![UiA University of Agder logo]

Fishing for more data: Exploratory stock assessment of the data-limited brown crab (*Cancer pagurus*) stock in Norway

JOHANNA BJÅNES MARCUSSEN

SUPERVISOR

Fabian Zimmermann
Ane Timenes Laugen
Guldborg Søvik

Summary

Formal stock assessment is a key step in managing fisheries sustainably and preventing overexploitation (Hilborn, 2011). However, the majority of the world's exploited stocks are considered data-limited and lack formal stock assessments (Ovando et al., 2021). Several attempts to find simple, low-cost, and generic solutions for the assessment of data-limited stocks have been made, still, individual adjustments are required to provide useful information for management purposes (Dowling et al., 2019).

To date, there has not been any systematic quantification of stock size of brown crab (*Cancer pagurus*) in Norway. The present study has therefore evaluated the applicability of data on brown crab in Norway in describing the stock development qualitatively or quantitatively, providing the potential basis for a formal assessment. Data from the reference fishery (Woll et al. 2006), landings data and fisheries-independent data were standardized using statistical models to derive indices that provide indicators of stock development. The indices also serve as inputs for stock assessment models. The available data were applied to stock assessment designed for data-limited stocks and their applicability was evaluated. The following methods were tested: (i) length-based indicators, (ii) length-based spawning potential ratio (LBSPR), and (iii) Surplus production in continuous time (SPiCT).

Few major spatial differences among the trends in length structures or abundance were found. Overall, the available data and the tested methods indicate stable trends, suggesting that the brown crab fishery has been sustainable. Vestlandet emerges as an exception with indications of overfishing, as the assessment unit does not meet the length-specific targets and the relative abundance has been abruptly declining since 2015. Despite relatively long time series, the applicability of the data in LBSPR and SPiCT is strictly limited. This highlights that data-limited methods have specific data demands, both quantitatively and qualitatively, that can make them unsuitable for specific cases, such as non-conventional shellfish fisheries. Furthermore, each method comes with specific assumptions that need to be fulfilled and often require sufficient contrast in time series, constraining their utility for stocks with stable trends. This underlines the challenges with the application of generic stock assessment methods, and a need for easily adjustable methods for shellfish and particularly decapods.

Sammendrag

Bestandsvurderinger er et viktig steg i retning bærekraftig fiske og for å unngå overfiske (Hilborn, 2011), likevel er hovedandelen av verdens fiskebestander regnet som data-fattige og ikke formelt vurdert (Ovando et al., 2021). Dette gjelder også den norske taskekrabbebestanden, til tross for at det er en kommersielt viktig art. Hovedutfordringen ved bestandsvurderinger er ofte begrenset og varierende datakvalitet, samt generiske metoder som ikke passer til alle arter eller bestander (Dowling et al., 2019). Derfor må hver enkelt bestand vurderes individuelt.

Denne studien har analysert et utvalg av tidsserier på taskekrabbe i Norge og evaluert deres egnethet til å beskrive endringer i bestanden og anvendelighet i formelle bestandsvurderingsmetoder. Siden utbredelsen av norske krabbebestanden strekker seg over ti breddegrader var det forventet romlig variasjon i bestanden og fiskepresset, derfor ble bestanden delt i fem områder. For å lage indekser som er representative for trender i bestanden, ble fangst- og lengdedata fra referansefiske (Woll et al. 2006) og landingsdata fra kommersielle fiskere standardisert ved hjelp av statistiske modeller. I tillegg fungerer indeksene som hovedinput i bestandsvurderings metoder. I et mylder av metoder som er tilpasset data-fattige bestander ble følgende metoder valgt: (i) length-based indicators, (ii) length-based spawning potential ratio (LBSPR), og (iii) surplus production in continuous time (SPiCT).

Indeksene indikerte stabile trender og en god tilstand i de fleste bestandene, noe som antyder at dagens fiske er bærekraftig. Vestlandet skilte seg ut fra de andre områdene ved at det er indikatorer på overfiske. Der ble det observert en klar nedgang i bestanden fra 2015 og området innfridde ikke de lengdebaserte målene. Selv om det eksisterer lange tidsserier var anvendeligheten av dataene i de to mest avanserte metodene, LBSPR og SPiCT, svært begrenset. Dette poengterer at også metoder tilpasset data-fattige bestander har svært spesifikke datakrav. I tillegg presterer de valgte metodene dårligere hvis det ikke er kontrast i dataene, noe som ytterligere begrenser bruken, særlig for bestander med stabile trender. Oppsummert understreker dette utfordringene ved å anvende generiske metoder, og behovet for metoder som enkelt kan tilpasses til skalldyr.

Table of content

Preface

1 Introduction

1.1 Stock assessment and sustainable fisheries management

Overfishing is a leading threat to marine ecosystems and sustainable development (Lotze et al., 2006; Worm et al., 2009). Sustainable and effective management of exploited stocks requires biological and catch information to assess the abundance or biomass of the targeted stock, as well as the fishing mortality rates. To inform fisheries management and enable harvest strategies that maintain profitable fleets without jeopardizing sustainability, several quantitative assessment methods have been developed. However, many of them require substantial data on the fishery as well as fisheries-independent data (Kelly and Codling, 2006; Bentley, 2015). This poses a major challenge, as the majority of worldwide stocks are data-limited.

Although data limitations are often associated with developing countries, there are still many commercially harvested fish stocks without a stock assessment in Europe, including in Norway. Because data-limited stocks are often in worse conditions than data-rich stocks, they are also in particular need of proper management and assessments (Costello et al., 2012; Hilborn et al., 2020). Unassessed stocks are often characterized by being mainly exploited as by-catch, or in small-scale or recreational fisheries, and are often of low economic values (Bentley and Stokes, 2009). Stocks with life history or ecology that deviates from a typical fish stock are also less often assessed than others, which is a particular problem for many shellfish species where e.g. aging is difficult or impossible (Smith and Addison, 2003; Sheridan et al., 2015). Proper assessment could also be hindered by low-quality data or the lack of resources (Dowling et al., 2019).

Several jurisdictions have developed systems to categorize stocks based on data available to reduce the number of unassessed stocks (Dichmont et al., 2016). The International Council for the Exploration of the Sea (ICES) categorizes stocks based on the available information, defining six categories (ICES, 2012): category 1 refers to the most data-rich stocks with full analytical stock assessments and forecasts; category 2 differs from the former by analytical assessments only being treated qualitatively as indicative of trends;

category 3 refers to stocks in which one or more available indices are indicative of trends; the last three categories (4,5, and 6) comprises stocks where the only available data are catch and/or landings data. Category 3-6 are considered data-limited and comprise half of the world's commercial stocks (Dichmont et al., 2016; Hilborn et al., 2020). In the myriad of assessment methods, ICES proposes a framework for the assessment of each of the six categories.

1.2 Catch data in assessments – state and trends of the population

There is a myriad of data-limited assessment methods, but most can broadly be categorized into length-based methods and surplus production models. While length-based methos require some length distribution information from the catches or surveys, surplus production models require time series of commercial catch data and indices are as inputs, typically scaling total biomass to catches. Using catch directly as an indicator of changes in population abundance has been attempted but may mask changes in the abundance index caused by other factors than actual changes in the abundance. The catchability of crustaceans is affected by different environmental factors, like which habitat (Addison and Lovewell, 1991), depth (Linnane et al., 2013), and area (Woll et al., 2006) is targeted, as well as when (Brown and Bennett, 1980; Woll et al., 2006). Additionally, fishery-specific factors can alter the catchability, such as gear type (Addison and Lovewell, 1991), bait (Chapman and Smith, 1978; Skajaa et al., 1998), and soak time (Bennett and Brown, 1979). Ideally, catch-per-unit-effort (CPUE) indices are based on highly resolved fisheries data that includes complete information on effort and fishing technology to account for technological creep or other changes in catch efficiency not linked to stock abundance (Honma, 1974; Beverton and Holt, 2012).

Fisheries data is often less than ideal, with varying reporting quality, inconsistencies in time, and missing information on relevant variables due to limited reporting requirements (Hilborn and Walters, 1992). This applies particularly to smaller-scale fisheries where reporting requirements and capacities are typically limited. Norwegian fisheries data illustrates this problem, as boats below 15 m were until mid-2022 only required to report the landings per boat trip without explicit information on fishing effort,

number of gear, etc. Such landings data is often considered to be too coarsely resolved to be indicative of trends in stocks. However, the characteristics of coastal shellfish fisheries may provide an exception, as effort (number of pots) is relatively constant per vessel and one fishing trip may provide an acceptable approximation to one haul (setting and emptying all pots once per landing). For the Norwegian pot fishery for *Nephrops*, it has been shown that landings per boat trip may provide an adequate proxy for catch rates (Zimmermann et al., 2022). Consequently, in addition to estimating a CPUE index based on data from the reference fishery, a landing-per-unit-effort (LPUE) index based on landing data was produced to explore the utility of the approach for brown crab (*Cancer pagurus,* Linnaeus 1758) and cross-validate the two indices. "Cross-validation" is used here to describe the comparison of indices or model estimates from two independent data sets, a recommended quality check to evaluate results in stock assessment (Kell et al., 2021).

1.3 Size data– insight in population dynamics

Fishing mortality not only affects the abundance of the stock but also changes its size and age distribution. Increased mortality through fishing truncates the demographic structure of a population, an effect that is compounded by fishing selectivity due to fishing gear used, fishing strategies, fish behaviour, or regulation. Selectivity in a fishery is unavoidable, and inevitably affects the compositions at both population and community level (Garcia et al., 2012). Selectivity can occur when fishers target specific sites or favourable traits possessed by the individuals to maximise profit, such as sex or size (Zhou et al., 2010; Zimmermann and Heino, 2013). To reduce fishing impacts, certain restrictions directly or indirectly related to body size have commonly been used: size limits (Halliday and Pinhorn, 2002), effort relating to gear technology and quantity (Kennelly and Broadhurst, 2002), as well as temporal and spatial restrictions (Dunn et al., 2011). Minimum landing size (MLS) is an example of a size-selective measure aimed at avoiding overexploitation by protecting the immature portion of the stock.

Size is linked to important population dynamics parameters. For instance, there is a positive size-fecundity relationship for brown crabs, meaning larger female crabs produce relatively more eggs than smaller crabs (Tallack, 2007). For maintaining reproductive potential and resilience to maladaptation, it is essential to maintain a

diverse length structure in a population (Walsh et al., 2006; Moland et al., 2010). Fishing has always demographics effects that tend to remove the oldest, largest individuals disproportionally, and when the fishing pressure is high the mean size is assumed to decrease (Trenkel et al., 2007). Size-selective fishing mortality which targets large individuals further truncates the size distribution (Berkeley et al., 2004). Changes in size structure and associated life-history parameters can therefore inform about the status of a population and changes in mortality it experiences. Analysis of length structures could therefore be indicative of trends in abundance and is an important component in stock assessment (Rochet and Trenkel, 2003).

1.4 Where do we draw the lines – defining stock

Since all stock assessments require adjustments to the targeted stock (Dowling et al., 2019), knowledge of the population dynamics and the ecology of the targeted species and the specific stock, as well as the fishery and available data, is essential. Drawing stock boundaries — which ideally should be based on biological criteria — is found to be challenging. Begg et al. (1999) propose two ways of separating a stock. One way of setting the boundaries could be to place them where there is low(er) genetic exchange between populations of a species. The second way of differentiating between stocks is by phenotypic variation, caused by prolonged separation in different environmental regimes.

By the first definition, the whole Northeast Atlantic brown crab stock may be viewed as one stock, as there is little genetic variation across areas (Ungfors et al., 2009). The low genetic variation could be caused by high connectivity, due to the prolonged planktonic larval stage (dispersal phase) (Palumbi, 2003). Length at sexual maturation could be a trait to distinguish stocks in the Northeast Atlantic, based on the second definition, as it varies throughout its distributional range, from 100 to 133 mm carapace width (Haig et al., 2016, and references within). Despite the wide geographical range in Norway, the length at sexual maturation (112 mm) is the same along the whole coast, while the age when the crabs reach that length increases with latitude (Bakke et al., 2018).

Even though the crab stock appears to be uniform by the first definition, stocks are managed differently throughout the species' distributional range in the Northeast Atlantic. Therefore, a third approach is often applied, separating stocks into management units due to historic reasons, fisheries dynamics, management practicalities, and national jurisdictions.

1.5 Brown crab fishery in Norway

Brown crab is commercially exploited throughout Western Europe, from France to Norway, with the biggest catches in Scotland. The brown crab fishery is one of the most important shellfish fisheries in Europe, both in terms of tonnage and value. In the last ten years, the total annual catches in Europe have been stable at 40,000-50,000 tonnes (FAO, 2019). The Norwegian crab fishery began in the early 1900s and peaked for the first time in the 1940-50s with catches of 7,000 to 9,000 tonnes annually. Annual commercial landings have been stable at approximately 5,000 tonnes after they peaked, a second time in 2007 (Figure 1). Today the fishery spans the coast from Skagerrak in the south to Lofoten in the north; the main fishing area is centered in mid-Norway, from Stad to Trøndelag. Additionally, there are substantial unreported catches, including from a large recreational fishery, crabs used as bait in the wrasse fishery, crabs caught as by-catch, and fishery-induced mortality caused by abandoned gear, known as ghost fishing. In 2021 more than 450 boats participated in the commercial fishery for brown crab (vessels with more than 100 kg landings reported). The fishery is concentrated from July till November (Woll et al., 2003), with the majority of the participants being multiple target fishers (i.e. fishing for more than one species). The fishing fleet is dominated by vessels smaller than 11 m with limited reporting duties, severely limiting data availability to sales slips from landings. Until 2022, only vessels larger than 15 m were obliged to use electronic logbooks. Logbook data, thus, exists only from a few larger vessels and is not representative.

Currently, the Norwegian crab fishery is open-access and largely unregulated except for a minimum landing size of 11 and 13 cm from the border to Sweden to Rogaland and north of Rogaland (~59°30'N) respectively. Additionally, mandatory escape gaps >80 mm (70 mm for commercial fishers) protect the smallest and immature crabs the whole year,

between the border to Sweden and Tysfjord in Nordland County (Høstingsforskriften §29). There is no ban on landing berried crabs (female crabs with external roes) or crabs with soft shells that have recently molted (ecdysis, the event of shell changing). However, these are rarely landed due to low economic value. Recreational fishers are restricted to 20 pots or fyke nets per fisher and vessel. Unlike the recreational European lobster fishery (*Hommarus gammarus,* Linnaeus 1758*)*, there is no obligation of registration in the recreational crab fishery.
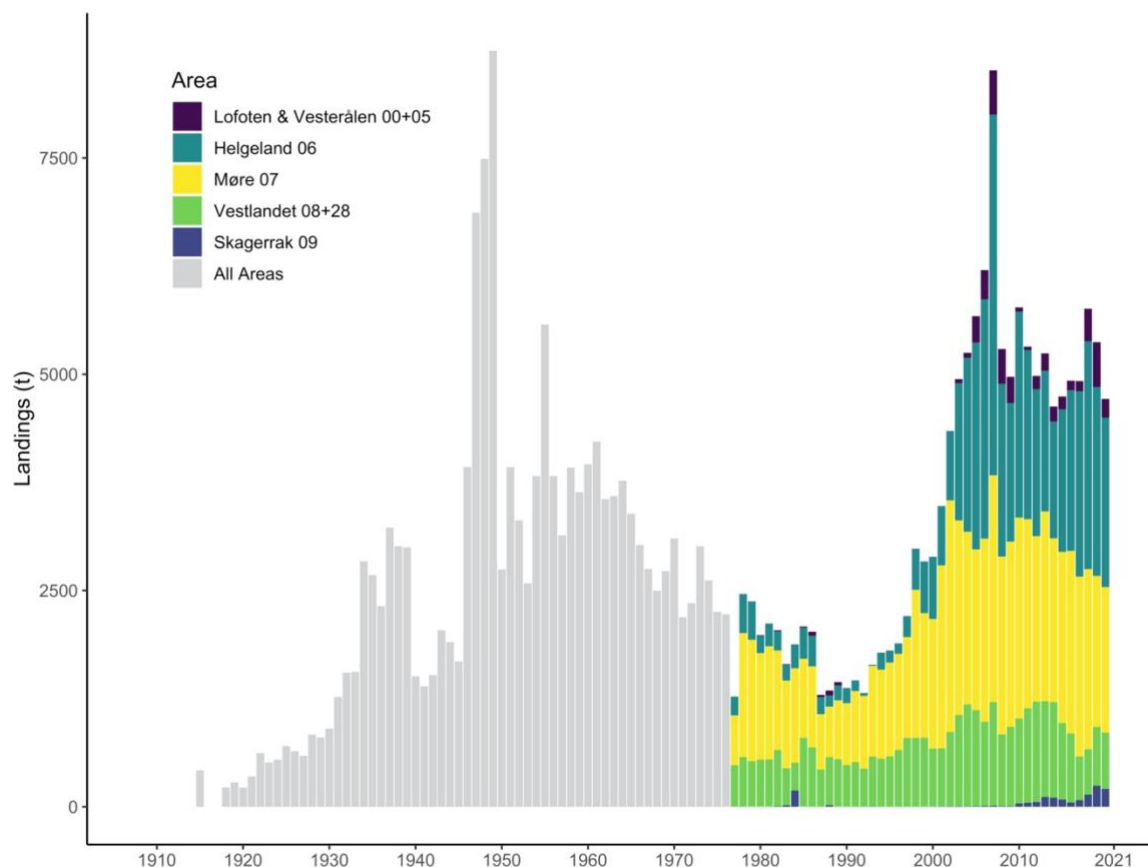
Figure 1 Annual brown crab (*Cancer pagurus*) landings, given by weight in tonnes, in the different areas since 1976 and aggregated by all areas from 1900. Data were retrieved from SSB.no (1900-1975) and from the Norwegian Directorate of Fishery (1976-2021).

1.6 Aims

To date, no formal assessment of the stock of brown crab in Norway has been conducted. Previous analyses and advice were mostly based on descriptive statistics, inferring trends and status from changes in mean catch rates and length distributions (Woll et al., 2006;

Søvik et al., 2017; Zimmermann et al., 2020). Consequently, we know little to nothing about the status of the stock in relation to the current fishing pressure. The main aim of this study was therefore to lift the brown crab out of the group of unassessed species, towards a knowledge-based fisheries management. A key step towards establishing an assessment framework is to evaluate all data available and explore the application of formal assessment methods to the currently available data. The thesis focussed on the main data sources available, including commercial landing, a 20-year time series of self-reported catch and length data from the reference fishery, and two surveys from the Institute of Marine Research (IMR), including one conducted as part and for the purpose of this study. To explore the state and trend in the crab stock in relation to fishery, spatially and temporally, I explored and evaluated the applicability of data from the two fishery-dependent time series, and cross-validated them against each other as well as against fisheries-independent data. Covering the distribution of the crab stock in Norway, the reference and landings time series are thought to be representable for the crab fishery along the Norwegian coast. The potting survey, covering one of the most important fishing areas, was conducted to evaluate the catch rates and composition data in the reference fishery. The gill- and fyke-net survey, on the other hand, was investigated as an existing time series that could potentially provide a fisheries-independent stock index.

Estimating a crab index using data from an existing survey (gill – and fyke net survey, by IMR) would be preferred to establishing a new independent potting survey, from both an economical perspective and since the survey time series is available back in time. A recent study has shown that existing survey data can be suitable to produce abundance indices for brown crab despite not being a target species (Mesquita et al., 2021). However, it was unclear whether the atypical design of the survey and gear types used were suitable to provide representative abundance information. Therefore, to evaluate whether the gill- and fyke net provides relevant information on the crab stock and cross-validate the time series, the potting survey parallel to the gill- and fyke net survey was conducted in August 2021, and subsequently analysed for the spatial patterns in catch rates between and within the two surveys.

CPUE and LPUE indices could, if standardized properly, be representative for relative abundance and indicative for trends in the population. To investigate trends in the stock at a national as well as regional scale, I estimated standardized CPUE and LPUE indices

based on the reference time series and the commercial landings time series. Furthermore, to evaluate the representativeness of the catch rates observed in the reference fishery, I compared them also against catch rates from the fishery-independent potting survey conducted in 2021 and, lastly, I compared the reference time series to the landings time series to cross-validate stock indices estimated from both time series. These were both evaluated as stock indicators on their own as well as used as input for stock assessment. To assess and explore the applicability of quantitative assessment methods to the available data, I then applied three stock assessment methods to the stock and evaluated the sensitivity to key assumptions, parameters and input data. Combined, the thesis 1) provides an overview of the available data on brown crab in Norway and their suitability for assessment purposes, 2) explores whether key assessment methods could be applied and accepted given the available data and information, and 3) presents and discusses the status of the stock and observed trends as suggested by the different data sources and assessment methods.

2 Methods

2.1 Study species

The brown crab is distributed from the north coast of Africa in the south to northern Norway in the north (Ungfors, 2008), and is as versatile in habitat preference as in geographical distribution. They are most common in subtidal areas and deeper waters down to 100 meters but have been sighted down to 400-500 meters (d'Acoz, 1999; Neal and Wilson, 2008; Bakke et al., 2019). The species is long-lived with a life expectancy of about 10-20 years (Sheehy and Prior, 2005; Ungfors, 2008), and habitat preferences vary throughout the lifecycle. The knowledge of which habitats they prefer during their first years is sparse and limited, but juveniles are found in the littoral zone, and when they reach the age of maturity they migrate to deeper waters. Brown crabs reach functional maturity at about 112 mm, which corresponds to an age of maturation of about five years, depending on the location as age and length at maturity vary with latitude (Bakke et al., 2016). In the adult stages of the life cycle, they find shelter as they bury themselves in the sandy bottom. The theoretical measure of maximum length, the asymptotic length based on the von Bertalanffy growth model, varies from 199 mm to 280 mm, depending on method and area (Chapman, 1994; Sheehy and Prior, 2005; Klaoudatos et al., 2013). Estimating the age of decapods is challenging and time-consuming (Sheridan et al., 2015), and length is therefore a more applicable measurement when assessing stock structure in decapods. This partly limits the opportunities of applicable assessment methods to length-based or biomass methods.

2.2 Study area

The Norwegian crab stock is likely to constitute one stock based on the definition of phylogenetic similarity (Bakke et al., 2018). Nonetheless, the abundance and the fishing pressure are expected to vary throughout its distribution range, and by using a higher spatial resolution than the entire Norwegian coast, these patterns can be revealed. In addition, data availability varies substantially among areas, largely following the degree

of fishing activity. Focusing on data-richer areas with consistent time series may therefore provide better information and avoid potential biases introduced by data-poorer areas, as well as averaging-out effects when combining areas with possibly very different trends in fishing pressure. To explore the role of spatial structure when assessing the stock and to quantify the spatial variation, units needed to be defined.

