Data limited fisheries: Incorporating expert knowledge into stock assessment

Anna Chrysafi

DENVI Doctoral Programme in Interdisciplinary Environmental Sciences
Organismal and Evolutionary Biology Research Programme
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

ACADEMIC DISSERTATION

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in Auditorium 2041, Biocenter 2, Viikinkaari 5, Helsinki on the 28th of May 2019 at 12 o'clock noon.

Author's present address:

Faculty of Biological and Environmental Sciences
P.O. Box 64
00014
University of Helsinki
Finland

email: anna.chrysafi@helsinki.fi

ISBN 978-951-51-5210-7 (paperback)

ISBN 978-951-51-5211-4 (PDF)

Helsinki 2019

Next print Oy

Thesis supervisor: Associate Professor Anna Kuparinen
Department of Biological and Environmental Science
University of Jyväskylä
Finland

Thesis advisory committee: Research Professor Jaakko Erkinaaro
Natural Resources Institute Finland
Finland

Adjunct Professor Päivi Haapasaari
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

Senior Scientist Samu Mäntyniemi
Natural Resources Institute Finland
Finland

Thesis pre-examiners: Associate Professor Athanasios Tsikliras
School of Biology
Aristotle University of Thessaloniki
Greece

Senior Scientist Outi Heikinheimo
Natural Resources Institute Finland
Finland

Thesis opponent: Professor Brian MacKenzie
National Institute of Aquatic Resources
Technical University of Denmark
Denmark

Custos: Professor Sakari Kuikka
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

The thesis consists of a summary and the following Articles, referred to in the text by their Roman numerals:

I Chrysafi A. and Kuparinen A. (2016). Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. Environmental Reviews 24: 25-38

II Chrysafi A., Cope J.M. and Kuparinen A. (2019). Eliciting expert knowledge to inform stock status for data-limited stock assessment. Marine Policy 101: 167-176

III Perälä T., Chrysafi A. and Vanhatalo J. Calibrating expert assessments using hierarchical Gaussian process models. In revision

IV Chrysafi A. and Cope J. Testing methods of determining relative stock abundance priors when setting catch recommendations using data-limited approaches. Submitted

Author's contribution to the Articles:

I A. Chrysafi was responsible for the original idea, carried out the literature review and had the main responsibility for formulating the structure of the Article. The authors wrote the Article jointly.

II A. Chrysafi was responsible for the original idea and had the main responsibility for designing and implementing the expert elicitation. J. Cope provided fishery data and assisted in the development of the elicitation tool and the selection of the experts. A. Chrysafi built the elicitation tool and processed the data. The results were interpreted jointly by A. Chrysafi and J. Cope. A. Kuparinen provided assistance with the statistical analyses. The authors wrote the Article jointly.

III A. Chrysafi provided the data elicited from experts and the original idea. T. Perälä had the main responsibility for formulating the models, carrying out the inference and testing the models in a simulation setting. T. Perälä, A. Chrysafi and J. Vanhatalo wrote the article jointly.

IV A. Chrysafi was responsible for the original idea and had the main responsibility for carrying out the model performance testing. J. Cope provided assistance with Stock Synthesis. A. Chrysafi processed the data and the results were interpreted jointly by A. Chrysafi and J. Cope. The authors wrote the Article jointly.

Contents

Abstract

Stock assessment is a critical step in fisheries management, since it directly estimates reference points that help determine whether a population's size is acceptable and subsequently, to set harvest levels. Therefore, many international agreements require that all exploited aquatic populations are assessed quantitatively. However, for the majority of the worlds' harvested fish stocks, data is lacking. Such fisheries are often referred to as data-poor or data-limited and are a major challenge for stock assessment scientists and fisheries managers, since the traditional approaches to stock assessment cannot be implemented. The necessity to assess the status of all fisheries, led to the development of models tailored to data-limited situations.

In this thesis, I first introduced the characteristics of data-limited fisheries, and then described the various quantitative indicators and models developed to assess them, some of which are widely used in real assessment schemes. I reviewed the approaches by their input requirements and their biological realism. Compared to the models used to assess data-rich stocks, models tailored to data-limited stock assessment contain a large degree of uncertainty and therefore, I recommended further exploration of the existing data-limited approaches.

This thesis continued by focusing on a particular group of data-limited assessment methods, which are based on stock reduction analysis. Although such models can cope with low data availability, at the same time, they are particularly sensitive to the misspecification of relative stock status (expressed as the current biomass level relative to virgin biomass), a critical input requirement. However, stock status is unavailable for the majority of data-limited stocks. Therefore, I explored different sources of information used to estimate stock status under such circumstances. First, I considered the use of fisheries experts' opinion and presented a method to elicit expert knowledge using a novel, user-friendly on-line application. To evaluate the experts' ability to predict stock status, I compared the elicited distributions to stock statuses derived from data-rich models. In this work, I explored the performance of experts with different levels of experience in stock assessment, since scientific expertise is not evenly distributed throughout the world, and quantified how well they performed relative to each other. The results indicated that the true stock status is the most significant factor accounting for bias in expert opinions, followed by their experience level.

Nevertheless, expert opinions are often used to inform management decisions and this thesis revealed that for data-limited stock assessment, expert elicited stock status priors potentially can be highly biased, leading to highly biased harvest recommendation levels. A way to overcome this issue is by calibrating expert judgment. To achieve this, my coauthors and I developed a hierarchical Bayesian model for expert calibration. The model's main assumption is that experts' biases vary as a function of the true value of the parameter, as identified in the expert elicitation experiment. Experts' bias function was explicitly modeled, following the supra-Bayesian approach, using Gaussian processes to construct the prior, and the

results of the expert elicitation experiment were used as calibration data to infer the posterior. The constructed models were tested both with simulated data and with the expert elicitation results. The tested models for expert judgment calibration, substantially improved stock status predictions compared to those that were uncalibrated and in comparison to vague uniform guesses, thereby demonstrating the value of calibration in minimizing expert bias.

In the last article included in this thesis, uncalibrated and calibrated expert opinion derived stock status priors were compared to productivity and susceptibility (PSA) vulnerability scores and catch trend-derived (Boosted regression trees; BRTs) stock status priors. Furthermore, the performance of each of these methods was evaluated and compared to a commonly used prior that assumes a stock is at $B_{40\%}$ (i.e. 40% of the virgin biomass). First, I evaluated the degree of bias in estimating true stock status and then, the effect of bias on the estimation procedure of overfishing limits (OFLs) in the specific assessment models for ten data-rich stocks. All, with the exception of fisheries experts with no experience in stock assessment, provided more accurate priors about stock status than the $B_{40\%}$ rule. Experts with experience in stock assessment produced particularly informative and accurate priors, exemplifying their important role in the assessment procedure. Based on the performance evaluation and the data requirements for constructing a stock status prior, I recommended a procedure for selecting the most appropriate prior(s).

1. Introduction

1.1. Exploitation of wild populations

As an obligation of their predominantly omnivorous lifestyles, humans have always exploited wild populations of plants, animals and fungi. Both terrestrial and aquatic organisms have been targeted for consumption and other needs. After the industrial revolution, the pressure on natural resources has intensified and has led to dramatic declines owned to exponentially increased exploitation rates. There is mounting evidence that overexploitation has led to the direct demographic extinction of many populations and species (Burney and Flannery 2005). This loss of both biological and genetic diversity is alarming, however, resource managers have been reluctant to accept these effects on harvested populations (Allendorf et al. 2008). Wild foods form a significant portion of the total food basket for households from agricultural, hunter, gatherer and forager systems (Bharucha and Pretty 2012), proving their importance for food security (Erskine et al. 2015). In addition, harvesting wild populations supports human communities financially, for example, global marine fisheries offer employment for 260 million people, encompassing full-time and part-time jobs in the direct and indirect sectors, with 22 million of those being small-scale fishers (Teh and Sumaila 2013). Therefore, it is vital that we learn how to manage wild populations in a sustainable manner, in order to protect the ecological systems that support our societies. Thus, scientific research exploring population dynamics and their interactions with human exploitation emerged.

1.2. The development of fishery science

Fishery science has been recognized as a scientific discipline since the middle of the 18th century, when Norwegian authorities requested that scientists explain the observed fluctuations in Atlantic cod catches. At that time, the western and "developed" world believed that natural resources were unlimited. On the contrary, indigenous people have known for thousands of years that overly intense exploitation can result in population declines (Jennings et al. 2001) but this is not the story of this thesis. It was the sentiment of Thomas Huxley, the late president of the Royal Society during the Great International Fishery Exhibition held in England in 1883, that the cod, the herring fishery and probably all the great sea fisheries are inexhaustible (Jennings et al. 2001). Unfortunately, at some point in history it became apparent that this statement is incorrect and that it is necessary to identify the effects of harvesting on wild populations and thus, many laboratories for fisheries science were established (Cushing 1988; Smith 1994).

The International Council for the Exploration of the Sea (ICES) was the result of this united effort to examine the effects of fishing on wild populations and was established in 1902 in Copenhagen, Denmark. Moreover, after World War II other new scientific and regulatory bodies were created, such as the International Commission for North-West Atlantic Fisheries (ICNAF) and the Inter-American Tropical Tuna Commission (IATTC) (Jennings et al. 2001). At the same time that fisheries scientists began collecting data on exploited stocks, Lotka (1907) developed the foundation of population dynamics theory, which evolved and grew under the care of subsequent theorists (Lewis 1942; Leslie 1945; Leslie 1948). In the field of fisheries science, the first steps towards sustainable exploitation were taken early in the 1930's when the fundamental idea behind maximum sustainable yield (MSY) was invented. Thereafter, scientists began exploring not only the fluctuations in fish populations, but the causes of flux as well (Russel 1931; Hjort et al. 1933; Graham 1935). A few decades later, the publication of Beverton and Holt's (1957) textbook "On the dynamics of exploited fish populations" brought with it a tremendous revolution in fisheries science and the grounds for modern fisheries stock assessment and management. The Beverton-Holt model for stock-recruitment relationship is still widely used today exemplifying how progressive this work was.

Since then, fishery science has developed into an interdisciplinary field, requiring fisheries managers to balance competing objectives such as, fishing at MSY and aiming for the maximum economical yield while minimizing by-catch and the risk of overfishing (Ricard et al. 2012). With so many goals in mind, it is very challenging to make appropriate recommendations and in the last few decades, fisheries management has received substantial criticism, as overfishing has remained a global problem (Worm et al. 2006; Link 2010).

As a result of the criticism, the perceived global fisheries crisis and new directions in fisheries management, new more intuitive management routines, such as harvest control rules (HCR), have been put in place (Kvamsdal et al. (2016). Figure 1 illustrates the basic steps involved in current fisheries management schemes within the European Union and ICES's contribution of scientific advice. This thesis focuses on the step of stock assessment.

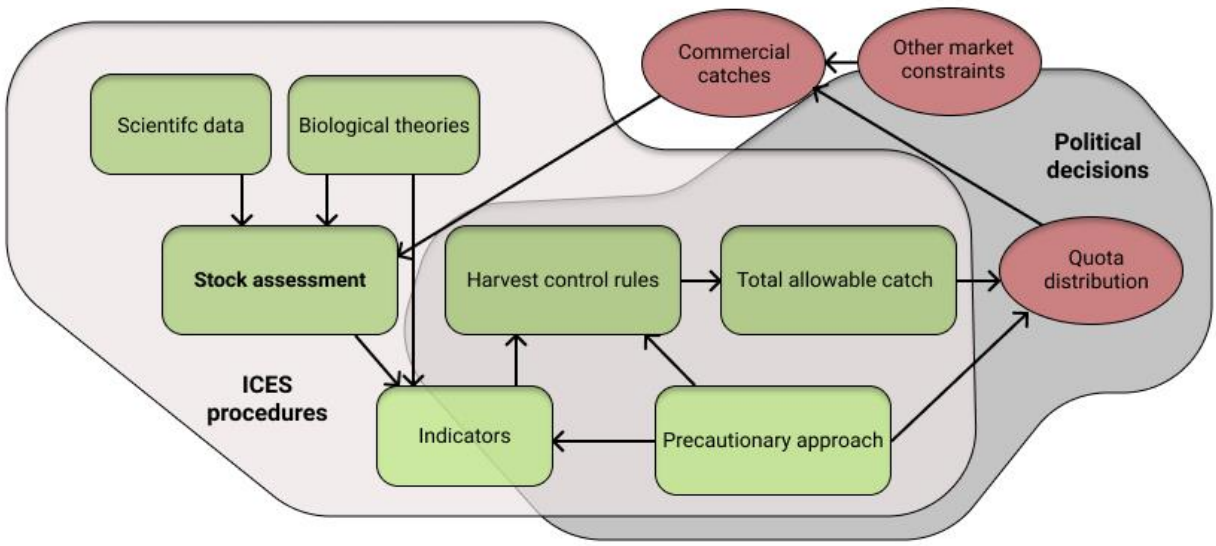

Figure 1: The current fisheries management system within the EU. The pink shaded area depicts ICES related activities and the grey shaded area covers the political domain. Arrows suggest information/impulse flows. Adapted from Kvamsdal et al. (2016).

