INSIGHTS ON THE DECISION TABLE

A Case Study of the Bay of Fundy Sea Scallop Fishery

by

Sarah Meng Li

Supervised by Dr. Joanna Mills Flemming

Submitted in partial fulfillment of the requirements for the degree
of Bachelor of Science: Combined Honours in Statistics and
Computer Science

at

Dalhousie University
Halifax, Nova Scotia
April 2019

Abstract

This thesis describes all aspects of the preparation of a decision table for the Bay of Fundy sea scallop fishery. We first describe inshore scallop fisheries in Canada, and the usual monitoring and fishery management processes for them. Next we examine the precautionary approach used by Canadian fisheries to ensure their sustainability. A state-space assessment model (SSAM) is then introduced in a frequentist setting as a tool for obtaining estimates (and predictions) of scallop biomass. In an effort to provide scientific advice for fishery management, we construct a decision table for Scallop Production Area (SPA) 4 using this model. The current SSAMs account for population dynamics along with measurement and process error, but neglect spatial sources of variability. We are working to enhance the SSAM framework to incorporate data from scientific surveys with fine-scale spatial information. We close by discussing several open questions regarding how to modify decision table to capture outputs from this upgraded model.

Acknowledgements

I would like to express my profound gratitude to the people who, in one way or another, have contributed to the writing of this honour thesis.

I would like to acknowledge the guidance of my thesis supervisor, Dr. Joanna Mills Flemming, without whose support, dedication, and patient advice this thesis would not have been possible. She has opened my eyes to applied statistical research, and introduced me to the exciting field of fisheries sciences.

Special thanks should also be paid to David Keith and Jessica Sameoto, who have generously shared their knowledge and years of experience from working in scallop fisheries. The staff of the Department of Mathematics and Statistics have played a crucial part in making my undergraduate experience rewarding and enjoyable. I am very grateful to their hard work. Senior students in Joanna's research group have always provided inspiration, support, and laughter. I especially want to thank Raphael McDonald, Kim Whoriskey, Ethan Lawler, and Yihao Yin.

I want to thank Sofie Yuan and Lauchlan Toal for accepting to review my thesis. They provided insightful feedback and suggestions that allow me to improve the readability of the content.

Finally, I want to thank my friends and my parents, for their unwavering support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

In this chapter, we provide a brief introduction to inshore scallop fisheries. References for this material include [8] and [12]. For a concise introduction about scallops, we also refer the reader to [2] which we loosely follow here.

1.1.1 Fishery science

Fishery science is about understanding the population dynamics, tracking the harvesting process, and developing effective fishery regulations to best preserve healthy and sustainable stock status. Fish stocks are assessed using biological and fisheries data. To describe the stock, biological data include age structure, age at first spawning, fecundity, ratio of males and females, natural mortality (M), fishing growth rate, spawning behavior, critical habitats, migratory habits, and food preferences. Fisheries data often include the type of fishery (e.g., commercial versus recreational), gear type (longline, road and reel, nets, etc.), pounds of fish caught, fishing effort, and the time and geographic location of the best catches. Stock assessment models combine both biological and fisheries data to estimate total biomass of the stock. These models make it possible to assess the current status and condition of the stock as well as to predict how stocks will respond to varying levels of fishing pressure in the future [1].

Overfishing has been a global concern as it reduces fish stocks, jeopardizes employment opportunities, and causes serious social, economic, and environmental problems. Recently, state-space assessment models (SSAMs) are well-accepted as standard deterministic and parametric stock assessment models [1]. It follows that if we can estimate the biomass of the stock and understand its biology, then we can estimate how many fish can be safely removed from the stock in order to ensure a sustainable resource.

1.1.2 Incentives for sea scallop fishery

Scallop harvesting is an important economic component of the commercial fishery in Maritimes Region as it contributes a substantial portion to the total value of fish landings, which in turn provides significant socio-economic benefits to individuals and communities throughout rural Nova Scotia and New Brunswick [8].

However, weak management measures failed to regulate fishing efforts at a sustainable level. For example, catches at Bay of Fundy increased drastically in 1989, but little concern was put into protecting small scallops so that they could grow to larger sizes or spawn in the subsequent years. Correspondingly, scallop stocks were quickly reduced in the fall of 1989. As a result, shrinking stocks and higher fishing costs led to a severe economic crisis as available stocks were no longer able to supply an active fleet with scallop harvest [8].

In response to the depletion of stock status each year, limited entry and strict criteria for issuing fishing licences were introduced to control harvesting capacity. In 1996, it was recognized by the federal court that there were offshore licenses and four classes of inshore scallop licenses (detailed definitions can be found in the Glossary): Full-Bay Fleet, Mid-Bay Fleet, Upper-Bay Fleet, and Inshore East of Baccaro Fleet [12]. This regulation limits access to fishing areas and enhances fishery management.

In an effort to record commercial biomass and monitor stock status, the Bay of Fundy fleets were required to participate in a Dockside Monitoring Program (DMP). During a fishing trip, a monitoring document (logbook) was kept. Upon ending the trip, data were entered into the database of Maritime Region Fisheries Information System (MARFIS) kept by Department of Fisheries and Oceans Canada (DFO). In 1999, the industry implemented an inshore port sampling protocol. Sampling equipment was specified and measuring processes were defined. All data is to be entered and verified in a spreadsheet format in a timely manner. This strategy provided information on the catch profile of scallops landed and was used to monitor the percentage of small juvenile scallops within the catch. On top of catch rate and biomass, industries were also interested in determining the impact that clappers had on scallop fishery. Analogously, guidelines for a Clapper Event were created for samples to be collected and submitted to DFO Science.

1.1.3 Fishing locations

For inshore fleets, Bay of Fundy is the most active scallop harvesting area in the Maritimes region. All scallop fishing in the Bay of Fundy takes place in Scallop Fishing Areas (SFA) 28A, 28B, 28C, and 28D as shown by Figure 1.1. These areas can be further subdivided into areas that are closed by either regulation or variation orders that restrict fishing during all or part of the year. However, some scallop biological production areas may fall within the boundaries of SFA. Hence, Scallop Production Area (SPA) was created as a broader definition for scallop fishing locations [8].

Correspondingly, scallop stocks in the Bay of Fundy are managed based on rules established specifically for each SPA. It allows a more detailed evaluation of the traditional and non-traditional (i.e., marginal) fishing areas through the Bay of Fundy [8].

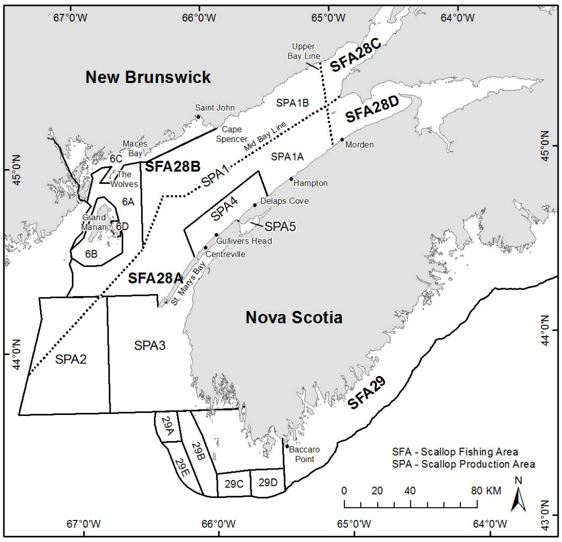

Figure 1.1: Bay of Fundy Scallop Fishing Areas (SFAs) and Scallop Production Areas (SPAs)

1.1.4 Management measures

Different management measures and controls were established to ensure sustainability of the inshore scallop fishery. For example, “limited entry” means that for any of the inshore scallop fleets, no new additional licenses are available; “seasonal/area closure” refers to areas in Bay of Fundy that have regulated seasons as a result of an agreement with the lobster fishery sector. Any changes in seasonal fishing time are discussed through the advisory process to better suit biological and economic objectives [8].

Several measures were adopted to regulate fishing efforts in favour of juvenile scallops. “Minimum shell height size” means any scallop less than or equal to the specified size may not be retained for shucking. “Meat count” means only larger scallops can be retained for shucking, and the count is expressed as the number of scallops whose meat weight is a minimum of 500 g. Note that the quantitative standards vary across regions, and can be modified by DFO through Variation Order [8].

Moreover, regulations for quota allocations across license holders are set, with Dockside Monitoring Program (DMP) and Vessel Monitoring System (VMS) launched to provide up-to-date and accurate data for DFO database [8].

1.2 Science in stock assessment

1.2.1 Biological analysis

The sea scallop, Placopecten magellanicus, is a member of the class Bivalvia. It is laterally compressed and its body is completely enclosed by two shells which are connected by a dorsal hinge [2]. Scallops can be found in the Northwest Atlantic from Cape Hatteras to Labrador and they live in relatively shallow water (i.e., variable depth range of about 10–120 m). They aggregate in patches and harvestable concentrations known as beds. The natural extent of these beds is determined by favourable local conditions, such as water temperature, food availability, substrate type, as well as spawning and settlement success [8].

Many scallops are hermaphrodites (having female and male organs simultaneously), altering their sex throughout their lives; while others exist as dioecious species (males and females are separate); and a few are protoandrous hermaphrodites (males when young then switching to female) [4]. At the age of two, they usually become sexually active, but do not contribute significantly to egg production until the age of four. Every year, spawning begins in late August to early September when eggs and sperm are released into the water, and fertilized ova sink to the bottom. After several weeks, the immature scallops hatch and the larvae drift in water for almost a month before settling to bottom to grow. Newly-settled sea scallops attach themselves to gravel, shells, and other objects via byssal threads to avoid being swept away by the current. Byssus is eventually lost with adulthood, transitioning almost all scallop species into free swimmers [8, 11].

Most species of scallops can move to avoid predators, by propelling themselves through the use of the adductor muscles that open and close their shells. Swimming occurs by first taking in water through the open valves, then closing the valves and propelling water from the corner of its hinge with strong force, which in turn pushes the scallop forward [3]. Some species of scallops can be found on the ocean floor attached to objects by byssal threads. Others be found on the ocean floor, moving with an extendable foot located between their valves or burrowing themselves in the sand by extending and retracting their feet [4].

