weather in offshore areas may have caused some sampling bias. However, the limited search effort that was possible in offshore areas, together with information from a simultaneous sightings survey, as well as information from fishing vessels further offshore indicated that the distribution of minke whales in the area was indeed, very coastal during this period. Therefore, we conclude that the sampling distribution was representative of the distribution of minke whales within the Icelandic continental shelf area. Differential digestion rates of different prey items is a potential source of error in stomach contents analysis (Hyslop 1980, Pierce et al. 1991). For example the proportion of small planktonic crustaceans in the diet may be underestimated due to higher rates of digestion (Gannon 1976). One way to address this problem is to compare the overall results to those of a subsample of stomachs containing only virtually undigested food. Comparison of stomachs containing fresh prey (n=51) with those containing digested prey (n=177) did not reveal significant differences in diet composition ($\chi^2(4) = 6.5967$, p=0.1588). This indicates that differential digestion rates of prey species did not introduce any appreciable bias in the study although some underrepresentation of krill cannot be ruled out. The close similarity in the results using different measures (RW or WF) for diet composition also indicates absence of significant bias due to differential

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
66   change?
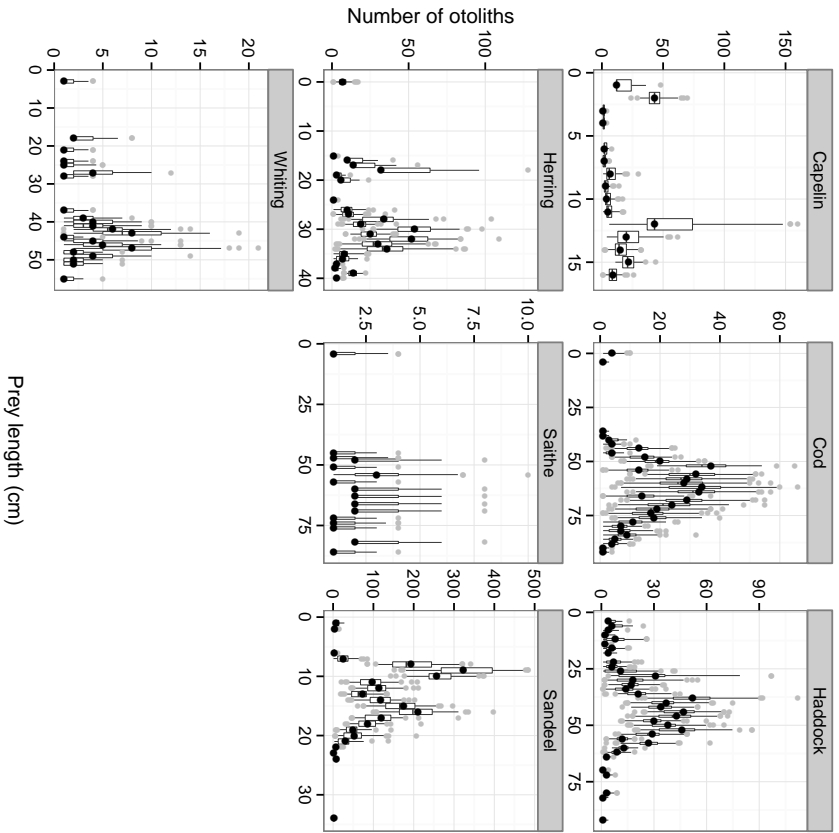
Figure I.6: Boxplots of length distributions of the most common fish species identified in the fore-stomachs of minke whales sampled in 2003–2007 according to direct measurements (undigested fish) and calculations from otolith lengths. The point estimate is indicated by a black dot and other boxplot parameters as in Figure I.3.
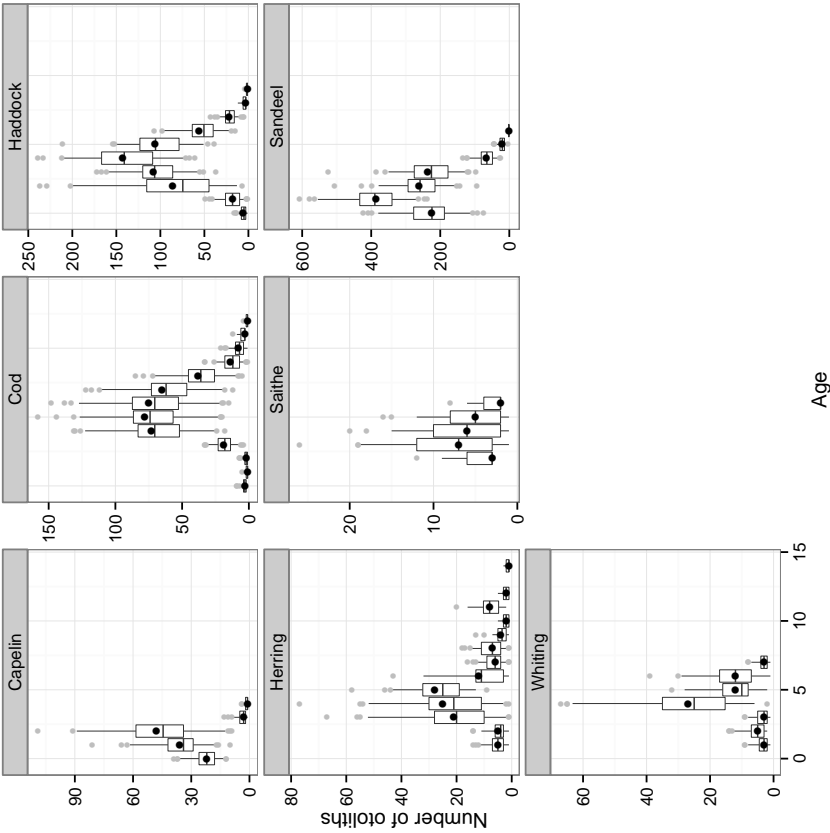
Figure I.7: Boxplots of age distributions of the most common fish species identified in the fore-stomachs of minke whales sampled in 2003 2007. Boxplot parameters as in Figure I.6.

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
68   change?

Table I.5: Best model for prey lengths. The Intercept term denotes the amount consumed by immature whales in Area 1.

	Estimate	Std. Error	t value
(Intercept)	-4.05	2.43	-1.67
Species:			
Capelin	11.32	2.48	4.56
Cod	60.04	3.94	15.25
Haddock	40.42	3.00	13.46
Herring	31.88	3.09	10.30
Saithe	58.21	9.94	5.85
Sandeel	15.67	2.31	6.79
Whiting	45.36	3.28	13.81
Maturity:			
Mature	0.98	0.94	1.04
Pregnant	0.22	1.05	0.21
Location:			
Area 2	-1.36	1.36	-1.00
Area 3	-0.15	1.50	-0.10
Area 4	-1.58	1.47	-1.07
Area 5	-0.88	1.93	-0.45
Area 6	-2.31	1.22	-1.90
Area 8	0.34	2.12	0.16
Area 9	3.60	0.97	3.72
Area 10	-0.54	1.33	-0.40

digestion rates. This similarity also indicates that WF, a rather crude index of diet composition applied because of the limitations of previously available data, might be a good alternative to the more labour intensive RW for monitoring of diet. However, further validations are required to evaluate the general applicability of the WF index.

I.4.2 Prey quantities and diversity

The median volume of stomach contents in this study is within the range reported off Norway (Nordøy and Blix 1992) and the small proportion of empty stomachs is similar to studies of minke whales from other areas (Mitchell 1974, Larsen and Kapel 1981, Kasamatsu and Hata 1985, Haug et al. 1997, Tamura and Fujise 2002). The maximum amount of food remains retrieved from a single fore stomach (106 kg) is to our knowledge the highest reported for this species. Not unexpectedly, length of the whale appeared to have the largest predictive power for the magnitude of stomach contents but significant geographical variation was also detected. The geographical variation might however be confounded by temporal changes in food availability as the sampling years were not evenly distributed on sampling areas.

This study confirms earlier findings on the euryphagous and predominantly piscivorous nature of northern hemisphere minke whales (Larsen and Kapel 1981, Jonsgård 1982, Tamura and Fujise 2002, Haug et al. 2002, Pierce et al. 2004, Windsland et al. 2007). While the range of species taken by minke whales was large, the majority of the whales (59%) had recently fed on only one prey species. This differed significantly according to a north-south division ($\chi^2(1) = 15.97$, $p < 0.001$), where proportion of whales with only one prey species was 70% in the south while it was 43% in the north. Low diversity of prey found in individual stomachs is typical for the so called swallowing type of baleen whales (Nemoto 1959) feeding on swarming zooplankton and schooling fish and several variations of this methods have been described for minke whales in Icelandic waters (Bertulli 2010). The prey diversity reported here is somewhat higher than that reported from Norway and Japan, where 69% and 90% respectively of the minke whales had only one prey species in their forestomachs (Tamura and Fujise 2002, Windsland et al. 2007).

I.4.3 Overall diet composition

Sandeel was the most important prey species overall contributing nearly 50% to the diet. Pelagic fish (herring and capelin) and gadoids each accounted for

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
70   change?

over 20% of the diet and krill around 8%. Dominance of sandeel in minke whale
diet has also been observed in the North Sea (Olsen and Holst 2001, Pierce
et al. 2004). Although gadoids have been reported as prey of minke whales
in several other areas of the North Atlantic (Sergeant 1963, Larsen and Kapel
1981, Horwood 1990), their contribution to the diet is usually less than that
reported here, particularly off north Iceland. However, comparable and even
higher values for gadoids have been observed in some areas off northern Norway
(Jonsgård 1982, Haug et al. 1997, Windsland et al. 2007).

According to this study, the prey of minke whales varies in size from few
millimetres (planktonic crustaceans and fish fry) to 92cm long gadoids. Similar
wide ranges have been documented for this species from Norway (Windsland
et al. 2007) where gadoids up to 85cm were found in minke whale stomachs.
Prey of this size contradicts the general idea of baleen whales feeding primarily
on planktonic crustaceans and small schooling pelagic fish species and it could
be argued that these large codfish are taken accidentally together with other
targeted prey. However, no evidence for this could be detected in the present
study. Thus in 36 % of the stomachs where gadoids were found, they constituted
the main prey and in 25 % they were the only prey group including a stomach
containing over 100 adult cods, mainly 5 to 7 years old. This clearly shows
that these large gadoids are targeted by minke whales rather than engulfed
accidentally.

The absence of 0-group herring in the diet off Iceland is in contrast to results
from Norwegian waters where they constitute an appreciable part of the diet
in some areas Haug et al. (1995). This can be explained by the difference in
spawning times of the two herring stocks (Dragesund et al. 1980, Jakobsson and
Stefánsson 1999). Off Norway, herring spawns in late winter and spring, while
in Icelandic waters herring spawns in mid summer not reaching a consumable
size until winter.

I.5   Geographical variation

The pronounced geographical variation in prey species composition is broadly
consistent with the known distribution of the prey species in Icelandic waters.
Thus, sandeel and herring are most abundant off the south, southeast and
southwest coasts (Jonsson and Palsson 1992, V. and K. 2009) while capelin
is mostly restricted to northern Icelandic waters during summer (Vilhjálmsson
1994) which is reflected in the minke whale diet composition. The gadoid species
found in the minke whale stomachs are distributed throughout Icelandic coastal

waters during summer (Jonsson and Palsson 1992). Haddock has traditionally
been most abundant off the southern and western coast of Iceland. However,
the distribution has expanded northwards in recent years (Astthorsson et al.
2007) so that present abundance is no less in North Icelandic coastal waters.
The summer distribution of cod covers all Icelandic coastal waters. Although
the species has been found at depths of over 600m it is most common at 100-
250 m (Jonsson and Palsson 1992). The relatively low proportion gadoids in
the diet in the southern areas during the first half of the period (when sandeel
was relatively abundant) could indicate a preference for sandeel when both are
available.

Higher diversity of the diet in the northern Icelandic areas is similar to that
found for harbour porpoises in Icelandic waters (Víkingsson et al. 2003). In
both these cases low prey diversity in the southern areas seems to be related
to the cetaceans utilization of the seasonally abundant high-energy fish species
capelin (porpoises in late winter) and sandeel (minke whales during summer).

I.5.1   Temporal changes in diet and the ecosystem

The present results differ considerably from those obtained in the only previ-
ous systematic study on minke whale diet in Icelandic waters conducted during
1977-1997 (Sigurjónsson et al. 2000). Although the previous data were limited,
it seems clear that significant changes have occurred in the diet of minke whales
in Icelandic waters between the two periods. The difference in sampling strate-
gies might contribute to the difference in diet composition between the present
study, which was designed to be representative of real distribution, and the
previous one (Sigurjónsson et al. 2000) which was based on commercial catches
up to 1985 and strandings. However, considering only comparable areas and
months (Figure I.2b) large differences are still apparent. Thus, in the more
recent period (2003-7), the proportion of large, benthic fish (mostly haddock
and cod) amounted to 30% (WFO) or sixfold that of the earlier period (1977-
1984). The contribution of herring to the diet had also increased from 1 to 9
% (Figure I.2b). The contribution of planktonic crustaceans (krill) and capelin
decreased from 45% to 17% and 36% to 22% respectively. Thus, between these
two sampling periods, krill, and the cold-water species capelin seem to have been
largely replaced by herring, sandeel and gadoids in the more recent period. Haug
et al. (2002) and (Tamura and Fujise 2002) described temporal dietary changes
in minke whales off Norway and Japan, respectively which they attributed to
changes in the ecosystem. The sampling scheme in the present study was not
designed to study inter-annual variation or trends and thus, the seasonal dis-

Chapter I Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
72   change?

tribution of the samples was somewhat different between years. However, this is unlikely to be a source of bias in this respect as no indications were seen from the data of within-season variability in the diet. The results of this study, therefore suggest that in recent years appreciable changes have occurred in the diet composition of minke whales in Icelandic coastal waters. Explanations of these changes may be found in other components of the ecosystem. Iceland is located at a dynamic front between warmer and more saline Atlantic water masses coming from the south and cold Polar water masses originating in the Arctic Ocean. While the area is in general highly productive, there is also considerable variability in environmental conditions particularly over the shelf north of Iceland where the Polar Front separates the contrasting Atlantic and Polar water masses. During the past decade, environmental conditions in Icelandic waters have been characterized by unusually high temperatures (surface and bottom) and salinity (MRI 2012a). This warming of the marine environment during the last decade appears to have led to pronounced changes in distribution and abundance of several fish species in Icelandic waters. Thus, southern gadoids such as saithe, whiting and haddock as well as monkfish (Lophius piscatorius Linnaeus, 1758) have shown increase in abundance and extended distribution to the waters north of Iceland Astthorsson et al. (2007), Solmundsson et al. (2010). Corresponding changes have also occurred for pelagic fish with the advancement of warm water species such as mackerel (Scomber scombrus Linnaeus, 1758) and blue whiting into Icelandic waters (MRI 2012b). Haddock is noteworthy in this respect, as in addition to the northward expansion of the species distribution, the estimated biomass in Icelandic waters increased threefold between 2000 and 2007 (MRI 2012b). This increased availability of haddock was clearly reflected in the diet of minke whales where the species hardly occurred during 2003-2004 but constituted an appreciable part of the diet in 2005-2007 (Figure I.5). Concomitant with these environmental changes, the cold-water species capelin appears to have retreated to a significant extent from Icelandic continental waters. Capelin was major component of the minke whale diet according to the samples collected during 1985-1997, but its contribution to the diet was much less during 2003-7. This is in accordance with changes in distribution and abundance of capelin in Icelandic and adjacent waters according to fish surveys in recent years (Astthorsson et al. 2007, MRI 2012b, Pálsson et al. 2012). The summer distribution has shifted away from the Icelandic continental shelf towards north and to the coast of East Greenland. Unfortunately, the coverage of this area in the most recent cetacean sightings survey (TNASS) was very poor (Pike et al. 2009) and insufficient to test the hypothesis whether minke whales have, to some extent followed this retreat of capelin from the shelf area.

   Sandeel was the single most important prey type overall during 2003-7. It also constituted a significant proportion of the diet according to the pre-2000 data where it was probably underestimated because of skewed sampling towards the northern areas. Unfortunately, specific research into sandeel abundance in Icelandic waters was not initiated until 2006. Indirect evidence (fish stomach samples, seabird breeding success) had then suggested a decline in abundance over the previous two decades, though with large inter-annual variations. The proportion of sandeel in the diet of minke whales decreased markedly during the sampling period (2003-2007) as did the number of sandeel detected in the stomachs of cod and haddock according to MRI's groundfish surveys (V. and K. 2009, Pálsson and Björnsson 2011). Available data on minke whale and sandeel abundance in Faxaflói, SW Iceland, indicate that there might be a connection between the densities of the two species, although the nature of the sandeel data does not allow a statistical comparison. Thus, up to around 2002-3, densities of both species were considerably higher than in subsequent years and fluctuations in recent years appear to be synchronous. The variation in sandeel abundance, has been particularly pronounced in Faxaflói Bay where the highest values have been obtained (V. and K. 2009). Abundance was particularly low in Faxaflói in 2007, considerably higher in 2008, but decreased again in 2009. Aerial surveys conducted in Faxaflói Bay in 2007, 2008 and 2009 Pike et al. (2011a), show similar patterns with considerably higher densities of minke whales in 2008 than in 2007 and 2009. Also, the decline to very low levels of sandeel abundance in Ingólfshöfði, SE Iceland, could explain the drastic decline in minke whale abundance in this area in the recent surveys (2007-2009). The Icelandic summer spawning stock of herring was growing rapidly during the sampling period mainly due to strong year classes in 1999 and 2000 (MRI 2012b). In addition, the herring fishery season during the winter 2006/7 was unusual in that most of the catch was taken off the south coast MRI (2007) as opposed to the east coast in previous years and west coast in later years MRI (2012b). This indicates high abundance of herring off south Iceland in 2006-2007, which appears to be reflected in the diet composition of minke whales.
   The summer distribution and abundance of two species of krill in northern and eastern Icelandic waters in during 1995-2007 was described by Skúladóttir et al. (2009). Although the data series does not extend back to the period of the previous minke whale study, it indicates that the more recent samplings (2003-2007) were conducted during a period of relatively low krill abundance (Skúladóttir et al. 2009). The krill distribution pattern seems to reflect whale densities to some extent. In particular, high abundance of humpback whales in the 1995 and 2001 surveys (Paxton et al. 2009) might be explained by krill

Chapter I   Recent changes in the diet composition of common minke whales
(*Balaenoptera acutorostrata*) in Icelandic waters – A consequence of climate
74   change?

abundance in those areas. Two other predominantly krill eating cetaceans have
changed distribution around Iceland in recent years. A recent northward shift
in distribution has been observed for blue whales (Balaenoptera musculus Lin-
naeus, 1758) in Icelandic coastal waters (MRI unpublished data) and a change
in distribution of fin whales (Balaenoptera physalus Linnaeus, 1758) has been
attributed to a change in enviromental conditions in the Irminger Sea (Víkings-
son and Valdimarsson 2006). In general, the krill densities were considerably
lower in 2007 than in 2001 and 1995 which may help explain its decreased im-
portance in the diet and possibly also the decreased minke whale abundance in
2007.

I.6   Conclusion

After the turn of the century, large changes have occurred in the Icelandic conti-
nental shelf ecosystem. Long-term studies of trophic patterns of cod in Icelandic
waters have indicated lowered growth rates in recent years due to shortage of
the cod's principal prey capelin and northern shrimp (Pandalus borealis Krøyer,
1838) (Pálsson and Björnsson 2011). Many of these changes appear also to be
unfavourable to minke whales, notably less abundance of important prey species
such as sandeel, capelin and krill. Although there is evidence of minke whales
having reacted to this by switching to other prey species, s.a. haddock, cod
and herring, minke whale abundance has also decreased significantly during
this period. Detailed examinations of body condition of 190 minke whales sam-
pled during 2003-2007 have not revealed any instances of severe malnutrition
(Christiansen et al. 2013). Such instances would be expected if the decline in
abundance was due to a real population decline, as a result of food shortage.
Thus, it seems more likely that the decline is the result of a shift in distribu-
tion from Icelandic coastal waters as a result of a decline in the availability of
preferred food types.

I.7   Acknowledgements

II

Paper II

A bootstrap method for estimating bias and variance in statistical fisheries modelling frameworks using highly disparate data sets

Bjarki Þór Elvarsson, Lorna Taylor, Verena M. Trenkel, Vojtech Kupca and Gunnar Stefánsson

Abstract

Statistical models of marine ecosystems use a variety of data sources
to estimate parameters using composite or weighted likelihood functions
with associated weighting issues and questions on how to obtain variance
estimates. Regardless of the method used to obtain point estimates, a
method is required for variance estimation. A bootstrap technique is in-
troduced for the evaluation of uncertainty in such models, taking into
account inherent spatial and temporal correlations in the data sets, which
are commonly transferred as assumptions from a likelihood estimation
procedure into Hessian–based variance estimation procedures. The tech-
nique is demonstrated on a real data set and the effects of the number
of bootstrap samples on estimation bias and variance estimates studied.
Although the modelling framework and bootstrap method can be applied
to multispecies and multiarea models, for clarity the case study described
is of a single species and single area model.

II.1 Introduction

Statistical models consolidate data from various sources by using them simultaneously to estimate parameters. The importance of using all data in a single model has been emphasised by several authors (Demyanov et al. 2006, Methot 1989) and although the benefits are clear, it is certainly not without problems, including the question of variance estimation, model mis-specification and weighting of all data sources (Francis 2011, Stefansson 2003, Maunder and Punt 2012). In the context of complex population dynamics models of exploited marine species, multiple data sources with widely different properties are routinely used in the estimation process.

Variance estimates of parameters in nonlinear models have commonly been derived from the inverted Hessian matrix at the optimum, when the method of least squares (or maximum likelihood) is employed for parameter estimation. Alternatively the Jacobian matrix of the residuals can be used. Several conditions need to be satisfied for statistical inference, e.g. confidence statements to hold in the finite-sample case. First, the model needs to be correct. Second, variance assumptions i.e. homoscedasticity and knowledge of the ratios of variances in individual data sets, need to be appropriate.

Methods of estimating variances in fish stock assessment models have been discussed and evaluated by many authors including Gavaris et al. (2000), Gavaris and Ianelli (2001), Magnusson et al. (2012) and Patterson et al. (2001). When the distributional properties of the data are not well understood or the models are incorrect, Hessian–based approaches have been seen to fail in several examples in fishery science (Patterson et al. 2001). Although this may seem to contradict the theoretical statements, the assumptions e.g. in Jennrich (1969) include independence of observations, a unique minimum, identically distributed errors and of course the results are only asymptotic. Any of these assumptions may fail. It follows that for problems in fishery science one cannot assume a priori that a Hessian–based method will give reasonable results. For example, disregarding correlation structure when present has been found to potentially lead to incorrect conclusions in single–species assessments, sometimes with serious consequences (Myers and Cadigan 1995). Similarly, multimodal likelihood functions have been seen in real applications (Richards 1991) and typically correspond to incorrect model assumptions that are not detected with traditional analysis (Stefansson 2003) but may potentially be detected if histograms of bootstrap parameter estimates also become multimodal (see example in Hannesson et al. 2009).

Many of the limitations of the Hessian–based approaches have been met

by alternative methods. In particular models developed using the Bayesian framework (as discussed in eg. Punt and Hilborn 1997) provide an elegant formulation of uncertainty as posterior distributions of the quantity of interest. In all but trivial cases the posterior distribution must be estimated numerically with methods such as Markov chain – Monte Carlo. With the commoditization of computers in conjunction with the development of frameworks such as BUGS (Spiegelhalter et al. 1996) and ADMB (Fournier et al. 2012), the Bayesian framework has become popular alternative to Hessian–based uncertainty methods. The attraction of the Bayes inference stems, to some degree, from the ability to include prior belief/knowledge into the model as explicit distributions. Various sources (eg. Chen et al. 2000, Millar 2002) suggest, however, that considerable care must be taken when choosing model priors to avoid misspecification and suggest a suite of robust priors applicable in fisheries model setting.

Alternative frequentist approaches to Hessian–based parameter variance estimation include bootstrap methods (Efron 1979, Efron and Tibshirani 1994). The simplest bootstrap method assumes that the data are independent measurements without correlation. However, semi–parametric approaches have also been developed to sample residuals from a model, possibly from a distribution (parametric bootstrap) or with a known correlation structure (Davison and Hinkley 1997).

This paper demonstrates a novel use of bootstrapping to address complex and disparate data issues. The approach is generic, but it has special application to statistical models of (multiple and interacting) marine populations such as those developed within the Gadget framework. Gadget is a statistical age–length structured modelling environment originally proposed by Stefansson and Palsson (1998), combining concepts from several earlier methods (Gavaris 1988b, Methot 1989, Tjelmeland and Bogstad 1989, Bogstad et al. 1992), described in Begley (2004) and subsequently used in multiple fisheries applications (e.g. Björnsson and Sigurdsson 2003, Taylor et al. 2007, Lindstrøm et al. 2009). The protocol used in Gadget to estimate likelihood component weights and optimise model parameters is described in detail in Taylor et al. (2007) and the weighting protocol is based on that described in Stefansson (1998) and Stefansson (2003).

In the following sections the development of an elementary sampling unit used in the bootstrap is described. The methodology is applied to a Gadget model for cod in Icelandic waters (the standard model from Taylor et al. 2007) and contrasted to a more traditional Hessian–based approximation of variance.

II.2   Development of an elementary sampling unit

Statistical fisheries models may involve the use of a large number of data from
a variety of sources. Every sample from each data source can be classified ac-
cording to sampling location and time. A model such as Gadget operates on
certain time-steps and also uses some spatial units. Within any modelled spatio–
temporal unit there will normally be several data samples. For any bootstrap
method the first question is therefore what the sampling unit should be. A unit
of measurement in marine studies tends to be based on a single fish and elemen-
tary resampling might bootstrap on individual fish (as in e.g. Gudmundsdóttir
et al. 1988). Doing this assumes that all individually measured fish are inde-
pendent which is invalid for several reasons (Hrafnkelsson and Stefansson 2004,
Pennington and Volstad 1994b). Resampling entire fish samples (as is done by
Singh et al. 2011) can potentially be used to account for this intra–haul correla-
tion. Appropriate analyses of variance can correspondingly be used to evaluate
these effects (Helle and Pennington 2004, De Croos and Stefansson 2011) and
when combining samples, alternatives to simple sums or means may be needed
for aggregation (Babak et al. 2007). However, considering samples as units may
not be quite enough, since fish at close geographic locations will also tend to be
similar due to a fine–scale spatial structure which can not be easily modelled
(e.g. Stefánsson and Pálsson 1997b).

   In addition to the sampling unit problem, one needs to take into account
the variety of data sources. Biological samples from commercial catches may
be collected on a fine temporal and spatial scale whereas scientific surveys are
typically only conducted once or twice a year and different surveys may or may
not overlap spatially. Other data sets such as species composition of stom-
ach contents or tagging experiments may be collected at completely different
resolutions to age or length data.

   Here the proposed sampling unit is based on spatial structure on the Ice-
landic coastal shelf developed by Taylor (2003), shown in figure II.1 where the
area within the gridlines are referred to as **subdivisions**. The spatial struc-
ture is based mainly on bathymetry, hydrography and species assemblages with
some further disaggregation defined by fishing regulations. In this context an
elementary sampling unit is all data collected inside a subdivision within a
time period of interest. In this context subdivisions and elementary sampling
units are therefore used interchangeably. In order to reduce correlations between
the elementary sampling units aggregations are made. For example, to remove
within–sample correlations between length groups (Hrafnkelsson and Stefansson
2004), only (combinations of) entire length samples are used, rather than lengths

of individual fish. Similarly, data are aggregated within the fairly large spatial areas and the shortest time–step is at least one month. This should eliminate intra–haul correlations (Pennington and Volstad 1994b) and those correlations between age–groups (Myers and Cadigan 1995) which are related to local shoals or small feeding patches.

To generate input files for Gadget a second aggregation method is applied on the elementary sampling units, that is all data from a particular subdivision, which varies somewhat depending on the data source. Some data types, e.g. length distributions, are simply added whereas others, such as mean length-at-age, may go through a computational mechanism involving age-length keys. A description of a fisheries data base, which is able to handle data aggregations in this manner, can be found in Kupca and Sandbeck (2003) and Kupca (2004).
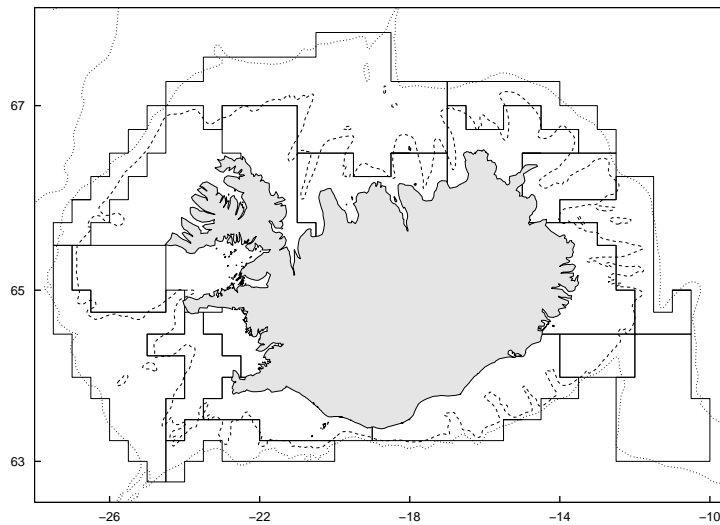
Figure II.1: The spatial structure of data storage on the Icelandic coastal shelf along with 200m (broken line) and 500m depth (pointed line) contours. These areas are referred to as "subdivisions". A given timeperiod, timestep size and subdivision is referred to as a "elementary sampling unit".

Here the fundamental idea is the aggregation of elementary sampling units in the creation model inputs. These sets of elementary sampling units can

therefore be sampled (with replacement) before aggregation, with each resample leading to a new model input data set. A typical model run for parameter estimation based on such a resampled data set will result in a resampled parameter estimate. The collection of all such estimates form a bootstrap sample. The procedure could be called a spatio-temporal block bootstrap with unequal block size.

II.3 A fisheries example

II.3.1 The setting

The example marine system used in this paper is based on cod in Icelandic waters (fig. II.1) with an approach very similar to Taylor et al. (2007). The model consists of two stock components of cod, i.e. mature and immature cod in a single area. Modelling maturity enables the calculation of spawning stock biomass and allows for different weight–length relationships to be used for immature and mature fish.

Two fixed station surveys are used to monitor the stock, in spring and autumn, providing population indices as well as biological samples. Landings information is available from official data bases and raw biological data (length distributions, age compositions) along with survey data in the Marine Research Institutes (MRI) databases (see e.g. ICES 2011, Palsson et al. 1989, Sigurdsson et al. 1997, Taylor et al. 2007, for a description of data and surveys). The technical details of the model are described in appendix II.6.

II.3.2 The data set and parameters

The model is a parametric and deterministic forward population dynamics simulation model. A single simulation results in a complete population structure, including predictions of all data sets, as described in Begley (2004) and Taylor et al. (2007) and a corresponding evaluation of a (negative log–)likelihood function (sums of squares in the present paper).

With the exception of landings data, data sets are only used in the likelihood components. For simplicity, landings data are used directly in the population models, whereby the populations are simply reduced in numbers to be in accordance with the corresponding landed weight. Note that in the approach proposed here the landings data are not resampled.

An overview of the datasets and model parameters used in this case study is shown in Tables II.1 and II.2 respectively.

II.3.3   Estimation protocol

The weights on the likelihood components are calculated for each model (i.e. each bootstrap run), according to the protocol described in appendix II.6.2 with arbitrary starting parameters. This is a two stage estimation method, where the error variances, within a data set, are estimated by increasing the weight on that particular component of the total sum of squares, followed by a final minimisation using those inverse variances as weights. For a full description of this procedure refer to the appendix.

The bootstrapping approach consists of the following:

- The base data are stored in a standardized data base:

    - Time aggregation: 3 months
    - Spatial aggregation: subdivision
    - Further disaggregation is based on a range of categories including fishing gear, fishing vessel class, sampling type (e.g. harbour, sea and survey). A full listing of data types used in the case study can be found in table II.1, these data are stored subdivision dis-aggregated to allow for use in a bootstrap.

- To bootstrap the data, the list of subdivisions, depicted in fig. II.1, required for the model is sampled (with replacement) and stored. For a multi–area model one would conduct the resampling of subdivisions within each area of the model.

- The list of resampled subdivisions is then used to extract data (with replacement so the same data set may be repeated several times in a given bootstrap sample).

- For a single bootstrap Gadget model, the same list of resampled subdivisions is used to extract each likelihood dataset i.e. length distributions, survey indices and age–length frequencies are extracted from the same spatial definition.

- A Gadget model is fitted to the extracted bootstrap dataset using the estimation procedure described above.

