Spatial modelling for monitoring and management of marine metapopulations

Gavin Fay

A dissertation
submitted in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2012

Reading Committee:
André E. Punt, Chair
Timothy E. Essington
Ray Hilborn

Program Authorized to Offer Degree:
School of Aquatic and Fishery Sciences

University of Washington

Abstract

Spatial modelling for monitoring and management of marine metapopulations

Gavin Fay

Chair of the Supervisory Committee:
Professor André E. Punt
School of Aquatic and Fishery Sciences

Accounting for spatial complexity provides a diversity of challenges for natural resource management. Challenges arise from uncertainties in spatial stock structure, heterogeneity in environmental and ecological constraints, and from spatial differences in exploitation and management actions. Simulation frameworks were used to analyse methods that account for spatial heterogeneity when (1) identifying spatial population trends, (2) evaluating the performance of monitoring designs among multiple sites, (3) estimating spatial distribution of sea lion prey, (4) examining energetic implications of foraging strategies given uncertainty, and (5) determining how spatial dynamics of exploited populations and fishing fleets affects performance of harvest strategies.

Linear state-space models using the Kalman filter were developed to estimate trends in pup production for Steller sea lions (*Eumetopias jubatus)*. Models assuming spatial correlation in trend among rookeries were more robust to stock structure assumptions, and estimated trends and abundance even given missing data. Relative gains in performance when optimising monitoring designs for pup production were evaluated using Canonical Correspondence Analysis and by assessing how frequently optimal designs out-performed random designs. Optimal allocation of monitoring effort depended on the metric of interest. Poorest performance occurred with suboptimal designs for current trend and individual rookery numbers. Bayesian hierarchical models were used to characterise the spatial distribution of fish species in the Gulf of Alaska. Spatial autocorrelation was prevalent in all species, with estimates of abundance generally lower than those obtained using models ignoring spatial correlation.

Individual-based modelling of sea lion foraging examined how spatial persistence of prey and the choice of foraging strategy impacted the ability of foragers to meet energetic requirements. Tradeoffs and interactions were observed among model components, successful strategies involved either low uncertainty about prey distribution, or placed substantial emphasis on previous experience. The performance of harvest strategies for the blue eye trevalla (*Hyperoglyphe antarctica*) fishery in southeast Australia was evaluated using Management Strategy Evaluation. Appropriate weighting of spatial data was required to meet management objectives, although uncertainties regarding natural

mortality and stock-recruitment steepness dominated variation in performance. In summary, spatial variability ought to be considered when modelling and managing marine resources, however appropriate scales of response are necessary.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

"Everything starts somewhere, though many physicists disagree. But people have always been dimly aware of the problem with the start of things. They wonder how the snowplough driver gets to work, or how the makers of dictionaries look up the spelling of words."

Terry Pratchett

Such are the feelings I have when thinking about trying to begin to thank the very great many people who have helped me in the creation of this dissertation, and during my time in graduate school. The community that is the School of Aquatic and Fishery Sciences (SAFS) at the University of Washington is an exceptional one, full of wise and wonderful people, with faculty and staff dedicated to exemplary learning and research, and a student body that have been my colleagues and friends during this journey.

I feel truly lucky to have been able to study and work with my advisor, André Punt. Thank you, André, for sharing your brilliance, your patience, and for being so generous with your time. I am truly grateful. As members of my reading committee, Tim Essington and Ray Hilborn provided me with timely, rigorous feedback on drafts of my chapters, and always pushed me to dig a little deeper in terms of assimilating model results into a bigger chapter. I thank the other members of my committee, Dave Beauchamp, Tom Gelatt, Tony Smith, and my GSR, Dee Boersma. Tom provided humour and excellent insight on Steller sea lions. Dave was invaluable in providing guidance for the bioenergetics modelling and I thank him for very rigorous reviews of my work. Tony not only guided me on my thesis work and facilitated my visits to Hobart, but also gave constant support and encouragement in all areas of my life, and I am deeply thankful for our many chats during my time in Hobart, whether they were in the office or at a show.

Dee went above and beyond what was necessary as my Graudate School Representative, and I am thankful for her enthusiasm and keen questioning of my work.

The group known as the Punt lab have become fast friends and admired professional colleagues. Teresa A'mar, Jason Cope, Melissa Haltuch, and Doug Kinzey, I wish we could work in the same office together every day. My sincere thanks also to Charlotte Boyd, John Brandon, Tommy Garrison, Felipe Hurtado, Toshi Kitakado, Dusanka Poljak, Ingrid Spies, Ilona Stobutzki, Nan-Jay Su, Cody Szuwalski, Jim Thorson, Rod Towell, Sheng-Ping Wang, Chantel Wetzel, and Motoki Wu. It has been fantastic working with you all.

Countless others within the SAFS community have helped me in ways they may not know. I would like to thank in particular the members of the Essington, Hilborn, Francis, Gallucci and Van Blaricom labs, and also to my soccer pals, the Dorsal Fins United. Particular thanks to David Armstrong, Anne Beaudreau, Amanda Bradford, Trevor Branch, Alison Carter, John Field, Bob Francis, Sarah Gaichas, Don Gunderson, Donna Hauser, Mark Henderson, Allan Hicks, Mary Hunsicker, Neala Kendall, Erin McClelland. Jen McIntyre, Michelle Lander, Josh London, Arni Magnusson, Kristin Marshall, Carey McGilliard, Carolina Minte-Vera, Ivonne Ortiz, Lauren Rogers, Alex da Silva, Ian Stewart, Ian Taylor, Jodie Toft, Juan Valero, Eric Ward, and Alex Zerbini.

Funding during my graduate studies was provided by NOAA Fisheries' National Marine Mammal Laboratory, the NOAA Northwest Fisheries Science Center, SAFS, the UW Center for Quantitative Science, and the CSIRO Marine and Atmospheric Research and Wealth from Oceans Research Flagship. Thank you to SAFS graduate advisers Nichole Byrne Lau, April Wilkinson, Lin Murdock, Scott Schafer, and Addi Daisley, who provided assistance with so many things, and who all at one time or another made sure I got my paperwork in order. Thanks also to many at CSIRO in Hobart for support, most notably David Smith, Geoff Tuck, Campbell Davies and Meredith Prendergast. Support

and flexibility during the final stages was also provided by Jason Link and Laura Shulman.

My colleagues at CSIRO and the NOAA Northeast Fisheries Science Center have also helped me in many ways. Thanks to the SESSF assessment team, none of whom I believe are Cylons: Jemery Day, Malcolm Haddon, Neil Klaer, Rich Little, Robin Thomson, Geoff Tuck, Judy Upston, Sally Wayte, and Athol Whitten. Thanks also to Natalie Dowling, Alan Williams, Ingrid van Putten, Beth Fulton, Mike Fuller, and Jessica Nilsson. My new friends and colleagues in the EcoAP group are thanked for providing a supportive atmosphere as I finished things up while ramping into new work. Particular thanks to Mike Fogarty, Jason Link, Scott Large, and Robert Gamble.

To my dear friends and family, I thank you for always championing my endeavours, and always being there for me. My Seattle friend-base, including the many people of FIUTS, I hold you all dear, and without you this experience at the UW would not have been half what it was. The price to pay for knowing excellent people is that we tend to secrete ourselves in every corner of the globe and so do not see each-other as often as we would like. No matter, I know where to find you. Special thanks to Anna Bhatia, Sachin Bhatia, Katie Boyes, Sarah Brown, Diane Camenisch, Bill and Debbie Chiles, Nathan Fay, Nicola Finch, Jennie Findlay, Dave Holmberg, Shannon McCluskey, Michael Roberts, Meghann Rother, Jordan Vincent, and Anne Williams.

Thanks to Kaia and Atticus for providing constant joy and companionship, even when they refused to do anything but sit on my work/computer/head.

To my Mum and Dad, I have nothing but love and appreciation for everything you do for me. You always support me in everything I do. Thank you.

Thanks especially to Michelle, for joy, laughter, and all things indescribable. Love you.

Introduction

"Everything is related to everything else, but near things are more related than distant things"

Waldo Tobler

All organisms are spatial, but some organisms are more spatial than others[1]. Tobler's (1970) first law of geography succinctly explains spatial autocorrelation, a pervasive feature when describing ecological systems. Spatial heterogeneity, or patchiness, arises from spatial and temporal variation in the distribution of abundance of vital resources, as well as in ecological processes (e.g. Levin 1976, 1992, Wiens 1989, Legendre and Legendre 1998, Kritzer and Sale 2006). Indeed, spatial structure has become widely recognized among ecologists as an important component of ecological systems (Legendre 1993). Adoption of such concepts in analysis of ecological problems has been relatively recent, however, and many ideas continue to be developed. In the Introduction to a special feature on spatial dynamics in the journal *Ecology* in 1994, Kareiva emphasized the need for ecologists to experimentally test, recognize and incorporate understanding of spatial patterns into research, invoking the work of science fiction writer Gene Rodenberry when declaring "Space: the final frontier for ecological theory" (Kareiva 1994).

The fields of island biogeography, landscape ecology, and the metapopulation ideas have all become entrenched perspectives for terrestrial ecologists, conservation agencies, and resource managers. The concept of metapopulation dynamics (Levins 1970, Gilpin and

---

[1] Many apologies to George Orwell.

Hanski 1991), where a population is viewed as a "population of populations" has gained particular traction from a modeling perspective, notably in cases where issues of concern relate to species persistence and extinction at local scales, though metapopulation theory has embraced more complicated (and perhaps more realistic) instances (Burgman et al. 1993). There has been much development in the application of the metapopulation concept to the marine environment (Kritzer and Sale 2006), though this has proved somewhat less tangible and muddled in definition. Acknowledgement of the importance of spatial structure in driving the ecology of marine organisms, particular from a resource management perspective, has tended to lag somewhat behind the terrestrial sphere. This is likely due to the opacity of the medium – the importance of habitat fragmentation and connectivity is easy to see and explain in a terrestrial environment, less so in an aquatic environment where boundaries among habitat and driving structures are harder to quantify. Nevertheless, acknowledgement of the importance of spatial structuring in marine populations is rapidly becoming common.

Accounting for spatial complexity when assessing, monitoring, and managing populations of marine organisms leads to a diversity of challenges (Cianelli et al. 2008, Lorenzen et al. 2010). Furthermore, the importance and implications of accounting for spatial heterogeneity within marine ecosystems are often poorly understood. Mismatches in spatial scale between resource dynamics and assumptions related to providing advice leading to (or not) management actions have been implicated in many well-publicised failures in marine resource management (e.g. Hilborn et al. 2005). Challenges can arise due to uncertainties in the nature of spatial 'stock' structure, spatial heterogeneity in the environmental and ecological constraints acting on those populations, and the results of spatial differences in exploitation or other management actions. These uncertainties tend to be compounded because a large number of hypotheses regarding the spatial components of dynamics of the populations in question are often equally supported by the available data. However, it is likely that different assumptions regarding spatial

structure will lead to markedly different estimates of stock status and the response to management actions.

Spatial heterogeneity is often critically important to resource managers. In addition to the natural resource exhibiting spatial pattern, the actions that the manager can take, and those components of the system that the management actions affect can also be spatially structured: including localized distribution of fishing communities, and spatial structure in governance and management systems (e.g. management boundaries based on jurisdictional regimes). Obtaining information from ecological systems of management concern also presents an opportunity for spatial heterogeneity. The budgetary constraints of monitoring agencies often means monitoring data tend to be restricted in space. The challenges of relying on fishery-dependent data to provide indicators of system trends are well-documented (e.g. Rose and Kulka 1999), and provide examples of the importance of understanding spatial complexity, both in resource dynamics and in data generating processes.

This dissertation focuses on issues related to accounting for spatial heterogeneity when modelling and managing marine metapopulations, and how an understanding of important spatially-varying processes can contribute to understanding the behaviour of policy-relevant system properties. Specific goals, which form the focus of the individual chapters, include: (a) the development of tools to account for and identify spatially-varying population trends; (b) evaluation of the efficacy of monitoring designs to estimate population parameters and spatially-varying quantities of interest such as trends and abundance; (c) estimation of spatial characteristics of marine fish species distributions, (d) examination of how the spatial distribution of forage items, and behaviours related to spatial utilisation when foraging, enable a marine predator to obtain its energy budget; and (e) how spatial dynamics of both exploited populations and fishing fleets impact the performance of data-limited fisheries harvest strategies. Insight into the research areas of interest are achieved using mathematical modelling, enabling a

synthesis of current information and providing the necessary platforms for rigorous testing of assumptions regarding real world dynamics, and the hypotheses of interest.

Common to all but one of the chapters is the use of simulation modelling. Monte Carlo simulation methods are used frequently to evaluate analysis methods in terms of their ability to estimate various quantities of interest to management accurately and precisely (e.g. Kirkwood 1981, de la Mare 1986, Punt et al. 2002). The use of alternative 'operating models', which represent different assumptions about the dynamics of the 'real-world' can be a powerful tool in the analysis of management actions when faced with uncertainty. Simulation methods can also be used to test tools developed for estimating quantities of interest, because the dynamics governing the 'true' underlying system are known, allowing the performance of the estimator to be quantified. These methods have proved powerful for understanding the relationships and feedback mechanisms among the dynamics of natural systems, estimators, and management frameworks (Butterworth and Punt 1999, Smith et al. 1999, Bunnefeld et al. 2011). While the work undertaken for this dissertation has implications for a wider suite of marine systems, the research foci are the western Alaskan stock of Steller sea lions (*Eumetopias jubatus*), and the population of blue eye trevalla (*Hyperoglyphe antarctica*) exploited off Southeast Australia.

The western stock of Steller sea lions in Alaska declined by over 80% since the early 1970s, with many hypotheses existing regarding the true cause of the decline (Trites and Larkin 1996, Loughlin 1998, National Research Council 2003, Trites and Donnelly 2003), and recent indications suggesting some level of population increase in parts of the range (e.g. Fadely et al. 2006, Fritz and Gelatt 2011). The timing and extent of change in population numbers of the western stock has not been, and does not continue to be, geographically uniform, indicating that factors affecting population processes are expressed differentially through the range of the population. Results of population modelling studies (e.g. York et al. 1996, Holmes and York 2002, Fay and Punt 2006,

Winship and Trites 2006, Wolf et al. 2006) show that adequate fits to the available data for Steller sea lions can be obtained for a wide range of possible mechanisms regarding the causes of population change. Approaches based on the metapopulation concept aggregated the Steller sea lion population into a number of distinct regions, each composed of groups of rookeries that have exhibited similar historical trends. Studies examining diet of Steller sea lions also seem to corroborate the existence of some regional structure within the overall population (Merrick et al. 1997, Sinclair and Zeppelin 2002). However, exceptions and deviations from these regional patterns (in trend and in diet) do occur, and there appears to be general agreement among researchers that factors affecting population change in Steller sea lions are somewhat rookery-specific (e.g. National Research Council 2003, Call and Loughlin 2005, Fadely et al. 2005). Desirable properties of identifying and estimating population trends should capture spatial variability, allow for individual rookery behavior, and properly account for the processes leading to the available data on population abundance.

The blue eye trevalla is a high-valued species in Australia's Southern and Eastern Scalefish and Shark fishery (SESSF). The fishery for this long-lived, late-maturing species is characterized by a large number of gear types operating in a range of areas, with uncertainty in stock structure, apparent spatial and seasonal variability in availability of different age classes, and low levels of sampling effort across the fishery (Smith and Wayte 2002). Scientific advice for management in the SESSF takes the form of a Recommended Biological Catch (RBC) for each species under quota management (including blue eye trevalla) for the entire fishery to inform the setting of the Total Allowable Catch (TAC) (Smith et al. 2008). The SESSF has recently adopted a harvest strategy framework as a basis for setting RBCs. This framework is based on a set of harvest control rules (HCRs), with the decision as to which rule to apply for a particular stock being dependent on the type of information available on stock status. It is not clear how the tier framework of HCRs performs for species within the SESSF, and specifically

for the species managed using the data-poor HCRs, and how best to cope with possibly conflicting information from multiple areas and gear types.

Chapter 1 uses linear state-space modelling methods, based on the Kalman filter, to estimate time-series of abundance for a group of populations where the trend in abundance varies spatially, in addition to temporally. The models developed are applied to pup count data from the western stock of Steller sea lions (SSL) in Alaska, which consists of a number of breeding populations (rookeries) that have exhibited changes in population size that have not been geographically uniform. Identification of spatially-varying trends in abundance is of importance to managers of species of conservation interest (such as endangered Steller sea lions), not only for identifying populations for which additional management action may be necessary, but also to learn how these trends may relate to stressors on the populations that also may have a spatial component. State-space estimation using a Kalman filter is particularly suited to situations where there are 'missing' data, i.e. when observations are not available for every time step. The available data for Steller sea lions can be viewed as incomplete, with populations not being surveyed every year, and often survey coverage lacking within those years where data exist. The chapter develops two models, one that assumes a regional 'stock' structure, and one that relaxes this assumption, and instead allows for spatial correlation in individual population trends. The performance of these models (and their associated estimation procedures) are evaluated using a simulation testing framework, to address how the ability of the models to estimate trend, population size, and the parameters governing population change varies with data quantity, values for the parameters governing population dynamics, and the level of spatial correlation in trend among populations. As spatial connection among populations and stock structure is not generally known, this chapter also examines model performance when the population structure of the estimation models is mis-specified. The two models are also fitted to pup count data from 47 Steller sea lion rookeries in the western Alaska stock for the period 1978-2010,

resulting in estimates of the time-series of pup production, current trends, and the values for the parameters governing the extent of process and observation error. Estimates are compared among several configurations of the models, and the results interpreted with respect to the findings obtained during simulation testing.

Managers frequently have limited resources available for monitoring purposes. In the case of spatially-varying populations, careful choices must be (and are) made regarding the spatial and temporal allocation of monitoring effort. Chapter 2 builds on the analyses of Chapter 1 by evaluating tradeoffs associated with different spatial and temporal configurations of monitoring designs, in terms of estimating key parameters and population quantities. The simulation framework developed in Chapter 1 for pup production at Steller sea lion rookeries is used to determine how, given a fixed quantity of monitoring effort, the choice of monitoring design impacts model performance for various performance measures. The analyses then examine tradeoffs associated with optimising the monitoring design for the different performance measures, by determining how frequently the 'best' design for a particular performance measure will perform when estimating other quantities. Canonical Correspondence Analysis (CCA) is used to synthesise the multivariate nature of the simulation output, to determine how patterns in the spatial coverage of monitoring designs are associated with estimation success.

Prey resources of marine predators such as Steller sea lions tend to be spatially aggregated and patchily distributed. The spatial distribution of suitable prey to the foraging success of sea lions is important, as this distribution will combine with physiological and behavioural mechanisms to determine how much food is available to an individual sea lion. Aggregations of such resources (marine fish) are also the focus of commercial and recreational fisheries. There is a need to understand how spatial autocorrelation in data used for assessing trends and abundance of such species impacts these estimates. Chapter 3 uses Bayesian hierarchical modelling methods to fit spatial models of fish density to data from the Gulf of Alaska biennial/triennial bottom trawl

survey. Species are chosen for analysis based on their importance as prey items in the diet of Steller sea lions. The models (spatial Generalized Additive Mixed Effects Models, GAMMs) integrate uncertainty due to observation and process error, allow for differences in spatial clustering behaviour by species and among years, and account for the impact of bathymetry on prey density. The results of spatial models are compared to those obtained when no consideration is made for spatial autocorrelation, to determine how accounting for spatial autocorrelation impacts estimates of abundance and trends for the fish species analysed.

When foraging, marine predators must make decisions related to how to spatially allocate foraging effort. Spatial distribution of prey resources may be of particular importance to central-place foragers, such as Steller sea lions, that must return to terrestrial haul-out sites. Changes in the spatial nature of prey resources, such as those estimated in Chapter 3, provide additional complications for foraging sea lions in addition to the need to obtain sufficient food to meet requirements for growth and reproduction while minimising predation risk. Understanding how the spatial distribution of prey resources contributes to optimal foraging strategies for Steller sea lions requires a synthesis of information on sea lion bioenergetics, foraging behavior, and the nature of the spatial distribution of sea lion prey. Chapter 4 develops a simulation framework that couples models of foraging behaviour to a bioenergetics model of Steller sea lions. These models are used to evaluate likely consequences of foraging strategies and the distribution of prey on the forager's ability to meet the energy budget. The results of Chapter 3 are used to develop simulated prey fields with spatial characteristics that mimics those observed from fish survey data. A spatially-explicit foraging model is linked to a model of bioenergetics for Steller sea lions foraging from a central-place haul-out site. Parameter values for the models are obtained from the literature on studies of satellite telemetry, diet composition, and sea lion physiology. A suite of foraging strategies are considered, involving ways to make decisions that determine the observed behaviour during a trip, including: when to leave

the haul-out to begin a foraging trip, how long to forage before returning to the haul-out, and how to use available (and previously obtained) information to move to locate prey resources. The efficacy of various foraging strategies given changes in the spatial distribution of prey is evaluated by comparing the probability distributions of expected energy gain from a foraging trip given a set of assumptions for the simulation model. The simulations are used to determine what set of foraging strategies are successful given uncertainty in the prey distribution and the values for model parameters. Finally, summary statistics of the prey fields and foraging trips are calculated to evaluate whether emergent indicators of sea lion foraging behaviour and prey availability, such as those that might be collected from monitoring surveys and telemetry studies, can be used to identify successful foraging strategies, or distinguish among strategies, by accounting for the observed variability in energy gain.

The performance of fisheries harvest strategies is likely dependent on scales of spatial structure, both in dynamics of the resource, and in fishery governance structure. Harvest control rules (HCRs), key components of fisheries management strategies, are used to calculate recommended catch levels given estimates of present stock status or levels of fishing mortality. Spatial variability, either in the population dynamics, fishery operations, or in data collection, has the potential to impact HCR performance as this can drive variability in indicators used for stock assessment. Chapter 5 uses a Management Strategy Evaluation (MSE) approach to evaluate the performance of HCRs for blue eye trevalla, by assessing how well they perform at satisfying SESSF management targets and objectives. The HCRs rely solely on information from the age structure of the catch. Several versions of the HCRs are tested, varying in the reference points used to determine management actions, and in the way spatial variability is accounted for when setting catch limits. The performance of the harvest strategies is evaluated when there is mis-specification of the values for biological parameters, in addition to testing the implications of spatial uncertainty. Methods for dealing with spatial data when applying

the estimators include pooling the data (assuming no spatial structure) and appropriate weighting of spatially-disaggregated estimates of fishing mortality.

Chapter 1: Methods for estimating spatial trends in Steller sea lion pup production using the Kalman filter

1.1. Introduction

Spatial heterogeneity in the population trend of a species can reflect both differentiation resulting from population structuring (stock structure), as well as geographic differences in factors affecting population processes such as growth and survival (Lande et al. 1999, Walters and Martell 2004, Kritzer and Sale 2006). Frequently, the degree of spatial organization and the connectivity among units within the overall population are poorly known. However, for species that demonstrate fidelity to specific breeding sites, such as pinnipeds or seabirds, a degree of population structure can be identified. Spatial differentiation in trend for these individual breeding populations can be characterised by the degree of connectivity among populations, either through dispersal among subpopulations, or by spatial correlation in factors affecting recruitment, growth, and survival (e.g. Burgman et al. 1993, Hanski 1999). Identification of spatial differences in population trends may be hampered by stochasticity in the system dynamics among areas (process error), by imprecise measurements taken from the system (observation error), and by incomplete sampling regimes. Appropriate methods for estimating abundance and population trends should account for spatial differences in trends to provide useful information for species management and conservation.

State space modelling, a method of time-series analysis, has become increasingly popular when modelling population dynamics because it provides a means to directly account for both process and observation error when analysing time series data (e.g. Lindley 2003, Gudmundsson 2004, Holt and Peterman 2004, Walters 2004, de Valpine and Hilborn 2005, Ward et al. 2007). State space models operate by simultaneously considering two separate models, a process model that governs the dynamics and variability of the system

being estimated (the state), and an estimation model that details the relationship between the state, and the observations made of the state.

Many methods exist for the estimation of the parameters and properties of these models (e.g., de Valpine and Hastings 2002, de Valpine 2005). Bayesian estimation methods, requiring intensive numerical methods to obtain estimates of parameter values and associated uncertainty, have frequently been employed to integrate across the process error variation, particularly for non-linear models (Meyer 1999, Millar and Meyer 2000, Newman and Lindley 2006). In contrast, the Kalman filter provides an analytical means for estimating the state variables of a time-series model when the system dynamics are linear, the observations are linearly related to the state variables, and both process and observation errors are Gaussian. As with Bayesian analyses, the Kalman filter integrates over the state-space of population trajectories to calculate the likelihood of population dynamics parameters (de Valpine 2005), as opposed to more 'traditional' "errors-in-variables" methods, which maximise the joint-likelihood of the states and population parameters (Ludwig and Walters 1981, Richards et al. 1997). Parameter estimation using the Kalman filter has been applied to state-space formulations of a suite of commonly-used population models for fisheries and wildlife applications, including simple random-walks with drift, biomass dynamic-pool production, delay-difference, length- and age-structured models (Collie and Walters 1991, Sullivan 1992, Schnute 1994, Kimura et al. 1996), and analysis of catch-effort data (Reed and Simons 1996). These models have been applied to data for a diverse range of species, including California sea otters (*Enhydra lutris*), Yellowstone grizzly bears (*Ursus arctos horribilis*), Icelandic cod, Pacific cod (*Gadus macrocephalus*), lingcod (*Ophiodon elongatus*), yellowtail flounder (*Limanda ferruginea*), and eastern Bering Sea/Aleutian Island rockfish (Gudmundsson 1994, Lindley 2003, Spencer and Ianelli 2005).

State space estimation using a Kalman filter is particularly suited to situations where there are 'missing' data, that is when observations are not available for every time step

for which an estimate of state are needed, because process error is propagated within the system at all time steps, affecting the distribution (and hence the uncertainty) of the estimate of the state vector. Such a property is desirable, because 'missing' data is a common feature of datasets for natural populations. Spencer and Ianelli (2005) used a Kalman filter to estimate interactions within a multispecies fisheries complex, using a data set that frequently did not have survey biomass estimates for all species in every year considered.

Examples of process variation include environmentally-induced factors operating on survival, growth, and recruitment (Sullivan 1992). These factors are likely to display spatial correlation, resulting in spatially-correlated impacts of such factors on the population. Enabling spatial correlation in the state variables within a state-space modelling framework implicitly facilitates the incorporation of these factors. For example, Collie and Walters (1991) used a Kalman filter to predict stock sizes of six subpopulations of yellowtail flounder where parameters were shared among stocks, in an effort to use the information available among stocks to reduce uncertainty about stock size and productivity for each individual subpopulation, compared to approaches that modelled these subpopulations separately.

This chapter develops, tests, and applies state-space modelling methods based on the Kalman filter to estimate trends, abundance, and the parameters governing system dynamics for a metapopulation where there exist multiple observations over time at a number of discrete population units, which may display spatial correlation in population trend. The models are designed based on an application to the western Alaska stock of Steller sea lions (*Eumetopias jubatus*), which is comprised of a large number of individual breeding populations (rookeries), and for which trends in pup production at rookeries have been shown to exhibit spatial patterns over the geographical range of the stock (York et al. 1996, National Research Council 2003, Call and Loughlin 2005, Fay and Punt 2006). The models developed aim to accurately estimate both abundance and

trend, while recognising these spatial differences, and properly account for both process and observation error.

Two models are developed: Model #1 assumes that individual populations (rookeries) belong to regions, with region-specific time-varying trends, and allows for population-specific deviations from those trends. The second model (Model #2) differs from Model #1 in that a regional 'stock structure' is not assumed; rather populations in geographic proximity to each other are assumed to behave similarly. Spatial correlation in the individual population trends is assumed, again with the possibility for deviations from these correlated trends. The performance of the two models and their associated estimation procedures is assessed using simulated data under various assumptions regarding the magnitudes of observation and process error, the extent to which the dataset is 'complete' (missing counts at rookeries, numbers of years of data), and the degree to which the model used for estimation is mis-specified. The models developed are then fitted to pup count data from 47 Steller sea lion rookeries in Alaska for the period 1978-2007, to demonstrate their application, with results interpreted given the insights gained from the simulation analyses.

1.2. Methods

This section describes the modelling approaches and their associated estimation procedures, then provides the specifications for the simulations designed to evaluate model performance, and finally outlines how the approaches are applied to the pup count data for the western Alaska stock of Steller sea lions.

1.2.1 Models and Estimation methods

The Kalman filter is a recursive procedure for estimating the distribution of a vector of state variables (summarised by their means, and variance-covariance matrix) through the

model space (in this case, time). Appendix A details the equations required for application of the Kalman filter, with notation primarily following that of Harvey (1989). The Kalman filter makes a prediction of the state variables based on the model of the state-space process (which includes recognition of the impact of process errors) given a distribution for the initial state vector, and values for the parameters governing the state update and observation processes. The estimate of the state vector is then updated by comparing the prediction of the state vector with the available data at that time-step. This procedure is then applied throughout the time period for which data are available. As the Kalman filter provides estimates of the state based on all data observed up until and including the time the observation of the state is made, a backward-smoother can be applied to obtain estimates of the state variables based on all of the data, including those from future time-steps. Application of the Kalman filter also enables calculation of a likelihood function, which allows for estimation of the parameters of the model.

The Kalman filter requires that the model describing the process be written in state-space form. For this case-study, the natural logarithms of the number of pups at each rookery through time are modelled as linear functions of the log-abundance the previous year, and by region/rookery-specific time-dependent trends subject to annual process errors, with no time-dependency in the variance of these process errors. Model #1 considers an annual regional trend as part of the state space, with rookery-specific deviations from this trend. Model #2 assumes that the similarity in trends among rookeries within years is due to spatial covariance in the rookery-specific trends. The number of pups at a rookery is expected to be highly correlated through time as pup production in a given year is a function of both the pregnancy rate and the number of mature females in the population, which have a high survival rate. The annual trends and deviations from these trends considered in the models therefore integrate the biological complexities of pregnancy rates, and juvenile and adult survival, into a single process.

1.2.1.1. Process models

Model #1: region-specific trends

Model #1 considers two sources of process error: annual changes to a region-specific trend, assumed to be subject to a random walk, with the trend in a given year being determined by the trend in the previous year modified by an error term, and annual rookery-specific deviations from the regional trend:

$$\begin{aligned} \ln P^A_{r,y+1} &= \ln P^A_{r,y} + \beta^A_y + \eta^A_{r,y} & \eta^A_{r,y} &\sim N(0, \sigma^2_{r^A}) \\ \beta^A_{y+1} &= \beta^A_y + \eta^A_{\beta,y} & \eta^A_{\beta,y} &\sim N(0, \sigma^2_{\beta^A}) \end{aligned} \tag{1.1}$$

where $\ln P^A_{r,y}$ is the natural logarithm of the pup production in year y at rookery r, which is located in region *A*,

β^A_y is the trend for region *A* during year *y*,

$\eta^A_{r,y}$ is the deviation in trend for rookery *r* in region *A* during year *y*,

$\eta^A_{\beta,y}$ is the change in regional trend for region *A* during year *y*,

$\sigma^2_{r^A}$ is the variance of the rookery-specific process errors for region *A*, and

$\sigma^2_{\beta^A}$ is the variance of the annual changes in the regional trend for region *A*.

The process error variance terms, $\sigma^2_{r^A}$ and $\sigma^2_{\beta^A}$, are specified as being region-specific. However, it is also possible to assume that these parameters are common across regions, that is the same parameter values are used for each region.

Model #2: spatial correlation in trend among rookeries

Model #2 does not assume a regional structure, and instead allows for spatial correlation in rookery-specific trends, such that rookeries closer to each other are assumed to exhibit similar trends in pup production. The rookery-specific state dynamics are governed by the following equation:

$$\begin{aligned} \ln P_{r,y+1} &= \ln P_{r,y} + \beta_{r,y} + \eta_{r,y} \qquad & \eta_{r,y} &\sim N(0,\sigma_r^2) \\ \beta_{r,y+1} &= \beta_{r,y} + \gamma_{r,y} & \boldsymbol{\gamma}_y &\sim \mathbf{MVN}(0,\mathbf{W}) \end{aligned} \tag{1.2}$$

where $\beta_{r,y}$ is the trend for rookery *r* during year *y*,

$\eta_{r,y}$ is the deviation from the trend for rookery *r* during year *y*,

σ_r^2 is the variance of the rookery-specific process errors,

$\boldsymbol{\gamma}_y$ is the vector of deviations in the trend for all rookeries during year *y*,

$\mathbf{W}$ is the variance-covariance matrix for the change in trends at all rookeries during year *y*, with covariance between rookeries *i* and *j*:

$$W_{i,j} = \begin{cases} \sigma_\beta^2 + \psi^2 & \text{if } i = j \\ \sigma_\beta^2 \rho_{i,j} & \text{if } i \neq j \end{cases}$$

σ_β^2 is the maximum spatially-correlated variance of annual changes in trend,

ψ^2 is the non-spatial "nugget" variance of the annual changes in trend, and

$\rho_{i,j}$ is the spatial correlation between rookeries *i* and *j*.

This formulation for the process error variance allows for spatial dependence among rookeries in trend, with the degree of spatial correlation among rookeries depending on

the values for the $\rho_{i,j}$'s. The "nugget" variance, ψ^2, allows for a non-spatial component to the variance in the trends.

The spatial correlation in trend among rookeries is assumed to depend on the distance between rookeries, with the expectation that rookeries close to each other will have more similar trends than rookeries located further apart. The functional form relating spatial correlation with distance followed either an exponential, Gaussian, or spherical process:

$$\rho_{i,j} = \begin{cases} e^{\left(\frac{\ln(\varphi_{100})D_{ij}}{100}\right)} & \text{if exponential} \\ e^{-\left(\frac{D_{ij}\sqrt{-\ln(\varphi_{100})}}{100}\right)^2} & \text{if gaussian} \\ \begin{cases} 0 \text{ if } D_{ij} \geq (1/\varphi_s) \\ 1\text{-}1.5\varphi_s D_{ij} + 0.5\left(\varphi_s D_{ij}\right)^3 \text{ otherwise} \end{cases} & \text{if spherical} \end{cases} \qquad (1.3)$$

where D_{ij} is the distance (in km) between rookeries *i* and *j*,

φ_{100} is the extent of correlation at a distance of 100km, and

φ_s is the rate at which the correlation declines with distance for the spherical function.

Application of the Kalman filter requires that the state variables for the model be placed in vector form, with the state vector in year *y* for Model #1, being the $\ln P^A_{r,y}$'s for each rookery and the region-specific trends, β^A_y. The state vector for Model #2 differs from that for Model #1 by including the rookery-specific trends, $\beta_{r,y}$.

1.2.1.2. Observation model

The observation model describes the manner in which the available observations (counts of pups at rookeries) are related to the state variables. This is the same for both models, and is a linear function of the state variables with additive Gaussian observation error, i.e. for Model #1:

$$y_{r,y} = \ln P^{A}_{r,y} + \varepsilon_{r,y} \qquad \varepsilon_{r,y} \sim N(0, \sigma^2_{\varepsilon}) \tag{1.4}$$

where $y_{r,y}$ is the natural logarithm of the pup count at rookery *r* for year *y*,

$\varepsilon_{r,y}$ is the annual rookery-specific observation error term, and

σ^2_{ε} is the observation error variance.

Missing data can occur in two ways: (a) there could be years for which there are no observations, and (b) in any given year a subset of rookeries might not have been observed.

1.2.1.3. Parameter estimation

Recursive application of the Kalman filter and backward smoother (Appendix A) results in an estimated time series for the distribution of the state vector, given values for the model parameters governing the process and observation errors, and an initial distribution for the state ($\mathbf{a}_0$ and $\mathbf{P}_0$). The model is initiated 10 years prior to the year in which the first observation is available. The elements of the expected value for the initial state vector $\mathbf{a}_0$ are set to the natural logarithms of the first observations of pup production at each rookery, while the initial regional/rookery-specific trends are set to zero to reflect uncertainty in the numbers and the trend in this initial year. The initial variance-covariance matrix, $\mathbf{P}_0$, is set to an identity matrix multiplied by 100,000. Propagation of

the Kalman filter for 10 years before the first data point with this initial variance-covariance matrix allows for uncertainty in the co-covariance among rookeries and trends, and essentially interprets the elements of the initial state vector $\mathbf{a}_0$ as being highly uncertain, an appropriate description given that no observations have been made.

Given a data set, and model configuration, the negative of the logarithm of the likelihood function (Equation A.10) based on the application of the Kalman filter is minimised using the AD Model Builder (Fournier et al. 2012) software package, to obtain estimates for the model parameters, the time series of the region- or rookery-specific trends, and estimates of the pup production by rookery for each year of the modelled period.

1.2.2. Simulation testing

The simulation analyses were conducted over a 25-year period, and all considered a group of 15 populations (rookeries), with spatial/regional arrangement as per the eastern range of the western Alaska stock of Steller sea lions (Table 1.1, Figure 1.1). This resulted in a population with four regions. The full set of 47 rookeries (Table B.2) was not chosen for the simulations for computational reasons. One of either Model #1 or Model #2 was chosen to represent the true underlying dynamics, and, given a set of values for the model parameters and a monitoring scheme, were used to generate a time series of pup production at each of the 15 rookeries each year given the process error terms, from which a set of observations were made. One of the estimation models was then fit to these data, and the resulting estimates of the model parameters and derived quantities (annual regional/rookery trends and abundances) compared to the true values to assess the performance of the estimator.

Exploratory analyses used a range of values for the parameters of the operating model (results not shown). For ease of presentation, however, a single set of parameter values was chosen as a ‘base-case’ scenario for the simulations, with alternative scenarios aimed

at identifying model behaviour given departure from this base-case. Table 1.1 lists the values for the model parameters, and the various levels of data quality, used for the "true" model in the base-case and other scenarios. The degrees of data coverage range from a 'complete' dataset (counts at all rookeries in all years) to a subset of years for which data are available, with incomplete coverage within those years. The levels of coverage were chosen to mimic the extent of data coverage found in the actual pup count data for Steller sea lions in Alaska, both in the more data-intensive areas, and over the entire range.

Spatial connection among populations and stock structure is not generally known, and so a set of the simulations examined the impact of mis-specifying the estimation model being tested (Table 1.1). Two approaches were considered: 1) the true Model is #1, but Model #2 is used for estimation, and 2) both the true and estimation models are Model #1, but the allocation of populations to regions when applying the estimator is incorrect. Under the second approach, allocation of rookeries to regions in the estimation was achieved randomly, but with the correct number of regions, with each region containing at least two rookeries. Both approaches consider the case where a regional stock structure exists, but that structure is unknown, the first not assuming a stock structure, and the second assuming a stock structure that is wrong.

The performances of the estimation models were evaluated in terms of their ability to estimate the model parameters governing the extent of process/observation error, the annual regional/rookery-specific trends, the annual numbers by rookery, and the annual total numbers across rookeries/regions.

The ability to estimate the model parameters was assessed by calculating the error of estimate (EE) for each of the parameters for each simulation:

$$EE_{i,j} = E_{i,j} - T_{i,j} \tag{1.5}$$

where $E_{i,j}$ is the estimated value for parameter *j* from simulation *i*, and

$T_{i,j}$ is the true value of parameter *j* used in simulation *i*.

The relative error of estimate was not chosen, because the true values for the parameters were close to or equal to zero for some of the scenarios (Table 1.1).

The ability to estimate the annual numbers at individual rookeries was evaluated by calculating a time series of mean absolute relative errors (MARE) for each simulation:

$$MARE_{i,y} = mean_r \left(\frac{100 \left| P_{E,i,r,y} - P_{T,i,r,y} \right|}{P_{T,i,r,y}} \right) \tag{1.6}$$

where $P_{E,i,r,y}$ is the estimated number of pups at rookery *r* during year *y* for simulation *i*, and

$P_{T,i,r,y}$ is the true value for the number of pups at rookery *r* during year *y* for simulation *i*.

Calculating the mean absolute relative error across rookeries for each year informs how well the numbers are estimated at a given rookery in that year. That is, low values suggest that on average the numbers at individual rookeries are estimated very well. In this respect, the mean is preferable to the median, as a metric based on the median would not be sensitive to poor estimation of numbers at a small number of rookeries.

The estimation performance of the annual rookery/region-specific trends was determined in a similar manner to the rookery numbers, i.e. calculating a time series of the mean absolute errors (MAE) for each simulation:

$$MAE_{i,y} = mean_r \left[\left| \left(\beta_{E,i,r,y} - \beta_{T,i,r,y} \right) \right| \right] \tag{1.7}$$

As with the performance metrics for the parameters, the percent relative error of estimate was not chosen as many of the trends may be close to or actually zero.

The percent relative error for the estimate of the total number of pups across *all* rookeries by year was also calculated for each simulation:

$$\%REE_{i,y} = \frac{100\left(\sum_r P_{E,i,r,y} - \sum_r P_{T,i,r,y}\right)}{\sum_r P_{T,i,r,y}} \tag{1.8}$$

Unlike the MARE of the individual rookery numbers, this metric is not concerned with how well pup numbers are estimated at individual rookeries, just the total pup production across all rookeries.

The various performance metrics are summarised across simulations for each operating / estimating model combination and dataset coverage type.

1.2.3. Estimation of Steller sea lion rookery production

Several configurations of Models #1 and #2 were fit to pup count data from 47 rookeries in the western Alaska stock of Steller sea lions for the period 1978-2010 (Figure 1.1). Table B.2 summarises the information for each of six regions (Western Aleutians, Central Aleutians, Eastern Aleutians, Western Gulf, Central Gulf, Eastern Gulf) to which rookeries were allocated when applying Model #1. These regions represent portions of the Steller sea lion range within which individual rookeries have demonstrated similarity in population trend, and also reflect regional differences in diet diversity (York et al. 1996, Sinclair and Zeppelin 2002, Call and Loughlin 2005). Configurations of Model #1 (Table 1.2) varied the regional/global nature of the process error variance terms. Configurations of Model #2 (Table 1.3) involved different functional forms for the spatial correlation function, and whether or not to include the "nugget" variance term in the

estimation. Where possible, the relative change in Akaike's Information Criteria (AIC, Akaike 1974, Burnham and Anderson 2002) was used to compare among the different configurations for each of the two models. Comparison between the two models using AIC is not possible because the likelihood functions are different because the dimension of the state vector differs between the two models.

1.3. Results

1.3.1. Simulation testing

1.3.1.1. Model #1 performance

The distributions for the MAEs for the trends, MAREs for individual population abundances, and %REE for the total pup numbers in the final year obtained when Model #1 is used as both the operating and estimating model are shown in Figure 1.2 (a-c), for the relevant scenarios described in Table 1.1. Figure 1.2 also summarises the distribution of the EE for each of the model parameters (Fig. 1.2d-f).

1.3.1.1.1. "Trends and abundance"

In general, the annual trends (summarised by the MAE for the trend in the final year of the simulations) are estimated more poorly than the rookery abundances (summarised by the MARE for the number of pups in the final year; Fig. 1.2a and b). The ability to estimate annual trends declines (MAE increases) as the value for the standard deviation of the trend, σ_β, increases and the trends become more variable (Fig. 1.2a, "hi trend error"). However, the ability to estimate rookery numbers appears insensitive to the value for σ_β. Estimation of the trend and the numbers at each rookery deteriorates as the true values for both the observation and rookery-specific process error standard deviations, σ_ε and σ_r, increase (Fig. 1.2a and b, scenarios "hi obs err" and "hi rookerr").

Unsurprisingly, the distribution for %REE of total numbers becomes wider, yet the estimates remain unbiased, as the true value for the standard deviation of observation error, σ_ε, increases (Fig. 1.2c, scenario "hi obs err"). Variability in the error of the estimate of total numbers appears to be most sensitive to the true value for the observation error standard deviation, rather than that for the process error terms (Fig. 1.2c). This suggests that the model is able to track abundance by adjusting the rookery-specific process errors even when the annual trend is estimated incorrectly.

1.3.1.1.2. "Parameters"

Given a full data set, the standard deviation of the observation error, σ_ε is generally well-estimated (Fig. 1.2d). The precision of the estimates of the rookery error, σ_r appear to depend on the true value for both the extent of observation and rookery error, with the distribution of the EE widening considerably with increasing values of σ_ε (Fig. 1.2e, scenario "hi obs err").

Estimates of the standard deviation of the trend, σ_β, appear unbiased (Fig. 1.2f). The distribution of the EE for σ_β narrows when the true values for the other two variance terms are low (Fig. 1.2f, scenarios "lo obs err" and "lo rook err"). Estimates of σ_β become less precise as σ_ε increases, while the width of the distribution of the EE for σ_β also appears to depend on the true value for σ_r, with a wider distribution with increasing σ_r (Fig. 1.2f, scenario "hi rook err"). There appears to be little sensitivity in the estimate of the trend error standard deviation to the true value for the rookery-specific error at low values for the trend error standard deviation (results not shown).

These results suggest (perhaps unsurprisingly) that the model has difficulty distinguishing among the two types of process error, particularly given high values for the observation

error. This is reflected in Figure 1.2 by more imprecise estimates for the two process error standard deviations when the true value for the observation error standard deviation is high.

1.3.1.1.3. "Lower data quality"

A reduction in data coverage results in a poorer ability to estimate both trend and abundance (Fig. 1.2). For scenarios where there are 12 years of data for each rookery, whether these data are spread over 20 years ("20y 12ypr"), or occur in complete surveys during those 12 years ("12y 12ypr") appears to have little impact on the MARE of the rookery numbers in the final year of the simulation (Fig. 1.2b). However, spreading these data over 20 years instead of 12 unsurprisingly results in better estimates of the regional trends in the final year (Fig. 1.2a). Despite this decline in performance, the ability to estimate current total abundance is relatively unchanged between the base-case and the 12 years-per-rookery scenario (Fig. 1.2c). Reduction of the data coverage to just six years per rookery results in further degradation in the estimates of trends and abundance. In these cases, it is clearly more advantageous to spread the data over more years (*i.e.* have missing data within years), because the distribution of the MARE of rookery numbers in the final year is narrower for this scenario (Fig. 1.2b, compare "20y 6ypr" with "6y 6ypr"). Reducing the amount of data coverage also increases the extent to which the distribution of the %REE for total abundance in the final year widens as the true value for the process error standard deviations increase (results not shown).

The most pronounced differences in the estimates of trend and abundance between the data-poor and the base-case scenarios occur earlier in the time series (results not shown), with appreciable negative bias in total abundance when there are only six years of data per rookery. This negative bias remains (but is not as pronounced) in the final year estimates (Fig. 1.2c, "20y 6ypr" and "6y 6ypr").

Reducing the data coverage unsurprisingly impairs the ability to estimate the model parameters (reflected in markedly wider distributions for the EE; Fig. 1.2c-d). Estimates of the process error standard deviations (σ_r and σ_β) are negatively biased when there are only 6 years of data per rookery, and these data are obtained in the same years (Fig. 1.2e-f, "6y 6ypr"). A reduction in data coverage also increases the sensitivity of the estimation of the rookery process error standard deviation to the true value for the observation error standard deviation (results not shown).

1.3.1.2. Model #2 performance

The results obtained during simulations when the operating model and estimation model are both Model #2 are summarised in Figure 1.3 (scenarios as defined in Table 1.1). The presentation of results is the same as for Model #1, except that Figure 1.3 also contains summaries of the distributions for the EE of the spatial correlation parameter φ_{100}, and the nugget, ψ.

In general, the model parameters, rookery numbers, and total abundance are estimated less precisely for Model #2 than for Model #1 (Figs 1.2 and 1.3). This is unsurprising given that Model #2 has more parameters than Model #1.

3.1.2.1. "Trends and abundance"

On average, Model #2 estimates the annual trends more poorly than Model #1, reflected by higher values for the MAE of the trends in the final year (Figs 1.2a and 1.3a). However, the median absolute error of the trend (rather than the mean, as reported by MAE) is generally lower than that for Model #1 (results not shown). These differences are most likely a consequence of rookeries within regions offering multiple opportunities to estimate the same trend under Model #1, whereas the trends under Model #2 are rookery-specific. Changes in the ability to estimate the trend with scenario for Model #2

are much the same as described for Model #1 in Section 1.3.1.1.1 (Fig. 1.3a). There appears to be little sensitivity of the MAE of the trend to the true value for the extent of spatial correlation (Fig. 1.3a, scenarios "no phi", "lo phi", "hi phi"). The ability to estimate the trend improves (MAE decreases) from the base-case when there is no nugget (non-spatial) component to the trend, but one is estimated ("no nug" in Fig. 1.3a).

The results for rookery numbers and total abundance are also very similar to those for Model #1. That is, the MARE of the rookery numbers and the width of %REE distribution for total abundance are most sensitive to the true value of the observation error, σ_ε (Fig. 1.3b-c). In contrast to Model #1, the MARE of rookery-specific numbers for Model #2 appears to be insensitive to the value for σ_r (compare "lo rook err" and "hi rook err" scenarios in Figs 1.2b and 1.3b). There is little sensitivity of the ability to estimate rookery numbers and total abundance to alternative values for the spatial correlation parameter, φ_{100}.

3.1.2.2. "Parameters"

As for Model #1, the observation error standard deviation is generally estimated well (Fig. 1.3d). In general, the estimate of the rookery-specific process standard deviation, σ_r, is imprecise (Fig. 1.3e). The width of the distribution of the EE for σ_r is narrower than for the base-case with either low or high values for the spatial correlation parameter φ_{100} (Fig. 1.3e, compare base-case with scenarios "lo phi" and "hi phi"). This also occurs with low true values for the trend standard deviation σ_β, which did not occur for Model #1. The estimate for σ_r is negatively biased, but precise in the absence of observation error, concordant with positive bias in the estimate of the observation error standard deviation (Fig. 1.3d scenario "lo obs err").

Estimates of the trend standard deviation, σ_β, are positively biased, except when the true value for σ_β is high, there is no nugget, or when spatial correlation in trend among rookeries is low (Fig. 1.3f, scenario "hi trenderr", "no nug", "lo phi").

Estimates of φ_{100} tend to be negatively biased (Fig. 1.3g). The extent of bias appears to be sensitive to the true value of σ_β, with more bias when σ_β is low, but unbiased (and precise) when σ_β is high. This is unsurprising, because greater variability in the trend is likely easier to detect. Consequently, spatial correlation (or a lack thereof) among rookeries may be more evident given the larger effect. As for many of the parameters, estimates of φ_{100} display positive bias when the true value is low, and negative bias when the true value is higher.

The estimate for the nugget ψ, tends to be negatively biased (Fig. 1.3h), with improved estimation performance as the spatial trend process error standard deviation increases (i.e., as the ratio of ψ to σ_β decreases). The precision of the estimate of ψ is sensitive to the true values for all of the other parameters. The bias of the estimate of ψ is lower when there is no spatial correlation (φ_{100} is zero, Fig. 1.3h "no phi"). In this case, ψ is merely an additive term to σ_β, therefore estimation of the two parameters is confounded, corroborated by the unbiased (but imprecise) estimates of σ_β in the absence of spatial correlation in trend (Fig. 1.3f "no phi").

1.3.1.2.3. "Lower data quality"

As for Model #1, reducing the data coverage impacts the ability to estimate trend and abundance. However, the mean change in performance in terms of the ability to estimate the trend in the final year is greater for Model #2 than it was for Model #1 when the data

coverage is reduced (compare Figs 1.2a and 1.3a). Unlike for Model #1, there is clearly a difference between having missing data within years of data rather than a few 'complete' years when estimating rookery numbers, even when there are 12 years of data per rookery (compare the interquartile range for "20y 12ypr" and "12y 12pyr"in Figs 1.2b and 1.3b). The degradation in the MARE in rookery numbers for the final year with missing data compared with the base-case is greater for Model #2 than for Model #1, with estimates of zero for the rookery-specific process error term, σ_r , in many cases for the most data-poor scenarios (Fig. 1.3, "20y 6ypr" and "6y 6ypr").

1.3.1.3. Impact of model mis-specification

Figure 1.4 summarises the impact of model mis-specification on the ability to estimate trends and abundance in the final year, given a full dataset, when the true dynamics are represented by Model #1. Fitting Model #1, but assuming an incorrect regional structure (Fig. 1.4, scenario "1") unsurprisingly results in poorer estimates of the trend and abundance than when Model #1 is based on the correct regional structure (Fig. 1.4, "base"). However, the degradation in performance in terms of estimation of rookery numbers and abundance is quite small (Fig. 1.4b). In terms of estimating both trend and abundance, better results are obtained by fitting Model #2 (Fig. 1.4, scenario "2") than assuming a regional structure that turns out to be incorrect. Even with only 6 years of data per rookery, trend estimation using Model #2 is on average still better than that obtained using Model #1 with an incorrect regional structure (compare scenarios "1" and "2, 6 ypr" in Fig. 1.4a). There is little difference in the MARE of rookery numbers and %REE of total pup production in the final year obtained by either fitting Model #2 or fitting Model #1 with an incorrect regional structure (Fig. 1.4b-c). However, as noted when the estimating model was not mis-specified, individual rookery numbers (and pup production) are poorly estimated when there are missing data (Fig. 1.4b-c, scenario "2, 6 ypr").

Unsurprisingly, parameter estimation is poorer when the estimation model is mis-specified; assuming an incorrect regional structure results in biased estimates of all of the model parameters, with large positive bias in the estimate of the rookery-specific process error standard deviation, σ_r (results not shown). This is presumably a result of having to account for an assumed average trend structure that is in conflict with the data. Fig. 1.5 shows the distribution for the estimate of φ_{100} across all the simulations when Model #2 is fitted to the data generated by Model #1. Despite the true regional structure being the same for all simulations, the distribution of estimates for this parameter is fairly broad, with values ranging from a correlation at 100km of 0.4 to almost 1 (Fig. 1.5). This suggests that Model #2 may be accounting for the regional structure not just with this parameter. However it is clear that low values for the correlation parameter are unsupported by data coming from populations with the given regional structure.

1.3.2. Application to Steller sea lions in Alaska

Tables 1.2 and 1.4 show the results of using several configurations of Model #1 that alter the degree to which the process error variances are global or regional to estimate pup production for Steller sea lions. Table 1.2 lists the maximum likelihood estimates for the standard deviations of the process and observation error terms, along with the ΔAIC value for each model configuration, while Table 1.4 lists the estimates of pup production in 2010 by region, and the estimates of the 2010 regional trend.

The complementary results from the application of Model #2 are shown in Tables 1.3 and 1.5, with different configurations here representing the choice of form for the spatial correlation function (exponential, Gaussian, or spherical) and whether or not to include the nugget in the trend variance. Note that the abundance estimates for Model #2 are reported by region for comparison with the results in Table 1.4, even though a regional

stock structure does not apply for this model. Figure 1.6 summarises the estimates of the trend by rookery in 2010 for the different configurations of Model #2.

The 'best' estimates obtained by the different configurations of both models provided fairly similar results in terms of current pup production (Tables 1.4 and 1.5). The Model #2 configurations with differing spatial correlation functions all resulted in nearly identical 2010 pup production estimates for all regions (Table 1.5). Model #1 configurations including region-specific σ_β led to lower estimates for 2010 pup production in the Western GoA, Eastern and Central Aleutians, while the configuration that assumed a global trend across the entire range resulted in lower estimates for the Western GoA, and higher estimates for the Central and Western Aleutians than other configurations. These differences are driven by estimation of negative current trends in the Aleutian regions and a positive trend in the Western GoA when the trends were allowed to vary by region, in contrast to a weakly negative trend when estimation is global (Table 1.4).

Parameter estimates were generally consistent among the different configurations (Tables 1.2 and 1.3). Results of the configurations of Model #1 that allowed the process error standard deviation terms to vary by region suggested that pup production for some regions are more variable than for others, reflected in estimates of σ_r ranging from 0.11 to 0.25 (for the Central Aleutians), as opposed to a value of 0.19 when σ_r was assumed to be the same for all regions. Similarly, configurations where the trend error standard deviation varied by region also indicated higher variability in the Central Aleutians region. The configuration of Model #1 suggested by AIC as the most parsimonious is based on region-specific rookery and trend process error standard deviations ("regional σ_r & σ_β" in Table 1.2). However, the evidence favouring this configuration over one with just the regional σ_r was not very strong (ΔAIC of 2.8 for the "regional σ_r", Table

1.2). In general, the configurations of Model #2 suggested more variable trends than Model #1 (trend error standard deviation of around 0.08 compared to 0.05). Annual changes in trend were estimated as being highly correlated in space (φ_{100}>0.95), suggesting a high degree of synchrony among populations in the trend changes. Estimates of trend and abundance were almost identical among the different configurations for the spatial correlation function (Table 1.3). This is reflected in the ΔAIC values, all of which were ≤2 (Table 1.3), suggesting little evidence for choosing among the functional form for the spatial correlation.

Estimates of current (2010) trend from both models are similar, indicating a "best estimate" increase in pup production throughout most of the range, with trends around zero for the Central and Eastern Gulf of Alaska (Model #1), and a decreasing trend in the Western Aleutians. The estimation intervals for these results are however, quite broad (Table 1.4; Fig. 1.6).

1.4. Discussion

1.4.1. The Kalman filter

The Kalman filter provides a more complete way to account for errors typically encountered in the observation of fisheries and wildlife resources than observation-error only methods (Spencer and Ianelli 2005). Advantages of the state-space modelling approach, and specifically the use of the Kalman filter, are that estimates of uncertainty for parameters and derived quantities of interest are based on the likelihood function for the data rather than on large-sample approximations (Lindley 2003), and the simplicity associated with accounting for missing observations (Harvey 1989). The production of a complete estimation covariance matrix for the state vector during application ensures that the cross-covariance structure for the components of the state vector is fully and easily specified, facilitating incorporation in further analyses. However, performance of the

Kalman filter is known to degrade when data are missing (Kimura et al. 1996, Holt and Peterman 2004, Sinopoli et al. 2004, Royer et al. 2005), and was confirmed here. However, bias in the estimates of total abundance was only demonstrated in the most extreme data-poor scenarios.

The simulation results described above suggest that the estimates of process error variance can be substantially biased. However, the low sensitivity of the abundance estimates to the values for the process error variance (Figs 1.2 and 1.3) suggests that the models are able to accommodate this bias, perhaps due to the inclusion of the two process error terms. Nevertheless, biases in estimates of process error variance can have important implications when making predictions about future population abundance and/or trends, common applications for population modelling exercises (e.g. Population Viability Analysis, PVA). However, Lindley (2003) found that bias in estimates of process error variance was reduced using the Kalman filter compared to estimates obtained from alternative estimators described by Holmes (2001) and Dennis et al. (1991).

Biomass estimates from a delay-difference production model using a Kalman filter were found to be positively biased when there is process error (Kimura et al. 1996). In contrast, the estimates of total abundance were in general unbiased in this study, and most sensitive to the value for observation error variance rather than the process error variance.

Separating process and observation error variances remains problematic using the Kalman filter, reflected in the sensitivity of the EE for the process error variance terms to the value for the observation error variance (Figs 1.2 and 1.3), although the observation error variance term itself was generally well estimated. De Valpine (2005) notes an increase in estimation performance of variance terms for the Kalman filter when the ratio among observation and process variances is known. In principle, multiple observations of individual populations at given time steps should provide additional information on

observation error. These data could perhaps be made available for Steller sea lions, as the reported counts are often the mean of several (2 or 3) counts by different observers at the same time (Fritz et al. 2008).

1.4.2. Should it be used (what is being done now)?

The application here demonstrates the suitability of the Kalman filter for estimation in the presence of both process and observation error, when the modelled system is linear, error terms are Gaussian and independent, and observations are linearly related to the state vector. However, Kimura et al. (1996) found that the Kalman filter performed well even when the assumption of independent process error vectors was violated. There are many alternative approaches to parameter estimation which generalise the Kalman filter to non-linear, non-normal error state-space models for biological time series, including the extended Kalman filter, the ensemble Kalman filter, non-linear state-space modelling, and Bayesian techniques such as Markov chain Monte Carlo (MCMC, Kitagawa 1981, Gronnevik and Evensen 2001, Calder et al. 2003, de Valpine 2005). Nevertheless, the linear model with random walk has been shown to perform well at tracking the dynamics of non-linear systems (Peterman et al. 2000). Indeed, Walters (2004) notes that the errors in biomass estimates associated with complex age-structured stock assessment models often mimic those computed by simple Kalman filter equations. For situations when age and length data may be relatively uninformative, and parameter estimation for such models relies heavily on fits to a time series of abundance (such as commercial catch per unit effort, CPUE) in which little confidence is placed, estimation of exploitable population trend using a Kalman filter is likely comparable to what might be expected from complex methods. Furthermore, the simple state-space model provides rapid integrated estimates of uncertainty that incorporate both process and observation error, features that are often computationally and time-intensive for more complex methods.