Scallop growth is characterized by measuring shell height (distance between hinge and the opposite ventral margin). Shell heights can range up to 20 cm but are rarely larger than 15 cm in fished areas. Age is determined from annual rings on the shell that result from the slowing or cessation of growth which usually occurs in the late winter. Rings can also form as a result of some trauma such as contact with fishing gears. These kinds of rings are referred to as shock marks and can be confused with the annual growth rings. Scallop growth can vary between locations, due to inter-annual differences in food availability [8]. Rapid growth occurs within the first several years, with an increase of 50 to 80% in shell height and quadrupled size in meat weight. Usually, scallops reach commercial size at about four to five years of age [11].

Natural mortality of sea scallop is high during its planktonic larval stage. At this stage, unfavourable environmental conditions can negatively affect population development. For example, currents can sweep larvae away from suitable habitats. Predation is another major threat to larvae as they are an important food source for many larger organisms. Once in their adult form, scallops are subject to predation by seastars, predatory snails, crustaceans, and some fish species [8].

1.2.2 Stock assessment

In consultation with industry and other stakeholders, DFO investigates methods to provide long-term advice on fishery management. During the stock assessment process, annual surveys are conducted throughout Bay of Fundy by commercial fishing vessels. For instance, an annual industry assessment survey of Scallop Fishing Area (SFA) 29 West of 65 °30’W was carried out under the department’s use of fish program, various at-sea sampling and survey work were conducted by or on behalf of fishers with DFO personnel, and stock status were recorded by at-sea observers. Where information was available, the amount and spatial distribution of bycatch (i.e., catch of other non-targeted species such as lobsters) were also collected [8].

Starting in 2013, stock status for SPA 1A, 1B, 3 and 4 (including 5) is being evaluated in terms of predicted biomass in the context of Precautionary Approach (PA). However, the current survey coverage is restricted by limited human resources, technical problems, and logistical concerns [8].

In the Bay of Fundy, there have been two major recruitment events since 1981. Population trends of scallops in the inshore areas of Bay of Fundy and southwest Nova Scotia are characterized by “boom and bust” cycles, where rapid increases in population sizes occurred but were then fished down. Until the next major recruitment takes place, fisheries scientists will focus on the existing commercial size scallops produced by low levels of recruitment in most areas. It is expected that sustainability will be achieved by keeping exploitation rate below the net productivity rate of the stocks, and biomass above the Upper Stock Reference Point (USR) [8]. The relationship between stock status and reference points such as USR will be further discussed in chapter 2.

1.3 Management issues

When evaluating inshore scallop on its stock status, all factors must be taken into account: impacts from all industry sectors (e.g., oil and gas, transport, etc), conflicts among different ocean users, and interests of different fishing sectors. Here we discuss several concerns in scallop surveys and resource management.

1. Species interactions: Annual surveys are conducted using scallop drags, which leave a footprint on the ocean floor and may affect other species living there. In addition, the use of drags in the scallop fishery will potentially impact fisheries of a range of other species. For example,

scallop fisheries and lobster fisheries have coexisted in some fishing areas, and scallop drags interact with fixed lobster traps. As a result, fishing participants in each sector have discussed through advisory process to handle conflicts and develop applicable fishing plans to protect the interests of both sides [8].

2. Juveniles protections: It is generally agreed upon by inshore scallop fleets that small scallops should be protected from overfishing for the following reasons. Firstly, this would provide a large number of animals for spawning. Secondly, when recruited to the fishery, scallops of larger size will maximize the yield as larger meats are worth a premium amount, resulting in larger landed values. When there was an above average number of juveniles confirmed by the sample data, industries initiated closures in the Bay of Fundy to protect specific areas [8].

3. Catch monitoring: The inshore scallop fleets are required to participate in DMP through a third party. However, there have been reports about violations of the hail time requirement, inconsistent dockside weight verification, and inaccurate landing information [8].

1.4 Objectives, strategies and tactics

The underlying principle of the fisheries management in Maritimes Region is that fishery is a common property resource to be managed for the benefit of all Canadians. Fishery management should be consistent with sustainable development of fisheries, Aboriginal rights to fish for various purposes, as well as socio-economic benefits to Canadian economy and fishing communities [8].

Overall, fisheries management is summarized by five comprehensive objectives as listed below [8].

Conservation objectives: maintain equilibrium of the ecosystem.

1. Productivity: do not cause unacceptable reduction in productivity so that components can play their roles in the functioning of the ecosystem.
2. Biodiversity: do not cause unacceptable reduction in biodiversity in order to preserve the structure and natural resilience of the ecosystem.
3. Habitat: do not cause unacceptable modification to habitat in order to safeguard both physical and chemical properties of the ecosystem.

Social, cultural, and economical objectives: ensure sustainable development.

1. Culture and Sustenance: respect Aboriginal and treaty rights to fish.
2. Prosperity: create the circumstances for economically prosperous fisheries.

Ensuring protection and promoting recovery of at-risk species is a national priority. The protection and recovery of species at risk falls within the general fishery decision-making framework. In resource management, Precautionary Approach (PA) is about being cautious when scientific information is uncertain, unreliable, or inadequate; and not using the absence of adequate scientific information as a reason to postpone or fail to take action to avoid serious harm to the resource [7]. In the next chapter, we examine the decision framework used by Canadian fisheries to ensure sustainability of key harvested stocks.

Chapter 2

Precautionary Approach in Canadian Fisheries

2.1 Fishery decision-making framework

In this chapter, we introduce basic concepts pertaining to a fishery decision-making framework. In particular, we describe reference points, stock status zones, and the management of uncertainty and risk. References for this material include [7] and [17].

Following is a list of primary components of the framework [7]:

1. Reference points and stock status zones (Healthy, Cautious, and Critical).
2. Harvest strategy and harvest decision rules.
3. Accounting for uncertainty and risk when developing reference points, as well as defining and implementing decision rules.

2.2 Reference point and stock status zones

The stock status zones are defined by a Limit Reference Point (LRP) and an Upper Stock Reference Point (USR). LRP marks the boundary between cautious and critical zones. When a fish stock falls below this boundary, there is a high probability that its productivity will be impaired and serious harm will occur. USR marks the boundary between healthy and cautious zones. When a fish stock falls below this boundary, removal rate must be progressively reduced in order to avoid serious harm to the stock. Note that the unit describing stock status will vary depending on the nature of the resource. LRP is based on biological criteria and established by Science through a peer-reviewed process. USR is developed by fishery managers informed by consultations with the fishery and other interests, with advice and input from Science [25].

In general, reference points will usually be determined using standard biomass and harvest metrics. However, for a number of stocks, such measurements are not available. In these cases, precautionary management actions should be based on empirical measurements of reproductive potentials, with the objective of avoiding serious harm to reproductive capacity of the stock [7].

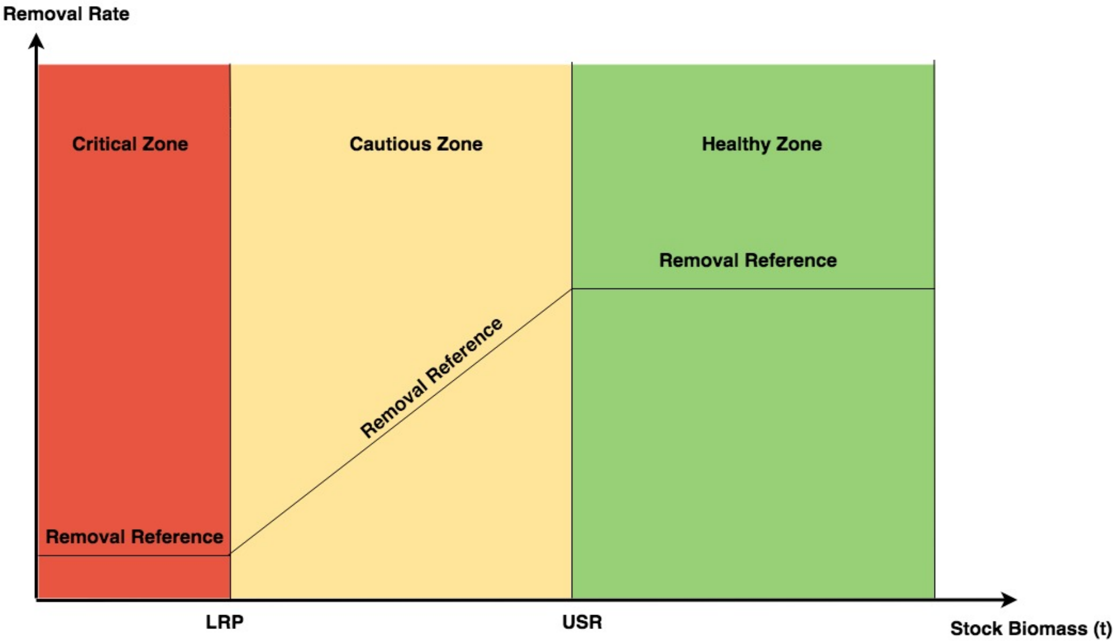

Figure 2.1: Stock Status

2.3 Uncertainty and risks

The treatment of uncertainty and risks is an important aspect of the decision framework when estimating stock status, setting reference points, as well as making and implementing management decisions. Both scientific uncertainty and uncertainty related to the implementation of a management approach must be explicitly considered. It is desirable to quantify scientific uncertainty as much as possible, as they can be used to assess the probability of achieving a target or the probability of a stock falling to a certain level under a specific management approach [7].

For example, uncertainty should be incorporated in the calculation of stock status and biological reference points [7]. In the population dynamics models to be discussed in chapter 4, unobserved stock states such as population biomass and natural mortality rates are predicted using state-space assessment models under a frequentist framework. This model handled identifiability issues caused by prior distributions which were assumed in the Baysian models [29]. In the decision table to be introduced in chapter 5, we perform simulation study and use the median of the outputs to eliminate effects of estimated parameters and outliers.

When using the framework to instruct commercial fishery, it is important to consider the impact from management actions on stock productivity. Based on the condition of stock status, reference points and management actions can be adjusted accordingly.

2.4 Applications of the decision framework

It is crucial to analyze stock status comprehensively before setting harvest decision rules. In the healthy zone, where economic considerations may prevail, stock reductions resulting from management actions are tolerated because there is a low probability of the stock falling to the critical zone [7].

In the critical zone, conservation concerns are paramount and there is no tolerance for preventable declines. In other words, management actions should promote stock growth; removals by all human sources must be kept to the lowest possible level to get the stock out of critical zone within a reasonable time frame [7].