1.3. The first steps in stock assessment

With the term fisheries stock assessment we mean the demographic analysis designed to determine the effects of harvesting on fish populations, while evaluating the potential consequences of alternative harvest policies (Methot and Wetzel 2013). Three model types have dominated stock assessment since the 60s: 1) methods reliant on catch-at-age data, 2) biomass models and 3) length-based approaches (Hilborn 1992). They are all derived from the general theory for catch and effort data analysis (Schnute 1985) but tailored to different input requirements. Virtual population analysis (VPA) set the grounds for catch-at-age models (Gulland 1965) and was further developed by Pope (1972). VPA requires a time-series of fishery catch-at-age data to reconstruct the virtual abundance of each annual cohort fished. Many variations of the basic concept of a catch-at-age model have been used at one time or another (Deriso et al. 1985; Gavaris 1988, Shepherd 1999) by different assessment bodies. Biomass models, also known as surplus production models, are based on Schaefer's (1954) work. Biomass models' basic input requirements are a time-series of an abundance index (e.g. effort) and total catches in weight for the same years. Both data sets are necessary to calibrate the production model and solve for a given population. Length-based models were developed as an alternative, to accommodate for situations that catch-at-age or abundance index data were unavailable. For example, many fish (e.g. tropical reef fish) are difficult to age, and length distribution data can be easier obtained. In general, length-based methods attempt to mimic catch-at-age analysis but they do not necessarily contain in their core a strong length-age relationship (Jones 19864; Schnute 1987; Fournier et al. 1990). Regardless of which approach one might use, depending on data availability, all consist of well-defined algorithms that utilize few specific data-types and produce outputs required in management decisions (Hilborn 1992).

1.4. Stock assessment today

With the development of VPA analysis as a starting point in the 60s, stock assessment has evolved into integrated assessment (IA) today (Fournier 1990; Hilborn et al. 2001a; Bull et al. 2005; Methot and Wetzel 2013). Advancements in computational power (e.g. Automatic Differentiation Model Builder (ADMB) software; Fournier et al.2012, Template Model Builder (TMB); Kristensen et al. 2016) have been essential for IA applications and driven much of the recent development in fisheries stock assessment. IA incorporates information from different sources, in an appropriate raw form, within the same framework. A strong element of IA is that the data do not come from different pre-analyses with different assumptions (which is the case with older forms of assessment models), thus ensuring consistency, as it is unlikely that the same processes will be modelled differently within the same model. Furthermore, given that all data are incorporated into the same analysis, it is easier to explore estimation sensitivities of the various data sources and assumptions (Maunder and Punt 2013). Examples of IA models are: Stock Synthesis (Methot and Wetzel 2013), which is widely used around the globe, earlier approaches such as Coleraine (Hilborn et al., 2000), CASAL (Bull et al., 2005) and MULTIFAN-CL (Fournier et al., 1998) and others such as Gadget (Begley and Howell 2004) and SAM (Nielsen and Berg 2014) that are used more in Europe.

In addition to IA, the other approaches to stock assessment, some of which discussed in section 1.3., are still used nowadays. Dichmont et al. (2016) reviewed the available stock assessment models used in the USA, and found the choice of a model being very region-specific, owed to several factors such as history in how assessments have been conducted and data availability. In general, though, the tendency is towards IA due to the advantages mentioned earlier. The level of biological complexity and efficiency in modern stock assessment models has significantly increased, but there is inherent uncertainty in model outputs (and will always be). Consequently, a whole field of research has evolved around developing and testing management procedures that are robust to this uncertainty, widely known as management strategy evaluation (MSE; e.g. Butterworth and Punt 1999; Smith et al. 1999, Ives and Scandol 2013; Punt et al. 2016; Nakatsuka 2017). Despite a dichotomy among scientists, whether or not stock assessment models can help achieve management goals, there is strong evidence that modern fisheries management and the implementation of IA, results in recovering stocks (Hilborn and Ovando 2014). IA models are indeed powerful and very useful, but also require high level of scientific expertise and long time-series of detailed fisheries dependent and independent data, which is largely lacking for the majority of exploited stocks (so called data-limited stocks). Thus, data-limited stocks are precluded from assessment schemes under data-driven models (Wiedenmann et al. 2013), which lead to the development of data-limited stock assessment approaches and the focus of this thesis.

1.5. Aim and structure of the doctoral thesis

In my thesis, I was interested in exploring data-limited fisheries stock assessment models. Costello et al. (2012) estimated that around 80% of the world's fisheries, representing more than half of the global catches in weight, are still unassesed, which are faring more poorly than those under assessment and management schemes (Hilborn and Ovando 2014), an issue of major concern for food security, social and ecological sustainability. In Article [I], I created a synthesis of recent data-limited assessment models that emerged as a result of the necessity to assess all exploited fish stocks. In this work (Article [I]), I first described the main characteristics of data-limited stocks, then presented the available quantitative approaches in order of increasing reliance on data, discussed the advantages and disadvantages of each approach and finally made suggestions for future developments. The focus of Article [II] was on specific data-limited assessment models that require information about stock status as input; stock status is expressed as B_t/B_0 where B_t is current biomass and B_0 virgin biomass. The goal was to try to imitate a real data-limited assessment, which used expert opinion as the stock status input. Stock status was inferred from expert knowledge because it is generally a model derived quantity and thus, unavailable for unassessed stocks. I provided the experts with amount and quality of data that would likely be available in a real situation. I evaluated expert performance with calibration data (from data-rich stocks) and was thus able to explore and quantify the bias in expert opinions and identify the variables that affect their performance. Given that the specific data-limited assessment models are very sensitive to mis-specifying stock status, in Article [III], the results of Article [II] were used to develop a model capable of calibrating expert opinions, in order to improve stock status predictions compared to the original ones (Article [II]). Gaussian process were used to describe the expert bias function and the alternative models were both simulation and real-data tested. Finally, in Article [IV] I evaluated how well expert opinions (both uncalibrated and calibrated) and other sources of stock status information performed in a) predicting stock status and b) in estimating overfishing levels (OFL), with the predicted stock statuses used as model input. Finally, I made overall recommendations for specifying stock status priors based on available resources and the performance evaluation results. Figure 2 presents the structure and connection between the articles included in this thesis.

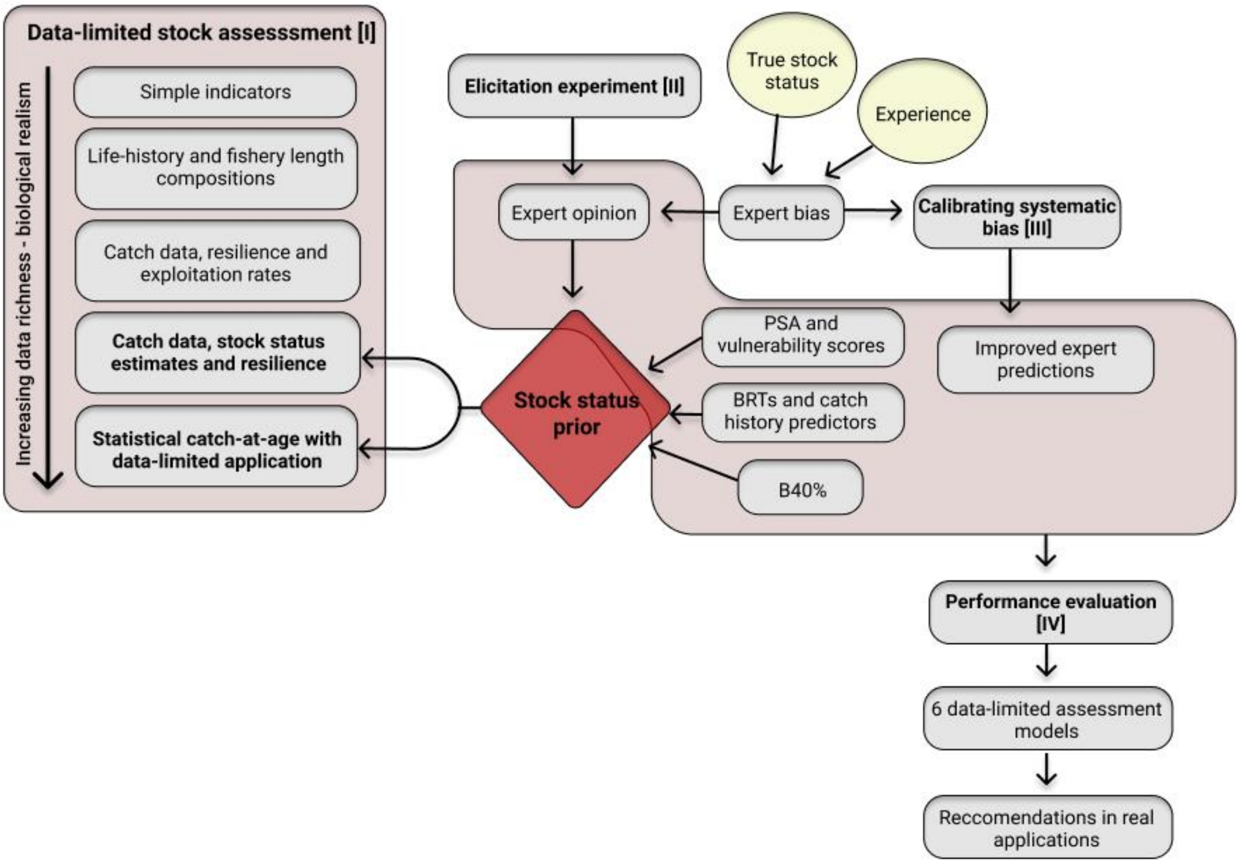

Figure 2: Flowchart of the thesis. Arrows suggest information/impulse flows.

2. Data-limited fisheries (Article [I])

2.1. What makes an exploited stock data-limited?

Dowling et al. (2015) describe a fishery as data-limited when a) a quantitative stock assessment cannot be undertaken due to limitations in the type and/or quality of available data, and/or 2) the best available information is inadequate to determine reference points, current stock status, and/or the exploitation of targeted stocks. Data-limitations persist in both old and newly developed fisheries, such as small-scale traditional practices and industrial deep-sea ventures respectively. Furthermore, they are not region-specific, as they can be found in both developed and developing countries, and not limited to specific gears and fishing methods.

Small-scale fisheries often consist of multi-species and multi-gear operations and therefore, this complex structure poses difficulties in using traditional single-species stock-assessment models (Stergiou et al. 2007). In most cases, little quantitative information is available for these fisheries in comparison with large-scale commercial ventures, due to their limited contribution to GDP, and, consequently, limited administrative interest in them (Salas et al. 2007).

New fisheries form the second main category of data-limited stocks. Industrial fishing activities have expanded in developing countries and the deep-seas, as demand for fish products continues to rise and many continental-shelf and epipelagic-oceanic fishes become overexploited (Worm and Branch 2012). Inevitably, whenever a new stock is discovered there will be high levels of uncertainty regarding its size, productivity and structure, potentially leading rapidly to overexploitation (which has occurred in the past with many deep-sea fisheries, Norse et al. 2012), stock collapse and fishery closures. For new fisheries stocks, landings-records and research-data are largely lacking, prohibiting the use of traditional stock-assessment models (Boyer et al. 2001).