The fishery reports spatial information linked to a statistical grid with predefined statistical areas and locations from the Norwegian Directorate of Fisheries. Using seven of the statistical areas as baselines, five assessment units were defined (Figure 2): (i) Skagerrak, (ii) Vestlandet, (iii) Møre, (iv) Helgeland, and (v) Lofoten & Vesterålen. The definitions of the assessment units are based on previous studies(Woll et al., 2006; Bakke et al., 2018), but slightly adjusted. Skagerrak corresponds to statistical area 09, corresponding to the Skagerrak and Kattegat. Vestlandet is defined as the statistical areas (08 and 28) situated west of the Skagerrak basin in the south and limited by Stad (62°N,5°E) in the north. The two areas are known to have smaller crabs in the catches than the two regions north of it (Woll et al., 2006) and were there for defined as a separate assessment unit. Møre and Helgeland are the region north of Stad but south of Lofoten & Vesterålen, corresponding to statistical areas 07 and 06, respectively. This vast region was kept as two separate assessment units. The data-poorer areas 00 and 05 were combined into the northernmost assessment unit Lofoten & Vesterålen, mainly because the data from these areas is too limited to allow for a finer spatial resolution. As it is on the border of the norther limit of the species distribution where the fishery has been more recent and dynamic, and to track the northward expansion of the species, the unit was kept separate from Helgeland

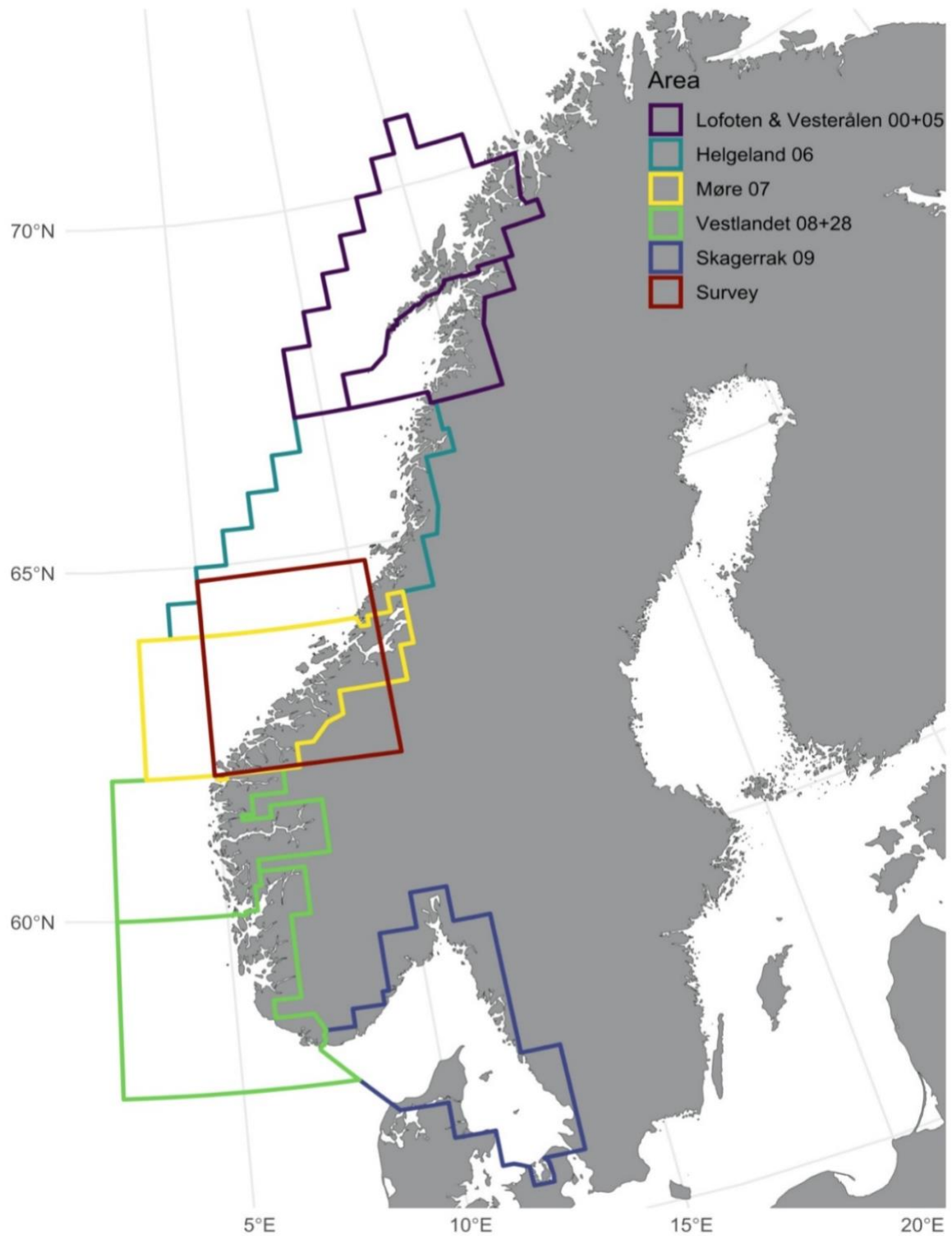**Figure 2** Map of Norway showing seven of the nine coastal statistical areas used by the Norwegian Directorate of Fisheries. Assessment units investigated in this study are based on the statistical areas: Skagerrak (blue, Area 09), Vestlandet (green, Area 08 and 28), Møre (yellow, Area 07), Helgeland (light-blue, Area 06), and Lofoten & Vesterålen (purple, Area 05 and 00). The red square marks the survey area in Figure 3

2.3 Data

2.3.1 Fishery-independent surveys

When assessing a stock both fishery-dependent and fishery-independent data are useful (Verdoit et al., 2003). IMR conducts an annual coastal gill- and fyke-net survey aimed at collecting information on demersal fish in coastal areas that are not covered by the coastal bottom trawl survey. Because substantial crab catches in both gear types have been observed, the question was raised whether the survey time series contains crab observations that are representative and consistent enough to estimate an annual index of crab abundance. To investigate the selectivity of the survey and cross-validate the data, an independent potting survey targeting brown crab was conducted in 2021 parallel to the gill- and fyke-net survey. To allow for a direct comparison with data reported by commercial reference fisheries and cross-validate data from target commercial fishing with data from random survey stations, areas where reference fishers are present were selected (Figure 3) and the same standardized crab pots as in the reference fishery used. In August 2021, the fishery-independent crab stock survey was conducted as a collaboration between the IMR and the Norwegian Directorate of Fisheries (for full sampling protocol see Appendix A).

A limitation of the gill- and fyke-net survey are measures put in place in 2016 to adapt sampling sites to avoid crab habitats and lift the gillnets above the seafloor, aimed at lowering the number of crabs in the catches. Surveys using passive gears can be useful tools for abundance estimates but only when suitable habitats of the relevant species are targeted (Haggarty and King, 2006). Therefore, two categories of stations were used in the potting survey: (i) stations located directly at the same place and time as a station of the gill- and fyke net survey; (ii) control stations randomly distributed in suitable crab habitat at bottom depths above 50 m in the vicinity of gill- and fyke- net stations (20 km radius). The comparison of station types, thus, provides an indication of how the gill- and fyke-net survey design to avoid crab habitats affects the observed catches compared to a randomized design.
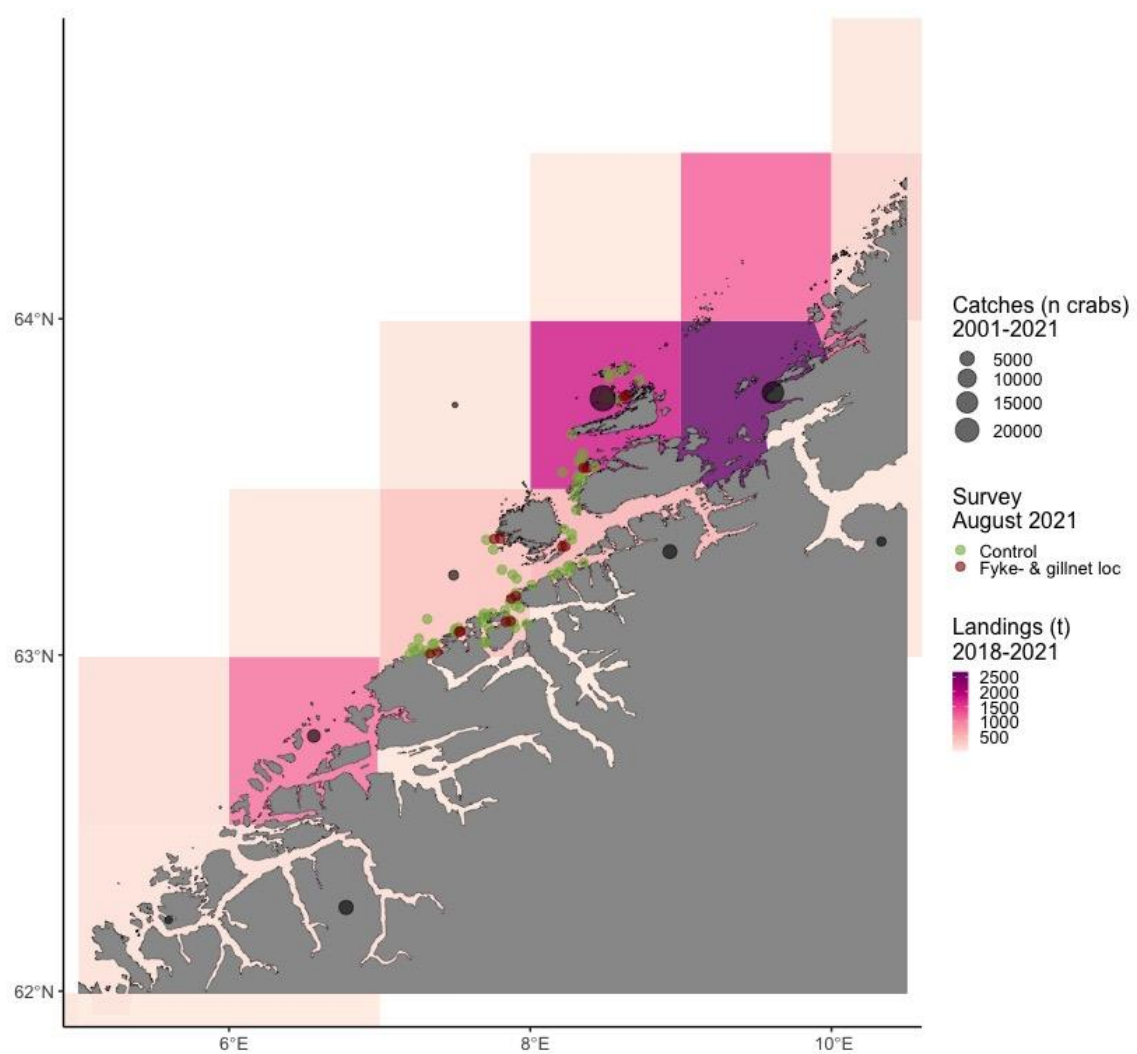
Figure 3 Close-up of the spatial distribution of the different data in Møre. The map includes landings from 2018-2021 and the number of crabs caught by the reference fishery throughout the time series (2001-2021), aggregated by each statistical location (centred in each statistical locations). Additionally, the two station types sampled during the potting survey in August 2021 are included. Stations parallel to IMRs gill- and fykenet survey (red) and control stations randomly distributed in 20 km radius of the former station (green).

2.3.2 The Reference fishery

To establish a time series with higher resolution data than available from landings data alone, providing information on catch rates and composition, Møreforskning established a network of crab fishers in 2001 who self-reported a subset of their catches. All fishers attending were provided with four standardized pots that were placed in the middle of a link of their own pots. Individual data on catches from each of the four pots, as well as the total landings (kg) from all the commercial pots, were reported in combination with the soak time of each haul. All crabs caught in the standardized pots were measured (carapace

width to the nearest ½ cm), sex determined and discards registered with the reason (undersized, berried, soft shell; for more detailed information see Woll et al. (2006). Initially, a relatively large number of fishers and information were collected to retrieve detailed information to evaluate the sampling program. After an evaluation of the program in 2003, Møreforskning decided to reduce the sampling from ten to five weeks, and from four weekly reports to two from 2004 onwards (Woll et al., 2006). A crucial bottleneck in volunteer sampling programs is the need for motivated participants that deliver data reliably, resulting in potential trade-offs between the number of participants and observations, data quality, and the costs (participants are compensated financially for their contribution). For economic reasons, and to increase the participation rate and data quality, the reporting was further reduced to a biennial cycle in 2015. To date, these data are the only available time series of catch and effort data as well as biological data, including length frequencies, of the brown crab fishery in Norway. The reason for this is mainly due to the limited reporting duties of boats under 15 m in Norway.

2.3.3 Landings slips

Until 1976, Norwegian landings are currently only available accumulated as national annual landings (Statistisk Sentralbyrå, 2022). From 1977 onward, the landings were reported by area and aggregated monthly. Since 2005, landings data are available per sales slip, i.e. when a boat lands the catch of one species at a landing site, and includes therefore detailed information on the vessel, landing site, statistical fishing area and location, the total landed weight, price, etc. (Fiskeridirektoratet, 2022). As the minimum requirement for LPUE standardization are vessel- and landings-specific information, only data from 2005 onward were included in the analysis.

2.4 Data analyses

2.4.1 Data exploration and validation

The potting survey deviates from the reference fishery by having a semi-random sampling design. Deviation between the catch rates indicates therefore an unquantifiable "fishing effect", implying that fishers target certain areas based on expectation and experience to maximize the catch rate. If such a fishing effect is evident it implies reduced applicability

of the reference fishery data, as observed catch rate may not be hyper-stable and not fully representative of the stock. To validate the reference fishery, catch rates from the reference fishery were compared with catch rates from the potting survey in the same area (Møre) and in the same period (August). To test if the catch rates, defined as number of crabs per pot, in the two data set were different, the data were analysed with a Generalized Additive Model (GAM) using the mgcv-package (Wood, 2017). To correct for different soak time (the number of hours the gear was fishing) and depth, the two variables were included as explanatory variables with a smoothing function:

$$Number\ of\ crabs\ per\ trap \sim Data\ set + s(Soaktime) + s(Depth),$$

with the two data sets as categorical variable, and the smooth functions (denoted as 's') are restricted to three degrees of freedom to avoid overfitting. Additionally, length distribution and sex composition were compared descriptively.

To evaluate whether the gill- and fyke-net survey would be suitable for estimating a crab abundance index, the catch rates were compared to catch rates in the potting survey. Similar to the comparison between survey and reference data, a possible deviation between catch rates in (i) the two station types of the potting survey and (ii) the two surveys will indicate an unquantifiable "fishing effect". Assuming that the control stations in the potting survey give a representative picture of crab density, given that the sampling design was semi-random and targeted suitable habitat, deviating catch rates between the two station types would indicate that the gill- and fyke-net survey does not target suitable habitat and is therefore not indicative of (changes in) crab density. First, the catch rates between the two station types were compared in a GAM to correct for different depths and soak times,

$$Number\ of\ crabs\ per\ trap \sim Station\ type + s(Soaktime) + s(Depth),$$

With the two station types sets as categorical variable, and the smooth functions restricted to three degrees of freedom to avoid overfitting. Second, the catch rates in the pots parallel to the gill- and fyke-net survey were compared descriptively and with linear regression to the catch rates in the gill- and fyke-net respectively.

2.4.2 Standardizing catch- and landings data

Maunder and Punt (2004) emphasize that although the majority of such CPUE standardization approaches are based on generalized linear models (GLM; Nelder and Wedderburn, 1972), generalized linear mixed models (GLMM; Pinheiro and Bates, 2006) are more appropriate in standardization of CPUE. By controlling for variance in certain variables rather than estimating their fixed parameter values, combining mixed-effects modelling allows for more flexible model construction, especially when combined with features of additive models into generalized additive mixed models (GAMMs) that allow to model nonlinearity in the relationship between continuous explanatory variables and the response. In order to standardize catch and effort data, the year must be included as one of the explanatory variables, as the year effect is assumed to reflect the changes in the abundance (Maunder and Punt, 2004).

The variables

Standardized indices of catches or landings per unit effort can provide a relative index of abundance to qualitative assess trends in a stock and to use as input in stock assessment models. To standardize catch rates, the annual change in catch rates needs to be isolated from all the potential influential variables that could affect the catch rates and partly explain the observed variation. Number of crabs per pot was set as the response variable in the catch-per-effort (CPUE) model, while weight of landed crabs per boat trip was used in the landings-per-effort (LPUE) model. The catch and landings data were explored to identify which variables to include. Based on data exploration, soak time indicated an asymptotic relationship with the catch rates. Similar patterns were found for the seasonal parameter (Season day). They were both included in the model as continuous fixed effects with a natural cubic spline (denoted as 'ns'), resulting in a generalized additive mixed-effects model (GAMM). All splines were restricted to three degrees of freedom to improve model convergence and ensure a mechanistically plausible shape of the smooth function (i.e. avoid overfitting and wobbly shapes due to e.g. data artifacts). The spline breaks down the variable into line segments, and the degrees of freedom control the number of segments. Soak time was calculated from the start and stop time of the soaking of the pots. Season day was defined as the day in each vessels' fishing season, in respect to the first day a vessel fished within a year (i.e. the first registered date in the landings data). All

three continuous variables were normalized (Z-score) to facilitate model convergence and comparability of estimated effect sizes. The model structures explaining the observed variation best were determined through backward model selection, removing fixed effects and comparing the resulting Akaike information criterion (AIC; Akaike, 1998).

Using mixed-effects modelling to account for the spatial and individual variation in the observed landings and catch rates, both variables were included as random effects. Spatial effects refer to the assessment units (Figure 1), and were included with three (Vestlandet, Møre, and Helgeland) and five levels (Skagerrak, Vestlandet, Møre, Helgeland, and Lofoten & Vesterålen) for the CPUE and the LPUE, respectively. Punt et al. (2001) emphasized that the vessel should be included as an explanatory variable if the fisher is thought to be the main influencer of catchability. The vessel and fisher effects were included as random effects since it resulted in a more suitable model structure. This models the individual catchability as unobserved processes with a normal distribution, reducing the number of parameters substantially compared to including them as categorical fixed effects. Registration number of vessels, representing 522 unique boats in the landing slips, was included in the LPUE standardization. In the CPUE standardization based on reference fishery data, we were able to identify that several fishers changed their vessel throughout the study period, making individual fisher a more suitable proxy of individual catchability effects than vessel name. This was corroborated in the analysis, as including fisher as an explanatory factor explained more of the variance than using vessel name. Fisher refers to the person reporting and named in the reference reports and was included with 38 unique levels in the CPUE standardization.

To allow for potential variation in spatial trends, modelling the area and fisher random effects as a first-order autoregressive process (AR1) in time was tested. Autoregressive modelling is a common concept in time series models, where the current predicted values is an autocorrelated regression, dependent on the past performance. This pattern can be absorbed in the model by including autoregressive function to the relevant variable, here time, allowing for the extraction of area specific trends. AR1 processes were applied to fisher effects in the CPUE standardization, and area and vessel in the LPUE standardization and compared to model structures without AR1 processes.

The standardization was done by fitting the reference fishery data and the landing slips data, respectively, to a GAMM model with log-link function and delta error structure, by using the *glmmTMB* function in the R-package with the same name (Brooks et al., 2017). The predicted year effect (hereafter, CPUE/LPUE index) was standardized to its mean and was plotted. All statistical analyses were implemented in the R statistical programming environment (Version 1.4.1106,R Core Team, 2021).

Model selection

The performance of all models and their goodness of fit were evaluated based on residual distribution using simulated residuals produced with the DHARMa package (Hartig, 2022) and AIC. The random effects structure as well as the fixed effects included in all models were selected by stepwise backward selection based on starting with all variables thought to explain/affect the effort. If the difference in AIC scores (Δ AIC) between two models was less than two, then the model with the fewest parameters was selected as it was considered to be more parsimonious (Burnham and Anderson, 2004).

Sensitivity analysis

To improve the fit of the models and remove noise in the data set caused by unbalanced sampling, typing errors, etc, the data was filtered. The effect of filtering conditions and threshold values were analysed through sensitivity analysis to ensure that the estimated indices are robust to specific filtering conditions and assumption. For the CPUE the sensitivity analysis was applied to the following filters: (i) a minimum number of observations per fisher and year, (ii) numbers of years each fisher is active in the fishery, (iii) starting year of the time series, and (iv) areas included. For the LPUE two additional filters were tested: (v) upper limit of catch per boat trip (kg), and (vi) a lower limit of annual catches. Sensitivity analysis and the final filtrating conditions used to estimate the indices are listed in Appendix B. The CPUE model was highly sensitive to filtrations on area and start year, but not sensitive towards the number of observations and active years. Similar to CPUE, the LPUE model was not sensitive towards the number of observations and active years, to the contrary it was highly sensitive to filtrations on annual catches. It was not sensitive towards which areas were included since specific area trends were already account for explicitly in the model.

2.4.3 Stock Assessment

Reference points are used in management to provide reference values against stock indicators, typically including stock sizes and fishing mortalities that maximize sustainable yields or where the stock productivity is put at risk. Maximum sustainable yield (MSY) is a favoured reference point in assessment, as it is easily interpretable for management bodies. MSY is a theoretical measure of the yield that could be harvested from the exploitable stock without compromising sustainability (Larkin, 1977). Unfortunately, it requires a great amount of biological input data, which is often absent in data-limited stocks. While category 1 and 2 stocks can be managed based on MSY, advice regarding stocks in categories 3-6 aims at sustainable catches while complying with a precautionary approach based on the best available data (ICES, 2021a)

There is a wide range of assessment methods, however there is substantial redundancy as most of them rely on a few classic equations and are iterations of the same basic concepts. Differences can typically be reduced to detailed specifications, optimization algorithms and implemented features. To avoid countless comparisons of near-identical versions of the same methods, it requires an approach that reduces them to archetypical implementations of the underlying assessment. Following the work conducted within WKLIFE (ICES, 2015), ICES suggests a few well-documented and maintained implementations of different length-based methods and surplus production models. I have used the ICES recommendations to guide my selection of which methods to test.

Length-based Indicators

Fishing does not only affect the abundance of a stock, but also the length distribution. The effect of mortality accumulates over time, as the cohort grows and ages. High fishing pressure can truncate the size distribution toward smaller animals , impacting population dynamics and productivity (Hixon et al., 2014). Size structures in a population can be used as a "snapshot" of the current status of the population or as an indicator for temporal changes. One of the length-based assessment methods are length-based indicators (hereafter indicators) that are compared to respective life-history parameters (hereafter reference points). The ratio between the two reports the status of the stock in an easily interpretable "traffic-light" manner. There are a variety of indicators and reference points,

I selected two that have been applied on brown crab stocks earlier (Mesquita et al., 2017), mean length and length of the largest individuals. Only landed crabs were included. ICES emphasizes that crustaceans should be evaluated by sex (ICES, 2015), and to evaluate spatial differences in the length structures, each area was estimated separately.