In retrospective forecasting analyses of sockeye salmon populations, Holt and Peterman (2004) found that Kalman filter models performed better than standard sibling models in being able to predict trends. As in the present study, Holt and Peterman (2004) noted high correlations in the values for annual trends among stocks, although they did not explicitly take advantage of these correlations during the estimation process. As with the results presented here, they also noted several stocks which exhibited behaviour that departed from the high correlations, suggesting the need to account for uniqueness of sites/stocks when modelling trends using simple models. The inclusion of an additional process error term here (the rookery-specific process error term, σ_r) incorporates the possibility for departure of a particular population from the expected trend. While not presented here, models that did not include the rookery-specific process errors did not fit the pup count data for the Alaskan Steller sea lions well.

Current assessment of trends in Seller sea lion abundance are generally based on either comparing the relative change in raw counts, or derived from log-linear regressions of count data, for sites consistently censussed over the time period considered (e.g. Fritz et al. 2008). The modelling methods presented here improve over these techniques by including all of the data and predicting trends and abundance at all sites, despite missing data, and by rapid calculation of uncertainty estimates that incorporate integrated impacts of both the observation process and stochasticity in population dynamics. The inclusion of each rookery as individual components of the state vector means that (possibly arbitrary) regional aggregation of data is unnecessary to make predictions, with recognition of spatial structure within the larger sea lion metapopulation accommodated by explicitly modelling spatial correlation in trends among rookeries (Model #2).

4.3. Recommendations for Steller sea lions (models and sampling schemes)

It is clearly advantageous to use Model #2 for estimation when the underlying stock structure is unknown (as is usually the case), because the simulations indicate that model

performance is better if correlation among subpopulations is inferred rather than when a rigid (and possibly incorrect) stock structure is assigned. However, the simulations also suggest that Model #2 is more prone to error and bias in estimates of abundance when there are missing data than is Model #1.

Fitting the models to the pup count data for Steller sea lions results in estimates of observation error variances with CVs of approximately 20-23%. No attempt was made to allow the observation error to vary through time and/or space, although the pup count data have been collected using a variety of methods over the length of the time series. Estimation results for Model #1 suggest that there are regional differences in variability, with the most parsimonious model attributing these differences to rookery-specific process errors in addition to regional differences in annual trend variability. The observation error standard deviation estimates were correlated with the estimates for the rookery process error standard deviation in the simulations, and so it is unclear as to whether the apparent differences in variability are due to the spatial/temporal variability in sampling. This suggests that it would be beneficial to obtain multiple counts for some rookeries in observation years, so as to better estimate the observation error. Recent surveys of pup production are more comprehensive than those earlier in the time series, with 96% and 89% coverage of all rookeries in 2005 and 2009 respectively, although the 2010 data only constitute 49% coverage. In general, there are geographic disparities in the levels of coverage, with rookeries in the Gulf of Alaska and Eastern Aleutians regions tending to have been counted more frequently than those in the Central and Western Aleutian Islands. This is unsurprising given the costs associated with assessing sea lion populations in these more remote areas. Addressing how the spatiotemporal allocation of data points influences the estimation of model parameters, trends, and abundance is a clear avenue for future work.

1.4.4. Status of Steller sea lions

The 2010 stock assessment for the western Alaskan population of Steller sea lions reports an estimate of pup production of 11,120 based on a 2005-2009 composite count (Allen and Angliss 2010). This number is slightly higher than the mean estimates for 2009 pup production found here for Model #1 (10,484) and Model #2 (10,614), for the model configurations resulting in the lowest AIC. However, the stock assessment estimate is within the 95% intervals for total 2009 pup production resulting from Model #2 (9,956-11,332) and only just outside the interval for Model #1 (9,889-11,118).

Reported estimates of current trends in pup production for the western Alaska population of Steller sea lions are the percent change by region from 2005 to 2009 (Fritz and Gelatt 2011). These estimates indicate appreciable increases in the Gulf of Alaska and Eastern Aleutians regions, and negative trends in the Central and Western Aleutians (Fritz and Gelatt 2011). Table 1.6 compares these results with the pup production trends obtained from the models presented here for the period 2005 to 2009, both for the trend rookeries used by Fritz and Gelatt (2011), and for the entire suite of rookeries from which data were fit to. The model results were somewhat consistent with the differences in raw counts for some areas. When comparing the same rookeries as in Fritz and Gelatt, both Models #1 and #2 were less optimistic regarding pup production trend for the Central Gulf of Alaska (Model #1 -5%, Model #2 -3%), and the Central Aleutians (-9% for both models). The models also did not predict as large an increase in pup production for the Eastern Gulf of Alaska as that cited by Fritz and Gelatt (2011). Both models suggested large declines in pup production in the Western Aleutians over the 2005-2009 period, with greater estimated decreases (Model #1 -34%, Model #2 -38%) than Fritz and Gelatt (2011). Consequently, the results from Model #2 suggested an overall increase of 5-6% in pup production over the entire western stock during this period, with Model #1 suggesting increases of 7-8%, compared to the increase of 10% reported by Fritz and Gelatt (2011) (Table 1.6). However, apart from differences in the Central Gulf of Alaska

and Western Aleutians, the changes in the raw data reported by Fritz and Gelatt (2011) are within the 95% confidence intervals for the changes in pup production from both models (Table 1.6).

The model frameworks presented here provide rapid appraisal of the status of pup production in Steller sea lions. While the simple linear models here do not explicitly account for population processes, positive trends in pup production are correlated with increased population size, higher survival rates, and increased breeding success, all of which correspond to either a larger or more productive mature sea lion population. Spatial correlation in the estimates of pup production trends are consistent with regional differences highlighted by other authors. The high level of spatial correlation in trend among rookeries obtained from model estimates is consistent with larger-scale regionality of population trend within the Steller sea lion metapopulation.

1.4.5. Conclusion

Although applied to Steller sea lions here, the modelling frameworks developed could be easily applied to any situation where multiple populations of a species are expected to exhibit some level of spatial correlation in population trend. The models are particularly suited to cases where more detailed population modelling is precluded by a lack of information other than census or count data, and where there are missing data for some sites/populations.

Table 1.1. Specifications for the operating models used for the simulations.

Scenario	operating model	estimating model	True values for model parameters σ_ε	σ_r	σ_β	φ_{100}	ψ	# years with data	# years data per rookery
Model #1									
base-case	1	1	0.2	0.1	0.05	-	-	25	25
no obs error	1	1	0.001	0.1	0.05	-	-	25	25
high obs error	1	1	0.5	0.1	0.05	-	-	25	25
no rookery process error	1	1	0.2	0.001	0.05	-	-	25	25
high rookery process error	1	1	0.2	0.2	0.05	-	-	25	25
low trend process error	1	1	0.2	0.1	0.01	-	-	25	25
high trend process error	1	1	0.2	0.1	0.1	-	-	25	25
20 years, 12 years per rooker	1	1	0.2	0.1	0.05	-	-	20	12
12 y, 12 ypr	1	1	0.2	0.1	0.05	-	-	12	12
20 y, 6 ypr	1	1	0.2	0.1	0.05	-	-	20	6
6 y, 6 ypr	1	1	0.2	0.1	0.05	-	-	6	6
Model #2									
base-case	2	2	0.2	0.1	0.05	0.5	0.01	25	25
no obs error	2	2	0.001	0.1	0.05	0.5	0.01	25	25
high obs error	2	2	0.5	0.1	0.05	0.5	0.01	25	25
no rookery process error	2	2	0.2	0.001	0.05	0.5	0.01	25	25
high rookery process error	2	2	0.2	0.2	0.05	0.5	0.01	25	25
low trend process error	2	2	0.2	0.1	0.01	0.5	0.01	25	25
high trend process error	2	2	0.2	0.1	0.1	0.5	0.01	25	25
20 y, 12 ypr	2	2	0.2	0.1	0.05	0.5	0.01	20	12
12 y, 12 ypr	2	2	0.2	0.1	0.05	0.5	0.01	12	12
20 y, 6 ypr	2	2	0.2	0.1	0.05	0.5	0.01	20	6
6 y, 6 ypr	2	2	0.2	0.1	0.05	0.5	0.01	6	6
no nugget	2	2	0.2	0.1	0.05	0.5	0	25	25
no φ_{100}	2	2	0.2	0.1	0.05	0.001	0.01	25	25
low φ_{100}	2	2	0.2	0.1	0.05	0.1	0.01	25	25
high φ_{100}	2	2	0.2	0.1	0.05	0.9	0.01	25	25
low φ_{100}, 12ypr	2	2	0.2	0.1	0.05	0.1	0.01	20	12
high φ_{100}, 12ypr	2	2	0.2	0.1	0.05	0.9	0.01	20	12
incorrect regional allocation	1	1[a]	0.2	0.1	0.05	-	-	25	25
incorrect model	1	2	0.2	0.1	0.05	-	-	25	25
incorrect model, 6 ypr	1	2	0.2	0.1	0.05	-	-	20	6

[a] allocation of rookeries to regions is incorrect.

Table 1.2. Parameter estimates and ΔAIC values for various configurations of Model #1 fitted to the Steller sea lion pup count data. Regional abbreviations are of the form: EGOA - Eastern Gulf of Alaska, CAI - Central Aleutian Islands.

Model #1 configuration	# σ_r's	# σ_β's	# regions	σ_ε	σ_r						σ_β						# estimated parameters	ΔAIC
					EGOA	CGOA	WGOA	EAI	CAI	WAI	EGOA	CGOA	WGOA	EAI	CAI	WAI		
base-case	1	1	6	0.21	0.194	0.194	0.194	0.194	0.194	0.194	0.050	0.050	0.050	0.050	0.050	0.050	3	12.2
regional σ_r	6	1	6	0.221	0.180	0.105	0.169	0.147	0.253	0.123	0.043	0.043	0.043	0.043	0.043	0.043	8	2.8
regional σ_β	1	6	6	0.205	0.193	0.193	0.193	0.193	0.193	0.193	0.032	0.039	0.033	0.019	0.133	0.000	8	5.8
regional σ_r & σ_β	6	6	6	0.218	0.184	0.108	0.171	0.148	0.239	0.127	0.033	0.042	0.033	0.020	0.121	0.000	13	0.0
single global trend	1	1	1	0.189	0.233	0.233	0.233	0.233	0.233	0.233	0.057	0.057	0.057	0.057	0.057	0.057	3	NA

Table 1.3. Parameter estimates and ΔAIC values for configurations of Model #2 fitted to the Steller sea lion pup count data.

spatial correlation form	estimate ψ?	σ_ε	σ_r	σ_β	φ_{100}	ψ	ΔAIC
exponential	Yes	0.233	0.124	0.076	0.985	3.80E-05	2.00
exponential	No	0.231	0.129	0.076	0.985	NA	0.00
gaussian	Yes	0.232	0.126	0.076	0.998	1.00E-04	1.80
gaussian	No	0.224	0.146	0.076	0.998	NA	0.60
spherical	Yes	0.233	0.124	0.076	0.986	4.10E-05	2.00
spherical	No	0.231	0.13	0.077	0.986	NA	0.20

Table 1.4. Estimated pup production and trend in 2010 by region (mean with 95% CI in parentheses), for the different configurations of Model #1 fitted to pup count data for Steller sea lions.

	estimated regional pup production in 2010 (mean & 95% CI)											
Configuration	Eastern GoA		Central GoA		Western GoA		Eastern Aleutians		Central Aleutians		Western Aleutians	
base	961	(688, 1342)	1475	(1043, 2114)	2756	(1566, 4850)	2620	(1733, 4010)	2563	(1810, 3643)	222	(94, 538)
regional σ_r	961	(682, 1354)	1531	(1127, 2086)	2790	(1638, 4753)	2584	(1759, 3820)	2574	(1750, 3813)	220	(111, 441)
regional σ_β	955	(690, 1321)	1479	(1054, 2104)	2750	(1594, 4746)	2608	(1748, 3933)	2610	(1846, 3708)	213	(107, 434)
regional σ_r & σ_β	956	(682, 1342)	1528	(1125, 2081)	2781	(1645, 4700)	2578	(1768, 3782)	2602	(1778, 3836)	208	(124, 352)
global trend	929	(679, 1269)	1465	(1050, 2088)	2345	(1310, 4197)	2568	(1704, 3936)	2630	(1878, 3707)	304	(134, 703)

	estimated 2010 trend (mean & 95% CI)											
	Eastern GoA		Central GoA		Western GoA		Eastern Aleutians		Central Aleutians		Western Aleutians	
base	0.051	(-0.132, 0.235)	-0.044	(-0.208, 0.119)	0.084	(-0.103, 0.27)	0.023	(-0.143, 0.189)	-0.038	(-0.192, 0.116)	-0.102	(-0.339, 0.134)
regional σ_r	0.050	(-0.112, 0.212)	-0.038	(-0.173, 0.098)	0.087	(-0.074, 0.248)	0.025	(-0.117, 0.167)	-0.046	(-0.188, 0.097)	-0.097	(-0.294, 0.1)
regional σ_β	0.043	(-0.094, 0.179)	-0.034	(-0.17, 0.101)	0.086	(-0.047, 0.218)	0.018	(-0.066, 0.103)	0.009	(-0.309, 0.327)	-0.117	(-0.16, -0.073)
regional σ_r & σ_β	0.044	(-0.094, 0.182)	-0.038	(-0.173, 0.097)	0.088	(-0.043, 0.219)	0.023	(-0.06, 0.105)	-0.008	(-0.313, 0.297)	-0.118	(-0.148, -0.088)
global trend	-0.02 (-0.167, 0.135)											

Table 1.5. Estimated pup production (mean and 95% CI) in 2010 by geographic region, for the different configurations of Model #2 fitted to pup count data for Steller sea lions.

		estimated regional pup production in 2010 (mean & 95% CI)											
spatial correlation form	estimate nugget?	Eastern GoA		Central GoA		Western GoA		Eastern Aleutians		Central Aleutians		Western Aleutians	
exponential	Yes	954	(675, 1346)	1547	(1073, 2305)	2510	(1493, 4223)	2637	(1757, 3983)	2563	(1795, 3669)	197	(97, 408)
	No	950	(674, 1339)	1544	(1072, 2298)	2510	(1493, 4221)	2637	(1759, 3981)	2564	(1798, 3664)	198	(96, 411)
gaussian	Yes	959	(679, 1354)	1544	(1070, 2306)	2515	(1499, 4218)	2645	(1765, 3992)	2576	(1805, 3686)	195	(96, 403)
	No	954	(680, 1339)	1525	(1063, 2278)	2528	(1508, 4240)	2670	(1784, 4023)	2572	(1812, 3660)	196	(95, 411)
spherical	Yes	954	(676, 1347)	1547	(1073, 2305)	2508	(1492, 4219)	2636	(1757, 3981)	2564	(1795, 3669)	197	(96, 408)
	No	951	(675, 1340)	1543	(1072, 2298)	2508	(1492, 4217)	2636	(1759, 3979)	2564	(1799, 3664)	198	(96, 411)

Table 1.6. Percentage change in pup production (mean and 95% CI) by geographical region from 2005 to 2009. Numbers taken from Fritz and Gelatt (2011) are for counts from 2005, and 2009 (see Table 6, Fritz and Gelatt 2011). Results for the 'best' (lowest ΔAIC) configurations of Models #1 and #2 are reported using the trend rookeries from Fritz and Gelatt (2011) (for comparison), and also for all rookeries considered. 'Total' column refers to the change in pup production for the entire western stock of Steller sea lions. Regional abbreviations are of the form: EGOA - Eastern Gulf of Alaska, CAI - Central Aleutian Islands.

	Region						
	EGOA	CGOA	WGOA	EAI	CAI	WAI	Total
Fritz and Gelatt (2011)	28	10	21	20	-5	-19	10
Fritz and Gelatt trend rookeries							
Model #1	22	-5	28	21	-9	-34	7
	(11.1,34.8)	(-11.1,2.3)	(13.5,44.4)	(9.8,33.4)	(-19.4,3.1)	(-43.0,-33.6)	(-4.0,18.7)
Model #2	21	-3	20	21	-9	-38	5
	(11.2,32.6)	(-11.5,5.5)	(6.4,35.7)	(8.9,34.1)	(-17.4,-0.3)	(-49.8,-22.3)	(-5.2,16.8)
all rookeries							
Model #1	23	-5	31	20	-8	-34	8
	(11.4,35.3)	(-11.5,2.5)	(16.0,47.5)	(8.7,31.8)	(-19.1,3.5)	(-43.0,-23.8)	(-3.3,19.7)
Model #2	21	-3	23	19	-8	-38	6
	(10.9,32.6)	(-12.5,8.8)	(9.0,39.4)	(7.8,32.4)	(-16.1,1.5)	(-49.8,-22.3)	(-4.3,18.7)

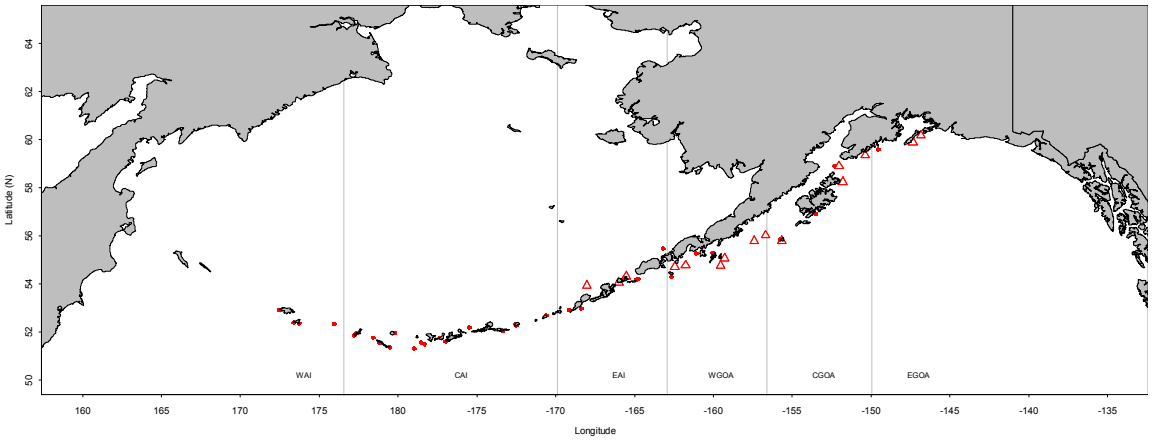


Figure 1.1. Locations of Steller sea lion rookeries and regions used for the analyses. Rookeries used for the bases of the simulation analyses are denoted by triangles.

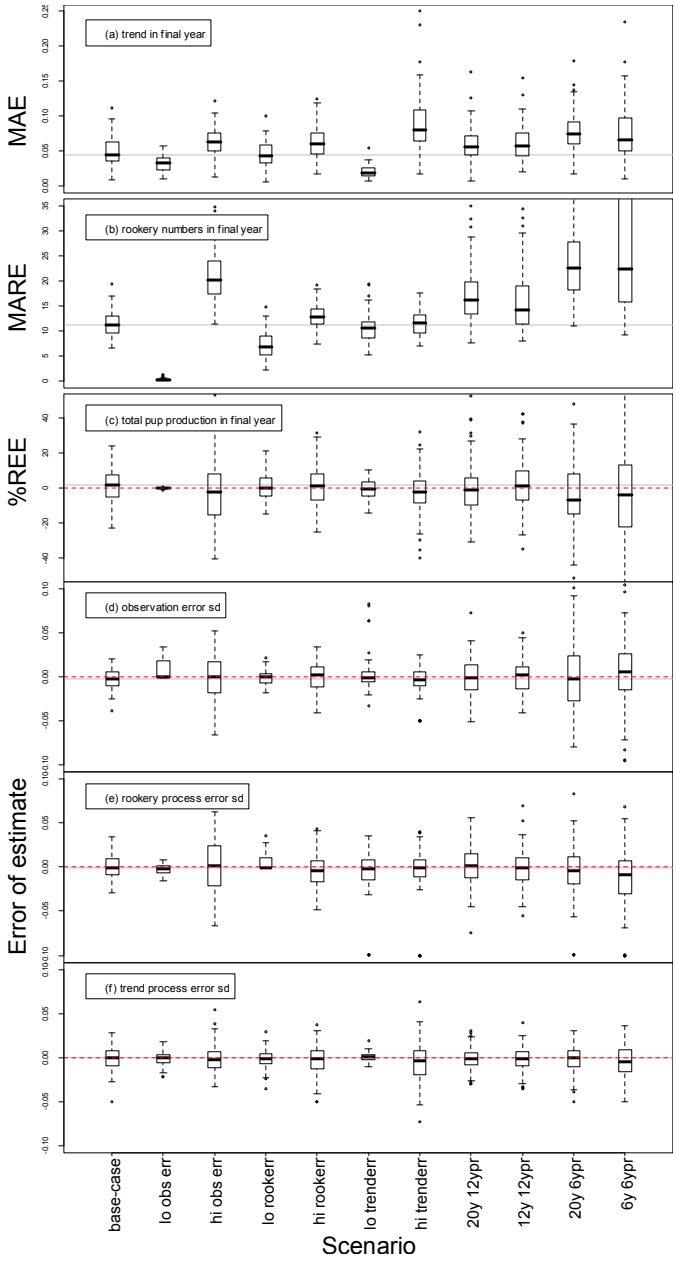


Figure 1.2. Distributions of the MAE of the trend, the MARE of rookery numbers, the %REE of total pup production in the final year, and the Error of Estimate (EE) for model parameters. Results are for simulations where the operating and estimating models are both Model #1. Red dotted lines indicate zero for the %REE and EE plots, and the grey horizontal line corresponds to the median value for the performance statistic under the base-case scenario.

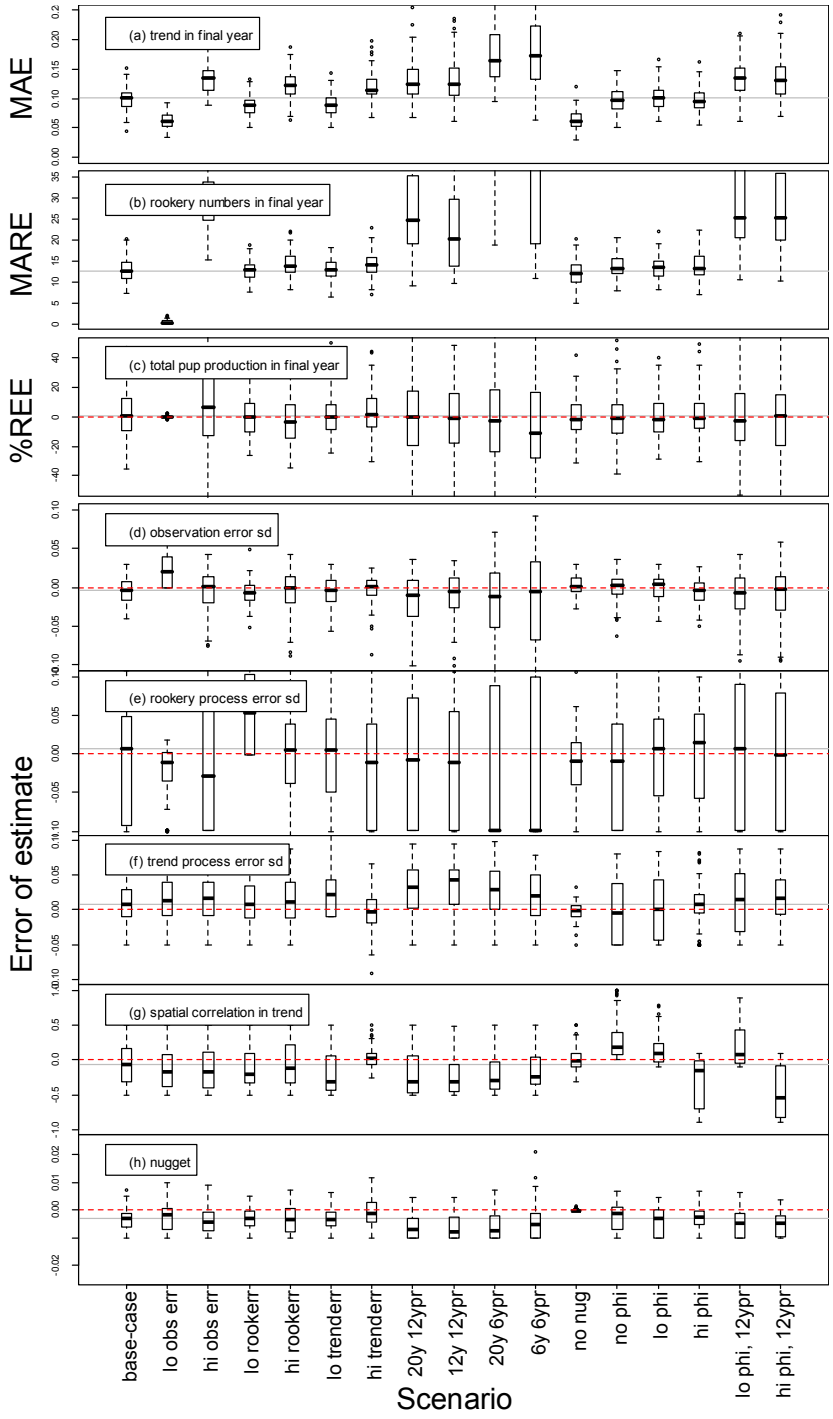


Figure 1.3. As for Figure 1.2, except that the results pertain to Model #2.

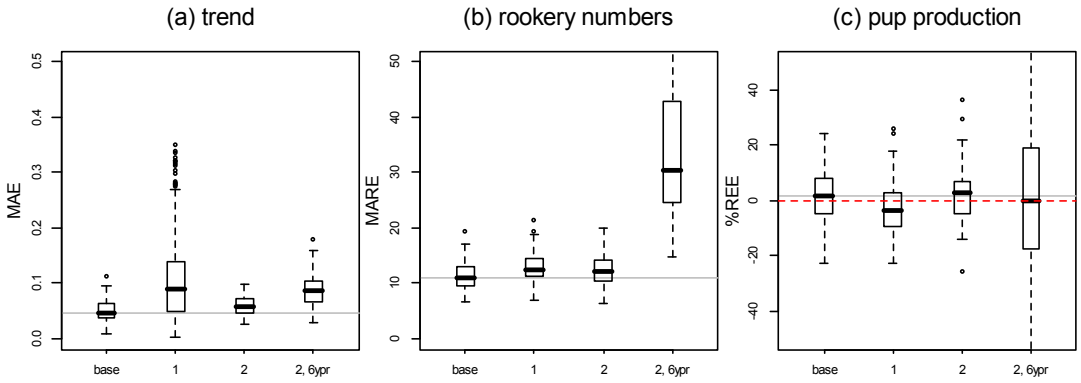


Figure 1.4. Simulation results for the pup numbers in the final year when the estimation model is mis-specified and Model #1 is the operating model. Scenario '1' : estimation model is Model#1, but rookeries are incorrectly assigned to regions, '2': estimation model is Model#2, no regional structure assumed. '2, 6ypr' : same as '2', except that there are missing data.

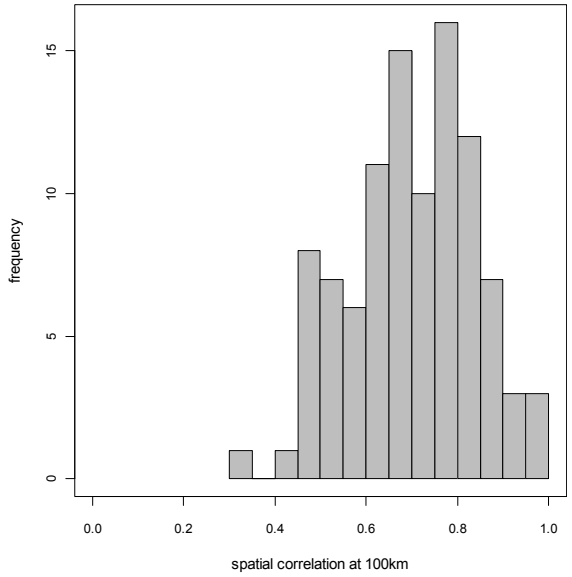


Figure 1.5. Distribution of estimates for φ_{100} when Model #2 is fitted to data where the "true" dynamics are governed by Model #1.

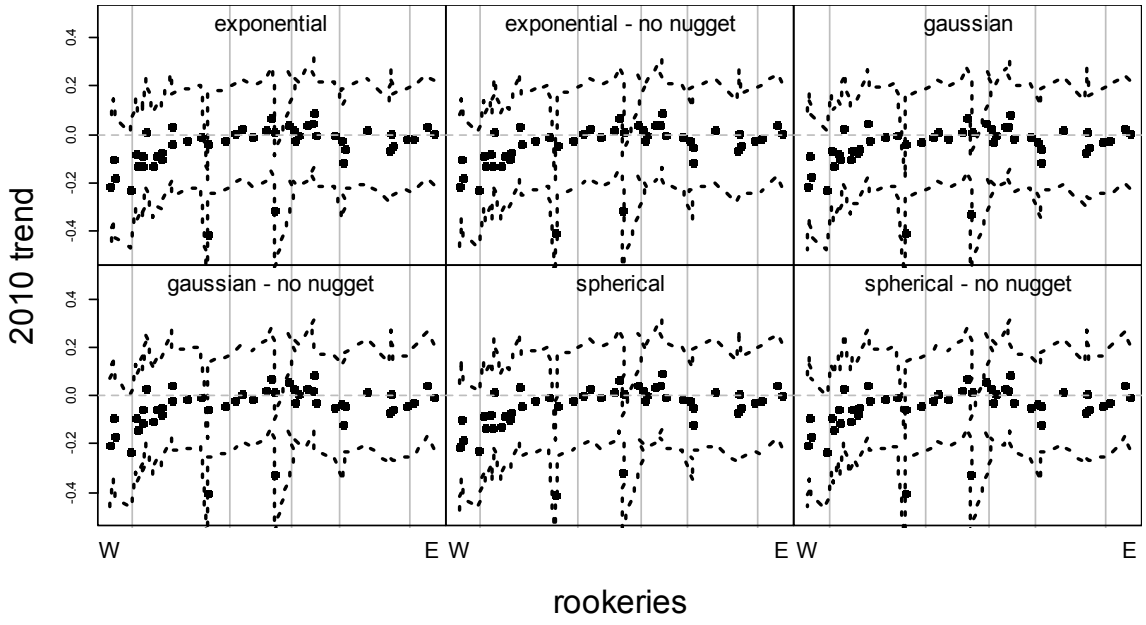


Figure 1.6. Estimates of the 2010 trend by rookery for the different configurations of Model #2. Vertical lines within panels delineate regions assigned in Model #1. Rookeries plotted west to east reading from left to right.

Chapter 2: Evaluating monitoring design performance for metapopulations with spatially varying trends

2.1. Introduction

Available effort for monitoring changes in population trends and abundance is often constrained by sampling costs and resources. Consequently, identifying monitoring strategies that allow for improved estimation of trends and abundance is crucial for resource management and efficient allocation of conservation efforts (Field et al. 2005, Wilson et al. 2006). Spatial structuring of changes in abundance within a population complicates the allocation of monitoring effort, as decisions must be made as to how sampling effort should be allocated spatially, in addition to tracking populations through time, as population trends likely vary geographically. In such instances, the importance of spatial considerations to survey design is well documented (e.g. Dobbie et al. 2008).

Evaluation of sampling designs frequently focuses on the use of statistical power analysis (e.g. Peterman 1990) to measure the effectiveness of a design to detect population trends, should they exist. While statistical power tends to increase with sampling effort, increases in sampling effort can be redundant when there is correlation among observations. Alternatively, the concept of optimal experimental design (e.g. Silvey 1980) views sample design as an optimisation problem, whereby designs are selected to achieve the maximum information (for instance the estimates of parameters and abundance with least bias or lowest RMSE), given certain design constraints, such as cost.

Chapter 1 developed models for estimating trends and abundance of spatially-correlated populations using count data, and included the consideration of instances where there are missing data. In this Chapter, we investigate how the spatial and temporal allocation of monitoring effort affects the performance of the Kalman filter-based models explored in Chapter 1 when there are a fixed number of data points. The results of the simulation

testing in Chapter 1 are further analysed to determine how spatio-temporal allocation of monitoring effort impacts model performance. The aim is to identify how the number of years sampled and percentage coverage of population units within years affect estimation of trend, abundance, and model parameters. This Chapter also investigates the interaction between the ability to estimate different parameters of interest and the choice of which of these parameters was optimised in the sampling design.

2.2. Methods

The bases for the analyses described in this Chapter are the configurations of Model #2 used for the simulations described in Chapter 1. This model assumes no regional stock structure; rather that there is spatial correlation in the annual changes in population trend. Pre-specified "true" sets of parameter values (Table 2.1) were used to generate multiple realisations of the operating model (Model #2, exponential form for the spatial correlation function, see Chapter 1). A suite of alternative monitoring designs (each defined by a specified allocation of data points spatially and temporally) were used to generate data sets, to which Model #2 was fit. A set of eight performance metrics (Chapter 1) were used to summarise the results of the simulations.

The results of the simulations were analysed in two ways. First, we evaluated the improvement in model performance that was obtained when monitoring designs were optimised for the individual performance measures. Tradeoffs in the ability of optimal monitoring designs to estimate other quantities of interest were identified by comparing the values for the performance measures obtained under optimal designs with those obtained from random monitoring designs. Second, we used a constrained ordination method to assess the degree to which the performance of the modelling framework was dependent on the spatial and temporal coverage pattern of the data.

2.2.1. Evaluating optimal monitoring strategies

Simulation scenarios considered alternative values for the operating model parameters, providing bracketing ranges for the degree of spatial autocorrelation among subpopulations, and the values estimated from fitting Model #2 to pup count data for western Alaska Steller sea lions from 1975-2010 (Table 2.1). For each scenario, 100 realisations of the operating model were generated for 15 subpopulations over 25 years. A complete data set (observations for each subpopulation for all years) was generated for each of the operating model realisations given a specified value for the observation error standard deviation. Selections from these complete data sets were then used to construct the data sets for the estimation model, given monitoring designs that differed in the spatial and temporal allocation of a fixed amount of sampling effort (90 observations over the 25 years and 15 subpopulations). 100 monitoring designs were constructed that represented four types of survey design:

1. 100% spatial coverage in all years for which data are available, six years of data,
2. observations in all 25 years, but each subpopulation is only counted six times (preference to sampling in time versus space),
3. observations restricted to 12 of the 25 years, with each subpopulation counted six times (intermediate between designs 1 and 2), and
4. observations in all 25 years, uneven sampling among subpopulations (5 subpopulations counted 12 times, the remaining 10 only 3 times each).

The 100 monitoring designs (Figure 2.1) represented a subset of the large number of ways of allocating the ninety available data points through time and space.

For each of the operating model realisations, the 100 survey designs were used to obtain 100 possible datasets (from the 'complete' dataset) that were then used to fit the

estimation model. The values for the eight performance metrics were calculated for each of the data set * survey design combinations.

For each of the eight performance metrics, an 'optimal' survey design for that metric was chosen, this being the survey design resulting in the lowest median of the absolute value of the particular performance metric (integrating across process and observation error). The percentage of times that the optimal design 'beat'[b] a survey design in which data points were randomly selected (without replacement) from the 25 years x 15 subpopulations for each of the 100 realisations of the operating model was calculated, both for the metric that the design was optimised for, and the remaining metrics to assess the value of optimising the survey design for a particular performance metric. High values for these percentages would indicate an impact of the spatial and temporal monitoring pattern on the ability to estimate the relevant quantities, with the distribution of the values for the performance metrics obtained when the survey design was optimised being different from that from random assignment of data points.

2.2.2. Assessing impacts of data coverage patterns

500 simulations were conducted using Model #2 as both the operating and the estimation model (Equation 1.2). The parameters of the operating model are detailed in Table 2.1. The data generated for each simulation were 90 data points, allocated randomly across the 25 years and 15 populations, with the only stipulation being that each population was counted at least twice during the 25 year time period. Each of the simulations therefore had 76% 'missing data', but a unique spatio-temporal coverage pattern for the data (of which those in Figure 2.1 are a subset).

---

[b] The optimal design was deemed to beat the random design if, for that operating model and full dataset realisation, the absolute value for the performance metric was lower.

The degree to which the variability in the values for the performance metrics obtained during the simulations was due to the pattern of data coverage was investigated using Canonical Correspondence Analysis (CCA, Ter Braak 1986, Legendre and Legendre 1998). CCA constrains an ordination of a data matrix (here, the values for performance measures for each simulation) based on an ordination of an additional matrix (here, the data coverage pattern by year). The goal is to quantify how the dominant patterns in an explanatory matrix contribute to ordination axes of the response metrics. This analysis aimed to identify how the data allocation pattern of monitoring designs contributed to estimation of model performance measures. The extent of data coverage in each year was used to constrain an ordination of the matrix of performance metrics. The data coverage for each year ranged from 0 to 100%, with the latter meaning that all populations were counted in that year. This implementation of CCA did not investigate how the data points were allocated by subpopulation over time and space, and simply considered the extent of coverage in each year with data. The resulting constrained ordination was then correlated with simple metrics from the data coverage pattern (such as the number of years with data).

2.3. Results

The general performance of the estimation framework was poorer than the base-case analyses presented in Chapter 1, due to the presence of 'missing data'. Integrating across monitoring designs, the distributions of values obtained for the eight performance metrics differed by scenario (Figure 2.2). Estimation of the trend and individual rookery numbers in the final year was better when the operating model parameters were based on the values estimated for western Alaska Steller sea lions (Fig. 2.2a-b) than for the other scenarios. This scenario had both higher levels of process error variability and greater extent of spatial correlation in the trend among rookeries (Table 2.1). The value for the trend process error standard deviation was better estimated under this scenario as were

the values for the observation error standard deviation and the spatial correlation parameter (Figures 2.2d, f, and g). There was little impact of changing the value for the spatial correlation parameter from the base-case scenario on the ability to estimate trends and abundance (Figure 2.2a-c, compare 'base-case' with 'spatial cor=0.9' and 'spatial cor=0.1').

Changes in the distributions of the various performance metrics given 'optimal' monitoring strategies depended on both the metric chosen to define optimality, and the scenario used as the basis for the operating model. Figure 2.3 illustrates the distributions obtained under the base-case scenario for the eight performance metrics (panels a-h) across all designs (left-most box in each panel) and the monitoring designs that resulted in the lowest median value for the performance metric (distributions of optimal designs plotted for each metric are the remaining columns).

Optimising the monitoring design for the final year trends, rookery numbers, and total abundance resulted in the greatest improvements in the values for these performance metrics (Figure 2.3a-c). Indeed, the percentage of times that the optimal monitoring design performed better than a random design for these three performance metrics, given that the metric of interest was that which the design had been optimised for, was greater than 85% for each of the scenarios (Tables 2.2 – 2.5, initial three bolded values along the diagonal). An exception was the performance of the optimal 'trend' design under the SSL-estimated parameters scenario (Table 2.3). Conversely, there appeared to be little gain in optimising the monitoring design for the model parameters, with the optimal designs resulting in values for the performance metrics that only performed better than a random design around half the time (Tables 2.2 – 2.5, remaining bolded values along the diagonals). This lack of improvement is reflected in Figure 2.3 (panels d-h) for the base-case scenario, with little change in the distribution of the values for the performance metric for the optimal designs compared to the distribution obtained when integrating across all monitoring designs.

Optimising monitoring designs for particular performance metrics also had an impact on the distributions obtained for performance metrics that the optimisation was not specific to. Under the base-case scenario (Table 2.1), optimising the monitoring design for the total abundance in the final year appears to be the best option as this design also results in estimates of the trend and individual rookery numbers that perform better than the random design at almost the same rate as when the design is optimised for these particular metrics (Table 2.2, 'total abundance' row, 1st and 2nd columns). Likewise, optimising for either trend or individual rookery numbers also leads to near-optimal rates of success for the other metric (Table 2.2, rows 'trend' and 'rookery #'s'), although the optimal design for these metrics does not appear to result in the best estimates of total abundance, with only a 65% rate of performance better than a random design compared to the 89% when the design is optimised for total abundance. Clearly then, there is more to gain by choosing a design that optimises the estimate of current total abundance over trend and individual rookery numbers for this scenario.

There is however, a tradeoff associated with optimising the design for abundance, as the values obtained for the estimates of observation and process error standard deviations are poorer than when the design is optimised for these parameters (Table 2.2, compare values in the 'total abundance' row with the bolded values in each column). The 95% intervals for the proportion of times the design beats the random one however, include 50%, meaning that the dropoff in performance is not that great. A greater decrease in the ability to estimate trend and abundance can be seen when the monitoring design is optimised for, say, observation error.

The losses and gains associated with optimising the design for different performance metrics unsurprisingly changes with the parameters of the operating model. There is little gain to be had in the estimate of current trend, regardless of what performance metric the survey design is optimised for when the operating model is based on the parameters estimated from Steller sea lion pup count data (Table 2.3, 'trend' column). This is

probably because the current trend is much better estimated for this scenario than for the others (Figure 2.2a). The results detailed in Chapter 1 indicate that this may be due to the absence of the nugget variance (see Figure 1.3a). As for the base-case scenario, optimising for trend results in near-optimal rates of success for estimation of individual rookery numbers, but very poor (mean 11%) rates of success at estimating current total abundance (Table 2.3, 'trend' row). The optimal monitoring design for current total abundance under this scenario also results in less than optimal estimates of individual rookery numbers (Table 2.3, 'total abundance' row). The optimal design for total abundance obtained when only the spatial correlation parameter is increased from the base-case scenario value does not do as well at estimating trends and rookery numbers (Table 2.4, 'total abundance' row). However, when the spatial correlation is low, the success rate of the optimal design for total abundance in estimating these metrics does approach that of the optimal designs for these metrics (Table 2.5, compare entries in 'total abundance' row with those from Table 2.2).

Despite not improving upon the choice of a random design, optimal designs for the trend process error standard deviation were almost as successful at estimating trends and individual rookery abundance in the base-case scenario and when the spatial correlation was 0.9 (Tables 2.2 and 2.4). However, the optimal design for this parameter did not prove successful at replicating optimal rates of selection when the operating model was based on the estimated parameters for Steller sea lions (Table 2.3), or when the spatial correlation was low (Table 2.5).

Given little spatial correlation in the trend (φ_{100} = 0.1), the optimal survey design for estimation of the spatial correlation parameter resulted in success rates for all three trend- and abundance-related performance metrics as good as when the survey design was optimised for these metrics (Table 2.5, 'φ_{100} row', first 3 columns). This was not observed for the other operating model scenarios. indeed Except for the estimation of

individual rookery numbers under the SSL parameter scenario, the optimal design for φ_{100} performed markedly poorly relative to a random design when estimating the trend and abundance metrics (Tables 2.2 - 2.4, 'φ_{100}' row).

The optimal designs for the final year trend, individual rookery numbers, and total abundance tended to favour sampling across space versus time (Figure 2.4, columns a-c). That is, the designs selected were (with the exception of total abundance under the φ_{100} = 0.1 scenario) all of the type that minimised the number of years with observations, with all subpopulations counted in those years with observations (Figure 2.4, columns a-c). In particular, these designs had a complete survey of the population in either the final, or penultimate, year. In contrast, the optimal designs for the observation error standard deviation had observations in all years, but sampling was allocated unevenly across subpopulations, in that a small number of subpopulations were sampled often, with only three observations at each of the other subpopulations during the time series (Figure 2.4, σ_{ε} column d). The optimal monitoring designs for the remaining model parameters varied between these two types, albeit for those designs where sampling through space was favoured, coverage in the final year was not as pervasive as for the trend and abundance optimal designs (Figure 2.4, columns e-h).

In summary, the largest gains in performance were obtained when optimising for the MAE in final year trend, the final year rookery numbers, and the total abundance in the final year, although the tradeoffs among the values for the performance metrics given different optimisations were more pronounced when the level of spatial correlation was high (both for the SSL estimated parameters and the 'φ_{100} = 0.9' scenarios). There was also a tendency for a negligible impact of survey design choice on the ability to estimate the model parameters (e.g. Figure 2.3,d-h).

2.3.2. Constrained ordination

The pattern of data coverage (number of rookeries surveyed each year) did not account for very much of the variation in the performance measures (13% of total variance) under the base-case values for the parameters of the operating model, with the first canonical correspondence axis (CCA1) explaining nearly all of the variability accounted for by the constraining matrix (97%, or 12.5% of the total variance). Figure 2.5 displays the positions of individual simulations along the first two canonical axes given the linear constraints. The extent of coverage in the most recent four years (years 22, 23, 24, and 25) had the highest loadings along these axes (Fig. 2.5, black arrows), and were correlated with changes in the value for the MARE of the rookery numbers in the final year, shown in Figure 2.5b by the gradation from red (poor estimates) to green (good estimates) of this metric. A similar pattern was observed for the MAE of the trend in the final year. However, CCA1 was not as correlated with the value for the MAE of the trend as for the MARE of rookery numbers (the colours of points in Fig. 2.5a are more scattered than in Fig. 2.5b). While the first canonical axis appeared to also explain much of the variation in the absolute value for the %REE of total pup production in the final year, the direction of change in the value for this metric was orthogonal to that of the MARE in rookery numbers, suggesting that simply having high levels of coverage in the recent years was not the only necessity for accurate estimates of this quantity. This result is consistent with the evaluation of survey designs, with the optimal designs for trend and individual rookery numbers not also being optimal for total abundance (Table 2.2).

Correlation of the number of years containing data for each simulation with the results of the constrained ordination revealed that poorer estimates of the final year trend and rookery abundance were associated with monitoring strategies that allocated data points over a large number of years (compare the direction of purple arrow headed '#yrs' in Fig. 2.5a-b to that of point colour change). This is unsurprising, because scenarios with high levels of coverage in the recent years have few observations in other years to achieve the

overall coverage level. The data coverage patterns did not appear to be correlated with the variability among the simulations in terms of the ability to estimate the model parameters, because there is little evidence of an ordering of the EE (gradation of green to red) along CCA1 and CCA2 in Figure 2.6. An exception would seem to be the spatial correlation parameter (Fig. 2.6d, "cor 100"), with better estimates for this parameter being associated with scenarios that spread the data over a number of years.

CCA results for scenarios when the true values for the spatial correlation were low (0.1 at 100km) or high (0.9 at 100km) were qualitatively similar to those for the base-case scenario (results not shown). A slightly greater proportion of the variance in the performance metrics was explained by the constrained ordination when the spatial correlation was low than under the base-case scenario (15.4%). The proportion explained by the constraints when the correlation was high was 10.1% of the total. This suggests that changes in estimation ability by simulation are less dependent on the pattern of the percentage coverage within years as the spatial correlation in trend among subpopulations increases.

2.4. Discussion

The choice of monitoring design clearly impacts the ability to estimate the final year trends, individual subpopulation numbers, and total metapopulation abundance, with tradeoffs in the performance of the estimation procedure depending on the particular metric that designs are optimised for. Choosing the optimal monitoring designs for estimating total abundance in the final year appears to also provide near optimal estimates for the final year trend and individual rookery abundance when spatial correlation in the population trend is low. However, the estimation of these metrics is suboptimal when the extent of spatial correlation is high. There was little gain in estimation ability over the selection of a random design when the survey design was

optimised for the estimation of the model parameters. The optimal monitoring designs for the recent trend and abundance performance metrics were typified by preferential sampling across space versus time, whereas those for the model parameters also included sampling designs that sampled across years, but which concentrated sampling on relatively few subpopulations (Figure 2.4).

The 100 survey designs examined in the evaluation of optimal designs for the different performance metrics are clearly a very small subset of the possible monitoring designs, even given the constraint of a fixed number (90) of data points throughout the monitoring period. However, the types of designs examined represent much of the range of possible designs in the allocation among subpopulations, space, and time.

Optimising the monitoring designs for the model parameters was generally not successful at improving the model estimates. This suggests that the errors in the parameter estimates (Figure 2.2d-h) cannot be overcome by careful allocation of monitoring effort, and that additional monitoring resources are required (see Figure 1.3, Chapter 1).

A lack of improvement in the estimation of the final year trend when the survey design was optimised for this metric and the operating model was based on the estimated values for Steller sea lions may be due to the fact that the MAE was generally substantially lower under this scenario than for the other operating model parameterisations (Figure 2.2a). However, a similar improved performance in the estimation of rookery numbers in the final year was also observed for this scenario (Figure 2.2b), yet the optimal designs for this metric were much better than a randomly-selected design (Table 2.3). The lack of a nugget term in the trend under this scenario may also be responsible; results of simulations in Chapter 1 showed that the MAE of the trend was most improved when there was no nugget component in the operating model (Figure 1.3a, Chapter 1).

A minmax strategy can be determined for each scenario by calculating, for each optimal design, the sum of the differences in performance rates for each metric when compared to that obtained with the optimal design for that metric. Such an approach weights improvement in performance equally among the performance metrics. Calculating this for the base-case scenario reveals that the optimal design which results in both the greatest improvement in performance and lowest reduction is that when optimising for the final year total abundance (Table 2.6, compare bolded value in 1st column with the other values in the same column), while that for the other scenarios is the design which is optimised for final year individual rookery numbers (Table 2.6, bolded values). The worst option using this criterion (italicised values in Table 2.6) for the base-case scenario is to choose the design optimal for estimating observation error.

The extent of spatial autocorrelation should have major implications for how sampling effort should be allocated to maximise prediction ability (Rhodes and Jonzén 2011). Despite assuming spatial and temporal independence in abundance, Carlson and Schmiegelow (2002) showed that the best strategy for maximising power was to sample many sites relatively infrequently. In the present study, there was less of a gain in trend estimation when spatial correlation in trend was high than when spatial correlation in trend was low when sampling was preferentially allocated across space versus time. Figure 2.4a shows that the optimal monitoring designs for trend estimation are similar between scenarios when the true spatial correlations are low and high, yet there is an appreciable difference in the success rate obtained when selecting these optimal designs over random designs (Tables 2.2 – 2.5, first column). The presence of the nugget variance in the simulations may have diminished the ability to identify differences in estimation ability with different values for the spatial correlation. Estimates for the value of the spatial correlation parameter for Steller sea lions were high, with correlation in trend greater than 0.98 at 100km for the analyses in Chapter 1 (e.g. Table 1.3). However, the optimal monitoring designs for estimation of trend and abundance for this scenario still

included relatively infrequent range-wide surveys, suggesting that the current practice of fairly comprehensive pup count surveys every other year or so might not be detrimental to the ability to estimate current trends and individual rookery abundance, given different methods of allocating the same set of monitoring resources.

The presence of substantial observation error in the simulations may have impacted the ability to gain much from optimising the sampling design – observation uncertainty is almost certainly prevalent when estimating abundance during monitoring programmes (e.g. Hilborn and Mangel 1997). Estimation of trends and abundance was improved in the absence of observation error (Chapter 1, Figure 1.3a-c). Rhodes and Jonzén (2011) noted that the presence of independent observation errors resulted in a diminished threshold to preferentially allocate sampling across time rather than space even in the presence of high spatial correlation. The influence of observation error may help to explain the result that the best strategies for estimating trend and abundance composed few survey years even when the spatial correlation in trend among subpopulations was high (e.g. Figure 2.4). Pollock et al. (2002) emphasise the double sampling approach as a means of estimating detectability, akin to estimating the observation error standard deviation in the present study. Evaluating the tradeoffs associated with allocating some of the monitoring effort to double counts at certain subpopulations, as a means of identifying the extent of observation error, might be worth pursuing (but evaluation of this option was beyond the scope of the present chapter).

Zimmerman (2006) found that optimal spatial network designs for abundance estimation (given known parameter values) were largely antithetical to designs that efficiently estimated covariance parameters, although designs that optimised prediction efficiency when covariance parameters were unknown showed similarities to those under both the previous scenarios depending on the situation, and depended on the strength of spatial correlation, in addition to other parameters. While Zimmerman was considering geostatistical estimation of a spatial point pattern, as opposed to a fixed spatial

configuration evolving over time (as is the case with the fluctuating subpopulations described here), the CCA results suggest that the percentage coverage of surveys by year did not explain the Monte Carlo variability in parameter estimation (Figure 2.6). Indeed, the 'optimised' designs demonstrated little benefit to parameter estimation by optimising for the model parameters.

The CCA results suggest that better estimates of current abundance and trends are associated with high levels of coverage in the recent years, as also noted by the choice of optimal designs (Figure 2.4), and also with a large number of years for which data are available. However, for Steller sea lions there are geographic disparities in the levels of coverage, with rookeries in the Gulf of Alaska and Eastern Aleutians regions tending to be counted more frequently than those in the Central and Western Aleutian Islands. This is unsurprising, given the cost associated with assessing sea lion populations in these more remote areas. Geographic allocation of data points may be more important to trend estimation than simply the number of years surveyed. Further simulation analyses could explore the ability of the models to estimate trends and abundance where the relative cost of surveying differs among subpopulations, perhaps by skewing the cost of counting a particular rookery to one end of the geographical range.

The methods presented here allow for simple examination of the potential impacts of spatio-temporal allocation of monitoring resources. In addition to not considering costs, the analyses also did not investigate the impact of relative subpopulation size on the various performance measures, That is, there was no accounting for the proportion of the total numbers monitored in a given year, just whether counts were obtained from subpopulations, irrespective of how these subpopulations contributed to the overall population. Clearly, in terms of overall abundance, it may be more beneficial to assign monitoring efforts to those subpopulations which make up the majority of the total population. However, this may mean that events such as declines or extinctions of smaller subpopulations may be identified less well. Distinguishing monitoring designs

that exhibit some preference for sampling subpopulations that constitute a larger proportion of the total population than others can be viewed as analogous to preferentially sampling sites with suitable habitat under a presence-absence study, to detect changes in site occupancy or community structure. In such an instance, Rhodes et al. (2006) found that the optimal strategy for detecting long-term occupancy rates or trends depended on whether occupancy was increasing or decreasing. Sampling high-quality habitat was optimal to detect negative trends. However a strategy that sampled intermediate-quality habitats was optimal for detecting positive trend. Habitat quality may also be more formally incorporated into the decision process as to where to sample for Steller Sea lions. Ban and Trites (2007) developed tools for defining haul-out (or rookery) quality in terms of the likelihood of sites supporting populations of Steller sea lions. The allocation of monitoring resources (given other constraints) could in theory be based on such metrics. Dobbie et al. (2008) review examples and methods for the inclusion of prior knowledge (or 'expert elicitation') in a design context.

The analyses considered here all displayed inter-simulation variability in relative population sizes and temporal changes in trends. That is, the optimal monitoring designs for each of the performance measures were not tailored to a particular realisation of the operating model. It may be that monitoring design ought to be adaptive in that the definition of what is optimal may depend on the actual state of the system at a given time, i.e. what the actual trends and relative population sizes are. Future applications of these analyses might be to simulation test the estimation ability of different monitoring designs given a known 'true' time series of state vectors. Extensions to the simulations with the estimated Steller sea lion parameters would be to use the estimated state vectors from Chapter 1 with which to evaluate the designs, both in the historical period that the models were fit to data from (during the period of Steller sea lion population decline) and using the final estimates of relative pup production by rookery and their trends (along with the associated estimated uncertainty). Advantages of this approach would be that comparison

could be made among designs with respect to a somewhat fixed notion of relative subpopulation size that reflects reality, and that integration across process and observation error could be done in a more coherent manner using the estimated variance-covariance matrices. It is unlikely however, that the evaluation of future possible designs would produce results substantially different from those described by the simulations presented here, unless alternative performance metrics were used in the definition of optimisation criteria. An additional approach might be to consider testing the implementation of adaptive monitoring designs, where the allocation of monitoring effort changes depending on the estimation of trends and abundance. Such an approach could also follow from the results of Chapter 1, where an estimate of the current state (magnitudes of trends and subpopulation sizes) of the Steller sea lion metapopulation is now available.

Evaluation of monitoring designs through repeated Monte Carlo simulations, as presented here, is a fairly straightforward approach to address issues related to the impacts of experimental design. Alternative, more complex and computer intensive methods, such as simulated annealing (Zhu and Stein 2006) and Stochastic dynamic programming (SDP) methods (Mangel and Clark 1988) have been used to address issues of defining optimal monitoring strategies for management of a collection of subpopulations (McDonald-Madden et al. 2008). The latter approach requires an objective function, which in the case presented here might involve an appropriate weighting of the various performance metrics. It is clear from the results of the simulations that any optimal strategy derived from a SDP approach will also depend heavily on such weightings. While SDP is attractive for providing exact optimal solutions, it can be time-consuming when the number of state variables is large, and can lack transparency over simple rule of thumb methods that attempt to capture simple properties of the behaviour which may often perform almost as well as the exact solutions (McDonald-Madden et al. 2008).

Ultimately, the value of specific monitoring schemes should be judged by the improvement they make to predictions (Hauser et al 2006). This suggests that it might be prudent to evaluate monitoring designs on the basis of the uncertainty associated with estimation in addition to the ability to obtain mean estimates of the quantities of interest accurately. It is also likely that the relative value of certain data allocation schemes will depend on the status of the population relative to thresholds that initiate or change management actions because the results of monitoring efforts are used to guide management actions. Certain outcomes from the monitoring design (e.g. low subpopulation size, or the ability to recognize negative population trends) may be more important than others. Thus, the monitoring design should be shaped by the (multiple) management objectives for the resource. Consequently, optimal monitoring designs for these particular population states can be expected to differ from those reported in this Chapter, where accurate and precise estimates of population parameters and current trends and abundance were used as the design objectives.

As demonstrated here, it is likely that the definition of what is considered an 'optimal' monitoring design will differ given different objectives. Simulation testing of the efficacy of alternative monitoring designs given specific objectives, as presented in this Chapter, provides a rapid and inexpensive way to evaluate the tradeoffs associated with selecting designs that are preferentially optimised for a particular objective, while incorporating uncertainty in both the population dynamics, and the observations obtained from the monitoring design. Combining this approach with knowledge regarding the state of the resource (as with the scenario using the SSL parameter estimates from Chapter 1) and costs associated with the various monitoring designs should provide practical advice for the consequences of alternative decisions regarding the allocation of monitoring resources.