2.2. Data-limited stock assessment

Many International agreements (Law of the Sea, UNCLOS 1982; Code of Conduct, FAO 1995; Magnuson-Stevens Act, NMFS 2007; Common Fisheries Policy, EC 2011) actively support the sustainable exploitation of aquatic resources, identifying it as an area of high priority and concern. Stock assessment is an invaluable tool to guide fisheries toward this goal, even data-limited fisheries, despite the fact that they are often incompatible with traditional models. Therefore, many alternative assessment models applicable to data-limited stocks have emerged in recent years. These approaches exist along a continuum of data accumulation and biological realism. As biological realism and data-richness increase, uncertainty decreases. The simplest of these approaches use catch time-series (Froese and Kesner-Reyes 2002; Anderson et al. 2012) to provide a stock status indicator based on predefined catch ratios, while the most complex are age-based (Cope 2013) and tailored for stocks with limited quantitative information.

The majority of data-limited methods fall between these two extremes. Fishery length-compositions are relatively easy and low-cost to collect and thus, many approaches use them. Some make inference based solely on length compositions (Froese 2004; Cope and Punt 2009; Froese et al. 2018), while others combine them with life history information (Kokkalis et al. 2014; Hordyk et al. 2014). A large proportion of these data-limited models and the focus of this thesis, are those that require input information about relative stock status. The majority of these models are modified versions of the classic stock reduction analysis (Kimura and Targart 1982; Kimura et al. 1984) and can produce estimates for future catches (MacCall 2009) and biological reference points (Dick and MacCall 2011; Martell and Froese 2013). Table 1 in Article [I] provides an overview of available quantitative models for data-limited stock assessment at the time of publication, their assumptions, limitations and examples of their application. Since the publication of Article [I], two additional based on SRA models have emerged. Froese et al. (2017) published an updated version of the model presented by Martell and Froese (2013), and Zhou et al. (2018) presented their own version of a similar model. Both follow the same principles and sensitivities as the models of interest and are therefore, also included in Article [IV].

2.3. Steps forward

An alternative approach to the practice of selecting a single model that fits the available data, is founded on the idea of data-richness continuum. The same models can be applied to both data-rich and data-limited stocks where the model learns from the stocks that have more data and can therefore make predictions for stocks with less data. Thus, for a particular fishery, we simply increase certainty in parameter estimation, as more data accumulate (Kuparinen et al. 2012; Bentley 2014). Moreover, data-limited stock assessment could further benefit within the hierarchical Bayesian framework that is flexible and effectively accounts for uncertainty. Within its context, the existing knowledge for a given population is encoded in a probabilistic manner and gets updated once/if additional data become available. The benefit of this approach is that low data availability does not mean compromised biological realism. However, it is computationally demanding, requires a high level of skills to operate and lengthy simulation time. In an ideal world, fisheries assessment should follow this approach, but in most cases, time, resources and skilled personnel are lacking. It is rather common that during workshops/meetings simpler methods are preferred over because they are easier to explain, easier to implement and run faster (Dichmont et al. 2016). Thus, data-limited models have a real value, as they fulfill the simpler models characteristics while they can also provide important insights into stock status when little quantitative information is available, until more data accumulate. A beneficial strategy is to apply multiple data-limited models that use a variety of data types and thus, might be more suitable than selecting a single model type that may not be appropriate (Carruthers et al. 2014; Berkson and Thorson 2015). A way to do this, is to combine the different data limited methods into ensambles and/or super-ensambles that utilize the best part of each model (Anderson et al. 2017; Rosenberg et al. 2018; Walsh et al. 2018). Many methods and approaches to application are available and to be able to identify the strengths and weaknesses of each of them, and thus translate them into appropriate HCRs and management strategies, it is vital to further explore them with performance evaluations and MSEs.

3. Focus on stock status and expert knowledge (Articles [II, III, IV])

3.1. Why stock status?

As was mentioned in section 1.5, the focus of this thesis are the data-limited models that require prior information about relative stock status. In these methods, stock status is a model input given as a distribution from which values are resampled or it is defined by upper and lower bounds. The user-specified stock status, the given catch data and biology of the stock solve for a particular biomass time-series based on which potential future yields are calculated. These models can be problematic, however, because the estimated yields are highly sensitive to stock status misspecification. Model performance evaluations (Wetzel and Punt 2011; Wiedenmann et al. 2013; Carruthers et al. 2014) have shown that under- or over-estimating stock status leads to under- or over- estimation of potential yields respectively. Such findings can have serious implication in data-limited fisheries management since, overestimating yields could potentially lead to over-exploitation or even stock collapse, without even realizing it, and underestimating yields can lead to substantial unnecessary revenue loss. However, the existing performance evaluations have only explored the effect of stock status miss-specification to a certain degree, requiring further exploration due to its critical role.

3.2. Where can one find information on stock status?

Stock status is typically a derived stock assessment output, so it is usually not available in data-limited situations. So what do fisheries scientists do when they have to use models that require stock status input? One option is to use information from data-rich species with similar exploitation patterns and/or biology in a "Robin Hood" manner (Punt et al. 2011). An interesting approach developed by Cope et al. (2015), constructed a stock status prior using Productivity-Susceptibility Analysis (PSA) vulnerability scores thus, requiring estimating vulnerability values for the stock of interest. More recently, Zhou et al. (2017) used boosted regression trees (BRTs) and catch history patterns to infer stock status that is suitable as a data-limited assessment model input option. In the fisheries literature, expert knowledge is also frequently recommended as a source of information (Berkson and Thorson 2015; Newman et al. 2015). However, to my knowledge, no peer-reviewed publications prior to those included in this thesis investigated the incorporation of expert knowledge into data-limited stock assessment. Therefore, I chose to explore aspects and possible implications of using expert opinion to define stock status (in Articles [II], [III], [IV]).

3.3. Expert knowledge elicitation

Expert elicitation has long been utilized and applied within various fields of research (Uusitalo et al. 2005; Roman et al. 2008; Zickfel et al. 2010). When time and resources are limited and the existing knowledge is inadequate, society calls on experts for advice (Burgman et al. 2012; Morgan 2014). The aim of elicitation is to formulate expert knowledge and beliefs about uncertain quantities into a probabilistic form (Garthwaite et al. 2005). In a typical elicitation experiment, the analyst collects the data and interprets them in the way he/she considers appropriate. Therefore, results can often be an analyst and expert view mixture (Mäntyniemi et al. 2013). Background information and assistance from the analyst are often provided to experts so they can make inference for an unknown quantity (O'Hagan et al. 2006). However, expert opinion has its shortcomings. Chief among these, as demonstrated by Tversky and Kahneman (1974), is the fact that both experts and lay people are sensitive to a host of psychological idiosyncrasies and subjective biases. People use heuristics leading to systematic biases such as conjunction fallacy, base rate neglect and miscalibration (Kynn 2008). Heuristics are 'rules of thumb' that are used to find quick solutions and are, at best, simplifications of the correct probability and at worst a 'guesstimate' of the answer. Nevertheless, experts' opinions are valuable, as they can have specialized knowledge obtained through training and experience, proven by their personal "track records" of efficient and effective application (Gullet 2000).

4. Methods

4.1. Eliciting expert knowledge

4.1.1. Experimental design

To imitate a real case of data-limited assessment, 18 data-rich stocks from the US. Pacific Fishery Management Council (PFMC) and Alaska Fisheries Science Center (AFSC) assessed with Stock Synthesis (SS; Methot and Wetzel 2013) or similar IA model were selected for my research and the available fishery data transformed to mimic those typical of a data-limited stock. Life history details for each species are provided in Table 1 in Article [II]. In addition to the assessed data-rich stocks, two simulated stocks representing the two main families (Sebastidae and Pleuronectidae) of the selected species were also used as control cases. The DLMtool package in R developed by Carruthers and Hordyk (2018) for MSE purposes was the platform for creating an exploitation history for the simulated stocks. To be able to explore different possibilities and variables that can potentially affect expert performance, four different data-richness scenarios were created (Figure 3) and six fisheries experts with different degrees of experience in stock assessment (experienced-novice-inexperienced), selected on the basis of their academic backgrounds, participated in the elicitation.

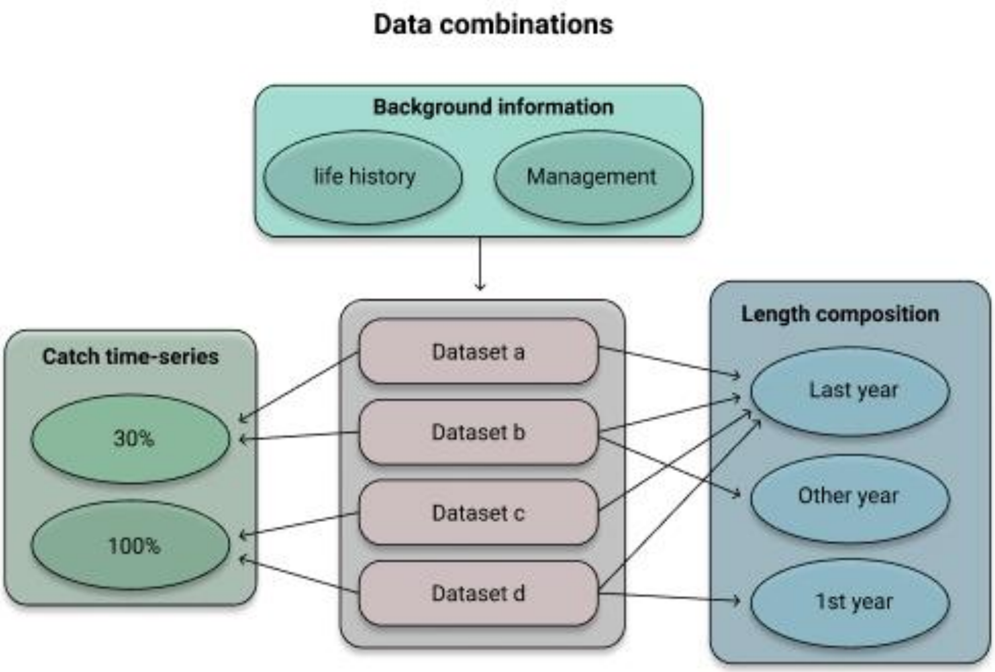

Figure 3: Constructed datasets. Background information was provided in all cases. The catch history was represented either by 30% of all the available years for a given stock or the entire exploitation history. Length compositions consisted of the 1st year, a mid-point year or the last year available within the exploitation period. Arrows suggest information/impulse flows.

4.1.2. Elicitation tool

For my purposes, I developed my own web-based application using the "Shiny" R package (Chang et al. 2015). Figure 4 is the user interface illustration of the constructed elicitation tool. The tool provided step-by-step instructions on how to use the application. Within the tool, the dataset and species selection action buttons produced the respective graphs of catch and length compositions combined with background information. Furthermore, scroll bars to adjust the mean and standard deviation of a beta distribution probability density function (PDF) were provided to capture the distribution that best fit expert perception. Additionally, a beta cumulative distribution function (CDF) was also drawn to help experts better visualize the desired distribution. Finally, the expert selected values were saved in a self-populated table after he/she had completed each dataset. Once the process was completed for all stocks and datasets, the final table of elicited stock status distributions was available for download as a comma-separated-values (csv) file and subsequently sent to me.