- The resampling process is repeated until the desired number of bootstrap samples are extracted.

When resampling, data are forced to remain in the correct year and time–step so resampling is based on sampling spatially the elementary data units within a given modelled unit of time and space. Thus, within a modelled spatial unit the bootstrap is a resampling of subdivisions. This implicitly assumes data contained within each area of the model to be independent and identically distributed. Indepence is justified by the definition of subdivions. Furthermore treating them as they were from the same distribution, i.e. bootstrap replicates, appears to have little negative effect when compared to more traditional methods (Taylor 2002).

The entire estimation procedure is repeated for each bootstrap sample. In particular, since the estimation procedure includes an iterative reweighting scheme, this reweighting is repeated for every bootstrap sample. The point of this is that the bootstrap procedure is no longer conditional on the weights. The procedure as a whole is quite computationally intensive but can easily be run in parallel, e.g. on a computer cluster.

In stark contrast to this, Hessian–based approaches usually only compute the Hessian at the final solution. Thus, they completely omit the effect of reweighting likelihood components when estimating uncertainty. Such methods are thus conditional on the weights obtained in a pre–estimation stage.

II.3.4   Application of the bootstrap procedure and its variants

The bootstrap procedure presented here is, as noted earlier, quite computationally demanding as the number of bootstrap samples increases. In this excercise 1000 bootstrap samples were chosen as the **baseline** simulation. This number of iterations was chosen as a practical upper limit, as a single optimisation run for a Gadget model takes a substantial amount of time. In addition to the baseline simulation two sensitivity tests are considered in the present case study. Here it is of considerable interest to study possible reduction in the number of bootstrap samples and other means to reduce the amount of calculations. An interesting comparison to the baseline simulation would be to reduce the number of bootstrap samples to 100 samples. A more thorough analysis of the effects of sample size is described in subsection II.3.6.

Another interesting sensitivity test would be a bootstrap procedure conditional on weights obtained at the pre–estimation stage, i.e. use the same (fixed) likelihood weights throughout the simulation. The reason for this comparison

is twofold, it is computational as the amount of calculations required would
be drastically reduced and also in relation to Hessian–based approaches. One
should note however that with this bootstrap the estimation is not the same
function of the data as the procedure where the weighting takes place for each
dataset. This may lead to inappropriate weights for a given dataset which in
turn can, as mentioned earlier, lead to inaccurate parameter estimates.

II.3.5   Hessian–based inference

For illustrative purposes the inferences arising from the bootstrap procedure
presented here is compared to a Hessian–based confidence interval (described
by Tinker et al. 2006, and references therein). In particular central differences
were used to calculate the needed second derivatives used to obtain an estimate
of the variance–covariance matrix and a multivariate delta method (Oehlert
1992) to obtain the confidence interval for derived biomass.

   The effects of samples size on the inferences obtained from the inverted Hessian matrix were studied by an artificial increase in measurements. The timestep
length was varied between 1, 2 and the baseline 3 months, with input files being
adjusted accordingly. The resulting CVs for the recruitment parameters were
estimated and the effects of the different step lengths contrasted. Similar analysis was conducted for the proposed bootstrap procedure but is, for the sake of
clarity, only discussed in connection to the Hessian based approach.

II.3.6   Number of bootstrap samples

With regards to the bootstrap procedure itself this study also examines the
effect of the number of bootstrap samples on the variance and bias estimates
using a retrospective bootstrap. For a sample number n, ranging from 25 to
1000 bootstrap samples, n vectors of parameter estimates from the baseline
bootstrap were sampled with replacement 100 times. From those 100 samples
the coefficient of variation (CV) was calculated for the mean and standard deviation of each parameter. Uncertainty in bias estimation is harder to quantify
in a similar way as parameter bias is often estimated close to zero.

II.3.7   Model output

Given the optimised parameter estimates it is possible to output a wide range
of descriptors of the model ecosystem as Gadget operates on and stores the
number in each age–length cell for each time-step of the model. For this study,

the estimated parameters along with a derived biomass trajectory (age 4+)
are considered. Comparisons of uncertainty estimates will be, as noted earlier,
made using the three bootstrap variants, i.e. both 1000 and 100 bootstrap
simulations with the iterative reweighting procedure applied to all bootstrap
samples and 1000 bootstrap simulations using fixed weigths, and the Hessian–
based approach. A schematic overview of all calculations performed here is
shown in figure II.2.
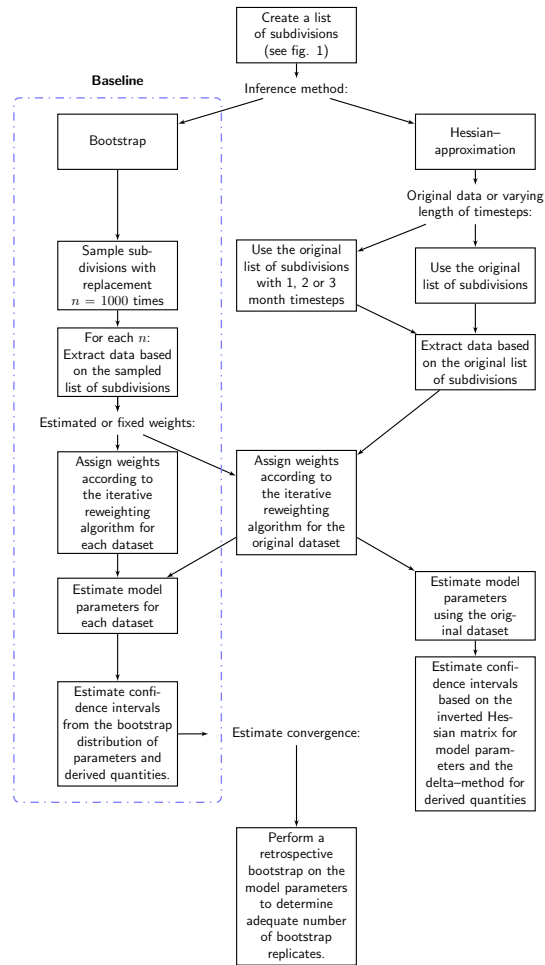
Figure II.2: A flowchart of the calculations performed. Boxes indicate action and unbounded text possible uncertainty estimation variants or decisions.

II.4   Results
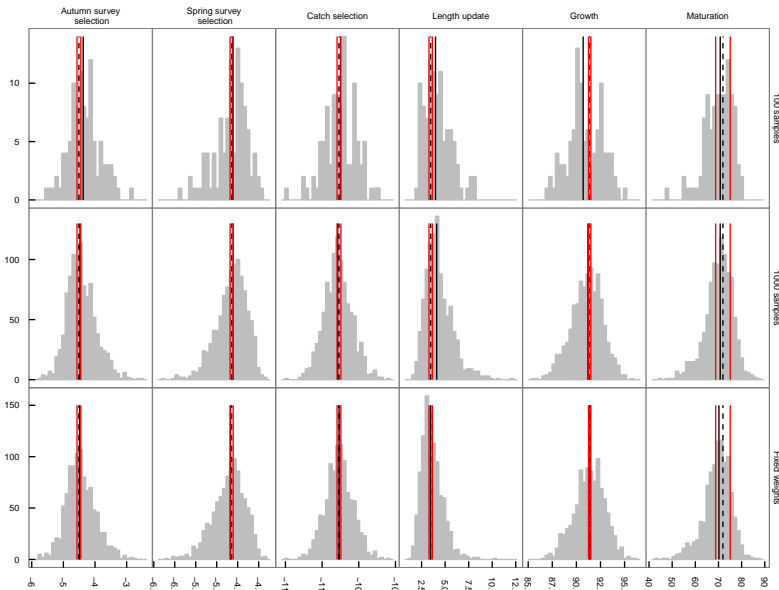
Figure II.3: Histograms of the estimated fleet selection parameter a_f for the three fleets (October survey, March survey, Commercial catch), β the parameter defining the length update matrix, k the growth rate and the maturity l_{50}. The parameter estimates were obtained from 1000 bootstrap samples, compared to a smaller number of bootstrap samples, 100, where for the two number of samples iterative weighting applied to all bootstrap samples. This is then all compared to 1000 bootstrap samples where in the parameter estimation the weighted likelihood function is conditioned on the original weights. The point estimate (black broken line) and bootstrap mean (black solid line) along with 95% confidence bound obtained from a Hessian–based approximation to the variance covariance matrix (red solid lines) are indicated.

The simplest model outputs are the point estimates of model parameters. Fig. II.3 gives histograms of bootstrap estimates of several parameters. It compares the distributions of those parameter estimates from 1000 bootstrap samples, either using reweighting for each dataset or fixed weights, to those

using only 100 samples with reweighting. For each parameter, the point estimate from the full data set, the median of the bootstrap estimates and 95% confidence intervals from a hessian-based approximation are indicated. The differences between the point estimate and the bootstrap mean can be seen to be relatively minor, i.e. there is no obvious sign of an estimation bias, in all cases except for the length update (see β in eq. II.3). It should be noted that the maturation parameters are correlated, affecting the relationship between the point estimate and bootstrap mean for the maturation. The different bootstrap methods exhibit similar distribution of parameter estimates with the exception of the length update where the bootstrap mean based on the original weights falls closer to the point estimate thus failing to detect bias in the length update.

Boxplots can be used to illustrate bootstrapped trajectories of various abundance or biomass measures. The estimate of the 4^+ biomass is shown in fig. II.4. It is seen that the main variation appears, in absolute terms, in the initial and final years, while only the final year shows considerable amount of variation in terms of CV. The initial and final years are of course considerably different from the intermediate ones, but in different ways. The number of fish in the initial year are part of the estimation procedure and therefore of a different nature when compared to subsequent years. Further, the survey starts in 1985 (with the model starting in 1984), which makes the initial conditions somewhat poorly determined. The final years are on the other hand poorly determined since there is relatively little information in the objective function for the younger year classes as they have only been surveyed for a few years.

The same effects are seen for estimated recruitment at age 1 (fig. II.5) where there is less variation in the earlier and intermediate years than the later years. As for the other parameters the Hessian–based confidence estimates are considerably smaller than those obtained using bootstrap methods. The CV of the Hessian–based approach roughly followed the same pattern as for those arising from the various bootstrap approaches but were generally around 12% of the corresponding bootstrap CV.

In figure II.6 CVs for the mean and standard deviation of the model parameters are shown as function of the number of bootstrap samples, n, where the separate panels show different groups of parameters. The CV–estimates appear to fall close to $\frac{1}{\sqrt{n}}$, as shown in the figure, and that most of them are less than 15% for 100 bootstrap samples. The initial conditions, that is the numbers at age in 1984, had a somewhat higher CV for the mean and standard deviation as the other parameter groups. The initial numbers at age 8 and 9 in 1984 in particular, showed a considerably higher CV for all sample sizes. Those two age

groups were, as noted earlier, poorly determined, and had a very low estimate compared to other intial numbers, as the corresponding year classes were only present in the data for the first few years of the model.

Hardly any biases were observed in this analysis. Notable exceptions were the length update parameter, shown in figure II.3, and the first two years the 4^+ biomass appeared to have a measurable bias. This was only detected in the bootstrap simulations where the iterative reweighting scheme was applied to all bootstrap samples. The fixed weight run and the Hessian–based approach failed to detect these differences.

The effects of the number of timesteps within a year can been seen in figure II.7. There the CV of recruitment is illustrated as a function the number of (intra–year) timesteps in the model. The number of timesteps appears to be inversely proportional to the CV size. These effects were not, when varying the timestep, observed when conducting similar analysis using the bootstrap.
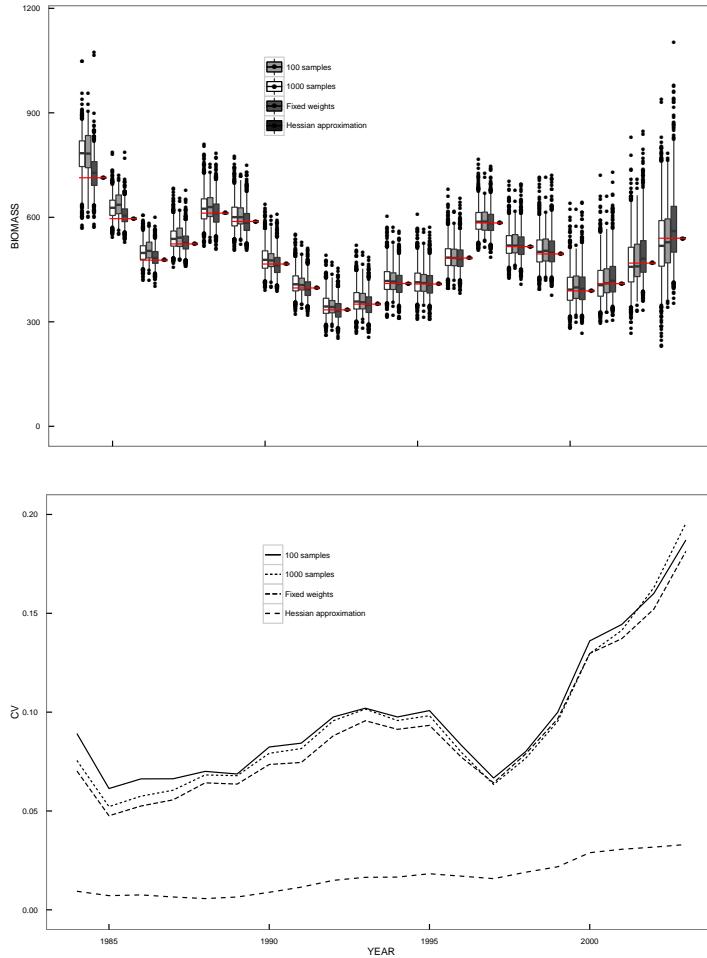
Figure II.4: Boxplot (top panel) of the end of year biomass for cod of age 4 and older
estimated on 1000 bootstrap samples, both using iterative weighting for each sample
and using the fixed weights for all samples, compared to 100 bootstrap samples. The
fixed weights were obtained using iterative weighting for the original dataset. The point
estimate is indicated by the central red line through the boxes. The box indicates the
interquartile range and the whiskers 95% confidence intervals. Any further outlying
data points are indicated as points. Bottom panel shows the estimated CV for the age
4^+ biomass using the same methods as above.
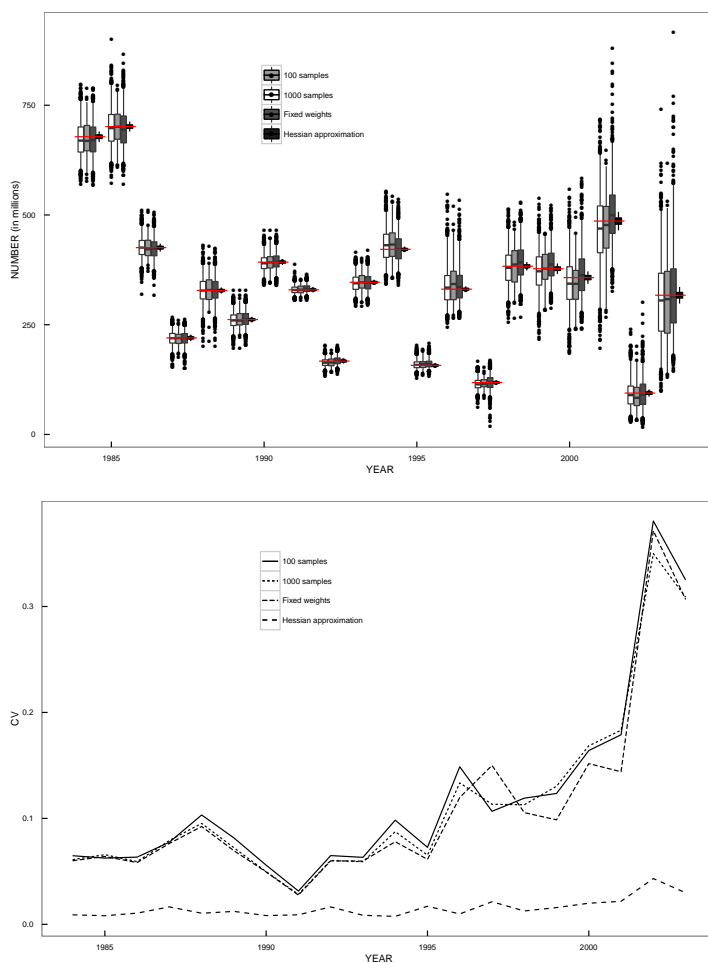
Figure II.5: Boxplot (top panel) of the number of recruits (age 1) in each year estimated by 1000 and 100 bootstrap models compared to 1000 bootstraps with fixed weights and a Hessian–based approximation to the 95% confidence interval. The point estimate indicated by a central red line through the boxes. Bottom panel shows the estimated CV for the recruitment using the same methods as above.
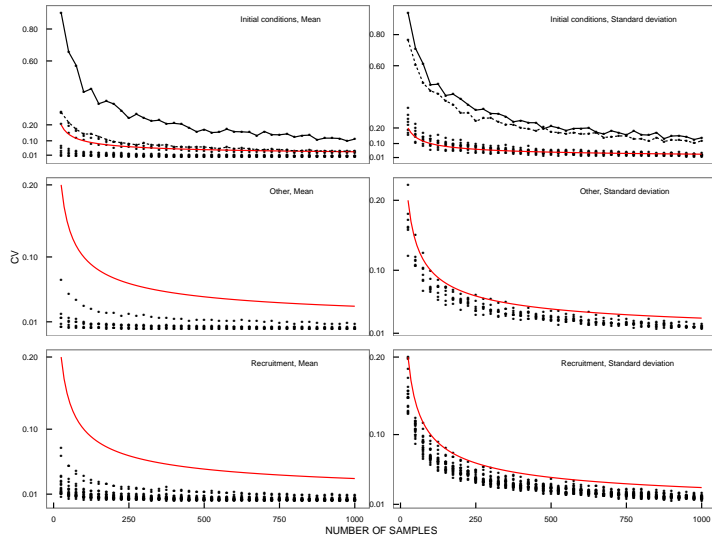
Figure II.6: Results of a restrospective bootstrap sampling on the parameter esti-
mates from the 1000 bootstrap samples, with iterative weighting applied to all samples.
This retrospective bootstrap studies the variation of the mean and standard deviation
of each parameter estimate by calculating the coefficient of variation (CV) as function
of the number of bootstrap samples, n, of both the mean and standard deviation (SD).
A point on the graph shows the CV of the mean (panels on the left hand side) or SD
(panels on the right hand side) for a particular parameter and number of samples,
n. The different panels contain the CVs of the initial number at age (ν_a in eq. II.6),
"Other" variables i.e. the variables which are shown in figure II.3, and yearly recruit-
ment shown in figure II.5 (R_y in eq. II.5). CV of the initial number at ages 8 and 9
are illustrated with solid and broken lines respectively. For comparison $1/\sqrt{n}$ is shown
(red solid line) on all panels.
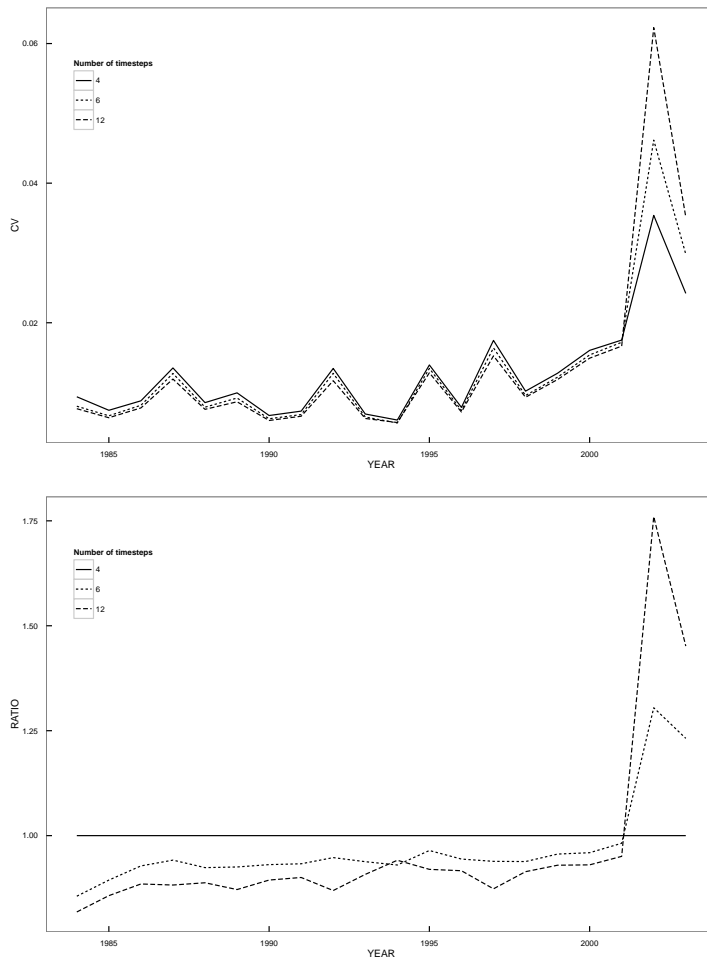
Figure II.7: The CV of recruitment arising from the inverted Hessian (upper panel) by year as the number of intra–year timesteps are increased. The bottom panel shows the ratio of the CV of the model with 3 month timesteps to the models with 2 months and 1 month steps.

II.5   Discussion

This paper has presented a novel bootstrap method suitable for models of population dynamics. Several modifications and alternatives to the original bootstrap methodology (Efron 1979, Efron and Tibshirani 1994) have been presented. For example, to account for correlations in simple non–replacement sampling schemes (as used for most questionnaires or "sample surveys"), without–replacement bootstraps and with–replacement bootstraps have been suggested along with somewhat more general resampling procedures for complex survey data (McCarthy and Snowden 1985, Gross 1980, Rao and Wu 1988, Sitter 1992). Theoretical assumptions and derivations behind these approaches do not easily extend to the present situation with disparate data sets, composite likelihoods in the estimation phase and last but not least the highly nonlinear population dynamics models used as a basis for obtaining predicted values and error sums of squares or likelihood functions. The "trick" in the current proposal is not a theoretical development but the methodology of having the bootstrap sampling unit y_i as a collection of all relevant datasets sufficiently aggregated such that they can be assumed to be independent.

Some of the modifications of the original bootstrap have been developed for marine surveys (Smith 1997) but this has been intended to reflect e.g. the sampling design used for the surveys and simple estimation of quantities such as a stratified mean. In the present setting the data need to go through an aggregation procedure to be used in a nonlinear population dynamics model and it is the output of this model which is of interest, not variances in the input. Thus there is a need for the bootstrap to mimic this aggregation procedure for the full data from raw data or finer-scale aggregates. This is the case with any population dynamics or assessment model, used in fisheries or other areas of resource harvesting particularly in a multispecies and multi–area context.

The methodology proposed here is certainly computationally intensive. However this is also the case for many other methods. For example, the MCMC evaluation of a Bayesian posterior involves a simulation of a correlated time–series whose stationary distribution is the posterior. This process is not trivially parallelizable over an arbitrary grid of computers (some of the difficulties are described in Wilkinson 2006). In comparison the bootstrap approach described here is fairly trivially distributed onto a computer cluster.

To make the bootstrap proposed here more feasible one could reduce the number of resampled datasets. Using 100 bootstrap replicates instead of 1000 yields satisfactory results in terms of variance estimation, allowing a drastic reduction in the computing time needed. Conditioning on the weights from the

original sample could further reduce the time needed but, judging by the results presented here, possible estimation biases may be harder to detect.

When compared to the bootstrap the Hessian–based approximation appears to underestimate the uncertainty by a factor of 8. This may seem contrary to previous results. Magnusson et al. (2012), using a simple catch–at–age simulation model, concluded that MCMC–method and the Hessian–based approach performed similarly. And recently in Stewart et al. (2012) an MCMC and a Hessian–based approach performed similarly for real applications. The notable difference between the model described here and the aforementioned approaches is the objective function used here and total number of data points (defined in appendix II.6.2) used in the estimation process. The objective function consists of simple sums of squares that ignores potential correlations and tends to exagerate the confidence level in the Hessian–approach as the number of datapoints increases. This is illustrated in figure II.7 where it appears that the main factor in determining the size of the CV is the number of datapoints in the input files. Scale changes, such as aggregating data to larger lengthgroups or increasing the size of the plus group by lowering the modelled maximum age, would in this case increase the size of the CV by simply reducing the number of datapoints. In contrast, to the approach used here, a multinomial model, where the degrees of freedom are estimated, is often employed on catch at age (eg. Trenkel et al. 2012) but length distributions, in the case of Icelandic cod, have serious distributional problems (Hrafnkelsson and Stefansson 2004). Future work on the model could potentially evaluate different distributional assumptions similar to those suggested above using the proposed bootstrap approach.

In this particular case study there were no discernible biases detected. Thus the consequences of the Hessian–based approach appear to be mostly restricted to narrower confidence intervals. However it is reasonable to assume that inconsistencies arising from conflicting data sources (eg. in Stefansson 2003, Schnute and Hilborn 1993) would not be detected without analysing the effects of their relative weights. On the other hand, incorrect variance estimates may directly affect how annual catches are set. This occurs for example if a harvest control rule were to be based on a probabilistic measure such as a that of a biomass not falling below a threshold or a TAC not deviating too much from a target.

It is of considerable interest to compare the proposed bootstrap method to MCMC methods used in the Bayesian framework. This is however outside of the scope of this study as it would require a considerable effort to adapt the Gadget framework to the Bayesian one. Future work could potentially focus on the evaluation of the two methodologies both on simulated datasets and for real applications similar as was done in Hannesson et al. (2009).

It is reassuring that the modelled years in which the greatest uncertainty is found are those where it is expected i.e. the initial year and then increasing towards the end of the modelled time period. The first year is the most data poor with no survey data or age–length compositions and towards the end of the time period there are fewer cohorts with data available for most ages.

The method described here is designed to alleviate several known problems with other methods of uncertainty estimation. Several issues remain, however. For example if a model is too "stiff" through fixing parameters or other assumptions, then this may not be detected here except in special cases. These considerations could be explored by different models, e.g. split the commercial fleet component by gears, which can be implemented within the Gadget framework. On a related note there is also a balance to be found between estimation errors due to too small size classes and distribution error caused by too large size classes (Vandermeer 1978). It is therefore of interest to investigate the effects of the choice of scale such as size class width but also time step (Drouineau et al. 2009). The relative merits of these models can then be evaluated using an approach similar to the one proposed here. Similarly different modelling approaches, such as the different data weighting discussed in Francis (2011) or Hu and Zidek (2002), can be also be compared using the bootstrap technique presented here. Ultimately, each reweighting scheme is a different method for obtaining a point estimate and the bootstrap is a perfectly general method to obtain variance estimates.

When designing an aggregated data base to be used for modelling several issues need to be taken into account. The most important statistical condition on the choice of the "data units" is that correlations between them should be minimal. On the other hand there also needs to be a fair number of them within each model area if the bootstrap mechanism is to provide some variation in results. For a given measurement type one can in many cases investigate spatial correlation or variograms to determine the distances at which those become negligible (Petitgas 2001). This can not easily be done for many data types, however (age-length tables, tagging experiments etc). In fact, the original reasoning for the areas used in this paper was ecological (Stefánsson and Pálsson 1997a, Taylor 2003) rather than based on spatial correlation, and it is likely that in most real situations data will be aggregated according either to such criteria or pragmatically into "statistical rectangles" of some form.

Simple bootstrap resampling usually assumes that the elementary data units, $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$, behave like idependently and identically distributed samples. Data in fisheries tends to be collected in a somewhat stratified manner, ranging from formal stratification to attempts to "spread out" sampling, across gears, time

and space. In the present setup this is simply ignored. This can be justified when the data are aggregated in a simple manner (through sums or averages) anyway since the bootstrap method then mimics the computation accordingly and/or when there are a large number of data units which can be viewed as representing a population of such units. In cases when one or a few of the subdivisions represent e.g. a spawning area and the intended analysis is stratified accordingly, this approach can clearly not be used since then the bootstrap resampling does not reflect the computational method in use. When such issues arise, whether with respect to fishing gear, space or other units, an appropriate approach is to include these elements in the model. For example the likelihood function can incorporate the various fishing gears, modelling each selectivity separately. The resampling then takes place separately for each gear.

Acknowledgements

II.6 Model description

II.6.1 Simulation model

The following describes the technical details of the Gadget model used in this case study. The model was previously described in Taylor et al. (2007) along with a statistical assessment of its fit to data. In the model the simulated quantity is the number of individuals, N_{alsyt}, at age $a = 1 \ldots 12$, in a lengthgroup l, representing lengths ranging between 4 and 140 cm in 2 cm lengthgroups, stock component s where $s = 0, 1$ denotes the mature and immmature stock component respectively, at year y which is divided into quarters $t = 1 \ldots 4$. The length of the timestep is denoted Δt. The population is governed by the

following equations:

$$N_{alsy,t+1} = \sum_{l'} G_l^{l'} \left[(N_{al'syt} - C_{fal'st})e^{-M_a \Delta t} + I_{al'syt} \right] \qquad \text{if } t < 4$$

$$N_{a+1,ls,y+1,1} = \sum_{l'} G_l^{l'} \left[(N_{al'sy,4} - C_{fal's,4})e^{-M_a \Delta t} + I_{al'sy,4} \right] \quad \text{if } t = 4 \ \& \ a < 12$$

$$N_{a,ls,y+1,1} = \sum_{l'} G_l^{l'} (N_{al'sy,4} - C_{fal'sy,4} +$$

$$N_{a-1,l'sy,4} - C_{f,a-1,l'sy,4})e^{-M_a \Delta t} \qquad \text{if } t = 4 \ \& \ a = 12$$

$$\text{(II.1)}$$

where $G_l^{l'}$ is the proportion in lengthgroup l that has grown $l' - l$ lengthgroups in Δt, C_{falsyt} denotes the catches by fleet $f \in \{S, A, C\}$, S, A and C denote the spring and autumn surveys and commercial fleets respectively[1], M_a the natural mortality at age a and I_{alsyt} denotes the movement from the immmature to the mature stock components. A short note on notation, here l is used interchangeably as either the lengthgroup or the midpoint of the length interval for that particular lengthgroup, depending on the context.

Growth

Growth in length is modeled as a two–stage process, an average length update in Δt and a growth dispersion around the mean update (as described in Stefansson 2005a). The average length update per time step is set according to a simplified form of the Von Bertanlanffy equation:

$$\Delta l = (l_\infty - l)(1 - e^{-k\Delta t}) \qquad \text{(II.2)}$$

where l_∞ is the terminal length and k is the annual growth rate. In the second step the growth is dispersed according to a beta–binomial distribution parameterised by the following equation:

$$G_l^{l'} = \frac{\Gamma(n+1)}{\Gamma((l'-l)+1)} \frac{\Gamma((l'-l)+\alpha)\Gamma(n-(l'-l)+\beta)}{\Gamma(n-(l'-l)+1)\Gamma(n+\alpha+\beta)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \qquad \text{(II.3)}$$

where α is subject to

$$\alpha = \frac{\beta \Delta l}{n - \Delta l} \qquad \text{(II.4)}$$

where n denotes the maximum length group growth and $(l' - l)$ the number of lengthgroups grown.

---

[1]The survey fleets catches are given a nominal catch to allow for survey age and length distribution predictions.

Recruitment and initial abundance

A fairly simple model is used for recruitment in this exercise. Recruitment enters to the population according to:

$$N_{1l0yt'} = R_y p_l \tag{II.5}$$

where t' denotes the recruitment time-step and R_y is the yearly recruitment. p_l is the proportion in lengthgroup l that is recruited which is determined by a normal density with mean according to the growth model and variance σ_y^2.

Equally simple is the formulation of initial abundance in numbers for each age group in lengthgroup l is:

$$N_{als11} = \nu_a q_l \tag{II.6}$$

where ν_a is the initial number at age a in the initial year and q_l the proportion at lengthgroup l which is determined by a normal density with a mean according to the growth model in equation II.2 and variance σ_a^2.

