

**Extending process and understanding for the
development of complex ecosystem models, with
application to the Chatham Rise Atlantis model.**

by

Vidette Louise McGregor, MSc

Submitted in fulfilment of the requirements for the Degree of
Doctor of Philosophy
Victoria University (2020)

Abstract

The Chatham Rise is a highly productive deep-sea ecosystem that supports numerous substantial commercial fisheries, and is therefore a likely candidate for an ecosystem based approach to fisheries management in New Zealand. This thesis describes model construction, calibration and validation, for the first end-to-end ecosystem model of the Chatham Rise, New Zealand. The work extends beyond what has previously been done for validating such models, and explores uncertainty analyses through bootstrapping the oceanographic variables, perturbing the model's initial conditions, and analysing species interaction effects, with the results further analysed with respect to known data gaps. This enables the inclusion of uncertainty in simulated scenarios using the Chatham Rise Atlantis model, thus providing an envelope of results with which to analyse and understand the likely responses of the Chatham Rise ecosystem. The model was designed with 24 dynamic polygons, 5 water column depth bins, 55 species functional groups, and used 12-hour timesteps. The transfer of energy was tracked throughout the system using nitrogen as the model's main currency. The model simulated the system from 1900–2015, preceded by a 35 year burn-in period. The model produced very similar biomass trajectories in response to historical fishing to corresponding fisheries stock assessment models for key fisheries species. Population dynamics and system interactions were considered realistic with respect to growth rates, mortality rates, diets and species group interactions. The model was found to be generally stable under perturbations to the initial conditions, with lower trophic level species groups having the most variability. The specification of the Spawning Stock Recruitment curve was explored, as it relates to the multi-species and ecosystem models within which it is now applied. Close attention needs to be given to population dynamics specific to multi-species interactions such as predation-release, in particular the Spawning Stock Recruitment curve. Potentially misleading dynamics under predation-release were identified, and the simple solution of applying a cap to recruitment when biomass exceeds virgin

levels was explored. The population dynamics of myctophids under fishing induced predation release were analysed with and without limiting recruitment to virgin levels. The effects were evident in several ecosystem indicators, suggesting unintentional mis-specification could lead to erroneous model results. It raises several questions around the specification of the Spawning Stock Recruitment relationship for multi-species models, and more generally, whether the concept of ‘virgin’ (or ‘unfished’) biomass should be reconsidered to reflect dynamic natural mortality and potentially changing unfished states. The model components that had knowledge gaps and were found to most likely to influence model results were the initial conditions, oceanographic variables, and the aggregate species groups ‘seabird’ and ‘cetacean other’. It is recommended that applications of the model, such as forecasting biomasses under various fishing regimes, should include alternatives that vary these components, and present appropriate levels of uncertainty in results. Initial conditions should be perturbed, with greater variability applied to species groups modelled as biomass-pools, and age-structured species groups that have little data available from the literature.

This thesis is dedicated to:

Sophia & Marcus

0.1 Acknowledgements

My children, Sophia and Marcus, for all the quality time playing games, kicking balls, baking, that kept me balanced. And for engaging in my dinner time conversations when they consisted of things like ‘my sharks all died today’, or ‘my ecosystem was taken over with giant crabs’. Also for engaging in my every-day statistics and mathematics lessons, for helping with the housework, and in general for being great kids!

Kush, our dearly departed cat, for agreeing to sit next to my laptop, not on it, while I was working.

Friends near and far for tolerating intermittent contact, especially Laura, Simon, Tracy and Jade.

All my family for being proud of my achievements, supporting me, and also tolerating intermittent contact sometimes. My brother, Cameron, for the text conversations when we hadn’t caught up properly for a while. Mum and Dad for raising me to believe hard work was worth it, and Grandad John for suggesting I find something I enjoy so it won’t feel like hard work (it still feels like hard work sometimes).

Ben for solo-patenting when I travelled, helping with the kids while I worked, and for the cooked meals.

My colleagues, particularly in the Population Modelling group, for bouncing ideas off and for some of the most entertaining conversations at morning tea. Other colleagues for lunch time games of soccer and touch rugby, that were much appreciated breaks.

Beth and Matt for supervising me as well as letting me find my own way. You’ve

both been an inspiration to me, and I've appreciated having such talented researchers to help guide me through this work.

Peter Horn (NIWA) for providing data inputs for the model, as well as filling in some of my knowledge gaps of biology and ecology. Thanks also for providing an excellent balance of scepticism and support. One of my favourite quotes was when you said 'Well I'll be blowed, I might have to accept this thing seems to be actually working!'

In relation to Chapter 2: Mark Hadfield for development of the ROMS model for oceanographic variables. Bec Gorton (CSIRO) for converting the ROMS variables into Atlantis model inputs. Ian Tuck for providing a comprehensive internal review of the manuscript. Cliff Law and Graham Rickard (NIWA) for help with ocean physics data and conversions. Matt Pinkerton for trophic level results from stable isotope analyses and visible band radiation at sea surface data. James Bell, Victoria University supervisor.

In relation to Chapter 3: NIWA for core funding under project FIFI1801 and supporting my PhD. Ian Doonan, Andy McKenzie, Kath Large (NIWA) and Malcolm Haddon (CSIRO) for discussing some of the ideas, Chris Francis (independent) for reviewing the initial draft. Ian Tuck, NIWA project leader. James Bell, Victoria University supervisor.

My supervisors, Dr. Matthew Dunn (NIWA) and Dr. Elizabeth Fulton (CSIRO), were both co-authors on the corresponding journal papers. They provided feedback on the analyses and proof-read drafts. Beth gave advice particularly earlier on in the research on Atlantis, and Matt reviewed the model dynamics for realism (in addition to Peter Horn (NIWA)), in particular around which species groups responded to oceanographic variability, and which were most influential or responsive when I perturbed the system. My Victoria University supervisor, Prof. James Bell reviewed

the ‘General Introduction’ and ‘Synthesising the process’ sections.

NIWA for supporting my PhD, and for funding under projects FIFI1501–FIFI1801.

Contents

0.1	Acknowledgements	4
1	General introduction	23
1.1	Model construction	26
1.2	Model calibration and fitting to history	29
1.3	Model sensitivity and skill	34
1.4	Scenarios for system understanding	36
1.5	Tool development	36
1.6	General methodology and approach	37
1.7	This thesis	39
2	From data compilation to model validation: A comprehensive analysis of a full deep-sea ecosystem model of the Chatham Rise	40
2.1	Abstract	40
2.2	Introduction	41
2.3	Methodological Approach	46
2.4	Model design	46
2.4.1	Model area	47
2.4.2	Time	48
2.4.3	Oceanography	49
2.4.4	Nutrients	50
2.4.5	Species groups	52
2.4.6	Predation	57
2.5	Calibration	60

2.6	Sensitivity analyses	62
2.6.1	Oceanography	62
2.6.2	Connectivity and influence	66
2.7	Fishing	70
2.7.1	Comparison with fisheries CPUE and Stock Assessment indices	72
2.8	Skill assessment	74
2.9	Bringing it together	76
2.10	Discussion	80
2.11	Conclusions	83
3	Spawning Stock Recruitment creates misleading dynamics under predation release in ecosystem and multi-species models	84
3.1	Abstract	84
3.2	Introduction	85
3.3	Simple population model with varying natural mortality	91
3.4	Fishing induced predation release	94
3.5	Ecological indicators	96
3.6	Discussion	104
4	Chaotic or stable? Exploring sensitivity to initial conditions of an end-to-end ecosystem model	109
4.1	Abstract	109
4.2	Introduction	110
4.3	Methods	112
4.3.1	Varying initial conditions	113
4.3.2	Characterising the system and its components	116
4.3.3	Modelling stability	120
4.4	Results	122
4.4.1	Variability from initial conditions	122
4.4.2	Characterisation	127
4.4.3	Additional natural mortality	129

4.4.4	Modelling stability	131
4.5	Discussion	144
5	Synthesising the process	149
6	References	163
7	Appendices	184
	Appendix A: Simulated biomass by species group from no-fishing model . .	184
7.0.1	Additional background:	184
	Appendix B: Size-at-age	187
7.0.2	Additional background:	187
	Appendix C: Proportion-at-age	190
7.0.3	Additional background:	190
	Appendix D: Observed vs estimated	193
7.0.4	Additional background:	193
	Appendix A: Ecosystem indicators - Kempton's Q	196
	Appendix B: Ecosystem indicators - Mean trophic level	197
	Appendix C: Ecosystem indicators - Biomass over catch	198
	Appendix D: Ecosystem indicators - Biomass pelagic over all	199
	Appendix E: Ecosystem indicators - Biomass trophic level 4 and higer over	
trophic level 3		200
	Appendix J: Biomass trajectories	202
	Appendix K: Parameters tuned during model calibration	208
	Appendix L: Experts consulted during model design and development . . .	208

List of Figures

1.1	The model process as applied for Atlantis models in the literature with approximate, qualitatively defined levels of development with respect to time. Earlier implementations of Atlantis focused more on the first 3 steps (framework, model development and calibration), while the last step (sensitivity analyses) has become more of a focus since 2016, with also a higher level of model skill assessments presented from this time.	31
2.1	Map of New Zealand with Chatham Rise marked, including 200 m, 500 m, and 1000 m isobaths.	45
2.2	Polygons as defined for CRAM with maximum depths for each polygon shown by colour (left) and depth layer bins (right).	49
2.3	Sea surface temperature (SST) ($^{\circ}$ Celsius) weekly averages for 1981–2017 with ROMS years 1996–2004 shaded blue (left) and mean SST by month (right) from the same data for 1981–2017 (black dashed line), with the subset from 1996–2004 (blue solid line), and additional historical SST data from 1961–1990, which were only available as monthly averages (orange solid line).	50
2.4	Nutrient cycle as modelled in Atlantis.	51
2.5	Summary of the proportion of prey groups in the diets of species functional groups (Tables 2.2 and 2.3) over model years 1900–2016 from the fished model where the proportion is by mg N consumed.	59

2.6	Biomass by trophic level with 95% confidence intervals from the 1900–2016 Chatham Rise Atlantis model simulation. The blue line is the fitted linear model to the median biomasses by trophic level, the slope which is in blue. The slopes of the linear models to fitted to the upper and lower 95% confidence interval limits are given in brackets.	62
2.7	Sea temperature (°C) from ROMS model outputs by day for each year 1996–2004 (dark blue line) and median sea temperature over all ROMS model years 1996–2004 (grey line).	64
2.8	Number of species groups for each simulation with one ROMS year repeated that went above (light blue) or below (midnight blue) the limits of bootstrapped ROMS simulations and for the Base Model where the years were repeated in order for the entire model simulation.	65
2.9	Keystoneness (y-axis) and relative overall effect (x-axis) for all age-structured species groups, with numbers giving keystone ranking (1 is the most influential using Equation 2.3). Colours indicate biomass proportion scaled by proportional change in biomass (Equation 2.4).	68
2.10	Responsiveness of age-structured species groups after 50 years of perturbation, as calculated in Equation 2.5.	69
2.11	Tonnes caught from Chatham Rise 1900–2014 for all species with top six species groups by total catch coloured separately.	70
2.12	Total tonnes caught by fishing fleet from the Chatham Rise 1900–2014. Descriptions for the fleet codes are in Table 2.7.	71

2.13	CRAM estimated spawning stock biomass (SSB) (black solid), stock assessment estimated SSB (red dot-dash), and CPUE (blue dash) where available for the hake (A), hoki (B), invertebrate scavengers (commercial) (primarily scampi) (C), ling (D), and orange roughy (E) . CPUE and stock assessment SSB were rescaled to match the mean of the CRAM estimated SSB. Inset boxplots show the range of values for the corresponding unscaled stock assessment SSB divided by the CRAM estimated SSB.	73
2.14	Skill assessment metrics MEF (A), RI (B) and Pearson’s correlation (C) for CRAM species groups that have trawl survey indices for abundance. Metric definitions in Equations 2.6–2.8. The black bars are the skill metrics with respect to single point estimates from the trawl survey. The orange bars are the skill metrics with respect to the trawl survey 95% confidence intervals. The grey horizontal lines in the MEF and RI figures mark the value for a perfect fit, which is 1 for both of these.	76
2.15	Keystoneness (x-axis) and responsiveness (y-axis) with numbers showing keystone ranking and colours how well each species group was informed and/or performed in the model (legend).	79
3.1	Beverton-Holt SSR curve with $h = 0.4$ (A, blue line) and $h = 0.8$ (B, green line), with B_0 and R_0 (black dot-dashed lines) and $20\%B_0$ and $R_{0.2}$ (grey dot-dashed lines).	88
3.2	Time-varying mortality (M) used for simple population model. The horizontal dashed line is the base mortality (base M) that would result in the population staying at B_0	92

3.3	Beverton-Holt SSR curves (A) and resulting biomass trajectories (B) from a simple population model using different steepness, h , ranging from 0.35–0.95 and time-varying mortality as in Figure 3.2. B_0 shown with vertical red dash line (A) and horizontal red dashed line (B); R_0 shown with horizontal red dashed line (A).	93
3.4	Beverton-Holt SSR curves with recruitment capped at R_0 (left) and resulting biomass trajectories (right) from a simple population model using different steepness, h , ranging from 0.35–0.95 and time-varying mortality as in Figure 3.2. B_0 shown with vertical red dash line (A) and horizontal red dashed line (B); R_0 shown with horizontal red dashed line (A).	93
3.5	Historical catches from 1900–2014 with top six species groups based on total catch coloured separately (McGregor et al., 2019b).	96
3.6	Hindcast biomass trajectory for the myctophids species functional group with different recruitment steepness (h) and no cap on recruitment (left) and with recruitment capped at R_0 (right).	96
3.7	Kempton’s Q calculated from Chatham Rise Atlantis model simulations with recruitment steepness set at 0.5 for myctophids, no cap on recruitment (left), recruitment capped at R_0 (right), and three catch scenarios: 1.) Zero catch; 2.) <i>Status quo</i> catch; 3.) Half catch, for the 2010–2016 hindcast period and 2016–2046 projection period. The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.	99

- 3.8 Mean trophic level of age-structured species groups calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (A, C), recruitment capped at R_0 (B, D), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (A, B) and h set at 0.9 for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (C, D). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small. 100
- 3.9 Biomass of age-structured species groups over catch calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (A, C), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (B, D) and h set at 0.9 for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small. 102
- 3.10 Biomass of pelagic fishes over biomass of all age-structured species groups from Chatham Rise Atlantis model simulations with recruitment steepness set at 0.5 for myctophids, no cap on recruitment (A), recruitment capped at R_0 (B), and three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2010–2016 hindcast period and 2016–2046 projection period. The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small. 103

3.11	Biomass ratio of trophic level 4 and higher over trophic level 3 from Chatham Rise Atlantis model simulations with recruitment steepness set at 0.5 for myctophids, no cap on recruitment (A), recruitment capped at R_0 (B), and three catch scenarios: 1.) Zero catch; 2.) <i>Status quo</i> catch; 3.) Half catch, for the 2010–2016 hindcast period and 2016–2046 projection period. The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.	103
4.1	Set 2 scalars used to perturb initial conditions, with scalars sampled from Normal distributions with $\mu = 0$ and $\sigma \in (0.025, 0.05, 0.1, 0.25)$ based on informance levels 1–4 respectively where 1: ‘Poorly specified’ (gold); 2: ‘Some data gaps and/or poor performance’ (magenta); 3: ‘Slight data gaps and/or poor performance’ (blue); 4: ‘No data gaps, performed well, abundance index available’ (green) (defined in McGregor et al. (2019b)).	115
4.2	Set 1 scalars used to perturb initial conditions for high keystone and biomass-pool species groups, with scalars sampled from the Normal distribution with $\mu = 0$ and $\sigma = 0.25$.	115
4.3	Illustration of primary (green arrows) and secondary (orange arrows) trophic connections for the species shaded blue.	118
4.4	Median (solid lines), and upper and lower quartiles (dot-dashed lines) for CVs of age-structured species groups (A) and biomass-pool species groups (B) from fished model runs (blue) and unfished model runs (orange). Density of CVs from 1970–2015 for age-structured species groups (C) and biomass-pool species groups (D).	124

4.5	Biomass trajectories from models with fishing included (blue lines) and no fishing (orange lines) for Ref Det (refractory detritus) (A) and hoki (B), with CVs from across the model runs by time from fished models (aqua asterisks) and unfished models (cerise asterisks) overlaid and using the right-hand axis.	125
4.6	Pearson’s correlation (A) and Spearman’s rank correlation (B) between initial condition scalars and changes in the perturbed model with respect to the base model for fished models (blue lines) fished species groups (solid line) and not-fished species groups (dashed line) and not fished models (orange lines) fished species groups (solid line) and not-fished species groups (dashed line).	126
4.7	Ecological indicators, mean trophic level of age-structured species groups (A), Kempton’s Q (B), and biomass ratio of pelagic fishes/all age-structured species groups (C) calculated from model simulations with fishing included (blue lines), and no fishing included (orange lines) from model years 1865–2015, which includes the burn-in period of 1865–1900.	128
4.8	Proportion of diet made up by top prey from the Chatham Rise Atlantis model (base) 1900–2015 model outputs.	129
4.9	Number of trophic level connections by species group for the Chatham Rise Atlantis model (base) 1900–2015 model outputs. A: Number of groups (frequency counts) by primary connections (green bars), secondary connections (blue bars), and tertiary connections (orange bars); B: Number of primary connections by species group (green bars), and number of secondary connections by species group (blue asterisks, and using right-hand axis).	129
4.10	Proportion of natural mortality (M) forced as additional mortality by species group from the Chatham Rise Atlantis model (base) 1900–2015 model outputs.	130

4.11	R^2 for GLMs fitted at each timestep to biomass CVs of all species groups that resulted from perturbing the initial conditions, using all model runs (A), only model runs with fishing (B), and only models without fishing (C), with bars coloured by explanatory variable. . . .	132
4.12	R^2 for GLMs fitted at each timestep to biomass CVs of biomass-pool (BP) species groups that resulted from perturbing the initial conditions, using all model runs (A), only model runs with fishing (B), and only models without fishing (C), with bars coloured by explanatory variable.	133
4.13	R^2 for GLMs fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using all model runs (A), only model runs with fishing (B), and only models without fishing (C), with bars coloured by explanatory variable.	134
4.14	Pearson's residuals for models fitted at each timestep to biomass CVs of all (ALL) species groups that resulted from perturbing the initial conditions, using all model runs, plotted against fitted values (A), TL (trophic level) (B), and ChaosAlt (C).	136
4.15	Pearson's residuals for models fitted at each timestep to biomass CVs of biomass-pool (BP) species groups that resulted from perturbing the initial conditions, using all model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), Number of primary trophic connections (D), and B_0 (virgin biomass) (E).	137
4.16	Pearson's residuals for models fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using all model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), and Number of primary trophic connections (D).	138

4.17	Pearson’s residuals for models fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using fished model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), and Number of primary trophic connections (D).	139
4.18	Pearson’s residuals for models fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using unfished model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), and Number of primary trophic connections (D), and informance (E). Informance levels 1–4 where 1: ‘Poorly specified’ (gold); 2: ‘Some data gaps and/or poor performance’ (magenta); 3: ‘Slight data gaps and/or poor performance’ (blue); 4: ‘No data gaps, performed well, abundance index available’ (green) (defined in McGregor et al. (2019b)). Informance level ‘1’ did not feature in the results as these data were dropped due to ‘NA’ values for other explanatory variables.	140
4.19	GLM effects from the final GLMs fitted to CVs from 1910–2015 for interaction term ChaosAlt:TL for ALL species groups (A), BP only species groups (B), and AS only species groups (C). ChaosAlt ‘A’ and ‘B’ perturbed all initial conditions by the same scalar for each run; ChaosAlt ‘C’ and ‘D’ perturbed initial conditions by uncertainty; ChaosAlt ‘E’ and ‘F’ perturbed initial conditions by keystoneity; ChaosAlt ‘A’, ‘C’, ‘E’ did not include fishing; ChaosAlt ‘B’, ‘D’, ‘F’ included fishing. Shading indicates the additional CV expected for each value of the interaction, with the darkest shading in each plot corresponding to the Max. CV (%) given in the top-right corner of the plot.	142

4.20	GLM effects from the final GLMs fitted to CVs from 1910–2015 for interaction term PrimCons: B_0 for BP only species groups (A), ChaosAlt:PrimCons for AS only species groups (B). ChaosAlt ‘A’ and ‘B’ perturbed all initial conditions by the same scalar for each run; ChaosAlt ‘C’ and ‘D’ perturbed initial conditions by uncertainty; ChaosAlt ‘E’ and ‘F’ perturbed initial conditions by keystone-ness; ChaosAlt ‘A’, ‘C’, ‘E’ did not include fishing; ChaosAlt ‘B’, ‘D’, ‘F’ included fishing. Shading indicates the additional CV expected for each value of the interaction, with the darkest shading in each plot corresponding to the Max. CV (%) given in the top-right corner of the plot.	143
5.1	The model process as applied for the Chatham Rise Atlantis model. The process begins in the centre of the spiral, then progresses towards the outer spiral, but can at any stage return to an earlier part of the process, and no part of the process is separate.	154

List of Tables

2.1	Sources of data for Oxygen, Nitrates, Ammonium and Silica. WOA, World Ocean Atlas; WOCE, World Ocean Circulation Experiment; NIWA, National Institute of Water and Atmospheric Research.	52
2.2	List of functional vertebrate groups for CRAM. Name is the species group name which is the same as the main species name for single-species groups but without punctuation. Lifespan is the assumed maximum number of years an individual in that group may live. Ben, benthic; Dem, demersal; invert, invertivore; pisc, piscivore.	53
2.3	List of functional invertebrate groups for CRAM. Name is the species group name which is the same as the species name for single-species groups. Description includes main species. Lifespan is the maximum number of years an individual in that group may live. Those groups with no value for lifespan are modelled as biomass pools and hence do not have a lifespan defined as this is only relevant when modelling numbers. Zoo, zooplankton; Invert comm, commercial invertebrates; herb, herbivore; scav, scavenger.	54
2.4	List of functional phytoplankton and algae groups for CRAM. Name is the species group name which is the same as the main species name for single-species groups. Description includes main species.	54
2.5	List of functional bacteria and detritus groups for CRAM. Name is the species group name which is the same as the main species name for single-species groups. Description includes main species.	54

2.6	Biological parameters assumed for age-structured species groups. VB, von Bertalanffy; M, instantaneous natural mortality rate; h, steepness value for the Beverton-Holt stock recruitment relationship. Length-weight parameters are: $W = aL^b$ (weight W in g, length L in cm). Where Reference is ‘Trawl db’ some data have been derived from the NIWA trawl survey database (see Mackay (2000)). Species group matches ‘Name’ in Tables 2.2 and 2.3 and are without punctuation.	56
2.7	Fishing fleets defined for Chatham Rise Atlantis model. Number of species groups is the number of species groups that have been caught by each fishing fleet; total catch is the total tonnes caught by each fishing fleet from 1900–2014.	71
3.1	Ecological indicators evaluated for hindcast simulations and catch scenario projections, with key references for each indicator.	97
3.2	Sensitivity (indicated by ✓) of ecological indicators to steepness and recruitment cap in the hindcast and/or catch scenario projections.	98
4.1	Key ecosystem indicators evaluated for responses to perturbing the initial conditions.	116
4.2	Explanatory variables offered to explain effects of perturbing the initial conditions, and whether these were defined for age-structured species groups or biomass-pool species groups, or all species groups. BP= biomass-pool species groups; ALL= all species groups; AS = age-structured species groups.	122
4.3	ChaosAlt definitions for perturbing the initial conditions, and including fishing in the model or not. ChaosAlt was offered as an explanatory variable to the GLMs.	122

4.4	Explanatory variables selected and corresponding r^2 values for GLMs fitted to ALL (all species groups) model CVs, BP (biomass-pool species groups) only model CVs, and AS (age-structured species groups) only model CVs, using model outputs from 1900–2015, with fished and unfished versions for AS. ChaosAlt=the set of runs, grouped by method for perturbing initial conditions and whether fishing was included or not; TL=trophic level; PrimCons=number of primary trophic connections; B_0 =virgin biomass; Inf=informance;	135
7.1	Parameters tuned and/or edited during model calibration.	208

Chapter 1

General introduction

The Chatham Rise, a submarine ridge running eastwards about 1000 km from New Zealand, is perhaps New Zealand's most productive fishing ground (Ministry for Primary Industries, 2019). It supports many commercially important fisheries, including several with MSC (Marine Stewardship Council) certification (Deepwater Group, 2018). The subtropical front (STF), a relatively broad permanent feature where warmer, more saline, and nutrient poor subtropical water from the north meets nutrient rich subantarctic water from the south, extends up the east coast of South Island, and then eastwards along Chatham Rise (Heath, 1985; Uddstrom and Oien, 1999). The demersal fish assemblage on Chatham Rise has the highest fish species richness in New Zealand waters (Leathwick et al. 2006). The range of habitats and depths, and the influence of the STF, are expected to provide a wide variety of foraging opportunities for demersal and pelagic organisms (Dunn et al., 2010; Stevens and Dunn, 2011). There is a wealth of data and knowledge on the Chatham Rise Ecosystem, from trawl (O'Driscoll et al., 2011) and acoustic surveys (Escobar-Flores et al., 2019), diet studies (e.g. Dunn et al., 2010; Stevens et al., 2011b), fisheries data (Ministry for Primary Industries, 2019), and oceanography (e.g. Rickard et al., 2016; Law et al., 2017; Hadfield et al., 2007). The body responsible for fisheries management in New Zealand, Fisheries New Zealand, is seeking to move away from single species management towards a more ecosystem approach, both to fulfil Fisheries Act obligations and Marine Stewardship Council (MSC) expectations (Ministry for Pri-

mary Industries, 2019; Marine Stewardship Council, 2014). The Chatham Rise is a good place to start for establishing an ecosystem approach to fisheries management in New Zealand, due to its commercial importance, and extensive data and knowledge currently available.

The need for ecosystem approaches to fisheries management is well established globally (Borja et al., 2016; Collie et al., 2016; Patrick and Link, 2015; Hilborn, 2011; Pikitch et al., 2004; Link, 2002). How to achieve it is still developing, but there is agreement that ecosystem models have the potential to be very helpful (Collie et al., 2016; Plagányi, 2007; Pikitch et al., 2004; Botsford et al., 1997). Marine ecosystems are complex, and they are challenging to study. In the somewhat famous words of Prof. John Shepherd (University of Southampton), ‘Managing fisheries is hard: it’s like managing a forest, in which the trees are invisible and keep moving around’. It follows that managing whole ecosystems is even harder because there are even more things moving around that we can’t see, which are interacting, and we can’t see that either. This is where models can help, as tools to explore complex systems, and to understand the implications of what we do.

While there is a lot of enthusiasm for ecosystem models and the roles they could and are playing in understanding ecosystem-wide effects, there are also sceptics of ecosystem models (Rose, 2012). There are many aspects of ecosystem models that provoke scepticism: they are complex; there are more parameters and processes than we have information and data to inform; model validation is yet to do justice to the complexities of the models; and there has been little progress in addressing uncertainty. There has been a lot of progress on the structure of ecosystem models (Broekhuizen et al., 1994; Baretta-Bekker, 1996; Walters et al., 1997; Murray and Parslow, 1997; Shin and Cury, 1999; Fulton, 2001; Christensen and Walters, 2004; Fulton et al., 2004a), and it is natural that the development of the models has preceded the development of validating them. Structurally, ecosystem models did not have to start from scratch; they incorporated theory already developed for other areas

of modelling, such as population dynamics, fisheries dynamics, and bio-physical processes. Fisheries models, which are well established for supporting the management of fisheries, have a solid base of theory and many decades of application. Much of fisheries theory still used today was developed by scientists in the 1950s such as Beverton and Holt (1957), Ricker (1954) and Schaefer (1957), with a greater emphasis on statistical methods by the end of last century (Kimura, 1980; Chang, 1982; Cerrato, 1990; Hilborn et al., 1992). Fisheries models are generally statistically optimised to fit data available, and integrated models will fit to multiple datasets simultaneously. With advancements in computing power, Bayesian approaches to fisheries models became available, with sample spaces explored using MCMC (Markov Chain Monte Carlo) algorithms. Bayesian approaches are statistical methods that allow for the incorporation of prior information with present data. MCMC algorithms are designed to explore sample space of a probability distribution. Models developed for fisheries are still improving and advancing —there are new methods like Vector Autoregressive Spatio-Temporal (VAST) that model fisheries dynamics spatially (Thorson, 2019), and multi-space and multi-species approaches are now available (Plagányi et al., 2014; Doonan et al., 2016). As fisheries models include more ecosystem dynamics, and ecosystem models become more stochastic, both branches of modelling can learn a lot from each other, and it seems likely the gap between them will lessen.

Atlantis is an end-to-end ecosystem modelling approach that can be used to create an environment in which different scenarios can be played out to test for different results and learn how a system may be reacting to changes within it (Audzijonyte et al., 2017b). Reviewed as one of the best modelling frameworks for exploring ‘what-if’ type questions (Plagányi, 2007), it includes the ability to compare social, conservation, and economic outcomes. With sufficient data, this modelling approach can be extremely useful for management strategy evaluation (Plagányi, 2007), and has been applied to multiple marine systems (from single bays to millions of square kilometres) in Australia, the United States, Europe, and South Africa (Savina et al., 2005; Fulton et al., 2007; Link et al., 2010; Ainsworth et al., 2015; Smith et al., 2015;

Sturludottir et al., 2018; Ortega-Cisneros et al., 2017). Atlantis is a deterministic simulation model such that for a given parameter set and model specification, the model outputs are identical. Atlantis models are too complex to statistically fit to observations, although subsets of key parameters can be estimated using statistical methods outside of the model. Analysing and understanding the model dynamics and potential weaknesses is essential before the model can be used to learn about the system.

This thesis largely focuses on end-to-end models using Atlantis, developed by Dr. Beth Fulton (CSIRO), with specific application to the Chatham Rise, New Zealand. The thesis follows the path of model construction, calibration, sensitivity and skill assessments, scenarios for system understanding, and tool development. The work extends beyond what has currently been done with respect to exploring and understanding model dynamics. Some components of the thesis, in particular Chapter 3, adds to our understanding of fundamental dynamics relating to ecosystem and multi-species models, affecting a broad range of models.

1.1 Model construction

Constructing an ecosystem model largely consists of specifying its structure, defining functional forms, and estimating parameters and initial conditions. The structure is defined early on in model development, and is influenced by the intended or likely use of the model, key components that are likely to be important for system dynamics, and availability of data and information. The structure relates to the physical structure of the model (boxes and depth layers in an Atlantis model), temporal (size of time-steps and timeframe covered), biological (specification of species functional groups, and age-structure within these), and harvest (specification of fishing fleets, and characteristics such as gear types, target species and catchabilities associated with them) (Audzijonyte et al., 2017a).

The spatial extent of the model should be sensible with respect to available data and information. For example, the Guam Atlantis model was restricted by depth to the shallower waters based on availability of data and information only extending patchily into the deeper waters (Weijerman et al., 2014). Within the defined model area, many systems have key areas of habitat that influence the spatial structure, or physical features that may influence currents and upwelling. For example, Link et al. (2010) based the physical structure on previously defined sub-regions with distinct geological, bathymetric, physical, chemical and biological properties, and at a finer scale within these to allow for hydrodynamic and biotic processes for the North-east United States Atlantis model. Similarly, the Guam Atlantis model was defined based on benthos, oceanographic, bathymetry and substrate, with consideration also of marine protected areas and survey data (Weijerman et al., 2014).

Timesteps in Atlantis models are often 12 hours to account for differences in behaviour between night and day (Ainsworth et al., 2011; Weijerman et al., 2014), although many of the reports and papers describing Atlantis models do not seem to make reference to the size of the time steps. It is possible to set it smaller than 12 hours which may be important if modelling a tidal area and wanting to focus on dynamics of the inter-tidal zone. The historical period of Atlantis models varies with Weijerman et al. (2014) starting at 1985, Ainsworth et al. (2011) extending back to 1980, Porobic et al. (2019) and Sturludottir et al. (2018) around 60 years, to some over 100 years (Ortega-Cisneros et al., 2017). The choice of historical period seems to depend largely on the lifespan of animals in the system, and the availability of data to inform the state of the historical system.

Species functional groups consist of biomass-pool species groups which are generally used to capture most invertebrate species, and age-structured species groups, which are typically used to model vertebrates species and key exploited invertebrate species such as cephalopods (Audzijonyte et al., 2017a). The specification of species

functional groups at the lower trophic level is similar between Atlantis models. All must have three types of detritus (labile detritus, refractory detritus and carrion), and most models have two types of bacteria (sediment and pelagic), and two types of phytoplankton (large and small). These functional groups are similar to the group structure used in typical NPZD (Nutrient-Phytoplankton-Zooplankton-Detritus) models (e.g. (Fasham et al., 1990) uses phytoplankton, zooplankton, bacteria, detritus and nutrients, and (Kishi et al., 2007) uses small and large phytoplankton, small, large and predatory zooplankton, and nutrients including nitrogen and silicate). Beyond these common species functional groups, Atlantis models diverge a little, although most will have a few zooplankton groups either split by size or function, several benthic invertebrate groups such as filter feeders, carnivores and grazers, and some primary producer groups such as macroalgae, microphytobenthos, seagrass and dinoflagellates. The higher trophic levels vary more, for example Weijerman et al. (2014) includes turtles as they are a tourism attraction as well as trophically connected; lobsters are a focus group in Porobic et al. (2019) due to the artisanal lobster fishery; the Icelandic model has four species groups of whales, with the potential to use this model to explore scenarios regarding whaling (Sturludottir et al., 2018); and the south eastern Australia Atlantis model has seven species groups of sharks, and shark fisheries in operation (Fulton et al., 2007). The number of species functional groups ranges from around 30 (Porobic et al., 2019; Ortega-Cisneros et al., 2017) to more than 90 (Ainsworth et al., 2015).

Earlier accounts of model construction for complex end-to-end models using Atlantis tended to reside in technical reports rather than published journal articles due to the enormity of technical details that are not relevant to a very wide audience. Four key examples of technical reports are Link et al., 2010; Ainsworth et al., 2011; Weijerman et al., 2014 and Ainsworth et al., 2015, all of which are around 200 pages long and include around 100 tables and figures. These earlier technical reports include many large tables with values used to inform the model, figures that could be condensed, and many details and equations that relate to Atlantis, and are not spe-

cific to the implementation being presented. The publication of an Atlantis Manual (Audzijonyte et al., 2017a) has helped with the latter as it now suffices to reference the manual, and only mention relevant specifics such as which functional forms were selected, and values of estimated parameters. More recent implementations of Atlantis models have tended to present succinct accounts of the model construction, such as Porobic et al., 2019; Sturludottir et al., 2018 and Ortega-Cisneros et al. (2017), with details such as parameter values available as supplementary information. Similarly, the presentation of the Chatham Rise Atlantis model construction and design in this thesis has been condensed into 14 pages, seven tables, and six figures. The concise and succinct account of model construction and development presented here communicates the key aspects of this process, at a level of detail that should retain interest to readers not necessarily reproducing the model, with the GitHub (a software development platform) repository (McGregor, 2018) providing the additional level of detail required for reproducibility. The newer minimalistic presentation of Atlantis models may downplay the enormity of developing such complex models for those not familiar with what is involved, but it also encourages the focus in the literature to progress from model development to other aspects such as model exploration and validation (Figure 1.1).

1.2 Model calibration and fitting to history

Model calibration is required when statistically fitting to data is not possible due to the complexity of the model. Statistical techniques used to find optimal parameter sets that best fit available data are only an option if there exists a globally optimal solution, and even in single species fisheries models, these can at times struggle with the optimiser producing solutions that are locally optimal, but not global. Complex end-to-end ecosystem models such as Atlantis may have multiple solutions, and it would be very difficult to write an optimising algorithm that would be effective over such a complex space. Hence, rather than seeking optimal solutions for Atlantis models, realistic solutions are sought through the process of calibration. The goals

of calibration are stated in Audzijonyte et al. (2017a) as:

1. Prevent all species present in the model from going extinct (unless they do so in the time series being fit to)
2. Have age structured groups grow such that size-at-age is reasonable (within 20% of initial conditions, typically)
3. For species with historical data are available, have the model recreate observations of abundance from surveys or assessments
4. For species with no historical data, the model should yield reasonable time series of abundance (especially under perturbation such as environmental forcing or fishing pressure)
5. Capturing observed spatial distributions

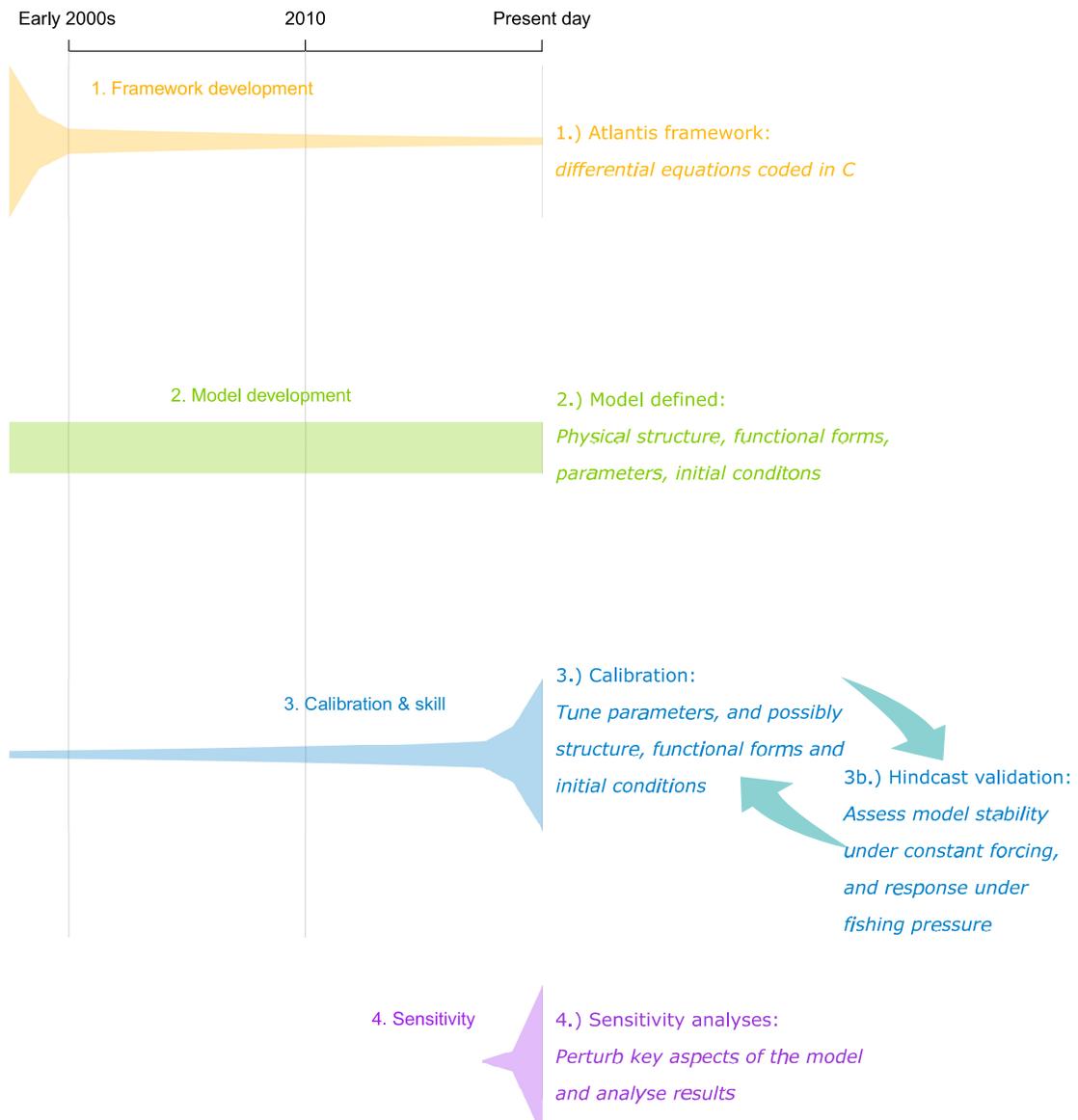


Figure 1.1: The model process as applied for Atlantis models in the literature with approximate, qualitatively defined levels of development with respect to time. Earlier implementations of Atlantis focused more on the first 3 steps (framework, model development and calibration), while the last step (sensitivity analyses) has become more of a focus since 2016, with also a higher level of model skill assessments presented from this time.

There has been general agreement on the calibration goals of model stability and minimal extinctions, from before Atlantis was Atlantis (Fulton et al., 2004b,a), to some of its earlier applications in the United States (Horne et al., 2010; Link et al., 2010; Ainsworth et al., 2011) and more recently developed models (Sturludottir et al., 2018; Porobic et al., 2019). The ability to replicate historic biomass trends (and/or catch trends if sufficient effort data are available) is often calibrated for once fishing is included in the model, with model outputs compared to relevant assessments of surveys (Link et al., 2010; Horne et al., 2010; Ainsworth et al., 2011; Weijerman et al., 2014; Ortega-Cisneros et al., 2017). Stable growth has been mentioned for some models, analysed using timeseries of weights for each age-class from the model output (Ainsworth et al., 2015; Weijerman et al., 2014). While this method for analysing growth does assess stability of growth, it does not easily assess realism of growth, especially if any burn-in period (period at the beginning of the model time that are not considered part of the model period due to the settling of initial dynamics that may be unrealistic) has been discarded. Analysing size-at-age would make the realism of growth more transparent and could be evaluated with respect to von Bertalanffy growth parameters (Von Bertalanffy, 1938) if these are available. Interestingly, realistic trophic connections or realised diets have only had sporadic mention for Atlantis models. These parameters were analysed for the southeast Australia Atlantis model (Fulton et al., 2007), the Iceland Atlantis model (Sturludottir et al., 2018), and the Chile Atlantis model (Porobic et al., 2019). Realised diets were mentioned in Weijerman et al. (2014) but did not feature in the results of analyses presented. Other models gave almost no mention to trophic connections for calibrating or assessing the model (Brand et al., 2007; Horne et al., 2010; Link et al., 2010; Ainsworth et al., 2011, 2015; Ortega-Cisneros et al., 2017). Brand et al. (2007) did suggest future work could calibrate the model to observed time series of diets such as from trawl surveys. Realised natural mortality rates only feature through the quest for stable densities by age-class. Similarly to growth rates, the stable density diagnostics do not assess realistic mortality rates if presented as time series because it is very difficult to assess the presence let alone the steepness of an exponential

decay curve. Weijerman et al. (2014) stated exponential decay curves were evident based on time series trajectories of densities for each age-class, although it was not clear in this presentation format.

The most commonly tuned parameters during calibration have been those relating to growth, the feeding functional parameters (clearance and handling time, although the latter is often referred to as growth, but is parameterised in Atlantis as 1/handling time) and non-predation natural mortality (Fulton et al., 2004b,a; Brand et al., 2007; Fulton et al., 2007; Link et al., 2010; Ainsworth et al., 2011, 2015). Recruitment parameters have also been key tuning parameters in at least two Atlantis models (Ainsworth et al., 2011, 2015; Ortega-Cisneros et al., 2017). It is unclear whether the realised Spawning Stock Recruitment (SSR) curves (these determine density-dependent dynamics relating expected recruitment to existing spawning stock biomass) were assessed in response to changing recruitment parameters. The SSR relationship is important in terms of the dynamics it produces when the model is not in equilibrium (such as when exploring scenarios). Tuning recruitment parameters during calibration without the examining the resulting SSR relationship could easily bring in unrealistic dynamics when the model is used to explore scenarios. Mentioning which parameters were tuned during calibration seems to have become a thing of the past, as the more recent Atlantis models published don't seem to mention them (Sturludottir et al., 2018; Ortega-Cisneros et al., 2017; Porobic et al., 2019; McGregor et al., 2019b).

Kaplan and Marshall (2016) define standards for end-to-end ecosystem models such as Atlantis:

- 1.) Biological functional groups should persist throughout the model run
- 2.) The model should achieve equilibrium under fixed environmental forcing with no fishing
- 3.) The hindcast biomass trends should be compared to survey and stock assess-

ment time-series

- 4.) Qualitative model comparisons to surveys and stock assessments
- 5.) The model should reproduce temporal dynamics of abundant species
- 6.) Most functional groups should qualitatively match expected productivity
- 7.) Natural mortality (M) as a function of age should be realistic
- 8.) Age and length structure with respect to age should match data
- 9.) Diet composition of the functional groups should match diet data

These standards seem to cover the breadth of model calibration goals from the literature, with the appropriate addition of realistic natural mortality, matching age and length structure to data, and matching expected productivity. Productivity is defined as FMSY (the fishing pressure expected to give the long-run maximum sustainable yield), but could be extended to encompass inspection of the Spawning Stock Recruitment curve, as these are generally defined for stock assessments (e.g. Bull et al., 2012, Fournier et al., 1998 and Methot Jr and Wetzel, 2013).

1.3 Model sensitivity and skill

Nearly 10 years ago, Fulton (2010) noted the handling of uncertainty of end-to-end ecosystem models to be both necessary and challenging, with the large numbers of parameters and feedback loops key reasons for the challenge. There have been recent advancements in assessing the skill of ecosystem models (Olsen et al., 2016), but so far only using rather simple metrics which do not do justice to the complexity of an end-to-end ecosystem model. Sensitivity analyses are still not generally carried out for ecosystem models (Hansen et al., 2019), even with advancements in computing power and data science. Skill assessments on their own assess a model's accuracy, but they do not assess its precision. A model output that is exact but wrong would not be useful, but neither would one where the results are so imprecise such that no

clear result can be ascertained. If accounting for uncertainty in the model yields no clear results, this is important to present. It may still be possible to subset out the uncertainty or tighten uncertainty bounds until results become coherent, but again, it would be important to report on this process.

Over the last few years, Atlantis models have been presented with greater efforts at assessing the skill of the model, often following Olsen et al. (2016), and accounting for at least some of the model's uncertainty. For example, Ortega-Cisneros et al. (2017) included a sensitivity analysis on zooplankton growth, and a skill assessment on biomass, catch and spatial distribution of the species for which there were sufficient observation data; Sturludottir et al. (2018) included a skill assessment using three metrics applied to biomass trajectories of five species functional groups and landings of 12 species functional groups, and included a sensitivity analysis on the feeding functional response parameters; and Hansen et al. (2019) perturbed recruitment, mortality, prey consumption and growth by +/- 25% for nine biomass-dominating key species. There has also been some progress on addressing uncertainties in Object-oriented Simulator of Marine Ecosystems (OSMOSE) models, an alternative ecosystem modelling framework based on agent-based representations (Shin and Cury, 1999). In OSMOSE, stochasticity is incorporated through the spatial distribution of species, and with the stochastic feeding algorithm that samples prey for a given predator-prey interaction based on spatial overlap, gape-size, and feeding preferences (Grüss et al., 2016a,b). There are still parameters that are tuned through calibration, and an initial state for the model needs to be defined. It is not apparent that sensitivities to initial conditions have been explored for any end-to-end ecosystem models in the literature.

1.4 Scenarios for system understanding

Tuning parameters for the calibration of the model in equilibrium runs the risk of inappropriately specifying the model for evaluating flow-on effects when we perturb the system. In a complex model, there are many ways realistic biomass trends and dynamics of individual species groups can be produced, and not all of these will be realistic. Perturbing the model in different ways and analysing the flow-on effects as they play out is one way of gaining understanding of the dynamics of the system, its parts as they relate to each other, and whether predicted responses are realistic/plausible, but there appears to have been little formal discussion on this in the literature (though it is a widely known approach amongst experienced modellers). The sensitivity analyses on zooplankton and phytobenthos growth in the model of Ortega-Cisneros et al. (2017) were analysed with respect to the model's sensitivity to these parameters, but the results were not analysed with respect to what could be learnt about the behaviour of the model. Whether the species groups that responded most strongly to these sensitivity analyses were sensible based on knowledge of the system, was not addressed. The sensitivity analyses on the Norwegian Sea Atlantis model did relate the results to dynamics of the system (Hansen et al., 2019), although the scope of the study was fairly limited. All perturbations were by 25%, thus no lessons could be learnt from analysing variability, as there was no randomness to the study design. Further, it was concluded by Hansen et al. (2019) that top predators responded more to perturbations to their parameters rather than to parameters of other species groups, but only two top predators in the model were perturbed, which is not a sufficiently large sample-size for drawing this conclusion.

1.5 Tool development

R is a software development environment specifically aimed at data and statistical analyses (R Core Team, 2013). There are some R packages for use during Atlantis model calibration (Audzijonyte et al., 2017a), which have been developed

to assist modellers with limited coding ability to develop Atlantis models without being limited to Excel and text editors to more manually explore model outputs. Given the multi-dimensional nature of complex ecosystem models, and the huge range of possible ways to explore the model, the more flexibility in analysing the model outputs and inputs the better. As an experienced R user from my work as a stock assessment scientist and my studies in Statistics, as well as training in software development, I developed fit-for-purpose R scripts throughout this work. The code I've developed is extensive, with approximately 1600 scripts. The R code specifically used for the analyses as presented in this thesis are available on github (<https://github.com/mcgregorv>). Additional scripts written for understanding particular aspects during calibration are also on github (<https://github.com/mcgregorv/AtlantisRscripts>), although there are many of them, many of an ad-hoc nature. Additional scripts were also written during development to solve a particular problem. For example, if a specific predator-prey interaction was failing to occur, I would have written *fprintf* statements in the Atlantis source code, then R scripts to analyse the resulting outputs, until I resolved the issue. Some of the more general scripts I've written have been used by other Atlantis modellers, especially those around predator-prey dynamics such as gape-size effects, and spatial and temporal overlap.

1.6 General methodology and approach

The Chatham Rise was chosen as the focus area of this research due to the large amount of data and information available, its many commercially important fisheries, and the appetite for an ecosystem approach to fisheries management of this system.

Studying the system was not the focus of this research, and Peter Horn (Fisheries Biologist, NIWA) gathered much of the data on the system. The focus of this research

was to develop the Atlantis model to best represent the Chatham Rise ecosystem for the purpose of supporting an ecosystem approach for fisheries management, to understand the dynamics, strengths and limitations of the model, and to test, explore and understand it. This work takes end-to-end ecosystem model testing, exploration and validation to levels much beyond those presented in the literature to date.