Table 2.1. Specifications used for the operating models during the simulations

Scenario	True values for model parameters					# of data points
	σ_ε	σ_r	σ_β	φ_{100}	ψ	
1. base-case	0.20	0.10	0.05	0.50	0.01	90
2. estimated SSL parameters	0.23	0.15	0.08	0.98	0.00	90
3. high φ_{100}	0.20	0.10	0.05	0.90	0.01	90
4. low φ_{100}	0.20	0.10	0.05	0.10	0.01	90

Table 2.2. Mean (and 95% interval based on repeated selection of alternative design) percentage of times that the optimal design for each performance measure 'beat' a random design, for the base-case operating model parameters. Bold values indicate where the performance measure looked at was the same as that which the design was optimized for.

performance measure	percentage that optimized design beat random design							
design is optimized for	trend	rookery #'s	total abundance	σ_ε	σ_r	σ_β	φ_{100}	ψ
trend	**93 (88,97)**	99 (97,100)	65 (57,72)	51 (42,59)	49 (42,57)	49 (41,57)	48 (40,57)	57 (47,65)
rookery #'s	92 (87,97)	**99 (97,100)**	64 (56,72)	52 (42,60)	48 (41,55)	50 (40,58)	49 (39,58)	56 (46,65)
total abundance	93 (88,97)	97 (93,100)	**89 (83,94)**	45 (36,54)	49 (41,56)	51 (43,59)	59 (50,68)	56 (49,64)
σ_ε	25 (19,31)	24 (18,31)	27 (20,34)	**63 (56,71)**	57 (49,65)	46 (38,54)	46 (39,55)	41 (33,48)
σ_r	46 (36,54)	39 (31,48)	66 (58,74)	56 (48,64)	**58 (52,66)**	46 (37,53)	46 (39,53)	56 (47,66)
σ_β	92 (88,97)	99 (96,100)	36 (27,45)	50 (41,56)	50 (43,58)	**58 (48,66)**	47 (39,55)	57 (50,66)
φ_{100}	23 (18,28)	21 (13,27)	43 (37,50)	59 (50,68)	51 (42,60)	44 (36,51)	**57 (50,66)**	40 (32,48)
ψ	92 (88,96)	99 (97,100)	35 (27,45)	49 (39,58)	50 (43,59)	58 (49,66)	47 (39,55)	**58 (49,66)**

Table 2.3. Mean (and 95% interval based on repeated selection of alternative design) percentage of times that the optimal design for each performance measure 'beat' a random design, for the operating model parameter values estimated for Steller sea lions. Bold values indicate where the performance measure looked at was the same as that which the design was optimized for.

performance measure	percentage that optimized design beat random design							
design is optimized for	trend	rookery #'s	total abundance	σ_ε	σ_r	σ_β	φ_{100}	ψ
trend	**64 (55,72)**	94 (89,98)	11 (6,16)	46 (40,53)	53 (45,61)	48 (41,58)	63 (56,70)	53 (45,60)
rookery #'s	62 (55,72)	**98 (95,100)**	64 (54,71)	54 (48,62)	51 (44,57)	49 (41,58)	67 (58,74)	46 (38,54)
total abundance	52 (44,60)	62 (53,70)	**81 (75,88)**	42 (34,48)	49 (41,56)	47 (39,57)	67 (60,74)	50 (43,58)
σ_ε	49 (41,56)	24 (16,31)	41 (34,48)	**58 (48,66)**	46 (39,55)	44 (36,51)	57 (49,64)	51 (44,59)
σ_r	54 (47,63)	32 (24,39)	42 (33,50)	51 (43,59)	**55 (48,63)**	51 (43,57)	63 (54,72)	55 (46,62)
σ_β	60 (53,67)	72 (64,79)	34 (26,41)	51 (44,58)	54 (47,62)	**58 (49,67)**	60 (51,71)	58 (50,65)
φ_{100}	62 (54,69)	98 (95,100)	25 (18,32)	46 (36,54)	51 (43,58)	49 (41,56)	**71 (63,80)**	48 (40,55)
ψ	51 (42,58)	28 (20,37)	51 (44,58)	52 (44,59)	51 (46,58)	48 (38,59)	68 (59,75)	**60 (53,66)**

Table 2.4. Mean (and 95% interval based on repeated selection of alternative design) percentage of times that the optimal design for each performance measure 'beat' a random design, with the spatial correlation at 100km equal to 0.9 (scenario 3, Table 2.1). Bold values indicate where the performance measure looked at was the same as that which the design was optimized for.

performance measure	percentage that optimized design beat random design							
design is optimized for	trend	rookery #'s	total abundance	σ_ε	σ_r	σ_β	φ_{100}	ψ
trend	**92 (87,97)**	99 (95,100)	56 (48,65)	49 (42,58)	49 (39,57)	51 (42,60)	56 (50,62)	53 (45,61)
rookery #'s	92 (88,96)	**99 (96,100)**	58 (49,66)	49 (40,58)	48 (41,57)	51 (43,59)	56 (49,64)	54 (46,63)
total abundance	62 (54,69)	70 (62,79)	**88 (81,94)**	49 (41,58)	45 (38,52)	57 (50,67)	57 (46,65)	57 (49,63)
σ_ε	42 (32,51)	45 (35,52)	58 (51,65)	**60 (51,67)**	54 (47,61)	46 (37,53)	50 (42,56)	50 (41,58)
σ_r	20 (13,26)	27 (19,35)	77 (69,84)	50 (42,58)	**47 (38,55)**	44 (35,51)	45 (37,52)	50 (41,58)
σ_β	94 (91,98)	97 (94,100)	10 (6,14)	36 (29,43)	52 (44,60)	**56 (50,63)**	62 (52,71)	57 (50,66)
φ_{100}	26 (18,33)	31 (23,40)	13 (8,18)	45 (38,51)	53 (44,61)	56 (49,62)	**59 (51,67)**	57 (49,66)
ψ	94 (89,99)	97 (94,100)	10 (5,15)	37 (31,44)	52 (45,60)	56 (49,64)	62 (52,70)	**58 (51,67)**

Table 2.5. Mean (and 95% interval based on repeated selection of alternative design) percentage of times that the optimal design for each performance measure 'beat' a random design, with the spatial correlation at 100km equal to 0.1 (scenario 4, Table 2.1). Bold values indicate where the performance measure looked at was the same as that which the design was optimized for.

performance measure	percentage that optimized design beat random design							
design is optimized for	trend	rookery #'s	total abundance	σ_ε	σ_r	σ_β	φ_{100}	ψ
trend	**93 (87,97)**	97 (94,100)	6 (2,10)	38 (31,45)	53 (45,61)	46 (38,54)	51 (44,58)	54 (45,62)
rookery #'s	91 (85,95)	**99 (97,100)**	60 (53,69)	58 (49,66)	53 (46,63)	56 (49,63)	54 (46,61)	54 (47,61)
total abundance	92 (88,96)	97 (94,100)	**87 (80,92)**	45 (38,52)	48 (40,57)	48 (40,57)	54 (47,61)	51 (43,59)
σ_ε	44 (36,54)	37 (30,45)	29 (20,38)	**61 (54,68)**	52 (44,59)	51 (43,60)	43 (34,53)	48 (40,55)
σ_r	31 (23,39)	27 (20,34)	27 (20,34)	55 (48,62)	**60 (52,67)**	60 (52,66)	46 (38,53)	45 (38,53)
σ_β	62 (54,71)	64 (55,73)	67 (60,75)	50 (44,56)	52 (45,58)	**56 (48,64)**	53 (45,61)	51 (44,58)
φ_{100}	92 (87,97)	97 (93,99)	86 (81,93)	44 (36,53)	48 (40,56)	48 (40,56)	**54 (47,62)**	51 (43,60)
ψ	64 (54,72)	72 (63,79)	87 (80,92)	52 (44,59)	52 (45,60)	46 (38,54)	55 (47,62)	**54 (44,62)**

Table 2.6. Evaluation of minmax strategies for each scenario. Values are the sums of the differences in performance rates over metrics compared to the optimal performance rate shown in Tables 2.2-2.5, for each of the optimised designs. Bolded values for each scenario show the optimal design with the highest value for this metric, indicating that the design resulted overall in the highest gain and lowest loss. Values in italics indicate the optimal design that did the worst.

performance measure	Scenario			
design is optimized for	1	2	3	4
trend	-233	-80	-247	-306
rookery #'s	-234	**-21**	**-245**	**-219**
total abundance	**-205**	-62	-267	-222
σ_ε	*-415*	*-142*	-347	-379
σ_r	-331	-109	-392	*-393*
σ_β	-255	-65	-288	-289
φ_{100}	-406	-62	*-412*	-224
ψ	-256	-103	-286	-262

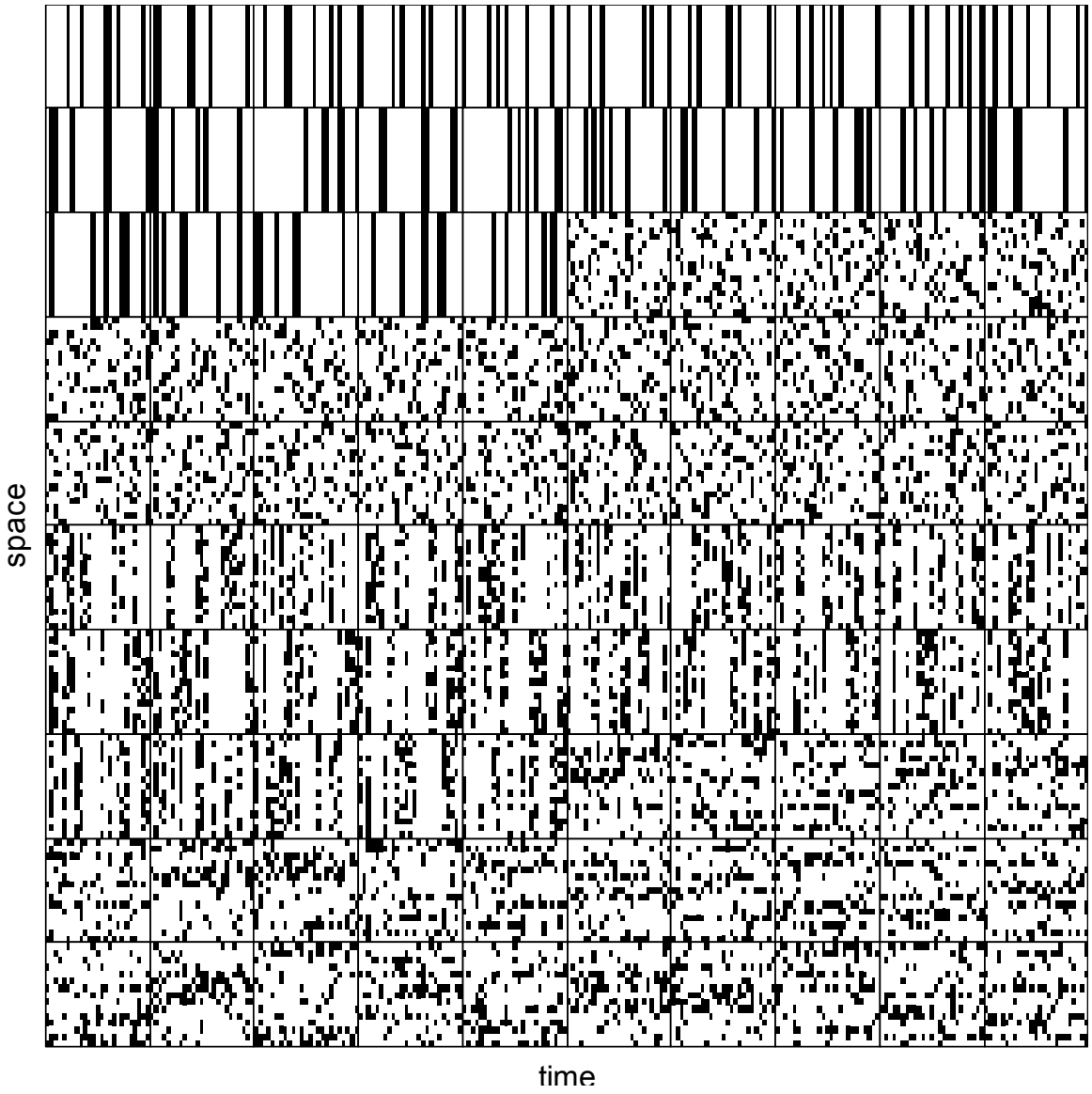


Figure 2.1. The monitoring designs. Designs 1-25 (the first two and an half rows) sample a few years, but count all subpopulations in those years. Designs 26-50 spread sampling over all years, with each subpopulation sampled 6 times. Designs 51-75 count each subpopulation 6 times, but sampling limited to 12 years during the time period (improved spatial coverage within sampled years). Designs 76-100 sample in all 25 years (10 subpopulations counted 3 times, 5 subpopulations counted 12 times).

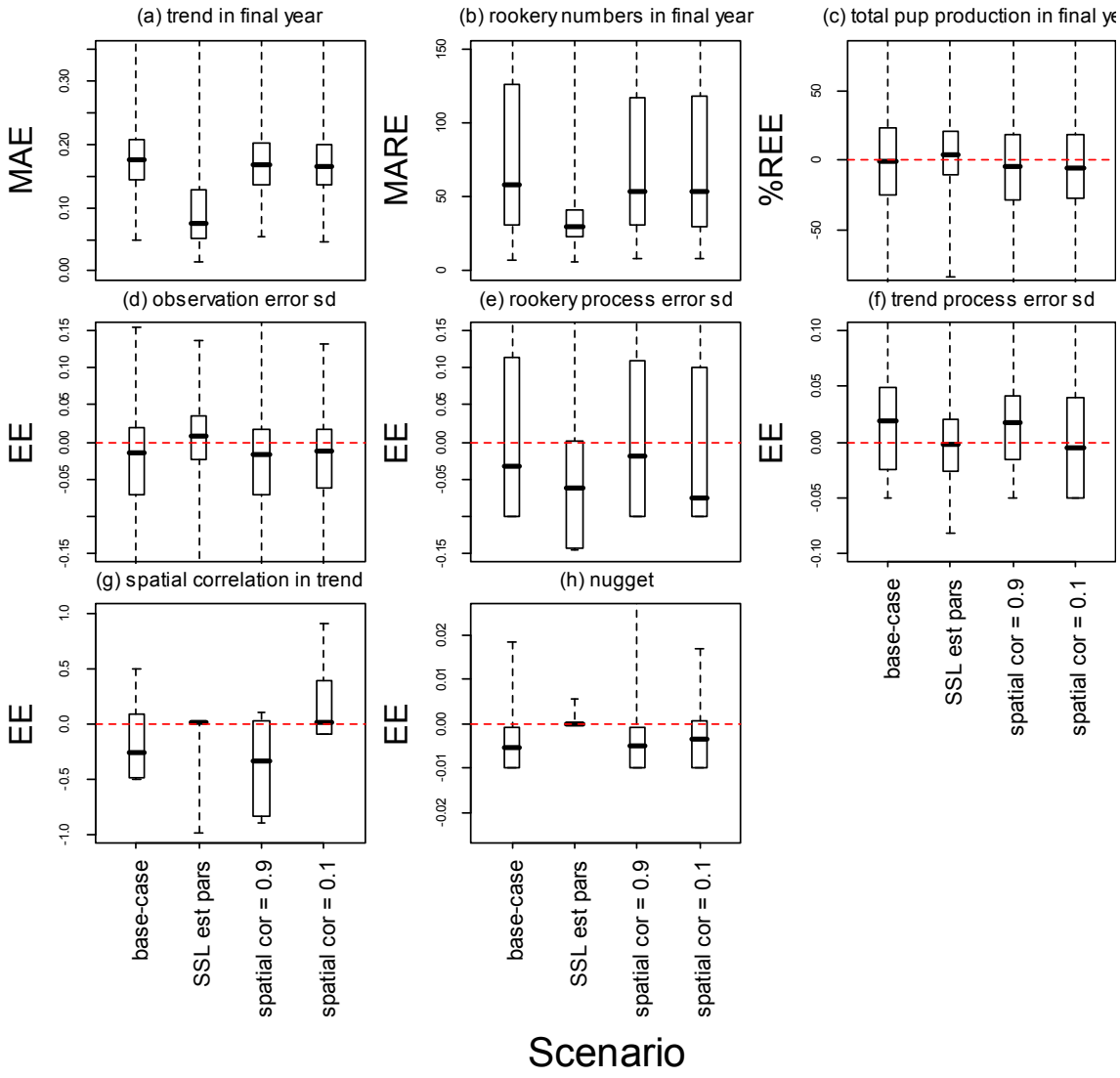


Figure 2.2. Distributions of the performance measures obtained from simulations for all the monitoring designs, for the four scenarios in Table 2.1.

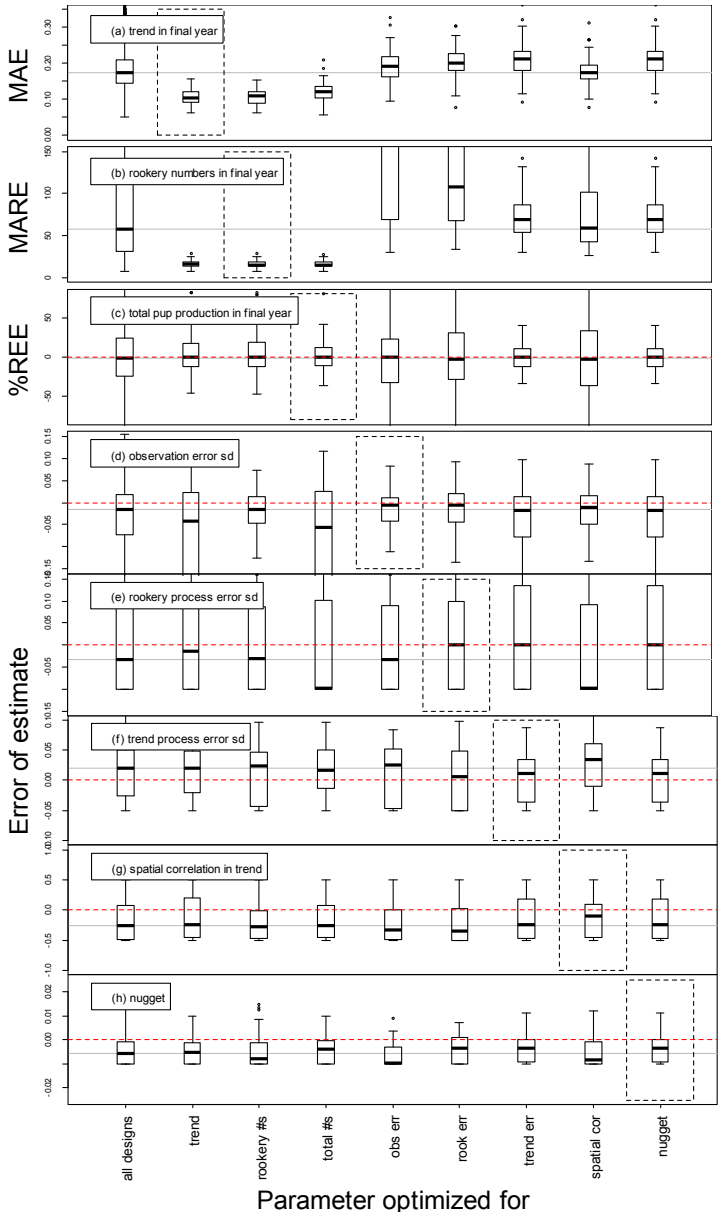


Figure 2.3. Distribution of the performance metrics obtained for all monitoring designs ("all designs") and when the optimal design was used (remaining columns). The spatial correlation parameter, φ_{100}, is 0.5 for this figure (base-case scenario in Table 2.1). Dashed polygons indicate where the performance measure looked at was the same as that which the design was optimized for. Ideally, the 2 metrics in panels a and b should be minimized while the remaining metrics should be as close to zero (red dashed line) as possible.

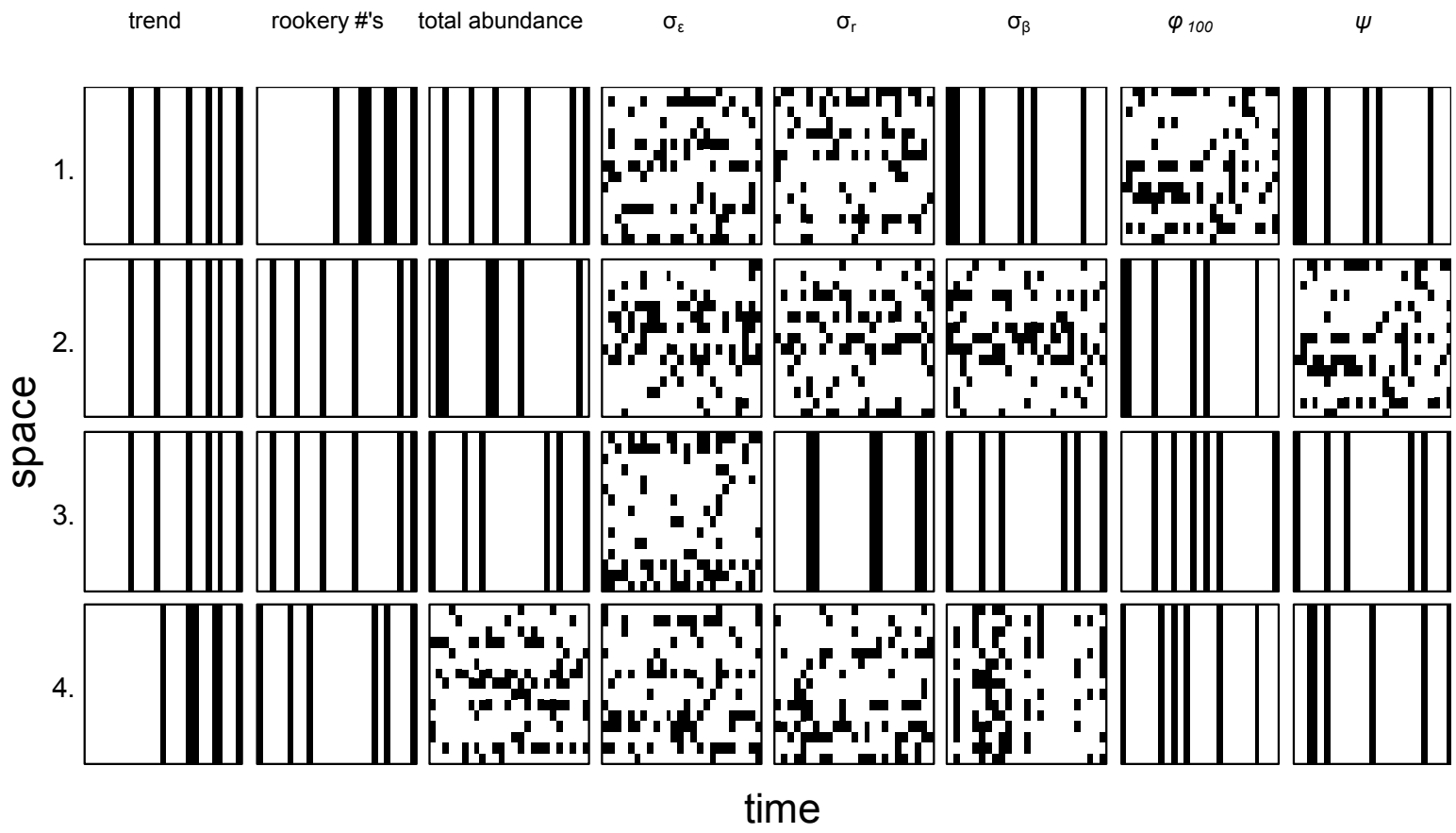


Figure 2.4. Optimal designs for each performance metric (columns) by scenario (rows). Scenario numbers as in Table 2.1.

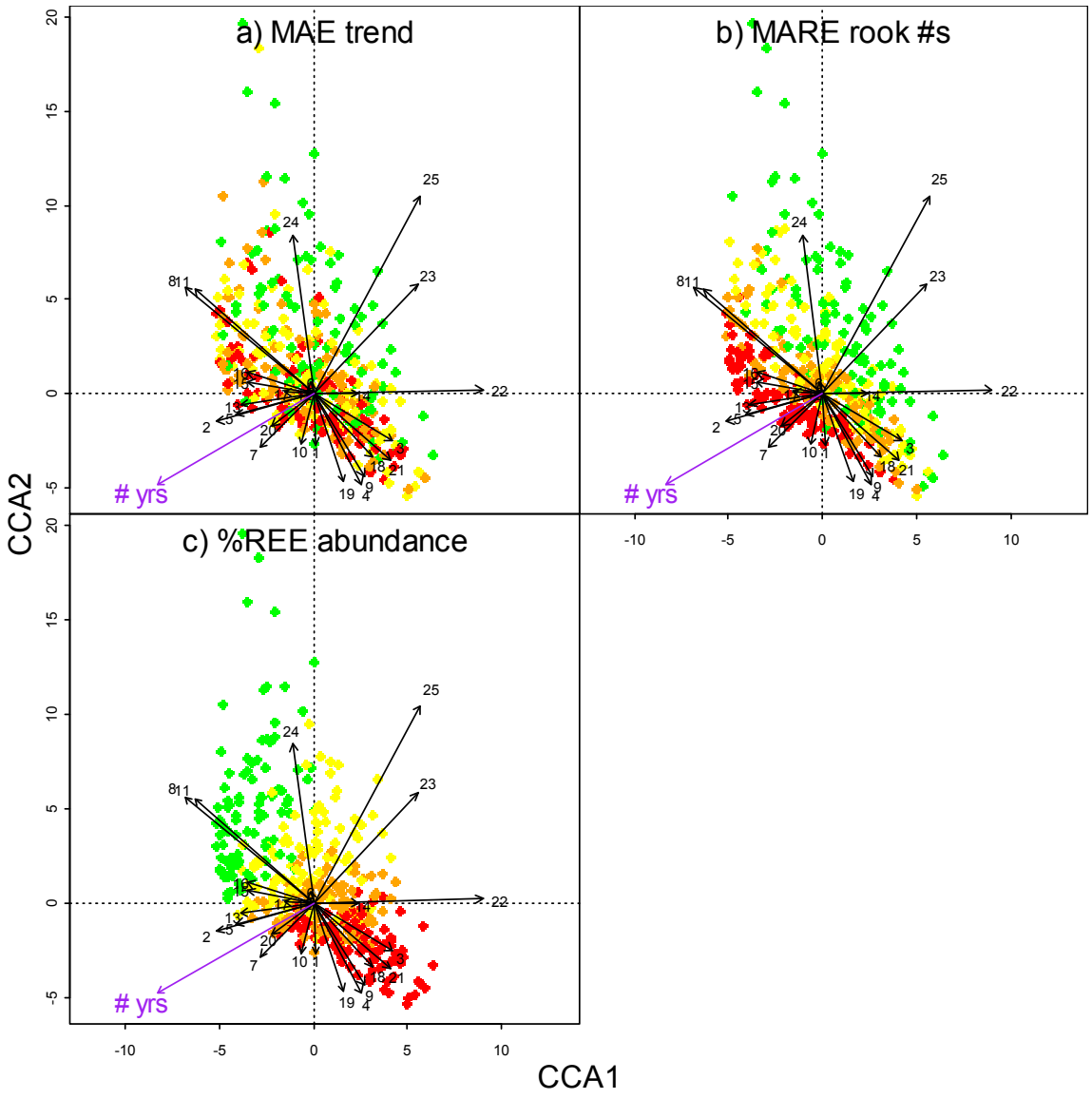


Figure 2.5. Results of the CCA analysis for Model #2 based on 90 data points. Colour of plotting characters reflects the magnitude of the absolute value for the appropriate performance metric (green indicates values close to zero – good performance, red indicates estimates far from zero – bad performance). Black arrows correspond to loadings on the axes of the coverage percentages for each simulation year, used to constrain the ordination. Purple '# yrs' arrow indicates the correlation of the summary statistic 'number of years with data' with the CCA axes.

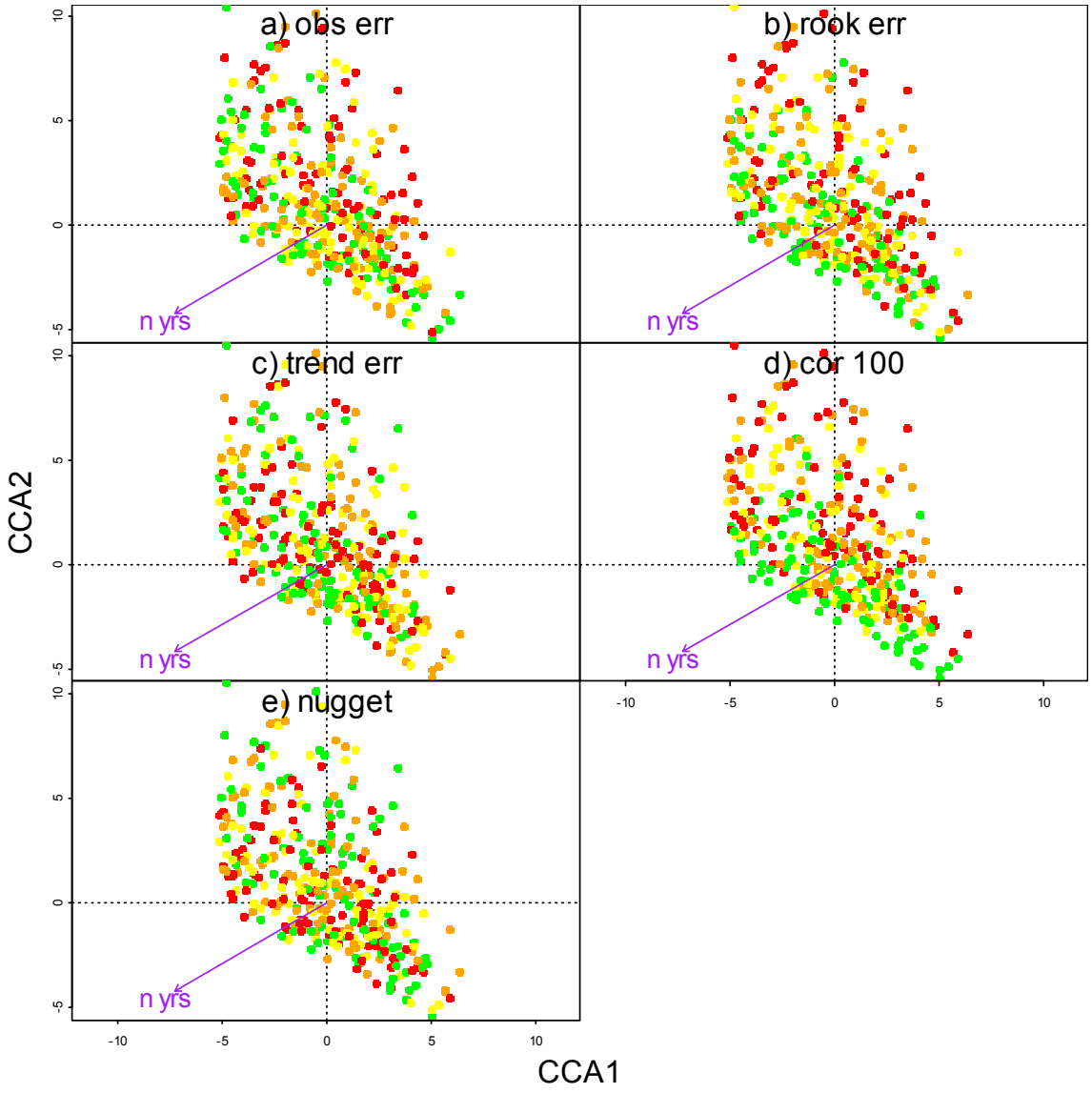


Figure 2.6. Results of the CCA analysis as described in Figure 2.5, but with colouring of data points related to the absolute value for the error of estimate for the model parameters.

Chapter 3: Characterising the spatial distribution of marine fish species using Bayesian hierarchical models

3.1. Introduction

The spatial distribution of marine species is driven by many factors, with aggregation behaviour leading to spatial autocorrelation in observations of marine fish abundance. The aggregation of species in response to habitat features and ecological interactions is a known driver of geographic distribution (MacArthur and Wilson 1967, Brown 1984, Levin 1992). Understanding the spatial distribution of prey species for marine predators such as Steller sea lions (*Eumetopias jubatus*) is key when considering the availability of prey. Characterisation of the spatial patterns of marine fish species, and the temporal variability inherent within these patterns is an important part of understanding how such species may interact with other ecosystem components, including predators.

Spatial autocorrelation in the distribution of abundance occurs because organisms aggregate and organise themselves in response to environmental and biological factors and covariates. Factors driving abundance of species can then be expected to be more similar at sites close to each other than at sites further apart because environmental characteristics and habitat are also spatially structured (Legendre 1993). Consequently, species abundance tends to be correlated in space. These habitat features are often dynamic in marine systems, as they reflect oceanographic variability, and so can be expected to be variable over both space and time (e.g. Bakun 1987, Rodríguez-Sánchez et al. 2002, Agostini et al 2006). The response of fish distribution to such oceanographic variability may not be well understood, so estimation of the types of spatial characteristics displayed by individual fish species from observed data can be useful for prediction.

Methods employed for characterising the properties of spatial distributions of fish have included geostatistical approaches such as variogram estimation and kriging (e.g. Petitgas 1993, Pelletier and Parma 1994, Rivoirard et al. 2000, Wieland and Rivoirard 2001, Mello and Rose 2005). These examples generally focus on using such methods to account for and/or remove spatial autocorrelation in observations when estimating abundance. The dangers of not accounting for spatial autocorrelation when analysing species abundance and distribution data are well recognized (Hoeting 2009). Consequently, it is desirable for a more formal incorporation of the implications of spatial processes into analyses.

In the ecological literature, a range of modelling methods has been applied to account for spatial autocorrelation, and are reviewed by Dormann et al. (2007). These methods include generalized linear models (GLMs, McCullagh and Nelder 1989), generalized additive models (GAMs, Hastie and Tibshirani 1990), generalized linear mixed effects models (GLMMs), and generalized additive mixed-effects models (GAMMs). Such models can take non-spatial covariates into account and also permit the use of non-Gaussian error structures during model fitting (often appropriate for over-dispersed count data). Of these, hierarchical modelling methods (including GLMMs) have become increasingly popular as methods for analyzing species distributional data (Cressie et al. 2009). When considering spatial autocorrelation, these methods have chiefly used random effects to account for spatial correlation among point-referenced data for fish and wildlife populations (e.g. Diggle et al. 1998, Agarwal et al. 2002, Thogmartin et al. 2004, Tvete et al. 2006), with a focus on determining the contribution of spatial processes to the observed data. Hierarchical methods are attractive in that prediction can take place at the level of the processes responsible for driving the data, while uncertainty about the between-group variability can be characterised.

This Chapter characterises the spatial distribution of nine fish groups (representing key Steller sea lion prey items) in the Gulf of Alaska (GoA) using a Bayesian hierarchical

modelling framework to fit GAMMs to trawl survey data. The mixed-effects modelling framework provides flexibility by allowing for inter-annual variability in the model parameters, and integrates the impact of spatial autocorrelation by including spatial random effects terms in the expected fish density. The posterior predictive distribution, obtained from the results, reflects the distribution from which unobserved data could be drawn, and can be used to characterise the spatial behaviour of the modelled fish species. The consequences of failing to account for spatial correlation are evaluated by comparing the results obtained from the spatial models to those from models when no spatial autocorrelation is considered.

3.2. Methods

GAMMs were fit to the available survey data for nine key Steller sea lion prey groups (Table 3.1). The prey groups were chosen because they represent the most frequently recorded components of sea lion diet in the Gulf of Alaska (Trites and Donnelly 2003, Sinclair et al. 2005). Separate analyses were conducted for each species; while this aided computation, it precluded an ability to account for interactions among species when predicting prey density in unobserved years. A Bayesian hierarchical estimation framework was used to make inference on the mean density of species and the parameters determining the spatial distribution for an 'unobserved' year. This section continues by first briefly describing the data available for model-fitting purposes, and then detailing the modelling framework.

3.2.1. Data

The data used were from the Gulf of Alaska bottom trawl survey (e.g. Wilkins 2007) for the years 1984, 1987, 1990, 1993, 1996, 1999, 2001, 2003, 2005, and 2007. The data consisted of the catch per unit effort (cpue) (in $kg.ha^{-1}$) of each prey group by haul for each year, with accompanying information on the depth and geographic location of each

haul. Data were limited to hauls between l45°W and 160°W (Figure 3.1), to make inference for the Gulf of Alaska (changes in oceanographic domain could lead to marked differences in spatial behaviour and density at gross regional scales), and to provide manageable computation times. Appropriate depth restrictions were also imposed on the data for each prey group (Table 3.1). The number of hauls available per year for each prey group ranged from 360 to 481.

3.2.2. Modelling framework

The observed cpue of prey group *j* at a given location (*s*) in year *y* was assumed to follow a zero-inflated negative binomial process, with the mean of both the negative binomial and the probability of an additional zero being modelled as functions of the mean density in year *y*, the depth at location *s*, and a stochastic component that introduces spatial autocorrelation. Zero-inflated models that account for the presence of additional zeroes above the expectation of distributions such as the Poisson and negative binomial have become prevalent in the literature as data frequently suggest this additional over-dispersion. Ainsworth (2007) details zero-inflation in the context of spatial modelling of count data.

The choice to use GAMMs to fit to the data described in Section *3.2.1* reflects the expectation that important covariates affecting fish density such as depth are likely to show a non-linear response. The cpue data for each species described in Section *3.2.1* were modelled as:

$$Y_y(\mathbf{s}_i) \sim \begin{cases} 0 & \text{with probability } \; p_{y,\mathbf{s}} \\ NegBin\left(\lambda_{y,\mathbf{s}}, \alpha_y\right) & \text{with probability } \; 1 - p_{y,\mathbf{s}} \end{cases} \tag{3.1}$$

where $Y_y(\mathbf{s}_i)$ is the observed cpue in year *y* at location $\mathbf{s_i}$, $p_{y,\mathbf{s}_i}$ is the probability of a non-structural zero density at location $\mathbf{s_i}$ in year *y*, $\lambda_{y,\mathbf{s}_i}$ is the mean of the negative

binomial distribution of density at location $\mathbf{s_i}$ in year y, and α_y is the overdispersion parameter for the negative binomial distribution in year y.

Zero-inflation was modeled by assuming that a latent binary process (e.g. Agarwal et al. 2002, Mendes et al. 2007) occurs at each observation site and takes the value 0 if the observation $Y(\mathbf{s})$ is a structural zero, or 1 if $Y(\mathbf{s})$ is an observation coming from a negative binomial distribution. As with the mean of the negative binomial distribution, the zero-inflation process is governed by a spatial process which determines the probability of a non-structural zero.

The $p_{y,\mathbf{s}}$'s and $\lambda_{y,\mathbf{s}}$'s in Equation 3.1 were modelled as functions of the location-specific depth and a gaussian spatial process, with these functions being additive with suitable link functions:

$$\begin{aligned} \text{logit}\left(p_{y,\mathbf{s}}\right) &= \overline{p}_y + \beta^p x_y(\mathbf{s}_i) + \sum_{k=1}^{K} Z_{x_i,k} b_k^p + w_{y,\mathbf{s}}^p \\ \log_e\left(\lambda_{y,\mathbf{s}}\right) &= \overline{\lambda}_y + \beta^\lambda x_y(\mathbf{s}_i) + \sum_{k=1}^{K} Z_{x_i,k} b_k^\lambda + w_{y,\mathbf{s}}^\lambda \end{aligned} \tag{3.2}$$

where $\overline{p}_y$ is the covariate-unadjusted mean for a non-structural zero in year y, with hyperprior $\overline{p}_y \sim \text{N}\left(\mu_p, \sigma_p^2\right)$, $\overline{\lambda}_y$ is the covariate-unadjusted mean for the density in year y, with hyperprior $\overline{\lambda}_y \sim \text{N}\left(\mu_\lambda, \sigma_\lambda^2\right)$, $w_{y,\mathbf{s}}^{p/\lambda}$ is the spatial error term at location $\mathbf{s}$ in year y, where:

$$\begin{aligned} \text{E}\left[w_{y,\mathbf{s}}^{p/\lambda}\right] &= 0 \\ Cov\left[w_{y,\mathbf{s}}^{p/\lambda}, w_{y,\mathbf{s}'}^{p/\lambda}\right] &= \left(\sigma_{w,y}^{p/\lambda}\right)^2 \exp\left(-\phi_y^{p/\lambda}\left|\mathbf{s}-\mathbf{s}'\right|\right) \end{aligned} \tag{3.3}$$

$\left(\sigma_{w,y}^{p/\lambda}\right)^2$ is the variance of the spatial process determining the $w_{y,\mathbf{s}}^{p/\lambda}$'s in year *y*, $\phi_y^{p/\lambda}$ is the range parameter of the spatial correlation function in year *y*, $x_y(\mathbf{s}_i)$ is the depth at location $\mathbf{s}_i$, $\beta_y^{p/\lambda} b_1^{p/\lambda},...,b_K^{p/\lambda}$ are the regression coefficients, with $b_k^{p/\lambda} \sim N\left(0,\sigma_{b,p/\lambda}^2\right)$, $\mathbf{Z}_{\mathbf{x}}$ is the depth spline matrix,

$$\mathbf{Z}_{\mathbf{x}} = \left[\begin{array}{c} \left|x_y(\mathbf{s}_i) - \kappa_k\right|^3 \\ 1 \le k \le K \end{array}\right]\left[\begin{array}{c} \left|\kappa_k - \kappa_{k'}\right|^3 \\ 1 \le k, k' \le K \end{array}\right]^{-1/2} \tag{3.4}$$

and $\kappa_1 < \kappa_2 < \kappa_k$ are fixed knots of the penalized splines for the effect of depth on the mean,

The spatial processes determining the $w_{y,\mathbf{s}}^{p/\lambda}$ for both the mean intensity and the zero-inflation were assumed to be stationary Gaussian random fields, independent of one another and in time, conditional on $\left(\sigma_{w,y}^{p/\lambda}\right)^2$ and $\phi_y^{p/\lambda}$. The parameters determining the covariance among observations were allowed to vary among years, as the degree of spatial correlation could be expected to change over time. Separate values for the spatial parameters were estimated for the zero-inflation and mean intensity. While allowing these parameters to vary complicates estimation, the data can then be used to determine the evidence for inter-annual variability, and reduces chances for overfitting. The annual values for $\left(\sigma_{w,y}^{p/\lambda}\right)^2$, $\phi_y^{p/\lambda}$, and α_y are treated as gamma-distributed random effects, with the parameters of the gamma-distributed hyperpriors (the mean and variance) being the quantities of interest, in that they provide information on the values of these quantities for an 'unobserved' year. Table 3.2 lists the distributional assumptions for these various time-varying parameters.

The effects of depth on the means of the probability of zero-inflation and the negative binomial were modelled as cubic thin-plate splines, with three knots each (Equation 3.4; c.f. Crainiceanu et al. 2007). Three knots appeared sufficiently flexible to capture the pattern of the depth effects obtained when comparing simpler models. The mixed-effects modelling framework permits the incorporation of penalized spline regression (e.g. Zhao et al. 2006, Crainiceanu et al. 2007). The GAM components ($b_1^{p/\lambda},...,b_K^{p/\lambda}$) were treated as random effects, with the degree of smoothness for each of the two functions estimated as the variance of the random effects for the knots, $\sigma^2_{b,p/\lambda}$.

The spatial autocorrelation terms ($w_{y,\mathbf{s}}^{p/\lambda}$) were defined as following a multivariate normal distribution with variance-covariance matrix $\mathbf{W}_y^{p/\lambda}$ (Equation 3.3), using an exponential form for the correlation function. Exponential forms for the spatial correlation function are the most frequently used in the literature, often produce results very similar to those resulting from other more complicated functions, and provide a somewhat simpler computational approach for estimation purposes. While the treatment of the approach for point-referenced data is intuitive, repeated evaluation of full conditional density functions during Bayesian estimation for large spatial data sets can be unstable, and computationally infeasible (e.g. Banerjee et al. 2005). In order to combat this, a subsampling strategy was incorporated into the model fitting approach. Using ideas proposed by Pardo-Igúzquiza and Dowd (1997), the density for $\mathbf{W}_y^{p/\lambda}$ was replaced by:

$$\prod_{i=1}^{n_y} f\left(w_{y,\mathbf{s}_i}^{p/\lambda} \mid w_{y,\mathbf{s}_j}^{p/\lambda}, \mathbf{s}_j \in \partial\mathbf{s}_i\right) \tag{3.5}$$

where $\partial\mathbf{s}_i$ is the set of 10 observations that are closest to $\mathbf{s}_i$. Pardo-Igúzquiza and Dowd (1997) recommend a fixed neighborhood size as opposed to the set of observations within

a fixed distance, and that 10-15 neighbors is suitable. The spatial random effects were therefore distributed as:

$$
\begin{aligned}
w_{y,\mathbf{s}_i}^{p/\lambda} &\sim N\left(\mu_{w,\mathbf{s}_i}^{p/\lambda}, \left(\sigma_{w,y}^{p/\lambda}\right)^2 \varsigma_{w,\mathbf{s}_i}^{p/\lambda}\right) \\
\mu_{w,\mathbf{s}_i}^{p/\lambda} &= \mathbf{v}_{s_i}^T \boldsymbol{\Sigma}_{s_i}^{-1} \mathbf{w}_{y,\partial\mathbf{s}_i}^{p/\lambda} \\
\varsigma_{w,\mathbf{s}_i}^{p/\lambda} &= 1 - \mathbf{v}_{s_i}^T \boldsymbol{\Sigma}_{s_i}^{-1} \mathbf{v}_{s_i} \\
\mathbf{v}_{s_i} &= \exp\left(-\phi_y^{p/\lambda} \mathbf{d}(\partial\mathbf{s}_i)\right) \\
\boldsymbol{\Sigma}_{s_i} &= \exp\left(-\phi_y^{p/\lambda} \mathbf{D}(\partial\mathbf{s}_i)\right)
\end{aligned}
\tag{3.6}
$$

where $\mathbf{w}_{y,\partial\mathbf{s}_i}^{p/\lambda}$ is the vector of spatial random effects corresponding to the observations in $\partial\mathbf{s}_i$, $\mathbf{d}(\partial\mathbf{s}_i)$ is the vector of distances between $\mathbf{s}_i$ and the members of $\partial\mathbf{s}_i$, and $\mathbf{D}(\partial\mathbf{s}_i)$ is the matrix of distances among the members of $\partial\mathbf{s}_i$.

3.2.3. Estimation

The estimated parameters of the model, along with their assigned prior distributions, are listed in Table 3.2. The estimated parameters included those of the hyperpriors mentioned above, as well as the parameters of the depth spline functions, $\beta^{p/\lambda}$ and $\sigma_{b,p/\lambda}^2$. Markov-Chain Monte Carlo (MCMC) techniques, implemented in WinBUGS (Spiegelhalter et al. 2000), were used to generate 150,000 samples from the posterior distribution. The first 50,000 of these samples were discarded to allow for "burn-in", and the remaining samples thinned to obtain a posterior sample of size 1,000. The resulting thinned chain was then assessed for lack of convergence using appropriate diagnostic statistics (statistics included those developed by Geweke (1992), Heidelberger and Welch (1983), and Raftery and Lewis (1992), and the extent of auto-correlation among the samples in the chain).

Weakly informative normal priors were assumed for the process means μ_p and μ_a (Table 3.2). The gamma-distributed hyperpriors governing the annual values for the over-dispersion, spatial variance, and spatial correlation terms were parameterized in terms of the mean and variance. Priors for the variance parameters were defined as uniform on the value for the standard deviation (Table 3.2, Gelman and Hill 2007). The mean of the annual over-dispersion parameters, μ_α, was assigned a weakly informative gamma distributed prior. For computational purposes, uniform priors were placed on the inverses of the means of the range of the spatial correlations, $\zeta_{\phi,1}^{p/\lambda}$ (e.g. Banerjee et al. 2005). The bounds on the priors for $1/\zeta_{\phi,1}^{p/\lambda}$ were selected so that, for a distance of 1km between locations, the minimum possible correlation was 5%, and that at a distance of 300 km between locations, the maximum allowed extent of correlation was 5% (Table 3.2).

To compare the results of the full model with those of analyses that did not account for spatial correlation, a version of the model that excluded the spatial random effects $w_{y,\mathbf{s}}^{p/\lambda}$ in the equations for the means (Equation 3.2) was also fit to the data.

3.3. Results

There was considerable variability among prey groups in the relative contributions of the spatial processes to the overall variability in both the mean density, $\lambda_{y,s}$, and the mean probability of a structural zero, $p_{y,s}$ (Table 3.3). Perhaps unsurprisingly given the data set used, the results for herring, salmon, smelt, eulachon and squid showed very high

levels of zero-inflation (μ_p^*[3], the probability of an 'extra' zero, is close to 1). Results for flatfish, cod, and both juvenile and adult pollock suggested on average almost no zero-inflation (Table 3.3, column μ_p^*). The 95% credible interval for μ_p^* was narrow for almost all prey groups, with only the results for eulachon, suggesting a moderate amount of uncertainty in the probability of zero-inflation at the average depth. Inter-annual variability in $\overline{p}_y$ was low for most species (Figure 3.2, Table 3.3 column σ_p^2), though the posterior for juvenile pollock suggested high or uncertain probabilities of structural zeroes for 1984, 1987, and 2005 (Figure 3.2). The proportion of additional zeroes for adult pollock and salmon also differed from the average in one or more years (Figure 3.2).

The posterior distributions for the annual values for the expected density given that an observation is not a structural zero (i.e. the mean (unadjusted by depth) of the negative binomial distribution), the exponential of $\overline{\lambda}_y$, showed more inter-annual variability than those for $\overline{p}_y$ (Figures 3.2 and 3.3). The results suggested a fairly high degree of temporal autocorrelation in the mean density, with the posteriors for density for several prey groups, in particular adult pollock, showing trends with time. The estimates for the density of herring in 1984 was higher and highly uncertain compared to the estimates for other years (Figure 3.3).

An example of the inter-annual variability in the estimated spatial distribution of fish density is shown in Figures 3.4 and 3.5, which display the posterior medians for 2001 and 2007, for the area around the Marmot Island Steller sea lion rookery off the east coast of

---

[3] μ_p^* is the transformed μ_p, in that $\mu_p^* = \exp(\mu_p) / \left(1 + \exp(\mu_p)\right)$

Kodiak Island. There are distinct differences in the estimated locations of high density areas between the two years, particularly for juvenile pollock, cod, salmon, and smelt. The median of the posterior predictive distribution for the same area, shown in Figure 3.6, retains some of the depth dependence in density evident from the results in Figures 3.4 and 3.5. The median values for the density from an 'unobserved' year (Fig. 3.6) do not show the spatial coherence of Figures 3.4 and 3.5, further indicating the variability in spatial distribution associated with the results (Fig. 3.6 is, effectively, a summary of many realizations of the spatial distribution, whereas Figs 3.4 and 3.5 summarise one realization each). In addition to the inter-annual differences in spatial distribution, Figures 3.4 and 3.5 also further demonstrate the inter-annual variability in the realized values for fish density, the means for which were shown in Figure 3.3.

Figure 3.7 summarizes how the posterior predictive distributions for the expected density change with depth for the different prey groups. These distributions take into account the impact of both inter-annual variability ($\sigma^2_{p/\lambda}$) and spatial variance $\left(\sigma^{p/\lambda}_{w,y}\right)^2$ in both the probability of a structural zero and the mean density, and show the impact of the GAM component on the model results. Results for all species are highly over-dispersed, with the positive tail of the distribution encompassing high values for density. Herring, smelt, and salmon show density declining with depth, with squid having the opposite trend (Fig. 3.7). The results for the other prey groups show a non-linear pattern for the change in expected density with depth, with a peak in density at an intermediate depth, and decreases in density at shallower and deeper depths. The changes in density with depth for herring and smelt are driven primarily by the probability of a structural zero, with almost no probability (but high uncertainty) of zero-inflation at depths less than 50 m, and a tight credibility interval around a probability of 1 for zero-inflation at greater depths.

The contribution of the spatial processes to the variability in the model predictions can be viewed by comparing the predictive distributions for the spatial variance $\left(\sigma_w^{p/\lambda}\right)^2$, to the posteriors for the inter-annual variance ($\sigma_{p/\lambda}^2$). The contribution of the mean spatial variance to the overall variance in mean density was less than that of inter-annual variance for five prey groups, and on the same order of magnitude for the others (Table 3.3, columns $\zeta_{w,1}^{\lambda}$ and σ_{λ}^2). Even when accounting for the inter-annual variability in the spatial variance ($\zeta_{w,2}^{\lambda}$), the contribution of the spatial process to the variability of the linear predictor was less or comparable to the inter-annual variability in the mean (Table 3.3, compare columns $\left(\sigma_{w,y}^{\lambda}\right)^2$ and σ_{λ}^2). In contrast, the contribution of the spatial process to the total variance in the linear predictor for the zero-inflation process was much greater than that for the mean density, with posterior predictive distributions for $\left(\sigma_w^{p}\right)^2$ having higher central tendencies, and encompassing larger ranges of values than those for σ_p^2 (Table 3.3, columns $\left(\sigma_{w,y}^{\lambda}\right)^2$ and σ_p^2). However, the logit transformation associated with the zero-inflation process meant that the influence of this variability on the mean probability of a structural zero was less than appears from the values for σ_p^2 and $\zeta_{w,1}^{p}$. This is mainly because the estimates of the zero-inflation probability are close to zero or one for all prey groups.

The distributions for the parameters governing the extent of spatial correlation (Figure 3.8), suggest appreciable uncertainty for almost all prey groups, particularly at the short distance of 2 km. Higher values for estimates of spatial correlation in the mean density than that in the structural zeroes were obtained for juvenile pollock, cod, and squid (Fig. 3.8), indicating that the density for these prey groups tended to be very similar among patches (given the same depth). The annual values for the spatial correlation terms were very similar to the posterior predictive distributions (Fig. 3.8, first column, all panels),

with little evidence for inter-annual variability. This is reflected in the summaries for the posterior distributions for the variance of the spatial range hyperpriors, $\zeta_{\phi,2}^{p/\lambda}$ in Table 3.3, which show low values for most prey groups. There was greater inter-annual variability in the spatial correlation parameters for the structural zeroes than the mean density (Fig. 3.8, compare shaded boxes [probability of an extra zero] for pollock and herring with unshaded boxes [mean density]). While there is uncertainty regarding the value for the $\phi_y^{p/\lambda}$'s, the distributions for these parameters are generally in areas of parameter space that suggest an appreciable degree of spatial autocorrelation at small spatial scales of less than 10km (Figure 3.8).

Not including spatial autocorrelation during the model fitting (i.e. setting $w_{y,\mathbf{s}}^{p/\lambda}$=0) gave different results compared to those for the full spatial model. In general, estimates of mean density were higher for all prey groups (besides herring) under the non-spatial model (Figure 3.3, blue bars). In addition, estimates of the probability of a structural zero tended to be less extreme than those for the spatial model (estimates less likely to be zero or one), as evidenced by the posteriors for the model parameters (Table 3.4) and the posteriors for the annual mean values for these probabilities (Figure 3.2, blue boxplots). The results for herring and flatfish were exceptions, with estimates mimicking those obtained for the spatial model (Fig 3.2). The estimates for the over-dispersion parameter for the negative binomial of fish density were greater for the non-spatial model than the spatial analyses (compare column α_y in Tables 3.3 and 3.4).

3.4. Discussion

Consistent notable differences were observed between the results obtained when spatial autocorrelation is accounted for than when it is not. A key result was that the non-spatial analyses led to higher estimates of density for almost all prey groups than did the spatial

analyses. The methods described here could be used as an alternative way of estimating abundance for the given species, as they account for temporal variability and the impact of depth on density, both of which are included in methods used by the NMFS to estimate abundance. Given the results obtained here, accounting for spatial autocorrelation is preferable and would lead to lower abundance estimates for these species groups.

The results of the spatial analyses showed appreciable differences in the influence of spatial autocorrelation by prey group. All prey groups displayed some degree of spatial autocorrelation in the density at the scale of a few kilometres (Fig. 3.8). There were also differences by prey group in the degree to which parameters governing the spatial processes changed over time. The inter-annual variability in the amount of spatial autocorrelation was generally low (Fig. 3.8, Table 3.3 columns $\zeta^{p}_{\phi,2}$ and $\zeta^{\lambda}_{\phi,2}$). The impact of the spatial process on the variability of the prey density however, did vary from year to year (Table 3.3, columns $\zeta^{p}_{w,2}$ and $\zeta^{\lambda}_{w,2}$). This suggests that the spatial extent of the factors responsible for organizing say, school size, are reasonably consistent from one year to the next, but the response of fish to these factors varies inter-annually.

Separate analyses were conducted for each species. Consequently, no correlations among species in density at locations were explicitly accounted for (but could, of course, arise anyway). While this assumption eased the computation and complexity of the estimation framework, it means that predicted distributions of sea lion prey density using these models could have more locations with diverse species compositions than observed in the data. In theory, it is feasible to conduct the analyses presented here on a multispecies basis, with covariance matrices describing the relationship among the likely densities of prey groups at a given observation site. However, the additional computation required would make these analyses quite prohibitive given the size of the data set. Alternative methods for dealing with spatial autocorrelation, such as multivariate CAR models (e.g. Carroll and Johnson 2008, Dormann et al. 2007), may need to be used. A simpler method

of accounting for the presence / absence of other prey groups might also be to include the other prey groups as additive covariates in the linear predictors (Equation 3.2).

The GAMMs presented here only included depth as a covariate affecting density besides the spatial processes, despite the survey data including more information (time of day, oceanographic variables, etc.). The decision not to include other covariates reflected the results of both preliminary analyses of the data and a desire to only consider physical non-changing factors to make predictions for the desired application useful. The estimated spatial autocorrelation terms $w_{y,\mathbf{s}}^{p/\lambda}$, and the parameters determining the values for these will therefore integrate the variability of the effects of other covariates on fish density, and so implicitly account for correlation in the data due to some of these covariates (which, for instance, sea surface temperature, also exhibit spatial structure) without explicitly modelling the relationships. This is advantageous for the application intended here (Chapter 4), because it avoids the need to pre-specify values for these additional covariates when making predictions. Integration of other covariates into the analyses might be useful when using these methods for abundance estimation. However, the methods in this Chapter would be useful when information from spatially structured covariates is not available.

The catchability of the survey is not likely to be the same for the different prey groups, as the data are from bottom trawls. Pelagic 'forage' fish may likely be under-represented in the survey – however, the presence of these prey groups in the data (albeit at low density) does give evidence for the patchy nature of these species' distributions. The data used in these analyses are by no means the only data available for this region, and, in principle, it should be possible to assimilate data from acoustic or pelagic trawl surveys (or indeed from fisheries) into the estimation framework.

Convergence problems for some of the prey groups were noted, particularly for the variance of the distribution for mean variance in the spatial random effects ($\zeta^{p}_{w,2}$, the inter-annual variability in the spatial variance, $\left(\sigma^{p}_{w,y}\right)^2$) for the zero-inflation component. This occurred most often for the prey groups with high probabilities (almost equal to one) of structural zeroes, and is likely due to confounding with the mean probability because of the logit transformation (high values for the variance were associated with very high (and very low) values for the untransformed mean of the linear predictor). That said, the contribution of these spatial variance terms did result in changes to the model predictions in some cases (i.e. spatial variability allowed for the occurrence of regions where the mean probability of zero-inflation was much higher or lower than the mean), which, given the posterior distributions plotted in Figures 3.4 and 3.5, appears reasonable.

Incorporating the spatial autocorrelation processes into the linear predictors of the GAMMs (Equation 3.2) implies an expectation that the mean density at neighbouring locations will be similar to each other, rather than the actual observations (Banerjee et al. 2005). This suggests that, when making predictions from the posterior, high values for say the over-dispersion parameter would likely reduce the influence of the spatial correlation. The low values for the mean overdispersion (μ_α) obtained from the spatial model (relative to those from the non-spatial model) indicates that this was probably not an issue.

Mendes et al. (2007) incorporated spatial autocorrelation into a Bayesian GLMM using a zero-inflated negative binomial model for hake recruits in Portugal. The approach presented here, where separate Gaussian random fields were assumed to be governing the zero-inflation and mean density processes, differs from that used by Mendes et al. (2007), and that of Tvete et al. (2006), who assumed that the same spatial process affected both components. While it might be reasonable to expect that areas for which the probability of an extra zero is higher than the mean should also suggest a lower than average mean

density, the converse, that areas for which the probability of an 'extra' zero are low should be expected to have higher than average density, is not necessarily intuitive. Examination of the degree of spatial correlation in the residuals obtained from fitting simple GLMs to the cpue data suggested that the modelling of separate spatial processes and inter-annual variability in the over-dispersion and spatial parameters was warranted. The posteriors for the spatial correlation of the two processes suggest that this was an apt choice for some prey groups (Fig. 3.8, juvenile pollock, adult pollock, cod, herring, and squid).

The intention of the analyses was to produce simulated sea lion prey fields that exhibited the same spatial characteristics as those observed in the survey data. The spatial modelling approach described above captures variability in the spatial distribution of prey, but does not provide insight into how these spatial distributions evolve over time, particularly at short time scales corresponding to individual or several foraging trips. In effect, some of the variation accounted for within these analyses encompasses that due to small-scale temporal changes in abundance; hauls taken in a given survey year are of course not conducted at exactly the same time. However, due to the necessities of survey design, hauls in spatial proximity to each other are generally also conducted at similar times. Therefore, at a scale on the order of sea lion foraging trips, which can last several days, the spatial autocorrelation in prey density accounted for in the estimation framework presented here likely captures the desired properties.

Table 3.1. Steller sea lion prey species groups used, with subsetting restrictions of the available survey data.

Prey groups	species	depth restriction	additional restrictions
Juvenile pollock	walleye pollock (*Theragra chalcogramma)*	< 200 m	avg wt ≤ 0.184 kg
Adult pollock	walleye pollock (*Theragra chalcogramma)*	< 350 m	avg wt > 0.184 kg
Cod	Pacific cod (*Gadus macrocephalus*)	< 350 m	
Salmon	Pacific salmon sp. (*Onchorynhcus sp.*)	< 250 m	
Herring	Pacific herring (*Clupea pallasi*)	< 300 m	
Eulachon	eulachon (*Thaleichthys pacificus*)	< 300 m	
Smelts	capelin (*Mallotus villosus*) Pacific sand lance (*Ammodytes hexapterus*)	< 300 m	
Flatfish	arrowtooth flounder (*Atheresthes stomias*) rock sole (*Lepidopsetta sp.*)	< 350 m	
Squid		< 500 m	

Table 3.2. Prior distributions for the model parameters, and hyperpriors for the random effects.