Mean length above length at first capture (L_{mean}) and mean length of the 5% largest individuals were calculated, plotted, and compared to their respective reference points. Length at first capture was calculated, as described by Mesquita et al. (2017), as the length at half the mode in the ascending part of a length frequency distribution. The L_{mean} indicator was compared to the $L_{F=M}$, a recognised proxy for the mean length at MSY (Jardim et al., 2015). $L_{F=M}$ is defined as the expected mean length when fishing mortality is equal to natural mortality, and is calculated based on the ratio between the natural mortality (M) and the growth rate (k) of the von Bertalanffy equation (Von Bertalanffy, 1938). Since the two parameters are uncertain, a M/k ratio equal to 1.5 is assumed in data-limited cases (Prince et al., 2015). The equation derived from Jardim et al. (2015) could therefore be applied to calculate the reference point,

$$L_{F=M} = 0.25L_{inf} + 0.75L_{c,}$$

where L_{inf} is the asymptotic size, the length that 50% of the population would reach if the population were unfished, and Lc is the aforementioned length at first capture. The mean length of the 5% largest individuals, ($L_{max5\%}$)(Probst et al., 2013) was compared to the reference point $0.9L_{inf,}$ (Miethe and Dobby, 2016). As the L_{inf} value varies a lot (190-240 mm), it was selected *a priori,* based on Mesquita et al. (2017) were L_{inf} equals 220 mm for both sexes and are derived from Chapman (1994). The ratio between the indicators and reference point was calculated and presented as an average of the three last years (2017, 2019, and 2021), for each sex and assessment units. To test the model's sensitivity towards the input values from L_{inf}, a sensitivity analysis was conducted.

Length-Based Spawning Potential Ratio

The second assessment method recommended by ICES for species in category 3 is length-based assessment for spawning potential ratio (LBSPR; Hordyk et al., 2016). Since estimates of local life history parameters (natural mortality and growth) are frequently

unavailable in data-limited stocks, and since LBSPR does not require these parameters, it is a specifically appealing assessment method. LBSPR uses maximum likelihood to estimate the ratio between fishing mortality and natural mortality (F/M), and length at first capture of 50% and 95% of the stock. These three parameters are then used to calculate the spawning potential ratio, an estimate of the stock status, defined as the proportion of unfished reproductive production left in a population under fishing pressure (Goodyear 1993; Walters and Martell, 2004). An unexploited stock is assumed to have a spawning potential ratio (SPR) of 100%. Under fishing pressure, the natural spawning potential is reduced to some proportion of the initial level. Simplifying assumptions are as common as they are essential in data-limited assessments, and the LBSPR model is no exception. LBSPR assumes that the (i) size composition data is representative of the exploited population in equilibrium, (ii) the growth is adequately described by the von Bertalanffy equation, (iii) normal distribution of the length-at-age, (iv) asymptotic selectivity, (v) across cohorts the growth rates remain constant, and (vi) the natural mortality is constant in all adult age classes. When applying LBSPR, these assumptions must be evaluated explicitly.

Size structures and SPR of an exploited stock could be described as a function of the ratio (F/M) of fishing (F) to natural mortality (M) and the two life-history ratios M/k and L_m/L_{inf}, where M is the rate of natural mortality, k is the von Bertalanffy growth coefficient, L_m is the size at maturity where 50% of the stock is mature, and L_{inf} is a parameter from the von Bertalanffy growth function, referring to the mean length of a crab in a given stock if it were to live infinitely (asymptotic size). This is the fundament of the LBSPR model (Beverton, 1963; Hordyk, 2015a). The input data in the LBSPR are the following: (i) Length at maturity, for both when 50% (L_{m50}) and 95% (L_{mat95}) of the stock is mature, corresponding to 112 mm (Bakke et al., 2018) and 159 mm (Woll et al., 2003), (ii) L_{inf} corresponding to 220 mm (Chapman, 1994), (iii) the M/k-ratio, which is assumed to be 1.5 for data-limited species (Prince et al., 2015), (iv) the variability of length-at-age (CVL_{inf}), which is normally assumed to be 10% when reliable length and age data is not available. LBSPR is applied to size composition data that are assumed to be representative of the commercial catches. Length data used in the LBSPR model were therefore based on the measurements of the carapace width from the reference fishery. The data were aggregated into 5 mm size classes. Only data from the Møre area were used for assessment purposes, due to the uniform length structure in the stock and sufficient sample size.

Reported discards were excluded to resemble the commercial landings. The analysis was conducted in R (version 1.4.2, R Core Team, 2022) by using the LBSPR package (Hordyk, 2021).

In lack of exhaustive life history data from Norwegian waters and to reduce the scope of this study, life-history parameters were obtained from various stocks in the whole distribution range of the species without a systematic approach.

Surplus Production in Continuous Time

Surplus production models are a common assessment method in data-limited fisheries, where mortality and age, and length structure of the population are not known (Beverton and Holt, 1957; Punt et al., 2013). Surplus production model in continuous time (SPICT; Pedersen and Berg, 2017) is one of the assessment methods acknowledged by ICES for category 3 stocks (ICES, 2015). SPiCT is using a reformulated Pella-Tomlinson production model that includes a shape parameter to skew the production curve to either side, with a symmetric Schaefer model as one specific form (given a shape parameter of 2). A key feature compared to other implementations is that SPiCT is not discrete but continuous in time, modelling within-year dynamics of a population by splitting the year into shorter segments. SPiCT allows, furthermore, to define priors for parameter values to provide bounding probabilities.

We used the standardized LPUE index with area-specific trends (E. 2), and annual observations of catch covering the period 2005-2021 To narrow down and guide the model, we assumed values for three priors. A logistic growth form (Schaefer model) was assumed to describe the smoothed growth form of decapods and the relevant shape parameter was fixed (n=2). The intrinsic population growth parameter (r) was constrained through a lognormally distributed prior with mean of 0.46 and a standard deviation of 0.1 (SeaLifeBase, 2022). Since the stock was fished but likely not overexploited, the prior for the initial biomass (B_0) was set to 75% of the carrying capacity (K).

Unless otherwise stated, the default settings and priors of the SPiCT package (Pedersen and Berg, 2017) were used when fitting the model. The packages use the computation tool

Template Model Builder from the TMB package (Kristensen et al., 2016).The model was validated through the SPiCT library diagnostics, for empirical autocorrelation in the residuals and if they were normally distributed. To investigate the robustness of the model towards the input data and priors, we conducted a sensitivity analysis for three priors and alternative inputs from the CPUE index and annual catches from 1970-2021.

3 Results

3.1 The reference fisher 2001-2021

The catches in the reference fishery peaked during the three first years, corresponding to high participation rates (Figure 4). Since 2004 the catches have varied between 3,000-9,000 crabs annually. More than 150,000 crabs have been caught and measured by the reference fishers since 2001, however, only 476 berried females have been caught. Discard in the fishery has both interannual variations as well as spatial variation (Zimmermann et al., 2020). Most of the discard consists of crabs smaller than the MLS. In the three northern areas, male crabs dominate the catches, whereas a more even sex composition is observed in Vestlandet, with annual fluctuations (Appendix C). Size distributions varied little between the areas, with exception of Vestlandet where the crabs tend to be smaller.
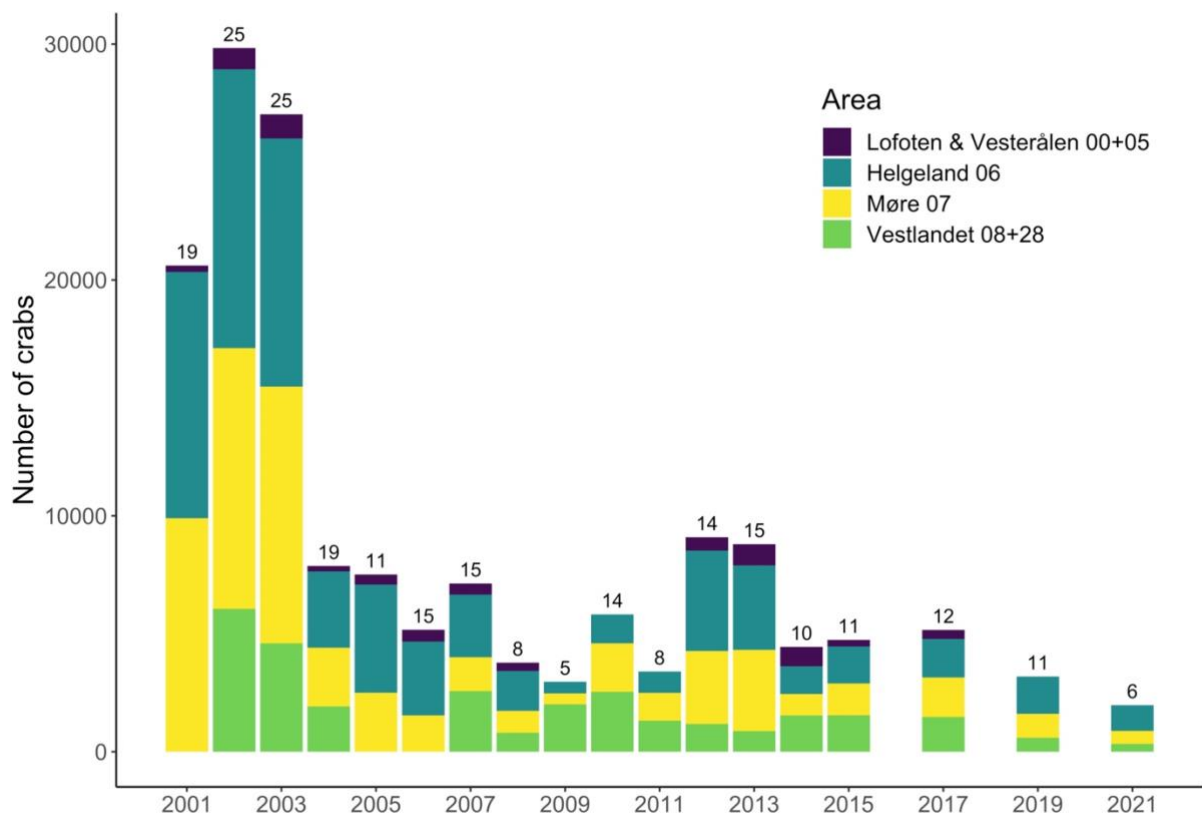

Figure 4 Annual catches in the reference fishery (2001 – 2021). Number of crabs caught and measured in the different areas. Numbers indicate number of fishers participating each year in all areas combined.

3.2 Data exploration and validation

3.2.1 Potting survey vs reference fishery

During the two-week potting survey, a total of 3079 crabs were caught, measured, and registered at 79 stations. The catch rates in potting survey and reference fishery were compared (**Error! Reference source not found.**) and the statistical analysis with a GAM showed no significant difference ($p > 0.05$) between the two catch rates when accounting for depth and soak time effects (Appendix D). This implies that the reference data is not different from randomized survey observations and, thus, likely not biased through potential fishing effects (e.g. fishers focusing on specific areas with the highest density, resulting in hyperstable catches). Subsequently, the reference data can be considered representative of changes in the stock. The validation is, however, limited to the area of Møre in August 2021, and is only based on two reference fishers. The length composition captured in the survey was normally distributed, while it was slightly right-skewed in the reference fishery (Appendix C).

3.2.2 Gill-and fyke-net vs potting survey

The catch rates in the two different station types in potting survey were compared and the statistical analysis with a GAM showed a significant ($p < 0.01$) difference between the two catch rates when accounting for depth and soak time effects (Appendix C). The catch rates at the stations parallel to the gill- and fyke-net survey were significantly smaller than at the control stations (**Error! Reference source not found.**). The low catch rates at the stations parallel to the gill- and fyke-net survey indicate successful avoidance of crab through choice of locations, indicating that gill- and fyke-net stations are mostly placed outside of main crab habitat and therefore not suitable as indicator of changes in crab abundance.

The catch rates of different gear types at the same location (i.e. a potting survey station taken parallel to a gill- and fyke-net station), were compared descriptively, and there was found to be a linear relationship between the pots and fyke- and gillnets, respectively (Figure 6). This implies that the among-station variation followed a similar trend independent of gear. However, the fact that the gill- and fyke-net survey actively avoids

crab habitat undermines the applicability of the latter survey as an adequate survey for making a crab abundance index.
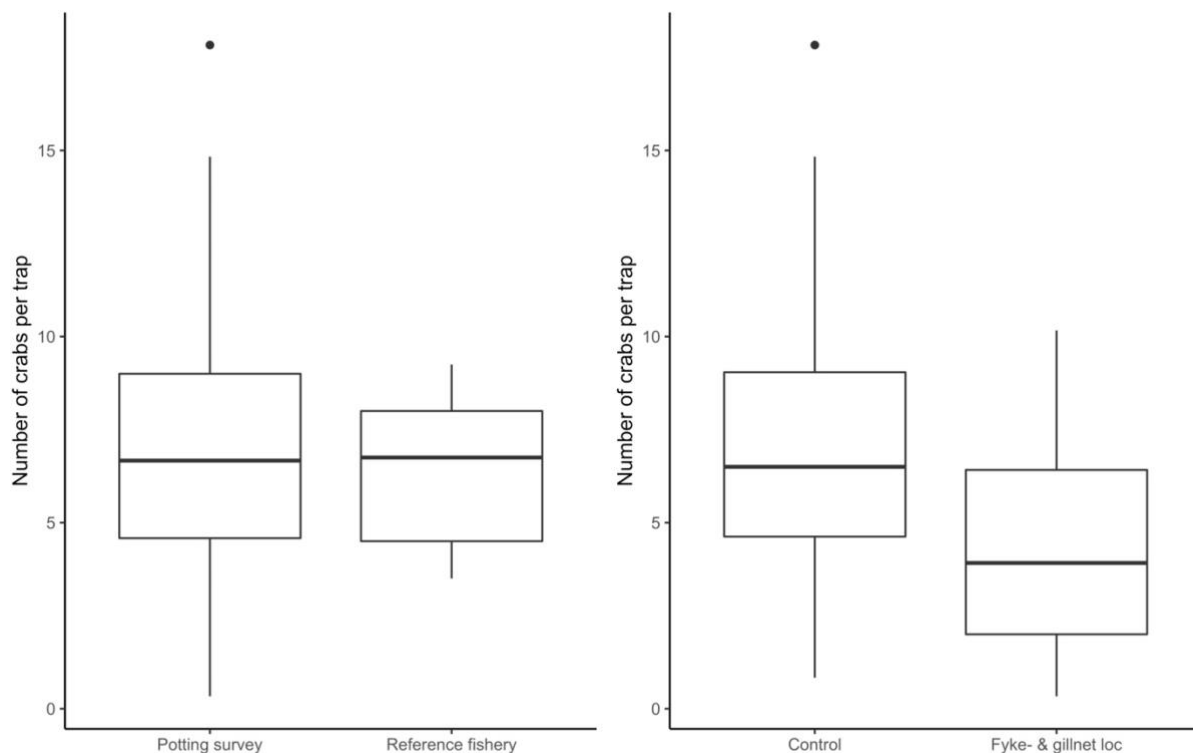
Figure 5 Catch rates compared between a) the potting survey and the reference fishery, where the data from the reference fishery consist of a subset of Møre in August 2021 which is temporally and spatially similar to the potting survey, and b) the two station types in the potting survey: (i) stations parallel to the gill- and fyke-net survey and (ii) control stations
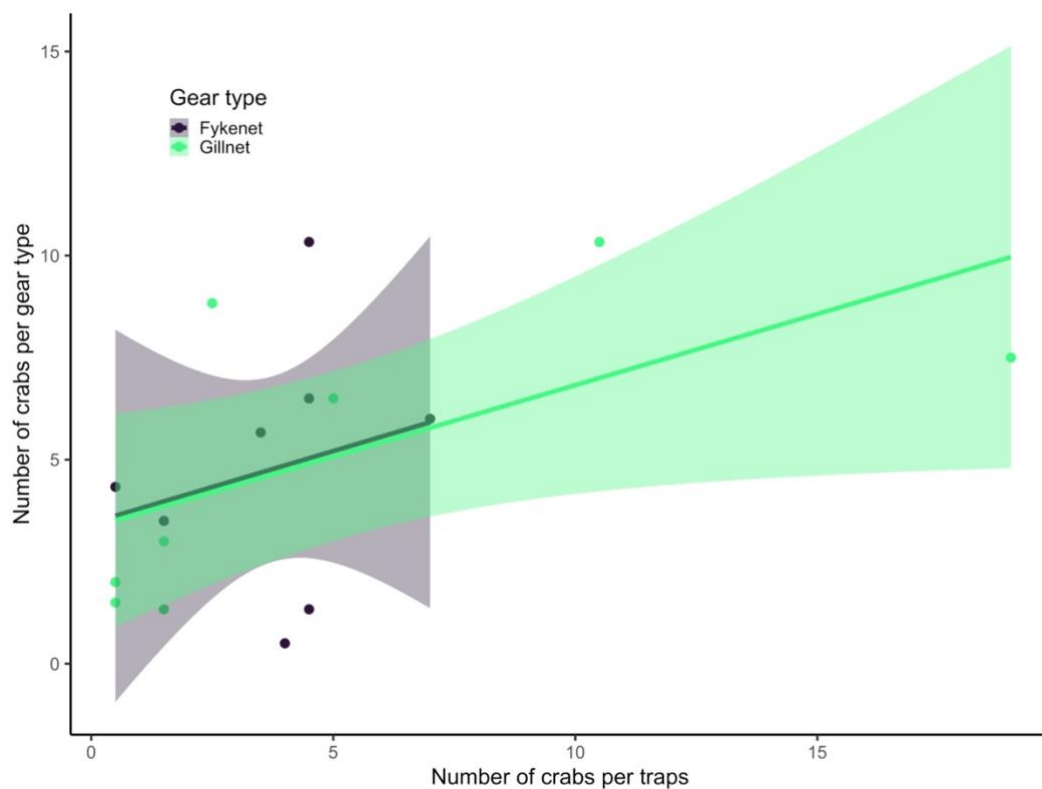
Figure 6 Catch rates from the two parallel stations during the surveys with pots and gill- and fyke-net. A linear regression with 95% confident interval is added to the data to highlight the relationship between the catch rates in the different gears (pots, gill- and fyke net.

3.3 Indices –reference fishery vs commercial landing

3.3.1 Standardization of catch/landing and effort

The catches and effort from the reference fishery (CPUE) and the landings and effort from the commercial fishers (LPUE) were successfully fitted to three models, E1 -E3. Models that included random effects captured a substantial part of the observed variation and performed well, especially when including random effects with a AR1 covariance structure over fisher and year in the CPUE model (E.1). Using fisher instead of vessel in the CPUE model improved the model, likely since several of the fishers changed boats during the time series, and using fishers, therefore, is a reduction in levels and more indicative of the (unobserved) individual catchability effects. No significant effect of depth on catch rates was found and based on AIC depth was excluded from the CPUE model in the model selection process.

$(E.1)$ $Number\ of\ crabs\ per\ trap \sim fYear + ns(Season\ day, df = 3)$
$$+\ ns(Soak\ time, df = 3) +\ (1|fArea) + ar1(fYear - 1|fFisher)$$

$(E.2)$ $Weight\ of\ crabs\ per\ trip \sim fYear + Julian\ day +\ ar1(fYear$
$$-\ 1|fArea) + ar1(fYear - 1|fVessel)$$

$(E.3)$ $Weight\ of\ crabs\ per\ trip \sim fYear + Julian\ day +\ (1|fArea)$
$$+\ ar1(fYear - 1|fVessel)$$

The LPUE model explaining the data best included random effects with a AR1 structure for both area and vessel in time (year) (E.2). Area effects with AR1 effects in time allow to predict LPUE indices with different trends for the different areas, a useful feature when testing area-specific surplus production models. The structure of E.2 was therefore used in the SPiCT assessment. The spatial resolution of the reference data was not sufficient to apply the same model structure. Therefore, to facilitate a direct comparison with the CPUE index the same random structure for area used in E.1 for the CPUE model was selected *a priori* for an alternative LPUE model (E.3) to generate CPUE and LPUE indices on global level (i.e. not per area) with the same random effects structure.

3.3.2 Trends in the abundance – national scale

An overall national trend in the Norwegian crab stock between 2001-2021, is presented in Figure 8. The two indices indicated no major changes in the relative abundance along the Norwegian coast, and all observed changes lie within the variability of the two indices. The two indices, CPUE and LPUE, showed very similar trends on a national scale, (Figure 7), suggesting the landings data to be adequate for further assessment purposes. The deviation between the two indices has not been tested statistically and is solely based on visual observation of the trends. The year effect was significant in all years across all models (Table 1).

Figure 7 Trends in the standardized catch- and landing-per-effort (CPUE and LPUE) indices (E.1 and E.3), which are thought to be indicative of changes in the relative abundance of the total brown crab (*Cancer pagurus)* stock in Norway. The lines are predictions of the year effect from the two GAMM models that were fitted to reference fishery data and the commercial landings resulting in the CPUE and LPUE. The shaded areas are the 95% confidence intervals. The index is standardized to its mean.