Maturation

Two stage maturity is modeled and represented by the two stock components. The movement between the two components is formulated as

$$I_{alsyt} = \begin{cases} \sum_{l'} N_{al'0y,t-1} \times m_{l'}^l & \text{if s} = 1 \text{ and t} > 1 \\ \sum_{l'} N_{al'0y-1,4} \times m_{l'}^l & \text{if s} = 1 \text{ and t} = 1 \\ -\sum_{l'} N_{al'0y,t-1} \times m_{l'}^l & \text{if s} = 0 \text{ and t} > 1 \\ -\sum_{l'} N_{al'0y-1,4} \times m_{l'}^l & \text{if s} = 0 \text{ and t} = 1 \end{cases} \tag{II.7}$$

where $s = 0$, as noted above, denotes the immature stock component. and $m_{l'}^l$ is the proportion of immatures that mature between the lengths l and l' defined as:

$$m_{l'}^l = \frac{-\alpha G_{l'}^l (l - l') e^{-\lambda(l'-l_{50})}}{1 + e^{-\lambda(l'-l_{50})}} \tag{II.8}$$

Fleet operations

Catches are simulated based on reported total landings and a length based suitability function for each of the three fleets, commercial fleet and the autumn and spring survey. Total landings are assumed to be known and the total biomass

is simply offset by the landed catch. The catches for lengthgroup l , fleet f at year y and time-step t are calculated as

$$C_{flsyt} = E_{ft} \frac{S_f(l) N_{lsyt} W_{ls}}{\sum_{s'} \sum_{l'} S_f(l') N_{l's'yt} W_{l's'}}$$

(II.9)

where E_{ft} is the landed biomass at time t and $S_f(l)$ is the suitability of length-group l by fleet f defined as:

$$S_f(l) = \frac{1}{1 + e^{(-a_f - b_f l)}}$$

(II.10)

The weight, W_{sl}, at lengthgroup l is calculated according to the following stock component specific length – weight relationship:

$$W_{sl} = \mu_s e^{\omega_s l}$$

(II.11)

II.6.2 Observation model

In Gadget data are assimilated using a weighted log–likelihood function. Here four types of data enter the likelihood, length-based survey indices, length distributions from survey and commercial fleets, age – length distribution from the survey and commerical fleets and maturity at length for 3 year olds.

Survey indices

The survey indices are defined as the total number of fish caught in a survey within a certain length interval. The intervals used here are $16 - 25$ cm, $26 - 38$ cm and larger than 38 cm. These intervals are chosen such that they roughly represent age 1, age 2 and age 3^+.

For each length range g the survey index is compared to the modelled abundance at year y and time-step t using:

$$l_{gf}^{\text{SI}} = \sum_y \sum_t (\log I_{gfy} - (\log q_f + \log \widehat{N_{gyt}}))^2$$

(II.12)

where

$$\widehat{N_{gyt}} = \sum_{l \in g} \sum_a \sum_s N_{alsyt}$$

The above formulation assumes that the length-based indices are independent and uniform selectivity by survey gear. This is seen as a fair assumption as little

correlation is observed in the data for the first two lengthgroups. With regards to survey selectivity this has been estimated to be fairly constant during the model time period (Gudmundsson 2013). For further implementation details see Taylor et al. (2007).

Maturity at length

Length at maturity comparison uses the number of mature cod that are age 3 observed in the Icelandic ground fish survey. The observed proportions are compared to the modelled proportion using sum of squares:

$$l^{\mathrm{M}} = \sum_y \sum_t \sum_l (\pi_{lyt} - \hat{\pi}_{lyt})^2 \qquad (\mathrm{II}.13)$$

where

$$\pi_{lyt} = \frac{\sum_a O_{al1yt}}{\sum_a \sum_{l'} \sum_s O_{alsyt}}$$

and

$$\hat{\pi}_{lyt} = \frac{\sum_a N_{al1yt}}{\sum_a \sum_{l'} \sum_s N_{alsyt}}$$

i.e. the observed and modelled proportions mature respectively in length group l, year y and timestep t.

Fleet data

Length distributions are compared using 2 cm lengthgroups for both commercial and survey fleets using

$$l_f^{\mathrm{LD}} = \sum_y \sum_t \sum_l (\pi_{lyt} - \hat{\pi}_{lyt})^2 \qquad (\mathrm{II}.14)$$

where f denotes the fleet where data was sampled from and

$$\pi_{lyt} = \frac{\sum_a \sum_s O_{alsyt}}{\sum_a \sum_{l'} \sum_s O_{alsyt}}$$

and

$$\hat{\pi}_{lyt} = \frac{\sum_a \sum_s N_{alsyt}}{\sum_a \sum_{l'} \sum_s N_{alsyt}}$$

i.e the observed and modelled proportions in lengthgroup l respectively at year y and timestep t. Similarly age – length data are compared using 4 cm length groups:

$$l_f^{\mathrm{AL}} = \sum_y \sum_t \sum_a \sum_l \sum_s (\pi_{falsyt} - \hat{\pi}_{falsyt})^2 \qquad \text{(II.15)}$$

where

$$\pi_{alyt} = \frac{\sum_s O_{alsyt}}{\sum_a \sum_{l'} \sum_s O_{alsyt}}$$

and

$$\hat{\pi}_{alyt} = \frac{\sum_s N_{alsyt}}{\sum_a \sum_{l'} \sum_s N_{alsyt}}$$

Iterative re–weighting

The total objective function used the modelling process combines equations II.12 to II.15 using the following formula:

$$l^{\mathrm{T}} = \sum_g \sum_{f \in \{S,A\}} w_{gf}^{\mathrm{SI}} l_{gf}^{\mathrm{SI}} + \sum_{f \in \{S,A,C\}} \left(w_f^{\mathrm{LD}} l_f^{\mathrm{LD}} + w_f^{\mathrm{AL}} l_f^{\mathrm{AL}}\right) + w^{\mathrm{M}} l^{\mathrm{M}} \qquad \text{(II.16)}$$

where $f = S, A$ or C denotes the spring survey, autumn survey and commercial fleets respectively and w's are the weights assigned to each likelihood components.

The weights, w_i, are necessary for several reasons. First of all it is used to prevent some components from dominating the likelihood function. Another would be to reduce the effect of low quality data. It can be used as an *a priori* estimates of the variance in each subset of the data.

Assigning likelihood weigths is not a trivial matter, has in the past been the most time consuming part of a Gadget model. Commonly this has been done using some form of 'expert judgement'. General heuristics have recently been developed to estimated these weights objectively. Here the iterative re–weighting heuristic introduced by Stefansson (2003), and subsequently implemented in Taylor et al. (2007), is used.

The general idea behind the iterative re-weighing is to assign the inverse variance of the fitted residuals as component weights. The variances, and hence the final weights, are calculated according to the following algorithm:

1. Calculate the initial sums of squares (SS) given the initial parameterization for all likelihood components. Assign the inverse SS as the initial weight for all likelihood components.

2. For each likelihood component, do an optimization run with the initial SS for that component set to 10000. Then estimate the residual variance using the resulting SS of that component divided by the degrees of freedom (df^*), i.e. $\hat{\sigma}^2 = \frac{SS}{df^*}$.

3. After the optimization set the final weight for all components as the inverse of the estimated variance from the step above (weight $= 1/\hat{\sigma}^2$).

The number of non-zero data-points (df^*) is used as a proxy for the degrees of freedom. While this may be a satisfactory proxy for larger datasets it could be a gross overestimate of the degrees of freedom for smaller datasets. In particular, if the survey indices are weighed on their own while the yearly recruitment is estimated they could be over-fitted. In general problems such as these can be solved with component grouping, that is in step 2 the likelihood components that should behave similarly, such as survey indices, should be heavily weighted and optimized together. This approach is used here for the spring and autumn survey indices.

II.6.3 Fit to data

A detailed discussion of the model fit to the datasets listed in Table II.1 is beyond the scope of this discussion. This analysis can be found in Taylor et al. (2007). However for illustrative purposes a comparison between survey indices is shown in Figure II.8. The model appears to follow the general features of the survey indices, in particular the age 3+ groups. The indices representing age one are in general seen to be more variable than for the latter groups.
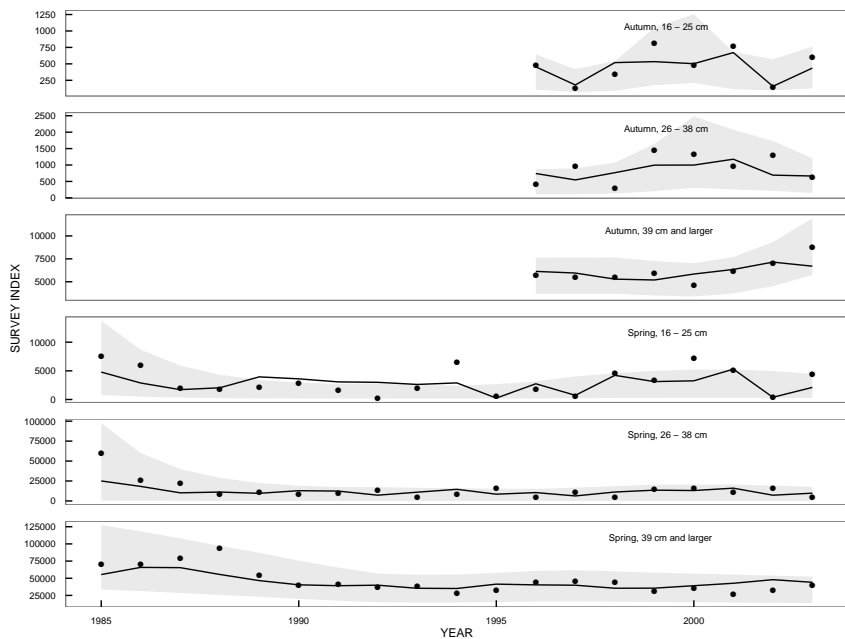
Figure II.8: Survey indices (points) by year compared to the model fit (lines) split by
season and length groups. The length groups roughly represent the one, two and older
agegroups. 95% confidence intervals from a loess smoother are shown for comparison.

Origin	Timespan	Length group size	Num. data-points	Likelihood function
	Length distributions:			
March Survey	1st quarter, 1985 − 2003	2 cm	1292	See eq. II.14
October Survey	4th quarter, 1995 − 2003	2 cm	558	See eq. II.14
Commercial catches	All quarters, 1984 − 2003	2 cm	5202	See eq. II.14
	Age − length frequencies			
March Survey	1st quarter, 1989 − 2003	4 cm	6120	See eq. II.15
October Survey	4th quarter, 1995 − 2003	4 cm	3348	See eq. II.15
Commercial catches	All quarters, 1984 − 2003	4 cm	31042	See eq. II.15
	Survey indices			
March Survey	1st quarter, 1985 − 2003	−	57	See eq. II.12
October Survey	4th quarter, 1995 − 2003	−	24	See eq. II.12
	Ratio of immature:mature at age 3 by length group			
March Survey	1st quarter, 1985 − 2003	2 cm	1672	See eq. II.13

Table II.1: Overview of the likelihood data used in the model. Survey indices are calculated from the length distributions and are disaggregated ("sliced") into three groups which correspond roughly to age 1, age 2 and age 3+ (as in Taylor et al. 2007). Number of datapoints refer to aggregated data used as inputs in the Gadget model and represent the original dataset. All data can obtained from the Marine Research Institute, Iceland.

Description	Notation	Comments	Formula
Natural mortality	M_a	Fixed at 0.2 for ages 3 to 9 while 0.5, 0.35 for the first two ages and 0.3, 0.5 and 0.7 for ages $10-12$	See eq. II.1
Growth function	k	l_∞ is fixed at 120 cm	See eq. II.2
Growth implementation	β	n is fixed at 15 lengthgroups	See eq. II.3
Fleet selection	a_f	b_f fixed at 0.09 for the survey fleets while 0.19 for the commercial fleet	See eq. II.10
Maturity ogive	λ, l_{50}	l_{50} in cm	See eq. II.7
Number of recruits by year	R_y	$y \in [1984, 2003]$. σ_y^2, i.e. variance in recruitment length, based on length distributions obtained in the spring survey	See eq. II.5
Initial abundance at ages $2-11$ in 1984	η_a	$a \in [2, 11]$. σ_a^2, i.e. variance in initial length at age a, based on length distributions obtained in the spring survey	See eq. II.6
Survey catchability	q_f	Intercept term in a log–linear relationship with abundance. The slope term is assumed to be 1 for all indices	See eq. II.12
Length–weight relationship	μ_s, ω_s	Different values by stock component, estimated outside of the model	See eq. II.11

Table II.2: An overview of the estimated parameters in the model. For those parameters with fixed values a description of how these values were derived can be found in Taylor et al. (2007) and references therein.

III

Paper III

RGadget: a R–package for development, testing and analysis of Gadget models

Bjarki Þór Elvarsson, Guðmundur Þórðarson and Gunnar Stefansson

Abstract

Various frameworks and software packages have been developed in re-
cent years to aid in the process of assessing marine resources. One in
particular, the Gadget framework, allows for the development of statis-
tically testable models of marine multi–species ecosystems. It has been
popular in cases where data on the resource is scarce. Gadget's flexi-
bility does however allow the user to define quite varied output leaving
something to be desired when it comes to standardised presentation of
results.

This paper summarises the development and functionality of RGadget,
a R-package built to aid in the development and testing of models built
using the Gadget framework. RGadget includes tools to build a model
skeleton, a heuristic to assign weights to disparate dataset used in the
estimation procedure, model diagnostics and an independent age-length
based simulation model that mimics key processes in Gadget. Tools useful
in a management context such as yield per recruit and forward projections
are also described.

III.1 Introduction

A typical fisheries stock assessment involves contrasting data and some representation of the marine ecosystem through a likelihood function. When the assessment model demands little in terms of data the assessment of the fit is fairly straightforward. For instance the standard stock production model (Pella and Tomlinson 1969) requires two timeseries of data, catches and an index of abundance or biomass, as input. Measurements of the goodness of fit are then only based on the index.

This comparison is however not as straightforward when multiple datasets and sources are contrasted in a single models. Models, such as those developed using the Gadget framework (described in Begley and Howell 2004), are typically based on a number of disparate and conflicting data sources. This gives rise to various issues when assessing the goodness of fit that are often connected to the:

- **Types of data**: data on different aspects of the ecosystems, e.g. indices of abundance and length distributions

- **Sources of data**: survey sampling or sampling from commercial operations.

As commonly done these issues are dealt with in the Gadget framework by using a weighted likelihood function. The weights are necessary for several reasons. Notably, they can be used to prevent some components from dominating the likelihood function, to reduce the effect of low quality data, and as *a priori* estimates of the variance in each subset of the data.

Choosing these weights is, however, not trivial. As an example one might attempt to emphasize a long time series of survey indices while reducing the effect of sporadic age measurements. The issue here is what constitutes a *high* and *low* weight, as that may vary between datasets and models. A more objective weighting scheme, *iterative reweighting*, was introduced in Stefansson (2003) and implemented for cod in Taylor et al. (2007). This methodology and its implementation in R is described further in the following sections.

Another important, but often neglected, point is a standard presentation of results to aid the identification of possible issues and increase familiarity for potential reviewers. The presentation falls essentially in two categories: First how well the model describes the observation, i.e. graphical representation and individual likelihood scores by data set. As an example in an effort to standardise the diagnosis of output from stock synthesis (Methot Jr and Wetzel 2013) a

specialised R–package, r4ss (Taylor et al. 2014), has been developed. The second
point concerns derived results, such as next year's qouta. This second point has
been a topic for discussion in various settings (e.g. Butterworth and Punt 1999)
and frameworks such as FLR (Kell et al. 2007) have been developed to take into
account various sources of variation in management strategy evaluations.

The appropriateness of the model assumptions is often tested using simu-
lation methods such as those described in Magnusson et al. (2012). This can
also be done using more detailed operating models (e.g. IWC 2008, and related
work) to which the management strategy is tested. Methot Jr and Wetzel (2013)
introduces a simulation model inspired by stock synthesis written in R (R Devel-
opment Core Team 2011) that allows an end to end evaluation of management
strategy under various assumptions of model error. These approaches can also
be used to investigate hypothesis regarding stock structure, as illustrated in
Elvarsson (2015).

Here a specialized R–package, named RGadget, is introduced. RGadget,
as the name suggests, is designed to compliment the Gadget framework by
providing aid for the development of Gadget models and facilitate testing of
said models. This includes an objective procedure to assign weights to likelihood
components, routines to collect output and assess the general fit of the model
to data, an independent implementation of Gadget in R and other useful tools.
This paper is structured into five different components based on the functionality
of RGadget:

- A general overview of the package and how it aids in the process of build-
  ing a typical Gadget model. This summarises the core functionality of
  RGadget upon which the subsequent items are built.

- The assignment of weights to the various likelihood components. This
  part details the iterative reweighting heuristic and procedures related to
  this approach.

- Methods to illustrate goodness of fit. This part introduces methods to
  collect and illustrate how one would present fit to various likelihood com-
  ponent graphically.

- Presentation of model results, that is derived population estimates, xpro-
  jection of stock status, and yield per recruit.

- A brief description on an independent implementation of Gadget in R.

The methodology implemented in RGadget is illustrated using tusk (*brosme
brosme*) in Icelandic waters (described in ICES 2014).

III.2   The tusk example

Tusk is a cod-like fish species found in the North Atlantic. It is fairly common
on the Icelandic continental shelf area and typicall found in depths ranging from
30 to 500 m. It can be fairly large, with a maximum recorded length of 120
cm, while the mean length in the fishery is typically around 55 cm (ICES 2014).
Ageing of tusk caught in Icelandic recently begun anew after a series validation
excercises (Thordarson pers. comm) but otoliths have been sampled since 2000.

The model for tusk is described by Þórðarson et al. (2011) but generally
follows the same structure as the model for cod described in Paper II, section
II.6. The data used by the model are described in Table III.1 with the estimated
parameters given in table III.2.

Origin	Time-span	Length group size	Num. data-points
	Length distributions:		
March Survey	1$^{\text{st}}$ quarter, 1985 – 2013	1 cm	1084
Longlines	All quarters, 1982 – 2013	1 cm	2418
	Age – length frequencies		
March survey	1$^{\text{st}}$ quarter, 1989 – 2013	1 cm	1711
Longline	All quarters, 1994 – 2013	1 cm	1079
	Survey indices		
March survey	1$^{\text{st}}$ quarter, 1985 – 2013	Table III.3	203

Table III.1: Overview of the likelihood data used in the model. Survey indices are
calculated from the length distributions and are dis-aggregated ("sliced") into seven
groups (Table III.3). Number of data-points refer to aggregated data used as inputs
in the Gadget model and represent the original data-set. All data can obtained from
the Marine Research Institute, Iceland.

III.3   Structure of Rgadget

The stock assessment process involves a detailed scrutiny of the applicability of
a particular model as a basis for a management descision. Typically the appli-
cability for a particular stock assessment model is evuluated based on a series
of comparisons to observations on the resource in question(e.g. ICES 2014, and
similar documents). As illustrated in figure III.1 two key aspects of a Gadget

Description	Notation	Comments	Formula
Natural mortality	M_a	Fixed at 0.15 for ages 3 to 20	See eq. II.1
Growth function	K, L_∞	Estimated from age–length frequencies, L_∞ is fixed at 200cm	See eq. II.2
Growth implementation	β	n is fixed at 15 length-groups	See eq. II.3
Fleet selection	$b_f, l_{50,f}$	One set for each of the fleets (Survey and Long-line)	See eq. II.10
Number of recruits by year	R_y	$y \in [1982, 2012]$. σ_y^2, i.e. variance in recruitment length, based on length distributions obtained in the autumn survey.	See eq. II.5
Initial abundance at ages 3 – 20 in 1982	η_a	$a \in [3, 25]$. σ_a^2, i.e. variance in initial length at age a, based on length distributions obtained in the spring survey.	See eq. II.6
Survey catch-ability	q_f	Intercept term in a log–linear relationship with abundance. The slope term is assumed to be 1 for all indices.	See eq. II.12
Length–weight relationship	μ_s, ω_s	Different values by stock component, estimated outside of the model	See eq. II.11

Table III.2: An overview of the estimated parameters in the model.

model need careful consideration prior to parameter estimation. These are a description of the key processes modelled through species and fleet interactions and the spatial scale, and how this links to observed data. In Kupca (2006) the development of specialized tools to deal with processing data for ecosystem models, such as those developed using Gadget, is described. RGadget complements these tools by adding functionality to read in and directly manipulate the Gadget input and settings file structure.

RGadget does this by defining a specialized Gadget – model class within R which mimics Gadget's file structure, which is described in Begley (2005), and broadly follows the dimensions decribed in figure III.1. Area and time files define the time period and spatial units of the model. Stock files define the stocks considered by the model and their interactions (if any) and fleet operations are defined similarly.

Rgadget can also control Gadget using system calls[1] and set up the optimizer for a likelihood fit. Gadget-models can produce output by a number of dimensions which is defined through specialised "printers" (see chapter 9 of Begley 2005). Output can be aggregated according to desired dimensions but digesting the output can however be an excercise in bookkeeping. To alleviate these problems a printer control file can be automatically generated based on the model setup using Rgadget.

RGadget aids in the process of compiling datasets for these comparison in three ways:

- By reading in the input data of an Gadget model

- By preparing the specialized printers based on the model particulars.

- By compiling the data and importing into R

Furthermore higher level functionality has been created to combine the observations to the simulated output. In addition to file manipulation, the creation of basic Gadget-setting files can be scripted directly from R eliminating a major source of errors.

---

[1] This has been tested under various *nix variants and windows.
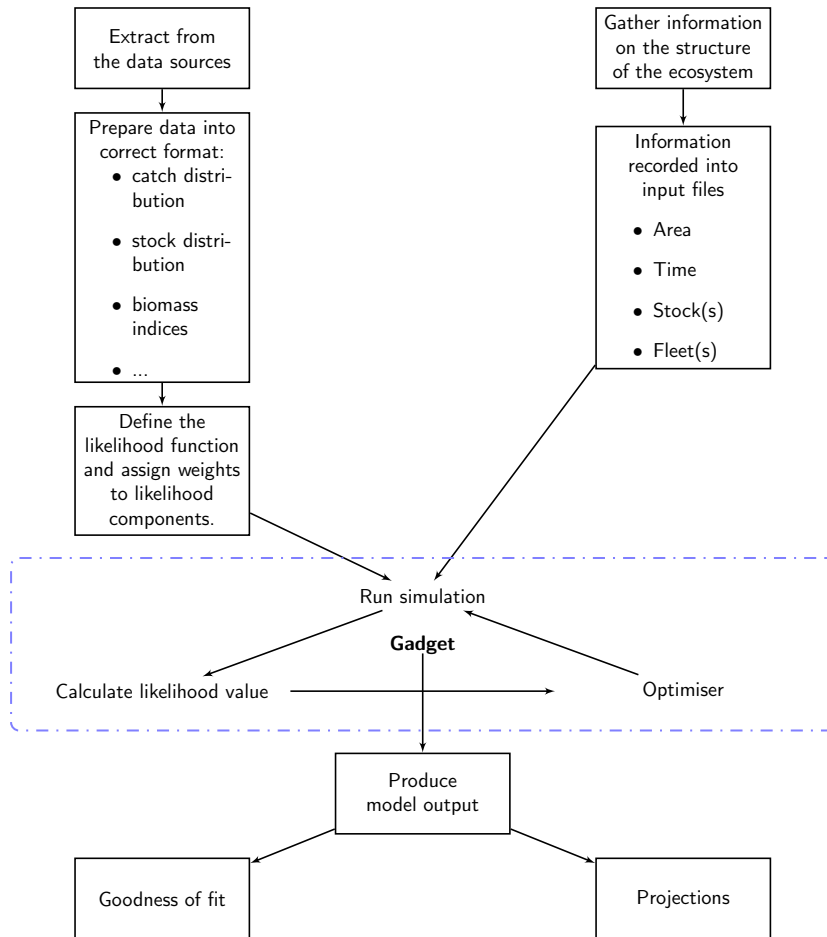
Figure III.1: A overview of a typical model estimation procedure within Gadget.

III.4   Model fit

III.4.1   Iterative reweighting

One of the main aims of a multi-species model, such as those implemented using Gadget, is to estimate values of selected unknown parameters. The likelihood function serves as a general measure of how well a model with a given set of parameters fits data and parameter estimation is therefore undertaken by maximizing the likelihood function over values of the unknown parameters.

The form of the likelihood function for a particular model and data set will vary depending on the nature of the data. Since fisheries data come from various sources, a large number of different likelihood functions have been implemented in Gadget. When such different data sources are combined in on analysis, the likelihood function becomes a product of the likelihood function for each data set. The individual pieces are referred to as likelihood components.

As is common practice, maximum likelihood estimation of parameters is implemented in Gadget through minimizing the negative log–likelihood. The negative log–likelihood function will referred to as the objective function. Thus the objective function serves as a measure of the discrepancy between the output of the model and measurements.

As noted in the introduction, several components enter the objective function in any single estimation. Therefore the objective function becomes a weighted sum of several components:

$$l = \sum_i w_i l_i$$

The weights, w_i, are necessary for several reasons. Notably, they can be used to prevent some components from dominating the likelihood function, to reduce the effect of low quality data and as *a priori* estimates of the variance in each subset of the data.

In this setting the assignment of these weights is, as noted above, generally not trivial, except in the case of a weighted regression. In Taylor et al. (2007) an objective reweighting scheme for likelihood components is described for cod in Icelandic waters. A simple heuristic, where the weights are the inverse of the initial sums of squares for the respective component resulting in an initials score equal to the number of components, is sometimes used. This has the intutitive advantage of all components being normalised. There is however a drawback to this since the component scores, given the initial parametrisation, are most likely not equally far from their respective optima and this in turn results in a sub-optimal weighting.

The iterative reweighting heuristic (described first in Stefansson 2003, which is inspired by the weighted regression case) tackles this problem by optimising each component separately in order to determine the lowest possible value for each component. This is then used to determine the final weights. The reasoning for this approach is as follows:

Conceptually the likelihood components can be thought of as residual sums of squares (SS), and as such their variance can be esimated by dividing the SS by the degrees of freedom. The optimal weighting strategy is then inverse of the variance. Here the iteration starts with assigning the inverse SS as the initial weight, that is the initial score of each component when multiplied with the weight is 1. Then a series of optimisation runs for each component with the intial contribution for that component to the objective function is set to 10000, while other component contribute only 1. After this series of optimisation runs the inverse of the resulting mininum SS is multiplied by the effective number of datapoints and used as the final weight for that particular component.

The effective number of datapoints is used as a proxy for the degrees of freedom is determined from the number of non-zero datapoints. This is viewed as satisfactory proxy when the dataset is large, but for smaller datasets this could be a gross overestimate. In particular, if the survey indices are weighted on their own while the yearly recruitment is esimated they could be overfitted. If there are two surveys within the year Taylor et al. (2007) suggests that the corresponding indices from each survey are weigthed simultaneously in order to make sure that there are at least two measurements for each recruiting yearclass, this is done through component grouping. In general when there is a chance of overfitting, e.g. the model has flexibility to almost perfectly predict the observations it worth while to consider grouping together related datasets.

III.4.2   Investigating model fit

To give an illustration of an application of the iterative reweighting heuristic, it was applied to the tusk example. Component grouping for the tusk example, which is described in table III.3, was used for two reasons. First, as indicated above, when survey indices are emphasised the model has the possibility to follow the data almost perfectly. Therefore the survey indices are emphasized as a single group. Second reason is lack of data on growth, as there age reading for tusk in Icelandic waters only just begun (Thordarson pers. com.). In this application age readings are grouped with length distributions by either survey or commercial catches depending of the origin of the sample. The rationale here is to group together data from similar orgins to prevent unwanted biases.

Component group	Data set
Sind	SI 20-39
-	SI 40-69
-	SI 70-110
Survey	LDs
-	ALKs
Catch	LDc
-	ALKc

Table III.3: Likelihood component groups for the tusk model.

Appart from stability in the model optimization, additional analyses can be obtained from the interim steps of the reweighting procedure. Table III.4 compiles the individual likelihood scores for all components of the likelihood when the fit to a particular dataset is emphasized. Thus this table adds additional information on the goodness of fit by both illustrating how far the final model estimate is from the optimum fit for a particular data set, and by identifing the potential inconsistencies between data sets. This inconsistency can be detected by studying the component scores when one component is emphasized.

	ALKc	ALKs	LDc	LDs	SI 20-39	SI 40-69	SI 70-110
Catch	7787	9158	26720	50630	15.91	36.19	83.25
Survey	8706	4944	154900	23030	14.46	36.22	99.91
Sind	11990	10780	113400	74260	1.797	4.335	17.75
Final	7676	4993	22480	15950	2.641	4.817	21.03

Table III.4: An example fit diagnostics for tusk in Icelandic waters from the iterative reweighting procedure. Rows indicate the weighting group, as described in table III.3, and columns the final score from the data set when a particular weighting group is emphasized.

RGadget provides tools to illustrate the various types of fit to data. A general overview plot is illustrated in Figure III.2 that shows an example of a squared residual fit to length distributions, and age – length distributions, by time step. In this particular case this figure illustrates potential outliers in length distributions and also the periods where age data are not available. In general this figure can provide insights into potential data discrepancies and aid in the identification of potential causes of a poor fit.
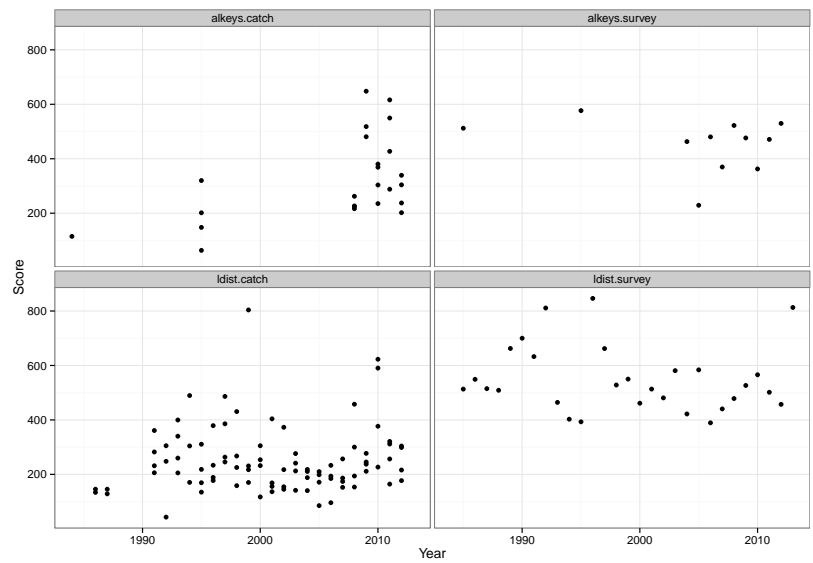
Figure III.2: Likelihood summary plot for the tusk example. The points inticate the
score at a particular time–step facetted by the liklihood component.

In addition to the general overview of the model fit, fit individual data sets can be produced and illustrated in a number of ways. RGadget provides rudimentary plots for many data types. An illustration of how a survey index is processed and plotted can be seen in Figure III.3, where the observed and fitted survey index for tusk are shown. Another example of how length distributions are processed is given in Figure III.4, where the observed and fitted survey length distributions for tusk are compared by year.
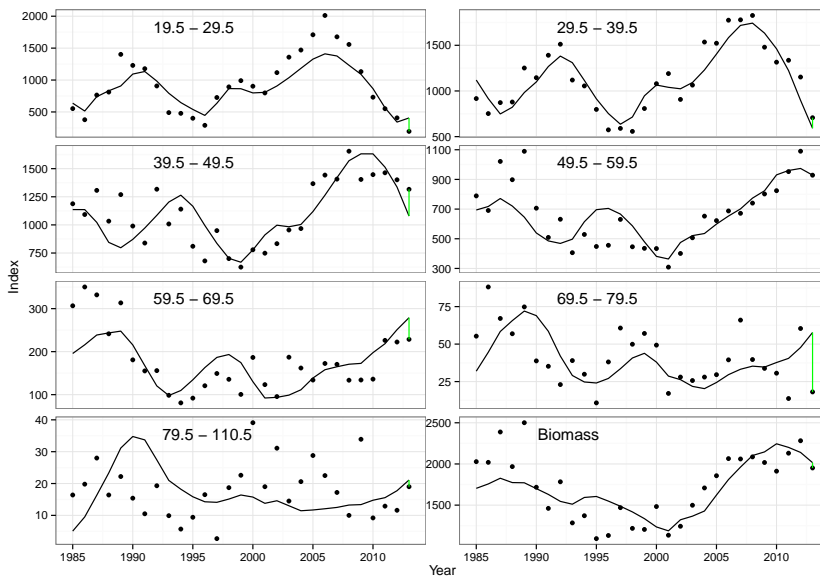
Figure III.3: Surveyindices by lengthgroup for the tusk example. Points indicate the observed values and lines fitted value as a function of year. Green line indcates the deviation of the fit from the observation in the terminal year.
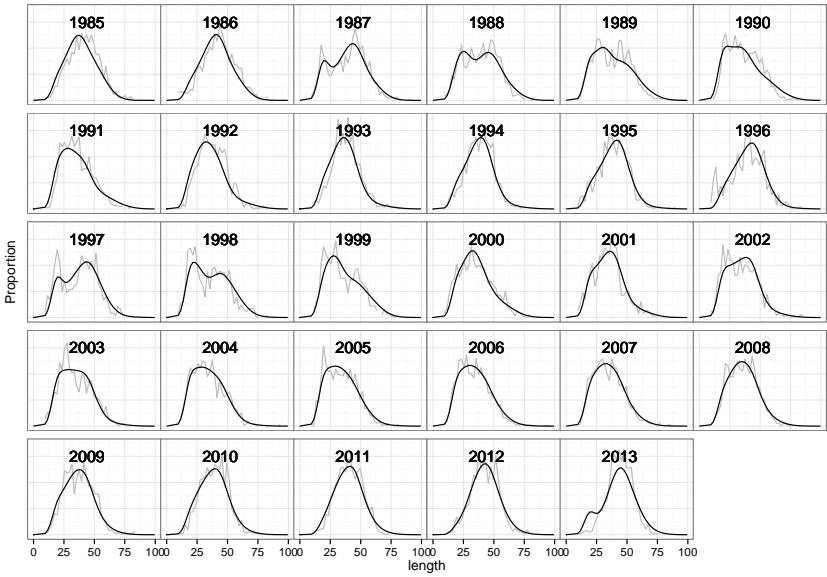
Figure III.4: Length distributions from the survey fleet for the tusk example. Points represent observed lenghtdistributions while the solid lines fitted.
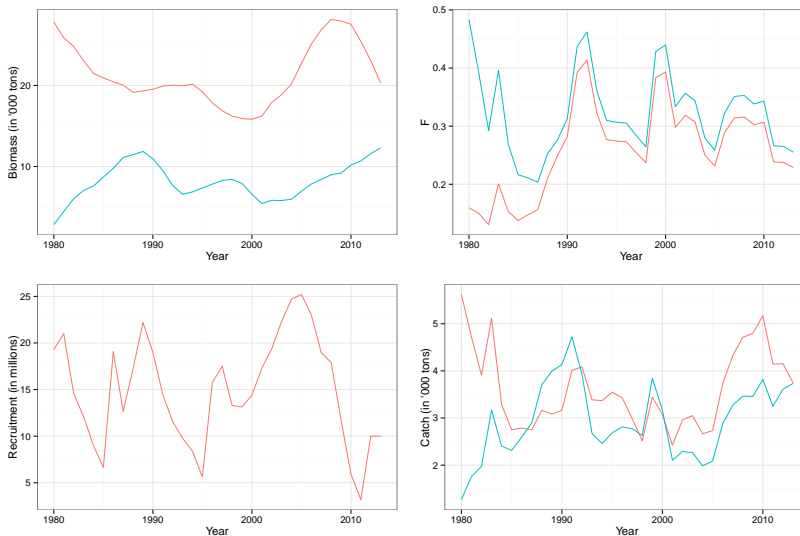
Figure III.5: An illustration of a typical management overview of stock status (here tusk). The top left panel shows the estimated biomass , top right the fishing mortality, bottom left recruitment and bottom right catches. In all panels the trajectories corresponding to the immature (red) and mature (blue) stock components.

