Doktoravhandlinger ved NTNU, 2023:73

Cian Kelly

Assimilation of real-time measurements with an individual-based model for estimation of geographical distribution and abundance of Norwegian herring

Doktoravhandling

NTNU
Norges teknisk-naturvitenskapelige universitet
Avhandling for graden
philosophiae doctor
Fakultet for informasjonsteknologi og
elektroteknikk
Institutt for teknisk kybernetikk

NTNU
Kunnskap for en bedre verden

Cian Kelly

Assimilation of real-time measurements with an individual-based model for estimation of geographical distribution and abundance of Norwegian herring

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Engineering Cybernetics

ISBN 978-82-326-6763-5 (printed ver.)
ISBN 978-82-326-6907-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

IMT-report 2023-07-W

Doctoral theses at NTNU, 2023:73

Printed by NTNU Grafisk senter

Summary

The fishing industry faces both challenges and opportunities in the wake of the green transition in fisheries.[1] Challenges stem from attempts to reduce the carbon footprint of activity while simultaneously providing food security. Opportunities arise where reducing fuel consumption can both drive down operating costs and increase the value of landings, a positive sum result for stakeholders. Technological innovation have reduced costs in the past and likely will in the future. However, innovations have also led to negative effects on fish stocks, and fisheries management is key to mitigating undesirable outcomes. To facilitate both economic and environmentally sustainable harvesting, real-time information about the state of the fish stock is advantageous. Sophisticated modelling approaches can be used to predict ocean resources at high spatial and temporal resolutions, and provide insight into the state of ecosystems. This thesis proposes a model-based estimation approach to predicting real-time geographical distribution and abundance of fish stocks. Providing recurrent predictions of fish distributions has the potential to reduce the time vessels spend searching for suitable fishing grounds and decrease fuel consumption in the fishing industry.

The model-based estimation approach is inspired by work done by Jens Glad Balchen, the founder of the Department of Engineering Cybernetics (ITK) at NTNU. He was an advocate of the power of modelling and estimation theory to predict ocean resources for harvesting. In that vein, we combine a model commonly used in fisheries research, called an Individual-Based Model (IBM), with an estimation procedure familiar in cybernetics, called an Ensemble Kalman Filter (EnKF). We chose the commercially important Norwegian Spring Spawning Herring (NSSH) as the model stock based on discussions with fishers participating in the project. The IBM forecasts the spawning migration of NSSH, with behaviour of individuals driven mainly by the Norwegian coastal current. The model structure was calibrated using survey data and compared with catch data from the Norwegian fishing fleet. After finding a suitable agreement between model output and observations, random disturbances were added to the IBM, accounting for unmodelled phenomena and other uncertainties in the model structure. This includes annual variations in migration patterns, reflected in variations in fishing grounds and search routes for fishing vessels. Incorporating random disturbances produced an ensemble of instances of the NSSH migration, with the variability within the ensemble representing the uncertainty in the model predictions. For compatibility with the EnKF, each realization of the IBM was mapped to a density field forming a state space for the EnKF correction step, and corrections were applied to the field based on

[1] https://www.frontiersin.org/research-topics/41245/green-transition-and-sustainability-in-fisheries

incoming measurements. The IBM was then adjusted to account for the applied corrections. The model predictions were improved using the assimilation procedure, which was demonstrated using a twin model simulation experiment. In addition, we proposed a machine learning method for generating synthetic observations using vessel positional systems as input and normalized fish densities as target output. This increased the number of measurements available to correct the IBM.

Data Assimilation reduces model uncertainty and provides information on unobserved areas with high catch potential during the fishing season. The model-based estimation approach is therefore being integrated into a web-based Decision Support System (DSS) developed during this project. We discuss how fishers have contributed to the development of the system through questionnaires and project meetings. The DSS can both integrate the real-time estimates from modelling work and facilitate systematic data capture by fishing vessels. Furthermore, data from fishing vessels can improve model structure and predictions by providing input to the assimilation procedure. Future work may include more state variables in the IBM, consider alternative representations of the underlying model and estimation procedure or find new sources of measurements. In addition, one may extend the migration model to include the winter migration from open sea to winter stay areas.

Preface

This thesis is submitted in partial fulfilment of the requirements for the degree of Philosophiae Doctor (Ph.D.) at NTNU - the Norwegian University of Science and Technology. The work has been carried out at the Department of Engineering Cybernetics (ITK) from January 2020 to January 2023. My supervisors have been Morten Omholt Alver, Finn Are Michelsen, Jeppe Kolding and Øystein Varpe. The work is part of the FishGuider project, an innovation-based project for the Industrial Sector. Funding has been provided by the Norwegian Research Council (grant number 296321). The remainder of funding and operational costs of the project FishGuider were managed the project partners: NTNU, the North Atlantic Institute for Sustainable Fishing (NAIS), University of Bergen and SINTEF Ocean AS.

NTNU, Trondeim, Norway	24.02.2023
Location	Date

Cian Kelly

Name

Acknowledgements

I am grateful to my supervisors for providing unwavering support and feedback throughout the course of my doctoral studies. I had regular, weekly meetings with my main supervisor at ITK, Morten Omholt Alver, which were important and I don't take for granted, given this valuable time spent discussing and refining the project work. Thoughtful feedback on written material was always provided. Similarly, my co-supervisor Finn Are Michelsen at SINTEF Ocean provided much input into written work and we had regular meetings to keep each other in the loop. I am grateful to co-supervisors Jeppe Kolding and Øystein Varpe at the University of Bergen who provided ideas for project work and feedback on all written material.

Although not supervising directly, I am grateful to the Fishguider project members whom offered much insight. The project administrator Are Prytz Berset lent his knowledge of the fishing industry and ideas for the project work when we held meetings. Karl Johan Reite also provided updates from the operational side of the project. Arnt Inge Nygård at the fishing company Nybonia Hav AS offered perspectives from the fishing industry side. On that note, I thank the fishermen who responded to the questionnaire detailed in this work.

I visited the Danish Technological University in Copenhagen for a research stay during my studies and I'm thankful to those who discussed project work with me, including Stefan Neuenfeldt, Fletcher Thompson, Kjetil Thorvaldsen and Paco Rodriguez.

List of Articles

1. Kelly, C., Michelsen, F. A., Kolding, J., and Alver, M. O. (2022). Tuning and Development of an Individual-Based Model of the Herring Spawning Migration. Frontiers in Marine Science 8, 754476. doi:10.3389/fmars.2021.754476

2. Kelly, C., Michelsen, F. A., and Alver, M. O. (2023). An ensemble modelling approach for spatiotemporally explicit estimation of fish distributions using Data Assimilation. Fisheries Research 261, 106624. doi: https://doi.org/10.1016/j.fishres.2023.106624

3. Kelly, C., Michelsen, F. A., and Alver, M. O. (In prep). Model-based estimation of fish catch potential using assimilation of synthetic measurements from a neural network driven by fishing vessel activity data.[2]

4. Kelly, C., Michelsen, F. A., Reite, K. J., Kolding, J., Varpe, Ø., Berset, A. P., and Alver, M. O. (2022). Capturing big fisheries data: Integrating fishers' knowledge in a web-based decision support tool. Frontiers in Marine Science 9, 1051879. doi:10.3389/fmars.2022.1051879

[2]In preparation in February 2023.

Contents

Abbreviations

AIS Automatic Identification System

ANN Artificial Neural Network

CG Centre of Gravity

CPUE Catch per Unit Effort

DSS Decision Support System

EnKF Ensemble Kalman Filter

GIC Global Index of Collocation

ITK Department of Engineering Cybernetics

NAIS North Atlantic Institute for Sustainable Fishing

NSSH Norwegian Spring Spawning Herring

PDF Probability Density Function

Chapter 1

Introduction

1.1 Importance of the fishing industry in Norway

The modern fishing fleet is of key importance to Norwegian economic activity, particularly in rural coastal regions (Ølmheim 2021). This is unsurprising given the extensive 2500 km of coastline that wraps from the Barents Sea in the north, to the North Sea further south (Elde et al. 2018). In particular, Northern Norway is a disproportionately active community in wild fisheries, with 9% of the Norwegian population, but 38% of wild captures. Approximately 70% of marine traffic in Northern Norway is from fishing fleet activity alone (Stoeva et al. 2022). Along with agriculture, marine fisheries have been traditionally very important in Norway. Archaeological evidence shows dried cod has been traded as a commodity in Lofoten since the 12th century.[1]

In modern times, the export market for Norwegian seafood is much larger than the domestic market, where 95% of Norwegian seafood is consumed abroad (including farmed fish), with Germany being the most important of the 140 international markets (Elde et al. 2018). To meet international demand, both the catch volume and value must be high. For illustration, in 2021, total catch of groundfish and pelagic species was 1 and 1.5 million tonnes, and the total catch value was 14 and 9 billion kroner respectively (Ølmheim 2021). The operating margins of Norwegian groundfish and pelagic fleets have been increasing for decades (see Figure 1.1).

Broadly, the changes to the fishing industry can be described as capital (mainly vessel equipment) gradually replacing labour (Hannesson 2007). Since the 1950s,

[1] See: https://www.skrei.net/exhibit/the-beginning-of-the-stockfish-trade/

the labour force has decreased by 85%, and the value of catch has doubled, with higher returns for larger, better equipped vessels (>28m), while governmental support has all but ceased (Fisheries Directorate 2021, Ølmheim 2021). These long-term trends in the industry are primarily due to the continuous proliferation of new technologies, particularly in the second half of the 21st century (Hamre and Nakken 1971). Introduction of innovations to the industry generates a positive feedback loop, where the adoption of new technologies leads to higher catch values, allowing more capital investment, facilitating the purchase of better equipment, and further increasing catch values (Hamre and Nakken 1971). It is important to note smaller fishers or fisheries that don't have the access to capital to invest in new technologies tend not to reap these benefits to the same extent (Hannesson et al. 2010). Regardless, modern technologies have been instrumental in increasing catch value in the industry.

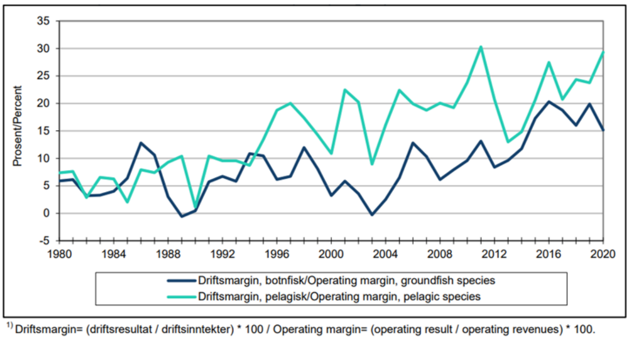


Figure 1.1: Operating margins for Norwegian groundfish and pelagic fleets from 1980 to 2019. Figure courtesy of the Norwegian Directorate of Fisheries (Ølmheim 2021).

1.2 Historical uptake of fishing technologies

The earliest developments in fishing activity began when fish surplus to nutritional requirements was traded. The invention of preservation techniques such as drying, salting and smoking preserved fish against rotting, providing the initial impetus to trade fish as a commodity (Jennings et al. 2001). In Norway, drying has been a historically important preservation technique.[2] Crucially, the development of modern onboard refrigeration and freezing technologies, such as refrigerated sea water, allowed the preservation of high quality fresh catch during long voyages and increased landing prices (Dellacasa 1987, Jennings et al. 2001, Nakken 2008).

The gradual mechanization of the fishing industry also had a significant effect. The first industrial revolution brought the steam engine and with it a boost in towing

[2]See: https://www.skrei.net/exhibit/the-beginning-of-the-stockfish-trade/

power, which increased hauling speeds, the distances travelled for fishing and the size of the gear available (Jennings et al. 2001). Uptake of diesel engines in the early 1900s reduced journey times of boats, while electric lighting allowed fishing at night and in darker winter months (Jennings et al. 2001, Nakken 2008). Mechanized fish pumps increased the volume of catch loaded onboard, while mechanical winches automated the setting and hauling of gear, both reducing manual labour (Jennings et al. 2001). The power block, a type of winch used in purse seining, was introduced in the 1960s allowing gear to be hauled onboard without the assistance of auxiliary vessels, reducing the manpower and time needed for fishing operations significantly[3] (Hamre and Nakken 1971, Gordon and Hannesson 2015). In the 1950s, non-rotting synthetic fibres facilitated the production of larger, pressure resistant fishing nets that trebled catch efficiency in cod fisheries (Gordon and Hannesson 2015). Finally, the total engine power of fishing vessels has increased gradually over time (Fisheries Directorate 2021).

The adaptation of communication technologies has been particularly important in more recent times. The telegraph, the first near real-time communications technology, offered frequent updates on the state of the fishery as early as the 1840s, before real-time weather reports and news of fish migrations became available through the introduction of radio transmitters and receivers to fisheries around the 1930s (Hannesson et al. 2010, Gordon and Hannesson 2015). The implementation of echosounders, specifically the Norwegian SIMRAD system in 1947, allowed detection of mid- and deep-water concentrations of fish and extended the spatial extent for searching and exploiting offshore grounds (Nakken 2008, Gordon and Hannesson 2015). In the 1960s, sonar was adapted to detect large shoals of fish by their characteristic shape through horizontal profiling, without relying on fish being underneath the vessel (Jennings et al. 2001, Gordon and Hannesson 2015). Moreover, both echosounders and sonar improved the actual operation of fishing gears, which were becoming mechanized, reducing the manual labour needed to haul nets (Nakken 2008). In the 1960s radar was introduced as a navigation tool, allowing more advanced planning of operations on a large spatial scale (Gordon and Hannesson 2015). The rapid developments of improved electronic navigation and telecommunication technology quickly became standard in the industrialized fisheries (Gordon and Hannesson 2015).

To conclude this section, the adoption of preservation, gear and communication technologies have had a profound effect on the fishing industry, increasing the length of fishing trips, the value of landings, the quality of those landings and the overall spatial extent of fishing operations (Figure 1.2). Technological development has been driven from the bottom up through individual adoption, but also

[3] See: https://www.fao.org/fishery/en/equipment/powerblock

through collaboration with Norwegian institutions. For example, the development of the SIMRAD echosounder is considered a successful example of strategic collaboration between the Norwegian Defence Institute and the Institute of Marine Research in Bergen (Handegard et al. 2021). Collaborative projects are key to achieving economic and environmental sustainability in wild fisheries in modern times.

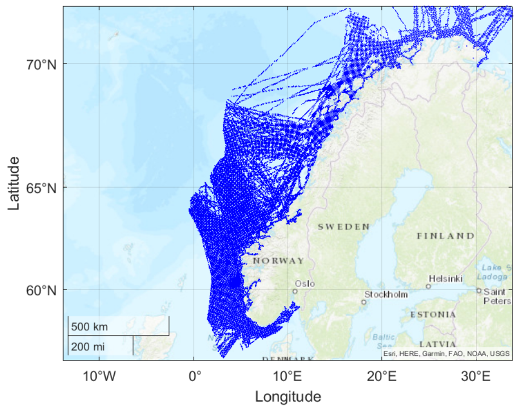


Figure 1.2: Individual positions of 186 vessels targeting Norwegian Spring Spawning Herring from September to December 2020. This illustrates the spatial extent of coastal and offshore fishing operations. Data is courtesy of the Norwegian Coast Guard.[4]

1.3 Sustainable harvesting of marine resources

1.3.1 Fisheries management

The steady march of technology is not all good news for marine fisheries. In the fishing industry, productivity gains from new technology are often countered by the reduction in fish resources (Hannesson et al. 2010). Illustratively, the cumulative impact of technological innovation, especially mechanical hauling, led to increased catch rates and the collapse of the Norwegian herring stock in the 1970s (Fiksen and Slotte 2002, Gordon and Hannesson 2015, Standal and Asche 2018). To create value in a sustainable manner international cooperation combined with government policies to control the cumulative effects of fishing on shared resources are key (Squires 1994, Ølmheim 2021).

To reduce the likelihood of similar collapses (and reduce overcapacity), Norwegian governance has introduced many policy measures since the herring collapse.

[4]https://www.kystverket.no/en/

These include individual vessel quotas, a tradeable quota system, decommissioning schemes, real-time closures of fishing grounds, discard bans and harvest control rules (Squires 1994, Hannesson 2007, Gullestad et al. 2015; 2014, Standal and Asche 2018, Ølmheim 2021). Gullestad et al. (2017) describes the modern Norwegian quota setting system as the annual regulatory cycle. This cycle involves negotiations with other states, regulatory meetings at the Norwegian Directorate of Fisheries and advice from international organisations such as ICES.[5] The Norwegian regulatory cycle is highly regarded internationally, with Norway rating high on compliance to the UN Code of Conduct for Responsible Fisheries (Gullestad et al. 2017). However, uncertainties owing to changes in the dynamics of fish populations and the introduction of new exploitation methods, mean rational precautionary approaches to management will remain important in the future.

1.3.2 Food security

During the green revolution of the 1960s, the breeding of high yield varieties of wheat is estimated to have reduced starvation from 60 to 14 percent between 1960 to 2000 (Swaminathan 2009). Such measures were needed at this time to feed booming populations requiring sustenance. Food security still remains an issue. By the year 2100, the human population is projected to reach approximately 10 billion people (Figure 1.3). Population growth presents a stark challenge for global food supply. For 4.5 billion people, more than 15% of their protein intake comes from fish, making it an essential consideration for food security (Béné et al. 2015). Fish is a a source of fatty acids and micronutrients that are important to brain development and cognition (Beveridge et al. 2013, Béné et al. 2015). Intuitively, the benefits of fish consumption should raise consumer demand in increasingly health conscious societies. To meet the nutritional needs, irrespective of consumer demands, is likely to present a challenge for already heavily exploited global fisheries (Beveridge et al. 2013).

1.3.3 Energy use

Increasing energy prices and volatility in prices are becoming more common due to factors such as declining supply of fossil fuels, geopolitical conflicts and energy intensive human lifestyles (Pelletier et al. 2014). Globally, between 30 and 50% of fisheries costs arise from fuel usage (Parker and Tyedmers 2015). It is estimated that the global average anthropogenic greenhouse gas emissions per tonne of fish landed grew by 21% between 1990 and 2011 (Parker et al. 2018). The IPCC recommends authorities to take "actions that limit global warming to close to 1.5°C would substantially reduce projected losses and damages related to climate change in human systems and ecosystems" (Pörtner et al. 2022). Governmental actions,

[5] International Council for the Exploration of the Sea: https://www.ices.dk/Pages/default.aspx

such as carbon taxes, may nudge the fishing industry to rapidly adapt fuel usage to limit carbon footprint and their own costs.

A life cycle assessment of Norwegian seafood products found that herring frozen and shipped in bulk produced the lowest emissions, producing 0.7 kg CO_2 per kg of edible product (Ziegler et al. 2013). In the same study, the carbon footprint of demersal species was rated higher than that of pelagic species. Other work has shown Norwegian Sea purse seiners as relatively fuel efficient, using approximately 0.1 kg of fuel per kg of fish caught (Schau et al. 2009). This illustrates how both the stock targeted and gear used influence fuel consumption, and will likely influence the development of new fisheries and fishing techniques in future.

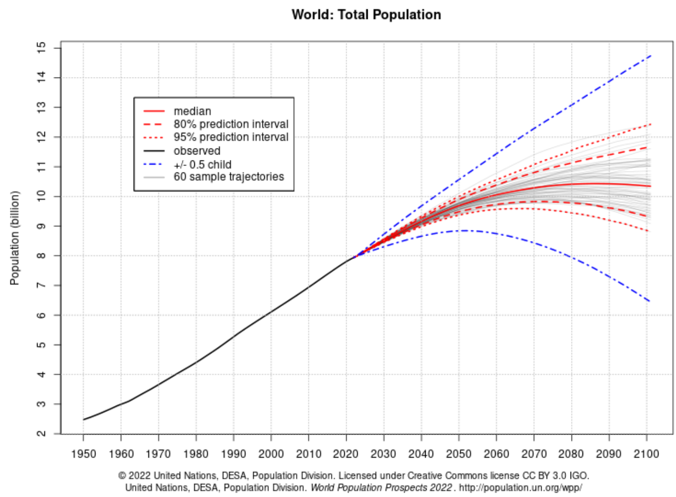


Figure 1.3: World population estimates from 1950 to 2100, calculated based on historic demographic trends.[6]

1.3.4 Fish population dynamics

It's also important to consider the future trends in fish population dynamics, particularly the impact of climate variability on marine ecosystems. Historic responses of fisheries to climate change have been studied extensively, and shifts in thermal niches are the most notable effect. Gvoždík (2018) defines thermal niches as the "range of body temperatures maintaining positive population growth". Studies have pointed to significant thermal niche shifts in plankton and mackerel with

[6]Source: United Nations, Department of Economic and Social Affairs, Population Division (2022). Custom data acquired via website:
https://population.un.org/wpp/Graphs/Probabilistic/POP/TOT/900.

smaller effects in demersal species (Beaugrand et al. 2002, Perry et al. 2005, Bruge et al. 2016). The shifts are predominantly northwards in the northern hemisphere, with the shift in overall preferences of marine species averaging 0.23 degrees celsius for non-tropical species (Cheung et al. 2013). One of the most significant average extensions was for warm-water copepods, showing a shift of 10 degrees latitude (Beaugrand et al. 2002). This suggests lower trophic levels may be more sensitive to warming.

There are several predicted effects of thermal niche shifts. Firstly, the lack of synchrony between pelagic species and plankton phenologies (recurrent annual life cycle events) may lead to trophic mismatch, decoupling relationships between predators and prey and reassembling food webs (Edwards and Richardson 2004, Pinsky et al. 2020). Secondly, such shifting in geographical niches is related faster life cycles and smaller body sizes (Perry et al. 2005). Thirdly, research suggests marine species are more sensitive to climate change than terrestrial species. Although marine species have better capacities to colonize new regions, they tend to have less behavioural adaptations and to be more physiologically sensitive to disturbances (Pinsky et al. 2019).

The changes in distribution range may also depend on time of year. Reductions in sea-ice extent are predicted to lead to more efficient fish foraging and seasonal expansions of fish to higher latitudes (Varpe et al. 2015, Langbehn and Varpe 2017). This could mean that migratory fish such as herring will move further north during summer, but back south during winter. Modelling and monitoring the effects of climate disturbances on ecosystems will help predict and respond to consequences of changing climates.

1.4 Modern monitoring of marine ecosystems

Possessing advanced knowledge of fish stocks will aid in transitioning to environmentally and economically sustainable harvesting. Thus, there is increasing interest in monitoring fish stocks using data from commercial vessels. Indices of abundance from fishing fleets are an important source of information for stock assessment (Hilborn and Walters 2015). Reference fleets are particularly useful for analysing data from non-commercially important species, where there are no research surveys to scientifically study stocks. An article by Jones et al. (2022) demonstrates how high resolution data from the US reference fleet has contributed to abundance indices for several stocks, while footprints of fishing vessels can inform planning of offshore wind projects. A Norwegian reference fleet programme found that gathering species and age composition data from fishing vessels is a cost-effective method of sampling and producing abundance time series for cod, haddock and redfish (Hatlebrekke et al. 2021).

In addition to biological samples from commercial vessels, fishing effort can be estimated at a high spatial resolution using satellite positional systems (Natale et al. 2015). For example, the fishing activity of tuna purse seiners has been estimated in this way (Bez et al. 2011). Analysing such information can screen for illegal fishing activity, map global footprint of effort and estimate abundance of fish stocks (de Souza et al. 2016, Kroodsma et al. 2018, Adibi et al. 2020). Global Fishing Watch is a large initiative that aims to increase transparency of marine activity through such analysis[7].