3.3.3 Trends in the abundance – regional scale

The variation between sites was captured by the random intercept in the indices of the area variable. This variation was found in both the LPUE and CPUE indices (Figure 8). By comparing the random intercepts between the two indices, the same area effect was found for Møre and Vestlandet. This pattern was not observed for the area of Helgeland. The fact that the two indices do not cohere in Helgeland does not necessarily imply inadequacy of the commercial data in this area but may be rather a result of few fishers present and high uncertainties in Helgeland. This is partly captured by the large variance observed in the random intercept of this area in the CPUE. The high uncertainties in the CPUE area estimates should therefore be taken into consideration.
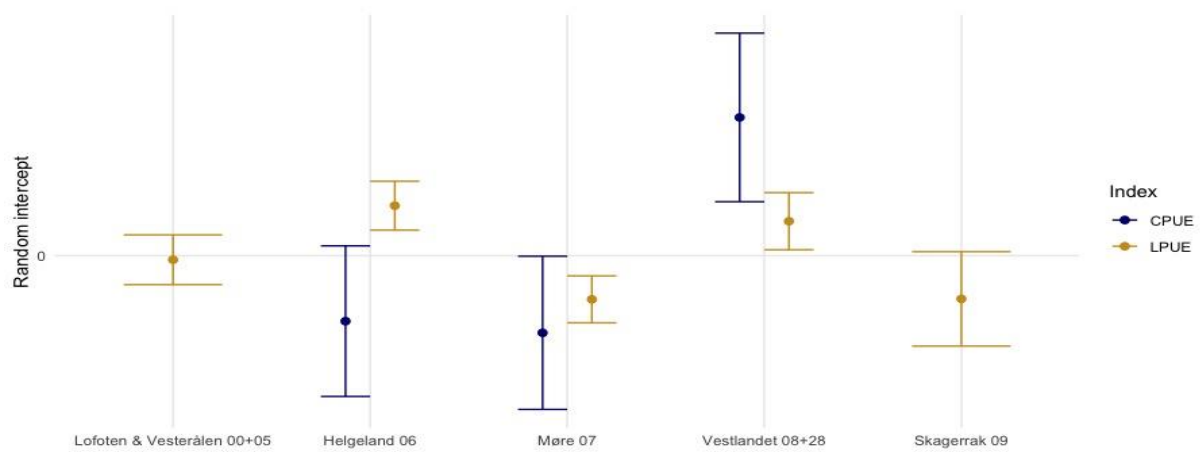
Figure 8 The predicted effect of area on the catch and landing rates. Represented by the random intercept from the two catch- and landing- per-unit-effort standardisation models. The two models fitted the reference fishery data and the commercial landing data to two GAMM models (E.1 and E.3), where the area variable was included with random effect structure, with three and five levels, respectively. The points are the mean effect, and the line are representative for the variance.

The model that includes region-specific trends in time resulted in the better fit based on AIC. Additionally, it allowed for estimating spatially separated indices of abundance in the five assessment units. Looking at trends in each of the five assessment units, differences emerge (Figure 9). In Vestlandet, an abrupt decline in the relative abundance was observed in 2015-16, before the relative abundance stabilized. In the same time span, the LPUE index indicates an abrupt increase in Lofoten & Vesterålen. However, the variability is high due to few boats participating, hence, smaller sample sizes the last years in this region. The abrupt increase comes after a period of relative abundance below initial levels. In Møre and Helgeland, the indices indicate no changes since 2005. In the southernmost region, Skagerrak, there was an increase after 2012, followed by a slight decrease after 2014. However, the uncertainty in the estimates is especially large due to the low number of observations, and all the changes in the index are within the 95% confidence intervals.
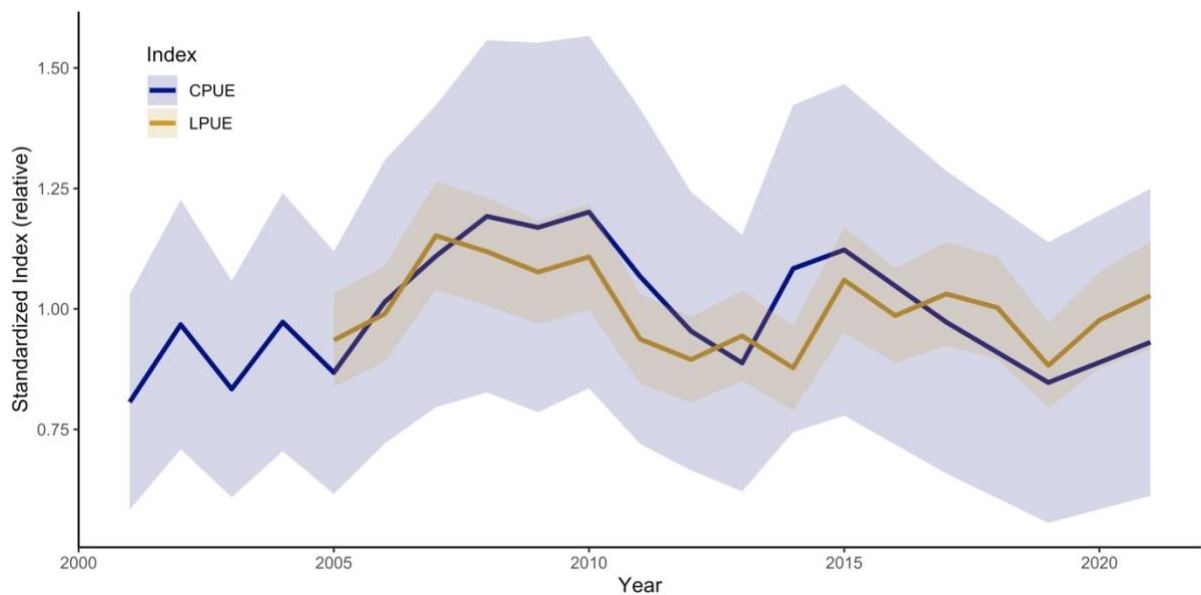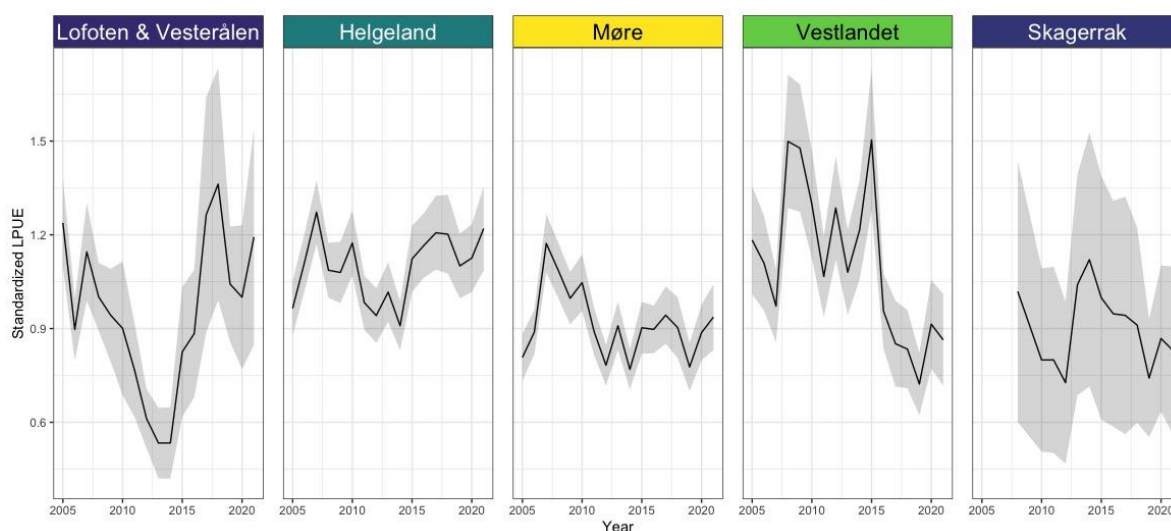
Figure 9 Trends in the standardized landing-per-effort (LPUE) index (E.2), which are thought to be indicative of changes in the relative abundance in each of the assessment units (Lofoten & Vesterålen, Helgeland, Møre, Vestlandet, and Skagerrak). The lines are predictions of the year effect from the GAMM model that was fitted to commercial landings resulting in the LPUE. The grey areas are 95% confidence intervals. The index is standardized to its mean.

3.3.4 Variables affecting the catch rates

Other variables besides the year effects were included in the models to standardize the catch and landings rates and extract the annual changes. A wide range of soak times were observed in the catch data (20- 300 hours). Soak time is known to have an asymptotic relationship with catch rates, where catch rates increase until a certain point and then start to decrease. The variable was included in the model with smoothing function with three degrees of freedom (hereafter segments). The predicted effect on the CPUE was significant in the second segment and absent in the first and third segment (Table 1). In other words, soak time increases the catches after a certain time, but the increase levels off. Season day increased the catch rates, modestly but significantly in the third segment (Table 1), implying an increase in catches later in the individual season. The Julian day variable in the LPUE standardization indicates a slight but significant reduction in catch rates during the year, suggesting a tendency of reduced catch rates linked to fishing mortality (fishing-out effect) and seasonal shifts in the behaviour and distribution of crabs.

Table 1 GAMM models fitted to the reference fishery data and the commercial landings data, to derive standardized catch-per-unit-effort indices that provide indicators of stock development. The three catch (CPUE) and landings-per-unit effort (LPUE) standardization are result from the three equations E.1-E.3. Soak time and Season day were given as smoothers with three degrees of freedom in the models.

Predictor	CPUE (E. 1)			LPUE (E. 3)			LPUE (E. 2)		
	Estimate	*SE*	*p-value*	*Estimate*	*SE*	*p-value*	*Estimate*	*SE*	*p-value*
Year									
2001	1.879	0.145	<0.01	-	-	-	-	-	-
2002	2.063	0.140	<0.01	-	-	-	-	-	-
2003	1.911	0.141	<0.01	-	-	-	-	-	-
2004	2.066	0.144	<0.01	-	-	-	-	-	-
2005	1.976	0.152	<0.01	6.138	0.053	<0.01	6.236	0.101	<0.01
2006	2.106	0.151	<0.01	6.195	0.051	<0.01	6.196	0.107	<0.01
2007	2.196	0.148	<0.01	6.347	0.050	<0.01	6.120	0.106	<0.01
2008	2.271	0.159	<0.01	6.318	0.051	<0.01	6.323	0.105	<0.01
2009	2.256	0.170	<0.01	6.279	0.050	<0.01	6.312	0.101	<0.01
2010	2.287	0.159	<0.01	6.308	0.051	<0.01	6.233	0.106	<0.01
2011	2.156	0.169	<0.01	6.141	0.050	<0.01	6.094	0.105	<0.01
2012	2.051	0.158	<0.01	6.094	0.050	<0.01	6.032	0.104	<0.01
2013	1.980	0.157	<0.01	6.148	0.051	<0.01	6.087	0.103	<0.01
2014	2.244	0.197	<0.01	6.075	0.051	<0.01	6.071	0.104	<0.01
2015	2.225	0.161	<0.01	6.264	0.052	<0.01	6.251	0.108	<0.01
2016	-	-	-	6.191	0.051	<0.01	6.170	0.106	<0.01
2017	2.044	0.171	<0.01	6.236	0.051	<0.01	6.234	0.111	<0.01
2018	-	-	-	6.208	0.054	<0.01	6.229	0.107	<0.01
2019	1.943	0.178	<0.01	6.081	0.051	<0.01	6.057	0.101	<0.01
2020	-	-	-	6.181	0.052	<0.01	6.159	0.103	<0.01
2021	2.032	0.178	<0.01	6.232	0.055	<0.01	6.201	0.107	<0.01
Soak time				-	-	-	-	-	-
1	0.154	0.053	<0.01						
2	0.588	0.115	<0.01						
3	0.258	0.165	0.118						
Season day				-	-	-	-	-	-
1	0.036	0.029	0.205						
2	0.063	0.043	0.142						
3	0.271	0.065	<0.01						
Julian Day	-	-	-	-0.062	0.002	<0.01	-0.062	0.002	<0.01

3.4 Stock Assessment

3.4.1 Length-based indicators

Two size-based indicators were compared to their respective reference point, where the ratios are limits that are indicative of the status of the stock. L_{mean} was compared L_{mf}, and $L_{max5\%}$ was compared to L_{inf90}, and the ratios should be equal to or above 1. The indicator-reference ratios are indicative of catches according to F_{MSY} and a minimum limit for conservation of the largest individuals, respectively. Additionally, the LBIs represent trends in the length composition through the time series, 2001-2021. The L_{mean} were relatively stable throughout the time series (Figure 10) but were generally below the target reference point ($L_{m=f}$) for both sexes, except both sexes in Møre and males in Helgeland (Table 2). The $L_{max5\%}$ of both males and females were consequently below the reference point throughout the time series, with exception of males in Møre some years. Without applying any statistical tests, the crabs caught were generally smaller in Vestlandet.
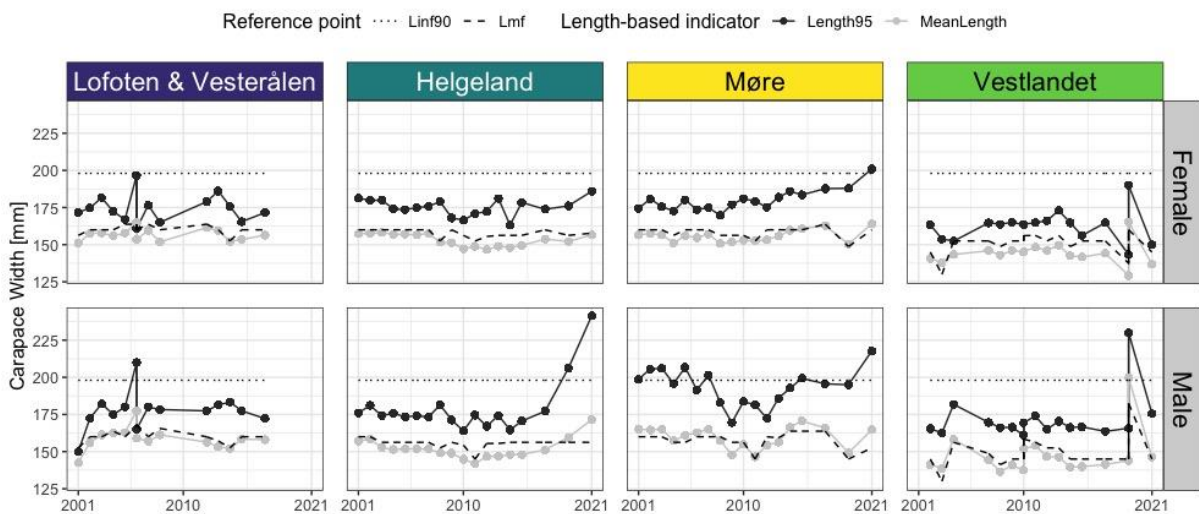
Figure 10 Annual changes in the size structures, expressed through the length-based indicators for each of the assessment units and for both sexes. The indicators mean length (grey) and mean length of the 5% largest individuals (black), and their respective reference points, L_{mf} (dashed line) and L_{inf90} (stiped line). $L_{m=F}$ refers to the mean length when natural mortality equals fishing mortality and L_{inf90} refers to 90% of the asymptotic length.

In the literature, the asymptotic length, L_{inf}, of crabs vary from 190-240 mm, and in this study L_{inf} was predetermined to 220 mm. Since L_{inf} is an input in both reference points, the ratio between the indicator and the reference point is likely to be influenced by the L_{inf}. Therefore, a sensitivity analysis was conducted (Appendix D). For the ratio between indicator for the largest 5% to the 90% of the L_{inf}, it indicated limited sensitivity for in terms of lifting the ratio above the limit target of 1. For Vestlandet, none of the L_{inf} values tested would lift the ratio to 1, this excludes input values as the unreliable source. For Møre, males and females would be above the threshold for the last three years if the L_{inf} were reduced to 210 mm and 200 mm, respectively. For the ratio, L_{mean} / $L_{M=F}$, it indicated high sensitivity to the input value (L_c), indicated by following the mean trend. This variation could be caused by sampling design rather than actual changes in the length composition.

Table 2 The ratio between the length-based indicators and their respective reference points. The indicator mean length (L_{mean}) was compared to the reference point length were F=M ($L_{M=F}$). The target ratio is 1 and is presented as an average for the three last years (2017, 2019, and 2021). Length-based indicators as a ratio between mean length (L_{mean}) and the reference point length were F=M, and between mean length of the 5% largest, 90% of asymptotic length. Green and grey indicate ratios above and below the target, respectively.

Area	Number of crabs	L_{mean} / $L_{M=F}$		$L_{max5\%}$ / L_{inf90}	
	Male/Female	Males	Females	Males	Females
Lofoten & Vesterålen*	488/1356	0.988	0.978	0.782	0.780
Helgeland	3092/11368	1.027	0.975	0.943	0.810
Møre	3740/ 8308	1.031	1.006	0.908	0.867
Vestlandet	3896/ 3388	0.996	0.961	0.780	0.754

*No fishers present in 2019 and 2021

3.4.2 Length-Based Spawning Potential Ratio

Length data from Møre from 2001-2021 were fitted to a LBSPR model, and the model did not converge. To date, there is no formal validation or checkpoints for acceptance of the model, other than checking for major trends in the selectivity, as this is a violation of the assumption of asymptotic selectivity (K. Ono, personal communication, February 2022). A major drawback of LBSPR is the assumption of logistics selectivity, a trait common in trawl, but not in pots, gill- and fyke- nets. Therefore, the observed variation in the selectivity estimations throughout the time series were no surprise. Since there is an MLS for crabs the selectivity is rather dome-shaped than logistic, and to date, there are no implemented functions in the LBSPR- package to account for this. In fisheries with dome-shaped selectivity LBSPR is likely to overestimate F/M and underestimate SPR (Hordyk et al., 2015). Furthermore, assuming asymptotic growth is a slight violation of the assumptions as the growth of crustaceans is stepwise and not asymptotic since they grow by changing shells. Taking the violation of the assumptions of logistic selectivity and growth into account, the result from the LBSPR should be applied with consideration.

From the reference fishery data in Møre, the selectivity at length when 50% (SL_{50}) of the exploitable part of the stock is likely to be caught was estimated to be 130 mm for males, which confirms that the MLS at 130mm determines most of the size selectivity for males. For females SL_{50} was estimated to be 120 mm, indicating that females below the MLS are landed. The estimated selectivity at length in the males reveals a decreasing trend between 2003-2008 before it stabilizes (Figure 11). Trends in the selectivity at length imply a potential violation of the assumption of asymptotic selectivity. The estimates of the F/M are relatively low, ranging between 1.0 and 1.5. The indications of a low fishing pressure correspond with the relatively high SPR (0.5-0.6). The even lower estimates of the F/M for females (0.2-0.8) correspond with the even higher SPR (>0.6).
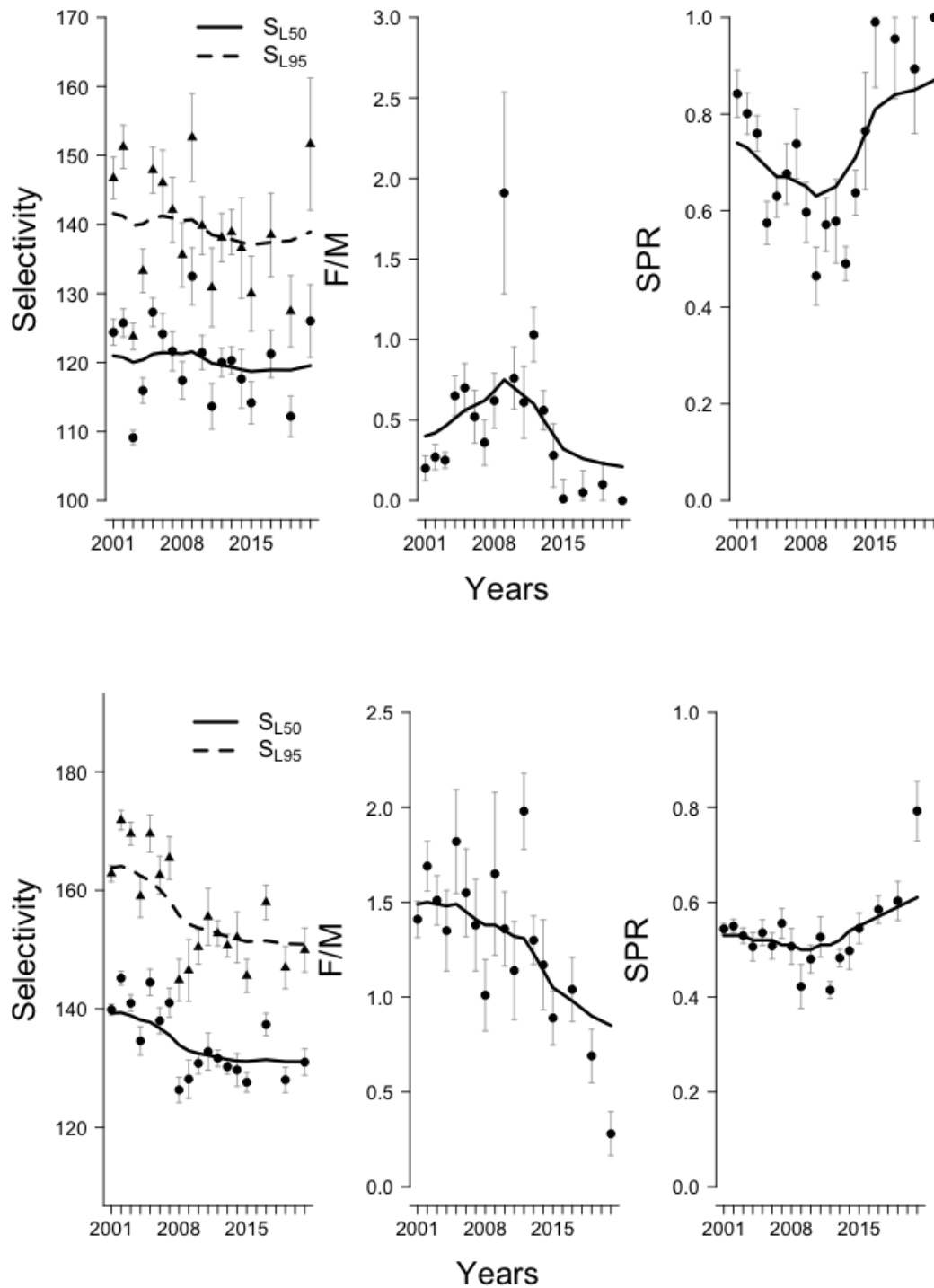
Figure 11 Result from the LBSPR in Møre for both sexes. The three upper panels are the results for males and the lower panels for females between 2001-2021. Selectivity refers to the length at when 50% (SL50) and 95% of the crabs are caught (SL95). The F/M is the ratio between fishing mortality and natural mortality. SPR refers to the Spawning potential ratio. It is assumed that an unexploited stock has a SPR equal to 1. Stocks that are underexploited has a SPR > 0.4, exploited optimally has a 0.2>SPR>0.4, and overexploited has a SPR < 0.2. As a default the LBSPR-package adds a smoother function to the estimated points.

3.4.3 Surplus Production in Continuous Time

The SPiCT model was successfully fitted to the time series of total catches and LPUE index. Results from the diagnostic tests indicated no major violation of most acceptance criterium. In more detail, the model did not converge, all model parameters were finite. However, the uncertainties in the outputs are high, and the model should therefore be rejected based on the acceptance criteria of reasonable low uncertainties. Despite this, some of the results are still shown, and could be indicative of trends, but should not be given too much weight. The mean trends of relative biomass (B/B_{msy}) and relative fishing mortality (F/F_{msy}) in all areas are relatively stable above B_{msy}, and below F_{msy}, with exception of Lofoten & Vesterålen which drops below B_{msy} between 2010 and 2016 (Figure 12).