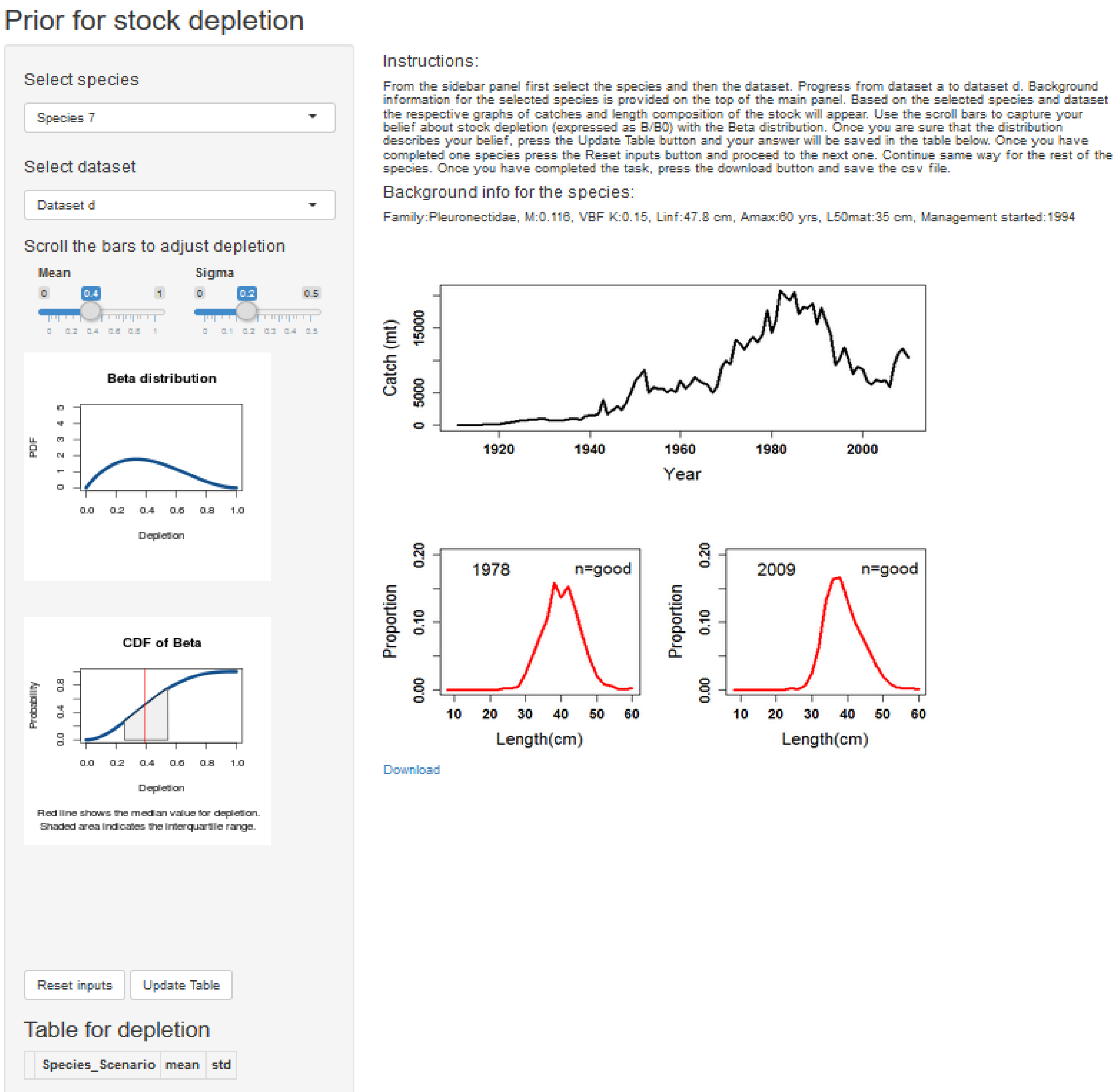

Figure 4: Stock status elicitation tool user interface

4.2. Expert calibration with Gaussian processes

In Article [II] the degree of bias in expert judgment and the variables that affect their performance were identified with the use of calibration data. The next step in this work was to address whether experts' probability assessments could be improved by correcting for their bias. Therefore, in Article [III], a fully Bayesian model for combining expert assessments using the supra-Bayesian approach was presented (French 1980, Lindley and Singpurwalla 1986). Here, the analyst builds an explicit model for the expert biases and updates his/her beliefs about the system under study using the Bayes' rule. Experts' opinions are assumed to consist of subjective biases and are described by their mean estimate, m, and uncertainty estimate s_{ji}. Expert mean estimate m of an unknown parameter, x, (here being stock status) was subsequently assumed to depend on the true parameter value as observed in Article [II]. The experts' ability to correctly assign the mean estimate was also affected by an unknown bias function, b(x) which was also assumed to depend on the true parameter value. The bias function was modelled using Gaussian processes (GP; Williams and Rasmussen 2006). In the constructed model, thee different forms for the bias function were tested: 1) additive bias, 2) logit additive bias, and 3) a marginally uniform prior for the expert's mean estimate. Figure 5 illustrates the model structure used in Article [III].

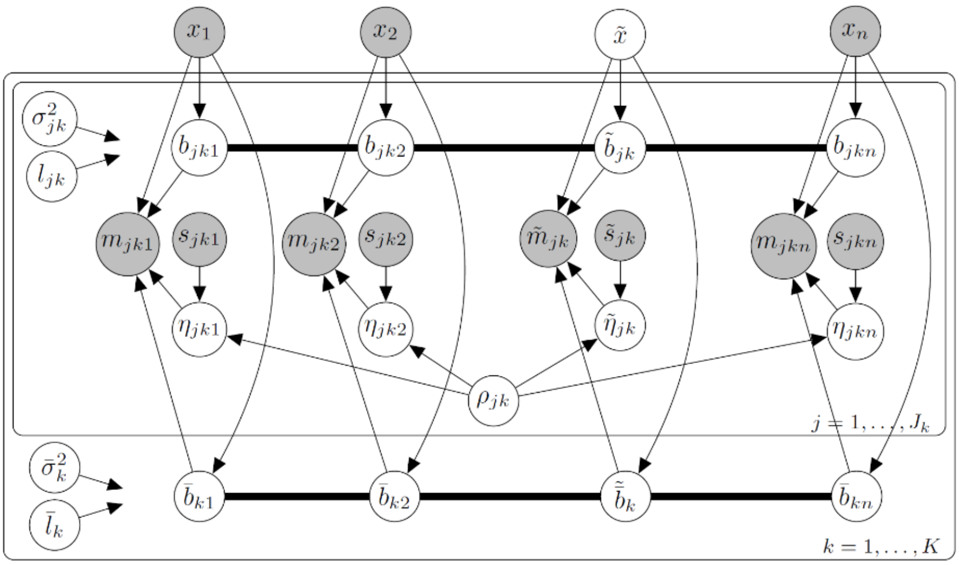

Figure 5: Graphical representation of the model. Gray circles denote the observed nodes, and white circles denote the unknown variables and functions. The variable $\tilde{x}$ denotes the parameter of interest in system for which we do not know the true value and variables x with gray background correspond to calibration systems. The thick black lines denote GP with undirected links between all pairs of latent variables where $\bar{b}_{ki}$ denotes the ι'th group wise latent variable of group k and b_{jki} is the ι'th latent variable of expert j in group k. The inner panel includes the expert wise GP and the outer panel includes the group wise GPs (Assumption A5). Under assumption A5* (prior independence between experts) the outer panel is removed from the model. Graph created by T. Perälä and used here with granted permission.

4.3. Alternative formulas to derive stock status

4.3.1. PSA vulnerability scores

Cope et al. (2015) developed a prior for relative stock status using PSA vulnerability scores (which also rely on expert knowledge). The authors used data from U.S. west coast ground fish stocks as the basis of their analysis both for constructing the prior and consequently testing its' performance. To achieve this, they used the "best available scientific information" (BASI) approach, which uses data-rich stock assessments and decreases the amount of data to imitate data-limited scenarios. The performance evaluation, showed that the constructed stock status prior was more accurate than the common application of stock status assumed at $B_{40\%}$ and improved performance overall.

4.3.2. BRTs and catch history

Using the RAM Legacy database (Ricard et al. 2012), Zhou et al. (2017) developed a BRTs model to correlate stock status to catch data using different predictors in the exploitation trends. Linear regression of scaled catch was the most important predictor, including regression coefficients for the entire exploitation history, the subseries before and after the maximum catch, and in recent years. From all the tested potential predictors, eight were able to explain about 80% of the variation in stock status. Due to the requirement of contrast in the data, the method makes better predictions when the stock is fished down to below half of the B_0. The authors of this approach, demonstrated that the BRT model outperformed existing catch-based stock status estimators (Froese and Kesner-Reys 2002; Rosenberg et al. 2014).

4.4. Performance evaluation

4.4.1. Expert bias from elicitation experiment

Relative error was used to quantify individual expert bias compared to model- and simulation-based stock status values and defined as

$$\mathrm{RE} = \frac{\mathrm{S_e} - \mathrm{S_m}}{\mathrm{S_m}} \qquad (1)$$

where RE is relative error, S_e is the distribution equating to the expert's estimation of stock status, and S_m is the true stock status. The bias was described with the median RE whereas, the imprecision of expert performance was described with the RE interquartile range (IQR). In addition, Table 2 in Article [II] shows the categorical variables tested in linear models as potential explanatory variables of expert performance. Backward stepwise model selection and the Akaike information criterion (AIC) and AIC with a correction for finite sample sizes (AICc) were used for final model selection (Montgomery et al. 2012).

4.4.2. Expert bias calibration

The three bias forms constructed in Article [III] were evaluated by estimating experts' bias functions with a simulation study. Expert elicitation data with varying levels of bias and uncertainty were simulated and thereafter, used to estimate the true simulated b(x). Models' performance in predicting an unknown $\tilde{x}$ in a new system was tested in both the simulation study and with the expert elicitation data from Article [II]. In the real data case, leave-one-out cross validation was used. In simulation studies, the tests were conducted at 10 equally spaced values for x in the interval [0.001, 0.999]. The predictive performance was compared using log predictive distribution statistics (Vehtari and Ojanen 2012) for both simulated and real data analyses. Furthermore, the posterior predictive densities were approximated with a kernel density estimation since the posterior inference was conducted with Markov Chain Monte Carlo (MCMC) methods. In the simulation studies, the root mean squared error (RMSE) between the median of the posterior distribution and the true value were also calculated.

4.4.3. Comparison of stock status priors

The methods to construct (or improve) a stock status prior (hereafter referred as SSPMs), described in sections 4.1-4.3, where evaluated for 10 data-rich fish stocks and applied to 6 assessment models that require input on stock status. The platform for the evaluation was the DLMtool package, which was also used in Article [II]. Selecting the specific R package for the evaluation, was due to its' flexibility. The DLMtool is easy to use and requires only a single input file that consequently is used to apply different data limited methods (DLMs), based on data availability. In addition, it is not limited to the build-in DLMs, the user can create his/her own DLM tailored to specific study cases. The SSPMs were tested in the following package build in DLMs: a) the Depletion Corrected Average Catch (DCAC; MacCall 2009), b) the Depletion-Based Stock Reduction Analysis (DB-SRA; Dick and MacCall 2011) and c) the Catch trend Surplus Production MSY (SPMSY; Martel and Froese 2013). Additionally the following d) the updated Monte Carlo version (CMSY; Froese et al. 2017) of Martel and Froese (2013), e) the Optimized Catch-Only Method (OCOM; Zhou et al. 2018) and f) Simple Stock Synthesis (SSS, Cope 2013) were coded for implementation in the DLMtool. All models were modified accordingly to create new DLMs that fit the DLMtool input requirements. The scenarios tested are described in Table 2 of Article [IV] and the modifications in each assessment model and input requirements in Table 3 of Article [IV].

The predicted stock status from each SSPM was compared to model derived stock status point estimate that was treated to describe the truth, representing the best available scientific information. A sample of 1000 retained runs from each DLM and all case-scenarios formed the basis for the performance evaluation. The estimated OFLs were compared to the official assessment OFLs. The subsequent overall performance metric was used:

$$OPA_{c,m} = \sum_{1}^{n} |\, 0.5 - Prob\_over_{c,m}| \quad (2)$$

Where Prob_over is the probability of overestimating the OFL relative to the official OFL, n is the total number of species, c is the case and m is the DLM. This metric is used to summarize results across species groups (stocks≤$B_{40\%}$ and stocks >$B_{40\%}$) and is based on the idea that overestimating and underestimating the OFL are equally undesirable behaviors of an assessment model. In addition, to be able to identify the direction of poor model performance, as in a data-limited situation it is reasonable to assume that an underestimation is more desirable, the following metric was also applied:

$P_{c,m} = \sum_{1}^{n}\left(Prob\_over_{c,m}\right)$ (3)

The overall performance metrics described in equations (2) and (3) are estimated relative to the official assessment OFLs which are used as a needed common reference point for all tested cases in the DLMs. The reference case 9, with the unbiased prior, represents the best possible stock status input. Therefore, the performance of each individual DLM for cases 1-8 is evaluated relative to the performance of case 9 (i.e, giving that method the correct stock status).