Finally, there are many cutting-edge projects seeking to reap novel observations of marine ecosystems. The Centre for Research-based Innovation in Marine Acoustic Abundance Estimation and Backscatter Classification (CRIMAC) is developing algorithms for processing and classification of acoustic data, with applications to fisheries at centre stage (Handegard et al. 2021). Sea kayaks are Unmanned Survace Vehicles (USVs) that can be rigged with an echosounder for assessment of small fish living in shallow waters in coastal regions e.g. sandeels (Totland and Johnsen 2022). Scantrol deep vision is an example of a product being developed, in cooperation with scientific institutions, to automatically record species and size information on fishing vessels *in situ* (Handegard et al. 2021). Similarly, the Fishguider DSS described in Chapter 5 is proposed as a method for capturing information from fishing vessels and delivering decision support services.

1.5 Model-based estimation of marine resources

Jens Balchen, the founder of the ITK department at NTNU, had strong views on the importance of using model-based estimation for analysing marine ecosystems, believing mathematical modelling "... is the most effective single tool to help understand the basic internal mechanisms of ocean subsystems and the interaction between such subsystems." (Balchen 2000). He wrote in the 1980s of the importance of "estimating the resources of fish and managing the utilization of these resources" (Balchen 1980). He also stressed the importance of validating state and parameter estimation of mathematical model with true measurements for: "Recursive estimation of fish aggregate quality and location" (Balchen 1981). His Ocean Bio Model project in 1975 reflected this vision, where he integrated models of hydronamics, nutrients, phytoplankon, and migration behaviour (Balchen 1980; 2000, Breivik and Sand 2009). Balchen (2000) suggested four reasons for modelling fish dynamics (paraphrased):

1. To simulate and study the behaviour of the entire system.

[7] See: https://globalfishingwatch.org/

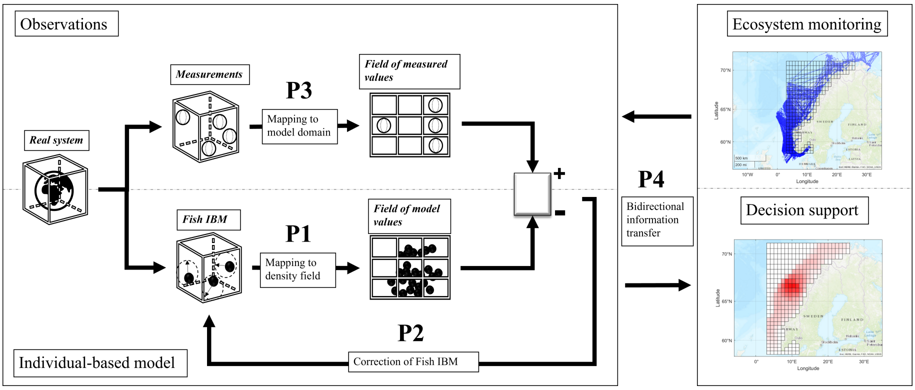


Figure 1.4: Block diagram illustrating the proposed model-based estimation system with the relevant paper contributions to the research objectives (labelled P1:P4 for related Paper X in the List of Articles). It shows the bidirectional information transfer between the modelling and estimation loop (left block) and fishing vessel activity (right block).

2. To estimate model variables and parameters not directly observable.
3. To estimate ocean resources for control of harvesting.
4. To control the behaviour of individuals in a biological population.

This thesis echoes the importance of such objectives, where we contribute methods for: estimating fish migration patterns (1), assimilating real-time fisheries observations for estimation of unobserved state variables (2), and estimating catch potential of ocean regions (3). Although a curious research idea, controlling the behaviour of individual fish was beyond the scope of our work (4). Ultimately, this thesis argues that Balchen's vision of model-based estimation is becoming more feasible as access to new measurement systems become available (Figure 1.4).

1.6 Objectives and contributions of the thesis

1.6.1 Objective 1: To develop an IBM of the NSSH spawning migration

To test the model-based estimation system, a model was needed as a proof-of-concept. Following discussions with fishers participating in the project, NSSH was chosen as the model stock. The spawning migration was modelled given its large spatial extent and the many hypotheses about the drivers of the migration available (Fernö et al. 1998). It also overlaps with a key period in the fishing season, and thus understanding the development of the migration can aid in planning fishing operations. IBMs are a suitable tool for simulating migrations, given one can explicitly represent fine-scale interactions between individuals and their surrounding environment. In addition, IBMs have been developed to estimate migration patterns of many pelagic fish species (Barbaro et al. 2009, Politikos et al. 2015, Boyd et al. 2020).

Contribution

This objective was met by developing an IBM of the NSSH spawning migration which was coupled to an ocean model developed at SINTEF, allowing individuals to respond to local environmental conditions. Although the implementation was based on local interactions, many interesting large-scale patterns emerged that were stable across simulations. Many of these patterns overlapped with patterns observed from fishers logbooks and those described during research surveys. The high spatial and temporal resolution of the model adds to its potential operational utility.

1.6.2 Objective 2: To implement a Data Assimilation approach for strengthening IBM estimates

Any IBM of fish migratory behaviour is subject to multiple sources of uncertainty, from misrepresentation of complex interactions to lack of knowledge of true fish behaviour. In addition, IBMs are trained on historical data sources that are subject to their own uncertainties. Thus, the IBM is prone to overfitting based on both the modellers intuitions and idiosyncrasies in the available datasets. To overcome this limitation, one may incorporate uncertainties through modelling an ensemble of model instances and subsequently, correcting model states based on available measurements. This approach is known as Data Assimilation. Data Assimilation can improve model predictions and strengthen the utility of the IBM developed as part of Objective 1.

Contribution

This objective was met by extending the herring IBM to include uncertainties in the orientation and speed of individuals. The corrections of model states were calculated using an EnKF and a twin model simulation experiment determined the capacity to correct model states with a variable number of measurements. One challenge was adapting the IBM for this setup, so we proposed an algorithm for mapping between state estimates in the IBM and the EnKF. This was similar to an approach described in Cocucci et al. (2022) as randomized redistribution. Ultimately, we showed that the model predictions could be improved by using the EnKF setup, especially in cases where the forecast IBM was highly inaccurate.

1.6.3 Objective 3: To assimilate synthetic measurements derived from fishing vessel activity

In order to effectively correct the IBM as described in Objective 2, one needs access to a large array of available measurements. The data available from surveys or catch logs are extremely sparse. As measurement sources are currently quite limited, deriving novel patterns from available data can supplement the assimilation procedure. We aimed to train a machine learning algorithm to predict NSSH densities based on vessel activity, convert predictions to synthetic measurements and input these measurements to the assimilation procedure. We also aimed to suggest how assimilated fields can inform catch potential of ocean regions.

Contribution

This objective was met by developing a method for utilizing output from a neural network as synthetic measurements to correct the IBM in the EnKF setup. The neural network predicted relative densities of NSSH based on fusion of vessel activity and electronically recorded catch logs. Assimilation of synthetic measurements drove model scenarios with dissimilar forecast states towards common spatial patterns. Tuning of the observation uncertainty parameter influenced the degree to which the model states were altered. In the case of inaccurate model forecasts, the IBM was improved with assimilation of synthetic measurements, relative to the control scenarios. Using an occurrence-based metric for catch potential, we illustrated how assimilation of synthetic measurements improved predictions of catch areas. Assimilated fields can suggest catch potential of unexplored territories during the fishing season.

1.7 Thesis outline

- Chapter 1 has given a brief background to the challenges faced by the Norwegian fishing industry in achieving economic and environmentally sustainable harvesting. It has also laid out the rationale behind the model-based estimation approach in aiding with these challenges.

- Chapter 2 introduces IBMs as tools for modelling complex systems and their application in fisheries ecology. It elaborates on the background to the herring IBM before describing the model setup and calibration.

- Chapter 3 describes the limitations of a single realization of the IBM, and how one can improve IBM estimates using Monte Carlo simulations and Data Assimilation. It then lays out the main steps involved in adapting the IBM to the Data Assimilation procedure.

- Chapter 4 discusses the capacity to relate vessel activity data to patterns in stock distributions that can be used as input to modelling work. It is illustrated how we generated synthetic observations from a neural network using vessel activity as input. Furthermore, it is shown how these measurements are integrated with the Data Assimilation procedure to predict catch potential of NSSH.

- Chapter 5 illustrates how DSS can aid fishing vessels in strategic, tactical and operational decision-making. It presents the progress in development of the Fishguider DSS tool during this project.
- Chapter 6 concludes and reflects on the contribution of this thesis. It also discusses alternative representations of the model and estimation systems, further analysis of spatial patterns and future collaboration possibilities.

Chapter 2

Individual-based modelling of herring migrations

2.1 Modelling complex systems

Complex systems that exhibit non-linear dynamics are ubiquitous in nature from networks of neurons to ecological systems (Strogatz 2001). To understand complex systems, key properties must be captured in mathematical models, and often, modelling simple interactions lead to complex dynamics. Famously, John Conway demonstrated how a "game of life", which simulated real-life processes like births, deaths and survival, could produce stable, symmetric population-level patterns (Gardner 1970). In Williams and Martinez (2000), it is illustrated how feeding behaviours such as cannibalism and omnivory can be explained in complex trophic food webs through assignment of simple rules such as randomly drawn niche values. Similarly, Cocucci et al. (2022) shows how COVID-19 transmission patterns can be modelled through simple individual-level attributes in compartmental models that count susceptible, infected, and recovered individuals (SIR models). Studies of small-world networks show how variable propagation speed of infectious diseases is influenced by local connectivity (Watts and Strogatz 1998)

Mechanistic models are used where causal mechanisms are derived from human observation of the phenomenon of interest. This is in contrast to machine learning methods which predict outcomes of complex mechanisms without need of understanding the phenomenon (Baker et al. 2018). One can effectively transition from theory to study of complex systems through a series of stages of calibration and validation of mechanistic models (Baker et al. 2018). Mechanistic models come in many flavours, depending on the scales being studied, the nature of interactions and so on. For example, spatially explicit landscape models (SELMs) deal at a coarse level of aggregated cohorts, with limited depictions of individuals, and disturbances on a large scale (e.g. droughts) are forcing the model (Perry and Enright 2006). Eulerian models express dynamics of population in terms of differential equations to represent time, space, age, length and weight in a continuum. They are useful in modelling lower trophic levels where processes such as advection dominate life stages, as in copepods, fish larvae and capelin (Reed and Balchen 1982, Alver et al. 2016). At a finer reolution, complex adaptive systems (CAS) are used to model conditional actions (if/then rules) through sequential sets of rules. Adaptation and evolution of agents can be understood using CAS (Holland 2006).

2.2 Individual-based modelling

Individual-Based Models (IBMs), simulate the complex interactions between individuals and their environment, and produce emergent behaviours at the population level (Grimm and Railsback 2005). IBMs are related to the field of CAS, where they similarly examine system level changes associated with changing capabilities of agents (Railsback 2001). IBMs may also be referred to as Agent-based models (ABMs), although ABMs tend to be used in simulating human systems as opposed to IBMs usage in simulating non-human systems (Bonabeau 2002, Grimm et al. 2006, McLane et al. 2011, DeAngelis and Grimm 2014). Finally, one can characterize IBMs as microscopic models and opposite approaches as macroscopic models, reflecting the focus on local interactions in IBMs (Bonabeau 2002).

Historically, the development of individual-based models was a response to the issue that most models at the time didn't distinguish between organisms location, meaning they were spatially inexplicit (Huston et al. 1988). Models were limited to describing populations by single undifferentiated variables such as population size. IBMs were considered a distinct approach in ecological modelling in 1988, and in the 90s their usage expanded due to their explanatory power (Huston et al. 1988, Grimm 1999). The advance in computational power allowed movement towards modelling of individual units in ecosystems, using local interactions encoded in numerical simulations, which were difficult to incorporate into previous analytic models (Grimm et al. 2006). This allows variation at a finer resolution (Figure 2.1). Using differentiated state variables, such as position and velocity, one can incorporate more sophisticated dynamics in ecosystems and explore spatial variability, local interactions and movement (DeAngelis and Mooij 2005).

There are two contrasting approaches to IBMs, depending on the motivation. On one hand, there are pragmatic approaches, where although the model is informed by theory, it is not meant to prove or disprove hypotheses from classical theory. On the other hand, a paradigmatic approach makes explicit reference to claims in

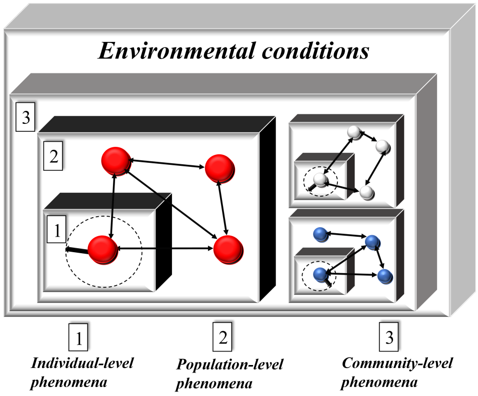


Figure 2.1: An illustration of how individual dynamics and interactions between individuals are at the core of population-level and community-level phenomena. Environmental conditions both faciliatate and constrain process at all levels. This conceptualization is adapted from (Huston et al. 1988).

theoretical ecology (Grimm 1999). For example, Huse et al. (2002) illustrates how the directed behaviour of a critical mass of individuals can redirect targeted movements of schools of herring, supporting the adopted-migrant hypothesis. One can also explore hypotheses related to simulating movement through complex landscapes, competition, community dynamics and evolutionary processes (DeAngelis and Mooij 2005, Pe'er and Kramer-Schadt 2008).

Conservation planning is an example of the pragmatic approach, where IBMs are used to assess the response of individuals to changes in habitat, especially critical ones (McLane et al. 2011). By predicting the consequences of changes to ecological systems, one may minimize adverse consequences (Stillman et al. 2015). This thesis will explore a pragmatic approach to IBMs given the applied nature of the project work.

2.3 Individual-based models in fisheries ecology

Balchen (2000) suggests that the individual dynamics of fish are extremely complex on short time scales, where schools of fish or "big fish" should be modelled instead. These aggregate units, now referred to as super-individuals, can take the average expected motion or other properties of groups of individuals. Super-individuals are initialized in grid cells simulating external conditions before the model is stepped forward in time. Functions are called to communicate information to individuals about these external conditions (Grimm and Railsback 2005). Differential or difference equations are used to calculate these steps forward in time. In this way, IBMs are useful in modelling spatiotemporal dynamics of fish (Giske et al. 1998).

Bauer and Klaassen (2013) describes migratory behaviour as the persistent and directional movement with distinct departing and arrival behaviours. The primary motivations for fish to migrate *en masse*, at large scales, are to reproduce, forage and avoid predators (Tamario et al. 2019). For example, mackerel is highly mobile in distribution area of annual migratory cycles and spawning areas range from the west coast of Portugal to the southwest coast of Norway.[1] IBMs have been applied to simulate migratory behaviour of many fish species including capelin, anchovy and mackerel (Barbaro et al. 2009, Tu et al. 2012, Politikos et al. 2015, Boyd et al. 2020). Our work was informed by such models for the purpose of spatially and temporally explicit predictions of herring migration patterns.

2.4 Background to the herring IBM

NSSH is a migratory pelagic stock mainly distributed along the Norwegian, Faroese and Icelandic coast, migrating vast distances during its life cycle (Dragesund 1970). It is one of the most commercially valuable stocks in the North Atlantic (Touzeau et al. 2000). Additionally, lack of information about spatial distributions has led to unsustainable harvesting historically (Fernö et al. 1998). Through project meetings and conferences attended by fishers involved in this project, it was concluded that we should prioritize model input for herring decision support (elaborated on in Chapter 5). Therefore, it's advantageous from both the fishing industry and management point of view to maximize information available about the stock.

In Paper 1 we developed a novel mechanistic IBM of the NSSH spawning migration, which predicts the geographic distribution of NSSH over short time increments. There were two reasons we focused on large-scale predictions. Firstly, we don't have access to data on individual histories of fish, such as swimming speed, body condition, and other biological characteristics. We thus focused on assimilating fish densities, as described in Chapters 3 and 4. Secondly, real-time estimates are useful for informing vessels of which fishing grounds have high catch potential throughout the season. Therefore, the model simulated large scale patterns of distribution, driven primarily by coupling fish behaviour to ocean states.

[1] See: https://www.hi.no/en/hi/temasider/species/mackerel

2.5 Development of the herring IBM

2.5.1 Model setup

The herring IBM was coupled to the SINMOD ocean model developed at SINTEF (Slagstad and McClimans 2005). At each time step k, individuals in the IBM accessed environmental conditions at their local SINMOD grid cell. The IBM was developed with difference equations that stepped forward the movements of each individual, where position $\mathbf{p}$ was updated on a continuous grid:

$$\mathbf{p}[k+1] = \mathbf{p}[k] + \Delta t \left(\mathbf{v}_f[k] + \mathbf{v}_c[k]\right) \tag{2.1}$$

where:

$$\mathbf{v}_f[k] = -\Phi \mathbf{v}_c[k] + \mathbf{v}_b[k] \tag{2.2}$$

where $\mathbf{v}_b$ is a vector with the horizontal velocity components of an individual fish in the x and y directions, based on behavioural cues. Similarly, $\mathbf{v}_c$ is a vector with the horizontal current velocity components in the x and y directions. The intended swimming velocity $\mathbf{v}_b$ is thus tempered by the prevailing current $\mathbf{v}_c$, leading to a realized velocity of $\mathbf{v}_f$. This formulation reflects the fact that the spawning migration proceeds counter to the Norwegian coastal current (Slotte and Fiksen 2000), and thus the model term $\Phi \mathbf{v}_c$ adds a counter-current component to the horizontal speed controlled by the parameter Φ.

The vector $\mathbf{v}_b$ was calculated as:

$$\mathbf{v}_b[k] = r_b[k] \left(\begin{bmatrix} \cos(\theta[k]) \\ \sin(\theta[k]) \end{bmatrix} \right) \tag{2.3}$$

where:

$$\theta[k] = f(\nabla T[k], \nabla D[k]) \tag{2.4}$$

where $\nabla T[k]$ and $\nabla D[k]$ are the temperature and bathymetry gradients, and the intended swimming speed is r_b. The the right term in Equation 2.3 is a unit vector determining the orientation of the individual, while r_b is the magnitude of the movement. The NSSH spawning migration develops southward alongside the continental slope (Slotte and Fiksen 2000). Additionally, herring are physostomous with an open swim bladder, which facilitates more rapid vertical movements, and vertical escape is considered central in predator avoidance (Blaxter 1985, Nøttestad 1998, Langård et al. 2014). The modelled response to ∇D incorporated this theory. NSSH avoid low temperatures and higher temperatures are associated with superior body condition and thus the response to ∇T encodes this response (Fernö et al. 1998). More details on the functional mechanisms are provided in Paper 1.

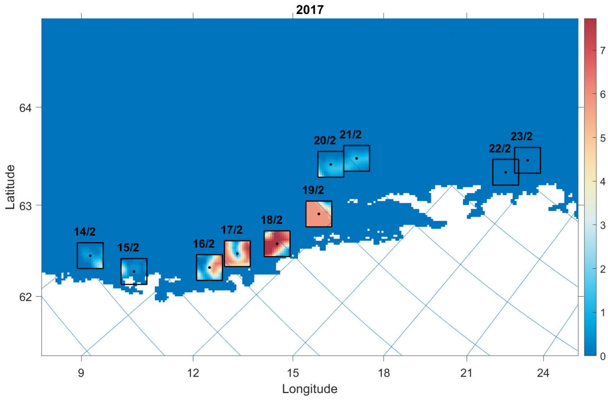


Figure 2.2: Transformed survey values used to compare against model values on specified dates in 2017. Black dots indicate the centre of mass of trawl positions on the date transcribed above the box. The black lines demarcate the outer boundary of cells included for the comparison. The colourmap indicates estimated number of individuals in grid cell j.

2.5.2 Parameter calibration

Survey data from herring research surveys were used to tune model parameters. The survey is carried out for approximately 14 days, beginning in the south of Norway and proceeding northwards (Figure 2.2). This offers a short time window where we tested the IBM estimates. The values from the survey were spatial abundance estimates based on trawl and acoustic sampling, so these were normalized to values that were equivalent to IBM output for comparison. The IBM output was the number of individuals in each 4 km^2 SINMOD grid cell. The model was iterated for several years (2015-2020) to optimize parameters and minimize deviations between these transformed observation values and the model predictions.

2.5.3 Comparison to catch data

Once the parameters were calibrated, the model was compared to independent logbook catch data for years simulated (2015-2020). For qualitative comparison, Figure 2.3 show catches take place along branches of the migration where the model predicts higher densities in 2016. Survey distribution also corroborate findings in early February with observations of high densities around 66-67 degrees latitude (Slotte et al. 2016). However, there is interannual variability. For example, both survey and catch observations show a more offshore distribution in 2015. Figures from 2015-2020 are included in the supplementary material of Paper 1.

Quantitatively, spatial indices of the development of the model were used to gauge

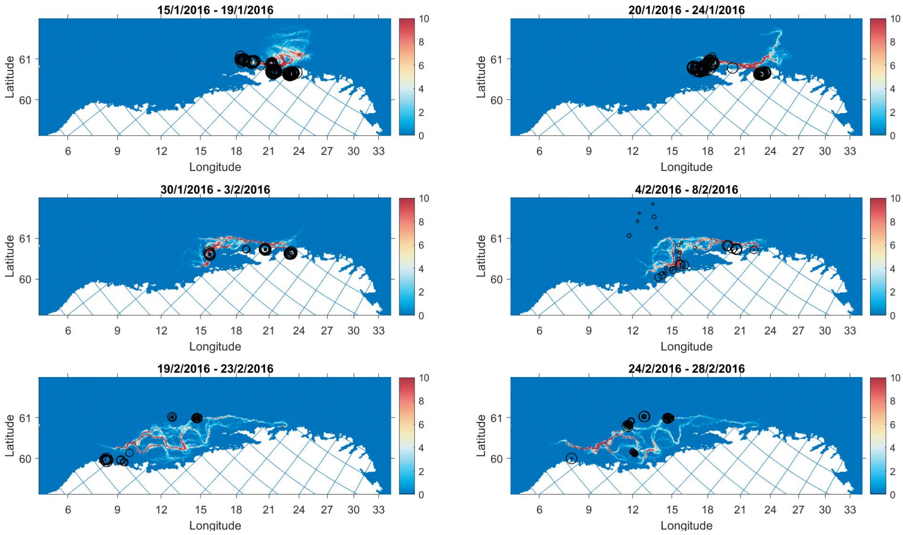


Figure 2.3: Model output (colourmap) with catch points (black circles) overlayed for selected periods in 2016. Size of circles are scaled according to the catch weight in kg. The colourmap gives the average number of individuals in grid cell j for a 5 day period.

the model performance. The centre of gravity (CG) measured the latitude and longitude points weighted by the density of individuals at each position. The global index of collocation (GIC) was used an index of the overlap between the model and observation spatial distributions. The comparisons were over 5 day time windows (Figure 2.4).

2.6 Discussion

The IBM modelled an undifferentiated mass of fish. Further work on the model may include age-class structure, bioenergetics, energy budgets or other configurations that were beyond the scope of the project work. These may add information for theoretical simulations. For example, one may study the effect of counter-current migration patterns on energy expenditure of individual fish.