Model development and calibration consumed much of the first two years of this research. During the initial model development, I learnt out how the model would use the inputs and what was required, and my colleague, Peter Horn, produced data from the literature or in some cases from his own experience, that could be used either directly as model inputs or to inform them. Some inputs went almost directly into the model, such as catch histories, or spatial distributions, with only minor changes to formatting and conversion of units. Others required more modelling effort, such as diets and recruitment. Diets were available from the predators perspective (what proportion of the diet is made up of a given prey), but enter the model from the perspective of the prey (what proportion of the prey may be eaten by a given predator). Further, as there are many dynamics that impact on the realised diets of predators in Atlantis, (including spatial and temporal overlap, gape-sizes and growth rates), obtaining appropriate realised diets went a long way into model calibration. Recruitment parameters for the Spawning Stock Recruitment curve were initially set using stock assessment parameters where they were available, or were estimated using C++ Algorithmic Stock Assessment Laboratory (CASAL) models (Bull et al., 2012) where there were sufficient data to do so, or approximated based on similar species. Adjustments to these values were initially required due to the quasi-version of Beverton-Holt recruitment in Atlantis. Investigations into this quasi-functional form lead to 1.) re-specifying the relationship in the Atlantis source code, and 2.) exploring the effects of the Spawning Stock Recruitment relationship in a multispecies context (the subject of Chapter 3).

1.7 This thesis

Chapters 2–4 are self-contained journal articles, two of which are published, and the final one is submitted. These are very concise and succinct accounts of the work undertaken for this thesis. Each chapter has its own Introduction section, while the General Introduction provides more general context for the work, and discusses aspects of the literature as they form the context for the overall thesis. The final chapter, ‘Synthesising the process’ brings together lessons learnt from all chapters, and discusses how the process applied to this work might look with the added benefit of hindsight. I give recommendations for the development, exploration and validation of complex ecosystem models into the future.

Chapter 2

From data compilation to model validation: A comprehensive analysis of a full deep-sea ecosystem model of the Chatham Rise

The following chapter is sourced from the open-source peer-reviewed article: McGregor V.L., Horn P.L., Fulton E.A., Dunn M.R. 2019. From data compilation to model validation: a comprehensive analysis of a full deep-sea ecosystem model of the Chatham rise. PeerJ 7:e6517 DOI 10.7717/peerj.6517.

PeerJ is an open access publisher, and all content is published under a CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

2.1 Abstract

The Chatham Rise is a highly productive deep-sea ecosystem that supports numerous substantial commercial fisheries, and is a likely candidate for an ecosystem based

approach to fisheries management in New Zealand. I present the first end-to-end ecosystem model of the Chatham Rise, which is also to be best of my knowledge, the first end-to-end ecosystem model of any deep-sea ecosystem. I describe the process of data compilation through to model validation and analyse the importance of knowledge gaps with respect to model dynamics and results. The model produces very similar results to fisheries stock assessment models for key fisheries species, and the population dynamics and system interactions are realistic. Confidence intervals based on bootstrapping oceanographic variables are produced. The model components that have knowledge gaps and are most likely to influence model results were oceanographic variables, and the aggregate species groups ‘seabird’ and ‘cetacean other’. I recommend applications of the model, such as forecasting biomasses under various fishing regimes, include alternatives that vary these components.

2.2 Introduction

The goal of incorporating a holistic approach to understanding the system-wide repercussions of how we manage our marine resources is admirable and ambitious (Long et al., 2015; Link and Browman, 2017). Ecosystem Based Management (EBM) requires a range of tools, often including ecosystem models (Smith et al., 2017; Stecken and Failler, 2016). Within ecosystems there are many processes at play, and the models developed to support EBM vary in scope and complexity (Plagányi, 2007; Fulton, 2010; Collie et al., 2016). End-to-end ecosystem models that can deal with bottom-up and top-down system controls have become popular for exploring scenarios involving human induced impacts including fishing and climate change (Rose, 2012).

The body responsible for fisheries management in New Zealand, Fisheries New Zealand, is seeking to move away from single species management towards a more

ecosystem approach, both to fulfil Fisheries Act obligations and Marine Stewardship Council (MSC) expectations (Ministry for Primary Industries, 2019; Marine Stewardship Council, 2014). The Chatham Rise is the location of several nationally important MSC certified fisheries (Deepwater Group, 2018), and a growing understanding of trophic interactions exists there (Stevens et al., 2011b; Dunn et al., 2009).

Chatham Rise is a submarine ridge running eastwards for about 1000 km from the east coast of South Island, New Zealand, rising up from depths of about 3000 m, to about 50 m at the western end, and sea level around the Chatham Islands at the eastern end (Figure 2.1). The subtropical front (STF), a relatively broad permanent feature where warmer, more saline, and nutrient poor subtropical water from the north meets nutrient rich subantarctic water from the south, extends up the east coast of South Island, and then eastwards along Chatham Rise (Heath, 1985; Uddstrom and Oien, 1999). The demersal fish assemblage on Chatham Rise has the highest fish species richness in New Zealand waters (Leathwick et al. 2006). The range of habitats and depths, and the influence of the STF, are expected to provide a wide variety of foraging opportunities for demersal and pelagic organisms (Dunn et al., 2010; Stevens and Dunn, 2011).

The Chatham Rise is perhaps New Zealand's most productive fishing ground. It supports substantial commercial fisheries for finfish and invertebrates, with notable examples being: trawl fisheries for hoki (*Macruronus novaezelandiae*), orange roughy (*Hoplostethus atlanticus*), hake (*Merluccius australis*), and black and smooth oreos (*Allocyttus niger*, *Pseudocyttus maculatus*); a longline fishery for ling (*Genypterus blacodes*); and a potting fishery for rock lobster (*Jasus edwardsii*) (Ministry for Primary Industries, 2019).

Analyses of trawl survey series and commercial fishery catch rates have shown that marked variations over time have occurred in the relative abundance of some common species on Chatham Rise, e.g., hoki, hake, orange roughy, scampi (*Metanephrops*

challengeri), and rock lobster (Maunder and Starr, 1995; Dunn et al., 2008; Stevens et al., 2017). Some factors driving these fluctuations have been identified (i.e., high exploitation levels, variation in recruitment) (Ministry for Primary Industries, 2019), but there will certainly be other physical and biological factors that will influence animal behaviour and survivability, resulting in changes to the ecosystem. A knowledge of how particular biological and ecological changes could affect the abundance and distribution of species will usefully inform the management of those species.

In an ecosystem, nothing exists independently. When assessing biological risks, it is difficult to conceptualise risk to the whole system. A system-level model within which different scenarios can be explored is an extremely valuable tool for gaining conceptual understanding of economic and biological risks for a whole system, as well as for individual parts.

Atlantis is an end-to-end ecosystem modelling approach that can be used to create an environment in which different scenarios can be played out to test for different results and learn how a system may be reacting to changes within it (Audzijonyte et al., 2017b). Reviewed as one of the best modelling frameworks for exploring ‘what-if’ type questions (Plagányi, 2007), it includes the ability to compare social, conservation, and economic outcomes. With sufficient data, this modelling approach can be extremely useful for management strategy evaluation (Plagányi, 2007), and has been applied to multiple marine systems (from single bays to millions of square kilometres) in Australia, the United States, Europe, and South Africa (Savina et al., 2005; Fulton et al., 2007; Link et al., 2010; Ainsworth et al., 2015; Smith et al., 2015; Sturludottir et al., 2018; Ortega-Cisneros et al., 2017). Atlantis is a deterministic simulation model such that for a given parameter set and model specification, the model outputs are identical. Atlantis models are too complex to statistically fit to observations, although subsets of key parameters can be estimated using statistical methods outside of the model. Analysing and understanding the model dynamics and potential weaknesses is essential before the model can be used to learn about

the system.

In this chapter, I describe the first end-to-end ecosystem model for the Chatham Rise, New Zealand (hereafter referred to as CRAM). I present analyses of the model, comparing its state and dynamics to current knowledge. I identify and assess the likely influence of current knowledge gaps and uncertainties.

In developing such models, knowledge gaps become evident, and we are provided with the opportunity to analyse the importance of these gaps, thus guiding direction of future research. The model was assessed for single species dynamics and inter-species connectivity. I conducted a skill assessment on species groups for which we have surveys capable of indexing abundance, and compared biomass trends as the model responded to historical fishing for species groups that have stock assessments or reliable catch per unit effort (CPUE) indices. I simulated changes in biomass for each species group and analysed responses throughout the system. This latter part formed the basis for analysing influence and importance of knowledge gaps, and where a species group performed poorly in the skill assessment it often highlighted a knowledge gap.

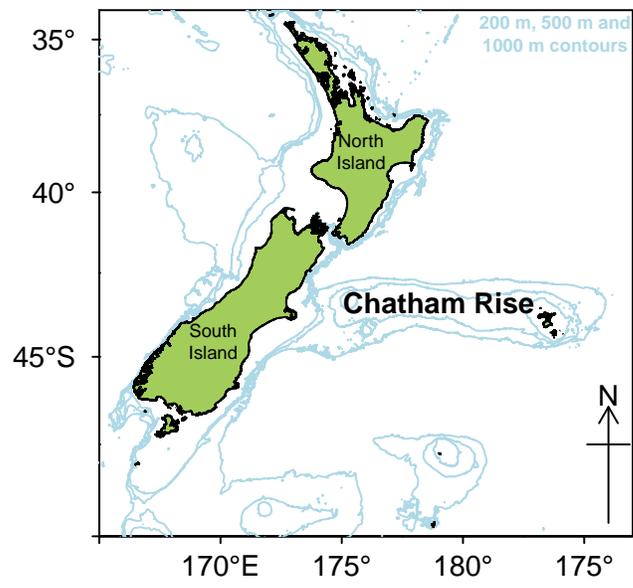


Figure 2.1: Map of New Zealand with Chatham Rise marked, including 200 m, 500 m, and 1000 m isobaths.

2.3 Methodological Approach

The process of developing this model was not linear, but rather iterative and incremental. There were five main stages to the development, each of which was re-visited until I was satisfied with the performance of the model and our understanding of its dynamics. The main stages can be summarised as:

- 1.) Data and model inputs were collated and defined.
- 2.) The base historical model was calibrated without fishing such that this model had stable biomass trajectories over the 1900–2016 model period, realistic diets, growth rates, natural mortalities.
- 3.) Sensitivity analyses were carried out with respect to oceanographic variables and simulations aimed at understanding connectivity and influence between the species functional groups.
- 4.) Fishing was included in the model using forced catch removals.
- 5.) Skill assessment and comparisons to abundance indices and biomass estimates were carried out.

Sections 2.4–2.8 cover each of these five main stages, followed by Section 2.9: ‘Bringing it together’, which discusses some of the implications of the models’ performance, dynamics and data gaps.

2.4 Model design

An Atlantis model simulates the ecosystem through time, calculating each new state based on the previous state and the events of the current timestep. This section describes the physical, biological, ecological, and fishing components of the Chatham Rise Atlantis Model. Further details on Atlantis can be found in the Atlantis user manual (Audzijonyte et al., 2017a).

2.4.1 Model area

The Chatham Rise Atlantis model area comprises waters from the shore-line around Chatham Islands (but excluding estuaries on the islands) to depths of 1300 m along the Chatham Rise, New Zealand (Figure 2.2). The western boundary of the area is defined as the 400 m contour on the western edge of the Mernoo Gap, a trough that separates the Chatham Rise from the coastal shelf off the mid east coast of South Island. This area boundary includes the core trawl survey strata of depths 200–800 metres and the additional depth down to 1300 metres covered by some additional strata added to the trawl survey in more recent years (Stevens et al., 2017).

An Atlantis model requires the modelled region to be split into polygons and depth layers. Each polygon/depth layer is referred to as a cell. The intention of the splits is to capture important aspects of the region but at a simplified level such that modelling the region over many years becomes possible. If I was modelling a smaller temporal scale, I may have considered a finer spatial scale. The polygons within the modelled area are referred to as dynamic polygons, and these are surrounded by non-dynamic polygons which define the boundary conditions for the modelled domain.

Several investigations of fish communities or fish species richness indicated that the division of the Chatham Rise into polygons for Atlantis modelling should occur primarily based on depth categories, with the northern and southern slopes separated (owing to the different water masses and fish communities to the north and south of the STF), and with some longitudinal differentiation as well. Species communities were found to group in adjacent depth-defined strata, but with differences between depths on the northern and southern Rise, as well as some longitudinal differentiation (Tuck et al., 2009).

A large amount of data on the abundance and distribution of demersal fish and invertebrate species has been collected from the series of trawl surveys of depths 200–800 m on Chatham Rise in January annually from 1992 to 2014 (Livingston

et al., 2002, Stevens et al., 2017). Some of the more recent surveys in the series also included strata to depths of 1300 m (Stevens et al., 2017). The survey area was stratified by depth, latitude, and longitude. It was logical, therefore, to base the Atlantis model polygon boundaries on the trawl survey strata boundaries. This area definition is also helpful for informing the model spatially based on trawl surveys. Consequently, the model area was divided into 24 dynamic polygons based on bottom depth bins (< 200 m, 200–400 m, 400–600 m, 600–800 m, 800–1300 m), with bins deeper than 400 m separated into northern and southern Rise polygons, and with longitudinal separation (where trawl survey strata allowed) aimed at producing western, central, and eastern polygons. The dynamic polygon area is surrounded by 6 additional non-dynamic polygons which allows for the exchange of water, nutrients and biota into and out of the dynamic model domain. The final configuration of the dynamic and non-dynamic polygons is shown in Figure 2.2.

All model polygons are further divided into water column depth layers, with depth bins defined to align with trawl survey strata, ranging from one layer in some near-shore polygons to five layers for the deepest polygons. Depth layers are also defined in Figure 2.2. Each box also contains one epibenthic (consisting of epibenthic invertebrates and habitat species functional groups) and one sediment layer.

2.4.2 Time

The model was run with a 35 year burn-in period (1865–1900) followed by a 115 year modelled period (1900–2015). The burn-in period allows for the model to adjust from potentially unstable initial conditions due to uncertainty of some of the parameters and age distributions for the age resolved groups, to a state that is more stable. A 35 year period was chosen as it covered initial fluctuations of most functional groups in the model. All results presented here are from the modelled period 1900–2015. The model used 12 hour timesteps to allow for changes in temperature, light and feeding patterns between night and day.

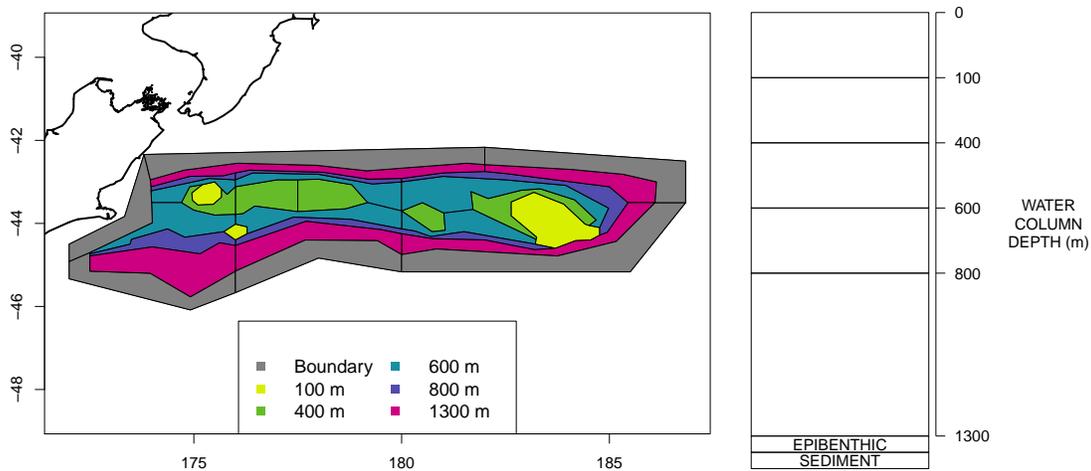


Figure 2.2: Polygons as defined for CRAM with maximum depths for each polygon shown by colour (left) and depth layer bins (right).

2.4.3 Oceanography

Salinity, temperature and water exchange between cells were forced in the Atlantis model using outputs from a ROMS (Regional Oceanographic Modelling System) model (Hadfield et al., 2007) that covered years 1996–2004. Water currents across each box face cause the horizontal movement of nutrients (such as ammonia and nitrate) available to primary producers. The speed and direction of currents influence the spatial distribution of plankton groups. Water temperatures influence biological processes such as respiration (Hoegh-Guldberg and Bruno, 2010). Based on sea surface temperatures (SST), the ROMS years (1996–2004) look to be fairly representative of those properties from 1961–2017 (Figure 2.3). The base model presented here repeated the available ROMS variables as a nine-year cycle. Averaging the ROMS variables was not sensible due to the water exchange between cells, as these change every 12-hour timestep in strength and direction, and averaging them could easily result in implausible physical dynamics. I ran sensitivities varying the order of ROMS years or repeating one ROMS year to help understand the effects of inter-annual oceanographic variability on this model.

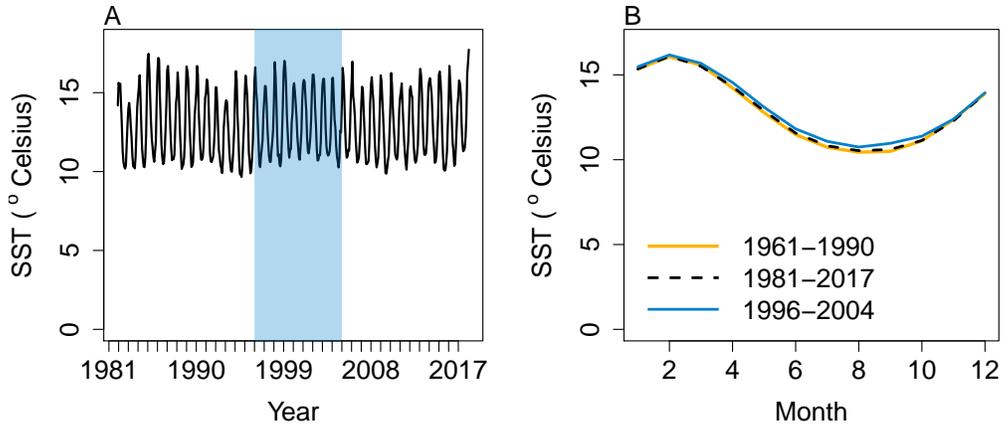


Figure 2.3: Sea surface temperature (SST) ($^{\circ}$ Celsius) weekly averages for 1981–2017 with ROMS years 1996–2004 shaded blue (left) and mean SST by month (right) from the same data for 1981–2017 (black dashed line), with the subset from 1996–2004 (blue solid line), and additional historical SST data from 1961–1990, which were only available as monthly averages (orange solid line).

2.4.4 Nutrients

Atlantis models use nitrogen, an important and often limiting nutrient in marine systems (Moore et al., 2013), to track the transfer of energy throughout the system. While carbon is also important (and is still tracked in Atlantis models), it is not limiting in marine systems, which is why nitrogen is the currency used in Atlantis. The nitrogen cycle can be seen in Figure 2.4. When biomass pools are tracked in the model, they are done so in mg N m^{-3} . When a fish (for example) eats another fish, it is nitrogen that is transferred up the food chain, with some nitrogen going to detritus and carrion, thus providing nitrogen to micro-organisms and filter feeders to fuel the cycle over again.

Nutrient data

Oxygen (O_2), nitrates (NO_3), ammonium (NH_4^+) and silica (SiO_2) were simulated in the model, and required spatially defined initial conditions (values for each cell in the model domain). Table 2.1 has a summary of the data sources for these nutrients.

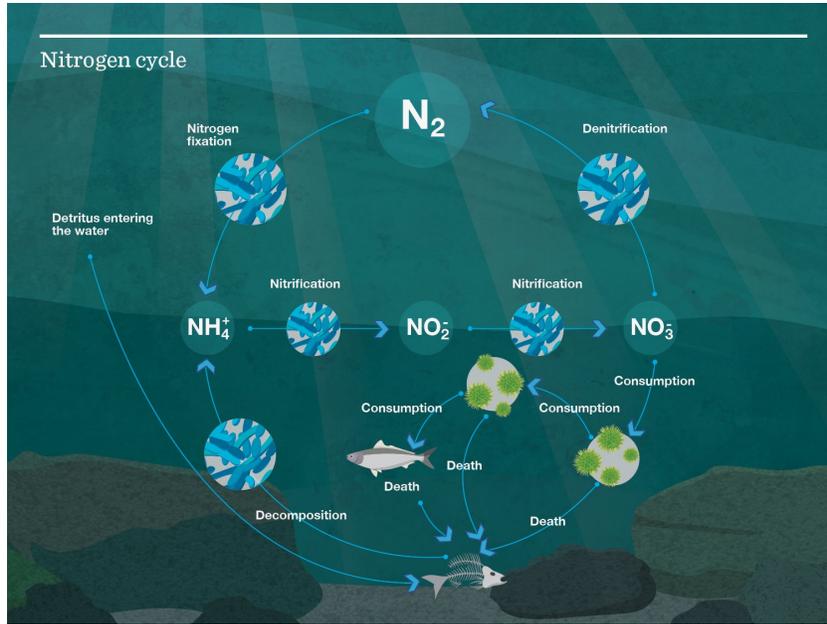


Figure 2.4: Nutrient cycle as modelled in Atlantis.

I used values from the World Ocean Atlas (WOA) for initial conditions for nitrate values down to 500 m, oxygen down to the full model depth of 1300 m, and silica down to 1300 m. The WOA contains objectively analysed climatological fields of *in situ* oxygen, temperature, salinity, and some nutrients (Locarnini et al., 2013; Zweng et al., 2013; Garcia et al., 2013a,b). $NO_3^- \mu mol/m^3$ were converted to $mg N/m^3$ by multiplying by 14 as the molecular mass of nitrogen is $14 g/mol$.

World Ocean Circulation Experiment (WOCE) (Deutsches Ozeanographisches Datenzentrum, 2006) data were used for nitrates at depths greater than 500 m, which were not covered by WOA. WOCE data were also used to compare values for oxygen, to inform initial conditions for silica, and to compare with salinity, temperature and chlorophyll *a*.

Ammonium values were available from NIWA (National Institute of Water and Atmospheric Research) oceanographic surveys, but only down to 50 m. This limited depth range was not too concerning as ammonium is a small component of the nitrogen budget.

Table 2.1: Sources of data for Oxygen, Nitrates, Ammonium and Silica. WOA, World Ocean Atlas; WOCE, World Ocean Circulation Experiment; NIWA, National Institute of Water and Atmospheric Research.

Variable	Source	Depth	Latitude	Longitude
Oxygen	WOA	1300 m	42–47 S	172 E –170 W
Oxygen	WOCE	1300 m	42.5 S	180 E
Nitrate	WOA	500 m	42–47 S	172 E –170 W
Nitrate	WOCE	1300 m	42.5 S	180 E
Silica	WOCE	1300 m	42.5 S	180 E
Ammonium	NIWA survey	0–50 m	43–46 S	172 E –180 E

2.4.5 Species groups

CRAM uses 55 functional groups to model the biological processes. Of these 55 groups, 15 vertebrates, and one invertebrate comprised single species; all other groups comprised two or more species. The main component species of the groups are shown in Tables 2.2 –2.5. Functional groups were defined based on broadly similar form, habitat, and diet, and also considered which species or groups of species may be important to investigate on their own. All vertebrate groups and five invertebrate groups were modelled with age-structure using up to 10 age-classes and varying number of years per age-class, depending on the longevity of the primary species in the group. Within each age-class, the model simulated numbers of individuals and the average weight (mg N) of individuals within each age class. Weights were split into structural (S_N) and reserve (R_N) components following the definition in Broekhuizen et al. (1994) where reserve weight is the part that can be used during periods of starvation, which includes flesh, fat, reproductive components and other soft tissue. Primary producers and remaining invertebrate groups were modelled as biomass pools (mg N m^{-3}) with no age-structure.

Table 2.2: List of functional vertebrate groups for CRAM. Name is the species group name which is the same as the main species name for single-species groups but without punctuation. Lifespan is the assumed maximum number of years an individual in that group may live. Ben, benthic; Dem, demersal; invert, invertivore; pisc, piscivore.

Name	Main species	Lifespan (years)
Baleen whales	Southern right whales (<i>Eubalaena australis</i>)	80
Basketwork eel	Basketwork eels (<i>Diastobranchus capensis</i>)	30
Baxters dogfish	Baxter's dogfish (<i>Etmopterus baxteri</i>)	50
Ben fish deep	Four-rayed rattail (<i>Coryphaenoides subserrulatus</i>)	20
Ben fish shal	Oblique banded rattail (<i>Coelorinchus aspercephalus</i>)	10
Black oreo	Black oreo (<i>Allocyttus niger</i>)	120
Bollons rattail	Bollons' rattail (<i>Caelorinchus bollonsi</i>)	20
Cetacean other	Primarily sperm & pilot whales & dolphins	30
Dem fish pisc	Giant stargazer (<i>Kathetostoma giganteum</i>)	20
Elasmobranch invert	Primarily skates & dogfish	20
Elasmobranch pisc	Primarily semi-pelagic sharks	50
Epiben fish deep	Spiky oreo (<i>Neocyttus rhomboidalis</i>)	100
Epiben fish shal	Common roughy (<i>Hoplostethus atlanticus</i>)	10
Ghost shark	Dark ghost shark (<i>Hydrolagus novaezealandiae</i>)	20
Hake	Hake (<i>Merlucciidae</i>)	30
Hoki	Hoki (<i>Macruronus novaezealandiae</i>)	20
Javelinfish	Javelinfish (<i>Coelorinchus australis</i>)	10
Ling	Ling (<i>Genypterus blacodes</i>)	30
Lookdown dory	Lookdown dory (<i>Cyttus traversi</i>)	30
Mackerels	Slender jack mackerel (<i>Trachurus murphyi</i>)	30
Orange roughy	Orange roughy (<i>Hoplostethus atlanticus</i>)	120
Pelagic fish lge	Southern bluefin tuna (<i>Thunnus thynnus</i>)	20
Pelagic fish med	Barracouta (<i>Thyrsites atun</i>)	10
Pelagic fish sml	Myctophids (<i>Myctophidae</i>)	4
Pinniped	NZ fur seal (<i>Arctocephalus forsteri</i>)	20
Reef fish	Blue cod (<i>Parapercis colias</i>)	20
Seabird	Seabirds & shorebirds	20
Seaperch	Seaperch (<i>Helicolenus</i> spp.)	50
Shovelnosed dogfish	Shovelnosed dogfish (<i>Deania calcea</i>)	40
Smooth oreo	Smooth oreo (<i>Pseudocyttus maculatus</i>)	100
Spiny dogfish	Spiny dogfish (<i>Squalus acanthias</i>)	30
Warehou	Silver, white & blue warehou	20

Table 2.3: List of functional invertebrate groups for CRAM. Name is the species group name which is the same as the species name for single-species groups. Description includes main species. Lifespan is the maximum number of years an individual in that group may live. Those groups with no value for lifespan are modelled as biomass pools and hence do not have a lifespan defined as this is only relevant when modelling numbers. Zoo, zooplankton; Invert comm, commercial invertebrates; herb, herbivore; scav, scavenger.

Name	Description	Lifespan (years)
Arrow squid	Arrow squid	2
Benthic Carniv	Benthic carnivores (0.9–90 mm)	
Carniv Zoo	Planktonic animals (size 2–20 cm)	
Cephalopod other	Squid & octopus	2
Deposit Feeder	Detritivores and benthic grazers (0.9–200 mm)	
DinoFlag	Dinoflagellates	
Filter Other	Non-commercial benthic filter feeders (size 0.9–200 mm)	
Gelat Zoo	Salps, ctenophores, jellyfish	
Invert comm herb	Paua & kina (size 0.9–140 mm)	10
Invert comm scav	Primarily scampi & crabs (size 0.9–500 mm)	14
Meiobenth	Benthic organisms (size 0.1–1 mm)	
MesoZoo	Planktonic animals (size 0.2–20 mm)	
MicroZoo	Heterotrophic plankton (size 20–200 μm)	
Rock lobster	Rock lobster	12

Table 2.4: List of functional phytoplankton and algae groups for CRAM. Name is the species group name which is the same as the main species name for single-species groups. Description includes main species.

Name	Description
Diatoms	Diatoms (large phytoplankton)
Macroalgae	Macroalgae (0.9 –2000 mm)
Microphytobenthos	Unicellular benthic algae (2–200 μm)
Pico-phytoplankton	Small phytoplankton (0.2–9 μm)

Table 2.5: List of functional bacteria and detritus groups for CRAM. Name is the species group name which is the same as the main species name for single-species groups. Description includes main species.

Name	Description
Carrion	Dead and decaying flesh
Labile detritus	Organic matter that decomposes at a fast rate
Pelagic bacteria	Pelagic bacteria
Refractory detritus	Organic matter that decomposes at a slow rate
Sediment bacteria	Sediment bacteria

Initial conditions and biological parameters for species groups

Initial biomasses for each species group were estimated using a single species stochastic stock assessment model, CASAL (Bull et al., 2012). Biomass estimates for the entire Chatham Rise were derived by using known biological parameters and a catch history to project back from an absolute abundance estimate in 2003. Values of relative abundance were available for most species groups from trawl surveys conducted annually from 1992 to 2014 (see O’Driscoll et al. (2011)). For each survey, these abundance estimates were converted to absolute values using trawl catchability quotients (specific to each group) derived by Peter Horn’s (NIWA) expert opinion, as fisheries scientists with experience dating back more than 30 years. Estimated absolute abundance for each group in 2003 (the midpoint of the survey series) was taken as the mean from all the survey estimates. For each species group, the initial biomass estimate was distributed across polygons in proportion to the survey series estimates (i.e., the mean proportion of total biomass by polygon over the survey series). The distribution of biomass by depth layer in each polygon was derived using Peter Horn’s (NIWA) expert opinion. Where there was no available catch history (e.g. seabirds), or no useful estimates of relative abundance from the trawl surveys (e.g. rock lobster), initial biomasses (and their distribution by model polygon) were estimated using Peter Horn’s (NIWA) expert opinion. For age-structured groups, initial biomass estimates were assigned to age-classes using estimates of instantaneous natural mortality (M). Initial average weights at age were calculated using Von Bertalanffy growth and length-weight conversion parameters. Values used for these parameters are in Table 2.6. Weights at age were split into reserve and structural components using ratio $R_N : S_N = 2.5 : 1$. This allows for an individual’s body mass to decrease by approximately 70% before starving, which is within the 60–80% range suggested by Broekhuizen et al. (1994).

All age-structured groups were modelled with Beverton-Holt recruitment, the steepness (h) values for which are in Table 2.6. These values are not ever well known, and scenarios explored using this model should consider sensitivities for these.

Table 2.6: Biological parameters assumed for age-structured species groups. VB, von Bertalanffy; M, instantaneous natural mortality rate; h, steepness value for the Beverton-Holt stock recruitment relationship. Length-weight parameters are: $W = aL^b$ (weight W in g, length L in cm). Where Reference is ‘Trawl db’ some data have been derived from the NIWA trawl survey database (see Mackay (2000)). Species group matches ‘Name’ in Tables 2.2 and 2.3 and are without punctuation.

Species group	VB Growth			Length-weight		M	h	Reference
	Linf (cm)	K	T_0	a	b			
Arrow squid	35	2.4	0	2.90E-02	3	4.6	0.8	Ministry for Primary Industries (2019)
Baleen whales						0.01	0.5	
Basketwork eel	47.3	0.283	-1.294	2.35E-03	3.25	0.19	0.8	Trawl db
Baxters dogfish	64.4	0.06	-2.97	5.95E-03	3.068	0.08	0.3	Irvine et al. (2006a)
Ben fish deep	36	0.3	-1.1	7.28E-03	2.632	0.2	0.8	Stevens et al. (2010), Trawl db
Ben fish shal	38	0.3	-1.1	2.35E-03	3.25	0.2	0.8	Stevens et al. (2010), Trawl db
Black oreo	37	0.1	-2	7.80E-03	3.27	0.044	0.75	Ministry for Primary Industries (2019)
Bollons rattail	47.3	0.283	-1.294	2.35E-03	3.25	0.19	0.8	Stevens et al. (2010)
Cephalopod other	45	2.4	0	2.90E-02	3	4.6	0.8	
Cetacean other						0.033	0.5	
Dem fish pisc	69.8	0.17	-0.53	1.50E-02	3.01	0.19	0.8	Sutton (1999), Ministry for Primary Industries (2019)
Elasmobranch invert	150.5	0.095	-1.06	2.68E-02	2.933	0.135	0.3	Ministry for Primary Industries (2019)
Elasmobranch pisc	84.7	0.1065	-4.56	1.50E-03	3.334	0.09	0.3	Irvine et al. (2006b)
Epiben fish deep	35.3	0.07	-0.5	2.83E-02	2.9322	0.05	0.75	Stewart and Smith (1994), Trawl db
Epiben fish shal	24	0.18	-0.3	2.65E-02	2.9126	0.2	0.8	Trawl db
Ghost shark	97	0.09	-1.17	2.02E-03	3.274	0.35	0.3	Ministry for Primary Industries (2019)
Hake	95.9	0.279	0.05	2.00E-03	3.288	0.19	0.8	Horn (2013)
Hoki	100.8	0.164	-2.16	4.79E-03	2.89	0.275	0.75	McKenzie (2016), Ministry for Primary Industries (2019)
Invert comm herb	155	0.15	0	3.00E-05	3.303	0.15	0.8	Breen et al. (2003)
Invert comm scav	50	0.25	0	3.73E-04	3.145	0.2	0.8	Tuck (2016)
Javelinfish	51.2	0.216	-1.618	1.38E-03	3.13	0.35	0.8	Stevens et al. (2010)
Ling	135.2	0.105	-0.72	1.07E-03	3.336	0.14	0.84	McGregor (2015)
Lookdown dory	50	0.075	-1	2.35E-02	2.97	0.15	0.8	Stewart and Smith (1994), Ministry for Primary Industries (2019)
Mackerels	74.25	0.111	-0.811	2.38E-02	2.7671	0.3	0.7	Cubillos et al. (1998), Kochkin (1994)

Orange roughy	37.2	0.065	-0.5	9.21E-02	2.71	0.045	0.75	Ministry for Primary Industries (2019)
Pelagic fish lge	182	0.205	0	1.88E-02	3.0078	0.2	0.8	Fournier et al. (1990), Ministry for Primary Industries (2019)
Pelagic fish med	85.2	0.298	-0.45	7.40E-03	2.94	0.3	0.7	Horn (2002), Ministry for Primary Industries (2019)
Pelagic fish sml	7	0.8	0	1.30E-02	2.81	1.58	0.7	Young et al. (1988), Trawl db
Pinniped							0.07	0.5
Reef fish	51.7	0.087	-1.7	1.91E-02	2.9818	0.14	0.8	Ministry for Primary Industries (2019)
Rock lobster	85	0.15	0	4.16E-03	2.935	0.12	0.8	Ministry for Primary Industries (2019)
Seabird							0.11	0.5
Seaperch	45.6	0.08	-0.8	7.77E-03	3.22	0.07	0.8	Paul and Horn (2009), Ministry for Primary Industries (2019)
Shovelnosed dogfish	106.4	0.106	-0.384	1.58E-03	3.192	0.13	0.3	Clarke et al. (2002), Trawl db
Smooth oreo	46	0.07	-1.5	3.05E-02	2.885	0.063	0.75	Ministry for Primary Industries (2019)
Spiny dogfish	104.8	0.093	-3.17	1.30E-03	3.2639	0.2	0.3	Hanchet (1986), Beentjes and Stevenson (2009)
Warehou	53.1	0.37	-0.88	8.28E-03	3.214	0.25	0.8	Horn and Sutton (1996), Ministry for Primary Industries (2019)

2.4.6 Predation

Simulated predation was a four step process that occurred within each cell and at each timestep. From the predator’s perspective the steps modelled can be summarised as: 1.) Am I allowed to eat it? (specified using a prey-availability matrix) 2.) Is it in the same place at the same time as me? 3.) Does it fit in my mouth? 4.) How much can I eat? Full details are in the Atlantis User’s Guide (Audzijonyte et al., 2017a). Step 4 uses a feeding functional response, of which there are 12 options currently available in Atlantis. I have applied the Holling Type II functional response to all age-structured species groups in this model, thus influencing the amount of prey consumed by prey abundance, and the predators search rate and handling time.

Diets of each species group were summarised in categories Algae, Bacteria, Bird, Cetacea, Coelenterate, Crustacean, Detritus, Echinoderm, Elasmobranch, Microzooplankton, Mollusc, Phytoplankton, Polychaete, Teleost, and Tunicate similar to that done in the diet study of Stevens et al. (2011a) (Figure 2.5). While this summary misses the temporal, spatial, age and size components of the predator-prey interactions, it is useful to check overall diets. For example, warehou and smooth oreos eat mostly salps (tunicates) as expected; Baxter's dogfish eat mostly fish, crustaceans, molluscs and tunicates as expected; and invertebrate herbivores (kina and paua) eat mostly algae, although they should also eat some phytoplankton, which they do but it is lost in the detail.

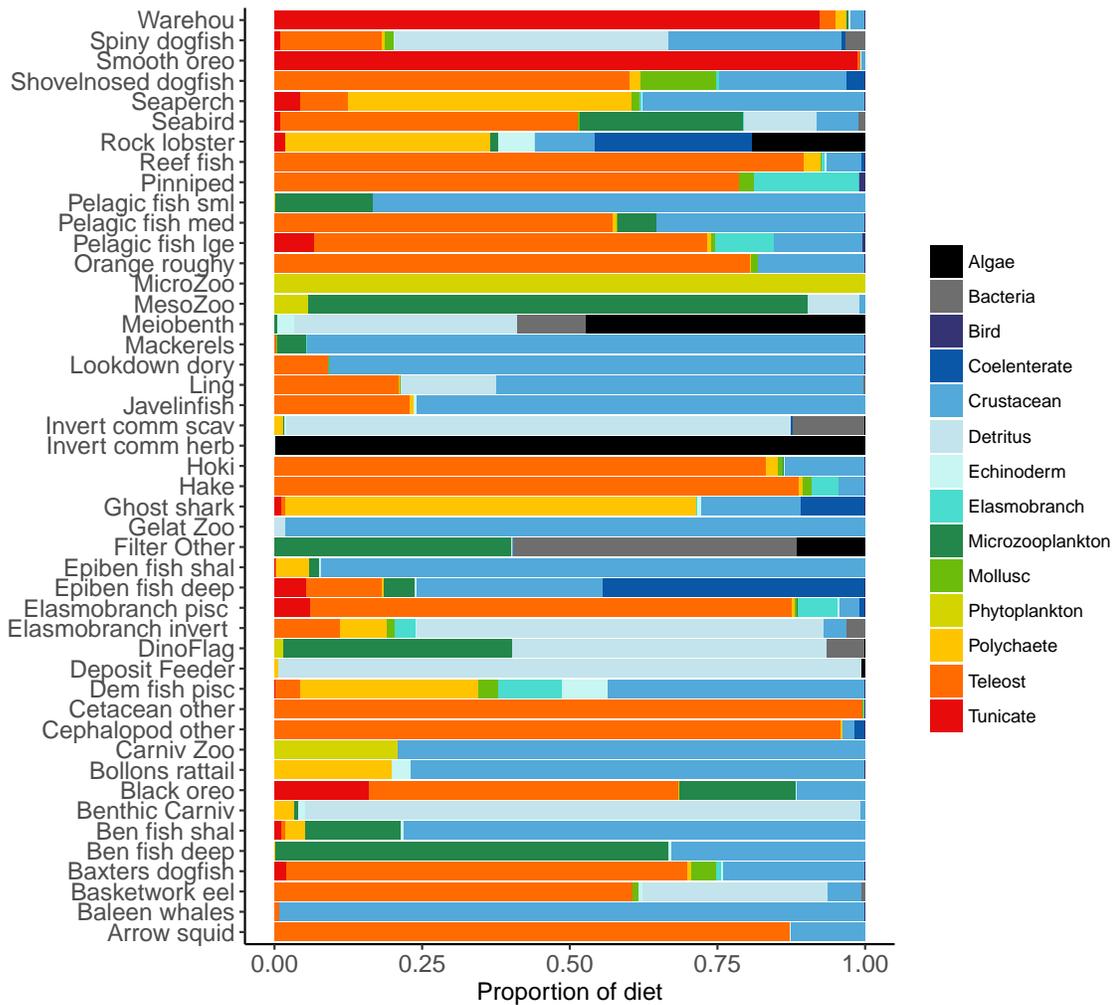


Figure 2.5: Summary of the proportion of prey groups in the diets of species functional groups (Tables 2.2 and 2.3) over model years 1900–2016 from the fished model where the proportion is by mg N consumed.

2.5 Calibration

Calibration of the model included ensuring stable biomass trajectories when applying no fishing; realistic realised diets; realistic growth and mortality (size-at-age and proportions-at-age); and biomass decreasing with increasing trophic level following the PREBAL (pre-balance) (Link, 2010) guidelines.

Biomass trajectories should reach a quasi-equilibrium when modelled with constant oceanography and no fishing (Kaplan and Marshall, 2016). While oceanography is not constant in our non-fishing model as it changes by year (Section 2.4.3), most of the age-structured groups should still be fairly stable. This stability was generally the case; all biomass trajectories remained within CVs of 20% over the simulated 1900–2016 model period, except for invertebrate scavengers (commercial) and seaperch. Invertebrate scavengers (commercial) are primarily scampi, and they are likely responding to changes resulting from the oceanographic variables. Biomass trajectories for all age-structured groups from the un-fished model are in Appendix A. Seaperch biomass was trending downward initially, but they seem to have reached an equilibrium by about 1950, with expected growth and mortality rates.

Atlantis simulates growth rates of age-structured groups as a function of consumption. If growth is too slow, there may be insufficient food available, the feeding search rate could be too low or handling time too high, and the reverse of these when growth is too fast. Simulated growth rates of age-structured species groups were assessed by comparing the simulated size-at-age with those expected based on growth curve estimates from the literature (Table 2.6). The overlaid simulated and ‘observed’ figures were generally very similar (Appendix B). For each species group, I estimated CVs required to satisfy the hypothesis that the modelled size-at-age were not significantly different from the ‘observed’ with probability of 0.95. The required CVs were all less than 30% except for epibenthic fish (deep and shallow), invertebrate herbivore (commercial), invertebrate scavenger (commercial), ling, rock lobster

and small pelagic fishes. For all these groups, the first age class, and sometimes the first few, were larger in size than expected. Deep epibenthic fish were larger than expected at all age classes, but for all other groups the characteristic of larger than expected size at age had been remedied by the time they were adults.

Natural mortality in the model consists of mortality intrinsic within the model from predation, starvation, and light, oxygen or nutrient deprivation, and additional forced mortality. The latter was applied for modelled species groups that would not otherwise suffer sufficient natural mortality within the model, such as those that have little known predation. Age-structured simulated natural mortality rates from the stable base model were compared to estimates of M from the literature where available (Table 2.6) by comparing the proportions-at-age. The overlaid simulated and ‘observed’ figures were generally very similar (Appendix C), although rock lobster and invertebrate herbivore commercial (primarily paua and kina) had slightly more mortality in the model, and demersal piscivores, epibenthic fish small, pelagic fish medium, and warehou had slightly less mortality.

I summarised biomass by trophic level for the base model as at 2016 on a log-scale, and biomass reduced with increasing trophic level with a fitted slope of -1.6 (Figure 2.6). This slope was close to the recommended range of PREBAL of (-1.5, -0.5). The biomass at trophic level 4 was slightly higher in this summary than in the model, because the summary was based on adult trophic level and many of the fish species are trophic level 4 as adults, but lower as juveniles. The use of adult trophic levels resulted in the biomass of the juveniles for these fish adding to the level 4 biomass whereas in the model they were perhaps functioning as a level 3.

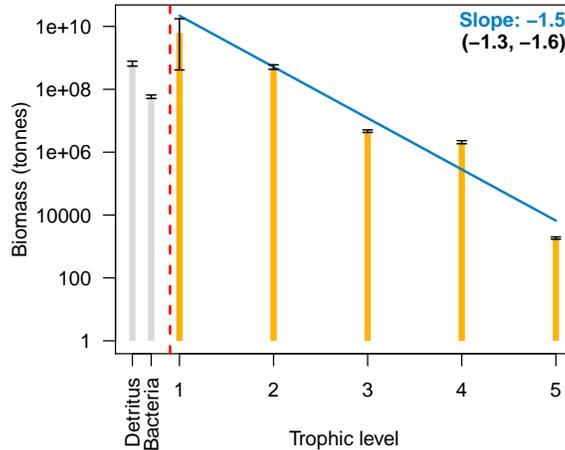


Figure 2.6: Biomass by trophic level with 95% confidence intervals from the 1900–2016 Chatham Rise Atlantis model simulation. The blue line is the fitted linear model to the median biomasses by trophic level, the slope which is in blue. The slopes of the linear models to fitted to the upper and lower 95% confidence interval limits are given in brackets.

2.6 Sensitivity analyses

2.6.1 Oceanography

Oceanographic variables from a ROMS (Regional Oceanographic Modelling System) model for years 1996–2004 were used to define temperature, salinity and flux (water exchange). As my model spanned more than these years, I needed to recycle the ROMS variables in some way. The purpose of this section has two parts: 1.) establishing confidence intervals for our model simulations with respect to oceanographic variability; 2.) assessing the effect of repeating oceanographic variables from any one year, and whether these take the model outside of the established confidence intervals.

To retain realistic within-year dynamics, the ROMS variables from each year were kept together as a unit, and the years covered by the ROMS model were considered the samples. I ran two sets of simulations: the first sampled ROMS years at random with replacement for each model year simulated (bootstrapped the ROMS years) and

repeated this for 50 model runs; the second repeated one ROMS year for all model years simulated and did a separate model run for each of the nine ROMS years. In both cases, the 2003 ROMS was repeated for a 35-year burn-in period, followed by a 50 year simulation. The 2003 ROMS was chosen for the burn-in period as this year had the closest sea temperatures to the means from all ROMS years (Figure 2.7). Bootstrapping the ROMS years was used to establish confidence intervals with respect to between-year oceanographic variability. Repeating each ROMS year in turn was testing the effect of multiple years being different to the other years in some consistent way, such as cooler or warmer.

The established biomass confidence intervals were fairly narrow for most species groups, with CVs $< 10\%$. Of the exceptions, diatoms had the highest CV of 79%, followed by carnivorous zooplankton (46%), labile detritus (23%), sediment bacteria (13%), invertebrate scavengers (commercial) (12%), refractory detritus (12%), meso-zooplankton (11%) and pelagic bacteria (11%). That these groups were found to be most sensitive to oceanographic variability in the model is a plausible and sensible result.

The years with cooler sea temperatures (1996, 1997, and 2004) when repeated for 50 years produced the most species groups that went above the established biomass confidence intervals, with the on average warmer years (1999, 2000, and 2001) having the most species groups that went below (Figure 2.8). These species groups affected by warmer or cooler years had quite a bit of overlap, with meso-zooplankton, meiobenthos, and black oreo most often affected. All of the species groups that went lower in warm years also went higher in cool years. The reverse was not true; three species groups (arrow squid, labile detritus, and ghost shark) went higher in the cool years, but not lower in the warm years.

Years 2003 and 1998 were closest to the average sea temperatures and had the least number of species groups outside the bootstrap confidence intervals. The Base Model that repeated the ROMS from all nine years in order for the entire model simulation

had 16 species groups that exceeded the bounds at some point (less than the warm years) and six species groups that went below the bounds at some point (less than the cool years) (Figure 2.8).

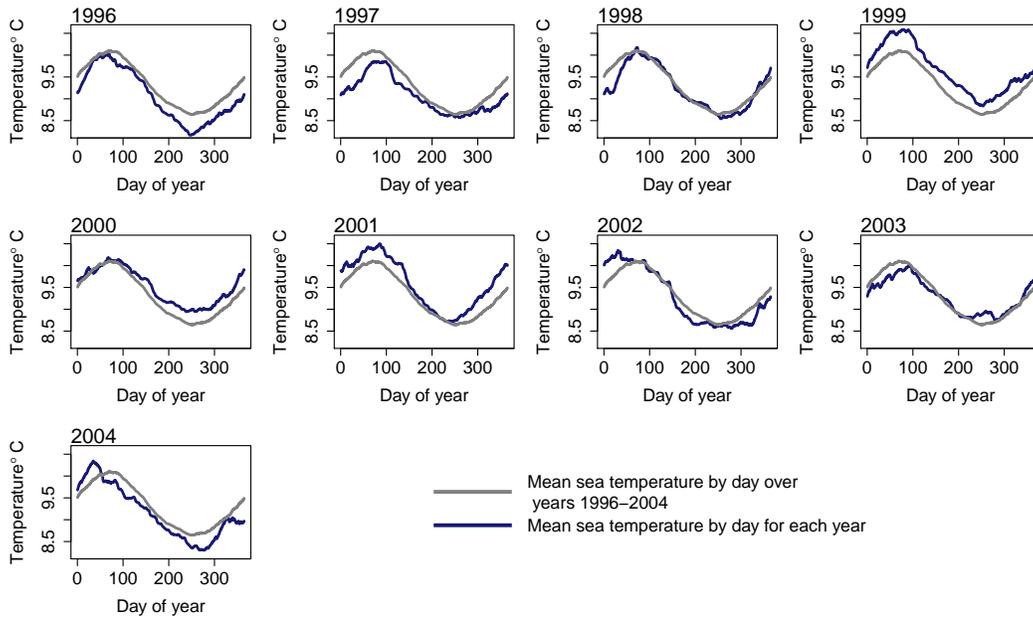


Figure 2.7: Sea temperature ($^{\circ}\text{C}$) from ROMS model outputs by day for each year 1996–2004 (dark blue line) and median sea temperature over all ROMS model years 1996–2004 (grey line).

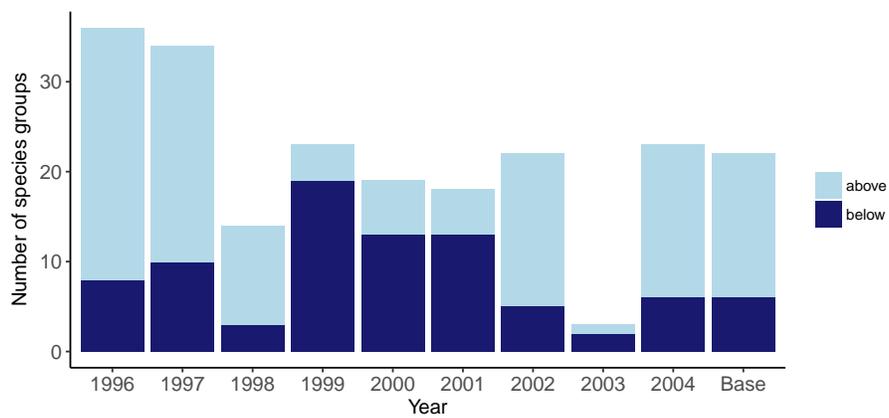


Figure 2.8: Number of species groups for each simulation with one ROMS year repeated that went above (light blue) or below (midnight blue) the limits of bootstrapped ROMS simulations and for the Base Model where the years were repeated in order for the entire model simulation.