5. Results

5.1. Evaluating expert knowledge

Both bias and imprecision of expert elicited opinion were explained by true stock status and experience level. The other variables tested were insignificant. For bias (median RE), true stock status primarily explained the variability and the interactions between true stock status and expert level explained it to a limited degree. For imprecision (IQR), true stock status was the main explanatory variable, while expert level explained it to a lesser degree. The interactions between true stock status and expert level explained the variability to a very limited degree. Due to the higher IQR range, experienced and novice experts tended to include true stock status in the constructed prior distributions more often than the inexperienced ones. Furthermore, inexperienced experts were more biased and exhibited overconfidence in comparison to experts from the other two levels in general (Figure 6). Regardless of their background, experts followed a consistent pattern in over- or under-estimating stock status that appears to depend on the true stock status (e.g., overestimated status for low true stock status; underestimated for high true stock status).

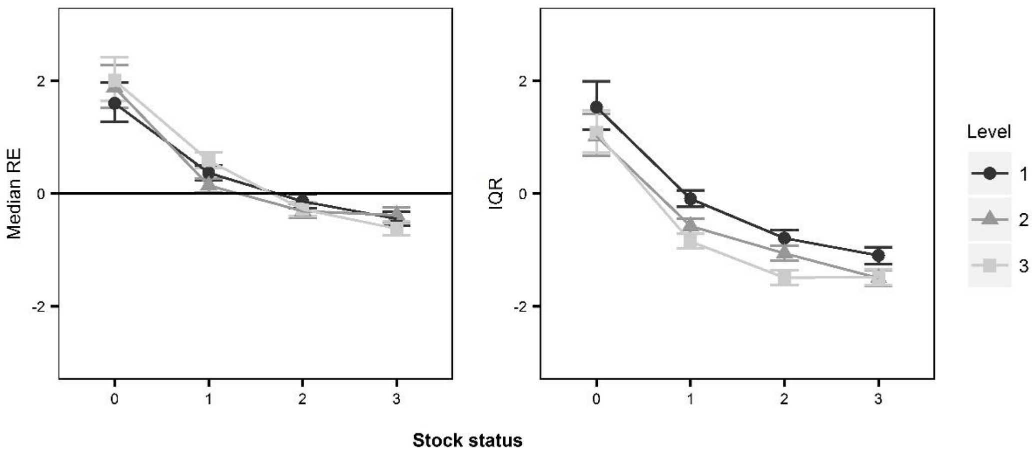

Figure 6: Fitted values (points) and 95% CI (bars) of the expert level effect (Levels 1, 2 and 3) on the different levels of stock status (Levels 0, 1, 2 and 3) in logarithmic scale for median RE and IQR.

For an expert's judgment to be well calibrated, the elicited value (or the median in the case of a distribution) should lie on the diagonal when plotted against the true probabilities for an event (Lichtenstein et al. 1982). To evaluate expert calibration, the median assessed value for stock status was plotted against true status (Figure 7). Locally Weighted Scatterplot Smoothing (LOWESS) was used to inspect expert's judgment calibration and it clearly revealed a shift in the opposite direction of the perfect calibration diagonal. Especially inexperienced experts tended to assign too high values for low stock status stocks, a pattern that is also clear in the effect of expert level-stock status interaction (Figure 6).

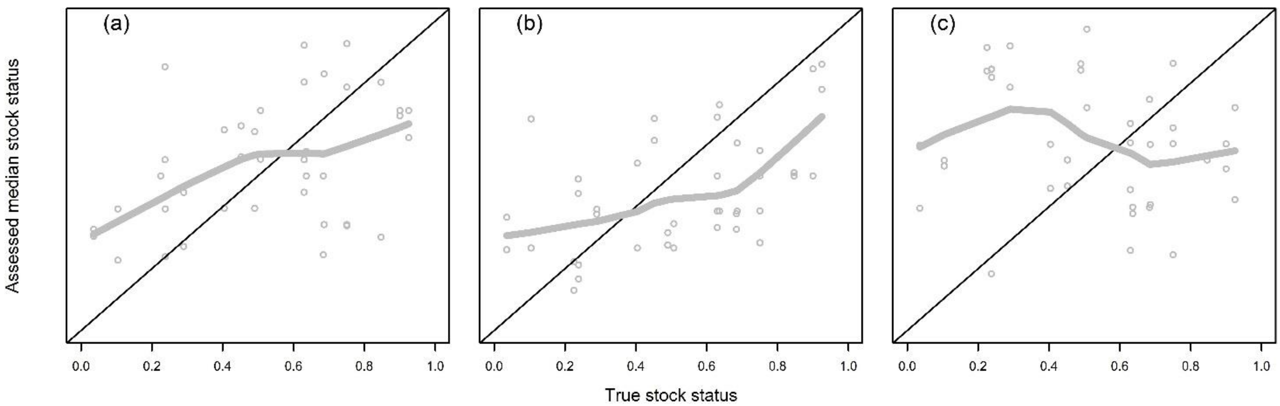

Figure 7: Expert calibration for Dataset d. Assessed mean probability for stock status compared to stock true status. (a) Experienced experts, (b) Novice experts and (c) Inexperienced experts. The diagonal line indicates perfect expert performance and below or above the diagonal line, indicates under and over-estimation of stock status respectively. Locally weighted scatterplot smoothing (LOWESS) is illustrated with the gray line.

5.2. Expert calibration with Gaussian processes

For all the potential models for the bias form tested, in the simulation experiment, the bias function was inferred with satisfactory accuracy. The total amount of data points, the noise in the data and the maximum bias level varied in the simulation testing. With an increased number of data points and decreased noise in the data, the accuracy of bias inference increased and the opposite was observed with larger maximum bias. For the additive bias model, only small differences in performance were observed among the different models (hierarchical and non-hierarchical) thus, performing marginally better than the rest. In the simulation study, a similar pattern was also found when tried to infer an unknown parameter for a new system with smaller difference between the three bias models compared to the bias inference. For the real case data from Article [II], model comparison was performed with leave-one-out cross-validation log predictive density and the coverage of the posterior distributions. Overall Model 3 with hierarchical prior for b(x) combined with only experts' mean produced better predictions relative to the rest of the tested models (Table1 and Table 2 in Article [III]).

5.3. Best performing SSPMs

BRT and PSA derived stock status priors followed the same pattern, as described in Article [II] for fisheries experts, where overestimation was observed for stocks with low stock status and underestimation for stocks with high stock status. The most biased priors came from an inexperienced in stock assessment expert, followed by the common $B_{40\%}$ assumption. On the opposite side, an experienced in stock assessment expert made the best stock status predictions. Furthermore, the expert bias calibration model improved predictions both for novice and inexperienced experts for some cases, while performed poorly for others. BRT and PSA derived stock status, resulted in similar predictions with BRT method being more accurate for stocks with high true status (Figure 1 and Table 4 in Article IV).

The pattern in over and underestimating stock status consequently affected the estimation of OFLs, across species. In general, overestimation of OFLs was observed for stocks that have low stock status and underestimation of the OFLs for stocks that are above $B_{40\%}$ regardless of SSPM and DLM (Figure 8, panels b and d). The reference case with the true stock status resulted in an overall best performance with the least overestimation and underestimation of OFLs, confirming relatively good OFL estimations when an accurate prior is used. Across DLMs, an inexperienced in stock assessment expert seemed to perform the worst and resulted to the highest probabilities of overestimating OFLs. On the other side, an experienced in stock assessment expert due to making good stock status predictions resulted in the best overall performance relative to the reference case. Both the opinion pool and the expert calibration model improved the overall performance for novice and inexperienced experts and the BRT and PSA derived priors followed in performance. Lastly, the B40% rule was the second worse performing SSPM after the inexperienced expert (Figure 8).

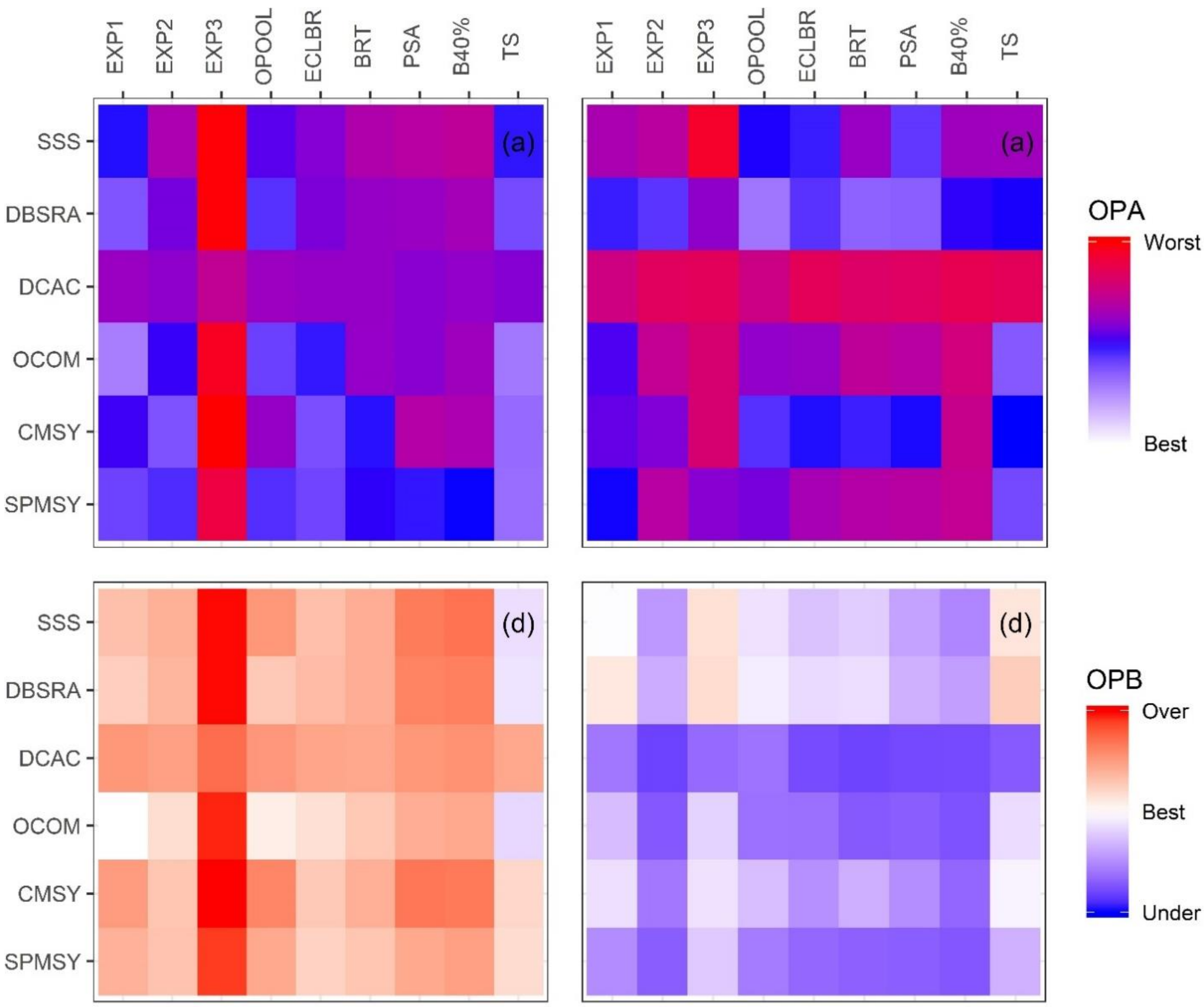

Figure 8: Overall performance evaluation of each SSPM for each DLM across all species. OPA is the performance metric of equation (2) and OPB is the performance metric of equation (3). Results in panels a) and c) are calculated from stocks with stock status≤$B_{40\%}$ and in panels b) and d) are calculated from stocks with stock status >$B_{40\%}$.