IBMs are subject to uncertainties from a number of sources including assumptions about the model structure and lack of knowledge of the real system. Additionally, the SINMOD ocean model itself has uncertainty caused by limits in model resolution, our knowledge of the processes resolved by the model, uncertainty in initial values, boundary conditions, parameter values, and inaccuracies in numerical implementations. Furthermore, climatic fluctuations can play an important role in survival of recruits and thus, biomass estimates. Therefore, understanding

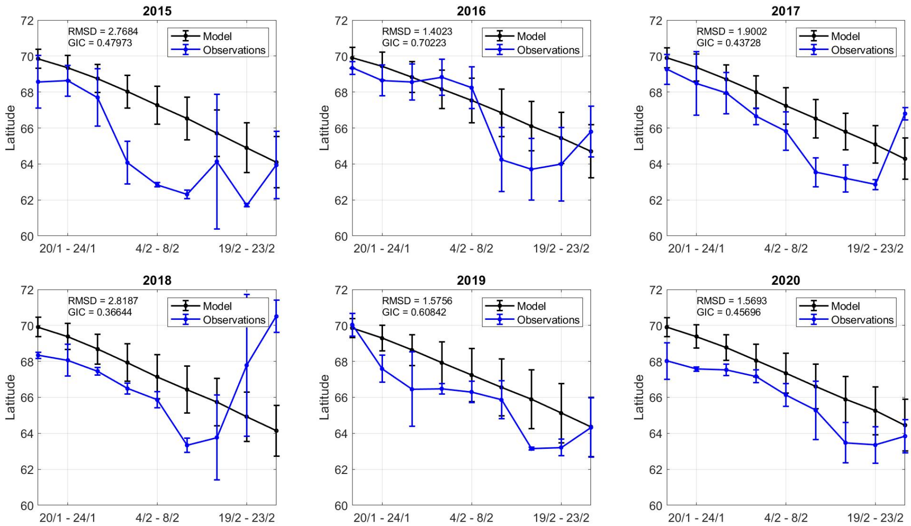


Figure 2.4: Model and observation comparison across all years. The x axis displays the time period (inclusive). The y axis displays the latitude centre point for catch and model values in the time period. The error bar shows the square root of the inertia values in each time interval.

how variability in environmental conditions can manifest as variability in the IBM structure is important further work.

Although we can improve SINMOD output and improve our understanding of the herring, it is not feasible to make a model that predicts the migration without uncertainty. Assimilation of observations from the real system will always be necessary for the model to provide useful real-time predictions (see Chapter 3).

2.7 Conclusion

Paper 1 presented a novel IBM of the NSSH spawning migration based on assumptions derived from theoretical studies. Particularly, this work corroborates the theory that memory- and gradient- based reactive mechanisms may drive the spawning migration of NSSH (Fernö et al. 1998). The model produces stable patterns from simple behaviour rules which show good agreement with survey and catch observations. The IBM is modelled with individual mechanisms, but is intended to produce large-scale estimates. Pragmatically, this IBM predicts near real-time estimates of the herring migration which can inform fishing activity.

Chapter 3

Model corrections using observations

3.1 Uncertainties in the IBM

In Paper 1, the IBM was calibrated based on the minimization of errors between the model and estimates from the survey. This produced a viable migration trajectory when compared with the spatial development of commercial catches. However, there are uncertainties in both the IBM and SINMOD model estimates, and there are gaps in our knowledge of true herring behaviour. Likewise, available measurements are prone to errors arising from sources such as size selectivity of nets and reporting inaccuracies.

As Evensen (2009) points out, the solution to a dynamical model is one of infinitely many realizations, and so we should consider the development of the Probability Density Function (PDF). Furthermore, Baker et al. (2018) explains that the utility of a model calibrated on historical data is limited in forecasting scenarios. Adding noise terms to model variables and parameters can be used to produce an ensemble of estimates through Monte Carlo experiments. States and parameters can be sequentially estimated when combining these model ensembles with measurements (Ward et al. 2016). The herring IBM can be considered one of an infinite number of possible migration scenarios. A large space of alternative migration dynamics can be envisioned given uncertainties owing to process noise, non-linearities in the model dynamics and high dimensionality of the model domain. These realized scenarios may emerge in the form of a more rapid rate of migration southward, or a trajectory that develops further offshore (Figure 3.1).

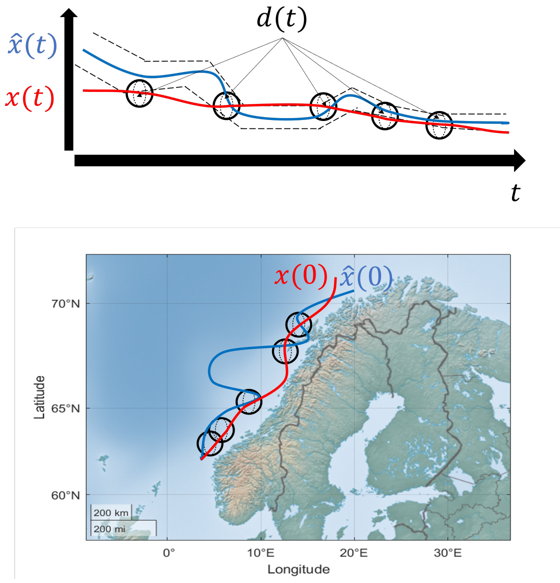


Figure 3.1: Conceptualization of the Data Assimilation procedure employed to predict the hypothetical true migration centre of mass (x), based on melding a forecast model ($\hat{x}$) and available measurements (d). Note that this a highly simplified illustration.

In Paper 2, we aimed to incorporate both model and observation uncertainties to improve the real-time estimates from the herring migration IBM. Specifically, we wished to develop a system where incoming real-time measurements are melded with corresponding model simulation output (Figure 3.1). This process of combining measurement and model values is known as Data Assimilation. There were two main objectives of this work. Firstly, to adapt the IBM to the Data Assimilation framework. Secondly, to test the capacity of the assimilation framework to correct unobserved model states.

Using this procedure one can predict the fish migration trajectory as it is developing, and control the divergence of the model from realistic trajectories. This has applications in real-time monitoring of the stock by fishing vessel and in studying novel migration patterns that may be related to shifts in thermal niches or recruitment dynamics. Finally, this represents the incorporation of a cybernetic approach uncommon in fisheries research.

3.2 Data Assimilation

Observations in fisheries are sparse and only give indirect information about the system while mathematical models diverge from realistic representations of fisheries systems owing to uncertainties. Data Assimilation controls model divergence through a set of mathematical techniques for combining model states with observations, sequentially in time, to provide the best possible estimate of the state of a physical system. It operates under the assumption that models or observations alone contain incomplete information in resolving the real system. One must therefore apply statistical correction terms (i.e. gains) to model estimates based on incoming measurements, melding both sources of information (Fu et al. 2011, Alver and Michelsen 2015). Data Assimilation has been used in estimation within fishery models, predictive ecology, the terrestrial carbon cycle and traffic simulation (Niu et al. 2014, Ward et al. 2016, Kieu et al. 2020). In addition, it is a crucial element in weather forecasting.

Theoretically, Data Assimilation is generally categorized as a Bayesisan estimation problem, where the repeated, sequential updating of model states is referred to as recursive Bayesian estimation. It steps forward a probability distribution function for a variable X. Bayesian estimation estimates the PDF $f(x)$, which is the probability that variable X will take on a particular value x. Bayes theorem proposes that the posterior PDF of a model state given measured states $f(x|d)$ is proportional to the PDF of the prior model state $f(x|d)$ times the conditional PDF of the measurement given the model state $f(d|x)$, also called the likelihood:

$$f(x|d) = \frac{f(x)f(d|x)}{f(d)} \tag{3.1}$$

This proposition considers the posterior state $f(x|d)$ the best possible estimate given the prior model states and available measurements. In other words, our certainty in the model estimate increases when we sequentially observe attributes of the real system. Formally, the Fokker–Planck equation is the analytical solution for stepping the probability distribution forward sequentially in time, but in high dimensional systems this is not feasible, so simplified representations are chosen. The Ensemble Kalman Filter (EnKF) is one such simplified representation.

3.3 The Ensemble Kalman Filter

The EnKF is a Data Assimilation method, initially developed by Evensen (2009), that uses Monte Carlo simulations to explicitly represent random process noise through simulation of N separate instances of the prediction model. Theoretically, if the set of N were infinitely large, it would perfectly represent the probability distribution. Computationally, we can represent the probability distribution with a high N, and treat the covariances $\mathbf{C}$ across the ensemble as a sufficient estimate of the probability distribution. The covariances $\mathbf{C}$ are estimates of the true covariances of the probability distribution of the model states, given the assumed process noise. When observations are available, a correction term is applied to each instance of the model based on **C** and observation error covariances (Evensen 2009). The EnKF is used for state and parameter estimation of non-linear systems e.g. atmospheric and ocean systems (Houtekamer and Mitchell 2001, Alver and Michelsen 2015). In Paper 2 we aimed to extend the IBM developed in Paper 1 to incorporate both model and observation uncertainties.

Generally, the EnKF works well in forecasting estimates of non-linear models, so it's a suitable approach for IBM or ABMs. Macro-scale estimates such as population size can be estimated from combining IBMs and Data Assimilation (Niu et al. 2014). Micro-scale adjustment of variables and parameters can be achieved using Data Assimilation to correct IBMs, although this is less explored territory. In an article in 2022, Cocucci et al. (2022) showed that attributes in SIR models, such as individuals susceptible to COVID-19 can be updated using assimilation. However, this presents a challenge as we usually don't have access to detailed measurements on micro-scale characteristics for fish such as length, body weight, condition factor, but rather macro-scale characteristics, such as geographical location and density of fish. We cannot link observations to specific individuals in the IBM, just to properties such as density. Therefore, as in Cocucci et al. (2022), we aimed to develop an algorithm in Paper 2 that mapped from microscopic to macroscopic states.

3.4 Data Assimilation framework for the herring IBM

3.4.1 Monte Carlo Simulations

The prediction model of the herring IBM was extended to N instances, representing process noise not accounted for in the single IBM. The full set of N instances is referred to as the ensemble and the approach for stepping each instance forward is known as Monte Carlo simulation. Concretely, position **p** and velocity **v** of individuals were extended from the single IBM to N ensemble members, notated by the state matrices **P** and **V**, both with N columns:

$$\mathbf{P} = \mathbf{P} + \Delta t \left(\mathbf{V} + \tilde{\mathbf{V}} \right) \tag{3.2}$$

In Paper 1, estimates of spatial indices were the focus, but with the EnKF setup, the mass of fish can be estimated when measurements are incorporated. Thus, biomass **B** of individuals was added as another state in this work. The biomass state **B** was forecasted as follows:

$$\mathbf{B} = \mathbf{B} - \Delta t \left(\tilde{\mathbf{B}} + \omega \right) \mathbf{B} \tag{3.3}$$

where Δt was the time increment, reduction in biomass was controlled by the constant parameter ω, and divergence in states **V** and **B** were caused by the stochastic errors $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{B}}$. The expected value $\mathrm{E}[\tilde{\mathbf{V}}] = \mathrm{E}[\tilde{\mathbf{B}}] = 0$. These errors produce prediction uncertainty in the system, representing errors in individuals migration direction, speed and mass leading to divergence in the model development (Figure 3.2). By adding stochastic noise representing the uncertainty in inputs and model dynamics, we get an ensemble spread which represents the covariance structure of the PDF of the model states, and such a covariance estimate is a prerequisite for computing the Kalman Gain in the EnKF.

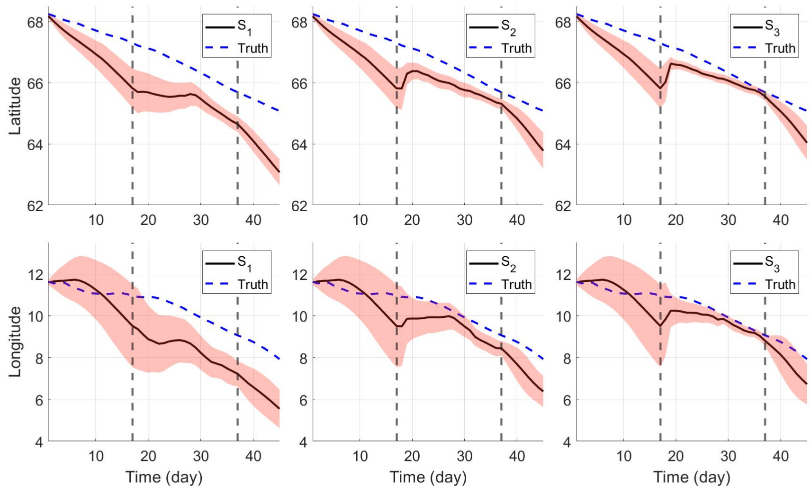


Figure 3.2: Time series of mean CG of latitude (first row) and longitude (second row) values during the simulation period, for three scenarios with increasing number of measurements from S_1 to S_3) and the true CG being estimated. The shaded areas show the standard deviation in spatial variation amongst ensemble members, arising from the addition of process noise.

3.4.2 Adapting the IBM to the Data Assimilation framework

The EnKF uses the error covariance structure of the ensemble forecast to calculate the correction term. However, the full covariance matrix $(n \times n)$ is too large to be explicitly calculated, and so an equivalent representation by Mandel (2006) was implemented:

$$\begin{aligned} \bar{X}^f &= \frac{1}{N}\sum_{i=1}^{N} X_i^f \\ A^f &= X^f - \bar{X}^f \end{aligned} \tag{3.4}$$

where X^f is the forecast state matrix mapped from the IBM states, $\bar{X}^f$ is the mean, and A are the model anomalies. The Kalman Gain is then calculated as follows:

$$\begin{aligned} HA &= HX^f - H\bar{X}^f \\ P &= \frac{1}{N-1} HA(HA)^T I_m + R \\ K &= L \odot \left(\frac{1}{N-1} A^f (HA)^T P^{-1} \right) \end{aligned} \tag{3.5}$$

where H is an $m \times n$ matrix that maps between model states and measured states, I_m is an $m \times m$ identity matrix, R is the $m \times m$ observation error covariance matrix, where each element on the diagonal is the variance of observation noise (Ω). The parameter Ω is an important parameter to tune, as it determines the strength of the final correction value. L is an $m \times N$ localization matrix which adds a penalty to model covariances that are distant from the measurment points. For a small ensemble and high dimensional system, localization is necessary to limit the impact of spurious correlations in the ensemble (Houtekamer and Mitchell 2005). Finally, K is the Kalman Gain, which is used to calculate the correction term. The analysis equation is calculated:

$$X^a = X^f + K(D - HX^f) \tag{3.6}$$

where the posterior estimate X^a is calculated based on the measurement matrix D, the forecast estimate X^f and the Kalman Gain K. One issue with using observations of fish densities is the non-negative nature of measurement values. Observations are usually perturbed with Gaussian noise to produce D of $m \times N$, but this causes instabilities in the IBM where there are low fish densities, and instead the N columns were treated as N replicates of the $m \times 1$ measurement vector d. Perturbations avoid the contraction of variance across the ensemble of model states. In Paper 2, we use an inflation factor to maintain variance:

$$X_z^a = \bar{X}^a + \psi \left(X_z^a - \bar{X}^a \right) \tag{3.7}$$

where z is the ensemble member and $\bar{X}^a$ is the mean of the analysis states.

The IBM states **P**, **V** and **B** represent information about individual fish in the IBM. This presented a challenge, as it required a mapping between micro-state representations of IBM states to macro-states of the EnKF. We developed a mapping function to minimize redistribution and maximize information retained. Mapping from the IBM to a density field is straightforward, where we use the position **P** and biomass **B** of individuals to generate X^f (see Algorithm 1 in Paper 2). However, on mapping from the field representation X^a back to the IBM, we are careful to avoid loss of information. We used a randomized redistribution method to assign X^a estimates back to **P** and **B** states, moving individuals to grid cells with non-zero density values and assigning new density values if necessary (see Algorithm 2 in Paper 2).

3.4.3 Twin model experiment

To test the setup described above, a twin model experiment was devised, where a twin IBM represented the true migration scenario and was used to generate synthetic measurements for correction of the model IBMs. The advantage of using a model as the true distribution is that we have full knowledge of the true state values at any given time. This allows us to infer the impact of corrections on unobserved states. The observation system devised was fixed for each scenario, so measurements were made at equal increments (once per day) in identical locations during the model time frame. The number of observations was varied in each scenario.

The twin model was similar to the herring IBM developed in Paper 1, but with a modified swimming speed parameter. The parameter was lower in the twin model, representing a true trajectory that develops more slowly than our prior model indicates. In Paper 2, we show the capacity of different scenarios to converge on this true model estimate (Figure 3.2 & 3.3).

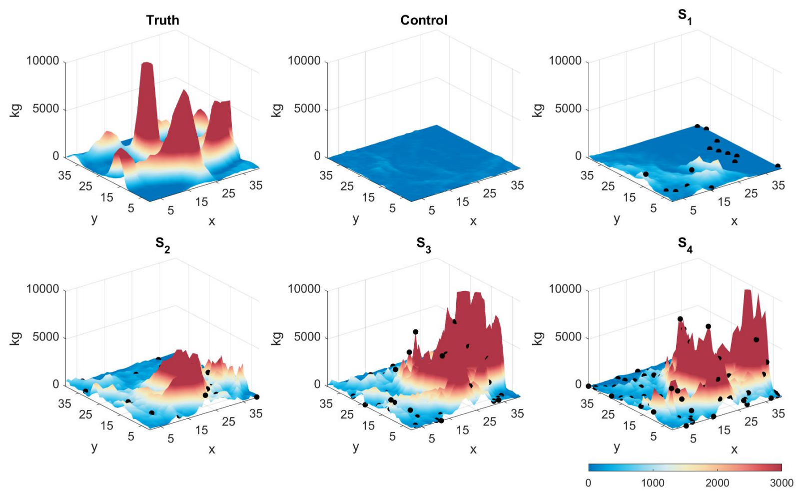


Figure 3.3: The local 3D representations of $\bar{X}^a$ for four scenarios with assimilation (S_1, S_2, S_3 and S_4), one scenario without assimilation (Control) and the twin model (Truth).

3.5 Discussion

Small ensemble sizes can lead to inaccurate covariance estimates, while larger ensembles are computationally heavy (Keppenne et al. 2008). An ensemble size of 100 was chosen, but this may need to be reduced in an operational setting, and simulations in our work suggest that the EnKF estimates are stable with at least 50 ensemble members. Fundamentally, the EnKF provides an appropriate solution for the purposes of this work, estimating parameters and states in the model that are not directly observable (Balchen 2000). Future work may sequentially estimate the swimming speed parameter to further improve the model predictions.

Although the IBM states were perturbed with Gaussian errors, upon simulation the distribution of the ensemble of states becomes non-Gaussian. However, while the EnKF implicitly assumes a Gaussian state-space, it provides good approximate solutions in cases where systems violate this assumption (Katzfuss et al. 2016). Additionally, for statistical consistency, measurements are usually perturbed with Gaussian noise. Perturbing measurements in our system produced an excessive amount of non-zero values. Instead, we used unperturbed measurements and applied an inflation factor to model states instead, as described in Evensen (2009). In Paper 3 we employed a deterministic EnKF, which better handles this problem (as described in Chapter 4).

3.6 Conclusion

The twin model experiment demonstrates that it is theoretically feasible to estimate the spatially explicit distribution and abundance of NSSH using the EnKF approach. Uncertainties in the development of the migration were included in extension of the IBM from Paper 1 to an ensemble of 100 instances. This was an appropriate number to represent the process noise in the model. We proposed a simple approach to mapping between IBM microscopic states and the EnKF macroscopic states, limiting the loss of information. This procedure was similar to randomized distribution in Cocucci et al. (2022). The model predictions were improved in this work, where an alternative migration of the NSSH was successfully estimated from simulated inaccurate priors. Ultimately, it is demonstrated that IBM estimates can be strengthened with Data Assimilation.

Chapter 4

Model-based estimation of catch potential

4.1 Fisheries-dependent data

Although showing the capacity for the Data Assimilation procedure to correct IBM estimates, real measurements were not utilized in Paper 3. To strengthen model estimates, it's valuable to incorporate measurements from the real fisheries system. As explored in Chapter 1 and later in Chapter 5, fishing vessels have access to breadth of unformalized information about marine ecosystems. In fisheries research, the term "fisheries-dependent data" is used to refer such data, reflecting the unsystematic structure of these datasets. However, fisheries-dependent data provide wide spatial coverage, long time series and variety in target species (Pennino et al. 2016).

For the purposes of this work, integrating fisheries-dependent data with the IBM controls its divergence from reality. In addition, the advantage of our model-based estimation approach is that we can account for observation errors. Thus, concerns of bias or errors in fisheries-dependent data can be somewhat mitigated. The final estimate is a melding of recurrent model and measurement values, and thus is less prone to overfitting idiosyncrasies in the data source.

Inferring patterns from nature is advantageous in modelling work, where clearly identifiable structures in nature or data extracted from nature contain hidden information and memory of system dynamics, which can be used to inform process variables and parameter estimates (Wiegand et al. 2003). This is known as pattern oriented modelling. Explicitly predicting occurrence of ubiquitous species using environmental factors is known as species distribution modelling, and are useful in identifying regions with large, stable, viable populations through spatial maps (Waldock et al. 2022). This chapter touches on pattern oriented and species distribution modelling, where we infer patterns from fishing activity data to use as observation input for the EnKF, thus predicting NSSH distribution patterns in real-time.

4.2 Relating fishing activity to stock distributions

The primary method of linking fishing activity to stock abundance is in the form of Catch per Unit Effort (CPUE) indices, which assumes the catch obtained per unit effort input (e.g. fuel consumption) is proportional to the abundance of fish scaled by a catchability parameter. Give survey data is costly and non-exhaustive, management advice often relies heavily on CPUE as an indirect measure of relative abundance of fish stocks (Hilborn and Walters 2015). CPUE is often imbalanced due to technological innovations that increase catchability (e.g. fish aggregating devices in tuna fisheries), or decreases due to decline in stock for reasons unrelated to fishing, and thus there have been efforts to standardize CPUE using statistical models, relating catch rates to geographical coordinates and other explanatory variables (Maunder and Punt 2004, Maunder et al. 2006). In CPUE estimation, zero catch values are often introduced as small perturbations or omitted, meaning that the information on non-catch events are not explicitly defined (Maunder and Punt 2004).

In more recent works, there has been a focus on improving effort estimates, through mining patterns from the vessel trajectories themselves. Many projects are explicitly analysing fishing vessel activity at the individual level, notably Global Fishing Watch, a global monitoring system that aims to increase transparency of human activities at sea, through monitoring of fishing effort.[1]. Several studies have used fishing vessel activity as input to classification algorithms which categorize large-scale movements such as searching, steaming and fishing (Bez et al. 2011, de Souza et al. 2016, May Petry et al. 2020). These algorithms use vessel specific data like catch, heading, speed and position from positional systems and catch logs, to discriminate between fishing and non-fishing events (de Souza et al. 2016). Furthermore, observers on board vessels may expertly validate and label these events (Walker and Bez 2010). The innovation of these methods are the high spatial and temporal resolution of predictions, which may be useful in dynamic monitoring of ocean resources (Kroodsma et al. 2018).

Related work from Adibi et al. (2020) used semantic trajectories and machine

[1]See: https://globalfishingwatch.org/about-%20us/

learning methods to relate environmental variables, daily catch reports and vessel activity to CPUE estimates. Similarly, in Paper 3, we used patterns from fishing activity as input to an Artificial Neural Network (ANN) and normalized catch data was set as target output. The predicted values were converted to synthetic measurements compatible with the assimilation setup, representing both presence and absence of NSSH. Assimilation of synthetic measurements was shown to strengthen the model predictions in cases when the model was inaccurate. Additionally, we converted the assimilated fields to representations representing catch potential of ocean areas, which is a useful input to decision support.

4.3 Assimilating synthetic data generated from neural network output

4.3.1 Synthetic measurements

Automatic Identification System (AIS) data for 186 vessels was accessed from the Norwegian Coast Guard for a total of six years from January 2015 to December 2020, covering the Norwegian Exclusive Economic Zone. The sample consisted of vessels primarily targeting NSSH with purse seiners and pelagic trawls, of both coastal and oceanic fleets. The data was interpolated over 10 minute intervals.