When the stock is in the Cautious or Healthy zone, management actions should be flexible, based on the stock status and the trajectory or rate of change in status, within the bounds of the appropriate removal rate. For example, if a stock is in the Cautious zone but has shown an obvious increase in stock abundance, then management actions should balance socio-economic and biological factors. It should promote the stock growth to the Health Zone within a reasonable time frame, with a certain level of risk tolerance. The focus of management actions will be shifted when a stock is in the Cautious zone and is declining sharply [7].

Utilization of the decision-making framework should involve all fishing participants, including fishermen, Provinces, Territories, Aboriginal people, scientists, and fishery managers etc. Under this framework, development of decision rules for scallop fisheries is a management responsibility and scientists are to provide advice and support the development. Harvest Rate Strategy with the associated harvest decision rules will take the form of a decision table. Ultimately, decision table could be constructed for each stock. If successfully implemented, this approach will facilitate stable and predictable commercial fishery and contribute to fishery sustainability [7].

Chapter 3

Statistical Inference

3.1 Overview

In this chapter, we introduce the frequentist inference approach. In particular, we introduce common methods such as maximum likelihood estimation, hypothesis testing, and confidence intervals. References for this material include [28, 10], and the lecture notes of STAT 3460–Intermediate Statistical Theory–taught by Dr. Edward Susko at Dalhousie University in the winter of 2019 [26].

Statistical inference is the process of using data analysis to deduce properties of an underlying probability distribution of a population. Given a hypothesis about a population, statistical inferences use data drawn from the population with some form of sampling, select a statistical model of the process, deduce propositions from the model, and make predictions based on the sampling data [28].

Some common forms of statistical proposition are listed below.

1. A point estimate is a particular value that best approximates some parameter of interest;

2. An interval estimate (e.g. a confidence interval or set estimate) is an interval constructed using a dataset drawn from a population so that under repeated sampling of the dataset, the interval would contain the

true parameter value with the probability at the stated confidence level;

3. Rejection of a hypothesis;

4. Clustering or classification of data points into groups.

Given some data, $\mathbf{y}$, to find a statistical model with parameter $\boldsymbol{\theta}$, we usually have five questions in mind.

1. What values of $\boldsymbol{\theta}$ are most consistent with $\mathbf{y}$?

2. Which of several alternative models are most consistent with $\mathbf{y}$?

3. What range of values of $\boldsymbol{\theta}$ are consistent with $\mathbf{y}$?

4. Is the model consistent with the data given any possible value of $\boldsymbol{\theta}$?

5. Does the data gathering process enable us to answer the previous questions as accurately and precisely as possible? If not, how to improve the integrity and accuracy of the data?

The first question can be answered by point estimation, and the second question can be answered by hypothesis testing. The third and fourth questions can be answered by interval estimation and model checking respectively. Uncertainty is inherent in the process of trying to learn about $\boldsymbol{\theta}$ from $\mathbf{y}$, and the fifth question is answered by experimental and survey design methods [28].

Frequentists and Bayesian approaches are two main classes of methods to answer questions 1 - 4, and they start from different basic assumptions. From the frequentist perspective, parameters are treated as having values that are fixed states of nature, and we want to learn about these parameters using data. There is randomness in our estimation of parameters, but not in the parameters themselves. From the Bayesian's point of view, parameters are treated as random variables, about which we want to update our beliefs in light of data. In other words, our beliefs are summarized by probability distributions of the parameters [10].

3.2 Frequentist approach

In this section, we focus on the frequentist inference approach, which is the underlying framework of the state-space assessment model introduced in the next chapter.

In the frequentist framework, we use probability to investigate what would happen under repeated replication of data sampling and the subsequent statistical analysis. In this approach, probability is all about how frequently events would occur. Following we introduce joint likelihood function, point estimation, hypothesis testing, and confidence intervals.

3.2.1 Joint likelihood function

Joint likelihood function is a main starting point for point estimation and statistical inferences, as it gives a comprehensive description of data with regards to relative parameters. We assumes θ is of one dimension to simplify the discussion below.

Definition 1. In discrete cases, if the observed sample values are $x_1, x_2, ..., x_n$, the probability of getting them $P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = f(x_1, x_2, ..., x_n; \theta)$. This is the value of the joint probability distribution of the random variables $X_1, X_2, ..., X_n$ at $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$. Since the sample values have been observed and are therefore fixed numbers, we consider $f(x_1, x_2, ..., x_n; \theta)$ as a value of a function of θ, and we refer to this function as the likelihood function [10].

An analogous definition applies when the random sample comes from a continuous population, and in this case $f(x_1, x_2, ..., x_n; \theta)$ is the value of the joint probability density of the random variables $X_1, X_2, ..., X_n$ at $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ [10].

3.2.2 Point estimation

For a population, we can conduct multiple sampling processes and retrieve sets of independently randomly distributed data. Given each set of data, we can propose a model with unknown parameters. Then with consideration of what do these parameters mean, it is often possible to come up with a way of getting reasonable parameter value guesses (i.e. estimations of unknown parameters) from the data. However, such model-by-model reasoning is time-consuming and unsatisfactory: how do we know if our estimation process is making good use of data? Therefore, we come up with a more general approach to deal with all models [28].

Among various estimation approaches such as the method of moments and least squares methods, maximum likelihood estimation stands out in terms of its practical utility and some nice probabilistic properties.

Maximum likelihood estimation

To start with, we give the intuition behind this approach. Parameter values that make the observed data appear relatively probable are more likely to be correct than parameter values that make the observed data appear relatively improbable. So the idea is to judge the likelihood of parameter values using $f_{\boldsymbol{\theta}}(\boldsymbol{y})$, the model's probability density function (p.d.f) according to the given value of $\boldsymbol{\theta}$, evaluated at the observed data. Because $\boldsymbol{y}$ is now fixed and we are considering the likelihood as a function of $\boldsymbol{\theta}$, it is usual to write the likelihood as $L(\boldsymbol{\theta}) \equiv f_{\boldsymbol{\theta}(\boldsymbol{y})}$. In fact, for theoretical and practical purposes, it is usual to work with log likelihood $l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$ [26]. The maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ is defined as

$$\hat{\boldsymbol{\theta}} = argmax_{\boldsymbol{\theta}} l(\boldsymbol{\theta})$$

Below is the formal definition of MLE. For easy discussions, we use a simplified case when θ is of one dimension, so the first order derivative on θ

is used directly instead of the partial derivative of $\boldsymbol{\theta}_j$, where j denotes the j_{th} element in $\boldsymbol{\theta}$.

Definition 2. If $x_1, x_2, ..., x_n$ are the values of a random sample from a population with the parameter θ, the likelihood function of the sample is given by $L(\theta) = f(x_1, x_2, ..., x_n; \theta)$ for values of θ within a given domain. Here, $f(x_1, x_2, ..., x_n; \theta)$ is the value of the joint probability distribution or the joint probability density of the random variables $X_1, X_2, ..., X_n$ at $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$. We refer to the value of θ that maximizes $L(\theta)$, $\hat{\theta}$, as the MLE of θ. Usually, $\hat{\theta}$ can be obtained by solving equation $\frac{dl(\theta)}{d\theta} = 0$, where $l(\theta)$ is a log-likelihood function [10].

Sometimes we need to examine the log-likelihood function itself to determine an MLE. Sometimes, MLE does not exist.

There is more to maximum likelihood estimation than just its intuitive appeal. To see this, we need to consider what constitutes a good estimation under replication of the data-gathering process. Repeating the estimation process results in a different value of $\hat{\theta}$ for each replication. We want these estimators to be consistent. This property is quantified as shown below [10].

Definition 3. If $E(\hat{\theta}) = \theta$, then $\hat{\theta}$ is unbiased. When $n \to \infty$, if $|E(\hat{\theta}) - \theta| \to 0$, then $\hat{\theta}$ is asymptotically unbiased.

Definition 4. If $\hat{\theta}$ is an unbiased estimator of the parameter θ and $var(\hat{\theta}) \to 0$ when $n \to \infty$, then $\hat{\theta}$ is a consistent estimator of θ.

Remark 1. Unbiasedness suggests that the estimator is accurate on average: a long-run average of $\hat{\theta}$ would tend towards the true value of the parameter. Low variance implies that any individual estimate is quite precise.

Theorem 1. *With regularity conditions, and in the large sample limit, MLE is the minimum variance unbiased estimator.*

Theorem 2. *With regularity conditions, MLE is consistent and asymptotically normal.*

This gives MLE following advantages in statistical analysis.

1. MLE provides a consistent approach to parameter estimation problems, and thus can be developed for a large variety of estimation situations. For example, they can be applied to censored data in survival analysis [26].
2. The consistency and asymptotic normality allows MLE to have desirable mathematical and optimality properties. For example, MLE becomes minimum variance unbiased as the sample size increases. They have approximate normal distributions and approximate sample variances that can be used to generate confidence intervals/regions and hypothesis tests for the parameters [6].

Given those nice statistical properties, there are several drawbacks of MLE. For the completeness of the discussion, we give a brief summary below.

1. The likelihood equations need to be specifically worked out for a given distribution and estimation problem. The mathematics is often non-trivial, particularly if confidence intervals for the parameters are desired.
2. The numerical estimation is usually non-trivial (i.e. solutions to systems of equations and solving complicated score functions may not be obvious). Except for a few cases where the maximum likelihood formulas are simple, we tend to rely on high quality statistical software to obtain MLE. Fortunately, high quality maximum likelihood software is becoming increasingly common.
3. The advantages of asymptotic normality may be diminished when the sample size is not large enough. In other words, MLE can be heavily biased for a small sample [6].

3.2.3 Hypothesis testing

Now consider the question of whether some defined restriction on θ is consistent with y. Suppose that we have a model defining a p.d.f., $f_\theta(y)$, for data vector y and that we want to test the null hypothesis $\boldsymbol{H}_0 : \theta = \theta_0$, where θ_0 is some specified value. We want to determine whether data could be reasonably generated from $f_{\theta_0}(y)$. In other words, assuming null hypothesis is true, how likely it is to obtain the effect y observed in your sample data [5, 28]?

P-value evaluates how well the sample data support the argument that the null hypothesis is true. It measures how compatible our data are with the null hypothesis. Accordingly, high p-value means our data are likely under the null hypothesis, so we fail to reject $\boldsymbol{H}_0$. Low p-value means our data are not likely given the null hypothesis, so there is a strong evidence against $\boldsymbol{H}_0$. Hence, the smaller the p-value, the stronger the evidence against the null hypothesis [26].