6. Discussion

Article [I] provided an overview of recent approaches developed to assess exploited fish stocks with limited quantitative information. This work raised the issues associated with the use of simplistic approaches to stock assessment and the uncertainty regarding model inputs and outputs, while affirming the necessity of developing such methods. Provided that for most data-limited cases, expertise, financial support and time are lacking, this thesis recommends further testing the available assessment methods and exploring how they behave relative to uncertain input requirements. Existing performance evaluations (Wetzel and Punt 2011; Carruthers et al. 2014) have covered various input sensitivities to an extent, but further testing of critical inputs, such as stock status, and model behavior is required. Therefore, this thesis explored model sensitivities and potential implications of misspecifying the stock status input requirement due to its peculiar nature. Usually, model input requirements such as F_{msy}/M ratio, B_{msy}/B_0 ratio (where F_{msy} is fishing mortality that results in MSY and M is natural mortality), life history and growth rate parameters can be easily obtained via meta-analyses, databases such as fishbase (Froese and Pauly 2016) or even by using other species with similar life history as a proxy. On the other hand, stock status is usually a data-rich model derived output and subsequently unavailable in data-limited cases. We could have perfect stock status information if one could count fish in the way we count the number of wolves in a pack, but this is impossible. In order to accommodate for this limitation, defining stock status has been based on arbitrary assumptions (e.g. the stock is on target at $B_{40\%}$), setting a stock status range based on catch trends or finally using expert knowledge, being the person/people responsible for performing the assessment. Given that people are biased by their nature and that data-limited assessments can rely heavily on their opinion, it is vital to explore how and to what degree expert opinions can potentially affect assessment outcomes and to take this source of bias into account when making management decisions.

The expert elicitation experiment of Article [II] revealed that the most important variables for expert performance were true stock status and stock assessment experience. In general, experts exhibited consistent overestimation for low stock status stocks, which could prove to be risky, provided that it could result in overestimated catch levels. On the other hand, experts made better predictions for intermediate true status stocks. This could be explained either, because it is easier to identify these cases or it could be due to using the adjustment and anchoring heuristic. Assuming a stock being at $B_{40\%}$ is the common assumption in such assessment models and therefore, experts might use this reference point as their "anchor". It has been demonstrated though, that adjustments around the anchor are insufficient to provide accurate predictions (Tversky and Kahneman 1974), as supported also in this study. Moreover, before conducting the elicitation, we expected that inexperienced experts would provide judgments that are more uncertain but the results demonstrated the opposite. The high levels of uncertainty found in experienced experts could be the result of the education fisheries stock assessment scientists receive, that is based on the precautionary approach (Hilborn et al. 2001b; González-Laxe F. 2005). In addition, people with strong quantitative background are aware of the fact, that even the highest quality data always contain uncertainty and thus, affecting model outputs regardless of model performance level (Magnuson and Hilborn 2007). Incorporating uncertainty is necessary as scientists have to and should assure resource users that the risks are tractable and manageable (Degnbol et al. 2006). Article [II] is a novel work with findings that were not previously documented in the available literature, and understanding them is necessary to if one wishes to use expert knowledge as a source of stock status information thus, proving vital to explore the effects of subjectivity and experience (or other not yet discovered variables) in estimating future catch levels.

Moreover, in Article [II] I developed an on-line elicitation tool using the "Shiny" R package. Many traditional approaches to expert elicitation (O'Hagan et al. 2006; Roman et al. 2008; Johnson et al. 2010; Zickfeld et al. 2010) can be time consuming, since they may require in-person meetings between the experts and analyst or extensive background materials. The application I constructed involves no such logistical limitations. In addition, all the required background information and tools to describe expert opinion in a probabilistic manner, were available within the same user interface. The application is an example of how a very flexible R package can be used for expert elicitation purposes. A simple method such the one developed in Article [II] has multiple benefits as it increases efficiency and logistical convenience, and as a result can decrease engagement and participation timeframes. A challenge with expert elicitation in general is that people have little motivation to participate because of the time consuming nature of completing the elicitation process (Voinov and Bousquet 2010).

The expert bias identified and quantified in Article [II] is still a limitation to the construction of an accurate stock status prior, but this work showed that experts can indeed provide informative and accurate priors on stock status, being an important advancement from the common assumption that a stock is at $B_{40\%}$. The next step in this work was to address the issue of whether experts' probability assessments can be improved. In Article [III], a hierarchical Bayesian model was presented to calibrate and combine multiple expert assessments following the fully Bayesian approach. Expert assessments were linked to the analyst's prior distribution with a likelihood function, to allow the calculation of a posterior distribution. Furthermore, different prior formulations for expert bias were proposed, built with hierarchical GPs, and then, demonstrated that these models can successfully infer expert bias and calibrate expert assessments before a real case study.

Using the results from Article [II], all alternative models improved the predictions of stock status compared to the uncalibrated expert assessments. Hierarchical model 3 that used only experts' mean value exhibited an overall superior performance compared to the rest of the models tested, individual assessments, uniform prior and explored pooling methods and was therefore selected for testing in Article [IV]. The model was built to accommodate for differences in calibration owing to the various experts' background and experience in stock assessment. A bias that does not change abruptly is indicative of an expert that provides accurate assessments consistently and therefore, predictions concerning the unseen x are improved based on that expert's assessments. On the other hand, the model neglects experts with less consistent prediction patterns. A caveat of this approach is that the calibration can be sensitive to experts that occasionally give inaccurate assessments as the model learns to trust that expert.

The expert elicited priors from Article [II] and the model of Article [III] were part of the performance evaluation conducted in Article [IV]. Expert raw and calibrated stock status priors were compared to other methods (BRTs and PSA) that have been explicitly developed to construct a stock status prior as an alternative to the $B_{40\%}$ rule. A general pattern of overestimating OFLs for stocks bellow $B_{40\%}$ and underestimating OFLs for stocks above $B_{40\%}$ was observed (Figure 8), due to the tendency to over- and underestimate stock status in the different SSPMs for stocks with low and high stock status respectively (Article [IV], Table 4). In this work, I demonstrated that when an expert with sufficient experience in stock assessment is responsible for specifying the stock status prior, the performance of the tested DLMs is comparable to when a very accurate prior is used. On the other hand, when experience in stock assessment is limited, the constructed priors can be highly biased and thus, using the model for calibrating experts' estimates developed in Article [III] is beneficial (and necessary), as it can improve predictions of stock status and the overall performance across DLMs (Figure 8). However, expert judgment is not always feasible or even desirable way to construct a stock status prior and therefore the other tested methods (BRT and PSA) can subsequently be utilized. These SSPMs, produced more biased priors and performed worse compared to an experienced expert relative to the reference case of using true stock status, but they are consistent in their bias and better alternatives than the $B_{40\%}$ assumption. Article [IV] demonstrated that the common $B_{40\%}$ assumption, is the second worst performing SSPM overall and either fails to provide a precautionary measure or is overly precautionary.

Stock status misspecification is a critical issue, as it strongly affects the estimation procedure, and therefore has been discussed extensively in the existing literature (Wetzel and Punt 2011; Wetzel and Punt 2015; Cope 2013; Carruthers et al. 2014) and lead to the alternatives approaches developed to define stock status explored in Article [IV]. However, there are no specific recommendations and guidelines on how to overcome the existing limitations and how to select an appropriate stock status prior, even though the existing pool of potential choices has expanded (Cope et al. 2014; Zhou et al. 2017; Article [II]). This deficiency is evident even in the ICES WKLIFE working group, which focuses on the assessment of data-limited stocks, in which has clearly been acknowledged that good guidance is required to make an appropriate choice for the stock status requirement (ICES 2015). Based on the findings of Article [IV] and the data requirements to construct a stock status prior, I created a decision tree that can provide assistance in specifying a stock status prior (Figure 9). Even though the results of Article [IV] suggest avoiding the $B_{40\%}$ assumption, this option is nevertheless included in the decision tree but its' use is recommended with caution and only if there is strong evidence that the stock status is indeed at $B_{40\%}$. The decision tree makes recommendations for appropriate course(s) of action for a given situation based on data availability in conjunction with the results of Article [IV] (Figure 8) and does not limit the user to specifying only a single prior. For example, different sources of information that exhibit consistent bias, such as an experienced expert or BRT and PSA derived priors, could be combined to create a single prior or used individually as model specification alternatives.

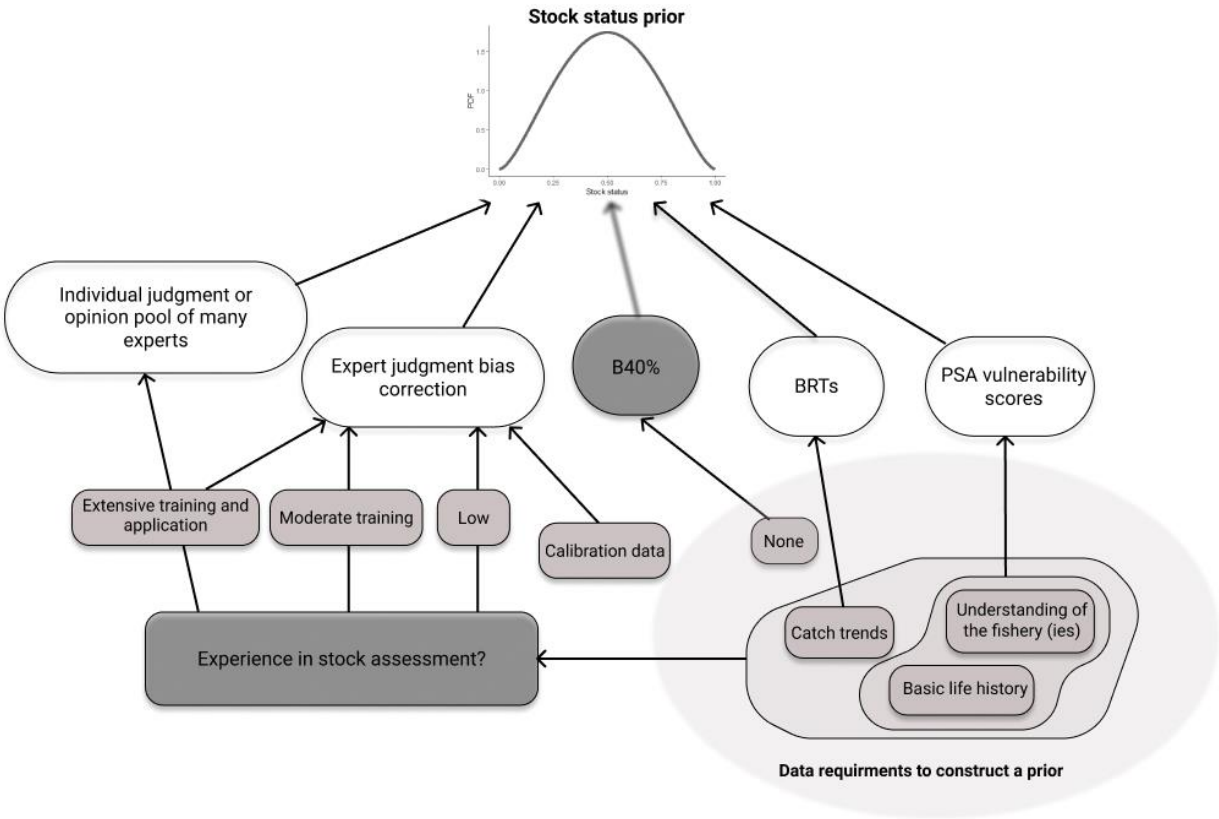

Figure 9: Alternative routes to create a stock status prior. B40% is presented here as it has been used in the past in lieu of any other information, but should be cautiously employed. Arrows indicate information/impulse flows.

7. Future directions and conclusions

In this thesis, the natural desire to answer questions arising from each subsequent article, lead to the interconnected Articles [II], [III] and [IV]. However, it was not possible to address all the questions and issues raised by this work and thus, these can become an avenue for future development and exploration. Before Article [II] no peer-reviewed publications on eliciting stock status from expert knowledge was available, and this work highlighted that an assumption that experts can fairly good specify stock status can be valid but also very false. Even though stock status is a highly sensitive input requirement to several data-limited models, it is often overlooked. Article [IV] revealed that when experts with limited to no experience in stock assessment set the stock status prior, it can cause highly overestimated OFLs and potentially lead to overfished stocks. The application of the calibration model developed in Article [III] significantly improved their biased judgments and subsequently lead to improved OFL estimates. However, there were instances that the calibration model performed poorly, but model prediction can further be improved with an increased size of calibration data, as shown in the simulation data experiments (Article [III]). An increased sample of calibration data would also assist in better understanding the biases in expert opinions and to explore different aspects of data thresholds and the value of information (Magnuson and Hilborn 2007) as discussed in Article [II].