Spatial, temporal and motion related features were calculated from AIS data for years 2015:2018, to generate a $p \times q$ matrix of input data v_i for a shallow ANN to predict normalized NSSH densities as output a_i, a $p \times 1$ vector of estimates. Each row of p represented the features in one 4 km^2 grid cell of daily activity of all vessels present. The target output t_i was the min-max normalized catch recordings of NSSH in the corresponding grid cell. The network minimized mean squared error between network outputs a_i and target outputs t_i:

$$F = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{4.1}$$

After minimizing F, ANN predictions a_i were converted to a $p \times 1$ measurement vector d. We calibrated the top 10% of percentiles of a_i as a threshold to determine non-zero catch values, where above this value d is a Gaussian random number is selected with mean and standard deviation based on t_i. Below this threshold, values were treated as absence with a value of zero for d (Figure 4.1). The vector d was calculated for the model simulation period of 15.01.2020 - 28.02.2020 and was assimilated with model forecast estimates for the same period. The transformation from a to d was based on the assumption that the filter can't really distinguish between high and low presence values, and furthermore, we needed a way to map filter outputs to reasonable values for model corrections. This is why we used a standard value perturbed with random Gaussian measurement noise.

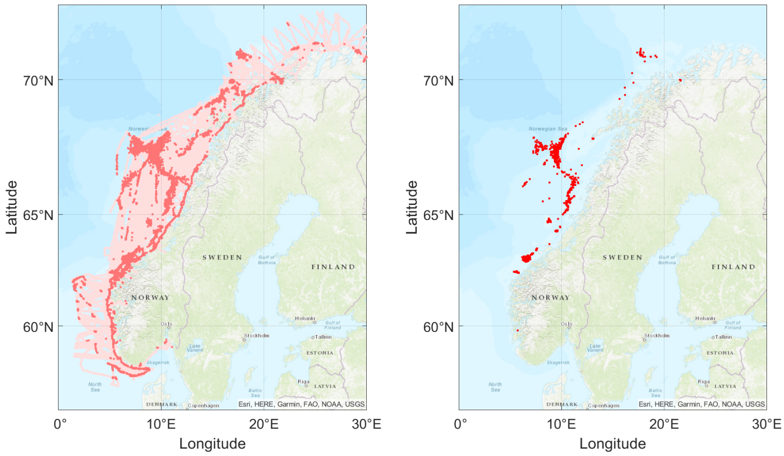


Figure 4.1: Illustration of the coverage of fishing vessels relative to the catch positions recorded in electronic logs. The left panel shows the syntetic measurements predicted by the neural network, with light red points indicating absence and dark red indicating presence. The predictions uses the independent AIS dataset from early January to late February 2020. The panel on the right shows electronically logged catch points for the same period.

4.3.2 Assimilation of synthetic measurements

One issue with using observations of fish densities is the non-negative nature of measurement values in d. Observations are usually perturbed with Gaussian noise, leading to an $m \times N$ matrix of measurement values D. Perturbing zero values leads to instabilities in the IBM, and so unperturbed observations are used in the EnKF. In Paper 2, an inflation factor was implemented to maintain variance in model. However, using synthetic measurements, we had access to approximately 2000 incoming measurements per day. For statistical consistency, this required a reformulation of the EnKF, and we used the variant known as the deterministic EnKF (Sakov and Oke 2008):

$$
\begin{aligned}
\bar{X}^a &= \bar{X}^f + K(d - H\bar{X}^f) \\
A^a &= A^f - \frac{1}{2}KHA^f \\
X^a &= A^a + \bar{X}^a
\end{aligned}
\tag{4.2}
$$

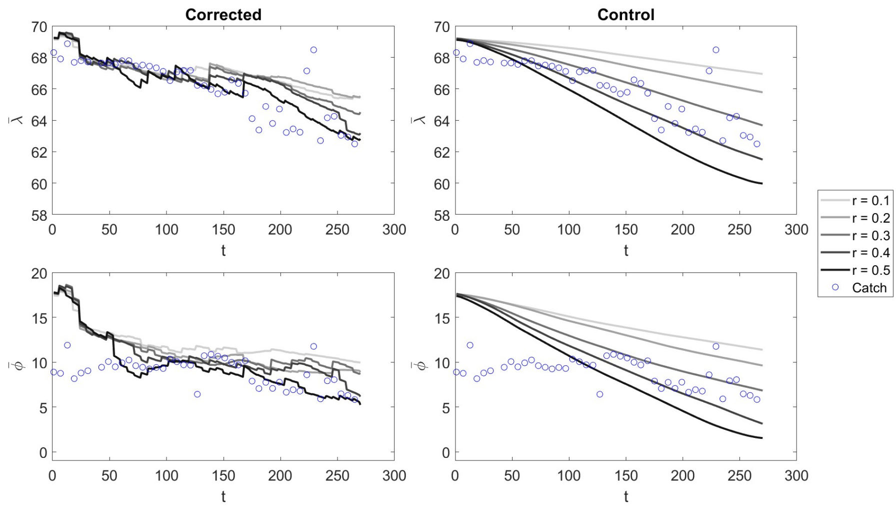


Figure 4.2: The $\bar{\lambda}$ and $\bar{\phi}$ values at each model time step for $\Omega = 10000$. The different r scenarios are shown ascending from light to dark grey lines, with the centre of gravity of catch values displayed at daily increments.

where X^f and A^f are calculated as in Equations 3.4, and K is calculated as in Equations 3.5. The Equations 4.2 replace Equation 3.6 in calculation of the X^a estimate, and this removes the need to inflate X^a as in Equation 3.7.

4.3.3 Simulations and Analysis

The model was simulated with a number of scenarios, based on a set of various standard deviations in observation errors (Ω) and swimming speed parameter values (r). Each simulated scenario was designed with the same random seeded numbers, so any variation was a function of these two parameters. The corrections based on d had a clear effect on the model spatial distribution when comparing against a control model without assimilation. The latitude and longitude centre of gravity ($\bar{\lambda}$ and $\bar{\phi}$) converged on similar values when there was a low Ω value (Figure 4.2). In comparison, there was less convergence with higher Ω values and $\bar{\lambda}$ and $\bar{\phi}$ developed more similarly to the control scenario with no assimilation.

4.3.4 Catch potential

In Paper 3, to demonstrate the ability of model output to inform fishing activity, we mapped X^a to ICES grid cells with approximately 55 km ^{2} resolution, summing the densities over these larger areas. The top 10% of cell were selected as presence of NSSH. Catch potential was calculated as the fraction of cells with at least one catch point with presence predicted by the model during the simulation period.

Compared to a strategy of choosing the top 10% of cells at random (28/280), even poor models performed almost twice as well. The best performances occurred for models with lower Ω values, which performed approximately 5 times better than the random strategy. In general, models that with a higher r values didn't respond as well to corrections, compared to models with lower r values. Figure 4.3 shows a visual illustration of catch potential with varying Ω values over a 5 day period, although we note that the metric used in Paper 3 is calculated daily.

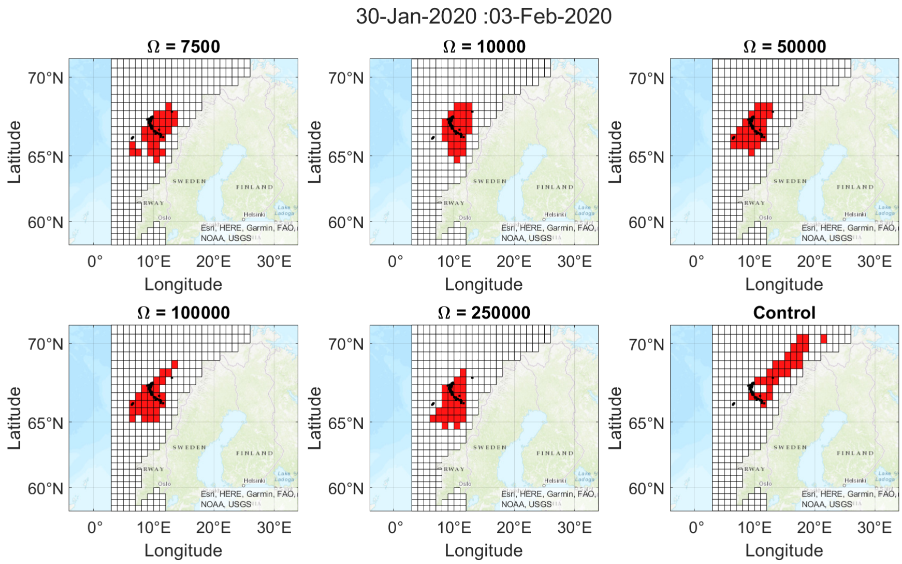


Figure 4.3: Illustration of catch potential calculated from $\bar{X}^a$ for different Ω values and a control scenario, where $r = 0.2$ from 30th of January 2020 to 3rd of February 2020. The black points are the NSSH catch locations during this period. The red squares show the top 10% of model cell values. Note that catch potential estimates were calculated daily in Paper 3.

4.4 Discussion

The assimilation approach presented in this article utilized vessel movement data to generate synthetic measurements using a neural network, providing a large array of measurements, predicting both absence and presence of fish, with which to correct model states. Given assimilated scenarios only have access to synthetic measurements in d trained on independent datasets, with no direct information on catches, this illustrates the power of using assimilation of indirect synthetic data to improve model predictions. The final filtering of model output has immediate utility in informing fishing activity.

However, scaling synthetic measurements according to neural network output a loses connection to absolute densities of fish. Further work is required to build a more accurate filter that can predict absolute values. For example, labelling activity along trajectories provides more refined information on vessel behaviour such as searching, steaming, pumping and fishing, and these features can improve predictions of densities (Adibi et al. 2020). Additionally, we exclusively used catch logs of NSSH in the ANN, but adding multispecies information on distribution of competitors, predators and prey may also improve predictions. The shallow ANN developed is a proof-of-concept for a more complex system of pattern recognition for complimenting modelling efforts.

We have shown that the model estimates can be strengthened with assimilation of synthetic measurements, but given we have incomplete knowledge of fish dynamics, scaling errors and other sources of uncertainty, it is challenging to relate predictions to real processes. For the pragmatic purpose of predicting catch potential, this is not a major issue, but when using such a model to test theoretical considerations in ecology, further work to include more individual states and parameters may be required. Regardless, the Data Assimilation procedure gives insight into the state and parameter values that better fit observations, so if models accurately represent real biological states and parameters, assimilation can sequentially estimate true states in nature.

4.5 Conclusion

The model-based estimation system presented in Paper 3 assimilates synthetic measurements generated from an ANN, with the IBM developed in Paper 1, using the assimilation procedure developed in Paper 2. Where we lack coverage of measurements over the model domain, we show how one can use vessel activity to synthesize measurement values for assimilation. Spatial indices show that the model responds to the assimilation of synthetic measurements, and with lower Ω values, the model is heavily altered. The corrected model forecasts outperform uncorrected forecast estimates when model output is converted to a catch potential metric, especially when the uncorrected forecast estimates are inaccurate. All model estimates perform better than picking fishing grounds at random. Further work can refine the ANN predictions, assess false positive rates in estimates and analyse the utility of the model-based estimation system in decision support.

Chapter 5

Decision support in the commercial fishing industry

5.1 Utilizing Fishers Knowledge

Globally, the spatial extent of fisheries is estimated at four times higher than agriculture (Kroodsma et al. 2018). Thus, knowledge from fishers is vast, both from personal experiences and information systems onboard (Stephenson et al. 2016). Therefore, there is interest in using this in decision-making processes. For example, in small scale Norwegian salmon fisheries, fishers put specific emphasis on passing on local ecological knowledge to next generations (Dyrset et al. 2022). At a larger scale, surveys of Norwegian fishers have been used to quantify the abandonment and pollution caused by fishing gears (Deshpande et al. 2019). Reference fleet programmes are effective at constructing abundance indices of species not surveyed by research vessels (Jones et al. 2022). Finally, positional information (such as AIS used in Chapter 4), is being used to monitor global fishing activity, both to map spatial distribution of fishing effort and screen for illegal fishing activity (de Souza et al. 2016).

However, there are headwinds in utilitzing fishers knowledges. Information is biased to ocean areas where catch takes place, in addition to such data being unsystematically gathered (Hind 2015, Karp et al. 2022). In addition, going from unstructured data, experiences and claims to legitimate and salient information for decision makers is challenging (Röckmann et al. 2015, Reite et al. 2021a). Still, there are possible paths to utilizing information with systematic protocol. As shown in Chapter 1, the modern Norwegian fishing fleets have high coverage of ocean areas and many vessels are equipped with advanced gear and communication technology. Additionally, as shown in Chapter 3 and 4, there are statistical methods that can incorporate the uncertainties in data sources to improve near real-time estimates of fish distributions. Decision support systems can provide the link between research activity and fishing activity to improve our knowledge base of marine ecosystems.

A key goal of the Fishguider project is to leverage such observations from fishing vessels to meet two objectives. Firstly, to supply information that reduces time and fuel spent searching for fishing grounds. Fuel consumption can reach as much as 70% of the annual costs, depending on the vessel and target species, and with better planning and routing, costs can be reduced (Reite et al. 2021a). Secondly, to systematically capture and organize information gathered by fishing vessels while at sea. Meeting these objectives can further collaboration between researchers and fishers, through bidirectional information transfer. Therefore, Paper 4 focused on compiling relevant work on utilizing fishers knowledge for decision support, and shows the work done as part of the Fishguider project in this context.

5.2 Decision support systems

Decision support systems (DSS) are mainly computer-based programs that integrate diverse knowledge sources in order to support complex decision- making processes (Truong et al. 2005, Bal Beşikçi et al. 2016, Granado et al. 2021). DSS have a broad range of applications including aiding manufacturers in delivering products to customers, informing decisions on complex activities within a large organization and improving healthcare delivery in clinical settings (Jacob and Pirkul 1992, Sala et al. 2019, Sutton et al. 2020). In the maritime context, the major applications have been in the shipping industry. They have been mainly applied to optimize routing and scheduling of vessels, which can limit total fuel use and avoid collisions between vessels (Granado et al. 2021).

The designs of DSS vary, but they can be reduced to two principal components common to all systems. Firstly, DSS require knowledge sources that can be used as input. Data-driven sources may come from remote-sensing, weather archive data, national and international databases (Iglesias et al. 2007, Lee et al. 2018). Model-driven knowledge sources can be obtained from sophisticated analysis of data sources (Bal Beşikçi et al. 2016). Secondly, there is a user interface that displays the relevant output in a cogent manner. Usually, the information from knowledge sources are displayed in a number of interactive layers, where the user can manually choose which information to draw for the decision (Granado et al. 2021). As these systems are designed to support rather than execute decisions, there can be several layers of information available for the user at all times (Figure 5.1).

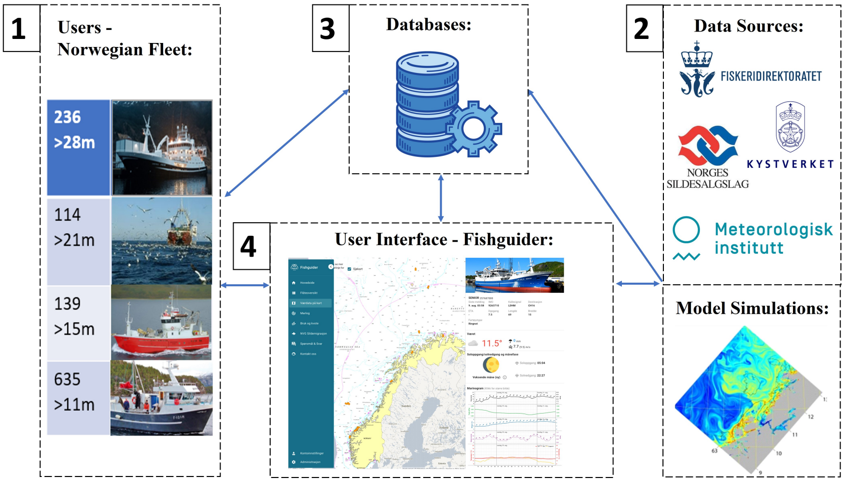


Figure 5.1: Conceptual figure of the components of the Fishguider DSS, where users (1) may both contribute and access information. Model simulations and data sources (2) can inform decisions and in the case of models, are improved through feedback. Structured international and national databases alsonprovide detailed information (3). The user interface (4) collects and integrates these sources for user.

In Paper 4, we summarize some of the main applications of DSS within the fisheries context in Table 1. There a wide range of inputs for DSS. For management applications, such as ecosystem-based approaches, interdisciplinary knowledge, trans-disciplinary partnerships, systems approaches, questionnaires and stock assessment are relevant inputs (for examples, see: (Lane and Stephenson 1998, Azadivar et al. 2009, Dowling et al. 2016). For work similar to our application, in supporting fishing activity and reducing fuel consumption of ships, common information sources used include remote sensing of ocean states, oceanographic modelling and simulations, catch data analysis and vessel speed and heading (for examples, see: (Iglesias et al. 2007, Lee et al. 2018, Reite et al. 2021b)). Considering multiple sources of inputs is preferable where there can be biases and errors in any one taken alone.

In supporting fishing activities, there are three levels of decisions which can be defined. Strategic decisions involve long-term planning of fishing based on the market situation and fishing possibilities (Reite et al. 2021b). Tactical decisions involve planning which fishing grounds to visit, the number of grounds to visit and in what order. Operational decisions involve immediate control of the vessels position, speed and heading and are made on short time scales, such as the manoeuvring of the vessel relative to fish schools (Haugen and Imsland 2019). Intuitively, the information sources will depend on the level of the decision-making. The focus of Paper 4 is the proof of concept for a DSS that can inform strategic and tactical decisions.

5.3 The Fishguider DSS tool

5.3.1 Pilot programme description

The Fishguider project began in 2019 as a science-industry research collaboration between NTNU, SINTEF, the University of Bergen on the research side, and the North Atlantic Institute for Sustainable Fishing (NAIS), an umbrella organisation of fishing companies. Consultation between industry and research partners led to the suggestion to build a DSS tool to reach objectives defined in Section 5.1.

The DSS designed in the Fishguider project is a proof of concept, with a 19 vessels from NAIS now utilizing the pilot system. The classes of the vessels are: 6 coastal vessels, 3 large coastal vessels, 7 ocean-going trawlers and 3 ocean-going purse seiners. A total of 16 vessels are above 21m in length. These classes determine the quotas and areas where the vessel operates. For example, ocean-going vessels cannot operate within fjords without special permission.

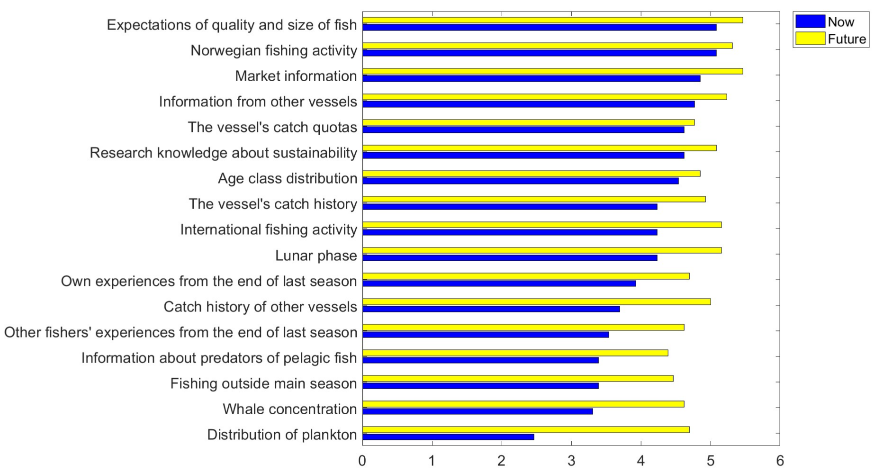


Figure 5.2: Important factors in strategic decision-making from the perspective of Fishguider partners. Blue bars indicate the average importance today, and the yellow bar indicates the average importance predicted in the future (N = 13).

5.3.2 Fishers knowledge in design of DSS tool

To gauge the factors that are important in strategic and tactical decision-making for herring and mackerel a questionnaire was delivered to 13 skippers of vessels in NAIS. It was conducted in Norwegian by phone and translated to English. Respondents were asked to rate the importance of factors when planning the fishing season in advance (strategic decisions) and factors involved when actively searching for fishing grounds (tactical decisions). Each factor was evaluated on a rating scale from 1 to 6, 1 being the lowest and 6 being the highest value. Items were further categorized based on their importance to fishers now and their potential importance in the future. This questionnaire was designed based on project meetings between NAIS and other participants.

The results of the questionnaire revealed that practical considerations such as the vessel's quota, catch history and Norwegian fishing activity are important to strategic decision making now, and are expected to be in the future (Figure 5.2). In terms of tactical decision making, fishers regard communication, market and weather forecasts, and knowledge of other vessels as of high importance (Figure 5.3).

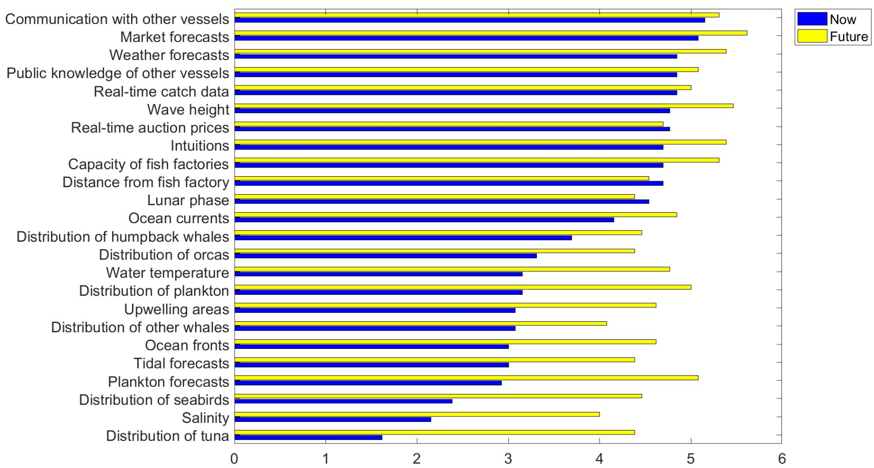


Figure 5.3: Important factors in tactical decision-making from the perspective of Fishguider partners. Blue bars indicate the average importance today, and the yellow bar indicates the average importance predicted in the future (N = 13).

5.3.3 Pilot version of the Fishguider DSS tool

The interface was built to include considerations from both the questionnaire and what was available. In Figure 5.4, a selection of layers pulled from the web portal are displayed. Below are the key developments in the user interface, based on the questionnaire and project meetings:

- Communication between vessels is facilitated through messaging apps.
- Weather forecasts from the meteorological institute.
- The lunar phase from the meteorological institute, which is considered a signal of migration initiation by fishers.
- Model-based estimates: Sea temperature, nitrate, calanus, and the fish migration model developed in Paper 1.