Parameter	Description	
Fixed effects		Prior
μ_p	mean of the distribution for the covariate-unadjusted mean for the probability of a non-structural zero	$Normal\left(0,1000\right)$
σ_p^2	variance of the distribution for the annual deviations in the covariate-unadjusted mean for the probability of a non-structural zero	$\left[Uniform\left(0,1000\right)\right]^2$
μ_λ	mean of the distribution for the covariate-unadjusted mean for the density	$Normal\left(0,1000\right)$
σ_λ^2	variance of the distribution for the annual deviations in the covariate-unadjusted mean for the density	$\left[Uniform\left(0,1000\right)\right]^2$
μ_α	mean of the distribution for the overdispersion parameter for the negative binomial distribution	$Gamma\left(0.1.0.1\right)$
σ_α^2	variance of the distribution for the over-dispersion parameter	$\left[Uniform\left(0,1000\right)\right]^2$
$\beta^{p/\lambda}$	mean trend of density with depth	$Normal\left(0,1000\right)$
$\sigma_{b,p/\lambda}^2$	variance of the distribution for the depth-related random effects	$\left[Uniform\left(0,1000\right)\right]^2$
$\zeta_{w,1}^{p/\lambda}$	mean of distribution for the variance for the spatial random effects	$\left[Uniform\left(0,1000\right)\right]^2$
$\zeta_{w,2}^{p/\lambda}$	variance of distribution for the variance for the spatial random effects	$\left[Uniform\left(0,1000\right)\right]^2$
$1/\zeta_{\phi,1}^{p/\lambda}$	inverse of the mean of the distribution for the range of the spatial correlation	$Uniform\left(0.334,99.499\right)$
$\zeta_{\phi,2}^{p/\lambda}$	variance of the distribution for the range of the spatial correlation	$\left[Uniform\left(0,1000\right)\right]^2$
Year-specific random effects		Hyperprior
$\overline{p}_y$	logit-transformed covariate-unadjusted mean for the probability of a non-structural zero	$Normal\left(\mu_p,\sigma_p^2\right)$
$\ln\left(\overline{\lambda}_y\right)$	natural logarithm of the covariate-unadjusted mean density in year *y*	$Normal\left(\mu_\lambda,\sigma_\lambda^2\right)$
α_y	over-dispersion parameter	$Gamma\left(shape=c_\alpha, rate=d_\alpha\right)$ $c_\alpha=(\mu_\alpha)^2/\sigma_\alpha^2, d_\alpha=\mu_\alpha/\sigma_\alpha^2$
$\left(\sigma_{w,y}^{p/\lambda}\right)^2$	variance of spatial random effects	$Gamma\left(shape=c_w^{p/\lambda}, rate=d_w^{p/\lambda}\right)$ $c_w^{p/\lambda}=(\zeta_{w,1}^{p/\lambda})^2/\zeta_{w,2}^{p/\lambda}, d_w^{p/\lambda}=\zeta_{w,1}^{p/\lambda}/\zeta_{w,2}^{p/\lambda}$
$\phi_y^{p/\lambda}$	range of spatial correlation	$Gamma\left(shape=c_\phi^{p/\lambda}, rate=d_\phi^{p/\lambda}\right)$ $c_\phi^{p/\lambda}=(\zeta_{\phi,1}^{p/\lambda})^2/\zeta_{\phi,2}^{p/\lambda}, d_w^{p/\lambda}=\zeta_{\phi,1}^{p/\lambda}/\zeta_{\phi,2}^{p/\lambda}$
$b_k^{p/\lambda}$	depth random effects	$Normal(0,\sigma_{b,p/\lambda}^2)$

Table 3.3. Posterior medians and 95% credibility intervals for the fixed effects model parameters for the full spatial model.

prey group	μ_p^*	σ_p^2	β_j^p	$\sigma_{b,p}^2$	$\zeta_{\phi,1}^p$	$\zeta_{\phi,2}^p$	ϕ_y^p	$\zeta_{w,1}^p$
juvenile	1.4e-11	0.032	-4.42	9.60	0.712	0.756	0.436	36.8
pollock	(0,8e-04)	(0.014,0.09)	(-15.5,22.1)	(0.17,65.26)	(0.133,2.72)	(0.072,3.78)	(4.9e-06,5.0)	(12.0,49.8)
adult	0.0272	0.60	-10.293	9.8	0.441	0.179	0.25	4.22
pollock	(0.002,0.23)	(0.25,1.33)	(-15.6,0.81)	(2.4,54.2)	(0.056,2.73)	(0.012,3.26)	(1.e-06,3.9)	(2.10,15.33)
	4.3e-09	0.123	1.91	2.33	1.049	0.374	1.017	25.37
cod	(0,0.02)	(0.038,0.948)	(-2.46,13.4)	(0.29,16.65)	(0.261,2.94)	(0.022,2.48)	(0.027,4.02)	(7.03,54.81)
	1	0.70	3.576	1.99	0.391	0.407	0.178	5.96
herring	(1,1)	(0.25,4.90)	(0.377,8.85)	(0.48,9.61)	(0.073,2.46)	(0.027,3.54)	(2e-07,3.79)	(3.79,10.74)
	0.98	0.54	-1.17	0.75	0.552	0.226	0.429	5.82
salmon	(0.90,1.00)	(0.17,2.83)	(-8.45,1.20)	(0.13,9.20)	(0.097,2.70)	(0.016,2.88)	(4.4e-05,3.4)	(2.99,22.35)
	1	0.120	-2.93	2.75	0.885	0.526	0.7536	23.35
smelt	(1,1)	(0.022,1.93)	(-13.7,5.39)	(0.32,23.48)	(0.194,2.80)	(0.033,3.17)	(0.004,3.91)	(5.82,75.73)
	0.99	0.47	11.54	2.81	0.4603	0.159	0.378	8.84
eulachon	(0.69,1.00)	(0.12,14.17)	(2.81,18.24)	(0.18,14.98)	(0.0814,2.65)	(0.015,2.33)	(2e-06,3.36)	(5.01,21.18)
	0.97	1.41	5.58	1.04	0.912	0.410	0.729	2.263
squid	(0.91,1.00)	(0.45,6.94)	(3.12,17.82)	(0.19,8.17)	(0.174,2.68)	(0.050,2.58)	(3.1e-04,3.9)	(0.874,8.44)
	3.7e-10	0.691	5.17	6.4	0.913	0.379	0.791	6.93
flatfish	(0,6.6e-04)	(0.072,4.247)	(1.14,24.1)	(1.3,55.8)	(0.189,2.81)	(0.053,2.85)	(0.016,4.17)	(1.25,36.38)

Table 3.3 continued

prey group	$\zeta^{p}_{w,2}$	$\left(\sigma^{p}_{w,y}\right)^2$	exp(μ_λ)	σ^2_λ	β^λ	$\sigma^2_{b,\lambda}$	$\zeta^{\lambda}_{\phi,1}$	$\zeta^{\lambda}_{\phi,2}$
juvenile pollock	436.1 (45.5,969)	29.06 (2.39,93.78)	0.0902 (0.006,0.549)	1.8 (0.6,12.4)	-4.02 (7.44,-1.05)	1.81 (0.47,10.15)	0.059 (0.025,0.29)	0.035 (0.011,0.33)
adult pollock	23.07 (4.16,481.2)	2.77 (0.0418,31)	6.57 (3.48,12.12)	1.9 (0.9,20.3)	0.7317 (-0.10,1.55)	0.202 (0.049,1.26)	0.95 (0.16,2.82)	0.471 (0.019,3.07)
cod	323.6 (41.5,960)	20.4 (0.373,86.2)	18.3 (11.5,28.8)	4.2 (1.8,12.7)	-0.279 (-0.84,0.18)	0.396 (0.086,2.59)	0.182 (0.065,2.78)	0.114 (0.014,2.68)
herring	8.55 (1.55,44.53)	5.33 (1.43,16.40)	0.212 (0.078,0.879)	0.70 (0.26,6.59)	-0.993 (-1.87,-0.22)	0.223 (0.022,1.89)	1.03 (0.23,2.80)	0.434 (0.037,3.42)
salmon	42.95 (1.73,555)	4.38 (0.074,39.8)	1.5 (1,2.3)	2.2 (1.1,5.4)	-0.139 (-0.47,0.05)	0.089 (0.016,0.60)	0.33 (0.11,2.66)	0.242 (0.024,3.18)
smelt	287.4 (12.6,939.1)	22.67 (1.75,99.09)	0.029 (0.016,0.063)	2.87 (0.98,16.07)	-0.0618 (-1.08,0.88)	0.138 (0.018,1.33)	0.71 (0.17,2.56)	0.188 (0.023,2.07)
eulachon	41.55 (1.43,456)	7.088 (0.172,39.1)	0.68 (0.28,1.64)	1.07 (0.49,4.29)	5.01 (3.67,6.03)	1.24 (0.47,6.51)	0.98 (0.24,2.75)	0.414 (0.033,2.44)
squid	3.690 (0.573,58)	1.647 (0.043,11.9)	0.038 (0.012,0.165)	1.67 (0.72,6.12)	3.14 (1.51,4.70)	0.87 (0.15,6.04)	0.214 (0.061,1.61)	0.155 (0.022,1.91)
flatfish	31.506 (0.993,755)	5.925 (0.30,61.86)	225.86 (178.0,282.3)	3.9 (1.8,9.4)	-1.451 (-1.88,-0.81)	1.09 (0.47,6.94)	0.66 (0.11,2.82)	0.271 (0.017,2.53)

Table 3.3 continued

prey group	ϕ_y^λ	$\zeta_{w,1}^\lambda$	$\zeta_{w,2}^\lambda$	$\left(\sigma_{w,y}^\lambda\right)^2$	μ_α	σ_α^2	α_y
juvenile pollock	0.000649 (7.7e-22,0.75)	2.73 (1.88,5.37)	3.585 (0.612,50.5)	2.16 (0.058,10.9)	7.56 (4.19,11.14)	0.315 (0.005,15.0)	7.53 (3.37,12.76)
adult pollock	0.767 (0.0052,3.7)	2.24 (1.71,2.72)	0.525 (0.0474,2.4)	2.138 (0.775,4.62)	2.57 (1.58,3.64)	0.0988 (0.007,1.14)	2.54 (1.32,4.07)
cod	0.205 (1.6e-07,3.5)	1.053 (0.95,1.20)	0.109 (0.016,0.51)	1.019 (0.401,2.04)	1.052 (0.78,1.422)	0.05066 (0.005,0.42)	1.012 (0.46,1.84)
herring	0.9468 (0.030,4.47)	2.71 (1.43,5.71)	7.003 (0.874,48.6)	1.735 (0.063,12.1)	1.198 (0.292,3.51)	0.1107 (0.0096,7.1)	1.1623 (0.070,4.41)
salmon	0.222 (2.1e-05,3.7)	1.032 (0.723,1.40)	0.563 (0.107,1.81)	0.859 (0.068,3.17)	0.2347 (0.066,0.65)	0.0333 (0.004,0.55)	0.1837 (0.002,1.11)
smelt	0.579 (0.0072,3.3)	2.26 (1.19,9.66)	7.17 (1.61,200.4)	1.309 (0.013,14.8)	1.317 (0.669,2.76)	0.1115 (0.0085,2.9)	1.243 (0.333,3.24)
eulachon	0.962 (0.018,3.71)	1.628 (0.947,5.72)	3.147 (0.872,68.6)	1.138 (0.021,9.60)	3.76 (1.88,7.36)	5.273 (0.115,55.8)	3.096 (0.28,13.54)
squid	0.0629 (1.6e-11,2.6)	0.797 (0.398,1.13)	0.2543 (0.051,1.31)	0.665 (0.054,2.13)	2.26 (1.01,3.50)	0.1575 (0.0071,5.0)	2.146 (0.725,4.52)
flatfish	0.558 (0.0016,3.7)	1.046 (0.842,1.49)	0.601 (0.191,2.28)	0.853 (0.088,3.42)	0.902 (0.709,1.11)	0.01579 (0.003,0.10)	0.896 (0.55,1.323)

Table 3.4. As for Table 3.3, but for the case when no spatial correlation is modelled in either the mean or zero-inflation components.

prey group	μ_p^*	σ_p^2	β_j^p	$\sigma_{b,p}^2$	μ_λ	σ_λ^2	β_j^λ	$\sigma_{b,\lambda}^2$
juvenile pollock	0.3162 (0.002,1.0)	0.70 (0.28,1.91)	-11.77 (-23.6,-1.9)	3.64 (0.053,43)	0.42 (0.12,1.7)	0.71 (0.37,1.23)	-2.79 (-5.2,-0.24)	1.78 (0.38,9.72)
adult pollock	0.079 (0.01,0.31)	0.97 (0.41,2.07)	2.75 (-0.75,5.4)	1.00 (0.18,9.8)	27.97 (21,39.6)	1.39 (0.72,2.40)	-0.509 (-1.3,-0.03)	0.099 (0.014,0.9)
cod	0.073 (0.02,0.16)	1.02 (0.36,2.35)	0.403 (-0.47,1.77)	0.48 (0.07,3.2)	19.4 (13.8,26)	1.72 (0.87,2.73)	-0.373 (-1.0,0.07)	0.149 (0.05,0.83)
herring	0.99 (0.97,1.00)	1.3 (0.6,3.2)	2.347 (0.94,3.76)	1.53 (0.51,8.5)	0.179 (0.04,0.6)	2.64 (0.77,89.7)	0.294 (-1.38,3.4)	0.78 (0.18,4.95)
salmon	0.84 (0.74,0.89)	1.15 (0.59,2.33)	-0.412 (-1.02,0.39)	0.174 (0.014,1.5)	2.6 (1.4,4.5)	2.22 (0.94,23.0)	-0.0488 (-0.6,0.77)	0.255 (0.05,1.79)
smelt	0.79 (0.34,0.94)	0.47 (0.19,0.96)	1.52 (-1.34,4.1)	1.04 (0.24,6.3)	0.0222 (0.006,0.07)	0.57 (0.29,1.15)	-1.85 (-4.2,-0.5)	0.36 (0.05,2.97)
eulachon	0.56 (0.31,0.77)	2.60 (0.84,15.1)	3.36 (0.73,5.08)	0.72 (0.19,4.3)	1.69 (0.94,3.6)	1.07 (0.52,2.25)	1.95 (0.97,3.8)	0.425 (0.08,2.95)
squid	0.20 (0.14,0.27)	1.70 (0.84,3.09)	-0.0151 (-0.48,0.67)	0.20 (0.06,1.2)	197.55 (155,353)	1.43 (0.77,2.31)	-0.435 (-1.0,0.02)	0.128 (0.013,1.1)
flatfish	1.3e-05 (3e-06,1.5e-04)	1.80 (0.82,13.3)	9.45 (6.41,11.9)	4.3 (1.8,20.6)	258.95 (205,326)	2.3 (1.2,4.0)	-2.73 (-3.3,-2.1)	1.04 (0.41,4.96)

Table 3.4 continued

prey group	μ_α	σ_α^2	α_y
juvenile pollock	13.6 (10.3,18.6)	1.43 (0.016,28)	13.3 (9.4,21)
adult pollock	5.43 (4.82,6.3)	0.127 (0.0054,1.7)	5.43 (4.3,7.1)
cod	2.85 (2.24,4.03)	0.952 (0.057,5.1)	2.79 (1.1,5.7)
herring	3.78 (2.16,6.59)	0.213 (0.007,7.3)	3.68 (1.6,7.2)
salmon	1.439 (0.98,2.26)	0.170 (0.009,2.9)	1.42 (0.4,3.3)
smelt	13.05 (9.16,16.8)	0.577 (0.009,66)	12.7 (7.0,22)
eulachon	6.21 (3.98,8.86)	3.15 (0.011,24)	5.87 (2.3,13)
squid	1.94 (1.58,2.46)	0.320 (0.076,1.3)	1.91 (0.8,3.6)
flatfish	1.79 (1.50,2.13)	0.148 (0.041,0.6)	1.754 (1.0,2.8)

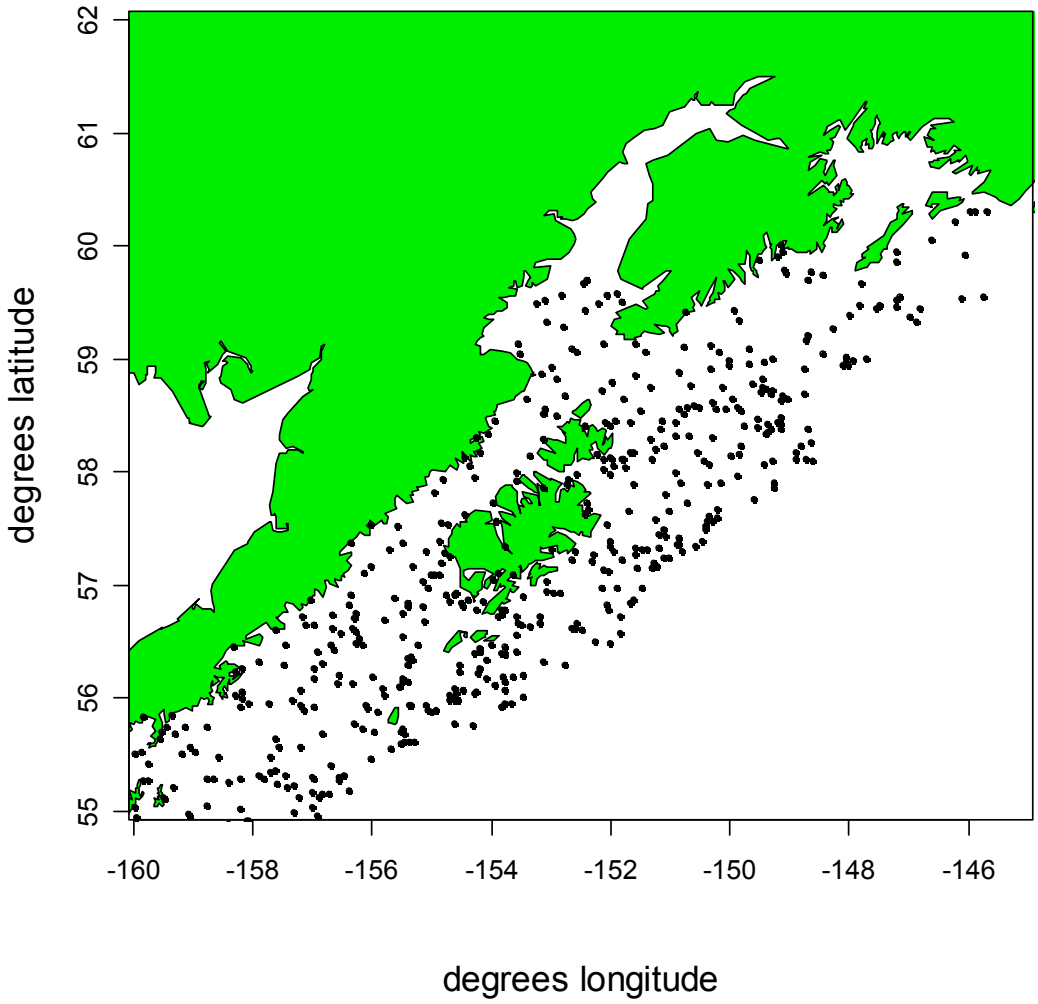


Figure 3.1. Subset of the Gulf of Alaska used in the analyses. Black points indicate locations of survey hauls for 2007.

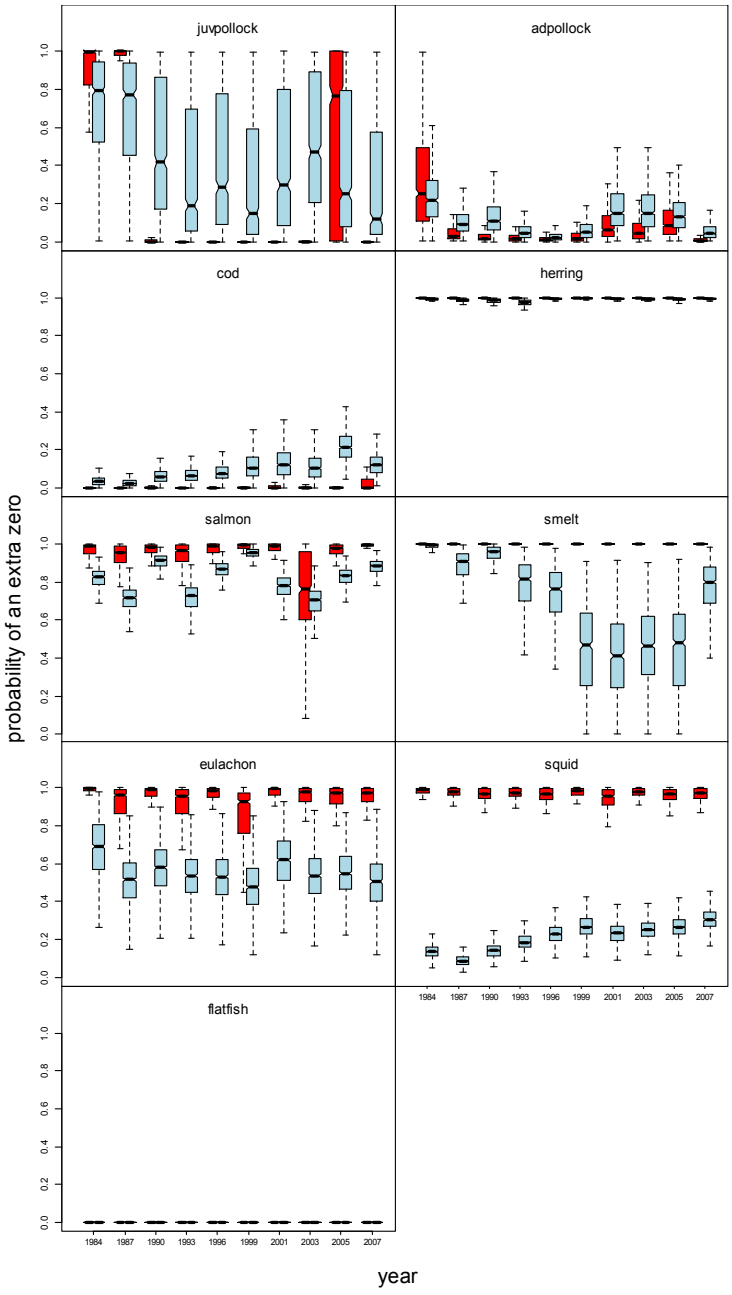


Figure 3.2. Posterior distributions for the annual values for the mean probability of an additional zero for the nine prey groups, for the full spatial model (red) and the non-spatial model (blue). Notches in boxes indicate the posterior median, width of boxes indicate the central 50% credibility interval, and the extent of the whiskers indicate the central 95% credibility interval.

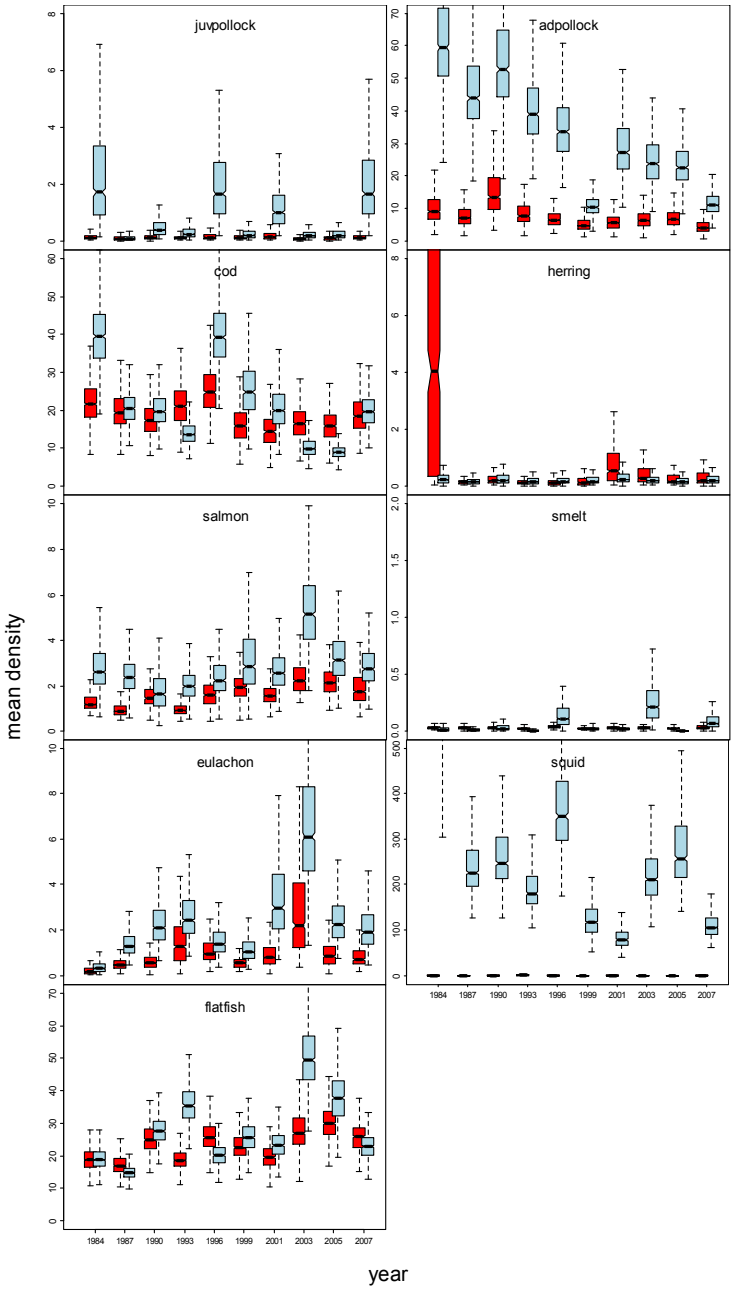


Figure 3.3. Posterior distributions for the annual values for the mean prey density ($kg.ha^{-1}$) for the nine prey groups, given that an observation is drawn from the appropriate negative binomial distribution. Red bars are for the full spatial model, blue bars indicate the values for the model with no spatial correlation. Boxes are plotted as for Fig. 3.2.

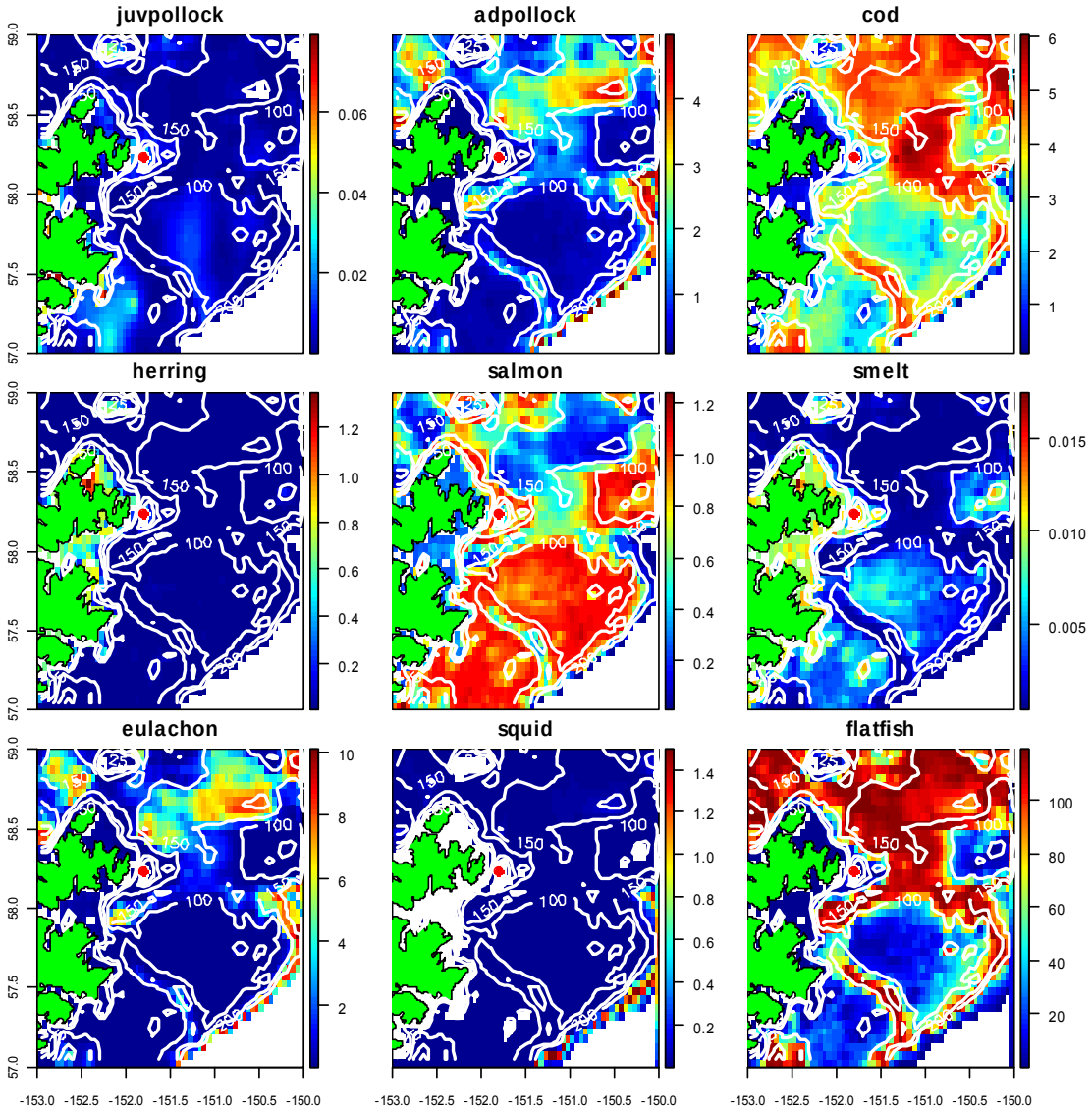


Figure 3. 4. Median of the posterior for density (kg.ha^{-1}) of prey groups around Marmot Island for 2001.

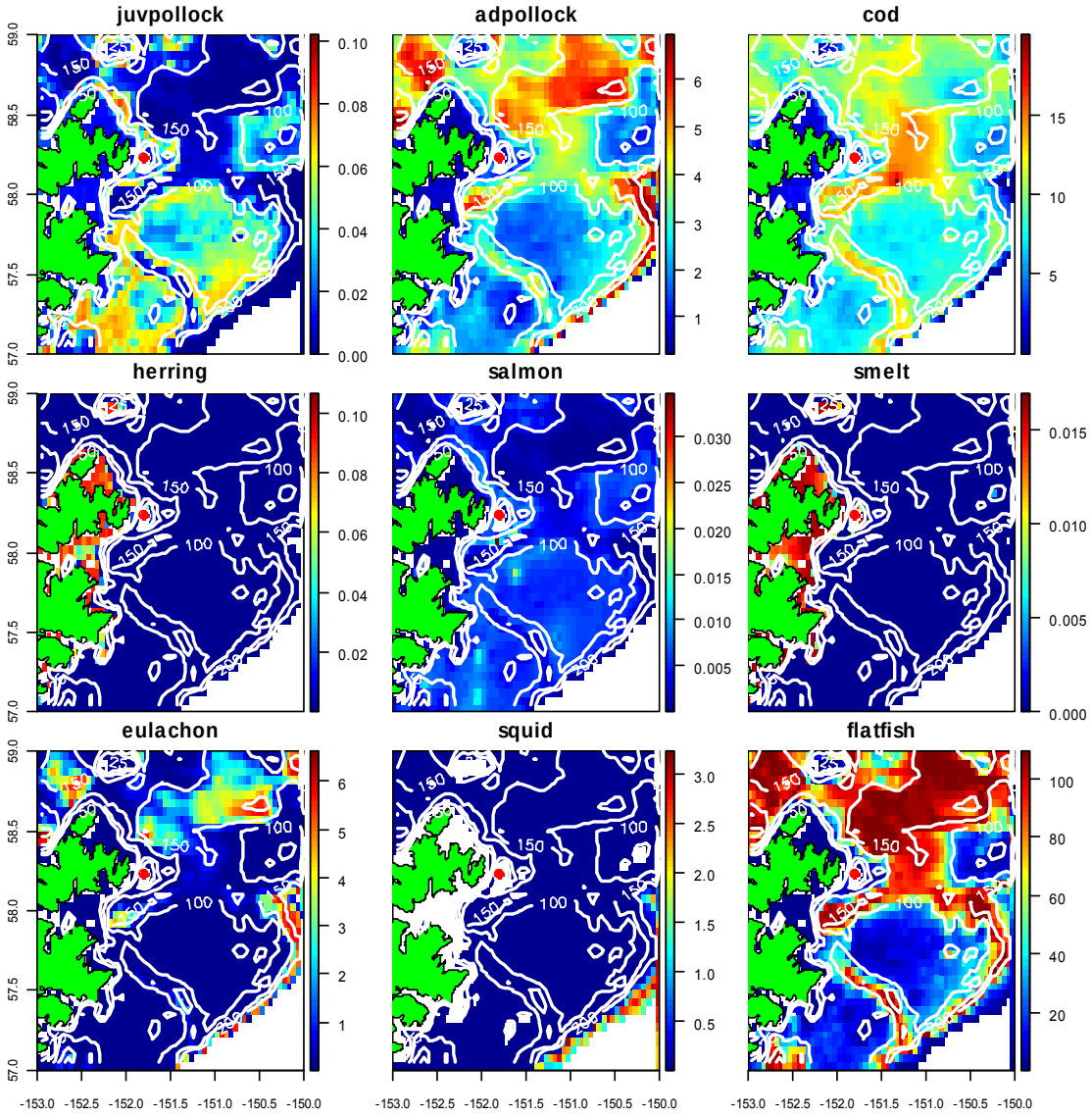


Figure 3.5. Median of the posterior for density (kg.ha^{-1}) of prey groups around Marmot Island for 2007.

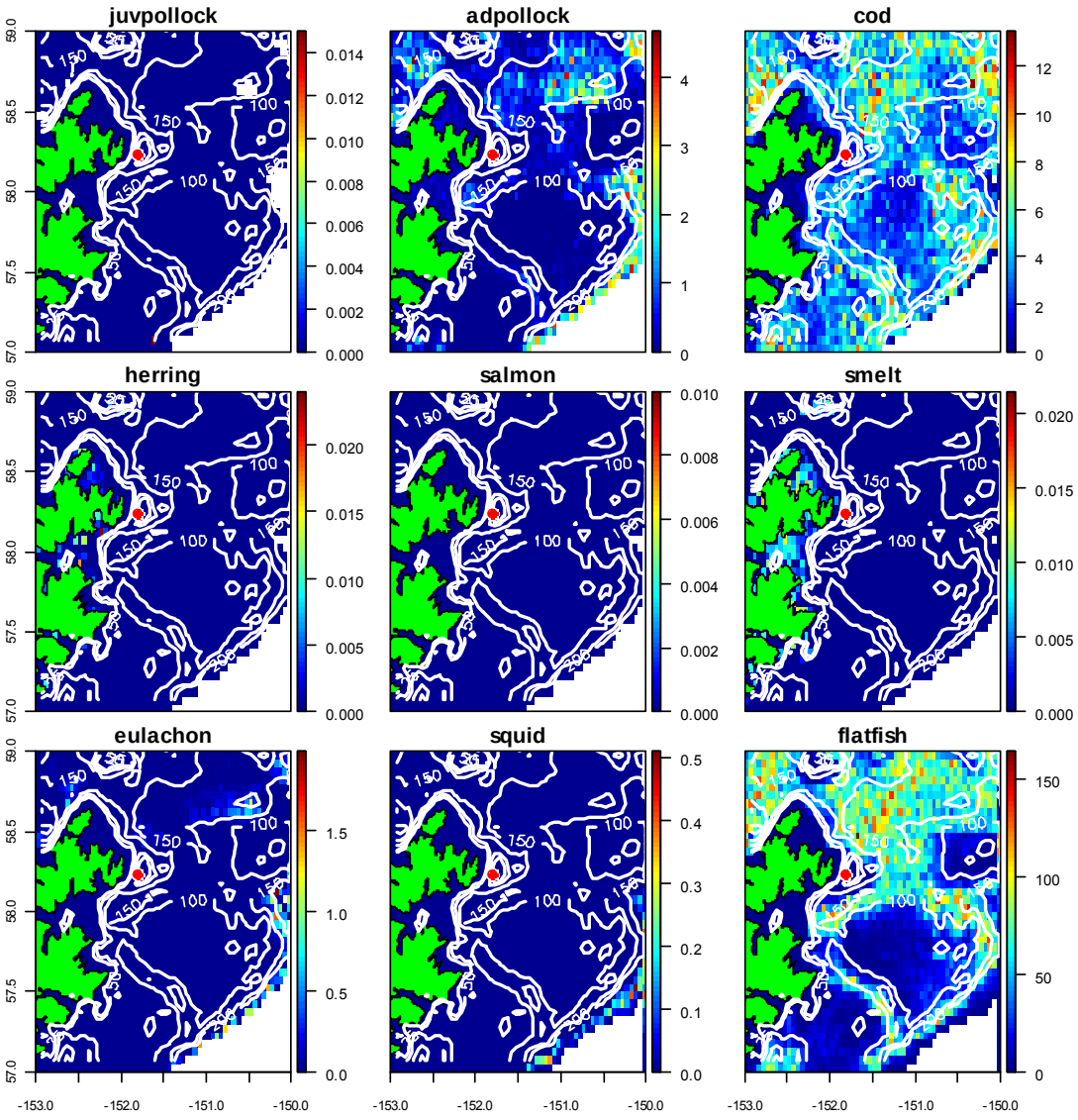


Figure 3.6. Median of the posterior predictive distribution for density (kg.ha^{-1}) of prey groups around Marmot Island.

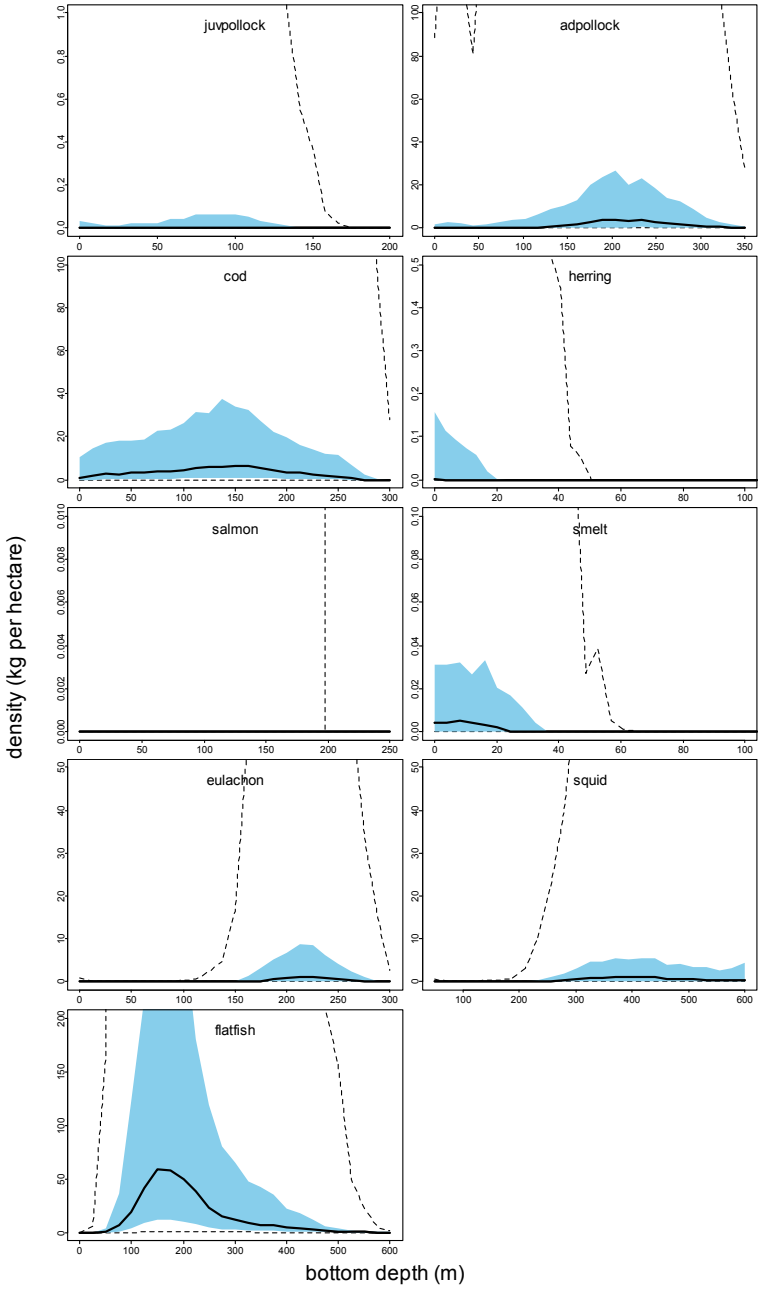


Figure 3.7. Posterior predictive distribution of the expected density (modified by zero-inflation) with depth. Solid black line represents the posterior median, shaded area indicates the central 50%, and dotted lines indicate the 95% credibility intervals.

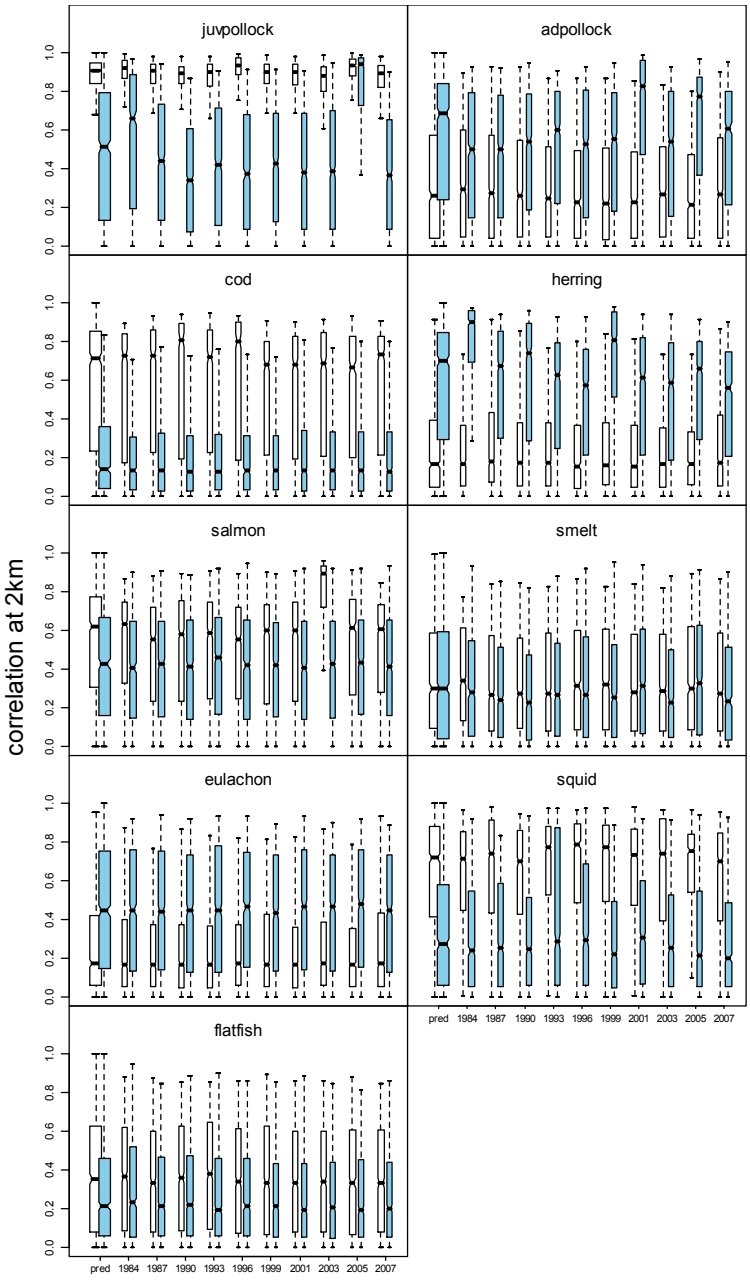


Figure 3.8. Posterior predictive distributions of the spatial range parameters, expressed as the correlation at a distance of 2km. white – correlation in mean of negative binomial, blue – correlation in probability of an extra zero. Boxes summarise the distribution as for Fig. 3.2.

Chapter 4: Linking foraging behaviour and bioenergetic requirements of Steller sea lions to the spatial distribution of their prey

4.1. Introduction

Many hypotheses for population change in the western stock of Steller sea lions in Alaska relate to prey distribution and diet composition (National Research Council 2003). These hypotheses state that the quantity and/or quality of food available to different segments of the Steller sea lion population has changed over time, due to changes in numbers and/or the geographic distribution of prey species, either as a result of environmental forcing factors, or as a result of directed fishing activity for these prey species. An evaluation of how such changes in the spatial nature of the prey base could impact Steller sea lions requires a synthesis of information on sea lion bioenergetics, foraging behaviour, and the nature of the spatial distribution of sea lion prey.

Steller sea lions appear to be able to respond to changes in prey distribution by adjusting their geographical distribution and foraging behaviour to take advantage of geographically predictable seasonal increases in the abundance of prey species (Raum-Suryan et al. 2002, Sinclair and Zeppelin 2002, Womble et al. 2005), and can also depress metabolic rates during periods of low food availability. Steller sea lion diet compositions appear to reflect general patterns of region-specific relative abundances of prey species(Sinclair et al. 2005). Sinclair and Zeppelin (2002) suggest that Steller sea lions rely on dense, predictable, prey patches, associated with bathymetric and temperature gradients, which is compatible with observations of sea lions feeding in the complex habitats associated with frontal systems in and around Aleutian passes (Fadeley et al. 2005).

Foraging patterns reflect tradeoffs between balancing various ecological and physical constraints, and the need to grow and survive to reproduce successfully and maintain fitness. Differences in foraging strategies among the life-history stages of Steller sea lions are perhaps expected given the diverse and changing nature of the various ecological and physiological constraints on growth and survival with life history status. Examination of foraging behaviour for Steller sea lions and other otariids indicates that persistence of prey patches is important for foraging efficiency, particularly when spatial use is constrained by factors such as the need to provision offspring (Bonadonna et al. 2001, Gende and Sigler 2006). For instance, females with dependent offspring will be limited in the distance they can travel from the haul-out (Trites and Porter 2002), because they must return regularly to deliver milk to the pup. In telemetry studies, lactating females remained close (often less than 20 km) to rookeries following the breeding season (Millette and Trites 2003). The location of rookeries and haul-outs in close proximity to predictable quantities of available and suitable prey is probably very important for lactating females and their offspring to meet energetic requirements. Pup growth was correlated with prey availability (CPUE around rookery; Andrews et al. 2002). The switching of mother/pup pairs from rookeries of birth to haul-out sites close to seasonal aggregations of prey reflects the importance of good foraging opportunities to lactating females (Womble et al. 2005).

The importance of the spatial distribution of suitable prey to foraging success of sea lions is clearly evident. This spatial distribution combines with physiological and behavioural mechanisms to determine how much food is available to an individual sea lion. Bioenergetics models are useful tools in evaluating the energetic requirements of organisms, given needs for maintenance, activity, and growth (e.g. Ney 1990, Brandt et al. 1992). Such models are frequently used to estimate expected consumption, or to identify areas likely to provide best conditions for growth. Bioenergetics models for Steller sea lions have been used to compare energy budgets among different life-history

stages (Winship et al. 2002), calculate the quantity of prey required (Winship and Trites 2003), and evaluate the impact of sea lion predation on a walleye pollock prey population (Cornick et al. 2006b). However, these models can also be employed to evaluate expected energy expenditure given assumptions regarding foraging strategies. Models of foraging behaviour often focus on identifying 'optimal' foraging strategies given a small set of decisions, and evaluate performance based on metrics such as the quantity of suitable food items obtained per unit time (e.g. Pyke 1984). Models employed for estimating utilisation of space in animals frequently have at their core an individual-based framework that accounts for the trajectory and movement behaviour of the animal in question, and assume that this behaviour is driven in response to some function of the habitat, such as prey availability (e.g. Matthiopoulos 2003, Nations and Anderson-Sprecher 2006). Coupling models of foraging behaviour to models of the bioenergetics would allow an evaluation of the likely consequences of both different behaviour patterns, and the spatial distribution of prey, on the ability of the forager to meet the energy budget.

This Chapter utilises information on the foraging behaviour of Steller sea lions and syntheses of data regarding the spatial distribution of important prey species to develop a simulation model of individual-based foraging trips that evaluates expected energy gains given prey distributions surrounding sea lion haul-out sites, and various hypotheses for foraging behaviour. Simulations are conducted for two life-history stages, juveniles and adult females, and are representative of summer foraging behaviour. Simulation output distributions of expected energy gain for sets of prey distribution and foraging strategy assumptions are compared to evaluate the relative impact of these assumptions on foraging success and which assumptions result in successful foraging. Foraging strategy parameters tested include the duration of the foraging trip, and how to make use of information from previous trips. Prey distribution parameters include density and patchiness of prey, detectability of prey, caloric value of prey, and persistence of prey

patches between foraging trips. Summary statistics of the prey fields and foraging trips are also calculated to investigate whether trip-scale indicators of foraging behaviour and prey availability can be used to identify successful strategies and account for observed variations in energy gain.

4.2. Methods

A spatially-explicit foraging model is linked to a model of bioenergetics for Steller sea lions (SSLs). The foraging model allows for an individual foraging sea lion to exploit a prey field from a central haul-out site given different values for key components of the foraging behaviour and prey distribution. Specifically, the foraging simulations compare energy gains given interactions among:

i. Sea lion age: (juvenile or mature female) - which determines some of the parameters of the bioenergetics model.

ii. Foraging behaviour: limited to foraging trip duration, the length of time hauled out between trips, and the weightings placed on previous experience versus local conditions when foraging.

iii. Prey availability: through density, detectability of prey, prey type (energy density of prey), spatial aggregation of prey, and persistence of prey fields among foraging trips.

4.2.1. Simulation Framework

The analyses are conducted using a trip-based individual model of Steller sea lion foraging, which occurs within a fixed foraging arena with a central haul-out site with a prey field, the spatial distribution of which is allowed to change between foraging trips. The general procedure for the simulations is the following:

1. draw a vector of parameter values determining the distribution of prey, prey persistence, and foraging strategy,

2. generate *N* (200) realisations of the spatial autocorrelation process for the prey,

3. calculate the initial prey distribution from the parameters and spatial process,

4. run the foraging simulation for a single sea lion for the *N* trips, updating the prey field at the beginning of each trip (to represent fish movement, growth, mortality, etc.),

5. repeat steps 2-4 *n* times with different stochastic realisations of the prey field time series,

6. repeat steps 1-5 given different values for the parameters governing the spatial distribution of prey, prey persistence, and the specifications of the foraging strategy.

The results of each set of *N* foraging trips are summarised in terms of the net energy gain over the last 100 trips (to allow for 'burn-in' of the foraging strategy), and a set of summary statistics is computed. 'Burn-in' is required to evaluate the expected energy gain given a foraging strategy in equilibrium rather than that from a naive forager.

4.2.2. Foraging arena

The foraging simulations take place in a 51 by 51 cell grid, with each grid cell being a square with edge of size 2km (i.e. the foraging arena is a square with edge of 102km), with the haul-out site being in the central grid location. Simulations therefore represent foraging around an island as opposed to a site positioned along a coastline.

Location of sea lion prey resources is associated with bathymetric profile (Chapter 3). However, for simplification, the foraging arena is modelled without depth, to more clearly understand how the parameters governing spatial correlation in prey distribution and prey persistence impact energy gain in foragers. A disadvantage of this approach is that it does not lead to trends in the spatial distribution of prey with distance from the haul-out site, for example as a result of increasing depth. An alternative would be to use a bathymetric profile designed to mimic the continental shelf in the Gulf of Alaska (e.g. by fitting nonlinear relationships of depth with distance from haul-out to bathymetric data from a set of sea lion haul-outs). Prey species of Steller sea lions are found at a range of depths relative to bathymetry around haul-outs along the continental shelf in the Gulf of Alaska (e.g. Chapter 3, Adams et al. 2008). Steller sea lions have been observed diving to depths greater than 200m, and 90% of seafloor within calculated home ranges of foraging adult females in Southeast Alaska was shallower than the maximum diving depth (Rehberg et al. 2009).

4.2.3. Model of prey distribution

The prey field is modelled to mimic the spatial characteristics of sea lion prey species in the Gulf of Alaska, and is guided by the results of the analyses in Chapter 3. While the results of Chapter 3 provide parameter estimates for determining prey distribution and density, it is necessary to know the catchability (of the survey) for each prey species, and also the multispecies functional response for sea lion predation given different densities of prey to develop prey fields for sea lions to exploit. While survey catchability for many prey species can be derived from stock assessments that integrate many types of data (e.g. Dorn et al. 2011), this is not so easy to obtain for forage fish species; although estimates of survey catchability could potentially be obtained from the results of ecosystem models (e.g. Aydin et al. 2007). Defining the functional response is more problematic because it is difficult to measure the ability of free-ranging Steller sea lions to find and exploit different densities of prey *in situ*.

To obtain a more tractable approach to addressing the objectives, the prey field is modelled as being a spatial distribution of patches, with patch size and spatial distribution modelled according to the spatial parameters in Chapter 3, but where prey density within a patch is modelled as a fraction of the maximum ingestion rate. Conducting simulations with different values for this prey density allows the impact of prey density on foraging strategy performance to be evaluated. In a similar fashion, the prey is modelled as having one of two energy densities, representing either a gadid prey base of walleye pollock (*Theragra chalcogramma*), or small schooling fish such as Pacific capelin (*Mallotus villosus*). Table 4.1 lists the parameter values used in the simulations. A particular realisation of the prey field is determined by selecting values for the probability of a grid cell containing a prey patch, the spatial correlation among grid cells, the patch density (as a fraction of the maximum ingestion rate), the energy density of the prey, and a value for the persistence of prey patches, which determines how the prey field changes between foraging trips (Figure 4.1).

Energy gain for foraging strategies would be expected to depend on how much the prey field changes (or is expected by the sea lion to change) between foraging trips. The evolution of the prey distribution $\mathbf{p}_t$ between foraging trips is governed by:

$$\begin{aligned} \text{logit}\left(\mathbf{p}_t\right) &= \text{logit}\left(\mu\right) + \mathbf{w}_t \\ \mathbf{w}_t &= \rho \mathbf{w}_{t-1} + \tilde{\mathbf{w}}_t \left(1-\rho_t^2\right)^{0.5} \end{aligned} \tag{4.1}$$

where $\mathbf{w}_t$ is the spatial deviation from the mean prey probability for the prey field at time *t*, ρ_t is the persistence of the prey field from the previous trip (ρ ^((trip time + haul-out duration)/24), where ρ is the persistence over 24 hrs), and $\tilde{\mathbf{w}}_t$ is a random spatial error vector, distributed according to a multivariate normal, with:

$$\begin{aligned} \mathrm{E}\left[\tilde{w}_{t,\mathbf{s}}\right] &= 0 \\ Cov\left[\tilde{w}_{t,\mathbf{s}}, \tilde{w}_{t,\mathbf{s}'}\right] &= \left(\sigma_w^p\right)^2 \exp\left(-\phi^p \left|\mathbf{s}-\mathbf{s}'\right|\right) \end{aligned} \tag{4.2}$$

The values for the spatial parameters σ_w^p and ϕ^p, as well as the mean probability of a cell containing a prey patch, μ (Table 4.1) are chosen from fits to Gulf of Alaska groundfish survey data using a GAMM approach (see Chapter 3 for details).

Evolution of the prey field occurs between trips and is not modelled to change during a foraging trip, either as a result of prey movement or decrement due to sea lion foraging. Prey patches exploited by Steller sea lions in Southeast Alaska were estimated to be sufficiently abundant to contain enough rations for many times the number of sea lions observed foraging (Sigler et al. 2009). To allow for comparison among scenarios (which vary the length of foraging trip duration and time ashore, dependent on the foraging strategy), the persistence parameter is defined over a 24 hour period, with the inter-trip persistence then being appropriately modified for the scenario being conducted, dependent on the length of the previous trip, and the haul-out duration.

4.2.4. Foraging strategies

Foraging strategies involve ways to make decisions that determine the behaviour during a trip, and include:

1. when to leave the haul-out to begin a foraging trip (haul-out duration),
2. foraging trip duration, and
3. how to use available (and previously obtained) information to move to locate prey resources.

Table 4.2 details the specifications of the scenarios, including the values of the parameters determining the foraging strategy. Values for the parameters were chosen to reflect the range of values reported in the literature. Sources are provided in Table 4.2. The amount of time spent between foraging trips is important, as the persistence of prey fields between trips is linked to the amount of time spent ashore. Metabolic requirements are also reduced while on land.

4.2.5. Foraging model

At a given time step in the foraging trip, foragers occupy a grid cell with location $\mathbf{x} = (x_1, x_2)^T$ on the (2-dimensional) foraging arena, and are able to move to another cell. A foraging trip consists of the following: leaving the haul-out, foraging, and homing to the haul-out site.

Leaving: The foraging trip begins at the haul-out site. Foraging trips begin following the pre-specified haul-out duration (T_{land}, Table 4.2).

Foraging: While foraging, the SSL moves randomly across the perceived resource gradient, with the expectation that the forager will move to cells with higher prey density[4], because SSLs are known to distribute themselves according to available prey and appear to be able to find and exploit areas of high productivity (Fadely et al. 2005). The ability of foragers to perceive the local conditions depends on a detectability parameter, which allows for uncertainty in the perceived prey field. Low detectability implies little ability to discern the location of available prey for movement purposes, whereas high detectability means that the forager is able to use the actual prey locations

[4] Prey density should perhaps be viewed as the expected consumption rate, given that available prey is modeled as a function of the maximum ingestion rate.

to determine movement. In this context, detectability is differentiated from a detection rate that might be used in a functional feeding response to determine consumption.

Movement at time *t* between the current location and neighbouring cells is a function of the current location, a movement distribution f_m, and the perceived resource density, f_R. The movement distribution is determined by the expected distance travelled given the time-step and the known range of swimming speeds for SSLs, and reflects the costs of movement with all else (prey density) being equal. It assumes no preference in the direction of travel, with position at time *t* distributed bivariate normal from position at time *t*-1:

$$f_m(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \Sigma_m) \propto \left|\Sigma_m\right|^{-0.5} \exp\left[-0.5\left(\mathbf{x}_t - \mathbf{x}_{t-1}\right)' \Sigma_m^{-1}\left(\mathbf{x}_t - \mathbf{x}_{t-1}\right)\right] \tag{4.3}$$

where $\Sigma_m = \sigma_m^2 \mathbf{I}$, with the values for σ_m^2 as in Table 4.2.

The location at time *t* is randomly sampled from a location distribution f_c, which is determined by:

$$f_c(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \Sigma_m, \mathbf{R}_t) \propto f_m(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \Sigma_m) f_R(\mathbf{x}_t \mid \mathbf{R}_t) \tag{4.4}$$

The resource distribution f_R is that distribution of prey density which influences SSL movement, which may or may not correspond to the actual prey density. That is, f_R is defined differently given assumptions about the foraging strategy of the SSL, and the detectability of prey. The resource distribution reflects the information on prey density that is available to the forager when it is making decisions about where to move to. The foraging strategy specifications in Table 4.2 indicate that individuals are modelled assuming a range of information levels, with a range of weightings for the relevant strength of information from current local conditions (modified by detectability) and memory from previous experience. In the latter case, information about f_R is retained

within and among foraging trips, such that the forager can utilise information from cells it has visited previously. Note that with low persistence of prey density between trips such information may not be correct.

$$f_R(\mathbf{x}_t \mid \mathbf{R}_t) = \theta f_R(\mathbf{x}_t \mid \mathbf{R}'_{t-1,t-2,\ldots,t=1}) + (1-\theta)\left(\alpha \mathbf{p}'_t + (1-\alpha)/2\right) / \sum_{'}\left(\alpha \mathbf{p}'_t + (1-\alpha)/2\right) \quad (4.5)$$

where θ is the weight assigned to previous experience versus local conditions,

$f_R(\mathbf{x}_t \mid \mathbf{R}'_{t-1,t-2,\ldots,t=1})$ is the visited resource distribution up until time t (prey density in each cell at the time that cell was last visited), and

α is the prey detectability, independent of location, which indicates how easy it is for the foraging sea lion to perceive local conditions.

A new location for each time step is then sampled from the location distribution (Figure 4.2). At each time-step within a foraging trip the energetic condition of the forager is calculated using the bioenergetics model given the updated location.

Homing: Homing to the haul-out site occurs when the remaining time in the trip is equal to that which it would take the sea lion to return to the haul-out moving at the average speed. Movement during homing is a function of both the movement distribution f_m, and the distance from the haul-out site. The sea lion always moves towards the haul-out (probability of moving to cells with distance from the haul-out greater than the current cell are set to zero), to force the SSL to travel to the haul-out site, rather than just being constrained in space by its presence.