III.5   Model results

In a typical stock assessment setting the manager is often interested in derived model quantities, i.e. not actual fit to data per se but in some sense the general status of the exploited resource. Commonly some sort of comparison of current state to a precautionary limit is made. ICES (2013) defines a number of those precautionary limits in terms of a single species stock assessment. RGadget provides tools to facilitate the development of these precautionary limits and general presentation of these results. These methods implement mainly the compilation of the stock status (recruitment, biomass estimates, etc), projection of the stock status (both deterministic and stochastic) and a yield per recruit function. Typical output used in the ICES assessment processs is shown in figure III.5.

Management reference are commonly derived from an assessment model. ICES (2013) suggest developing these on the basis of an analysis of the yield per

recruit of the resource in question. Yield-per-recruit is calculated by following
one year class through the fisheries and calculating the total yield from that
year class as function of fishing mortality of fully recruited fish. In the model,
the selection of the fisheries is length based, as illustrated for the tusk model in
Figure III.6, so only the largest individuals of recruiting year classes are caught
reducing mean weight of the survivors, more as fishing mortality is increased.
In contrast the with age based yield-per-recruit, where the same weights-at-age
are assumed in the landings independent of the fishing mortality even when the
catch weights are much higher than the mean weight in the stock. In general
YPR-curves estimated as in Gadget should give a more conservative estimates
(lower) of $F_{0.1}$ and F_{max} than equivalent age based assessment.

To properly estimate F_{msy} the stock status could be projected forward,
as function of fishing mortality, taking into account sources of variations (as
described in Francis and Shotton 1997) in addition to the identification of a
spawning stock recruitment process. To estimate the uncertainty Paper II in-
troduced an specialised bootstrap for fisheries models based on spatial units.
RGadget, in conjuction with Kupca (2006), allows for the use of this procedure
in a fairly automated fashion once the model has been built. This is a two step
process:

1. The DST^2 db system (as described in Kupca 2006) generates a list of
   bootstrapped spatial units and compiles the bootstrap replicates.

2. The Gadget model is fit to each replicate data set using the iterative
   reweighting heuristic.

RGadget handles the book–keeping for the second step and compiles summaries
of fit and biomass, and the uncertainty of those estimates.

Apart from the model error, although to some degree it can be analysed
using the diagnostic from the weighting heuristic, the biggest source of error
unaccounted for is process error, which mainly error involving stock recruit-
ment. In RGadget a special function sets up the forward projections based
on either on a particular parameter vector or the set of parameters from the
bootstrap. It has the option of doing a deterministic projections, where the
recruitment going forward is based on the average recruitment for the past few
years. Stochastic simulations are implemented to a degree and were applied by
Elvarsson and Þórðarson (2014). There variations in recruitment was modeled
as an autoregressive process with lag 1:

$$\bar{R}_y = \bar{R}_{y-1} + \epsilon_y \tag{III.1}$$
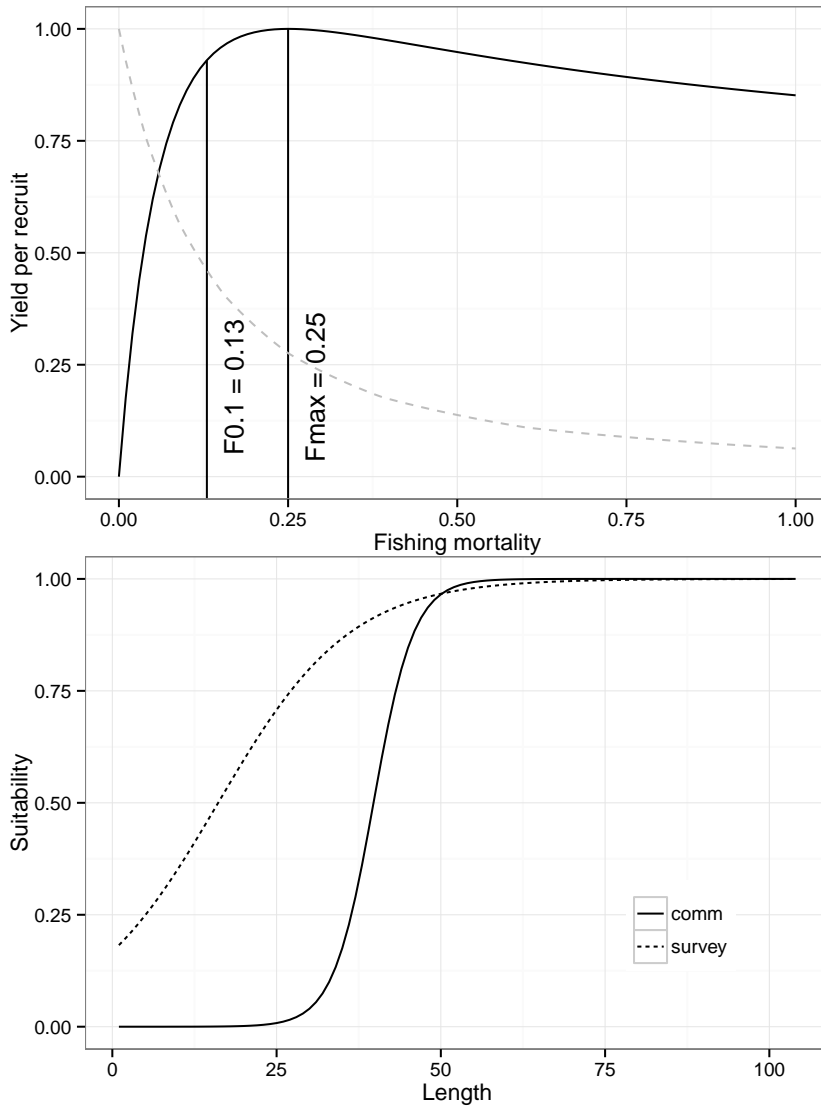
Figure III.6: Top panel shows a yield per recruit curve for the tusk example. The curve represents the biomass caught as a function of fishing moratility. The broken line respresents the proportion that reaches maturity and typical reference points (F_{max} and $F_{0.1}$) are indicated by vertical lines. Bottom show the selection by fleet, commercial and survey, as modelled in the tusk example.

where ϵ_y is a mean zero gaussian with variance σ^2. This is of course a crude approximation to the recruitment process that assumes that recruitment is not impaired during the projection period. A more general spawning–stock recruitment relationship allowing for variation in the number of recruits requires additions to the Gadget framework.

III.6   Age–length simulation model

The Rgadget simulator replicates some of the more commonly used features of Gadget, with use of some of these features optional. The most complex simulation model possible can consist of a number of substocks, where maturation can be modelled with a younger substock maturing into the older substock at a given age and timestep. The stock units are subject to growth (as described by Stefansson 2005b), natural mortality, and they live on a number of areas. A number of fleets can be defined where the harvesting scheme is user defined. Migration between the areas can be defined on a stock basis. Predation can also be implemented between stocks. For example the mature substock can predate upon the immature stock. As with Gadget, the fleets are modelled as predators and the same selection functions can be applied (as described in ch. 4 of Begley 2005). The order in which the population processes are modelled on each timestep is identical to that in Gadget. A RGadget simulation can be entirely created from a simple collection of R objects or it can be started from Gadget input files, although not all options of Gadget itself are supported.

III.7   Discussion

RGadget provides utilities for rapid developement of Gadget models and the subsequent analysis. Hopefully this will encourage the application of Gadget in new areas. During package development, parts of RGadget have been used in various settings. The R simulator has seen some published applications, when tagging was developed in Gadget (see Hannesson et al. 2008, for further details) and in a simulation study on fin whale stock structure was presented (Elvarsson 2015) . The iterative reweighting procedure has seen many applications during its developement (e.g. Elvarsson et al. 2014a, Elvarsson and Þórðarson 2014, Þórðarson and Elvarsson 2013, to name a few).

   Rapid model developmentis important for models of the complexity of a typical Gadget model. For instance scale changes can be fairly cumbersome

to implement and errorprone, while necessary to prevent unecessary noise in the input data (Demyanov et al. 2006). dst^2 (2004) saw the development of specialised database system (Kupca 2006) which allows for exactly this kind of analysis and flexible extraction of ecosystem data. This database system is currently being revised as part of MareFrame (2013). RGadget complements these developments by providing a scriptable generation of model settings for Gadget and digestion of results.

The "weights issue" is really only a major issue when there is *apparent inconsistency between the data sets*. Note that this is a very careful choice of wording. Basically, very different results arise from different choices of weights only if the model can not fit well to all data sets simultaneously. In actual fact what this means is that the model can not explain all the data sets at the same time, i.e. there is something missing from the model. For example, when a really good survey and a typical catchability trend in commercial CPUE data are modelled together they will seem to be inconsistent when a model using constant catchability in both – the inconsistency goes away when a time trend in q is added for the CPUE data.

It is not just important to get the weights right, it is equally important to find where these discrepancies lie. The easiest way to do that is probably by tabulating the component scores as illustrated in table III.4. When datasets are inconsistent the table can then be used to find where the model needs to be made more flexible.

The methodology implemented in RGadget for the weighted likelihood function is certainly not the only one possible. A number of possible approaches could be considered here, as an example Wang and Zidek (2005) select these weights using cross-validation whereas Francis (2011) proposes a weight selection method based on repetitive assignment of weights until the weights converge.

The utility of an independent Gadget-like simulation engine in R is considerable. Appart from the obvious consistency check, it can be used as a tool to test the feasibiltity of potential new features for Gadget. Additionally it could be used as a tool to study different assumptions of the interaction between the two substocks and harvesting strategies as done in Elvarsson (2015). As such the simulator could be used as teaching aid in fisheries science. Data can also be generated for complicated modelling frameworks such as Gadget. For example in (Hannesson et al. 2008), an earlier version of the simulator was used to simulate a tagging experiment designed to test Gadget's performance.

The development of RGadget echoes efforts of other similar projects. Notable examples include r4ss (Taylor et al. 2014) and to a lesser extent FLR (Kell et al. 2007) as these packages aim to standardise model results in the general stock

assements and management strategy evaluations process. Both RGadget and
r4ss are designed around their respective modelling frameworks, i.e. Gadget and
Stock synthesis. FLR appears to be aimed for a more general integration with
age based stock assessment model such as extended survivor analysis (XSA)
(Shepherd 1999) and State space Assessment Model (SAM) (Nielsen and Berg
2014). Howell and Bogstad (2010) developed a link between Gadget and FLR
for a harvest control rule evaluation. Further integration may be considered at
some stage and RGadget could in some sense be considered a first step in that
direction.

III.8   Acknowledgements

IV

Paper IV

Evaluating stock structure hypotheses using genetically determined close relatives: A simulation study on North Atlantic fin whales

Bjarki Þór Elvarsson

Abstract

Certain facets of the population dynamics of a species are hard to quantify, including stock structure. In particular geographical boundaries of stocks or populations are often hard to estimate. This document discusses the application of a recent tagging method, applicable when breeding populations overlap on feeding grounds. The tagging efficiency is augmented with information on genetically determined close relatives. The proposed tagging method is studied using simulations. Statistics which can be used to compare rivalling stock structure hypotheses, are introduced and contrasted. The simulation emulates competing stock structure hypotheses for North Atlantic fin whales (*Balaenoptera physalus*).

The results indicate that, in the case of North Atlantic fin whales, a considerable improvement can be made in terms discriminatory power using information on close relatives when compared to more conventional tag–recapture experiments.

IV.1   Introduction

When managing a marine resource a number of complicated processes interact. The management typically encompasses diverse economical and political objectives such as the maximization of the resource's yield and maintaining employment security. Politics and economics aside a rational utilization of the marine resource is often desired, e.g. how much of the resource can be reliably harvested sustainably (Baldursson et al. 1996). For the exploitation of fish and marine mammals managers have, historically, focused typically on issues pertaining to stock assessment while important questions regarding stock structure and distribution remain unanswered. However it is an established result that ignoring stock structure can lead to overexploitation of individual substocks.

Stocks have often been defined by management boundaries, set according to the distribution of the key species of commercial interest (Halliday and Pinhorn 1990). This could lead to separate advice given to the same stock depending on the management area, for obvious reasons. Therefore, when dealing with fish stocks, numerous methods have been employed to test the accuracy of stock definition and borders. Tagging, both mark recapture (Rayner 1940, Chenuil et al. 2000, Laurenson et al. 2005, Peakall et al. 2006, Hannesson et al. 2008) or satellite tracking (Mate et al. 2007, Víkingsson and Heide-Jörgensen 2014, Horton et al. 2011, Matthews et al. 2011), is commonly used to identify individual and stock movement between (and within) areas.

Mark–recapture experiments are however not always suited to track individual movements. For instance small fish may not survive the marking, marks do not survive moulting, are shed for other reasons or larvae may drift between populations. Other factors, not directly connected to the biology of the species can make traditional mark–recapture analysis impossible or ill suited to answer questions regarding stock boundaries. An example of such a case are tag – recapture experiments on baleen whales where a moratorium has been in place since 1986 preventing recapture. Satellite tracking methods are widely used in studies of animal population structure (eg. in Block et al. 2005, Claridge et al. 2009, Combreau et al. 2011, Kennedy et al. 2013, Mauritzen et al. 2002, Seegar et al. 1996). However, this method is difficult to apply to large cetaceans and despite decades of considerable effort in technical developement and trials satellite tracking of cetaceans rarely last longer than few months (eg. Mate et al. 2007, Silva et al. 2013, Víkingsson and Heide-Jörgensen 2014). Therefore satellite monitoring is currently not suitable to monitor annual migration patterns.

Even when tagging is possible it is not always sufficient to detect separate breeding populations. Breeding stocks can overlap on feeding grounds, as sug-

gested by an analysis of otolith classification or genetic structure (Reynolds and Templin 2004, Wennevik et al. 2008, Jónsdóttir et al. 2007). These results indicate that, if neglected, managing two (or more) separate breeding populations could have adverse effects, such as an overexploitation of one breeding population, without being detected by conventional assessment methods (Punt 2003). Genetic differences have been used to suggest separate breeding populations of marine mammals (Andersen et al. 1997, Bérubé et al. 2002, Parsons et al. 2006, Fontaine et al. 2007, Pampoulie et al. 2008).

Despite considerable efforts through decades, traditional population genetic studies have in many cases failed to give unequivocal answers to important questions concerning cetacean stock structure (Donovan 1991). A major obstacle for interpretations of these studies has been the fact that for most baleen whales the breeding grounds are unknown and sampling has thus been restricted to the summer feeding grounds. Furthermore, large baleen whales, such as the North Atlantic Fin whales, do not often exhibit sufficient genetic variability to detect separate breeding populations (Bérubé et al. 1998, Pampoulie et al. 2008). The possibility of two or more breeding populations can, however, not be ruled out as the hypothesized split occurred relatively recently and the stocks have not had time to detectably differentiate (Pampoulie et al. 2008).

In the absense of detectable genetic structures genetic tagging (as described by Palsbøll 1999) or other genetic methods such as information on close relatives could be used to answer question related to stock structure (Skaug 2001, Palsbøll et al. 2010, Nielsen et al. 2001), for instance the effect of larval drift between different breeding stocks (Planes et al. 2009). A simulation experiment of its potential application in management of marine mammals can be found in Økland et al. (2010), where managment units are defined for geographically segregated stocks using genetically determined close relatives. The authors, however, note that their method would hardly be applicable to stocks that overlap on feeding grounds whilst separate on breeding grounds, as is common for baleen whales, due low discriminatory power.

Here a simulation study of a (genetic) mark–recapture experiment, that has been augmented using information on genetically determined close relatives, is described. The simulation is based on a marine mammal population, the North Atlantic Fin whales, where it is assumed that the animals migrate between feeding and breeding grounds where they overlap to a varying degree. The tagging experiment aims to answer important management questions regarding stock structure. Relevant stock structure hypotheses are introduced and constrasted using three test statistics. The resulting analysis provides a power analysis of the comparison of the competing hypotheses as a function of sample size.

IV.2   Methods

IV.2.1   The setting

The North Atlantic fin whale has a wide distribution in the North Atlantic and in a recent NASS surveys the total abundance in the Central North Atlantic was estimated to be around 22.000 – 25.000 animals in 2001 (Víkingsson et al. 2009, Pike et al. 2008b). Commercial fin whale operations started in the late 19th century. However after a collapse in fin whale abundance (according catch series data) the Icelandic parliament issued a ban on all whaling activities in 1914 (Sigurjónsson and Gunnlaugsson 2006). When whaling resumed in 1948 the fin whale stocks in Icelandic waters had made a significant recovery (IWC 2008). Since 1986 a whaling moratorium has been in place within the IWC. In 2006 the Icelandic government lifted the ban on commercial whaling and began issuing an annual quota of 150 fin whales, while the actual takes have been considerably less (MRI 2013).

The fin whales in the North Atlantic are managed by the IWC using the Revised Management Procedure (RMP) (as described in IWC 2013, and references therein). In the implementation simulation trial (IST) for the fin whales (IWC 2008) the performance of several management variants were tested with respect to seven possible stock structure hypotheses.

Tagging studies in Icelandic waters on the North Atlantic Fin whales have suggested that fin whales exhibit some site fidelity from year to year (Víkingsson and Gunnlaugsson 2005). For management areas in and adjacent to the Irminger sea, that is East–Greenland (EG), West–Iceland (WI) and East–Iceland (EI), two different stock structures hypotheses have been suggested that could explain the fin whale distribution. The first is a mixing hypothesis, where it is assumed that the fin whales in the waters around Iceland and East–Greenland originate from three separate breeding stocks, with no dispersion (no sharing of genetic material) between the breeding grounds. It is assumed that the whales choose the next year's feeding ground independently, i.e. with no "memory" of where they were this year. The other hypothesis assumes that there is some degree of dispersion on the breeding grounds. Under this hypothesis whales choose their feeding ground based on last year's feeding ground and breed with individuals in adjacent substocks. An illustration contrasting the different feeding migrations patterns is shown in figure IV.1.

To compare the two stock structure hypotheses of dispersion vs mixing, a genetic tagging experiment (biopsy collection) in the East–Greenland subarea has been suggested. During which time, for a ten year period, a quota of 150
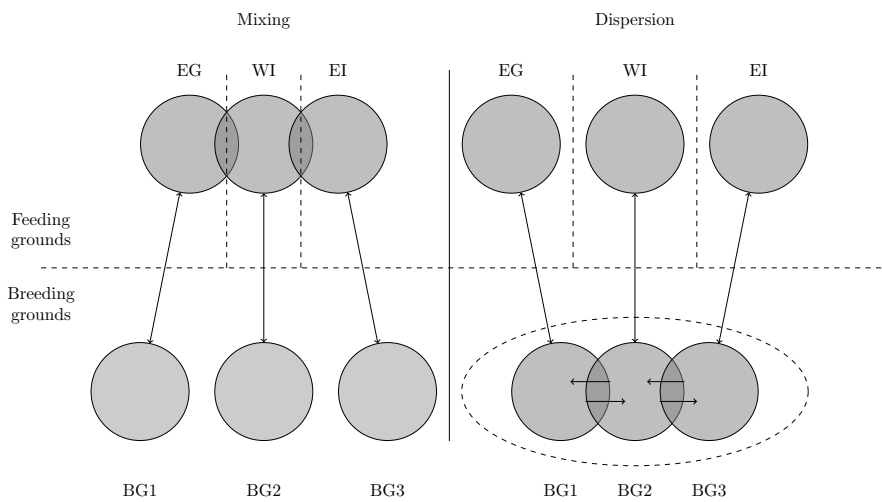
Figure IV.1: An overview of the competing stock structure hypotheses for the North Atlantic Fin whales in the Irminger sea. The figures illustrate three breeding components with unknown breeding grounds BG1, BG2 and BG3 that migrate to feeding grounds East of Greenland (EG), West (WI) and East of Iceland (EI). The figure on the left illustrates the mixing stock structure, where there is no gene sharing with some degree of overlap on the feeding grounds (shaded region). The figure on the right shows the dispersion hypothesis, i.e. genetic sharing between the breeding components, where arrows indicate movement between breeding components.

fin whales annually would be set in the West Iceland subarea. The tagging experiment is designed is order to reject the mixing hypothesis in favor of the dispersion hypothesis, if possible. The simulation model, and parameter settings, for this analysis is based on the results from implementation simulation trials for an RMP implementation of NA–fin whales (IWC 2010). The reason for this particular methodology relates to the types of data that can potentially be obtained, limited by the moratorium and the biology of the species.

Augmenting the genetic tagging experiment with information on close relative would potentially show greater contrast between the two hypothesis than a conventional tag–recapture experiment. It is expected that the sample size will increase, as methodology also tags the nearest relatives of the tagged individual. In addition to that the effects of the different hypotheses on the recaptures of related individuals are substantially different. The dispersion hypothesis effectively assumes a single breeding stock, which implies a greater number of relatives in all areas, whereas the mixing hypothesis assumes three, implying fewer relatives across areas.

IV.2.2   Simulation

The stock dynamics in this study were implemented in a computer program, `Rgadget` (Elvarsson et al. 2014b), set up in such a way as to closely mimic the dynamics of the Baleen II model as described in Punt (1999). In the analysis which follows comparison will be made on the basis of two possible stock structures, mixing or dispersal type, as shown in figure IV.1. When mixing dynamics are assumed, separate breeding stocks overlap (to some fixed degree) on the feeding grounds, while dispersion denotes the permanent migration between breeding stocks.

The general dynamics of the population is (as in Punt 1999) governed by

the following. The stock size is determined by the following equations:

$$N_{gj,t+1,0} = \frac{b_{j,t+1}}{2}$$

$$N_{gj,t+1,a} = e^{-M} \sum_{j \neq j'} \left[(1 - D_{jj'})(N_{gjt,a-1} - C_{gjt,a-1}) + \right.$$

$$\left. D_{j',j}(N_{gj't,a-1} - C_{gj't,a-1}) \right]$$

$$N_{gj,t+1,x} = e^{-M} \sum_{j \neq j'} \left[(1 - D_{jj'})(N_{gjt,x-1} - C_{gjt,x-1} + N_{gjtx} - C_{gjtx}) + \right.$$

$$\left. D_{j'j}(N_{gj't,x-1} - C_{gj't,x-1} + N_{gj't,x} - C_{gj'tx}) \right]$$

$$(IV.1)$$

where N_{gjta} is the number of animals of gender g, age $a < x$, $x = 25^+$ is the maximum age and stock j at the start of year t. C_{gjta} is the catch in numbers, b_{jt} the number of calves, $M = 0.08$ the natural mortality and $D_{jj'}$ is the stock dispersions from stock j to j'. Furthermore it is assumed that the stocks are at their equilibrium density, with a stock size of 7000 animals for each stock.

The recruitment to the stock is determined by the number of mature females in the stock population.

$$b_{jt} = B_j N_{fjt} \left[1 + A_j \left(1 - \left(\frac{N_{fjt}}{K_j} \right)^{z_j} \right) \right] \qquad (IV.2)$$

where B_j is the average number of births per mature female and year in stock j, A_j and z_j are the resilience and compensation parameters, $N_{fjt} = \sum_{a=a_m}^{x} N_{fjta}$ is the number of mature females in stock j, $a_m = 6$ the age of first parturition and K_j is the number of mature females in the pristine population. Values for the recruitment parameters are chosen such that the MSYR is 0.01 and is obtained at $0.72 K_j$.

Let α denote the mixing proportion, i.e. the proportion of a stock which migrates to a specific different feeding ground (see below), and let β denote the proportion of animals which "move" to a different stock.

The stock distribution on feeding grounds is defined by a mixing matrix $\mathfrak{V} = V_{jk}$, where V_{jk} denotes the proportion of stock j that migrates to area k annually. Under the mixing hypothesis it is assumed that the sub–stocks, even though separate on the breeding grounds, overlap on the feeding grounds. The feeding grounds have been split up into three distinct subareas. The subareas

represent the main feeding ground for each of the sub–stocks. While a simulation baseline $1 - 2\alpha$ of the stocks' individuals migrate to their own feeding ground, regardless of where they were last year, α migrate to each of the areas adjacent to their native feeding ground. The dispersion hypothesis assumes that individuals stray between sub–stocks while the sub–stocks migrate to a fixed feeding area. In general the stock overlap (mixing) is according to:

$$\mathfrak{V} = \left(\begin{array}{c|ccc} & \mathbf{C1} & \mathbf{C2} & \mathbf{C3} \\ \hline \mathbf{EG} & 1-\alpha & \alpha & 0 \\ \mathbf{WI} & \alpha & 1-2\alpha & \alpha \\ \mathbf{EI} & 0 & \alpha & 1-\alpha \end{array} \right) \tag{IV.3}$$

where the columns represent the breeding stocks (C1, C2, C3) and the rows feeding areas (EG, WI, EI).

Under the dispersion hypothesis the annual straying between three sub–stocks is only defined between adjacent stocks.

$$D_{ij} = d_{ij} \frac{K_i + K_j}{2K_i} \tag{IV.4}$$

where

$$d_{ij} = \begin{cases} \beta & \text{if i=1 and j} = 2 \\ \beta & \text{if i=2 and j} = 3 \\ 0 & \text{if } |i - j| > 1 \\ d_{ji} & \text{j<i} \end{cases}$$

Under the mixing hypothesis $\alpha = 0.05$ and $\beta = 0$, while under the dispersion $\alpha = 0$ and $\beta = 0.05$.

Commercial catches from the stocks are calculated in the usual manner:

$$C_{gjta} = \sum_k F_{gkt} V_{jkt} S_{ga} N_{gjta}$$

$$F_{gtk} = \frac{C_{gkt}}{\sum_{j'} V_{j'kt} \sum_{a'} S_{ga'} N_{gj'ta'}} \tag{IV.5}$$

where F_{gkt} is the harvest mortality, S_{ga} is the age–based gender specific selectivity formulated as Punt (1999):

$$S_{ga} = \frac{1}{1 + e^{-(a - a_{g50})/\delta_g}} \tag{IV.6}$$

and V_{jkt} is the proportion of stock j in area k at year t. The values of the gender specific selection parameters are:

Selection:	Male	Female
δ_f	0.57	1
a_{50}	3.6	4.5

Tagging can, as noted in the introduction, be used to estimate stock migrations. Although it is assumed here that all tagging will be made using skin biopsies to obtain genetic material the biopsies can also be used in a conventional mark–recapture analysis (Palsbøll 1999). The dynamics of the tagged sub–population in the simulations is the same as for the untagged population. For the sake of simplicity only a single tagging experiment, conducted in a single area, is considered in this analysis. The initial ($t = 0$) number of tagged animals is distributed across stocks j, ages a and genders g, according to the equation:

$$\mathfrak{T}_{gj0a} = \frac{N_{gj0a} * \phi_0}{\sum_{gja} N_{gj0a}}, \tag{IV.7}$$

where ϕ_0 is the total number of tagged animals. The expected number of animals recaptured is a function of the dynamics applied to the population, both tagged and untagged. The recaptures, $\hat{U}_t$, were considered to be distributed according to

$$f(\hat{U}_t) = \prod_t \frac{\Gamma(\frac{U_t}{\lambda} + \hat{U}_t)}{\Gamma(\hat{U}_t + 1)\Gamma(\frac{U_t}{\lambda})} \left(\frac{1}{1+\lambda}\right)^{\frac{U_t}{\lambda}} \left(\frac{\lambda}{1+\lambda}\right)^{\hat{U}_t}, \tag{IV.8}$$

i.e. a negative binomial distribution with mean U_t, the predicted number of animals recaptured by commercial whaling fleets, and variance $U_t(1 + \lambda)$. Here a negative binomial distribution for the tag-recaptures is used, instead of a more commonly used poisson model (eg. in Hilborn 1990), which is intended to allow for greater variation in recaptures due to clustering. In these simulations $\lambda = 2$.

IV.2.3   Comparing hypotheses

In the setting described above two different stock structure hypotheses are to be contrasted. To compare these stock structure hypotheses three potential methods of comparison are studied here.

- Time–trend analysis using regression.

- Total number of recaptured animals by area.

- Number of recaptured animals by area in relation to number of intra–related individuals within the catch.

To compare the two hypotheses using direct (genetic) tagging a Poisson regression model for a time trend in the recapture rates is often fitted. The Poisson regression deviance function used here is

$$D(\hat{U}, \mu) = 2 \sum_t \left\{ \hat{U}_t \log\left(\frac{\hat{U}_t}{\mu_t}\right) - (\hat{U}_t - \mu_t) \right\}$$

where $\log(\mu_y) = \mathfrak{a} + (\mathfrak{b} - M)t$ and M the natural mortality, as above. Natural mortality is used here to offset the expected number of recaptures due to the effects of M on the number of tags remaining in the population. The dispersion hypothesis is expected to have an increased recapture rate ($\mathfrak{b} > 0$) while the rate should be constant under the mixing hypothesis ($\mathfrak{b} = 0$). The model for positive trend can be compared, using a likelihood ratio test, with a model with no time trend using the following statitistic:

$$\mathfrak{D} = D(y, \mu) - D(y, \bar{\mu}) \tag{IV.9}$$

where $\bar{\mu}_t = \bar{\mathfrak{a}} - Mt$. This model form tends to be more liberal than a regression based on a negative binomial distribution as it would have a higher rejection rate for the null hypthesis. However it is used here due to estimation issues as the negative binomial would require an additional parameter.

In genetic tagging augmented with information regarding genetically determined close relatives (parent-offspring or half-siblings), such as described in Skaug (2001), a skin sample from a single whale can, in the case of NA–fin whales, effectively tag $2.5 - 3.5$ other whales, as shown in Gunnlaugsson (2011). Using information on close relatives time trends in occurrence at feeding grounds are expected to be harder to detect. Intuitively this can be explained by noting that with a dispersing stock relatives are already present at all three feeding grounds at the time of tagging. The total number of caught aninmals that are related to tagged individuals, i.e. the number of **effectively tagged** individuals, is

$$T_{..k} = \sum_t \sum_j T_{jkt} \tag{IV.10}$$

where T_{jkt} is the number of animals related to the tagged animals from stock j and caught in area k at time t $T_{..k}$ should, given a similar degree of dispersion and mixing, be somewhat higher for dispersing stocks than mixing, based on a similar argument as before.