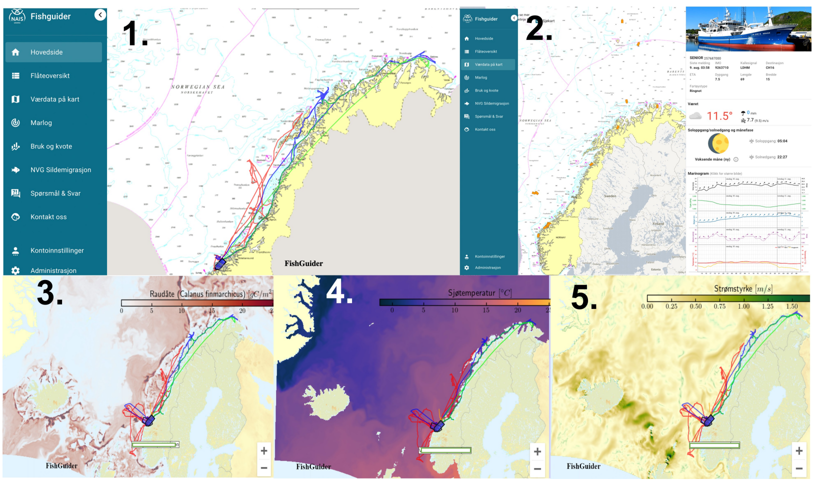


Figure 5.4: A selection of layers from the user interface in the pilot version of the Fishguider DSS tool. Layer 1 shows the main screen on the interface, layer 2 shows the weather forecast data from the meteorological institute and layers 3 - 5 show geographical distribution of copepods (*Calanus finmarchicus*), horizontal current velocities and nitrate. A beta version of the herring IBM is being integrated into the interface.

5.4 Discussion

The DSS tool seeks to provide a technical aid to a complex problem. The research knowledge contributed should be matched with the needs of the users in industry for widespread uptake (Röckmann et al. 2015). Otherwise, we may over-optimize on components of the model structure that have little practical value. Thus, fishers should play an active role in the development of such systems.

Bad weather, rough waves, icing, operation of fishing gear and preservation of catch are all variables that can create issues and variability in costs. Although the catch phase is only one part of the total values chain of fish food production, it is the dominant contributor to pollution (Schau et al. 2009). Further work should investigate more thoroughly which costs of fishing activity are most amenable to the model-based estimation solution provided here.

There are strengths and limitations of the questionnaire as a source of data. It's useful as a way to extract information about the experiential knowledge of fishers. However, their priorities in the "future" category are challenging to interpret since the questions didn't differentiate between the likelihood of achieving different solutions. Additionally, it's generally difficult to predict what will be useful in the future with accurate foresight.

5.5 Conclusion

The Fishguider DSS tool integrates various sources of information to help guide fishing operations. Model-based estimates of physical and biological variables are of interest to fishers involved in this project. The project seeks to increase the number of participants and invite feedback from users. This is particularly useful in the strengthening of the model-based estimation approach described earlier in earlier chapters of this work.

Chapter 6

Conclusion & Future Considerations

6.1 Conclusion

In Chapter 1, the background for the cybernetic approach to fisheries modelling was established, with the theoretical roots stemming from Jens Glad Balchen's material in the 1970s. The model-based estimation procedure Balchen put forward both has theoretical value in prediction of model states not directly observable and pragmatic value in estimation of ocean resources for harvesting. The challenges with estimation of fish dynamics lie in the sparsity of real measurements and non-linear dynamics arising from individual fish behaviour.

There are increasingly more projects, from CRIMAC to Global Fishing Watch, that seek to process and analyse data from commercial fishing vessels. We anticipate an explosion in the quantity of data that will become available. However, sparsity in the spatial distribution of measurements, as well as infrequency in sampling rate will remain a challenge, not to mention uncertainties from sampling bias to calibration of sensors.

Drawing on cybernetic methods of estimation, we have demonstrated how one can combine these new incoming data sources with models to correct and recursively estimate model states, while explicitly accounting for model and observation uncertainties. We illustrate both the challenges and opportunities of implementing model-based estimation to inform fishing activity.

6.2 Future considerations

6.2.1 Model structure

For the pragmatic purpose of our work, the focus was on spatiotemporally explicit estimates. Extension of the IBM to include more biological states is a possible route. Dynamic energy budget (DEB) modelling focuses on quantifying metabolic activity of organisms at an individual level, and assumes mechanisms are not species-specific (Sousa et al. 2010). Energy can be partitioned generally as storage for reserves, maintenance of body structure or reproduction (Kooijman 2009). It's primarily used to study life cycles of individuals based on chemical and physical principles, employing species-specific parameters to determine the allocation of energy. There is freely available matlab data for energetics of animal species, code for estimating parameters and DEB parameters.[1]Extending the model in such a way may allow testing of hypotheses related to the interplay between energetics and fish behaviour. During the thesis, we experimented with DEB states and parameters for herring, but didn't find an immediate use case. Quantifying energy consumed during migrations or the effect of energy consumption on migration patterns are two interesting paths.

The EnKF was used to estimate IBM states in this work, but it can also be used to estimate parameters, such as swimming speed. Simply extending the state matrix to include a vector of the perturbed parameter facilitates sequential parameter estimation, as described in Ward et al. (2016). Each instance of the parameter is corrected based on the covariance structure of the model state variables. As the parameter is not measured directly, the correction is indirect, but this method is still useful in sequentially calibrating the IBM.

6.2.2 Alternative representations

There are two ways in which the model representation may be altered. One may change either the underlying process model (the IBM) or the estimation model (the EnKF). Although we laid out the advantages of the IBM-EnKF approach, there are possible alternatives that may also be suitable for model-based estimation. For one, the IBM represents the behaviour mechanics of super-individuals at each simulation step and a coarser model that represents the dynamics without modelling super-individuals (e.g. Eulerian represenations) may produce similar predictions. No mapping between the IBM and the EnKF is necessary if densities are explicitly modelled. This comes at the cost of being less intuitive and losing the mechanics of individuals. Furthermore, IBMs are the preferred modelling setup for studying fish migrations in fisheries research.

[1]See DEB code at Add my pet: https://www.bio.vu.nl/thb/deb/deblab/add_my_pet/

As for the estimation procedure, a particle filter is an alternative approach that uses Monte Carlo methods to sample from complex high-dimensional distributions (Moral and Doucet 2014). The term particles is used to describe a large cloud of random samples, and at each simulation step, particles with high potential values are duplicated, while those with low values are removed. It is possible to apply this method to the sequential estimation of biological systems. The EnKF was used as it has been successfully used in estimation of high dimensional, non-linear systems.

6.2.3 Catch potential

For estimating catch potential, a simple transformation from model output to an occurrence index was employed. To further strengthen the predictive capacity of the ANN, the behaviour of vessels may be classified. An example of classification of vessel behaviour can be found in de Souza et al. (2016), where fishing activities were detected using vessel speed as an observation input to a Hidden Markov Model. Alternatively, AIS data can be passed to an unsupervised clustering algorithm that groups activities. Applying such filters to the vessel data may provide more refined information for generating synthetic measurements.

More attention may be paid to exactly how model results can be transformed into patterns that are useful for fishing vessels. Delving further into identifying empirical patterns from predictive models may yield useful findings (Gallagher et al. 2021). For example, the Fuzzy Kappa statistic can calculate spatial autocorrelations and find agreement between spatial maps of distributions (Hagen-Zanker 2009). In addition, Fuzzy Inference has been discussed in project meetings as a feasible approach for predicting large-scale spatial patterns of fish stocks.

6.2.4 Collaboration

Finally, it is important to emphasize that collaboration between industry, researchers, authorities and other Norwegian institutions are central to achieving solutions that are mutually beneficial. Operationalizing the model-based estimation system can assist decision-making for fishers, while at the same time automatically capturing and systematizing information for research analysis. It can thus achieve a positive sum solution that is mutually beneficial for researchers and fishers. Further establishing recurrent meetings and workshops between project participants is essential, as well as analysing the successes and failures of the modelling work.

Chapter 7

Articles

Paper 1

Kelly, C., Michelsen, F. A., Kolding, J., and Alver, M. O. (2022). Tuning and Development of an Individual-Based Model of the Herring Spawning Migration. Frontiers in Marine Science 8, 754476. doi:10.3389/fmars.2021.754476

frontiers in Marine Science

ORIGINAL RESEARCH
published: 13 January 2022
doi: 10.3389/fmars.2021.754476

Tuning and Development of an Individual-Based Model of the Herring Spawning Migration

Cian Kelly[1], Finn Are Michelsen[2], Jeppe Kolding[3] and Morten Omholt Alver[1]*

[1] *Department of Engineering Cybernetics, NTNU, Trondheim, Norway,* [2] *Fisheries and New Biomarine Industry, SINTEF Ocean, Trondheim, Norway,* [3] *Department of Biological Sciences, UiB, Bergen, Norway*

Norwegian spring spawning herring is a migratory pelagic fish stock that seasonally navigates between distant locations in the Norwegian Sea. The spawning migration takes place between late winter and early spring. In this article, we present an individual-based model that simulated the spawning migration, which was tuned and validated against observation data. Individuals were modelled on a continuous grid coupled to a physical oceanographic model. We explore the development of individual model states in relation to local environmental conditions and predict the distribution and abundance of individuals in the Norwegian Sea for selected years (2015–2020). Individuals moved position mainly according to the prevailing coastal current. A tuning procedure was used to minimize the deviations between model and survey estimates at specific time stamps. Furthermore, 4 separate scenarios were simulated to ascertain the sensitivity of the model to initial conditions. Subsequently, one scenario was evaluated and compared with catch data in 5 day periods within the model time frame. Agreement between model and catch data varies throughout the season and between years. Regardless, emergent properties of the migration are identifiable that match observations, particularly migration trajectories that run perpendicular to deep bathymetry and counter the prevailing current. The model developed is efficient to implement and can be extended to generate multiple realizations of the migration path. This model, in combination with various sources of fisheries-dependent data, can be applied to improve real-time estimates of fish distributions.

Keywords: individual, model, migration, observation, catch, tuning, comparison, spawning

OPEN ACCESS

Edited by:
Jie Cao,
North Carolina State University, United States

Reviewed by:
Chongliang Zhang,
Ocean University of China, China
Francisco Leitão,
University of Algarve, Portugal

****Correspondence:***
Cian Kelly
cian.kelly@ntnu.no

Specialty section:
This article was submitted to Marine Fisheries, Aquaculture and Living Resources, a section of the journal Frontiers in Marine Science

Received: *06 August 2021*
Accepted: *15 December 2021*
Published: *13 January 2022*

Citation:
Kelly C, Michelsen FA, Kolding J and Alver MO (2022) Tuning and Development of an Individual-Based Model of the Herring Spawning Migration. Front. Mar. Sci. 8:754476. doi: 10.3389/fmars.2021.754476

1. INTRODUCTION

Incomplete knowledge or inadequate access to time-sensitive spatial distributions can result in inefficient harvesting of fish stocks. This is especially true of migratory species that migrate vast distances for periods of their life cycle. Such species prove difficult to quantify, manage, and exploit given flexibility and variability in migration strategies (Fernö et al., 1998; Tamario et al., 2019). However, as fishing operations advance, fishing vessels attain access to more fine grain sources of data (acoustic, satellite etc.). Specifically, acoustic technology now makes use of a multi-beam system that can resolve multiple targets at once (Chu, 2011). Such data is an untapped resource for understanding the development of stocks throughout a fishing season, as it provides good coverage of stocks in real-time, all-year round (Pennino et al., 2016). One of the limitations of such data is bias toward presence data. Regardless, developments in technology and access to additional observations will likely improve our ability to quantify abundance and distribution.

The Norwegian spring spawning herring (NSSH) (*Clupea harengus L.*) is an Atlanto-Scandian herring that is mainly distributed along the Norwegian, Faroese, and Icelandic coast. NSSH is a schooling, migratory pelagic stock that move large distances during it's life cycle. The principal fishery for adult NSSH is along the western Norwegian coast prior to and during the spawning season (Dragesund et al., 1980). NSSH is one of the largest stocks in the entire Atlantic, and one of the most commercially valuable (Touzeau et al., 2000). Although the bulk of the revenues and employment are in Norway, this stock is also harvested by Iceland, Russia, the Faroe Islands, and the EU (Bjørndal et al., 2004). The lack of spatial information on abundance of this species has contributed to unsustainable harvesting in the past, specifically to unforeseen stock collapses (Fernö et al., 1998). For example, collapse of the stock in the 1960s has been attributed to overfishing resulting from advances in harvesting technology, suboptimal management and climactic fluctuations (Arnason et al., 2000; Toresen and Østvedt, 2000; Fiksen and Slotte, 2002; Bjørndal et al., 2004). The fishery closed in the 1970s to allow recovery of the stock (Bjørndal et al., 2004). One issue is that attaining reliable spatio-temporal data is difficult due to interannual fluctuations in abundance and changing migration patterns (Dragesund et al., 1980). Modelling migration patterns can improve time-sensitive estimates of stock distributions.

Bauer and Klaassen (2013) define migratory behaviour as being persistent and directional with distinct departing and arrival behaviours. The functions of such migratory behaviour include reproduction, feeding, and avoidance of predators (Tamario et al., 2019). NSSH conserve energy through overwintering in fjords in northern Norway, prior to moving southwards toward spawning grounds along the Norwegian coast in spring, before migrating westwards to feed offshore for the summer months (Varpe et al., 2005). During each installment of the annual cycle, some drivers are likely to supersede others. For example, during the overwintering period, movement is limited and energy is conserved. In contrast, the spawning migration takes place across a distance of approximately 800 km counter-current and is characterized by rapid energy depletion (Slotte and Fiksen, 2000).

Mechanistic models incorporate the main mechanisms by which discrete individual components in a system may behave through fundamental assumptions and equations. Differential/difference equations plus stochastic noise are common features used to explore variability in these models. The utility of a model is gauged through its capacity to match observations from the real system. Transitioning from theory to application of such models demands a series of stages of refinement through tuning/calibration and validation of the model (Baker et al., 2018). There is much theory about what drives the spawning migration of NSSH. There is a need to translate this theory into model output that provides estimates throughout the season.

Individual-based models (IBMs) are a class of mechanistic models that are built to explore the emergent properties at the population level, arising from individuals interacting with other individuals and their surrounding environment (Grimm and Railsback, 2005). IBMs have been used to predict spatial patterns of many migratory fish species during periods of their life cycle (Barbaro et al., 2009; Politikos et al., 2015; Boyd et al., 2020). Coupling models of physical oceanography with IBMs is an effective method for simulating the complex interactions between individuals and their local environment. Furthermore, physical models simulate the main environmental conditions that force individual behaviour and the physical transport of larval stages and prey items (Giske et al., 2001; Alver et al., 2016). In the case of many migratory pelagic species, environmental variables such as currents and temperature have been demonstrated to provide useful information for successful navigation of individuals between distant areas (Barbaro et al., 2009; Tu et al., 2012). There is evidence NSSH use similar mechanisms (Fernö et al., 1998; Slotte and Fiksen, 2000). For Icelandic capelin, current and temperature data, without the use of forcing terms, reproduced the observed migration route (Barbaro et al., 2009). Without using forcing terms, one can easily add noise to IBM components and extend simulations more efficiently. The novel use of multiple realizations of the IBM, together with observations, can improve estimates of the NSSH distribution in real-time. These estimates can support stakeholder decisions in the fishing industry. The IBM developed in this article shall be used in this way.

This paper describes the development of an IBM of the NSSH spawning migration, centred on an individuals response to environmental forcing. The focus is on modelling memory- and gradient- based reactive mechanisms (Fernö et al., 1998). This work also explores sensitivity of the migration to initial conditions, specifically initial location. Following, model densities are compared to observed patterns from 2015-2020 catch data using geospatial indices. Ultimately, the IBM was developed as a tool for comparison and correction with real-time observations, so this work focuses on model agreement with available observations, and where and how disagreements may be resolved. As mentioned before, this can support efficient, sustainable harvesting of NSSH. The description of the model is informed by IBM protocol developed by Grimm et al. (2006).

2. METHODS

2.1. Purpose and Structure of Model

The purpose of the model is to predict spatial patterns of abundance for the spawning migration along the Norwegian coast. This model provides discrete estimates across the model area that can be used to compare against concurrent observations. Furthermore, this model is designed to improve estimates when observations become available. A brief schedule of the main operations is presented in **Table 1**. The movement of individuals is modelled by changes in orientation and horizontal speed. In particular, the response of individuals to the Norwegian Coastal Current (NCC), along with temperature and depth gradients are modelled. The individuals also utilize knowledge of previous states, such as orientation angle, when moving.

TABLE 1 | Overview of the main components of the model algorithm with reference to associated sections and equations.

Overview of model algorithm
Input data: • Load environmental data from SINMOD (Section 2.2). • Load survey and catch data.
Initialization: • Initialize parameter values (**Table 3**). • Initialize position and orientation of individuals (Section 2.3). • Initialize individuals on 2D grid in mid-January.
Simulation: • Update date and time. • Access environmental values based on positions of individuals and current time (Section 2.3). • Pass environmental values to functions which calculate the individual's response to cues (Section 2.4.2). • Update horizontal speed, orientation and position respectively (Equation 6, 4 and 1).
Analysis: • Tuning using survey data (Section 2.5). • Comparison with catch data (Section 2.6).

2.2. Model System

Estimates of environmental conditions were loaded from SINMOD, a physical oceanographic model that is based on the primitive Navier–Stokes equations, and uses a z-coordinate grid (Slagstad and McClimans, 2005). The configuration used has 970 × 635 horizontal grid cells with 4km resolution and centres on the Norwegian Sea. The model is divided into 34 vertical depth layers. The IBM developed in this paper modelled individuals in a 2D environment where position was updated on a continuous horizontal plane in a Lagrangian approach (**Figure 1**). Environmental variables were calculated based on assumptions of the herring's depth preferences, described in section 2.4.1. State variables were updated at discrete time steps of Δt = 4h. Temperatures and current speeds were extracted from SINMOD output from 2015 to 2020, and along with the bathymetry field of the model area, drove changes in fish movements.

The model was developed in MATLAB, which is a matrix-based programming language that is suited for iterative analysis involving numerous matrix operations. Below, the development of the model is outlined in regards to the model system, the main equations, parameters and state variables. Thereafter, we explore tuning and comparison against observations of the real system.

When referring to vectors that can take on continuous values, x and y indices will be used, while the index j will indicate the discrete linear index of a grid cell, ranging from 1 to the number of elements in the grid. Finally, boldface characters denote vectors.

2.3. State Variables

The individual state variables used were position $\mathbf{p}$, orientation angle θ, horizontal speed r_b, and horizontal speed offset r_o. Position was updated at each discrete time k, with time step Δt:

$$\mathbf{p}[k+1] = \mathbf{p}[k] + \Delta t \left(\mathbf{v}_f[k] + \mathbf{v}_c[k]\right) \tag{1}$$

where:

$$\mathbf{v}_f[k] = -\Phi \mathbf{v}_c[k] + \mathbf{v}_b[k] \tag{2}$$

where $\mathbf{p}$ is a vector $[\mathbf{p}_x \ \mathbf{p}_y]^T$ with the x and y coordinates of the individual in the continuous space of the model grid and $\mathbf{v}_b$ is a vector $[\mathbf{v}_{bx} \ \mathbf{v}_{by}]^T$ with the horizontal velocity components of an individual fish in the x and y directions, based on behavioural cues. Similarly, $\mathbf{v}_c$ is a vector $[\mathbf{v}_{cx} \ \ \mathbf{v}_{cy}]^T$ with the horizontal current velocity components in the x and y directions. The superscript T denotes the transpose of the vector. The spawning migration proceeds counter to the NCC (Slotte and Fiksen, 2000). This is modelled by the term $-\Phi\mathbf{v}_c$ that adds a counter-current component to the horizontal speed controlled by the parameter Φ. An individual's realized swimming velocity $\mathbf{v}_f$ is composed of the counter current term and behavioural responses from $\mathbf{v}_b$. This formulation demands individuals respond to the prevailing current with higher priority, relative to other cues. To prevent unrealistic dynamics in the first-order approximation of velocity ($\mathbf{v}_f + \mathbf{v}_c$), the short Δt of 4h was used. The vector $\mathbf{v}_b$ was calculated as:

$$\mathbf{v}_b[k] = r_b[k] \left(\begin{bmatrix} \cos(\theta[k]) \\ \sin(\theta[k]) \end{bmatrix} \right) \tag{3}$$

where r_b is the horizontal speed of the individual in m s^{-1} and θ the orientation angle according to gradient cues. As indicated before, when $\|\mathbf{v}_c\|$ approaches zero, the individuals approach a speed of r_b with the orientation angle θ. The angle θ was updated as follows:

$$\begin{aligned} \theta[k+1] &= \alpha\theta[k] + (1-\alpha)G[k] \\ \theta[k] &= \angle \mathbf{v}_b[k] \end{aligned} \tag{4}$$

where α is a weighting parameter, and G is the angle of the vector inputs calculated from near-field gradients, as explored in section 2.4.2. It is likely NSSH base their movements on comparison of conditions from previous experience and present information when calculating their new orientation (Fernö et al., 1998). To account for this, α acted as a low-pass filter, avoiding erratic changes in θ, similar to a formulation by Føre et al. (2009). Furthermore, r_b was calculated as a random process with a deviation r_o from the cruising speed:

$$r_b[k+1] = \bar{r}_b + r_o[k+1] \tag{5}$$

where $\bar{r}_b$ was the cruising speed in m s^{-1} of the individual. The value was calibrated during the tuning procedure in section 2.5. The offset r_o was then calculated as a Gauss-Markov process with exponential auto-correlation. This meant r_o and r_b were correlated with recent values. The offset was included to model randomness in the horizontal speed. It was updated as follows:

$$r_o[k+1] = e^{-\beta\Delta t} r_o[k] + \sqrt{1-(e^{-\beta\Delta t})^2}\, \mathcal{N}(0, \sigma^2) \tag{6}$$

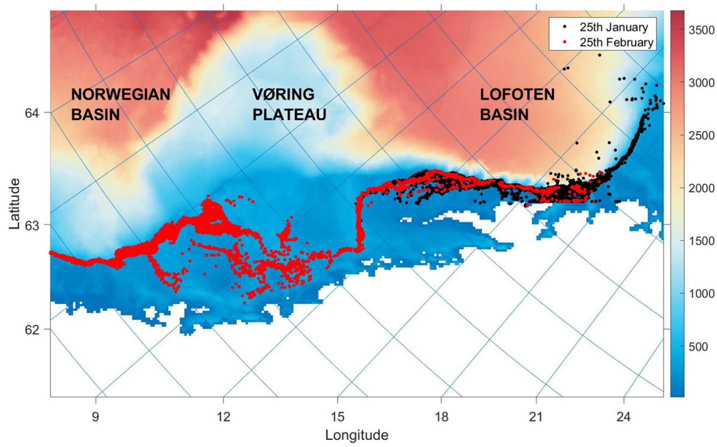


FIGURE 1 | The continuous model system with individuals plotted at two time stamps in the same simulation scenario. The colourmap and colourbar display the bottom depth of grid cell *j* in metres, while the lines of latitude and longitude extend from ticks along the y and x axes, respectively. The map displays the Norwegian coast and Norwegian Sea from 62 to 72 degrees north, including bathymetry features important in ocean circulation.

from a normal distribution $\mathcal{N}$ with zero mean, a standard deviation σ in swimming speed and shaping parameter for auto-correlation β.

2.4. Environmental Forcing

The spawning migration follows overwintering, a stationary period characterized by slow swimming speeds and low energy use. NSSH mainly spawn in southern regions of coastal Norway before migrating westwards to feed over summer. The spawning migration begins early to mid-January and spawning usually begins in late-February/early-March (Dragesund et al., 1980). Temperature and current data for the period from mid-January to the end of February (2015–2020) were loaded from SINMOD, along with the bathymetry. Below, the responses that play a role in the spawning migration are outlined.