4.2.6. Bioenergetics model

The bioenergetics model is a modification of that developed for Steller sea lions by Winship et al. (2002). The activity of sea lions and prey encounter rate from the foraging model are used to evaluate the energy gains and losses by individual foragers. Values for the parameters of the bioenergetics model were taken from Winship et al. (2002), and

recent literature on the metabolic cost of transport and diving in Steller sea lions and other pinnipeds (e.g. Rosen and Trites 1997, 2002a, 2002b, 2003), and the energetic composition of sea lion prey (e.g. Winship and Trites 2003, Kitts at al. 2004). The values used for the bioenergetics parameters, along with their sources, are listed in Table 4.3.

The net energy gain for a trip *i* depends on activity and foraging success, and is calculated by:

$$J_i = \sum_{t=1}^{T_{forage,i}} G_{\mathbf{x}_t} - T_{sea,i} BM^{sea} - T_{land} BM^{land} \tag{4.6}$$

where $BM_t^{land/sea}$ is the energy loss due to basal metabolism during time *t*:

$$BM_t^{land/sea} = A^{land/sea} MC_t \tag{4.7}$$

MC_t is the maintenance cost over the time step,

$A^{land/sea}$ is an activity multiplier dependent on whether the sea lion is at sea during time *t* or hauled out,

$T_{sea,i}$ is the total time spent at sea during trip *i,*

$T_{forage,i}$ is the time spent foraging during trip *i* (does not include time spent at sea when returning to the haul-out),

T_{land} is the haul-out duration, and

$G_{\mathbf{x}_t}$ is the expected energy gain from point **x** at time t.

MC_t is a function of body size (Table 4.3). While active metabolism is a function of the distance travelled and the time to cover that distance, the simulations considered a fixed multiplier to maintenance costs irrespective of distance travelled in a given time-step, because the foraging model (Section 4.2.5) restricts movement based on average swim speed. Locomotor costs for SSL are consistent with five-fold multipliers of maintenance

costs (Table 4.3, Rosen and Trites 2002b, Winship et al. 2002). $G_{\mathbf{x}_t}$ is the net of energy losses and gains associated with foraging at a particular location at time *t*, and is a function of the prey density at the location where the SSL is foraging (defined as the fraction of maximum ingestion rate), the energy content of that prey, the assimilation efficiency, and the digestive capacity. Metabolic costs associated with diving in search of food are assumed to be included in $BM_t^{land/sea}$. Metabolic rates of diving (foraging) events in Steller sea lions are similar to surface metabolic rates in summer (Fahlman et al. 2008, Svärd et al. 2009).

$$G_{\mathbf{x}_t} = E_{HIF} E_{dig,i} ED_i \tilde{B}_{\mathbf{x}_t} \tag{4.8}$$

where ED_i is the energy density of prey species *i*,

$E_{dig,i}$ is the digestive efficiency for prey species *i*,

E_{HIF} is the efficiency of utilisation of metabolizable energy,

$\tilde{B}_{i,\mathbf{x}_t}$ is the ingested biomass of prey species *i*:

$$\tilde{B}_{i,\mathbf{x}_t} = B_{i,\mathbf{x}_t} DC \tag{4.9}$$

$B_{i,\mathbf{x}_t}$ is the prey density at location **x**, and

DC is the digestive capacity (maximum biomass consumable per time-step).

An alternative would be to replace equation (4.9) with a Holling (1959) Type III functional response (Real 1977). Parameterisation of the functional response could be achieved from the results of feeding experiments with different prey densities. For example, Lovvorn et al. (2009) utilised such an approach in their parameterisation of a foraging model for spectacled eiders. Simultaneous observations of foraging behaviour and prey availability from acoustic surveys have been conducted for baleen whales, and Witteveen et al. (2008) quantify prey selection of capelin over walleye pollock for humpback whales in the Gulf of Alaska. Experimental results for captive Steller sea lions

do demonstrate that there is a relationship between prey density and feeding rate (Cornick et al. 2006a). However these results are for individual encounter rates, and do not reflect Steller sea lions foraging on schools of fish, the scale considered in the simulations. Assuming a constant ingestion rate for a given scenario provides a tractable way of comparing the effects of prey density on foraging success without the need to introduce further (arbitrary) assumptions.

4.2.7. Summary statistics

The results of the simulations, for each set of foraging strategies, are summarised using the distributions for: (a) the net energy gain (both for each trip and across the length of the trip time series), (b) the proportion of foraging trips that are successful (i.e. Equation 4.6 is positive), (c) the average (and maximum) distance from the haul-out during each trip, (d) the magnitude of spatial usage per trip (number of grid cells visited while foraging), and (e) the total amount of prey available within a fixed distance from the haul-out. The latter is calculated to identify whether the quantity of prey within a certain distance (e.g. 20km) correlates with foraging success. Quantities (a) and (b) indicate whether foraging is successful or not, (c) and (d) are typically reported or calculated when summarising the results of telemetry studies, and (e) could be enumerated from fisheries research surveys. The indicators of foraging success are then used to compare the results among the different foraging strategies and given the levels of persistence in the prey field between foraging trips. Quantities (c), (d), and (e), along with trip time, are then used as linear predictor terms in logistic regressions of foraging success on an individual trip basis, to evaluate how effective these summary statistics are at predicting whether foraging is likely to be successful or not. Defining foraging success purely in terms of a positive net energy gain over a trip cycle is admittedly a fairly low bar, as Steller sea lions have energetic requirements for growth and reproduction in addition to activity and maintenance costs. However, activity and maintenance costs (accounted for here) are the primary component of the energy budget for Steller sea lions, with

additional requirements (e.g. for growth) only being a small percentage of total energy requirements (Winship et al. 2002).

4.3. Results

Results from foraging scenarios are presented in terms of the median and central 95% simulation interval for the proportion of trips that were successful, and the total net energy over the last 100 trips of each simulation, for each combination of prey parameters, trip time, haul-out duration, prey persistence, prey density (relative to maximum ingestion rate), prey detectability, and weight placed on previous experience relative to (memory). For example, the top-left panel of Figure 4.3 shows how the distribution (line median, shading 95% interval) for the proportion of trips which are successful changes is related to the weight placed on memory when moving, for three values of prey detectability, given an ingestion rate of 100%, and low prey persistence of 0.1 $24h^{-1}$ when foraging trip time and haul-out duration are both 24 hours, and the prey base is modelled as for forage fish. Results shown are limited to ingestion rates of 100% and 50% of the maximum, as values lower than these appear unable to sustain successful foraging (almost all trips are unsuccessful). General results are discussed for a 'base-case' scenario with a forage fish prey base and foraging trip time and haul-out duration equal to 24 hours. Results for the other scenarios are then detailed relative to this scenario.

4.3.1. Base-case scenario

Foraging success depends on the prey detectability, ingestion rate, and how unvisited cells were treated when placing weight on memory. Higher values for prey detectability resulted in increased foraging success (Figure 4.3, compare lines within each panel). At low prey detectability, increasing the weight assigned to memory resulted in increases in the proportion of trips that were successful (Figure 4.3 shows increasing values for

proportion of successful trips along the x axis, weight assigned to memory). This was translated into increases of net energy gain, with values of 0.6 for the memory weight resulting in positive net energy gain at low detectability for maximum ingestion. At high prey detectability however, there is a clear tradeoff associated with placing weight on memory, with high foraging success when no weight is placed on memory, diminished foraging success at intermediate weights, and an increase when the amount of weight assigned to memory is high (Figure 4.3). This effect is most pronounced when ingestion rate is 50% of maximum, with intermediate weightings to the extremes resulting in negative net energy gains in the 24 hour trip / haul-out case (Figure 4.3d-f, j-l).

Extreme sensitivity to the way in which unvisited cells were treated when foraging using memory was also demonstrated. Two methods were considered when developing $f_R(\mathbf{x}_t \mid \mathbf{R}'_{t-1,t-2,\ldots,t=1})$, the visited resource distribution, 1) unvisited cells were assigned a value equal to the average of visited grid cells, and 2) unvisited cells were assigned a prey density of zero. Comparing these two methods (Figures 4.3 and 4.4) it appears that treating the unvisited cells as having a zero prey density is preferable, as both the proportion of trips that are successful and the net energy gain are higher (Figure 4.4 has higher values than Figure 4.3). The dip in foraging success with intermediate weights assigned to memory is also shifted to lower values when unvisited cells are assigned a value of zero, with increases in foraging success with a weight of 0.25 and higher (Figure 4.4).

There was marginal sensitivity of the results to the value for the prey persistence parameter. In general, higher values for prey persistence led to more successful foraging, and a larger range (wider distribution) for the proportion of trips successful and the net energy gain (e.g. Figure 4.4, compare interval width across panel columns).

4.3.2. Trip duration

Foraging success generally increased with trip duration. Given a fixed haul-out duration of 24 hours, both the proportion of trips that were successful and the total net energy gain over the 100 trips (standardised by 24 hr period) increased with the average trip duration (Figure 4.5). The additional activity costs of a 36 hr trip compared to the haul-out time therefore appeared to be balanced by the increase in biomass able to be encountered given the extended foraging period (Figure 4.5c). A foraging trip time of 12 hrs only resulted in positive energy gains when ingestion was at the maximum, persistence was either very high and considerable weight was placed on memory (Figure 4.5a), or when detectability was high (Figure 4.5a, dotted line).

Increasing the haul-out duration relative to the foraging trip time (foraging trip time 24h, haul-out duration 30h) resulted in slightly diminished foraging success (Figure 4.6, compare with Figure 4.3). Keeping the relative time between foraging trip duration and haul-out duration the same, but halving the 36h foraging time (foraging trip time 18h, haul-out duration 12h) also resulted in decreases in foraging success (Figure 4.7, compare with Figure 4.5c) although results from the (18h foraging 12h haul-out) scenario were almost identical to those in the base-case (24h foraging 24h haul-out).

4.3.3. Prey distribution

Foraging scenarios using the gadid prey base parameters perhaps unsurprisingly resulted in lower levels of foraging success (Figure 4.8) compared to the forage fish prey base. Although the gadid prey is less patchy (higher mean probability of a grid cell containing prey), the energy density of the gadid prey is half that of the forage fish. In addition, there appeared to be no increase in the proportion of trips successful with increasing weight assigned to memory. This difference in pattern is related to the parameters determining the spatial distribution of the prey and not solely due to the reduced energy density of the

gadid prey compared with the forage fish. The general trend of simulation results with respect to changes in the weight assigned to memory matched those in the forage fish case when the gadid energy density was used with the forage fish spatial distribution parameters(Figure 4.8c).

The frequency of foraging success with the gadid prey base increased when the foraging trip duration was increased to 36 hours, high weight was placed on memory, ingestion rate was at the maximum, and unvisited cells were assumed to have a prey density of zero (Figure 4.9a-c, g-i).

4.3.4. Body mass

Replacing the foraging sea lion with a juvenile of body mass of 120kg results in far fewer successful foraging trips, and less energy gain associated with those trips (Figure 4.10). This is a result of the increased mass-specific metabolic rate (Table 4.3). When foraging trip duration and haul-out time are both 24 hours, positive energy gain is only associated with high prey detectability unless unvisited cells are assigned a value of zero, and very high weight is placed on memory (Figure 4.11). Increasing the trip length results in an increase in foraging success as for the adult parameters. However positive energy gain is only associated with high weights to memory, when the persistence of the prey field between trips is high, and when unvisited cells are assigned a value of zero (Figure 4.11).

4.3.5. Correlations with Trip characteristics

Summary statistics of trip data (total prey within 20km of haul-out, mean and maximum distance from haul-out during foraging trip, trip length) were somewhat correlated with trip success. Across all scenarios for the forage fish prey base and adult sea lions, the amount of prey within 20km of the haul-out was not as good an indicator of foraging success as the maximum distance during the foraging trip, when the results of logistic regressions with the summary measures as predictors were compared using AIC (Table

4.4). In general, maximum distance from the haul-out and the total quantity of prey within 20km of the haul-out were more effective at distinguishing between successful and unsuccessful foraging trips than trip length and the degree of spatial usage (Figure 4.12), although logistic regressions including all five summary statistics as linear predictors was a better option (Figure 4.12f). The percentage of deviance explained by the individual term models was low (maximum of 13%, Table 4.4), with the model including all five summary statistics explaining 45% of the deviance. The degree to which summary statistics could explain whether individual trips were successful was highly dependent on foraging strategy. In particular, correlation of foraging success with metrics such as average and maximum distance tended to be higher when greater weights were assigned to memory while foraging (Figure 4.13). Similar dependencies of correlations between foraging success and summary statistics were also observed for other foraging strategy parameters, and for prey-related factors (prey density, prey detectability).

Summary statistics for juvenile foragers were less successful at identifying successful foraging, with low predicted probabilities of success for most successful foraging trips for all statistics (Figure 4.14). Results for gadid prey base were similar to those for the forage fish prey base (Figure 4.15), with the amount of prey within 20km appearing to perform best at distinguishing foraging success when the summary statistics were used in isolation.

4.4. Discussion

The simulations described in this paper synthesise a wide range of information and insight from various fields of study to evaluate how the spatial distribution of prey and foraging behaviour impacts the ability of Steller sea lions to obtain their energy budget. The spatial and temporal mismatch of data for the various components of the exercise limit the ability to estimate relationships among various factors directly, and so the

simulation framework provides a way to account for knowledge whilst at the same time recognising the uncertainty associated with linking the various hypotheses regarding interaction between species distribution, behaviour, and bioenergetics.

4.4.1. Steller sea lion foraging data: comparisons

Mean trip distance for juvenile Steller sea lions (>10 months, less than two years old) was 24.6 km (sd 57.2 km, Loughlin 2003). The analyses here did not consider life-history stage as a determinant to the foraging behaviour (although the bioenergetics parameters did determine what constituted foraging success). The distribution for average trip distance did of course depend on the foraging strategy parameters and prey field characteristics. Average trip distance across all the simulations for the forage fish prey base with trip length of 24h was 19.2 km, for successful foraging trips this was 14.8 km. The corresponding average trip distances with the gadid prey base were 19.8 km and 13.6 km. Using the juvenile bioenergetic parameters with the forage fish prey base, average trip distance for successful foraging trips reduced to 9.2 km. Mismatch between the simulation results and observed data could be due to differences in the prey exploited during foraging trips, further spatial restrictions on prey location, and alternative drivers of foraging behaviour such as managing tradeoffs associated with risk of predation from killer whales.

The values for haul-out and trip durations chosen for the simulations reflected those from observations of Steller sea lions (Table 4.2). Summer observations for immature sea lions indicate trip times in excess of 30 hours (Table 4.2, e.g. Trites et al. 2006), values that are consistent with parameters required in the simulations for successful foraging (positive energy gain). The frequency of foraging success increased with average trip duration for those scenarios tested, with energetic gains associated with increased foraging time offsetting the additional at-sea activity requirements. However, inspection of results for

trips with at-sea times in excess of 36 hours revealed a decrease in the proportion of trips that were successful.

Simulation results suggest that due to elevated metabolic costs, juvenile Steller sea lions need to rely on predictable prey, and place weight on previous experience when foraging. These results are consistent with movement of sea lions from rookeries in the breeding season to haul-out sites that are in the vicinity of high energy, predictable prey (Gende and Sigler 2006).

4.4.2. Prey field and bioenergetics: parameter uncertainty

The spatial depiction of prey fields in the simulations was simple, to enable the impact of foraging strategy to be separated from those of other factors. Further use of the results from Chapter 3 would entail incorporating variability in spatial parameters for the different prey types, modelling prey fields with multiple prey types, and incorporating the effects of bathymetry (and/or other covariates) on the spatial distribution of prey. Relationships for prey density with depth would likely further spatially restrict the prey field, with the exact impact dependent on how depth changed with distance from the haul-out, and the specific bathymetry used.

The simulations also did not consider uncertainty in the bioenergetics parameters. In particular, the relative values for the metabolic scalars for at-sea activity and the elevated metabolic rates for juvenile sea lions could be expected to change the overall levels of energy gain, and hence foraging success. The maximum ingestion rate used here (16% body mass) was taken from results with captive juvenile Steller sea lions (Rosen and Trites 2004). This is likely an upper limit for the digestive capacity of the adult sea lion, as stomach capacity tends to decrease in terms of % body size with increasing body size.

Ingestion rate naturally was an important determinant of energy gain. The simulations modelled ingestion at a constant rate that is the hourly proportion of the maximum daily

ingestion rate. Clearly this is an over-simplification, as individuals will eat until satiated. Alternative methods of modelling the dynamics of sea lion foraging could consider the digestive capacity, and allow for 'front-loading' of food upon encounter. Such a scenario might improve the success rate of some foraging strategies, as foraging success might depend more on individuals finding food as opposed to (possibly) needing to successively find food throughout the foraging trip. Reduced prey density was modelled as a function of the maximum ingestion rate. Consequently, the effective prey density modelled was different between the adult and juvenile simulations. Defining the simulated prey in terms of abundance rather than ingestion rate would also be complicated by non-linear relationships between prey density and consumption. Empirical data on the foraging efficiency of Steller sea lions with respect to prey encounter rate suggests that the prey density at which feeding efficiency is 50% of the maximum is at a prey density much less than 50% the prey density at which maximum foraging efficiency is observed (Cornick et al. 2006a). If the ingestion rate at each time step was reduced linearly with the distance travelled (assuming an average swimming speed of 4.7 $km.h^{-1}$), then high prey detectability did not result in successful foraging when not assigning weight to memory when foraging (Figure 4.16). In such a case, larger moves are penalised as the (instant) return at that time step is decreased because less time is available for foraging. Figure 4.16 shows the potential impact of accounting for a functional response in the prey encounter rate, although this could be modified to include a non-linear response (e.g. Cornick et al. 2006a), or by including the reduced expected prey encounter rate in the foraging model.

4.4.3. Alternative models and foraging strategies

State-space models with behavioural switching between two states (e.g. Jonsen et al. 2005), nominally representing searching and foraging, have seen increasing application in the fitting of movement models to data from remotely tagged animals, including pinnipeds. A common approach is to model animal tracks as correlated random walks

(CRW). Such approaches are frequently not mechanistic, i.e. the movement data are used to infer behaviour, which is associated with certain states, athough Hidden Markov Models have begun to be used as a framework for incorporating covariates into state transitions (e.g. Patterson et al. 2009). For example, a state associated with low movement or frequent dives might imply a bout of foraging on a patchily distributed resource. This type of inference is necessary because it is rare to have simultaneous data on the actual foraging activity and prey availability, although it has been used to estimate characteristics of foraging behaviour and distinguish foraging strategies (e.g. Breed et al. 2009). A dynamic state variable model applied to foraging for juvenile Steller sea lions in Prince William Sound was used to make inference about tradeoffs between foraging behaviour and predation risk (Frid et al. 2009), although this model did not consider uncertainty associated with prey resource distribution. Similarly, a population modelling framework including patch dynamics of sea lion prey was able to capture salient features of the Steller sea lion decline (Matthiopoulous et al. 2008), but did not address questions related to how foraging behaviour and uncertainty in prey distribution combine to determine sets of (un)successful foraging strategies. The methods employed in this chapter address foraging from a process-based perspective, that is the consequence of foraging movement (attaining locations of prey) is modelled as a driver of movement. This does not imply a statistical framework. An extension of the analyses here could be to use the foraging model to generate 'data' akin to that available from foraging individuals, and test whether methods frequently employed to estimate foraging behaviour (e.g. state-space models fitting CRWs) can be used to distinguish among the different strategies.

The limited set of foraging strategies presented here provided a means to evaluate the possible impacts of alternative prey dynamics on the performance of potential strategies. In principle, the tested foraging strategies could be expanded ad infinitum. For example, the energy condition of the forager could be calculated at each time step as a dynamic variable, representing current foraging success. Such a metric could then be utilised as a

parameter of the foraging strategy (such as a critical threshold for determining when to 'give up' and return to the haul-out). Also, foraging within the simulations was decidedly 'local', alternatives could be to pre-select preferred foraging locations at the start of the trip (either randomly or based on previous experience), and model movement toward these locations in the same way as done for homing. However, the results of the simulations presented in this chapter clearly demonstrate tradeoffs associated with different foraging strategies and how these respond to spatial and temporal dynamics of prey fields. Foraging behaviour in Steller sea lions is known to be very plastic, with sea lions being able to adjust both metabolic rates and behaviour in response to limited food intake.

4.4.4. Summary measures as proxies for foraging success.

The foraging simulations produced estimates for quantities such as space-use (home range), attendance time, average trip distance, etc., and allowed for evaluation of how these may change given the constraints on foraging discussed. Given that estimates of these quantities are available from satellite telemetry data, natural avenues for additional work that build on the material described here would involve addressing what monitoring schemes (or data choices) best enable the ability to detect both foraging behaviours, and whether sea lions would be able to achieve energy budgets. However, correlations of trip statistics with foraging success indicate that foraging strategy (e.g. weight assigned to memory, assumptions about how to deal with unvisited grid cells) has the potential to obscure relationships such that prediction of foraging success from such information is difficult.

4.4.5. Conclusion

Simulation results show that successful foraging can be achieved over a wide range of strategies. However, movement to locations where prey was previously located is most necessary when prey density is reduced, or when metabolic rates are high (e.g. for

juveniles). Observations of movement of juvenile sea lions to haul-outs close to predictable sources of high energy prey appear to be consistent with these findings. Correlating field data with foraging success can be expected to be hampered by unknown determinants of foraging behaviour. However, availability of a suite of data types from foraging, including information on prey biomass can likely lead to identification of instances where foraging may be successful, and suggest locations where management actions might be required.

Table 4.1. Specifications for the Prey fields. Values correspond to different scenarios.

Parameter	Symbol	Values	Source
Mean probability of prey patch	μ	0.97, 0.02	Chapter 3
Variance of spatial field	σ_w^p	2.77, 5.33	Chapter 3
Spatial correlation (at 2km) in probability of prey patch.	ϕ^p	0.25, 0.75	Chapter 3
Mean prey patch density (relative to maximum sea lion ingestion rate)	$B_{i,\mathbf{x}_t}$	0.1, 0.25, 0.5, 1.0	
Persistence ($24h^{-1}$)	ρ	0.1, 0.25, 0.5, 0.75, 0.9	
Energy density of prey (kJg^{-1})	ED_i	4 (gadids), 8.5 (small schooling fish)	Winship (2000) summarises multiple studies

Table 4.2. Movement parameters, and specifications for the foraging strategies.

Parameter	Symbol	Value	Source
Standard deviation of movement (h^{-1})	σ_m^2	2, 3.5	Average swimming speeds of 4.7 $km.h^{-1}$, 7.2$km.h^{-1}$ (Rosen and Trites 2002b, Cornick et al. 2006a)
Trip duration (h)		18, 24, 36	Merrick and Loughlin 1997, Trites et al. 2006, Loughlin et al. 2003
Haul-out time (h)	T_{land}	12, 24, 30	19h for adult females, 25h for winter young of the year (Merrick and Loughlin 1997)
Weight afforded to prior experience (versus moving randomly given local conditions)	θ	0, 0.25, 0.5, 0.75, 0.95	

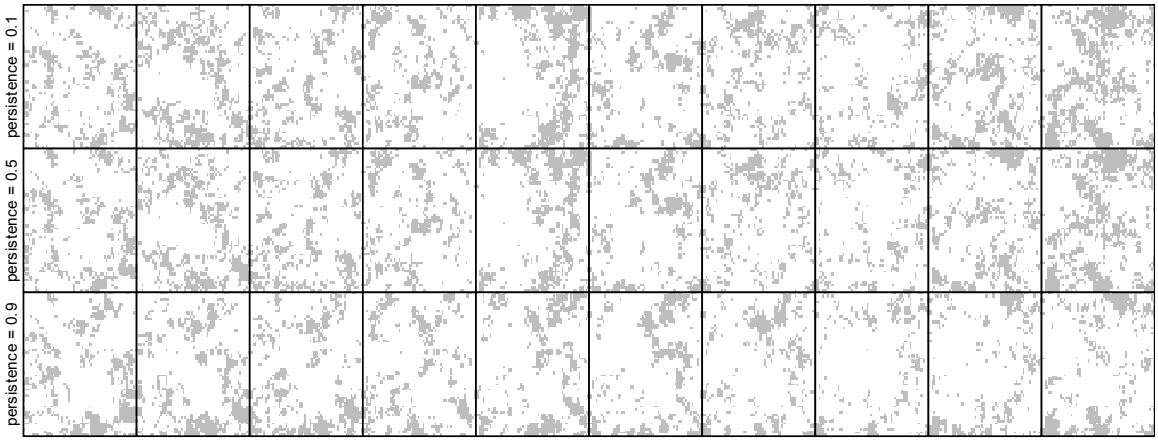
Table 4.3. Parameter values and sources for the bioenergetics model.

Parameter	Value		Source
	Juvenile	Mature	
Sea lion biomass	120 kg	250kg	Winship et al. 2001
Maintenance Metabolic Rate	Twice adult.	$292.88M^{0.75}$	Winship et al. 2002, Rosen and Trites 1997, Kleiber 1975
A^{land}, activity multiplier when hauled out	1.2	1.2	Winship et al. 2002
A^{sea}, activity multiplier for when at sea	4	4	Winship et al. 2002
E_{HIF}, heat increment of feeding	$-0.013\ ED_i+0.229$		Winship and Trites 2003
$E_{dig,i}$, digestive efficiency	$0.951 / [1 + EXP(-1.86(ED_i-2.1))]$		Rosen and Trites 2000, Winship and Trites 2003
Maximum ingestion rate (% body size)	16		Rosen and Trites 2004

Table 4.4. Logistic regression results of predicting foraging trip success with individual summary statistics, for the adult sea lion simulations foraging on forage fish prey.

Summary Statistic	Intercept	Slope	% deviance explained	ΔAIC
Prey within 20km of haul-out	-1.394	0.0286	12.2	74581
Average trip distance	1.648	-0.296	7.2	510638
Maximum trip distance	2.131	-0.210	13.1	0
Spatial usage (# grid cells visited)	1.634	-0.102	9.2	344051
Trip time	-1.822	0.0768	6.4	582560

(a) Forage fish



(b) gadid

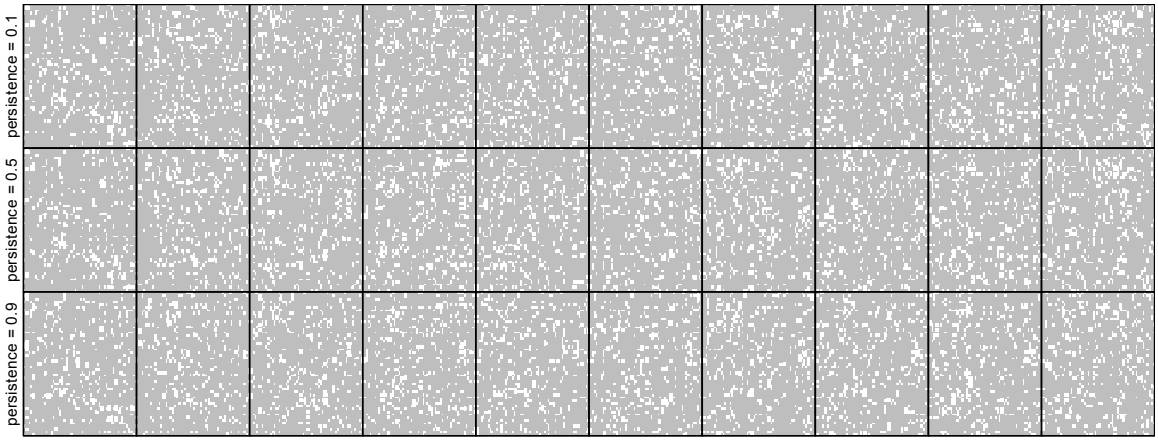


Figure 4.1. Example evolution of prey field over ten trips (columns) given different values for the persistence parameter for each prey type.

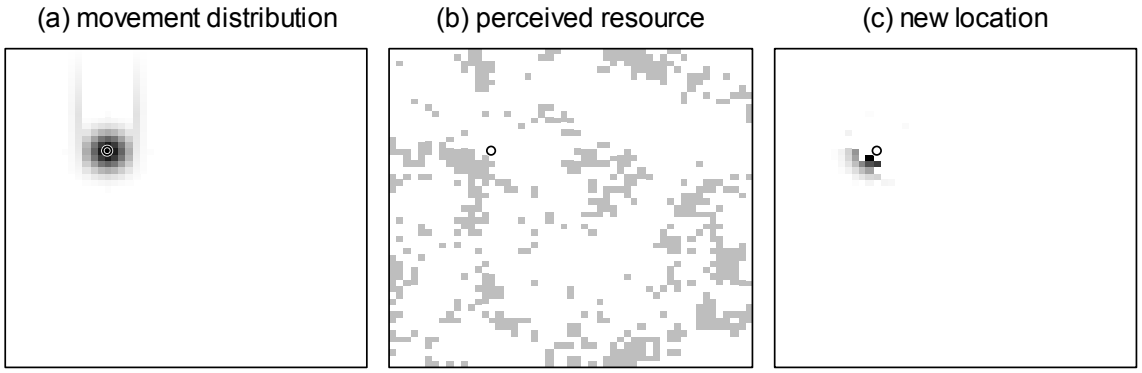


Figure 4.2. Example of movement step during foraging. At a given time step, the probability distribution of possible moves based on swimming speeds (a), is multiplied by the perceived resource distribution (b) to obtain the distribution for the new location (c). In this case, prey are perceived to be southwest of the current position and so the new location distribution contains possible moves only in this direction.

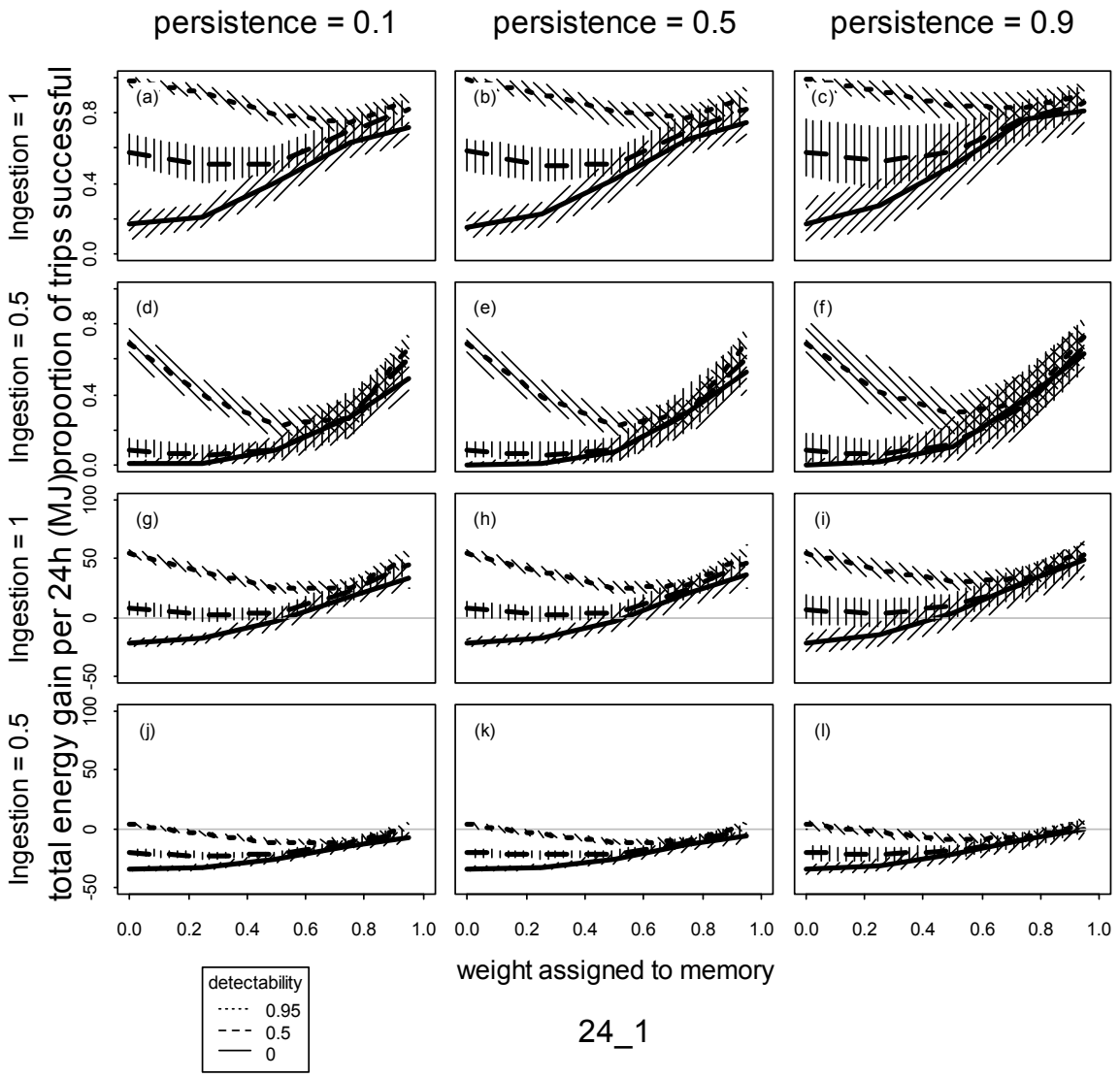


Figure 4.3. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), when unvisited cells are set to the average of those visited in the memory map given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a foraging trip time of 24h, and haul-out duration of 24h.

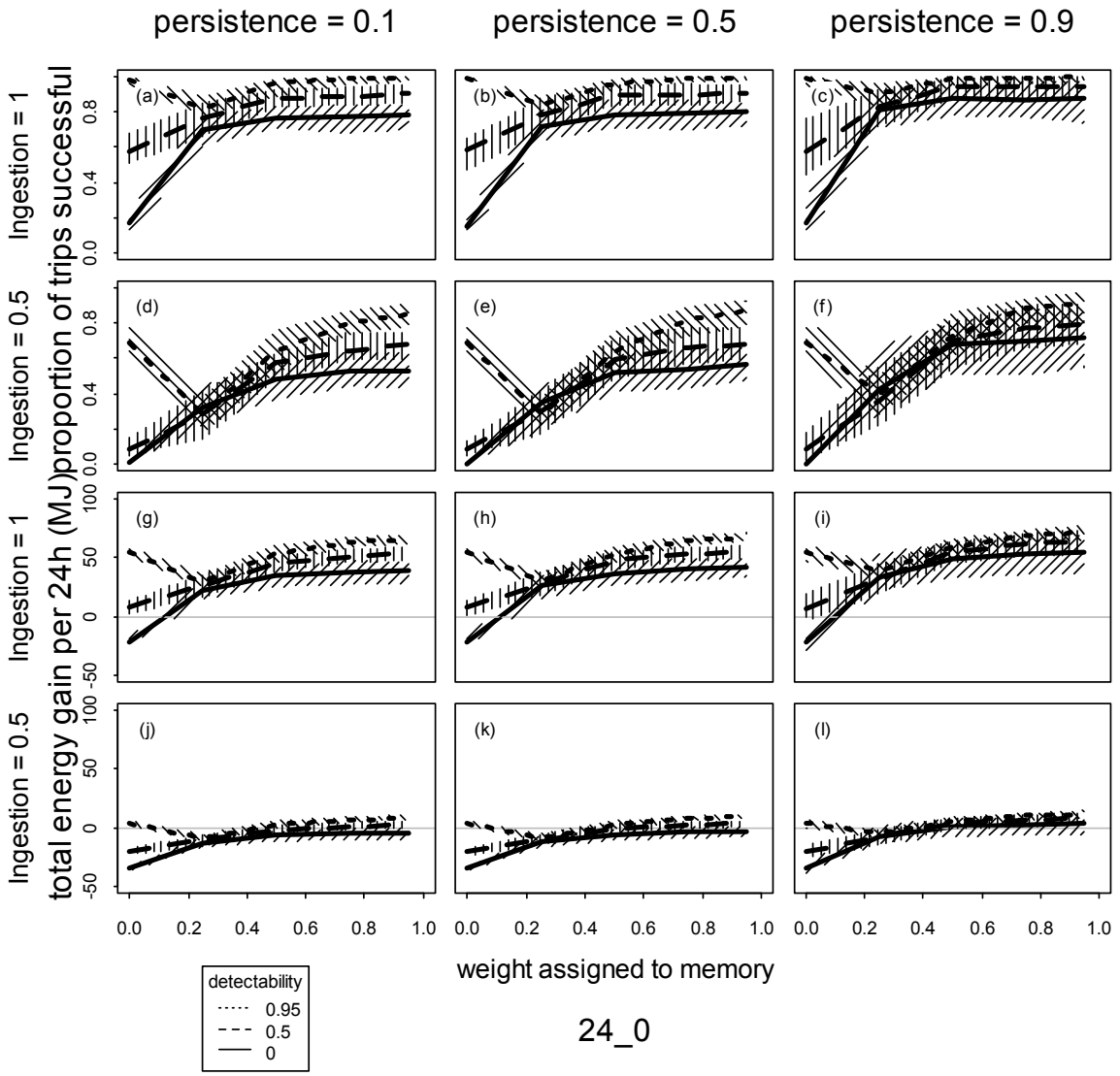


Figure 4.4. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a foraging trip time of 24h, haul-out duration of 24h, when unvisited cells are assumed to have a prey density of zero in the memory map.

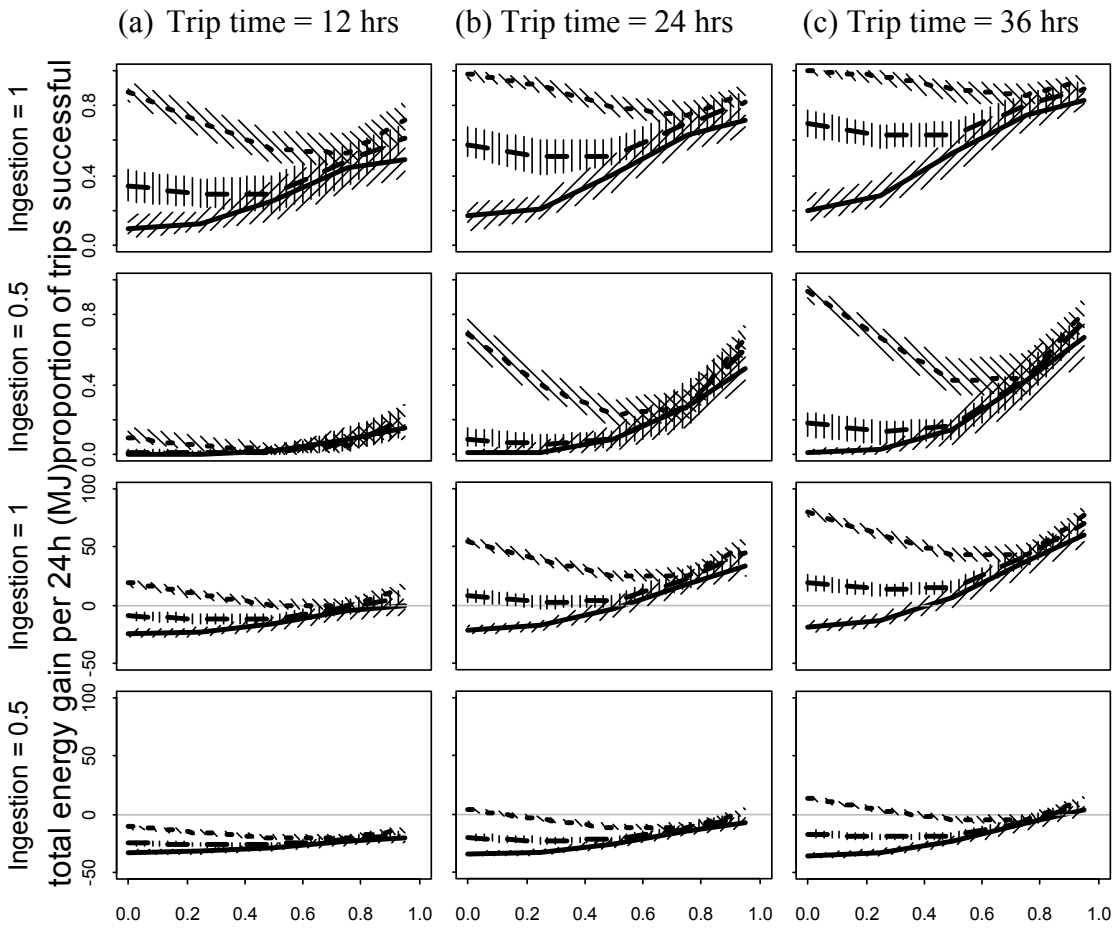


Figure 4.5. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), given trip times of 12, 24 and 36 hours (columns), prey detectability, prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a haul-out duration of 24h, prey persistence of 0.1, and when treating unvisited cells as average of those visted in the memory map.

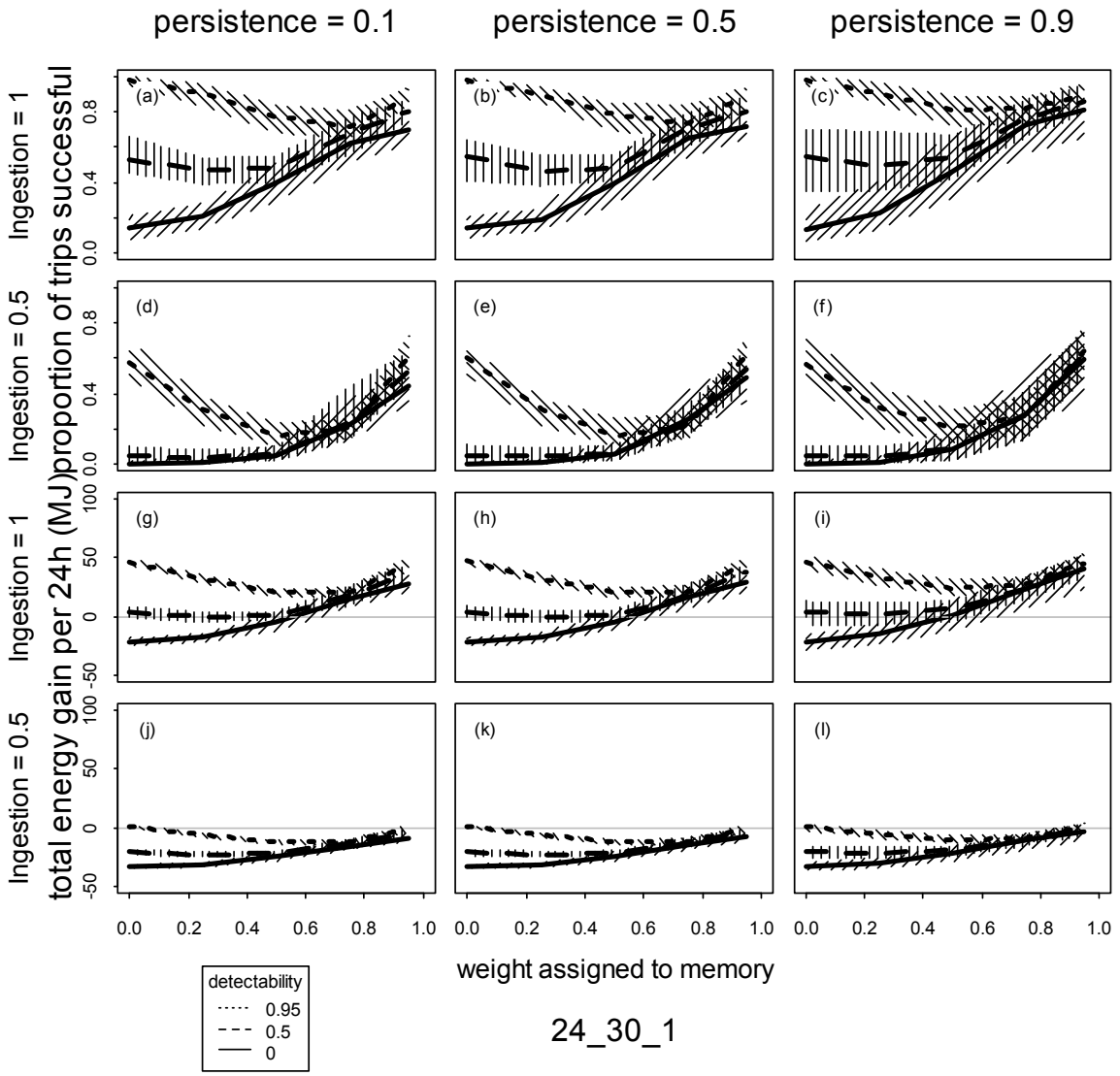


Figure 4.6. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a foraging trip time of 24h, haul-out duration of 30h, when treating unvisited cells as average of those visted in the memory map.

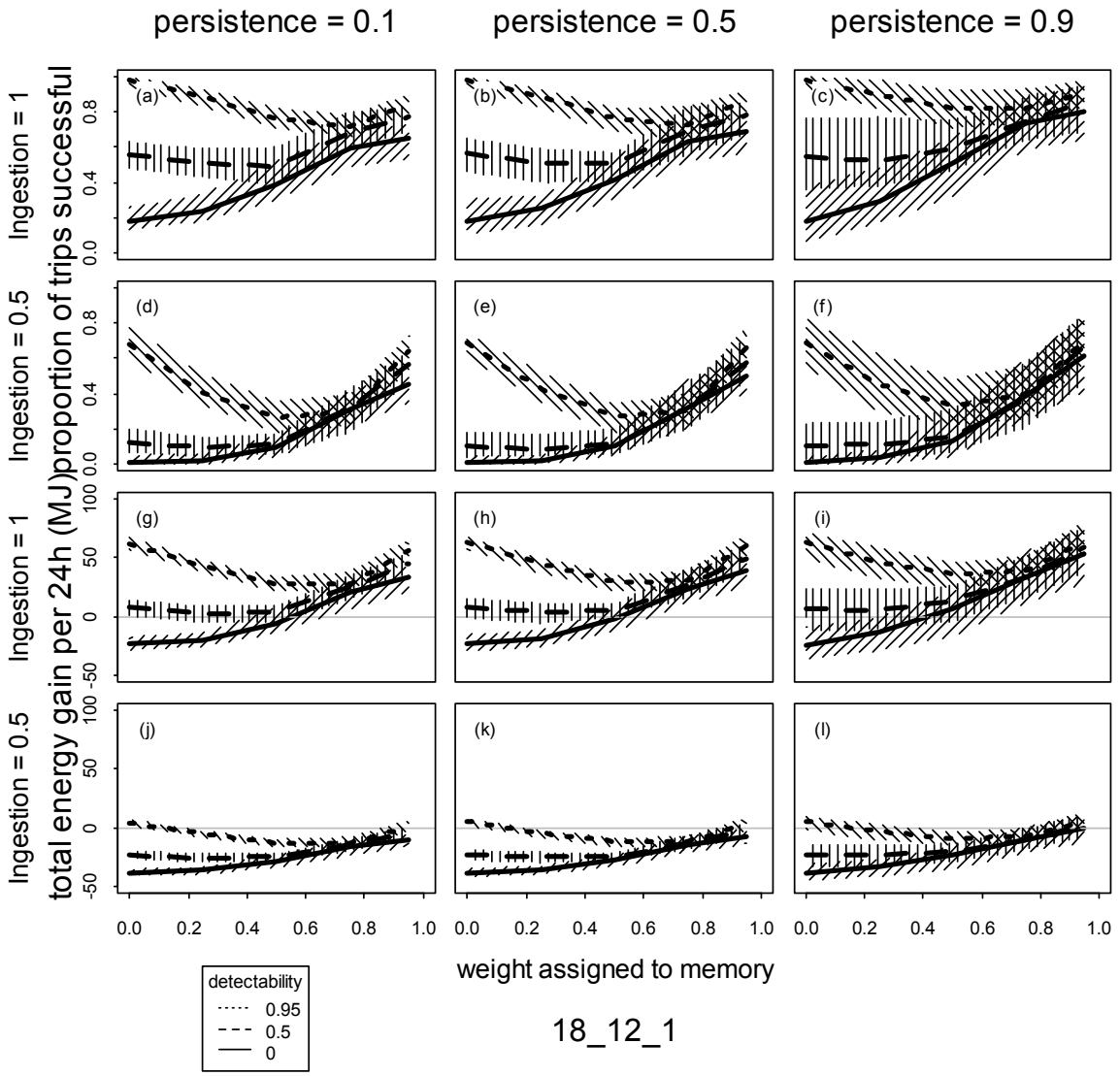


Figure 4.7. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a foraging trip time of 18h, haul-out duration of 12h, when treating unvisited cells as average of those visted in the memory map.

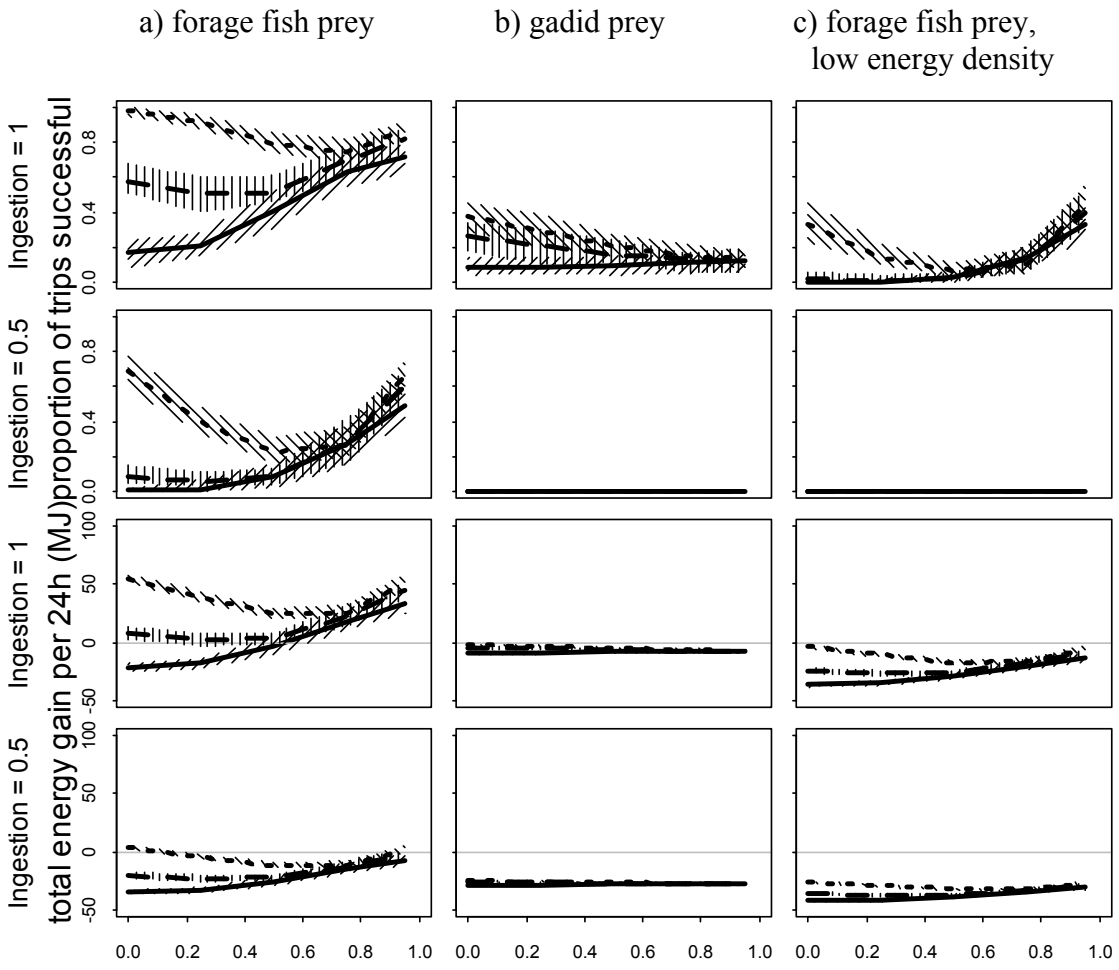


Figure 4.8. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), for a) the base scenario with forage fish prey base, b) gadid prey base, and c) forage fish spatial parameters but with gadid energy density. Values are shown for the range of prey detectability, prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a trip-time of 24h, haul-out duration of 24h, prey persistence of 0.1, and when treating unvisited cells as average of those visted in the memory map.

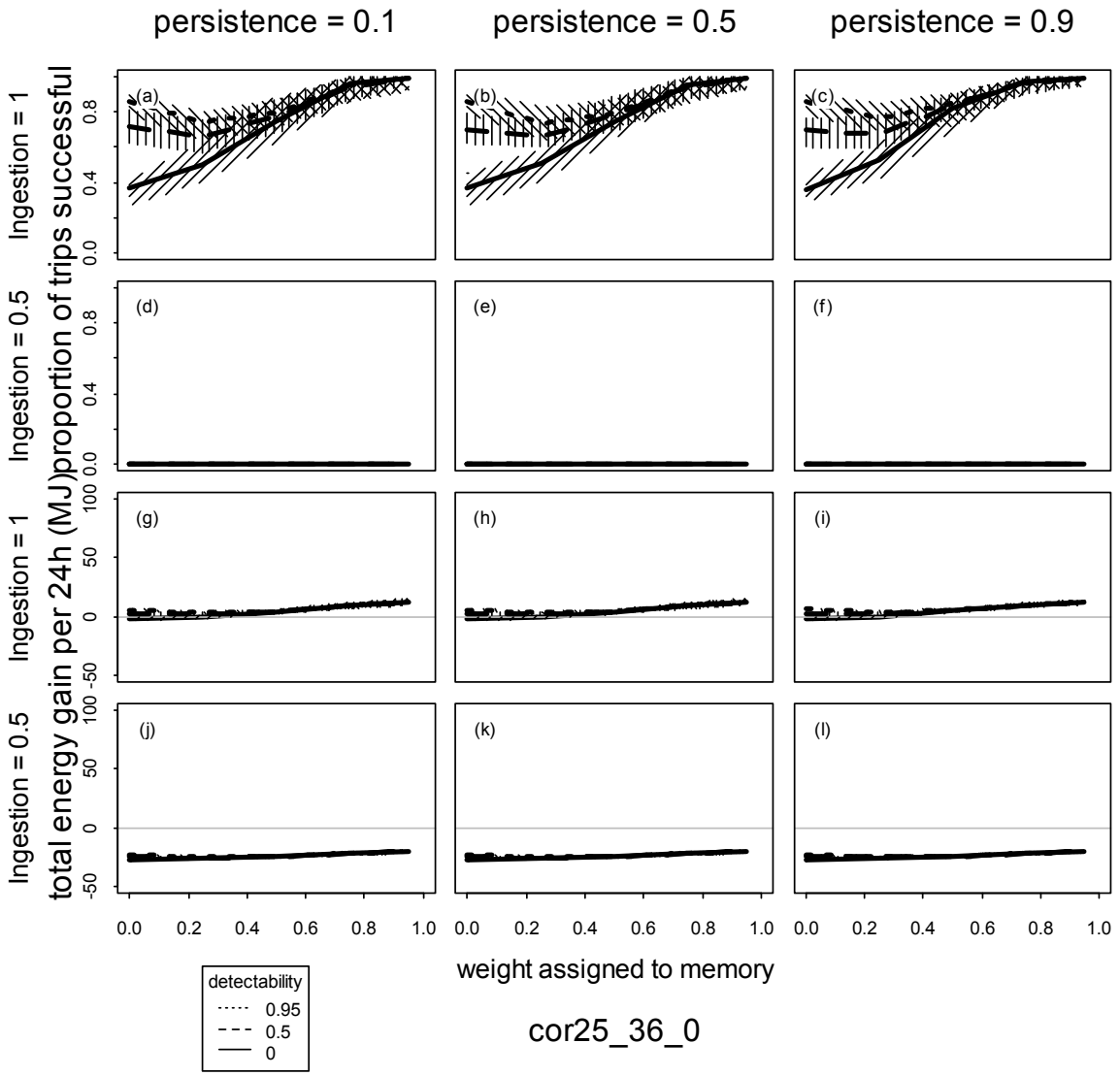


Figure 4.9. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), for the gadid prey base, given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a foraging trip time of 36h, haul-out duration of 24h, when unvisited cells are assumed to have a prey density of zero in the memory map.

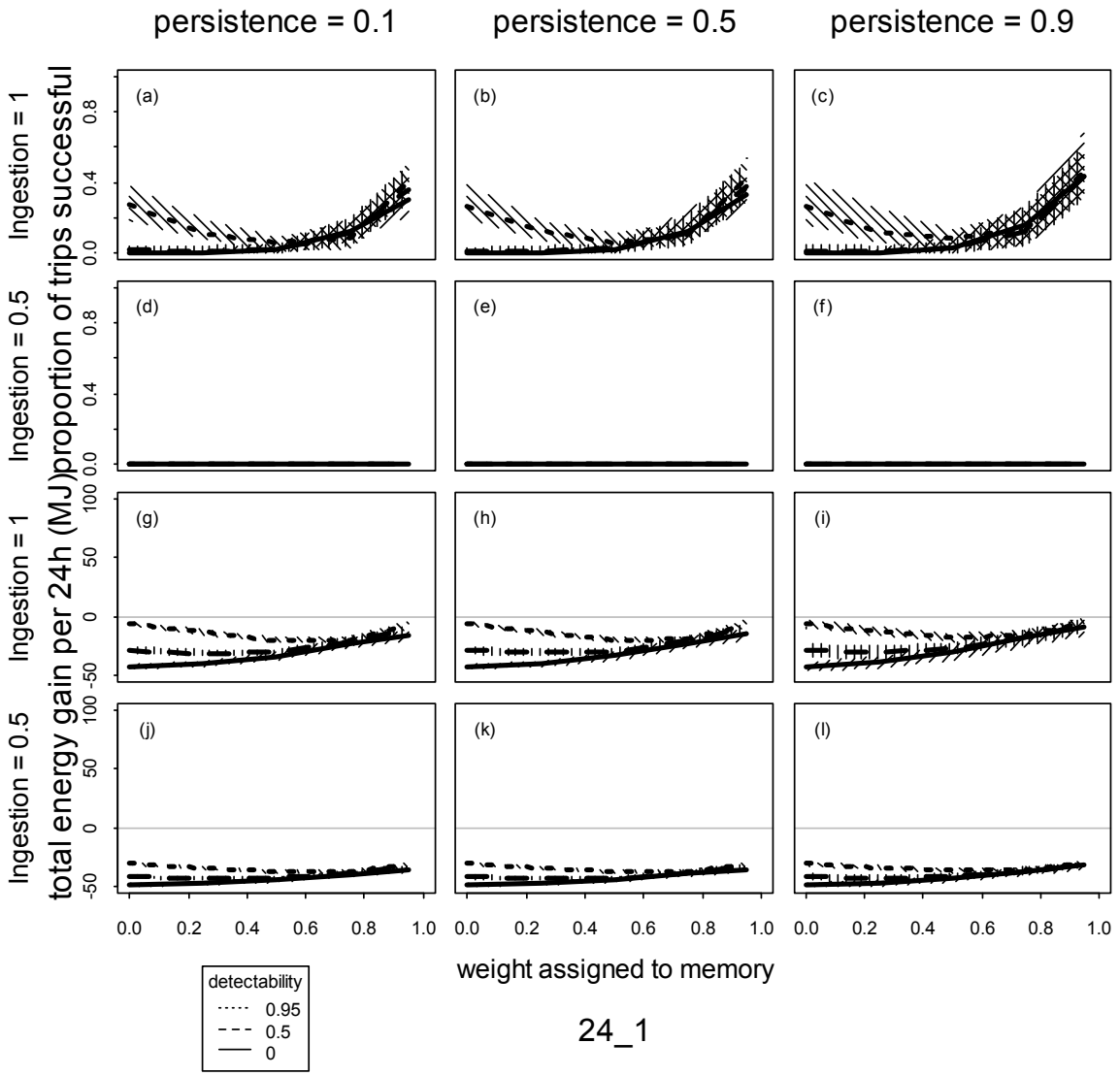


Figure 4.10. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), when foraging is by a 120kg juvenile sea lion, given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), given a foraging trip time of 24h, haul-out duration of 24h, when treating unvisited cells as average of those visted in the memory map.