Untagged whales caught are also a source of information regarding the stock structure. Consider the total number of animals caught in area k from stock

j, denoted by c_{jk}. This number is not observable but if one assumes only one genetic relation can be detected per individual within a stock, then the total number of possible detections is $\frac{c_{jk}(c_{jk}-1)}{2}$. One measure of the magnitude of genetic relatedness detected between all whales caught of stock j, denoted R_k, in area k is therefore:

$$R_{jk} = \frac{c_{jk}(c_{jk}-1)}{2n_j} \qquad (IV.11)$$

Note that R_{jk} tends to be smaller as n_j, the total number of individuals in stock j, grows larger. Furthermore the total number of relations detected in the total catch for the time period, which can be observed directly from the catches:

$$R_{.k} = \left(\sum_{j \in \mathfrak{J}} R_{jk} \right)$$

becomes smaller with fixed total abundance as the number of breeding stock decreases.

Using the information on related individuals one can augment equation IV.10 by calculating the following ratio for each area:

$$\rho_k = \frac{T_{..t}}{R_{.k}} \qquad (IV.12)$$

The above quantity should become larger for dispersing stocks as there is genetic interchange, even if the total number of effectively tagged individuals is similar.

For each of the stock structure hypothesis the number of simulated datasets per hypothesis was 100 for each number of tags. The number of tags in this experiment varied between 100 to 1500. The tag–recaptures were simulated using equation IV.8. Under the mixing hypothesis the stock proportions within the catch (c_j from equation IV.11) were simulated using a multinomial distribution parametrised by the expected value of number of individuals caught from each stock. Assuming the dispersion hypothesis of $R_{.k}$, as defined by equation IV.11, is constant. Here $R_{.k}$ is assumed to be a Poisson process with mean $\frac{c(c-1)}{2n}$ where c is the total catch and n is the total abundance of all stocks..

The distribution of the three test statistics was analysed and for the null hypothesis, which is in this case the mixing hypothesis, the rejection interval was chosen in such a way that it would have a rejection probability of 0.05. Using simulated data based on the alternative hypothesis, which is the dispersion hypothesis, the power of the test was calculated as a function of the number of tags.
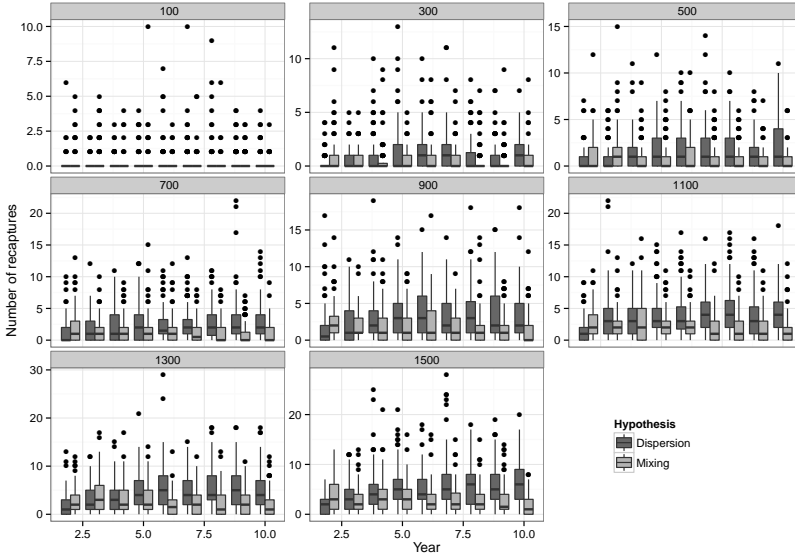
Figure IV.2: Boxplot of the simulated distribution of recaptures in West–Iceland under the two hypotheses, mixing and dispersion, faceted by the number of tags released in the East Greenland area.

IV.3   Results

The simulation method used here allows for a study of the underlying processes that govern the distribution of the various statistics discussed above. Figure IV.2 shows the simulated distribution of recaptures in the West-Iceland area. It illustrates that the recapture rate of contrasting hypotheses have different time trends, the mixing hypothesis has a somewhat downwards trend while the dispersion hypothesis a constant or slightly upwards trend. Similarly, the under the dispersion hypothesis, the range of recapture rates is higher than under the mixing hypothesis.

The tagged population is shown in figure IV.3, where the expected proportion of the tags remaining by area is illustrated. The contrasting hypotheses exhibit different properties, under the mixing hypothesis the proportion tagged spikes in the two areas around Iceland and gradually decreases in all areas due to natural mortality. In contrast the dispersion hypothesis the tagged popula-
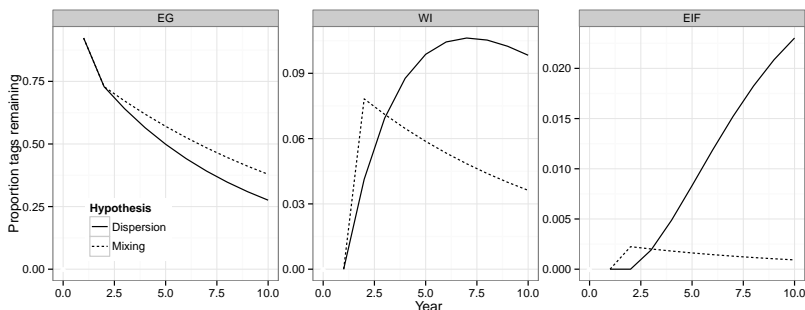
Figure IV.3: Expected proportion of tags remaining in the population by area as a function of year.

tion appears to peak in the WI area in the 7th year and it had not peaked in the 10th year in EIF. In both cases the expected proportion of tagged animals is greater under the dispersion hypothesis after the third year.

IV.3.1   Trend analysis

Figure IV.4 illustrates the discriminatory performance of a conventional tagging experiment. The Poisson regression model appears to have substantial problems with detecting the difference between the two hypotheses at these low numbers of tagged whales and mixing rates. It was observed that, even though the discriminatory power of the increases with number of tags as illustrated in table IV.1, the analysis of trend would at least require 500 tags to have a power greater than 0.25 and hardly goes above 0.4.

Trend is estimated to be significant under both hypotheses. For 100 tags 32% of the simulations the trend was significant and for 1500 tags up to 84% and 42% were significant for dispersion and mixing hypothesis respectively (Figure IV.4, lower panel).

IV.3.2   Total number of recaptured animals

The second metric, that is the total number of recaptured animals (direct or related to tagged animals) as described by equation IV.10 shows greater discriminatory power, as shown in figure IV.5. The number of recaptured animals has the ability to distinguish between the mixing and dispersion hypotheses
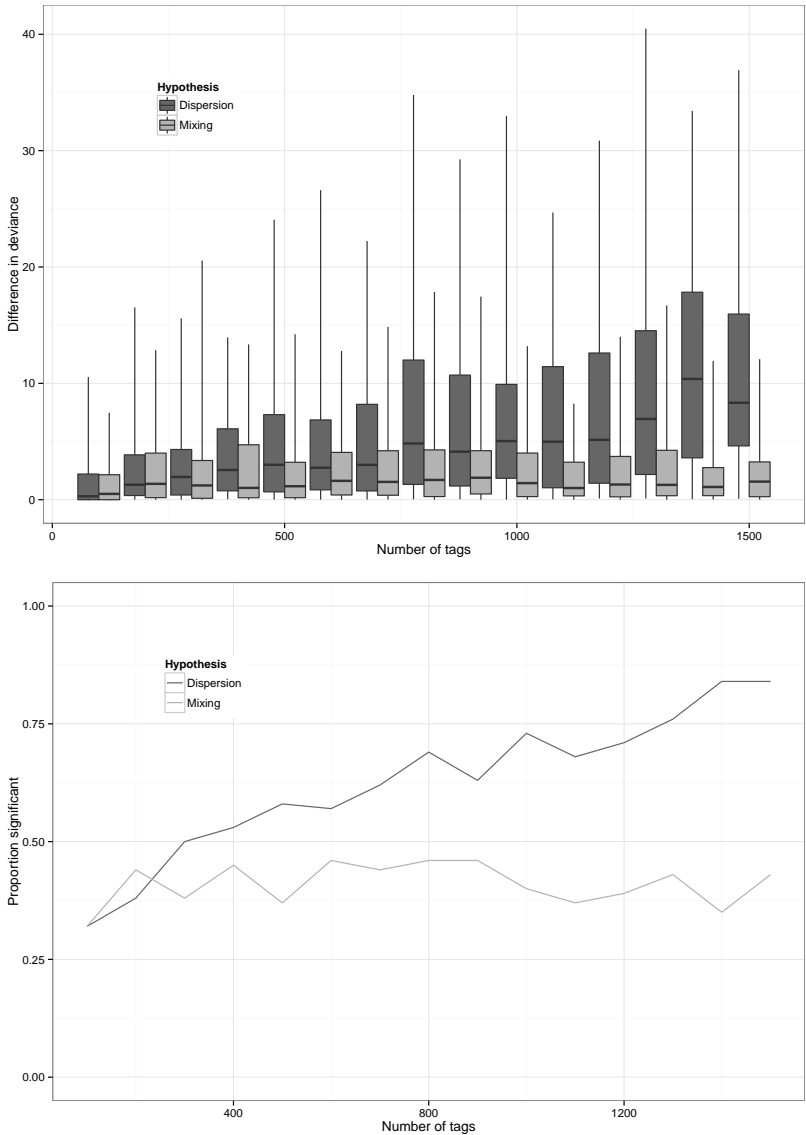
Figure IV.4: The top panel shows the distribution of the difference in deviance of a Poisson regression model of a constant rate of recaptures and a model with time trend by stock structure hypothesis. The distribution of this difference is illustrated using a modified boxplot, where the narrow vertical lines indicate the 95% quantile range and the boxes the interquartile range. The lower panel shose the proportion of simulations where trend in recaptures were estimated to be significant using the possion regression model.
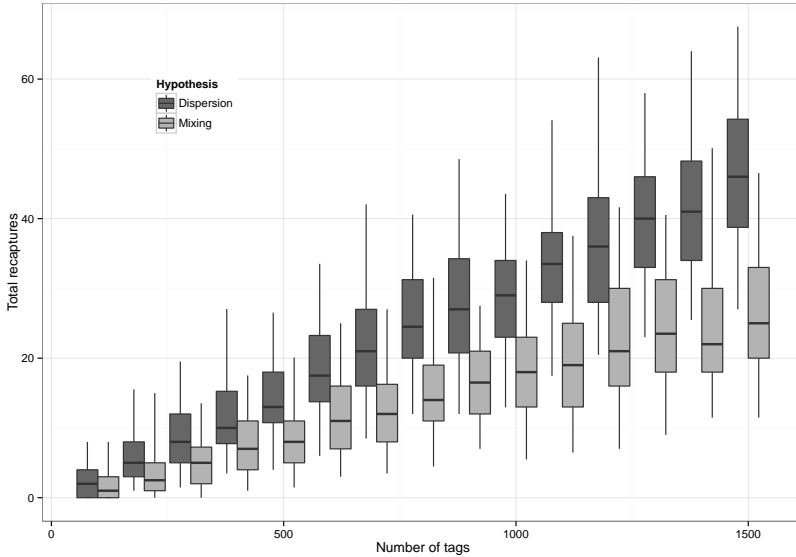
Figure IV.5: The distribution of the total number of recaptures by number of samples and stock structure hypothesis. The distribution is illustrated using a modified boxplot, where the narrow vertical lines indicate the 95% quantile range and the boxes the interquartile ranges.

with close to 800 effective tags ($\approx 200 - 400$ biopsy samples) with a power of 0.42 and level of $\alpha = 0.05$, as illustrated in table IV.1. The total number of recaptures was on average 62% higher for the dispersion hypothesis than it was for mixing for all number of tags.

IV.3.3   Total number of recaptured animals to relatives detected in the catch

The effects of the additional information on relatives detected in the catches are are illustrated in figure IV.6. It appears that the added information improves upon the detection power of the total recaptures. The results in table IV.1 show that the power exceeds 0.8 after 400 effective tags. The expected number of relatives detected is substantially different between hypotheses, under the dispersion hypothesis the number of relations detected in the catches was between
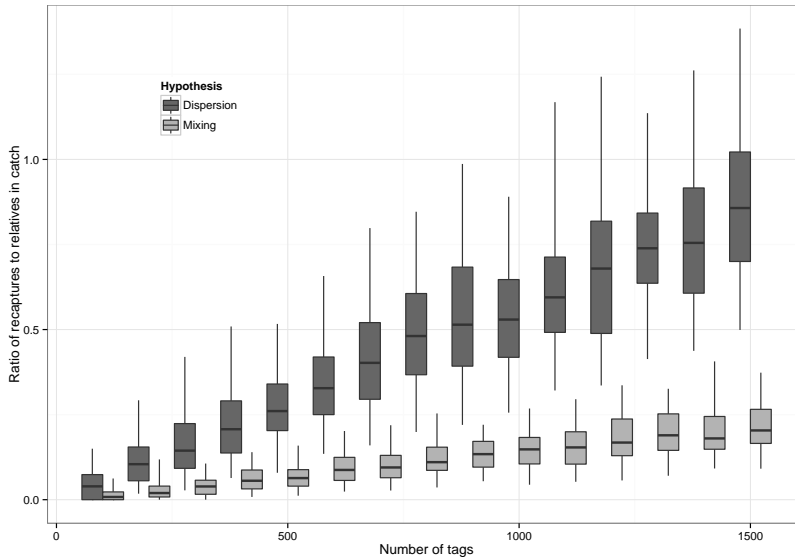
Figure IV.6: The distribution of ρ, the ratio of total recaptures to related individuals within the catch by stock structure hypothesis. This distribution is illustrated using a modified boxplot, where the narrow vertical lines indicate the 95% quantile range and the boxes the interquartile range.

39 to 68 whereas 120 upto 130 were detected under the mixing hypothesis.

IV.4   Discussion

The analysis presented here illustrates a potential use of genetically determined close relatives when determining management units for large baleen whales. It adds to an already described methodology illustrated in Økland et al. (2010) by allowing for stock overlap in the feeding grounds. The methods discussed here need not to be restricted to whale populations. Analysis of stock migrations using close-kin analysis have been used successfully for fish stock migrations, both spawning migrations (Peery et al. 2008) and larval distribution (Planes et al. 2009). The simulation experiment, as described here, is sufficiently general to be applied to management questions regarding many exploitable marine resources. The results could therefore be further extended to other situations,

Table IV.1: The probability of rejecting the mixing hypothesis conditioned on the
stock hypotheses using a direct tag-recapture experiment (time trend in tags), total
recaptures and ρ, the ratio of total recaptures to related individuals within the catch.
The rejection interval was chosen such that the probability of type I error was 5%

Number of tags	Rejection probability		
	Trend	Total recaptures	ρ
100	0.11	0.05	0.39
200	0.09	0.09	0.62
300	0.04	0.27	0.77
400	0.13	0.21	0.81
500	0.21	0.26	0.91
600	0.17	0.25	0.86
700	0.17	0.22	0.90
800	0.19	0.42	0.95
900	0.25	0.51	0.98
1000	0.25	0.35	0.97
1100	0.38	0.58	1.00
1200	0.34	0.41	1.00
1300	0.32	0.54	0.99
1400	0.58	0.59	1.00
1500	0.38	0.60	1.00

while the biology and abundance of the species would need some consideration.

When studying Fin whales in the North Atlantic the methods described here compare favorably to conventional mark–recapture methods simply through the increase in the effective number of samples. Genetically closely related individuals (matched pairs) may have split and dispersed over a longer period than the time lapse between the collection of the samples. This was not modeled here but would potentially give additional strength to the test described by equations IV.11 and IV.12.

Two sources of potential biases need to be mentioned regarding this study. First it is assumed that the simulated whale populations are at their carrying capacity. This assumption is made merely out of convenience, as the birth rates of the whale populations are not considered to influence the results dramatically. The simulated time period is relatively short compared to the lifespan of the whales, who enter the harvest at the average age of around 4 years and the production of the stock is rather low.

The second source of bias is the fact that the stock abundances are assumed to be known. This assumption could, when included into the simulation study, decrease the power of the tests as there is substantial uncertainty in the total abundance in the area. However this is not considered to influence the results substantially as the main driver of uncertainty is the model for recaptures. The expected number of recaptured individuals is proportional to the sub–stock abundance its fluctuations are seen to vary less than the simulated recaptures.

The level of the stock overlap, i.e. mixing rates, would potentially influence the results. At these really low levels the test statistics described in equations IV.10 and IV.12 perform considerably better than a regression model with a time trend. However it is expected that as the mixing rate increases the power of the time trend analysis would stay fixed while at the same time the power of the other two statistics would ecrease. The effect on management in this particular case would however be reduced as the qouta assigned in the area West of Iceland would increase under the mixing hypothesis (as noted in IWC 2012).

With close kin relationships such as utilized here one needs to take considerable care when deciding what level of close relation are included. On one hand it is a question of feasibility as the number of alleles needed for accurate detection of close kin increases substantially as the order of relation increases. On the other hand it depends on the stock structure hypotheses under scrutiny. In this case it is expected, as noted earlier, that in the case of the North Atlantic fin whale the effective number of tags lies between 2.5 and 3.5 per skin sample which could increase the power of the tests by a similar factor. However if one

were to study hypotheses such as male mediated DNA one would need to focus on mother–offspring pairs.

Information on close relatives can, when available give indications on stock movement or overlap. Analysis of the applicability of a methodology capable to determine close relatives have shown promising results for baleen whales, in particular minke whales (Skaug 2001). Data from biological samples obtained from catches of North Atlantic fin whales west of Iceland show promising results (Pampoulie et al. 2011) while further work needs to be done.

Genetic tagging has the potential to improve upon traditional tagging methods. In its most basic form it allows for a nonivasive and non-lethal traditional tagging experiment. This is of course of value for species under a low harvesting rate such a fin whales in the North Atlantic. By augmenting the genetic tagging data with genetically determined close relatives additional strength can be sought, as illustrated here. Furthermore, a possible source of data could be genetic material sampled from earlier periods. Gunnlaugsson (2011) illustrates the possibility of a relative of a whale caught in previous catch periods in Icelandic waters to be still alive in the most recent catches. Although this analysis considers a single tagging experiment in East Greenland with all catches in West Iceland the methodology described here is applicable to different settings. In addition methodological developements, such as those described and illustrated in Benónýsdóttir (2012), will allow for a better understanding of the overall distribution of baleen whales in the North Atlantic.

IV.5   Acknowledgements

V

Paper V

Adding data on species interaction: A case study on cod (*Gadus morhua*) and common minke whales (*Balaenoptera acutorostrata*) interactions in Icelandic waters

Bjarki Þór Elvarsson, Gísli Víkingsson and Gunnar Stefánsson

Abstract

Ecosystem approaches to fisheries management call for complicated models for multi–species interactions. However, data on many of these interactions are often a limiting factor as they require extensive analysis of stomach content data. Direct sampling is typically fraught with difficulties both due to technical and political reasons. Changes in the environment can also play a role as the environmental changes may affect species differently. In the case of cod in Icelandic waters one of the key uncertainties identified as the consumption by baleen whales, common minke whales in particular. Common minke whales are the most abundant baleen whale species in Icelandic waters. A recent study on the feeding ecology suggested a high degree of gadoid consumption. The results from this survey were assimilated into a model of Cod–Minke whales interactions in Gadget, a statistical framework for age – length based marine ecosystems models. A study is presented on what effect including minke whales into a model for cod has on the model characteristics.

V.1   Introduction

Typically when modeling a marine resource, the models are used to determine
satisfactory utilization levels, such as total allowable catch, for the exploited
resource. Due to this management perspective these models tend to focus solely
on a particular species and its associated fishery (Möllmann et al. 2013). Single
species models offer a simplistic view of the resource that can, while convenient,
fail to predict changes due to interaction with other species (e.g. due to fac-
tors suggested Víkingsson et al. 2014). In recent years there has been a call
for the inclusion of multi–species considerations and a move towards an ecosys-
tem based approach to management. These approaches attempt to account for
the single species models' shortcomings by capturing some or all interactions
between species, whether indirect or direct.

  Ecosystem considerations appear to be an obvious, and necessary, part of
the management of marine resource. Optimum fisheries management based
on ecosystem considerations is however an elusive target. Even if the model
sufficiently describes the target ecosystem Voss et al. (2014) illustrated that
the definition of the this optimum is dependent on the managers view of the
resource. Although the rational for ecosystem based management is strong its
adoption has been slow, which Möllmann et al. (2013) suggests is due to a lack
of strategy. However, estimation of the effects of species interactions on the
management of a commercially exploited resource is a fairly complicated and
data intensive excercise, as it requires data not directly observed by fisheries.
Dickey-Collas et al. (2013) note that the state of knowledge is not sufficient
to embark on ecosystem engineering and although in theory this might have
desirable effects on the status of key resources, as such an undertaking may
have undesirable and unforseen consequences.

  There is considerable uncertainty on the role of minke whales in the Icelandic
continental shelf ecosystem. Previous studies have indicated that cetaceans, and
in particular minke whales play an important role in the marine ecosystem by
consuming several times the total fishery landings (Sigurjónsson and Víkings-
son 1997). Initial attempts to include three species of cetaceans, minke, fin
and humpback whales, in a multispecies model indicate that their effect on the
development of the stocks of cod and capelin may be considerable (Stefánsson
et al. 1997). There was, however considerable uncertainty associated with this
estimate. One of the greatest sources of uncertainty regarding the effects on the
cod stock was associated with the very limited knowledge of the diet composi-
tion found in minke whales in Icelandic waters, cod age groups in particular. It
was therefore considered of prime importance for further development of multi-

species modelling in Icelandic waters to obtain data on the diet of minke whales and investigate multi-species interactions in in more detail, in particular those between minke whales and the cod stock.

Here a study of the statistical properties of models of species interactions are investigated. The effects of linking cod (*Gadus morrhua*) and the common minke whale (*Balaenoptera acutorostrata*) in Icelandic waters in a single model are used as case study in light of recently available data presented in Víkingsson et al. (2014). It illustrates the effects on various data sources and different weightings, as described in Paper III.

V.2   Materials and methods

The Icelandic continental shelf area is situated on a dynamic front between colder Artic current and the warmer Gulf stream current (Astthorsson et al. 2007). It features a complicated ecosystem and, as in many other areas in the North Atlantic, gadoid species, especially cod, are the main focus of the fishery. Species interaction in Icelandic waters has received considerable attention (e.g. Gislason and Ástþórsson 1997, Stefánsson and Pálsson 1997a).

V.2.1   Minke whales in Icelandic waters

The common minke whale (*Balaenoptera acutorostrata*) is the most abundant whale species in Icelandic continental shelf waters (Borchers et al. 2009, Pike et al. 2011a). Minke whales are found all around the Icelandic coast, principally in shallower waters (<500m). Direct measurements of the abundance of minke whales in Icelandic began in 1986 with an aerial survey covering the continental shelf around Iceland (Borchers et al. 2009). Since then four complete surveys have been conducted in the area, 1987, 2001, 2007 and 2009 (Pike et al. 2011a).

Significant fluctuations have been observed in the abundance of minke whales within the Central North Atlantic stock area, although confidence intervals are large (Pike et al. 2011a), potentially due to distributional shift (NAMMCO 2009). After a peak around 2001 there has been an appreciable decline in abundance in Icelandic coastal waters. Information on diet composition indicated a strong preference for sandeel (*Ammodytes sp.*) and a recent collapse in sandeel abundance coincided with a decrease in the number of minke whales in Icelandic waters (Víkingsson et al. 2014). Unfortunately no reliable estimate of sandeel abundance prior to the stock collapse exists. However auxiliary information can

be obtained from stomach content data from haddock (*Melanogrammus aeglefinus*) sampled during the annual survey in Iceland (Bogason pers. comm).

The history of minke whale exploitation dates back to early 1914. Although systematic registration of the catch was not initiated until 1974, the catches have been estimated to be less that 50 per year for the first 50 years (Sigurjonsson 1988). Prior IWC's whaling moratorium (around 1975 1985) the annual catch was around 200 animals. The export of minke whale products, initiated in 1976 may have facilitated fleet selection to larger whales

V.2.2   Cod in Icelandic waters

Data on the status the commercially exploited marine resources in Icelandic waters have been systematically collected since 1986 (eg. Pálsson et al. 1997). As cod is the species with the highest commercial value survey sampling schemes have been specially designed to reflect the distribution of cod.

In this study cod is considered to be a homogenous unit on the Icelandic continental shelf as frequently done when managing the stock. However there are indications that the stock structure may be more complicated (eg. Pampoulie et al. 2006, Taylor 2011).

V.2.3   Stock dynamics

The general stock dynamics modules in Gadget are described in detail in Begley and Howell (2004). The following describes stock dynamics of the common minke whale and cod as implemented within the Gadget framework, where species specific formulations are highlighted where appropriate. The model for cod is described in detail in Taylor et al. (2007) and subsequently Elvarsson et al. (2014a). In the model the simulated quantity is the number of individuals in a cell, $N_{gralsyt}$ where g denotes the species, r denotes the area in which the individuals within the cell live, a is the age, l the cells lengthgroup, s is the maturity status, where $s = 0$ denotes immature animals while $s = 1$ mature, at year y, ranging from 1960 to 2011 and t is the quarter within the year. Age ranges between a_{0g} and $a_{\max g}$ years while the length ranges from l_{0g} to $l_{\max g}$ cm, with δl_g cm length increments. The population is governed by the following

equations:

$$N_{gralsy,t+1} = \sum_{l'} G_{gl}^{l'} \big[(N_{gral'syt} - C_{fgral'st})e^{-M_{ga}\Delta t} +$$

$$I_{gral'syt} \big] \qquad \text{if } t < 4$$

$$N_{gra+1,ls,y+1,1} = \sum_{l'} G_{gl}^{l'} \big[(N_{gral'sy,4} - C_{fgral's,4})e^{-M_{ga}\Delta t} +$$

$$I_{gral'sy,4} \big] \qquad \text{if } t = 4 \ \& \ a < 12$$

$$N_{ra,ls,y+1,1} = \sum_{l'} G_{gl}^{l'} (N_{gral'sy,4} - C_{fgral'sy,4} +$$

$$N_{gra-1,l'sy,4} - C_{fgr,a-1,l'sy,4})e^{-M_{ga}\Delta t} \quad \text{if } t = 4 \ \& \ a = 12$$

$$(V.1)$$

where $G_l^{l'}$ is the proportion in lengthgroup l, of species g, that grows $l' - l$ lengthgroups in a timeperiod of length Δt, C_{falsyt} denotes the catches by fleet $f \in \{S_c, C_c, S_m, C_m\}$, $S.$ and $C.$ denote the survey and commercial fleets respectively, M_{ga} the natural mortality at age a of species g and $I_{gralsyt}$ denotes the movement from the immmature to the mature stock components. [1]

Growth in length is modeled as a two–stage process, an average length update in Δt and a growth dispersion around the mean update (as described in Stefansson 2005b). The average length update per time step is set according to a simplyfied form of the Von Bertanlanffy equation:

$$\Delta l_g = (l_{g,\infty} - l)(1 - e^{-k_g \Delta t}) \qquad (V.2)$$

where $l_{g,\infty}$ is the terminal length and k_g is the annual growth rate of species g. In the second step the growth is dispersed according to a beta–binomial distribution parametrised by the following equation:

$$G_{gl}^{l'} = \frac{\Gamma(n_g+1)}{\Gamma((l'-l)+1)} \frac{\Gamma((l'-l)+\alpha_g)\Gamma(n_g-(l'-l)+\beta_g)}{\Gamma(n_g-(l'-l)+1)\Gamma(n_g+\alpha_g+\beta_g)} \frac{\Gamma(\alpha_g+\beta_g)}{\Gamma(\alpha_g)\Gamma(\beta_g)} \quad (V.3)$$

where α_g is subject to

$$\alpha_g = \frac{\beta_g \Delta l_g}{n_g - \Delta l_g} \qquad (V.4)$$

where n_g denotes the maximum length group growth of the species within a timestep and $(l' - l)$ the number of lengthgroups grown.

---

[1] A short note on notation, here l is used interchangeably as either the lengthgroup or the midpoint of the length interval for that particular lengthgroup, depending on the context.

The minke whales stock components, i.e. immmature and mature, are assumed to have started at their mutual carrying capacity in 1960 when the model simulation starts. Birthrate of minke whales follows a variant of the Pella–Tomlinson spawning model (as in Punt 1999, Stefánsson et al. 1997):

$$b_t = BN_{ft}\left[1 + A\left(1 - \left(\frac{N_{ft}}{K_f}\right)^z\right)\right] \tag{V.5}$$

where B_j is the birthrate per mature female in the pristine population, N_{ft} number of mature females (assumed to be 50% of the mature population), A and z are determined by MSY level and rate, and K_f the carrying capacity of mature females.

Two types of recruitment approaches are used for cod, depending on the time period. For the data rich time period, that is 1984 and onwards, the total number of recruits is estimated per year. Recruitment enters to the population according to:

$$N_{1l0yt'} = R_y p_l \tag{V.6}$$

where t' denotes the recruitment time-step and R_y is the yearly recruitment. p_l is the proportion in lengthgroup l that is recruited which is determined by a normal density with mean according to the growth model and variance σ_y^2. For the period before 1984 the number of recruits per year is considered to be constant for all years, i.e. $R_y = R \ \forall y < 1984$.

Maturity is modeled and represented by the two stock components of each species. The movement between the two components is formulated as

$$I_{alsyt} = \begin{cases} N_{al0yt} \times m_x & \text{if s} = 1 \\ -N_{al0yt} \times m_x & \text{if s} = 0 \end{cases} \tag{V.7}$$

where $s = 0$, as noted above, denotes the immmature stock component and m_x is the proportion mature defined as:

$$m_x = \frac{1}{1 + e^{-\lambda_g(x-x_{50})}} \tag{V.8}$$

where x is either length or age for cod and minke respectively and x_{50} representes the mid-point of the maturity ogive.

Under the assumption that minke whales have a preference for pelagic fish stocks such as sandeel and capelin the fluctation in minke whale abundance is in the model explained by linking the abundance to indices of sandeel biomass by a migration matrix:

$$\begin{pmatrix} 1 & \delta_s m + m_c \\ 0 & 1 - (\delta_s m + m_c) \end{pmatrix} \tag{V.9}$$

where δ_s is the sandeel abundance index and m and m_c are the relative proportion of minke whales that migrate to the Icelandic continental shelf during the summer. Autumn migration is estimated similarly but in the opposite direction.

Catches are simulated based on reported total landings and a length based suitability function for each fleet, both commercial fleets and surveys. Total landings are assumed to be known and the total biomass (or in the case of minke whales, abundance) is simply offset by the landed catch. The catches for lengthgroup l, fleet f at year y and timestep t are calculated by:

$$C_{flsyt} = E_{ft} \frac{S_f(l)N_{lsyt}W_{ls}}{\sum_{s'}\sum_{l'} S_f(l')N_{l's'yt}W_{l's'}} \tag{V.10}$$

where E_{ft} is the landed biomass (or numbers with $W = 1$) at time t and $S_f(l)$ is the suitability of lengthgroup l by fleet f defined as:

$$S(l) = \begin{cases} \frac{1}{1-e^{-(\alpha+\beta l)}} & \text{For whaling and fishing operations} \\ p_0 + p_1 e^{-\frac{(\log(l/L)-p_2)^2}{p3}} & \text{For whaling operations in the early years} \end{cases} \tag{V.11}$$

where the whaling fleet went through a gradual shift from a dome-shaped selectivity to a logistic curve between the years 1970 to 1980.

Consumption by minke whales of cod is modeled in a similar fashion as fleets, i.e. through a length based suitability function. However the amount is determined based on the energy requirement of the predator, which is in this case minke whales. In the simple predator-prey relation between the species the prey availability to a predator becomes

$$\frac{F_{Ll}}{\sum_{l'} F_{Ll'} + \mathfrak{s}OA} \tag{V.12}$$

where $F_{Ll} = S_{Pp}(L,l)N_{pl}W_{pl}$ is the prey biomass of length l that the predator could consume,

$$S_{Pp}(L,l) = \frac{1}{1-e^{-(\mathfrak{a}+\mathfrak{b}l)}}$$

the suitability of a prey p for predator P, W_{pl} is the average weight for prey of length l, O is the density, per unit area, of other unspecified food sources and A is the size of the area. Here OA is defined such that enough food is available to minke whales other than cod.

The feeding level of the predator is

$$\Psi_L = M_L \frac{\sum_{l'} F_{Ll'} + \mathfrak{s}OA}{\sum_{l'} F_{Ll'} + \mathfrak{s}OA + HA} \tag{V.13}$$

where M_L is the maximum consumption of a predator of length L for a particular time step length Δt. It is determined by

$$M_L = m_0 L^{m_4} \Delta t \tag{V.14}$$

where m_0 and m_4 are user defined constants[2]. H is the density (biomass per area unit) of available food at which the predator can consume half maximum consumption. Note that H should reflect the ability of the predator to pursue its prey, which should factor in variables such as temperature and the cost of search. H determines how quickly the predators consumption reaches M_L. $H = 0$ would indicate that the predator would easily consume M_L of the available biomass. A larger value for H would indicate that prey is harder to find and therefore prey needs to be more abundant for the predator to reach M_L.