2.4.1. Environmental Values

Depth is an important variable as it influences the temperature and currents that an individual experiences. Environmental values were calculated as a linear combination of values taken from a depth layer in the upper water column (0–75 m) and the lower water column (75–300 m):

$$\mathbf{v}_c = (1-d)\mathbf{v}_{c_{l_1}} + d\mathbf{v}_{c_{l_2}} \tag{7}$$

$$T = (1-d)T_{l_1} + dT_{l_2} \tag{8}$$

where $\mathbf{v}_c$ is the horizontal current velocity, T is the temperature, l_1 and l_2 are the vertical indices of layers sampled in the upper and lower water column, respectively, d was the fraction of daylight at the sampled time and latitude (hours of daylight/24). This approximation allows variability in environmental values, rather than use of values from a constant depth layer. The choice of d as a variable reflected the individuals need to spend time in layers that provide light conditions which facilitate the capacity to school and sense local surroundings (Huse and Ona, 1996). The vertical indices l_1 and l_2 were chosen based on $\min(\|\mathbf{v}_c\|)$, as herring have the ability to choose depth layers with favourable current speeds (Nøttestad et al., 1996).

2.4.2. Environmental Cues

Apart from the direct response to current, individuals responded to the temperature and bottom depth gradients at their location. The orientation angle G was calculated as:

$$G[k] = w\angle\mathbf{G}_D[k] + (1-w)\angle\mathbf{G}_T[k] \tag{9}$$

where the angles $\angle\mathbf{G}_D$ and $\angle\mathbf{G}_T$ are the gradient dependent orientation angles calculated based on the depth and temperature gradients, respectively, and w is the weighting parameter on $\angle\mathbf{G}_D$.

Bottom depth: The NSSH spawning migration develops southward alongside the continental slope (Slotte and Fiksen, 2000). Herring are physostomous with an open swim bladder, which facilitates more rapid vertical movements (Blaxter, 1985; Nøttestad, 1998). Vertical escape is considered central in predator avoidance (Langård et al., 2014). For these reasons, $\mathbf{G}_D$ was used to direct individuals movements perpendicular to the direction of the depth gradient, in the southerly direction:

$$\mathbf{G}_D = \begin{cases} -\dfrac{\nabla\mathbf{D}[k]}{\|\nabla\mathbf{D}[k]\|}, & \text{if } D \geq 400 \\ \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right)\dfrac{\nabla\mathbf{D}[k]}{\|\nabla\mathbf{D}[k]\|}, & \text{otherwise} \end{cases} \tag{10}$$

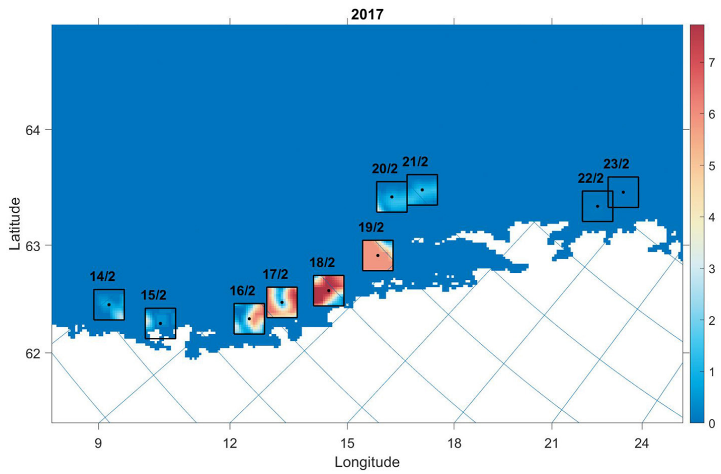


FIGURE 2 | Transformed survey values used to compare against model values on specified dates in 2017. Black dots indicate the weighted sum of trawl values on the date transcribed above the box. The black lines demarcate the outer boundary of cells included for the comparison. The colourmap indicates estimated number of individuals in grid cell *j*.

where $\nabla\mathbf{D}$ is the bottom depth gradient and D is the bottom depth at grid point j. The first case reorients individuals toward the coast, while the second case directs individuals southwards along isobaths. The y component of $\nabla\mathbf{D}$ is multiplied by sign(y) prior to the calculation in the second case.

Temperature: NSSH avoid low temperatures and higher temperatures are associated with superior body condition (Fernö et al., 1998). Therefore, individuals oriented toward higher temperatures, based on the local gradient. If temperatures reached an upper limit, the herring reoriented toward cooler waters based on the near field gradient. The function was formulated as below:

$$\mathbf{G}_T = \begin{cases} -\dfrac{\nabla\mathbf{T}[k]}{\|\nabla\mathbf{T}[k]\|}, & \text{if } T \geq 12 \\ \dfrac{\nabla\mathbf{T}[k]}{\|\nabla\mathbf{T}[k]\|}, & \text{otherwise} \end{cases} \tag{11}$$

where $\nabla\mathbf{T}$ is the temperature gradient calculated from the T field in Equation (8).

2.5. Tuning

To develop the individuals responses to environmental information, the parameters $[\Phi\ \alpha\ w]^T$ in Equations (2), (4), and (9) were tuned and then analysed. Simulations with 2017–2020 environmental values and 4 separate initialization scenarios were investigated. In order to model reasonable responses of individuals to their environment the model was tuned using a numerical optimization algorithm that took parameter values as input and minimized the deviation between the model and observed distributions at specific time stamps. NSSH survey data from the Norwegian Institute of Marine Research was used for this purpose (Salthaug et al., 2020). The survey data combines acoustic and trawl data to estimate abundance in predefined strata areas. To ensure consistency, 2017 and 2018 estimates were used for tuning, when the dates sampled were in the last two weeks of February. 2019 and 2020 data provided an independent data set to validate the optimization.

2.5.1. Tuning: Setup

To allow fine grain comparison on specified dates, the following transformation of strata values was carried out, converting observation estimates into numbers per grid cell j:

- A 12 km^2 sliding penalty used acoustic zero values to penalize areas sampled with low abundances. This meant that grid cells with strata values in close proximity to those with zero acoustic values were set to zero.
- A 20 km^2 sliding mean was then calculated to smooth out areas and maintain spatial patterns of densities.
- The weighted sum of trawl data was used to centre the comparison for single days from the survey. A 52 km^2 grid was drawn around the centre as a bin for comparison.
- The transformed observation values were then normalized and scaled according to the number of model individuals (**Figure 2**).

This procedure offered individuals specific objective functions for a set of time stamps on the migration path. Thus, an optimization routine was used to find the minimum f in Equation (12). This algorithm tuned the parameters $[\Phi\ \ \alpha\ \ w]^T$ and

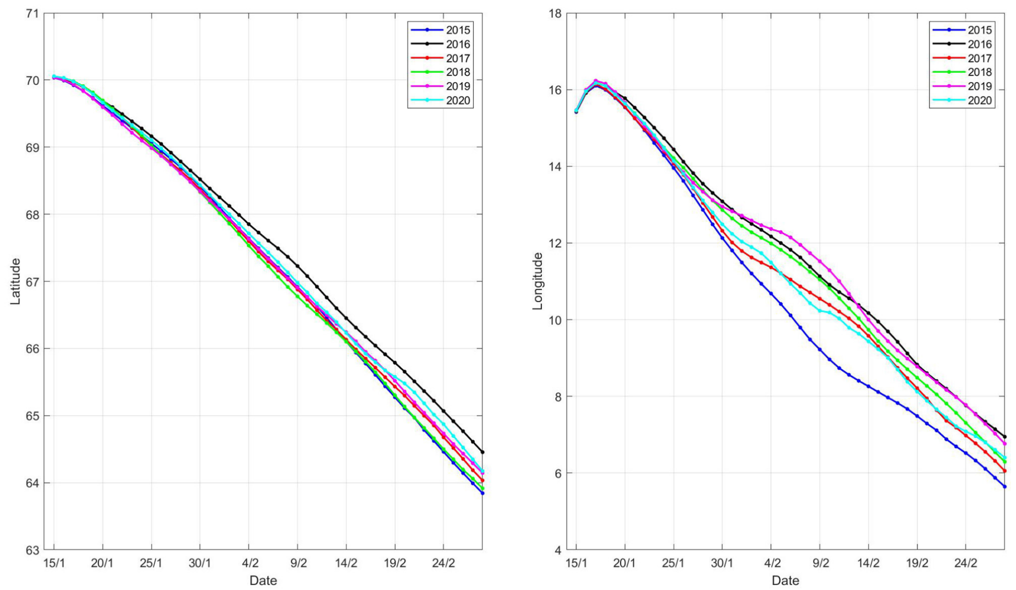


FIGURE 3 | Latitude and longitude centre points during the migration period from S_2. Each point was calculated from the weighted sum of individuals per grid cell j.

constrained them to between 0 and 1. The densities of individuals were compared with the fraction of modelled individuals at the sampled date as follows:

$$f = \frac{\sum_{i=1}^{y} \left(\sqrt{\frac{1}{N} \sum_{j=1}^{N} (x_j - \hat{x}_j)^2} \right)}{y} \tag{12}$$

where i is the year, y is the number of years, N is the number of cells, and x_j and $\hat{x}_j$ are the number of observation and model individuals in grid cell j, respectively. For simplicity, indices indicating day are omitted, where the Root Mean Square Deviation (RMSD) calculation (in parentheses) is executed on the relevant date (**Figure 2**).

2.5.2. Tuning: Simulations

One source of uncertainty is the intialization of the spawning migration. Given we can perform temporal comparisons above, 4 scenarios with different initial locations were selected, based on information from the NSSH survey and the Norwegian Directorate of Fisheries. They represented scenarios where the central mass of individuals were at variable distances from the coast and variable distances north (**Table 2**). This design tested the sensitivity of the migration to initial conditions. The probability of an individuals presence in grid cell j (p_j) was calculated from a gaussian radial basis function with the following equation:

TABLE 2 | Centre points for each scenario.

Scenarios	$\hat{c}$ latitude	$\hat{c}$ longitude
S_1	69	12
S_2	70	15
S_3	71	12
S_4	70	9

$$p_j = \exp\left(\frac{-\|\mathbf{c}_j - \hat{\mathrm{C}}\|^2}{2\rho^2} \right) \tag{13}$$

where $\mathbf{c}_j$ are floored x and y coordinates of grid cell j, $\hat{\mathrm{C}}$ is the floored centre point x and y coordinates, and ρ is a parameter that controls the spread around $\hat{\mathrm{C}}$, which was calibrated in conjunction with the optimization routine. This strategy allowed the fine grain control of initial distribution by proving spatial correlations in p based on distance from $\hat{\mathrm{C}}$. These simulations aimed not to fully describe the distribution prior to migration, but provided insights into how such variation can influence model output.

2.5.3. Tuning: Analysis

Normalized RMSD values between transformed survey values and model output was used to gauge the sensitivity of parameters across to initialization scenarios S_1 to S_4. The parameter values from one scenario was then selected for the remainder of the comparisons. The 2015–2020 SINMOD environmental values were coupled with the IBM to produce 6 distinct realizations.

To inspect the tuning results at high resolution (section 3.2), densities of individuals were post-processed. The number of individuals in each model grid cell j were summed to calculate density per grid cell. Following, a sliding mean calculation for 20 km^2 was used to derive spatial correlations in densities. To visualize the central tendency of the trajectory, the weighted sum of latitude and longitude points of individuals were computed for each day. Thereafter, acoustic density values from the NSSH survey were converted to the relative fraction along all transects for the sampled time. Each grid cell j along the transect was sampled for the number of model individuals. The model and observation values were interpolated within 50 m bins from 0 to 500 m based on the bottom depth in grid cell j. Finally, we calculated the fraction of model and observation values in each depth bin. This calculation can illustrate the spread of model and observation values off-coast. The 16-18th of February 2019 and 2020 were chosen for visual inspection.

TABLE 3 | Model parameter values.

Parameters	Description	Unit	Value
Δt	Time step	h	4
$\bar{r}_b$	Cruising speed	m s^{-1}	0.32
σ	Standard deviation in swimming speed	m s^{-1}	0.1
Φ	Weight of counter-current response	-	0.91
α	Weight of previous orientation angle	-	0.56
w	Weight of depth gradient cue	-	0.72
β	Gauss-Markov time constant	-	0.2
$\hat{\mathbf{c}}$ lat	Initialization centre point latitude	°lat	70
$\hat{\mathbf{c}}$ long	Initialization centre point latitude	°long	15
ρ	Spread around centre point	-	1.57

2.6. Geospatial Comparisons

The model output was compared with 2015–2020 catch data from Norwegian Fisheries Directorate in 5 day ranges. The catch data used were the x and y starting locations of trawling and associated catch weight in kg. The main purpose of this comparison was to investigate where and when the model deviates from observations and what this reveals about the capacity for the IBM to resemble realistic NSSH distributions. The initial model distribution was fixed in each year to focus on relative comparisons.

Observation and model densities were allocated 5 day windows, where catch and model positions and values were

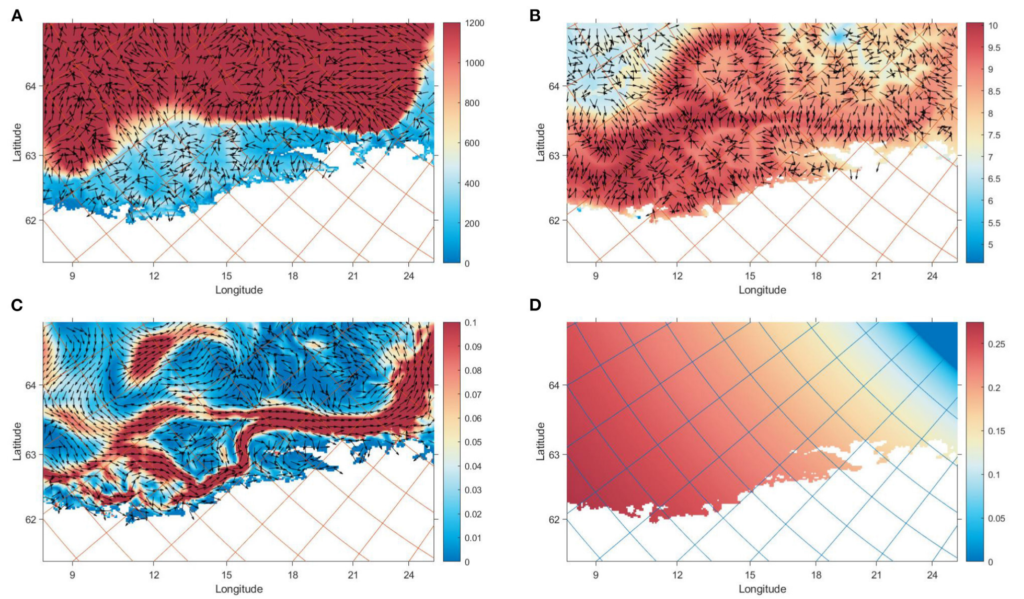


FIGURE 4 | Environmental values that NSSH utilized on the 30th January 2020 relative to their position on the grid. The maps cover the same area as **Figure 1**: (**A**) The bottom depth in m (colourmap) and gradient (arrows) **(B)** The temperature in °C (colourmap) calculated from Equation (8) and gradient (arrows) **(C)** The x and y components of $\mathbf{v}_c$ (arrows) calculated from Equation (7). The colourmap shows the magnitude $\|\mathbf{v}_c\|$ in m s^{-1} **(D)** The fraction of daylight d along the Norwegian coast (colourmap), which was used in Equations (7) and (8).

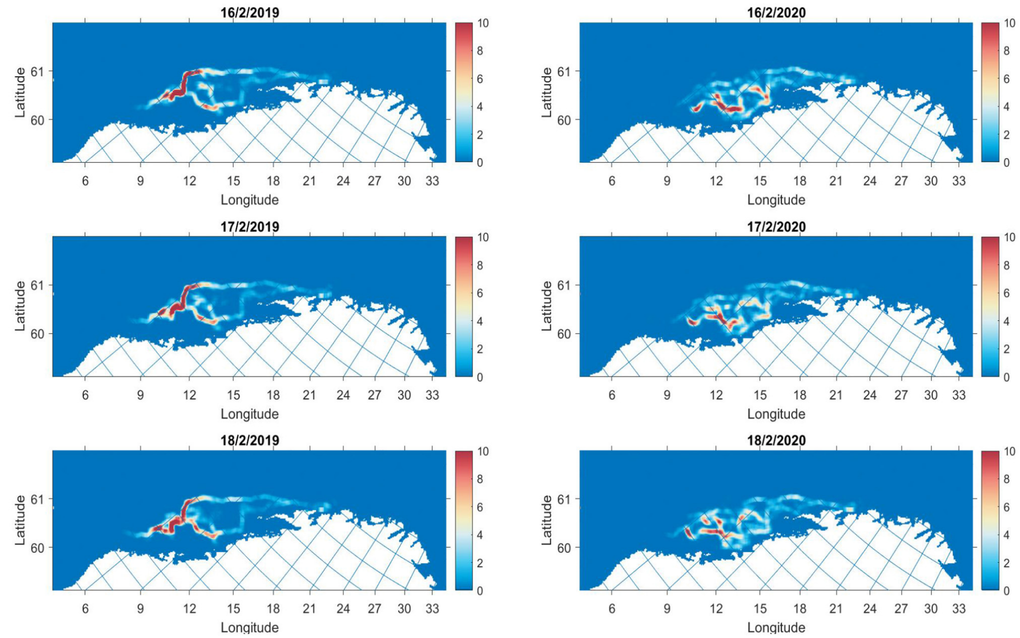


FIGURE 5 | Densities of individuals at 3 time stamps along the migration for 2019 (left column) and 2020 (right column). The colourmap shows the number of individuals in grid cell *j*.

described by their centre of mass. These were described as latitude and longitude centre points. Then, model output was analysed using geospatial comparisons, as described in Woillez et al. (2007). The global index of collocation (GIC) indicates how geographical distinct two ditributions are. It is based on the distance between the centre points and the variance around these centres (inertia). A value of 0 means there is no overlap, whilst a value of 1 implies identical distributions. The average GIC value for each year was used to score the model and observation overlap. In addition, RMSD values indicated the average error for the year. Using GIC and RMSD indices illuminate where the model and observations disagree. It also offers insight into potential limitations of comparing model and observation output.

3. RESULTS

3.1. Sensitivity Analysis

Of the 4 simulation scenarios run, 3 scenarios produced reasonable migration patterns. S_1, S_2, and S_3 performed relatively well (**Supplementary Material**). S_4 is far from continental slope, which is vital information for orientation and therefore couldn't minimize Equation (12) properly. The average Φ value, produced by the tuning process for the four scenarios, was quite high and displayed higher variability (0.78 ± 0.2) compared to α (0.59 ± 0.08) and w (0.65 ± 0.1). This illustrates the sensitivity of the response to the current in relation to starting point. The relatively high α value also shows the importance of fish retaining knowledge of previous states in the migration. These simulations highlight the centrality of the continental slope as a landmark in the migration and how individuals are likely to utilize it. Further comparisons from 2015 to 2020 were made using the fitted parameter values based on S_2 values (**Table 3**).

3.2. Scenario Analysis

3.2.1. Trajectory

There was variation in the trajectories of individuals in the migration, especially with regards to longitude (**Figure 3**). In all cases, the beginning of the migration is quite slow, reflecting the strong current magnitude around the Lofoten basin (**Figure 4C**). For example, in 2020, the Latitude centre point moves approximately 1 degree from 15/1 - 25/1, in comparison with 1.5 degrees from 9/2 - 19/2. In addition, the longitude centre is more consistent amongst years from 15/1 - 25/1, again reflecting the convergence of environmental cues at this stage, especially $\nabla \mathbf{D}$ (**Figure 4A**). There is divergence after this point due to inter-annual variation in T and $\mathbf{v}_c$. This demonstrates that environmental variability can produce distinct realizations of the IBM.

The spatial distribution of individuals illustrates the emergent structures from the simulation (**Figure 5**). For example, in the 2019 simulation, midway through the migration, emergence of

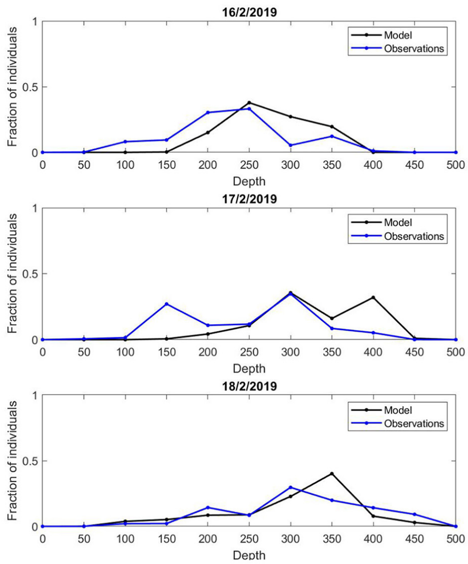


FIGURE 6 | Fraction of estimated model and observation values along acoustic transects at 3 time stamps in 2019 (left column) and 2020 (right column), with associated bottom depth.

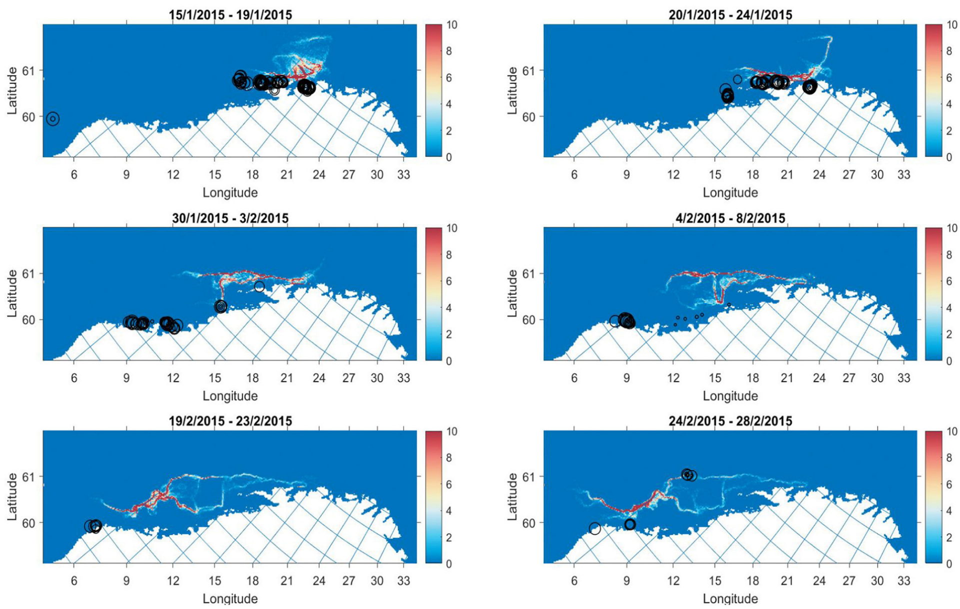


FIGURE 7 | Model output (colourmap) with catch points (black circles) overlayed for selected periods in 2015. Size of circle is scaled to the catch weight. The colourmap gives the average number of individuals in grid cell *j* for a 5 day period.