The process of hypothesis testing consists of four steps.

1. Formulate the null hypothesis $\boldsymbol{H}_0$ and the alternative hypothesis $\boldsymbol{H}_a$.
2. Identify a test statistic that can be used to assess the truth of the null hypothesis.
3. Compute the p-value.
4. Compare the p-value with an acceptable significance value α. If $p \leq \alpha$, it means the observed effect is statistically significant, so we reject the null hypothesis; otherwise, we fail to reject the null hypothesis.

3.2.4 Confidence intervals

Confidence interval describes the amount of uncertainty associated with a sample estimate of a population parameter. Recall the question of finding the range of values for the parameters that are consistent with the data. We can provide a range of values for any parameter θ that would have been accepted in a hypothesis test. For example, we can return a set of values for θ that would yield a p-value of more that 5% if used as a null hypothesis for the parameter. Such a set is known as a 95% confidence set for θ. If the set is continuous, then its upper bound and lower bound define a 95% confidence interval [28].

If we reject a hypothesis when the p-values is less than 5%, then we will reject the null hypothesis on 5% of occasions when it is correct and therefore accept it on 95% when it is correct. This means the corresponding confidence intervals must exclude the true parameter value on those 5% of occasions as well. Therefore, a 95% confidence interval has a 0.95 probability of including the true parameter value (with an infinite sequence of replicates of the data gathering and intervals estimation process) [28, 10].

Chapter 4

State Space Assessment Model

This chapter is based on identifiable state-space assessment models (SSAM) proposed by [29]. Here we give a brief introduction about the hierarchical SSAM. Understanding this model helps us generate a decision table within the frequentist framework.

SSAMs are prevalent in fisheries science for providing management advice when faced with noisy survey and commercial fishery data. A frequentist framework was suggested where inferences were made using Laplace approximation with automatic differentiation. The approach is implemented with an R package, Template Model Builder (TMB), to speed up the model-fitting process. In addition, this model handles some identifiability issues associated with SSAMs, and thus yields reliable inference and provides more conservative advice than models in a Bayesian framework [29].

In order to assess stock status for commercial species of fish and invertebrates, we need mathematical population dynamics models to interpret observed fisheries data collected through time [19]. The results of these assessments are used to set fishing quotas in an effort to protect stocks from over-exploitation and to quantify the progress of rebuilding stocks that have been reduced to below sustainable levels. Effective fisheries management relies on not only having appropriate data and model, but also accounting for all possible sources of uncertainty when making statistical inferences. These uncertainties result from the over-simplification of the assumed population dynamics and the observational nature of fisheries data. Note that such oversimplification is inevitable due to our limited understanding of the biological and environmental processes [29].

4.1 Model framework

The SSAM is defined by two stochastic processes $\boldsymbol{X}_t$ and $\boldsymbol{Y}_t$, with the index $t = 1, 2, ...$ denoting distinct time steps. The q-dimensional $\boldsymbol{X}_t$ process is often called the unobserved state sequence as it represents dynamic features of the system under study that are not directly observed. In fisheries science, $\boldsymbol{X}_t$ usually include the fish stock biomass, its abundance, and the fishing mortality rate. The r-dimensional $\boldsymbol{Y}_t$ process is directly observed and acts as a vector of response variables. In fisheries science, $\boldsymbol{Y}_t$ usually include the reported total commercial catch and survey indices corresponding to a standardized fishing effort. On top of $\boldsymbol{X}_t$ and $\boldsymbol{Y}_t$, there are some other variables observed in the survey, and they usually play the role of fixed covariates and thus will be omitted in the model notation [1].

4.2 Notations

The SSAM is a hierarchical model, and thus it can be conveniently represented by a set of equations, each representing a different level of hierarchy [1]. There are two hierarchical levels in the the model [29] we used to generate the decision table. The top level is a system of observation equations linking observed variables to unobserved variables and observation error. The second level consists of process equations describing the dynamics of the unobserved variables, along with process error [29]. Deeper levels of hierarchy can be added, such as prior distributions of model parameters if the SSAM is cast in a

Bayesian Framework [1]. All variables are indexed by a discrete time index t = 1,...,T referring to years.

Generally in fisheries science, the SSAM can be written as shown below:

$$E[\boldsymbol{Y}_t|\boldsymbol{X}_t] = g(\boldsymbol{X}_t|\boldsymbol{\theta})$$

$$E[\boldsymbol{X}_t|\boldsymbol{X}_{t-1}] = h(\boldsymbol{X}_{t-1}|\boldsymbol{\theta})$$

where functions g and h satisfying some regularity conditions with initial states $\boldsymbol{X}_0$ to be estimated, and $\boldsymbol{\theta}$ is a vector of unknown model parameters. $\boldsymbol{X}_t$ are considered as random effects, whereas $\boldsymbol{\theta}$ are considered fixed. Terminology-wise, $\boldsymbol{X}_t$ is said to be predicted, while $\boldsymbol{\theta}$ is said to be estimated [1].

Explanatory variables: $\boldsymbol{z_t}$

Let z_t denote a vector of observed quantities that are considered as fixed explanatory (or control) variables. Then $z_t = (N_t, C_t, g_t, g_t^R, r_t, CV_{\epsilon t}, CV_{\nu t})^\top$.

N_t : survey estimates of the biomass of recruitment size scallops.

C_t : commercial catch reported at the end of the year.

g_t : growth rates of commercial size scallops.

g_t^R : growth rates of recruitment size scallops.

r_t : ratio of commercial size scallops caught in lined and unlined survey drags, where the lined gear is meant to catch scallops of smaller size.

Response variables: $\boldsymbol{Y_t}$

Let Y_t denote a vector of observed quantities that are considered as response variables. Then $Y_t = (I_t, I_t^R, L_t)^\top$.

I_t : survey index of the biomass of commercial size scallops.

I_t^R : survey index of the biomass of recruitment size scallops.

L_t : survey estimates of the number of scallops that die due to natural causes.

Response variables: $\boldsymbol{X_t}$

Let X_t denote a vector of unobserved state variables, representing features of the fish stock that cannot be directly observed. Then $X_t = (B_t, R_t, m_t)^\top$.

B_t : population biomass of commercial size scallops (commercial biomass).

R_t : population biomass of recruitment size scallops (recruitment biomass).

m_t : natural mortality rate (death due to any cause other than fishing).

4.3 Observation and process equations

The SSAM used to generate the decision table in this thesis [29] is inherited from the model of [23]. It is cast within a frequentist approach, without the use of priors or constraints between parameters. Thus we consider $\boldsymbol{\theta} = (q_I, q_R, S, a, \chi, \sigma^2_\epsilon, \sigma^2_\nu, \sigma^2_\kappa, \sigma^2_\zeta, \sigma^2_\tau, \sigma^2_\phi, \sigma^2_\eta)$ as a vector of fixed effects, and $\boldsymbol{X}_{1:T}$ is considered as dynamic random effects to be predicted based on some estimate of $\boldsymbol{\theta}$.

For $t = 1, ..., T$, the observation equations are shown below.

$$I_t = q_I B_t \epsilon_t, \quad \epsilon_t \overset{ind}{\sim} ulN(\sigma^2_\epsilon) \tag{1}$$

$$I^R_t = q_R R_t \nu_t, \quad \nu_t \overset{ind}{\sim} ulN(\sigma^2_\nu) \tag{2}$$

$$L_t = m_t S\left(\frac{S}{2}N_{t-1} + \left(1 - \frac{S}{2}\right)N_t\right)\kappa_t, \quad \kappa_t \overset{ind}{\sim} ulN(\sigma^2_\kappa) \tag{3}$$

$$C_t = \frac{C_{t-1}}{B_{t-1}} B_t \left(\frac{B_{t-1}}{aB_1/2}\right)^\chi \zeta_t, \quad \zeta_t \overset{ind}{\sim} ulN(\sigma^2_\zeta) \tag{4}$$

Remark

1. The survey indices I_t and I^R_t of commercial and recruitment biomass are assumed to be proportional to their respective stock features (denoted by B_t and R_t respectively), both with multiplicative noise representing model and survey uncertainties.

2. $ulN(\sigma^2)$ denotes a log-normal distribution with unit mean and variance σ^2 on the natural logarithm scale.

3. q_t and q_R are two separate catchability coefficients [24]. The error terms ϵ_t, ν_t, and κ_t follow the log-normal distributions, where σ_ν^2 and σ_ϵ^2 are assumed to be constant through time.

4. In equation (4), index t starts at $t = 2$. Motivated by data simulation [29], this equation was adapted from [27]. In the 10,000 independent simulations, some samples featured a depleted stock (i.e., $C_t > B_t$ at some t). Such depletion resulted from the interaction of many random factors. In particular, the random walks for $\log R_t$ and $\log m_t$ occasionally produced too few recruits, which coincided with low levels for B_t. This phenomenon is more significant as the range of time series is extended. We assume the existence of an equilibrium biomass at which fishing effort remains at a constant level on average. Then the catch is assumed to be driven by effort-dynamics, rather than being fixed throughout the time. By simulating catch, the risk of stock depletion is drastically reduced as the fishing mortality is realistically adjusted in accordance with the stock population.

5. There are several parameters to estimate from data: a is the ratio of biomass at bio-economic equilibrium to biomass at maximum sustainable yield (which is assumed equal to half of the initial commercial biomass B_1); χ is the rate at which effort enters or exits fishery; and σ_ς^2 is the observation variance.

For $t = 2, ..., T$, the process equations are shown below.

$$B_t = \exp(-m_t)g_{t-1}(B_{t-1} - C_{t-1}) + \exp(-m_t)g_{t-1}^R R_{t-1}\tau_t, \quad \tau_t \overset{ind}{\sim} ulN(\sigma_\tau^2) \tag{5}$$

$$R_t = R_{t-1}\phi_t, \quad \phi_t \overset{ind}{\sim} ulN(\sigma_\phi^2) \tag{6}$$

$$m_t = m_{t-1}\eta_t, \quad \eta_t \overset{ind}{\sim} ulN(\sigma_\eta^2) \tag{7}$$

Remark

The process equations are used predominantly to construct the decision table

4.4 Model summary

In the frequentist approach, considering $\boldsymbol{\theta}$ as a fixed unknown parameter at the population level, the observation and process equations can be combined to define the joint likelihood of $\boldsymbol{Y}_{1:T}$ and $\boldsymbol{X}_{1:T}$ as shown below.