Given the feeding level and the prey availability the desired consumption of minke whales predators of length L is

$$\begin{aligned} C_{L,l} &= N_L \Psi_L \Big[\frac{F_{Ll}}{\sum_{l'} F_{Ll'} + \mathfrak{s}OA} \Big] \\ &= N_L M_L \frac{F_{Ll}}{\sum_{l'} F_{Ll'} + \mathfrak{s}OA + HA} \end{aligned} \tag{V.15}$$

The parameter settings for the consumption are based on Lindstrøm et al. (2009) and references therein.

Following Stefánsson et al. (1997) natural mortality of cod is here factored into two parts, one induced by the consumption of minke whales and the second due to other sources.

$$M_a = M_{0a} + M_{ma} \tag{V.16}$$

where M_{0a} is the baseline mortality due to other sources and M_{ma} is the mortality induced by minke whales at age a. Table V.1 gives the values of M_a for all a. In the case where consumption is not modeled $M_a = M_{0a}$. When consumption is modelled M_{ma} is assumed *a priori* to be 0.08 and the modeled total natural mortality is adjusted accordingly. This results in a M_{0a} of 0.12 for most age classes which is inline with a recent estimate of the natural moratility reported by Gudmundsson (2013).

---

[2]The numbering scheme is set to correspond to parameters in Begley and Howell (2004)

Table V.1: An overview of the estimated parameters in the model. For those parameter with fixed values a description of how these values were derived can be found in Taylor et al. (2007) and references therein.

Description	Notation	Comments	Formula
Natural mortality	M_{ag}	For cod: fixed at 0.2 for ages 3 to 9 while 0.5, 0.35 for the first two ages and 0.3, 0.5 and 0.7 for ages $10 - 12$. For minke whales 0.087	See eq. V.1
Growth function	$k_g, L_{\infty,g}$	Estimated from age–length frequencies, L_∞ is fixed at 140cm for cod and 1100 cm for minke whales	See eq. V.2
Growth implementation	β_g	n_g is fixed at 15 length-groups for cod and 30 for minke whales	See eq. V.3
Fleet selection	a_f, b_f	One set for each of the fleets (Spring and autumn survey, commercial cod fleet and whaling fleets	See eq. V.11
Maturity ogive	λ, l_{50} or a_{50}		See eq. V.8 where for whales $\lambda = 1/1.2$ and $a_{50} = 7$.
Number of recruits by year	R_y	$y \in [1982, 2012]$. σ_y^2, i.e. variance in recruitment length, based on length distributions obtained in the autumn survey.	See eq. V.6

Initial abundance at ages 3 – 20 in 1982	η_a	$a \in [3, 25]$. σ_a^2, i.e. variance in initial length at age a, based on length distributions obtained in the spring survey.	See eq. II.6
Survey catchability	q_f, β_f	Intercept term and slope in a log–linear relationship with abundance. The slope term is assumed to be 1 for odd numbered variants for all indices.	See eq. V.18
Length–weight relationship	μ_s, ω_s	Different values by stock component, estimated outside of the model	See eq. II.11
Carrying capacity	K_f	Estimated, stock assumed to start at K	See eq. V.5
Density dependence	A, z	Resilance parameters of the minke fecunity function, determined from MSYR and MSYL	See eq. V.5
Consumption	m_0, m_4	Parameters defining the energy requirement of a minke whale of length L. Using values from Lindstrøm et al. (2009)	See eq. V.14
Predation	$\mathfrak{a}, \mathfrak{b}, \mathfrak{s}$	Parameters defining prey suitability for minke whales	See eq. V.12

V.2.4 Observation model

In Gadget data are assimilated using a weighted log–likelihood function. Here five types of data enter the likelihood, absolute abundance and relative biomass indices, length distributions from survey and commercial fleets combined into a single likelihood, age – length distribution from from the survey, maturity at

Chapter V   Adding data on species interaction: A case study on cod (*Gadus morhua*) and common minke whales (*Balaenoptera acutorostrata*) interactions in Icelandic waters

158

length for all ages, and stomach contents from minke whales. The datasets used here is shown are table V.2.

Survey indices

Abundance estimates from Pike et al. (2011a) for minke whales enter the likelihood using the following equation:

$$l^{\text{SI}} = \sum_y \sum_t (\bar{N}y - \widehat{N_{yt}})^2 \tag{V.17}$$

where $\bar{N}_y$ is the observed abundance estimate and

$$\widehat{N_{yt}} = \sum_l \sum_a \sum_s N_{alsyt}$$

The survey indices for cod are defined as the total number of fish caught in a survey within a certain length interval. The intervals used here are $16 - 25$ cm, $26 - 38$ cm and larger than 38 cm. These intervals are chosen such that they roughly represent age 1, age 2 and age 3^+.

For each length range g the survey index is compared to the modelled abundance at year y and time-step t using:

$$l^{\text{SI}}_{gf} = \sum_y \sum_t (\log I_{gfy} - (\log q_f + \beta \log \widehat{N_{gyt}}))^2 \tag{V.18}$$

where

$$\widehat{N_{gyt}} = \sum_{l \in g} \sum_a \sum_s N_{alsyt}$$

For implementation details see Taylor et al. (2007) and subsequently the annex to Paper II.

Maturity at length

Length at maturity comparison uses either the number of mature males observed in the scientific survey of minke whales or the number of immature and mature at age 3 from the Icelandic groundfish survey. The observed proportions are compared to the modelled proportion using sum of squares:

$$l^{\text{M}} = \sum_y \sum_t \sum_l (\pi_{lyt} - \hat{\pi}_{lyt})^2 \tag{V.19}$$

where π_{lyt} and $\hat{\pi}_{lyt}$ are the observed and modelled proportions mature respectively in length group l, year y and timestep t.

Table V.2: Overview of the likelihood data used in the model. Survey indices are calculated from the length distributions and are dis-aggregated ('sliced') into three groups. Number of data-points refer to aggregated data used as inputs in the Gadget model and represent the original data-set. All data can obtained from the Marine Research Institute, Iceland.

Origin	Time-span	Length group size	Num. data-points	Likelihood function
Length distributions:				
March Survey	1st quarter, 1985 – 2003	2 cm	1845	See eq. V.20
October Survey	4th quarter, 1995 – 2003	2 cm	897	See eq. V.20
Commercial catches	All quarters, 1984 – 2003	2 cm	5916	See eq. V.20
Minke catches	All quarters, 1974–1985, 2003–2011	50 cm	335	See eq. V.20
Age – length frequencies				
March Survey	1st quarter, 1989 – 2003	4 cm	3856	See eq. V.21
October Survey	4th quarter, 1995 – 2003	4 cm	2019	See eq. V.21
Commercial catches	All quarters, 1984 – 2003	4 cm	9158	See eq. V.21
Commercial and survey catches (minke)	All quarters, 1977,1978,2003–2007	50 cm	119	See eq. V.21
Survey indices				
March Survey	1st quarter, 1985 – 2003	–	87	See eq. V.18
October Survey	4th quarter, 1995 – 2003	–	51	See eq. V.18
March Survey	2nd quarter, 1987,2001,2007 and 2009	–	4	See eq. V.18
Ratio of immature:mature at age 3 by length group				
March Survey	1st quarter, 1985 – 2003	2 cm	2045	See eq. V.19
Ratio of immature:mature minke whales in the scientific survey				
Scientific survey	2003 – 2007	50 cm	55	See eq. V.19
Stomach contents of minke whales in the scientific survey				
Cod	2003 – 2007	2 cm	121	See eq. V.22

Fleet data

Length distributions are compared using either 2 cm or 50 cm lengthgroups for cod and minke whales respectively for both commercial and survey fleets using

$$l_f^{\text{LD}} = \sum_y \sum_t \sum_l \sum_s (\pi_{lsyt} - \hat{\pi}_{lsyt})^2 \qquad \text{(V.20)}$$

where f denotes the fleet where data was sampled from. Similarly age – length data are compared using 4 cm (or 50 cm) length groups:

$$l_f^{\text{AL}} = \sum_y \sum_t \sum_a \sum_l \sum_s (\pi_{falsyt} - \hat{\pi}_{falsyt})^2 \qquad \text{(V.21)}$$

Stomach data

Stomach contents of minke whales are compared to modelled consumption in a similar manner as fleet data or by:

$$l_f^{\text{ST}} = \sum_y \sum_t \sum_l \sum_s (\pi_{lsyt} - \hat{\pi}_{lsyt})^2 \qquad \text{(V.22)}$$

where π_{lyt} and $\hat{\pi}_{lyt}$ are the observed and modelled proportions of stomachs with prey respectively in length group l, year y and timestep t.

V.2.5   Estimation procedure

The parameter estimation procedure was however split into two parts. First two single species models of cod and minke whales were fitted. The total objective function used the modelling process combines equations V.18 to V.21 using the following formula:

$$l^{\text{T}} = \sum_g \sum_{f \in \{S,A\}} w^{\text{SI}} l^{\text{SI}} + \sum_{f \in \{S,C\}} \left(w_f^{\text{LD}} l_f^{\text{LD}} + w_f^{\text{AL}} l_f^{\text{AL}} \right) + w^{\text{M}} l^{\text{M}} + l_f^{\text{ST}} \quad \text{(V.23)}$$

where $f = S$ or C denotes the survey and commercial fleets respectively and w's are the weights assigned to each likelihood components. These weights are estimated using the iterative reweighting described in Elvarsson et al. (2014b). For cod two likelihood component grouping variants in the iterative reweighting were considered and three model variants as illustrated in table V.3

In the second part the two single species model were connected by defining a predator prey relationship using a phased approach, alternating between

Variants	SI_i 4 - 17	SI_i 17 - 33	SI_i 33 - 140	SI_a 16 - 27	SI_a 27 - 39	SI_a 39 - 140
V1	G_1 F	G_1 F	G_1 F	-	-	-
V2	G_1 S	G_1 S	G_1 F	-	-	-
V3	G_1 F	G_1 F	G_1 F	G_2 F	G_2 F	G_2 F
V4	G_1 S	G_1 S	G_1 F	G_2 S	G_2 S	G_2 F
V5	G_1 F	G_2 F	G_3 F	G_1 F	G_2 F	G_3 F
V6	G_1 S	G_2 S	G_3 F	G_1 S	G_2 S	G_3 F

Table V.3: Model variants tested in the case of cod in Icelandic waters. $G_.$ denotes the weigthing group, F denotes a fixed $\beta = 1$ in equation V.18 and S the case where is β estimated.

Phase	Cod	Minke	Predation
1	0	0	1
2	1	0	0
3	0	0	1
4	1	0	0
5	0	1	0
6	1	0	0
7	0	1	0
8	1	0	0

Table V.4: Phasing used in the minimization procedure when estimating the link between minke whales and cod. The columns indicate the parameter set and the line the phase. 1 denotes that the parameter set is estimated and 0 not estimated.

predation parameters, cod stock parameters and minke stock parameters. This process was repated eight times and is described in table V.4.

As with other Gadget models a fair amount of results can be derived from the model. In particular the fit to each data set during the iterative refweighting can provide information on compromises made in the final optimisation. This is illustrated here by tabulating the sum of squres for each likelihood component (as described in Taylor et al. 2007). Comparions between model variants can be made either based on the scores or graphically based on fits to individual data sets.

V.2.6 Forward projections

To compare the model variants in terms of management the stock was projected forward 5 years in order to assess the differences in advice. Fishing mortality in the projections was set according to be the average mortality in the last 5 years for each variant. Recruitment in the projections were modeled as an $AR(1)$ process:

$$\bar{R}_y = \bar{R}_{y-1} + \epsilon_y \tag{V.24}$$

where ϵ_y is a mean zero gaussian with variance σ^2_{rec}. This relationship was based on estimated number of recruits before 2010. It is assumed that this approximation for the recruitment process is satisfactory as the cod spawning stock in Icelandic waters has increased steadily since the late 1990's and therefore the recruitment is not likely to be impaired during the forecast period. This increase is assumed to be the result of an implementation of improved management procedures for the Icelandic cod stock (as suggested by Anon 2009, and similar documents).

V.3 Results

V.3.1 Likelihood weights

The results from the iterative reweighting for the single species minke whaled model is shown in table V.5. It is fairly clear that the age–length and length distribution have a strong contrast with the whale abundance, as the residual sum of squares rises dramatically when the these distributions data are emphasised. These datasets are fairly disjoint in time. Age–length data from surveys are only available for whales caught in 2004 – 2007 and some older samples in 1977.

Length distributions however cover a longer timeperiod than the abundance
series and information on maturity.

	M	AL_{old}	LD	AL_{new}	$\hat{N}$
$\hat{N}$	132.40	0.47	8.74	0.92	1047000
M	74.76	0.44	9.73	1.03	13760000
AL_{old}	109.10	0.27	11.51	1.41	15800000
LD	180.50	0.70	5.22	1.11	23140000
AL_{new}	158.60	0.85	13.90	0.57	11760000
Final	90.29	0.38	6.19	0.70	1467000

Table V.5: Diagnostics for the single species model for minke whales. Table represents the individual likelihood component scores by dataset (columns) and component emphasized in the iterative reweighting (rows). Here $\hat{N}$ is the abundance data, M data on maturity, AL age length distribution and LD length distributions.

Similarly the resulting likelihood component scores from the single species cod model is shown in table V.6. The table presents the likelihood component scores from each compenent, and where the likelihood function differs between variants, the scores are shown for each variant. When estimating the slope for the fit to the survey indices (even number variants) the fit improves to the indices of survey group one and two. The fit appears to improve slightly more for the spring survey index, from ≈ 10 to ≈ 3 for group one and from ≈ 3.5 to ≈ 2 for group two, while the autumn survey improves slightly less, from ≈ 7 to ≈ 4.5 for group one and from ≈ 5.5 to ≈ 3.5 for group two. Grouping survey indices by survey fleet appears to put more emphasis on survey group three at the expense of groups one and two. In addition to the effects on the survey indices, grouping by survey fleet appears to have an adverse effect on the fit to proportion of matures, with a score between 75 to 90 for this grouping while a score around 46 where the grouping is based on length groups.

Table V.7 shows the percentage change in score when the models for cod and minke whales are combined. Minor changes are observed in the scores of individual components when predation on cod of minke whales added. For variants 4 and 5 a worse fit to maturity at age three from the spring survey, with an improvement of the fit other data series. Similarly variant 6 had scores improving for length and age–length distributions at the expense of survey indices.

Chapter V   Adding data on species interaction: A case study on cod (*Gadus morhua*) and common minke whales (*Balaenoptera acutorostrata*) interactions in Icelandic waters

164

Variant	Group	M_i	AL_a	AL_c	AL_i	LD_a	LD_c	LD_i	SI_i 4 - 17	SI_i 17 - 33	SI_i 33 - 140	SI_a 16 - 27	SI_a 27 - 39	SI_a 39 - 140
-	AL_a	266.90	0.06	3.38	0.27	0.04	9.99	0.17	11.82	5.19	10.84	4.15	10.20	2.75
-	AL_c	823.30	0.11	0.53	0.22	0.05	0.59	0.15	24.66	22.76	5.98	12.70	3.78	4.34
-	AL_i	701.60	0.07	3.44	0.11	0.04	10.46	0.21	8.55	5.89	8.29	2.95	6.72	2.75
-	LD_a	670.60	0.67	7.73	1.08	0.03	2.54	0.36	70.45	8.43	34.93	4.99	3.17	2.31
-	LD_c	427.60	0.44	4.45	0.73	0.05	0.25	0.28	99.64	19.77	28.07	10.11	4.19	1.28
-	LD_i	503.00	0.13	1.87	0.29	0.04	1.17	0.06	14.25	4.55	2.52	3.05	1.58	1.09
-	M_i	42.54	0.20	1.60	0.36	0.16	1.93	0.33	15.22	11.69	8.21	6.05	1.24	1.42
V1	G_1	307.80	0.36	7.30	0.50	0.08	9.21	0.21	3.38		2.05	0.24		0.24
V2	G_1	307.80	0.36	7.30	0.50	0.08	9.21	0.21	3.38		2.05	0.24	0.24	
V3	G_1	52.72	0.57	3.86	0.97	0.21	2.53	0.49	1.49	0.60	23.23	9.33	1.94	1.47
V3	G_2	307.30	1.02	8.94	2.08	0.23	7.35	0.43	79.32	46.24	0.82	0.80	4.80	1.34
V3	G_3	498.20	1.11	19.35	1.49	0.66	22.87	0.68	41.68	16.03	7.29	9.99	0.46	0.50
V4	G_1	51.80	0.14	1.59	0.26	0.10	1.14	0.18	0.00	4.17	3.90	3.43	1.42	1.20
V4	G_2	193.30	0.22	1.91	0.40	0.09	0.82	0.28	8.73	10.80	0.00	4.76	1.93	2.30
V4	G_3	542.30	0.69	10.22	0.95	0.08	11.13	0.18	25.77	15.65	6.34	7.75	0.43	0.97
V5	G_1	307.50	0.37	7.37	0.51	0.08	9.24	0.21	3.37	8.22	2.02	5.66	0.24	2.31
V5	G_2	133.70	0.18	8.35	0.55	0.13	10.08	0.49	64.98	1.41	71.72	0.92	14.31	0.58
V6	G_1	240.10	0.25	7.37	0.36	0.06	10.23	0.12	4.08	5.55	1.00	4.55	0.34	2.21
V6	G_2	209.60	0.36	1.82	0.62	0.14	0.80	0.37	25.49	15.65	7.98	7.06	3.31	0.00
V1	final	75.56	0.08	0.72	0.19	0.04	0.37	0.12	11.59		3.29	0.97		
V2	final	75.56	0.08	0.72	0.19	0.04	0.37	0.12	11.59		3.29	0.97	0.97	
V3	final	46.74	0.08	0.75	0.19	0.05	0.41	0.12	10.08	7.17	3.77	4.92	1.47	1.38
V4	final	46.34	0.08	0.76	0.18	0.05	0.41	0.12	3.14	4.41	2.14	3.41	1.31	1.44
V5	final	57.93	0.08	0.74	0.17	0.04	0.38	0.13	11.06	8.75	3.46	5.99	1.04	1.75
V6	final	90.06	0.08	0.74	0.18	0.04	0.34	0.12	4.53	5.69	1.88	4.38	1.20	1.42

Table V.6: Diagnostics for the single species model variants for Cod in Icelandic waters. Table represents the individual likelihood component scores by dataset (columns) and component emphasized in the iterative reweighting (rows). Here SI are the abundance indices, M data on maturity, AL age length distributions and LD length distributions.

Variant	M_i	AL_a	AL_c	AL_i	LD_a	LD_c	LD_i
V1	66.40	0.08	0.73	0.18	0.04	0.38	0.12
V3	46.62	0.08	0.75	0.18	0.05	0.44	0.12
V4	62.75	0.08	0.73	0.18	0.04	0.39	0.12
V5	68.15	0.08	0.73	0.18	0.04	0.38	0.12
V6	89.75	0.08	0.74	0.18	0.04	0.35	0.12

Variant	SI_i 4 - 17	SI_i 17 - 33	SI_i 33 - 140	SI_a 16 - 27	SI_a 27 - 39	SI_a 39 - 140
V1	11.79		3.38		0.86	
V3	9.43	6.30	4.66	4.50	1.89	1.62
V4	3.42	4.38	2.33	3.61	1.19	1.70
V5	10.86	7.96	3.55	5.83	0.89	2.03
V6	5.73	6.30	2.37	5.38	0.85	1.95

Variant	$\hat{N}$	AL_{old}	LD	AL_{new}	M
V1	1467000	0.38	6.20	0.70	90.28
V3	1467000	0.38	6.19	0.70	90.29
V4	1467000	0.38	6.19	0.70	90.29
V5	1467000	0.38	6.19	0.70	90.28
V6	1467000	0.38	6.19	0.70	90.31

Table V.7: Effects of adding minke whale consumption to the model for cod on the likelihood scores of individual components. Columns indicate the final component score for a particular dataset while the rows indicate the model variant.

V.3.2 Fit to data

The fit to the abundance indices for cod and minke whales is shown in figure V.1 for all variants with and without species interactions. In general little differences can be observed between the model variants. The model variants appear to consistently underestimate survey group one, i.e. proxy group for age 1, after 2000 and overestimate before. The model variants with predation of minke whales exhibit faster growth in survey group three that non-predation variants, coinciding with a substantial drop in minke whale abundance. Estimating catchability (odd number variants) appears to allow the model to predict better survey groups one and two. Figure V.2 illustrates individual likelihood component scores by year and model variant. All model variants have trouble with fitting the age data from the spring survey in the first few years, wich is apparent from a substantially higher component score in those years. Matu-

rity data appears to conflict with other data sources, as noted above, which is illustrated in figure V.2.

The fit to length distribution from commercial and survey operations is shown in Figure V.3. It appears that the model underestimates the number of small whales caught and overestimates the number of large whales caught in post moratorium period. The model replicates the shift in fleet selecion apparent from the data in pre-moratorium period that the whaling focussed on smaller whales than in later years. A comparison of the fitted age–length distribution to actual data is shown in Figure V.4. The largest discrepancies are in the oldest agegroup (25 and older) in the post–moratorium takes.

Modelled proportion consumed by minke as function of cod length group is shown in Figure V.5. Observed ratios in the stomachs of minke whales appear to be considerably consentrated to certain length groups (between 50 to 70 cm) that model does not predict to the same degree. In length groups greater that 80 cm the model overestimates the consumption of cod.

V.3.3   Model variables and biomass trajectories

Recruitment estimates for cod are shown in figure V.6. The difference between model variants, with and without minke whale predation, are seen to be minimal, while single species model variants estimate slightly higher number of recruits. Growth is similarly unaffected by the different model variants (figure V.7). Modelled growth of minke whales is fairly slow after the age of 5 as modelled average length at recruitment was estimated to ≈ 515 cm.

The suitability functions of of the fleets and predators implemented in the model are shown in figure V.8. The survey fleets have very similar selection of cod with $l_{50} \approx 50$ cm, while the commercial fleet focuses mainly on the larger cod with neglible selection below 50 cm. Minke whales in the model appear to target smaller cod, with l_{50} close to 25 cm. The whaling fleet selection is dome shape in the earlier years moving over to a knife-edged selection in the more recent years.

Model biomass, shown in figure V.9 appears to be minimally affected by different weighting schemes, appart from variant six without minke whale predation which simulates a higher biomass than the others. After the year 2000, when the abundance of minke whales in Icelandic waters starts to decrease, the biomass of cod appears to increase faster for variants with minke whale predation than those without. Annual fishing mortality by model variant, shown in figure V.10, has decreased during the years of the model. In all model variant a sharp decrease in fishing mortality is observed after 1993, the year that a
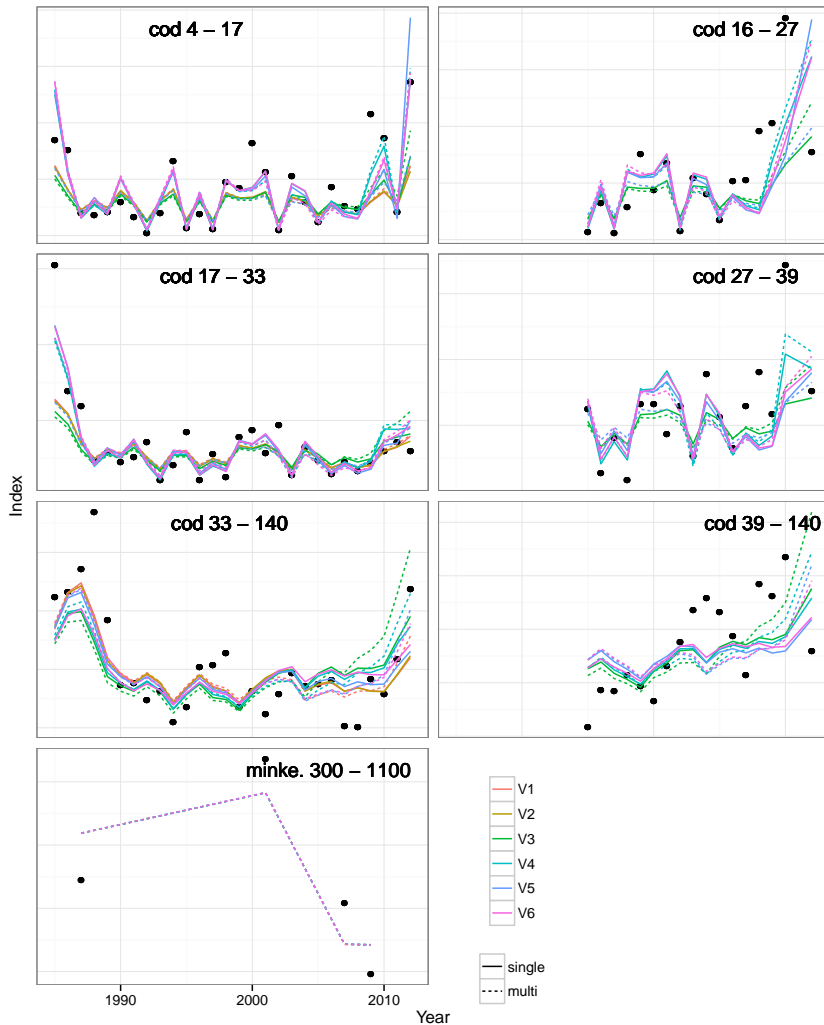
Figure V.1: Model fit to survey indices by model variant. Points represent observations while coloured lines (both solid and broken) indicate fit by model variant. The panels indicate the survey index fitted, panels on the left labelled cod are the indices from the spring survey while autumn survey on the right by lengthgroup. Bottom panel illustrates the fit to abundance for minke whales. Solid lines indicate model fit from single species variants while broken lines models with predation.

Chapter V   Adding data on species interaction: A case study on cod (*Gadus morhua*) and common minke whales (*Balaenoptera acutorostrata*) interactions in Icelandic waters

168
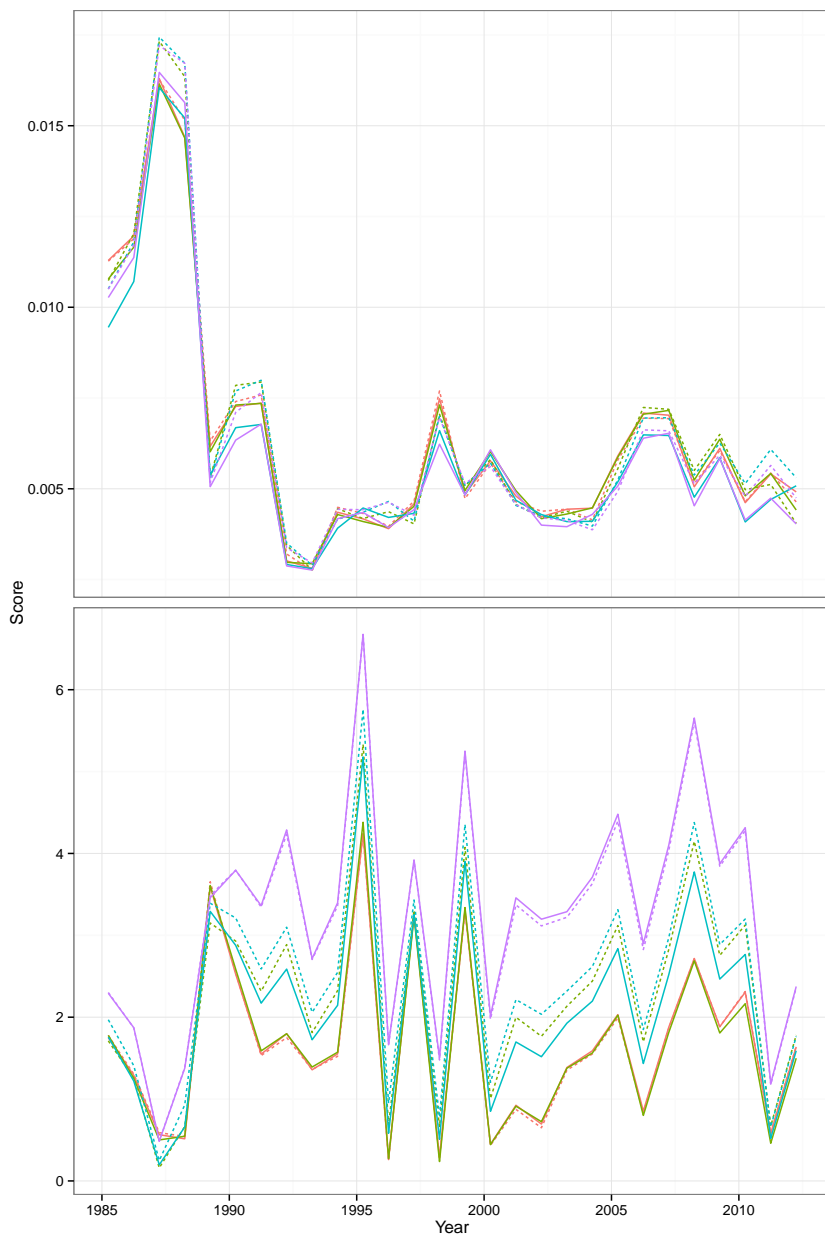
Figure V.2: Illustration of the likelihood component score and the contrast between the model variants. The top panel shows the residual sum of squares for the age–length data from the spring survey and the bottom figure the same for ratio of immatures at age 3 in the spring survey. Solid lines indicate model fit from single species variants while broken lines models with predation.
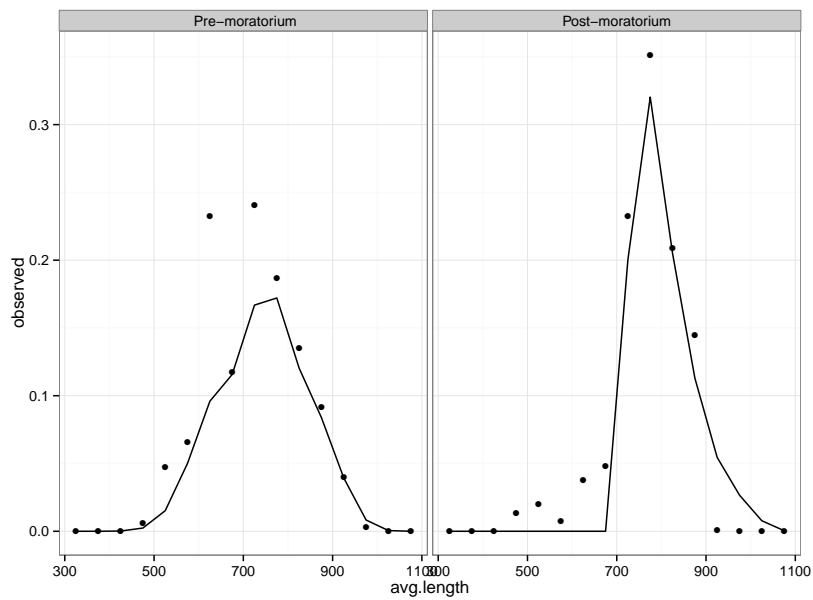
Figure V.3: Fit to length distributions of minke whales caught by commercial and survey whaling vessels before (1970 – 1984) and after (2003–present) the whaling moratorium was set in place.
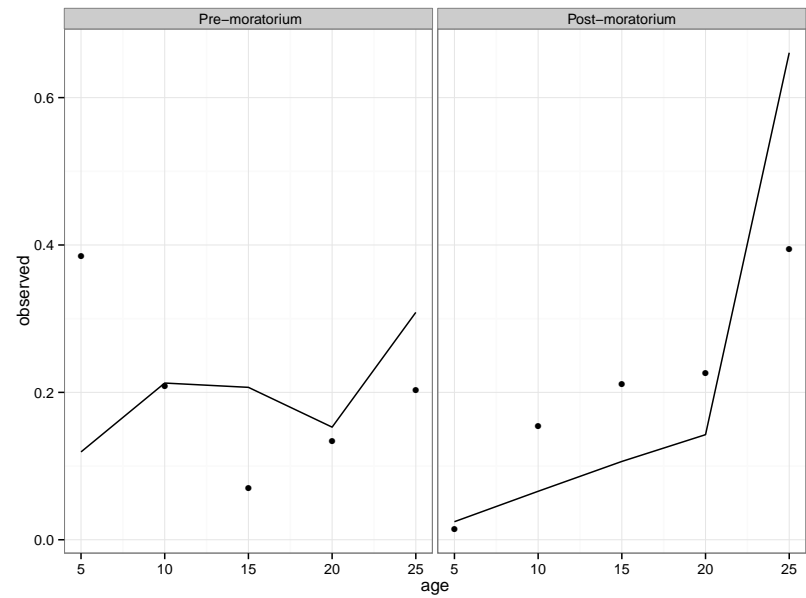
Figure V.4: Fit to age distributions of minke whales caught by commercial and survey whaling vessels before (1970 – 1984) and after (2003–present) the whaling moratorium was set in place.
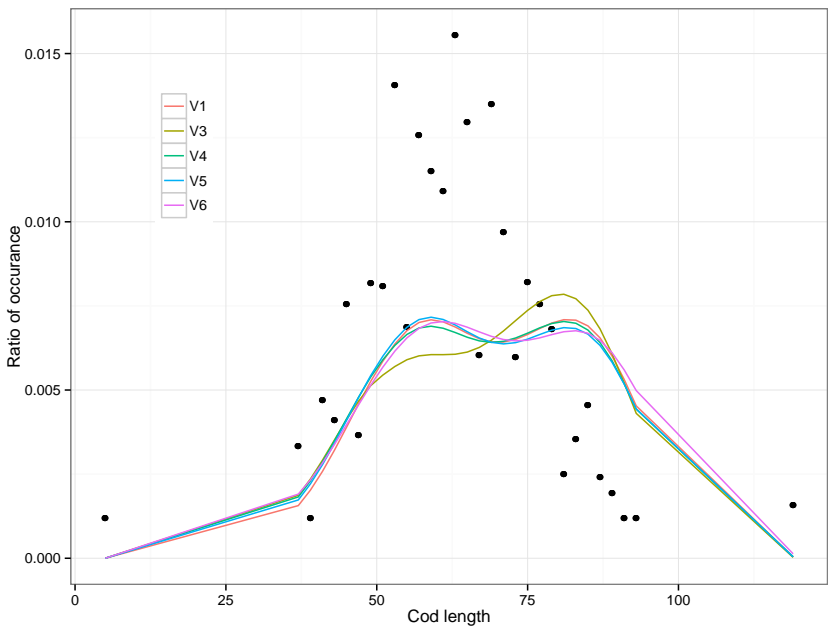
Figure V.5: The ratio of occurrance observed (points) of cod found in stomachs of minke whales in 2003–2007 as function of length group compared to model predictions (solid lines) by model variant.