2.6.2 Connectivity and influence

Understanding which species groups are most influential or responsive in the model is another test for realistic dynamics, and may be useful to help understand results of scenarios explored using this model in the future. To do this, we need to perturb each species group in turn, then assess the responses of the other groups in the system. This type of analysis is known as *pulse perturbations* (Pantus, 2007). For each age-structured species group, I ran two simulations, one with a small additional mortality and one larger; $M(\text{per year}) + (0.1, 0.005)$. I assessed responses of the groups with respect to the Base Model at the completion of 50-year simulations. I analysed the ‘keystoneness’ and responsiveness of the groups based on biomasses relative to the Base Model.

I calculated keystoneness using an adaption of the method in Libralato et al. (2006). It is a measure of the effect the group has on the rest of the system (change in biomass of the other species groups), that takes into account its proportion of the total biomass. For example, if two species groups have the same effect, but one has a large biomass and one a small biomass, the smaller would have a larger keystoneness. I used simulation outputs to estimate the total effect (ϵ) of each species group (Equation 2.1) which used the change in biomass of each group relative to the Base Model (Equation 2.2). The simulated change in biomasses ($S_{f,g}$) were used in place of the mixed trophic impact values calculated from mass balanced models and used by Libralato et al. (2006). As the additional mortality applied in my simulations caused larger and smaller changes to the focus groups, I scaled the focus groups’ biomass proportions by their change in biomass ($S_{f,f}$ in Equation 2.4). Hence, the resulting keystoneness allowed for the effect changing each group had on the other groups, the focus groups biomass as a proportion of the total, and the proportional change in biomass of the focus group relative to the base model.

$$\epsilon_f = \sqrt{\sum_{g \neq f}^G S_{f,g}^2} \quad (2.1)$$

$$S_{f,g} = \frac{B_{f,g} - B_{b,g}}{B_{b,g}} \quad (2.2)$$

$$\kappa_f = \log(\epsilon_f(1 - p_f)) \quad (2.3)$$

$$p_f = \frac{B_{b,f}}{\sum_{g=1}^G B_{b,g}} \times |S_{f,f}| \quad (2.4)$$

ϵ_f , effect group f has on the other groups

$S_{f,g}$, proportional change in biomass of group g when group f was reduced, relative to the Base Model

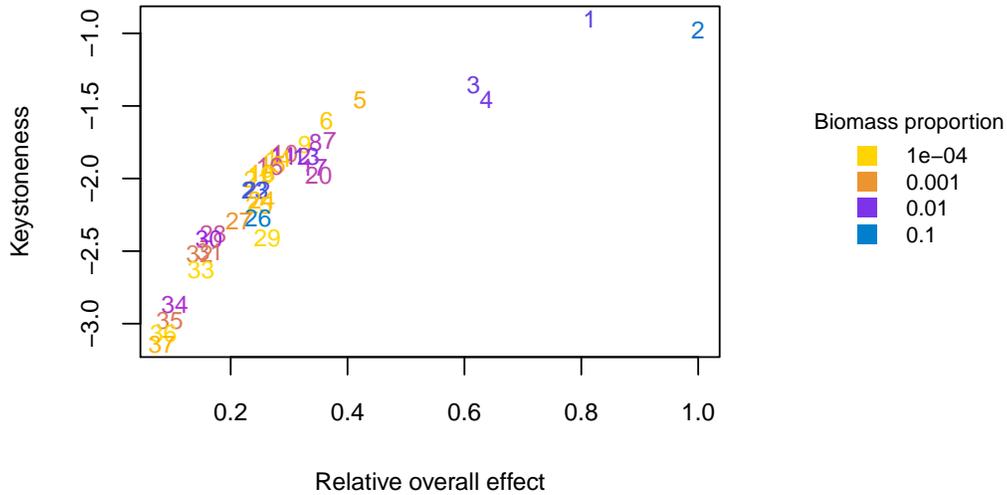
B_{b_g}, B_{b_f} , biomass in base model of group g, f

B_{f_g} , biomass of group g in model with group f mortality increased

κ_f , keystone-ness of group f

p_f , biomass proportion of group f

There were four species groups that stood out as having more effect than the other groups: orange roughy, hoki, pelagic fish small (primarily myctophids) and spiny dogfish. These species groups remain the top four for keystone-ness, but the order changes due to the proportional biomasses (Figure 2.9).



- | | | |
|-------------------------|----------------------|------------------------|
| 1 Hoki | 14 Cephalopod other | 27 Javelinfish |
| 2 Orange roughy | 15 Mackerels | 28 Baxters dogfish |
| 3 Spiny dogfish | 16 Bollons rattail | 29 Pelagic fish lge |
| 4 Pelagic fish sml | 17 Elasmobranch pisc | 30 Ghost shark |
| 5 Ben fish shal | 18 Basketwork eel | 31 Dem fish pisc |
| 6 Seabird | 19 Arrow squid | 32 Elasmobranch invert |
| 7 Pelagic fish med | 20 Lookdown dory | 33 Invert comm scav |
| 8 Epiben fish shal | 21 Invert comm herb | 34 Seaperch |
| 9 Cetacean other | 22 Epiben fish deep | 35 Ben fish deep |
| 10 Hake | 23 Smooth oreo | 36 Pinniped |
| 11 Ling | 24 Reef fish | 37 Rock lobster |
| 12 Shovel-nosed dogfish | 25 Baleen whales | |
| 13 Warehou | 26 Black oreo | |

Figure 2.9: Keystone-ness (y-axis) and relative overall effect (x-axis) for all age-structured species groups, with numbers giving keystone-ness ranking (1 is the most influential using Equation 2.3). Colours indicate biomass proportion scaled by proportional change in biomass (Equation 2.4).

I calculated responsiveness in a similar way to keystone-ness, but from the perspective of the response group (Equation 2.5).

$$R_g = \sqrt{\sum_{f \neq g}^G (m_{f,g}^2 \times p_f)} \quad (2.5)$$

R_g responsiveness of group g to increased mortality in all other groups

The most responsive group was pelagic fish small (primarily myctophids), followed by smooth oreo, invertebrate scavengers commercial (primary scampi), and pelagic fish medium (primarily barracouta) (Figure 2.10). The pelagic fish small species group ranked high for keystone-ness and responsiveness, and so may be most important and influential in scenarios explored with this model.

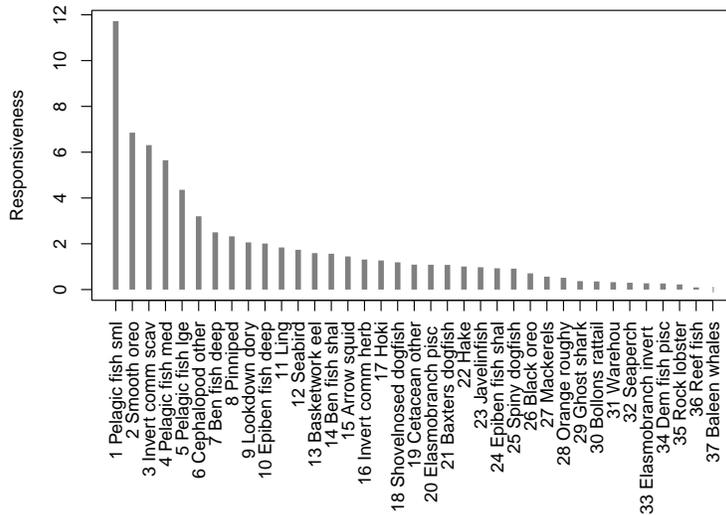


Figure 2.10: Responsiveness of age-structured species groups after 50 years of perturbation, as calculated in Equation 2.5.

2.7 Fishing

Most of the fisheries on the Chatham Rise became established after the mid-1970s, with the exception of the blue cod (*Parapercis colias*) (reef fish species group) fishery which extends back to the early 1900s. Individual catch histories are in Appendix D and Figure 2.11 presents a summary of catches from the Chatham Rise with the top six species by total catch shown in colour and the others combined into an 'other' category. Hoki had the largest total catch, followed by orange roughy, smooth oreo, ling, black oreo, then barracouta. Orange roughy comprised the largest individual fishery in the late-1970s–early-1990s after which it declined markedly; from the 1990s hoki was the dominant fishery.

The fisheries were modelled with six fleets, defined in Table 2.7. The demersal line fishery was dominant until mid–late 1960s when the demersal trawl fishery became dominant, catching approximately 70 000 tonnes per year (Figure 2.12). The historical catches from these fleets were forced in the model using spatially and temporally resolved inputs.

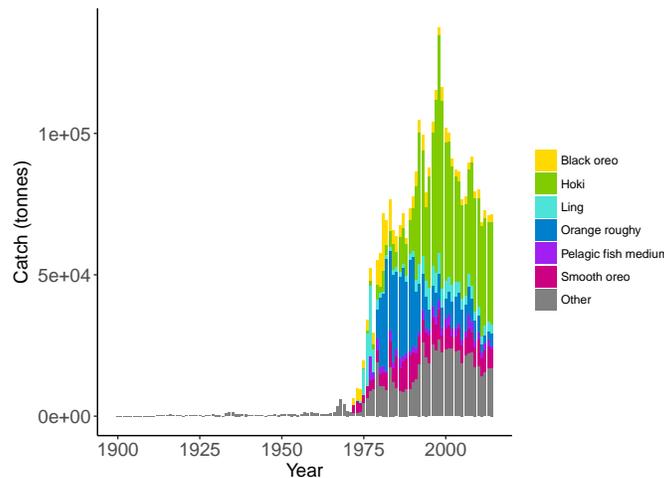


Figure 2.11: Tonnes caught from Chatham Rise 1900–2014 for all species with top six species groups by total catch coloured separately.

Table 2.7: Fishing fleets defined for Chatham Rise Atlantis model. Number of species groups is the number of species groups that have been caught by each fishing fleet; total catch is the total tonnes caught by each fishing fleet from 1900–2014.

Code	Description	Number of species groups	Total catch (t)
trawlDEM	trawl on demersals and mesopelagics	33	2 850 000
lineDEM	line on demersals and mesopelagics	16	1 200 000
snetDEM	setnet on demersals and sharks	6	45 700
potIVS	potting on lobster and blue cod	4	241 000
jigCEP	jig on squid	1	1 700
diveIVH	diving on paua and kina	2	158 000

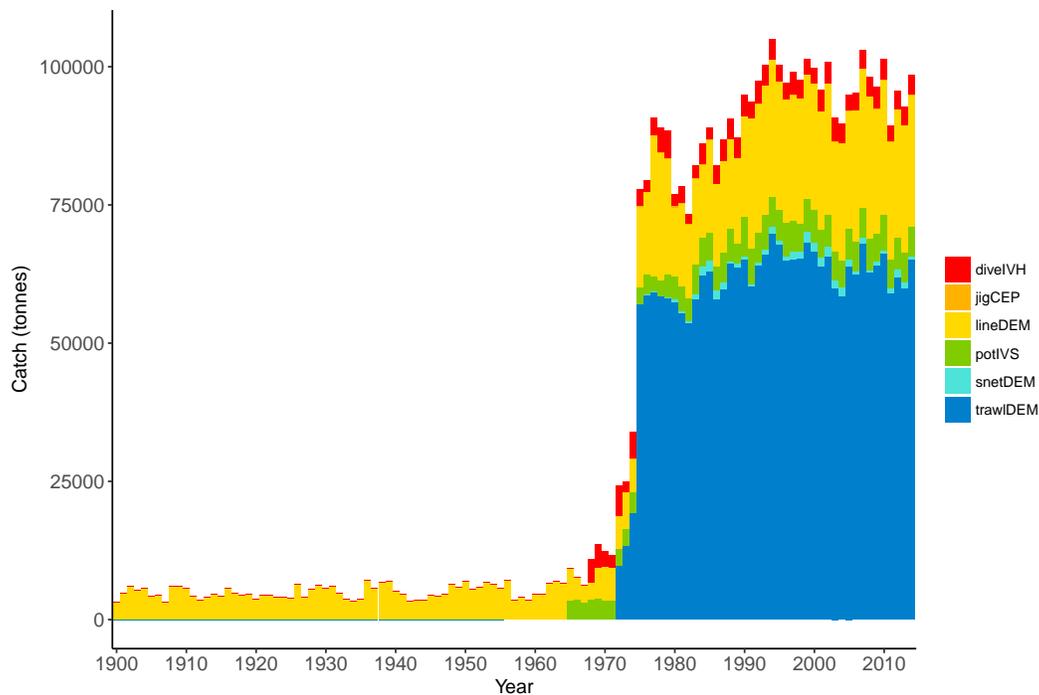


Figure 2.12: Total tonnes caught by fishing fleet from the Chatham Rise 1900–2014. Descriptions for the fleet codes are in Table 2.7.

2.7.1 Comparison with fisheries CPUE and Stock Assessment indices

CRAM model estimates of biomass trends for key fisheries species were compared to CPUE and/or stock assessment indices where these were available. The Atlantis model captures the main biomass trends of hoki in response to historical fishing (Figure 2.13). Hoki are the largest fishery on the Chatham Rise, and has one of the most complex stock assessment models in New Zealand, with multiple areas, intricately defined migration, and annual recruitment deviates (McKenzie, 2016). The Atlantis model results are very similar to the stock assessment model results for hake and ling, and although the stock assessment models for these are not as complicated as hoki, they still have between-year recruitment deviates (Horn, 2013; McGregor, 2015) that are not present in the Atlantis model. The species group ‘Invertebrate scavengers (commercial)’ is primarily scampi, and the matched increase in the late 1990s–early 2000s is particularly pleasing as catches were fairly constant over this time (Tuck, 2016), so the increase is coming from dynamics within the model. Orange roughy is a close match to the stock assessment, even though this stock assessment model also has between-year recruitment deviates (Dunn and Doonan, *ress*) that are not in the Atlantis model. The magnitude of the stock assessment biomasses (unscaled) are compared to the CRAM biomasses in the inset boxplots in Figure 2.13. Hoki, hake, and invertebrate scavengers (commercial) were all close to one, indicating matched magnitudes between the stock assessment and CRAM biomasses. Ling were generally less than one, indicating the CRAM biomasses were larger than the stock assessment biomasses. Orange roughy were greater than one, indicating CRAM biomasses were smaller than the stock assessment biomasses.

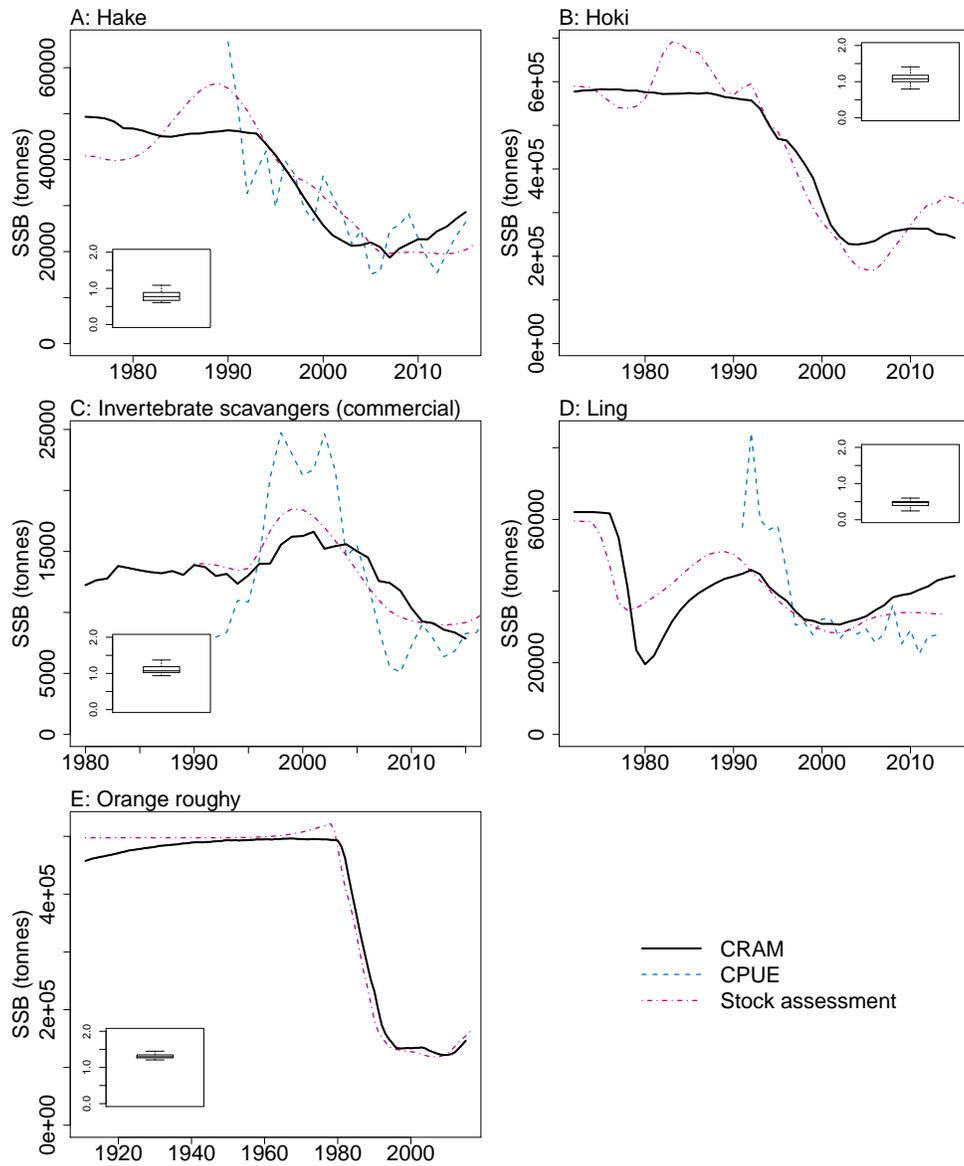


Figure 2.13: CRAM estimated spawning stock biomass (SSB) (black solid), stock assessment estimated SSB (red dot-dash), and CPUE (blue dash) where available for the hake (A), hoki (B), invertebrate scavengers (commercial) (primarily scampi) (C), ling (D), and orange roughy (E). CPUE and stock assessment SSB were rescaled to match the mean of the CRAM estimated SSB. Inset boxplots show the range of values for the corresponding unscaled stock assessment SSB divided by the CRAM estimated SSB.

2.8 Skill assessment

Quantitative skill assessments have become popular as part of assessing the performance of Atlantis models (Sturludottir et al., 2018; Ortega-Cisneros et al., 2017; Olsen et al., 2016). A quantitative skill assessment was carried out, comparing model biomass estimates with those from trawl surveys where available (O’Driscoll et al., 2011; Stevens et al., 2017). The trawl surveys target hoki, hake and ling, and as such the biomass indices are most reliable for these three species. The metrics selected were three of those suggested in Olsen et al. (2016) and Stow et al. (2009): Modelling efficiency (MEF) used to assess model predictions relative to the mean of the observations (Equation 2.6); Reliability index (RI) gives the average factor the model predictions differ from observations (Equation 2.7); Pearson’s correlation (r) assesses whether model predictions are correlated with observations (Equation 2.8). The full set of CRAM biomass trajectories with historic catches and trawl survey indices are in Appendix D.

$$MEF = \frac{\sum_{y=1}^Y (O_y - \bar{O})^2 - \sum_{y=1}^Y (O_y - P_y)^2}{\sum_{y=1}^Y (O_y - \bar{O})^2} \quad (2.6)$$

$$RI = \exp \sqrt{\frac{1}{Y} \sum_{y=1}^Y \left(\log \frac{O_y}{P_y} \right)^2} \quad (2.7)$$

$$r = \frac{\sum_{y=1}^Y (O_y - \bar{O}) (P_y - \bar{P})}{\sqrt{\sum_{y=1}^Y (O_y - \bar{O})^2 \sum_{y=1}^Y (P_y - \bar{P})^2}} \quad (2.8)$$

where

Y is the number of years for which there are observations,

O_y is the observed biomass in year y ,

P_y is the model biomass in year y

Each skill assessment metric was calculated using single point estimates from the trawl survey, and variants on RI and MEF were calculated allowing for the trawl survey estimated 95% confidence intervals. Both variants only penalised the skill metric for terms outside of the 95% confidence intervals of the trawl survey.

An MEF close to one indicates a close match between model predictions and observations, with zero indicating the mean of the observations is as close as the model predictions, and a negative value indicating the model predictions fit the observations worse than the mean of the observations. When the observed values are roughly stationary about the mean, as was the case for Ling, it is difficult for the predictions to improve on the mean of the observations. Ling stands out at approximately -2.5 when compared to the trawl survey point estimates, but as all the predicted points for ling sit within the 95% confidence interval, it receives a score of one when taking the bounds into account (Figure 2.14). Benthic invertivores (shallow) and lookdown dory are slightly negative with respect to the trawl survey point estimates.

A reliability index (RI) of one indicates the model predictions are exactly equal to the observations. RI greater than one (it cannot be less than one) indicates the factor by which observations are on average different to predictions. Since $\log(O/P)$ is equal to $-\log(P/O)$ and the RI squares these terms, an observation that is, for example, half the prediction will contribute exactly the same to this index as an observation that is twice a prediction. Hence, a RI of 2 indicates the observations differ from the predictions on average by 2, but these could be generally twice as big or half as big, or both. All groups had RIs between 1 and 1.5 (Figure 2.14), indicating the observations are at worse on average $1.5\times$ the predictions or $(2/3)\times$ the predictions.

A Pearson's correlation close to one indicates trends in the predictions vary with those in the observations, close to zero indicates there is little relationship between the trends, and negative indicates the predicted trends tend to be opposite from

the observed trends. Hake and hoki had good correlation, close to 0.8. The other groups were either close to zero or negative (Figure 2.14). These low correlations were neither surprising nor concerning as the trawl survey estimates for these groups tend to have high variability and high CVs which are not taken into account here. The trawl survey design is optimised for hoki and hake, and so there are tighter CVs on these trawl survey estimates.

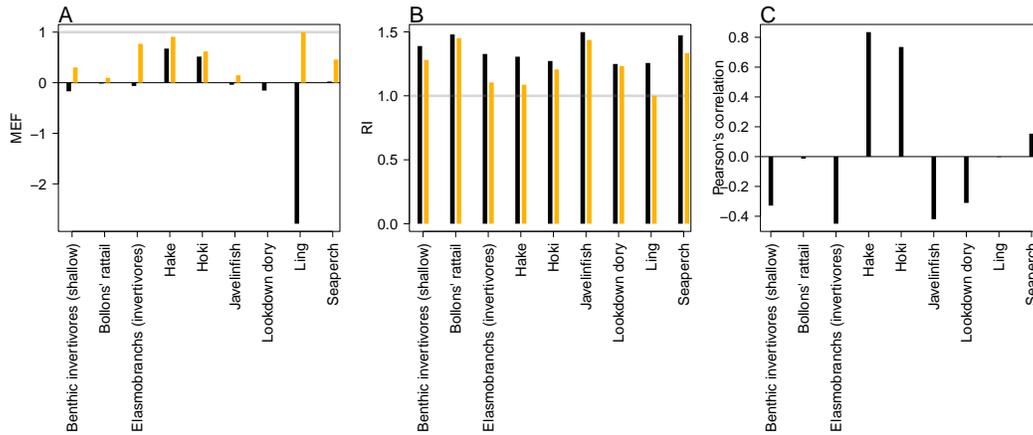


Figure 2.14: Skill assessment metrics MEF (A), RI (B) and Pearson's correlation (C) for CRAM species groups that have trawl survey indices for abundance. Metric definitions in Equations 2.6–2.8. The black bars are the skill metrics with respect to single point estimates from the trawl survey. The orange bars are the skill metrics with respect to the trawl survey 95% confidence intervals. The grey horizontal lines in the MEF and RI figures mark the value for a perfect fit, which is 1 for both of these.

2.9 Bringing it together

I qualitatively graded the species groups by how well they performed in the model and how well informed they were by data, information and other research (referred to as 'informance'). I compared these gradings with the keystone and responsiveness from Section 2.6.2. Figure 2.15 gives a visual guide for how well the most influen-

tial or responsive species groups did for informance and performance. While poor knowledge may not be concerning if paired with high responsiveness providing keystone-ness is low (since the effects may be more limited to this single species group), the triple of highly responsive, a keystone species, and poorly defined may need consideration for future scenarios.

The groups that were highest for keystone-ness and highest for informance and performance were hoki, orange roughy, benthic fish shallow (primarily oblique banded rattail), and hake. These species groups all have abundance indices available, biological parameters, diet information, and all perform well with respect to these in the model. Hoki, orange roughy and hake (groups 1, 2, and 10 for keystone-ness) have full stock assessments, which the model matches well. These species groups are important groups for fisheries and will likely feature strongly in any fisheries scenarios explored with this model.

Species groups Pelagic fish small (primarily myctophids) and Pelagic fish medium (primarily barracouta) were both high with respect to keystone-ness and responsiveness, and while both were fairly well defined, these had some areas of poor model performance and do not have abundance indices to compare. The estimated length at age 1 from CRAM for small pelagic fish is larger than expected. This size difference may be due to the size of recruits being larger than they should be, or the fish eating (and hence growing) more than they should in this first year. These age-1's are not so big that the effect transfers to the age-2's, as the age-2's are the correct size (Appendix B), so this is probably not influential on the model overall. Medium pelagics have slightly less natural mortality in the model than they should (Appendix C), and may be less responsive to fishing mortality as a result. Because this species group are 7th with respect to keystone-ness and high for responsiveness, they could affect scenario outcomes and are worth considering when analysing results. Medium pelagics make up approximately 1% of the age-structured biomass.

Spiny dogfish were third for keystone-ness, and low for responsiveness. This species group fit well to mortality and growth curves, but we do not have an index of abundance with which to compare the model simulated biomass in response to historical fishing. Spiny dogfish make up approximately 5% of the age-structured biomass.

Epibenthic fish shallow (primarily common roughy) were 8th for keystone-ness, but low for responsiveness. This species group compare reasonably well to the trawl survey abundance index (Appendix D), but have less natural mortality in the model than they should. Epibenthic fish shallow make up approximately 1% of the age-structured biomass.

Species groups ‘Seabird’ and ‘Cetacean other’ are both poorly defined and rank within the top 10 for keystone-ness, although lower for responsiveness. These species groups are both composite groups, with Seabird consisting of all sea and shore birds, and Cetacean other consisting primarily of sperm whales, pilot whales and dolphins (Table 2.2). Scenarios explored in the future may benefit from sensitivity analysis with respect to these two groups to understand their effect on the outcomes, or perhaps some more work to better define them.

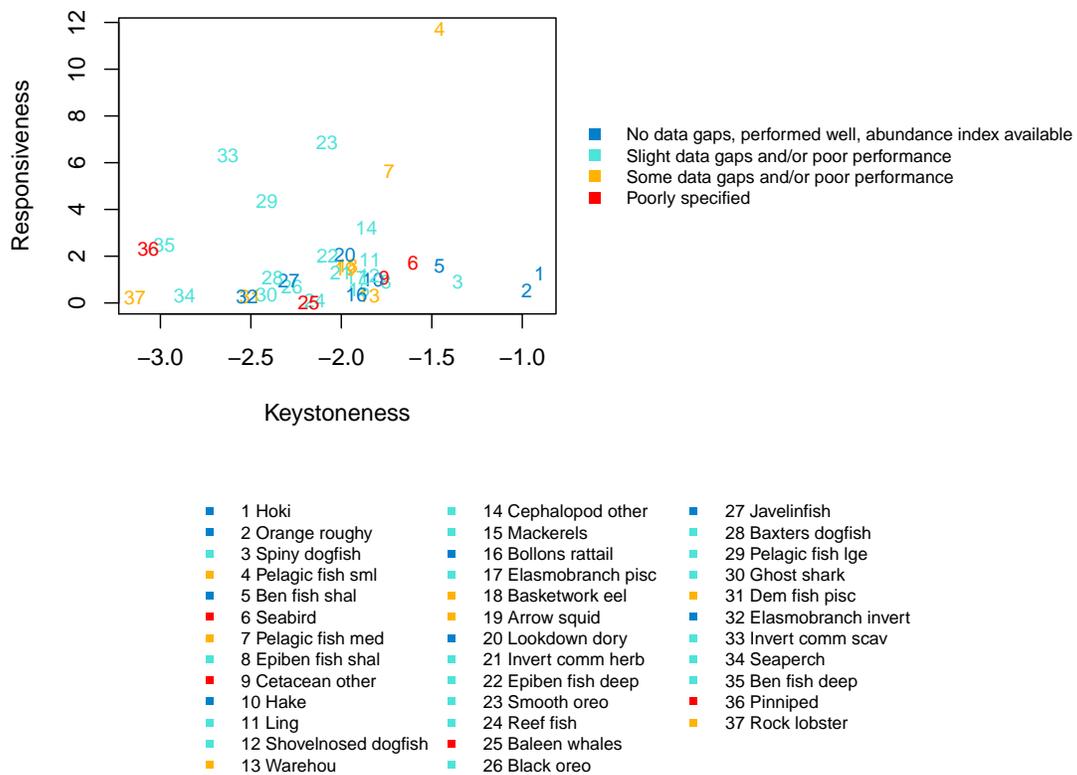


Figure 2.15: Keystoneness (x-axis) and responsiveness (y-axis) with numbers showing keystone ranking and colours how well each species group was informed and/or performed in the model (legend).

2.10 Discussion

Ecosystem-based fisheries management is most likely to be achievable with the best information and modelling available (Heymans et al., 2010). The Chatham Rise Atlantis model presented here uses the wealth of data and information available for the Chatham Rise and its fisheries, and one of the best ecosystem models for exploring ‘what-if’ type questions (Plagányi, 2007) and ecosystem-level management strategy evaluation (Fulton et al., 2014). This comprehensive ecosystem model with realistic population dynamics and flow-on effects has the potential to be a valuable tool for understanding potential system-wide responses to fisheries management strategies in one of New Zealand’s largest fishing grounds Ministry for Primary Industries (2019).

Some key aspects of this model performed convincingly well, such as responses of key fisheries species under fishing, realised diets, and the keystone rankings. That the key fisheries species results were very similar to the corresponding stock assessment results gives confidence that the model can respond to fishing in a way that is realistic, and that the ecosystem effects relative to these species are realistic. The stock assessment models fit data such as proportions at length and biomass indices with the help of between-year recruitment deviates, which are not present in the Chatham Rise Atlantis model. Conversely, the stock assessment models do not have time-varying natural mortality or growth rates, which are present in the Chatham Rise Atlantis model. As such, both modelling approaches achieve similar results but in very different ways. It is possible that the recruitment deviates in the stock assessments are proxies for the other ecosystem dynamics that the Atlantis model is able to capture (or vice versa). However, the Atlantis model is too complex to fit comprehensively to data and is entirely deterministic. Hence, the Chatham Rise Atlantis model’s ability to achieve the same results as the stock assessment models, that were fitted to data, is the best outcome.

Realistic diets and the influence of species groups on the rest of the ecosystem

are key to the model's potential to explore and gain understanding of flow-on and cascading effects. It may be possible, for example, for a species to have realistic growth rates, but it is not very useful in an ecosystem modelling context if they do so by eating the wrong things. While they might respond realistically to direct pressure such as fishing, the flow-on effects would not likely reflect reality. Due to the complex nature of the Atlantis model, the summary of realised diets, together with analysing the keystone and responsiveness, are appropriate for determining whether species interactions are generally realistic, at a level of complexity that can be comprehensible. The Chatham Rise Atlantis model has realistic diet summaries for all species groups, and the top keystone species groups were all those that made sense to NIWA biologists Peter Horn and Matthew Dunn. These positives are not to say the model could not benefit from further future work examining the realised diets at a finer scale - spatially, temporally, and by age-class.

It is worth noting that the trawl survey that provides much of our data particularly for diets, is a summer survey and always carried out in January (Stevens et al., 2017). Hence, if there are shifts in diet through out the year, these are unlikely to be known. Within-year shifts in diet are something we could explore with this model for plausibility. For example, given timing and location of known spawning events, we may be able to explore hypotheses that could connect these events to prey abundance and distribution, and predator behavior.

Exploring the models sensitivity to initial conditions, while not an insignificant amount of work, will be worth doing to add to our understanding of the models stability and persistence of dynamics. Initialisation uncertainty has not, to the best of our knowledge, been done for Atlantis or OSMOSE models, likely due to the enormous complexity and computing resources required for the task. Sensitivities to initial conditions have been explored using Ecopath (Essington, 2007) and Ecopath with Ecosim (EwE) (Steenbeek et al., 2018). We are in the early stages of developing an EwE (Ecopath with Ecosim) version of the Chatham Rise Atlantis

model at NIWA, and it may be more feasible to explore ranges of initial conditions within the EwE framework, with the possibility of then adapting the analyses to the Atlantis model. Sensitivities of high-ranking keystone species, such as spiny dogfish, would be simpler to implement and may produce greater understanding of the model.

While there are some knowledge gaps, I have identified those most likely to influence scenario outcomes through analysing how influential (keystoneness) and influenced (responsiveness) the species groups are on and to each other. The composite groups ‘cetacean other’ and ‘seabirds’ were highly influential while poorly specified. Two solutions would be to a.) split these groups into smaller groups that can be better specified; b.) run sensitivities with respect to these groups when exploring scenarios using this model. As option a.) would require more data than we currently have available, option b.) is the only currently viable option.

The oceanographic variables based on years 1996–2004 were found to be influential on the simulated biomasses of the species groups, and the order they were repeated changed the results, with CVs of up to nearly 80%. This influence suggests scenarios carried out using this model need to consider oceanographic variability in simulated results, using multiple runs with different oceanographic years repeated or changing the order. This influence from the oceanographic variables may be true for many ecosystem models, but I am unaware of similar analyses completed elsewhere. Further work understanding which species groups and/or spatial areas of the model are most affected by oceanographic variability might be helpful in understanding potential impacts on scenario results.

As Atlantis is spatially resolved, there is scope for a greater emphasis on the effects of features such as habitats, depth, and oceanographic features on responses to fisheries management scenarios. Kaplan et al. (2012) explored spatially resolved fisheries management scenarios using an Atlantis model of the California Current, including areas closed to bottom-contact fishing gear, and varying spatial manage-

ment specification relating to Marine Protected Areas (MPAs). In the Chatham Rise ecosystem, it may be that repeating cooler or warmer years such as carried out in this study could influence the spatial distribution of some species. These effects could in turn influence the range of plausible responses to fisheries management scenarios that have a spatial aspect, such as MPAs, the effects of different fishing gear, serial depletion of fishing grounds, and potential effects on by-catch species that may overlap spatially with species that are targeted by fisheries.

The Chatham Rise Atlantis model will most likely to initially be used for exploring scenarios that relate to fisheries. Fishing scenarios currently of interest include fishing at current levels, or at current limits; exploring historical prey shifts in response to fishing; effects of gear selectivity; and effects of area closures. There is interest from Fisheries New Zealand in these scenarios.

2.11 Conclusions

The analyses presented in this chapter are intended to set the stage for an understanding of how the model is specified and how it behaves, but it is not exhaustive. The model produces similar results to fisheries stock assessment models for key fisheries species, and the population dynamics and system interactions are realistic. Confidence intervals based on bootstrapping oceanographic variables were fairly narrow for most species groups, with diatoms, carnivorous zooplankton and labile detritus having the largest CVs. The species groups with the highest keystone-ness were orange roughy, hoki, pelagic fish small (primarily myctophids) and spiny dogfish. The model components that have knowledge gaps and are most likely to influence model results were oceanographic variables, and the aggregate species groups ‘seabird’ and ‘cetacean other’. I recommend applications of the model include alternatives that vary these components. It is expected that any future use of the model will add first to our understanding of the model, and then possibly to our understanding of the ecosystem.

Chapter 3

Spawning Stock Recruitment creates misleading dynamics under predation release in ecosystem and multi-species models

The following chapter is sourced from the open-source peer-reviewed article: McGregor, V. L., Fulton, E. A., Dunn, M. R., 2019a. Spawning stock recruitment creates misleading dynamics under predation release in ecosystem and multi-species models. PeerJ 7, e7308. DOI 10.7717/peerj.7308

PeerJ is an open access publisher, and all content is published under a CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

3.1 Abstract

Ecosystem and multi-species models are used to understand ecosystem-wide effects of fishing, such as population expansion due to predation release, and further cascading effects. Many of these models are based on fisheries models that focus on a single, depleted population, and may not always behave as expected in a multi-

species context. The Spawning Stock Recruitment relationship, a curve linking the number of juvenile fish to the existing adult biomass, can produce dynamics that are counter-intuitive and change scenario outcomes. I analysed the Beverton-Holt Spawning Stock Recruitment curve and found a population with low resilience when depleted becomes very productive under persistent predation release. To avoid implausible increases in biomass, I propose limiting recruitment to its unfished level. This modification allows for specification of resilience when a population is depleted, without sudden and excessive increase when the population expands. I demonstrate this dynamic and solution within an end-to-end ecosystem model, focusing on myctophids under fishing-induced predation release. I present one possible solution, but the specification of stock-recruitment models should continue to be a topic of discussion amongst multi-species and ecosystem modellers and empiricists going forward.

3.2 Introduction

The advent of multi-species and ecosystem models is beginning to extend and augment the fisheries management advice provided by single-species population models, as part of a global move towards ecosystem-based fisheries management (Pikitch et al., 2004, Skern-Mauritzen et al., 2016, Collie et al., 2014). A variety of ecosystem modelling approaches are now available (Plagányi, 2007), and whilst some are highly complex full-ecosystem (end-to-end) models, all have numerous simplifying assumptions to remain tractable, and many include components for population dynamics that were developed for single-species population models for fisheries. Examples of end-to-end models suitable for addressing broad questions include Atlantis (Fulton et al., 2011), EwE (Ecopath with Ecosim) (Christensen and Walters, 2004, Pauly et al., 2000), and GADGET (Globally Applicable Disaggregated General Ecosystem Toolbox) (Begley and Howell, 2004), and simpler models include multispecies size-based models (Blanchard et al., 2014, Pope et al., 2006) and Models of Intermediate

Complexity for Ecosystems (MICE) (e.g., Plagányi et al., 2014 Doonan et al., 2016), the latter of which are used for providing more tactical advice.

The model relating a population’s Spawning Stock Biomass (SSB) to subsequent recruitment, the Spawning Stock Recruitment (SSR) curve, is one component that has transferred from fisheries population models into ecosystem and multispecies models. Spawning stock is usually measured as biomass by weight, and recruitment as numbers at age 1, such that the relationship describes the number of individuals expected in recruitment for a given mature biomass. The Beverton-Holt (Beverton and Holt, 1957) is the most commonly used SSR curve (Shertzer and Conn, 2012), which assumes that recruitment increases with spawning stock size to an asymptote, and therefore it does implicitly assume some density-dependent ecosystem effect, including competition for resources, although the form of this effect is non-specific. While it is possible for some ecosystem models to model the larval stage in sufficient detail such that a SSR relationship is not assumed but is an emergent property (e.g. OSMOSE; Shin and Cury, 2004), most ecosystem and multispecies models do not include sufficient detail for this, and hence a SSR curve is still assumed.

The Beverton-Holt SSR (Equation 3.1) is defined such that when the SSB is at its unfished level (B_0), the number of recruits produced is R_0 . These parameters are not the upper limit for biomass or recruitment as these are also influenced by natural mortality. The asymptote of the Beverton-Holt SSR curve is greater than R_0 , hence if the SSB exceeds B_0 as it may when mortality is low, recruitment will exceed R_0 . Once recruitment joins a population, the amount retained for future years, and hence contributing to the SSB, will also be influenced by natural mortality.

$$R = \frac{\alpha S}{\beta + S} \tag{3.1}$$

where

R is the number of recruits

S is the SSB

When modelling a population with low resilience, the curve can be specified such that the initial part of the curve does not increase as steeply, thus when the population is depleted, it will take longer to recover. The resilience aspect will not affect the unfished state of the population providing the point (B_0, R_0) remains the same. In fisheries modelling, resilience at low biomass levels is defined using the term steepness (h), which is the proportion of R_0 recruited when SSB is 20% B_0 (Mace and Doonan, 1988, Francis, 1992, Haddon, 2001, Lee et al., 2012). From this definition, Haddon (2001) showed the derivation of α and β in terms of h , R_0 and B_0 (Equations 3.2 and 3.3).

$$\alpha = \frac{4hR_0}{5h - 1} \quad (3.2)$$

$$\beta = \frac{B_0(1 - h)}{5h - 1} \quad (3.3)$$

The initial slope of the curve is steeper when defined with a high h and shallower when defined with a low h . To preserve the populations unfished state, the curve passes through (B_0, R_0) for all values of h (Figure 3.1). When S is greater than B_0 , R will exceed R_0 (Equations 3.4–3.11). Both the preservation of the unfished state and the specification of resilience when the population is depleted are important aspects of the SSR for fisheries modelling whether in a multi- or single-species context. Unfortunately, it is these two aspects that lead to counter-intuitive model behaviour when a population expands beyond its unfished state. As the curve is defined to pass through the point (B_0, R_0) , a slow initial increase results in a steeper increase near (B_0, R_0) . Hence, when a populations biomass exceeds its unfished level (B_0), the effect of h reverses and a population defined to have low resilience when depleted will increase more rapidly as it passes B_0 . Changing the initial slope of the curve while keeping the recruitment asymptote unchanged, would require the virgin state of the population to change, and hence I have not explored it in this chapter.

$$R = \left(\frac{4hR_0S}{5h-1} \right) / \left(\frac{B_0(1-h)}{5h-1} + S \right) \quad (\text{Equations 3.2 \& 3.3 into 3.1}) \quad (3.4)$$

$$= \frac{4hR_0S}{B_0(1-h) + S(5h-1)} \quad (3.5)$$

$$\frac{R}{R_0} = \frac{4hS}{B_0(1-h) + S(5h-1)} \quad (3.6)$$

$$R > R_0 \text{ iff } \frac{R}{R_0} > 1 \quad (3.7)$$

$$1 < \frac{4hS}{B_0(1-h) + S(5h-1)} \quad (3.8)$$

$$B_0(1-h) + S(5h-1) < 4hS \quad (3.9)$$

$$B_0(1-h) < S(1-h) \quad (3.10)$$

$$B_0 < S \quad (3.11)$$

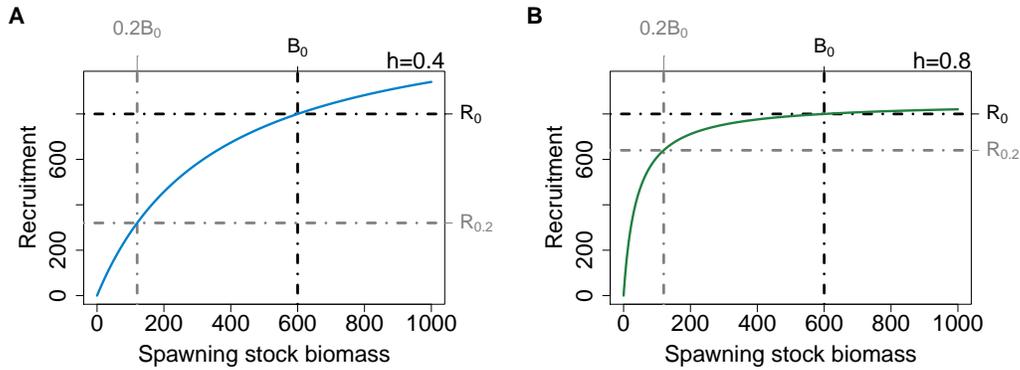


Figure 3.1: Beverton-Holt SSR curve with $h = 0.4$ (A, blue line) and $h = 0.8$ (B, green line), with B_0 and R_0 (black dot-dashed lines) and $20\%B_0$ and $R_{0.2}$ (grey dot-dashed lines).

The switching of the steepness effect has not been such a concern for fisheries assessment models due to stochasticity, statistical fitting to data, and depleted populations. When used in a stochastic population model for fisheries assessments, the SSR curve is combined with recruitment deviates that allow for between year variability. Both the deviates and unfished recruitment (R_0) are estimated using statistical methods to best fit data, hence reducing the concern of implausible recruitment dynamics (e.g. Bull et al. (2012), Fournier et al. (1998), Methot Jr and Wetzell

(2013)). When used for projections, the situation is similar to that in multi-species and ecosystem models, except the population modelled is likely in a depleted state due to fishing and unlikely to exceed its unfished state during the model projection, typically only around five years.

Understanding the impacts of fishing on ecosystems requires careful specification relating to flow-on or cascading effects as it is not just the direct effects on a single exploited species that we are capturing. Inherent in ecosystem and multispecies models is time-varying natural mortality, the sources of which may include predation, disease or starvation (e.g. Fulton et al., 2011, Christensen and Walters, 2004, Pauly et al., 2000, Begley and Howell, 2004). When natural mortality decreases such as from predation release resulting from fishing pressure, a population's biomass may increase beyond its unfished level (B_0). As I demonstrate here, it is possible for a population's biomass to increase to an unrealistic amount in response to fishing pressure on its' predators, and without careful inspection of the SSR curve, the cause of this could remain a mystery. Worse still, such increases could be accepted as valid responses and erroneously affect ecosystem indicators giving misleading results to fisheries scenarios.

A simple solution to large population increases under low steepness may be to cap recruitment at R_0 , such that the SSR relationship is the same for all steepness values when the population exceeds its virgin level. There may be other solutions, perhaps involving damping, switches or smoothers, but it seems a simple solution is a good place to start. There may also be other solutions using the Ricker SSR model (Ricker, 1954), although this model brings with it different assumptions around the SSR relationship. Similar to the Beverton-Holt curve, the Ricker model can be parameterised in terms of steepness, and the effect of steepness switches when biomass increases beyond B_0 , but it also has additional complexity due to all curves going below R_0 when biomass is greater than B_0 . It is the simple solution of capping recruitment at R_0 using the Beverton-Holt SSR curve, that is explored in this study.

While it may not be the best solution for all situations—there may be populations that require recruitment to vary when a population is larger than its virgin level, the simplicity of this solution makes exploration of its application clearer.

One of the most abundant and widespread groups of fish this SSR dynamic could affect in ecosystem or multi-species modelling are myctophids (also called lanternfishes). Myctophids make up roughly 65% of all deep-sea fish biomass in the oceans (Paxton et al., 1998), are a key prey of many commercially important fisheries species globally (Irigoiien et al. (2014), Koz (1995), Clark (1985), Collins et al. (2008), Young et al. (2010)) and have fairly low fecundity rates (Catul et al., 2011). These schooling forage fish are critical to a healthy marine ecosystem both in terms of energy flows within the food web, and nutrient recycling as they act as a biological pump through their feeding-motivated daily vertical migrations (Hernández-León et al., 2010). As such, myctophids may be influential in ecosystem indicators used to evaluate ecosystem-wide effects of fisheries scenarios within ecosystem and multispecies models.

In this chapter, I first explore the Beverton-Holt SSR curve under varying natural mortality with different steepness values using a simple population model. I then present fishing-induced predation release of myctophids using an end-to-end ecosystem model of the Chatham Rise, New Zealand (McGregor et al., 2019b). I compare resulting biomass trajectories from the end-to-end ecosystem model of myctophids under predation release using different steepness values, with and without capping recruitment. In addition, I calculate key ecological indicators from fishing scenarios with and without capping recruitment.

3.3 Simple population model with varying natural mortality

I explored the effect of Beverton-Holt SSR on population abundance with time-varying mortality and different steepness values by simulating a simple population (Equation 3.12). I used case study values of $B_0 = 600$ tonnes, $R_0 = 800$ thousand individuals and ran 13 simulations, each with a different h from 0.35–0.95. Base instantaneous natural mortality rate (M) was calculated to produce a constant population where the number of deaths match the number of recruits and held the population at B_0 (Equation 3.13), with N_0 the product of B_0 and the average weight of spawning stock individuals (set at 500 grams). Time-varying mortality was constructed using the sine curve to include a period of high mortality followed by a period of low mortality, and with random variability added using random variable $X \sim N(0, 0.005)$ (Equation 3.14 and Figure 3.2).

$$N_t = (N_{t-1} + R_{t-1})e^{-M} \quad (3.12)$$

where

N_t is the abundance in numbers at time t

R_t is the recruitment in numbers at time t

$$M = -\ln\left(\frac{N_0}{N_0 + R_0}\right) \quad (3.13)$$

$$M_t = 0.8M \times \sin\left(\frac{\pi t}{100}\right) + M + X \quad (3.14)$$

where t is the timestep (year), M is the base M ,

$X \sim N(0, 0.005)$

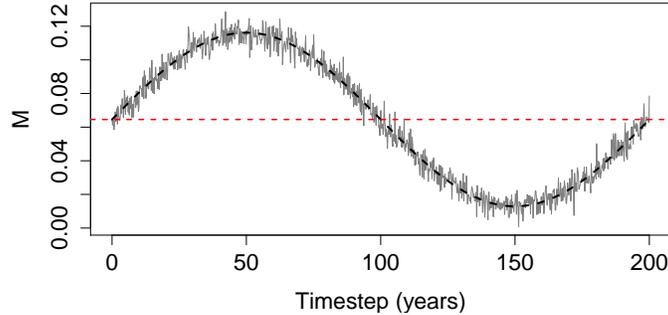


Figure 3.2: Time-varying mortality (M) used for simple population model. The horizontal dashed line is the base mortality (base M) that would result in the population staying at B_0 .

The effect of h on the simulated SSB was as expected when M was high and the population depleted; the effect was reversed when M was low and the population increased beyond B_0 (Figure 3.3). SSB trajectories were lower when SSR was parameterised with low steepness (h) during the period of high mortality, as is the intended effect of steepness—a population with low h is less resilient and if depleted it will take longer to recover. SSB trajectories during the period of low mortality were higher when SSR was parameterised with low h , which is not the intended effect of steepness. Thus, a population deemed to have the least resilience when depleted is the most productive when mortality is reduced. The SSR curve coloured to reflect different values of h shows the cause; recruitment is greater for a given SSB when $SSB > B_0$ (Figure 3.3).

I applied a cap to recruitment at R_0 and re-simulated the same simple population for each steepness value (Figure 3.4). The resulting biomass trajectories were identical to the un-capped version when mortality was high and $SSB < B_0$ because this part was not affected by the recruitment cap. During the period of low mortality, the trajectories almost converged to the same curve, rather than increasing further when using lower steepness. Hence, when applying a cap of R_0 to recruitment, the steepness value defined to reflect resilience when the population is depleted will be

inconsequential when the population increases beyond B_0 .