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{y}_{1:T}, \boldsymbol{x}_{1:T}) = p(\boldsymbol{y}_1|\boldsymbol{x}_1, \boldsymbol{\theta}) \prod_{t=w}^{T} p(\boldsymbol{y}_t|\boldsymbol{x}_t, \boldsymbol{\theta}) p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{\theta}) \tag{8}$$

Since only $\boldsymbol{y}_{1:T}$ is observed, $\boldsymbol{\theta}$ is estimated by maximizing the marginal log-likelihood as follows.

$$log\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{y}_{1:T}) = log \int \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{y}_{1:T}, \boldsymbol{x}_{1:T}) d\boldsymbol{x}_{1:T} \tag{9}$$

To approximate the high-dimensional integral, [29] used Laplace's method as implemented in the R package TMB, whose efficient implementation allows fast-fitting of the model and running 10,000 replications of the simulation study in a short amount of time. This has in turn sped up the generation process for the decision table.

Chapter 5

Decision Table

As introduced in chapter 1, the Bay of Fundy is fished by three separate scallop fishing fleets: Full Bay, Mid Bay, and Upper Bay. In our assessment, scallop removals included landings from all three inshore scallop fleets, as well as Food, Social, and Ceremonial (FSC) catch by scallop drag. Note that landed recreational and FSC catch by dip netting, diving, tongs, and hand are not available and thus are not accounted for in the assessment [17].

5.1 Commercial data and survey

In this assessment, scallops with a shell height of 80 mm and greater are referred to as commercial data, and scallops with a shell height of 65 to 79 mm are referred to as recruits, and are expected to grow to commercial size in the following year. Scallops less than 65 mm are defined as prerecruits [17]. As an important component of the assessment of scallop stock in the Maritimes region, commercial data is in the form of catch, effort, and location [18].

Three survey designs are used in the Bay of Fundy annual survey: simple random, stratified random, and sampling with partial replacement (SPR). For simple random, survey tows are distributed randomly within a survey area. For stratified random, survey tows are distributed proportionally to a number of strata and the tow replacement within each stratum is random. For SPR, a subset of tows in a given year is a repeat of tows in the previous year. Throughout the survey period, every live scallop and clapper (dead scallop with paired shells) are counted and used to estimate abundance and shell height frequencies. Detailed biological sampling is done on approximately every other tow or based on the judgement of the lead DFO scientist aboard. Each scallop from the sample is shucked, the meat is weighed to the tenth of a gram, and the shell height is measured. These data determine the meat weight-shell height relationships that are used to estimate biomass of the stock. Individual shells are also aged, but the age data of the current year was not immediately available [17].

5.2 Decision table architecture

The Decision table is also called the Harvest Scenario Table. It is used to review the status of scallop stocks by evaluating the annual catch level in terms of resulting exploitation (e), expected changes in biomass (%), probability of biomass increase, probability that after removal the stock will be above the Upper Stock Reference Point (USR), and above the Limit Reference Point (LRP) [20, 21]. We continue to assume a USR of 750 t and a LRP of 530 t.

Catch	e	% Change	Pr Increase	Pr > LRP	Pr > USR
80	0.1035	-7.0819	0.3962	0.8234	0.3866
100	0.1292	-9.5871	0.3636	0.7882	0.3558
120	0.1543	-11.7578	0.3278	0.7702	0.3196
140	0.1831	-16.2065	0.2596	0.7092	0.2544
160	0.2084	-18.4482	0.241	0.6842	0.2362
180	0.2347	-21.2374	0.21	0.6292	0.2026
200	0.2584	-22.9879	0.1806	0.6056	0.1764
220	0.2834	-25.3592	0.1506	0.5568	0.1466
240	0.3123	-29.1001	0.1242	0.4948	0.1198
260	0.3388	-31.9255	0.1026	0.4404	0.0992
280	0.3622	-33.8637	0.081	0.4078	0.0778
300	0.3917	-37.495	0.0596	0.3404	0.0574
320	0.4146	-39.3861	0.0432	0.2906	0.0422
340	0.4391	-41.7436	0.0354	0.2534	0.034
360	0.4646	-44.3334	0.0296	0.2054	0.0288
380	0.4917	-47.306	0.0206	0.1566	0.0194
400	0.5179	-50.044	0.015	0.1208	0.0138
420	0.5415	-52.2898	0.0082	0.0944	0.0082
440	0.5695	-55.3727	0.0066	0.068	0.0064
460	0.5961	-58.1894	0.0024	0.0454	0.0024
480	0.6218	-60.8296	0.0024	0.0298	0.0022
500	0.6469	-63.3911	0.003	0.0208	0.003

Table 5.1: Decision table for SPA 4 scallop data

5.3 Decision table interpretation

Harvest scenarios for 2016/2017, as well as the catches that correspond to predicted biomass changes in the following year 2017/2018, are presented in the

Table 5.1. Given different catch rates and based on data obtained from previous years, we estimate the resulting biomass next year. For example, a catch of 80 t corresponds to a predicted exploitation rate of 0.1035, and is predicted to result in an 7.0819% negative change in biomass. The probability of biomass increase is 39.62%. As a result, around 82.34% of time the biomass will stay above the LRP, and around 38.66% of time the biomass will stay above USR.

Based on different values of LRP and USR, given a sequence of catch rates, we will produce different tables. As discussed in chapter 2, LRP and USR are boundary lines define different stock status zone. LRP represents the stock status below which serious harm is occurring to the stock. At this stock status level, there may also be resultant impacts to the ecosystem, associated species and a long-term loss of fishing opportunities. Reasoning similarly, USR is the stock level threshold below which removals must be progressively reduced in order to avoid reaching the LRP. [7]. Here, our decision table can be used as a reference for fishery manager to adjust fishing quota to maintain the sustainability of scallop stock status.

5.4 Decision table generation

In this case study, we used the SSAM developed by [29] to construct a decision table, with data collected from the year of 2016/2017 issued by DFO. Next, we ran a number of simulations to estimate the resulting biomass given certain catch rates. Below is an algebraic description of the decision-table generating process. Note that for each catch rate, 5000 simulations are generated to yield B_{t+1}. R code for the state-space assessment model and the decision table generation can be found in the appendix.

Firstly, define a sequence of catch rates.

$$C_{t+1} := c(80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340,$$
$$360, 380, 400, 420, 440, 460, 480, 500)$$

Then the exploitation rate (e) corresponding to a certain catch rate is calculated by the following equations inherited from the SSAM. Note that B_{t+1} and m_{t+1} are population biomass and instantaneous natural mortality, respectively in year $t+1$. g_t is the growth rate of the portion of the population recruited to fishery in year t, and g_t^R is the growth rate of the portion of the population recruited to fishery in year $t+1$. C_{t+1} is the commercial catch in year $t+1$, which can be set to a sequence of fixed values within the decision table generator. The τ_{t+1} represents random process error associated with the model dynamics. The state-space structure of the model and the Frequentist methods for estimation were reviewed in [29].

$$e = median(\frac{C_{t+1}}{C_{t+1} + B_{t+1}})$$
$$B_{t+1} = (e^{-m_{t+1}} g_t(B_t - C_{t+1}) + e^{-m_{t+1}} g_t^R R_t)\tau_{t+1} \qquad \tau_{t+1} \sim ulN(\sigma_\tau^2)$$
$$m_{t+1} = m_t \eta_{t+1} \qquad \eta_{t+1} \sim ulN(\sigma_\eta^2)$$
$$R_t = R_{t-1}\phi_{t+1} \qquad \phi_{t+1} \sim ulN(\sigma_\phi^2)$$

The expected changes in biomass (%) is calculated by

$$\% \text{ change} = median(\frac{B_{t+1} - B_t}{B_t})$$

The probability of biomass increase can be calculated by

$$\text{Pr Increase} = \frac{\text{The number of simulations where \% change} > 0}{5000}$$

The probability that after removal the stock will be above the LRP = 530:

$$\text{Pr} > \text{LRP} = \frac{\text{The number of simulations where expected biomass} > \text{LRP}}{5000}$$

The probability that after removal the stock will be above the USR = 750:

$$\text{Pr} > \text{USR} = \frac{\text{The number of simulations where expected biomass} > \text{USR}}{5000}$$

Chapter 6

Conclusions and Open Questions

6.1 Summary

Fisheries scientists collect biological and fisheries data, and then build models to perform stock assessments. These models evaluate stock status and predict biomass, then provides fisheries managers with information required to regulate fish stocks [14, 9]. The Complex State Space Assessment Models (SSAMs) required in the field of fish stock assessment stimulate further development of both the underlying statistical theory and supporting software. Correspondingly, the advancement of statistical software empowers fisheries management and increases the accuracy and the quality of stock assessment. For example, new software such as Automatic Differentiation Model Builder (ADMB, [22]) and R package Template Model Builder (TMB, [15]) made it possible to use the Laplace approximation to implement these models efficiently. As a result, SSAMs become fully operational and now they are prevalent assessment tools in global fisheries stock management [1].

In this thesis, we discussed the background underlying Canadian scallop fisheries and the decision-making framework used in the sustainable fishery management. We then introduced the frequentist inference approach, based on which a state-space assessment model (SSAM) was developed [29] to assess the stock status and predict population biomass. Next, we applied the SSAM in a case study to construct a decision table, using data collected from the Bay of Fundy Scallop Production Area 4. Ultimately, this table will support scallop fishery management with predictions about stock status corresponding to different catch rates.

The discussion below follows from the proposal of Collaborative Research Team (CRT) project (2018-2021) titled “Towards Sustainable Fisheries: State Space Assessment Models for Complex Fisheries and Biological Data”. Funded by Canadian Statistical Sciences Institute (CANSSI), this project is led by Dr. Joanna Mills Flemming, with statistical and fisheries collaborators from various institutes in Canada, Europe, USA. This initiative allows statisticians together with marine biologists and fisheries managers to establish best practices for stock assessments in Canada.

6.2 Open questions

In fishery stock assessment, usually it is not possible to measure abundance directly, so we input catches reported by scientific survey vessels into a SSAM. This model accounts for factors (e.g., gear type, time of day, spatial changes in fishing) which are expected to affect catchability. The goal is to retrieve the most precise yearly indices of abundance. Then we use these indices with their estimated uncertainties as observations in the model. At present, time-series methods are incorporated in the SSAMs fitted under both Bayesian [23] and frequentist [29] frameworks. SSAMs account for population dynamics along with measurement and process error; however, they neglect spatial source of variability.