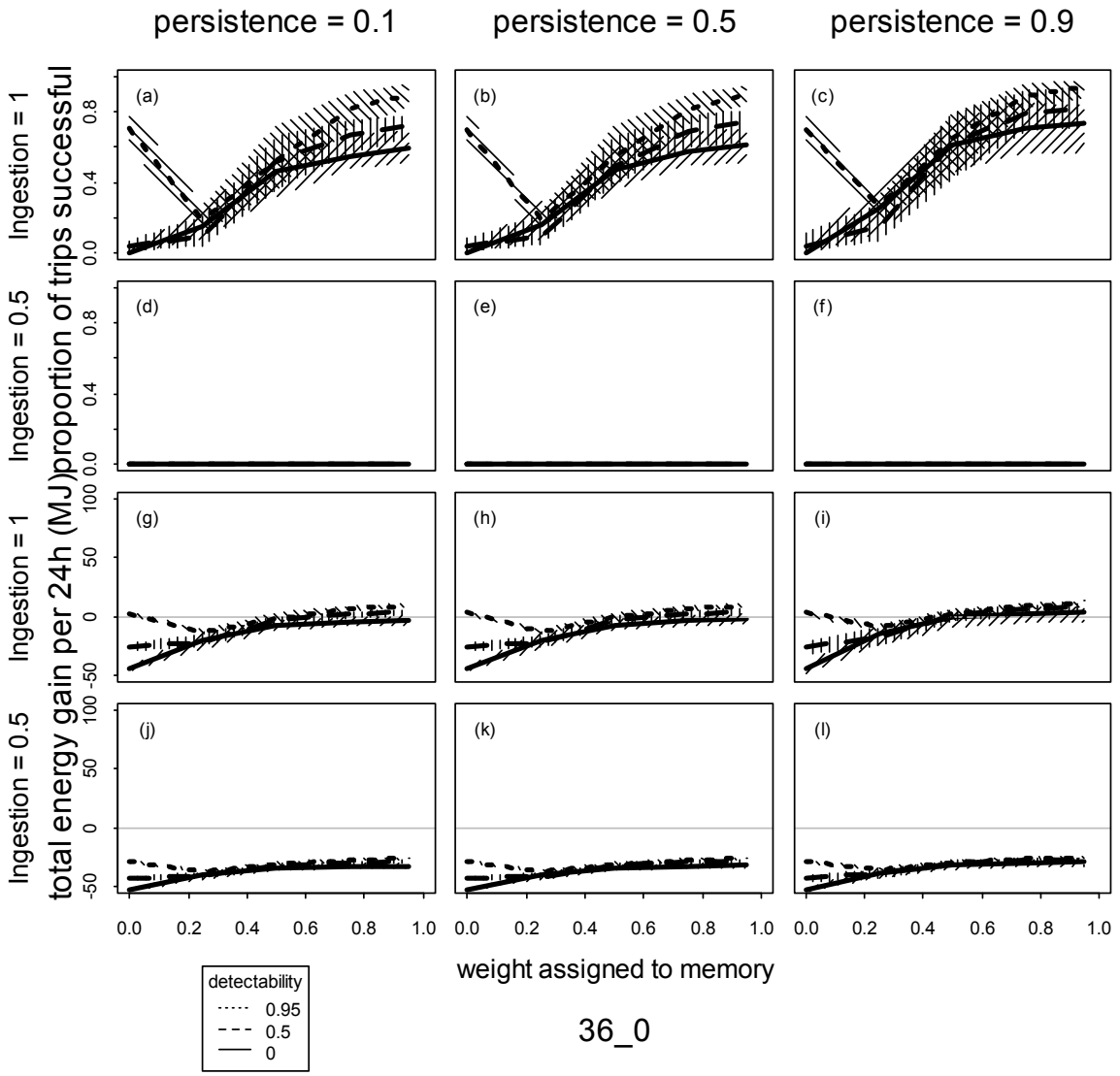


Figure 4.11. Median (line) and 95% distribution (shading) for the proportion of trips that are successful, and the net energy gain (per 24h), when foraging is by a 120kg juvenile sea lion, foraging trip time is 36 hours, and unvisited cells in the memory map are treated as having a value of zero, given prey detectability (lines), prey persistence (columns), prey density (Ingestion, rows) and weight assigned to memory when foraging (x-axis), and haul-out duration of 24h.

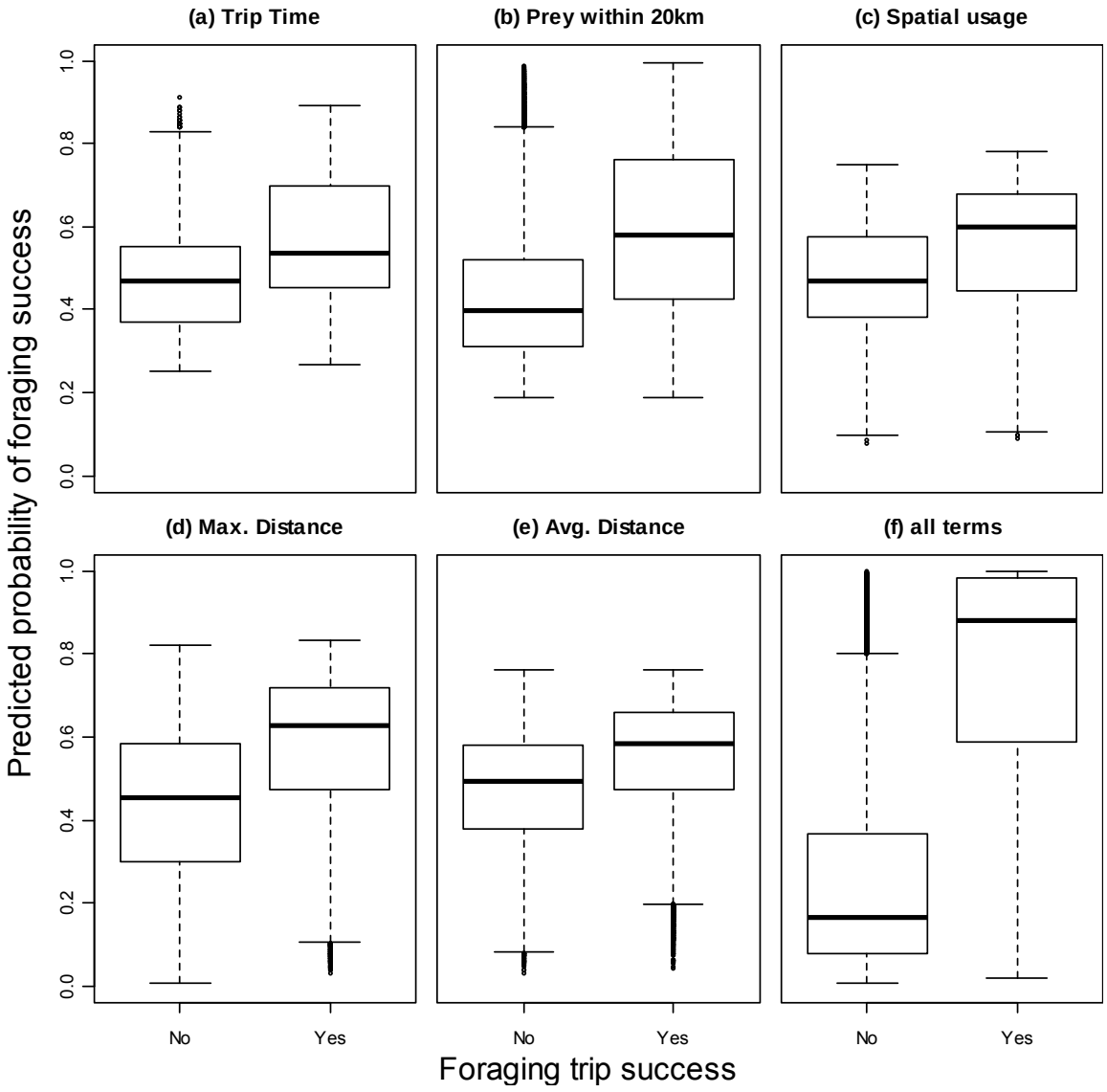


Figure 4.12. Predicted probabilities of trip foraging success when the true success is either true or false (positive net energy gain), from logistic regressions including individual trip summary statistics (a-e) and all statistics (f) as linear predictors. Good model performance would be a predicted probability of 0 for 'No' success and 1 for 'Yes'.

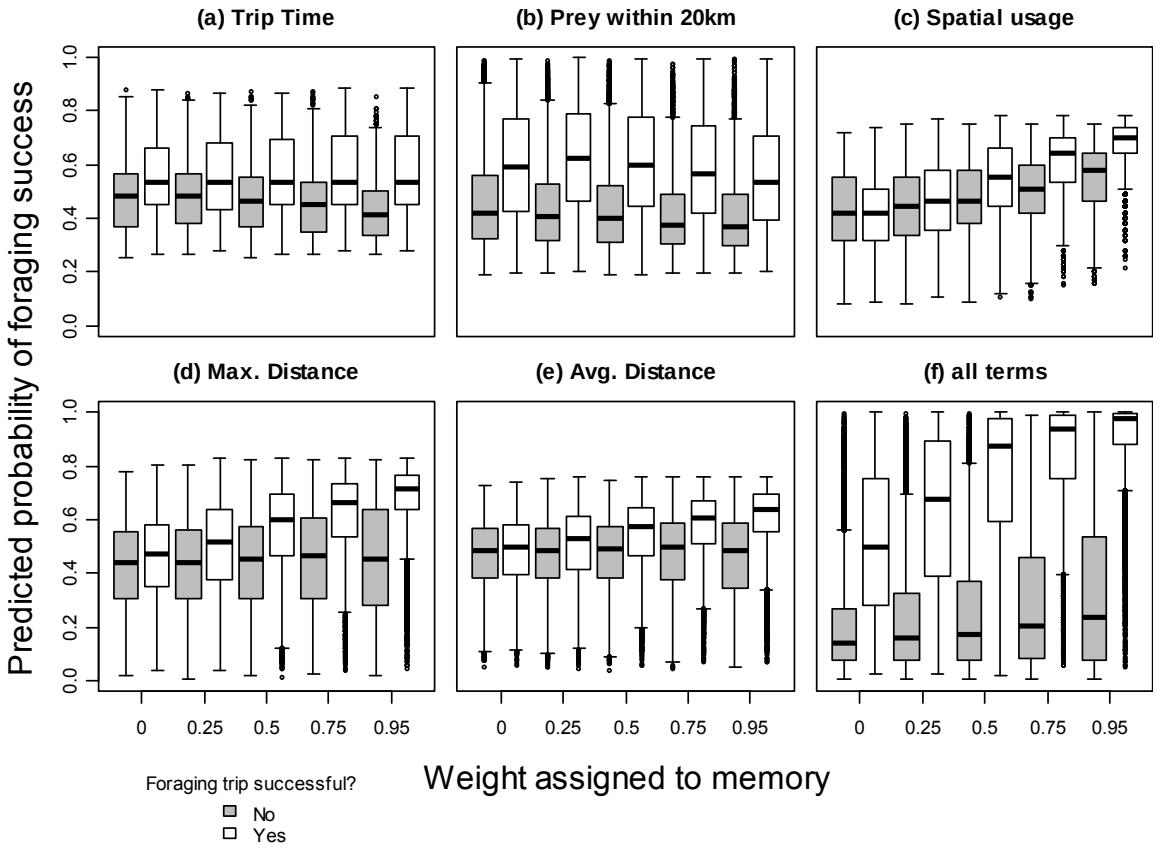


Figure 4. 13. As for 4.12 but split by the weight assigned to memory in the foraging strategy.

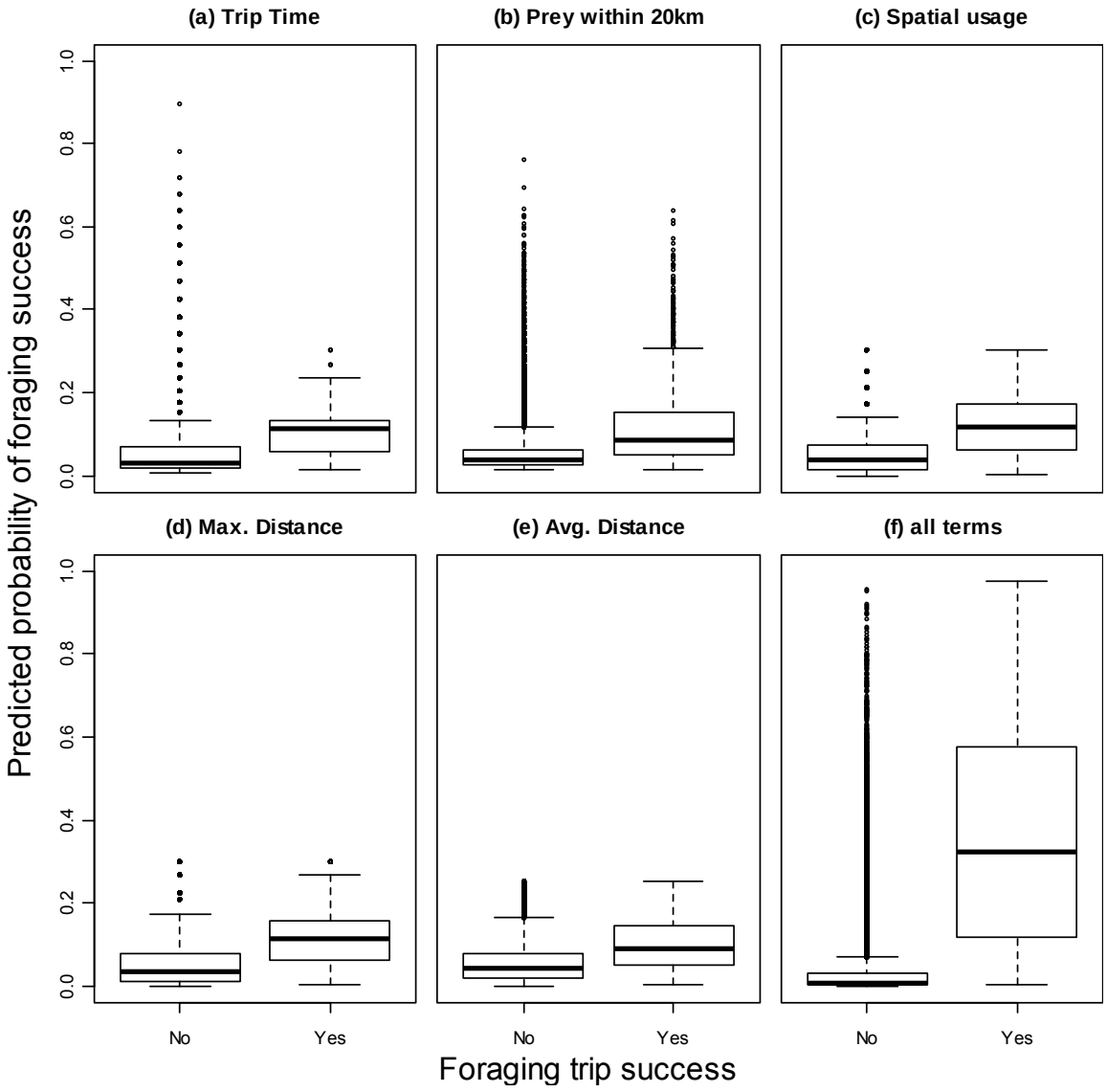


Figure 4.14. As for Fig. 4.12 but with juveniles.

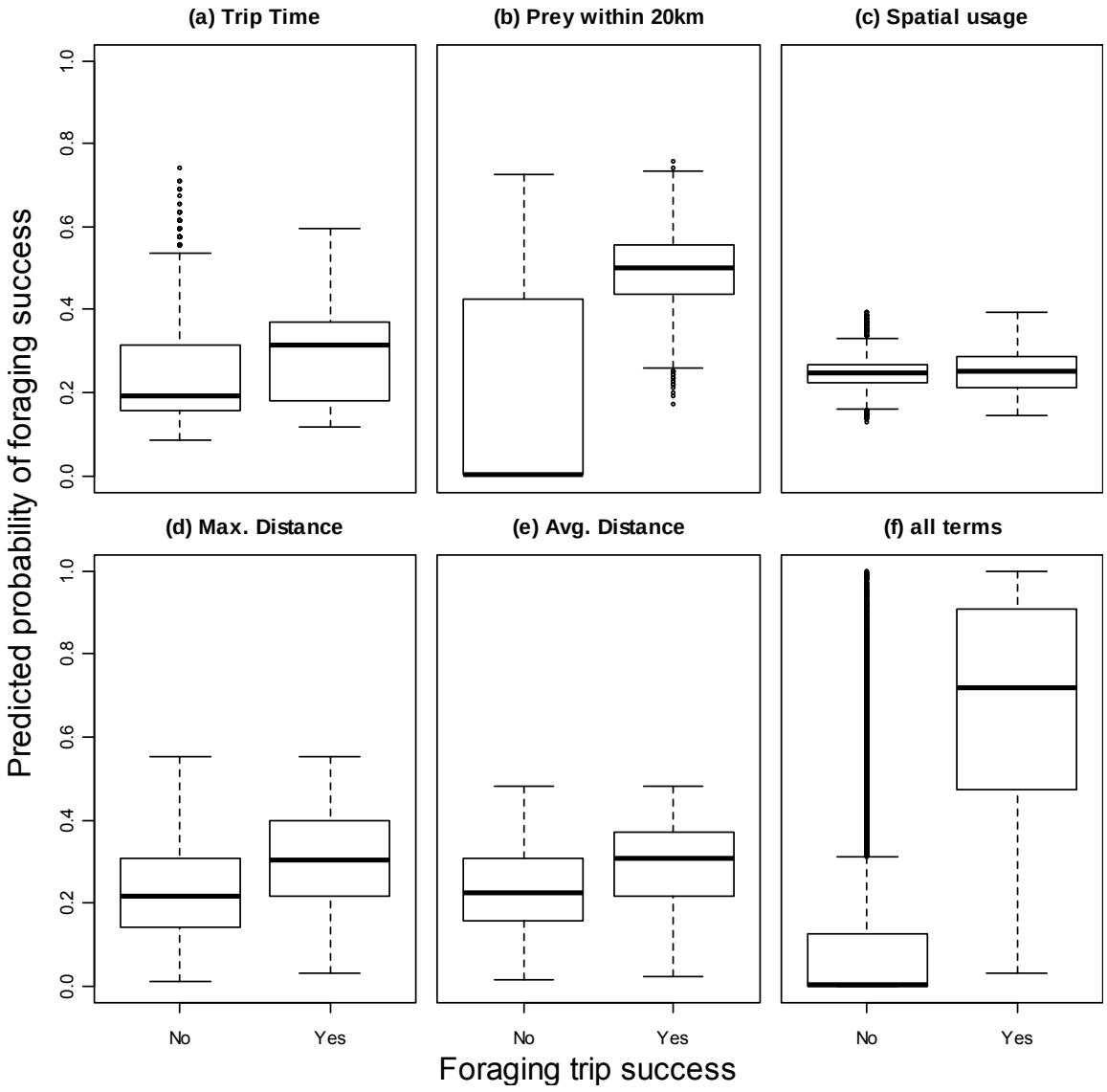


Figure 4.15. As for Fig. 4.12 but with the gadid prey base.

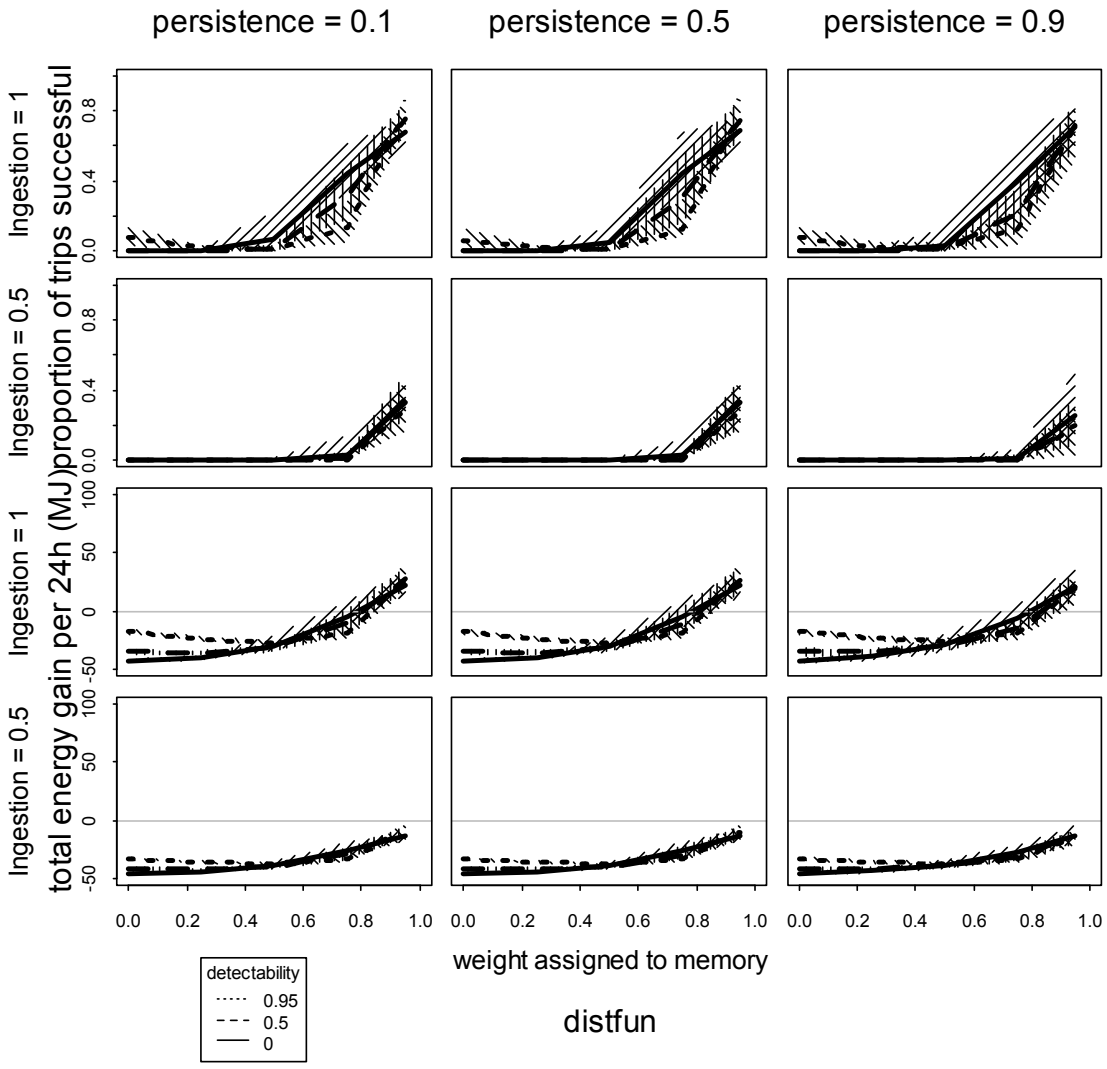


Figure 4.16. Foraging success when the prey encounter rate is modelled as a linear function of the distance travelled at each time-step, for the base scenario with forage fish prey base, a foraging trip time of 24h, haul-out duration of 24h, and when treating unvisited cells as containing prey density equal to the average of those visited in the memory map.

Chapter 5: Impacts of spatial uncertainty on performance of age structure-based harvest strategies for blue eye trevalla (*Hyperoglyphe antarctica*)

5.1. Introduction

Harvest strategies (also called Management Procedures) are well recognised as effective tools for conservation of natural resources and have been applied widely in fisheries management, principally in output control, data-rich fisheries (e.g. Butterworth et al. 1997, Butterworth and Punt 1999, Cooke 1999, Kell et al. 1999, 2005). Harvest strategies consist of the following components: data collection schemes, assessment methods, and harvest control rules (HCRs, Table 5.1). The latter translate indicators from stock assessments into specifications for management actions (e.g. Restrepo and Powers 1999). A successful HCR should provide an appropriate response to deviations from management targets, be robust to key uncertainties, and emphasise precautionary action given uncertainty. The latter point is particularly important for so-called 'data-poor' situations, when the reliability of indicators is likely questionable. The Management Strategy Evaluation (MSE) approach offers powerful, well-developed simulation-based method for comparing the performance of and testing the overall robustness of HCRs (e.g. Butterworth and Punt 1999, Smith et al. 1999, De Oliveira et al. 2008).

The blue eye trevalla (*Hyperoglyphe antarctica*) is a high-value species in Australia's Southern and Eastern Scalefish and Shark fishery (SESSF), caught primarily on the continental slope at depths of 300-600m (Haddon 2011) since the 1970s. The fishery for this long-lived, late-maturing species is characterised by multiple gear types (initially dropline, now mainly auto-longlining, trawl, with some gillnetting) operating in a range of areas, with uncertainty in stock structure (Hindell et al. 2005), spatial and seasonal variability in the availability of different age classes, and low levels of sampling effort across the fishery (Smith and Wayte 2002, Fay 2007). Scientific advice for management

in the SESSF takes the form of a Recommended Biological Catch (RBC) for each species under quota management (including blue eye trevalla), used to inform the setting of the Total Allowable Catch (TAC) by species (Smith et al. 2008). At present, the TAC for blue eye trevalla applies across all fishery sectors, because there are few measures in place to allocate the TAC spatially[5] (a separate TAC is applied for one sector of the fishery, the trawl fishery in the Great Australian Bight [GAB]).

The SESSF adopted a formal harvest strategy framework as a basis for setting RBCs in 2005 (Smith and Smith 2005, Smith et al. 2008). This framework is based on a tier system of HCRs, with tier assignment based on the type of information available to determine stock status. The tier framework is intended to foster the precautionary approach, in that as information quality decreases and progression through the tiers proceeds, HCRs ought to lead to lower RBCs, and result in higher average levels of spawning biomass. The SESSF harvest strategies specify a biomass level B_{LIM} (currently 20% of unfished spawning biomass), below which targeted fishing should cease, and a target biomass B_{TARG}. The HCRs specify a maximum fishing mortality rate that defines overfishing (F_{LIM}), and a target fishing mortality rate that defines optimum utilization (F_{TARG}). Increases in the uncertainty around stock status for the more data-poor tier levels is reflected in the application of discounts to (i.e. reductions of) the RBC, which are intended to achieve the same end as a decrease in the target fishing mortality rate as uncertainty about stock status increases.

[5] Catches by blue eye trevalla in the trawl fishery in the Great Australian Bight (GAB) are not included in the SESSF TAC, although the catch by other gears in this area are. This is not a major sector of the fishery, catches by trawl in the GAB have been at most on the order of 1-2% of the annual total catch for blue eye trevalla in the SESSF.

Many so-called "data-poor" fisheries lack the information to conduct rigorous quantitative stock assessments. In such cases, HCRs are required that can use simple indicators of stock status from limited data. These indicators may be contradictory among the different sectors of the fishery, either as a result of differences in gear selectivity, sampling protocol, catch-history driven changes, or spatial structure in the population dynamics. The need to address inconsistent data streams is well-documented in fisheries stock assessment (Richards 1991, Schnute and Hilborn 1993). However robustness to biases associated with such effects may be more important when the quantity of data available is low. In the SESSF, the data-poor "Tier 3" HCR has been applied to blue eye trevalla (e.g. Fay 2009). This HCR is designed for stocks for which there is no estimate of current biomass, but where an estimate of the current fishing mortality rate, F_{CUR}, is available, most frequently from the results of catch curve analysis applied to age composition data. The Tier 1 HCR (Smith et al. 2008) is applied when there is ample information to fit an integrated stock assessment model (e.g. Stock Synthesis, Methot 2007), but this is not the case for blue eye trevalla. As the SESSF harvest strategy framework was not tested before it's initial implementation in 2005, it is not clear how well the tier framework of HCRs performs, and indeed whether scientific advice for management is more precautionary for species managed using the Tier 3 HCR than would be the case had the species been data-rich and managed under Tier 1. Finally, it is not clear how best to cope with possibly conflicting information from multiple areas and gear types.

This chapter uses MSE to evaluate the performance of variants of the Tier 3 HCR for blue eye trevalla under scenarios that represent key uncertainties. Specifically, we examine how well these variants satisfy management targets and objectives while also considering risks of unfavourable outcomes. The variants of the Tier 3 HCR utilise different estimation methods and/or use different reference points (including assumptions regarding the stock-recruitment relationship). HCR performance is considered both when

there is no spatial structuring of the population or fishery, and when there are two regions in which the fishery operates, with uncertainty related to exploitation pattern and selectivity by region, and also given various assumptions regarding the spatial structure and degree of mixing of the fished stock between regions. The spatial scenarios enable testing of methods to account for situations where spatial variability in the data from the fishery has the potential to impact management performance.

5.2. Methods

5.2.1. Simulation protocol

Performance of the harvest strategy framework for blue eye trevalla is evaluated using a simulation modelling framework that incorporates feedback between the HCR and the population dynamics. Attention is focused in this section on describing the HCRs and the modifications made to their implementation, rather than describing the technical details of the operating model, which are provided in full in Appendix C. The general approach on which the operating model is based consists of tuning a spatial age-structured model to represent a set of scenarios for the dynamics of the blue eye trevalla population and fishery. The values for the parameters of the operating model are not based on the results of a stock assessment, as no model for the population dynamics exists at present for blue eye trevalla in the SESSF. Rather, values for parameters were either sourced from the literature, derived from previous analyses for this fishery, or selected to mimic characteristics of the available data for blue eye trevalla given hypotheses in the literature regarding the dynamics of blue eye trevalla.

The operating model is first projected over a historical period given the known catch history for blue eye trevalla, and age-composition data are generated given the known ‘true’ population dynamics represented by the operating model. The chosen variant of the HCR is then used to determine the RBC for the following year(s), given an estimation

method (catch curve analysis, the stock assessment) and the selected parameters governing the HCR. The RBC is then allocated to fleet and region within the operating model, the population size in the operating model is updated given this new catch, and additional data are generated. This assessment / population update cycle is repeated for 20 years, with annual assessment and updating of the RBC. A scenario is defined as the combination of a set of parameter values for the operating model, a data collection scheme, and a specific variant of the Tier 3 HCR. One hundred simulations were conducted for each scenario, each differing due to process error in the population dynamics (annual recruitment), observation error when generating the age-composition data, and error associated with implementation of the estimation method and application of the HCR6. A series of summary statistics are calculated at the end of the 20 year projection period. The summary statistics for each simulation are further analysed to derive a set of performance measures, which are used to compare results among scenarios.

5.2.2. Operating model

The operating model consists of an age-structured population dynamics model that can be parameterised to include spatial regions (with movement of fish among regions), and multiple fleets, to capture key dynamics for blue eye trevalla (Appendix C). Analyses detailed in this chapter consider two versions of the operating model: (a) a single population occupying a single region and exploited by a single fishing fleet, and (b) a population occupying two regions with movement between regions, and exploited by one or two fishing fleets (with different selectivity patterns). Several parameterisations of

---

[6] The results obtained from 100 simulations were almost identical to those obtained when 1,000 simulations were used to characterise a scenario for a subset of the scenarios described.

each version of the operating model are considered to investigate the implications of key uncertainties. The parameterisation of the operating models, along with the values for biological parameters for blue eye trevalla considered for the various scenarios, are given in Tables 5.2 and 5.3, and Figure 5.1. Values for the parameters determining growth and maturity were obtained from published literature on blue eye trevalla in Australia and New Zealand. Blue eye trevalla recruit into the fishery around age 2 yr and lengths of 40-50cm as they move from a pelagic phase on to the continental slope, and females mature at 12 years of age, at an average size of 72cm (Horn 1988, Baelde 1996, Smith and Wayte 2002). Blue eye trevalla are long-lived, with estimated rates of natural mortality of the order of 0.08yr-1 – 0.1yr-1 (Fay 2009, Smith and Wayte 2002). Larger spawning size fish appear to be only seasonally available to the fishery. Selectivity patterns estimated in previous stock assessment work (Fay 2007, 2009) were used to specify the patterns used in the simulations. Blue eye trevalla may move large distances, both during the pelagic phase and after settling onto the continental shelf and slope (Smith and Wayte 2002).

Scenarios using the two-region version of the operating model were designed to mimic general assumptions regarding the nature of the blue eye trevalla fishery, rather than the actual spatial structure. Two 'continental slope' regions with differing exploitation histories were assumed, with levels of stock mixing between the two regions, ranging from full mixing, in that the impacts of spatial variability are minimal, to almost no mixing, indicating a high degree of spatial structuring in the population. Spatial differences in population responses to exploitation are more likely to be observed under the limited mixing scenario. The regional catch histories used to drive the population dynamics before implementation of the HCRs (Figure 5.1f) are taken from the landings data from the relevant zones of the SESSF, with the geographic split in these data being catches taken east and west of Tasmania. The fishery for blue eye trevalla initially developed in the eastern area during the 1970s, primarily off Tasmania. Catches west of Tasmania however, are now on the order of those in the east (Figure 5.1f).

5.2.3. Harvest strategies

The harvest strategies consist of a data collection scheme, a method to estimate the current fishing mortality rate FCUR, and an HCR. Scenarios are limited to instances where the harvest strategy is applied every year of the projection period, consistent with the current practice of annual setting of TACs within the SESSF. The two forms for the Tier 3 HCR in Figure 5.2 are tested, with two methods for estimating FCUR. Variants of the HCR that utilise different reference points and have differing data requirements are implemented as outlined below.

5.2.3.1. Data and estimation methods

Data available for the Tier 3 analyses were limited to fishery-dependent age-composition data (i.e. no index of abundance or fishery-independent data), with an annual multinomial sample size of 100 allocated by fleet and region in the same proportions as the annual catch (i.e. the total annual sample size was 100). While this sample size is substantially lower than the number of otoliths that have been aged annually in recent years for blue eye trevalla (on the order of 500 per year), it is perhaps unreasonable to think that these data truly constitute a random sample of size equal to the number of aged otoliths (e.g. Miller and Skalski 2006, Candy 2008). A random sample of 100 represents a reasonably good sample size that might be hoped for from SESSF species. Evaluation of HCR performance when sample size is reduced or not randomly determined from the catch was not considered in this chapter. Four years of age-composition data were assumed to be available to the estimators each time they were applied (the most-recent four years). Two catch curve estimation methods were employed: (a) the Chapman and Robson (1960) catch curve estimator (CR), and (b) a multi-year equilibrium F age-structure based-estimator (MYEF). The estimators aim to estimate total mortality, Z, with estimation of F then achieved given an assumed value for the rate of natural mortality, M (denoted "assumed M" in Tables 5.2 and 5.3). For the CR method, catch curves were applied to the

annual age-composition data, with FCUR calculated as an inverse-variance weighted average of the annual estimates. In contrast, MYEF integrates over all years, therefore averaging over years is not required to obtain an estimate of FCUR. for this estimation method.

a) Chapman and Robson catch curve estimator (CR)

The CR estimator assumes that the population is in equilibrium, and that recruitment is constant over time. The estimate of Z, from a sample of the age composition for a given year is:

$$Z_y = \ln\left(\frac{1+\overline{a}_y - 1/n_y}{\overline{a}_y}\right) \tag{5.1}$$

where $\overline{a}_y$ is the mean age (above the recruitment age) of the sample and n_y is the sample size for year y. A single estimate of Z is required to calculate the RBC, and so weighted averages of estimates of Zy from the most recent four years of age-composition data were calculated, with weighting inverse to the variance estimate for each year:

$$Var(Z_y) \approx \frac{\left(1-e^{-Z_y}\right)^2}{n_y e^{-Z_y}} \tag{5.2}$$

Catch curve estimators are known to be sensitive to the age-range of the data used (Chapman and Robson 1960, Dunn et al. 2002). The recruitment age was set separately for each of the four years at that for which the numbers–at-age were greatest, with the maximum age being determined from the sample. The CR estimator assumes uniform selectivity for ages above the recruitment age, likely biasing estimates of vulnerable biomass when selectivity is actually dome-shaped.

(b) Multi-year equilibrium F age-structure based-estimator (MYEF)

Estimation of FCUR using MYEF involves fitting an equilibrium-based age-structured population dynamics model to the available age-composition data, with the population model being of the form:

$$N_a = \begin{cases} 1 & \text{if } a = 0 \\ N_{a-1} e^{-(s_{a-1}F+M)} & \text{if } 0 < a < 100 \\ \dfrac{N_{a-1} e^{-(s_{a-1}F+M)}}{\left(1 - e^{-(s_a F+M)}\right)} & \text{if } a = 100 \end{cases} \tag{5.3}$$

where the N_a's are the numbers-at-age, sa is the (estimated) selectivity at age (assumed to be asymptotic and to follow a logistic curve), F is the estimated rate of fishing mortality, M is the assumed rate of natural mortality, and a = 100 is modelled as a plus-group. The values for F and the parameters which define sa are determined by maximizing the following log-likelihood function:

$$\ln L = \sum_{y=Y-3}^{Y} n_y \sum_a O_{y,a} \ln\left(\frac{\tilde{N}_a}{O_{y,a}}\right) \tag{5.4}$$

where $O_{y,a}$ is the observed proportions in the sample by age in year y, ny is the sample size for year y, Y is the most recent year of data, and $\tilde{N}_a$ are the predicted proportions of catch-at-age:

$$\begin{aligned} \tilde{N}_a &= \bar{N}_a / \sum_{a'} \bar{N}_{a'} \\ \bar{N}_a &= \frac{N_a s_a F}{\left(s_a F + M\right)}\left(1 - \exp^{-(s_a F+M)}\right) \end{aligned} \tag{5.5}$$

Maximisation of Equation 5.4 was achieved using AD Model Builder (Fournier et al. 2012). Differences between MYEF and the CR estimator are that MYEF accounts for selectivity, data from all ages are used, and the likelihood is multinomial. Unlike the CR estimator, no averaging of annual mortality estimates is necessary to calculate the RBC under MYEF because F is calculated using all the available data simultaneously.

The scenarios outlined in Tables 5.2 and 5.3 include uncertainties related to applying the estimation methods. Importantly, the impact of assuming the incorrect value for M when conducting the estimation is examined.

5.2.3.2. Harvest Control Rules

Concern about the performance of the Tier 3 HCR was noted following implementation (Klaer et al. 2009). There was concern that the original nature of the calculation of RBCs for Tier 3 stocks (applying an appropriate multiplier to recent average catch levels, Figure 5.2b) could lead to a ratchet effect of continually increasing or decreasing catches, even though information suggests that the target level has been reached. A revised harvest control rule (Klaer et al. 2009), which shows consistency with the more data-rich tier rules in terms of reference points, was applied in 2008 (Figure 5.2c). Unlike the previous rule, this revised rule does not cap annual catch increases.

Each of the scenarios outlined in Table 5.2 were projected using three variants of the Tier 3 HCR (Figure 5.2), which differed either by adopting the 'old' or 'new' Tier 3 rule, and in the choice for the target and limit reference points:

T3a The shape of the HCR follows the 'old' rule (Figure 5.2b), with $F_{TARG} = M$,

T3b The shape of the HCR follows the 'new' rule (Figure 5.2c), with $F_{TARG} = 0.5M$ and $F_{LIM} = M$, and

T3c As for T3b, but with the reference points adopting a Tier 1-like approach with $F_{\text{TARG}} = F_{40}$, and $F_{\text{LIM}} = F_{20.}$ (F_{40} and F_{20} are the fishing mortality rates which will result in [under equilibrium age structure] spawning biomasses of 40% and 20% of unfished spawning biomass [corresponding to the B_{MSY} proxy and B_{LIM} under the SESSF Tier 1 HCR]).

The values for M used in the HCRs (and that used to calculate F) are the 'assumed M' values as detailed in Tables 5.2 and 5.3.

Empirical investigation suggests that the assumption $F_{MSY} \approx M$ is too optimistic for blue eye trevalla (Figure 5.3). Walters and Martell (2004) suggest $F_{msy} = cM$, with values for c including 0.8 in general, but 0.6 or less for commonly fished species (Walters and Martell 2004). For U.S. west coast groundfish species, the average is c = 0.62 (MacCall 2007), and so 0.5M was chosen for the analyses here to adopt a conservative estimate. Calculation of F40 and F20 under T3c depends on the values for the parameters of the stock-recruitment relationship (assumed to follow a Beverton-Holt form), and requires estimates of the steepness parameter h (Mace and Doonan 1988), information on growth and fecundity, and selectivity in addition to an estimate of M. The values for these reference points used in the HCR were calculated based on the estimates of selectivity from the estimators, an assumed value for h of 0.75 (irrespective of the true value for h), and the 'correct' values for growth and fecundity (Figure 5.1a-c). In contrast, T3a and T3b rely only on an estimate of M to calculate the RBC given FCUR. However, T3c more appropriately accounts for biology when determining the likely response to fishing.

Calculation of the RBC under T3a is achieved by applying the appropriate multiplier from Figure 5.2b to Ccur, defined as the average catch over the four years prior to the year for which an RBC is needed. Under T3b and T3c, the RBC is calculated by first obtaining FRBC given Figure 5.2c, and applying the formula:

$$RBC = C_{cur} YPR(F_{RBC}) / YPR(F_{cur}) \quad (5.6)$$

where $YPR(F)$ is the yield-per-recruit obtained given a fishing mortality F. Note that equation (5.6) allows for greater increases in catch than does T3a, the old Tier 3 HCR (maximum increase of 20% above the recent average, Figure 5.2b), if FCUR is estimated to be below the target level. Irrespective of this, the maximum allowable change in the catch (RBC) from one year to the next was restricted to 50%7.

Comparison of Tier 3 performance with that expected under Tier 1 is achieved by calculating the projected spawning stock biomass trajectories for a subset of the scenarios in Table 5.2 under the Tier 1 HCR. This involved generating additional data (cpue (initially 20 yrs, CV=0.3), and length composition (initially 10 years, ny=100)) during both the historical and projection periods, and applying Stock Synthesis 2 (Methot 2007) to this data set each year of the projection period. Results for Tier 1 HCR are simply shown for comparison purposes because the focus of this chapter is the Tier 3 HCRs.

5.2.3.3. Accounting for fleet/spatial structure

Uncertainty in spatial structure through the scenarios in Table 5.3 presents additional challenges when implementing the Tier 3 HCR. Estimates of the fishing mortality rate may differ by fleet and region, and so decisions must be made regarding what combinations of fleet and region are to provide the parameters used when calculating the RBC. The scenarios in Table 5.3 were crossed with the following options to investigate how performance given spatial structure changes with assumptions regarding how the data are analysed:

---

[7] Within the harvest strategy framework of the SESSF, a rule exists where the TAC cannot change by more than 50% from one year to the next.

1. spatial complexity is ignored, and a single analysis (CR, MYEF) is conducted using the data pooled across regions (added together as samples are allocated by region relative to the size of the catch) – this is how the Tier 3 control rule is applied in the SESSF at present;

2. the data from the two regions are analysed separately to obtain two estimates of current fishing mortality / stock status; these estimates are then weighted by the inverse of their variance estimates to obtain the RBC; and

3. separate analyses as in 2, but the maximum estimated F is used to calculate the RBC.

The variance estimates for FCUR are (primarily) driven by sample size, and so option 2 effectively weights the regional estimates by the current catch allocation. Option 3 is potentially more conservative because it bases the RBC on the parameters for the region with the highest estimated fishery mortality rate. However, this option can be expected to be more prone to inaccurate estimates of F that might result from low sample sizes.

5.2.4. Performance measures

Performance of the HCRs is evaluated using a set of performance measures (abbreviations in parentheses):

1. The median (over simulations) spawning stock status at the end of the projection period (final spawning biomass as a fraction of unfished spawning biomass, B_0), [median final depl.].

2. The inter-quartile range of the spawning stock status (relative to B_0) at the end of the projection period [IQR final depl.].

3. The probability of the spawning biomass being below the Tier 1 limit reference point (B_{20}) at the end of the projection period [P(final B < B_{LIM})].

4. The probability of the spawning biomass going below the Tier 1 limit reference point (B_{20}) at some point during the projection period [P(B < B_{LIM} anytime)].

5. The median of the average annual catch during the projection period [median (avg. TAC)].

6. The median (over simulations) of the CV of the annual catches during the projection period [median(CV TAC)].

7. The mean (over simulations) number of years for which the RBC is less than 4t[8] [mean(#yrs collapse)].

Performance measures 1-4 relate to the effect of implementing the HCR on spawning biomass, while measures 5-7 provide information regarding the catch performance of the HCR.

5.3. Results

The results of the simulations are displayed as boxplots of the performance measures across scenarios to compare among the HCRs and methods for obtaining *F* estimates.

[8] 4t is 1% of the total catch prior to implementation of the harvest strategy framework, this performance measure is intended to reflect the frequency of fishery collapse.

Simple linear models are used to evaluate the contribution of the different scenario specifications to the values for the performance measures. The scenario characteristics as defined in Tables 5.2 and 5.3, the catch curve estimation type, the choice of HCR, and (for the spatial analyses) the option for obtaining a single F estimate, were included as factors in the linear predictors of these models, fitted separately for the seven performance measures. Interaction terms involving some of the factors were also considered.

5.3.1. Non-spatial analyses

The performance of the three HCRs, given estimation using CR and MYEF, are compared in Figure 5.4 which aggregates the performance measures over all scenarios, i.e. the horizontal bar in Figure 5.4a is the median over all scenarios (Table 5.2) of the median final spawning depletion of the 100 simulations run for each scenario. T3a, the old Tier 3 HCR, led to levels of spawning biomass that are well below the Tier 1 target and limit biomass reference points (40 and 20% of unfished spawning biomass) at the end of the projection period, (Figure 5.4a) with high probabilities of dropping below B_{20} during the projection period (Figure 5.4d). The performance of the T3b and T3c HCRs varied considerably among scenarios, and the projections under these HCRs led to more variable catches (Figure 5.4f). However, the results for T3b and T3c were generally more optimistic regarding stock status than T3a (Figures 5.4a, 5.4c), although many scenarios still remained below the target biomass at the end of the projection. Comparison of performance for scenarios 1-3 suggests that, for these scenarios at least, the Tier 3 HCRs are not precautionary compared to the Tier 1 HCR, because the Tier 1 HCR led to higher relative biomass, a lower probability of dropping below the limit reference biomass, and lower, less variable annual catches (Figure 5.5).

The changes in performance with respect to stock status, the risk of dropping below the limit reference point, and magnitude in catch levels were largely independent of the

estimation method (CR versus MYEF), but were sensitive to the choice of HCR used (Table 5.4, T3b and T3c rows). Performance of the new Tier 3 HCR was also sensitive to the choice of reference points. T3c, the HCR based on the spawner-recruitment reference points, tended to produce higher relative spawning biomass, lower probabilities of dropping below the limit reference point, and lower median annual catches than T3b (Figure 5.4a, c-e). However, performance was variable among scenarios, and the probability of dropping below the limit biomass was very high for several of the scenarios (Figure 5.4c-d).

Variability in the values obtained for the performance measures was associated with life history characteristics and the initial stock status. Scenarios where the true value for steepness was low ("*h=0.3"*) resulted in lower final biomass, an increased probability of dropping below B_{LIM}, and increased variability in the annual catches (Table 5.4). Likewise, more productive stocks ("*h=1.0"*) resulted in higher final biomass levels and a lower probability of being below the limit reference point at the end of the projection period. Scenarios in which the initial (prior to implementation of the HCR) relative stock size was low ("InitDepl = 0.2") resulted in lower and more variable levels of catch. Under T3a, a higher initial stock size ("InitDepl = 0.75") led to higher catches, an increase in the final relative stock size and a lower probability of being below the limit reference point at the end of the projection period. However, an interaction between the initial stock size and the choice of HCR meant that under T3b and T3c, scenarios starting at high relative biomass resulted in large, unsustainable catches being taken, and a general poor performance of the HCR in terms of maintaining stock status, and near ubiquitous probability of ending the projection below the limit biomass (Table 5.4, "T3b/c : Init Depl interactions"). This change in behaviour among HCRs for the high initial stock size scenarios was largely a result of the difference in maximum allowable increases in catch (T3a has a cap of 20% increase versus 50% for T3b and T3c).

Aside from the initial stock size, the factor with the largest impact on the biomass-related performance measures in terms of magnitude was whether the assumed value for *M* was correct. When the assumed value for *M* used in the estimation method and HCR was less than the true value in the operating model, the HCR resulted in higher final biomasses, lower probabilities of dropping below the limit reference point, and lower average catches (Table 5.4, 'assumed *M* < true *M*'). Conversely, assuming a value for *M* greater than the true value resulted in an under-estimation of *F*, and consequently, outcomes with lower final relative biomass and a higher risk of dropping below the limit (Table 5.4, 'assumed *M* > true *M*').

5.3.2. Spatial analyses

The results for the 'spatial' two-region scenarios for T3c are shown in Table 5.5, and include results for the three options related to how to deal with the spatial data (see Section 5.2.3.3). Results for T3c are shown because this variant appeared to best satisfy management objectives in the non-spatial analyses described above. A decrease in the connectivity of the regions resulted in reductions in the probability of dropping below B_{20}, and increases in final spawning depletion for scenarios when the stock was initially at the biomass target (Table 5.5, *F* option 1). Higher rates of fishing mortality can be expected in the initially exploited region when the connectivity between regions is lower, which may be more easily detected by the estimators. The sensitivity of the performance measures to the degree of connectivity was lost when the initial spawning depletion was either at high or low levels (Table 5.5, scenarios 1-7). The initial status of the stock therefore appeared to be at least as important (if not more so) in determining the values for the performance measures as the connectivity among the regions. Indeed, the results of linear modelling to predict the values of performance measures suggested that the effect of the degree of mixing between regions was not important. The coefficients for the intermediate and limited mixing terms were not significant contributors to variation in

the values of the performance measures (Table 5.6, 'intermediate mixing' and 'limited mixing').

The performance of the HCR was relatively consistent between the intermediate (in which the average mixing rate is 20%) and the limited (5%) level of mixing. The age structure of fish mixing between regions appeared to have little impact on the value of the performance measures. HCR performance was more sensitive to the true shape of the selectivity curve and whether the true selectivity was variable among regions. In general, dome-shaped selectivity led to higher final relative spawning biomasses and lower probabilities of dropping below B_{20} (Tables 5.5 and 5.6, scenarios 8-9, 'Different selectivities by region' and 'Selectivity dome-shaped in both regions'). Lower catches resulted from selectivity being dome-shaped in both regions (Table 5.6, 'Selectivity dome-shaped in both regions').

Analysing the data by region and then choosing the maximum estimated F to set the RBC (Table 5.5, F option 3) led to the highest values for relative spawning stock biomass, and the lowest probabilities of dropping below the limit reference point (Figure 5.6). The relative performance of the different scenarios was very similar when data from both regions were aggregated or when regional F estimates were weighted by their variance (Table 5.5, F options 1 and 2), although results from option 2 appeared more variable. An exception was when movement between regions was limited to pre-recruits (Table 5.5, scenario 8). In this instance, aggregating the data and conducting a single analysis (F option 1) appeared to be a more conservative way to determine RBCs, because the relative biomass was well below B_{20} when regional estimates of F were weighted by the inverses of their variances.

As with the non-spatial analyses, a large proportion of the variation in the performance measures among scenarios could be attributed to the pre-specified value of the relative stock size prior to implementation of the harvest strategy, rather than the specifications

for the particular HCR implemented. The increase in final biomass and decrease in the risk of dropping below B_{20} associated with choosing the maximum regional F estimate were naturally associated with lower catches, but did not result in a decrease in variability in catches, nor a decrease in the relative frequency of fishery collapse (Figure 5.6f-g). This option also appeared to mitigate the change in performance associated with the spatial connectivity among regions, as the values for the performance measures did not change with decreasing connectivity as was the case when the regional data were aggregated prior to analysis (Table 5.5, compare scenarios 1, 2, and 5 between F options 1 and 3). This suggests the degree to which sampling error has an effect on the performance of the HCR: in scenario 1 the true exploitation rate was the same in both regions yet F option 3 resulted in higher final spawning stock biomass than F option 1.

Higher values for the final spawning stock biomass associated with application of T3b and T3c may not necessarily imply a satisfactory outcome. Figure 5.7 shows that higher relative spawning biomass may be a result of closing the fishery for a number of years following a series of successive increases (or decreases) in the RBC, rather than achieving the target stock size with sustainable catch levels . This is also reflected in the values for the mean number of years in which the fishery collapses (Figures 5.4g, 5.5g, 5.6g). These results suggest that for a species such as blue eye trevalla, the catch curve may be unable to detect changes in F, as catch trajectories show a delay in the response to changes in spawning biomass (Figure 5.7). This can be expected for a long-lived species, where there would presumably be considerable inertia in the age structure. As such, the estimates of F obtained may not be reflecting the current fishing mortality rate.

5.4. Discussion

Management based on rapid stock assessment is attractive for fisheries where there are limited data, and methods for such assessment, including catch curve analysis, are well-established (albeit also with well-known shortcomings related to unrealistic assumptions such as constant mortality rate and lack of recruitment variability). The MSE testing of the variants of the Tier 3 HCR presented here suggests that it is indeed possible to formulate HCRs based on the results of catch curves that achieve management objectives (i.e. maintain spawning stock biomass at or above target levels), despite some of these shortcomings. However, it is also clear that implementation of the Tier 3 HCR can result in undesirable system behaviour, and that outcomes can be sensitive to many of the known shortcomings of the associated assessment methods.

Assessing performance of the HCRs through their ability to conserve stock biomass may not be an appropriate choice. The spawning biomass trajectories in Figure 5.7 suggest that satisfactory outcomes for a scenario (for example, a low probability of being overfished) can be achieved with undesirable system properties (such as complete closure of the fishery following a ratcheting increase in catch), even for the new Tier 3 HCR. Wayte and Klaer (2010) and Smith et al. (2008) address issues related to the unresponsiveness and ratcheting behaviour of the Tier 3 HCR. These undesirable properties are likely to be more pronounced for longer-lived species because the estimates of F from the catch curve do not relate well to current conditions. Unresponsiveness in the Tier 3 HCR is also a consequence of restrictions on the magnitude of permitted changes in management actions (the RBC is only allowed to change by 50% in a given year even if the estimate of F changes dramatically). The results suggest that this behaviour appears to result in higher final biomass levels for stocks that are at low relative stock size prior to implementation of the HCRs compared

with those achieved for stocks that are at or above management targets (Tables 5.4 and 5.6). Increasing the time period over which the catch is averaged will mitigate the ratcheting effect of RBCs (concurrent increases or decreases), but doing so effectively downweights the influence of previous management actions.

Differences in the performance among the variants of the Tier 3 HCR appeared to be related to both the form of the HCR, and the values chosen for the reference points (e.g. Figure 5.4). Tier 3 HCRs that used the spawner-recruit-based reference points (T3c) resulted in the best performance (Figure 5.4). However, performance of the rules that used a target of 0.5*M* (T3b) were generally only marginally different than those using spawner-recruit-based reference points, even though the data requirements were markedly less (Table 5.4; Figure 5.4). As estimates of *M* tend to be very uncertain (with results being very sensitive to assuming the wrong value), including additional uncertainty associated with estimating the compensation of the spawner-recruit curve (steepness) is perhaps unnecessary. However, Figure 5.3 clearly shows that 0.5*M* is not necessarily an appropriate target rate of fishing mortality (when compared with Tier 1 reference points) for all instances (e.g. when steepness is low). Note that even the 'poorly' performing HCRs require an estimate of *M*, typically derived from longevity and growth information (e.g. Pauly 1980, Hoenig 1983, Jensen 1996). While such information generally tends to be available, the nature of a 'data-poor' fishery means that these estimates of *M* are likely to be uncertain. The CV of *M* estimates from the Pauly and Hoenig methods are 0.53 and 0.61 respectively (MacCall 2009), and therefore it might be unreasonable to assume greater certainty in *M* than this for a data-poor stock (MacCall 2009 recommends a CV of at least 0.5).

The results clearly demonstrate the need for careful application of 'common sense' when applying methods such as the Tier 3 HCRs. For example, the implications of selectivity being dome-shaped were that mortality was over-estimated, leading to specification of lower catches. However, it would be somewhat foolish to use this conservation of stock

biomass as a reason for implementation of the Tier 3 HCR, if selectivity is known to be dome-shaped. Having an accurate estimate for *M* appeared to be very important for HCR performance, with scenarios where the chosen value for *M* was higher than the true value resulting in high probabilities of dropping below the limit reference point. Similarly, scenarios for which the assumed value for *M* was lower than the true value were among the most conservative in terms of biomass relative to the unfished state at the end of the projection. These results were unsurprising, as the estimate of *F* is clearly negatively correlated with the value assumed for *M*. The impact of selectivity being dome-shaped is similar to that of under-estimating *M*, in that *F* is over-estimated (because the estimators assume selectivity to be asymptotic), resulting in lower RBCs and higher spawning stock biomasses.

Although the analyses examined the impact of collecting data from multiple regions, and where the regional allocation of catches was changing, the data were generated in proportion to the catch, with no over-dispersion or bias in the sample other than the stochasticity imposed on the data through sample size and multinomial sampling. The low sampling effort present in the actual blue eye trevalla data set, coupled with seasonal differences in availability, means that the age and length data for blue eye trevalla are not representative of the fishery as a whole. Indeed, the sample size of the age data for blue eye trevalla is such that pooling age data across years to obtain an age-length-key and then applying this to the year/region-specific length composition data is the most feasible means of estimating the age-composition of the catch (e.g. Klaer 2009). While the analyses investigating the impacts of region-specific selectivity went some way to addressing these questions, it is likely that incorporation of bias and non-representative sampling into the MSE framework would further degrade HCR performance. It is also not clear how the way in which future catches are allocated by region/fleet impacts the results of this chapter. The lack of difference between scenarios that varied in the degree of spatial connectivity can be attributed somewhat to the allowance of a shortfall in the

catch in one region (as a result of insufficient available biomass) to be taken in the other region if required.

Most fisheries and also fished populations exhibit spatial structure, which creates spatial heterogeneity in realised exploitation rates and biomass trends. The extent of this heterogeneity will depend on the level of mixing in the overall stock. However, most management agencies lack the ability (or rather, the infrastructure) to specify the TAC at the level of this spatial structure. HCRs that show robustness to spatial differences are therefore desirable. Disaggregating the data by fleet and region, analysing these data separately, and then choosing the maximum estimate of F to determine the RBC appeared to produce the most conservative results irrespective of the true nature of stock connectivity and fishery behaviour, and also resulted in the most consistent values for the performance measures among scenarios that varied in spatial structure. However, application of this version of the HCR led to unnecessarily low catches when the connectivity of the stock between regions was high, reflecting the impact that sampling error can be expected to have on HCR performance in such scenarios. The maximum F option would also be inappropriate if the maximum F estimate came from a sector of the fishery which was a minor component of the catch, as it would be more likely that such an F estimate would be both uncertain and not representative of the overall exploitation rate. Weighting fleet and regional estimates of F by their variance accounts for this if the data are collected proportionally with the catch, as was assumed here. If not, then additional rules, such as weighting the F estimates by the regional catch rather than by variance, will be needed. However, such methods will not accommodate the effects of dome-shaped selectivity on over-estimation of F. Spatial disaggregation of data that already have low sample size will result in more variable estimates of mortality than might be expected given population dynamics, particularly when constructing annual catch curves.

The adoption of a precautionary approach to management of exploited marine resources is increasingly called for, and it is clear that testing of harvest control rules (e.g. using MSE) is necessary to understand whether these rules can be expected to perform as intended. The analyses described focused on parameterisations of the operating model which mimic blue eye trevalla, but the system can be extended to examine the performance of the tier framework given different life histories. Indeed, a natural avenue for further extension of these analyses would be to examine whether the relative performance of the various Tier 3 HCRs depends on the life-history of the species of interest, and whether the various HCRs would need to be modified depending on true values for parameters such as steepness and natural mortality. Being long-lived and late-maturing, blue eye trevalla have a long generation time (>20 years). It is perhaps to be expected that the behaviour of HCRs would differ if a longer time frame was considered. However such time periods may not be useful for current managers, suggesting perhaps that alternative control rules or reference points might be more pertinent given slower life histories.

While improved performance and conservation of stock biomass is achieved under the T3b and T3c HCRs over that of the T3a HCR, the variability of the stock biomass, and of catches under this HCR are greater than that expected for a more data-rich scenario (e.g. integrated assessment using Stock Synthesis). This is unsurprising; data-poor methods should be expected to estimate quantities of interest with greater uncertainty than those for which more data are available. While the Tier 3 HCR based on reference points such as F_{40} and F_{20} (T3c) is more similar to the Tier 1 HCR, care should be taken regarding the ability to estimate F sufficiently well to apply this rule successfully. Application of these reference points under the Tier 3 HCR requires an estimate of the value for steepness, which cannot be obtained during the analysis and was assumed to be 0.75 regardless of the true value. Consequently, performance of the HCRs was poor when the true value for steepness was lower than this. However, the approach taken here is not

very different than that employed for data-rich scenarios in the SESSF, as estimation of steepness within stock assessments for this fishery is restricted to the assessments for tiger flathead (*Neoplatycephalus richardsoni*, e.g. Klaer 2010) and eastern gemfish (*Rexea solandri*, Little and Rowling 2009), due to a lack of contrast over time in spawning biomass levels for other stocks. In general, although the T3c HCR only relies on one source of current data from the fishery, application of this HCR requires estimates be available for the majority of biological parameters that would be included in a more formal stock assessment based on a statistical catch-at-age model. The performance of such estimators (e.g. Stock Synthesis) given solely age-composition data (i.e. no index of abundance from surveys or catch per unit effort), the robustness of such models to mis-specifications such as those investigated here, and comparison of performance with the Tier 3 HCRs warrants interest.