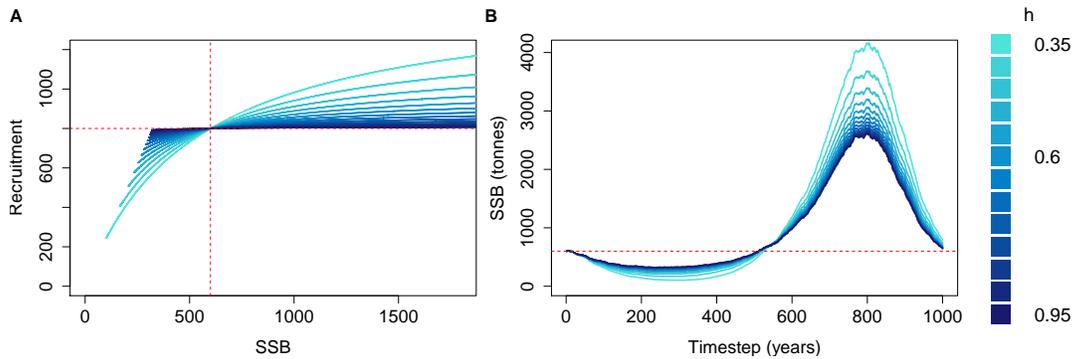


Figure 3.3: Beverton-Holt SSR curves (A) and resulting biomass trajectories (B) from a simple population model using different steepness, h , ranging from 0.35 –0.95 and time-varying mortality as in Figure 3.2. B_0 shown with vertical red dash line (A) and horizontal red dashed line (B); R_0 shown with horizontal red dashed line (A).

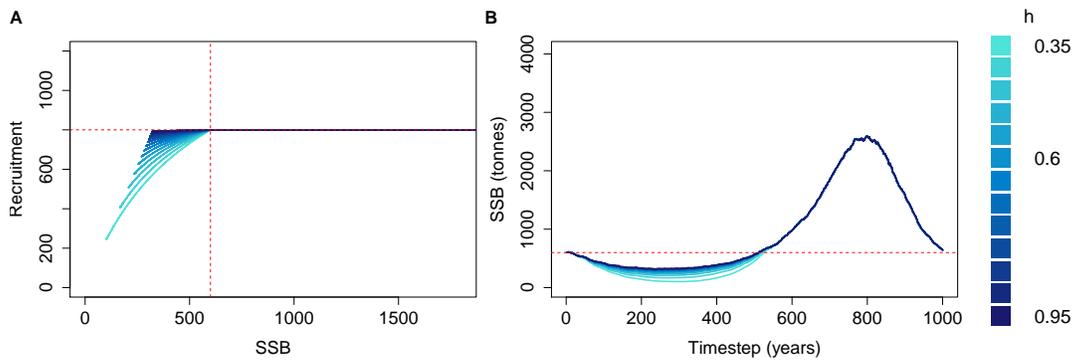


Figure 3.4: Beverton-Holt SSR curves with recruitment capped at R_0 (left) and resulting biomass trajectories (right) from a simple population model using different steepness, h , ranging from 0.35 –0.95 and time-varying mortality as in Figure 3.2. B_0 shown with vertical red dash line (A) and horizontal red dashed line (B); R_0 shown with horizontal red dashed line (A).

3.4 Fishing induced predation release

I tested the Beverton-Holt SSR dynamics within an end-to-end ecosystem model of the Chatham Rise, New Zealand, using Atlantis (McGregor et al., 2019b; Fulton et al., 2011). The Chatham Rise Atlantis model can be run with forced catches, allowing me to analyse the resulting model estimates of biomass. This model consists of 55 species functional groups, 37 of which were modelled with age-structure, and the remainder as biomass pools. Myctophids (modelled with age-structure) were selected as the case-study species group for recruitment dynamics as they are an abundant prey species of many key fisheries species in this area, including hoki (*Macruronus novaezelandiae*), black oreo (*Allocyttus niger*), orange roughy (*Hoplostethus atlanticus*), and medium pelagic fish (primarily barracouta (*Thyrsites atun*)). Myctophids are also an abundant prey species globally (Irigoiien et al., 2014), and have fairly low fecundity (Catul et al., 2011) suggesting resilience of their population when it is depleted to low levels may be low.

Catch histories for the Chatham Rise Atlantis Model are available as supplementary materials to McGregor et al. (2019b). Species with stock assessments have catch histories that were reconstructed under the review of technical fisheries working groups, including hoki (McKenzie, 2016), orange roughy (Dunn and Doonan, *ress*), hake (*Merluccius australis*) (Horn, 2013) and ling (*Genypterus blacodes*) (McGregor and Horn, 2015). Catch histories for species without a stock assessment were reconstructed using commercial catch and effort data from the 'Warehou' database administered by Fisheries New Zealand (1989–present day), the Fisheries Statistics Unit (FSU) database administered by NIWA (1974–1989), and from Annual Reports on Fisheries (1900–1973). While the reliability of these data varies across species and time, and the quality of data reporting likely improved with the introduction of New Zealand's Quota Management System (QMS) in 1986, these catch histories are a sufficiently accurate representation both in trend and magnitude.

I simulated the Chatham Rise Atlantis model with alternate SSR specification for myctophids consisting of all combinations of steepness $h \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$ and recruitment capped at R_0 or not — 10 model simulations in total. From the mid-1970s when many of the Chatham Rise fisheries became established (Figure 3.5), the effects of predation release on myctophids in the model become apparent (Figure 3.6). Their biomass increased beyond B_0 and increased more with lower recruitment steepness (h), with low h producing the greatest resulting biomass. When I capped recruitment at R_0 , the biomass was still able to increase as we may expect under predation release, but the increase was the same regardless of h . This modified dynamic allows for the specification of a species that is less resilient when depleted without it increasing more rapidly when released. The effect on the myctophid population dynamics was similar to the low-mortality response in the simple population model of the previous section in which I controlled time-varying natural mortality. The Atlantis model simulations gave a brief period of high mortality (near 1980), and this only affected the biomass when steepness was lowest and recruitment was capped at R_0 (yellow line, right plot, Figure 3.6). When recruitment was not capped, the low steepness simulation kept biomass above B_0 when the capped version was not able to do this.

Steepness was converted to α and β for each simulation, with the modified α and β used as parameters for the model. The cap was implemented within the Atlantis source code in the `Get_Recruits` function, within the `SSB_BevHolt_recruit` case (precase 19 in the Atlantis parameter input file). I calculated R_0 and B_0 using the initial conditions of the model, but assuming that the initial conditions were actually slightly below B_0 (I took $B_0 = S/0.9$). Then, if any recruitment R exceeded R_0 I overwrote it with R_0 . It would have been more efficient to calculate and store R_0 on model initialisation, but for the purpose of this work it was convenient to have all the code modified within one function.

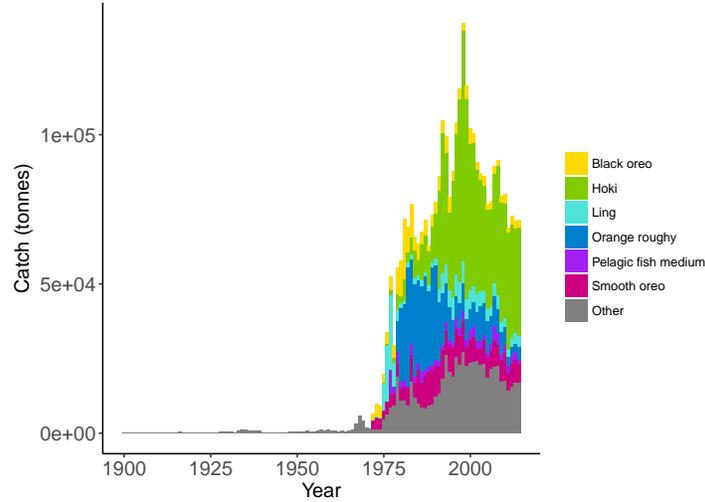


Figure 3.5: Historical catches from 1900–2014 with top six species groups based on total catch coloured separately (McGregor et al., 2019b).

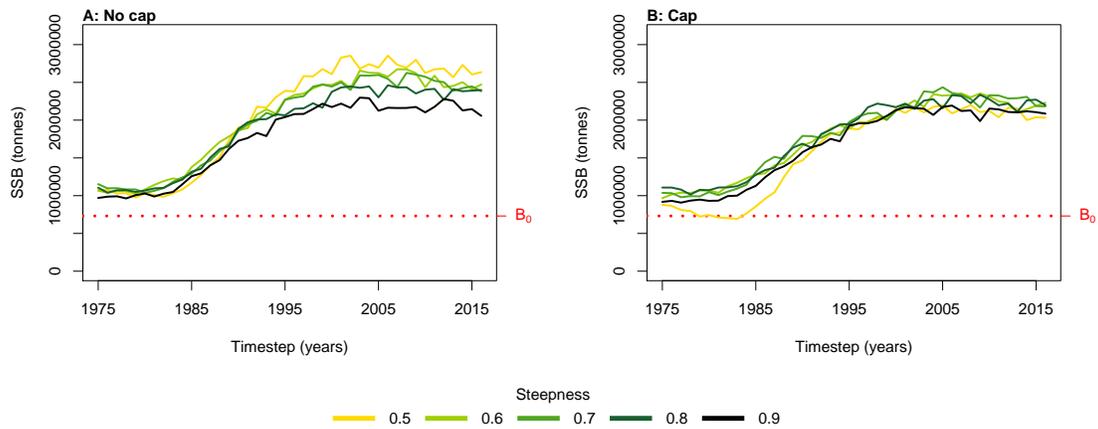


Figure 3.6: Hindcast biomass trajectory for the myctophids species functional group with different recruitment steepness (h) and no cap on recruitment (left) and with recruitment capped at R_0 (right).

3.5 Ecological indicators

I evaluated a range of ecological indicators intended to assess the effect of fishing on an ecosystem and tested for sensitivity in the response to recruitment steepness of myctophids and whether their recruitment was capped at R_0 . To do these tests, I used the outputs from the Chatham Rise Atlantis model, including estimated biomass, numbers-at-age, and size-at-age. The ecological indicators selected (descriptions and

references in Table 3.1) were those that have gained traction in the literature due to ease of application and usefulness, and which were also suited to the Chatham Rise (Tuck et al., 2009). Kempton’s species diversity index measures the number of species relative to the total biomass. I used the modification of Kempton’s species diversity index by Ainsworth and Pitcher (2006) which uses species groups rather than individual species and the 10th percentiles to eliminate any very small or large species groups in terms of biomass (Equation 3.15). All indicators were calculated for the hindcast period (1900–2016) for six models consisting of three steepness values (0.5, 0.7, 0.9) and either capped or not capped recruitment for myctophids, and for projections using all of these models under three catch scenarios (zero catch, *status quo* catch, half catch), giving 18 projections. *Status quo* catch was taken to be the average from the most recent five years (2012–2016) for each fished species group, and half catch was half the *status quo* catch.

Table 3.1: Ecological indicators evaluated for hindcast simulations and catch scenario projections, with key references for each indicator.

Indicator	References
Kempton’s species diversity index	Ainsworth and Pitcher (2006)
Mean trophic level	Pauly and Watson (2005), Shin et al. (2018)
Biomass/landings	Shin et al. (2010), Shin et al. (2018)
Landings/biomass of primary producers	Link (2005)
Biomass of pelagic fishes/biomass total	Link (2005)
Biomass trophic level 4 and above/biomass trophic level 3	Link (2005)

$$Q = \frac{0.8\psi}{\log(\rho_2/\rho_1)} \quad (3.15)$$

Q is the adapted Kempton’s species diversity index from Ainsworth and Pitcher (2006)

ψ is total number of functional species groups

ρ_1, ρ_2 are the lower and upper 10th percentiles in the cumulative abundance distribution

All indicators except landings/primary productivity had different responses to both recruitment steepness and recruitment cap in either or both the hindcast simulations and the catch scenario projections (summarised as ‘sensitivity’ in Table 3.2). These indicators may help understand the effect of fishing on the ecosystem - whether the effect is positive, negative, or perhaps caused by a shift in the system. The indicators focus on different aspects of the system, and hence may give different impressions on the state of the ecosystem. Imposing a cap on recruitment for myctophids generally suggested more favourable responses of the ecosystem to fishing, except for biomass over catch which was reduced. The full set of figures for ecosystem indicators are in Appendices A–E, with subsets in the text (Figures 3.7–3.11) that most effectively illustrate the analyses.

Table 3.2: Sensitivity (indicated by ✓) of ecological indicators to steepness and recruitment cap in the hindcast and/or catch scenario projections.

Indicator	Steepness	Recruitment cap
Kempton’s species diversity index	✓	✓
Mean trophic index	✓	✓
Biomass/catch	✓	✓
Landings/biomass of primary producers	-	-
Biomass of pelagic fishes/biomass total	✓	✓
Biomass trophic level 4 and above/biomass trophic level 3	✓	✓

Kempton’s Q, a measure of diversity as adapted for ecosystem models by Ainsworth and Pitcher (2006), suggested the models with no cap on myctophid recruitment were more likely to suffer a drop in diversity earlier (Figure 3.7, and Appendix A). The models with a cap on myctophid recruitment indicated the system would not recover in terms of diversity within the next 30 years in any of the three catch scenarios, but the uncapped models recovered providing fishing was at that the *status quo*. In the un-capped models, recovery occurred earlier when steepness was lower. It seems counter-intuitive the *status quo* catch scenarios would recover in terms of diversity and the reduced catch scenario (half catch) would not. Given the higher catch scenario brings recovery and the timing is earlier for lower myctophid steepness, this

counter-intuitive dynamic is likely due to the additional myctophids in the system due to predation release and the reversed effect of steepness.

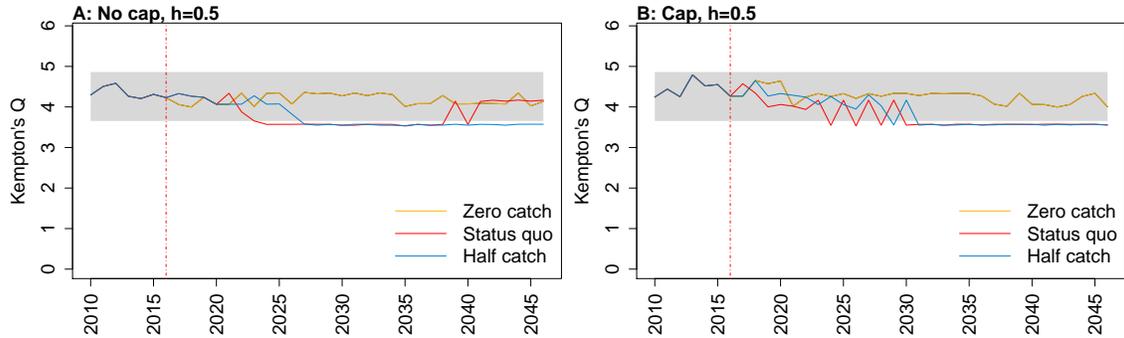


Figure 3.7: Kempton's Q calculated from Chatham Rise Atlantis model simulations with recruitment steepness set at 0.5 for myctophids, no cap on recruitment (left), recruitment capped at R_0 (right), and three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2010–2016 hindcast period and 2016–2046 projection period. The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.

Mean trophic level was calculated for the age-structured species groups in the model. This index was affected by steepness in the hindcast models with no cap on recruitment and the corresponding projections, although the effect was not linear. Mean trophic level was lowest when myctophids had highest steepness (0.9) and there was little difference between steepness of 0.5 and 0.7, although the latter was slightly higher (Figure 3.8 and Appendix B). There is some feedback within the model with medium and large pelagic fishes increasing in biomass with myctophids, and this is likely to come through in this trophic level indicator. All the models with capped recruitment gave higher mean trophic levels both in the hindcast and projections.

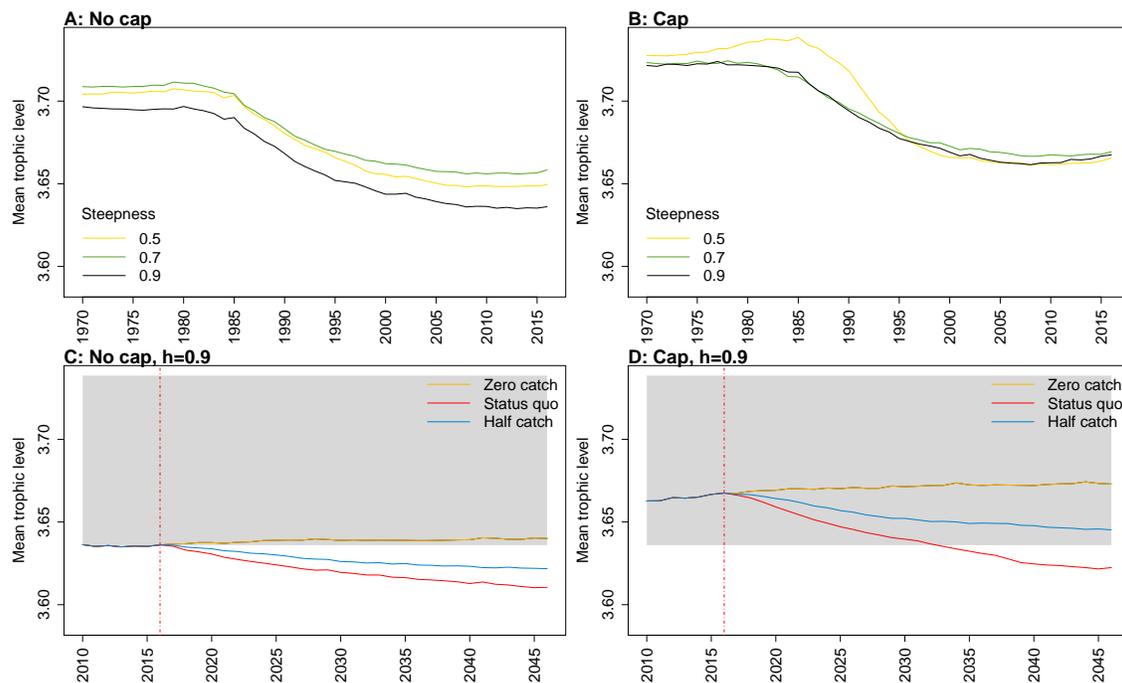


Figure 3.8: Mean trophic level of age-structured species groups calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (A, C), recruitment capped at R_0 (B, D), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (A, B) and h set at 0.9 for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (C, D). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.

Biomass over catch was calculated using the biomass of age-structured species groups as this covers the species that may be fished. This index was affected by recruitment steepness when recruitment was not capped, although the effect was fairly subtle, and perhaps dwarfed by the large and sudden increase in this ratio when catches were halved (Figure 3.9 and Appendix C). In all cases, halving the catches resulted in the biomass/catch ratio increasing suddenly, as expected. The ratio was highest when recruitment was not capped, and steepness was lowest. While this pattern was present in all hindcast and projections, it was most evident when steepness was 0.5 and recruitment was not capped. When recruitment was capped, there was a dip in the ratio in the 1980s–mid-1990s in the low-steepness model. This reduction due to low steepness is a more plausible effect as I would expect there to be less myctophids when they are defined with low steepness. As I have not included biomass-pool species groups in this index, the main prey groups of myctophids were excluded. Hence, a reduction in myctophids was not offset in the index by an increase in their prey.

While biomass of pelagic fishes over total biomass exceeded previous levels under *status quo* fishing for all recruitment scenarios, it was highest when steepness was lowest (0.5) and recruitment was not capped (Figure 3.10 and Appendix D). For this recruitment scenario, the ratio for both *status quo* and half catch projections went above the previous highest level by approximately 20%. This effect is a direct response to the higher myctophid biomass as the low-steepness recruitment model expands under predation release.

Ratio of biomass at trophic level 4 and higher over biomass at trophic level 3 was sensitive to recruitment cap and steepness in the projections. The ratio was reduced when steepness was lower, and recruitment was not capped. In all *status quo* projections with no recruitment cap, the ratio went below the previous lowest level, and was lower when steepness was lower (Figure 3.11 and Appendix E). In all projections with recruitment capped, the ratio remained within pre-fishing bounds.

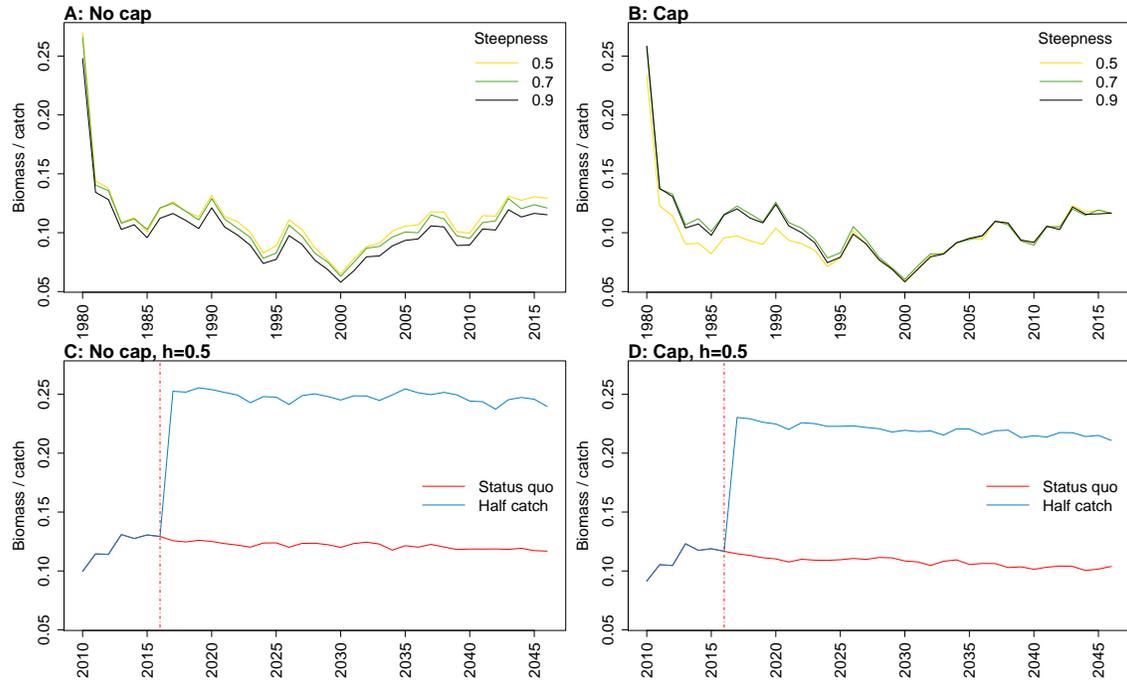


Figure 3.9: Biomass of age-structured species groups over catch calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (A, C), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (B, D) and h set at 0.9 for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.

The warning and limit reference points for this indicator as defined by Link (2005) were concerned with the ratio becoming too high, with the suggested control rule to alleviate fishing on trophic level 3 species. As fishing pressure on the Chatham Rise is high on trophic level 4 species and level 3 species, this ratio becoming too low may also be of interest or concern as it relates to the balance of the ecosystem.

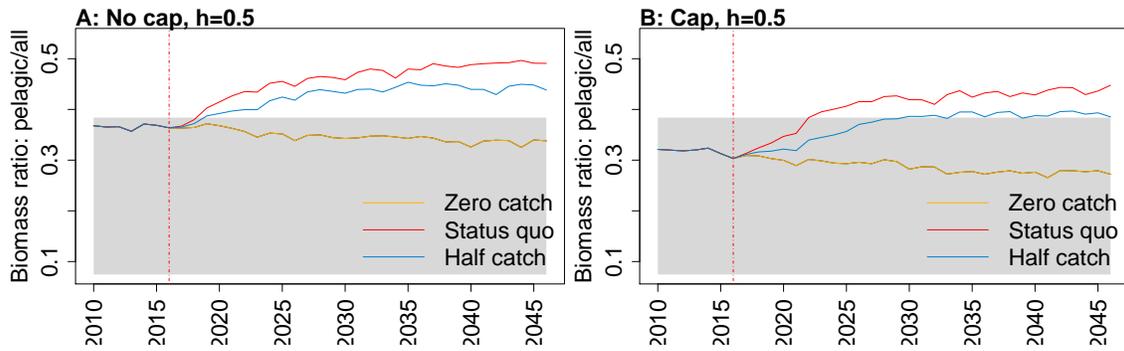


Figure 3.10: Biomass of pelagic fishes over biomass of all age-structured species groups from Chatham Rise Atlantis model simulations with recruitment steepness set at 0.5 for myctophids, no cap on recruitment (A), recruitment capped at R_0 (B), and three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2010–2016 hindcast period and 2016–2046 projection period. The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.

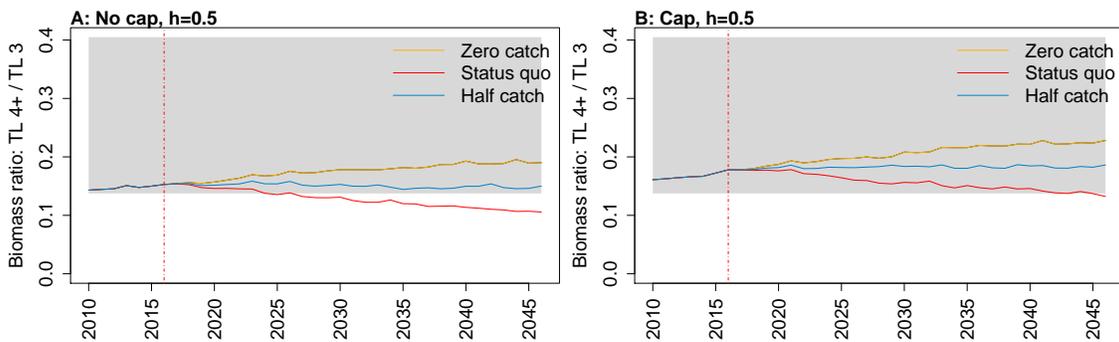


Figure 3.11: Biomass ratio of trophic level 4 and higher over trophic level 3 from Chatham Rise Atlantis model simulations with recruitment steepness set at 0.5 for myctophids, no cap on recruitment (A), recruitment capped at R_0 (B), and three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2010–2016 hindcast period and 2016–2046 projection period. The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970 when fishing was non-existent or very small.

3.6 Discussion

SSR parameter ‘steepness’ that reflects a population’s resilience when depleted, is an influential and often unknown parameter in fisheries stock assessment models and is often used in ecosystem and multi-species models (Lee et al., 2012). Regardless of whether we define resilience using the fisheries term steepness or not, this attribute of the SSR curve still exists and is important for specifying the dynamics of a modelled population. It is important to be able to define the slope of the initial increase of the recruitment curve without changing the level of recruitment expected when there is no fishing (R_0).

I found the effect of changing steepness in the Beverton-Holt SSR model is reversed when SSB is greater than the unfished biomass. This effect matters when modelling populations where mortality may be reduced such as from predation release, as this has the potential to allow a population to increase beyond its unfished level, as demonstrated in this chapter. Using the Beverton-Holt SSR model, a highly productive population with a high steepness value would become less productive when SSB exceeds unfished biomass than a population that has a low steepness value, even though both are defined with the same unfished recruitment and biomass.

When recruitment is capped at its unfished level, a population may still increase beyond its unfished state under reduced mortality, but it will not increase more with lower steepness. Hence, steepness will only affect a population when it is depleted. This modification is an improvement on the effect of steepness being reversed, but there may be an alternative that allows for a population with low steepness to continue to be less productive, for situations when this dynamic is appropriate.

I explored the ecosystem wide responses to fishing with varying steepness and capping or not capping recruitment for the myctophids species group. This group are abundant in marine ecosystems including the Chatham Rise, and they are prey

to many key fisheries species. Even though I was only changing recruitment in one species group, this did result in changes in ecosystem indicators. Scenarios modelled without a cap on myctophid recruitment were more likely to suggest a shift in the ecosystem through increased biomass of pelagic fishes with respect to the total biomass, and more likely to suggest the systems diversity would recover if fishing remained high. Scenarios modelled with a cap on myctophid recruitment were less likely to suggest reduced average trophic level in the system in response to fishing, and the diversity of the system took longer to decline.

The results of ecosystem indicators with respect to positive or negative effects of fishing do not tell us much about whether recruitment should be capped or not, but they do flag some situations that may be in danger of misleading results if we do not. The shift in the system due to an increase in pelagic species biomass beyond previous levels is one example of potentially misleading results. It does not make sense for the expansion of pelagic fishes to be more extreme when they are specified as less resilient. The recovery of diversity under the highest catch scenario when recruitment was not capped and not in the lower catch scenario, with earlier recovery for lower steepness, could be concerning.

The decision of whether to cap recruitment or not –or to apply a different solution, may be different across models, population dynamics, and perhaps intended model use. There may be information in data or the literature to inform reasonable or likely population growth in such situations as predation release, in which case these may be used to inform this decision. If there is nothing available as a guide, then sensitivity analyses are a sensible option, at least to ascertain the possible responses of the modelled system to an explored scenario. This approach is not unique to the specification of the SSR curve; it holds for all model development and exploration. It is the responsibility of the modeller to understand the dynamics of the model they are creating, not just from comparing biomass trends to observations, but actually understanding the dynamics relating to flow-on effects, and the implications that

come with every parameterisation and specification of their model. Where there are uncertainties in the model, results from explored scenarios need to reflect these.

There is traction in the literature for controlling fishing effort on different trophic levels such as through balanced harvesting (Jacobsen et al. (2014), Reid et al. (2016), and critiqued in Froese et al. (2015)). Key to this idea is the response of prey species under predation release, as it aims to replace this predation mortality with fishing mortality, thus taking more total catch with little negative effect to the ecosystem. An example is the study Smith et al. (2011) where comparisons are drawn between scenarios in which lower trophic levels are not fished and those where they are. As I have shown in this study, such comparisons may be misleading. Careful consideration needs to be taken for the population dynamics under scenarios involving predation release.

Key to this work is that natural mortality varies with respect to time in ecosystem and multi-species models, and this may change how we think of ‘unfished’ biomass and recruitment. While we can adjust the SSR model to cope with increased biomass beyond its unfished level, it does raise the question ‘should we?’ Perhaps when biomass increases beyond unfished biomass, B_0 and R_0 should change, and hence the specification of the SSR curve should also change. But would this additional dynamic complexity help a model in its ability to fit to or replicate historical data, and predict future trends? There are times in single species fisheries models where natural mortality varies with time and the SSR relationship remains fixed (Nielsen and Berg (2014), Deroba and Schueller (2013), Johnson et al. (2014)). The species in focus is most likely kept below its unfished level due to fishing mortality, so the steepness effect for biomass greater than unfished biomass is irrelevant. However, the meaning of unfished biomass and recruitment perhaps should be reconsidered.

The purpose of the SSR model for fisheries models is to allow for dynamics at the larval stage that are not otherwise captured in the model. In reality, while there is

evidence of SSR relationships (e.g. Ye (2000), Wahle (2003)), recruitment between years can also be highly variable due to environmental variability (Haddon, 2001). Hence, the SSR used in a stochastic model such as Stock Synthesis (Methot Jr and Wetzel, 2013) or CASAL (Bull et al., 2012) would provided the *expected* recruitment, not the *actual* recruitment. The expected recruitment resulting from the SSR would then be combined with yearly recruitment deviates that allow for the recruitment each year to be higher or lower than the expected recruitment. Hence, for a stock assessment model, the cause of variation in recruitment is not pursued, but the effect is allowed for. There is potential with ecosystem models to model the variability in recruitment by linking it to the dynamics of the model. For example, an Atlantis model can include temperature effects on spawning success, and can also modify the actual recruitment each year based on food availability through the condition of animals in the spawning population (Audzijonyte et al., 2017a). The availability of these dynamics in ecosystem models could be useful for assisting ecological sampling design techniques aimed at understanding SSR and associated variability.

While some ecosystem models (e.g. OSMOSE, Shin and Cury (2004)) do not require a SSR curve, as the relationship is emergent due to the explicit handling of larval dynamics in the model, these models are in the minority. Ecopath with Ecosim models require a recruitment function to relate recruitment to the adult numbers, biomass and food consumption (Blanchard et al., 2002), although some of the density dependent dynamics dealt with through a spawning stock recruitment relationship that are also modelled more explicitly. These explicit dynamics include changes in adult feeding rates and fecundity; changes in duration of the juvenile stage and hence in total time exposed to relatively high predation risk; and changes in juvenile foraging time (and hence exposure to predation risk) with changes in juvenile feeding rates (Christensen and Walters, 2004; Walters et al., 2000). Many size-based models apply a spawning stock recruitment relationship to account for the density dependent dynamics (Andersen et al., 2016). Gadget has two options for recruitment: either a number of recruits are added each year, and the parameters determining the

numbers added each year are obtained through optimisation (statistically fitting to data); or a spawning stock recruitment relationship is applied (Begley and Howell, 2004). MICE models such as CASAL2 are closely related to fisheries models and apply a spawning stock recruitment relationship (e.g. Doonan et al. (2016)).

Models sufficiently complex to have emergent relationships will not always be appropriate, particularly when smaller or minimum realistic models are desired—for example for rapid or tactical assessments. There are strong arguments for using the simplest model sufficient for its purpose, and this extends even to ecosystem modelling and has seen the growth in popularity of approaches such as MICE (Models of Intermediate Complexity for Ecosystem assessments (Plagányi et al., 2014)). This objective means that for the foreseeable future it is likely we will remain in need of SSR models with consistent dynamics under varying natural mortality and recruitment resilience. I propose one such variant here, but SSRs should continue to be a topic of discussion amongst ecosystem modellers and empiricists going forward.

Chapter 4

Chaotic or stable? Exploring sensitivity to initial conditions of an end-to-end ecosystem model

4.1 Abstract

Ecosystem models require the specification of initial conditions, and these initial conditions have some level of uncertainty. It is important to allow for uncertainty when presenting model results, because it reduces the risk of errant or non-representative results. It is crucial that model results are presented as an envelope of what is likely, rather than presenting only one instance. I perturbed the initial conditions of the Chatham Rise Atlantis model developed in this thesis. I found the model dynamics were not chaotic, as trajectories generally converged, but some species groups were more sensitive to initial conditions than others. I fitted Generalised Linear Models to the sensitivity of species groups to initial conditions, offering a range of possible explanatory variables. I found trophic level explained much of the contrast, with lower trophic level species groups generally more sensitive. I recommend that in any set of scenarios explored using this model that associated uncertainty analysis include perturbations of the initial conditions, with greater changes applied to species groups that were least well informed in model development.

4.2 Introduction

The Chatham Rise is perhaps New Zealand’s most productive fishing ground, and fishing of many fish species and some invertebrate species began from the mid-1970s (Ministry for Primary Industries, 2017). These activities, and the notoriety of deep marine systems as sensitive or vulnerable to disturbance (Clark and Rowden, 2009; Norse et al., 2012) makes the area an interesting one in terms of defining effective ecosystem based management. To assist this goal, an end-to-end ecosystem model has been developed. One of the challenges in developing an end-to-end ecosystem model is specifying the initial conditions. In the Chatham Rise Atlantis Model (McGregor et al., 2019b), the initial conditions were specified to reflect the ecosystem in its unfished, or virgin state. We have varying levels of understanding of the components of this ecosystem in its unfished state, and as such, there are varying levels of confidence around the estimates for the initial conditions, with all components having some level of error. Hence, an important step towards understanding the dynamics and implications of this model is to explore its sensitivity to changes in the initial conditions. If we change the initial conditions slightly, does the model produce very different results - i.e., is it chaotic? Or, if it is not chaotic, are some aspects of the model more stable than others?

The effects of uncertainty in the initial conditions seems to have received little attention in the development of end-to-end ecosystem models to date. In a review paper, Payne et al. (2015) found generally marine ecosystem models have not explicitly addressed uncertainty of initialisation, and more recently, Hansen et al. (2019) noted it is not something that has been done for Atlantis models. Payne et al. (2015) speculated as to the likely effects of initialisation uncertainty in end-to-end models such as Atlantis, noting long-lived species might dampen effects, and short-lived

species may amplify them.

Stability, chaos, and the importance of initial conditions do, however, feature within ecological theory. There have been studies looking at what characteristics of an ecosystem are linked with stability, both from a theoretical perspective, and from observation. May (1972) showed mathematically, that while complex and diverse systems can be stable, this does not hold true if either there are a large number of connections or if the connections are too strong. However, there seem to be exceptions to this rule, as later studies have shown. Roberts (1974) argued that most systems in practice appear to be more stable with more connections —contrary to the mathematical analysis of May (1972). Roberts (1974) showed if only feasible solutions are included in the analyses, such that no species may have a negative population, larger systems are actually more stable. May (1975) examined non-linear difference equations with respect to chaotic, cyclic and stable biological dynamics. Other aspects subsequently shown to increase stability of ecosystems include negative pairwise correlations (Tang and Allesina, 2014), species dispersal (Allesina and Tang, 2012), modularity (subsets of closely connected components) (Grilli et al., 2016), predator-prey relationships (Tregonning and Roberts, 1979), a high proportion of weak interactions (Olsen et al., 2016, May, 1972), and spatial structure (Fulton, 2001).

The Chatham Rise Atlantis model (McGregor et al., 2019b) is spatially defined, although at a fairly coarse scale, using 24 dynamic polygons, and 5 water column depth layers. Species are modelled using 55 species functional groups, which include species of bacteria, detritus, phyto-plankton, invertebrates, fish, sharks, cetaceans and birds. Some species groups were modelled as biomass pools, and some with age-structure, using numbers-at-age and mean weight-at-age. For many of the species, we have estimates of biomass, growth rates, age of maturity, natural mortality, spatial distributions, and diets, although some species have more knowledge gaps than others. McGregor et al. (2019b) characterised the species functional groups by key-

stoneness, responsiveness, and informance, and these attributes may relate to stability. Keystoneness and responsiveness were quantified using model simulations perturbing a single species functional group at a time, and analysing the flow-on effects to the rest of the system. Keystoneness measures the effect changes in biomass of a species group has on the rest of the system; responsiveness measures how responsive a species group is to changes in biomass of other species groups within the system. Informance was a qualitative measure used to reflect both how well informed each species group was, and how well it performed in the model —were key dynamics such as growth, mortality rates, diets, and responses to fishing, all realistic based on what we ‘know’? While these analyses were carried out to provide insight into the model’s strengths and weaknesses, and highlight which gaps are likely to be most influential in model results, they are also useful for further analyses such as carried out in this study —for perturbing initial conditions in a meaningful way based on likely uncertainties, and to add further context when analysing the results.

This study goes beyond the question of whether the Chatham Rise Atlantis model is unstable, and analyses what features of the Chatham Rise ecosystem, and how I have modelled it, affect this answer. Stability for the purpose of this study is with respect to sensitivity of the model to its initial conditions, rather than stability of the models dynamics with respect to time. I discuss the likely impacts of my findings for future use of this ecosystem model. I highlight areas of potential future research with respect to model development, and to support decisions relating to the sustainable use of the Chatham Rise marine ecosystem resources.

4.3 Methods

The analyses presented here were carried out in three main sections: 1.) the initial conditions of the model were perturbed, and the resulting model simulations were compared; 2.) components of the modelled system were characterised with respect to

attributes that may affect stability; and 3.) correlations between model component attributes and responses to perturbations of the initial conditions were analysed, thus linking the first two sections.

4.3.1 Varying initial conditions

I varied the initial conditions for the number-at-age variables of age-structured species groups, and the biomass of biomass-pool species groups. For the biomass-pool species groups, biomass is the only option to perturb; age-structured species groups could have errors in the specification of numbers-at-age and/or size-at-age, both of which affect the biomass-at-age. In a stock assessment model, size-at-age (or growth rates) are generally the same with respect to time, whereas there is likely a difference in numbers with respect to time (especially before fishing compared to after fishing). Hence, a different virgin biomass in a stock assessment model would generally be made up of a different number of fish, rather than the same number of fish but a different size. To align with this, I perturbed numbers rather than size for the age-structured species groups. The resulting number of variables to perturb was 361, of which 341 were numbers-at-age, 18 were the nitrogen content of biomass-pool species groups, and 2 were the silicate content of biomass-pool species (diatoms and microphytobenthos). The numbers-at-age of all age-classes for a given species group were scaled by the same amount for each simulation, such that the proportions-at-age were preserved, thus preserving M (instantaneous natural mortality) in the initial conditions. My intention was to explore the model's sensitivity to its initial conditions, not uncertainty around estimated parameters. Age-structured species groups were modelled with between 2 and 10 age-classes, and were perturbed by applying one scalar for all age-classes of a given species group. This method significantly reduced the number of scalars required to 57, of which 37 were for scaling numbers-at-age for age-structured species groups.

Initially, I perturbed all initial conditions using the same scalar for all variables

within each model run. The scalars I used were 0.5, 0.8, 0.9, 0.95, 1.05, 1.1, 1.2, 1.5. These scalars were chosen to cover a range from slight (+/- 5%) to extreme (+/- 50%) errors in the initial conditions.

Because shifting all initial conditions by the same amount may not give an indication as to how robust or sensitive the model is to mis-specification of the initial conditions where changes could vary in direction and magnitude, I next simulated multiple model runs, with the initial conditions scaled with some random variability. I scaled the initial conditions of each variable, sampling the scalar for each from a normal distribution, $N(0, \sigma)$ with σ chosen based on how large I assumed a plausible change could be.

In total, I ran three sets of 35 simulations, and repeated each set with and without fishing.

Set 1: All up or down. All initial conditions were scaled (numbers for age-structured, biomass for biomass-pool) with the same scalar for each run;
scalars $\in \{0.5, 0.8, 0.9, 0.95, 1.05, 1.1, 1.2, 1.5\}$

Set 2: High uncertainty. All initial conditions were scaled (numbers for age-structured, biomass for biomass-pool), with the scalars sampled from normal distributions with $\mu = 0$ and σ set based on the informance ratings defined in McGregor et al. (2019b) (Figure 4.1). Biomass-pool species groups were assumed poorly specified as these were not ranked in McGregor et al. (2019b), but would likely come out as poorly specified if they had been assessed more formally.

Set 3: High keystone species. These runs only scaled the initial conditions of species groups likely to be most influential on the system. The species groups that ranked in the top 10 for keystoneity in McGregor et al. (2019b), and all biomass pool species groups were scaled using normally distributed scalars sampled with $\mu = 0$ and $\sigma = 0.25$, giving 95% confidence intervals of $\approx +/ - 0.5$. All other species groups were unchanged (Figure 2.9).

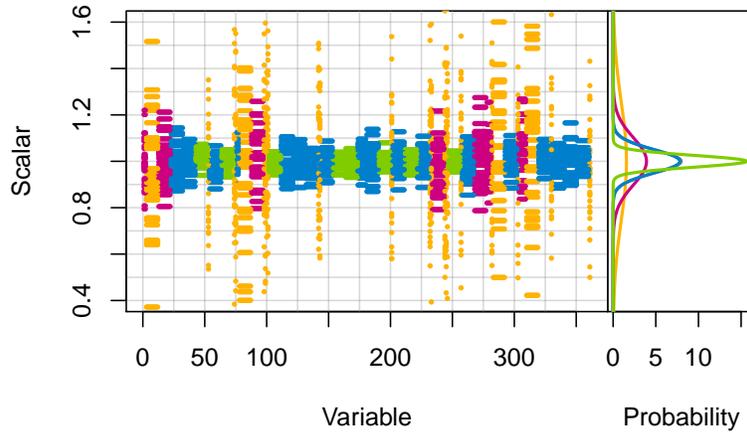


Figure 4.1: Set 2 scalars used to perturb initial conditions, with scalars sampled from Normal distributions with $\mu = 0$ and $\sigma \in (0.025, 0.05, 0.1, 0.25)$ based on informance levels 1–4 respectively where 1: ‘Poorly specified’ (gold); 2: ‘Some data gaps and/or poor performance’ (magenta); 3: ‘Slight data gaps and/or poor performance’ (blue); 4: ‘No data gaps, performed well, abundance index available’ (green) (defined in McGregor et al. (2019b)).

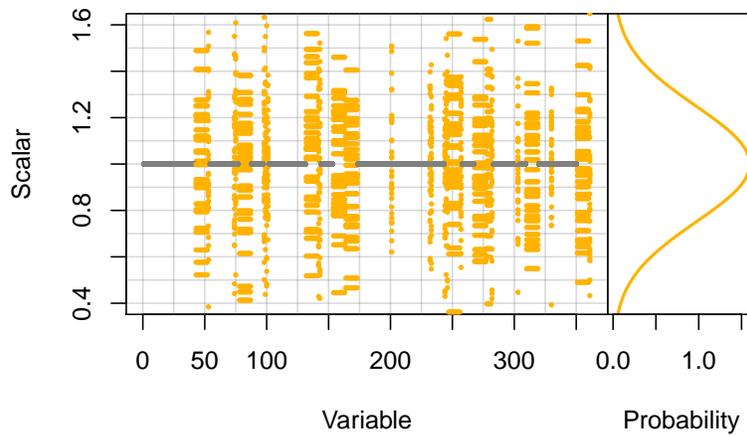


Figure 4.2: Set 1 scalars used to perturb initial conditions for high keystone and biomass-pool species groups, with scalars sampled from the Normal distribution with $\mu = 0$ and $\sigma = 0.25$.

I compared each model run with the base run at each year, and tested for correlation between the initial condition scalars and the amount each model run differed from the base model. At each year, I used a pairwise comparison at the level of model

run and species group. I calculated Pearson’s correlation and Spearman’s rank correlation to test for linearity in the relationship. If Spearman’s rank correlation is high, and higher than Pearson’s correlation, the relationship is monotonic, but not linear.

4.3.2 Characterising the system and its components

I calculated a subset of the ecosystem indicators analysed for the base model in McGregor et al. (2019a) (Table 4.1) at each timestep for all model simulations. Mean trophic level, diversity and the ratio of pelagic to total biomass were chosen as they responded to fishing scenarios for the Chatham Rise Atlantis model (McGregor et al., 2019a), but didn’t require catch for the calculation (so I could apply them to model runs here with or without fishing included). I analysed the results for consistencies or discrepancies in shifts of the ecosystem reflected through these indicators, with particular focus on the response of the system when heavy fishing became established during the mid-1970s.

Table 4.1: Key ecosystem indicators evaluated for responses to perturbing the initial conditions.

Indicator	References
Mean trophic level	Pauly and Watson (2005); Shin et al. (2018)
Diversity (modified Kempton’s Q)	Ainsworth and Pitcher (2006)
Biomass of pelagic fishes/biomass total	Link (2005)

If certain species groups appear to be more stable than others, I wanted to be in a position to investigate whether the more stable species groups have shared characteristics —for example, are there links between sensitivities to changes in the initial conditions and how connected each species group is in the system, how abundant they are, or how long they live, or some combination of these.

To do this, I characterised species groups based on the base model so I could test for links between these attributes of the species groups and sensitivities to changes

in the initial conditions. I considered keystone-ness, trophic level, biomass, animal size, lifespan, background mortality, number of trophic connections, and proportion of most dominant ('top') prey. All but the final three of these indices were available from McGregor et al. (2019b). The proportion of diet made up by most dominant prey, number of trophic connections, and the proportion of natural mortality that was made up of additional mortality were calculated for this study using R version 3.4.3.

Proportion of top prey

For each species group, I calculated the proportion of its diet that was made up by its top prey species group based on biomass consumed from the base model. This calculation was to classify the extent to which each species group was eating as a specialist or generalist as they are modelled. It is possible for a predator to perform in the model as more of a specialist due to aggregation of species into groups—they could predate on several prey species that are modelled in the same species group. For each species group, I summed the prey eaten over the entire model region and all modelled years 1900–2015, then selected the largest proportion of prey to classify the species group.

Number of trophic connections

The number of primary trophic connections was calculated for each species group (these are immediate predators or prey). A trophic connection was deemed to be either a prey or predator of the species group. Because sometimes a predator may eat a very small, negligible amount of a prey, I included a cut-off at 1%, such that a prey or predator was not included in the connections count if they made up less than 1% of the total weight of prey consumed or predation pressure applied, respectively. I calculated the number of secondary connections as the number of species that are predated on or by any of the primary connections (prey or predators of the prey or predators) (Figure 4.3), and the number of tertiary connections was calculated similarly.

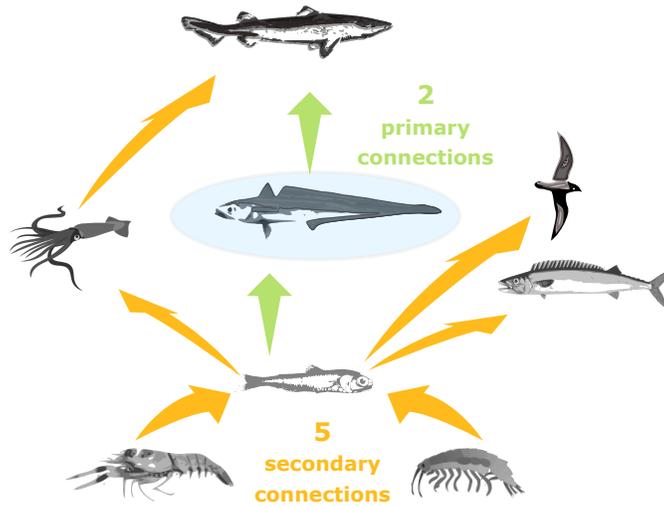


Figure 4.3: Illustration of primary (green arrows) and secondary (orange arrows) trophic connections for the species shaded blue.

Additional natural mortality

There is the option in Atlantis to apply additional natural mortality either as a quadratic term, which is density dependent, or as a linear term (Audzijonyte et al., 2017a). The balance between additional natural mortality and mortality coming from dynamics within the model may affect the model's stability. Higher levels of additional mortality reduce the strength of connections in the model, with 100% additional mortality effectively resulting in parallel single species models. Additional mortality was required for some species groups in the Chatham Rise Atlantis model that did not suffer sufficient natural mortality through predation, starvation or disease in the model to match estimates of mortality from the literature. For all age-structured species groups in this model, linear rather than quadratic mortality was applied because this is a close approximation to instantaneous natural mortality (M). When M is small, as it is when applied at small time-steps, e^{-M} can be approximated by $1 - x$ using the first two terms of its Taylor series expansion. Hence, if we take N_t to be the number of individuals at timestep t and $N_{t+\delta}$ to be the number

at timestep $t + \delta$ where δ is small we get

$$N_{t+\delta} = N_t e^{-M_\delta} \approx N_t (1 - M_\delta) \quad (4.1)$$

As linear mortality, m_L , is applied at every timestep (12 hours for this model), we can use M_δ to approximate m_L . This approximation is, however, complicated by a temperature effect which is applied to m_L in Atlantis. Additional mortality in Atlantis is assumed to be related to metabolic rates, and hence are temperature dependent. The temperature effect is applied as a scalar (T_{corr}) calculated as a function of the current water temperature (T) (in a given cell at a given time) relative to a base temperature, set at 15 °C (Equation 4.2).

$$T_{corr} = 2^{(T-15)/10} \quad (4.2)$$

As temperature varies spatially and temporally, so does the scaled m_L . I calculated the additional applied mortality for each species group based on their spatial distribution, m_L values, and temperature corrections, using the median, upper and lower quartile, and 95% confidence intervals for the applied additional mortality to reflect the variability of temperature spatially and temporally. These were calculated for both juveniles and adults as m_L and spatial distributions were defined separately for these life stages.

Total realised mortality rates were estimated from the model by fitting an exponential decay curve to the proportions-at-age. By running the model with no fishing, the realised mortality consisted entirely of natural mortality, including sources within the model such as predation, as well as additional mortality from m_L . I then compared the total realised natural mortality with the range of additional mortality to estimate what proportion of natural mortality was coming from dynamics within the model, and what proportion was forced. I produced a weighted average for each species group that combined the proportions for adults and juveniles, weighted by the numbers of adults and juveniles respectively.