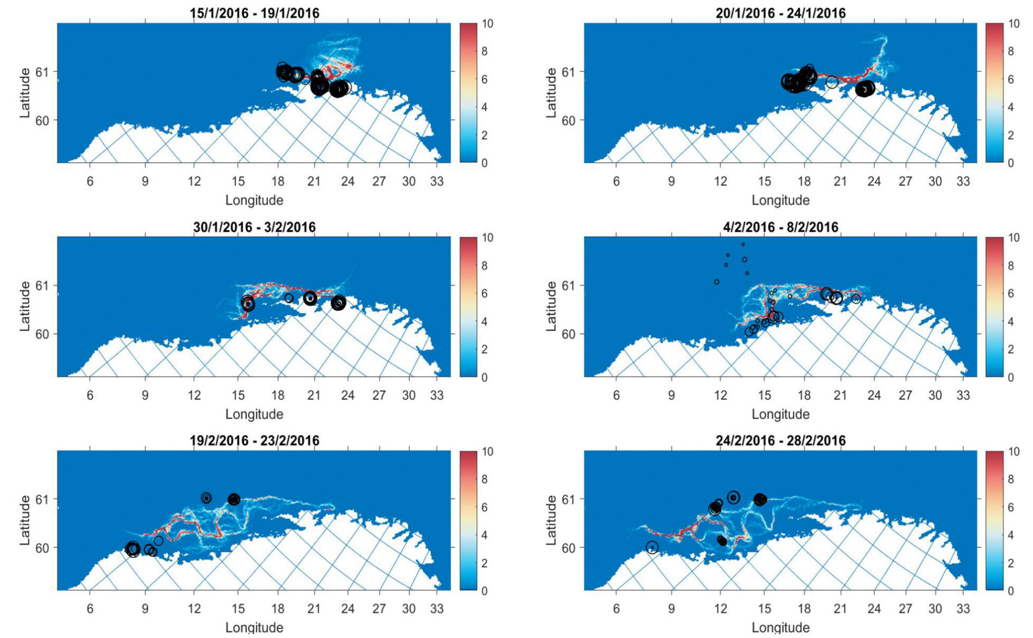


FIGURE 8 | Model output (colourmap) with catch points (black circles) overlayed for selected periods in 2016. Size of circles are scaled according to the catch weight in kg. The colourmap gives the average number of individuals in grid cell *j* for a 5 day period.

two branches of migration trajectories appeared. One branch extends alongside the continental slope, whilst another curves down toward the coast at approximately 67 degrees latitude and 9 degrees longitude. The 2020 simulation shows similar branching but there is more movement between branches. Toward the end of the migration the individuals push further off-coast, with high densities at 64 degrees latitude and 6 degrees longitude. The two branches join at this point and form a continuous tail that is prominent at the end of February.

3.2.2. Survey Comparison

The high resolution comparison of densities of individuals along acoustic transects with acoustic estimates was useful. It revealed that the model predicts a high spread off coast, not bunching individuals at one particular location (**Figure 6**). Due to the design of the model, densities are high before tailing off at deep bathymetry ($D \geq 400$m), a pattern present in the acoustic data also.

3.3. Comparison With Catch Data

3.3.1. Qualitative Analysis

In 2/3 of the years there is good agreement between model and catch values (2016, 2017, 2019, and 2020). However, there is intra- and inter-annual variation that can be the result of changes in both fish and vessel behaviour. Below, two simulations are presented with poorer and better agreement between model and observations, respectively (**Figures 7**, **8**).

In 2015, the model migration proceeds southerly in a manner that appears slower than the development of catch during this period, particularly from the end of January to the beginning of February, where there are catches in traditional spawning areas at a very early stage of the migration (**Figure 7**). Survey estimates of abundance from this year were uncertain, where a shift in strong NSSH year classes and immigration from off-coast areas in early February listed as possible explanations for discrepancies (Slotte et al., 2015). Given the comparisons here, it seems there is immigration from off-coast regions.

In contrast, the 2016 catches take place along branches of the migration where the model predicts higher densities (**Figure 8**). The development of catches overlap with the model evolution. Survey distribution also corroborate findings in early February with observations of high densities around 66–67 degrees latitude (Slotte et al., 2016). The figures from 2017 to 2020 are included in the **Supplementary Material**.

3.3.2. Quantitative Analysis

The centre points for fishing activity is in the northern fjords in mid-January, and the simulations were designed around these starting points. The longitude varies more in this period, suggesting that that the variability is mainly off-coast. In 2018,

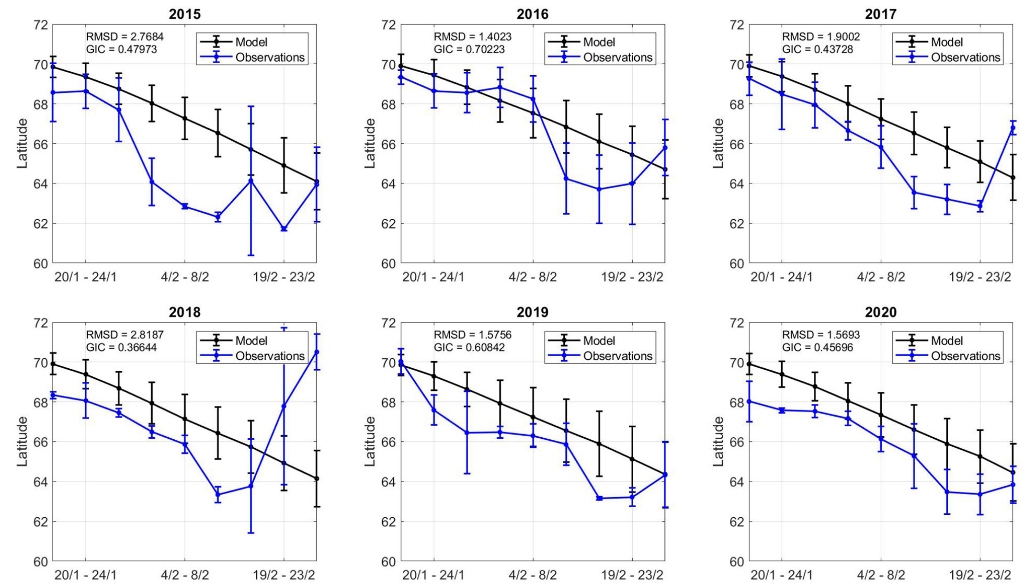


FIGURE 9 | Model and observation comparison across all years. The x axis displays the time period (inclusive). The y axis displays the latitude centre point for catch and model values in the time period. The error bar shows the square root of the inertia values in each time interval.

there was an high off-coast component of catches at the beginning of the season. The least overlap (GIC) and the largest error (RMSD) in location comes from this time (**Figures 9**, **10**). Toward the end of the 2018 season many catches shift to northern regions. The simulation in 2015 deviated from catches also, for reasons which are described in the previous section.

In general, early February showed the highest GIC values, with lowest RMSD values, suggesting the model dynamics in this period provide reliable geospatial estimates. Both the model and observations shift southward during the simulation time frame. The southern evolution of the catch and the model relate to how NSSH migrates. The spatial variation in centre points (inertia) varies more in observations than model estimates. There are many time periods where there appear catch points spread in space (19/2-23/2 2018) and others when catches are concentrated in one area (4/2-8/2 2018). As the model is physics-based, it has a more constant spatial distribution (inertia and centre points), although there are local differences (**Figures 7**, **8**).

4. DISCUSSION

In general, the model showed good agreement with survey and catch observations (**Figures 9**, **10**). Emergent properties of the migration trajectory overlap with vessel catch data. Regardless, there are limitations to modelling behaviour at such low resolutions. There are many sources of uncertainty not fully resolved in this model. Therefore, in conjunction with discussion about the simulations in this paper, we shall detail how models can be improved in future work to improve estimates.

4.1. Model Development

The model formulation required many steps of refinement to give reasonable output. In Equations (1) and (2), the response of the individual was formulated to account for the physical response directly against the prevailing current. Removing the counter-current response led to passive drifting northwards. Running simulations with environmental variables sampled close to the surface resulted in similar drifting patterns, as the magnitude of $\mathbf{v}_c$ is very high close to the surface, especially around the Lofoten basin (**Figure 4C**). A formulation that incorporated vertical conditions was important to model, as in Equations (7) and (8).

It is difficult to decouple the low resolution effects of local environmental values on model states, as they are spatially correlated in the horizontal plane. However, the combined effect of reacting to gradients from deep bathymetry, and current patterns seemed the most consistent properties amongst simulations (**Figure 4**). This study shows that memory- and gradient based reactive mechanisms can be used to model the migration of pelagic fish species such as NSSH (Fernö et al., 1998).

The energetic states of individuals were not included in this study, but may provide more insight into variations in observed patterns. Further work in this project aims to incorporate

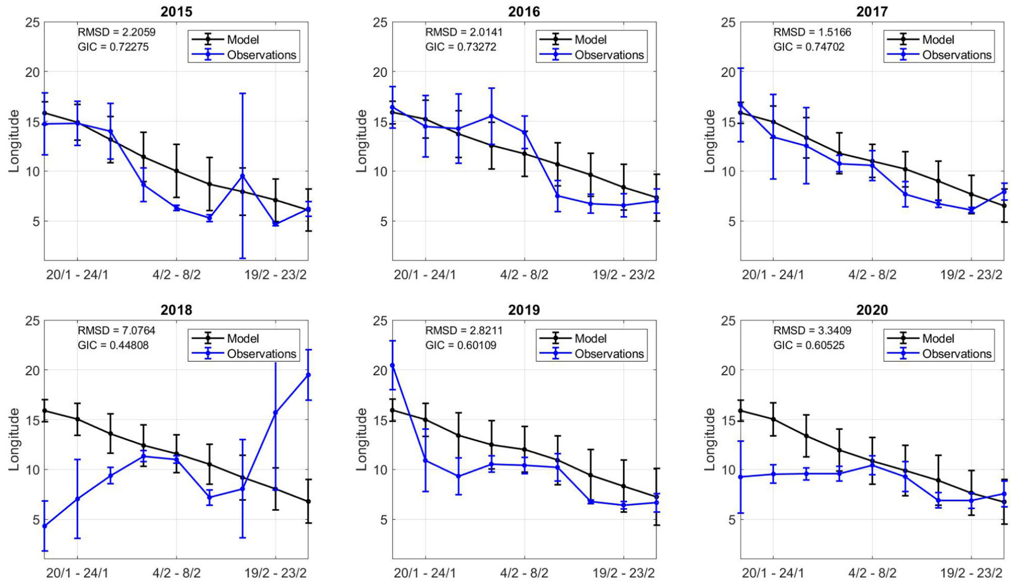


FIGURE 10 | Model and observation comparison across all years. The x axis displays the time period (inclusive). The y axis displays the longitude centre point for catch and model values in the time period. The error bar shows the square root of the inertia values in each time interval.

Dynamic Energy Budget theory, that can be used to model energetic costs of migration (Kooijman, 2000). Thus, state-dependent selection of spawning grounds can be explored, testing the effect of body length and condition on choice of area (Slotte and Fiksen, 2000). Natural mortality may also be included here. Interactions between individuals was explored, such as attraction and repulsion, but deemed difficult to model at the 4 km^2 scale. There may be a case to model interactions between schools, where information transfer could potentially give access to novel information, but this is beyond the scope of the project at this stage.

It should be mentioned that the ocean models that force the behaviour of the IBM are subject to their own limitations. Ocean model outputs have uncertainty caused by limits in model resolution, our knowledge of the processes resolved by the model, uncertainty in initial values, boundary conditions, parameter values, and inaccuracies in numerical implementations (Lermusiaux et al., 2006). Additionally, climatic fluctuations can play an important role in survival of recruits and thus, biomass estimates (Toresen and Østvedt, 2000). Therefore, understanding how variability in climate can manifest as variability in individual state and parameter values is important moving forward.

4.2. Initialization and Tuning

The initialization procedure presented above functioned to generate a generic distribution that was applied to all model scenarios. This provided a realistic initial distribution as input for the tuning procedure. In reality, there is much uncertainty around the initial distribution of NSSH. This may be explained by changes in winter stay areas, which has historic precedent (Dragesund, 1970). Thus, in future work, we shall model initial distributions based on prescient knowledge of winter stay areas and recent catch observations preceding the model run.

Tuning the migration model against survey data allowed for high resolution, time-sensitive estimates of fish densities. There was good agreement between tuned models and independent datasets. Nevertheless, developing the means to calibrate the model was challenging. One difficulty was the progression of the migration in comparison to surveys. The survey starts in southern Norway and progresses northward, the inverse of the NSSH migration. This meant estimates at the beginning and end of survey largely consisted of acoustic absence. In the course of this work, many objective functions for tuning of different resolutions were provided for the individuals and there is room for further experimentation here.

The model was trained using data that has its own biases (e.g., gear selectivity) which may bias the model tuning as a result. Additionally, fisheries data is sparse in space and time. Immigration and recruitment will also sporadically alter fish abundance. As such, one model realization cannot encapsulate the full variation in fish distributions. Adding noise terms to model parameters can be used to produce an ensemble of estimates through Monte Carlo experiments. State and parameter estimates can be obtained effectively in this way (Ward et al., 2016).

4.3. Comparison With Catch Data

The catch data served to judge the performance of the model across the time frame and highlight discrepancies. This shows what the model can and cannot resolve, and at what scales. The physics- based model used assumptions of responses to the prevailing current and gradients to reproduce the likely migration at large-scale. In contrast, catch data reflects which locations were targeted based on market value, quotas, weather conditions, etc. This is variability the model cannot capture. The model output is useful in predicting relevant catch locations in space and time at low resolutions (**Figures 9**, **10**). It gives insight into where the development of model densities is matched with fishing activity and where it does not. The output will have operational use when it is combined with real-time fisheries-dependent data, as described below. The novel use of real-time observations can give time-sensitive estimates that are more reliable than use of model or observation values in isolation.

4.4. Future Developments

In future, information on vessel activity will be incorporated to improve observation quality. Classifying vessels in terms of search vs. fishing activity can provide fine-grain, continuous information on the behaviour of fishermen. Thus, one can improve coverage of observations and incorporate absence data. Finally, using a data assimilation procedure, we can improve estimates by combining a model, such as the one developed here, with observation data. Such estimates require expansion of the model to include uncertainties in initial conditions, environmental variables, parameters, etc. We have developed one realization of the model here, but modelling such uncertainties produces multiple realizations, from which one can calculate the most probable estimate of the model state (Evensen, 2009).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: https://www.fiskeridir.no/Tall-og-analyse/AApne-data/AApne-datasett/elektronisk-rapportering-ers.

AUTHOR CONTRIBUTIONS

CK: wrote the manuscript, developed the model, processed data, ran the simulations, and produced figures. FM: helped with accessing and processing data, feedback on written work and simulations. JK: input on herring research, feedback on written work. MA: helped with conception of the project, aided with setting up the model, gave constant feedback and suggestions on every aspect of the project. All authors contributed to the article and approved the submitted version.

FUNDING

The work is part of the FishGuider project, which was funded by the project participants and the Norwegian Research Council (Project Number 296321). North Atlantic Institute of Sustainable Fishing (NAIS) is an industry partner that funds 50% of this project.

ACKNOWLEDGMENTS

We greatly acknowledge the support, input and feedback from the project participants: NTNU, SINTEF, NAIS, and UiB. We also acknowledge useful input from discussions with Øystein Varpe at UiB.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2021.754476/full#supplementary-material

REFERENCES

Alver, M. O., Broch, O. J., Melle, W., Bagøien, E., and Slagstad, D. (2016). Validation of an Eulerian population model for the marine copepod Calanus finmarchicus in the Norwegian Sea. *J. Mar. Syst.* 160, 81–93. doi: 10.1016/j.jmarsys.2016.04.004

Arnason, R., Magnusson, G., and Agnarsson, S. (2000). The Norwegian spring-spawning herring fishery: a stylized game model. *Mar. Resour. Econ.* 15, 293–319. doi: 10.1086/mre.15.4.42629328

Baker, R. E., Peña, J.-M., Jayamohan, J., and Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biol. Lett.* 14:20170660. doi: 10.1098/rsbl.2017.0660

Barbaro, A., Einarsson, B., Birnir, B., Sigurðsson, S., Valdimarsson, H., Pálsson, O., Sveinbjörnsson, S., and Sigurðsson, o. (2009). Modelling and simulations of the migration of pelagic fish. *ICES J. Mar. Sci.* 66, 826–838. doi: 10.1093/icesjms/fsp067

Bauer, S., and Klaassen, M. (2013). Mechanistic models of animal migration behaviour - their diversity, structure and use. *J. Anim. Ecol.* 82, 498–508. doi: 10.1111/1365-2656.12054

Bjørndal, T., Gordon, D. V., Kaitala, V., and Lindroos, M. (2004). international management strategies for a straddling fish stock: a bio-economic simulation model of the Norwegian spring-spawning herring fishery. *Environ. Resour. Econ.* 29, 435–457. doi: 10.1007/s10640-004-1045-y

Blaxter, J. H. S. (1985). The herring: a successful species? *Can. J. Fish. Aquat. Sci.* 42, s21–s30. doi: 10.1139/f85-259

Boyd, R., Sibly, R., Hyder, K., Walker, N., Thorpe, R., and Roy, S. (2020). Simulating the summer feeding distribution of Northeast Atlantic mackerel with a mechanistic individual-based model. *Progr. Oceanogr.* 183:102299. doi: 10.1016/j.pocean.2020.102299

Chu, D. (2011). Technology evolution and advances in fisheries acoustICS. *J. Mar. Sci. Technol.* 19:2. doi: 10.51400/2709-6998.2188

Dragesund, O. (1970). Factors influencing year-class strength of norwegian spring spawning herring (clupea harengus l.). *Fiskeridirektoratets Skrifter Serie Havundersøkelser* 15, 381–450.

Dragesund, O., Hamre, J., and Øyvind, U. (1980). Biology and population dynamics of the Norwegian spring- spawning herring. *Rapports et Procès-Verbaux des Rèunions* 177, 43–71.

Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter*. Berlin: Springer.

Fernö, A., Pitcher, T. J., Melle, W., Nøttestad, L., Mackinson, S., Hollingworth, C., et al. (1998). The challenge of the herring in the Norwegian sea: making optimal collective spatial decisions. *Sarsia* 83, 149–167. doi: 10.1080/00364827.1998.10413679

Fiksen, Ø., and Slotte, A. (2002). Stock environment recruitment models for norwegian spring spawning herring (clupea harengus). *Can. J. Fish. Aquat. Sci.* 59, 211–217. doi: 10.1139/f02-002

Føre, M., Dempster, T., Alfredsen, J. A., Johansen, V., and Johansson, D. (2009). Modelling of Atlantic salmon (Salmo salar L.) behaviour in sea-cages: a Lagrangian approach. *Aquaculture* 288, 196–204. doi: 10.1016/j.aquaculture.2008.11.031

Giske, J., Huse, G., and Berntsen, J. (2001). Spatial modelling for marine resource management, with a focus on fish. *Sarsia* 86, 405–410. doi: 10.1080/00364827.2001.10420482

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecol. Model.* 198, 115–126. doi: 10.1016/j.ecolmodel.2006.04.023

Grimm, V., and Railsback, S. F. (2005). *Individual-Based Modeling and Ecology*. Princeton: Princeton University Press, 3–21.

Huse, I., and Ona, E. (1996). Tilt angle distribution and swimming speed of overwintering Norwegian spring spawning herring. *ICES J. Mar. Sci.* 53, 863–873. doi: 10.1006/jmsc.1996.9999

Kooijman, S. A. L. M. (2000). *Dynamic Energy and Mass Budgets in Biological Systems, 2 Edn*. Cambridge: Cambridge University Press.

Langård, L., Fatnes, O., Johannessen, A., Skaret, G., Axelsen, B., Nöttestad, L., et al. (2014). State-dependent spatial and intra-school dynamics in pre-spawning herring Clupea harengus in a semi-enclosed ecosystem. *Mar. Ecol. Progr. Ser.* 501, 251–263. doi: 10.3354/meps10718

Lermusiaux, P. F., Chiu, C., Gawarkiewicz, G., Abbot, P., Robinson, A., Miller, R. N., et al. (2006). Quantifying uncertainties in ocean predictions. *Oceanography* 19, 90–103. doi: 10.5670/oceanog.2006.93

Nøttestad, L. (1998). Extensive gas bubble release in Norwegian spring-spawning herring (Clupea harengus) during predator avoidance. *ICES J. Mar. Sci.* 55, 1133–1140. doi: 10.1006/jmsc.1998.0416

Nøttestad, L., Aksland, M., Beltestad, A., Fernö, A., Johannessen, A., and Arve Misund, O. (1996). Schooling dynamics of norwegian spring spawning herring (*Clupea harengus* L.) in a coastal spawning area. *Sarsia* 80, 277–284. doi: 10.1080/00364827.1996.10413601

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 J. Mar. Sci.* 73, 2302–2310. doi: 10.1093/icesjms/fsw062

Politikos, D., Somarakis, S., Tsiaras, K., Giannoulaki, M., Petihakis, G., Machias, A., and Triantafyllou, G. (2015). Simulating anchovy's full life cycle in the northern Aegean Sea (eastern Mediterranean): a coupled hydro-biogeochemical–IBM model. *Progr. Oceanogr.* 138, 399–416. doi: 10.1016/j.pocean.2014.09.002

Salthaug, A., Stenevik, E., Vatnehol, S., Anthonypillai, V., Ona, E., Anthonypillai, V., et al. (2020). *Distribution and Abundance of Norwegian Spring Spawning Herring During the Spawning Season in 2020*. Technical Report, Institute of Marine Research, Bergen, Norway.

Slagstad, D., and McClimans, T. A. (2005). Modeling the ecosystem dynamics of the barents sea including the marginal ice zone: i. physical and chemical oceanography. *J. Mar. Syst.* 58, 1–18. doi: 10.1016/j.jmarsys.2005.05.005

Slotte, A., and Fiksen, Ø. (2000). State-dependent spawning migration in Norwegian spring-spawning herring. *J. Fish Biol.* 56, 138–162. doi: 10.1111/J.1095-8649.2000.TB02091.X

Slotte, A., Johnsen, E., Pena, H., Salthaug, A., Utne, K., Anthonypillai, V., et al. (2015). *Distribution and Abundance of Norwegian Spring Spawning Herring During the Spawning Season in 2015*. Technical Report, Institute of Marine Research, Bergen, Norway.

Slotte, A., Salthaug, A., Utne, K., Ona, E., Vatnehol, S., and Pena, H. (2016). *Distribution and Abundance of Norwegian Spring Spawning Herring During the Spawning Season in 2016*. Technical Report, Institute of Marine Research, Bergen, Norway.