We begin by considering the integrated modelling problem: how can commercial landings which involve no spatial information be combined with scientific surveys that have fine-scale spatial information within the same model? Statisticians involved in the CRT project are working to develop a spatio-temporal model for the scallop fishery, which can reliably predict biomass into 1 to 2 years. Resulting estimates and predictions will be at a much finer spatial resolution than previously available ones. Meanwhile, we are hoping that advancements in computational power and estimation techniques will empower such integrated models, which could analyze as much relevant data as possible. As a result, a decision table should be implemented to fully describe the predictions produced by the spatio-temporal models, and make the numerical values meaningful to fishery managers.

As a result, there are some concerns we need to address in order to carry out a spatio-temporal stock status analysis.

1. Data concerns: Given catch reported from a Scallop Production Area per year, how to allocate them to different locations in the area? As survey data are provided via the Dockside Monitoring Program, is it possible to incorporate spatial details into the reporting system? If so, is there a way to best monitor the data collection process to ensure the validity and integrity of the reported data? In addition, in other fisheries where scallops are considered bycatch, how could we partition the reported bycatch into different locations?

2. Model concerns: At this point, it is still unclear whether it is possible to develop a spatio-temporal model that converges and produces reliable biomass predictions. Once a model is proposed, simulation study is needed to show that results are not unduly affected by outliers and violations of underlying assumptions. However, it is uncertain that the current available statistical software will support the efficient implementation of such a simulation study.

3. Decision table concerns: The current decision table is of two dimensions. Given a catch value, a row displays predicted values for various referenced

characteristics of the stock. Stock managers will be able to use the decision table as a scientific reference to decide the Total Allowable Catch for the next year in order to maintain a healthy stock status. Without an explicit spatio-temporal model in place, we do not know what form the decision table will take. Will it become an integrated multi-page reference book, where each page contains a decision table dedicated to one location or an area, or will there be some scale-up to larger area(s) a priori.

Moreover, the current decision table is generated in the R environment with a number of simulations of data. Apparently, integrating spatial information will complicate the model, but how will that affect the generation of a new decision table? If the dimensions of the decision table turn out to be multidimensional, can we continue to run the simulation in R, or should we switch to different programming environment? Correspondingly, what impact will it have on the efficiency of the table generation?

Bibliography

[1] William H Aeberhard, Joanna Mills Flemming and Anders Nielsen. “Review of state-space models for fisheries science”. In: *Annual Review of Statistics and Its Application* 5 (2018), pp. 215–235.

[2] Andrew R Brand. “Scallop ecology: distributions and behaviour”. In: *Developments in Aquaculture and Fisheries science*. Vol. 35. Elsevier, 2006, pp. 651–744.

[3] JY Cheng, IG Davison and ME DeMont. “Dynamics and energetics of scallop locomotion”. In: *Journal of Experimental Biology* 199.9 (1996), pp. 1931–1946.

[4] Gilman Arthur Drew. *The habits, anatomy, and embryology of the Giant Scallop,(Pecten tenuicostatus, Mighels)*. 6-7. 1906.

[5] Minitab Blog Editor. *How to Correctly Interpret P Values*. Apr. 2014. URL: `https://blog.minitab.com/blog/adventures-in-statistics-2/how-to-correctly-interpret-p-values`.

[6] James J Filliben. *1.3.6.5.2. Maximum Likelihood*. Oct. 2013. URL: `https://www.itl.nist.gov/div898/handbook/eda/section3/eda3652.htm`.

[7] Fisheries and Communications Branch Oceans Canada. *A fishery decision-making framework incorporating the precautionary approach*. Mar. 2009. URL: `http://www.dfo-mpo.gc.ca/reports-rapports/regs/sff-cpd/precaution-eng.htm`.

[8] Fisheries and Communications Branch Oceans Canada. *Maritimes Region - 2015.* Feb. 2017. URL: `http://www.dfo-mpo.gc.ca/fm-gp/peches-fisheries/ifmp-gmp/scallop-petoncle/scallop-petoncle2015-toc-eng.htm`.

[9] Joanna Mills Flemming. *Towards Sustainable Fisheries: State Space Assessment Models for Complex Fisheries and Biological Data.* 2018. URL: `http://www.canssi.ca/canssi-collaborative-research-team-projects/towards-sustainable-fisheries-state-space-assessment-models-for-complex-fisheries-and-biological-data/`.

[10] John E Freund, Irwin Miller and Marylees Miller. *John E. Freund's Mathematical Statistics: With Applications.* Pearson Education India, 2004.

[11] Deborah R Hart and Antonie S Chute. *Essential fish habitat source document. Sea scallop, Placopecten magellanicus, life history and habitat characteristics.* DIANE Publishing, 2004.

[12] Antonio Hervás, Lee Murray and Christina Annand. *MSC SUSTAINABLE FISHERIES CERTIFICATION FBSA Canada Full Bay Sea Scallop Fishery.* Apr. 2013.

[13] St. Andrew's House. *Fish recruitment.* Dec. 2009. URL: `https://www2.gov.scot/Topics/marine/marine-environment/ecosystems/population/recruitment`.

[14] Eyolf Jul-Larsen. *Management, co-management or no management?: Major dilemmas in southern African freshwater fisheries. Case studies.* Vol. 422. Food & Agriculture Org., 2003.

[15] Kasper Kristensen et al. "TMB: automatic differentiation and Laplace approximation". In: *arXiv preprint arXiv:1509.00660* (2015).

[16] Carol Lalli and Timothy R Parsons. *Biological oceanography: an introduction.* Elsevier, 1997.

[17] L Nasmith, J A Sameoto and A Glass. *Scallop Production Areas in the Bay of Fundy: Stock Status for 2015 and Forecast for 2016.* 2016.

[18] L Nasmith et al. *Scallop Production Areas in the Bay of Fundy: Stock Status for 2013 and Forecast for 2014.* 2014.

[19] Terrance J Quinn and Richard B Deriso. *Quantitative fish dynamics.* Oxford University Press, 1999.

[20] *Research Document - 2014/016.* Apr. 2018. URL: `http://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2014/2014_016-eng.html`.

[21] *Research Document - 2015/016.* Apr. 2018. URL: `https://waves-vagues.dfo-mpo.gc.ca/Library/363979.pdf`.

[22] Hans J Skaug and David A Fournier. "Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models". In: *Computational Statistics & Data Analysis* 51.2 (2006), pp. 699–709.

[23] Stephen J Smith and Brad Hubley. "Impact of survey design changes on stock assessment advice: sea scallops". In: *ICES Journal of Marine Science* 71.2 (2013), pp. 320–327.

[24] Stephen J Smith and Mark J Lundy. *Scallop production area 4 in the Bay of Fundy: stock status and forecast.* Canadian Science Advisory Secretariat, 2002.

[25] Stephen J Smith, Jessica A Sameoto and Craig J Brown. "Setting biological reference points for sea scallops (Placopecten magellanicus) allowing for the spatial distribution of productivity and fishing effort". In: *Canadian journal of fisheries and aquatic sciences* 74.5 (2016), pp. 650–667.

[26] Edward Susko. *Maximum Likelihood.* Jan. 2019.

[27] James T Thorson et al. “A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment”. In: *Canadian Journal of Fisheries and Aquatic Sciences* 70.12 (2013), pp. 1829–1844.

[28] Simon N Wood. *Core statistics*. Vol. 6. Cambridge University Press, 2015.

[29] Yihao Yin et al. “Identifiable state-space models: A case study of the Bay of Fundy sea scallop fishery”. In: *Canadian Journal of Statistics* 47.1 (2019), pp. 27–45.

Glossary

bycatch In the fishing industry, bycatch is a fish or other marine species that is caught unintentionally while catching certain target species and target sizes of fish, crabs etc. Bycatch can be either of a different species, the wrong sex, or undersized or juvenile individuals of the target species. 7

byssal threads are strong, silky fibers extending from the muscular foot, used to attach to a firm support, such as a rock [4]. 5

Clapper Event occurs when the number of clappers (dead paired shells) in a regular fishing tow makes up 25 percent or more of the total catch in that tow. Certain information should be collected about that tow and any similar tows in that area[8]. 3

Dockside Monitoring Program (DMP) All inshore scallop license holders must hail out to the Interactive Voice Recognition (IVR) center at the start of the trip and must hail in to a Dockside Monitoring Company (DMC) before they land with their scallops. All inshore scallop license holders are required to submit monitoring documents or logbooks to the DMC for entry into the DFO (Fisheries and Oceans Canada) database. The weight of scallops landed is verified by a Dockside Observer [8]. 3, 4

Full-Bay Fleet Scallop fishing of full-bay fleet is its primary activity. Full Bay scallop license holders are able to fish scallops anywhere in the Bay of Fundy, and the fleet has traditionally been based in Digby and Nova Scotia [12, 17]. 2

hail time requirement The Bay of Fundy fleets are required to hail out before leaving port to an Interactive Voice Recognition Centre (IVR) to start their fishing trip. All inshore fleets are required to hail in to a DMC at least two hours prior to returning to port [8]. 8

Harvest Rate Strategy is the approach taken to manage the harvest of a stock and is a necessary element of any fishery plan [7]. 13

Inshore East of Baccaro Fleet have access to SFA 29 East of 65°30' west longitude and a small scallop fishery in SFA 29 West of 65°30' west Longitude to the 43°40' north latitude line. Scallops are not the primary fishery (mainly lobster) for the licence holders [12]. 2

juvenile is an individual organism that has not yet reached its adult form, sexual maturity or size. 8

Limit Reference Point (LRP) marks the boundary between the cautious and critical zone. When a fish stock falls below this zone, there is a high probability that its productivity will be so impaired that serious harm will occur [7]. 10

maximum likelihood estimator (MLE) begins with the mathematical expression known as a likelihood function (or log-likelihood function) of the sample data. The likelihood of a set of data is the probability of obtaining that particular set of data given the chosen probability model. This expression contains the unknown parameters. Those values of the parameter that maximize the sample likelihood are known as the maximum likelihood estimators (MLE). 17