4.3.3 Modelling stability

I analysed variability as a result of perturbing the initial conditions for each species group by calculating the coefficient of variation (CV) of biomass between the runs. A species group with high variability between model simulations relative to the mean biomass from all simulations would have a high CV, and hence be considered highly sensitive to perturbations to the initial conditions.

I analysed the effects of perturbing the initial conditions by fitting a GLM (Generalised Linear Model) to the CV for the biomass of each species group across model runs. I used Atlantis model outputs following a 35-year burn-in period, to match the burn-in used in McGregor et al. (2019b). Variables from characterising the species groups (Table 4.2) were offered as possible explanatory variables, using a step-wise selection algorithm, with each iteration selecting the variable (or pair of interaction variables) that explained the largest proportion of the null deviance. This process was repeated until the additional deviance explained was less than 10%. This cut-off value was selected to limit the number of explanatory variables selected, while retaining most of the explained null deviance. I initially explored untransformed, and log (base 10) and cubed root transformations of the response variable (CV), with all modelled using the Gaussian distribution. The analyses presented here used the cubed root transformation as I found this produced greater homogeneity of residuals with respect to the fitted values.

I could not model the biomass-pool species group CVs with respect to all attributes, as some attributes had not been analysed for biomass-pool groups (e.g. Keystone and Response), and some attributes relate to individuals, such as maximum size and instantaneous mortality. Hence, I fitted three versions of the GLM: 1.) limited the species groups included in the analyses to species with age-structure in order to consider the full list of explanatory variables; 2.) retained all species groups, but limited the explanatory variables offered; 3.) limited the species groups to biomass-pool species groups, with the limited the explanatory variables offered.

Table 4.2 gives the full list of explanatory variables offered for biomass-pool (BP), age-structured (AS), and all-species (ALL) versions of the model. All possible paired interaction terms were also offered. PropByTopPrey was dropped from BP models as nearly half (8/17) of the biomass-pool species groups were not predators, and this variable only applies to predators.

I fitted the GLM to model outputs for each year (1900–2015) to test for temporal shifts in the effects (a separate GLM was fitted at each year). To allow for influence from the method of perturbing the initial conditions (all up or down, based on keystone-ness, or based on uncertainty), I included this (‘ChaosAlt’) as a potential explanatory variable. I also explored splitting out the fished model runs from the unfished, or including this within ChaosAlt (Table 4.3).

I fitted a summary GLM for each version (ALL, AS, BP species; with/without fishing included), using a subset of the years simulated by the models, where the explanatory variables selected for models fitted at each timestep were roughly consistent. I used these summary models to explore the effects of the selected explanatory variables on between-run CVs. I analysed the residuals to check for trends or biases in the model fits, and present these as well as the effects of selected explanatory variables.

Table 4.2: Explanatory variables offered to explain effects of perturbing the initial conditions, and whether these were defined for age-structured species groups or biomass-pool species groups, or all species groups. BP= biomass-pool species groups; ALL= all species groups; AS = age-structured species groups.

Variable	Model BP	Model ALL	Model AS	Type
1. Informance			✓	Categorical
2. TL	✓	✓	✓	Continuous
3. Keystone			✓	Categorical
4. Response			✓	Categorical
5. NumL1cons	✓	✓	✓	Categorical
6. Lifespan	✓	✓	✓	Categorical
7. propAdM			✓	Continuous
8. propJuvM			✓	Continuous
9. B0	✓	✓	✓	Continuous
10. PropByTopPrey		✓	✓	Continuous
11. Linf			✓	Continuous
12. ChaosAlt	✓	✓	✓	Categorical

Table 4.3: ChaosAlt definitions for perturbing the initial conditions, and including fishing in the model or not. ChaosAlt was offered as an explanatory variable to the GLMs.

ChaosAlt	Description	Included fishing
A	All up or down	
B	All up or down	✓
C	Based on uncertainty	
D	Based on uncertainty	✓
E	Based on keystoneity	
F	Based on keystoneity	✓

4.4 Results

4.4.1 Variability from initial conditions

Some species groups diverged while others converged, with biomass-pool groups more likely to have persistent high CVs between model runs (Figure 4.4). Age-structured groups typically had very tight CVs of less than 10% from 1970 onwards, particularly with fishing turned on (Figure 4.4). Fishing reduced the between-model CVs for some age-structured species groups, such as for hoki, but the effects of fishing were not ap-

parent in any biomass-pool species groups—in biomass trajectories or CVs between model runs (Figure 4.5 for two examples; Appendix A for the full set of figures.). Responses of age-structured species groups to fishing were generally consistent across model runs. These responses included direct effects of fishing on a species (such as hoki (*Macruronus novaezelandiae*), hake (*Macruronus novaezelandiae*), orange roughy (*Hoplostethus atlanticus*) and ling (*Genypterus blacodes*)), and predation-release responses (such as cephalopods and pelagic fish). Exceptions were invert comm herb (primarily paua (*Haliotis iris*) and kina (*Evechinus chloroticus*)), invert comm scav (primarily scampi (*Metanephrops challengeri*)), dem fish pisc (primarily giant stargazer (*Kathetostoma giganteum*)), and seaperch (*Helicolenus spp.*), which all gave varied responses with fishing included in the model.

The scalar applied to the initial conditions remained correlated to the difference in the perturbed model to the base model at about 60% (Spearman’s) and 45% (Pearson’s) for the not-fished species groups through to the completion of the model runs (Figure 4.6). The difference between the Spearman’s and Pearson’s correlations suggests that while the relationship remained fairly monotonic, it was not linear. The fished species groups retained very little correlation to the initial condition scalars, with Pearson’s correlation less than 10% 25 years into the model (1890), and Spearman’s correlation less than 10% by 1915 (Figure 4.6).

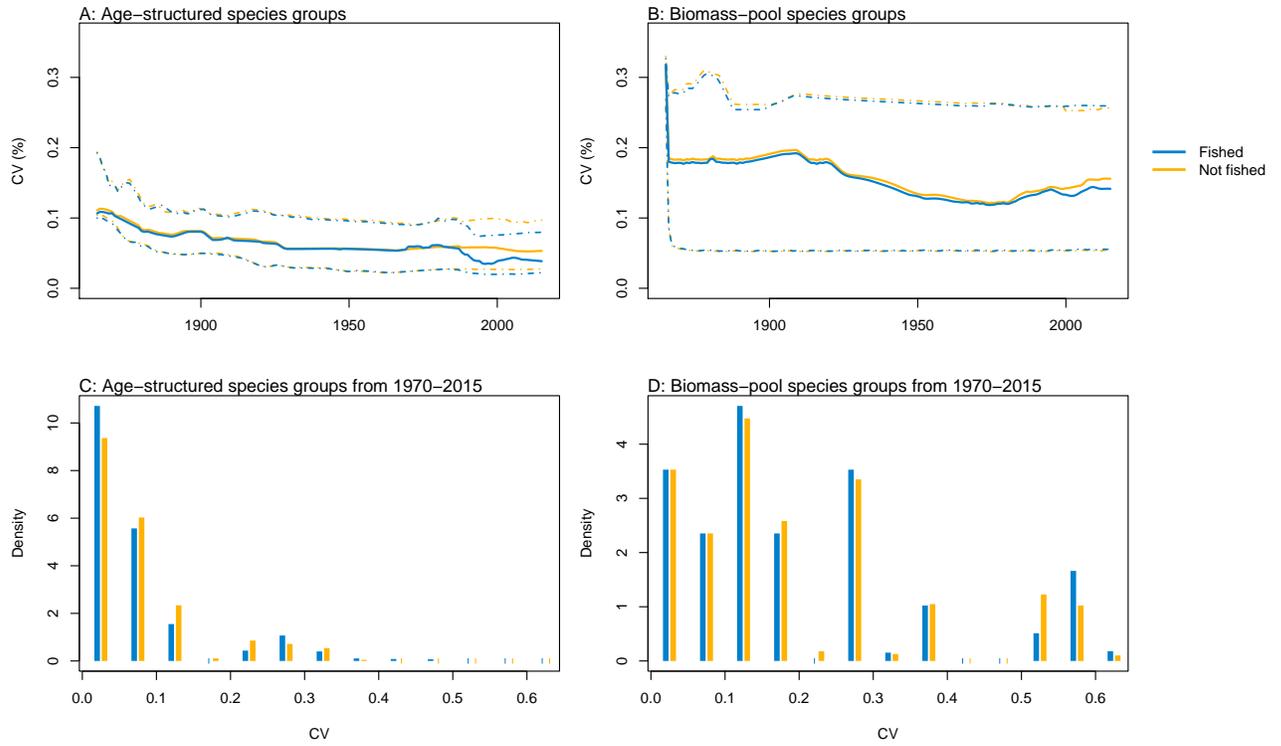


Figure 4.4: Median (solid lines), and upper and lower quartiles (dot-dashed lines) for CVs of age-structured species groups (A) and biomass-pool species groups (B) from fished model runs (blue) and unfished model runs (orange). Density of CVs from 1970–2015 for age-structured species groups (C) and biomass-pool species groups (D).

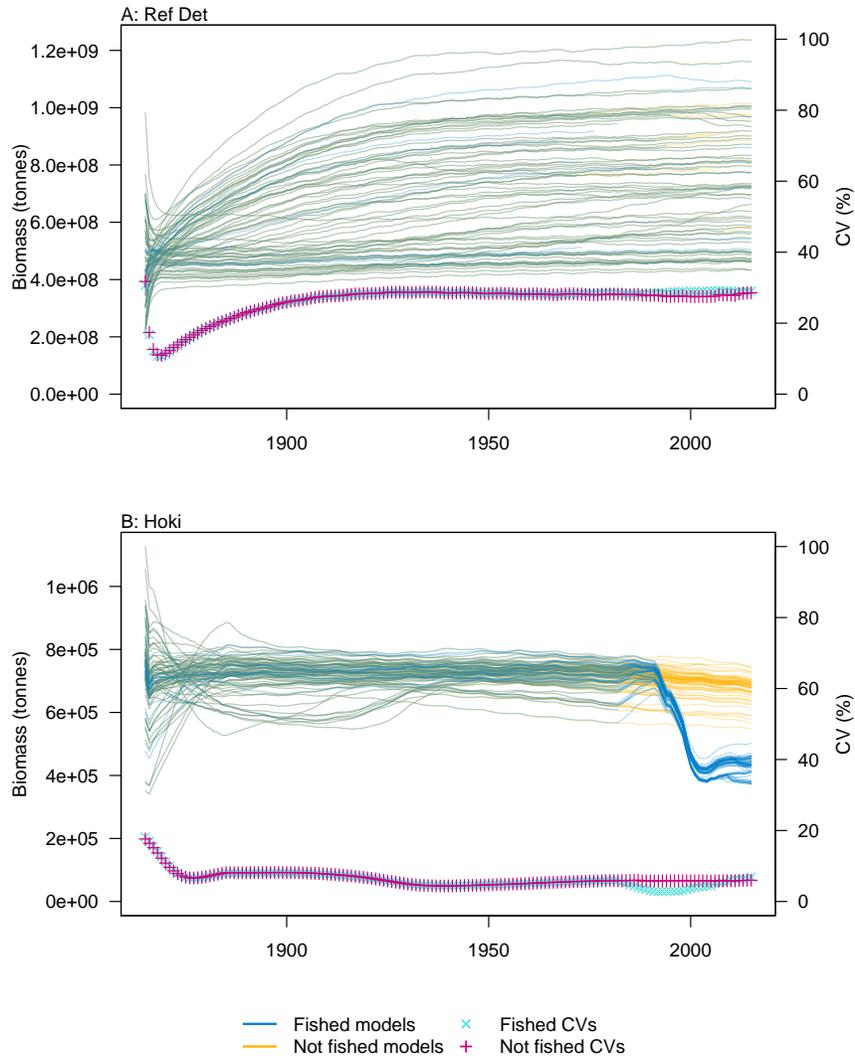


Figure 4.5: Biomass trajectories from models with fishing included (blue lines) and no fishing (orange lines) for Ref Det (refractory detritus) (A) and hoki (B), with CVs from across the model runs by time from fished models (aqua asterisks) and unfished models (cerise asterisks) overlaid and using the right-hand axis.

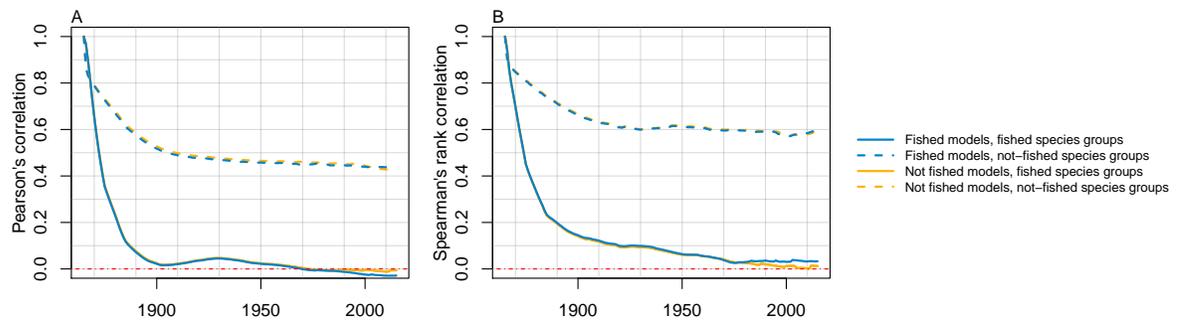


Figure 4.6: Pearson's correlation (A) and Spearman's rank correlation (B) between initial condition scalars and changes in the perturbed model with respect to the base model for fished models (blue lines) fished species groups (solid line) and not-fished species groups (dashed line) and not fished models (orange lines) fished species groups (solid line) and not-fished species groups (dashed line).

4.4.2 Characterisation

Ecological indicators

Ecological indicators demonstrated variability from the perturbed initial conditions that generally neither converged nor diverged throughout the model simulations. However, the responses to heavy fishing from the mid-1970s were consistent across runs, with a decline in mean trophic level, a slight increase in diversity, and an increase in the ratio of pelagic biomass over total biomass (Figure 4.7). There was a slight decline in mean trophic level from 1900–2015 in some of the unfished models, although the decline was approximately 0.02 of a trophic level over 100 years, so rather small.

Proportion of top prey

Some diets consisted almost entirely of one species group, but many others did not have a dominant species with the ‘top’ prey making up less than 50% of the diet, and there was quite an even spread in between, with top prey making up around 50–70% of many diets (Figure 4.8). Not all species groups predate, which is why some species groups (such as sediment bacteria, macroalgae) do not have a highest proportion of prey.

Number of trophic connections

The number of primary connections ranged from 1 through to 30, and with fairly even spread in between (Figure 4.9). Most species groups were almost fully connected by the third level, and all species groups had at least 43 tertiary connections, of the 55 available species groups. Many of the species groups had more than 20 secondary connections, and those with fewer secondary connections generally had fewer primary connections. The number of secondary and tertiary connections are unlikely to be informative for stability between runs as there is little contrast.

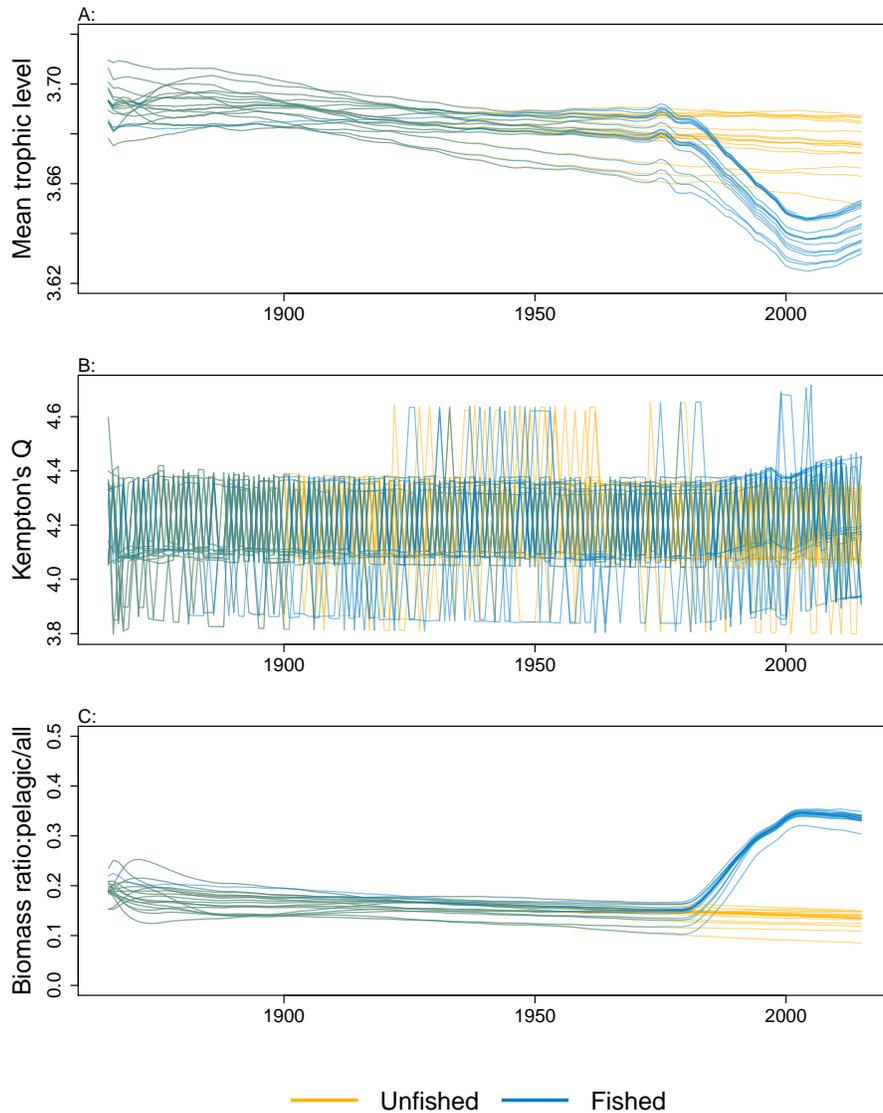


Figure 4.7: Ecological indicators, mean trophic level of age-structured species groups (A), Kempton's Q (B), and biomass ratio of pelagic fishes/all age-structured species groups (C) calculated from model simulations with fishing included (blue lines), and no fishing included (orange lines) from model years 1865–2015, which includes the burn-in period of 1865–1900.

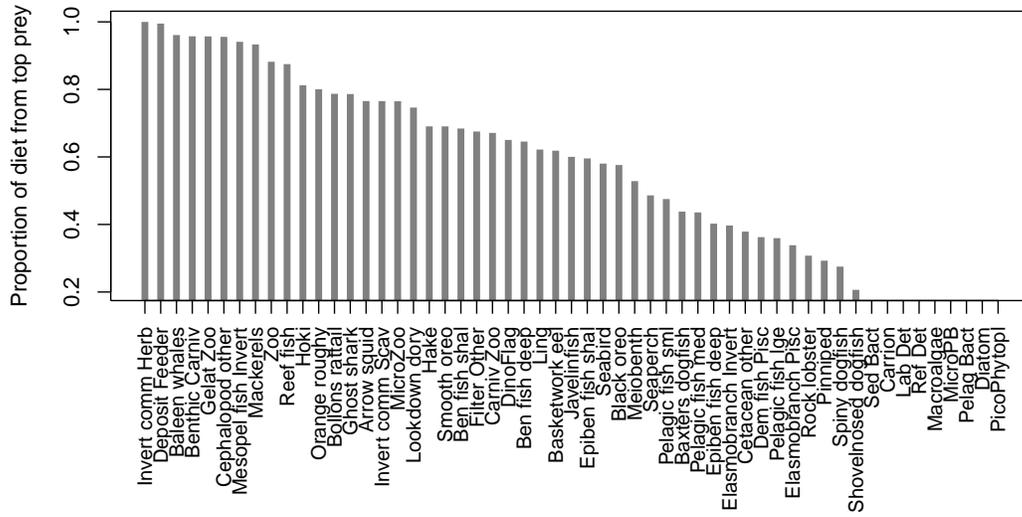


Figure 4.8: Proportion of diet made up by top prey from the Chatham Rise Atlantis model (base) 1900–2015 model outputs.

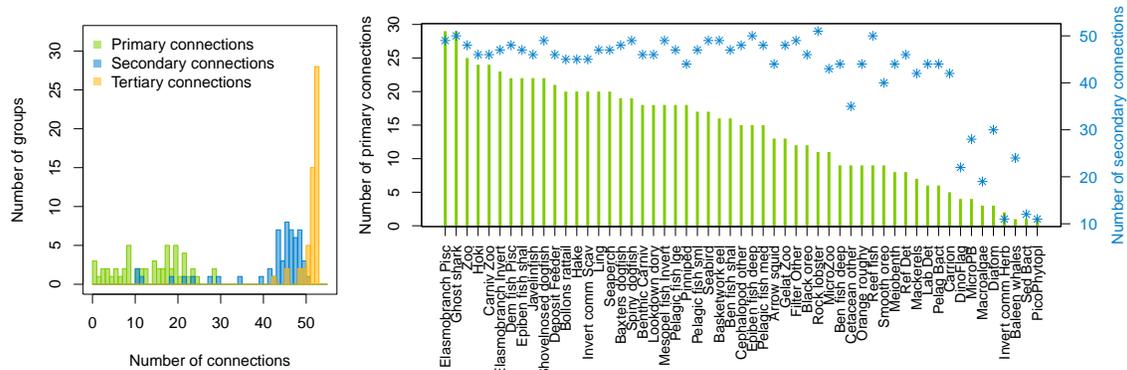


Figure 4.9: Number of trophic level connections by species group for the Chatham Rise Atlantis model (base) 1900–2015 model outputs. A: Number of groups (frequency counts) by primary connections (green bars), secondary connections (blue bars), and tertiary connections (orange bars); B: Number of primary connections by species group (green bars), and number of secondary connections by species group (blue asterisks, and using right-hand axis).

4.4.3 Additional natural mortality

The proportion of natural mortality forced with additional mortality through the mL term ranged from just over 0.8 for spiny dogfish down to zero for several species

(Figure 4.10). While baleen whales, cetacean other, pinnipeds and seabirds all have zero additional mortality through mL , this does not mean their populations are entirely constrained due to mortality within the model, as these groups all migrate out of the model and their populations are restrained on re-entry into the model domain. Pelagic fish small (primarily myctophids), arrow squid, cephalopods other, and invert comm scav (primarily scampi) have all their natural mortality from sources such as predation within the model. Just over half (19/37) of the age-structured species groups had more than 80% of natural mortality forced as adults. Most age-structured species groups (31/37) had less than 50% of natural mortality forced for juveniles.

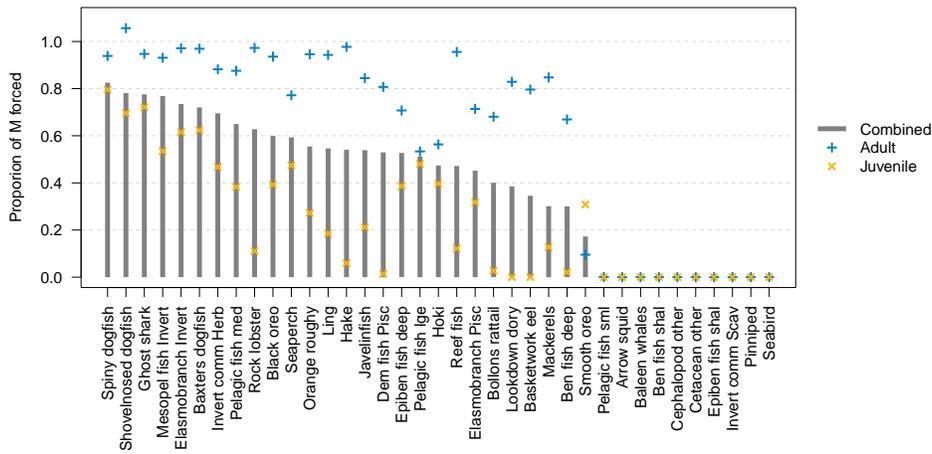


Figure 4.10: Proportion of natural mortality (M) forced as additional mortality by species group from the Chatham Rise Atlantis model (base) 1900–2015 model outputs.

4.4.4 Modelling stability

GLMs fitted at each timestep

The models fitted at each timestep (year) with all species groups combined (ALL) selected the interaction term ChaosAlt (the way in which the initial conditions were perturbed) and trophic level, and explained just under 50% of the null deviance (Figure 4.11). The ChaosAlt:trophic level interaction term was also the most important explanatory variable for biomass-pool (BP) only species group models (Figure 4.12), and age-structured (AS) only species group models (Figure 4.13). BP models consistently selected a second term; the interaction of the number of primary trophic connections and virgin biomass (B_0) and explained between 50% and 60% of the null deviance (Figure 4.12). The AS models had different explanatory variables selected at different timesteps, and these were also influenced by whether fishing was included in the models. The interaction between ChaosAlt and the number of primary trophic connections was the most consistently selected second explanatory variable for AS models (Figure 4.13). The BP and AS models seemed to have a shift at around 1910. Explanatory variables selected prior to 1910 did not include ChaosAlt, but trophic level was important as an interaction with virgin biomass and the number of primary trophic connections for BP models, and trophic level and informance for AS models (Figures 4.12 and 4.13).

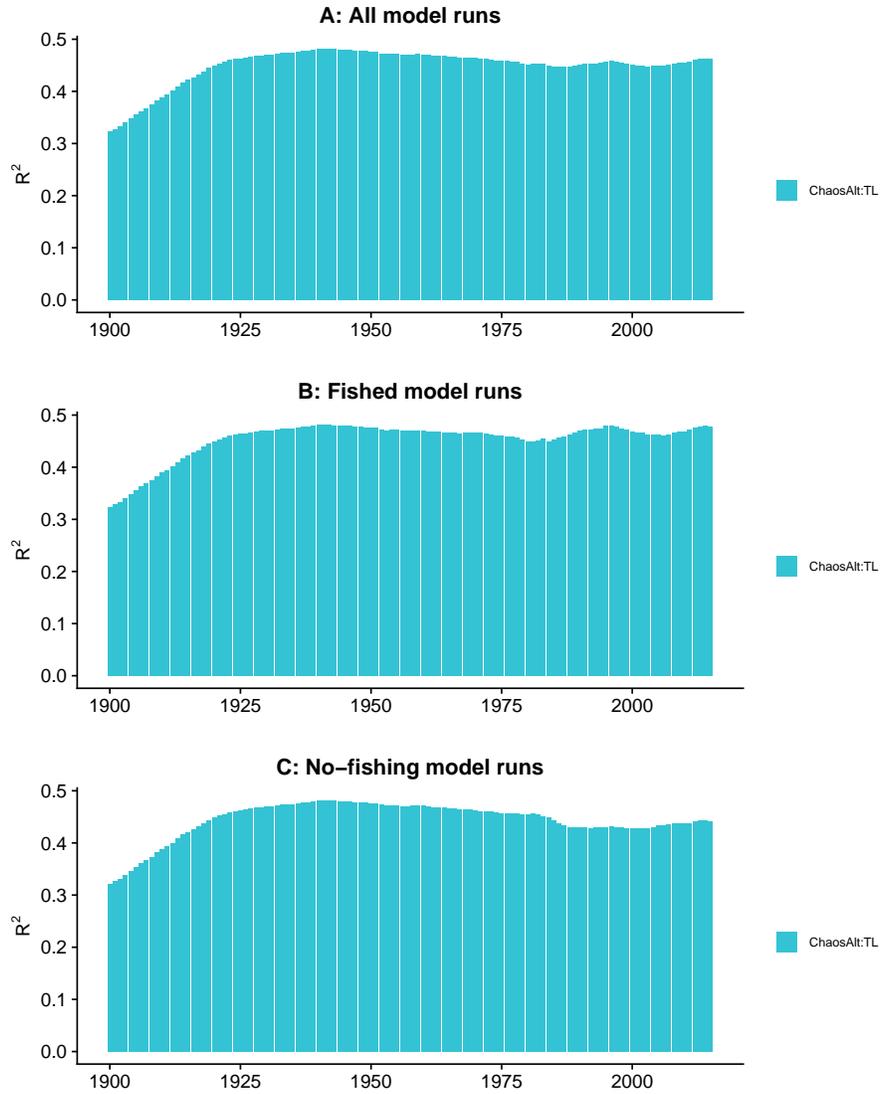


Figure 4.11: R^2 for GLMs fitted at each timestep to biomass CVs of all species groups that resulted from perturbing the initial conditions, using all model runs (A), only model runs with fishing (B), and only models without fishing (C), with bars coloured by explanatory variable.

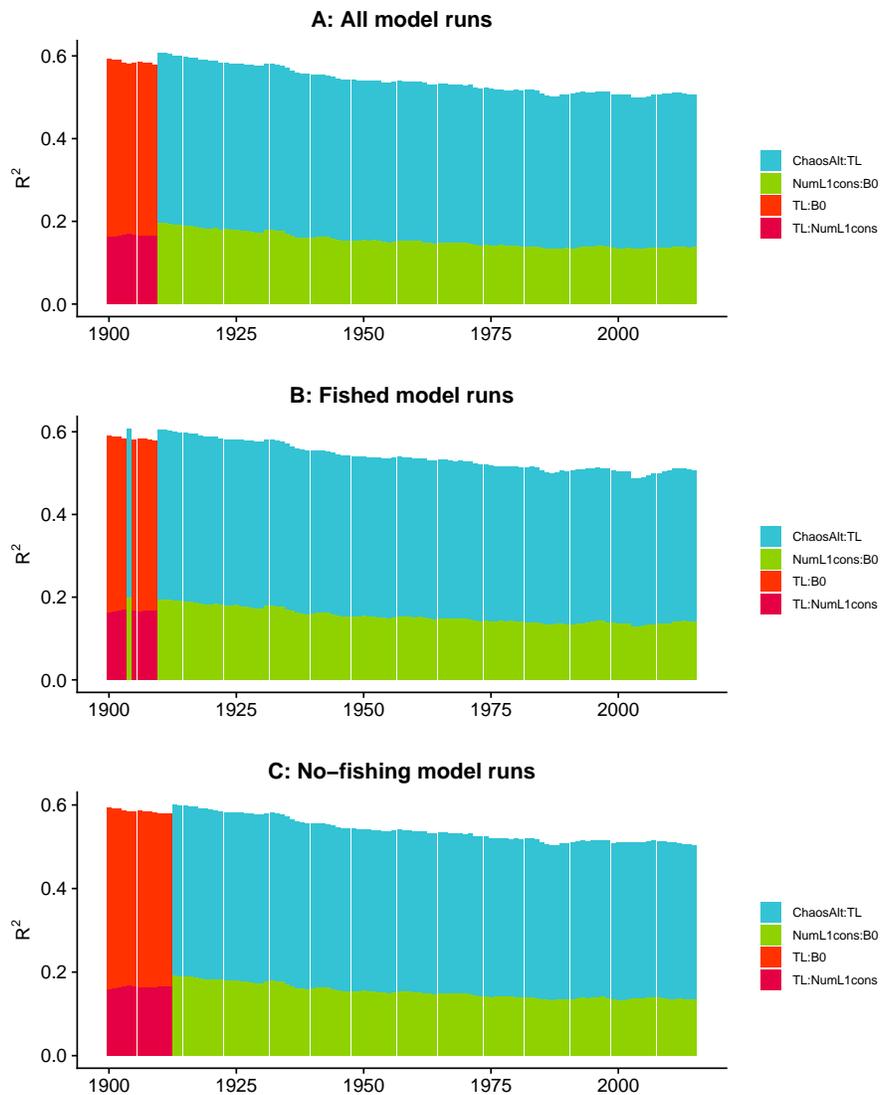


Figure 4.12: R^2 for GLMs fitted at each timestep to biomass CVs of biomass-pool (BP) species groups that resulted from perturbing the initial conditions, using all model runs (A), only model runs with fishing (B), and only models without fishing (C), with bars coloured by explanatory variable.

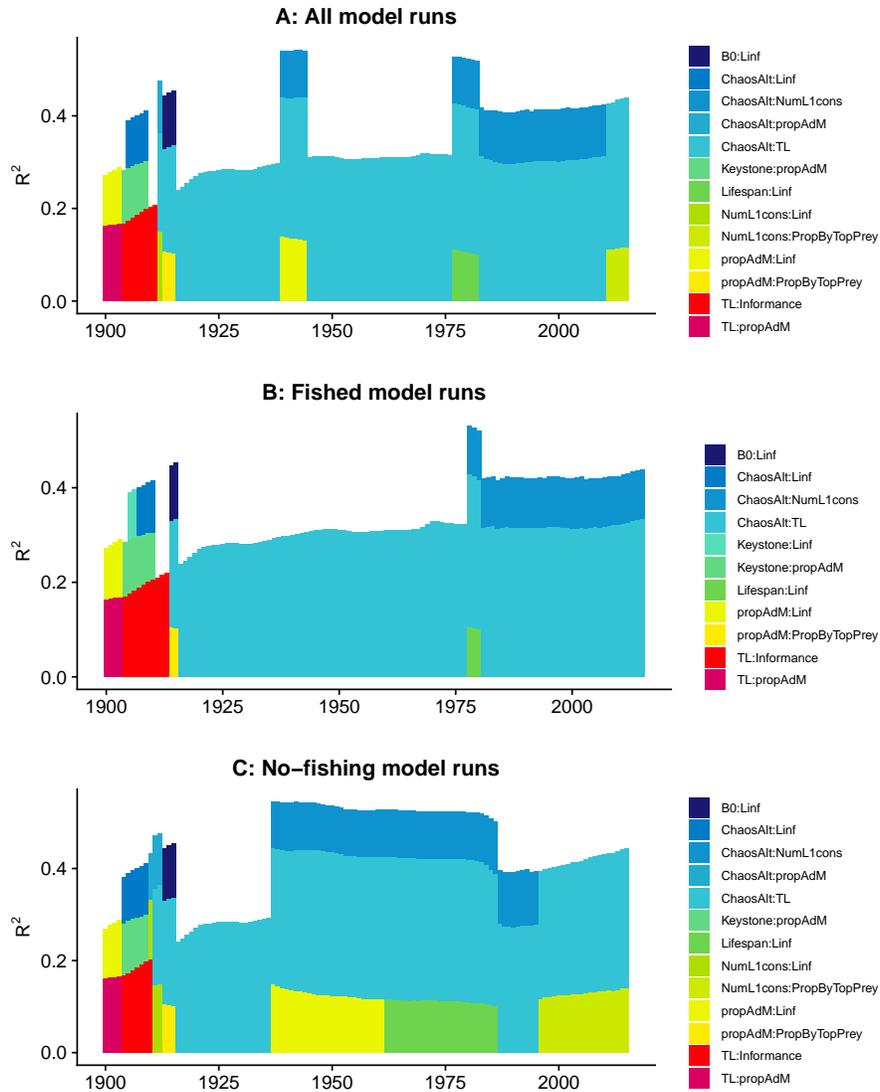


Figure 4.13: R^2 for GLMs fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using all model runs (A), only model runs with fishing (B), and only models without fishing (C), with bars coloured by explanatory variable.

Final GLMs

The GLMs fitted to all data from 1910–2015 selected similar explanatory variables to the GLMs fitted at each timestep (Table 4.4). The interaction term ChaosAlt:TL was selected first for all models, and was the only term selected for the ALL model. The BP only species model also selected the interaction term NumL1cons: B_0 . The AS species only model selected interaction ChaosAlt:NumL1cons whether fishing was included or not, and a third term, interaction NumL1cons:Informance was selected for the unfished AS model.

The Pearson’s residuals generally showed no concerning patterns against fitted values or explanatory variables for the final GLMs (Figures 4.14–4.18). One exception was the residuals with respect to B_0 for the BP model, which suggested decreasing errors with increasing B_0 , and a possible outlier (Figure 4.15).

Table 4.4: Explanatory variables selected and corresponding r^2 values for GLMs fitted to ALL (all species groups) model CVs, BP (biomass-pool species groups) only model CVs, and AS (age-structured species groups) only model CVs, using model outputs from 1900–2015, with fished and unfished versions for AS. ChaosAlt=the set of runs, grouped by method for perturbing initial conditions and whether fishing was included or not; TL=trophic level; PrimCons=number of primary trophic connections; B_0 =virgin biomass; Inf=informance;

Model	ChaosAlt:TL	PrimCons: B_0	ChaosAlt:PrimCons	PrimCons:Inf	Total r^2
ALL	0.47				0.47
BP	0.38	0.53			0.53
AS	0.32		0.44		0.44
AS (fished)	0.33		0.45		0.45
AS (unfished)	0.31		0.44	0.54	0.54

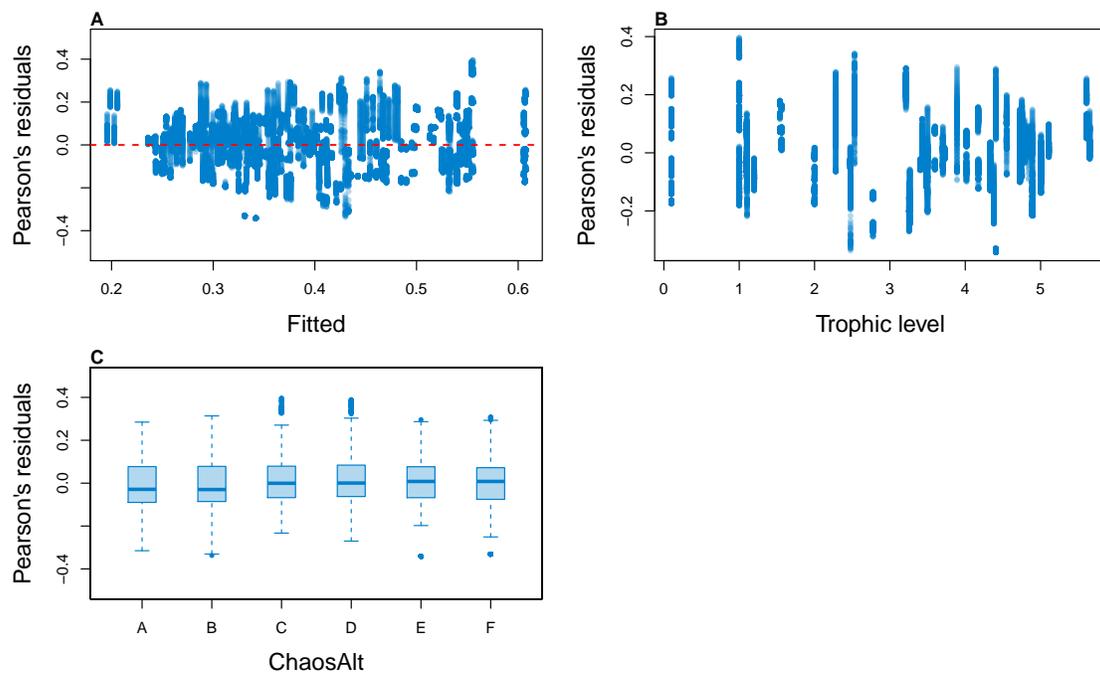


Figure 4.14: Pearson's residuals for models fitted at each timestep to biomass CVs of all (ALL) species groups that resulted from perturbing the initial conditions, using all model runs, plotted against fitted values (A), TL (trophic level) (B), and ChaosAlt (C).

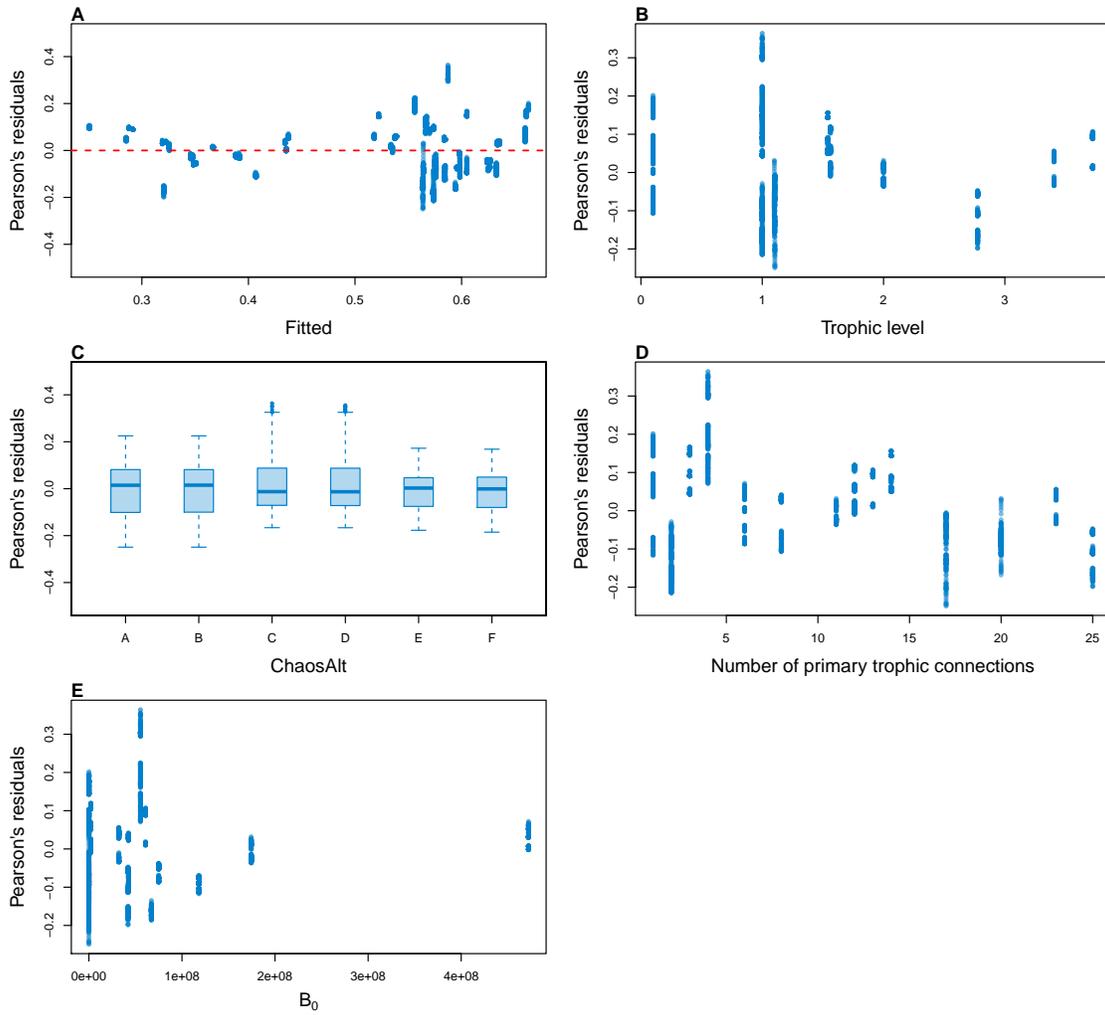


Figure 4.15: Pearson's residuals for models fitted at each timestep to biomass CVs of biomass-pool (BP) species groups that resulted from perturbing the initial conditions, using all model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), Number of primary trophic connections (D), and B_0 (virgin biomass) (E).

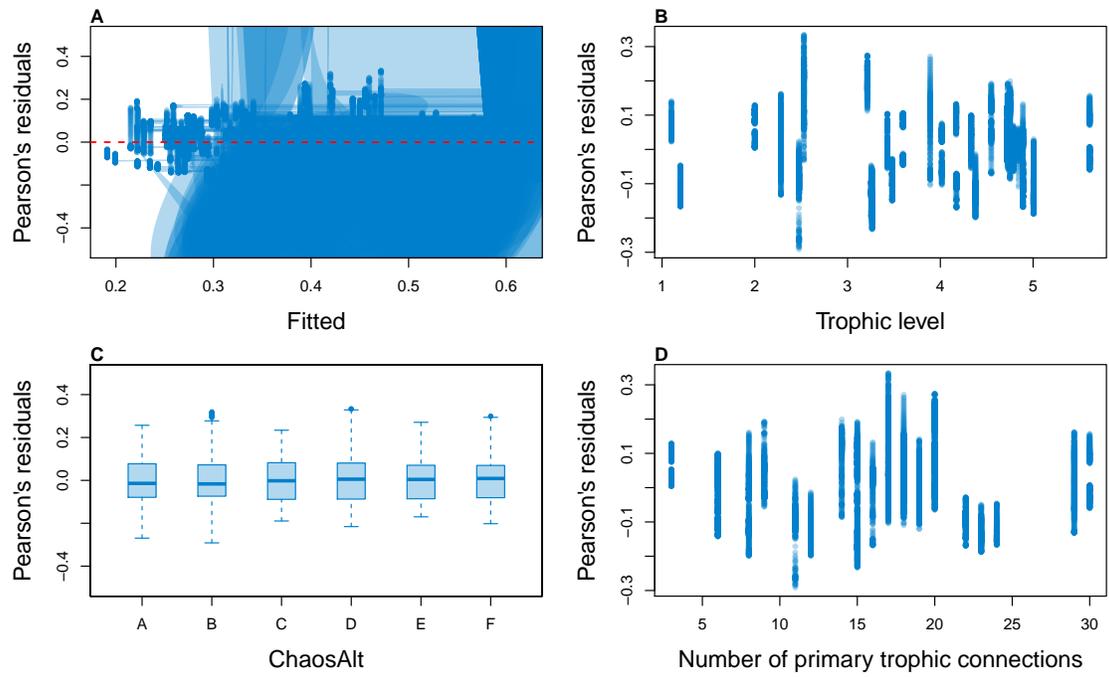


Figure 4.16: Pearson's residuals for models fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using all model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), and Number of primary trophic connections (D).

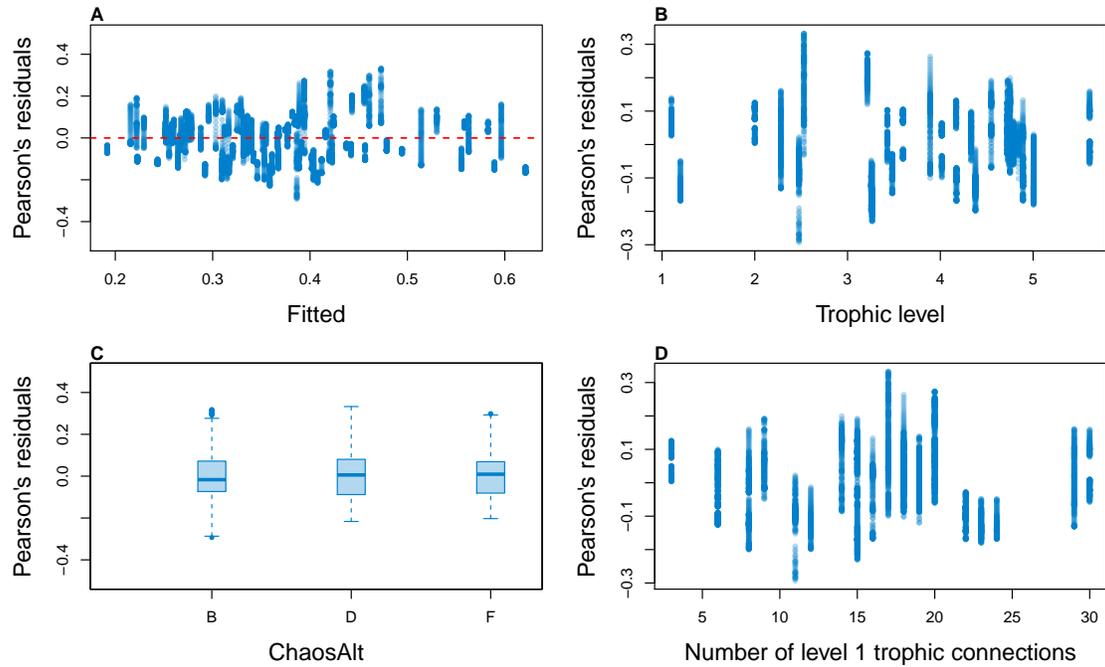


Figure 4.17: Pearson's residuals for models fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using fished model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), and Number of primary trophic connections (D).

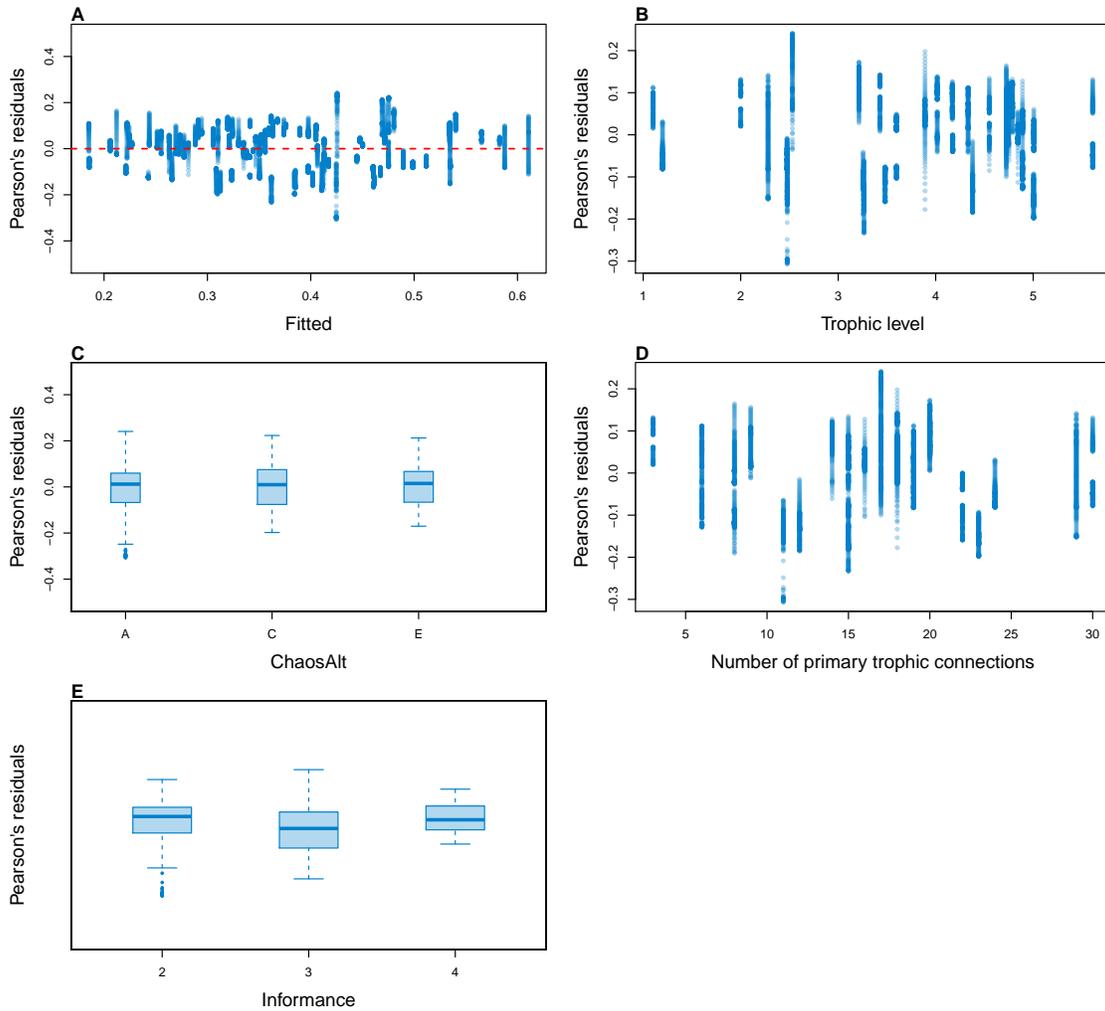


Figure 4.18: Pearson's residuals for models fitted at each timestep to biomass CVs of age-structured (AS) species groups that resulted from perturbing the initial conditions, using unfished model runs, plotted against fitted values (A), TL (trophic level) (B), ChaosAlt (C), and Number of primary trophic connections (D), and informance (E). Informance levels 1–4 where 1: 'Poorly specified' (gold); 2: 'Some data gaps and/or poor performance' (magenta); 3: 'Slight data gaps and/or poor performance' (blue); 4: 'No data gaps, performed well, abundance index available' (green) (defined in McGregor et al. (2019b)). Informance level '1' did not feature in the results as these data were dropped due to 'NA' values for other explanatory variables.