The desired F to be estimated is the current rate of fishing mortality, whereas the annual catch curve integrates over the fishing mortality rates experienced by the stock for the length of time reflected in the age structure, which may either not correspond well with recent trends in F, or, if data are noisy, may impede estimation of F. Poor ability to estimate F may mean a lack of ability to accurately discriminate between the target and limit reference points used in the HCR. This may be particularly important when the values for F_{40} and F_{20} are similar. Successful implementation of a HCR relies on being able to distinguish between values for stock indicators that ought to result in changes in management action. Approximate confidence intervals for the current rate of fishing mortality on blue eye trevalla based on application of the MYEF estimator to data from the auto-longline fishery are wider than the range of values for F over which changes in management actions are indicated given the HCR (Fay 2009, Figure 5.3).

Precaution with respect to Tier 1 is not explicitly built into the Tier 3 HCR at present, as the quantities for F_{TARG} and F_{LIM} are the same as for Tier 1 (even though their estimates may be different). Any conservation of stock biomass under the Tier 3 HCR relative to

Tier 1 would arise from the behaviour of the rules, rather than an explicit specification. Additional modifications to the Tier 3 HCR such that there is equivalence of risk with the Tier 1 HCR (same probability of biomass dropping below the limit) could involve the choice of alternative reference points (e.g. $F_{TARG} < F_{40}$), the application of a discount to the RBC for being at a less data-rich tier level (Smith et al. 2008), or perhaps application of current HCRs with a more conservative value for F_{CUR}, based on some percentile of the confidence interval of the estimate. Further simulation testing to address the efficacy of such approaches is clearly warranted, and is a suitable candidate for future work that would include exploring the tradeoffs among the risk to the stock, average spawning stock size, and the magnitude of catches for different approaches.

Table 5.1. List of acronyms used in the chapter with their definitions.

Acronym	Definition
SESSF	Southern and Eastern Scalefish and Shark Fishery
MSE	Management Strategy Evaluation
HCR	Harvest Control Rule
RBC	Recommended Biological Catch
TAC	Total Allowable Catch
CR	Chapman and Robson (1960) estimator
MYEF	Multi-year Equilibrium F estimator
C_{cur}	Current catch (average over last four years)
YPR(F)	Yield-per-recruit given fishing mortality rate F
T3a	Old Tier 3 HCR (Figure 5.2b)
T3b	New Tier 3 HCR (Figure 5.2c), F_{TARG}=0.5M, F_{LIM}=M
T3c	New Tier 3 HCR, $F_{TARG}=F_{40}$, $F_{LIM}=F_{20}$
F_{CUR}	Current fishing mortality rate
F_{TARG}	Target fishing mortality rate
F_{LIM}	Limit fishing mortality rate
F_{40}	Fishing mortality rate that would (on average) result in spawning biomass of 40% unfished.
F_{MSY}	Fishing mortality rate that produces maximum sustainable yield
B_0	Unfished spawning biomass
B_{curr}	Spawning biomass prior to implementation of HCR
B_{LIM}	Limit biomass reference point
B_{TARG}	Target biomass reference point, % unfished spawning biomass

Table 5.2. Parameterisation of the operating model for the non-spatial scenarios. h is the steepness parameter of the spawner-recruit relationship, B_{curr}/B_0 is the spawning biomass relative to unfished prior to implementation of the HCRs, and n_A is the annual sample size for the age composition data.

#	Scenario	Type of selectivity curve	true M	h	B_{curr}/B_0	n_A	assumed M
1	base-case	asymptotic	0.08	0.75	0.40	100	0.08
2		asymptotic	0.08	0.75	0.20	100	0.08
3		asymptotic	0.08	0.75	0.75	100	0.08
4		asymptotic	0.12	0.75	0.40	100	0.12
5		asymptotic	0.12	0.75	0.20	100	0.12
6		asymptotic	0.12	0.75	0.75	100	0.12
7		asymptotic	0.18	0.75	0.40	100	0.18
8		asymptotic	0.18	0.75	0.20	100	0.18
9		asymptotic	0.18	0.75	0.75	100	0.18
10	low steepness	asymptotic	0.08	0.30	0.40	100	0.08
11		asymptotic	0.08	0.30	0.20	100	0.08
12		asymptotic	0.08	0.30	0.75	100	0.08
13	high steepness	asymptotic	0.08	1.00	0.40	100	0.08
14		asymptotic	0.08	1.00	0.20	100	0.08
15		asymptotic	0.08	1.00	0.75	100	0.08
16	dome-shaped selectivity	dome-shaped	0.08	0.75	0.40	100	0.08
17		dome-shaped	0.08	0.75	0.20	100	0.08
18		dome-shaped	0.08	0.75	0.75	100	0.08
19	assume wrong M	asymptotic	0.08	0.75	0.40	100	0.05
20		asymptotic	0.08	0.75	0.40	100	0.12
21		asymptotic	0.12	0.75	0.40	100	0.08
22		asymptotic	0.12	0.75	0.40	100	0.18

Table 5.3. Parameterisation of the operating model for the spatial scenarios. 'Full' connectivity between regions implies single stock dynamics, 'intermediate' has 20% annual movement rate from one region to the other, while the 'limited' scenario only has a 5% annual movement rate. The movement patterns are as shown in Figure 5.1d.

#	Scenario	Type of selectivity curve		true M	h	B_{curr}/B_0	n_A	assumed M	connectivity	movement
		region 1	region 2							pattern
1	base-case, full mixing	asymptotic	asymptotic	0.08	0.75	0.40	100	0.08	full	constant
2	intermediate mixing	asymptotic	asymptotic	0.08	0.75	0.40	100	0.08	intermediate	constant
3		asymptotic	asymptotic	0.08	0.75	0.20	100	0.08	intermediate	constant
4		asymptotic	asymptotic	0.08	0.75	0.75	100	0.08	intermediate	constant
5	limited connectivity	asymptotic	asymptotic	0.08	0.75	0.40	100	0.08	limited	constant
6		asymptotic	asymptotic	0.08	0.75	0.20	100	0.08	limited	constant
7		asymptotic	asymptotic	0.08	0.75	0.75	100	0.08	limited	constant
8	movement declines with age	asymptotic	asymptotic	0.08	0.75	0.40	100	0.08	intermediate	pre-recruit
9	movement increases with age	asymptotic	asymptotic	0.08	0.75	0.40	100	0.08	intermediate	adult
10	dome-shaped selectivity	dome-shaped	dome-shaped	0.08	0.75	0.40	100	0.08	intermediate	constant
11	differing selectivities	asymptotic	dome-shaped	0.08	0.75	0.40	100	0.08	intermediate	constant

Table 5.4. Coefficients estimated from the linear models for the non-spatial analyses, by performance measure. Numbers shown represent terms that were assessed to be significant at the $\alpha = 0.05$ level in a full model that included all terms listed.

Linear predictor term	Performance measure						
	median (B_{final}/B_0)	IQR(B_{final}/B_0)	$P(B_{final}<B_{lim})$	$P(B_{proj}<B_{lim})$	median (avg TAC)	median (CV TAC)	# yrs collapse
Base intercept (CR, T3a, InitDepl=0.4, *h*=0.75, asymptotic Sel)		0.17	0.79	0.87	721	0.16	
MYEF							
T3b	0.14	0.10	-0.17	-0.13	-209	0.96	5.9
T3c	0.17	0.15	-0.22	-0.15	-249	0.95	5.4
assumed *M* > true *M*	-0.14	-0.26	0.35	0.22		0.54	4.6
assumed *M* < true *M*	0.56		-0.64	-0.75	-468	0.85	5.7
Initial depletion = 0.75	0.39	0.24	-0.59	-0.64	476	0.20	
Initial depletion = 0.2				0.12	-394	0.62	4.8
h = 1.0	0.15		-0.17	-0.08			
h = 0.3	-0.14	-0.15	0.32	0.11		0.39	3.6
true *M* = 0.12							
true *M* = 0.18					150		-1.3
T3b : Init Depl = 0.75 interaction	-0.58	-0.51	0.90	0.84	1279		-2.4
T3c : Init Depl = 0.75 interaction	-0.60	-0.52	0.94	0.87	1108		
T3b : Init Depl = 0.2 interaction	0.19		-0.32			-0.30	-2.8
T3c : Init Depl = 0.2 interaction	0.23		-0.34				
dome-shaped selectivity : T3a interaction							
dome-shaped selectivity : T3b	0.12		-0.20	-0.14	-216		
dome-shaped selectivity : T3c							

Table 5. 5. Values for the performance measures for the spatial analyses given estimation under MYEF, application of the T3c HCR. 'F option' is the option used to obtain a single FCUR estimate given the fleet and regional data (Section 5.2.3.3): 1) aggregate fleet and regional data, 2) obtain fleet and region-specific F estimates, weight by inverse variance, and 3) separate analyses as in 2, but use the maximum estimated F to calculate the RBC. Scenario numbers as in Table 5.3.

		Performance measure						
F option	Scenario #	median (B_{final}/B_0)	IQR(B_{final}/B_0)	P($B_{final}<B_{lim}$)	P($B_{proj}<B_{lim}$)	median (avg TAC)	median (CV TAC)	# yrs collapse
1)	1	0.12	0.36	0.61	0.77	491	0.99	3.9
	2	0.24	0.35	0.42	0.69	479	0.95	2.6
	3	0.46	0.24	0.09	0.90	114	1.58	7.1
	4	0.00	0.00	0.97	0.98	1,944	1.48	5.3
	5	0.21	0.34	0.50	0.74	478	1.03	3.8
	6	0.47	0.22	0.07	0.86	90	1.56	7.2
	7	0.00	0.00	0.94	0.97	2,122	1.41	4.6
	8	0.21	0.42	0.47	0.58	458	1.02	4.1
	9	0.14	0.34	0.61	0.73	511	0.92	3.2
	10	0.47	0.44	0.26	0.33	288	1.00	2.6
	11	0.33	0.39	0.32	0.54	404	0.96	2.6
2)	1	0.15	0.42	0.55	0.72	436	1.04	4.1
	2	0.15	0.33	0.56	0.75	481	0.97	3.3
	3	0.44	0.22	0.09	0.87	109	1.53	6.8
	4	0.00	0.00	0.99	0.99	2,006	1.46	5.1
	5	0.19	0.42	0.56	0.71	479	1.05	3.7
	6	0.45	0.25	0.14	0.88	98	1.53	6.7
	7	0.00	0.00	0.97	0.97	2,161	1.47	5.2
	8	0.00	0.07	0.83	0.91	494	1.39	7.0
	9	0.07	0.36	0.60	0.73	508	0.92	3.3
	10	0.37	0.41	0.29	0.45	344	0.92	2.2
	11	0.22	0.40	0.46	0.60	443	0.94	2.9
3)	1	0.38	0.28	0.24	0.47	353	0.95	2.6
	2	0.35	0.36	0.26	0.48	367	0.98	2.9
	3	0.53	0.16	0.05	0.87	71	1.72	8.9
	4	0.00	0.11	0.83	0.84	1,899	1.27	3.5
	5	0.39	0.41	0.29	0.48	310	1.05	3.0
	6	0.52	0.18	0.02	0.83	46	1.86	10.2
	7	0.00	0.06	0.81	0.82	1,834	1.27	3.9
	8	0.62	0.23	0.01	0.07	135	1.12	3.5
	9	0.45	0.34	0.21	0.39	322	0.90	1.8
	10	0.59	0.27	0.12	0.22	205	1.02	2.9
	11	0.49	0.35	0.14	0.30	282	1.10	3.3

Table 5.6. Coefficients estimated from the linear models for the spatial analyses. Numbers shown represent terms that were assessed to be significant ($p<0.05$) in a full model that included all terms listed. Base intercept values were: *MYEF*, T3c HCR, spatial *F* option 1 (obtain a single *F* estimate with all data), full mixing between regions, movement constant with age, initial depletion = 0.4, and h=0.75.

Linear predictor term	Performance measure						
	median (B_{final}/B_0)	IQR(B_{final}/B_0)	P($B_{final}<B_{lim}$)	P($B_{proj}<B_{lim}$)	median (avg TAC)	median (CV TAC)	# yrs collapse
base intercept	0.19	0.37	0.51	0.71	468	0.97	3.5
F option 2, wt regional *F* estimates by variance							
F option 3, choose highest regional *F*	0.15		-0.21	-0.21	-141		
Initial depletion = 0.75	-0.25	-0.34	0.49	0.29	1562	0.39	1.4
Initial depletion = 0.2	0.22	-0.16	-0.36	0.23	-344	0.63	4.6
Different selectivities by region							
Selectivity dome-shaped in both regions	0.23		-0.21	-0.31	-145		
intermediate mixing							
limited mixing							
juveniles move only		-0.12				0.19	1.8
adults move only							

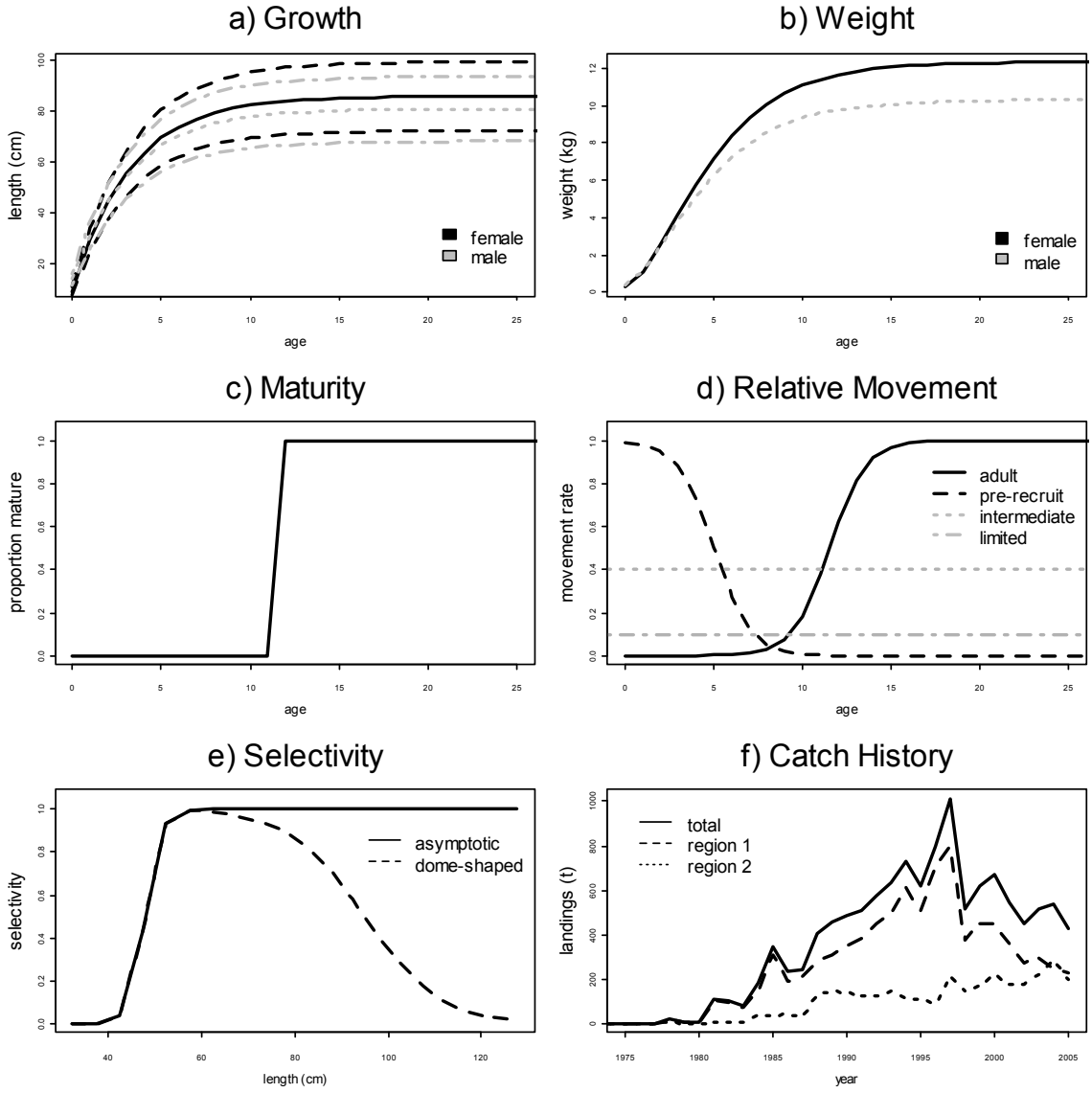


Figure 5.1. Biological and fishery-related parameters. a) Growth panel shows mean (solid) and 95% interval (dashed) of distributions for length-at-age. d) Age-specific pattern of relative movement rate for adults and (dashed) pre-recruit, along-with intermediate and limited levels of mixing. f) Total catch time series used in the non-spatial analyses, along with the regional catch histories used in the spatial analyses.

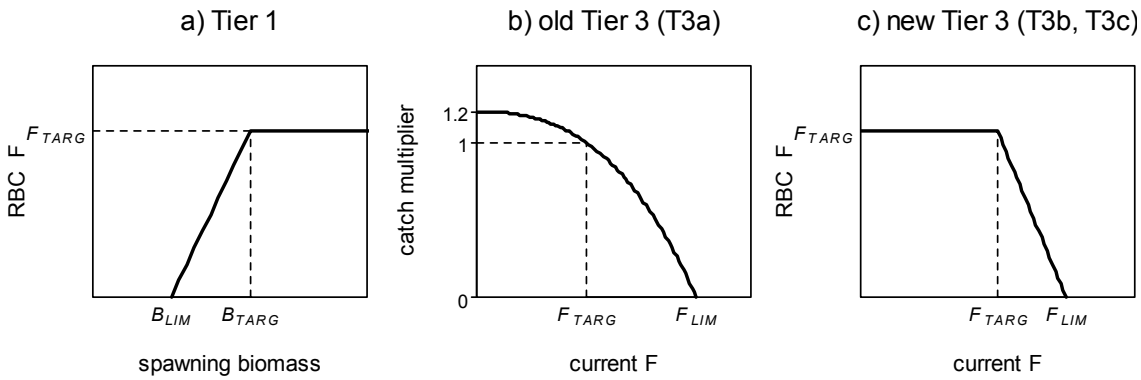


Figure 5.2. Forms for the Harvest Control Rules (HCRs) for Tier 1 (top-left panel), old Tier 3 (T3a, top-right panel), and new Tier 3 (T3b and T3c, bottom-left panel). Solid lines indicate the derived values for the RBC rate of fishing mortality (panels a and c) or the multiplier to the current catch (panel b), given the estimated value for the stock indicator on the x axis. Dashed lines indicate the operation of the HCRs at the target reference points.

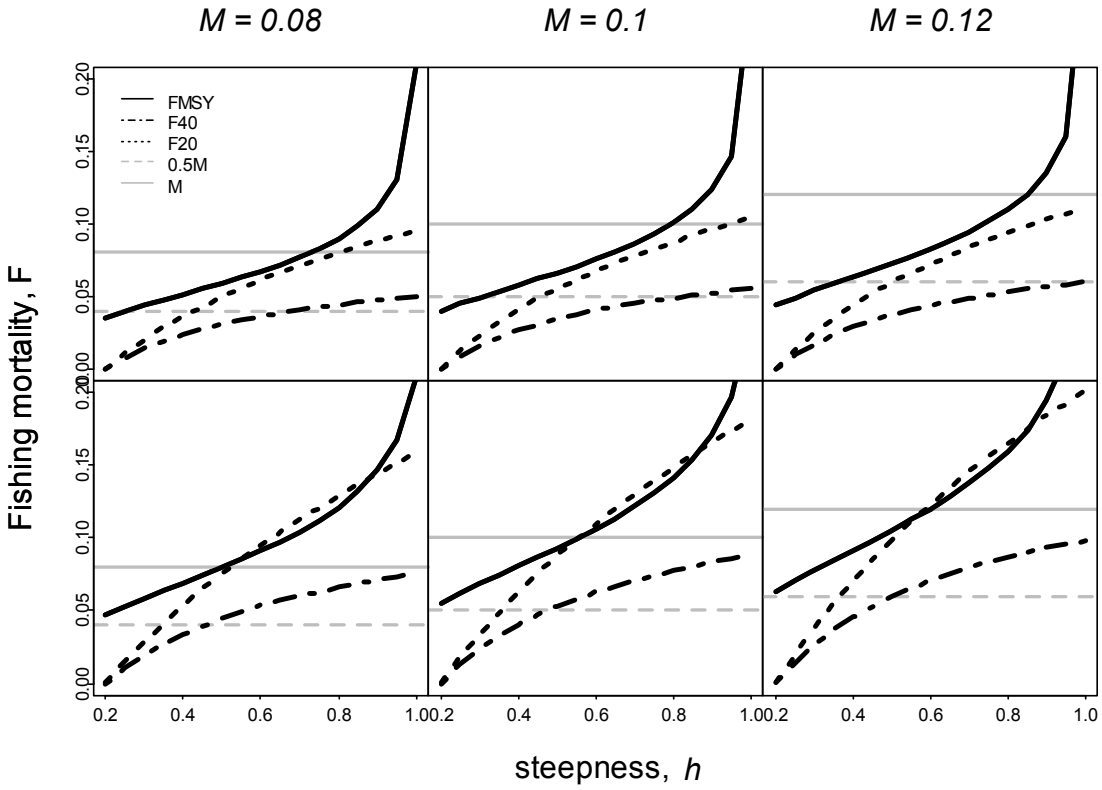


Figure 5.3. Relationship between the values for the target and limit reference points, and those for *M*, and *h*. Top row of panels corresponds to an age-at-maturity of 12 yrs, as used in the analyses. The bottom row of panels shows how the relationship changes when age-at-maturity equals 6 yrs.

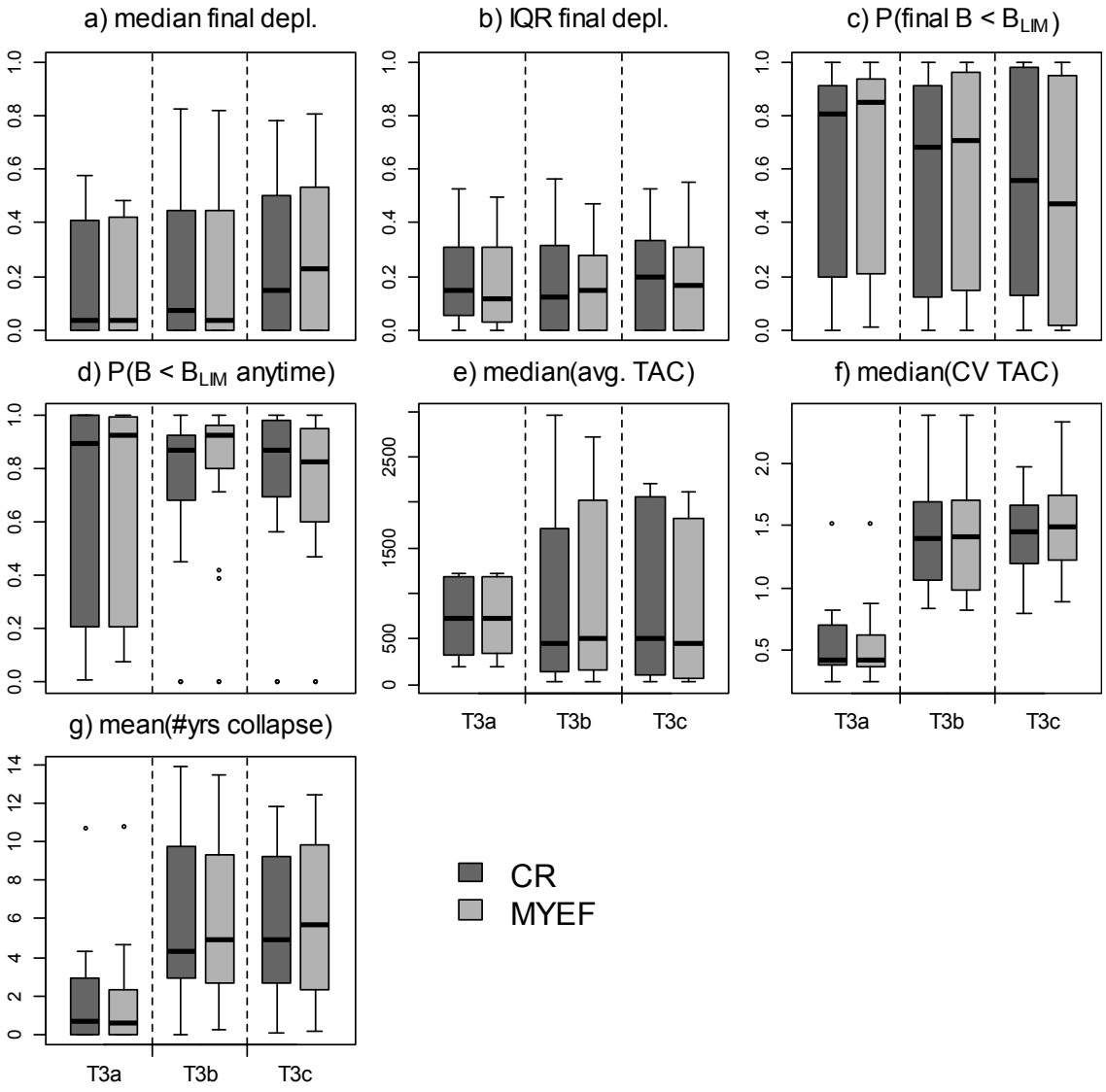


Figure 5. 4. Distribution of the performance measures across scenarios for the non-spatial analyses, for the two estimation methods, CR and MYEF, given application of the three HCRs, T3a, T3b and T3c.

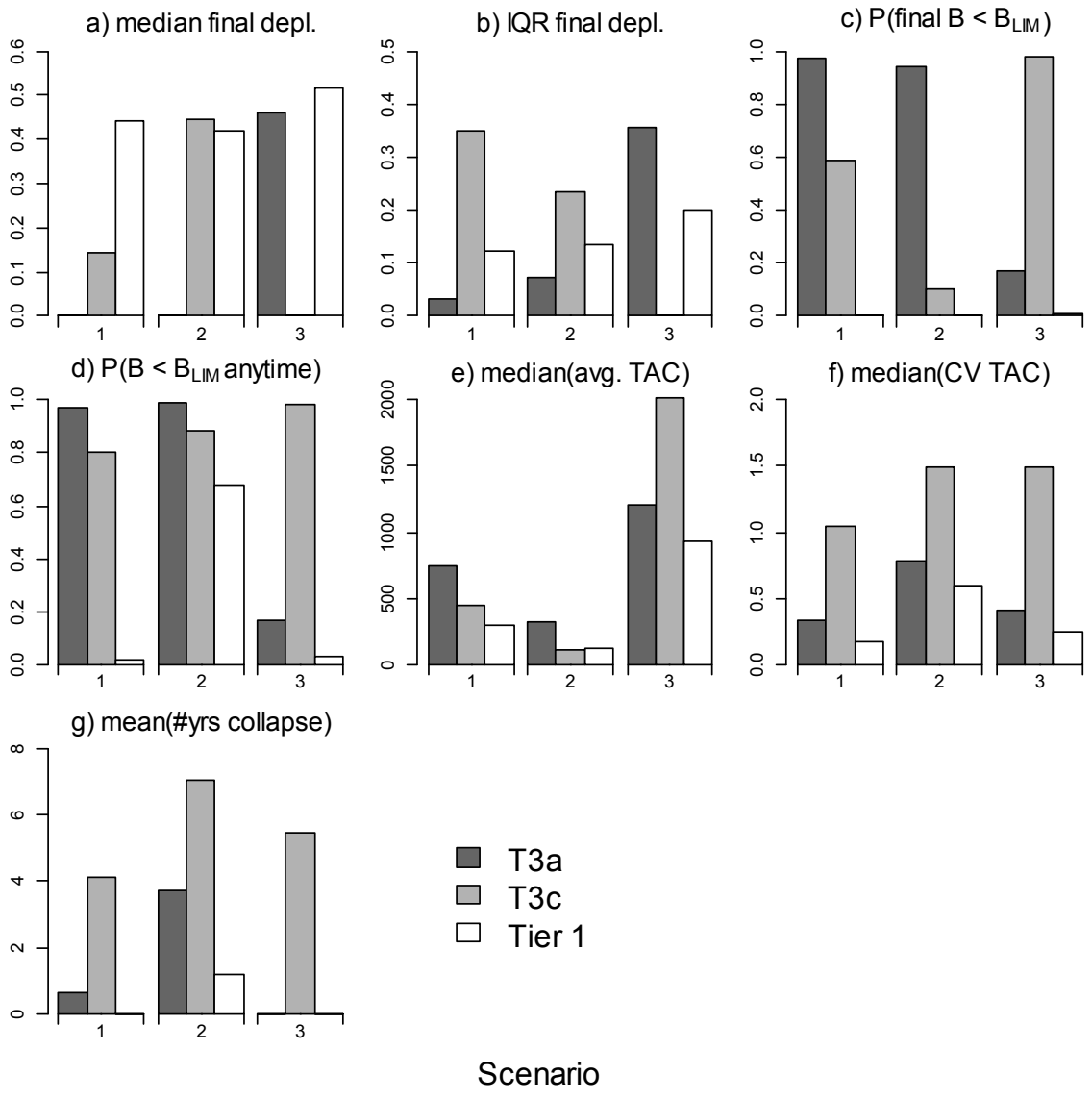


Figure 5.5. Performance of T3a, T3c based on MYEF, and the Tier 1 HCR, for three of the non-spatial scenarios. Numbers correspond to scenarios listed in Table 5.2.

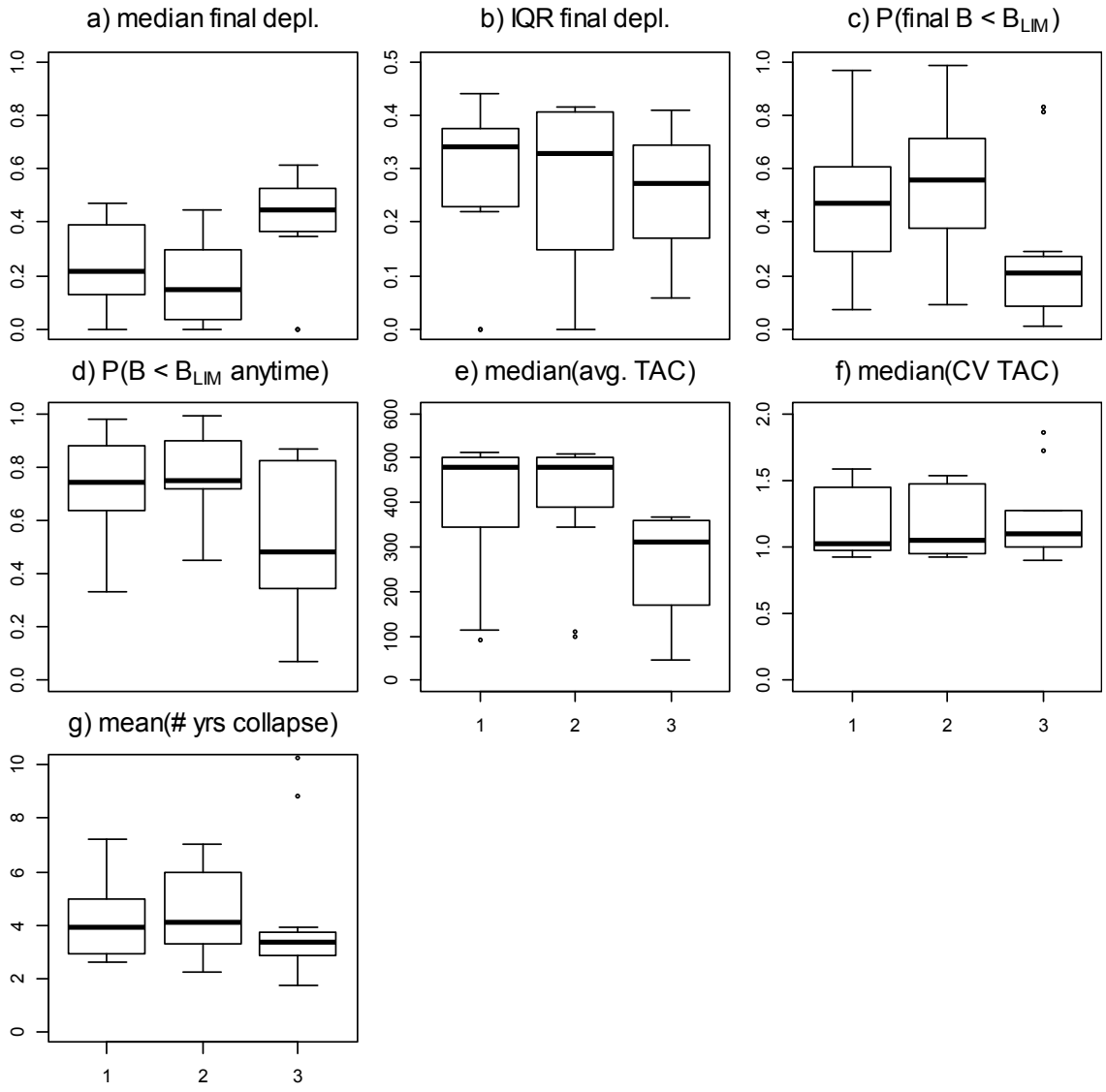


Figure 5.6. Distribution of the performance measures across scenarios for the spatial analyses, for the different options of choosing the annual *F* estimate (1 = aggregate data, 2 = analyse by region, weight estimates by variance, 3 = analyse by region and choose the maximum estimated F). Estimation is using MYEF with the T3c HCR.

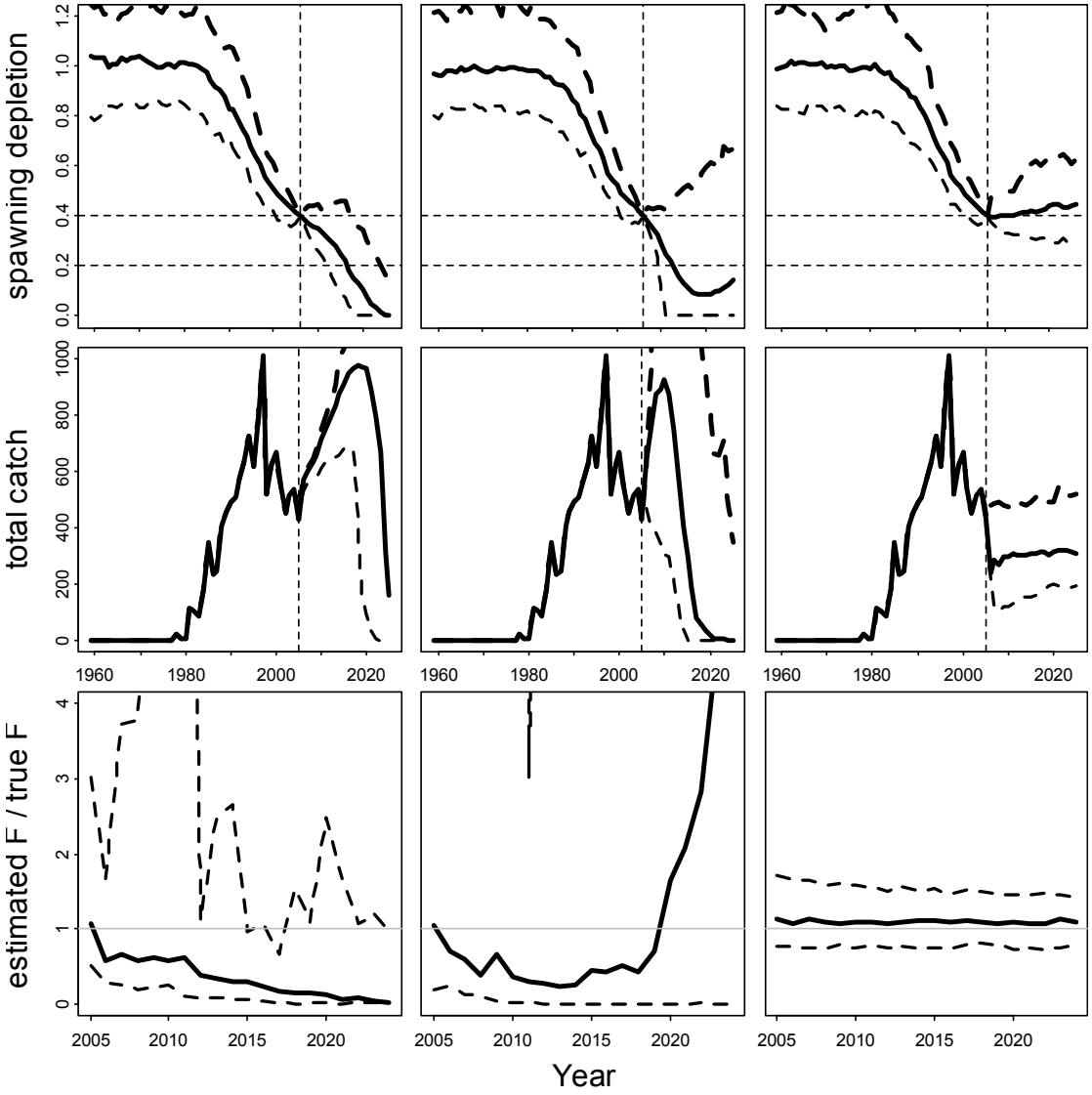


Figure 5.7. Trajectories of the relative spawning biomass and catch, along with the distribution for the value of the estimated *F* compared to the true value, for scenario 1 of the non-spatial analyses for (left) T3a, (center) T3c with MYEF and (right) Tier 1. Plotted are the median (solid line) and central 95% intervals (dashed lines) from the simulations. The dash-dot vertical line indicates the start of implementation of the harvest strategy, horizontal black dotted lines indicate the target (B_{40}) and limit (B_{20}) reference points.

Conclusions

Spatially-explicit modelling approaches are increasingly being applied in marine systems, with spatial issues currently being viewed as one of the frontiers in the development of quantitative models for fisheries management (Hilborn 2012). There are many model-based approaches for dealing with spatial heterogeneity, with diversity in approach reflecting the general purpose of needing to account for spatial processes. Linear state-space models were able to successfully estimate population trends and abundance that varied in space (Chapter 1), and showed robustness to uncertainty regarding spatial structure when spatial autocorrelation in trend among populations was accounted for. Optimal spatial and temporal allocation of monitoring effort (Chapter 2) differed depending on the estimated quantity of interest, although optimising monitoring designs for total abundance appeared to perform well when estimating other quantities. Identifying tradeoffs in monitoring design performance through analyses such as those in Chapter 2 could be used to guide the design of monitoring plans for marine species. Including spatial autocorrelation in GAMMs for estimating the density and distribution of Steller sea lion prey species (Chapter 3) changed estimates of abundance and temporal trend for some species, and revealed interannual variability in the values for parameters governing spatial characteristics. Foraging success of Steller sea lions (Chapter 4) depended on the spatial distribution of prey species, but this was also determined by the treatment of spatial information by the foraging sea lion. Foraging strategies that placed high value on information from previously visited locations were often the most successful, particularly when the detectability of prey species was low. This was particularly the case for simulations of juvenile sea lions, as increased metabolic requirements compared to adults resulted in fewer successful foraging events. The performance of data-poor Harvest Control Rules for blue eye trevalla (Chapter 5) varied, given uncertainty in spatial structure in population dynamics and fishery operation.

Management objectives for the blue eye trevalla fishery were better met when spatial estimates of fishing mortality were appropriately weighted for calculating catch quotas.

Spatial structure arises in ecological data either due to the process producing the data being spatial or because the process depends on one or more causal variables that are spatially structured (Legendre 1993). In many cases, most likely, it will be both. The spatial autocorrelation estimated for Gulf of Alaska fish species in Chapter 3 probably reflects spatial structuring in habitat-determining factors such as primary production and temperature. In cases where data on these variables are unavailable, the statistical models of spatial correlation used in Chapter 3 can approximate these processes. Spatial correlation in pup production among Steller sea lion rookeries (Chapter 1) will be due to both similarities in life-history traits, and spatial correlation in environmental conditions impacting population parameters (e.g. prey availability, predation pressure, etc.). The implications of these processes also depend on scale.

The analyses presented in this dissertation highlight the importance of spatial heterogeneity at different scales, and for different purposes; be it geographic differences in population trend, patchiness in resource distribution, behavioural uses of spatial information, and broad management scales. Modelling approaches, such as those presented in this dissertation, can be used to identify which data might be of use in further understanding system processes, and to evaluate the efficacy of potential management actions. No single modelling framework will likely be appropriate, as numerous pieces of the puzzle can probably be better explained by focussed attention. This is not a problem restricted to the analysis of spatial heterogeneity.

The models used in Chapter 1 are quick to implement, are able to accommodate missing data, and could be applied to a wide range of marine populations. The methods of Chapter 1 do rely on the assumption of linear systems with Gaussian error terms. It is unlikely that these assumptions will be true, although these types of models can perform

adequately when they are not upheld. Degenerate results (e.g. zero values for estimated observation or process error) were obtained when the amount of missing data was high. The models developed in Chapter 1 did not consider demographic processes, and therefore subsumed information on population processes into annual trend and error terms. While this approach is practical from a data availability perspective, it is perhaps restrictive in terms of testing mechanistic hypotheses for population change. In more data-rich cases, the systems of equations, state variables, and observation equations could be expanded beyond numbers at sites through time. Often however, enumeration of population size and trend is a first step toward understanding population status.

Choices about spatial and temporal allocation of monitoring effort clearly impacted the ability to estimate the final year trends, individual subpopulation numbers, and total metapopulation abundance (Chapter 2). Tradeoffs in the performance of the estimation procedure depended on the metric that designs were optimised for. Ultimately, evaluation of the efficacy of monitoring designs ought to include both the relative benefits of doing so with respect to management objectives, population status, and the costs associated with options. Additional work following from Chapter 2 could account for the costs associated with surveying various populations.

Notable differences were observed in Chapter 3 between the results obtained when spatial autocorrelation was accounted for and when it was not. A key result was that the non-spatial analyses led to higher estimates of mean density for almost all prey groups than did the spatial analyses. Challenges were observed using trawl survey data to monitor small schooling fish, and additional data types such as acoustic monitoring and pelagic trawl surveys could improve these estimates. Chapter 3 did not formally compare the performance of the methods. Simulation testing could be used to help identify types of species for which the method can be expected to perform well, given the types of available data. Analyses to tackle these issues might require additional assumptions associated with sampling bias.

The foraging simulations in Chapter 4 synthesised a range of information and insight from several fields of study. The spatial and temporal mismatch of available data limited the ability to estimate relationships among factors directly; so the simulation framework facilitated learning, whilst also recognising the uncertainty associated with linking hypotheses for interactions between species distribution, behaviour, and bioenergetics. Of the many caveats discussed in Chapter 4, the plasticity of sea lion physiology and behavioral response is perhaps the most challenging to overcome. While it is true that the simulations considered a single prey type, of uniform density within patches, it is unlikely that these constraints would impact general results. Similarly, the models could always be improved by additional data on some of the parameters. Steller sea lions are known to exhibit both physiological and behavioural responses to changes in foraging success; indeed one would expect this to be the case. Consequently, the parameter space governing options for foraging strategies could be much more complicated than that tested.

The foraging simulations of Chapter 4 utilised prey fields that were characteristic of the results obtained in Chapter 3 for sea lion prey species in the Gulf of Alaska. The exact results, and indeed the predicted spatial distribution maps from Chapter 3, were not used as the interest was in determining general behaviour from the simulations rather than conditioning to a particular location/time, etc. However, it could be possible to estimate foraging parameters based on data from telemetry and prey distribution (or proxies thereof) directly. The number of state variables involved in models combining these processes would be particularly large as integration would be required over prey distribution at each time step in addition to assumptions for behavioural parameters. This approach is probably not feasible from a computational and sampling perspective, and might only provide insight at specific sampling sites.

Chief among caveats associated with Chapter 5 include the ability to successfully estimate reference points and fishing mortality rates given the data currently available.

The analyses did not address issues related to biases occurring from spatial mismatches in the data collection and the fishery operation. Mismatches in data collection within the SESSF occur because of the limited resources available for both port and onboard sampling, although the allocation of monitoring effort under the Integrated Scientific Monitoring Plan (ISMP) has become increasingly more representative of the distribution of catches both seasonally and spatially. Nevertheless, species such as blue eye trevalla remain problematic as sampling effort must be allocated among gear types, and availability of fish to the fishery appears variable. Optimising such monitoring schemes is especially difficult given the need to obtain information on many species and stocks simultaneously, from numerous locations, fishing gears, and times of year. Further analyses that include bias due to non-representative sampling would demonstrate impacts on harvest control rule performance.

A simulation testing approach appears to be the most appropriate for resolving questions related to the efficacy of monitoring designs, the performance of assessment methods, the relative merits of foraging behaviors, and the performance of fisheries management strategies. While these methods are occasionally computationally intensive, they are distinctly less costly, potentially more rigorous, and more repeatable than experimental approaches. Integrative approaches that synthesise modelling and empirical information, along with simulation frameworks and estimation methods are simply too useful to ignore.

The proliferation of MSE as a tool to evaluate the performance of management strategies at both tactical and strategic levels within the fisheries science community ought not be surprising. Indeed, within the marine fisheries community MSE is almost becoming viewed as de rigueur (Bunnefeld et al. 2011). Application of MSE within fisheries management has mainly focused on testing the performance of assessment methods and harvest control rules. However, the general framework is ideally suited for questions related to the suitability of monitoring designs (e.g. Punt and Fay 2006) and the

development of monitoring needs with fishery (or alternative resource use) development (Dowling et al. 2008). Indeed, MSE provides an ideal framework for evaluating the design of monitoring programs in concert with the performance of subsequent management strategies. Applications of MSE that explicitly include spatial structure have generally focussed on implications of the scale/boundaries for management actions, and the efficacy of marine protected areas (e.g. Mapstone et al. 2008, McGilliard et al. 2011).

Challenges associated with expanding to spatially-resolved models are similar to those for implementing ecosystem models and minimal realistic models, in that data required to condition such approaches are often not available. In such instances, bracketing uncertainty using a suite of plausible alternatives for say spatial structure (such as that presented in Chapter 5), is a feasible approach (e.g. Plagányi and Butterworth 2011). The spatial resolution of currently applied end-to-end marine ecosystem models (e.g. Fulton et al. 2011) not only provides the capacity for addressing issues associated with spatial processes, but also highlights the growing concern to address spatial problems.

Including behavioural responses to successful foraging would lend itself to an agent-based-modelling approach. Agent-based models (ABMs) are increasingly being used to evaluate emergent properties of system behaviour as a result of behavioural choices by individuals (or groups of individuals). As such, ABMs can provide an opportunity for further development of simulation-based scenario evaluation. Applications of ABMs for marine resource management are rare (but see McDonald et al. 2008, Thorson et al 2012). One of the main challenges of ABMs is amount of type of data required to parameterise them (the various inputs needed for Chapter 4 exemplifies this). Formal decision analysis approaches utilising MSE that extends to ABMs will be challenging to implement because of the (much) increased state-space that must be integrated across, including appropriate sources of uncertainty. Application of ABMs in this context can be expected to share with MSE the question of what levels of uncertainty are appropriate to consider. Scenario testing seems more likely feasible than formal statistical evaluation.

Spatially-disaggregated data are frequently available. However the budgetary constraints of agencies charged with managing marine species means that data are unlikely to be consistent temporally, and among species / stocks of concern. Analyses for discerning spatial processes in data-limited situations are common (e.g. Cope and Punt 2009), and method development can be expected to continue to expand, particularly because of demands from policy-makers to increasingly accommodate spatial concerns within the decision-making toolkit, and the burgeoning recognition of spatial structuring among and within marine populations. Whether this increased recognition necessitates the inclusion of such spatial structuring when modelling and making management decisions may well depend on the situation at hand. Lorenzen et al. (2010) advocate participatory approaches to spatial governance that identify appropriate scales that are adaptive in implementation at their core. Mathematical modelling techniques, and in particular the use of simulation tools for scenario evaluation are ideal for identifying when such methods may be appropriate and successful (or not).

A key question that might (or ought) be asked when developing methods with a spatial focus, is it worth the effort to go spatial? (Sanchirico and Wilen 1999). Applying the models of Chapter 1 to Steller sea lion pup production, assuming single mean population trends across the western Alaska stock resulted in poorer fits to the data. Monitoring schemes that did not include spatial components did not perform well in terms of estimating population parameters, trends, and abundance (Chapter 2). Estimates of fish density, and trends in abundance in Chapter 3 differed depending on whether spatial autocorrelation was accounted for. The results of foraging models in Chapter 4 suggested that mean prey density was not a good predictor of foraging success given patchiness of prey resource and foraging strategies that relied on spatial information. Accounting for spatial heterogeneity in fishing mortality was an important driver of the success of harvest strategies for management of blue eye trevalla (Chapter 5). However, these analyses also demonstrated that while spatial dynamics were important to consider when

developing effective harvest control rules, mis-specification of biological factors such as natural mortality and the steepness of the stock-recruit relationship was even more crucial. Chapter 5 therefore cautions that while accounting for spatial heterogeneity can be important for modelling and monitoring, it must be placed in context of other concerns, and indeed in some cases might not be very important at all. Proponents of ecosystem-based fisheries management highlight the fact that once larger scale problems (high fishing mortality) are 'under control', additional processes and ecosystem drivers become relatively more important for successful system management. These additional complexities perhaps include environmental and ecological factors, but also no doubt spatial factors as well. In general, the need to model spatial variability is driven by whether spatial variability occurs on a scale that is relevant to either monitoring or management.

Spatial concerns have become more entrenched in marine resource management, following policy directives and improved ecological understanding. Spatial management has been associated with fisheries recoveries (Worm et al. 2009), and current moves toward adopting ecosystem approaches to management (or ecosystem-based management) highlight spatial considerations when understanding system processes and making management decisions (Francis et al. 2007). Statistical models that account for or attempt to explain spatial heterogeneity are being embraced by fisheries ecologists (e.g. Ciannelli et al. 2008), and implications of spatial structure in population dynamics and resource use are being incorporated into models used to support fisheries stock assessment and management, protected species conservation, and biodiversity assessment. The development of methods for dealing with spatial processes when managing marine resources will likely continue, particularly in the light of decisions such as those under the US. National Ocean Policy to explicitly include Coastal and Marine Spatial Planning as a priority in marine management.

Modelling approaches provide means for testing the efficacy of management actions. However, a diversity of tools will be needed to tackle questions related to different scales, species types, and different combinations of ecosystem services and different species. All organisms are spatial, but some organisms are more spatial than others.

Bibliography

Adams, C. F., R. J. Foy, D. S. Johnson, and K. O. Coyle. 2008. Seasonal changes in pelagic fish biomass around the Chiswell Island Steller sea lion rookery in 2003. Fisheries Research 93: 179-185.

Agarwal, D. K., A. E. Gelfand, and S. Citron-Pousty. 2002. Zero-inflated models with application to spatial count data. Environmental and Ecological Statistics 9: 341-355.

Agostini, V. N., R. C. Francis, A. B. Hollowed, S. D. Pierce, C. Wilson, and A. N. Hendrix. 2006. The relationship between Pacific hake (*Merluccius productus*) distribution and poleward subsurface flow in the California Current System. Canadian Journal of Fisheries and Aquatic Sciences 63: 2648-2659.

Ainsworth, L. 1990. Models and methods for spatial data: detecting outliers and handling zero-inflated counts. Ph.D. Thesis. Simon Fraser University, British Columbia. xiii+98pp.

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

Allen, B. M., and R. P. Angliss. 2010. Alaska marine mammal stock assessments, 2010. U.S. DepArtment of Commerce, NOAA Technical Memorandum NMFS-AFSC-223, 292 p.

Andrews, R. D., D. G. Calkins, R. W. Davis, B. L. Norcross, K. Peijnenberg, and A. W. Trites. 2002. Foraging behavior and energetics of adult female Steller sea lions. Steller Sea Lion Decline: Is It Food II. Alaska Sea Grant College Program. AK-SG-02-02.

Aydin, K., S. Gaichas, I. Ortiz D. Kinzey, and N. Friday. 2007. A comparison of the Bering Sea, Gulf of Alaska, and Aleutian Islands large marine ecosystems through food web modeling. U.S. Dep. Commer., NOAA Technical Memorandum NMFS-AFSC-178. 298p.

Baelde, P. 1996. Biology and dynamics of the reproduction of blue-eye trevalla, *Hyperoglyphe Antarctica* (Centrolophidae), off Tasmania, southern Australia. Fishery Bulletin 94: 199-211.

Bakun, A. 1987. Monthly variability in the ocean habitat off Peru as deduced from maritime observations, 1953 to 1984. Pages 46-74 in D. Pauly and I. Tsukayama (eds) The Peruvian anchoveta and its upwelling ecosystem: three decades of change.

ICLARM studies and reviews 15. Instituto del Mar del Peru (IMARPE), Callao, Peru; Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ), GmbH, Eschborn, Federal Republic of Germany; and International Center for Living Aquatic Resources Management, Manila, Philippines. 351p.

Ban, S. and A. W. Trites. 2007. Quantification of terrestrial haul-out and rookery characteristics of Steller sea lions. Marine Mammal Science 23: 496-507.

Banerjee, S., B. Carlin, and A. E. Gelfand. 2005. Hierarchical modelling and analysis for spatial data. Chapman and Hall. London.

Bonadonna, F., M. A. Lea, O. Dehorter, and C. Guinet. 2001. Foraging ground fidelity and route-choice tactics of a marine predator: the Antarctic fur seal Arctocephalus gazelle. Marine Ecology Progress Series 223: 287-297.

Brandt, S. B., D. M. Mason and E. V. Patrick 1992. Spatially-explicit models of fish growth-rate. Fisheries 17: 23-33.

Breed, G. A., I. D. Jonsen, R. A. Myers, W. D. Bowen, and M. L. Leonard. 2009. Sex-specific, seasonal foraging tactics of adult grey seals (*Halichoerus grypus*) revealed by state-space analysis. Ecology 90: 3209-3221.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. The American Naturalist 124: 255-279.

Bunnefeld, N., E. Hoshino, and E.J. Milner-Gulland. 2011. Management strategy evaluation: a powerful tool for conservation? Trends in Ecology and Evolution 26: 441-447.

Burgman, M. A., S. Ferson, and H. R. Akçakaya. 1993. Risk assessment in Conservation biology. Chapman and Hall, London.

Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodel inference: A practical-theoretic approach, 2nd edition. Springer-Verlag.

Butterworth, D. S., K. L. Cochrane, and J. A. A. De Oliveira. 1997. Management Procedures: a better way to manage fisheries? The South African experience. Pages 83-90 in: Pikitch, E. L., D. D. Huppert, M. P. Sissenwine (Eds.). Global trends: Fisheries Management. American Fisheries Society, Symposium 20, Bethesda, Maryland.

Butterworth, D. S. and A. E. Punt 1999. Experiences in the evaluation and implementation of management procedures. ICES Journal of Marine Science 56: 985-998.

Calder, C., M. Lavine, P. Müller, and J. S. Clark. 2003. Incorporating multiple sources of stochasticity into dynamic population models. Ecology 84: 1395-1402.

Call, K. A. and T. R. Loughlin. 2005. An ecological classification of Alaskan Steller sea lion (*Eumetopias jubatus*) rookeries: a tool for conservation/management. Fisheries Oceanography 14 (S1): 212–222.

Candy, S. G. 2008. Estimation of effective sample size for catch-at-age and catch-at-length data using simulated data from the Dirichlet-multinomial distribution. CCAMLR Science 15: 115-138.

Carlson, M. and F. Schmiegelow. 2002. Cost-effective sampling design applied to large-scale monitoring of boreal birds. Conservation Ecology 6: 11.

Carroll, C. and D. S. Johnson. 2008. The importance of being spatial (and reserved): assessing northern spotted owl habitat relationships with hierarchical Bayesian models. Conservation Biology 22: 1026-1036.

Chapman, D. G. and D. S. Robson. 1960. The analysis of a catch curve. Biometrics 16: 354-368.

Ciannelli, L., P. Fauchald, K. S. Chan, V. N. Agostini, and G. E. Dingsør. 2008. Spatial fisheries ecology: Recent progress and future prospects. Journal of Marine Systems 71: 223-236.

Collie, J. S. and C. J. Walters. 1991. Adaptive management of spatially replicated groundfish populations. Canadian Journal of Fisheries and Aquatic Sciences 48: 1273-1284.

Cooke, J.G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. ICES Journal of Marine Science 56: 797-810.

Cope, J. M. and A. E. Punt. 2009. Drawing the lines: resolving fishery management units with simple fisheries data. Canadian Journal of Fisheries and Aquatic Sciences 66: 1256-1273.

Cornick, L. A., S. D. Inglis, K. Willis, and M. Horning. 2006a. Effects of increased swimming costs on foraging efficiency of captive Steller sea lions: Evidence for

behavioral plasticity in the recovery phase of dives. Journal of Experimental Marine Biology and Ecology 333: 306-314.

Cornick, L. A., W. Neill, and W. E. Grant. 2006b. Assessing competition between Steller sea lions and the commercial groundfishery in Alaska: A bioenergetics modelling approach. Ecological Modelling 199: 107-114.

Crainiceanu, C. M., D. Ruppert, and M. P. Wand. 2007. Bayesian analysis for penalized spline regression using Win BUGS. Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 40.

Cressie, N., C. A. Calder, J. S. Clark, J. M. V. Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecological Applications 19: 553–570.

de la Mare, W. K. 1986. Fitting population models to time series of abundance data. Reports of the International Whaling Commission 36: 399-418.

De Oliveira, J. A. A., L. T. Kell, A. E. Punt, B. A. Roel, and D. S. Butterworth. 2008. Managing without best predictions: the Management Strategy Evaluation framework. Pages 104-134 in: Advances in Fisheries Science; 50 years on from Beverton and Holt. Blackwell Publishing, Oxford.

de Valpine, P. 2005. Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-Gaussian state-space models. Bulletin of Marine Science 70: 455-471.

de Valpine, P. and A. Hastings. 2002. Fitting population models incorporating process noise and observation error. Ecological Monographs 72: 57-76.

de Valpine, P. and R. Hilborn. 2005. State-space likelihoods for nonlinear fisheries time-series. Canadian Journal of Fisheries and Aquatic Sciences 62: 1937-1952.

Dennis, B., P. L. Munholland, and J. M. Scott. 1991. Estimation of Growth and Extinction Parameters for Endangered Species. Ecological Monographs 61: 115-144.

Diggle, P. J., J. A. Tawn, and R. A. Moyeed. 1998. Model-based geostatistics (with discussion). Applied Statistics 47: 299-350.

Dobbie, M. J., B. L. Henderson, and D. L. Stevens, Jr. 2008. Sparse sampling: Spatial design for monitoring stream networks. Statistics Surveys 2: 113-153.

Dormann, C. F., J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. D. Kissling, I. Külf, R. Ohlemüller, P. R. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30: 609-628.

Dorn, M. K. Aydin, S. Barbeaux, M. Guttormsen, K. Spallinger, and M. Wilkins. 2011. Assessment of the Walleye Pollock Stock in the Gulf of Alaska. *In*: Stock assessment and fishery evaluation report for groundfish resources of the Gulf of Alaska. Prepared by the Gulf of Alaska Groundfish Plan Team, North Pacific Fishery Management Council, P.O. Box 103136, Anchorage, AK 99510. North Pacific Fisheries Management Council, Anchorage, AK.

Dowling, N. A., D. C. Smith, I. Knuckey, A. D. M. Smith, P. Domaschenz, H. M. Patterson, and W. Whitelaw. 2008. Developing harvest strategies for low-value and data-poor fisheries: Case studies from three Australian fisheries. Fisheries Research 94: 380-390.

Dunn, A., R. I. C. C. Franics, and I. J. Doonan. 2002. Comparison of the Chapman-Robson and regression estimators of *Z* from catch-curve data when non-sampling stochastic error is present. Fisheries Research 59: 149-159.