Mid-Bay Fleet Mid Bay license holders can only fish for scallops on the northern side of the Mid Bay line (Figure 1), and traditionally the fleet has consisted mainly of New Brunswick-based vessels with multiple licenses for different species [12, 17]. 2

minimum variance When the sample size grows to infinity, the estimator has the smallest variance, and thus the narrowest confidence interval, of all estimators of that type. 19

minimum variance unbiased estimator The estimator for the parameter θ of a given distribution that has the smallest variance of all unbiased estimators for θ is called the minimum variance unbiased estimator, or the best unbiased estimator for θ [10]. 18

observation equations are equations of a SSM specifying how observed quantities depend on unobserved states, typically including an (observation) error term [1]. 23

p-value is the probability of obtaining an effect at least as extreme as the one in your sample data, assuming the truth of the null hypothesis [5]. 20

plankton are the diverse collection of organisms that live in large bodies of water and are unable to swim against a current [16].The individual organisms constituting plankton provide a crucial source of food to many large aquatic organisms, such as fish and whales. Technically the term does not include organisms on the surface of the water or those that swim actively in the water. 6

Precautionary Approach (PA) In resource management, the PA is about being cautious when scientific information is uncertain or inadequate, and not using the absence of adequate scientific information as an excuse to postpone or fail to take action to avoid serious harm to the resource [7]. 7, 9

process equations are equations of a SSM specifying how unobserved states at a given time depend on past states, typically including an (process) error term [1]. 23

recruitment Changes in fish populations are primarily driven by variability in recruitment. Recruitment can be defined as the number of fish surviving to enter the fishery or to some life history stage such as settlement or maturity [13]. 7

state-space assessment models are a class of hierarchical statistical models specifying dynamic unobserved random variables (states) and their link to observed random variables [1]. 2, 12, 22

stock assessment models are statistical models used to infer the current state of fish stocks and predict trajectories of their key features [1]. 1

unbiased When the sample size grows to infinity, the average value of the parameter estimates will be approximated to the population value. 19

Upper Stock Reference Point (USR) marks the boundary between the healthy and cautious zone. When a fish stock falls below this point, the removal rate must be progressively reduced in order to avoid serious harm to the stock [7]. 7, 10

Upper-Bay Fleet Upper bay license holders fish east of the Upper Bay line, and are often multi-species vessels based in either Nova Scotia or New Brunswick. [12, 17]. 2

Vessel Monitoring System (VMS) are mandatory for the Full Bay, Mid Bay, and Upper Bay Fleets that fish in the Bay of Fundy. The eligible East of Baccaro license holders with access to some scallop fishing areas must have VMS [8]. 4