Higher trophic level was found to be associated with lower biomass CVs for all models and ChaosAlts (Figure 4.19). CVs were generally lower for ChaosAlt ‘A’ and ‘B’, which were the model runs with all initial conditions shifted up or down and by the same scalar within each run. ChaosAlt ‘C’ and ‘D’, with initial conditions perturbed based on species group uncertainty, generally had slightly higher CVs across trophic levels (Figure 4.19). This effect was also apparent in the interaction with primary connections in the AS model (Figure 4.20). Biomass CVs were found to decrease with increased B_0 and with increasing number of primary connections for biomass pool species group (Figure 4.20). The number of primary connections had the opposite effect for age-structured species groups, with more primary connections correlated with larger biomass CVs, although these CV effects were smaller (max. 11%) than for biomass pool species groups (max. 18%) (Figure 4.20).

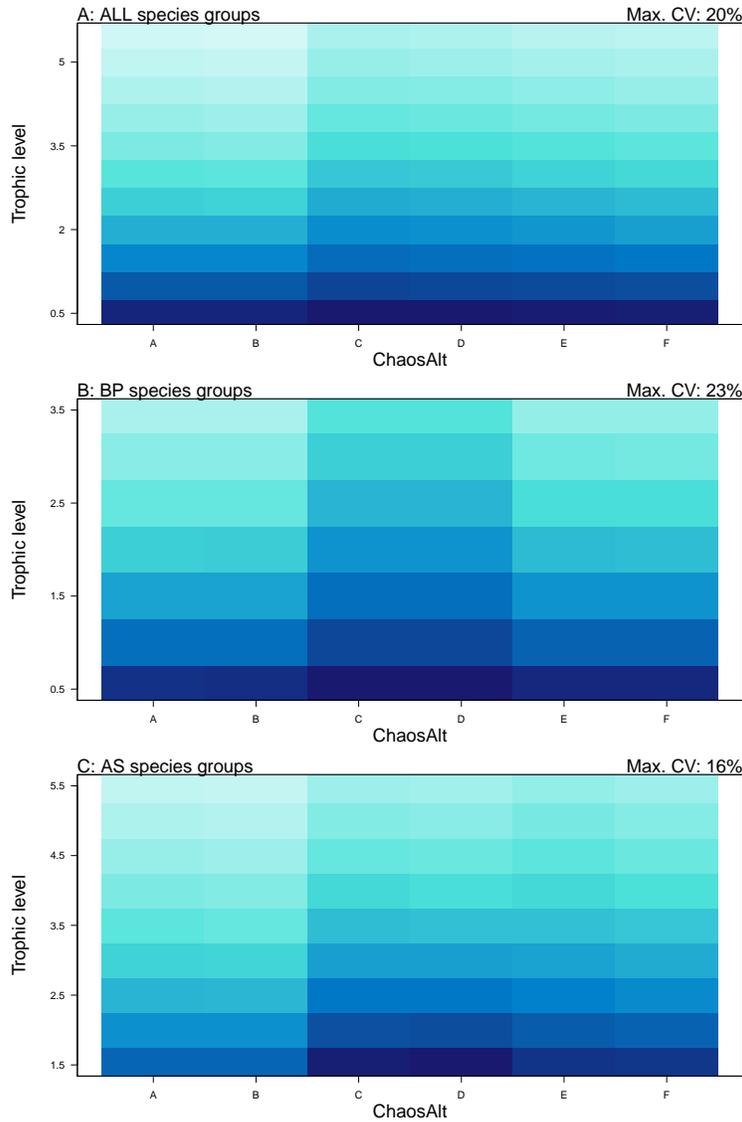


Figure 4.19: GLM effects from the final GLMs fitted to CVs from 1910–2015 for interaction term ChaosAlt:TL for ALL species groups (A), BP only species groups (B), and AS only species groups (C). ChaosAlt ‘A’ and ‘B’ perturbed all initial conditions by the same scalar for each run; ChaosAlt ‘C’ and ‘D’ perturbed initial conditions by uncertainty; ChaosAlt ‘E’ and ‘F’ perturbed initial conditions by keystone-ness; ChaosAlt ‘A’, ‘C’, ‘E’ did not include fishing; ChaosAlt ‘B’, ‘D’, ‘F’ included fishing. Shading indicates the additional CV expected for each value of the interaction, with the darkest shading in each plot corresponding to the Max. CV (%) given in the top-right corner of the plot.

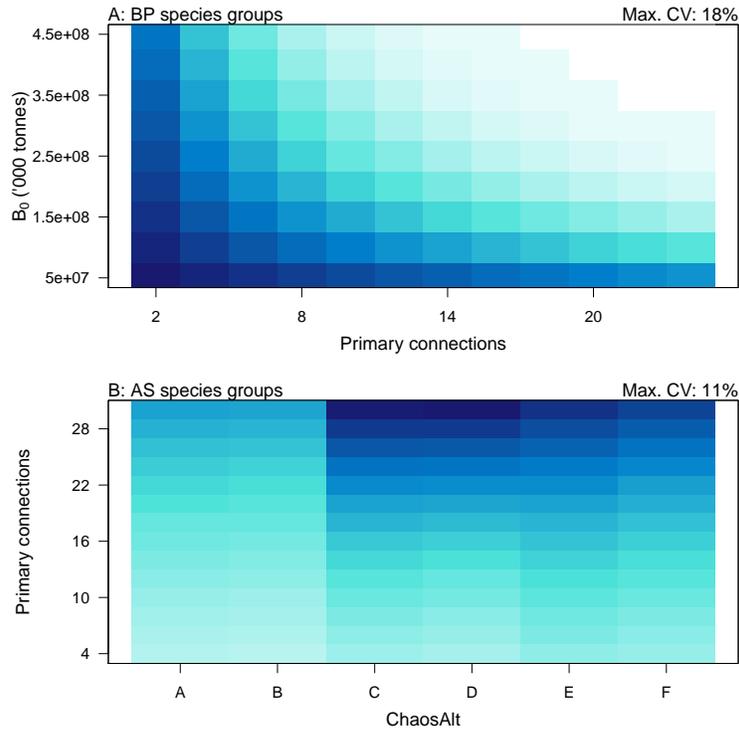


Figure 4.20: GLM effects from the final GLMs fitted to CVs from 1910–2015 for interaction term PrimCons: B_0 for BP only species groups (A), ChaosAlt:PrimCons for AS only species groups (B). ChaosAlt ‘A’ and ‘B’ perturbed all initial conditions by the same scalar for each run; ChaosAlt ‘C’ and ‘D’ perturbed initial conditions by uncertainty; ChaosAlt ‘E’ and ‘F’ perturbed initial conditions by keystone-ness; ChaosAlt ‘A’, ‘C’, ‘E’ did not include fishing; ChaosAlt ‘B’, ‘D’, ‘F’ included fishing. Shading indicates the additional CV expected for each value of the interaction, with the darkest shading in each plot corresponding to the Max. CV (%) given in the top-right corner of the plot.

4.5 Discussion

Analysing sensitivities to initial conditions is an important part of developing complex models (Rabier et al., 1996; Rosati et al., 1997; Payne et al., 2015; Cheung et al., 2016). If small perturbations to the initial conditions produce vastly different results, this may make interpreting results from the model challenging. Accounting for model uncertainty provides an envelope of model results, which tells us about the range of plausible outcomes rather than one possible instance. It is when the envelope is so wide that no result can be ascertained that the model can be frustratingly un-useful, and it is important we are aware when this is the case. For example, if scenarios exploring reduced fishing effort improved the general state of the ecosystem in some model runs, and deteriorated it in others, with all runs equally plausible, then we would be left none the wiser. It would be misleading to present results of only a subset or even a singular model run that does not adequately reflect the range of plausible outcomes.

I found the Chatham Rise Atlantis model was not chaotic, in that I could perturb the initial conditions by small, and even quite large (up to 50%) changes, and the model produced very similar results with respect to biomass trends and ecosystem indicators. The age-structured species groups typically had very tight CVs by the end of the burnin period, which is a reassuring result given Atlantis models are still fairly novel in the literature and have not yet undergone testing for chaotic dynamics. While the values of ecosystem indicators did not converge over the model time series, the response to fishing was consistent across model runs, suggesting system dynamics were consistent under perturbed initial conditions. This result puts us in a position to simulate scenarios using the Chatham Rise Atlantis model, including uncertainty of the initial conditions, and obtain an envelope of results with which to analyse and understand the likely responses of the Chatham Rise ecosystem.

While the system as a whole generally agreed within the range of results pro-

duced, the biomasses of some species groups varied between model runs more than others. The dynamics of some species groups appeared hyperstable as they promptly converged, while others retained variability between the runs, and for some the variability increased. I found the species groups that were more likely to have high biomass CVs (coefficient of variations) were those of lower trophic levels. In nature, we expect to see more variability in the abundances of lower trophic level species, but most relevant field experts would likely suggest those patterns derive from variability within the environment (Dippner et al., 2000, 2001; Molinero et al., 2008), which I am not applying in this study. If I combined varying the initial conditions with bootstrapping of the oceanographic variables, as carried out in McGregor et al. (2019b), we would likely see even greater variability in the lower trophic levels.

The effects of uncertainty from the oceanographic variables explored in McGregor et al. (2019b) had a greater range than those from perturbing the initial conditions. Similarly to the initial conditions uncertainty, the species most affected were also lower trophic levels. Diatoms had the highest CV at 79%, followed by carnivorous zooplankton CV at 46%. The effects of uncertainty from specification of the spawning stock recruitment relationship as it was applied to the small pelagic fish species group were explored in McGregor et al. (2019a). These effects were seen right through to the ecosystem indicators, and would likely be more evident if more than one species group were directly affected in the study. Further work exploring the effects of parameter uncertainty ought to be carried out with the Chatham Rise Atlantis model, and these effects compared to those from the initial conditions, oceanographic variables and spawning stock recruitment already explored. One key area of parameters to explore is the predator/prey interactions, including the feeding functional response form and parameters. The effects of these will likely be noticed more as the model is used and hence taken away from its calibrated balanced state, where we might expect to see more variability in prey abundance.

An aspect not explored in this study is the effect of initialisation uncertainty on

the model realised diets. There is plenty of scope for Atlantis diets to vary as they are the result of spatial and temporal overlap, gap-size limits, growth rates, feeding functional response, availability of other prey, predation from other predators, habitat refuge and prey preferences. As ecosystem models are generally developed to help understand flow-on effects within a system, understanding the effect of uncertainties on the species interactions could be important, and we recommend future work considers this aspect.

Another aspect of the trophic level effect on variability is the way in which I have modelled the species groups in the Chatham Rise Atlantis model. First, we have the difference between species modelled as biomass-pools and those modelled with age-structure. Biomass-pool representations are more dynamic as there is little/no delay structure built in—growth is pooled across its many forms (reproductive, somatic and otherwise), so can effectively occur instantaneously, unlike in age-structured groups where maturity may take years and specific events like spawning are restrained. Given biomass-pool groups are also generally lower trophic level (with naturally higher levels of productivity and turnover), the GLM fitted to CVs of all species groups could pick up trophic level as an explanatory variable that also accounts for this group structure. Within the age-structured species groups, trophic level could also be confounded with the proportion of additional mortality. The additional forced mortality would likely be a stabilising attribute, and the proportions applied were greater for the higher trophic level species, as these were the ones with less predation mortality in the system. That the stabilising aspect filters down through the trophic levels, with the lower trophic levels retaining variability, could suggest the extent to which this is a top-down controlled system.

The method used to perturb the initial conditions was found to be important in explaining the CVs. The runs based on keystone-ness did not result in the highest CVs, even though these runs perturbed the initial conditions of the species groups expected to have the greatest impact on the rest of the system. While Paine (1969)

suggested keystone species have a stabilising effect on a system, it was more recently suggested to be more complicated than that (Mills et al., 1993). In this study, the possible stabilising effect of keystone species could be due to additional mortality applied to some of the high keystone species, and hence exerting a stabilising effect on the system. The runs perturbed based on uncertainty produced the greatest CVs. The effects of other explanatory variables, such as higher CVs for lower trophic levels, were consistent regardless of the method used to perturb the initial conditions. Hence, the method was not influential in how the system responded, only in how strongly it responded. It is possible the latter difference would diminish with a greater number of runs simulated for each set. In future simulations, perturbing the initial conditions based on uncertainty would seem appropriate, and should encompass the variability we would expect to see from other methods of perturbation.

One of the age-structured species groups that was most sensitive to initial conditions was the invert comm scav group (primarily scampi). When I account for uncertainty from initial conditions, the response of this group to heavy fishing is inconclusive. The heavy fishing on the system from the mid-1970s (Ministry for Primary Industries, 2017), some of which was targeted on scampi (Tuck, 2016), could easily be positive or negative for scampi based on these model results, and CVs for this species group remained high at just over 20%. In the base Chatham Rise Atlantis model (McGregor et al., 2019b), scampi were shown to respond to fishing in a very similar way to the fisheries stock assessment estimated biomass. The results here illustrate that the base model result for this species group, while convincing as it matched the fisheries models so well, was actually only one of many plausible results using this ecosystem model.

In the quest to provide meaningful and realistic results to simulations explored using complex ecosystem models, with high levels of uncertainty, we need to produce result envelopes, not single trajectories. It is important we move in the direction of simulating many instances of the model that account for its uncertainties, to

understand how likely a given response is, and avoid presenting what may be errant or non-representative results. We know there is uncertainty in defining initial conditions of ecosystem models, so varying the initial conditions to reflect this uncertainty in model results is crucial. It is not the only area of uncertainty; there are many. Given the complexity of these models, exploring all possible uncertainties explicitly is unlikely to be tractable. It may be possible, however, to address subsets of uncertainty that encompass the broader range of the uncertainty of the model by targeting its key dynamics. The key dynamics of an ecosystem model generally consist of growth, recruitment, mortality, trophic connections, environmental effects, and initial state. Three of these (growth, mortality and trophic connections) relate directly to predation and consumption, and we could vary the feeding response function to explore the effects of uncertainties in these dynamics. Initial conditions were the topic of this study, and uncertainty from environmental effects were explored through bootstrapping the oceanographic variables in McGregor et al. (2019b). Recruitment/productivity remains, for which we could vary the spawning stock recruitment parameters. The specifics of varying these will vary between models and systems, but accounting for uncertainty with respect to four main categories: 1.) initial conditions; 2.) environmental; 3.) feeding functional response; 4.) productivity/recruitment, is likely to cover the broad range for most systems and models.

Chapter 5

Synthesising the process

The development of complex ecosystem models, such as Atlantis, is not a linear process, but rather iterative and incremental. This extends to the entire process from initially building the model, to being in a position to use it. Atlantis models presented in the literature historically focused most heavily on model development and calibration (Link et al., 2010; Ainsworth et al., 2011; Weijerman et al., 2015). More recent applications have also focused on formal skill assessment following the recommendations of Olsen et al. (2016) and sensitivity analyses that provide greater insight into the dynamics of the model (Ortega-Cisneros et al., 2017; Sturludottir et al., 2018; Porobic et al., 2019). This study extends beyond what is currently done in the literature with respect to exploring and understanding model dynamics, and adds to the understanding of fundamental dynamics relating to multi-species and ecosystem models. This synthesis chapter summarises the findings of the previous chapters, then presents the path of model development, exploration and validation, drawing on the benefits of hindsight gained through this study, and highlighting areas of understanding that remain less clear. There is still a lot of development yet to happen in this space.

Chapter 2 presented the Chatham Rise Atlantis model, from data compilation to model validation. This chapter synthesizes two years of work, and relates to ‘Model construction’, ‘Model calibration and fitting to history’, ‘Model sensitivity and skill

assessment’, ‘Scenarios for system understanding’, and ‘Tool development’. The model was not just validated in this work; it was explored. In the model validation, I assessed model results with observations from surveys and assessment, looking at biomass time-series in response to historical fishing; size-at-age were compared to von Bertalanffy curves defined from the literature; proportions-at-age were compared to exponential decay curves defined from the literature; realised diet summaries were compared to those from the literature. To further understand model dynamics, I bootstrapped the oceanographic variables that are forced in the model, then analysed which species groups responded more or less across all bootstrap runs, and then by warmer or cooler years. I perturbed each species group in turn and analysed the flow-on effects in the model to gain understanding of how the dynamics were connected. I analysed data gaps in the model with respect to the flow-on effects and gave recommendations for sensitivities recommended when this model is used for exploring scenarios.

Chapter 3 explored a structural component of ecosystem models that has transferred from fisheries modelling, the Spawning Stock Recruitment function. The dynamics of this relationship have only previously been given serious thought when relating to the biomass of a depleted biomass, with the dynamic of interest being the rate of recovery. I focused the work on the most commonly used Spawning Stock Recruitment curve, the Beverton-Holt. There are three main attributes of this curve; 1.) the steepness of the curve when the population’s biomass is less than 20% of its virgin level (referred to as ‘steepness’); 2.) the expected recruitment when biomass is at its virgin level; 3.) the asymptotic value of the curve. We generally have the most information about the virgin state of fish stocks as fisheries models are designed to estimate back to virgin biomass and recruitment based on the time series of abundance and fishing available. We generally have less information about the depleted part of the curve, as the only way of really getting information on this is to deplete the population below 20% which is what we are trying to avoid doing. So sensitivities on steepness to understand the repercussions of different values on model results are

often used. We might want to do something similar in a multispecies or an ecosystem context, as it could be influential on model results here too. Set steepness too high, and simulations may predict the ecosystem would recover much faster than would actually be likely. Get it too low, and species may be more prone to extinction than they would in reality. When we don't know whether it should be high or low, we need to present results that reflect this uncertainty. However, in varying the steepness, we still want to keep the virgin state of the population the same. Retaining the virgin population state and varying steepness means the asymptote changes, and it does not take too much exploration to find it changes in the opposite direction to the steepness change. A population defined to have low steepness when depleted will have much higher asymptotic recruitment. But this doesn't matter if the population can't exceed its virgin state, right? Well yes, and no. The virgin state of a population is linked to an assumed level of natural mortality. If this level of natural mortality remains the same, then without fishing effort, the population will end up at its virgin level. Constant natural mortality is not a feature of ecosystem or multispecies models. It is possible that in the realm of time-varying natural mortality, the concept of virgin state needs to change. While I discuss the concept of virgin state with respect to time-varying natural mortality in the chapter, the analyses presented assume the virgin state remains the same. I propose the simplest solution, which was to cap recruitment at its virgin level, and I explored the effects of doing so both on a simple population model, and within the Chatham Rise Atlantis model. The simplest solution was chosen to demonstrate the effects, but it may not always be an appropriate solution. There are aspects of the Ricker Spawning Stock Recruitment model that may be more appropriate, as this curve reaches its highest level of recruitment when biomass is greater than the virgin level (and the maximum is higher if steepness is low), but the curve then descends back towards zero, thus tending the population back down below its virgin state. It is also quite plausible for a population to recover slowly when depleted and have recruitment potential well above its virgin level—in which case no change to the Beverton-Holt curve would be necessary. My concern was not just in the specification of the Spawning Stock Recruitment for a species, but

also for sensitivities on the steepness parameter, where exploring a range of steepness values without the effect on the asymptote was otherwise not possible.

In chapter 4, my analyses focused on the Chatham Rise Atlantis model's sensitivity to initial conditions. We know there are uncertainties around the initial conditions, and the implications of these should be considered in model results. In Atlantis, there are many thousands of parameters for which initial conditions are specified (several hundred just for species function groups, without considering spatial structure), and all of them will have some level of error both in magnitude and direction. Different combinations of perturbing the initial conditions may have different effects on the model outputs, so we need to account for that, but the number of simulations required to explore this entire space of uncertainty is enormous. Hence, my approach was to perturb the initial conditions in a way most likely to be representative of the uncertainties we are trying to account for. Initially, I perturbed all the species initial conditions by the same magnitude and direction for each simulation. This approach allowed me to perturb a large number of initial conditions (all those for species functional groups), but it did not offer any insight into the effect of changes in multiple directions. In the second set I perturbed the initial conditions of the species groups that were likely to be most influential on the rest of the system - using the keystone ranking of Chapter 2. The third approach I used was to perturb the initial conditions by random amounts that accounted for both magnitude and direction, and were more likely to be larger for species groups that were poorly informed. Each of the three sets had 35 runs with fishing included, and 35 runs without fishing; enough runs to allow me to analyse variance. My first curiosity was whether the model was chaotic —from the work I had carried out so far, I could not predict the outcome to this result. Would I adjust the initial conditions slightly and find the model produced very different biomasses and dynamics? While the results did not produce generally very different biomasses and dynamics, there were still interesting aspects; I noticed some biomass trajectories converged very quickly, while others continued along separate paths. This difference prompted the second

part of this chapter which modelled the variability between runs using species characteristics such as lifespan or trophic levels as explanatory variables.

The path of model development as presented in this thesis can be thought of in seven main stages. Figure 5.1 shows the stages in a spiral, reflecting the iterative and incremental nature of the process. The smaller spirals that overlap with the large spiral reflect the possibility of returning to the drawing board at any stage, and that all of the stages are connected. Some activities likely occurred prior to these stages, such as deciding a model was needed and for what purpose, and communicating with relevant stakeholders, interest groups and experts. It would be beneficial for the modeller to be involved in the pre-model development activities, such as was the case for Porobic et al. (2019), as then the modeller gains more understanding of the purpose of the model and has input on selecting an appropriate model framework. It is not always the case, as project inception can occur prior to modeller involvement, which was the situation here. The decision to develop an Atlantis model of the Chatham Rise for the purpose of assisting an Ecosystem Approach to Fisheries Management had been made prior to my involvement in the project.

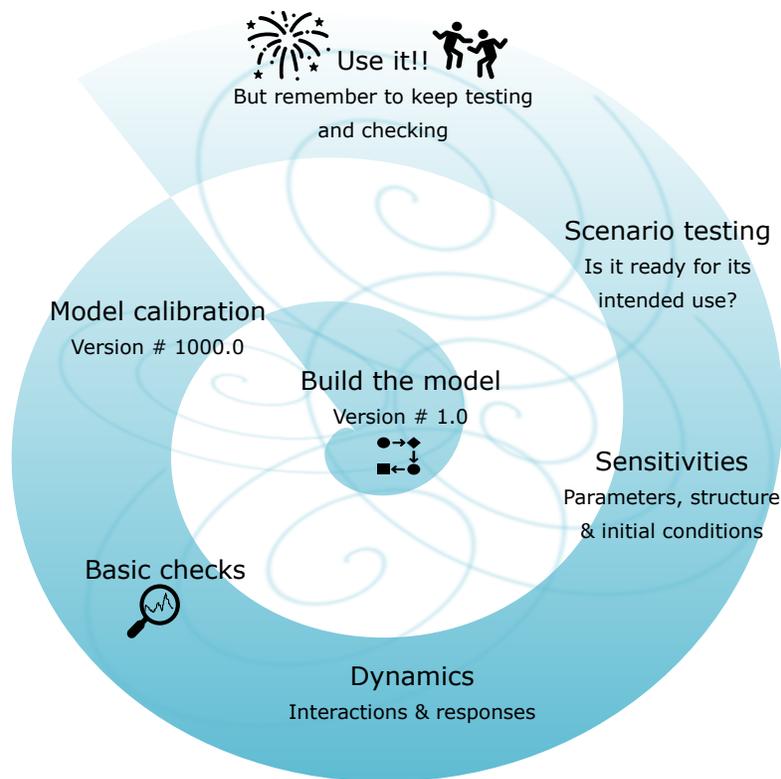


Figure 5.1: The model process as applied for the Chatham Rise Atlantis model. The process begins in the centre of the spiral, then progresses towards the outer spiral, but can at any stage return to an earlier part of the process, and no part of the process is separate.

The stages are:

Stage 1: Build the model, largely featured in Chapter 2, but it could be revisited at any time—even during model use. The initial build includes model design as well. Defining the physical structure of the model, deciding on timestep size, defining species functional groups, and all their parameters and functional forms, defining fisheries, defining prey availabilities, feeding parameters and functional forms, migration parameters, and spatial distributions. Initially, it is a matter of putting it all there and seeing what happens. The initial build of the model is a significant amount of work, and yet is unlikely to produce much in the way of realistic dynamics—the whole system could crash within its first timestep, for example. It requires patience, persistence, and perhaps optimism to develop and then calibrate a complex ecosystem model. One lesson for this stage: Think carefully about the spatial structure and the size of timesteps as they can't be easily changed without creating a new model. Most other aspects can be changed, including group structure.

Stage 2: Model calibration, links very closely with stage 1. The initial hurdle is for the model to persist. A large imbalance in the nutrients, for example, will cause the model to crash. Early on, the model dynamics can be very unrealistic—for example, all the sharks could go extinct, or crabs could grow to 5 times their size and proceed to take over the ecosystem. One lesson from calibrating a complex ecosystem model: multiple model runs save time and lead to better model decisions. Even if focusing on a component over short time periods, such as nutrient dynamics, still always do multiple runs. They don't have to be huge, but several parameters varied by several scalars will tell you so much more about what is going on so much more quickly than individual runs. During model calibration I wrote an R function to change a given parameter for a given set of species groups, by a vector shifts that were either absolute or proportional. This was used for exploring parameter space during calibration. Matching this, I wrote an R function to read in multiple runs and compare the outputs, either summarised over space, by polygon, or by cell (layer of a polygon). Following from this, I wrote a wrapper to run all combinations of runs for a

given set of species groups and a given vector of changes on a given set of parameters.

Stage 3: Basic checks overlaps with stage 2 (and hence stage 1) as these checks are one of the things informing the calibration process, especially once the model is not crashing. Species groups were checked for growth by inspecting realised growth curves, and mortality through exponential decay curves fitted to the realised proportions-at-age. Too much mortality and too much growth may not show up in biomass trajectories if these balance each other out, but their dynamics in the system will be different. The effects of fishing could change due to size selectivities, and it will affect predator-prey interactions due to gape-sizes. The diets take a lot of attention. I wrote approximately 50 R-scripts focused on predation, analysing spatial and temporal overlap, gape sizes, growth rates of the predator, mortality rates of the prey, comparing realised diets with the literature. Much of the work on diets included becoming familiar with the Atlantis source code, in particular the function ‘Eat()’. When I created a new *fprintf* statement in Atlantis, I often wrote a corresponding R script to analyse the outputs, as these were often multi-dimensional. There was a lot of repetition in basic checks and calibrations, which I’m sure could benefit from automation. There is the idea that going through this process helps the modeller to really learn and understand their model. I suspect the benefits of automation would outweigh the disadvantages of model familiarity. It would allow the modeller time to explore a well-functioning model with sensitivities. There is progress in this area, with some colleagues in Europe working on optimisation routines for subcomponents of Atlantis models, with the intention of running them using High Performance Computing. It is possible that calibrating complex ecosystem models will become a thing of the past, with calibration algorithms developed instead. There is risk involved in automating more of the development and calibration process as the potential for blindly creating complex models increases. Hopefully, standards for transparently and coherently assessing the dynamics of complex ecosystem models will develop ahead of processes for automating their development.

Step 4: Dynamics involved in using the model system in some way and analysing how it responded. As is typically done, applying fishing should reduce the biomass of some affected species. In more recent years, there have been examples of comparing the biomass trends with historical data using a skill assessment (Olsen et al., 2016; Ortega-Cisneros et al., 2017; Sturludottir et al., 2018). I took this a little further by comparing data outputs to trawl survey estimates, stock assessment estimates, and CPUE (catch per unit effort) estimates of abundance. The latter two allowed me to compare the ecosystem dynamics of these species with the single species dynamics. It was interesting they matched so well, as the stock assessment was using between-year recruitment deviates, which are very influential on the resulting stock size, but fixed natural mortality, whereas the Atlantis model did not have recruitment deviates, but did have time-varying mortality and growth rates through the ecosystem dynamics. The interaction dynamics have not been explored in this way for Atlantis models, although was based on theory behind some of the Ecopath with Ecosim outputs (Libralato et al., 2006). This exercise was a communication interface between the model dynamics and the ecosystem experts. While I did not know which groups should be most influential (although myctophids were not a surprise), when I presented the results to colleagues familiar with the Chatham Rise ecosystem, they confirmed the results to be realistic. Had the results been implausible, I would have needed to find out why. Turning the keystone equation around to calculate responsiveness was new from all ecosystem modelling, as was analysing the informance of each species group with respect to their keystone and responsiveness ratings.

Step 5: Sensitivities is the area I would argue is most lacking in complex ecosystem models. Sensitivity analyses can be analytical or empirical (Pantus, 2007), although for complex ecosystem models such as Atlantis, analytical is not a viable option. Some fairly tentative examples of empirical sensitivity analyses exist of perturbing a selection of parameters for a selection of species (Sturludottir et al., 2018; Ortega-Cisneros et al., 2017; Hansen et al., 2019), but not analysed as resulting uncertainty in model results. I would expect this level of sensitivity analyses to be

explored during model calibration and basic checks. Sensitivities that account for uncertainties require a more robust statistical approach, where perturbations have some random variability. One lesson from this step is to do it earlier in the process. When I presented biomass responses to fishing in Chapter 2, they should have included uncertainty. This was particularly apparent in scampi which was presented as a 'surprisingly good match' in Chapter 2 only to find out it was more complicated than that in Chapter 4. Once I accounted for uncertainty in the initial conditions, the result disappeared in the noise of many different and equally plausible results. It would have made sense to combine the oceanography variability with sensitivity to the initial conditions, and it is my recommendation this is done when in Step 7: Using the model. I suspect there is much more we can and should do to address uncertainty, and I imagine we will see more development in accounting for uncertainty in results of complex ecosystem models in the near future.

Step 6: Scenario testing involves testing components of a scenario for plausible responses. For example, are changes in the model dynamics reflected through ecosystem indicators? This is useful to know ahead of running scenarios if outcomes will be assessed using ecosystem indicators. It also suggests more development work may be required if these indicators are unrealistic or unresponsive. The analyses I presented assessing ecosystem indicators in response to fishing, when evaluating the spawning stock recruitment curve, was also informative for scenario testing. Some focused ecosystem response analyses exist in the literature: Masi et al. (2017) tested responses of ecosystem indicators to fishing in the Gulf of Mexico Atlantis model, and Ortega-Cisneros et al. (2017) presented a similar analyses but comparing an Atlantis model and an EwE model. Assessing the realism of responses relating to climate change scenarios is another example of scenario testing. Ortega-Cisneros et al. (2018) explored model responses to temperature increases on species groups prior to full climate change exploration using the Southern Benguela Atlantis model. This could be taken further by exploring the effects as they flow-on from nutrients and primary productivity. It may be possible to analyses these flow-on effects with

respect to Earth System Models such as those analysed for New Zealand in Law et al. (2017).

The model process for complex ecosystem models to date has generally consisted of building the model, calibrating the model, some basic checks, the fisheries dynamics tested, very little on sensitivities, with almost no uncertainty in model results, almost no scenario testing, but quite a bit of model use. This thesis urges greater emphasis on model dynamics, sensitivities, presenting uncertainties in model results, and scenario testing.

Scenarios explored using the Chatham Rise Atlantis model need to consist of multiple runs that vary the initial conditions based on uncertainty of the species functional groups and sample the ROMS variables for each year simulated. Varying the initial conditions based on the uncertainty levels of the species groups will bring in more variability for the seabird and cetacean groups found to be poorly defined and influential, however further perturbing these groups perhaps through their productivity parameters could be explored. Scenarios should be explored with reduced primary production to account for potentially lower levels, particularly in the spring and autumn blooms, expected with climate change at the mid-and end of this century (Law et al., 2017).

Recommendations for understanding results from complex ecosystem models:

- Always understand what in the model has produced a result before considering what the result may tell us about the system. There are so many ways a result could occur in a complex model.
- Don't fall into the trap of accepting a model result just because you like it. It is just as important to understand the results you like, as the results you don't like.
- Always test results with uncertainty. It is potentially misleading to present

results from single model runs as they could be flukes or non-representative

- Incorporate random variability in uncertainty analyses, varying both direction and magnitude.
- Ideally an ecosystem model should have two lead scientists; one an expert of the model, and one an expert of the system. Ecologists or biologists are more likely to understand the dynamics as they exist in the real world, and a modeller is more likely to understand the mathematical equations and the statistical analyses required for the abstraction of these dynamics into a computer simulation. The ecologist can describe how they would sample the real system and what they would expect to see, and the modeller can replicate this by sampling the modelled system and comparing what they see.
- Always present multiple indicators. Biomasses and catches only show part of the picture, and generally don't reflect trophic interactions. Ecosystem indicators come closer. Realised diets can be presented as a timeseries which can be great for identifying predator/prey switching.
- Remember there are multiple dimensions. Results are often summarised over space and time, but remember these are summaries and you may be missing some key dynamics. Predator/prey dynamics all occur at a fine scale, so the cause of prey-switching could be due to a shift in spatial distribution. Fine if the switch is plausible, but if not then more work on this part of the model may be required.
- Look at species population dynamics in different ways. For example, numbers by age class will tell you about shifts in mortality, and size by age class will tell you about shifts in food availability. Both are important aspects in understanding responses of a species group to a scenario, and may help understand overall results too.
- Remember assumptions. There will be too many to remember each individu-

ally, but the general areas of assumption (e.g. predator-prey response parameters, productivity, migration)

Recommendations for empiricists and the modelling community going forward

- Rethink the application of theory developed for exploited single-species fish stocks to multi-species models. The Spawning Stock Recruitment curve is one, and this raised the idea that ‘virgin state’ of a population should be reconsidered when natural mortality varies with time.
- Predator/prey feeding response curve functional form and parameters could benefit from some more attention. These are often changed during model calibration, but model calibration is often focused on the model dynamics in or close to equilibrium (due to the goal of stability under constant forcing). The real purpose of the feeding response curve is to describe feeding behaviour at different levels of prey availability, and this won’t be tested at equilibrium.
- Development or exploration of statistical techniques for analysing multi-dimensional results that include uncertainty. Perhaps Brownian Motion for introducing random variability, and Martingale Processes to separate the pattern from the noise, or perhaps a simple difference of means test.
- Develop programs to automate at least some of the calibration and basic checks parts of the process. Artificial Intelligence may assist in these developments.

Recommendations for fisheries and marine resource managers using ecosystem models in the future

- Be extremely wary of results that are presented without uncertainty bounds.
- Complex ecosystem models are learning tools, where scenarios are explored to help understand the dynamics of the system. To be useful, results from ecosystem models need to be thought of differently to results from single-species

fisheries models. For example, rather than seeking a probability of falling below a threshold, we may be seeking the extent of impact.

- Single species reference points such as F_{MSY} require a bit more thought in a multi-species context because they don't allow for interactions between species. For example, a reduction of stock for one exploited species may change the population dynamics and hence reference points of another exploited species.

Ecosystem modelling is exciting because of the potential of what it can offer, however, with this excitement comes the risk of haste—presenting results and outputs from models that are not ready, or that are not well understood. On the other hand, waiting for the perfect model would likely mean waiting forever. So, as a modelling community, it is finding a balance between cautious optimism of where we are now, and ambitious optimism as we strive for where we can be in the future.

Chapter 6

References

- Ainsworth, C. H., Kaplan, I. C., Levin, P. S., Cudney-Bueno, R., Fulton, E. A., Mangel, M., Turk-Boyer, P., Torre, J., Pares-Sierra, A., and Morzaria-Luna, H. N. (2011). Atlantis model development for the northern gulf of california.
- Ainsworth, C. H. and Pitcher, T. J. (2006). Modifying kempton’s species diversity index for use with ecosystem simulation models. *Ecological Indicators*, 6(3):623–630.
- Ainsworth, C. H., Schirripa, M. J., and Morzaria-Luna, H. N. (2015). An Atlantis ecosystem model for the Gulf of Mexico supporting integrated ecosystem assessment. *NOAA technical memorandum NMFS-SEFC*, NMFS-SEFSC-676.
- Allesina, S. and Tang, S. (2012). Stability criteria for complex ecosystems. *Nature*, 483(7388):205.
- Andersen, K. H., Jacobsen, N. S., and Farnsworth, K. D. (2016). The theoretical foundations for size spectrum models of fish communities. *Canadian Journal of Fisheries and Aquatic Sciences*, 73(4):575–588.
- Audzijonyte, A., Gorton, R., Kaplan, I., and Fulton, E. A. (2017a). Atlantis users guide part i: General overview, physics & ecology.
- Audzijonyte, A., Pethybridge, H., Porobic, J., Gorton, R., Kaplan, I., and Fulton,

- E. A. (2017b). Atlantis: a spatially explicit end-to-end marine ecosystem model with dynamically integrated physics, ecology and socio-economics modules.
- Baretta-Bekker, J. (1996). European regional seas ecosystem model-i. *Oceanographic Literature Review*, 3(43):271.
- Beentjes, M. P. and Stevenson, M. L. (2009). Inshore trawl survey of Canterbury Bight and Pegasus Bay, May-June 2008 (KAH0806). *New Zealand Fisheries Assessment Report*, 57:105 p.
- Begley, J. and Howell, D. (2004). An overview of gadget, the globally applicable area-disaggregated general ecosystem toolbox. ICES.
- Beverton, R. J. and Holt, S. J. (1957). *On the dynamics of exploited fish populations*. Springer Science & Business Media.
- Blanchard, J., Pinnegar, J., and Mackinson, S. (2002). *Exploring marine mammal-fishery interactions using 'Ecopath with ecosim': modelling the Barents Sea ecosystem*. Citeseer.
- Blanchard, J. L., Andersen, K. H., Scott, F., Hintzen, N. T., Piet, G., and Jennings, S. (2014). Evaluating targets and trade-offs among fisheries and conservation objectives using a multispecies size spectrum model. *Journal of Applied Ecology*, 51(3):612–622.
- Borja, A., Elliott, M., Andersen, J. H., Berg, T., Carstensen, J., Halpern, B. S., Heiskanen, A.-S., Korpinen, S., Lowndes, J. S. S., Martin, G., et al. (2016). Overview of integrative assessment of marine systems: the ecosystem approach in practice. *Frontiers in Marine Science*, 3:20.
- Botsford, L. W., Castilla, J. C., and Peterson, C. H. (1997). The management of fisheries and marine ecosystems. *Science*, 277(5325):509–515.
- Brand, E. J., Kaplan, I. C., Harvey, C. J., Levin, P. S., Fulton, E. A., Field, J. C., and Hermann, A. J. (2007). A spatially explicit ecosystem model of the california current's food web and oceanography.

- Breen, P. A., Kim, S. W., and Andrew, N. L. (2003). A length-based Bayesian stock assessment model for the New Zealand abalone *Haliotis iris*. *Marine and Freshwater Research*, 54(5):619–634.
- Broekhuizen, N., Gurney, W., Jones, A., and Bryant, A. (1994). Modelling compensatory growth. *Functional Ecology*, 8(6):770–782.
- Bull, B., Francis, R., Dunn, A., McKenzie, A., Gilbert, D., Smith, M., and Bian, R. (2012). CASAL (c++ algorithmic stock assessment laboratory): Casal user manual v2.30–2012/03/21. *NIWA Technical Report*, 135.
- Catul, V., Gauns, M., and Karuppasamy, P. (2011). A review on mesopelagic fishes belonging to family myctophidae. *Reviews in Fish Biology and Fisheries*, 21(3):339–354.
- Cerrato, R. M. (1990). Interpretable statistical tests for growth comparisons using parameters in the von bertalanffy equation. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(7):1416–1426.
- Chang, W. Y. (1982). A statistical method for evaluating the reproducibility of age determination. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8):1208–1210.
- Cheung, W. W., Frölicher, T. L., Asch, R. G., Jones, M. C., Pinsky, M. L., Reygondeau, G., Rodgers, K. B., Rykaczewski, R. R., Sarmiento, J. L., Stock, C., et al. (2016). Building confidence in projections of the responses of living marine resources to climate change. *ICES Journal of Marine Science: Journal du Conseil*, page fsv250.
- Christensen, V. and Walters, C. J. (2004). Ecopath with ecosim: methods, capabilities and limitations. *Ecological modelling*, 172(2-4):109–139.
- Clark, M. R. (1985). The food and feeding of seven fish species from the campbell plateau, new zealand. *New Zealand Journal of Marine and Freshwater Research*, 19(3):339–363.

- Clark, M. R. and Rowden, A. A. (2009). Effect of deepwater trawling on the macro-invertebrate assemblages of seamounts on the chatham rise, new zealand. *Deep Sea Research Part I: Oceanographic Research Papers*, 56(9):1540–1554.
- Clarke, M., Connolly, P., and Bracken, J. (2002). Age estimation of the exploited deepwater shark *Centrophorus squamosus* from the continental slopes of the Rockall Trough and Porcupine Bank. *Journal of Fish Biology*, 60(3):501–514.
- Collie, J. S., Botsford, L. W., Hastings, A., Kaplan, I. C., Largier, J. L., Livingston, P. A., Plagányi, É., Rose, K. A., Wells, B. K., and Werner, F. E. (2014). Ecosystem models for fisheries management: finding the sweet spot. *Fish and Fisheries*.
- Collie, J. S., Botsford, L. W., Hastings, A., Kaplan, I. C., Largier, J. L., Livingston, P. A., Plagányi, É., Rose, K. A., Wells, B. K., and Werner, F. E. (2016). Ecosystem models for fisheries management: finding the sweet spot. *Fish and Fisheries*, 17(1):101–125.
- Collins, M. A., Xavier, J. C., Johnston, N. M., North, A. W., Enderlein, P., Tarling, G. A., Waluda, C. M., Hawker, E. J., and Cunningham, N. J. (2008). Patterns in the distribution of myctophid fish in the northern scotia sea ecosystem. *Polar Biology*, 31(7):837–851.
- Cubillos, L., Alarcón, R., Vilugran, A., Sepúlveda, A., George-Nascimento, M., Araya, M. Medina, J., Zambrano, M., Guzman, L., Martinez, J., Peñailillo, R., and Gili, Z. Young, V. A. V. B. L. M. L. (1998). Validación de los métodos aplicados en la determinación de edad y crecimiento, y determinación de la mortalidad del jurel en la zona centro sur. *Informe Final Proyecto FIP, 95–10*, page 170 p.
- Deepwater Group (2018). Deepwater Groups fisheries certification programme. Technical report, <http://deepwatergroup.org/certification/> [Accessed 7 June 2018].
- Deroba, J. J. and Schueller, A. M. (2013). Performance of stock assessments with misspecified age-and time-varying natural mortality. *Fisheries Research*, 146:27–40.

- Deutsches Ozeanographisches Datenzentrum, Bundesamt Fur Seeschiffahrt Und Hydrographie, G. (2006). The world ocean circulation experiment (woce) global hydrographic climatology.
- Dippner, J. W., Hänninen, J., Kuosa, H., and Vuorinen, I. (2001). The influence of climate variability on zooplankton abundance in the northern baltic archipelago sea (sw finland). *ICES Journal of Marine Science*, 58(3):569–578.
- Dippner, J. W., Kornilovs, G., and Sidrevics, L. (2000). Long-term variability of mesozooplankton in the central baltic sea. *Journal of Marine Systems*, 25(1):23–31.
- Doonan, I., Large, K., Dunn, A., Rasmussen, S., Marsh, C., and Mormede, S. (2016). Casal2: New zealands integrated population modelling tool. *Fisheries Research*, 183:498–505.
- Dunn, M., Anderson, O., and Doonan, I. (2008). An evaluation of stock status for orange roughy on the east and south chatham rise in 2008. *New Zealand Fisheries Assessment Report*, 65:30.
- Dunn, M. and Doonan, I. (in press). Assessment of the Chatham Rise orange roughy stocks for 2017. *New Zealand Fisheries Assessment Report*.
- Dunn, M., Horn, P., Connell, A., Stevens, D., Forman, J., Pinkerton, M., Griggs, L., Notman, P., and Wood, B. (2009). Ecosystem-scale trophic relationships: diet composition and guild structure of middle-depth fish on the chatham rise. Technical report, Final Research Report for Ministry of Fisheries Research Project ZBD2004-02 Objectives 1–5.
- Dunn, M. R., Griggs, L., Forman, J., and Horn, P. (2010). Feeding habits and niche separation among the deep-sea chimaeroid fishes harriotta raleighana, hydrolagus bemisi and hydrolagus novaezealandiae. *Marine Ecology Progress Series*, 407:209–225.

- Escobar-Flores, P. C., Ladroit, Y., and O'Driscoll, R. L. (2019). Acoustic assessment of the micronekton community on the chatham rise, new zealand, using a semi-automated approach. *Frontiers in Marine Science*, 6:507.
- Essington, T. E. (2007). Evaluating the sensitivity of a trophic mass-balance model (ecopath) to imprecise data inputs. *Canadian Journal of Fisheries and Aquatic Sciences*, 64(4):628–637.
- Fasham, M. J., Ducklow, H. W., and McKelvie, S. M. (1990). A nitrogen-based model of plankton dynamics in the oceanic mixed layer. *Journal of Marine Research*, 48(3):591–639.
- Fournier, D., Sibert, J. R., Majkowski, J., and Hampton, J. (1990). Multifan a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian journal of fisheries and aquatic sciences*, 47(2):301–317.
- Fournier, D. A., Hampton, J., and Sibert, J. R. (1998). Multifan-cl: a length-based, age-structured model for fisheries stock assessment, with application to south pacific albacore, thunnus alalunga. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(9):2105–2116.
- Francis, R. I. C. C. (1992). Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*hoplostethus atlanticus*) on the chatham rise, new zealand. *Canadian Journal of Fisheries and Aquatic Sciences*, 49(5):922–930.
- Froese, R., Walters, C., Pauly, D., Winker, H., Weyl, O. L., Demirel, N., Tsikliras, A. C., and Holt, S. J. (2015). A critique of the balanced harvesting approach to fishing. *ICES Journal of Marine Science*, 73(6):1640–1650.
- Fulton, E., Smith, A., and Smith, D. (2007). Alternative management strategies for

southeast australian commonwealth fisheries: stage 2: quantitative management strategy evaluation.

- Fulton, E. A. (2001). *The effects of model structure and complexity on the behaviour and performance of marine ecosystem models*. PhD thesis, University of Tasmania.
- Fulton, E. A. (2010). Approaches to end-to-end ecosystem models. *Journal of Marine Systems*, 81(1-2):171–183.
- Fulton, E. A., Link, J. S., Kaplan, I. C., Savina-Rolland, M., Johnson, P., Ainsworth, C., Horne, P., Gorton, R., Gamble, R. J., Smith, A. D. M., and Smith, D. C. (2011). Lessons in modelling and management of marine ecosystems: the atlantis experience. *Fish and Fisheries*, 12(2):171–188.
- Fulton, E. A., Parslow, J. S., Smith, A. D., and Johnson, C. R. (2004a). Biogeochemical marine ecosystem models ii: the effect of physiological detail on model performance. *Ecological modelling*, 173(4):371–406.
- Fulton, E. A., Smith, A. D., and Johnson, C. R. (2004b). Biogeochemical marine ecosystem models i: Igbema model of marine bay ecosystems. *Ecological Modelling*, 174(3):267–307.
- Fulton, E. A., Smith, A. D., Smith, D. C., and Johnson, P. (2014). An integrated approach is needed for ecosystem based fisheries management: insights from ecosystem-level management strategy evaluation. *PLoS One*, 9(1):e84242.
- Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O., Zweng, M., Reagan, J., and Johnson, D. (2013a). World ocean atlas 2013, volume 3: Oxygen. *Mishonov Technical Ed.; NOAA Atlas NESDIS 73*, page 27.
- Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O., Zweng, M., Reagan, J., and Johnson, D. (2013b). World ocean atlas 2013, volume 4: Nutrients. *Mishonov Technical Ed.; NOAA Atlas NESDIS 73*, page 25.
- Grilli, J., Rogers, T., and Allesina, S. (2016). Modularity and stability in ecological communities. *Nature communications*, 7:12031.

- Grüss, A., Harford, W. J., Schirripa, M. J., Velez, L., Sagarese, S. R., Shin, Y.-J., and Verley, P. (2016a). Management strategy evaluation using the individual-based, multispecies modeling approach osmose. *Ecological Modelling*, 340:86–105.
- Grüss, A., Schirripa, M. J., Chagaris, D., Velez, L., Shin, Y.-J., Verley, P., Oliveros-Ramos, R., and Ainsworth, C. H. (2016b). Estimating natural mortality rates and simulating fishing scenarios for gulf of mexico red grouper (*epinephelus morio*) using the ecosystem model osmose-wfs. *Journal of Marine Systems*, 154:264–279.
- Haddon, M. (2001). *Modelling and Quantitative Methods in Fisheries*. Statistics/Biology. Taylor & Francis.
- Hadfield, M. G., Rickard, G. J., and Uddstrom, M. J. (2007). A hydrodynamic model of chatham rise, new zealand. *New Zealand Journal of Marine and Freshwater Research*, 41(2):239–264.
- Hanchet, S. M. (1986). *The Distribution and Abundance, Reproduction, Growth, and Life History Characteristics of the Spiny Dogfish, Squalus Acanthias (Linnaeus), in New Zealand: A Thesis Submitted for the Degree of Doctor of Philosophy at the University of Otago, Dunedin, New Zealand*. PhD thesis, University of Otago.
- Hansen, C., Drinkwater, K. F., Jähkel, A., Fulton, E. A., Gorton, R., and Skern-Mauritzen, M. (2019). Sensitivity of the norwegian and barents sea atlantis end-to-end ecosystem model to parameter perturbations of key species. *PloS one*, 14(2):e0210419.
- Heath, R. (1985). A review of the physical oceanography of the seas around New Zealand—1982. *New Zealand journal of marine and freshwater research*, 19(1):79–124.
- Hernández-León, S., Franchy, G., Moyano, M., Menéndez, I., Schmoker, C., and Putzeys, S. (2010). Carbon sequestration and zooplankton lunar cycles: could we be missing a major component of the biological pump? *Limnology and Oceanography*, 55(6):2503–2512.