Chapter V   Adding data on species interaction: A case study on cod (*Gadus morhua*) and common minke whales (*Balaenoptera acutorostrata*) interactions in Icelandic waters

172
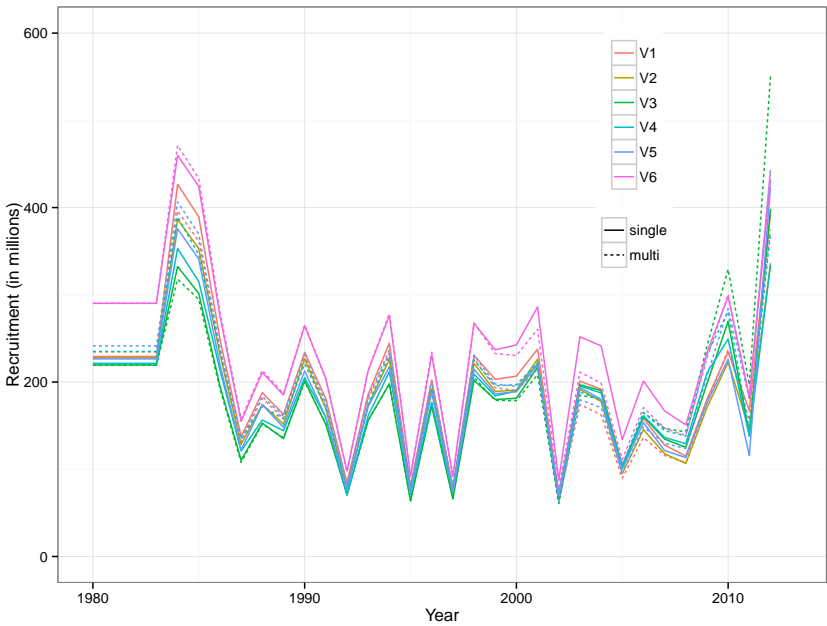
Figure V.6: Estimates of number of recruits by model variant for cod.
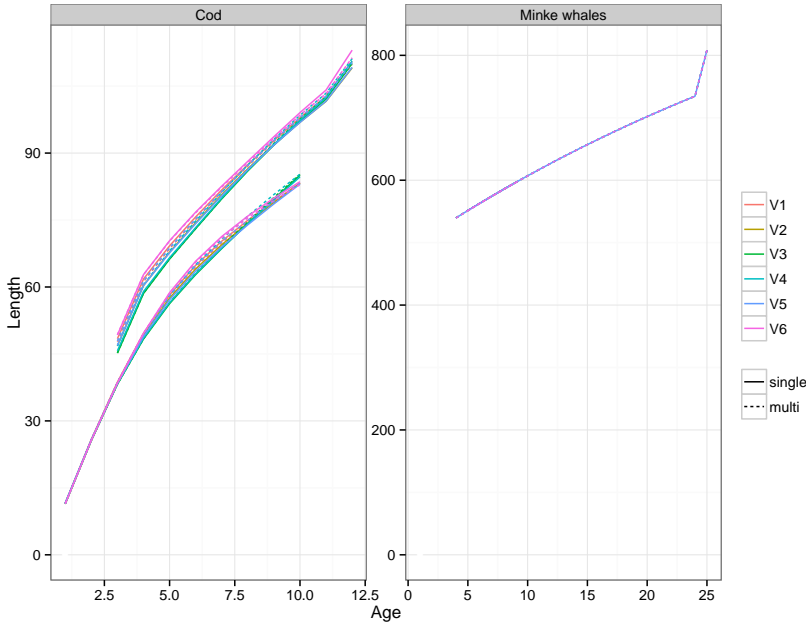
Figure V.7: Average modelled length at age for immature and mature cod, and minke whales in 2013 by model variant.
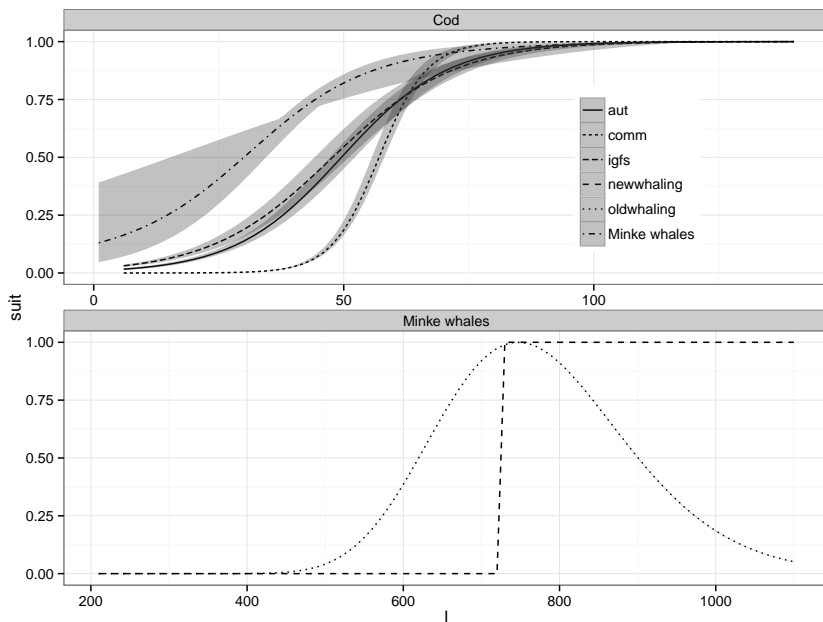
Figure V.8: Model selection curves, length–based, by prey for the survey and commercial fleets, and minke whales. Shaded region indicate the range estimated by the model variants. Legend refers to the autumn and spring surveys, `aut` and `igfs` respectifely, commercial fleets, `comm`, whaling operations (`oldwhaling` and `newwhaling` and predation of minke whales.
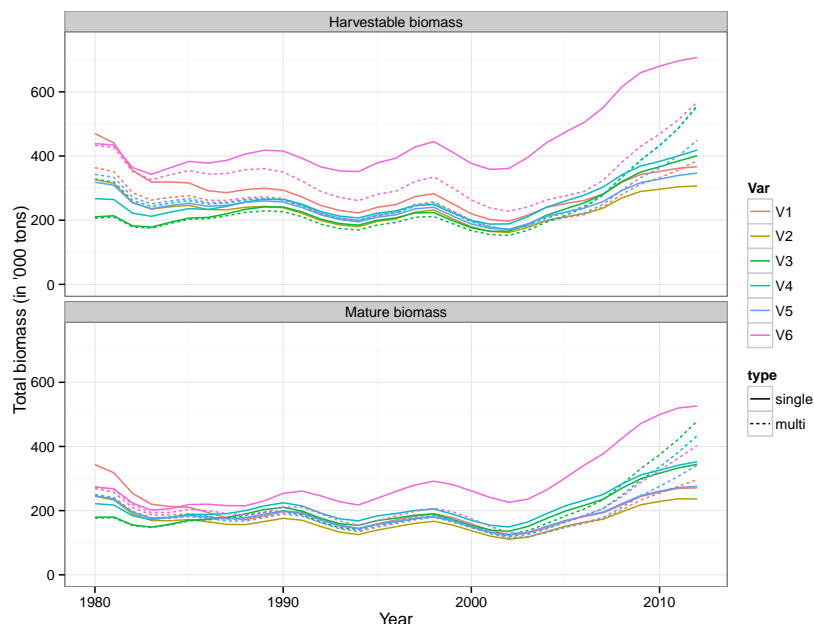
Figure V.9: Estimated historical harvestable and mature biomass of cod as a function of year by model variant. Solid lines indicate model variants without minke whale predation and broken lines those with predation.

harvest control rule was implemented for cod, and peaked again in 2000, after a period of overestimation of the cod stock.

Projected status of the mature biomass, which is shown in figure V.11, illustrates further these differences between models with and without predation of minke whales. Both models predict that the mature biomass will rise sharply until 2018, and then stabilise at different levels. Appart from variant six other variant that ignore minke whale predation stabilise at a lower biomass level around 500 thousand tons while those that include minke whale predation around 750 thousand tons.

Total consumption in terms of cod biomass is shown in figure V.12 where the annual consumption is seen to have fallen substantially from 2000 till 2006, from about 4–6 thousand tons to around 2 thousand tons. Consumption then increased slightly and ranged between 2.8 and 3.3 thousand tons. The estimated mortality induced by minke whale predation is estimated around 0.03 (ranging
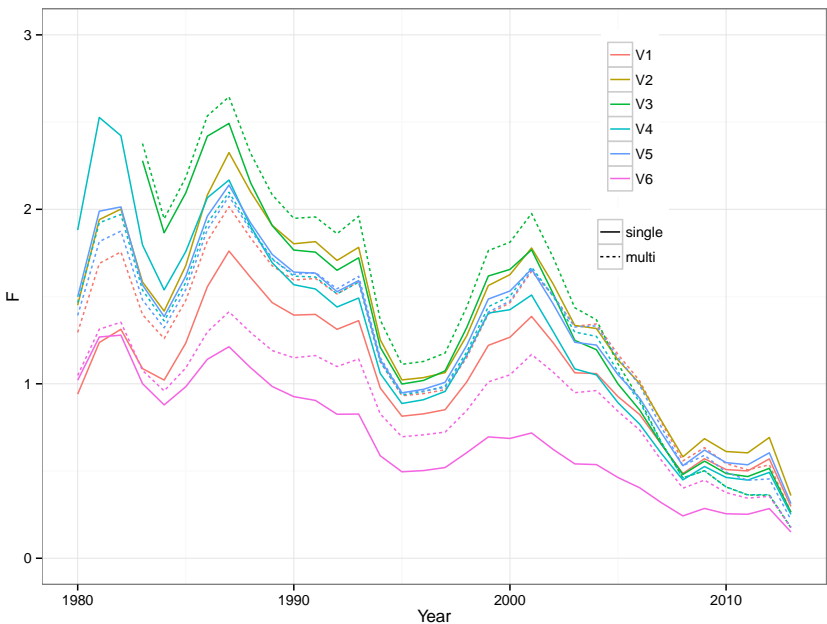
Figure V.10: Modelled fishing mortality of mature cod as a function of year by model variant. Solid lines indicate model variants without minke whale predation and broken lines those with predation.
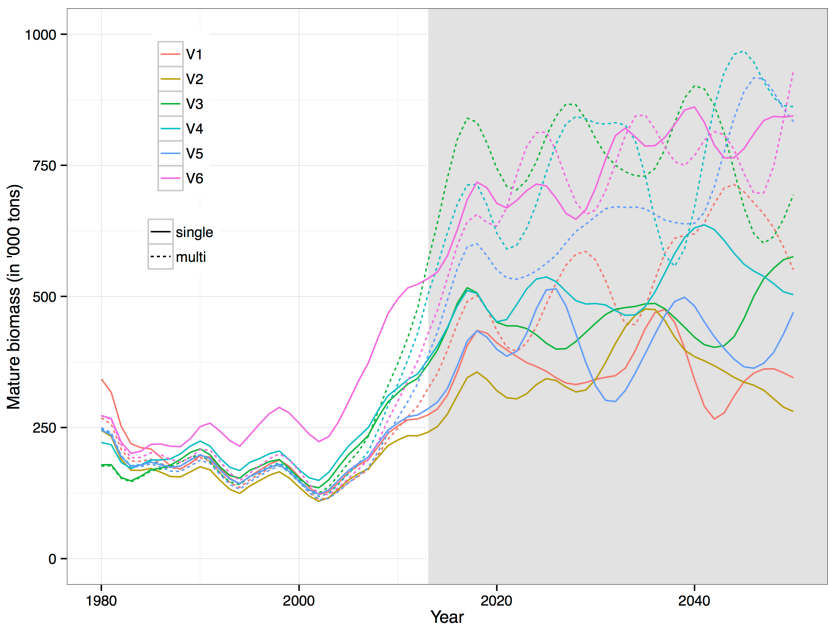
Figure V.11: Projections of mature biomass of cod as a function of year by model variant. Solid lines indicate model variants without minke whale predation and broken lines those with predation. Shaded region indicates the projection period.
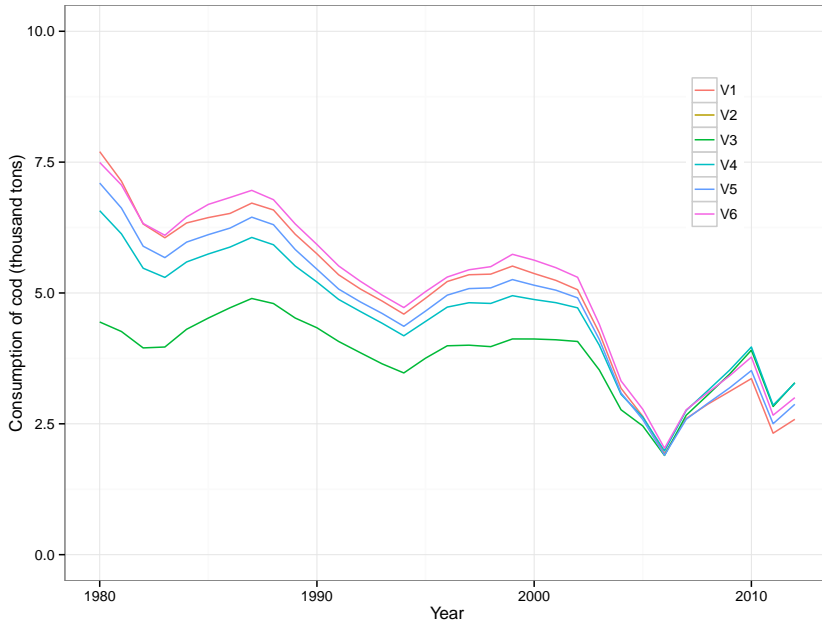
Figure V.12: Total consumption of cod, both mature and immature, by minke whales as a function of year and by model variant.

between 0.027 and 0.032) in 1990, increased slightly to around 0.037 (between 0.033 and 0.04) in 2000. Since 2000 the mortality has decreased substantially and is estimated less that 0.01 (between 0.008 and 0.009) in the terminal year.

V.3.4   Discussion

The analysis presented here provides the first attempt to link cod in Icelandic waters to a large mammalian predator, minke whales, in a statistical age-length based model. It provides a basis to answer some of the questions set out with in Stefánsson et al. (1997) and subsequently MRI (2003). Furthermore it provides insight into the restraints on the model by adding interaction between species, particularly species with such disparate data available as cod and minke whales, has on the overall fit to individual datasets.

   The abundance of minke whales has been observed to have decreased substantially in the last decade, both by whalers and whalewatcher. Here the major

driver in the abundance is considered to be linked to sandeel biomass, based on observations from Víkingsson et al. (2014). Additional drivers for minke whale abundance in Icelandic waters have not been excluded. For instance capelin has been identified as a important part of the diet of baleen whales (Sigurjónsson and Víkingsson 1997) and a shift in capelin summer distribution away from Icelandic waters has been observed in recent years (Pálsson et al. 2012). Pre-liminary work of including capelin in a multi–species model has however been unsuccessful due to the nature of the data (Taylor 2011) but work is on-going.

The effects of minke whales on cod is surprisingly small, judging by the fit to data. The difference between model variants with minke whales predating cod and those with no predation becomes apparent when projecting the stock status. When interpreting stock projections for cod, considerable care needs to be taken as the changes in ecosystem that cause the departure of minke whales, which are not included in the model could also negatively affect cod. The consumption of minke whales estimated here may seem minor in comparison with fleet catches. However most of the consumption is focused on juvenile cod within the models, while the commercial fishery targets larger fish. The mortality related to whale predation is also smaller by an order of magnitude. This may contradict results in Paper I which indicated substantially bigger consumption than estimated here. Here the form of the suitability function may play a role, as more dome–shaped function may be more appropriate. It cannot, however, be ruled out that the data on consumption are inconsistent with other data used with respect to the model.

As with all statistical models, access to representative data is necessary to estimate the processes of interest. Here for the first time information on the feeding ecology of minke whales in Icelandic waters is analysed. Data available from whaling operations before the moratorium, although limited, suggest con-siderable changes in the diet. Therefore the effects suggested by the model may not be as pronounced.

It is interesting to see the effects of different weighting schemes on the fit to data. Comparing this with the uncertainty estimated in Elvarsson et al. (2014a) the model variant seems to alternate between to close local minima, although in most cases within the interquartile range. This difference in model variants, particularly the weight grouping, might explain the difference between the pre-conditioned weights and estimated weights run in the uncertainty estimation, as described in II.

Comparing the results for biomass and recruitment to the one obtained in Taylor et al. (2007) (and subsequently Elvarsson et al. (2014a)) it is of some con-cern the overall changes between those two estimates. The level of recruitment

is estimated now to be considerably lower, although the general trends and variations appear to be the same. This may be due differences in fleet composition in recent year, where longliners have replaced gillnets to some degree.

A sensible next steps could include a wide range of analyses. Minor changes could include modifying minke whale selection to be changed to be more dome shaped, test different levels of depletion and productivity, and alter the assumptions of their energy requirements. Major changes could include continued developement of the two area model of cod described in Taylor (2011), inclusion of other species such as sandeel and capelin, which are seen as a large part of minke whale diet. Fleet composition will need some scrutiny as well.

V.4   Acknowledgements

Bibliography

Akaike, H. 1974. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723.

Andersen, L., Holm, L., Siegismund, H., Clausen, B., Kinze, C., and Loeschcke, V. 1997. A combined dna-microsatellite and isozyme analysis of the population structure of the harbour porpoise in danish waters and west greenland. *Heredity*, 78(3):270–276.

Anon 2002. Council regulation (ec) no. 2371/2002 of 20 december 2002 on the conservation and sustainable exploitation of fisheries resources under the common fisheries policy. Technical report, Council of the European Communities, Brussels.

Anon 2009. Report of the ad hoc group on icelandic cod hcr evaluation (agicod).

Arnott, S. A. and Ruxton, G. D. 2002. Sandeel recruitment in the north sea: demographic, climatic and trophic effects. *Marine Ecology Progress Series*, 238:199–210.

Astthorsson, O., Gislason, A., and Jonsson, S. 2007. Climate variability and the icelandic marine ecosystem. *Deep Sea Research Part II: Topical Studies in Oceanography*, 54(23-26):2456–2477.

Babak, O., Hrafnkelsson, B., and Palsson, O. 2007. Estimation of the Length Distribution of Marine Populations in the Gaussian-multinomial Setting using the Method of Moments. *Journal of Applied Statistics*, 34(7-8):985–996.

Baldursson, F. M., Daníelsson, A., and Stefánsson, G. 1996. On the rational utilization of the icelandic cod stock. *ICES Journal of Marine Science: Journal du Conseil*, 53(4):643–658.

Baranov, F. I. 1918. On the question of the biological basis of fisheries. *Nauchnyi Issledovatelskii Ikhtiologicheskii Institut, Izvestiya*, 1:81–128.

Begley, J. 2004. Gadget User Manual. Technical Report 120, Marine Research Institute.

Begley, J. 2005. Gadget User Guide. *Marine Research Institute, Reykjavík, Iceland) 90 Marine Research Institute Report 120*.

Begley, J. and Howell, D. 2004. An overview of Gadget, the globally applicable area-disaggregated general ecosystem toolbox. *ICES CM*, pages 1–15.

Benónýsdóttir, S. 2012. Statistical approach to relatedness analysis in large collections of genetic profiles an application to a dna-registry of fin whales. Master's thesis, University of Bergen, Norway.

Bertsekas, D. 1999. *Nonlinear programming*. Athena Scientific, 2nd edition.

Bertulli, C. 2010. *Minke whale (Balaenoptera acutorostrata) and white-beaked dolphin (Lagenorhynchus albirostris) feeding behaviour in Faxaflói bay, southwest Iceland*. PhD thesis, MSc thesis, University of Iceland.

Bérubé, M., Aguilar, A., Dendanto, D., Larsen, F., Notarbartolo di Sciara, G., Sears, R., Sigurjónsson, J., URBAN-R, J., and Palsbøll, P. 1998. Population genetic structure of north atlantic, mediterranean sea and sea of cortez fin whales, *Balaenoptera physalus* (linnaeus 1758): analysis of mitochondrial and nuclear loci. *Molecular Ecology*, 7(5):585–599.

Bérubé, M., Urbán, J., Dizon, A., Brownell, R., and Palsbøll, P. 2002. Genetic identification of a small and highly isolated population of fin whales (*Balaenoptera physalus*) in the sea of cortez, mexico. *Conservation Genetics*, 3(2):183–190.

Beverton, R. and Holt, S. 1957. On the dynamics of exploited fish populations. UK Min. Agric. *Fish., Fish. Invest.(Ser. 2)*, 19:533.

Björnsson, H. and Sigurdsson, T. 2003. Assessment of golden redfish (*Sebastes marinus* L.) in Icelandic waters. *Scientia Marina*, 67(Suppl 1):301–314.

Block, B. A., Teo, S. L., Walli, A., Boustany, A., Stokesbury, M. J., Farwell, C. J., Weng, K. C., Dewar, H., and Williams, T. D. 2005. Electronic tagging and population structure of atlantic bluefin tuna. *Nature*, 434(7037):1121–1127.

Bogstad, B., Hauge, K., and Ulltang, Ø. 1997. MULTSPEC–a multi-species model for fish and marine mammals in the Barents Sea. *Journal of Northwest Atlantic Fishery Science*, 22:317–341.

Bogstad, B., Tjelmeland, S., Tjelda, T., and Ulltang, O. 1992. Description of a multispecies model for the Barents Sea (multspec) and a study of its sensitivity to assumptions on food preferences and stock sizes of minke whales and harp seals. Technical report, SC/44/O 9.

Borchers, D., Pike, D., Gunnlaugsson, T., and Víkingsson, G. 2009. Minke whale abundance estimation from the nass 1987 and 2001 aerial cue-counting surveys taking appropriate account of distance estimation errors. *NAMMCO Sci. Publ.*, 7:95 – 110.

Butterworth, D. and Punt, A. 1999. Experiences in the evaluation and implementation of management procedures. *ICES Journal of Marine Science: Journal du Conseil*, 56(6):985–998.

Chen, Y., Breen, P., and Andrew, N. 2000. Impacts of outliers and misspecification of priors on bayesian fisheries-stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(11):2293–2305.

Chenuil, A., Crespin, L., Pouyaud, L., and Patrick, B. 2000. Movements of adult fish in a hybrid zone revealed by microsatellite genetic analysis and capture-recapture data. *Freshwater Biology*, 43(1):121–131.

Christensen, V. and Walters, C. 2004. Ecopath with Ecosim: methods, capabilities and limitations* 1. *Ecological Modelling*, 172(2-4):109–139.

Christiansen, F., Víkingsson, G. A., Rasmussen, M. H., and Lusseau, D. 2013. Minke whales maximise energy storage on their feeding grounds. *The Journal of experimental biology*, 216(3):427–436.

Claridge, A. W., Mills, D. J., Hunt, R., Jenkins, D. J., and Bean, J. 2009. Satellite tracking of wild dogs in south-eastern mainland australian forests: implications for management of a problematic top-order carnivore. *Forest Ecology and Management*, 258(5):814–822.

Combreau, O., Riou, S., Judas, J., Lawrence, M., and Launay, F. 2011. Migratory pathways and connectivity in asian houbara bustards: evidence from 15 years of satellite tracking. *PloS one*, 6(6):e20570.

Corana, A., Marchesi, M., Martini, C., and Ridella, S. 1987. Minimizing mul-
    timodal functions of continuous variables with the "simulated annealing" al-
    gorithm corrigenda for this article is available here. *ACM Transactions on
    Mathematical Software (TOMS)*, 13(3):262–280.

Davison, A. and Hinkley, D. 1997. *Bootstrap Methods and Their Application.*
    Cambridge University Press.

De Croos, M. and Stefansson, G. 2011. A case study of sampling strategies
    for estimating the length composition of commercial catches: The sri lankan
    shrimp trawl fishery. *Crustaceana, 84,* 12(13):1581–1591.

Demyanov, V., Wood, S. N., and Kedwards, T. J. 2006. Improving ecological
    impact assessment by statistical data synthesis using process-based models.
    Journal of the Royal Statistical Society Series C, 55(1):41–62.

Dickey-Collas, M., Engelhard, G. H., Rindorf, A., Raab, K., Smout, S., Aarts,
    G., van Deurs, M., Brunel, T., Hoff, A., Lauerburg, R. A., et al. 2013.
    Ecosystem-based management objectives for the north sea: riding the for-
    age fish rollercoaster. *ICES Journal of Marine Science: Journal du Conseil*,
    page fst075.

Donovan, G. 1991. A review of iwc stock boundaries. *Rept. Int. Whal. Commn.,
    Special*, 13:39–68.

Dragesund, O., Hamre, J., and Ulltang, Ø. 1980. Biology and population dy-
    namics of the norwegian spring-spawning herring. *Rapports et Procès-Verbaux
    des Réunions du Conseil International pour l'Exploration de la Mer*, 177:43–
    71.

Drouineau, H., Mahévas, S., Bertignac, M., and A., F. 2009. Assessing the
    impact of discretisation assumptions in a length-structured population growth
    model. *Fisheries Research*, 91(2-3):160–167.

dst^2, editor 2004. *Final Report: dst^2: Development of structurally detailed sta-
    tistically testable models of marine populations*, volume 118. Marine Research
    Institute, Reykjavik.

Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. *The Annals
    of Statistics*, 7(1):1–26.

Efron, B. and Tibshirani, R. 1993. *An introduction to the bootstrap*, volume 57.
    CRC press.

Efron, B. and Tibshirani, R. J. 1994. *An Introduction to the Bootstrap.* Chapman & Hall/CRC.

Elvarsson, B. Þ. and Þórðarson, G. 2014. An application of gadget to a data limited stock: An assessment of ling in icelandic waters. Presented at ICES WGDEEP.

Elvarsson, B., Taylor, L., Trenkel, V., Kupca, V., and Stefansson, G. 2014a. A bootstrap method for estimating bias and variance in statistical fisheries modelling frameworks using highly disparate datasets. *African Journal of Marine Science*, 36(1):99–110.

Elvarsson, B. T., Þórðarson, G., and Stefánsson, G. 2014b. Rgadget: a r–package for development, testing and analysis of gadget models. Unpublished manuscript.

Elvarsson, B. . 2015. Evaluating stock structure hypotheses using genetically determined close relatives: a simulation study on north atlantic fin whales. *ICES Journal of Marine Science: Journal du Conseil*, 72(2):661–669.

Fontaine, M., Baird, S., Piry, S., Ray, N., Tolley, K., Duke, S., Birkun, A., Ferreira, M., Jauniaux, T., Llavona, Á., et al. 2007. Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in old world waters. *BMC biology*, 5(1):30.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., and Sibert, J. 2012. Ad model builder: using automatic differentiation for statistical inference of highly parameterized complex non-linear models. *Optimization Methods and Software*, 27(2):233–249.

Francis, R. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(6):1124–1138.

Francis, R. and Shotton, R. 1997. risk in fisheries management: a review. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(8):1699–1715.

Frøysa, K., Bogstad, B., and W Skagen, D. 2002. Fleksibest–an age-length structured fish stock assessment model. *Fisheries Research*, 55(1-3):87–101.

Fulton, E. A., Smith, A. D., and Punt, A. E. 2005. Which ecological indicators can robustly detect effects of fishing? *ICES Journal of Marine Science: Journal du Conseil*, 62(3):540–551.

Fulton, E. A., Smith, A. D. M., and Smith, D. C. 2008. Alternative Management Strategies for Southeast Australian Commonwealth Fisheries: Stage 2: Quantitative Management Strategy Evaluation.

Gannon, J. E. 1976. The effects of differential digestion rates of zooplankton by alewife, alosa pseudoharengus, on determinations of selective feeding. *Transactions of the American Fisheries Society*, 105(1):89–95.

Gavaris, S. 1988a. An adaptive framework for the estimation of population size. *CAFSAC Res. Doc*, 88(29):12.

Gavaris, S. 1988b. An adaptive framework for the estimation of population size. *CAFSAC Res. Doc*, 88(29):12.

Gavaris, S. and Ianelli, J. N. 2001. Statistical issues in fisheries stock assessment. *Scand. J. Statistics: Theory and Appl., 29*, pages 245–272.

Gavaris, S., Patterson, K. R., Darby, C. D., Lewy, P., Mesnil, B., Punt, A. E., Cook, R. M., Kell, L. T., O'Brien, C. M., Restrepo, V. R., Skagen, D. W., and Stefansson, G. 2000. Comparison of uncertainty estimates in the short term using real data. ICES (Int. Counc. Explor. Sea) Document C.M. 2000/V:03.

Gislason, A. and Ástþórsson, Ó. S., editors 1997. *Fjölstofnarannsóknir 1992 – 1995 (Multi–species research 1992–1995)*, volume 57 of *Fjölrit Hafrannsóknarstofnunar*. Marine research institute.

Gislason, A., Petursdottir, H., Astthorsson, O. S., Gudmundsson, K., and Valdimarsson, H. 2009. Inter-annual variability in abundance and community structure of zooplankton south and north of iceland in relation to environmental conditions in spring 1990–2007. *Journal of plankton research*, 31(5):541–551.

Graham, M. 1935. Modern theory of exploiting a fishery, and application to north sea trawling. *Journal du Conseil*, 10(3):264–274.

Gross, S. 1980. Median estimation in sample surveys. *Proceedings of the Section on Survey Research Methods*, pages 181–184.

Gudmundsdóttir, Á., Steinarsson, B. Æ., and Stefánsson, G. 1988. A simulation procedure to evaluate the efficiency of some otolith and length sampling schemes. *ICES (Int. Counc. Explor. Sea) C.M. 1988/D:14*, 500:14.

Gudmundsson, G. 2013. Fish stock assessment by time series analysis. Presented at ICES NWWG.

Gulland, J. A. 1965. Estimation of mortality rates. In *Annex to Arctic Fisheries Working Group Report (meeting in Hamburg*, volume 3 of *ICES C.M. Gadoid Fish Comm*, page 9.

Gunnlaugsson, T. 2011. Relatedness between samples quantified and an optimal criterion for match detection approximated. *J. Cetacean Res. Manage*, 12(3).

Gunnlaugsson, T. and Sigurjónsson, J. 1990. Nass-87: Estimation of whale abundance based on observations made onboard icelandic and faroese survey vessels. *Reports of the International Whaling Commission*, 40:571–580.

Halliday, R. and Pinhorn, A. 1990. *The delimitation of fishing areas in the northwest Atlantic*. Northwest Atlantic Fisheries Organization.

Hannesson, S., Jakobsdottir, A., Begley, J., Taylor, L., and Stefansson, G. 2008. On the use of tagging data in statistical multispecies multi-area models of marine populations. *ICES Journal of Marine Science*, 65(9):17–62.

Hannesson, S., Jakobsdottir, A., Begley, J., Taylor, L., and Stefansson, G. 2009. On the use of tagging data in statistical multispecies multi-area models of marine populations. *ICES Journal of Marine Science*, 65(9):1762–1772.