Appendix A

R Code

```r
library(TMB)
compile("SS2v4.cpp") # alternative model
dyn.load(dynlib("SS2v4"))

#* Skip this line
{
  source('SS2v4simul.r') # simulate data according to
   ↪ alternative model

  simulate<- FALSE
  ### setup and design as in paper
  NY <- 26 # time series length, as in BoF real data
  yearsvec <- 1:NY
  year <-1992:2017

  gt <- rep(1.3,NY) # 1.2*rlnorm(NY,meanlog=-(0.1^2),sdlog=0.1)
  gRt <- rep(1.8,NY) #
   ↪ 1.8*rlnorm(NY,meanlog=-(0.1^2),sdlog=0.1)
```

```
Nt <-
  ↪  exp(16.8+as.numeric(arima.sim(model=list(ar=c(1.2,-0.5)),n=NY,sd=0.3))

sigmatau <- 0.1 # Bt
sigmaphi <- 0.1 # Rt
sigmam <- 0.1 # mt
sigmaepsilon <- 0.1 # It
sigmaupsilon <- 0.1 # Jt
sigmakappa <- 0.1 # Lt
sigmaC <- 0.1 # Ct
qI <- 0.4
qR <- 0.2
S <- 0.5
a <- 1.6
chi <- 0.1

theta.true <- c(sigmatau,sigmaphi,sigmam,
                sigmaepsilon,sigmaupsilon,sigmakappa,sigmaC,
                qI,qR,S,a,chi)
length.theta <- length(theta.true)

B1 <- 1000   # initial Bt
R1 <- 0.2*B1 # initial Rt
m1 <- 0.2    # initial mt
C1 <- 0.2*B1 # initial Ct



### simulate
if( simulate == TRUE ) {
```

```
    dat <- SS2v4simul(gt=gt,gRt=gRt,Nt=Nt,
                      B1=B1,R1=R1,m1=m1,C1=C1,

                       ↪  sigmatau=sigmatau,sigmaphi=sigmaphi,sigmam=sigmam,

                       ↪  sigmaeps=sigmaepsilon,sigmaups=sigmaupsilon,
                      sigmakap=sigmakappa,sigmaC=sigmaC,

                       ↪  qI=qI,qR=qR,S=S,a=a,chi=chi,seedvalue=NULL)
    str(dat) # all vectors of length NY
    any(dat$Bt<=dat$Ct) # stock depleted?
  } else {}
}

### fit SSM
if( simulate == TRUE ){
  datalist <- list('I'=dat$It,'IR'=dat$Jt,'L'=dat$Lt,'C'=dat$Ct,
                   'g'=dat$gt,'gR'=dat$gRt,'N'=dat$Nt)

} else {
  datalist<-
   ↪  DDspa4.2017.dat[c("I","IR","clappers","C","g","gR","N")]
  # datalist<-
   ↪  DDspa4.2017.dat[c("I","IR","clappers","C","g","gR","N")]
  names(datalist)[[3]]<- "L"
  datalist<- lapply(datalist,tail,NY)
}
# When enable_catcheq is set to 1, the fitted model is the full
 ↪  alternative model;
```

```
# when enable_catcheq is set to 0, equation(9) is ignored and a reduced model is fitted.
datalist<-c(datalist,list(enable_catcheq = 0))
#log_a = 0, log_sigma_C = 0, log_chi = -1
parlist <- list('log_sigma_tau'=0,'log_sigma_phi'=0,'log_sigma_m'=0,
                'log_sigma_epsilon'=0,'log_sigma_upsilon'=0,
                'log_sigma_kappa'=0,'log_sigma_C'=0,
                'log_q_I'=-1,'log_q_R'=-1,
                'log_S'=0,'log_a'=0,'log_chi'=-1,
                # 'log_B'=rep(log(max(dat$It)*10),NY),
                # 'log_R'=rep(log(max(dat$Jt)*10),NY),
                'log_B'=rep(log(max(datalist$I)*10),NY),
                'log_R'=rep(log(max(datalist$IR)*10),NY),
                'log_m'=rep(log(0.5),NY))

#A switch to adjust the model depending on the setting: wheteher enable_catch is equal to 0 or not.
if(datalist$enable_catcheq == 0){
  maplist = list(log_a = as.factor(NA),log_sigma_C = as.factor(NA),
                 log_chi = as.factor(NA))
}else{
  maplist = list()
}

obj <- MakeADFun(data=datalist,parameters=parlist,
                 random=c('log_B','log_R','log_m'),
                 map = maplist,
```

```
                 DLL="SS2v4",silent=T)

system.time(opt <- try(nlminb(start=obj$par,obj=obj$fn,gr=obj$gr,
                              control=list(eval.max=1000,iter.max=1000))
# ^ less than 1 sec
opt$message # converged properly?


### look at fixed param estimates
system.time(rep <- sdreport(obj,bias.correct=F))
# ^ less than 1 sec
summ.rep <- summary(rep)

summ.rep[(length.theta+3*NY+1):(2*length.theta+3*NY),]
### plot predicted randeff with CI as colored envelope
pred.Bt <- summ.rep[dimnames(summ.rep)[[1]]=='B',1]
pred.logBt <- summ.rep[dimnames(summ.rep)[[1]]=='log_B',1]
se.pred.logBt <- summ.rep[dimnames(summ.rep)[[1]]=='log_B',2]
lb.ci.Bt <- exp(pred.logBt-1.96*se.pred.logBt) # 95% CI lower bound
ub.ci.Bt <- exp(pred.logBt+1.96*se.pred.logBt) # 95% CI lower bound

pred.Rt <- summ.rep[dimnames(summ.rep)[[1]]=='R',1]
pred.logRt <- summ.rep[dimnames(summ.rep)[[1]]=='log_R',1]
se.pred.logRt <- summ.rep[dimnames(summ.rep)[[1]]=='log_R',2]
```

```
lb.ci.Rt <- exp(pred.logRt-1.96*se.pred.logRt) # 95% CI lower bound
ub.ci.Rt <- exp(pred.logRt+1.96*se.pred.logRt) # 95% CI lower bound

pred.mt <- summ.rep[dimnames(summ.rep)[[1]]=='m',1]
pred.logmt <- summ.rep[dimnames(summ.rep)[[1]]=='log_m',1]
se.pred.logmt <- summ.rep[dimnames(summ.rep)[[1]]=='log_m',2]
lb.ci.mt <- exp(pred.logmt-1.96*se.pred.logmt) # 95% CI lower bound
ub.ci.mt <- exp(pred.logmt+1.96*se.pred.logmt) # 95% CI lower bound
# ###########Plot for the reduced model#########
pred.Bt.Reduced <- pred.Bt
lb.ci.Bt.Reduced <- lb.ci.Bt
ub.ci.Bt.Reduced <- ub.ci.Bt
pred.Rt.Reduced <- pred.Rt
lb.ci.Rt.Reduced <- lb.ci.Rt
ub.ci.Rt.Reduced <- ub.ci.Rt
pred.mt.Reduced <- pred.mt
lb.ci.mt.Reduced <- lb.ci.mt
ub.ci.mt.Reduced <- ub.ci.mt

colmed <- c('#2b05ff') # color code for pred
colenv <- paste0(colmed,'30') # color for envelope

par(mfrow=c(3,1))
# Bt
library(scales)
```

```

plot(year,pred.Bt.Reduced,type='o',col=colmed[1],pch=0,
     xlab='Years',ylab=expression(italic(B[t])),
     main=expression('Predicted commercial
      ↪  biomass'~italic(B[t])),
     ylim=c(100,6000))
grid(nx=NA,ny=NULL,equilogs=F)
polygon(c(year,year[NY:1]),c(lb.ci.Bt.Reduced,ub.ci.Bt.Reduced[NY:1]),
        col=colenv[1],border=NA,xpd=F)

# lines(year,pred.Bt.Full,type='o', col='red')
#
 ↪  polygon(c(yearsvec,yearsvec[NY:1]),c(lb.ci.Bt.Full,ub.ci.Bt.Full[NY:1])
#         col=alpha('red',0.4),border=NA,xpd=F)

legend("topright", col = c(colmed[1], 'red'), lty = 1, bty="o",
 ↪  legend=c("Reduced", "Full"))



# Rt
plot(year,pred.Rt.Reduced,type='o',col=colmed[1],pch=0,
     xlab='Years',ylab=expression(italic(R[t])),
     main=expression('Predicted recruitment
      ↪  biomass'~italic(R[t])),
     ylim=c(1,4500))
grid(nx=NA,ny=NULL,equilogs=F)
polygon(c(year,year[NY:1]),c(lb.ci.Rt.Reduced,ub.ci.Rt.Reduced[NY:1]),
        col=colenv[1],border=NA,xpd=F)
# lines(yearsvec,pred.Rt.Full,type='o', col='red')
```

```r
# polygon(c(yearsvec,yearsvec[NY:1]),c(lb.ci.Rt.Full,ub.ci.Rt.Full[NY:1])
#          col=alpha('red',0.4),border=NA,xpd=F)

legend("topright", col = c(colmed[1], 'red'), lty = 1, bty="o", legend=c("Reduced", "Full"))

# mt
plot(year,pred.mt.Reduced,type='o',col=colmed[1],pch=0,
     xlab='Years',ylab=expression(italic(m[t])),
     main=expression('Predicted natural mortality rate'~italic(m[t])),
     ylim=c(0,0.3))
#,xlim = c(1992,2017)


grid(nx=NA,ny=NULL,equilogs=F)
polygon(c(year,year[NY:1]),c(lb.ci.mt.Reduced,ub.ci.mt.Reduced[NY:1]),
        col=colenv[1],border=NA,xpd=F)
# lines(yearsvec,pred.mt.Full,type='o', col='red')
# polygon(c(yearsvec,yearsvec[NY:1]),c(lb.ci.mt.Full,ub.ci.mt.Full[NY:1])
#          col=alpha('red',0.3),border=NA,xpd=F)
legend("topright", col = c(colmed[1], 'red'), lty = 1, bty="o", legend=c("Reduced", "Full"))
# ###########Plot for the full model##########
# pred.Bt.Full <- pred.Bt
# lb.ci.Bt.Full <- lb.ci.Bt
# ub.ci.Bt.Full <- ub.ci.Bt
```

```
# pred.Rt.Full <- pred.Rt
# lb.ci.Rt.Full <- lb.ci.Rt
# ub.ci.Rt.Full <- ub.ci.Rt
# pred.mt.Full <- pred.mt
# lb.ci.mt.Full <- lb.ci.mt
# ub.ci.mt.Full <- ub.ci.mt
#
# colmed <- c('#2b05ff') # color code for pred
# colenv <- paste0(colmed,'30') # color for envelope
#
# par(mfrow=c(3,1))
# # Bt
# library(scales)
#
# plot(year,pred.Bt.Full,type='o',col=colmed[1],pch=0,
#      xlab='Years',ylab=expression(italic(B[t])),
#      main=expression('Predicted commercial
 ↪  biomass'~italic(B[t])),
#      ylim=c(100,6000))
# grid(nx=NA,ny=NULL,equilogs=F)
#
 ↪  polygon(c(year,year[NY:1]),c(lb.ci.Bt.Full,ub.ci.Bt.Full[NY:1]),
#          col=colenv[1],border=NA,xpd=F)
#
# # lines(year,pred.Bt.Full,type='o', col='red')
# #
 ↪  polygon(c(yearsvec,yearsvec[NY:1]),c(lb.ci.Bt.Full,ub.ci.Bt.Full[NY:1])
# #          col=alpha('red',0.4),border=NA,xpd=F)
#
```

```
# legend("topright", col = c(colmed[1], 'red'), lty = 1,
 ↪ bty="o", legend=c("Reduced", "Full"))
#
#
# # Rt
# plot(year,pred.Rt.Full,type='o',col=colmed[1],pch=0,
#   xlab='Years',ylab=expression(italic(R[t])),
#   main=expression('Predicted recruitment
 ↪ biomass'~italic(R[t])),
#   ylim=c(1,4500))
# grid(nx=NA,ny=NULL,equilogs=F)
#
 ↪ polygon(c(year,year[NY:1]),c(lb.ci.Rt.Full,ub.ci.Rt.Full[NY:1]),
#    col=colenv[1],border=NA,xpd=F)
# # lines(yearsvec,pred.Rt.Full,type='o', col='red')
# #
 ↪ polygon(c(yearsvec,yearsvec[NY:1]),c(lb.ci.Rt.Full,ub.ci.Rt.Full[NY:1])
# #   col=alpha('red',0.4),border=NA,xpd=F)
#
# legend("topright", col = c(colmed[1], 'red'), lty = 1,
 ↪ bty="o", legend=c("Reduced", "Full"))
#
# # mt
# plot(year,pred.mt.Full,type='o',col=colmed[1],pch=0,
#   xlab='Years',ylab=expression(italic(m[t])),
#   main=expression('Predicted natural mortality
 ↪ rate'~italic(m[t])),
#   ylim=c(0,0.3))
# #,xlim = c(1992,2017)
```

```
#
#
# grid(nx=NA,ny=NULL,equilogs=F)
#
 ↪  polygon(c(year,year[NY:1]),c(lb.ci.mt.Full,ub.ci.mt.Full[NY:1]),
#          col=colenv[1],border=NA,xpd=F)
# # lines(yearsvec,pred.mt.Full,type='o', col='red')
# #
 ↪  polygon(c(yearsvec,yearsvec[NY:1]),c(lb.ci.mt.Full,ub.ci.mt.Full[NY:1])
# #          col=alpha('red',0.3),border=NA,xpd=F)
# legend("topright", col = c(colmed[1], 'red'), lty = 1,
 ↪  bty="o", legend=c("Reduced", "Full"))

# ############END BoF_SSM############

# ###########Decision-table generation###########
# sigma_eta: random effects of natural mortality rate (m_tplu)
sigma_eta<-summ.rep['sigma_m',1]

# sigma_tau: random effects of population biomass of commercial
 ↪  size scallops (b_tplus)
sigma_tau<-summ.rep['sigma_tau',1]

# And sigma for recruits
sigma_phi <- summ.rep['sigma_phi',1]

# results store the computed b_tplus from each one of the 5000
 ↪  simulations
results<-c()
```

```

# g_t: growth rates of commercial size scallops (mean number
 ↪  per tow) from 2017
g_t <- 1.078598

# gR_t:growth rates for recruitment size scallops from 2017
gR_t <-1.612625

# B_t: population biomass of commercial size scallops (metric
 ↪  tons) from 2017
B_t <- 7.454024e+02

# R_t: population biomass of recruitment size scallops (metric
 ↪  tons) from 2016
R_t <- 2.428874e+01


# m_t: natural mortality rate
m_t <- 5.118057e-02

# Generate 5000 times of simulation with random effects
 ↪  sigma_eta and sigma_tau and sigma_phi

#results store the computed b_tplus from each one of the 5000
 ↪  simulations
n_sims <- 5000
catch <- seq(80,500,by=20)
n_catch <- length(catch)
LRP <- 530 # Limit Reference Point (LRP), will vary by area
```

```
USR <- 750 # Upper Reference Point (USR), will vary by area
sims_res <- NULL; sum_stats <- NULL #  A couple of lists to
 ↪  store all the data...
b_tplus <- NA; eta <- NA; tau <- NA ;phi <- NA; m_tplus <- NA;
 ↪  R_est <- NA; B_process <-NA

for (j in 1:n_catch)
{
  for(i in 1:n_sims)
  {
    # eta, tau, and phi follows a log-normal distribution with
     ↪  unit mean and standard deviation sigma_eta, sigma_tau,
     ↪  and sigma_phi
    eta[i] <- rlnorm(1,-(sigma_eta^2)/2,sigma_eta)
    tau[i] <- rlnorm(1,-(sigma_tau^2)/2,sigma_tau)
    phi[i] <- rlnorm(1,-(sigma_phi^2)/2,sigma_phi)
    m_tplus[i] <- m_t*eta[i]
    R_est[i] <- R_t*phi[i]

    #b_tplus <- (exp(-m_tplus)*g_t*(B_t -
     ↪  Ct_plus)+exp(-m_tplus)*gR_t*R_t)*tau
    # Break the formula this into 2 lines
    B_process[i] <- exp(-m_tplus[i])*g_t*(B_t -
     ↪  catch[j])+exp(-m_tplus[i])*gR_t*R_est[i]
    b_tplus[i] <- (B_process[i])*tau[i]
  } # end for(i in n_sims)
  # Get this into a data frame and calculate all the other
   ↪  interesting pieces
```

```
  sims_res[[as.character(catch[j])]] <- data.frame(B_next =
   ↪  b_tplus,B_process = B_process,R_est = R_est,m =
   ↪  m_tplus,eta=eta,tau=tau,phi=phi)
  sims_res[[as.character(catch[j])]]$exp <- catch[j]/(catch[j] +
   ↪  b_tplus) # exploitation rate
  sims_res[[as.character(catch[j])]]$Per_change <- 100*(b_tplus
   ↪  - B_t)/B_t # Percent change in biomass next year
  sum_stats[[as.character(catch[j])]]$Exploit <-
   ↪  median(sims_res[[as.character(catch[j])]]$exp)
  sum_stats[[as.character(catch[j])]]$Med_Pre_change <-
   ↪  median(sims_res[[as.character(catch[j])]]$Per_change)

  sum_stats[[as.character(catch[j])]]$P_increase <-
   ↪  length(which(sims_res[[as.character(catch[j])]]$B_next >
   ↪  B_t))/n_sims
  sum_stats[[as.character(catch[j])]]$P_above_LRP <-
   ↪  length(which(sims_res[[as.character(catch[j])]]$B_next >
   ↪  LRP))/n_sims
  sum_stats[[as.character(catch[j])]]$P_above_USR <-
   ↪  length(which(sims_res[[as.character(catch[j])]]$B_next >
   ↪  USR))/n_sims

  sum_stats$catch <- rownames(sum_stats)
}
# Now unwrap the summary stats lists into a data frame with the
 ↪  results
sum_stats <- do.call(rbind.data.frame,sum_stats)
sum_stats$catch <- rownames(sum_stats)
```