A novel way to achieve this, is to develop a calibration data tool with a common reference baseline, using an online application as the one developed in Article [II] and import data directly from a database, e.g. the RAM legacy database (Ricard et al 2012). Such a readily available calibration data tool could be used in real stock assessment scenes, and subsequently apply the expert bias calibration model to construct a “bias corrected” stock status priors. This tool would provide a large amount of calibration data (without the need to construct them from scratch) which subsequently would result in improvement of the calibration model’s predictive capacity, thereby, becoming less sensitive to the infrequent misjudgments made by typically trustworthy/reputable/accurate experts. It is my personal belief that a calibration model, like the one developed in Article [III], could be very beneficial and should be utilized when expert bias is systematic and predictions can be improved. There may potentially be ethical issues regarding whether or not expert opinion should be corrected, but given that we know people are biased by definition, why not to do so if it allows predictions of unknown quantities to be improved?. This work has filled a knowledge gap in data-limited stock assessment and further explored the implications associated with stock status prior misspecification. Finally, I hope that this work as a whole and the constructed framework for selecting appropriate stock status priors, can assist fisheries assessors to make effective decisions for data-limited assessment while taking under consideration the associated limitations.

8. Acknowledgments

I would like to thank everyone who has been part of my doctoral studies, which has truly been a journey where one fortunate introduction has led to another. First, I would like to thank my MSc thesis supervisor, Prof. Henrik Gislason of DTU Aqua, for the opportunity to work with and learn from him about fisheries science and in particular, for his introduction to the world of data-limited fisheries. These early steps lead to my PhD research proposal, which was ultimately funded by the University of Helsinki. As such, I would like to thank the Doctoral Programme in Interdisciplinary Environmental Sciences (DENVI) for believing in me, my project, and for offering me the privilege of a four-year-long funded position. This work would not have been possible if it was not for my PhD thesis supervisor, Assoc. Prof. Anna Kuparinen, who supported me from the beginning and enthusiastically wrote and submitted the research proposal with me, even though we had not yet met in person. I would also like to thank Anna for her guidance and for allowing me the freedom to explore my own interests and steer my work in the direction I saw fit.

The first year of my doctoral studies was exploratory. During that time, I was learning more about my topic and the questions I wanted to answer. For inspiration, I visited the University of Washington and Prof. André Punt, who I would like to thank for welcoming me into his lab group and for making me feel like a genuine member of the team. During that visit, I made many new friends, deepened my understanding of fisheries science, and of course, it was then that I was introduced to my co-author Dr. Jason Cope. Jason, I would like to thank you for all the support, guidance, and mentoring you have offered me over the years. You heartily embraced my ideas and with your constructive feedback, my work reached a level that I am very proud of. I am grateful for all the help you have offered during those challenging moments struggling with Stock Synthesis (ha!) and for the data-limited stock assessment knowledge you have so generously given me. I would also like to thank my co-authors Dr. Tommi Perälä and Asst. Prof. Jarno Vanhatalo for supporting my idea for correcting experts' biased opinions and for developing a model that I never could have created alone.

Many thanks to the members of my advisory committee, Dr. Jaakko Erkinaro, Adj. Prof. Päivi Haapasaari, and Dr. Samu Mäntyniemi for their constructive comments at different stages of my research. Special thanks to Samu for his contribution during the early stages of my studies and for helping me write the research proposal. In addition, special thanks to my good friends Maisa Nevalainen, Kelsey La Mere and Dunja Jusufovski for being in this journey together, helping each other in the highs and lows of this education. Also thanks to the friends and colleagues from the FEM group and its affiliates for the enjoyable discussions during meetings, lunches, and other social events. Many thanks to former DENVI coordinator, Dr. Anni Tonteri, and current coordinator, Dr. Karen Sims-Huopaniemi, for their help and guidance through the occasional complicated administrative issues and to the YEB doctoral school for funding my travel abroad on many occasions. A warm thank you to Prof. Sakari Kuikka for his assistance in finalizing this thesis and helping me to navigate the many steps required to make it to graduation and to Kelsey, for brilliantly editing the thesis summary.

Last but not least, I would like to thank my family for supporting me all these years and for encouraging me to pursue my dream of becoming the next Jacques Cousteau (I didn’t, but close enough, haha). Without them I would not be who I am nor where I am today. To my husband Olli, thank you for being there for me. My life with you has shaped my personality and values and I am grateful to share it with you and our wonderful little girl, Emma.

9. References

Allendrof F.W., England P.R., Luikart G., Ritchie P.A and Ryman N. (2008). Genetic effects of harvest on wild animal populations. Trends in Ecology and Evolution 23 (6): 327-337

Anderson S.C., Branch T.A., Ricard D., Lotze H.K. (2012). Assessing global marine fishery status with a revised dynamic catch-based method and stock assessment reference points. ICES Journal of Marine Science 69: 1491-1500

Anderson S. C., Cooper A. B., Jensen O. P., Minto C., Thorson J. T., Walsh J. C., … Osio, G. C. (2017). Improving estimates of population status and trend with superensemble models. Fish and Fisheries: 171, 732–741

Begley J., and Howell D. (2004). An overview of Gadget, the Globally applicable Area-Disaggregated General Ecosystem Toolbox. ICES Conference and Meeting Documents 2004/FF: 13. http://www. hafro.is/gadget/

Bentley N. (2014). Data and time poverty in fisheries estimation: potential approaches and solutions. ICES Journal of Marince Science, doi:10.1093/icesjms/fsu023

Berkson J. and Thorson J. T. (2015). The determination of data-poor catch limits in the United States: is there a better way? ICES Journal of Marine Science 72(1): 237-241

Beverrton R. J. H. and Holt S.J. (1957). On the Dynamics of Exploited Fish Populations. Gt. Britain, Fishery Invest., Ser. II, Vol. XIX. 533 pp.

Bharucha, Z. and Pretty, J. (2012). The roles and values of wild foods in agricultural systems. Phil Trans R Soc B 365: 2913–2926

Boyer D.C., Kirchner C.H., McAllister M.K., Staby A., Staalesen B.I. (2001). The orange roughy fishery of Namibia: Lessons to be learned about managing a developing fishery. South African Journal of Marince Science 23: 205-221

Bull B., Francis R. I. C. C., Dunn A. McKenzie A., Gilbert D. J. and Smith M. H. (2005), CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2.07-2005/08/21, 2005 NIWA Technical Report 127

Burgman M., Carr A., Godden L., Gregory R., McBride M., Flander L., Maguire L. (2011). Redefining expertise and improving ecological judgment. Conservation Letters 4: 81-87.

Burney D.A., Flannery T.F. (2005), Fifty millennia of catastrophic extinctions after human contact. Trends in Ecology and Evolution 20: 395-401

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

Carruthers T.R., Punt A., Walters C., MacCall A., McAllister M., Dick E., Cope J. (2014). Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research 153: 48-68

Carruthers T. R. and Hordyk A.R. (2018). The Data-Limited Methods Toolkit (DLMtool): An R package for informing management of data-limited populations. Methods in Ecology and Evolution: 9:2388-2395

Chang W., Cheng J., Allaire J.J., Xie Y., McPherson J. (2015). Shiny: Web Application Framework for R. R package version 0.11.1., <URL: http://CRAN.R-project.org/package=shiny

Cope J.M. and Punt A.E. (2009). Length-based reference points for data-limited situations: Applications and restrictions. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science 1:169-186

Cope J. (2013). Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fisheries Research 142: 3-14

Cope M.J., Thorson J.T., Wetzell C.R., DeVore J. (2015). Evaluating a prior on relative stock status using simplified age-structured models. Fisheries Research 171: 101-109.

Costello C., Ovando D., Hilborn R., Gaines S. D., Deschenes O., Lester S.E. (2012). Status and solutions for the world's unassessed fisheries. Science 338: 517-520

Cushing D.H. (1988). The study of stock and recruitment. Gulland J.A. Fish Population Dynamics. Wiley, New York. pp.105-128

Degnbol P., Gislason H., Hanna S., Jentoft S., Nielsen J.R., Svendrup-Jensen S., Wilson D.C.(2006). Painting the floor with a hammer: Technical fixes in fisheries management. Marine policy 30: 534-543

Deriso, R.B., Quinn T.J., Neal P.R. (1985). Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. 42, 815–824

Dichmont C.M., Deng R., Punt A.E., Brodziak J., Chang Y.J., Cope J.M., Ianelli J.N., Legault C.M., Methot R.D., Porch C.E., Prager M.H., Shertzer K. (2016). A review of stock assessment packages in the United States. Fisheries Research 183: 447-460 (2016)

Dick E.J and MacCall A.D. (2011). Depletion-Based Stock reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research 110: 331-341

Dowling N.A., Dichmont C.M., Haddon M., Smith D.C., Smith A.D.M., Sainsbury K. (2015). Empirical harvest strategies for data-poor fisheries: A review of the literature. Fisheries Research, in press.

EC (2011) Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the Common Fisheries Policy. EC, Brussels, COM/2011/0425 final - 2011/0195 (COD). Available at: http://eur- lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0425:FIN:EN:PDF

Erskine W., Ximenes A., Glazebrook D., da Costa M., Lopes M., Spyckerelle L., Williams R. and Nesbitt H. (2015). Food Security 7 (1): 55-65

FAO (1995). Code of Conduct for responsible Fisheries. FAO, Rome, Italy. Available at: http://www.fao.org/docrep/005/v9878e/v9878e00.HTM

Fournier D. A., 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 by using data for southern bluefin tuna (Thunnus maccoyii). Can. J. Fish. Aquat. Sci. 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: 2105–2116

Fournier D.A., Skaug H.J., Ancheta J., Ianelli J., Magnusson A., Maunder M.N., Nielsen A. and Sibert J. (2012) AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models, Optimization Methods and Software, 27:2, 233-249, DOI: 10.1080/10556788.2011.597854

French, S. (1980). “Updating of Belief in the Light of Someone Else’s Opinion.” Journal of the Royal Statistical Society. Series A (General), 143(1): 43–48.

Froese R, Kesner-Reyes K (2002) Impact of fishing on the abundance of marine species. ICES Document CM 2002/L:12, 15 pp

Froese R. (2004). Keep it simple: three indicators to deal with overfishing. Fish and Fisheries 5: 86-91

Froese R. and Pauly D. Editors. (2016). FishBase. Wolrd Wide Web electronic publication. www.fishbase.org (01/2016)

Froese R., Demirel N., Coro G., Kleisner K.M., Winker H. (2017). Estimating fisheries reference points from catch and resilience. Fish and Fisheries 18: 506-526

Froese, R., Winker H., Coro G., Demirel N., Tsikliras A. C., Dimarchopoulou D., Scarcella G., Probst W. N., Dureuil M., and Pauly D. A. (2018). A new approach for estimating stock status from length frequency data. ICES Journal of Marine Science 75: 2004–2015

Garthwaite P.H., Kadane J.B., O’Hagan A. (2005). Statistical methods for eliciting probability distributions. Journal of the American Statistical Association 100: 680-701

Gavaris S. (1988). An adaptive framework for the estimation of population size. Can.Alt. Fish. Sci. Adv. Comm. Res. Doc. 88/29

González-Laxe F. (2005). The precautionary principle in fisheries management. Marine Policy 29: 495-505

Graham M. (1935). Modern theory of exploiting a fishery, an application to North Sea trawling. Journal de Conseil International pour L’Exploration de la Mer 10: 264-274

Gulland, J.A. (1965). Estimation of mortality rates. Annex to Arctic Fisheries Working Group Report (meeting in Hamburg, January 1965). ICES, C.M.1965, Doc. No. 3 (mimeographed).