Furthermore, based on the sensitivity analysis (Appendix E) the model is found to be strongly driven by the selected input values. The estimated values of the shape parameter (n) and the intrinsic growth rate (r) were, not surprisingly, driven by the input of the n and r. The estimated MSY and K were mainly driven by the catch input. Lack of contrast in the input data could be the cause of the high uncertainty observed in Møre and Helgeland, or it could be caused by a violation of one of the assumptions. Either way, this rejects the acceptance of the SPiCT assessment, and thereby the application of the results into reliable advice for management bodies.
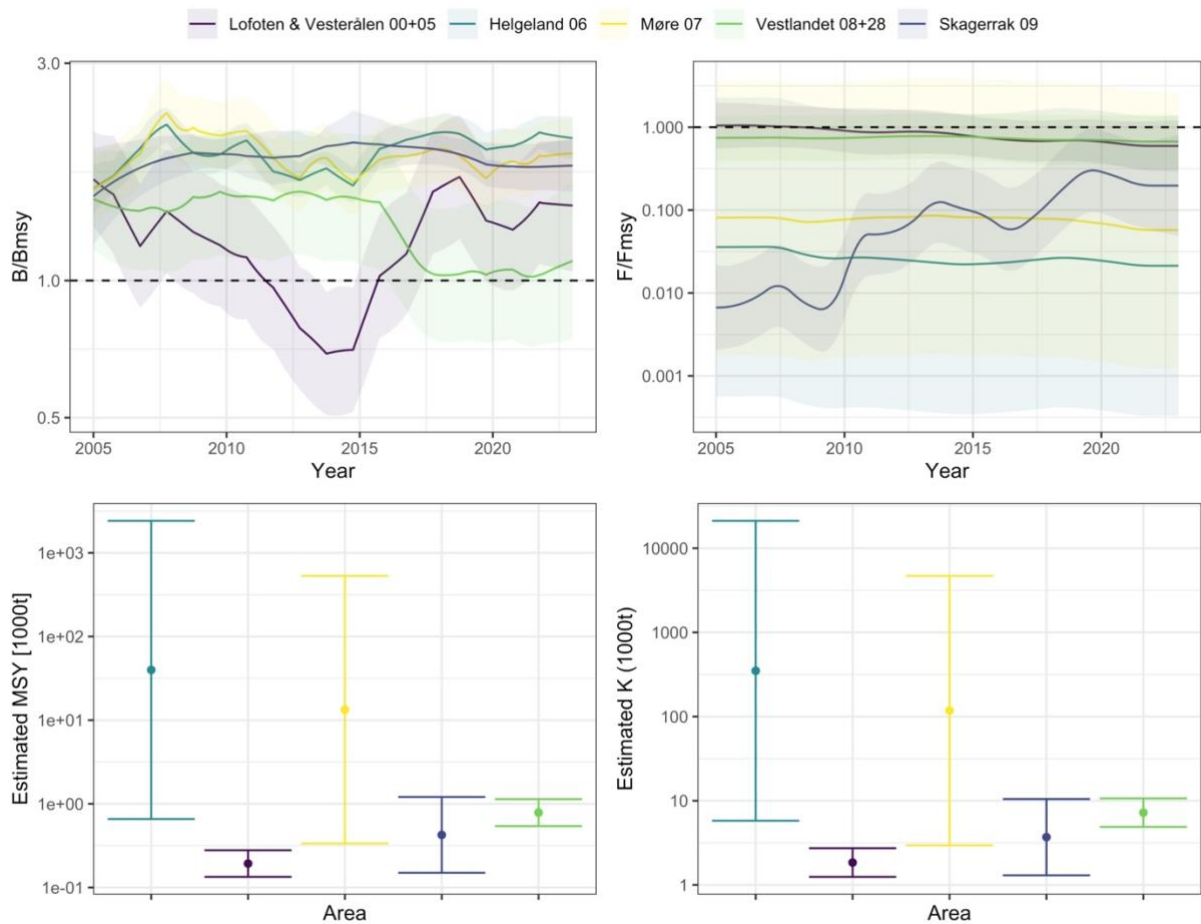
Figure 12 Output of the SPiCT model, run with the landings-per-effort index from 2005-2021. The top left (t.l) panel shows the relative biomass with the 95% confidence interval (CI) and the stiped line represent the B_{msy} reference point. T.r panel shows the relative fishing mortality, here the 95 % CI is very wide. B.l panel shows the predicted maximum sustainable yield (MSY) for each area, note the wide CI for Helgeland and Møre The b.r panel shows the estimated carrying capacity (K) with the following priors: b/k = 0.75, n=2 and r=0.46. Y-axis in all figures (except the F/Fmsy-plot) are at log-scale.

4 Discussion

Data-limited stocks describe a wide range of fish stocks that suffer from a lack of data quantity or quality, or data that complicates or prevents the use of typical assessment methods. Data limitations can take many forms and include short or inconsistent time series, no fisheries-independent data, coarsely resolved fisheries data, illegal or unreported fishing, and insufficient individual data. In addition, biological or fisheries characteristics outside the norm can cause additional challenges, a common situation for shellfish species. Evaluating all available data, their potential and limitations, and their applicability in existing assessment methods is a key step towards establishing a formal assessment for currently unassessed stocks. This study has successfully explored the applicability of two fishery-dependent time series for stock size of the brown crab along the Norwegian coast and validated them by comparing them 1) against each other, on a national and a regional scale, and 2) with a fishery-independent survey conducted in the most important crab fishing grounds in Norway in 2022.

Another key step in establishing a formal assessment is standardization of catch-per-unit-effort, producing indices of relative abundance of the stock that can serve as qualitative indicators of stock status as well as provide required inputs for assessment methods such as surplus production models. From two fishery-dependent datasets, CPUE and LPUE indices were estimated at a national scale. Since the Norwegian brown crab stock is distributed over more than ten degrees of latitude, regional variance was expected. Exploring regional variation within the CPUE standardization was constrained by the low number of fishers from specific areas and locations that resulted in an insufficient spatial resolution. To detect spatial dynamics in a statistical framework as used here, the dataset would need to cover all relevant fishing areas with multiple fishers to allow for disentangling area-specific trends from fisher effects and other variation. In contrast, a regional separation of trends was detected in the commercial landing data. This was possible by utilizing the desired sampling properties described by Woll et al (2006), as landings data encompass the entire commercial fleet and include therefore a large number of fishers and observed landings per area. Based on the LPUE indices, a decline in the southern areas, Skagerrak and Vestlandet, and an increase in Lofoten & Vesterålen were observed, while Møre and Helgeland showed no clear trend.

Formal assessment of data-limited stocks in general and crustaceans specifically (Smith and Addison, 2003), require explicit adjustment to available methods, plenty of assumptions, and reliable biological inputs. As a result, the simplest methods are often the ones performing best when biological input data are uncertain and assumption after assumption is violated (Cousido-Rocha et al., 2022). The simplest method applied here, LBI, indicated stable trends with a mean length below target levels, except for both sexes in Møre and males in Helgeland. Furthermore, it showed that all the areas were below the target for the 5 % largest, the absence of largest individuals in the catches, implies a length truncation, a known signal of accumulated fishing mortality (Miethe and Dobby, 2015). This is not in line with the knowledge of the fishery that suggests relevant fishing pressure but no overfishing. The method's sensitivity for the input value, L_{inf}, was tested, and supports the concern of high sensitivity towards the input values. With exception of the indicator mentioned above, the three assessment methods (the mean LBI, LBSPR, and SPiCT) are conclusive of a good state and stable trend in Møre (Table 3).

Table 3 Overview of the trends derived from the area-specific LPUE standardization (LPUE), and the three assessment methods: (i) Length-based indicator (LBI), (ii) Length-based spawning potential ratio (LBSPR) and (iii) Surplus production model in continuous time (SPiCT). The assessment is conducted based on data from two time series ranging from 2001-2021 and 2005-2021.

Area	LPUE	LBI	LBSPR	SPICT
Lofoten & Vesterålen	Increasing	Below target	Not assessed	Below B_{MSY}
Helgeland	Stable	Within targets for mean length (males), but below target for max 5%	Not assessed	Stable, above B_{MSY} and below F_{msy}
Møre	Stable	Within targets for mean length, but below target for max 5%	Female – 0.4> SPR >0.2	Stable, above B_{MSY} and below F_{msy}
Vestlandet	Declining	Below targets	Not assessed	Decline, and close to B_{MSY} since 2016
Skagerrak	Declining	No length data	Not assessed	Above B_{MSY} and below F_{MSY}

4.1 Fishing for data - cross-validating the available data

Information on the state of the stock can be derived from fishery- and/or fishery-independent data. Fishery-independent surveys are particularly useful as they provide high-quality data through standardized procedures (Pennino et al., 2016), however, they are costly and therefore comparatively few stocks are monitored with targeted surveys on an annual basis, typically the most valuable commercial stocks. These surveys can generate useful information on non-target species, as has been shown also for brown crab (Mesquita et al., 2021) but because such surveys are not designed for the species of interest, the coverage is often limited spatially or in terms of survey gear and methods, introducing spatial and temporal biases. On the other side, fishery-dependent data are often readily available and cover the distribution of the fishery (which may or may not be identical with the distribution of the targeted species), and despite lower data quality they are essential in obtaining required data for formal assessment. Ideally, stock assessment therefore integrates fisheries-dependent and fisheries-independent data, including cross-validation of the two data sources. Such cross-validation can be used to evaluate how representative and robust estimated indices are and can reveal fishery effects and biases in the fishery-dependent data.

When passive gear is used, catch rates are affected by a variety of biological (behaviour, mobility), ecological (food availability), and environmental parameters (temperature, currents) (Green et al., 2014), and may only provide useful abundance estimates when adequate habitats are targeted (Haggarty and King, 2006). Because the gill- and fyke-net survey is actively avoiding suitable habitats for crabs, it limits drastically the applicability of that time series as an adequate proxy for the size of the crab stock, despite our results indicating comparable trends in catch rates across gear types. Consequently, the gill- and fyke-net survey in its current form cannot be recommended as a monitoring tool for brown crab.

The fishery-independent potting survey aimed to validate the fishery-dependent reference fishery in one of the most important crab fishing areas along the coast. The reference fishery was successfully cross-validated in Møre, showing that the observed catch rates in the reference data are not different from those in a randomized survey design for the surveyed areas. To ensure comparability and reproducibility in the survey,

attempts were made to keep catchability constant so the survey catch rates can be used as a proxy for species abundance (Staby et al., 2020). This was done by selecting the stations randomly, using standardized gear that is identical with the one used in the reference fishery, and correcting for soak time and depth. The results presented here confirm therefore that the reference data is sufficiently representative to inform about changes in the stock. In this study, soak time and depth effects were also accounted for when comparing the two data sets, in contrast to other studies that have cross-validated similar data sources by only comparing catch rates directly against each other (Starr and Vignaux, 1997).

However, despite confirming the representativeness of the reference data through cross-validation with a fisheries-independent survey, its use for assessment purposes was found to be limited. After 20- years of the reference fishery program, it is, therefore, time for an evaluation of the data collection program that has been increasingly challenging to maintain due to its reliance on volunteers. Volunteer logbook programs require sufficient dedicated participants, and therefore the sampling frequency from the reference fishery was reduced to a biennial cycle in 2015 to increase the participation rate and quality of the data. Despite the aim of increasing the participation rate, only six had delivered data in time for this study (May 2022). This shows that the combination of reduced sampling frequency and decreasing participation has led to a situation where the data quantity has become too limited to provide a reliable indicator of stock status. Consequently, area-specific information often depends entirely on a single fisher, making it difficult to disentangle whether observed differences are area or fisher effects. After the fourth year of the reference program, Woll et al. (2006) stated that smaller samples from a higher number of fishers, are the only way to reduce the variance of the mean CW in all regions. Brown crab is a slow-growing animal, and this "low" temporal resolution might be adequate to cover potential temporal changes in the size distribution. However, if there is only one fisher present in each assessment units, it is impossible to separate area and fisher effects, and subsequently fisher-dependent variation might constitute the main influential factor causing spatial variance. I, therefore, support the advice from Woll et al. (2006) and recommend a larger number of samples from more fishers to improve the representativeness of the data, rather than an independent reference fleet. Because information on catch rates will be increasingly available through electronic logbooks, the sampling design in the future should focus more on collecting information on individuals

(e.g. length, sex, molting stage and berried females) that are as good and representative as possible.

4.2 How much – standardizing catch and landing data

Standardized time series of commercial catch rates are a key component of many stock assessments, providing indices of relative changes in abundance to complement time series of total landings. Assessment methods based on catch only (i.e. that do not require any abundance indices) exist but have been found to produce imprecise estimates and are not considered reliable, especially for under-exploited stocks (Free et al., 2020). Even though it is often difficult to translate CPUE or LPUE indices into actual stock abundance (Harley et al., 2001), they provide important information about stock dynamics if properly standardized (Campbell, 2004). Standardization is conditional on sufficient information in the commercial data to fully account for effort and technological parameters affecting catch efficiency. This is currently lacking for a large segment of the Norwegian fleet as boats below 15 m have not been required to report their fishing activity in high-resolution logbooks but only as landing slips without effort information.

However, indices in the shellfish fishery can provide useful information even with limited effort information. As mentioned in the *Nephrops* study, the day-to-day variation in effort (i.e. pots and soak time) is limited as the number of pots set each day is relatively constant for each specific fisher (Zimmermann et al., 2022), in contrast to e.g. trawl fisheries where effort can vary drastically among fishing trips. This relative consistency in fishing effort and behaviour makes it possible to absorb a substantial proportion of individual effort by a vessel/fisher random effect. This assumption holds for relatively short time series as used here (2005 – 2021) but may break down over longer periods due to technological creep. Gear type is known to alter the catchability (Addison and Lovewell, 1991), and increased catch efficiency over time caused by the development of new technology in the gear results in substantial bias if not properly accounted for, as shown in a similar fishery for European lobster in Norway (Kleiven et al., 2022). Additionally, other fishery-specific factors can alter the catchability, such as bait (Chapman and Smith, 1978), and soak time (Bennett and Brown, 1979). They were both accounted for by using standardized gear and including soak time as a covariate in the standardization of CPUE indices.

As concluded in the section above, the landings data are more applicable as area-specific indices for abundance than the CPUE data from the reference fishery. However, high resolution data for effort is lacking, as the landing data is limited to the weight of crabs per boat trip. For the future application, instead of using the landing data, electronic logbooks could provide higher resolution of the effort data after the implementation of the mandatory reporting duty system for all boats. From 1st of July 2022, all boats between 11 and 15 m are obliged to report position and all fishing activity including a continuous report of catches and soak time. This will be implemented for boats between 10 and 10.99 m and below 10 m stepwise, respectively by 1st March 2023 and 1st of January 2024. The catch of each individual operation should be recorded, including catches in passive gears, however, the reporting of catches from passive gear will have a break-in period and not be enforced during the first period of the implementation. (Fiskeridirektoratet, 2021). Increased reporting duty is a good step towards lifting the brown crab out of the category of data-limited stocks. However, including correct information on soak time and number of pots in the electronic logbook system that is being phased in for boats below 15 m will be crucial to model these effects explicitly and facilitate standardization of CPUE indices.

The results here showed very similar trends for the area component as well as the national trend for the CPUE and LPUE indices, confirming through cross-validation that the two time series are likely indicative for changes in the population. However, the uncertainty is large, and questions remain. Even though the CPUE and LPUE indicate no major changes (Figure 7) throughout the time series in Norway, this needs to be evaluated with care and is not necessarily proof of a healthy stock. For instance, the link between catch rates and changes in abundance is not clear and may be biased by hyperstability in the stocks (Erisman et al., 2011; Ward et al., 2013). Hyperstability is the phenomenon where the catch rates remain high or stable while the actual abundance is declining (Crecco and Overholtz, 1990; Hilborn and Walters, 2013), for instance through spatial contraction of a fishery towards areas with the highest density, schooling effects or – as in this case here – because of passive gear that attracts catches from a potentially large area. Furthermore, aggregating trends to a national level may average out local trends and potentially mask area-specific fluctuations or declines. Significant changes in the LPUE index were found when estimating time-varying trends for the different assessment units. Because the model with area-specific trends performed better, it is likely that there are

different trends in the different areas that could not be captured in the CPUE standardization due to limited sample size and spatial resolution.

A lot of stock assessments only exist in the grey literature, and to date, there has only been conducted one formal assessment on brown crab in Europe (Mesquita et al., 2021). They produced an abundance index based on crab catches in a dredge and trawl time series in the North Sea, where brown crab is a frequent by-catch. After a stable increase since 2008, an abrupt decline was observed from 2016 in both time series. Assuming the area-specific trends in LPUE reveal an accurate picture of the stock status, the declining trend observed in Vestlandet is concerning and may hint at a decline similar to what has been observed in other European countries. Another exploratory assessment from Scotland (2013-2015) indicated that most assessment units in Scotland were fished close to or above F_{msy}. They concluded therefore that the current fishing pressure should be lowered to obtain the optimal yield (Mesquita et al., 2017). In France, a decline in national and regional CPUE indices has been observed since 2016 (ICES, 2021b). The lack of further studies to compare with the results presented here emphasizes the major knowledge gap on the state and trend of the various brown crab stocks. My thesis, despite being exploratory, contributes therefore to filling some of these gaps and provides a basis for improving the monitoring of brown crab in Norway.

4.3 Stock assessment

Applying the three selected assessment methods (LBI, LBSPR and SPiCT) to the currently available data from the reference fishers and commercial landings showed mixed results, both in terms of stock status and model performance. As the only area assessed with all methods, Møre will be the basis for further discussion. Despite being data-limited methods, all approaches used ran into challenges with assessing the data-limited brown crab stock of Norway. Specifically, the generic assessment methods and their assumptions were often not fully compatible with the specific characteristics of the fishery and the available data. This underlines the need for individual adjustments and high expertise of the species and the data when conducting these analyses, as stated by Dowling et al. (2019).

When looking at all the three assessment methods and the LPUE index, it is unambiguous that the trends in Møre over the period with available data have been stable, providing little contrast. Although a stable status is good news for management and fishers, this is likely the cause of suboptimal performance in all methods. Specifically, the need for contrast is one of the main drawbacks of surplus production models (Hilborn and Walters, 1992), as they ideally require time series that cover a wide range of biomass levels for the model to gauge accurately the correlated parameters for carrying capacity and population growth. They perform therefore better for a time series that includes phases of heavy exploitation, compared to a stock that has only been lightly exploited, as observed for the crab stock in Møre. LBSPR has a tendency of underestimating harvest rates (Pons et al., 2020), because it depends on a signal in the length composition of the largest individuals to estimate the fishing mortality (Hordyk et al. 2015). This is supported by Mesquita et al. (2021), as they found a decline in the abundance estimate of the crabs in the North Sea that was not reflected in the length composition of the population. Consequently, a clear signal in the length composition caused by fishing may only be visible when the stock is heavily depleted. In long-lived species, the size structures and dynamics are less likely to change drastically over time, masking a decline when not using a sufficiently long time series. The time perspective of 20 years available for this study might not be adequate, taking into account that the brown crab is a long-lived species.

Length-based assessment methods are sensitive to biological input data (Rudd et al., 2019), and LBI and LBSPR are no exceptions. The latter is specifically sensitive to underestimates of the L_{inf} (Hordyk et al 2015). In a recent study by Kell et al. (2022), the performance of LBI in different scenarios was evaluated, and it was found to be robust towards life-history parameters. When comparing it to other models, Cousido-Rocha et al. (2022) suggested LBI to be more adequate than LBSPR when important biological input data (e.g. L_{inf} and M) were uncertain. Regarding the input data of brown crab in Norway, as well as the Northeast Atlantic, the L_{inf} input is uncertain. In this study, I used the same values for both sexes and areas, with an L_{inf} of 220 mm as used in an LBI study in Scotland (Mesquita et al., 2017). Due to limited scope of this study and limited information from the Norwegian stock, only published life-history estimates of brown crab from other areas were used despite geographical inconsistency. For the future application of length-based methods to the Norwegian brown crab, producing stock-

specific parameter estimates, a systematic approach to selecting the life history estimates and a thorough sensitivity analysis should be considered.

To conduct length-based stock assessment effectively it is crucial to collect size distribution data at a suitable spatial and temporal scale, as the sex, size, and moult composition vary throughout the year, and the sex composition is linked to which areas are targeted (Tully et al., 2002). Woll et al. (2006) found increased rates of females in exposed off-coast areas in Norway. This might be because females migrate to deeper waters when they are ready to spawn (October-February), while males tend to be more stationary in shallow waters (Bennett and Brown, 1983; Karlsson and Christiansen, 1996). Additionally, an off-coast migration pattern with smaller crabs close to the shore and larger crabs in deeper waters has been found (Brown and Bennett, 1980). Since the brown crab fishery takes place in the second half of the year, the catch composition could be influenced by season, a factor that needs to be accounted for. Additionally, the length composition could vary in space and, thus, not be captured by the current coarse spatial resolution (statistical fishing areas and location).

In a recent and extensive review of different catch- and length-based assessment methods, Length-based integrated mixed effect (LIME; Rudd and Thorson, 2018) was found to perform better than LBSPR for long-lived species and stocks with moderately depletion levels (Chong et al., 2019; Pons et al., 2020). A major benefit of the LIME is that it is not equilibrium-based and could therefore be a good fit for the crab data. Simulation testing of methods before applying them is good practice (Cope, 2009) but was outside the scope of this study. The applied length-based assessment methods were therefore selected *a priori*, solely based on the ICES guide. It is therefore highly recommended to test alternative methods such as LIME in the future. Nevertheless, it is likely that the LBSPR method has the potential to serve as an assessment method for brown crab, but the challenges remain, including due to the inflexible selectivity definition (logistic "trawl-type" selectivity). However, the gear selectivity of pots restricted by opening funnels as well as regulation through MLS may require a more flexible selectivity function, and implementation of dome-shaped selectivity as a feature in the LBSPR-package would contribute to an increased applicability of this method (Dowling et al., 2019)

Unreported catches from crabs caught as by-catch or in the recreational fishery result in underestimated fishery mortality rates (Lodge et al., 2007). In the similar fishery for European lobster, recreational fishers accounted for 65% of the catches in the study area (Kleiven et al., 2012) , and similar numbers are not unlikely for the brown crab. Currently, the input for the biomass model is restricted to index and catches from commercial landing, which is likely to underestimate MSY and B_{msy} (Omori et al., 2016). However, the same study, emphasised that surplus production models would still provide adequate estimates of trends in the relative biomass and fishing mortality as long as the catch and effort in the unreported fishery remain stable. Interpreting results of SPiCT, when only the commercial landings are included, should be done with caution. If the true impacts of fisheries are to be evaluated the total extraction of the brown crabs is essential. Zimmermann et al. (2020) highlights recreational fishery and unreported bait used in the wrasse fishery as the sources to substantial extraction of brown crab in Norway. Future studies should aim to include and explore these data to obtain a robust knowledge-based basis for stock status and trends.