- Heymans, J. J., Howell, K. L., Ayers, M., Burrows, M. T., Gordon, J. D., Jones, E. G., and Neat, F. (2010). Do we have enough information to apply the ecosystem approach to management of deep-sea fisheries? An example from the West of Scotland. *ICES Journal of Marine Science*, 68(2):265–280.
- Hilborn, R. (2011). Future directions in ecosystem based fisheries management: A personal perspective. *Fisheries Research*, 108(2-3):235–239.
- Hilborn, R., Walters, C. J., et al. (1992). Quantitative fisheries stock assessment: choice, dynamics and uncertainty. *Reviews in Fish Biology and Fisheries*, 2(2):177–178.
- Hoegh-Guldberg, O. and Bruno, J. F. (2010). The impact of climate change on the worlds marine ecosystems. *Science*, 328(5985):1523–1528.
- Horn, P. (2002). Age estimation of barracouta (*Thyrsites atun*) off southern New Zealand. *Marine and Freshwater Research*, 53:1169–1178.
- Horn, P. (2013). Stock assessment of hake (*Merluccius australis*) on the Chatham Rise (HAK 4) and off the west coast of South Island (HAK 7) for the 2012–13 fishing year. *New Zealand Fisheries Assessment Report*, 31:62.
- Horn, P. L. and Sutton, C. P. (1996). Validated ages, growth, and productivity parameters for silver warehou (*Seriolella punctata*) off the south and east coasts of south island, new zealand. *New Zealand Journal of Marine and Freshwater Research*, 30(3):301–312.
- Horne, P. J., Kaplan, I. C., Marshall, K. N., Levin, P. S., Harvey, C. J., Hermann, A. J., and Fulton, E. A. (2010). Design and parameterization of a spatially explicit ecosystem model of the central california current. Technical report.
- Irigoiien, X., Klevjer, T. A., Røstad, A., Martinez, U., Boyra, G., Acuña, J., Bode, A., Echevarria, F., González-Gordillo, J. I., Hernandez-Leon, S., Agusti, S., Aksnes, D. L., Duarte, C. M., and Kaartvedt, S. (2014). Large mesopelagic fishes biomass and trophic efficiency in the open ocean. *Nature communications*, 5:ncomms4271.

- Irvine, S. B., Stevens, J. D., and Laurenson, L. J. (2006a). Comparing external and internal dorsal-spine bands to interpret the age and growth of the giant lantern shark, *Etmopterus baxteri* (Squaliformes: Etmopteridae). *Environmental Biology of Fishes*, 77(3-4):253–264.
- Irvine, S. B., Stevens, J. D., and Laurenson, L. J. (2006b). Surface bands on deep-water squalid dorsal-fin spines: an alternative method for ageing *Centroselachus crepidater*. *Canadian journal of fisheries and aquatic sciences*, 63(3):617–627.
- Jacobsen, N. S., Gislason, H., and Andersen, K. H. (2014). The consequences of balanced harvesting of fish communities. *Proceedings of the Royal Society B: Biological Sciences*, 281(1775):20132701.
- Johnson, K. F., Monnahan, C. C., McGilliard, C. R., Vert-pre, K. A., Anderson, S. C., Cunningham, C. J., Hurtado-Ferro, F., Licandeo, R. R., Muradian, M. L., Ono, K., et al. (2014). Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. *ICES Journal of Marine Science*, 72(1):137–150.
- Kaplan, I. C., Horne, P. J., and Levin, P. S. (2012). Screening california current fishery management scenarios using the atlantis end-to-end ecosystem model. *Progress in Oceanography*, 102:5–18.
- Kaplan, I. C. and Marshall, K. N. (2016). A guinea pig’s tale: learning to review end-to-end marine ecosystem models for management applications. *ICES Journal of Marine Science*, 73(7):1715–1724.
- Kimura, D. K. (1980). Likelihood methods for the von bertalanffy growth curve. *Fishery bulletin*, 77(4):765–776.
- Kishi, M. J., Kashiwai, M., Ware, D. M., Megrey, B. A., Eslinger, D. L., Werner, F. E., Noguchi-Aita, M., Azumaya, T., Fujii, M., Hashimoto, S., et al. (2007). Nemuroa lower trophic level model for the north pacific marine ecosystem. *Ecological Modelling*, 202(1-2):12–25.

- Kochkin, P. (1994). Age determination and estimate of growth rate for the Peruvian jack mackerel, *Trachurus symmetricus murphyi*. *Journal of Ichthyology*, 34(3):39–50.
- Koz, A. (1995). A review of the trophic role of mesopelagic fish of the family myctophidae in the southern ocean ecosystem. *CCAMLR Science*, 2:71–77.
- Law, C. S., Rickard, G. J., Mikaloff-Fletcher, S. E., Pinkerton, M. H., Behrens, E., Chiswell, S. M., and Currie, K. (2017). Climate change projections for the surface ocean around new zealand. *New Zealand Journal of Marine and Freshwater Research*, pages 1–27.
- Lee, H.-H., Maunder, M. N., Piner, K. R., and Methot, R. D. (2012). Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research*, 125:254–261.
- Libralato, S., Christensen, V., and Pauly, D. (2006). A method for identifying key-stone species in food web models. *Ecological Modelling*, 195(3-4):153–171.
- Link, J. S. (2002). What does ecosystem-based fisheries management mean. *Fisheries*, 27(4):18–21.
- Link, J. S. (2005). Translating ecosystem indicators into decision criteria. *ICES Journal of Marine Science*, 62(3):569–576.
- Link, J. S. (2010). Adding rigor to ecological network models by evaluating a set of pre-balance diagnostics: a plea for prebal. *Ecological Modelling*, 221(12):1580–1591.
- Link, J. S. and Browman, H. I. (2017). Operationalizing and implementing ecosystem-based management. *ICES Journal of Marine Science*, 74(1):379–381.
- Link, J. S., Fulton, E. A., and Gamble, R. J. (2010). The northeast us application of atlantis: a full system model exploring marine ecosystem dynamics in a living marine resource management context. *Progress in Oceanography*, 87(1-4):214–234.

- Livingston, M., Bull, B., Stevens, D., and Bagley, N. (2002). A review of hoki and middle depth trawl surveys of the Chatham Rise, January 1992–2001. *NIWA Technical Report*, 113.
- Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., and Seidov, D. (2013). World ocean atlas 2013, volume 1: Temperature. *Mishonov Technical Ed.; NOAA Atlas NESDIS 73*, page 44.
- Long, R. D., Charles, A., and Stephenson, R. L. (2015). Key principles of marine ecosystem-based management. *Marine Policy*, 57:53–60.
- Mace, P. M. and Doonan, I. (1988). *A generalised bioeconomic simulation model for fish population dynamics*. MAFFish, NZ Ministry of Agriculture and Fisheries.
- Mackay, K. (2000). Database documentation: trawl. Technical report, NIWA Internal Report.
- Marine Stewardship Council (2014). MSC Fisheries certification requirements and guidance v2.0. page 528p.
- Masi, M., Ainsworth, C., and Jones, D. (2017). Using a gulf of mexico atlantis model to evaluate ecological indicators for sensitivity to fishing mortality and robustness to observation error. *Ecological indicators*, 74:516–525.
- Maunder, M. and Starr, P. (1995). Rock lobster standardised cpue analysis. new zealand fisheries assessment research document 95/11. 28 p. *Unpublished report held in NIWA Greta Point library, Wellington*.
- May, R. M. (1972). Will a large complex system be stable? *Nature*, 238(5364):413.
- May, R. M. (1975). Biological populations obeying difference equations: stable points, stable cycles, and chaos. *Journal of Theoretical Biology*, 51(2):511–524.

- McGregor, V. (2015). Stock assessment of ling (*Genypterus blacodes*) on the Chatham Rise (LIN 3&4) for the 2014–15 fishing year. *New Zealand Fisheries Assessment Research Document*, 82.
- McGregor, V. and Horn, P. L. (2015). *Factors affecting the distribution of highly migratory species in New Zealand waters*.
- McGregor, V. L. (2018). CRAM data and code for exploration and validation of historic Atlantis model of Chatham Rise.
- McGregor, V. L., Fulton, E. A., and Dunn, M. R. (2019a). Spawning stock recruitment creates misleading dynamics under predation release in ecosystem and multi-species models. *PeerJ*, 7:e7308.
- McGregor, V. L., Horn, P. L., Fulton, E. A., and Dunn, M. R. (2019b). From data compilation to model validation: a comprehensive analysis of a full deep-sea ecosystem model of the chatham rise. *PeerJ*, 7:e6517.
- McKenzie, A. (2016). Assessment of hoki (*Macruronus novaezelandiae*) in 2015. *New Zealand Fisheries Assessment Report*, 01:88.
- Methot Jr, R. D. and Wetzel, C. R. (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142:86–99.
- Mills, L. S., Soulé, M. E., and Doak, D. F. (1993). The keystone-species concept in ecology and conservation. *BioScience*, 43(4):219–224.
- Ministry for Primary Industries (2017). Report from the fisheries assessment plenary. stock assessments and yield estimates. *Ministry of Fisheries Report*.
- Ministry for Primary Industries (2019). Report from the fisheries assessment plenary. stock assessments and yield estimates. *Ministry of Fisheries Report*.

- Molinero, J. C., Ibanez, F., Souissi, S., Buecher, E., Dallot, S., and Nival, P. (2008). Climate control on the long-term anomalous changes of zooplankton communities in the northwestern mediterranean. *Global Change Biology*, 14(1):11–26.
- Moore, C., Mills, M., Arrigo, K., Berman-Frank, I., Bopp, L., Boyd, P., Galbraith, E., Geider, R., Guieu, C., Jaccard, S., Jickells, T., La Roche, J., Lenton, T., Mahowald, N., Marañón, E., Marinov, I., Moore, J., Nakatsuka, T., Oschlies, A., Saito, M., Thingstad, T., Tsuda, A., and Ulloa, O. (2013). Processes and patterns of oceanic nutrient limitation. *Nature Geoscience*, 6(9):701.
- Murray, A. G. and Parslow, J. S. (1997). Port phillip bay integrated model.
- Nielsen, A. and Berg, C. W. (2014). Estimation of time-varying selectivity in stock assessments using state-space models. *Fisheries Research*, 158:96–101.
- Norse, E. A., Brooke, S., Cheung, W. W., Clark, M. R., Ekeland, I., Froese, R., Gjerde, K. M., Haedrich, R. L., Heppell, S. S., Morato, T., et al. (2012). Sustainability of deep-sea fisheries. *Marine policy*, 36(2):307–320.
- O’Driscoll, R., MacGibbon, D., Fu, D., Lyon, W., and Stevens, D. (2011). A review of hoki and middle-depth trawl surveys of the Chatham Rise, January 1992–2010. *New Zealand Fisheries Assessment Report*, 47:814.
- Olsen, E., Fay, G., Gaichas, S., Gamble, R., Lucey, S., and Link, J. S. (2016). Ecosystem model skill assessment. yes we can! *PloS one*, 11(1):e0146467.
- Ortega-Cisneros, K., Cochrane, K., and Fulton, E. A. (2017). An atlantis model of the southern benguela upwelling system: Validation, sensitivity analysis and insights into ecosystem functioning. *Ecological modelling*, 355:49–63.
- Ortega-Cisneros, K., Cochrane, K. L., Fulton, E. A., Gorton, R., and Popova, E. (2018). Evaluating the effects of climate change in the southern benguela upwelling system using the atlantis modelling framework. *Fisheries Oceanography*, 27(5):489–503.

- Paine, R. T. (1969). A note on trophic complexity and community stability. *The American Naturalist*, 103(929):91–93.
- Pantus, F. J. (2007). Sensitivity analysis for complex ecosystem models.
- Patrick, W. S. and Link, J. S. (2015). Myths that continue to impede progress in ecosystem-based fisheries management. *Fisheries*, 40(4):155–160.
- Paul, L. and Horn, P. (2009). Age and growth of sea perch (*Helicolenus percooides*) from two adjacent areas off the east coast of South Island, New Zealand. *Fisheries Research*, 95(2-3):169–180.
- Pauly, D., Christensen, V., and Walters, C. (2000). Ecopath, ecosim, and ecospace as tools for evaluating ecosystem impact of fisheries. *ICES Journal of Marine Science*, 57(3):697–706.
- Pauly, D. and Watson, R. (2005). Background and interpretation of the marine trophic index as a measure of biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1454):415–423.
- Paxton, J. R., Eschmeyer, W. N., and Kirshner, D. (1998). *Encyclopedia of fishes*. Singapore: Academic Press.
- Payne, M. R., Barange, M., Cheung, W. W., MacKenzie, B. R., Batchelder, H. P., Cormon, X., Eddy, T. D., Fernandes, J. A., Hollowed, A. B., Jones, M. C., et al. (2015). Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES Journal of Marine Science*, 73(5):1272–1282.
- Pikitch, E. K., Santora, C., Babcock, E. A., Bakun, A., Bonfil, R., Conover, D. O., Dayton, P., Doukakis, P., Fluharty, D., Heneman, B., et al. (2004). Ecosystem-based fishery management.
- Plagányi, É. E. (2007). *Models for an ecosystem approach to fisheries*. Number 477. Food & Agriculture Org.

- Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R. D., Thorson, J. T., Fulton, E. A., Smith, A. D., et al. (2014). Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. *Fish and Fisheries*, 15(1):1–22.
- Pope, J. G., Rice, J. C., Daan, N., Jennings, S., and Gislason, H. (2006). Modelling an exploited marine fish community with 15 parameters—results from a simple size-based model. *ICES Journal of Marine Science*, 63(6):1029–1044.
- Porobic, J., Fulton, E. A., Parada, C., Frusher, S., Ernst, B., and Manríquez, P. (2019). The impact of fishing on a highly vulnerable ecosystem, the case of juan fernández ridge ecosystem. *PloS one*, 14(2):e0212485.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rabier, F., Klinker, E., Courtier, P., and Hollingsworth, A. (1996). Sensitivity of forecast errors to initial conditions. *Quarterly Journal of the Royal Meteorological Society*, 122(529):121–150.
- Reid, D. G., Graham, N., Suuronen, P., He, P., and Pol, M. (2016). Implementing balanced harvesting: practical challenges and other implications. *ICES Journal of Marine Science*, 73(6):1690–1696.
- Rickard, G. J., Behrens, E., and Chiswell, S. M. (2016). Cmp5 earth system models with biogeochemistry: An assessment for the southwest pacific ocean. *Journal of Geophysical Research: Oceans*, 121(10):7857–7879.
- Ricker, W. E. (1954). Stock and recruitment. *Journal of the Fisheries Board of Canada*, 11(5):559–623.
- Roberts, A. (1974). The stability of a feasible random ecosystem. *Nature*, 251(5476):607.

- Rosati, A., Miyakoda, K., and Gudgel, R. (1997). The impact of ocean initial conditions on ENSO forecasting with a coupled model. *Monthly weather review*, 125(5):754–772.
- Rose, K. A. (2012). End-to-end models for marine ecosystems: Are we on the precipice of a significant advance or just putting lipstick on a pig? *Scientia Marina*, 76(1):195–201.
- Savina, M., Grist, E., Boschetti, F., Fulton, E., and McDonald, A. (2005). Implementation of the Atlantis ecological model in the Westernport scoping study.
- Schaefer, M. B. (1957). Some considerations of population dynamics and economics in relation to the management of the commercial marine fisheries. *Journal of the Fisheries Board of Canada*, 14(5):669–681.
- Shin, Y. and Cury, P. (1999). Osmose: a multispecies individual-based model to explore the functional role of biodiversity in marine ecosystems. *Ecosystem approaches for fisheries management, University of Alaska Sea Grant, Fairbanks*, pages 593–607.
- Shin, Y.-J. and Cury, P. (2004). Using an individual-based model of fish assemblages to study the response of size spectra to changes in fishing. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(3):414–431.
- Shin, Y.-J., Houle, J. E., Akoglu, E., Blanchard, J. L., Bundy, A., Coll, M., Demarcq, H., Fu, C., Fulton, E. A., Heymans, J. J., et al. (2018). The specificity of marine ecological indicators to fishing in the face of environmental change: A multi-model evaluation. *Ecological Indicators*, 89:317–326.
- Shin, Y.-J., Shannon, L. J., Bundy, A., Coll, M., Aydin, K., Bez, N., Blanchard, J. L., Borges, M. d. F., Diallo, I., Diaz, E., et al. (2010). Using indicators for evaluating, comparing, and communicating the ecological status of exploited marine ecosystems. 2. setting the scene. *ICES Journal of Marine Science*, 67(4):692–716.

- Skern-Mauritzen, M., Ottersen, G., Handegard, N. O., Huse, G., Dingsør, G. E., Stenseth, N. C., and Kjesbu, O. S. (2016). Ecosystem processes are rarely included in tactical fisheries management. *Fish and Fisheries*, 17(1):165–175.
- Smith, A. D., Brown, C. J., Bulman, C. M., Fulton, E. A., Johnson, P., Kaplan, I. C., Lozano-Montes, H., Mackinson, S., Marzloff, M., Shannon, L. J., et al. (2011). Impacts of fishing low-trophic level species on marine ecosystems. *Science*, 333(6046):1147–1150.
- Smith, D. C., Fulton, E. A., Apfel, P., Cresswell, I. D., Gillanders, B. M., Haward, M., Sainsbury, K. J., Smith, A. D., Vince, J., and Ward, T. M. (2017). Implementing marine ecosystem-based management: lessons from australia. *ICES Journal of Marine Science*, 74(7):1990–2003.
- Smith, M. D., Fulton, E. A., and Day, R. W. (2015). Using an atlantis model of the southern benguela to explore the response of ecosystem indicators for fisheries management. *Environmental Modelling & Software*, 69:23–41.
- Stecken, M. and Failler, P. (2016). Ecosystem approach to fisheries and marine ecosystem modelling: review of current approaches. *Journal of Fisheries & Livestock Production*, 4(4).
- Steenbeek, J., Corrales, X., Platts, M., and Coll, M. (2018). Ecosampler: A new approach to assessing parameter uncertainty in ecopath with ecosim. *SoftwareX*, 7:198–204.
- Stevens, D., Hurst, R., and Bagley, N. (2011a). Feeding habits of New Zealand fishes: a literature review and summary of research trawl database records 1960 to 2000. *New Zealand Aquatic Environment and Biodiversity Report*, 85.
- Stevens, D., O’Driscoll, R., Ballara, S., and Ladroit, Y. (2017). Trawl survey of hoki and middle-depth species on the Chatham Rise, January 2016 (TAN1601). *New Zealand Fisheries Assessment Report*, 8:131.

- Stevens, D., Smith, M., Grimes, P., Devine, J., Sutton, C., MacGibbon, D., and Maolagáin, C. (2010). Age, growth, and maturity of four New Zealand rattail species. *NZ Aquat. Environ. Biodiv. Rept*, (59):39.
- Stevens, D. W. and Dunn, M. R. (2011). Different food preferences in four sympatric deep-sea macrourid fishes. *Marine biology*, 158(1):59–72.
- Stevens, D. W., Hurst, R. J., and Bagley, N. W. (2011b). *Feeding habits of New Zealand fishes: a literature review and summary of research trawl database records 1960 to 2000*. Ministry of Fisheries.
- Stewart, B. and Smith, D. (1994). Development of methods to age commercially important dories and oreos. *Corporation, Project 91/36*.
- Stow, C. A., Jolliff, J., McGillicuddy Jr, D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A., Rose, K. A., and Wallhead, P. (2009). Skill assessment for coupled biological/physical models of marine systems. *Journal of Marine Systems*, 76(1-2):4–15.
- Sturludottir, E., Desjardins, C., Elvarsson, B., Fulton, E. A., Gorton, R., Logemann, K., and Stefansson, G. (2018). End-to-end model of Icelandic waters using the Atlantis framework: Exploring system dynamics and model reliability. *Fisheries Research*, 207:9–24.
- Sutton, C. (1999). Ageing methodology, growth parameters, and estimates of mortality for giant stargazer (*Kathetostoma giganteum*) from the east and south coasts of the South Island. New Zealand Fisheries Assessment Research Document 99/15. 19 p. *Unpublished report held in NIWA library, Wellington, 22*.
- Tang, S. and Allesina, S. (2014). Reactivity and stability of large ecosystems. *Frontiers in Ecology and Evolution*, 2:21.
- Thorson, J. T. (2019). Guidance for decisions using the vector autoregressive spatio-temporal (vast) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research*, 210:143–161.

- Tregonning, K. and Roberts, A. (1979). Complex systems which evolve towards homeostasis. *Nature*, 281(5732):563.
- Tuck, I. (2016). Characterisation and a length-based assessment model for scampi (*Metanephrops challengeri*) on the Mernoo Bank (SCI 3).
- Tuck, I. D., Cole, R., and Devine, J. A. (2009). *Ecosystem indicators for New Zealand fisheries*. Ministry of Fisheries.
- Uddstrom, M. J. and Oien, N. A. (1999). On the use of high-resolution satellite data to describe the spatial and temporal variability of sea surface temperatures in the New Zealand region. *Journal of Geophysical Research: Oceans*, 104(C9):20729–20751.
- Von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws. ii). *Human biology*, 10(2):181–213.
- Wahle, R. A. (2003). Revealing stock–recruitment relationships in lobsters and crabs: is experimental ecology the key? *Fisheries Research*, 65(1-3):3–32.
- Walters, C., Christensen, V., and Pauly, D. (1997). Structuring dynamic models of exploited ecosystems from trophic mass-balance assessments. *Reviews in fish biology and fisheries*, 7(2):139–172.
- Walters, C., Pauly, D., Christensen, V., and Kitchell, J. F. (2000). Representing density dependent consequences of life history strategies in aquatic ecosystems: Ecosim ii. *Ecosystems*, 3(1):70–83.
- Weijerman, M., Fulton, E. A., Kaplan, I. C., Gorton, R., Leemans, R., Mooij, W. M., and Brainard, R. E. (2015). An integrated coral reef ecosystem model to support resource management under a changing climate. *PloS one*, 10(12):e0144165.
- Weijerman, M., Kaplan, I. C., Fulton, E. A., Gordon, B., Grafeld, S., and Brainard, R. (2014). Design and parameterization of a coral reef ecosystem model for guam.

- Ye, Y. (2000). Is recruitment related to spawning stock in penaeid shrimp fisheries? *ICES Journal of Marine Science*, 57(4):1103–1109.
- Young, J., Bulman, C., Blaber, S., and Wayte, S. (1988). Age and growth of the lanternfish *Lampanyctodes hectoris* (*Myctophidae*) from eastern Tasmania, Australia. *Marine Biology*, 99(4):569–576.
- Young, J. W., Lansdell, M. J., Campbell, R. A., Cooper, S. P., Juanes, F., and Guest, M. A. (2010). Feeding ecology and niche segregation in oceanic top predators off eastern australia. *Marine Biology*, 157(11):2347–2368.
- Zweng, M., Reagan, J., Antonov, J., Locarnini, R., Mishonov, A., Boyer, T., Garcia, H., Baranova, O., Johnson, D., D.Seidov, and Biddle, M. (2013). World ocean atlas 2013, volume 2: Salinity. *Mishonov Technical Ed.; NOAA Atlas NESDIS 73*, page 39.

Chapter 7

Appendices

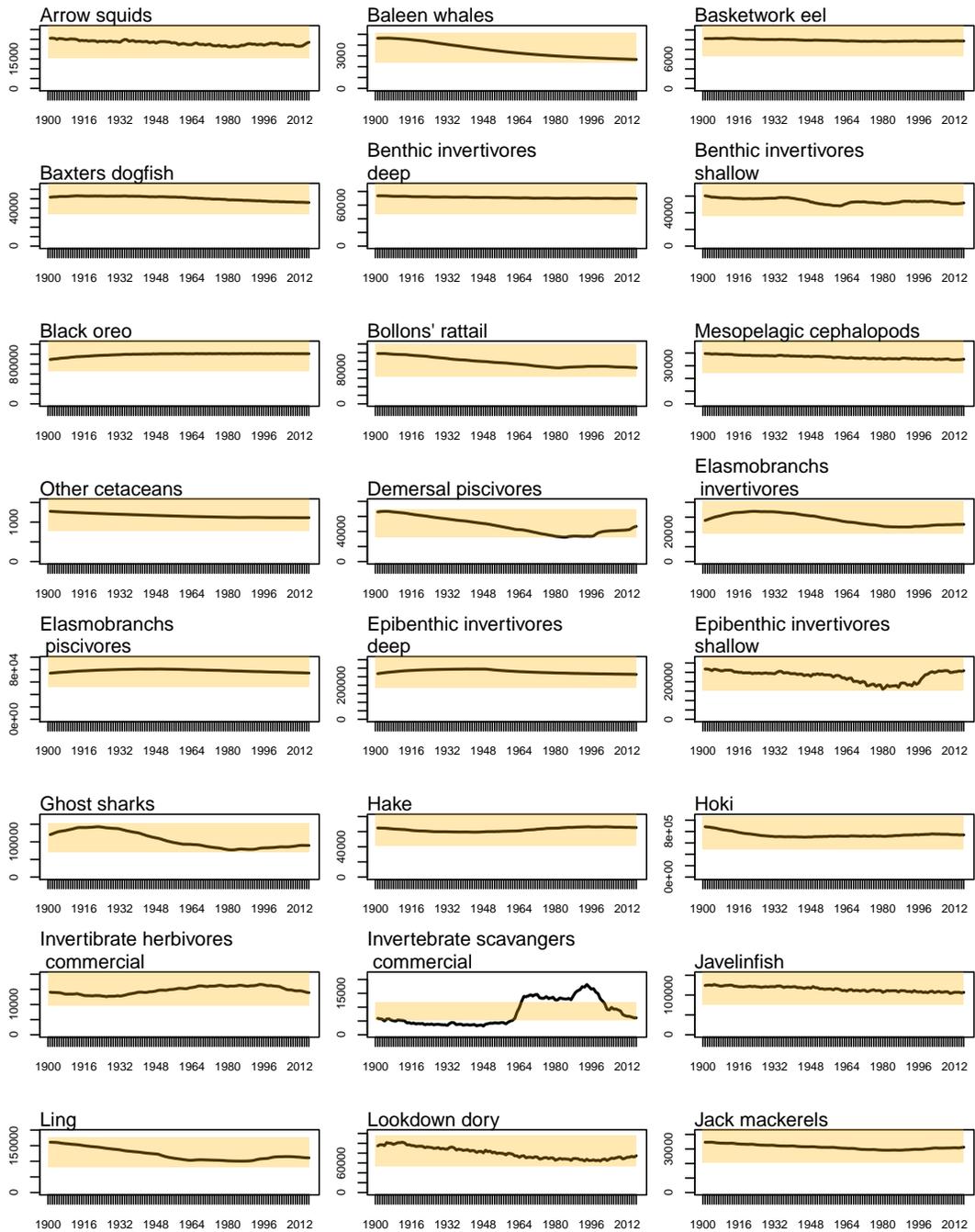
Appendix A: Simulated biomass by species group from no-fishing model

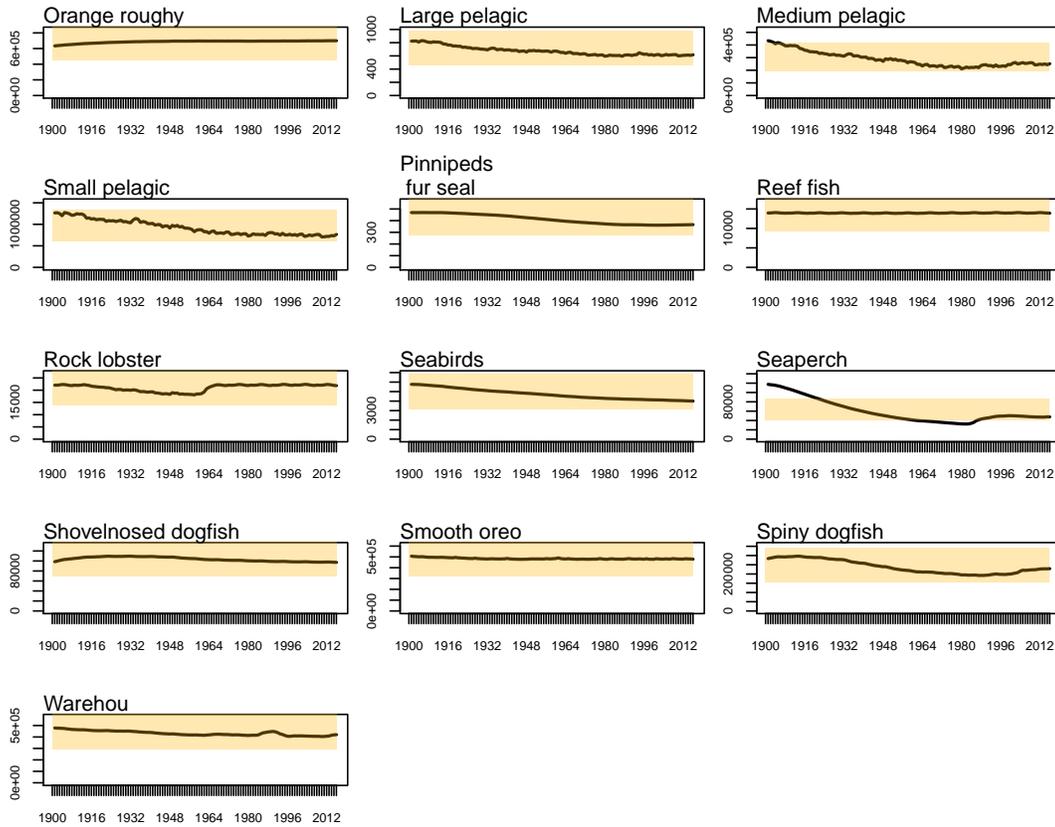
Simulated biomass from the un-fished model (black line) with 95% confidence intervals based on 20% CVs (Coefficient of Variation) shaded orange by species group.

7.0.1 Additional background:

Biomass trajectories were referred to in Chapter 2. These are from the base Chatham Rise Atlantis model were produced to check the model's stability, following the recommendations of Kaplan and Marshall (2016). The base model presented in this study has oceanographic variables repeated on a nine year cycle, rather than just one year, and this could bring in some instability. The fluctuating biomass of the invertebrate scavengers commercial species group (primarily scampi) could be due to oceanographic variability. Most of the biomass trajectories remain within the 20%CV bands over the 1900–2015 timespan simulated. The timespan inspected is over 100 years due to the long life-span of some of the species, such as orange roughy (*Hoplostethus atlanticus*) and the oreos (*Allocyttus niger*, *Pseudocyttus maculatus*) that live for more than 100 years. There was a burn-in period of 35 years, which

doesn't come close to the lifespan of many species, but did seem to be long enough for the system dynamics to settle.





Appendix B: Size-at-age

Size-at-age using values based on literature (Table 2.6) where available (orange shaded shows 95% confidence intervals using CV 10%) and from CRAM simulated years 1900–2015 (boxplots).

7.0.2 Additional background:

These figures are referred to in Chapter 2. They show the size-at-age of age-structured species groups from the base Chatham Rise Atlantis model outputs from 1900–2015, compared to the von Bertalanffy curves with parameters from the literature. The variability from the model outputs are shown in the distributions of the boxplots. As the base model was designed to be fairly stable, there is not much variability in most groups, and almost none in some (apparent with flat horizontal lines rather than boxes). Evaluating base Atlantis models for stability in growth has been done before (e.g. Weijerman et al. (2014)), but only as a time series, which cannot be easily compared to a von Bertalanffy curve. The outputs used from Atlantis for these figures were the average weights (reserve mg N and structural mg N) for each age class in each cell at each timestep. I converted these to grams using the conversion used in Atlantis (Equation 7.1).

$$W_{grams} = (R_N + S_N) \times 20 \times 1e - 3 \times 5.7 \quad (7.1)$$

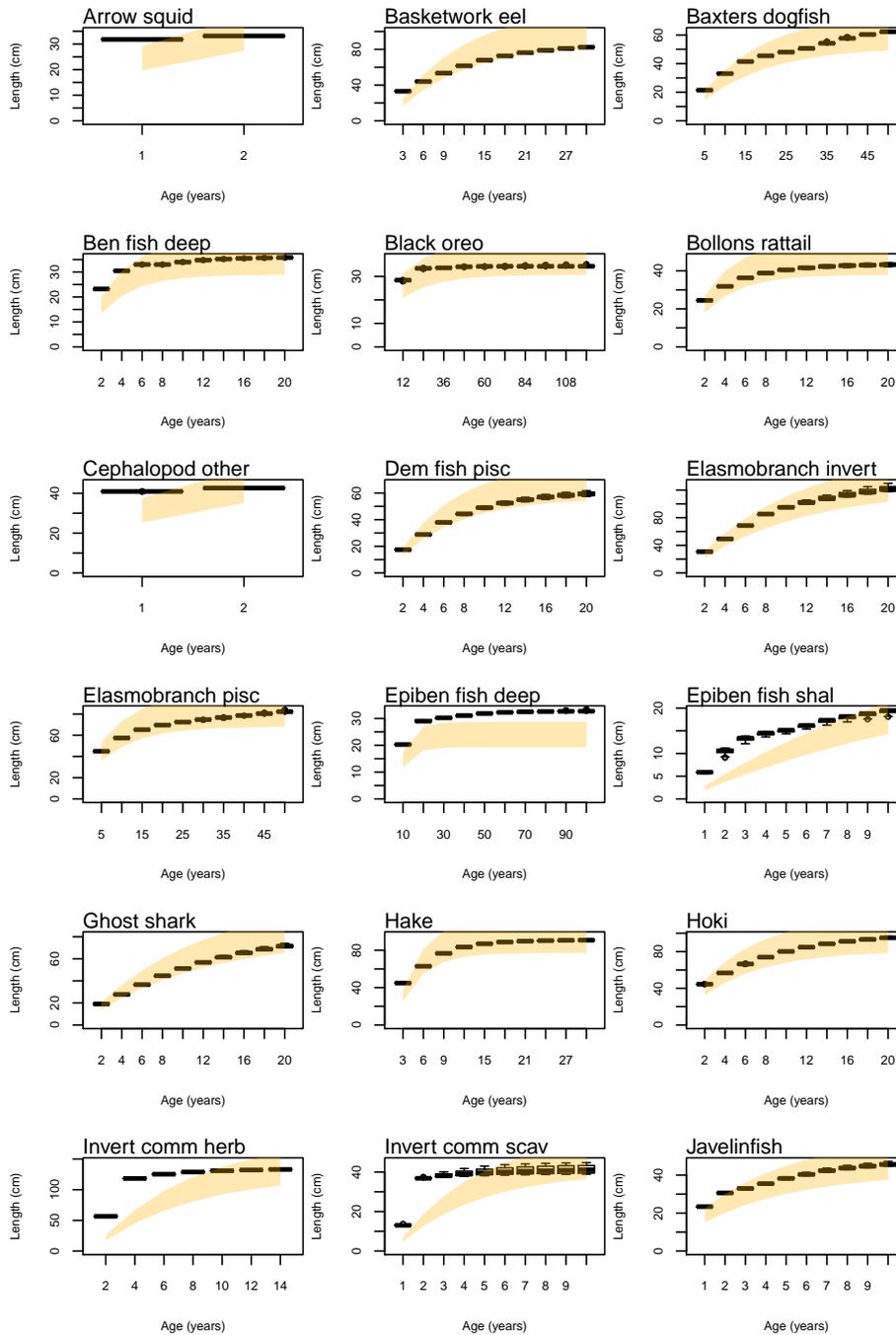
where

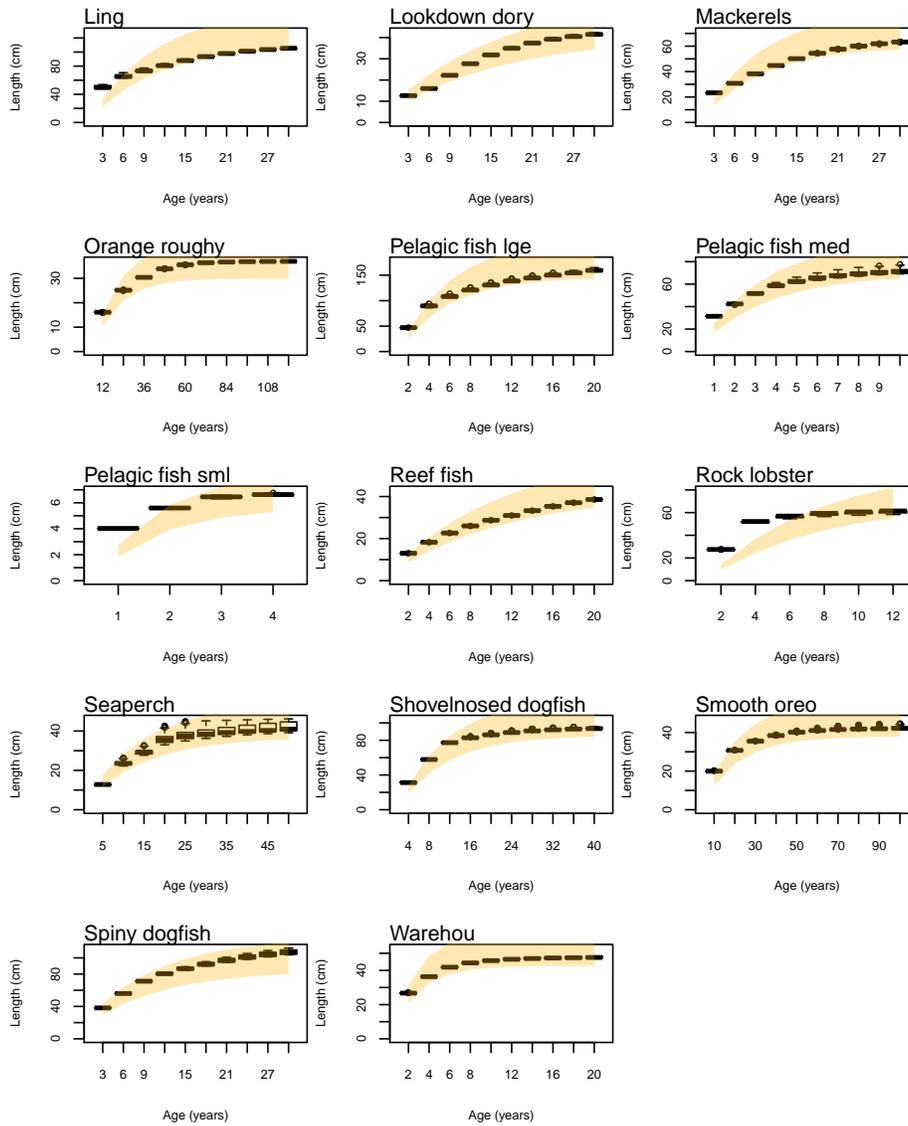
20 is the wet-dry weight conversion,

1e-3 converts mg to g,

5.7 is the Redfield ratio for converting nitrogen to carbon.

The weights were then converted to length using $L = aW^b$, with a and b parameters from the literatures.



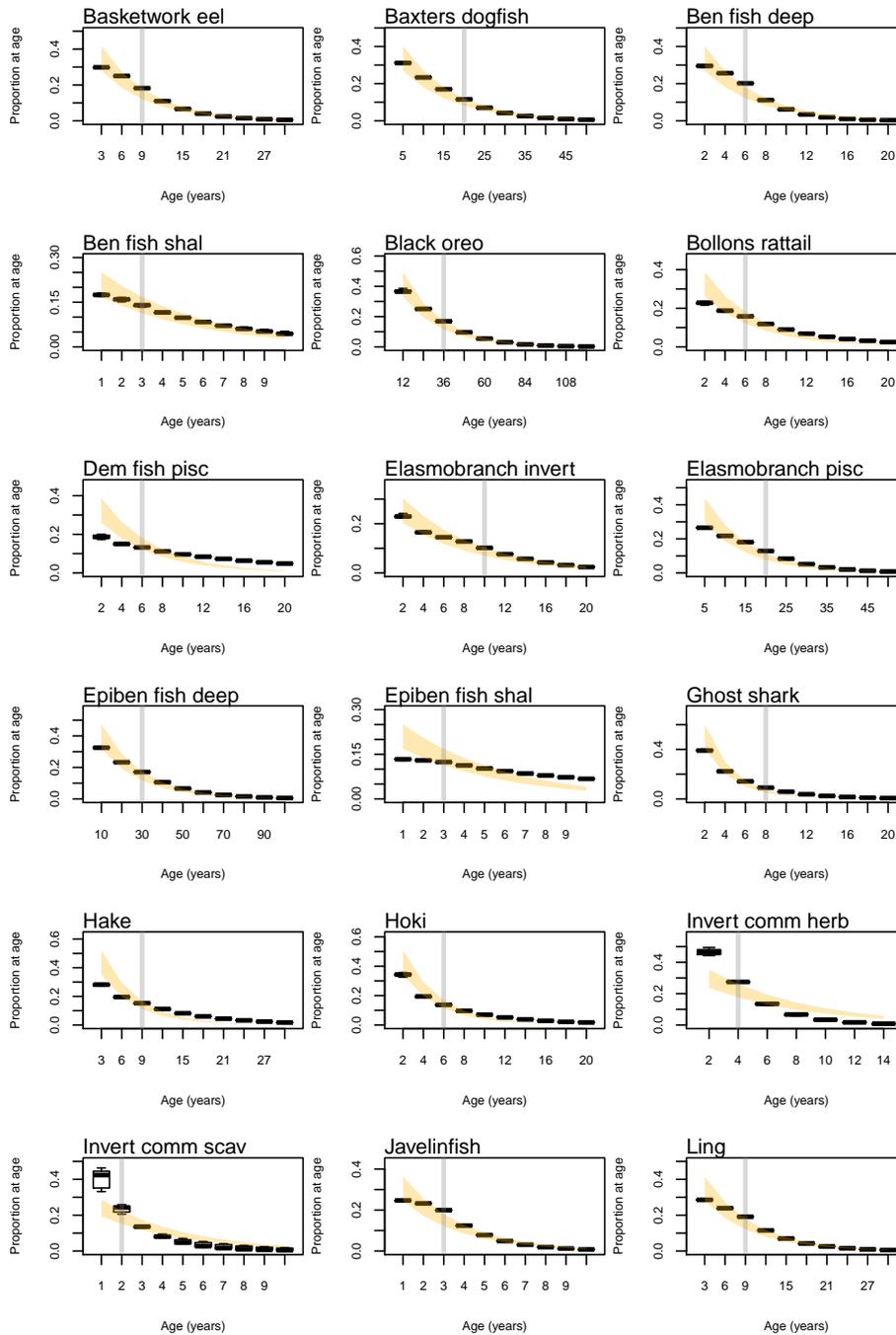


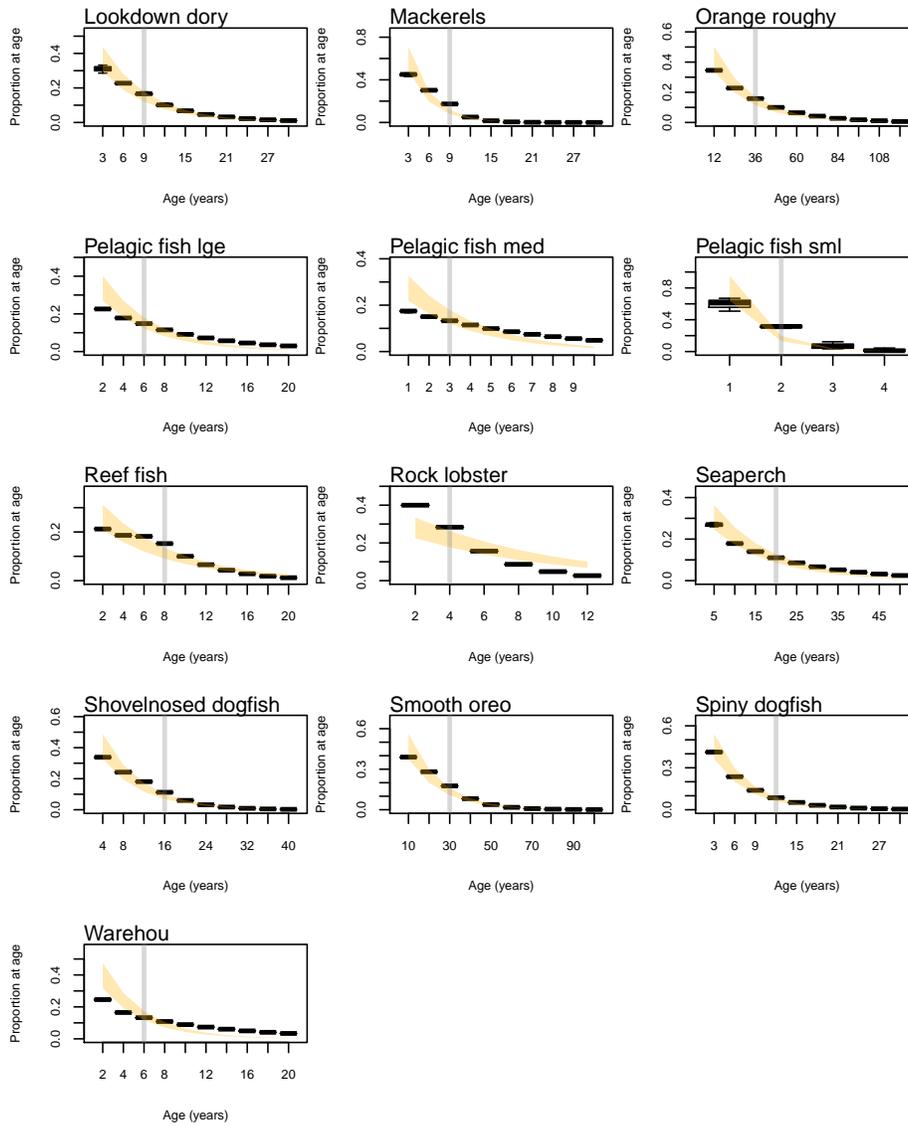
Appendix C: Proportion-at-age

Proportions at age using M based on literature (Table 2.6) where available (orange shaded shows 95% confidence intervals using CV 10%) and from CRAM simulated years 1900–2015 (boxplots).

7.0.3 Additional background:

These figures are referred to in Chapter 2. They show the proportions-at-age of age-structured species groups from the base Chatham Rise Atlantis model outputs from 1900–2015, compared to the exponential decay curve using M (instantaneous natural mortality) values from the literature. As these are from the base model with no fishing, they reflect natural mortality in the model, and hence should be similar to the curve from the literature if I have adequately captured the population dynamics of each species. The variability from the model outputs are shown in the distributions of the boxplots. As the base model was designed to be fairly stable, there is not much variability in most groups, and almost none in some (apparent with flat horizontal lines rather than boxes). Evaluating exponential decay curves was mentioned by Weijerman et al. (2014), but showed as timeseries plots of densities by age class, within which the decay curve is not apparent and its shape is not clear.