Fadely, B. S., B. W. Robson, J. T. Sterling, A. Greig and K. A. Call. 2005. Immature Steller sea lion (*Eumetopias jubatus*) dive activity in relation to habitat features of the eastern Aleutian Islands. Fisheries Oceanography 14(S1): 243-258.

Fadely, B., L. Fritz, R. Ream, R. Towell, J. Sterling, C. Stinchcomb, W. Perryman, and T. Gelatt. 2006. Contrasting Western Steller Sea Lion and Northern Fur Seal Population Trends in Alaska. AFSC Quarterly Report Feature (January-February-March 2006). 8 p.

Fahlman, A., C. Svärd, D. A. S. Rosen, D. R. Jones, and A. W. Trites. 2008. Metbaolic costs of foraging and the management of O_2 and CO_2 stores in Steller sea lions. Journal of Experimental Biology 211: 3573-3580.

Fay, G. 2007. Tier 3 calculations for blue eye trevalla (*Hyperoglyphe antarctica*) using data up to and including 2005. Pages 524-531 in: Tuck, G. N. (Ed.). Stock assessment for the Southern and Eastern Scalefish and Shark Fishery 2006-2007, Volume 1: 2006. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Fay, G. 2009. Tier 3 calculations for blue eye trevalla (*Hyperoglyphe antarctica*) using data up to and including 2007. Pages 211-225 in: Tuck, G. N. (Ed.). Stock assessment for the Southern and Eastern Scalefish and Shark Fishery 2008. Part 2. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Fay, G. and A. E. Punt. 2006. Modeling spatial dynamics of Steller sea lions (*Eumetopias jubatus*) using maximum likelihood and Bayesian methods: evaluating causes for population decline. Pages 405-433 in: A. W. Trites, S. K. Atkinson, D. P. DeMaster, L. W. Fritz, T. S. Gelatt, L. D. Rea, and K. Wynne (Eds.). Sea lions of the world. Alaska Sea Grant College Program, University of Alaska Fairbanks.

Fay, G., A. E. Punt, and A. D. M. Smith. 2009. Operating model specifications. Pages 125-133 in: Wayte, S. E. (Ed.). Evaluation of new harvest strategies for SESSF species. CSIRO Marine and Atmospheric Research, Hobart and Australian Fisheries Management Authority, Canberra.

Field, S. A., A. J. Tyre, and H. P. Possingham. 2005. Optimizing allocation of monitoring effort under economic and observational constraints. Journal of Wildlife Management 69: 473-482.

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27: 233-249.

Francis, R. C., M. A. Hixon, M. E. Clarke, S. A. Murawski, and S. Ralston. 2007. Ten commandments for ecosystem-based fisheries scientists. Fisheries 32: 217-233.

Frid, A., J. Burns, G. G. Baker, R. E. Thorne. Predicting synergistic effects of resources and predators on foraging decisions by juvenile Steller sea lions. Oecologia 158: 775-786.

Fritz, L., and T. Gelatt. 2011. Surveys of Steller Sea Lions in Alaska, June-July 2010. Memorandum for the Record. U.S. Department of Commerce. 31 January 2011. 31 p.

Fritz, L., M. Lynn, E. Kunisch, and K. Sweeney . 2008. Aerial, ship, and land-based surveys of Steller sea lions (*Eumetopias jubatus*) in Alaska, June and July 2005-2007. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-183. 70 p.

Fulton, E. A., J. S. Link, I. C. Kaplan, M. Savina-Rolland, P. Johnson, C. Ainsworth, P. Horne, R. Gorton, R. J. Gamble, A. D. M. Smith, and D. C. Smith. 2011. Lessons in modelling and management of marine ecosystems: the Atlantis experience. Fish and Fisheries 12: 171-188.

Gelman, A. and J. Hill. 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge.

Gende, S. M. and M. F. Sigler. 2006. Persistence of forage fish 'hot spots' and its association with foraging Steller sea lions (*Eumetopias jubatus*) southeast Alaska. Deep-Sea Research Part II-Topical Studies in Oceanography 53: 432-441.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Pages 169–193 in: J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith. (Eds.). Bayesian Statistics 4. Oxford University Press, Oxford.

Gilpin, M. and I. Hanski. 1991. Metapopulation Dynamics: Empirical and Theoretical Investigations. Academic Press, London.

Gronnevik, R. and G. Evensen. 2001. Application of ensemble-based techniques in fish stock assessment. Sarsia 86: 517-526.

Gudmundsson, G. 1994. Time-series analysis of catch-at-age observations. Applied Statistics-Journal of the Royal Statistical Society Series C 43: 117-126.

Gudmundsson, G. 2004. Time series analysis of abundance indices of young fish. ICES Journal of Marine Science 61: 176-183.

Haddon, M. 2011. Catch rate standardizations 2010 (data from 1986 to 2009). Pages 44-179 in: Tuck, G. N. (Ed.). Stock assessment for the Southern and Eastern Scalefish and Shark Fishery 2010. Part 2. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Hanski, I. 1999. Metapopulation ecology. Oxford University Press. Oxford, United Kingdom.

Harvey, A. C. 1989. Forecasting, structural time series models, and the Kalman filter. Cambridge University Press, Cambridge.

Hastie, T. J. and R. Tibshirani. 1990. Generalized additive models. Chapman and Hall, London.

Hauser, C. E., A. R. Pople, and H. P. Possingham. 2006. Should managed populations be monitored every year? Ecological Applications 16: 807-819.

Heidelberger, P. and P. D. Welch. 1983. Simulation run length control in the presence of an initial transient. Operations Research 31: 1109–1144.

Hilborn, R. 2012. The evolution of quantitative marine fisheries management 1985-2010. Natural Resource Modeling 25: 122-144.

Hilborn, R. and M. Mangel. 1997. The ecological detective. Confronting models with data. Princeton University Press, Princeton.

Hilborn, R., J. M. Orensanz, and A. M. Parma. 2005. Institutions, incentives, and the future of fisheries. Philosophical Transcations of the Royal Society B 360: 47-57.

Hindell, J., P. Hamer, H. McPartlan, and S. Robertson. 2005. Preliminary assessment of the utility of otolith microchemistry, otolith shape analysis, and mitochondrial DNA analyses in stock discrimination of blue-eye trevalla (*Hyperoglyphe antarctica*) from Australian shelf waters and offshore seamounts and New Zealand. Final report to Fisheries Research and Development Corporation Project No. 2003/045. Primary Industries Research Victoria, Queenscliff.

Hoenig, J.M. 1983. Empirical use of longevity data to estimate mortality rates. Fishery Bulletin 82: 898-903.

Hoeting, J. A. 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. Ecological Applications 19: 574-577.

Holling, C. 1959. The components of predation as revealed by a study of small mammal predation of the European pine sawfly. Canadian Entomology 91: 293-320.

Holmes, E. E. 2001. Estimating risks in declining populations with poor data. Proceedings of the National Academy of Sciences of the United States of America 98: 5072-5077.

Holmes, E. E. and A. E. York. 2002. Using age structure to detect impacts on threatened populations: A case study using Steller sea lions. Conservation Biology 17: 1794-1806.

Holt, C. A., and R. A. Peterman. 2004. Long-term trends in age-specific recruitment of sockeye salmon (*Oncorhynchus nerka*) in a changing environment. Canadian Journal of Fisheries and Aquatic Sciences 61: 2455-2470.

Horn, P. L. 1988. Age and growth of bluenose, *Hyperoglyphe antractica* (Pisces: Stromateodei) from the lower east coast North Island, New Zealand. New Zealand Journal of Marine and Freshwater Research 22: 369-378.

Jensen, A. L. 1996. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Canadian Journal of Fisheries and Aquatic Sciences 53: 820-822.

Jonsen, I. D., J. Mills-Flemming, and R. A. Myers. 2005. Robust state-space modeling of animal movement data. Ecology 86: 2874-2880.

Kareiva, P. 1994. Space: The Final Frontier for Ecological Theory. Ecology 75: 1.

Kell, L. T., C. M. O'Brien, M. T. Smith, T. K. Stokes, and B. D. Rackham. 1999. An evaluation of management procedures for implementing a precautionary approach in the ICES context for North sea plaice (*Pleuronectes platessa* L.). ICES Journal of Marine Science 56: 834-845.

Kell, L. T., M. A. Pastoors, R. D. Scott, M. T. Smith, F. A. Van Beek, C. M. O'Brien, and G. M. Pilling. 2005. Evaluation of multiple management objectives for Northeast Atlantic flatfish stocks: sustainability vs. stability of yield. ICES Journal of Marine Science 62: 1104-1117.

Kimura, D. K., J. W. Balsiger, and D. H. Ito. 1996. Kalman filtering the delay-difference equation: practical approaches and simulations. Fishery Bulletin 94: 678-691.

Kirkwood, G. P. 1981. Estimation of stock size using relative abundance data - a simulation study. Reports of the International Whaling Commission 31: 729-735.

Kitagawa, G. 1981. A nonstationary time series model and its fitting by a recursive filter. Journal of Time Series Analysis 2: 103-116.

Kitts,D. D., M. D. Huynhl, C. Hu, and A. W. Trites. 2004. Season variation in nutrient composition of Alaskan walleye pollock. Canadian Journal of Zoology 82:1408-1415.

Klaer, N.L. 2009. Yield and total mortality values and Tier 3 estimates for selected shelf and slope species in the SESSF. Pages 158-210 in: Tuck, G. N. (Ed.). Stock assessment for the Southern and Eastern Scalefish and Shark Fishery 2008. Part 2. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Klaer, N. L. 2010. Tiger flathead (*Neoplatycephalus richardsoni*) stock assessment based on data up to 2008. Pages 164-189 in: Tuck, G. N. (Ed.). Stock assessment for the

Southern and Eastern Scalefish and Shark Fishery 2009. Part 1. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Klaer, N. L., S. E. Wayte, A. E. Punt, L. R. Little, A. D. M. Smith, R. B. Thomson, and G. N. Tuck. 2009. Simulation testing of alternative Tier 3 assessment methods and control rules for the SESSF. Pages 42-70 in: Wayte, S. E. (Ed.). Evaluation of new harvest strategies for SESSF species. CSIRO Marine and Atmospheric Research, Hobart and Australian Fisheries Management Authority, Canberra.

Kleiber, M. 1975. The fire of life: an introduction to animal energetics. Huntington, NY, Robert Krieger.

Kritzer, J. P., and P. F. Sale. 2006. Marine metapopulations. Academic Press, San Diego.

Lande, R., S. Engen, and B.-E. Sæther. 1999. Spatial scale of population synchrony: environmental correlation versus dispersal and density regulation. The American Naturalist 154: 271-281.

Legendre, P. 1993. Spatial autocorrelation: Trouble or new paradigm? Ecology 74: 1659–1673.

Legendre, P. and L. Legendre. 1998. Numerical ecology. 2nd English edition. Elsevier.

Levin, S. A. 1976. Population dynamics models in heterogeneous environments. Annual Review of Ecology, Evolution, and Systematics 7: 287-310.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73: 1943-1967.

Levins, R. 1970. Extinction. Pages 77-107 in: M. Gerstenhaber (Ed.). Some mathematical questions in biology. American Mathematical Society, Providence RI.

Lindley, S. T. 2003. Estimation of population growth and extinction parameters from noisy data. Ecological Applications 13: 806-813.

Little, L. R. and K. Rowling. 2009. Eastern Gemfish (*Rexea solandri*) stock assessment based on 2008 survey data. Pages 167-208 in: Tuck, G. N. (Ed.). Stock assessment for the Southern and Eastern Scalefish and Shark Fishery 2008. Part 1. Australian Fisheries Management Authority and CSIRO Marine and Atmospheric Research, Hobart.

Lorenzen, K., R. S. Steneck, R. R. Warner, A. M. Parma, F. C. Coleman, and K. M. Leber. 2010. The spatial dimension of fisheries: putting it all in place. Bulletin of Marine Science 86: 169-177.

Loughlin, T. R. 1998. The Steller sea lion: a declining species. Biosphere Conservation 1: 91-98.

Loughlin, T. R., J. T. Sterling, R. L. Merrick, J. L. Sease, A. E. York. 2003. Diving behavior of immature Steller sea lions (*Eumetopias jubatus*). Fishery Bulletin 101: 566-582.

Lovvorn, J. R., J. M. Grebmeier, L. W. Cooper, J. K. Bump, and S. E. Richman. 2009. Modeling marine protected areas for threatened eiders in a climatically changing Bering Sea. Ecological Applications 19: 1596-1613.

Ludwig, D. and C. J. Walters 1981. Measurement errors and uncertainty in parameter estimates for stock and recruitment. Canadian Journal of Fisheries and Aquatic Sciences 38: 711-720.

MacArthur, R. H. and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton.

MacCall, A. 2007. Appendix D - Depletion-adjusted average catch. Pages 27-30 in: Rosenberg, A., D. Agnew, E. Babcock, A. Cooper, C. Mogensen, R. O'Boyle, J. Powers, G. Stefansson, and J. Swasey. Setting Annual Catch Limits for U.S. fisheries: An expert working group report. Lenfest Ocean Program. Washington D.C.

MacCall, A. 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES Journal of Marine Science 66: 2267–2271.

Mace, P. M. and I. J. Doonan. 1988. A generalized bioeconomic simulation model for fish population dynamics. New Zealand Fisheries Assessment Research Document 88/4. MAFFish Fisheries Research Centre, Wellington.

Mangel, M. and C. W. Clark. 1988. Dynamic modelling in behavioural ecology. Princeton Univesity Press, Princeton.

Mapstone, B. D., L. R. Little, A. E. Punt, C. R. Davies, A. D. M. Smith, F. Pantus, F. D. McDonald, A. J. Williams, and A. Jones. 2008. Management strategy evaluation for line fishing in the Great Barrier Reef: balancing conservation and multi-sector fishery objectives. Fisheries Research 94: 315-329.

Matthiopoulos, J. 2003. Model-supervised kernel smoothing for the estimation of spatial usage. Oikos 102: 367-377.

Matthiopoulous, J., S. Smout, A. J. Winship, D. Thompson, I. L. Boyd, and J. Harwood. 2008. Getting beneath the surface of marine mammal – fisheries competition. Mammal Review 38: 167-188.

McCullagh, P. and J. A. Nelder. 1989. Generalized linear models (2nd edition). Chapman and Hall, London.

McDonald, A. D., L. R. Little, R. Gray, E. Fulton, K. J. Sainsbury, and V. D. Lyne. 2008. An agent-based modelling approach to evaluation of multiple-use management strategies for coastal marine ecosystems. Mathematics and Computer Simulations 78: 401-411.

McDonald-Madden, E., P. W. J. Baxter, and H. P. Possingham. 2008. Subpopulation triage: how to allocate conservation effort among populations. Conservation Biology 22: 656-665.

McGilliard, C. R., R. Hilborn, A. MacCall, A. E. Punt, and J. C. Field. 2011. Can information from marine protected areas be used to inform control-rule-based management of small-scale, data-poor stocks? ICES Journal of Marine Science 68: 201-211.

Mello, L. G. S. and G. A. Rose. 2005. Using geostatistics to quantify seasonal distribution and aggregation patterns of fishes: an example of Atlantic cod (*Gadus morhua*). Canadian Journal of Fisheries and Aquatic Sciences 62: 659-670.

Mendes, J. M., K. F. Turkman, and E. Jardim. 2007. A Bayesian hierarchical model for over-dispersed count data: a case study for abundance of hake recruits. Environmetrics 18: 27-53.

Merrick, R. L. and T. R. Loughlin. 1997. Foraging behavior of adult female and young-of-the-year Steller sea lions in Alaskan waters. Canadian Journal of Zoology 75: 776-786.

Merrick, R. L., M. K. Chumbley, and G. V. Byrd. 1997. Diet diversity of Steller sea lions (*Eumetopias jubatus*) and their population decline in Alaska: a potential relationship. Canadian Journal of Fisheries and Aquatic Sciences 54: 1342–1348.

Methot, R. D. 2007. User manual for the Integrated Analysis program Stock Synthesis 2 (SS2). Model Version 2.00b.

Meyer, R. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. Canadian Journal of Fisheries and Aquatic Sciences 56: 37-52.

Millar, R. B. and R. Meyer. 2000. Non-linear state space modelling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. Applied Statistics 48: 327-342.

Miller, T. J. and J. R. Skalski. 2006. Integrating design- and model-based inference to estimate length and age composition in North Pacific longline catches. Canadian Journal of Fisheries and Aquatic Sciences 63: 1092–1114.

Millette, L. L. and A. W. Trites 2003. Maternal attendance patterns of Steller sea lions (*Eumetopias jubatus*) from stable and declining populations in Alaska. Canadian Journal of Zoology 81: 340-348.

National Research Council. 2003. The decline of the Steller sea lion in Alaskan waters: Untangling food webs and fishing nets. Committee on the Alaska Groundfish Fishery and Steller Sea Lions. National Research Council of the National Academies. The National Academies Press. Washington D.C. 204 pp.

Nations, C. S. and R. C. Anderson-Sprecher. 2006. Estimation of animal location from radio telemetry data with temporal dependencies. Journal of Agricultural, Biological, and Environmental Statistics 11: 87-105.

Newman, K. B. and S. T. Lindley. 2006. Accounting for demographic and environmental stochasticity, observation error, and parameter uncertainty in fish population dynamics models. North American Journal of Fisheries Management 26: 685-701.

Ney, J. J. 1990. Trophic economics in fisheries - Assessment of demand-supply relationships between predators and prey. Reviews in Aquatic Sciences 2: 55-81.

Pardo-Igúzquiza, E. and P. A. Dowd. 1997. AMLE3D: A computer program for the inference of spatial covariance parameters by approximate maximum likelihood estimation. Computers and Geosciences 23: 793-805.

Patterson, T. A., M. Basson, M. V. Bravington, and J. S. Gunn. 2009. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. Journal of Animal Ecology 78: 1113-1123.

Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. Journal du conseil / Conseil international pour l'exploration de la mer. 39: 175-192.

Pelletier, D. and A. M. Parma. 1994. Spatial distribution of Pacific halibut (*Hippoglossus stenolepis*): an application of geostatistics to longline survey data. Canadian Journal of Fisheries and Aquatic Sciences 51: 1506-1518.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47: 2-15.

Peterman, R. M., B. J. Pyper, and J. A. Grout. 2000. Comparison of parameter estimation methods for detecting climate-induced changes in productivity of Pacific salmon (*Oncorhynchus* spp.). Canadian Journal of .Fisheries and Aquatic Sicences 57: 181–191.

Petitgas, P. 1993. Geostatistics for fish stock assessments: a review and an acoustic application. ICES Journal of Marine Science 50: 285-298.

Plagányi, E. E., D. S. Butterworth. 2011. The Scotia Sea krill fishery and its possible impacts on dependent predators – modelling localized depletion of prey. Ecological Applications 22: 748-761.

Pollock, K. H., J. D. Nichols, T. R. Simons, G. L. Farnsworth, L. L. Bailey, and J. R. Sauer. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics 13: 105-119.

Punt, A.E. and G. Fay. 2006. Can experimental manipulation be used to determine the cause of the decline of western stock of Steller sea lions (*Eumetopias jubatus*)? Pages 435-454 in: A.W. Trites, S. Atkinson, D.P. DeMaster, L.W. Fritz, T.S. Gelatt, L.D. Rea, and K. Wynne (Eds.). Sea lions of the world. Alaska Sea Grant College Program, University of Alaska Fairbanks.

Punt, A. E., A. D. M. Smith, and G. Cui. 2002. Evaluation of management tools for Australia's South East Fishery. 2. How well do commonly-used stock assessment methods perform? Marine and Freshwater Research 53: 631-644.

Pyke, G. H. 1984. Optimal foraging theory: a critical review. Annual Review of Ecology and Systematics. 15: 523-575.

Raftery, A. E. and S. Lewis. 1992. How many iterations in the Gibbs sampler? Pages 763–773 in: J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith. (Eds.). Bayesian Statistics 4. Oxford University Press, Oxford.

Raum-Suryan, K. L., K. W. Pitcher, D. G. Calkins, J. L. Sease, and T. R. Loughlin. 2002. Dispersal, rookery fidelity, and metapopulation structure of Steller sea lions (*Eumetopias jubatus*) in an increasing and a decreasing population in Alaska. Marine Mammal Science 18: 746-764.

Real, L. A. 1977. The kinetics of functional response. The American Naturalist 111: 289-300.

Reed, W. J. and C. M. Simons. 1996. Analyzing catch-effort data by means of the Kalman filter. Canadian Journal of Fisheries and Aquatic Sciences 53: 2157-2166.

Rehberg, M. J., R. D. Andrews, U. G. Swain, and D. G. Calkins. 2009. Foraging behaviour of adult female Steller sea lions during the breeding season in Southeast Alaska. Marine Mammal Science 25: 588-604.

Restrepo, V.R. and J. E. Powers. 1999. Precautionary control rules in US fisheries management: specification and performance. ICES Journal of Marine Science 56: 846-852.

Rhodes, J. R. and N. Jonzén. 2011. Monitoring temporal trends in spatially structured populations: how should sampling effort be allocated between space and time? Ecography 34: 1040-1048.

Rhodes, J. R., A. J. Tyre, N. Jonzén, C. A. McAlpine, and H. P. Possingham. 2006. Optimizing presence-absence surveys for detecting population trends. Journal of Wildlife Management 70: 8-18.

Richards, L. J. 1991. Use of contradictory data sources in stock assessments. Fisheries Research 11: 225-238.

Richards, L. J., J. T. Schnute, and N. Olsen. 1997. Visualizing catch-age analysis: a case study. Canadian Journal of Fisheries and Aquatic Sciences 54: 1646-1658.

Rivoirard, J., N. Bez, P. Fernandes, K. Foote, and J. Simmonds. 2000. Geostatistics for estimating fish abundance. Blackwell Science, Oxford.

Rodríguez-Sánchez R., D. Lluch-Belda, H. Villalobos, and S. Ortega-García. 2002. Dynamic geography of small pelagic fish populations in the California Current System on the regime time scale (1931–1997). Canadian Journal of Fisheries and Aquatic Sciences 59: 1980-1988.

Rose, G. A., and D. W. Kulka. 1999. Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. Canadian Journal of Fisheries and Aquatic Sciences 56: 118–127.

Rosen, D. A. S. and A. W. Trites. 1997. Heat increment of feeding in Steller sea lions, *Eumetopias jubatus*. Comparative Biochemistry and Physiology Part A: Physiology 118: 877-881.

Rosen, D. A. S. and A. W. Trites. 2000. Digestive efficiency and dry-matter digestibility in Steller sea lions fed herring, pollock, squid, and salmon. Canadian Journal of Zoology 78: 234-239.

Rosen, D. A. S. and A. W. Trites. 2002a. Changes in metabolism in response to fasting and food restriction in the Steller sea lion (*Eumetopias jubatus*). Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology 132: 389-399.

Rosen, D. A. S. and A. W. Trites. 2002b. Cost of transport in Steller sea lions, *Eumetopias jubatus*. Marine Mammal Science 18: 513-524.

Rosen, D. A. S. and A. W. Trites. 2003. No evidence for bioenergetic interaction between digestion and thermoregulation in Steller sea lions *Eumetopias jubatus*. Physiological and Biochemical Zoology 76: 899-906.

Rosen, D. A. S. and A. W. Trites. 2004. Satiation and compensation for short-term changes in food quality and availability in young Steller sea lions (*Eumetopias jubatus*). Canadian Journal of Zoology 82: 1061-1069.

Royer, F., J.-M. Fromentin, and P. Gaspar. 2005. A state-space model to derive bluefin tuna movement and habitat from archival tags. Oikos 109: 473-484.

Sanchirico, J. N., and J. E. Wilen. 1999. Bioeconomics of spatial exploitation in a patchy environment. Journal of Environmental Economics and Management 37: 129-150.

Schnute, J. T. 1994. A general framework for developing sequential fisheries models. Canadian Journal of Fisheries and Aquatic Sciences 51: 1676-1688.

Schnute, J. and R. Hilborn. 1993. Analysis of contradictory data sources in fish stock assessment. Canadian Journal of Fisheries and Aquatic Science 50: 1916-1923.

Sigler, M. F., D. J. Tollit, J. J. Vollenweider, J. F. Thedinga, D. J. Csepp, J. N. Womble, M. A. Wong, M. J. Rehberg, and A. W. Trites. 2009. Steller sea lion foraging response to seasonal changes in prey availability. Marine Ecology Progress Series 388: 243-261.

Silvey, S.D. 1980. Optimal design: an introduction to the theory for parameter estimation. Chapman and Hall, London.

Sinclair, E. H., S. E. Moore, N. A. Friday, T. K. Zeppelin, and J. M. Waite. 2005. Do patterns of Steller sea lion (*Eumetopias jubatus*) diet, population trend and cetacean occurrence reflect oceanographic domains from the Alaska Peninsula to the central Aleutian Islands? Fisheries Oceanography 14(S1): 223-242.

Sinclair, E. H. and T. K. Zeppelin. 2002. Seasonal and spatial differences in diet in the western stock of Steller sea lions (*Eumetopias jubatus*). Journal of Mammalogy 83: 973-990.

Sinopoli, B., L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry. 2004. Kalman filtering with intermittent observations. IEEE Transactions on Automatic Control 49: 1453-1464.

Smith, A. D. M. and D. C. Smith. 2005. A harvest strategy framework for the SESSF. Report to the Australian Fisheries Management Authority, Canberra.

Smith, A. D. M. and S. E. Wayte (eds). 2002. The South East Fishery 2001. Fishery Assessment Report compiled by the South East Fishery Assessment Group. Australian Fisheries Management Authority, Canberra.

Smith, A. D. M., K. J. Sainsbury, and R. A. Stevens. 1999. Implementing effective fisheries management systems – management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science 56: 967-979.

Smith, A. D. M., D. C. Smith, G. N. Tuck, N. Klaer, A. E. Punt, I. Knuckey, J. Prince, A. Morison, R. Kloser, M. Haddon, S. Wayte, J. Day, G. Fay, F. Pribac, M. Fuller, B. Taylor, and L. R. Little. 2008. Experience in implementing harvest strategies in Australia's south-eastern fisheries. Fisheries Research 94: 373-379.

Spencer, P. D. and J. N. Ianelli. 2005. Application of a Kalman filter to a multispecies stock complex. Pages 613-634 in: G. H. Kruse, V. F. Gallucci, D. E. Hay, R. I. Perry, R. M. Peterman, T. C. Shirley, P. D. Spencer, B. Wilson, and D. Woodby (Eds.). Fisheries assessment and management in data-limited situations. Alaska Sea Grant College Program. AK-SG-05-02.

Spiegelhalter, D., A. Thomas, and N. Best. 2000. WinBugs version 1.3 user manual. Medical Research Council Biostatistics Unit, Cambridge. http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml.

Sullivan, P. 1992. A Kalman filter approach to catch-at-length-analysis. Biometrics 48: 237-257.

Svärd, C., A. Fahlman, D. A. S. Rosen, R. Joy, and A. W. Trites. 2009. Fasting affects the surface and diving metabolic rates of Steller sea lions (*Eumetopias jubatus*). Aquatic Biology 8: 71-82.

Ter Braak, C. J. F. 1986. Canonical Correspondence Analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167-1179.

Thogmartin, W. E., J. R. Sauer, and M. G. Knutson. 2004. A hierarchical spatial model of avian abundance with applications to cerulean warblers. Ecological Applications 12: 1766-1779.

Thorson, J. T., I. J. Stewart, and A. E. Punt. 2012. Development and application of an agent-based model to evaluate methods for estimating relative abundance indices for shoaling fish such as Pacific rockfish (*Sebastes* spp.). ICES Journal of Marine Science 69: 635–647.

Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46: 234-240.

Trites, A. W. and C. P. Donnelly. 2003. The decline of Steller sea lions *Eumetopias jubatus* in Alaska: a review of the nutritional stress hypothesis. Mammal Review 33: 3-28.

Trites, A. W. and P. A. Larkin. 1996. Changes in the abundance of Steller sea lions (*Eumetopias jubatus*) in Alaska from 1956 to 1992: how many were there? Aquatic Mammals 22: 153-166.

Trites, A. W. and B. T. Porter. 2002. Attendance patterns of Steller sea lions (*Eumetopias jubatus*) and their young during winter. Journal of Zoology 256: 547-556.

Trites, A. W., B. P. Porter, V. B. Deecke, A. P. Coombs, M. L. Marcotte, and D. A. S. Rosen. 2006. Insights into the timing and weaning and the attendance patterns of lactating Steller sea lions (*Eumetopias jubatus*) in Alaska during winter, spring, and summer. Aquatic Mammals 32: 85-97.

Tvete, I. F., G. Storvik, and B. Natvig. 2006. Modelling Pollock egg counts from the western Gulf of Alaska by a zero-inflated Bayesian hierarchical space-time model. UIO note. University of Oslo, Norway.

Walters, C. J. 2004. Simple representation of the dynamics of biomass error propagation for stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 61: 1061-1065.

Walters, C. J. and S. J. D. Martell. 2004. Fisheries Ecology and Management. Princeton University Press, Princeton.

Ward, E. J., R. Hilborn, R. G. Towell, and L. Gerber. 2007. A state-space mixture approach for estimating catastrophic events in time series data. Canadian Journal of Fisheries and Aquatic Sciences 64: 899-910.

Wayte, S.E. and N. L. Klaer. 2010. An effective harvest strategy using improved catch-curves. Fisheries Research 106: 310-320.

Wieland, K. and J. Rivoirard. 2001. A geostatistical analysis of IBTS data for age 2 North Sea haddock (*Melanogrammus aeglefinus*) considering daylight effects. Sarsia 86: 503-516.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3: 385-397.

Wilkins, M. E. 2007. Bottom trawl survey of groundfish resources in the Gulf of Alaska. Cruise Results – Gulf of Alaska Biennial Bottom Trawl Survey. 2007-01. NMFS / AFSC / RACE / GOA.

Wilson, K. A., M. F. McBride, M. Bode, and H. P. Possingham. 2006. Prioritizing global conservation efforts. Nature 440: 337-340.

Winship, A. J. 2000. Growth and bioenergetic models for Steller sea lions (*Eumetopias jubatus*) in Alaska. MSc thesis, University of British Columbia, Vancouver.

Winship, A. J. and A. W. Trites. 2003. Prey consumption of Steller sea lions (*Eumetopias jubatus*) off Alaska: How much prey do they require? Fishery Bulletin 101: 147-167.

Winship, A. J. and A. W. Trites. 2006. Risk of extirpation of Steller sea lions in the Gulf of Alaska and Aleutian Islands: a population viability analysis based on alternative hypotheses for why sea lions declined in western Alaska. Marine Mammal Science 22: 124-155.

Winship, A. J., A. W. Trites, and D. G. Calkins. 2001. Growth in body size of the Steller sea lion (*Eumetopias jubatus*). Journal of Mammalogy 82: 500-519.

Winship, A. J., A. W. Trites, and D. A. S. Rosen. 2002. A bioenergetic model for estimating the food requirements of Steller sea lions *Eumetopias jubatus* in Alaska, USA. Marine Ecology Progress Series 229: 291-312.

Witteveen B. H., R. J. Foy, K. M. Wynne, and Y. Tremblay. 2008. Investigation of foraging habits and prey selection by humpback whales (*Megaptera novaengliae*) using acoustic tags and concurrent fish surveys. Marine Mammal Science 24: 516-534.

Wolf, N., J. Melbourne, and M. Mangel. 2006. The method of multiple hypotheses and the decline of Steller sea lions in western Alaska. pp. 275–293 in: I. Boyd, S. Wanless, and C. J. Camphusen (eds.) Top predators in marine ecosystems. their role in monitoring and management. Cambridge: Cambridge University Press.

Womble, J. N., M. F. Willson, M.F. Sigler, B.P. Kelly, and G.R. VanBlaricom. 2005. Distribution of Steller sea lions *Eumetopias jubatus* in relation to spring-spawning fish in SE Alaska. Marine Ecology Progress Series 294: 271-282.

Worm, B., R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, E. A. Fulton, J. A. Hutchings, S. Jennings, O. P. Jensen, H. K. Lotze, P. M. Mace, T. R. McClanahan, C. Minto, S. R. Palumbi, A. M. Parma, D. Ricard, A. A. Rosenberg, R. Watson, and D. Zeller. 2009. Rebuilding global fisheries. Science 325: 578-585.

York, A. E., Merrick, R. L. and T. R. Loughlin. 1996. An analysis of the Steller sea lion metapopulation in Alaska. Pages 259-292 in: McCullough, D.R. (Ed.). Metapopulations and Wildlife Conservation. Island Press, Covela.

Zhao, Y., J. Staudenmayer, B. A. Coull, and M. P. Wand. 2006. General Design Bayesian Generalized Linear Mixed Models. Statistical Science 21: 35-51.

Zhu, Z. and M. L. Stein. 2006. Spatial sampling design for prediction with estimated parameters. Journal of Agricultural, Bioogical, and Enviromental Statistics 11: 24-44.

Zimmerman, D. L. 2006. Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction. Environmetrics 17: 635-652.

Appendix A: The Kalman filter and Likelihood

The Kalman filter is a recursive procedure which estimates a state vector $\boldsymbol{\alpha}_y$, characterized by an estimate of its mean and variance-covariance matrix, given an initial value for the state vector and the variance-covariance matrix, and the various model parameters.

Equation (1.1) is put in the following form to apply the Kalman filter and estimate the state variables given values for the variance of process and observation error:

$$\underline{\boldsymbol{\alpha}}_{y+1} = \mathbf{T}\underline{\boldsymbol{\alpha}}_y + \mathbf{R}\underline{\eta}_y \tag{A.1}$$

where $\underline{\boldsymbol{\alpha}}_y$ is the state vector in year *y,* which for Model #1 is:

$$\boldsymbol{\alpha}_y = \left(\ln P_{1,y}^A, \ln P_{2,y}^A, ..., \ln P_{n,y}^A, \beta_y^1, ..., \beta_y^{N_A}\right)^T$$

and for Model #2:

$$\boldsymbol{\alpha}_y = \left(\ln P_{1,y}, \ln P_{2,y}, ..., \ln P_{n,y}, \beta_{1,y}, \beta_{2,y}, ..., \beta_{n,y}\right)^T$$

where n is the number of rookeries,

A is the region to which rookery r is assigned,

N_A is the number of regions,

$\underline{\eta}_y$ is a vector of process errors during year y, with the $\underline{\eta}_y$ having distribution $\underline{\eta}_y \sim N(0, \mathbf{Q})$, where the first n elements of the diagonal of $\mathbf{Q}$ equal σ_r^2, and for Model #1, the last N_A elements of the diagonal of $\mathbf{Q}$ equal to σ_β^2, and for Model #2, the lower-right quadrant of $\mathbf{Q}$ is equal to $\mathbf{W}$ as defined in equation (1.2).

T and **R** are (time-invariant) update matrices describing how the state variables and process error terms relate to each other. For the analyses here, **R** is the identity matrix, and **T** is set up such that equation (A.1) reflects equations (1.1) and (1.2). The state vector $\underline{\boldsymbol{\alpha}}_y$ therefore consists of the logarithms of pup abundance for year *y* for each rookery, and the annual trends $\beta_{r,y}$ for each region or rookery, depending on whether Model #1 or Model #2 is being used for estimation.

Placing equation (1.4) in vector form for application of the Kalman filter gives:

$$\underline{\mathbf{y}}_y = \mathbf{Z}\underline{\boldsymbol{\alpha}}_y + \underline{\varepsilon}_y \tag{A.2}$$

where $\underline{\mathbf{y}}_y$ is the vector of the natural logarithms of observed pup counts at rookeries in year *y*,

Z is a matrix which describes how the observations relate to the state vector, which for Model #1 with one region of five rookeries is:

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

and for Model #2 is the identity matrix with dimensions *n* x *n*.

$\underline{\varepsilon}_y$ is the vector of observation errors, with distribution $\underline{\varepsilon}_y \sim N(0, \mathbf{H}_y)$, where for the case of five rookeries:

$$\mathbf{H}_y = \begin{pmatrix} x_{1,y} & 0 & 0 & 0 & 0 \\ 0 & x_{2,y} & 0 & 0 & 0 \\ 0 & 0 & x_{3,y} & 0 & 0 \\ 0 & 0 & 0 & x_{4,y} & 0 \\ 0 & 0 & 0 & 0 & x_{5,y} \end{pmatrix} \tag{A.3}$$

The non-zero elements of $\mathbf{H}_y$ are defined by:

$$x_{r,y} = \begin{cases} \sigma_\varepsilon^2 \text{ if } y_{r,y} > 0 \\ \infty \text{ otherwise} \end{cases} \tag{A.4}$$

Equation (A.4) accommodates 'missing' values, that is, years in which observations were not made at all rookeries, by treating the missing observation as being highly uncertain.

Given a distribution for the state vector at the start of time-step *y*-1, the Kalman filter proceeds by forecasting the distribution for the state vector at the start of time-step *y* using the dynamics (one of equations 1.1 or 1.2), and then updating this distribution based on the observed data for time-step *y*. With a specified distribution for the initial state, this forecasting-updating process is applied recursively to compute distributions for the state vector (captured by its mean and variance-covariance matrix) for each time-step. The Kalman filter propagates the expected value $\mathbf{a}_y$ of the state vector $\boldsymbol{\alpha}_y$, and the uncertainty of estimate, $\mathbf{P}_y$, given the process error and measurements of the state vector. At each time step, the estimate of the state vector and corresponding covariance matrix is updated through two calculations: first, a *prediction* is made of the new state vector given the data to the previous time-step, the system dynamics, and the magnitude of process error, $\mathbf{a}_{y|y-1}$ and $\mathbf{P}_{y|y-1}$; second, the prediction is compared to the observations at the current time step and the prediction is *updated* given the new data to obtain an estimate contingent on the full data to that time step, $\mathbf{a}_{y|y}$, and $\mathbf{P}_{y|y}$. The update is based on how

well $\mathbf{a}_{y|y-1}$ predicts the observation vector in year *y*, $\mathbf{y}_y$. The residual between the observation and its prediction is assumed to be observation error and little change is made to the estimate of the state if the measurement is very uncertain and the state estimate is relatively precise. Conversely, the residual contains considerable information about errors in the state estimate and a large correction could be made to the state estimate if the uncertainty in the measurement is small and that in the state estimate is large.

Prediction

The predicted estimate of the mean of the state vector $\boldsymbol{\alpha}_y$, $\mathbf{a}_y$ is:

$$\mathbf{a}_{y|y-1} = \mathbf{T}\mathbf{a}_{y-1|y-1} \tag{A.5}$$

where $\mathbf{a}_{y|y-1}$ is the mean of the estimate of the state vector in year *y* given the data up to and including year *y*-1, and

$\mathbf{a}_{y|y}$ is the mean of the estimate of the state vector in year *y* given all the data up to and including year *y*.

Similarly, the predicted covariance matrix of the estimation error is given by:

$$\mathbf{P}_{y|y-1} = \mathbf{T}\mathbf{P}_{y-1|y-1}\mathbf{T}^T + \mathbf{R}\mathbf{Q}\mathbf{R}^T \tag{A.6}$$

where $\mathbf{P}_{y|y-1}$ is the estimated covariance matrix in year *y* given the data to year *y*-1, and

$\mathbf{P}_{y|y}$ is the estimation error covariance matrix in year *y* given all the data up to and including year *y*.

Update

Following the prediction, $\mathbf{a}_{y|y-1}$ and $\mathbf{P}_{y|y-1}$ are updated based on the available data in year *y*:

$$\mathbf{a}_{y|y} = \mathbf{a}_{y|y-1} + \mathbf{P}_{y|y-1}\mathbf{Z}'\mathbf{F}_y^{-1}\left(\mathbf{y}_y - \mathbf{Z}\mathbf{a}_{y|y-1}\right) \tag{A.7}$$

$$\mathbf{P}_{y|y} = \mathbf{P}_{y|y-1} - \mathbf{P}_{y|y-1}\mathbf{Z}'\mathbf{F}_y^{-1}\mathbf{Z}\mathbf{P}_{y|y-1} \tag{A.8}$$

where $\mathbf{F}_y$ is the variance-covariance matrix for year *y*:

$$\mathbf{F}_y = \mathbf{Z}_y\mathbf{P}_{y|y-1}\mathbf{Z}_y^T + \mathbf{H}_y \tag{A.9}$$

The quantity $\mathbf{P}_{y|y-1}\mathbf{Z}'\mathbf{F}_y^{-1}$ is often referred to as the Kalman Gain, $\mathbf{K}_y$.

Parameter estimation

Recursive application of the Kalman filter results in an estimated time series of the distribution of the state vector given an initial distribution ($\mathbf{a}_0$ and $\mathbf{P}_0$), values for the process and observation error variances ($\sigma^2_{r^A}$, $\sigma^2_{\beta^A}$, and $\sigma^2_{\varepsilon^A}$), and the other parameters of the model (ψ^2 and φ_{100}, Model #2). Since the distribution for $\mathbf{y}_y$ contingent on all the data to the previous time step, $\mathbf{Y}_{y-1}$, is normal ($\mathbf{y}_y \mid \mathbf{Y}_{y-1} \sim N(\mathbf{Z}\mathbf{a}_{y|y-1}, \mathbf{F}_y)$), the maximum likelihood estimates of the parameters of the model can be obtained by minimising the negative log-likelihood function:

$$-\ln L(\mathbf{y} \mid \theta) = 0.5\sum_y \left[\ln\left(\left|\mathbf{F}_y\right|\right) + \left(\mathbf{y}_y - \mathbf{Z}\mathbf{a}_{y|y-1}\right)\mathbf{F}_y^{-1}\left(\mathbf{y}_y - \mathbf{Z}\mathbf{a}_{y|y-1}\right)^T \right] \tag{A.10}$$

Backward smoothing

Following application of the Kalman filter, a fixed interval smoother (Harvey 1990, Appendix A) is applied, beginning with the final quantities $\mathbf{a}_{Y|Y}$ and $\mathbf{P}_{Y|Y}$, and moving backwards through the state. This backward smoother conditions the state space estimates on all of the data (i.e. $\mathbf{a}_{y|Y}$ where $Y > y$). A fixed interval smoother (Harvey 1989) was applied to the pup production estimates:

$$\begin{aligned} \mathbf{a}_{y|Y} &= \mathbf{a}_{y|y} + \mathbf{P}_y^* \left(\mathbf{a}_{y+1|Y} - \mathbf{T}\mathbf{a}_{y|y} \right) \\ \mathbf{P}_{y|Y} &= \mathbf{P}_{y|y} + \mathbf{P}_y^* \left(\mathbf{P}_{y+1|Y} - \mathbf{P}_{y+1|y} \right) \mathbf{P}_y^{*T} \end{aligned} \tag{A.11}$$

where:

$$\mathbf{P}_y^* = \mathbf{P}_{y|y} \mathbf{T}^T \mathbf{P}_{y+1|y}^{-1} \tag{A.12}$$

Appendix B: Data tables

Table B.1. Subset and spatial configuration of rookeries used as the basis for the simulations

#	Region	Rookery	Latitude (ºN)	Longitude (º W)	great circle distance (km) between rookeries														
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Eastern Gulf	Seal Rocks	60.16	146.84	0	62	402	582	592	1071	1169	1253	1475	1513	1746	1821	2164	2220	2443
2		Wooded (Fish)	59.88	147.34	62	0	342	521	532	1009	1108	1191	1414	1451	1685	1760	2103	2159	2382
3	Central Gulf	Outer (Pye)	59.34	150.38	402	342	0	192	190	686	777	860	1082	1121	1349	1424	1765	1821	2043
4		Marmot	58.23	151.80	582	521	192	0	70	494	588	671	894	932	1164	1239	1582	1638	1861
5		Sugarloaf	58.89	152.04	592	532	190	70	0	509	593	675	896	936	1161	1235	1575	1632	1854
6		Chirikof	55.78	155.66	1071	1009	686	494	509	0	117	195	411	446	687	764	1110	1164	1391
7		Chowiet	56.01	156.69	1169	1108	777	588	593	117	0	84	306	344	579	655	1001	1055	1281
8	Western Gulf	Lighthouse Rocks	55.78	157.41	1253	1191	860	671	675	195	84	0	223	261	496	572	917	972	1198
9		Atkins	55.05	159.29	1475	1414	1082	894	896	411	306	223	0	43	277	354	699	753	980
10		Chernabura	54.75	159.55	1513	1451	1121	932	936	446	344	261	43	0	246	323	668	721	948
11		Pinnacle Rocks	54.77	161.76	1746	1685	1349	1164	1161	687	579	496	277	246	0	77	423	477	704
12		Clubbing Rocks	54.71	162.45	1821	1760	1424	1239	1235	764	655	572	354	323	77	0	346	400	627
13	Eastern Aleutians	Akun (Billings Head)	54.29	165.53	2164	2103	1765	1582	1575	1110	1001	917	699	668	423	346	0	57	281
14		Akutan (Cape Morgan)	54.06	165.99	2220	2159	1821	1638	1632	1164	1055	972	753	721	477	400	57	0	228
15		Bogoslof	53.93	168.03	2443	2382	2043	1861	1854	1391	1281	1198	980	948	704	627	281	228	0

Table B.2. Names and locations of rookeries used in the estimation of pup production, and number of years with counts at each rookery during the period 1978-2010.

#	Region	Rookery	Latitude (ºN)	Longitude	(º W/E)	# counts
1	Eastern Gulf	Seal Rocks	60.16	146.84	W	17
2		Wooded (Fish)	59.88	147.34	W	14
3		Chiswell Islands	59.60	149.57	W	6
4	Central Gulf	Outer (Pye)	59.34	150.38	W	16
5		Sugarloaf	58.89	152.04	W	14
6		Ushagat	58.91	152.37	W	5
7		Marmot	58.23	151.80	W	20
8		Two-headed	56.91	153.55	W	8
9		Chowiet	56.01	156.69	W	12
10		Chirikof	55.78	155.66	W	15
11		Nagai Rocks	55.83	155.79	W	5
12	Western Gulf	Lighthouse Rocks	55.78	157.41	W	6
13		Atkins	55.05	159.29	W	16
14		Chernabura	54.75	159.55	W	13
15		The Whaleback	55.28	160.08	W	6
16		Jude	55.26	161.10	W	6
17		Pinnacle Rock	54.77	161.76	W	12
18		Clubbing Rocks	54.71	162.45	W	9
19		South Rocks	54.30	162.69	W	4
20	Eastern Aleutians	Sea Lion Rock (Amak)	55.46	163.20	W	7
21		Aiktak	54.18	164.85	W	5
22		Ugamak (and Round)	54.23	164.79	W	13
23		Akun (Billings Head)	54.29	165.53	W	11
24		Akutan (Cape Morgan)	54.06	165.99	W	10
25		Bogoslof	53.93	168.03	W	17
26		Ogchul	53.00	168.40	W	9
27		Adugak	52.91	169.18	W	11
28	Central Aleutians	Yunaska	52.69	170.61	W	12
29		Seguam (Saddleridge)	52.35	172.57	W	12
30		Seguam (Turf Point)	52.26	172.52	W	6
31		Amlia (Sviechnikof Harbor)	52.03	173.40	W	7
32		Kasatochi	52.19	175.52	W	12
33		Adak (Lake Point-Cape Yak	51.62	176.99	W	9
34		Kanaga (Ship Rock)	51.78	177.35	W	5
35		Gramp Rock	51.48	178.34	W	10
36		Tag	51.56	178.58	W	10
37		Ulak (Hasgox Point)	51.32	178.98	W	11
38		Semisopochnoi	51.96	179.77	E	8
39		Amchitka (East Cape)	51.37	179.47	E	7
40		Amchitka (Column Rocks)	51.54	178.82	E	8
41		Ayugadak	51.76	178.41	E	9
42		Kiska (Lief Cove)	51.95	177.34	E	11
43		Kiska (Cape St. Stephen)	51.88	177.21	E	10
44	Western Aleutians	Buldir	52.34	175.90	E	8
45		Agattu (Cape Sabak)	52.38	173.72	E	9
46		Agattu (Gillon Point)	52.40	173.36	E	6
47		Attu (Cape Wrangell)	52.91	172.47	E	6

Appendix C: Operating Model Specifications

The operating model consists of an age-structured population dynamics model, a data-generation module, and a component to allow future projections of the population model given input from estimation methods and HCRs. The operating model can be appropriately dimensioned and parameterised to account for several spatial regions and multiple fleets in order to capture the key dynamics for blue eye trevalla. The specifications are a simplified version of those used to evaluate management strategies for a variety of species within the SESSF (Fay et al. 2009).

C.1 Population dynamics

The operating model includes one or more regions; population dynamics operate at the level of the fish stock, with a single stock occupying one or more regions. Fishing fleets operate in one or more regions.

C.1.1 Abundance dynamics

The number of animals of sex s and age a in region r at the start of year t, $N^r_{s,a,t}$ is given by:

$$N^r_{s,a,t} = \begin{cases} \tilde{N}^r_{s,a-1,t-1} & \text{if } 1 \le a < x \\ \tilde{N}^r_{s,x-1,t-1} + \tilde{N}^r_{s,x,t-1} & \text{otherwise} \end{cases} \tag{C.1}$$

where $\tilde{N}^r_{s,a,t}$ is the number of animals of sex s and age a in region r following mortality (all sources) and movement during year t:

$$\tilde{N}^r_{s,a,t} = \bar{N}^r_{s,a,t} + \sum_{r' \neq r} \bar{N}^{r'}_{s,a,t} X^{r',r}_{s,a} - \sum_{r' \neq r} \bar{N}^r_{s,a,t} X^{r,r'}_{s,a} \tag{C.2}$$

$$\bar{N}^r_{s,a,t} = N^r_{s,a,t} e^{-M} (1 - u^r_{s,a,t}) \tag{C.3}$$

$X^{r',r}_{s,a}$ is the proportion of animals of sex s and age a moving from region r' to region r,

M is the rate of natural mortality,

$u^r_{s,a,t}$ the exploitation rate (due to all fleets) on animals of sex s and age a in region r during year t:

$$u^r_{s,a,t} = \sum_f \tilde{u}^{f,r}_t s^f_{s,a} \tag{C.4}$$

where:
$$\tilde{u}^{f,r}_t = \frac{C^{f,r}_t}{\sum_s \sum_l w_{l,s} S^f_l \sum_a \Phi_{l,s,a} N^r_{s,a,t} e^{-0.5M}} \tag{C.5}$$

$C^{f,r}_t$ is the retained catch by fleet f in region r during year t,

S^f_l is the selectivity of fleet f on animals in length bin l,

$\Phi_{l,s,a}$ is the proportion of fish of sex s and age a in length bin l,

$s^f_{s,a}$ is the selectivity of fleet f on animals of sex s and age a,

$w_{l,s}$ is the mean weight of a fish of sex s in length bin l, and

x is the maximum age (treated as a plus-group).

C.1.2 Selectivity

The sex- and age-specific selectivity pattern for fleet f is calculated from the inputted length-specific selectivity pattern:

$$s^f_{s,a} = \sum_{l=1}^{N_L} S^f_l \Phi_{l,s,a} \tag{C.6}$$

$$\Phi_{l,s,a} = \begin{cases} \tilde{\Phi}\left(\dfrac{L^{hi}_l - \bar{L}_{s,a}}{\sigma_{L_{s,a}}}\right) & \text{if } l = 1 \\ \tilde{\Phi}\left(\dfrac{L^{lo}_{l+1} - \bar{L}_{s,a}}{\sigma_{L_{s,a}}}\right) - \tilde{\Phi}\left(\dfrac{L^{lo}_l - \bar{L}_{s,a}}{\sigma_{L_{s,a}}}\right) & \text{if } 1 < l < N_L \\ 1 - \tilde{\Phi}\left(\dfrac{L^{lo}_l - \bar{L}_{s,a}}{\sigma_{L_{s,a}}}\right) & \text{if } l = N_L \end{cases} \tag{C.7}$$

where L_l^{hi} and L_l^{lo} are upper and lower limits of length bin *l*,

N_L is the number of length bins,

$\tilde{\Phi}$ is the standard normal cumulative distribution function,

$\bar{L}_{s,a}$ is the mean length of a fish of sex *s* and age *a* in the middle of the year, and

$\sigma_{l_{s,a}}$ is the input standard deviation of the length of a fish of sex *s* and age *a*.

C.1.3 Growth

The mean length-at-age by sex is calculated by:

$$\bar{L}_{s,a} = L_{\infty,s}\left(1-\exp\left[k_s\,(a-t_{0,s})\right]\right) \tag{C.8}$$

where $L_{\infty,s}$, k_s, and $t_{0,s}$ are the input growth parameters for animals of sex *s*.

Weight-at-length is governed by a length-power relationship:

$$w_{L,s} = \alpha_s\,(L)^{\beta_s} \tag{C.9}$$

where α_s and β_s are the input parameters of the weight-length relationship for sex *s*, and $L_l = L_l^{lo} + 0.5\left[L_l^{hi} - L_l^{lo}\right]$ is the average length of animals in length bin *l*.

C.1.4 Recruitment

The annual recruitments (by region) are log-normally distributed about an underlying Beverton-Holt stock-recruitment relationship (SRR):

$$N_{s,0,t}^r = 0.5\,R_t^r\,e^{\varepsilon_t - 0.5\sigma_R^2} \tag{C.10}$$

$$R_t^r = \lambda_t^r\left(\frac{4hR_0SB_t}{SB_0\,(1-h)+SB_t\,(5h-1)}\right) \tag{C.11}$$

where ε_t is the recruitment residual for year t [$\varepsilon_t \sim N(0, \sigma_R^2)$],

R_0 is the number of age-0 animals at pre-exploitation equilibrium,

h is the steepness of the stock-recruitment relationship,

SB_0 is the spawning biomass at pre-exploitation equilibrium (when recruitment equals R_0),

σ_R is the standard deviation of the recruitment residuals, and

λ_t^r is the expected fraction of the number of age-0 animals assigned to region r during year t:

$$\lambda_t^r = \tilde{SB}_t^r / SB_t \tag{C.12}$$

The total spawning biomass during year t is given by:

$$SB_t = \sum_r \tilde{SB}_t^r = \sum_r \sum_{a=1}^{x} N_{\text{fem},a,t}^r \tilde{w}_{\text{fem},a} f_a \tag{C.13}$$

where f_a is the fraction of females of age a that are mature, and

$\tilde{w}_{\text{fem},a}$ is the weight at age of a female of age a at the start of the year:

$$\tilde{w}_{\text{fem},a} = \sum_L \Phi_{L,\text{fem},a} w_{L,\text{fem}} \tag{C.14}$$

C.1.5 Movement

The probabilities of moving among regions are determined by:

$$X_{s,a}^{r',r} = \frac{\bar{X}_{s,a}^{r',r}}{\sum_{r'=1}^{N_{reg}} \bar{X}_{s,a}^{r',r}} \tag{C.15}$$

where $\bar{X}_{s,a}^{r',r}$ is the average probability of an animal of sex s and age a moving from region *r'* to region *r*:

$$\bar{X}_{s,a}^{r',r} = \begin{cases} T_s^{r',r} m_{s,a} & \text{if } r' \neq r \\ 1 - \sum_{r' \neq r} T_s^{r',r} m_{s,a} & \text{otherwise} \end{cases} \tag{C.16}$$

$T_s^{r',r}$ is the maximum average probability of moving from region *r'* to region *r*, with $T_s^{r'+} = 1$,

$m_{s,a}$ is the relative age-specific movement rate for an animal of sex *s*, and

N_{reg} is the number of regions.

C.1.6 Initial Conditions

The initial (*t*=1) numbers at age for each sex by region are determined by solving the following set of linear equations:

$$\mathbf{N}_{s,1} = \left(\mathbf{I} - \mathbf{G}_s\right)^{-1} \tilde{\mathbf{R}}_{s,1} \tag{C.17}$$

where $\mathbf{N}_{s,1}$ is a vector of length $N_{reg}(x+1)$ containing the initial age-structure for animals of sex *s*,

$\tilde{\mathbf{R}}_{s,1}$ is the corresponding vector of recruits with elements:

$$\tilde{R}_{s,a,1}^{r} = \begin{cases} 0.5\,\lambda_0^r\, R_0 & \text{if } a = 0 \\ 0 & \text{if } 1 \leq a \leq x \end{cases} \tag{C.18}$$

where λ_0^r is the fraction of recruits allocated to region *r* in equilibrium, the value for which is solved for in order to satisfy equation (C.12), and

$\mathbf{G}_s$ is a square transition matrix with dimensions of length the same as $\mathbf{N}_{s,1}$, describing the mortality and movement pattern, the elements of which are obtained from the equations for the population update:

$$G_{s,p,q}^{r_p,r_q} = \begin{cases} X_{s,a_p}^{r_q,r_p} e^{-M} & \text{if } a_p = a_q - 1 \\ X_{s,a_p}^{r_q,r_p} e^{-M} & \text{if } a_p = a_q = x \\ 0 & \text{otherwise} \end{cases} \quad \text{(C.19)}$$

a_p is the age associated with row p,

a_q is the age associated with column q,

r_p is the region associated with row p, and

r_q is the region associated with column q.

C.2 Generating Age-composition Data

The observed catch-at-age proportions by region, sex and fleet are a multinomial sample of size $n_{s,t}^{f,r}$ from the true catch-at-age proportions. The proportion of the catch that is of age a during year t for fleet f and sex s in region r is:

$$p_{a,s,t}^{f,r} = \frac{\tilde{C}_{s,a,t}^{f,r}}{\sum_a \tilde{C}_{s,a,t}^{f,r}} \quad \text{(C.20)}$$

where:

$$\tilde{C}_{s,a,t}^{f,r} = I^{f,r} \sum_L S_L^f \Phi_{L,s,a} N_{s,a,t}^r e^{-0.5M} \quad \text{(C.21)}$$

and $I^{f,r}$ is an indicator equal to 1 if fleet f operates in region r, and zero otherwise.

The individual $n_{s,t}^{f,r}$'s by fleet and region are derived from a multinomial sample of the total annual age–composition with relative probability given by $C_t^{f,r}$ and sample size 100. As such, the ageing samples are (on average) proportionally allocated by fleet and region with respect to the annual catch.

C.3 Allocation of TAC by fleet and region during projection period

For each year of the projection period, the catches for each fleet and region are calculated using a multinomial allocation of the total TAC for that year (with sample size equal to the TAC). The expected proportions of the catch for each fleet/region are:

$$p_{C,t}^{f,r} = \frac{C_t^{f,r}}{\sum_{f'}\sum_{r'} C_t^{f',r'}} = \frac{\tilde{p}_{C,t}^{f,r}}{\sum_{f'}\sum_{r'} \tilde{p}_{C,t}^{f',r'}}$$
$$\tilde{p}_{C,t}^{f,r} = \frac{\xi^{f,r} + \psi^{f,r}\left(B_t^{f,r}\right)^{\varsigma^{f,r}}}{\sum_{f'}\sum_{r'}\left[\xi^{f',r'} + \psi^{f',r'}\left(B_t^{f',r'}\right)^{\varsigma^{f',r'}}\right]} \qquad \text{(C.22)}$$

where $\xi^{f,r}$, $\psi^{f,r}$, and $\varsigma^{f,r}$ are the parameters of the relationship between biomass distribution and catch allocation, and

$B_t^{f,r}$ is the vulnerable biomass in region *r* for fleet *f* in the middle of year *t*:

$$B_t^{f,r} = I^{f,r}\sum_s\sum_L w_{L,s} S_L^f \sum_a \Phi_{L,s,a} N_{s,a,t}^r e^{-0.5M} \qquad \text{(C.23)}$$

The values of the parameters $\xi^{f,r}$, $\psi^{f,r}$, and $\varsigma^{f,r}$ are determined by fitting the multinomial model in equation (C.22) to the (known) historical catch proportions by fleet and region.

Vita

Gavin Fay was born in the steel city of Sheffield, England in 1978. In 2000, Gavin earned a Bachelor of Science (with Honours) degree in Marine Biology from the University of Stirling. In 2004, Gavin earned a Master of Science from the University of Washington in Aquatic and Fishery Sciences. His thesis, titled "A Bayesian stochastic metapopulation model for Steller sea lions in Alaska" was supervised by Dr. André E. Punt. The delights associated with living in Seattle proved very strong, and in 2012 Gavin earned a Doctor of Philosophy at the University of Washington in Aquatic and Fishery Sciences, also under the supervision of Professor Punt. Gavin enjoys music and the outdoors, and loves exploring new places, particularly with Michelle, Kaia, and Atticus. From 2008-2011, Gavin and family lived in Hobart Tasmania, where Gavin worked as a fisheries modeller at CSIRO Marine and Atmospheric Research. Gavin now lives in New England, where he works as a postdoctoral fellow at NOAA's Northeast Fisheries Science Center.