Gullet, W. (2000). The precautionary principle in Australia: policy, law and potential precautionary EIAs. Risk, Health, Safety and Environment 11: 93–124.

Hilborn R. (1992) Current and future trends in fisheries stock assessment and management, South African Journal of Marine Science, 12:1, 975-988, DOI:10.2989/02577619209504756

Hilborn, R., M. Maunder, A. Parma, B. Ernst, J. Payne, and P. Starr. (2001a). Coleraine: A generalized age-structured stock assessment model. User’s manual 2.0. Fish. Res. Inst. Univ. Wash. Rep. 0116.

Hilborn R., Maguire J.J., Parma A.M., Rosenberg A.A. (2001b). The precautionary approach and risk management: Can they increase the probability of successes in fishery management? Canadian Journal of Fisheries and Aquatic Sciences. 58: 99-107. DOI: 10.1139/cjfas-58-1-99

Hilborn R., Maunder M. N., Parma A., Ernst B., Payne J., and Starr P. J. (2000). Documentation for a general age-structured Bayesian stock assessment model: code named Coleraine. Fisheries Research Institute, University of Washington, FRI/UW 00/01. Software available at http://fish.washington.edu/research/coleraine/

Hilborn R, Ovando D. (2014) Reflections on the success of traditional fisheries management Ices Journal of Marine Science. 71: 1040-1046. DOI: 10.1093/icesjms/fsu034

Hjort J., Jahn G., Ottestad P. (1933). The optimum catch. Hvalradets Skrifter 7: 92-127

Hordyk A., Ono K., Valencia S., Loneragan N., Prince J. (2014). A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. ICES Journal of Marine Science, doi:10.1093/icesjms/fsu004

ICES (2015). Report of the Fifth Workshop on the Development of Quantitative Assessment Methodologies based on Life-history Traits, Exploitation Characteristics and other Relevant Parameters for Data-limited Stocks (WKLIFE V), 5–9 October 2015,Lisbon, Portugal. ICES CM 2015/ACOM:56, 157 pp

Ives M. C. and Scandol J.P. (2013). BIOMAS: A bio-economic modelling and assessment system for fisheries management strategy evaluation. Ecological Modelling 249: 42-49

Jennings S., Kaiser M.J., Reynolds J.D. (2001). Marine fisheries ecology. Blackweel Science Ltd

Jones R. (1984). Assessing the effects of changes in exploitation pattern using length composition data (with notes on VPA and cohort analysis). FAO. Fish. tech. Pap; 256: x + 118 pp. ams

Johnson S.R., Tomlison G.A., Hawker G.A., Granton J.T., Grosbein H.A., Feldman B.M. (2010). A valid and reliable belief elicitation method for Bayesian priors. Journal of Clinical Epidemiology 63: 370- 383.

Kimura, D., Tagart, J. (1982). Stock reduction analysis, another solution to the catch equations. Canadian Journal of Fisheries & Aquatic Sciences 39, 1467–1472.

Kimura, D.K., Balsinger, J.W., and Ito, D.H. 1984. Generalized stock reduction analysis. Canadian Journal of Fisheries & Aquatic Sciences **41**(9): 1325–1333. doi:10.1139/f84-162.

Kokkalis A., Thygesen U.H., Nielsen A., Andersen K.H. (2014). Limits to the reliability of size-based fishing status estimation for data-poor stocks. Fisheries Research, doi: 10.1016/j.fishres.2014.10.007

Kristensen K., Nielsen A., Berg C.W., Skaug H. and Bell B. (2016). TMB: Automatic differentiation and laplace approximation. Journal of Statistical Software 70 (5)

Kvamsdal S. F., Eide A., Ekerhovd N. et al. (2016). Harvest control rules in modern fisheries management. Elementa: Science of the Anthropocene 4: 000114. doi: 10.12952/journal.elementa.000114

Kuparinen A., Mäantyniemi S., Hutchings J.A., Kuikka S. (2012). Increasing biological realism of fisheries stock assessment: towards hierarchical Bayesian methods. Environmental Reviews 20: 135-151

Kynn M. (2008). The ‘heuristics and biases’ bias in expert elicitation. Journal of the Royal Statistical Society A 171(1): 239-264.

Lindley D.V. (1982). The improvement of probability judgments. Journal of Royal Statistical Society 145: 117-126

Link JS. (2010). Ecosystem-Based Fisheries Management: Confronting Tradeoffs. New York: Cambridge University Press.

Leslie P.H. (1945). On the use of matrices in certain population mathematics. Biometrika 33(3): 183-212

Leslie P.H. (1948). Some further notes on the use of matrices in certain population mathematics Biometrika, 53 (1948), pp. 213-245

Lewis E.G. (1942). On the generation and growth of a population. Sankhya 6: 93-96

Lichtenstein S., Fuschhoff B., Phillips L. (1982). Judgment under uncertainty: heuristics and biases. eds Kahneman D., Slovic P., Tversky A. (Cambridge Univ Press, New York), pp 306–334.

Lotka A. J. (1907). Relation between birth rates and death rates. Science N.S. 26: 21-22

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

Magnusson, A., and Hilborn, R. 2007. What makes fisheries data informative? Fish and Fisheries 8: 337–358.

Martel S. and Froese R. (2013). A simple method for estimating MSY from catch and resilience. Fish and Fisheries 14: 504-514

Methot, R.D., Jr., and Wetzel, C.R. (2013). Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res.142: 86–89.

Montgomery D.C., Peck E.A., Vining G.G. (2012). Introduction to Linear Regression Analysis. Wiley, Chichester, UK.

Morgan M.G. (2013). Use (and abuse) of expert elicitation in support of decision making for public policy. PNAS 111 (20): 7176-7184.

Mäntyniemi S., Haapasaari P., Kuikka S., Parmanne R., Lehtiniemi M., Kaitaranta J. (2013). Incorporating stakeholders' knowledge to stock assessment: Central Baltic herring. Canadian Journal of Fisheries & Aquatic Sciences **70**: 591–599.

Nakatsuka S. (2017). Management strategy evaluation in regional fisheries management organizations − How to promote robust fisheries management in international settings. Fisheries Research 187: 127-138

Newman, K., Buckland, S.T., Morgan, B., King, R., Borchers, D.L., Cole, D., Besbeas, P., Gimenez, O., Thomas, L. (2015). Modelling Population Dynamics. New York, NY. Springer-Verlag.

Niels A and Berg C.W (2014). Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research 158: 96-101

NMFS. Magnuson-Stevens Fishery Conservation and Management Act. Department Of Commerce, U. S. O. A. (2007).

Norse E.A., Brooke S., Cheung W., Clark M.R., Ekeland I., Froese R., Gjerde K.M., Haedrich R.L., Heppel S.S., Morato T., Morgan L.E., Pauly D., Sumaila R., Watson R. (2012). Sustainability of deep-sea fisheries. Marine Policy 36:307-320

O'Hagan A., Buck C.A., Daneshkhah A., Richard Eiser J., Garthwaite P.H., Jenkinson D.J., Oakley J.E., Rakow T. (2006). Uncertain Judgments: Eliciting experts' probabilities. Wiley, Chichester, UK.

Pope J.G. (1972). An investigation of the accuracy of Virtual Population Analysis using cohort analysis. ICNAF Res. Bull. 9: 65–74.

Punt A.E., Smith D.C., Smith A.D.M. (2011). Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES Journal of Marine Science 68: 972-981

Punt A. E., Butterworth D. S., de Moor C. L., De Oliveira J. A. A., & Haddon M. (2016). Management strategy evaluation: Best practices. Fish and Fisheries 17: 303–334.

Roman HA, et al. (2008) Expert judgment assessment of the mortality impact of changes in ambient fine particulate matter in the U.S. Environmental Science &Technology 42(7):2268–2274.

Rosenberg A. A., Fogarty M. J., and Cooper A. B., et al. (2014) Developing new approaches to global stock status assessment and fishery production potential of the seas, (Vol. 1086). Fisheries and Aquaculture Circular No. 1086. Rome: FAO

Rosenberg A. A., Kleisner K. M., Afflerbach J., Anderson S. C., Dickey-Collas M., Cooper A. B., ... Ye, Y. (2018). Applying a new ensemble approach to estimating stock status of marine fisheries around the world: Estimating global fisheries status. Conservation Letters, 11(1), [e12363]. DOI: 1 0.1111/conl.12363

Russel E.S. (1931). Some theoretical considerations on the overfishing problem. Journal de Conseil International pour L'Exploration de la Mer 6: 3-20

Salas S., Chuenpagdee R., Seij J., Charles A. (2007). Challenges in the assessment and management of small-scale fisheries in Latin America and Caribbean. Fisheries research 87: 5-16

Shepherd J.G. (1999). Extended survivors analysis: An improved method for the analysis of catch-at-age data and abundance indices. ICES Journal of Marine Science, 56: 584–591

Schaefer M. B. (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bull. Inter-Am. trap. Tuna Commn 1: 27-56.

Schnute J. (1985). A general theory for analysis of catch and effort data. Can. J. Fish. aquat. Sci. 42(3): 414-429.

Schnute J. (1987). A general fishery model for size-structured fish population. Can.J. Fish. Aquat.Sci. 44: 924-940

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

Stergiou K.I., Moutopoulos D.K., Tsikliras A.C., Papaconstantinou C. (2007). Hellenic marine fisheries: A general perspective from the national statistical service data. Current state of the fisheries sector.

UNCLOS (1982) United Nations Convention on the Law of the Sea. 1883 UNTS 3. Available at: http://www.un.org/depts/los/convention_agreements/texts/unclos/UNCLOS-TOC.htm

Teh L.C.L and Sumaila U.R. (2013). Contributions of marine fisheries to worldwide employment. Fish and Fisheries 14 (1): 77-88

Tversky, A. and Kahneman, D. (1974). "Judgment under uncertainty: Heuristics and Biases." Science, 185(4157): 1124–1131

Uusitalo L., Kuikka S., Romakkaniemi A. (2005). Estimation of Atlantic salmon smolt carrying capacity of rivers using expert knowledge. ICES Journal of Marine Science 62: 708-722.

Vehtari, A. and Ojanen, J. (2012). "A survey of Bayesian predictive methods for model assessment, selection and comparison." Statistics Surveys, 6: 141–228

Voinov A., Bousquet F. (2010). Modelling with stakeholders. Environmental Modelling & Software 25: 1268-1281

Walsh J.C., Minto C., Jardim E., Anderson S.C., Jensen O.P., Afflerback J.....Cooper A. B. (2018). Trade-offs for data-limited fisheries when using harvest strategies based on catch-only models. Fish and Fisheries 19:1130-1146

Wetzel C.R. and Punt A.E. (2011). Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fisheries Research 110: 342-355

Wetzel C.R. and Punt A.E. (2015). Evaluating the performance of data-moderate and catch-only assessment methods for U.S. west coast groundfish. Fisheries Research, 171: 170-187. http://dx.doi.org/10.1016/j.fishres.2015.06.005

Wiedenmann J., Wilberg M.J., Miller T.J. (2013). An evaluation of harvest control rules for data-poor fisheries. North American Journal of Fisheries Management 33: 845-860

Williams, C. K. and Rasmussen, C. E. (2006). Gaussian processes for machine learning. MIT Press.

Worm B., Barbier E.B., Beaumont N., Duffy J.E., Folke C. et al. (2006). Impacts of Biodiversity Loss on Ocean Ecosystem Services. Science 314: 787–790

Worm B. and Branch T. (2012). The future of fish. Trends in Ecology and Evolution 27: 594-599

Zhou, S., Punt, A. E., Ye, Y., Ellis, N., Dichmont, C. M., Haddon, M., Smith, D. C. et al. (2017). Estimating stock depletion level from patterns of catch history. Fish and Fisheries, xx: 1–10

Zhou S., Punt A.E., Smith A.D.M., Ye Y., Haddon M., Dichmont C.M., Smith D.C. (2018). An optimized catch-only assessment method for data poor fisheries. ICES Journal of Marine Science 73 (3): 964-976

Zickfeld K, Morgan MG, Frame DJ, Keith DW (2010) Expert judgments about transient climate response to alternative future trajectories of radiative forcing. PNAS 107(28): 12451– 12456.