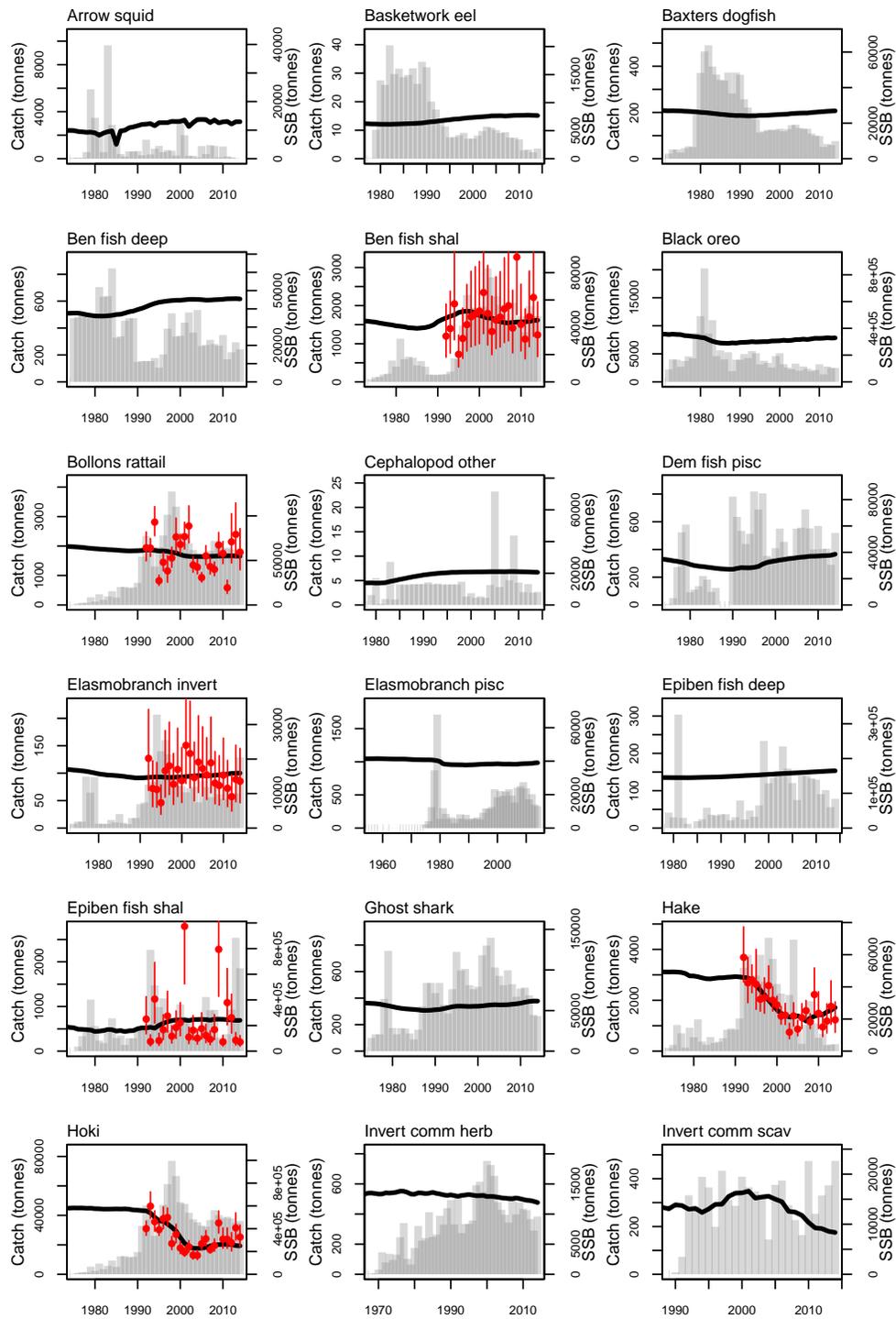


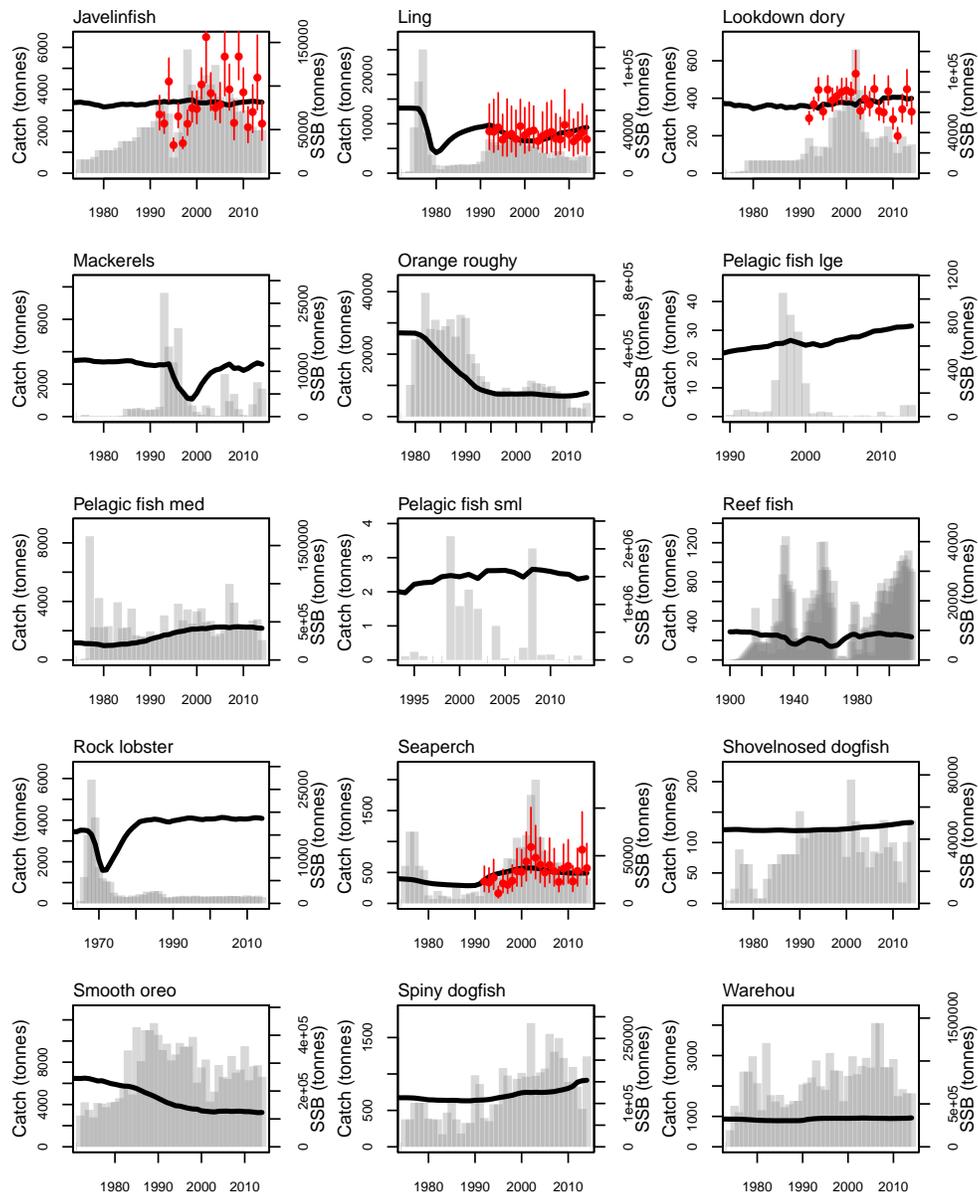
Appendix D: Observed vs estimated

Observed biomass estimated from trawl surveys (red), estimated biomass from CRAM (black) and forced catch history (grey) for all groups with trawl survey estimates.

7.0.4 Additional background:

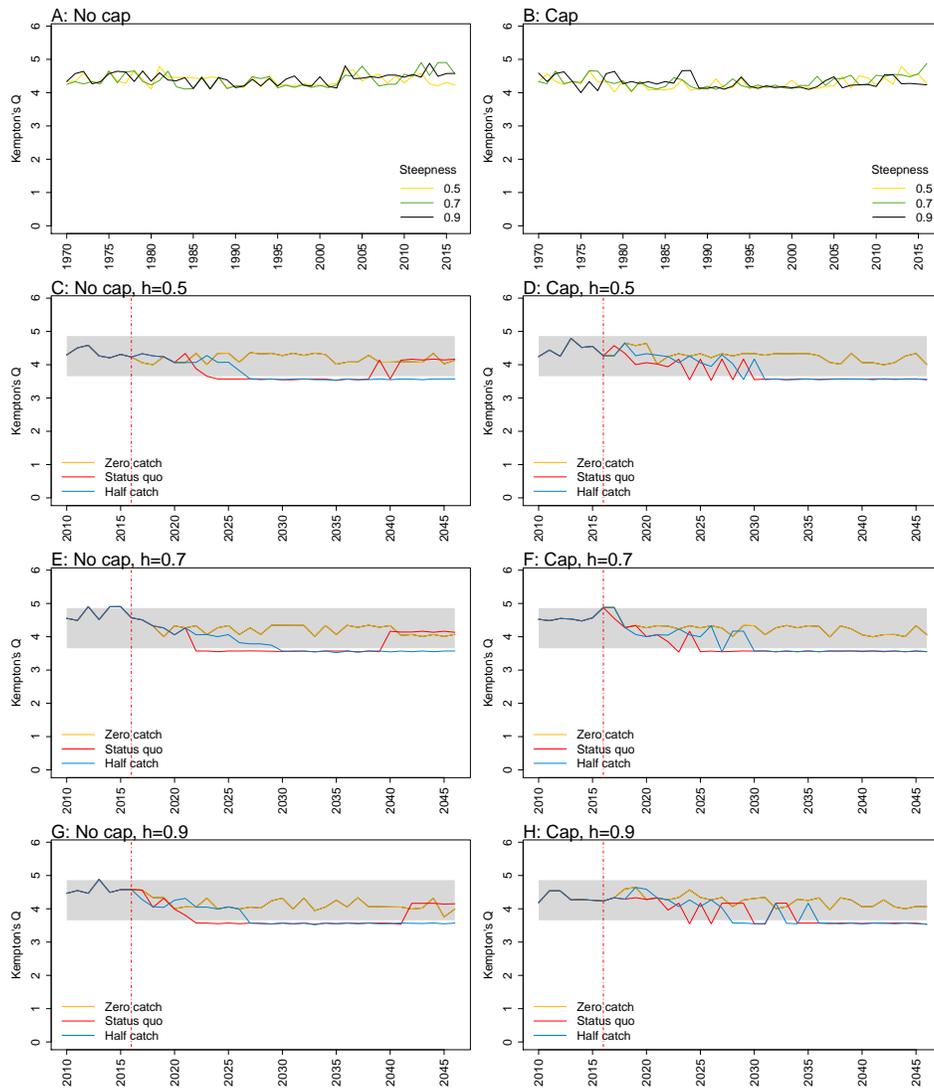
These figures are referred to in Chapter 2. They show biomass trajectories from the base Chatham Rise Atlantis model with fishing included, overlaid with both the forced historic catches and the trawl survey estimated abundance where available. The trawl survey is designed for hoki (*Macruronus novaezelandiae*) and hake (*Merluccius australis*) and these have the tightest CVs match the Atlantis model outputs well. Ling (*Genypterus blacodes*) and seaperch (*Helicolenus*) also match well, although the survey CVs are larger. I did not have to do much tuning of the model once the base model had been calibrated without fishing. Once the species groups had appropriate growth rates, mortality rates, and trophic interactions, the responses to fishing followed to be appropriate. Rock lobsters (*Jasus edwardsii*) were an exception, as they crashed under fishing in the initial runs. I eased off the predation on rock lobsters until they no longer crashed under fishing, but still responded.





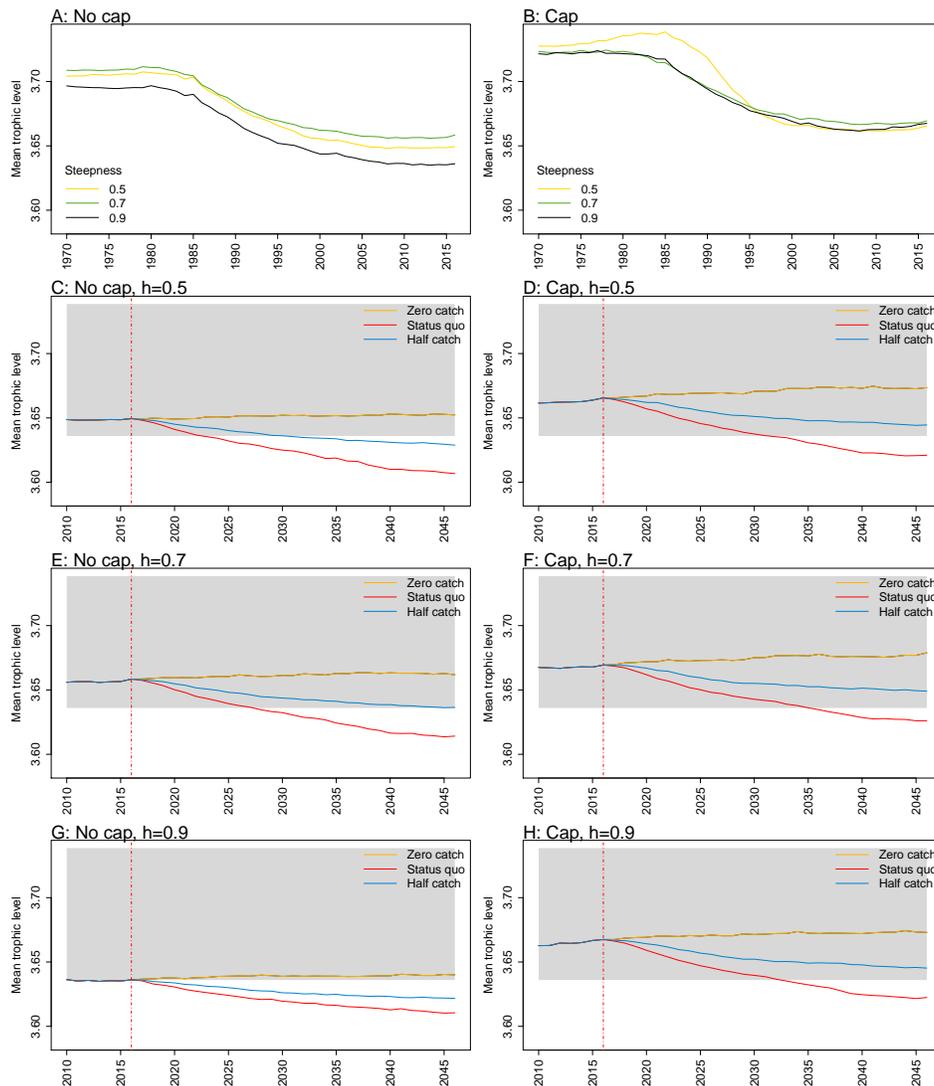
Appendix E: Kempton's Q

Kempton's Q calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (left), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (top) and h set at 0.5 (second to top), 0.7 (second to bottom), and 0.9 (bottom) for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970.



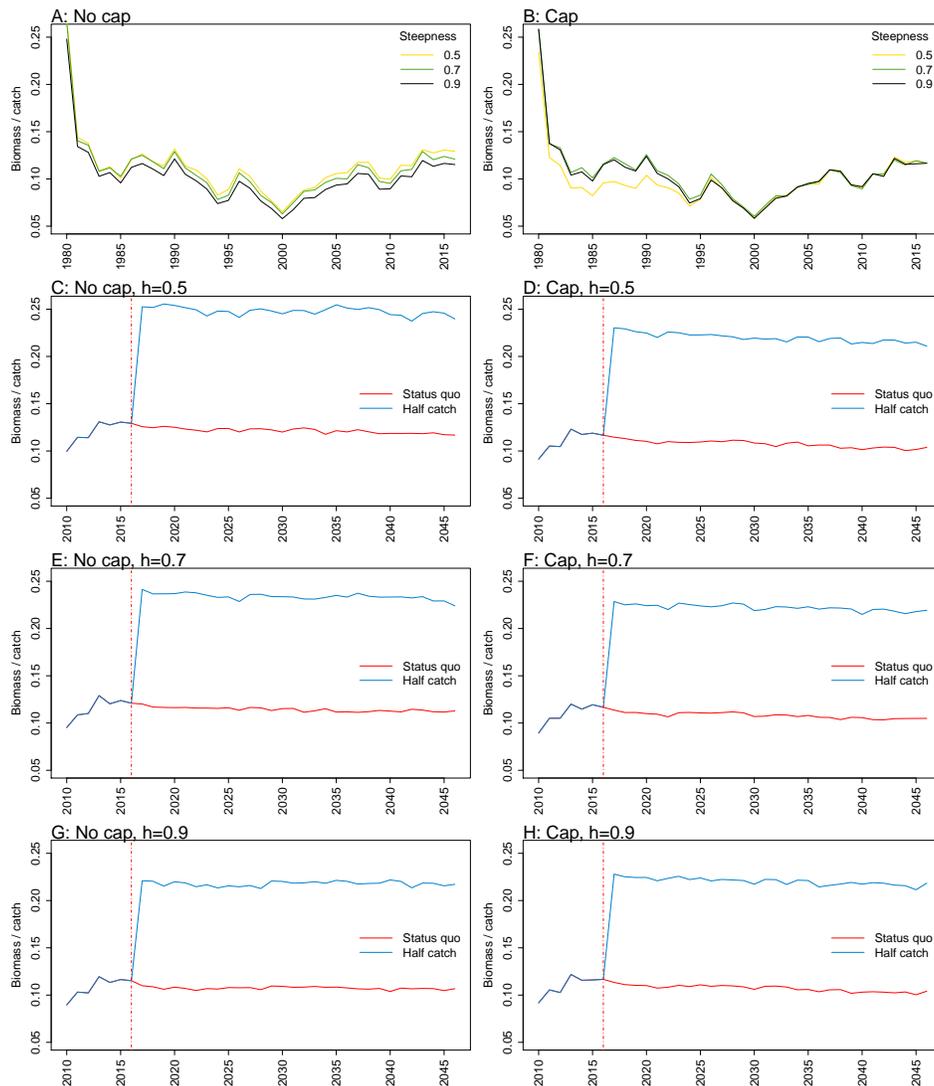
Appendix F: Mean trophic level

Mean trophic level of age-structured species groups calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (left), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (top) and h set at 0.5 (second to top), 0.7 (second to bottom), and 0.9 (bottom) for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970.



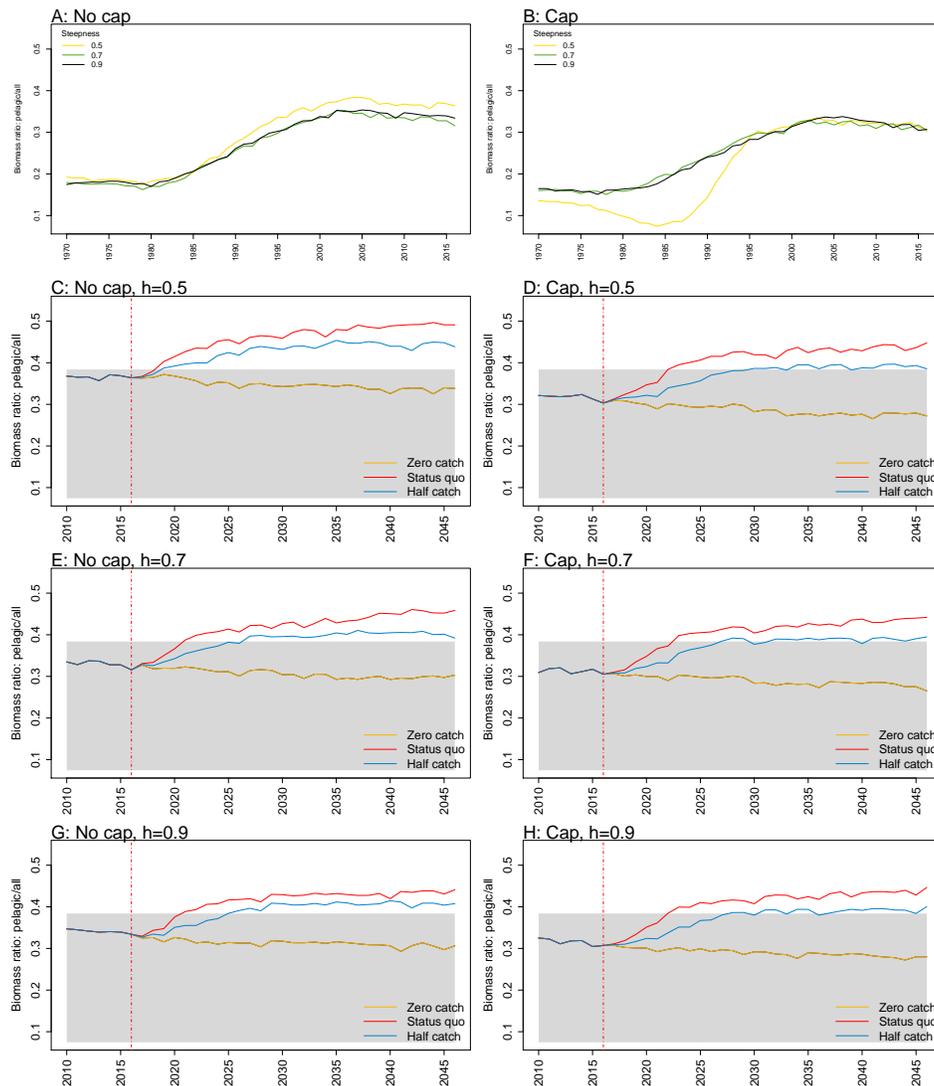
Appendix G: Biomass over catch

Biomass of age-structured species groups over catch calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (left), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (top) and h set at 0.5 (second to top), 0.7 (second to bottom), and 0.9 (bottom) for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970.



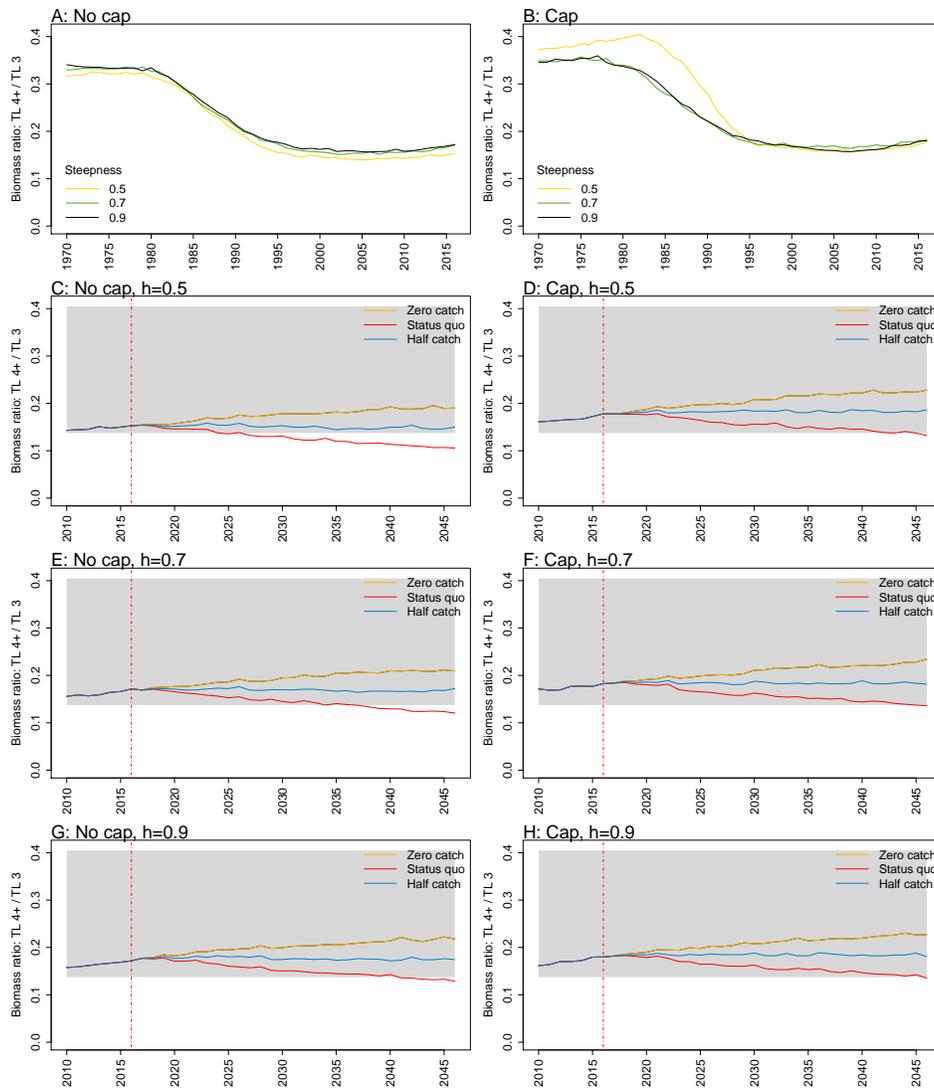
Appendix H: Biomass pelagic over all

Biomass of pelagic fishes over biomass of all age-structured species groups calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (left), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (top) and h set at 0.5 (second to top), 0.7 (second to bottom), and 0.9 (bottom) for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970.



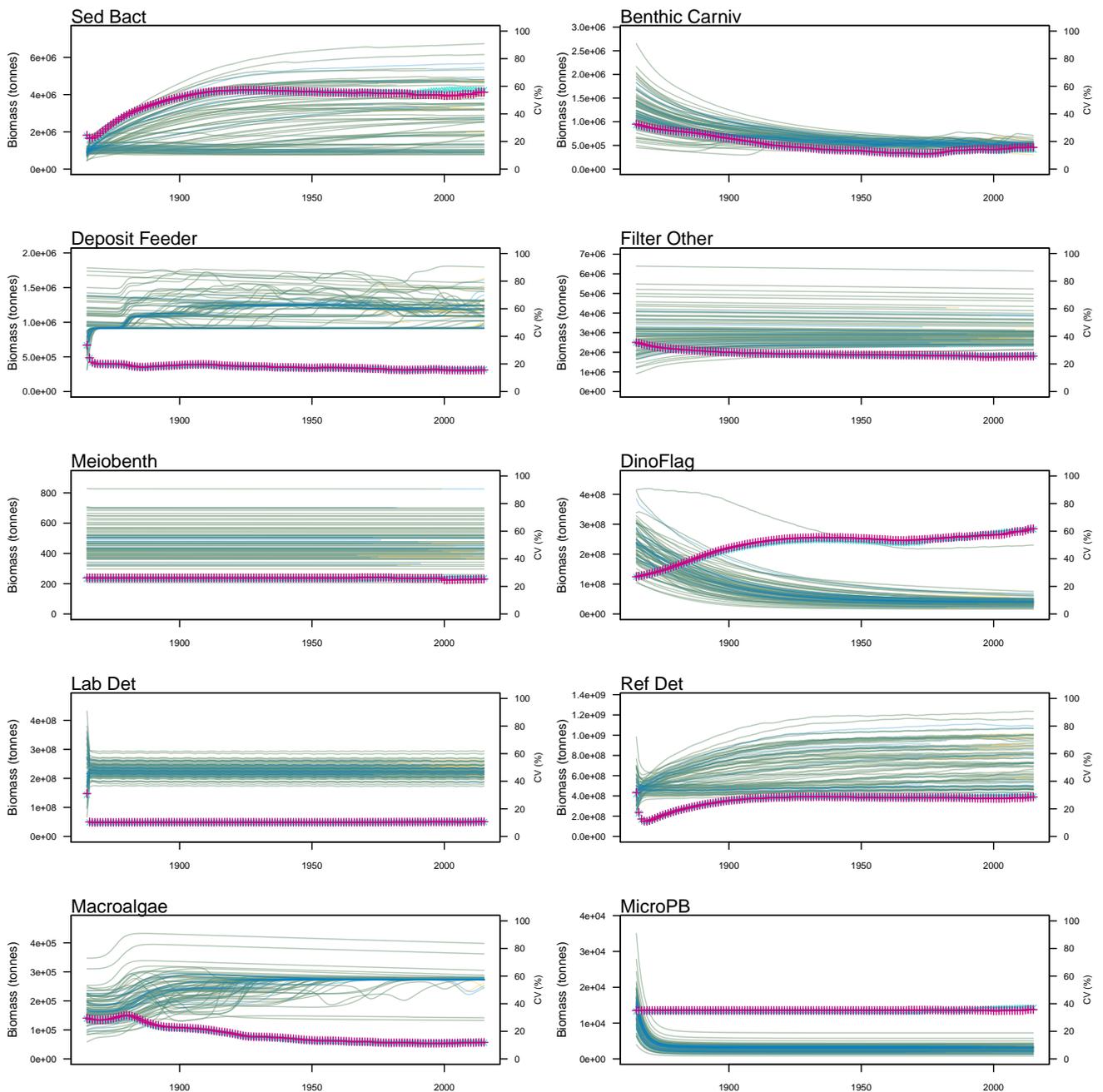
Appendix I: Biomass trophic level 4 and higher over trophic level 3

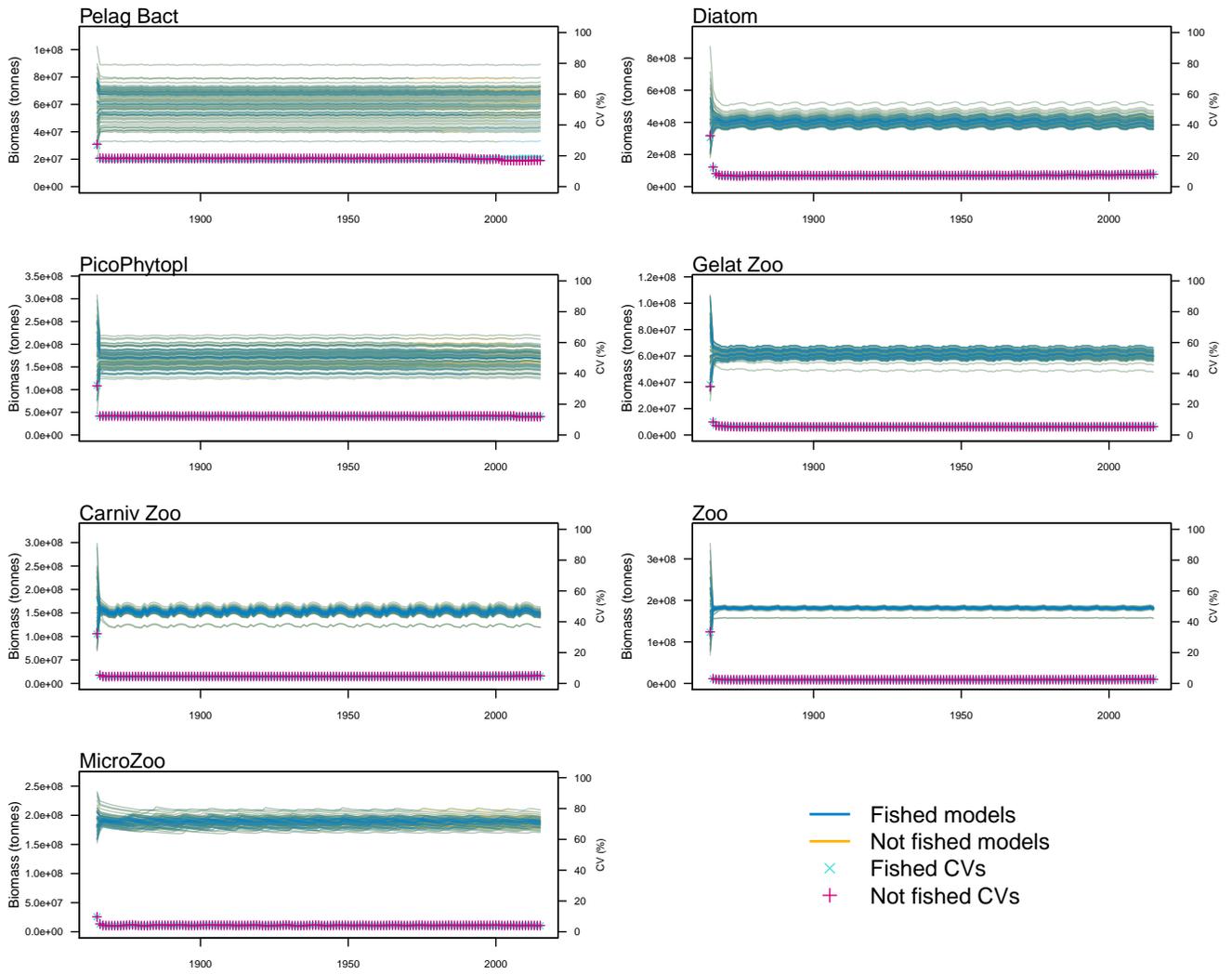
Biomass ratio of trophic level 4 and higher over trophic level 3 calculated from Chatham Rise Atlantis model simulations with no cap on recruitment (left), recruitment capped at R_0 (right), recruitment steepness values $h \in (0.5, 0.7, 0.9)$ for the 1970–2016 hindcast period (top) and h set at 0.5 (second to top), 0.7 (second to bottom), and 0.9 (bottom) for myctophids, with three catch scenarios: 1.) Zero catch; 2.) *Status quo* catch; 3.) Half catch, for the 2016–2046 projection period (bottom). The red dashed vertical line marks the last hindcast year, 2016. The grey region shows the range of values from the period 1900–1970.

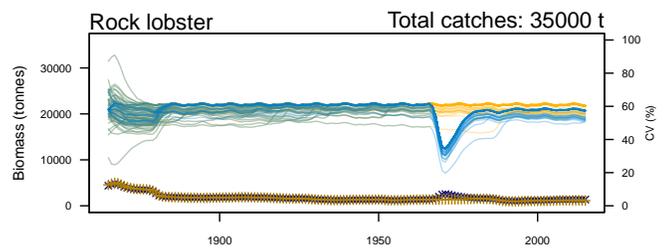
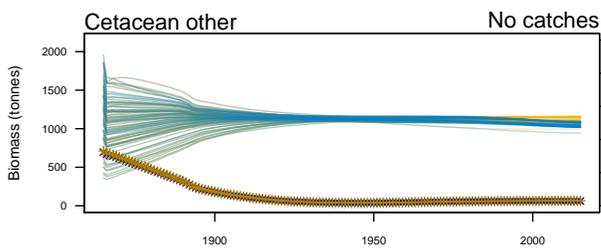
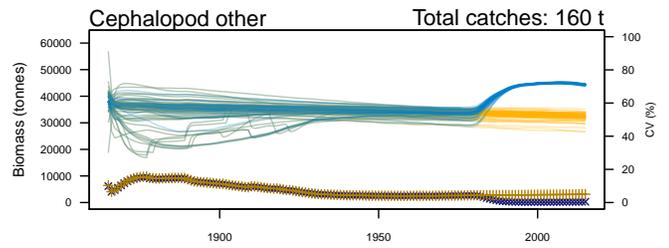
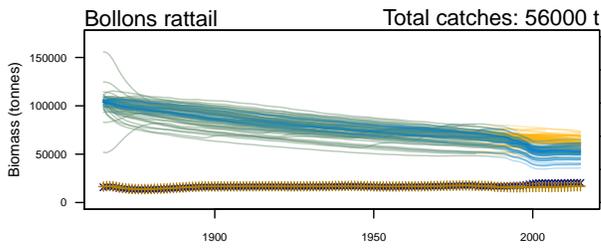
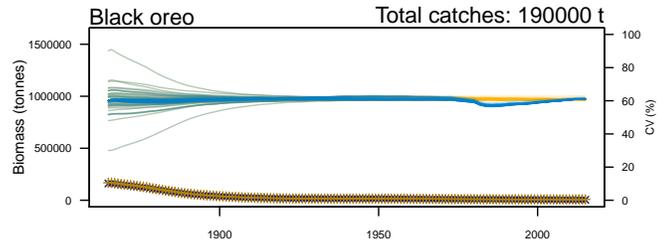
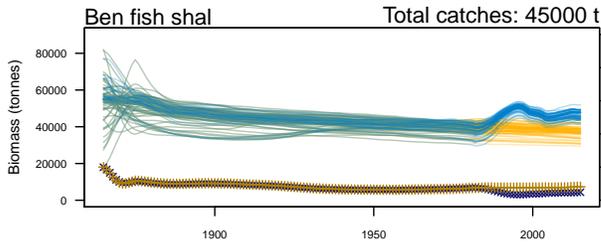
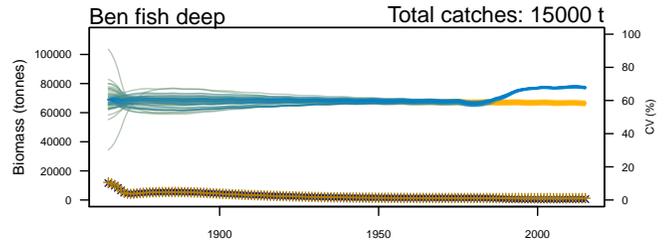
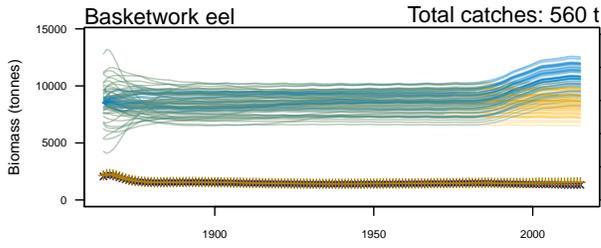
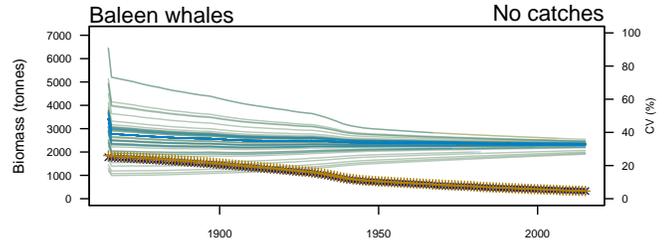
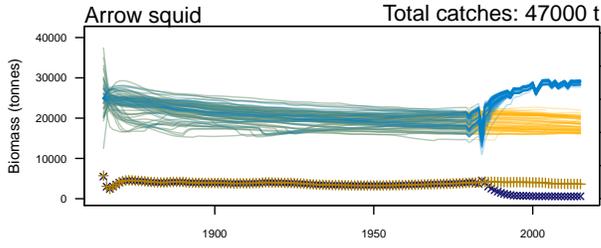


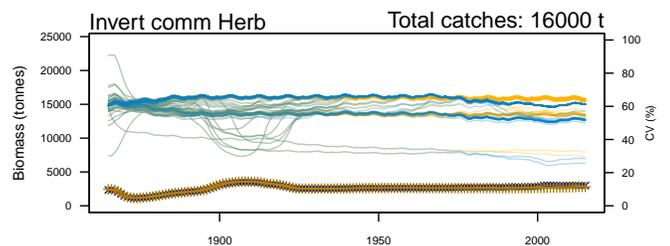
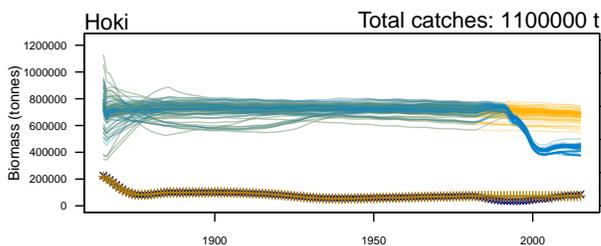
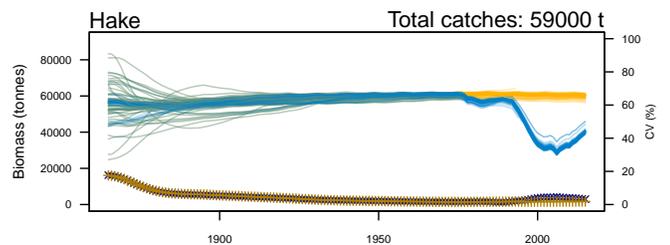
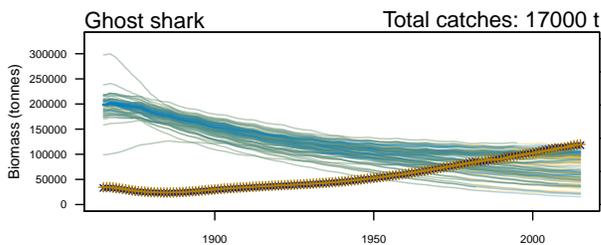
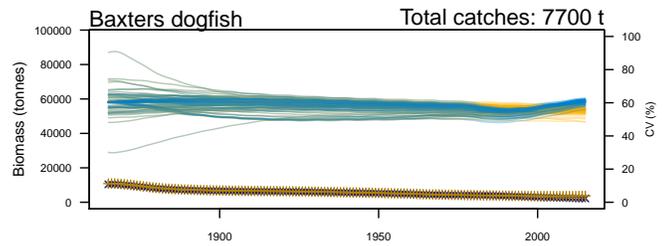
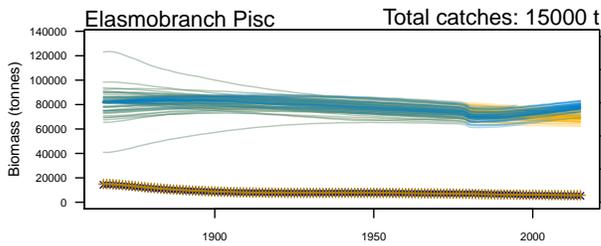
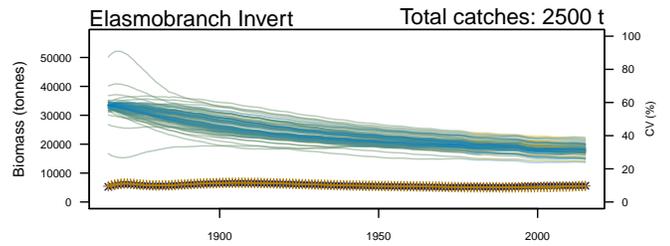
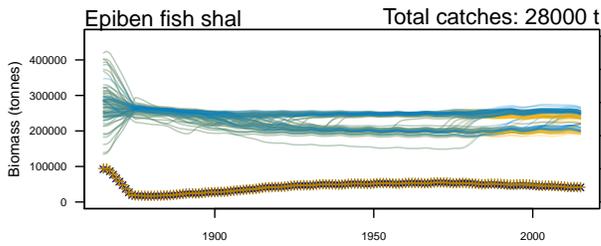
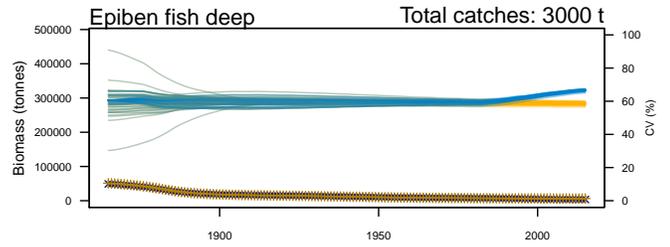
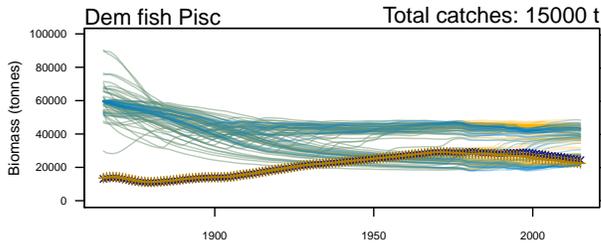
Appendix J: Biomass trajectories

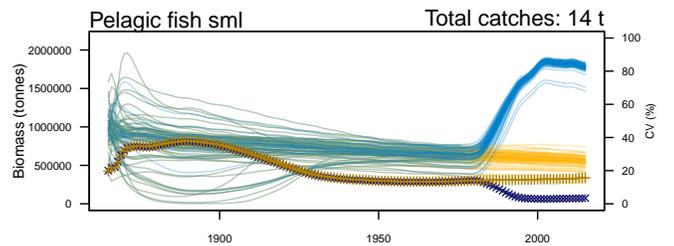
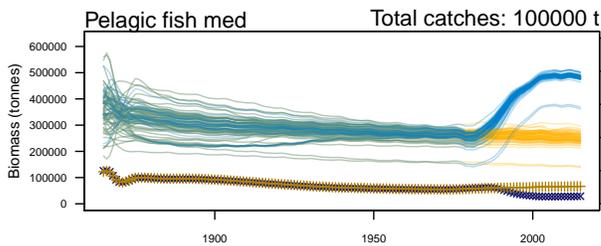
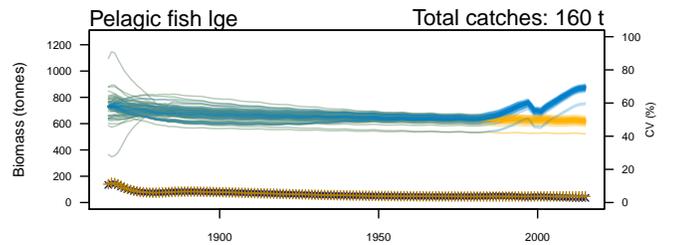
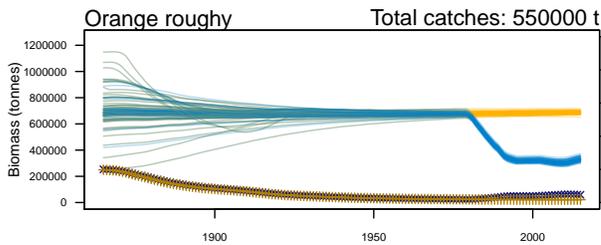
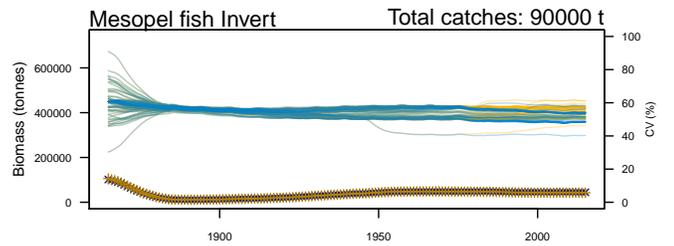
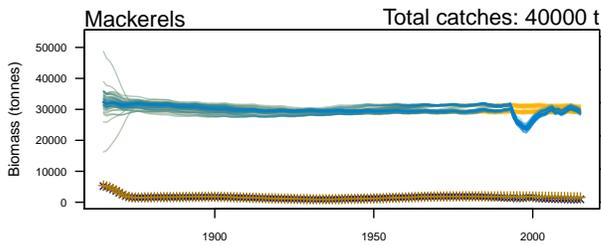
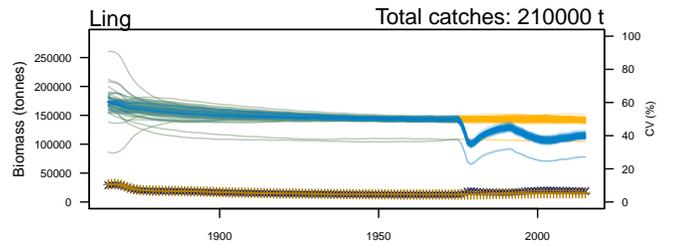
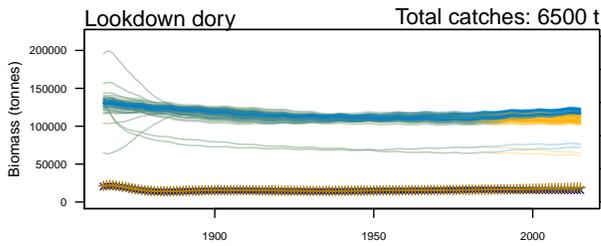
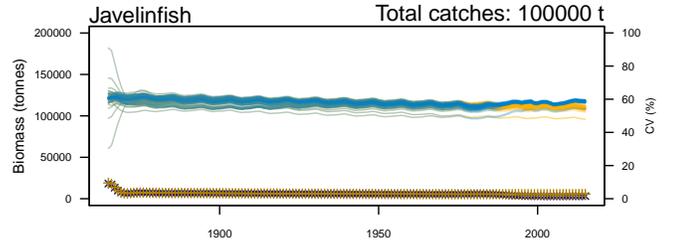
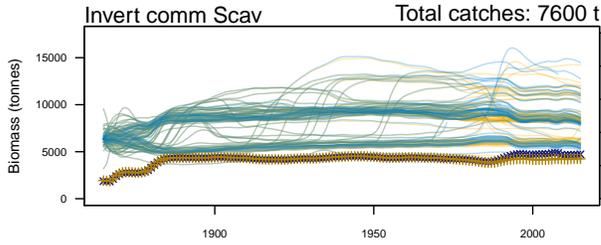
Biomass trajectories from models with fishing included (blue lines) and no fishing (orange lines) for each species group, with CVs from across the model runs by time from fished models (aqua asterisks) and unfished models (cerise asterisks) overlaid and using the right-hand axis.

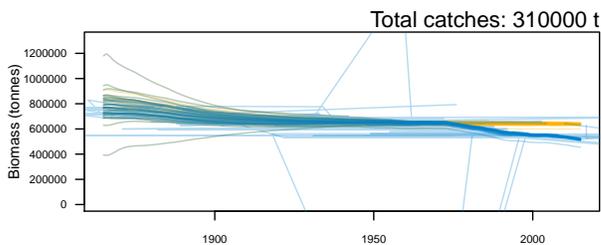
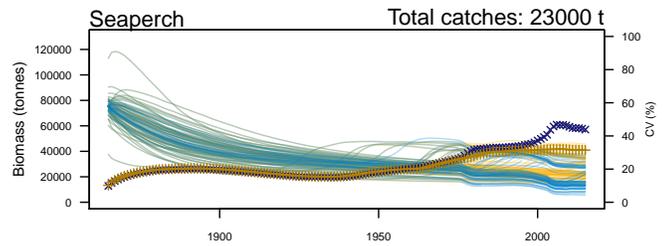
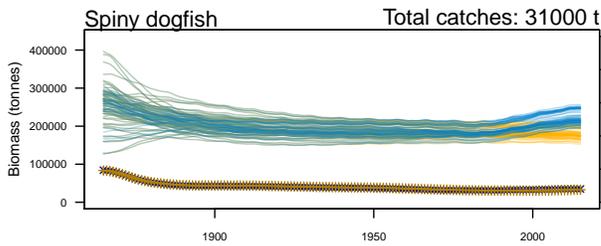
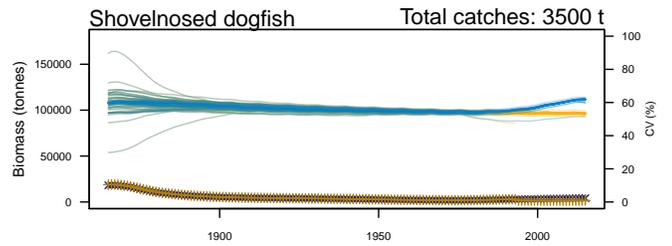
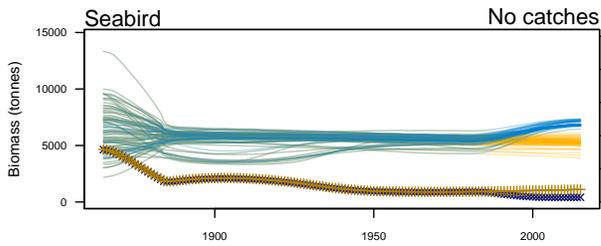
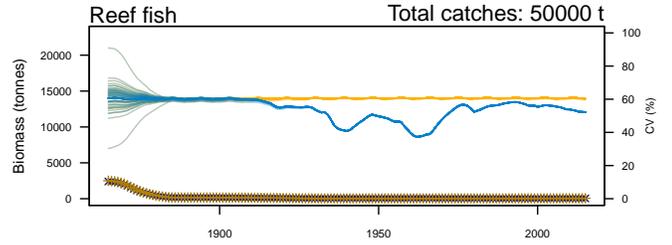
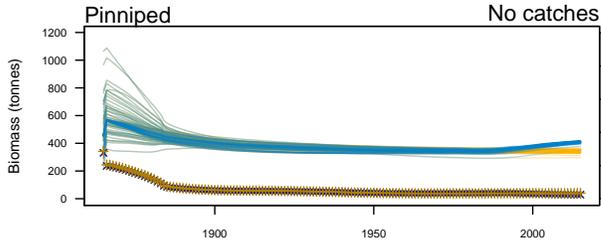












Appendix K: Parameters tuned during model calibration

Table 7.1: Parameters tuned and/or edited during model calibration.

Parameter	Description	General purpose of tuning
pPREY	prey availability	Tuned for realistic realised diets, growth rates and mortality rates
mum (biomass pool species)	Growth of biomass pool species groups	Tuned to provide sufficient food for predators
mQ (biomass pool species)	Quadratic (density dependent) additional mortality	Tuned to limit biomass expansion in cells with less predation pressure
mL (age-structured species groups)	Linear additional mortality	Adjusted to obtain realistic decay curves in the absence of fishing
mum, ht, C (age-structured species, predacse 8)	Maximum growth, handling time and clearance rate used in Holling Type II feeding response	Explored and adjusted, but only with respect to feeding rates and growth rates in the base case
KLP, KUP (age structured species groups)	Upper and lower gape sizes	Explored and adjusted with respect to size-at-age of predator and their prey, and obtaining realistic realised diets
KI (primary producers)	Light saturation	Explored and adjusted with respect to light levels and food required
FSM, FSMG (migratory species)	mortality and growth applied on re-entering the model	Adjusted with respect to growth and mortality curves of migrating species groups

Appendix L: Experts consulted during model design and development

Peter Horn (NIWA, now retired) was the main expert on the system and data inputs
 Beth Fulton (CSIRO) provided feedback and assistance on the Atlantis modelling
 Matt Dunn (NIWA) also provided information and feedback regarding species and dynamics of the system

Mark Hadfield (NIWA) developed the ROMS model

Bec Gorton (CSIRO) converted the ROMS model outputs into Atlantis model inputs

Suzanne Rosier (NIWA) and Dave Frame (VUW) helped me understand light attenuation in water and the atmosphere

Graham Rickard and Cliff Law (NIWA) helped with accessing nutrient data and converting units

Rosemary Hurst (NIWA) provided feedback on the model particularly with respect to the fisheries

Moira Decima and Andres Gutierrez Rodriguez (NIWA) provided guidance on zooplankton growth rates

Ian Tuck (NIWA) provided feedback on the scampi population dynamics, based on the stock assessment which he has been the lead modeller on

Andy McKenzie (NIWA) provided feedback and migratory information on hoki, based on the stock assessment which he has been the lead modeller on.