4.4 Fumbling in the dark - Implication for future sustainable management

The brown crab stock seems to be uniform in terms of phenotypic and genetic composition across borders on a national and a regional scale (Ungfors, 2007; Bakke et al., 2018). Despite this, spatially uneven ecosystem dynamics, regulation, and fishing pressure (Woll et al., 2006), may result in different abundances and population dynamics. Several aspects of the crab biology and population structure are not known, and therefore separation of assessment and management units cannot be drawn solely based on biology. The boundaries of assessment units in this study were based on statistical fishing areas that are equal for all species caught, and were chosen partly based on other studies on crab in Norway (Woll et al., 2006; Bakke et al., 2018) and by exploratory plotting of population composition traits. The specific border remain, however, largely arbitrary, and the units may be inappropriate to fully capture regional dynamics. The same trends were observed in Møre and Helgeland, providing an argument for merging the two assessment units. The generally smaller sizes of the crabs in Vestlandet indicates that this area should remain a separate management unit, however due to a lack of length information it was

not possible to evaluate a potential southern limit of this unit and whether it can be separated from the Skagerrak area.

The Norwegian coast stretches from temperate to arctic climate zones, and ecosystem composition, coastal topography, and resource exploitation vary considerable from south to north. Species inhabiting northern Atlantic waters are prone to changes in productivity caused by climate change, as the North Atlantic is the fastest-changing water body. Whereas the majority of cold-water species in their southern distribution are likely to suffer negatively from warming waters, the same species may benefit at the northernmost fringe (Kjesbu et al., 2021). Because brown crab is a temperate species and Norway represents the northern part of its range, the species is prone to be affected positively by higher temperatures, although other effects of climate change such as ocean acidification may counteract these increases. This study found a tendency of a slight increase in the relative abundance in Lofoten & Vesterålen from 2015, but the data from this area is limited to few fishers and the increase could be caused by other drivers such as economic dynamics of the fishery. As a small-scale coastal fishery, crab fishers are highly dependent on landing and processing facilities in close vicinity to their fishing grounds, as it is economically not sustainable if catches need to be transported over large distances. Changes in the availability of processing plants for crab can therefore have strong effects on the fishing activity in an area. The recent increase in Lofoten & Vesterålen is largely the product of a major dip in the catches between 2008-2015, which results in SPiCT in stock estimates below target values (B_{msy}). In absence of fishing-independent data, commercial data from a relatively newly developed and highly variable fishery such as found in the northernmost (and, possibly, southernmost) areas is likely very uncertain and may provide biased signals, hence a precautionary approach should be taken. This is compounded by a potentially higher vulnerability of the most northern crabs, since moulting frequencies are less frequent at higher latitudes, meaning they are older when they reach maturity (Bakke et al., 2018). Slower growth and later maturity decrease the productivity and resilience, and subsequently the northern population more prone to overexploitation.

5 Conclusion and recommendation

The evaluation of data and assessment methods conducted in this master thesis confirmed that information on stock status and trends can be derived from available time series, yet their limitations restrict the suitable assessment methods and result in a largely inconclusive picture on a national scale. Trends emerge when dividing the stock into specific stock areas, highlighting the importance of the spatial scale in establishing a formal stock assessment and arguing for separate management units. While most assessment units showed largely stable trends above target levels, Vestlandet was found to be the only area where the population might be on the limit of sustainable exploitation. The relative abundance showed a negative trend and length-based indicators of a healthy stock are below targets. Whether this is due to insufficient biological input data estimating the reference points or a truncation in the size distribution caused by many years of fishing is still unknown, and a precautionary approach should be emphasized in further management of that area.

The applied length-based and catch-based assessment methods did not perform satisfactorily and could likely not (yet) be accepted as stock assessments. These issues are common when stocks are not partly or fully depleted and contrast over the available time series is lacking (Pons et al., 2020). The reference fishery provides useful length-information but it is spatially limited and suffers from declining and variable participation that severely limits it utility to a national index with high uncertainty. The fishery scientist's dream of an annual fishery-independent survey would be optimal to produce representative, reliable data, but is likely not feasible for this stock because it's too time-consuming and resource intensive. A recommended and straight-forward way to improve the data availability is therefore to implement mandatory digital logbooks for vessels smaller than 11 meters, including essential effort data (e.g. number of pots). This is scheduled to occur in the course of the next two years, and will likely lower the uncertainties and facilitate the estimation of area-specific CPUE indices, providing a better indicator of changes in catch rates and abundance. It will, however, be important to be able to link these logbook data with the existing time series to maintain a longer-term perspective on the stock dynamics. Furthermore, the mandatory reporting of logbooks calls for a re-thinking of the reference fishery, possibly focusing more on the

collection of representative individual data. As emphasized throughout the thesis, fishing pressure and changes in the stock may be detectable through length structures in the stock. Although we were not able to find any clear changes in the length compositions in the available time series, it is still a relatively cost- and time-efficient data collection method that may provide useful information to monitor stock status. It is important to emphasize that all recommendations mentioned above remain limited as long as the (as of now, unknown) unreported removals from recreational fishery, by-catches (wrasse fishers) and ghost fishing are not known and included in an assessment.

Based on my analysis, I make the following recommendations to close crucial data and knowledge gaps and move towards establishing an assessment framework for brown crab in Norway:

1. As the commercial fleet is present in all areas, better information from the commercial fleet would improve the reliability of the LPUE index as an adequate proxy for abundance. This will be partly achieved by the implementation of mandatory logbook systems for all commercial boats. This measure will be introduced step-wise from 2022-2024. An integration of the new logbook data with existing CPUE and LPUE indices should be explored to estimate a stock index that is as consistent as possible in the future. However, the mandatory reporting relies on properly reporting the relevant effort data. This could then be applied to the existing CPUE and LPUE indices and extrapolated back in time.

2. Even though the reference time series has served as an indispensable source for information on catch rates and composition, the constantly decreasing number of participants in combination with lower frequency of data collection makes the result more and more unreliable. The results presented here call for a re-evaluation of the procedure and the future use of the reference data.

3. The crab stock in Norway is presumably one stock considering the low genetic and phenotypic variation, but exploitation rates vary spatially. There is a need for appropriate separation of the stock into assessment/management units by other criteria than solely based on practicalities represented by the statistical areas. Further research of the spatial differences of the stocks and appropriate assessment units as well as the data requirements for spatially explicit assessment are therefore needed.

4. To cross validate the fishery-dependent data, there is a need for fishery-independent surveys. Since the gill- and fyke-net survey was found to be inadequate for estimating crab abundance due to active avoidance of suitable crab habitat, a survey with a more adequate design is therefore needed. Cross-validation of commercial fishery data in a few commercially important areas could be sufficient when comparing catch rates, as shown with the survey. However, the validity of length structures obtained from reference data has not been confirmed.

5. The information on catches and LPUE are limited to officially reported commercial landings. Because there is a substantial unreported outtake of crabs, from recreational fishers, unreported crabs used as bait in the wrasse fishery and crabs caught in lost gears (ghost fishing), the currently available data covers only a fraction of the total landings and may therefore bias any assessment. Specifically, the lack of complete landings data limits the performance of surplus production models, and thereby our understanding of the state of the stock and possibility to give explicit advice for management purposes.

6. Even though all the applied methods indicate stable trends in all of the assessment units except of Vestlandet, a precautionary principle calling for a cautious approach to management in the absence of clear indicator should be considered due to the current data limitations and high uncertainty. Declining trends and comparatively small mean size in the area of Vestlandet, mirroring patterns of a decline in other stocks in the North Sea, should be given special attention in management of the stock and calls for a future fishery-independent survey. Especially when taking signals of a decline in other stocks in the North Sea.

7. As long as it is not feasible to manage the stock based on through output regulation (i.e. a total allowable catch), alternative management approaches (e.g. input control through gear restrictions) and conservation measures (MPAs, preventing ghost fishing) should be considered. Subsequently, there is a need for further research on the impact of ghost fishing and potential benefits of MPAs. The latter may be complicated by ecological interactions, especially with European lobster that can outcompete brown crab.

6 References

Addison, J., and Lovewell, S. 1991. Size composition and pot selectivity in the lobster (*Homarus gammarus*) and crab (*Cancer pagurus*) fisheries on the east coast of England. ICES Journal of Marine Science, 48: 79-90.

Akaike, H. 1998. Information theory and an extension of the maximum likelihood principle. *In* Selected papers of hirotugu akaike, pp. 199-213. Springer.

Bakke, S., Buhl-Mortensen, L., and Buhl-Mortensen, P. 2019. Some observations of *Cancer pagurus* Linnaeus, 1758 (Decapoda, Brachyura) in deep water. Crustaceana, 92: 95-105.

Bakke, S., Larssen, W. E., Woll, A. K., Søvik, G., Gundersen, A. C., Hvingel, C., and Nilssen, E. M. 2018. Size at maturity and molting probability across latitude in female *Cancer pagurus*. Fisheries Research, 205: 43-51.

Begg, G. A., Friedland, K. D., and Pearce, J. B. 1999. Stock identification and its role in stock assessment and fisheries management: an overview. Fisheries Research, 43: 1-8.

Bennett, D. B., and Brown, C. G. 1979. The problems of pot immersion time in recording and analyzing catch-effort data from a trap fishery [Crustaceae, crabs lobster] Rapports et Proces-Verbaux des Reunions, 175: 186-189.

Bennett, D. B., and Brown, C. G. 1983. Crab (*Cancer pagurus*) migrations in the English Channel. Journal of the Marine Biological Association of the United Kingdom, 63: 371-398.

Bentley, N. 2015. Data and time poverty in fisheries estimation: potential approaches and solutions. ICES Journal of Marine Science, 72: 186-193.

Bentley, N., and Stokes, K. 2009. Moving fisheries from data-poor to data-sufficient: evaluating the costs of management versus the benefits of management. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science, 1: 378-390.

Berkeley, S. A., Hixon, M. A., Larson, R. J., and Love, M. S. 2004. Fisheries sustainability via protection of age structure and spatial distribution of fish populations. Fisheries, 29: 23-32.

Beverton, R. J. H., and Holt, S. J. 1957. On the dynamics of exploited fish populations. Fisheries Investigations Series II. Marine Fisheries, Great Britain Ministry of Agriculture, Fisheries and Food, 19.

Beverton, R. J. H., and Holt, S. J. 2012. On the dynamics of exploited fish populations, Springer Science & Business Media.

Brooks, M., Kristensen, K., van Bentham, K., Magnusson, A., Berg, C., Nielsen, A., Skaug, H., et al. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal of Animal Ecology, 8: 378-400.

Brown, C. G., and Bennett, D. B. 1980. Population and catch structure of the edible crab (*Cancer pagurus*) in the English Channel. ICES Journal of Marine Science, 39: 88-100.

Burnham, K. P., and Anderson, D. R. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research, 33: 261-304.

Campbell, R. A. 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fisheries Research, 70: 209-227.

Chapman, C. 1994. Assessments on Crab and Lobster (Scotland). FRS Marine Laboratory, Aberdeen, UK.

Chapman, C., and Smith, G. 1978. Creel catches of crab, *Cancer pagurus* using different baits. ICES Journal of Marine Science, 38: 226-229.

Chong, L., Mildenberger, T. K., Rudd, M. B., Taylor, M. H., Cope, J. M., Branch, T. A., Wolff, M., et al. 2019. Performance evaluation of data-limited, length-based stock assessment methods. ICES Journal of Marine Science, 77: 97-108.

Cope, J. M. 2009. Issues and advances in data-limited stock assessment: experimentation through simulation. University of Washington.

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

Cousido-Rocha, M., Cerviño, S., Alonso-Fernández, A., Gil, J., Herraiz, I. G., Rincón, M. M., Ramos, F., et al. 2022. Applying length-based assessment methods to fishery resources in the Bay of Biscay and Iberian Coast ecoregion: Stock status and parameter sensitivity. Fisheries Research, 248: 106197.

Crecco, V., and Overholtz, W. J. 1990. Causes of density-dependent catchability for Georges Bank haddock *Melanogrammus aeglefinus*. Canadian Journal of Fisheries and Aquatic Sciences, 47: 385-394.

d'Acoz, U. 1999. Inventaire et distribution des crustacés décapodes de l'Atlantique nord-oriental, de la Méditerranée et des eaux continentales adjacentes au nord de 25° N. Patrimoines Naturels (Muséum National d'Histoire Naturelle, Service du Patrimoine Naturel, Paris), 40: i-x+ 1-383.

Dichmont, C. M., Punt, A. E., Dowling, N., De Oliveira, J. A., Little, L. R., Sporcic, M., Fulton, E., et al. 2016. Is risk consistent across tier-based harvest control rule management systems? A comparison of four case-studies. Fish and Fisheries, 17: 731-747.

Dowling, N. A., Smith, A. D. M., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury, K., Wilson, J. R., et al. 2019. Generic solutions for data-limited fishery assessments are not so simple. Fish and Fisheries, 20: 174-188.

Dunn, D. C., Boustany, A. M., and Halpin, P. N. 2011. Spatio-temporal management of fisheries to reduce by-catch and increase fishing selectivity. Fish and Fisheries, 12: 110-119.

Erisman, B. E., Allen, L. G., Claisse, J. T., Pondella, D. J., Miller, E. F., and Murray, J. H. 2011. The illusion of plenty: hyperstability masks collapses in two recreational fisheries that target fish spawning aggregations. Canadian Journal of Fisheries and Aquatic Sciences, 68: 1705-1716.

FAO 2019. FAO yearbook. Fishery and Aquaculture Statistics 2019. Retrived from: http://www.fao.org/fishery/species/2627/en [11.05.2022].

Fiskeridirektoratet 2021. Utvidelse av rapporteringsplikten for alle fiskefartøy. Retrieved from: https://www.fiskeridir.no/Yrkesfiske/Rapportering-paa-havet/utvidelse-av-rapporteringsplikten-for-alle-fiskefartoy [14.05.2022].

Fiskeridirektoratet 2022. Fiskeridirektoratets landings- og sluttseddelregister.

Free, C. M., Jensen, O. P., Anderson, S. C., Gutierrez, N. L., Kleisner, K. M., Longo, C., Minto, C., et al. 2020. Blood from a stone: performance of catch-only methods in estimating stock biomass status. Fisheries Research, 223: 105452.

Garcia, S. M., Kolding, J., Rice, J., Rochet, M.-J., Zhou, S., Arimoto, T., Beyer, J., et al. 2012. Reconsidering the consequences of selective fisheries. science, 335: 1045-1047.

Green, B. S., Gardner, C., Hochmuth, J. D., and Linnane, A. 2014. Environmental effects on fished lobsters and crabs. Reviews in Fish Biology and Fisheries, 24: 613-638.

Haggarty, D. R., and King, J. R. 2006. CPUE as an index of relative abundance for nearshore reef fishes. Fisheries Research, 81: 89-93.

Haig, J. A., Bakke, S., Bell, M. C., Bloor, I. S., Cohen, M., Coleman, M., Dignan, S., et al. 2016. Reproductive traits and factors affecting the size at maturity of *Cancer pagurus* across Northern Europe. ICES Journal of Marine Science, 73: 2572-2585.

Halliday, R., and Pinhorn, A. 2002. A review of the scientific and technical bases for policies on the capture of small fish in North Atlantic groundfish fisheries. Fisheries Research, 57: 211-222.

Harley, S. J., Myers, R. A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Canadian Journal of Fisheries and Aquatic Sciences, 58: 1760-1772.

Hartig, F. 2022. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.5. https://CRAN.R-project.org/package=DHARMa.

Hilborn, R. 2011. Future directions in ecosystem based fisheries management: a personal perspective. Fisheries Research, 108: 235-239.

Hilborn, R., Amoroso, R. O., Anderson, C. M., Baum, J. K., Branch, T. A., Costello, C., De Moor, C. L., et al. 2020. Effective fisheries management instrumental in improving fish stock status. Proceedings of the National Academy of Sciences, 117: 2218-2224.

Hilborn, R., and Walters, C. J. 1992. Stock and recruitment. *In* Quantitative Fisheries Stock Assessment, pp. 241-296. Springer.

Hilborn, R., and Walters, C. J. 2013. Quantitative fisheries stock assessment: choice, dynamics and uncertainty, Springer Science & Business Media.

Hixon, M. A., Johnson, D. W., and Sogard, S. M. 2014. BOFFFFs: on the importance of conserving old-growth age structure in fishery populations. ICES Journal of Marine Science, 71: 2171-2185.

Honma, M. 1974. Estimation of overall effective fishing intensity of tuna longline fishery: Yellowfin tuna in the Atlantic Ocean as an example of seasonally fluctuating stocks. Far Seas Fisheries Research Laboratory Bulletin (Japan).

Hordyk, A. 2021. LBSPR: Length-Based Spawning Potential Ratio. R package version 0.1.6. https://CRAN.R-project.org/package=LBSPR.

Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015. 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, 72: 217-231.

Hordyk, A. R., Ono, K., Prince, J. D., and Walters, C. J. 2016. A simple length-structured model based on life history ratios and incorporating size-dependent selectivity: application to spawning potential ratios for data-poor stocks. Canadian Journal of Fisheries and Aquatic Sciences, 73: 1787-1799.

Høstingsforskriften 2021. Forskrift om gjennomføring av fiske, fangst og høsting av viltlevende marine ressurser. *(*FOR-2021-12-23-3910). Lovdata, Retrieved from: https://lovdata.no/forskrift/2021-12-23-3910/§29 [15.05.2022].

ICES 2012. ICES Implementation of Advice for Data-limited Stocks in 2012 in its 2012 Advice. ICES Document CM 2012/ACOM:, 68: 42.

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).

ICES 2021a. Advice on fishing opportunities. In Report of the ICES Advisory Committee, 2021. ICES Advice 2021, section 1.1.1. https://doi.org/10.17895/ices.advice.7720.

ICES 2021b. Working Group on the Biology and Life History of Crabs (WGCRAB; outputs from 2019 meeting). ICES Scientific Reports, 3: 68.

Jardim, E., Azevedo, M., and Brites, N. M. 2015. Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. Fisheries Research, 171: 12-19.

Karlsson, K., and Christiansen, M. F. 1996. Occurrence and population composition of the edible crab (*Cancer pagurus*) on rocky shores of an islet on the south coast of Norway. Sarsia, 81: 307-314.

Kell, L. T., Minto, C., and Gerritsen, H. D. 2022. Evaluation of the skill of length-based indicators to identify stock status and trends. ICES Journal of Marine Science.

Kell, L. T., Sharma, R., Kitakado, T., Winker, H., Mosqueira, I., Cardinale, M., and Fu, D. 2021. Validation of stock assessment methods: is it me or my model talking? ICES Journal of Marine Science, 78: 2244-2255.

Kelly, C. J., and Codling, E. A. 2006. 'Cheap and dirty' fisheries science and management in the North Atlantic. Fisheries Research, 79: 233-238.

Kennelly, S. J., and Broadhurst, M. K. 2002. By-catch begone: changes in the philosophy of fishing technology. Fish and Fisheries, 3: 340-355.

Kjesbu, O. S., Sundby, S., Sandø, A. B., Alix, M., Hjøllo, S. S., Tiedemann, M., Skern-Mauritzen, M., et al. 2021. Highly mixed impacts of near-future climate change on stock productivity proxies in the North East Atlantic. Fish and Fisheries.

Klaoudatos, D., Conides, A., Anastasopoulou, A., and Dulčić, J. 2013. Age, growth, mortality, and sex ratio of the inshore population of the edible crab, *Cancer pagurus* (Linnaeus 1758) in South Wales (UK). Journal of Applied Ichthyology, 29: 579-586.

Kleiven, A. R., Espeland, S. H., Stiansen, S., Ono, K., Zimmermann, F., and Olsen, E. M. 2022. Technological creep masks continued decline in a lobster (*Homarus gammarus*) fishery over a century. Scientific Reports, 12: 3318.

Kleiven, A. R., Olsen, E. M., and Vølstad, J. H. 2012. Total catch of a red-listed marine species is an order of magnitude higher than official data. PLoS One, 7: e31216.

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), 1-21. doi:10.18637/jss.v070.i05.

Larkin, P. A. 1977. An epitaph for the concept of maximum sustained yield. Transactions of the American Fisheries Society, 106: 1-11.

Linnane, A., McGarvey, R., Hoare, M., and Hawthorne, P. 2013. The importance of conserving recruitment pulses in rock lobster (*Jasus edwardsii*) fisheries where puerulus settlement is low or highly sporadic. Marine Biology Research, 9: 97-103.

Lodge, M. W., Anderson, D., Løbach, T., Munro, G., Sainsbury, K., and Willock, A. 2007. Recommended Best Practices for Regional Fisheries Management Organizations: Report of an independent panel to develop a model for improved governance by Regional Fisheries Management Organizations, Chatham House.

Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., Kidwell, S. M., et al. 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. science, 312: 1806-1809.

Maunder, M. N., and Punt, A. E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research, 70: 141-159.

Mesquita, C., Dobby, H., Pierce, G. J., Jones, C. S., and Fernandes, P. G. 2021. Abundance and spatial distribution of brown crab (*Cancer pagurus*) from fishery-independent dredge and trawl surveys in the North Sea. ICES Journal of Marine Science, 78: 597-610.

Mesquita, C., Miethe, T., Dobby, H., and Mclay, A. 2017. Crab and lobster fisheries in Scotland: results of stock assessments 2013–2015. Scottish Marine and Freshwater Science, 8: 1990-1991.

Miethe, T., and Dobby, H. 2015. Selection of length-based indicators for shellfish stocks and fisheries. Working Document, ICES report WKLIFE V 2015 ICES CM 2015/ACOM: 56.

Miethe, T., and Dobby, H. 2016. Testing length-based indicators in harvest control rules (HCR) for shellfish stocks and fisheries. Working Document in Report of the ICES workshop on the development of quantitative assessment methodologies based on life-history traits, exploitation characteristics and other relevant parameters for data-limited stocks category 3-6 (WKLIFE VI), 3-7 October 2016 Lisbon, Portugal.

Moland, E., Olsen, E. M., and Stenseth, N. C. 2010. Maternal influences on offspring size variation and viability in wild European lobster *Homarus gammarus*. Marine Ecology progress series, 400: 165-173.

Neal, K., and Wilson, E. 2008. *Cancer pagurus*. Edible crab.

Nelder, J. A., and Wedderburn, R. W. 1972. Generalized linear models. Journal of the Royal Statistical Society: Series A (General), 135: 370-384.

Omori, K. L., Hoenig, J. M., Luehring, M. A., and Baier-Lockhart, K. 2016. Effects of underestimating catch and effort on surplus production models. Fisheries Research, 183: 138-145.