Härkönen, T. 1986. *Guide to the otoliths of the bony fishes of the northeast Atlantic*. Danbiu.

Haug, T., Gjøster, H., Lindstrøm, U., Nilssen, K., and Røttingen, I. 1995. Spatial and temporal variations in northeast atlantic minke whale balaenoptera acutorostrata feeding habits. *Developments in marine biology*, 4:225–239.

Haug, T., Lindstrøm, U., Nilssen, K., Røttingen, I., and Skaug, H. 1996. Diet and food availability for northeast atlantic minke whales balaenoptera acutorostrata. In *ICES Council Meeting Papers. 33*, page 33.

Haug, T., Lindstrøm, U., Nilssen, K., and Skaug, H. J. 1997. On the variation in size and composition of minke whale balaenoptera acutorostrata forestomach contents. *Journal of Northwest Atlantic Fisheries Science*, 22:105–114.

Haug, T., Lindstrøm, U., and Nilssen, K. T. 2002. Variations in minke whale (balaenoptera acutorostrata) diet and body condition in response to ecosystem changes in the barents sea. *Sarsia: North Atlantic Marine Science*, 87(6):409–422.

Helle, K. and Pennington, M. 2004. Survey design considerations for estimating the length composition of the commercial catch of some deep-water species in the northeast Atlantic. *Fisheries Research*, 70(1):55–60.

Hilborn, R. 1990. Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(3):635–643.

Hirst, D., Storvik, G., Aldrin, M., Aanes, S., and Huseby, R. B. 2005. Estimating catch-at-age by combining data from different sources. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(6):1377–1385.

Hooke, R. and Jeeves, T. 1961. Direct Search Soulution of Numerical and Statistical Problems. *ACM Journal*, 8:212–229.

Horton, T., Holdaway, R., Zerbini, A., Hauser, N., Garrigue, C., Andriolo, A., and Clapham, P. 2011. Straight as an arrow: humpback whales swim constant course tracks during long-distance migration. *Biology letters*, 7(5):674–679.

Horwood, J. 1990. *Biology and exploitation of the minke whale.* CRC.

Howell, D. and Bogstad, B. 2010. A combined Gadget/FLR model for management strategy evaluations of the Barents Sea fisheries. *ICES Journal of Marine Science: Journal du Conseil*.

Hrafnkelsson, B. and Stefansson, G. 2004. A model for categorical length data from groundfish surveys. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(7):1135–1142.

Hrafnkelsson, B. and Stefánsson, G. 2004. A model for categorical length data from groundfish surveys. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(7):1135–1142.

Hu, F. and Zidek, J. V. 2002. The weighted likelihood. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 30(3):pp. 347–371.

Hyslop, E. 1980. Stomach contents analysisa review of methods and their application. *Journal of fish biology*, 17(4):411–429.

ICES 2013. General context of ices advice. Technical report, ICES.

ICES, C. 2011. Report of the north western working group. Technical report, International Council for the Exploration of the Sea.

ICES, C. 2014. Report of the working group on the biology and assessment of deep-sea fisheries resources. Technical report, ICES.

IWC 2002. Annual report of the international whaling commission 2001.

IWC 2008. Report of the first intersessional rmp workshop on north atlantic fin whales. Report 3, International Whaling Commission.

IWC 2010. Report of the 2nd intersessional workshop of the north atlantic fin whale *Implementation. J. Cetacean Res. Manage*, 11 (SUPPL. 2):587–627.

IWC 2012. Report of the sub–committee on the revised management procedure (rmp). *Journal of Cetacean Research and Management*, 13:88–95.

IWC 2013. The revised management procedure.

IWC 2015. Report of the intersessional rmp workshop on north atlantic fin whales. Report 3, International Whaling Commission.

Jakobsson, J. and Stefánsson, G. 1999. Management of summer-spawning herring off iceland. *ICES Journal of Marine Science: Journal du Conseil*, 56(6):827–833.

Jennrich, R. I. 1969. Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics*, 40(2):633–643.

Jónsdóttir, I., Marteinsdottir, G., and Campana, S. 2007. Contribution of different spawning components to the mixed stock fishery for cod in icelandic waters. *ICES Journal of Marine Science: Journal du Conseil*, 64(9):1749–1759.

Jonsgård, A. 1982. The food of minke whales (balaenoptera acutorostrata) in northern north atlantic waters. *Report of the International Whaling Commission*, 32:259–262.

Jonsson, G. and Palsson, J. 1992. Íslenskir fiskar (icelandic fishes, in icelandic). *Fjolva utgafan, Reykjavik*.

Jónsson, J. 1994. Fisheries off iceland, 1600-1900. In *ICES marine science symposia. Copenhagen[ICES MAR. SCI. SYMP.]. 1994*.

Kasamatsu, F. and Hata, T. 1985. Notes on minke whales in the okhotsk seawest pacific area. *Report of the International Whaling Commission*, 35:299–304.

Kell, L., Mosqueira, I., Grosjean, P., Fromentin, J., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M., Poos, J., et al. 2007. FLR: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science: Journal du Conseil*, 64(4):640.

Kennedy, A., Zerbini, A. N., Vasquez, O. V., Gandilhon, N., Clapham, P., and Adam, O. 2013. Local and migratory movements of humpback whales (megaptera novaeangliae) satellite-tracked in the north atlantic ocean. *Canadian Journal of Zoology*, 92(1):9–18.

Klinowska, M. 1991. *Dolphins, porpoises and whales of the world: the IUCN Red Data Book*. Iucn.

Kupca, V. 2004. A standardized database for fisheries data. ICES (Int. Counc. Explor. Sea) Document C.M. 2004/FF:15.

Kupca, V. 2006. Dst^2 data warehouse. Technical report, Marine Research Institute, Reykjavík, Iceland.

Kupca, V. and Sandbeck, P. 2003. dst^2 Datawarehouse structure and data import. In *dst^2: Development of structurally detailed statistically testable models of marine populations*, volume 98 of *Technical Report*, pages 37–46. Marine Research Institute, Reykjavik.

Larsen, F. and Kapel, F. O. 1981. Collection of biological material of minke whales off west greenland, 1979. *Rep. int. Whal. Commn*, 31:279–85.

Laurenson, C., Johnson, A., and Priede, I. 2005. Movements and growth of monkfish *Lophius piscatorius* tagged at the shetland islands, northeastern atlantic. *Fisheries Research*, 71(2):185–195.

Lindstrøm, U., Haug, T., and Nilssen, K. T. 1997. Diet studies based on contents from two separate stomach compartments of northeast atlantic minke whales balaenoptera acutorostrata. *Sarsia*, 82(1):63–68.

Lindstrøm, U., Smout, S., Howell, D., and Bogstad, B. 2009. Modelling multi-species interactions in the Barents Sea ecosystem with special emphasis on minke whales and their interactions with cod, herring and capelin. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56(21-22):2068–2079.

Link, J., Fulton, E., and Gamble, R. 2010. The Northeast US Application of ATLANTIS: A full system model exploring marine ecosystem dynamics in a living marine resource management context. *Progress In Oceanography*.

Livingston, P. A. and Methot, R. D. 1998. Incorporation of predation into a population assessment model of eastern bering sea walleye pollock. *Fishery stock assessment models*, pages 663–678.

Lotka, A. J. 1925. *Elements of physical biology*. Williams and Wilkins, Baltimore, Md.

Magnusson, A., Punt, A. E., and Hilborn, R. 2012. Measuring uncertainty in fisheries stock assessment: the delta method, bootstrap, and mcmc. *Fish and Fisheries*.

Magnússon, K. 1995. An overview of the multispecies vpa theory and applications. *Reviews in Fish Biology and Fisheries*, 5(2):195–212.

MareFrame 2013. Mareframe: Co-creating ecosystem-based fisheries management solutions. Technical report, FP7-KBBE-2013-7.

Mate, B., Mesecar, R., and Lagerquist, B. 2007. The evolution of satellite-monitored radio tags for large whales: One laboratory's experience. *Deep Sea Research Part II: Topical Studies in Oceanography*, 54(3):224–247.

Matthews, C., Luque, S., Petersen, S., Andrews, R., and Ferguson, S. 2011. Satellite tracking of a killer whale (orcinus orca) in the eastern canadian arctic documents ice avoidance and rapid, long-distance movement into the north atlantic. *Polar biology*, 34(7):1091–1096.

Maunder, M. N. and Punt, A. E. 2012. A review of integrated analysis in fisheries stock assessment. *Fisheries Research*.

Mauritzen, M., Derocher, A. E., Wiig, Ø., Belikov, S. E., Boltunov, A. N., Hansen, E., and Garner, G. W. 2002. Using satellite telemetry to define spatial population structure in polar bears in the norwegian and western russian arctic. *Journal of Applied Ecology*, 39(1):79–90.

McCarthy, P. J. and Snowden, C. B. 1985. The bootstrap and finite population sampling. *Vital Health Stat 2*, 95:1–23.

Methot, R. D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. *Amer. Fish. Soc. Symp.*, 6:66–82.

Methot Jr, R. D. and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.

Millar, R. B. 2002. Reference priors for bayesian fisheries models. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(9):1492–1502.

Mitchell, E. 1974. *Preliminary report on Newfoundland fishery for minke whales (Balaenoptera acutorostrata)*. Office of the International Commission on Whaling.

Möllmann, C., Lindegren, M., Blenckner, T., Bergström, L., Casini, M., Diekmann, R., Flinkman, J., Müller-Karulis, B., Neuenfeldt, S., Schmidt, J. O., et al. 2013. Implementing ecosystem-based fisheries management: from single-species to integrated ecosystem assessment and advice for baltic sea fish stocks. *ICES Journal of Marine Science: Journal du Conseil*, page fst123.

MRI 2003. A programme for a two year feasibility study on cetaceans in icelandic waters.

MRI 2007. Nytjastofnar sjávar 2006/2007. aflahorfur fiskveiðiárið 2006/2007. state of marine stocks in icelandic waters 2006/2007. prospects for the quota year 2006/2007.

MRI 2012a. Þættir úr vistfræði sjávar 2011. environmental conditions in icelandic waters 2011.

MRI 2012b. Nytjastofnar sjávar 2011/2012. aflahorfur fiskveiðiárið 2012/2013. state of marine stocks in icelandic waters 2011/2012. prospects for the quota year 2012/2013.

MRI 2013. Nytjastofnar sjávar 2013/2014. aflahorfur fiskveiðiárið 2013/2014. state of marine stocks in icelandic waters 2013/2014. prospects for the quota year 2013/2014.

Myers, R. A. and Cadigan, N. G. 1995. Statistical analysis of catch-at-age data with correlated errors. *Canadian journal of fisheries and aquatic sciences(Print)*, 52(6):1265–1273.

NAMMCO 1997. Nammco annual report 1997.

NAMMCO 2009. Nammco annual report 2008.

Nemoto, T. 1959. Food of baleen whales with reference to whale movements. *Sci. Rep. Whales Res. Inst*, 14:149–290.

Neve, P. 2000. The diet of the minke whale in greenland–a short review. *Minke Whales, Harp and Hooded Seals: Major Predators in the North Atlantic Ecosystem*, pages 92–96.

Nielsen, A. and Berg, C. W. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research*, 158(0):96 – 101. SI: Selectivity.

Nielsen, R., Mattila, D., Clapham, P., and Palsbøll, P. 2001. Statistical approaches to paternity analysis in natural populations and applications to the north atlantic humpback whale. *Genetics*, 157(4):1673–1682.

Nordøy, E. and Blix, A. S. 1992. Diet of minke whales in the northeastern atlantic. *Report of the International Whaling Commission*, 42:393–398.

Oehlert, G. 1992. A note on the delta method. *American Statistician*, pages 27–29.

Økland, J., Haaland, Ø., and Skaug, H. 2010. A method for defining management units based on genetically determined close relatives. *ICES Journal of Marine Science: Journal du Conseil*, 67(3):551–558.

Olsen, E. and Holst, J. C. 2001. A note on common minke whale (balaenoptera acutorostrata) diets in the norwegian sea and the north sea. *Journal of Cetacean Research and Management*, 3(2):179–184.

Þórðarson, G. and Elvarsson, B. Þ. 2013. Exploratory gadget stock assessment of greenland halibut (*Reinhardtius hippoglossoides*) in subareas v, vi, xii, and xiv. Presented at ICES NWWG benchmark meeting.

Þórðarson, G., Elvarsson, B. Þ., and Kristinsson, K. 2011. Statistical estimation of likelihood components and bootstrap approach for estimating uncertainty in the gadget stock assessment of golden redfish and tusk. Progress report.

Palsbøll, P., Zachariah Peery, M., and Berube, M. 2010. Detecting populations in the "ambiguous" zone: kinship-based estimation of population structure at low genetic divergence. *Molecular Ecology Resources*, 10(5):797–805.

Palsbøll, P. J. 1999. Genetic tagging: contemporary molecular ecology. *Biological Journal of the Linnean Society*, 68(1-2):3–22.

Pálsson, Ó. K. and Björnsson, H. 2011. Long-term changes in trophic patterns of iceland cod and linkages to main prey stock sizes. *ICES Journal of Marine Science: Journal du Conseil*, 68(7):1488–1499.

Pálsson, O. K., Gislason, A., Guðfinnsson, H. G., Gunnarsson, B., Ólafsdóttir, S. R., Petursdottir, H., Sveinbjörnsson, S., Thorisson, K., and Valdimarsson, H. 2012. Ecosystem structure in the iceland sea and recent changes to the capelin (mallotus villosus) population. *ICES Journal of Marine Science: Journal du Conseil*, 69(7):1242–1254.

Pálsson, Ó. K., Gíslason, Á., Gunnarsson, B., Guðfinnsson, H. G., Valdimarsson, H., and Pétursdóttir, H. 2014. Meginþættir í vistkerfi íslandshafs og breytingar á lífsháttum loðnu. *Náttúrufræðingurinn.*

Palsson, O. K., Schopka, S. A., Stefansson, G., and Steinarsson, B. A. 1989. Icelandic groundfish survey data used to improve precision in stock assessments. *J. Northw. Atl. Fish. Sci.*, 9:53–72.

Pálsson, Ó. K., Steinarsson, B. Æ., Jónsson, E., Guðmundsson, G., Jénsson, G., Stefánsson, G., Björnsson, H., and Schopka, S. A. 1997. Icelandic groundfish survey. *ICES CM 1997/Y:29.*

Pampoulie, C., Danielsdottir, A., Berube, M., Palsbøll, P., Arnarson, A., Gunnlaugsson, T., Olafsdottir, D., Øien, N., Witting, L., and Vikingsson, G. 2008. Lack of genetic divergence among samples of the north atlantic fin whale collected at feeding grounds; congruence among microsatellite loci and mtdna in the new icelandic dataset. *IWC SC/60/PFI11.*

Pampoulie, C., Elvarsson, B. Þ., Ólafsdóttir, D., rg Hauksdóttir, S., Skírnisdóttir, S., Ólafsson, K., Magnúsdóttir, S., Chosson, V., Halldórsson, S. D., Ólafsdóttir, D., Gunnlaugsson, Þ., Daníelsdóttir, A. K., and Víkingsson, G. A. 2011. Mother-fetus pair analyzes in the atlantic fin whale (*Balaenoptera physalus*): an insight into the migration pattern of the species? *IWC SC/63/RMP1 presented to the IWC Scientific Committee.*

Pampoulie, C., Ruzzante, D. E., Chosson, V., Jörundsdóttir, T. D., Taylor, L., Thorsteinsson, V., Daníelsdóttir, A. K., and Marteinsdóttir, G. 2006. The genetic structure of atlantic cod (gadus morhua) around iceland: insight from microsatellites, the pan i locus, and tagging experiments. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(12):2660–2674.

Parliament, E. 2008. Directive 2008/56/ec of the european parliament and of the council of 17 june 2008 establishing a framework for community action in the field of marine environmental policy (marine strategy framework directive). *Official Journal of the European Union L*, 164:19–40.

Parsons, K., Durban, J., Claridge, D., Herzing, D., Balcomb, K., and Noble, L. 2006. Population genetic structure of coastal bottlenose dolphins (tursiops truncatus) in the northern bahamas. *Marine Mammal Science*, 22(2):276–298.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V., Skagen, D. W., and Stefansson, G. 2001. Estimating uncertainty in fish stock assessment and forecasting. *Fish and Fisheries*, 2(2):125–157.

Paxton, C., Burt, M., Hedley, S., Víkingsson, G., Gunnlaugsson, T., and Desportes, G. 2009. Density surface fitting to estimate the abundance of humpback whales based on the nass-95 and nass-2001 aerial and shipboard surveys. *NAMMCO Sci. Publ*, 7:143–159.

Peakall, R., Ebert, D., Cunningham, R., and Lindenmayer, D. 2006. Mark–recapture by genetic tagging reveals restricted movements by bush rats (rattus fuscipes) in a fragmented landscape. *Journal of Zoology*, 268(2):207–216.

Peery, M. Z., Beissinger, S. R., House, R. F., Bérubé, M., Hall, L. A., Sellas, A., and Palsbøll, P. J. 2008. Characterizing source-sink dynamics with genetic parentage assignments. *Ecology*, 89(10):2746–2759.

Pella, J. J. and Tomlinson, P. K. 1969. *A generalized stock production model*. Inter-American Tropical Tuna Commission.

Pennington, M. and Volstad, J. H. 1994a. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. *Biometrics*, pages 725–732.

Pennington, M. and Volstad, J. H. 1994b. Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. *Biometrics*, 50:1–8.

Petitgas, P. 2001. Geostatistics in fisheries survey design and stock assessment: models, variances and applications. *Fish and Fisheries*, 2(3):231–249.

Pierce, G., Boyle, P., et al. 1991. A review of methods for diet analysis in piscivorous marine mammals. *Oceanography and Marine Biology: An Annual Review*, 29.

Pierce, G., Santos, M., Reid, R., Patterson, I., and Ross, H. 2004. Diet of minke whales balaenoptera acutorostrata in scottish (uk) waters with notes on strandings of this species in scotland 1992-2002. *Journal of the Marine Biological Association of the UK*, 84(6):1241–1244.

Pike, D., Gunnlaugsson, T., Elvarsson, B., and Víkingsson, G. 2011a. Correcting perception bias for icelandic aerial surveys, 2007 and 2009. *NAMMCO SC/18/AESP/08*.

Pike, D., Gunnlaugsson, T., and Víkingsson, G. 2008a. T-nass icelandic aerial survey: Survey report and a preliminary abundance estimate for minke whales. *Paper IWC SC/60/PFI12, 29pp.*

Pike, D., Paxton, C., Gunnlaugsson, T., and Víkingsson, G. 2009. Trends in the distribution and abundance of cetaceans from aerial surveys in icelandic coastal waters, 1986-2001. *NAMMCO Sci. Publ*, 7:117–142.

Pike, D. G., Gunnlaugsson, T., Elvarsson, B. Þ., and Víkingsson2, G. A. 2011b. Correcting perception bias for icelandic aerial surveys, 2007 and 2009. Technical report, NAMMCO.

Pike, D. G., Gunnlaugsson, T., Víkingsson, G. A., and Mikkelsen, B. 2008b. Estimates of the abundance of fin whales (balaenoptera physalus) from the t-nass icelandic and faroese ship surveys conducted in 2007. *IWC SC/60/PFI13-revised*.

Plagányi, É. 2007. *Models for an ecosystem approach to fisheries*. Food & Agriculture Org.

Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D., et al. 2014. Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. *Fish and Fisheries*, 15(1):1–22.

Planes, S., Jones, G., and Thorrold, S. 2009. Larval dispersal connects fish populations in a network of marine protected areas. *Proceedings of the National Academy of Sciences*, 106(14):5693–5697.

Punt, A. 1999. A full description of the standard baleen ii model and some variants thereof. *J. Cetacean Res. Manage*, 1:267–276.

Punt, A. E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. *Fisheries Research*, 65(1):391–409.

Punt, A. E. and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the bayesian approach. *Reviews in Fish Biology and Fisheries*, 7(1):35–63.

R Development Core Team 2011. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rao, J. N. K. and Wu, C. F. J. 1988. Resampling inference with complex survey data. *Journal of the American Statistical Association*, 83(401):231–241.

Rayner, G. 1940. *Whale marking: progress and results to December 1939*. University Press.

Recchia, C. and Read, A. 1989. Stomach contents of harbour porpoises, phocoena phocoena (l.), from the bay of fundy. *Canadian Journal of Zoology*, 67(9):2140–2146.

Reynolds, J. and Templin, W. 2004. Detecting specific populations in mixtures. *Environmental biology of fishes*, 69(1):233–243.

Richards, L. 1991. Use of contradictory data sources in stock assessments. *Fisheries research*, 11(3-4):225–238.

Sæmundsson, B. 1932. *Spendýrin (Mammalia Islandiae)*. Bókaverslun Sigfúsar Eymundssonar, Reykjavík. 437 pages.

Sæmundsson, B. 1939. *Mammalia*. The Zoology of Iceland IV:1-37. Bókaverslun Sigfúsar Eymundssonar, Reykjavík.

Schnute, J. T. and Hilborn, R. 1993. Analysis of contradictory data sources in fish stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, 50(9):1916–1923.

Seegar, W. S., Cutchis, P. N., Fuller, M. R., Surer, J., Bhatnagar, V., and Wall, J. G. 1996. Fifteen years of satellite tracking development and application to wildlife research and conservation. *Johns Hopkins APL Technical Digest*, 17(4):401–411.

Sergeant, D. 1963. Minke whales, balaenoptera acutorostrata lacepede, of the western north atlantic. *Journal of the Fisheries Board of Canada*, 20(6):1489–1504.

Shepherd, J. 1999. Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. *ICES Journal of Marine Science: Journal du Conseil*, 56(5):584–591.

Sigurdsson, T., Hjorleifsson, E., Bjornsson, H., and Palsson, O. K. 1997. Fall groundfish survey in icelandic waters (stofnmaeling botnfiska a Islandsmidum haustid 1996). Technical Report 61, Marine Research Institute. (in Icelandic).

Sigurjonsson, J. 1988. Operational factors of the icelandic large whale fishery. *Report of the International Whaling Commission*, 38:327–333.

Sigurjónsson, J., Galan, A., and Vikingsson, G. 2000. A note on stomach contents of minke whales (balaenoptera acutorostrata) in icelandic waters. *Minke Whales, Harp and Hooded Seals: Major Predators in the North Atlantic Ecosystem*, pages 82–90.

Sigurjónsson, J. and Gunnlaugsson, T. 2006. Revised catch series and cpue for fin whales taken from the early modern whaling land stations in iceland. *IWC SC/58/PFI4*.

Sigurjónsson, J. and Víkingsson, G. 1997. Seasonal abundance of and estimated food consumption by cetaceans in icelandic and adjacent waters. *Journal of Northwest Atlantic Fishery Science*, 22:271–287.

Silva, M. A., Prieto, R., Jonsen, I., Baumgartner, M. F., and Santos, R. S. 2013. North atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: Building up energy reserves for the journey? *PloS one*, 8(10):e76507.

Sinclair, M., Arnason, R., Csirke, J., Karnicki, Z., Sigurjonsson, J., Skjoldal, H. R., and Valdimarsson, G. 2002. Responsible fisheries in the marine ecosystem. *Fisheries Research*, 58(3):255 – 265.

Singh, W., Thordarson, G., Haputhantri, S., and Stefansson, G. 2011. Identification of year classes in length distributions.

Sitter, R. R. 1992. A resampling procedure for complex survey data. *Journal of the American Statistical Association*, 87(419):755–65.

Skaug, H. J. 2001. Allele-sharing methods for estimation of population size. *Biometrics*, 57(3):750–756.

Skaug, H. J., Gjøsæter, H., Haug, T., Nilssen, K. T., and Lindstrøm, U. 1997. Do minke whales (balaenoptera acutorostrata) exhibit particular prey preferences. *J. Northw. Atl. Fish. Sci*, 22:91–104.

Skúladóttir, U., Gíslason, A., Bragason, G., and Brynjólfsson, S. 2009. Útbreiðsla ljósátu á norður-austurmiðum að sumarlagi (distribution of euphausiids, north and east of iceland during summer). *Hafrannsóknir, fjölrit*, 145:66–74.

Smith, S. J. 1997. Bootstrap confidence limits for groundfish trawl survey estimates of mean abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(3):616–630.

Smith, T. D. 2002. A history of fisheries and their science and management. *Handbook of Fish Biology and Fisheries, Volume 2: Fisheries*, pages 61–83.

Solmundsson, J., Jonsson, E., and Bjornsson, H. 2010. Phase transition in recruitment and distribution of monkfish (lophius piscatorius) in icelandic waters. *Marine biology*, 157(2):295–305.

Spiegelhalter, D., Thomas, A., Best, N., and Gilks, W. 1996. Bugs 0.5: Bayesian inference using gibbs sampling manual (version ii). *MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK*.

Stefansson, G. 1998. Comparing different information sources in a multispecies context. In Funk, F., II, T. Q., Heifetz, J., Ianelli, J., Powers, J., Schweigert, J., Sullivan, P., and Zhang, C., editors, *Fishery Stock Assessment Models: Proceedings of the international symposium; Anchorage 1997*, 15th Lowell Wakefield Fisheries Symposium, pages 741–758.

Stefansson, G. 2003. Issues in multispecies models. *Natural Resource Modeling*, 16(4):415–437.

Stefansson, G. 2005a. Growth models in population simulations. In *Final Report: dst^2: Development of structurally detailed statistically testable models of marine populations*, volume 118 of *Technical Report*, pages 73–83. Marine Research Institute, Reykjavik.

Stefansson, G. 2005b. Growth models in population simulations. *Final Report DST2: Development of Structurally Detailed Statistically Testable Models of Marine Populations, Technical Report*, 118:73–80.

Stefánsson, G. and Pálsson, Ó. 1997a. BORMICON. boreal migration and consumption model. *Marine Research Institute Report. 58. 223 p.*

Stefansson, G. and Palsson, O. 1998. A framework for multispecies modelling of Arcto-boreal systems. *Reviews in Fish Biology and Fisheries*, 8(1):101–104.

Stefánsson, G. and Pálsson, Ó. K. 1997b. Statistical evaluation and modelling of the stomach contents of icelandic cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences*, 54(1):169–181.

Stefánsson, G., Sigurjónsson, J., and Víkingsson, G. 1997. On dynamic interactions between some fish resources and cetaceans off iceland based on a simulation model. *J. Northwest Atl. Fish. Sci.*, 22:357–370.

Stewart, I. J., Hicks, A., Taylor, I., Thorson, J. T., Wetzel, C., and Kupschus, S. 2012. A comparison of stock assessment uncertainty estimates using maximum likelihood and bayesian methods implemented with the same model framework. *Fisheries Research.*

Tamura, T. and Fujise, Y. 2002. Geographical and seasonal changes of the prey species of minke whale in the northwestern pacific. *ICES Journal of Marine Science: Journal du Conseil*, 59(3):516.

Taylor, I., Stewart, I., Hicks, A., Garrison, T., Punt, A., Wallace, J., Wetzel, C., Thorson, J., Takeuchi, Y., Monnahan, C., and other contributors. 2014. *r4ss: R code for Stock Synthesis*. R package version 1.22.1.

Taylor, L. 2002. Stock dynamics and long term assessment of haddock (*Melanogrammus aeglefinus*) in icelandic waters. Presented at ICES North-Western working group (available as http://www.hi.is/ gunnar/papers/ltdyn.pdf).

Taylor, L. 2003. Definition of areas in Icelandic waters. In dst^2: *Development of structurally detailed statistically testable models of marine populations*, volume 98 of *Technical Report*, pages 222–230. Marine Research Institute, Reykjavik.

Taylor, L. 2011. Gadget models of cod–shrimp interactions in icelandic waters. Technical report, Technical Report RH–03–2011, Science Institute, University of Iceland.

Taylor, L., Begley, J., Kupca, V., and Stefansson, G. 2007. A simple implementation of the statistical modelling framework Gadget for cod in Icelandic waters. *African Journal of Marine Science*, 29(2):223–245.

Tinker, M. T., Doak, D. F., Estes, J. A., Hatfield, B. B., Staedler, M. M., and Bodkin, J. L. 2006. Incorporating diverse data and realistic complexity into demographic estimation procedures for sea otters. *Ecol Appl*, 16(6):2293–312.

Tjelmeland, S. and Bogstad, B. 1989. *MULTSPEC: The manual.* Institute of Marine Research, Bergen.

Tomilin, A. 1957. Mammals of the ussr and adjacent countries. *Vol. IX. Cetacea. Moscow, Soviet Union (English translation, 1967, Israel Program for Scientific Translations, Jerusalem, Israel).*

Trenkel, V. M., Bravington, M. V., and Lorance, P. 2012. A random effects population dynamics model based on proportions-at-age and removal data for estimating total mortality. *Canadian Journal of Fisheries and Aquatic Sciences*, 69(11):1881–1893.

UN 2002. Johannesburg declaration on sustainable development. Report, United Nations.

V., B. and K., L. 2009. Rannsóknir á sandsíli (an initiation of sandeeel monitoring in iceland). *Hafrannsóknir*, 145:36–41.

Vandermeer, J. 1978. Choosing category size in a stage projection matrix. *Oecologia*, 32(1):79–84.

Víkingsson, G., Elvarsson, B. Þ., Ólafsdóttir, D., Chosson, V., Sigurjónsson, J., and Galan, A. 2014. Recent changes in the diet composition of common minke whales (balaenoptera acutorostrata) in icelandic waters. – consequence of climate change? *Marine Biology Research*, 10(2):138–152.

Víkingsson, G. and Heide-Jörgensen, M. P. 2014. Movement of minke whales, tracked by satellite in icelandic waters 2001-2011. *Marine Mammal Science (In Press).*

Víkingsson, G., Ólafsdóttir, D., and Sigurjónson 2003. Geographical and seasonal variation in the diet of harbour porpoises (*Phocanema phocanema*) in icelandic coastal waters. *NAMMCO Sci. Publ*, 5:243–270.

Víkingsson, G. and Valdimarsson, H. 2006. Hvalir og umhverfisþættir (whales and the environment). *Hafrannsöknir, fjölrit*, 125:30–33.

Víkingsson, G. A. and Gunnlaugsson, Þ. 2005. Stock structure of fin whales (balaenoptera physalus) in the north atlantic–indications from non-genetic data. *IWC/SC/57/PFI3. Available from the International Whaling Commission (http://www. iwcoffice. org/).*

Víkingsson, G. A., Pike, D. G., Desportes, G., Øien, N., Gunnlaugsson, T., and Bloch, D. 2009. Distribution and abundance of fin whales (balaenoptera physalus) in the northeast and central atlantic as inferred from the north atlantic sightings surveys 1987-2001. *NAMMCO Sci. Publ*, 7:49–72.

Víkingsson, G. A., Pike, D. G., Schleimer, A., Valdimarsson, H., Gunnlaugsson, T., Silva, T., Elvarsson, B. Þ., Mikkelsen, B., Øien, N., Desportes, G., Bogason, V., and Hammond, P. 2015. Distribution, abundance and feeding ecology of rorqual cetaceans in icelandic waters: have recent environmental changes had an effect? *Frontiers in Ecology and Evolution.* Accepted for publication.

Vilhjálmsson, H. 1994. *The Icelandic Capelin Stock: Capelin, Mallotus villosus (Müller) in the Iceland-Greenland-Jan Mayen area.* Marine Research Institute.

Volterra, V. 1928. Variations and fluctuations of the number of individuals in animal species living together. *J. Cons. Int. Explor. Mer*, 3(1):3–51.

Voss, R., Quaas, M. F., Schmidt, J. O., and Hoffmann, J. 2014. Regional trade-offs from multi-species maximum sustainable yield (mmsy) management options. *MARINE ECOLOGY PROGRESS SERIES*, 498:1–12.

Walsh, S. J. and Morgan, M. J. 2004. Observations of natural behaviour of yellowtail flounder derived from data storage tags. *ICES Journal of Marine Science: Journal du Conseil*, 61(7):1151–1156.

Wang, X. and Zidek, J. V. 2005. Selecting likelihood weights by cross-validation. *Ann. Statist*, 33(2):463–500.

Wennevik, V., Jørstad, K., Dahle, G., and Fevolden, S. 2008. Mixed stock analysis and the power of different classes of molecular markers in discriminating coastal and oceanic atlantic cod (gadus morhua l.) on the lofoten spawning grounds, northern norway. *Challenges to Marine Ecosystems*, pages 7–25.

Wilkinson, D. J. 2006. Parallel bayesian computation. *Statistics textbooks and monographs*, 184:477.

Windsland, K., Lindstrøm, U., Nilssen, K., and Haug, T. 2007. Relative abundance and size composition of prey in the common minke whale diet in selected areas of the northeastern atlantic during 2000-04. *J. Cetacean Res. Manage*, 9:167–178.