Ovando, D., Hilborn, R., Monnahan, C., Rudd, M., Sharma, R., Thorson, J. T., Rousseau, Y., et al. 2021. Improving estimates of the state of global fisheries depends on better data. Fish and Fisheries, 22: 1377-1391.

Palumbi, S. R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. Ecological applications, 13: 146-158.

Pedersen, M. W., and Berg, C. W. 2017. A stochastic surplus production model in continuous time. Fish and Fisheries, 18: 226-243.

Pennino, M. G., Conesa, D., López-Quílez, A., Muñoz, F., Fernández, A., and Bellido, J. M. 2016. Fishery-dependent and -independent data lead to consistent estimations of essential habitats. ICES Journal of Marine Science, 73: 2302-2310.

Pinheiro, J., and Bates, D. 2006. Mixed-effects models in S and S-PLUS*, Springer science & business media.

Pons, M., Cope, J. M., and Kell, L. T. 2020. Comparing performance of catch-based and length-based stock assessment methods in data-limited fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 77: 1026-1037.

Prince, J., Hordyk, A., Valencia, S. R., Loneragan, N., and Sainsbury, K. 2015. Revisiting the concept of Beverton–Holt life-history invariants with the aim of informing data-poor fisheries assessment. ICES Journal of Marine Science, 72: 194-203.

Probst, W. N., Kloppmann, M., and Kraus, G. 2013. Indicator-based status assessment of commercial fish species in the North Sea according to the EU Marine Strategy Framework Directive (MSFD). ICES Journal of Marine Science, 70: 694-706.

Punt, A., Smith, D., Thomson, R., Haddon, M., He, X., and Lyle, J. 2001. Stock assessment of the blue grenadier *Macruronus novaezelandiae* resource off south-eastern Australia. Marine and Freshwater Research, 52: 701-717.

Punt, A. E., Huang, T., and Maunder, M. N. 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. ICES Journal of Marine Science, 70: 16-33.

R Core Team 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rochet, M.-J., and Trenkel, V. M. 2003. Which community indicators can measure the impact of fishing? A review and proposals. Canadian Journal of Fisheries and Aquatic Sciences, 60: 86-99.

Rudd, M. B., and Thorson, J. T. 2018. Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1019-1035.

Rudd, M. B., Thorson, J. T., and Sagarese, S. R. 2019. Ensemble models for data-poor assessment: accounting for uncertainty in life-history information. ICES Journal of Marine Science, 76: 870-883.

SeaLifeBase 2022. *In* World Wide Web electronic publication.Retrieved from: www.sealifebase.org, version (04/2022). Ed. by M. Palomares, and D. Pauly.

Sheehy, M., and Prior, A. 2005. Analysis of stock age structure and population parameters in edible crab, *Cancer pagurus*, using lipofuscin age-pigment: data for resource management. Marine Fisheries Research and Development Final Rep. MF0225. Department of Environmental and Rural Affairs, London, UK.

Sheridan, M., Officer, R., O'Connor, I., and Lordan, C. 2015. Investigating the Feasibility of using Growth Increments for Age Determination of Norway Lobster (*Nephrops norvegicus*) and Brown Crab (*Cancer pagurus*). Journal of Crustacean Biology, 35: 495-498.

Skajaa, K., Fernö, A., Løkkeborg, S., and Haugland, E. K. 1998. Basic movement pattern and chemo-oriented search towards baited pots in edible crab (*Cancer pagurus*). Hydrobiologia, 371: 143-153.

Smith, M. T., and Addison, J. T. 2003. Methods for stock assessment of crustacean fisheries. Fisheries Research, 65: 231-256.

Staby, A., Aglen, A., Gjøsæter, H., and Fall, J. 2020. Akustisk mengdemåling av sei og kysttorsk i Finnmark-Møre høsten 2020

Starr, P. J., and Vignaux, M. 1997. Comparison of data from voluntary logbook and research catch-sampling programmes in the New Zealand lobster fishery. Marine and Freshwater Research, 48: 1075-1080.

Statistisk Sentralbyrå 2022. Fiskeri (avslutta i Statistisk sentralbyrå). Retrived from: https://www.ssb.no/statbank/table/05463/tableViewLayout1/ [05.02.2022].

Søvik, G., Jenssen, M., Hjelset, A. M., and Krogness, C. 2017. Ressursundersøkelse av taskekrabbe langs norskekysten. Rapport fra fangstregistreringer i 2001-2015. Rapport fra havforskningen.

Tallack, S. M. L. 2007. Size–fecundity relationships for *Cancer pagurus* and Necora puber in the Shetland Islands, Scotland: how is reproductive capacity facilitated? Journal of the Marine Biological Association of the United Kingdom, 87: 507-515.

Trenkel, V. M., Rochet, M.-J., and Mesnil, B. 2007. From model-based prescriptive advice to indicator-based interactive advice. ICES Journal of Marine Science, 64: 768-774.

Tully, O., Robinson, M., Addison, J., Bell, M., Eaton, D., Smith, M., Elson, J., et al. 2002. Collection and evaluation of assessment data for the European edible crab (*Cancer pagurus*) stocks. Final report for the European Commission. Contract: 234.

Ungfors, A. 2007. Sexual maturity of the edible crab (*Cancer pagurus*) in the Skagerrak and the Kattegat, based on reproductive and morphometric characters. ICES Journal of Marine Science, 64: 318-327.

Ungfors, A. 2008. Fisheries Biology of the Edible Crab (*Cancer pagurus*) in the Kattegat and the Skagerrak: Implications for Sustainable Management*,* Department of Marine Ecology, University of Gothenburg.

Ungfors, A., McKeown, N. J., Shaw, P. W., and André, C. 2009. Lack of spatial genetic variation in the edible crab (*Cancer pagurus*) in the Kattegat–Skagerrak area. ICES Journal of Marine Science, 66: 462-469.

Verdoit, M., Pelletier, D., and Bellail, R. 2003. Are commercial logbook and scientific CPUE data useful for characterizing the spatial and seasonal distribution of exploited populations? The case of the Celtic Sea whiting. Aquatic Living Resources, 16: 467-485.

Von Bertalanffy, L. 1938. A quantitative theory of organic growth (inquiries on growth laws. II). Human biology, 10: 181-213.

Walsh, M. R., Munch, S. B., Chiba, S., and Conover, D. O. 2006. Maladaptive changes in multiple traits caused by fishing: impediments to population recovery. Ecology letters, 9: 142-148.

Ward, H. G., Askey, P. J., and Post, J. R. 2013. A mechanistic understanding of hyperstability in catch per unit effort and density-dependent catchability in a multistock recreational fishery. Canadian Journal of Fisheries and Aquatic Sciences, 70: 1542-1550.

Woll, A. K., Fossen, I., van der Meeren, G., and Tveite, S. 2003. Preliminary results from a resource study of edible crab (*Cancer pagurus*) in Norway 2001-2003. ICES.

Woll, A. K., van der Meeren, G. I., and Fossen, I. 2006. Spatial variation in abundance and catch composition of *Cancer pagurus* in Norwegian waters: biological reasoning and implications for assessment. ICES Journal of Marine Science, 63: 421-433.

Wood, S. 2017. Package 'mgcv'. R package version, 1: 729.

Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., Fogarty, M. J., et al. 2009. Rebuilding global fisheries. science, 325: 578-585.

Zhou, S., Smith, A. D., Punt, A. E., Richardson, A. J., Gibbs, M., Fulton, E. A., Pascoe, S., et al. 2010. Ecosystem-based fisheries management requires a change to the selective fishing philosophy. Proceedings of the National Academy of Sciences, 107: 9485-9489.

Zimmermann, F., and Heino, M. 2013. Is size-dependent pricing prevalent in fisheries? The case of Norwegian demersal and pelagic fisheries. ICES Journal of Marine Science, 70: 1389-1395.

Zimmermann, F., Jenssen, M., Nedreaas, K. H., Søvik, G., Hjelset, A. M., and Bakke, S. 2020. Kunnskapsgrunnlaget for taskekrabbe langs norskekysten. Rapport fra havforskningen.

Zimmermann, F., Kleiven, A. R., Ottesen, M. V., and Søvik, G. 2022. Inclusion of recreational fishing in data-limited stocks: a case study on Norway lobster (*Nephrops norvegicus*) in Norway. Canadian Journal of Fisheries and Aquatic Sciences, 99: 1-10.

Appendix A

Protocol Potting Survey August 2022

Study Area

The most important fishing grounds, where reference fishers were present were targeted, which are illustrated in Figure 3. Inside the study area one to station types were defined as: (i) stations parallel to the IMR gill- and fyke-net survey station and (ii) a control station, randomly dispersed within a 20km radius of the former station. Each station was evaluated upon arrival, to secure the right depth interval and suitable bottom substrate.

Sampling

Each day ten links, where two were placed parallel (spatially and temporally) to the gill- and fyke-net survey and eight were placed around the two former station. To ensure comparable results between the pot survey and the gill- and fyke-net survey, the pots at the parallel stations were set and raised at the same time as the gill/fyke nets. Soak time ranged from 24-48 hours. Each link consisted of six baited pots with 25- 30-meter distance (Figure A.1), identically to the ones used by the reference fishers. The standardized trial pots of the project were moulded black polyethylene, measuring 80 cm × 35 cm around and 31 cm high, with entrances on each short side of the pot (Woll et al., 2006). As a standard bait 0.5 – 1 kg of fresh saithe *(Pollachius virens)* was used per pot. To get a representable picture of the entire length composition, the pots were without escape gaps. Therefore, an application for special fishing approval was sent to the Norwegian Directorate of Fisheries.
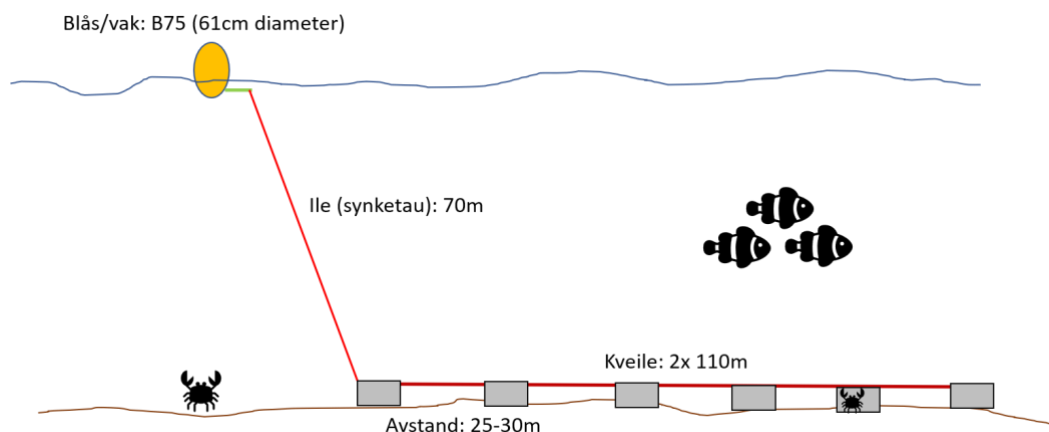

Figure A.1 Schematic illustration of the pot-link setup.

Measurements

All crabs caught were sorted, counted, and measured per station and pot. The carapace width (CW) was measured at the widest part of the carapace, to the nearest millimetre, with an electric calliper (Figure A.2). The sex was determined by visual observation of the width of the abdomen, where females have a wide round-shaped abdomen, and males have a narrow triangle-shapes abdomen. The shell condition was determined by visual inspection of the individual, based on colour, epi-growth, and wear of the shell and dactyls. The criteria used to determine stages were defined as: (1) Clean and soft shell, (2) New, hard shell with no epi-growth, pointy dactyls, no dark spots (3) Hard shell with some epi-growth, slightly rounded dactyls, some dark spots, (4) Hard, dark shell, with a lot of epi-growth and rounded dactyls.

Figure A.2 Measuring the carapace width at the widest part of the carapace

Appendix B

Sensitivity analyses CPUE and LPUE standardization

Sensitivity analyses were conducted on each of the two models (CPUE and LPUE) to detect potential source to noise in the models which could be filtered out to optimize the model without compromising it. The final selected filtration values, that did not affect the model substantially are summarized in Table B.1.
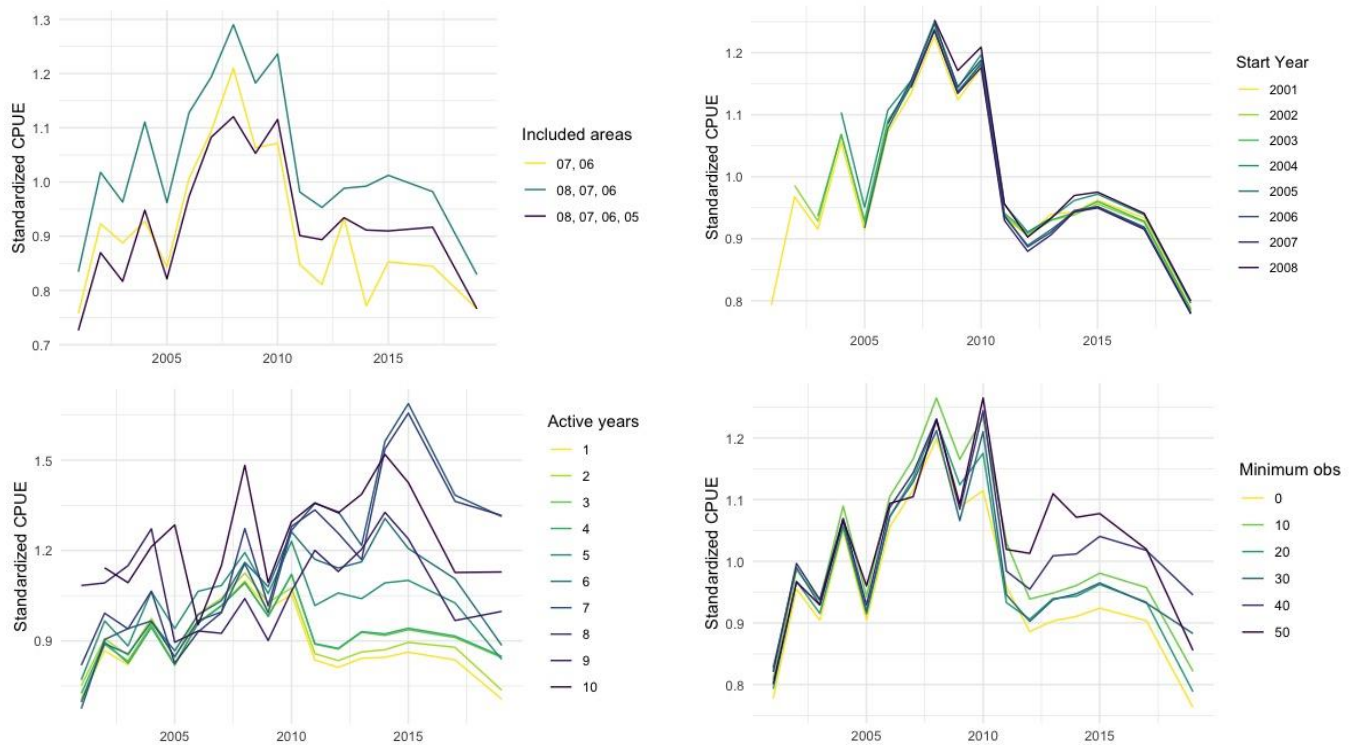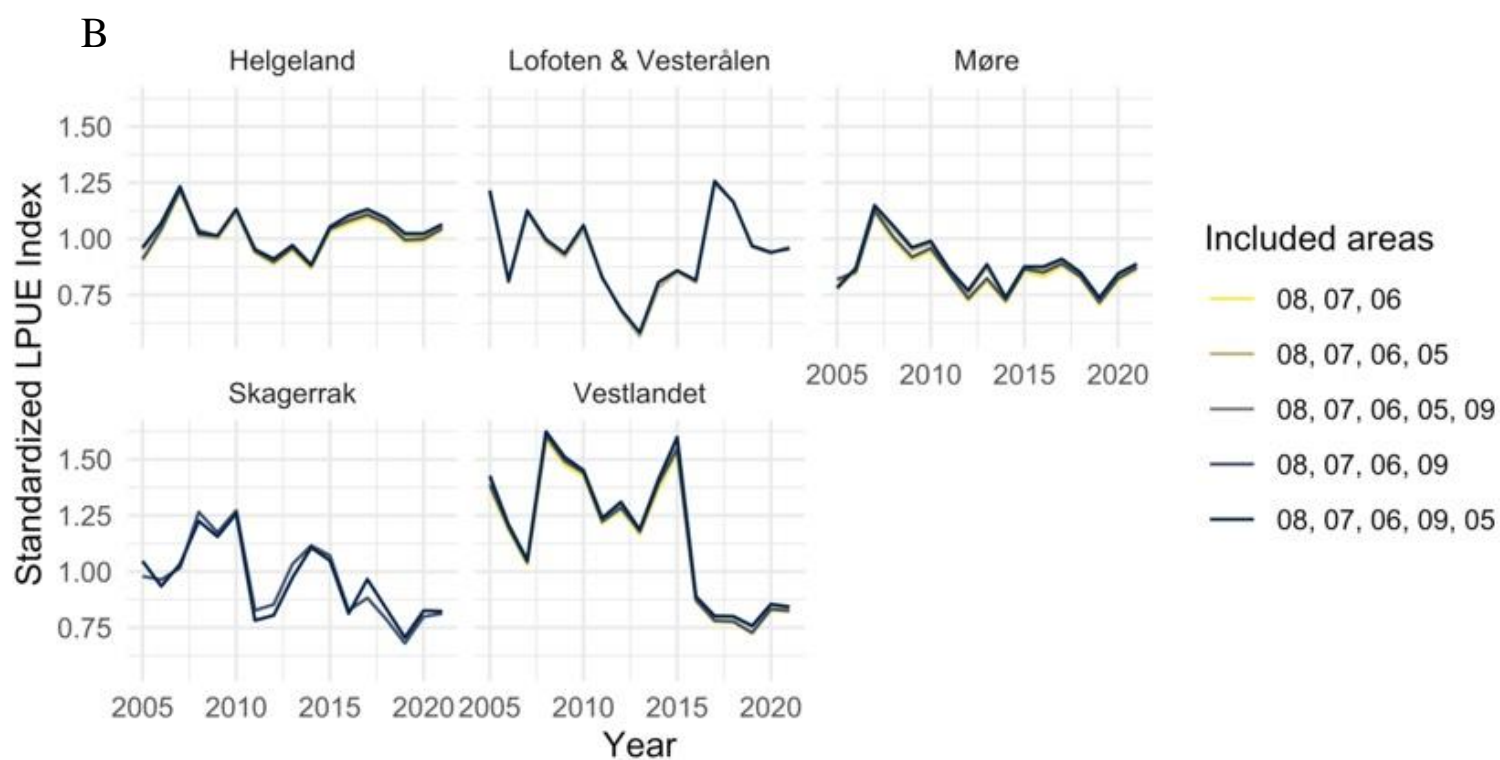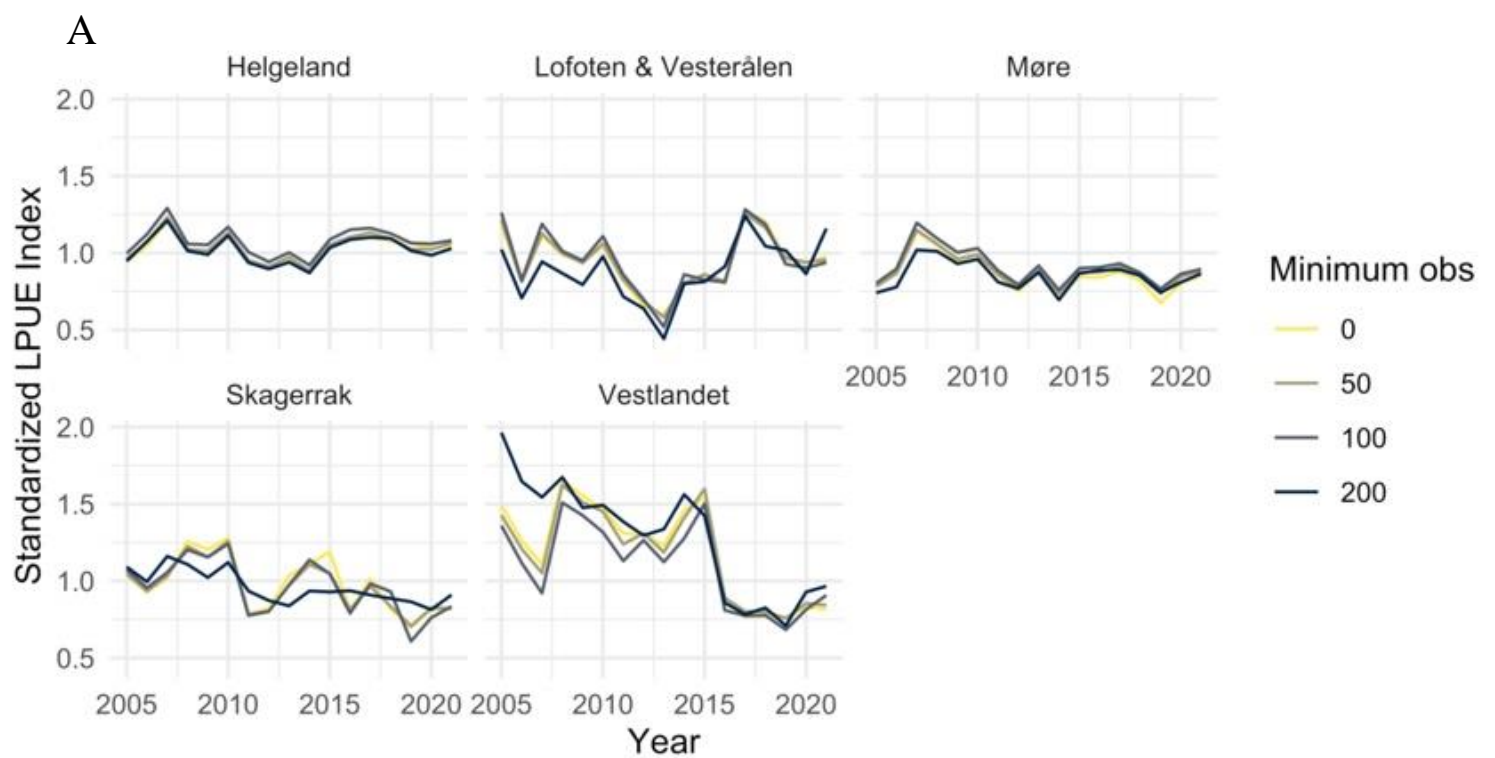

Figure B.1 Sensitivity analysis of the catch-per-unit-effort standardisation (CPUE; E.1). The figure shows changes in the CPUE model according to filtration of the four relevant parameters: (a) areas included, (b) which year the index was calculated from, (c) number of how many years each fishers has been active in the reference fishery, and (d) number of annual observations from each fishers. The colours indicate which values that were tested in each of the four sensitivity analysis.
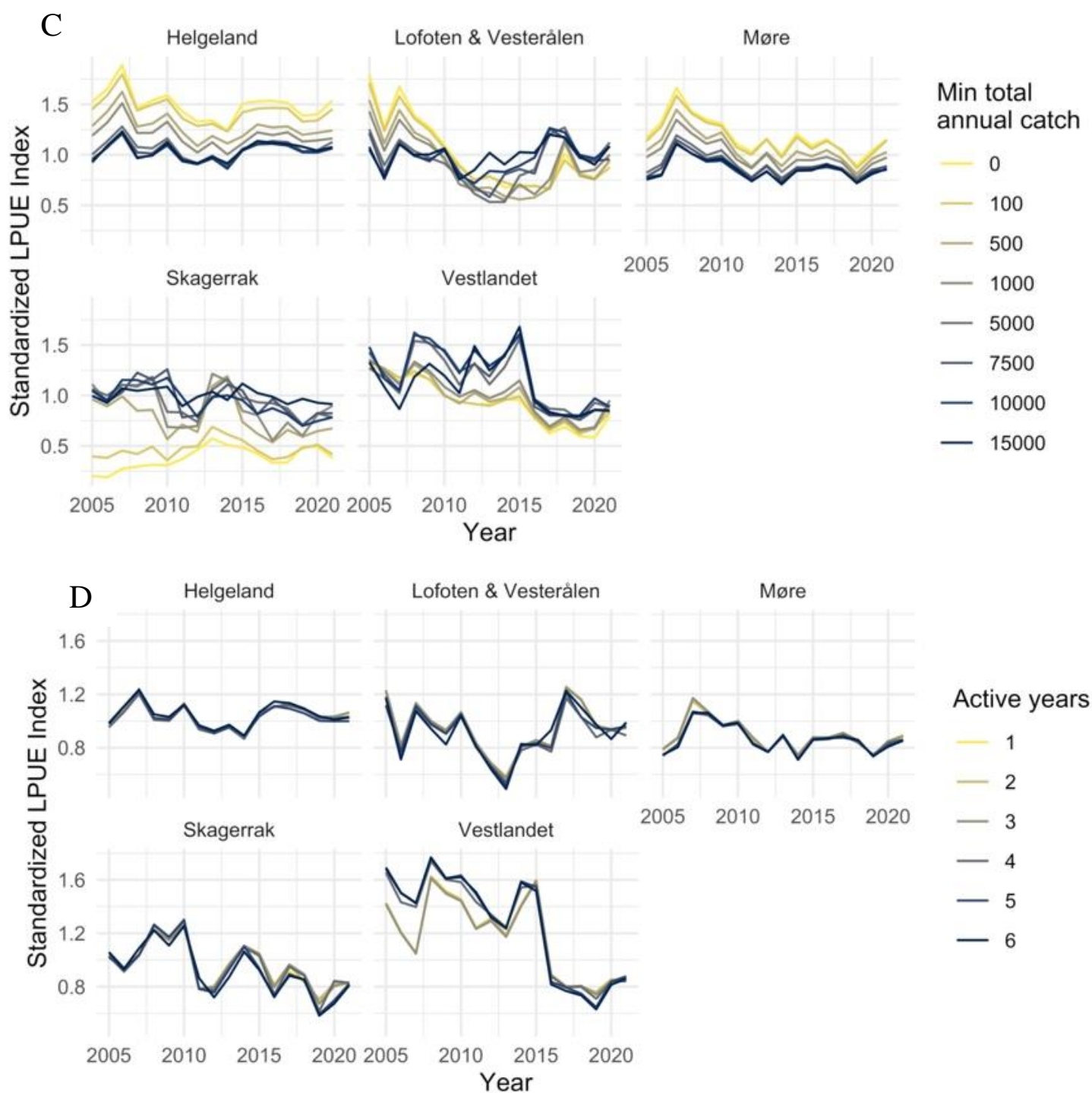
A

B

Figure B.2 Sensitivity analysis of the landing-per-unit-effort standardisation, with area specific trends (LPUE; E.2). The figure shows changes in the LPUE model according to filtration of the four relevant parameters: (A) minimum observations reported form each fisher, (B) included areas, (C) minimum annual landing of each fisher, and (D) number of how many years each fisher has been active in the commercial fishery. The colours indicate which values that were tested in each of the four sensitivity analyses.

Table B.1 Filter and threshold values used in the Catch-per-unit effort index (CPUE) and Landings-per-unit-effort index (LPUE) to balance the data. The values are the result of sensitivity analysis. Numbers of years active refers to number of years the fishers or the vessel has been active in the fishery, for the CPUE and LPUE respectively. Minimum observations indicate the number of observations reported from each fishers each year.

Filter	Threshold	
	CPUE	LPUE
Number of years active	2	2
Minimum observations	20	50
Areas included	Vestlandet (08+28), Møre (07), and Helgeland (06)	Skagerrak (09), Vestlandet (08+28), Møre (07), Helgeland (06), and Lofoten & Vesterålen (00+05)
Upper limit of Catch per boat trip (kg)	-	4550
Lower limit of annual catches (kg)	-	5000

Appendix C
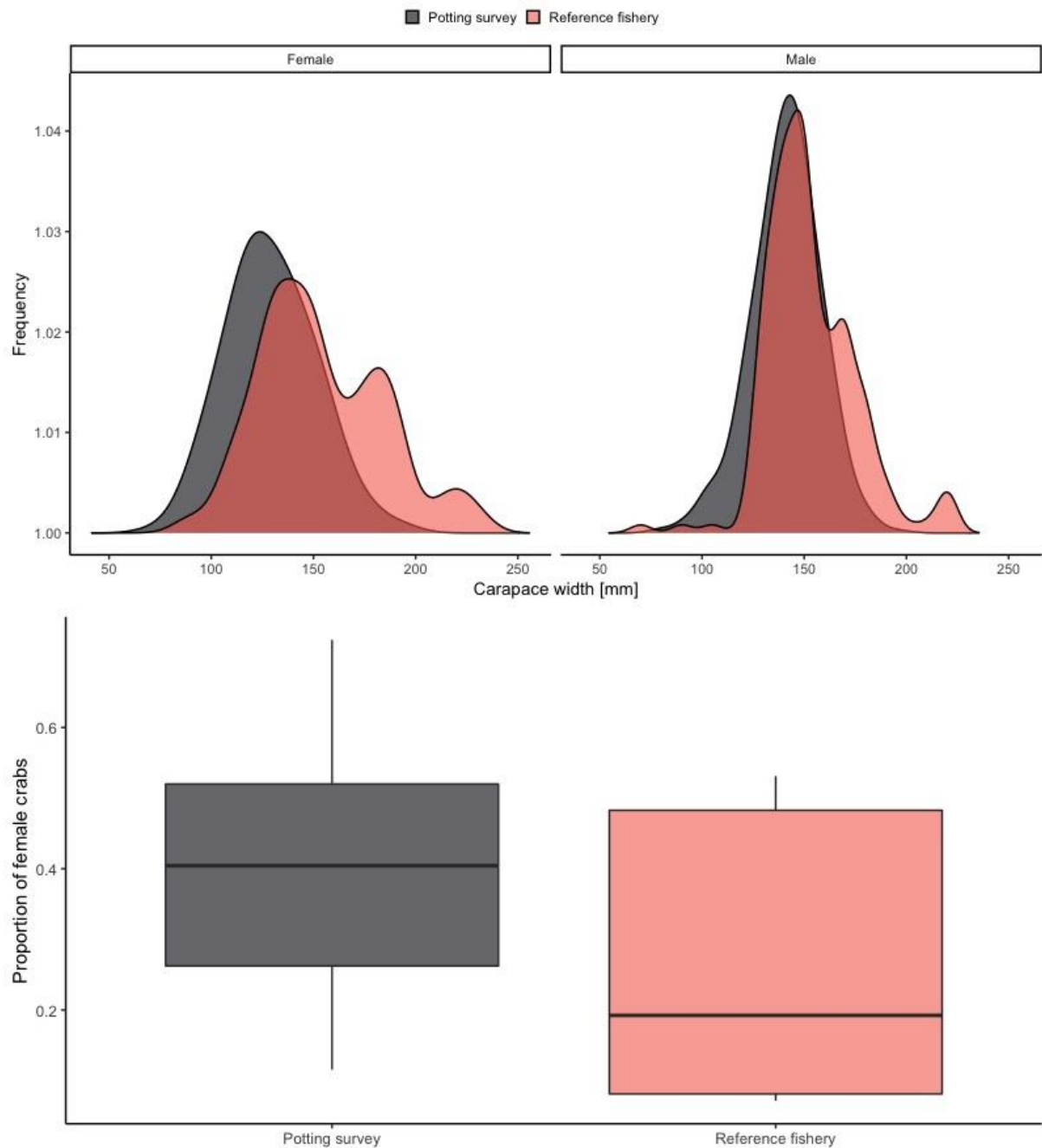
The Potting Survey vs. The Reference Fishery

Figure C.1 Differences in populations structures in the reference fishery compared to the potting survey. Upper panel compares the length distribution observed in the catches and the lower panel shows the sex composition. To ensure that the data are directly comparable the data are exclusively area Møre in August 2021, and the same pots are used in both surveys.

Table C.1 Statistical cross validation of the reference fishery by comparing catch rates with the potting survey. Two separate general additive models (with gaussian distribution) were used for testing whether catch rates in the two cases were significant different: (i) the two data sets, potting survey and reference fishery, and (ii) the two station types used for the survey, one that were parallel to the gill- and fyke-net station and one control station. The modelled corrects for depth and soak time, with a spline is added to two covariates, denoted as 's'.

Catch rates in the two data sets: potting vs. reference				Catch rates in the two station types: control vs parallel			
Formula: Number of crabs per pot ~ Data set + s(Depth) + s(Soak time)				Formula: Number of crabs per pot ~ Station type + s(Depth) + s(Soak time)			
Parametric coefficients							
Parameter	Estimate	SE	p	Parameter	Estimate	SE	p
Intercept	5.108	1.825	0.010	Intercept	4.616	0.877	<0.001
Data set: Potting survey	2.084	2.032	0.308	Station: Control	2.357	0.985	0.029
Smoothing terms							
Parameter	edf	F	p	Parameter	edf	F	p
Depth	1.128	3.599	0.073	Depth	1	3.883	0.053
Soak time	1.201	0.110	0.762	Soak time	1	1.986	0.169

R^2 adj: 0.214, n=72, Deviance explained: 6.7% R^2 adj:0.100, n=79, Deviance explained: 13.5%

Appendix D

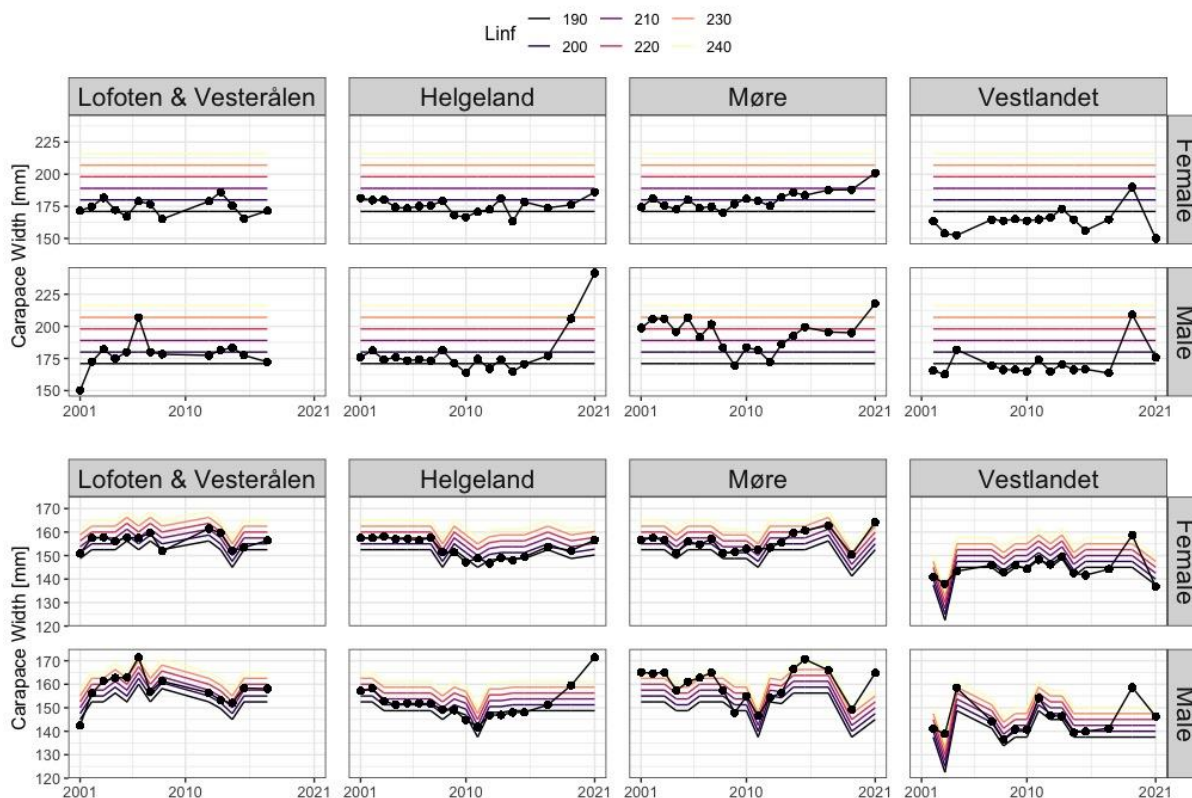## Sensitivity Analyses Length-Based Indicators

Figure D.1 Sensitivity analysis of the asymptotic length (L_{inf}) value, used as an input for the two reference points: 90% of the L_{inf} (upper panel) and length at were fishery mortality equals natural mortality (lower panel). The colours indicate the different L_{inf} values, ranging from 190-240, which is consistent with the span of Linf values in the literature.

Appendix E

Sensitivity Analysis for Surplus Production in Continuous Time
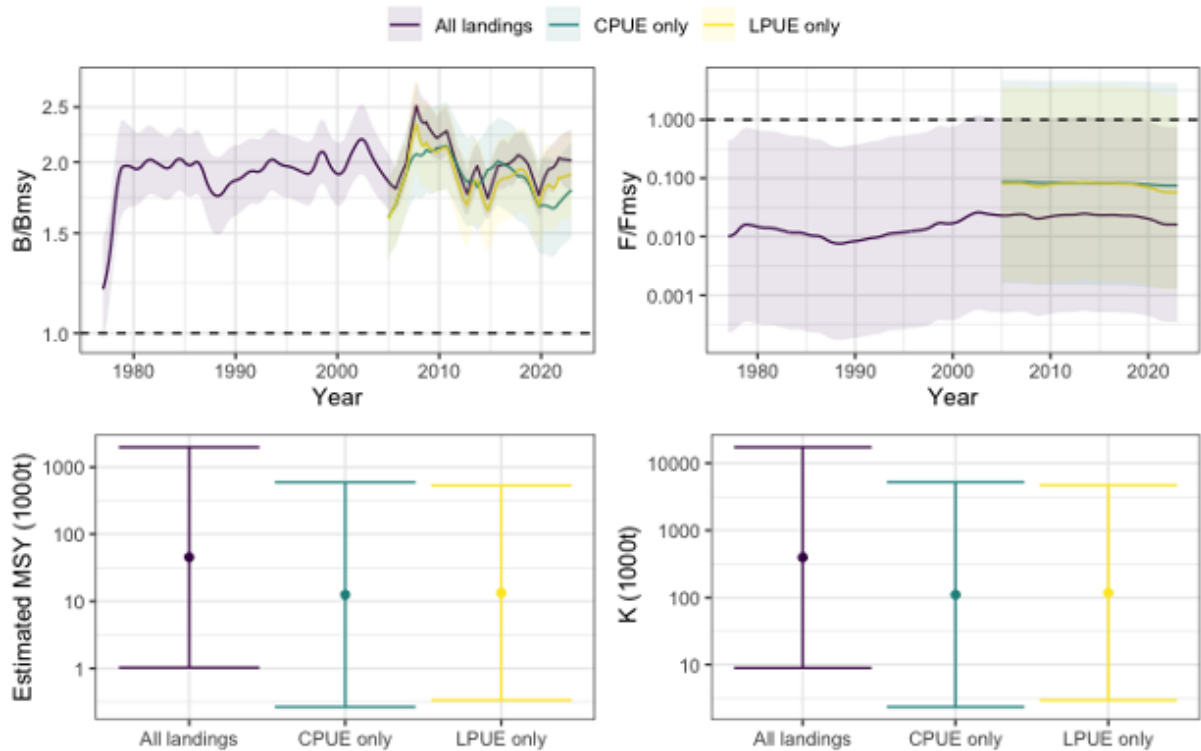

Figure E.1 Sensitivity analysis of the input data regarding both catch and index. (i) "All landings" include all available catch data, back to 1970, in combination with the LPUE index, (ii) "CPUE only" include catches from 2005 and the CPUE index, while the (iii) "LPUE only" includes commercial catch data from 2005 in combination with the LPUE index. The latter option was assumed a priori to best fulfill the criterium of the SPiCT model of representing the commercial landings. Since the variation was high in all cases the prior assumption was used in the final model. Note that all plots (except B/$_{Bmsy}$,Year) are log-scaled.
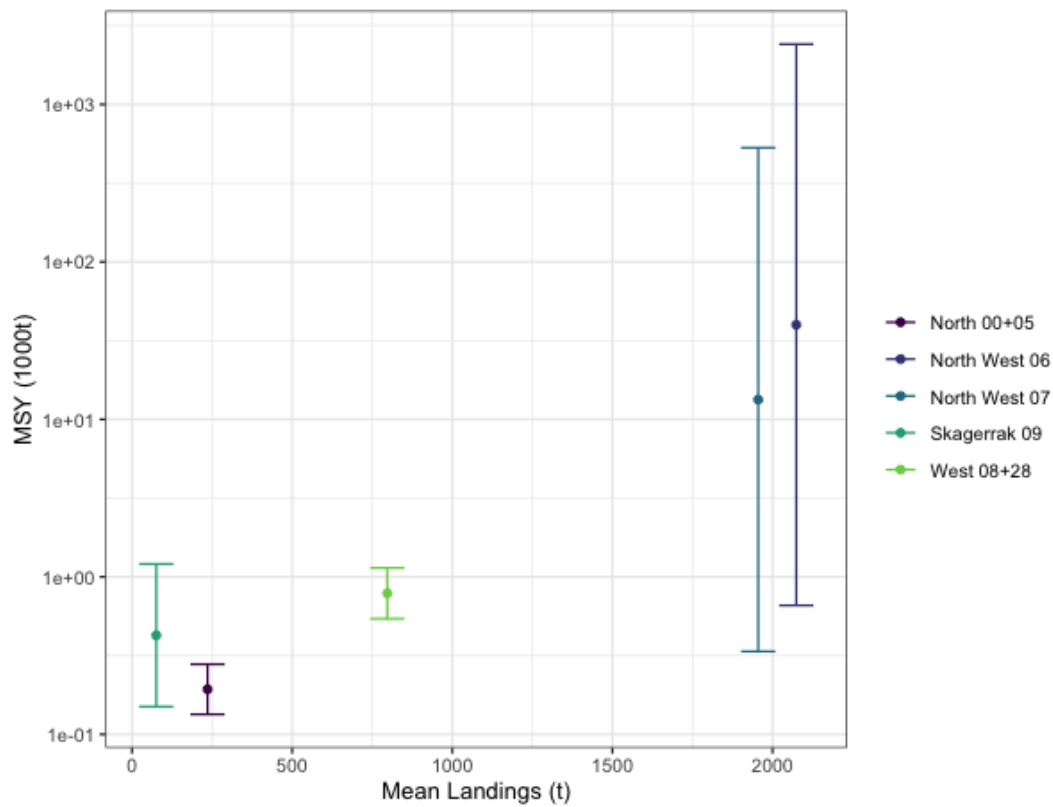
Figure E.2 The relationship between the estimated maximum sustainable yield (MSY), and the mean landings in each of the five assessment units. The plot indicates that the MSY estimates are mainly driven by the landings. MSY is given as 1,000 tonnes at log-scale and the mean landings are the mean annual landings in each are, given in tonnes.
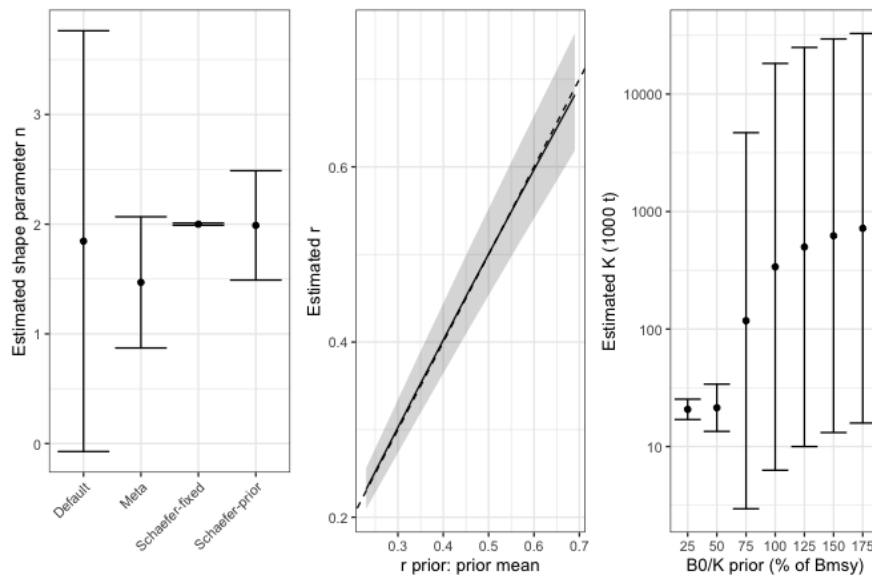

Figure E.3 Sensitivity of the three estimated values for shape (n), instinct growth rate (r), and carrying capacity (K) to the prior input values: shape, instinct growth rate and the ratio between initial biomass and carrying capacity (b_0/k), respectively. A pattern (i.e linear) indicate that the estimated value is dependent on the input value. The figures indicate that the model is specifically sensitive to R input. The b/k input seems to have a tipping point/sensitivity threshold (50-75%). The version of this SPiCT model is run for area Møre, with the landings-per-unit-effort index and landings from 2005-2021 as inputs. The following priors are used b/k = 0.75, n=Schaefer, and r=0.46.