Tamario, C., Sunde, J., Petersson, E., Tibblin, P., and Forsman, A. (2019). Ecological and evolutionary consequences of environmental change and management actions for migrating fish. *Front. Ecol. Evol.* 7:271. doi: 10.3389/fevo.2019.00271

Toresen, R., and Østvedt, O. (2000). Variation in abundance of norwegian spring-spawning herring (clupea harengus, clupeidae) throughout the 20th century and the influence of climatic fluctuations. *Fish Fisheries* 1, 231–256. doi: 10.1111/j.1467-2979.2000.00022.x

Touzeau, S., Lindroos, M., Kaitala, V., and Ylikarjula, J. (2000). Economic and biological risk analysis of the norwegian spring-spawning herring fishery. *Ann. Oper. Res.* 94, 197–217. doi: 10.1023/A:1018973217951

Tu, C.-Y., Tseng, Y.-H., Chiu, T.-S., Shen, M.-L., and Hsieh, C.-H. (2012). Using coupled fish behavior-hydrodynamic model to investigate spawning migration of Japanese anchovy, *Engraulis japonicus*, from the East China Sea to Taiwan: Spawning migration model of Japanese anchovy. *Fish. Oceanogr.* 21, 255–268. doi: 10.1111/j.1365-2419.2012.00619.x

Varpe, Ø., Fiksen, Ø., and Slotte, A. (2005). Meta-ecosystems and biological energy transport from ocean to coast: the ecological importance of herring migration. *Oecologia* 146, 443. doi: 10.1007/s00442-005-0219-9

Ward, J. A., Evans, A. J., and Malleson, N. S. (2016). Dynamic calibration of agent-based models using data assimilation. *R. Soc. Open Sci.* 3, 150703. doi: 10.1098/rsos.150703

Woillez, M., Poulard, J.-C., Rivoirard, J., Petitgas, P., and Bez, N. (2007). Indices for capturing spatial patterns and their evolution in time, with application to European hake (Merluccius merluccius) in the Bay of Biscay. *ICES J. Mar. Sci.* 64, 537–550. doi: 10.1093/icesjms/fsm025

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Paper 2

Kelly, C., Michelsen, F. A., and Alver, M. O. (2023). An ensemble modelling approach for spatiotemporally explicit estimation of fish distributions using Data Assimilation. Fisheries Research 261, 106624.
doi: https://doi.org/10.1016/j.fishres.2023.106624

Fisheries Research 261 (2023) 106624

Contents lists available at ScienceDirect

Fisheries Research

journal homepage: www.elsevier.com/locate/fishres

An ensemble modelling approach for spatiotemporally explicit estimation of fish distributions using data assimilation

Cian Kelly [a,*], Finn Are Michelsen [b], Morten Omholt Alver [a]

[a] Department of Engineering Cybernetics, NTNU, Trondheim, Norway
[b] Department of Fisheries and New Biomarine Industry, SINTEF Ocean, Trondheim, Norway

A R T I C L E I N F O

Handled by Jie Cao

Keywords:
Individual-based model
Ensemble
Data assimilation
Corrections
State estimation
Decision-support

A B S T R A C T

This article presents a novel method for estimating large scale spatiotemporal distribution patterns of fish populations modelled at the individual level. A single realization of an individual-based model calibrated on historic data has weak predictive capacity, given the underlying uncertainties faced when modelling a relatively small cluster of individuals operating in a high dimensional spatial plane. By incorporating real-time data sources to update these models, we can improve their predictive capacity. When correcting estimates from a large population of individuals, we don't have access to information about individual histories, such as information derived from tagging data. We propose mapping individuals to derived density matrices, which can be corrected using conventional data sources which describe a mass of individuals e.g. catch data. An ensemble of derived states are used as forecast inputs to an assimilation procedure, that calculates an analysis state matrix of the same form. An individuals' position and biomass values are updated based on the analysis values. To assess the effect of corrections, we setup a simulation experiment to explore the impact the number of measurement points has on the updated spatiotemporal distribution. The measurement points were sampled from derived states of a twin model that resembles the original model. The output of the twin model serves as the true distribution. With an increasing number of measurement points the centre of mass of the modelled distribution converges on the true distribution and the two distributions increase in overlap. Additionally, the absolute error between model and true values decreases. This estimation method, applied to individual-based models and coupled with real-time fisheries data, can improve spatially explicit estimates of fish distributions.

1. Introduction

Individual-based models (IBMs) simulate interactions between a population of model individuals and their surrounding environment (Grimm and Railsback, 2005). IBMs capture large scale phenomena with simple interactions. Complexity arises from modelling bottom-up processes, rather than imposing population level parameters such as birth and death rates (DeAngelis and Grimm, 2014). It is the individuals local input information that produces unique responses. Infection transmission models in epidemiology demonstrate this. Contact rates and transmission probabilities vary in accordance with the unique behaviour of individuals. The social network of the individual matters too (Koopman and Lynch, 1999; Buchwald et al., 2020). For these reasons population-level features are not a simple sum of parts. The subtle differences between individuals alter system behaviour over time. Differences arise as individuals update state variables, such as position and velocity, at frequent time intervals (DeAngelis and Grimm, 2014). In this way, internal states represent the integration of past and present input over time. Incrementing states forward in time, in distinct simulation scenarios, can explain the evolution of higher level phenomena. For example, an individual fish's response to temperature and current explains variation in migratory routes (Barbaro et al., 2009; Tu et al., 2012).

These properties make IBMs attractive explanatory tools. However, IBMs have weak predictive capacity at a precise location and time, and are of limited operational use. As Baker et al. (2018) has pointed out, mechanistic models rely on oversimplified assumptions that are narrow in nature and limited in broad predictive power. Models are tuned once using historical data, validated on an independent dataset, before forecasting future estimates. This is useful for points trained on the historical data, but as the model progresses, states diverge from reality owing to uncertainties (Ward et al., 2016; Kieu et al., 2020). We consider model

* Corresponding author.
E-mail address: cian.kelly@ntnu.no (C. Kelly).

https://doi.org/10.1016/j.fishres.2023.106624
Received 28 April 2022; Received in revised form 5 October 2022; Accepted 17 January 2023
Available online 20 January 2023

simplifications, intialization values, mechanistic assumptions, parameters and inputs as main sources of uncertainty. Integrating real-time, real-world observations to correct model states is a way of controlling divergence. Data Assimilation updates model states based on real-time observations. It operates under the assumption models or observations alone cannot resolve the real system (Fu et al., 2011). It estimates the real system by applying a statistical correction term to model estimates (Alver and Michelsen, 2015. Data Assimilation has been applied successfully to applications in fishery models, predictive ecology, the terrestrial carbon cycle, traffic simulation, amongst other areas (Niu et al., 2014; Kieu et al., 2020).

The Ensemble Kalman Filter (EnKF) is a Data Assimilation method initially developed by Evensen (1994). It is used for state and parameter estimation of non-linear systems e.g. atmospheric and ocean systems (Houtekamer and Mitchell, 2001; Alver and Michelsen, 2015). The EnKF simulates separate instances of a model in a Monte Carlo setup where instances diverge over time due to random perturbations which represent the uncertainty in the model and its inputs. The divergence in model states is used to calculate error statistics. When observations are available, a correction term is applied to each instance of the model, based on these error statistics (Evensen, 2009). Although the EnKF implicitly assumes Gaussian distributions of prior states, it is effective in approximate estimation of states in non-linear systems that violate this assumption, which is often the case (Katzfuss et al., 2016).

Assimilating measurement data with IBM output can vastly improve predictions. This has been explored in the case of population-level estimates (Niu et al., 2014). Here we focus on high dimensional spatially explicit patterns of abundance in fish distributions. Real-time integration of available observations has the potential to facilitate the goal of time sensitive decision-support for stakeholders in the fishing industry. There are two main challenges to achieving this objective. Firstly, at large spatial scales, we currently don't have access to measurement data that describe individual fish with unique identities tracked through time to compare with IBM output directly. Secondly, the real spatially explicit distribution of fish stocks at any given time is highly uncertain, due to the sparsity of observations.

We propose a novel approach for correcting the IBM that is compatible with measurement sources that don't preserve an individual fishes identity, such as catch data. This method maps IBM output onto a two dimensional spatial grid, where derived density estimates are utilized as prior states in the EnKF. With minimal manipulation, the analysis estimate is remapped to the IBM individuals. The EnKF is advantageous for this purpose, as the IBM model mechanics are not altered directly, avoiding degeneracy of the model structure (Katzfuss et al., 2016). Additionally, the EnKF is suitable for assimilation when we don't fully understand the sources of errors.

To address the issue of the true underlying distribution, we use a twin model experiment to simulate observations. That is, we simulate an altered IBM and treat it as the true migration pattern. The IBM is based on the spawning migration of the Norwegian Spring Spawning Herring, as described in Kelly et al. (2022). The model IBM mechanics are extended from the single realization described, to an ensemble of estimates, through addition of stochastic perturbations to model components. In the true scenario, the deterministic realization is simulated alongside the ensemble of models, then sampled for measurement points. The measurement values are assimilated with each instance of the ensemble. We then investigate the impacts of measurements on the model distribution, given we have full knowledge of the true geographical distribution. The convergence of the ensemble on the true distribution indicates the capacity to correct the IBM. Spatial indices were used to measure this convergence and scenarios run in this study examined the influence of number of observations on spatial patterns.

With improvements in technology, observations will become less sparse and thus, our capacity to correct models shall improve (Fu et al., 2011). For example, acoustic technology today involves use of a multi-beam system that can resolve multiple fish at once (Chu, 2011). Additionally, studies have shown it is possible to classify fishing activity with high precision, from available vessel data at an individual level, such as position, speed and turning angle of boats (Bez et al., 2011; de Souza et al., 2016). Assimilating such sources of fisheries data with spatially explicit model predictions can improve our collective understanding of dynamics of large scale fish distribution patterns.

2. Materials and methods

The purpose of the model is to improve spatiotemporal estimates of fish distributions through integration of observations when they become available. The following description primarily focuses on two aspects: 1) Modifying the IBM to make it compatible with the EnKF procedure for assimilating data. 2) Setup of the twin model experiment to analyse the impact of measurements on the model (Fig. 1).

2.1. Ensemble of IBM trajectories

The IBM prediction model developed in Kelly et al. (2022) of a single model trajectory of herring is reproduced here for completeness, with the following set of difference equations at each time step k:

$$\mathbf{p}[k] = \mathbf{p}[k-1] + \Delta t\left(\mathbf{v}_f[k-1] + \mathbf{v}_c[k-1]\right) \tag{1}$$

$$\mathbf{v}_f[k-1] = -\Phi\mathbf{v}_c[k-1] + \mathbf{v}_b[k-1] \tag{2}$$

$$\mathbf{v}_b[k-1] = \mathbf{r}_b\left(\begin{bmatrix}\cos(\boldsymbol{\theta}[k-1])\\ \sin(\boldsymbol{\theta}[k-1])\end{bmatrix}\right) \tag{3}$$

$$\boldsymbol{\theta}[k-1] = f(\nabla T[k-1], \nabla D[k-1]) \tag{4}$$

where $\mathbf{p}$ is the vector of positions, $\mathbf{v}_c$ are the horizontal current components vector at the individuals position in m s^{-1}, Φ is a parameter that controls the response to the current, $\mathbf{r}_b$ is the swimming speed of the individual and $\boldsymbol{\theta}$ is the angle of orientation, which is a function temperature and bathymetry gradients (T and D). This configuration allows the individual to respond with a higher priority to the horizontal components of the prevailing current.

As Evensen (2009) notes, the solution to a dynamical model is one of an infinite many realizations, and for meaningful solutions, we must consider the time series of the probability density function. The IBM modelled one realization of the herring migration pattern, optimized based on a narrow set of assumptions (Kelly et al., 2022). Numerous alternative realizations are possible, given uncertainties in model evolution over time. Here, we add random perturbations to the IBM state variables to generate a set of N divergent instances sequentially in time. This generates N trajectories of the original IBM, which are held in memory and updated independently at each time step.

Position $\mathbf{p}$ and velocity $\mathbf{v}$ of individuals were extended from the single IBM to N instances, notated by the state matrices $\mathbf{P}$ and $\mathbf{V}$, both with N columns. Additionally, biomass $\mathbf{B}$ of individuals in kg is added as another state here, where each individual was treated as a mass of fish (also referred to as a superindividual):

$$\mathbf{P}[k] = \mathbf{P}[k-1] + \Delta t(\mathbf{V}[k] + \widetilde{\mathbf{V}}[k]) \tag{5}$$

$$\mathbf{B}[k] = \mathbf{B}[k-1] - \Delta t(\widetilde{\mathbf{B}}[k] + \omega)\mathbf{B}[k-1] \tag{6}$$

where Δt was the time increment, reduction in biomass was controlled by the constant parameter ω, and divergence in states $\mathbf{V}$ and $\mathbf{B}$ were caused by the stochastic errors $\widetilde{\mathbf{V}}$ and $\widetilde{\mathbf{B}}$. The expected value E $[\widetilde{\mathbf{V}}]$ = E $[\widetilde{\mathbf{B}}]$ = 0. These errors produce prediction uncertainty in the system, representing errors in individuals migration direction, speed and mass and were formulated as follows:

$$\widetilde{\mathbf{V}}[k] = \widetilde{\mathbf{R}}[k]\begin{bmatrix}\cos(\widetilde{\boldsymbol{\Theta}}[k])\\ \sin(\widetilde{\boldsymbol{\Theta}}[k])\end{bmatrix} \tag{7}$$

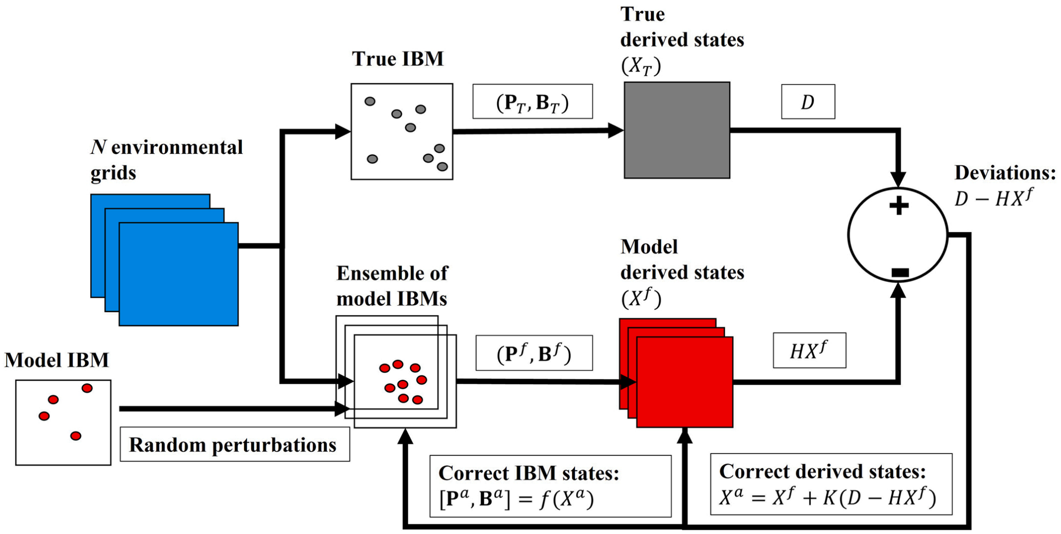


Fig. 1. A conceptualization of the assimilation of data using the twin model experiment.

$$\widetilde{\mathbf{R}}[k] = \alpha_1 \widetilde{\mathbf{R}}[k-1] + \alpha_2 \left(e_{sx1} \mathcal{N}\left(0, \epsilon_R^2\right)_{1xN} \right) + \mathcal{N}\left(0, \sigma_R^2\right)_{sxN} \tag{8}$$

$$\widetilde{\boldsymbol{\Theta}}[k] = \alpha_1 \widetilde{\boldsymbol{\Theta}}[k-1] + \alpha_2 \left(e_{sx1} \mathcal{N}\left(0, \epsilon_\Theta^2\right)_{1xN} \right) + \mathcal{N}\left(0, \sigma_\Theta^2\right)_{sxN} \tag{9}$$

$$\widetilde{\mathbf{B}}[k] = \alpha_1 \widetilde{\mathbf{B}}[k-1] + \alpha_2 \left(e_{sx1} \mathcal{N}\left(0, \epsilon_B^2\right)_{1xN} \right) + \mathcal{N}\left(0, \sigma_B^2\right)_{sxN} \tag{10}$$

where temporally correlated, slowly varying errors were controlled by the parameters α_1 and α_2, e_{sx1} is an $sx1$ vector of ones, where s is the number of individuals and ϵ_R, ϵ_Θ and ϵ_B represent standard deviation in speed, angle and biomass for each of the ensemble members. Applying these errors cause the N columns to diverge, creating an ensemble of random realizations.The standard deviations were calibrated to maintain spread between ensemble members and limit the severity of corrections when data was assimilated. This formulation is similar to system noise modelled in ocean models that use temporal autocorrelation of random noise to account for errors in representation of certain processes (Keppenne et al., 2008). In this case, we represent the errors in the individuals state matrix, resulting from uncertainties in the evolution of migration patterns. In addition to the temporally autocorrelated ensemble noise, spurious gaussian noise is added to each individual with mean of zero and standard deviations of σ_R, σ_Θ and σ_B. These individual noise components account for uncertainties in the migration of individual fish, regardless of ensemble member.

2.2. Data Assimilation framework

Before assimilation, the forecast IBM position $\mathbf{P}^f$ and biomass $\mathbf{B}^f$ from Equation (5) and (6) are converted to derived estimates:

$$X^f = f(\mathbf{P}^f, \mathbf{B}^f) \tag{11}$$

where X^f is an nxN grid of density values, with each cell representing the sum of the biomass of all individuals within that grid cell.

The EnKF uses the error covariance structure of the ensemble forecast X^f to calculate the correction term. However, the full covariance matrix is too large to be explicitly calculated here. We employ an equivalent implementation by Mandel (2006), which avoids the calculation of the full error covariance matrix and derives directly the prediction error covariance matrix in the observation space:

$$A = X^f - \frac{1}{N}\left(X^f e_{Nx1}\right) e_{1xN} \tag{12}$$

$$HA = HX^f - \frac{1}{N}\left(\left(HX^f\right) e_{Nx1}\right) e_{1xN} \tag{13}$$

$$P = \frac{1}{N-1} HA(HA)^T I_m + R \tag{14}$$

$$K = L \odot \left(\frac{1}{N-1} A(HA)^T P^{-1} \right) \tag{15}$$

where H is an mxn matrix that contains ones at m measured states, I_m is an mxm identity matrix, R is the mxm observation error covariance matrix, where each element on the diagonal is the variance of observation noise (σ_O^2), L is an mxN localization matrix and finally, K is the Kalman Gain, which is used to calculate the correction term. Localization adds a penalty to model covariances that are distant from the measurment point. For a small ensemble and high dimensional system, localization is necessary to limit the impact of observations (Houtekamer and Mitchell, 2005). The operator ($\odot$) is the Schur product, an elementwise operation acting on all covariance values. The full L matrix was calculated from a radial basis function:

$$L_{ij} = \begin{cases} 0, & \text{if } \| g_i - g_j \|^2 > c \\ \exp\left(\dfrac{- \| g_i - g_j \|^2}{2\rho^2} \right), & \text{otherwise} \end{cases} \tag{16}$$

where we calculate the euclidean distance between the xy grid coordinate for each model grid cell g_i and the measured grid cell g_j. The value is calculated for all model coordinates ($i = 1{:}\ n$) and measurement point coordinates ($j = 1{:}\ m$). When i equals j, the value of L equals one, and as i moves away from j there is exponential decline in the value of L. To controls spatial correlations around the measurement point, the constant parameter ρ is used. In addition, to avoid spurious correlations, a cut-off point c sets distant spatial covariances to zero.

The analysis estimate X^a is calculated as follows:

$$X^a = X^f + K(D - HX^f) \quad (17)$$

where D is the mxN measurement matrix. The standard EnKF adds $\mathcal{N}(0,R)$ realizations of observation errors to generate observational perturbations. However, in this study, where we are sampling from a non-negative distribution with mostly zero values, perturbation of observations led to inaccuracies in the posterior field. Instead, the columns of D are treated as replicates of the original measurement vector. Treating the observations as deterministic contracts the variance across the ensemble in the analysis estimate (Burgers et al., 1998). To compensate for this contraction in spread between columns of X^a, an inflation factor ψ was used to replace the analysis estimates, as mentioned in Evensen (2009):

$$X_z^a = \overline{X}^a + \psi\left(X_z^a - \overline{X}^a\right) \quad (18)$$

where z is the index for the ensemble member and $\overline{X}^a$ is the ensemble mean of the analysed derived states.

Following assimilation, the IBM is modified to reflect updated grid cell biomass values and this is achieved with minimum manipulation of the underlying model structure. The derived analysis estimate X^a is converted back to IBM states:

$$[\mathbf{P}^a, \mathbf{B}^a] = f(X^a) \quad (19)$$

where $\mathbf{B}^a$ and $\mathbf{P}^a$ are the analysis biomass and position values for individuals, calculated from X^a.

2.3. Data assimilation adapted for the IBM

In this section, we detail how the conversion between IBM and EnKF states was achieved in Equations (11) and (19). Mapping from individual representations to derived density states means aggregating information from individuals into a grid representation that describes geographical distribution and abundance. Cocucci et al. (2022) describes this as a transition between micro- and macro-states. To achieve this mapping, the forecasted states are derived individual by individual, as shown in Algorithm 1, until the X^f matrix is furnished with an ensemble of density fields.

Algorithm 1. Algorithm for mapping from IBM states to forecast states X^f in Equation (11).

```
for e ← 1 to N do
    for i ← 1 to s do
        1. Find cell coordinates xy of the individual from P^f(i, e)
        2. Find biomass b of the individual from B^f(i, e)
        3. Add derived value to forecast state matrix: X^f(x, y, e) = X^f(x, y, e) + b;
    end
end
```

Mapping from the high dimensional analysis field to relatively fewer individuals is more challenging, and Algorithm 2 was designed to maximize the retention of density values, while limiting adjustments to the IBM. Each cell is checked for the analysis estimate and if it is greater than zero, the value is assigned as individual biomass, divided evenly amongst individuals at that position. If the analysis estimate is greater than zero, but there are no individuals present, one individual positioned in a cell with a zero analysis estimate is randomly moved there (assuming there is an individual available to move).

Algorithm 2. Algorithm for mapping from analysis states X^a to IBM states in Equation (19).

(*continued on next column*)

(*continued*)

```
for e ← 1 to N do
    for c ← 1 to n do
        1. Find xy coordinates of cell from linear index c
        2. Find derived density d of analysis matrix: d = X^a(x, y, e)
        3. Check the number of individuals si at coordinates from P^f(:, e)
        4. if si > 0 & d > 0 then
            a) Each individual retains the prior position: P^a = P^f;
            b) Divide d evenly amongst si individuals in cell, so for each individual: B^a = d/si;
           end
        5. if si == 0 & d > 0 then
            a) Randomly move an individual from a cell where density is zero: P^a = xy;
            b) Set the biomass of moved individual to the density in this cell: B^a = d;
           end
    end
end
```

This method is similar to the randomized redistribution described in Cocucci et al. (2022), where individuals are moved between categories where needed and attributes are updated. In this case, $\mathbf{P}^a$ and $\mathbf{B}^a$ are estimated from the macro- to micro-state mapping. This mapping conserves density estimates with higher priority than individual histories, given real fisheries observations are of aggregated individuals.

2.4. Twin model development

The observations used in this study were synthetically generated using a twin model, which represents the true distribution here. Twin model design has been used to give insight into capacity to correct model components with few observed variables (Simon and Bertino, 2009. Specifically, we are testing the data assimilation procedure and observability of the system in our setup. Here, we observe a derived variable from the twin model (X_T), which is a density field with dimensions $nx1$. This was sampled in the assimilation procedure to furnish the D matrix in Equation (17). The samples were taken from a predefined grid along the Norwegian coast (Fig. 2). Like the model IBM, these values were derived from individual state variables:

$$X_T = f(\mathbf{P}_T, \mathbf{B}_T) \quad (20)$$

where $\mathbf{P}_T$ and $\mathbf{B}_T$ were position and biomass of twin model individuals. Unlike the model IBM, the twin individuals were stepped forward with no feedback from the assimilation procedure. The twin IBM was updated using the same dynamics as the main IBM, with the exception of the swimming speed $\mathbf{r}_b$ in Equation (3), which was reduced in the twin model. This hypothetical scenario represents a situation where the model overestimates the true migration speed of the herring.

2.5. Model Simulation

The environmental conditions were obtained from a run of the physical-biological ocean model SINMOD (Slagstad and McClimans, 2005 set up in a domain with 4 km horizontal resolution covering the Norwegian and Arctic Seas. The same grid resolution was used for the derived states, where $n = 941 \times 620$. The IBM modelled individuals in a 2D environment where position was updated on N continuous horizontal planes. The s individuals initialized in each ensemble member had their position $\mathbf{P}$ centred in an area in Northern Norway in mid-January. The biomass $\mathbf{B}$ states for each ensemble were initialized from a Gaussian distribution with mean μ_B and standard deviation Σ_B. These values were divided among individuals based on their proximity to the centre point of the starting position. The model was simulated for a period of 45 days during the herring spawning migration. The time increment Δt was 4 h, for a total of 270 time steps. The simulation period was split into assimilation and non-assimilation periods. The assimilation period operated from day 18–37. During the assimilation period, corrections were performed once per day. This left a period prior- and post-assimilation for the states to diverge from the ensemble mean. Model parameters were calibrated to stabilize the assimilation

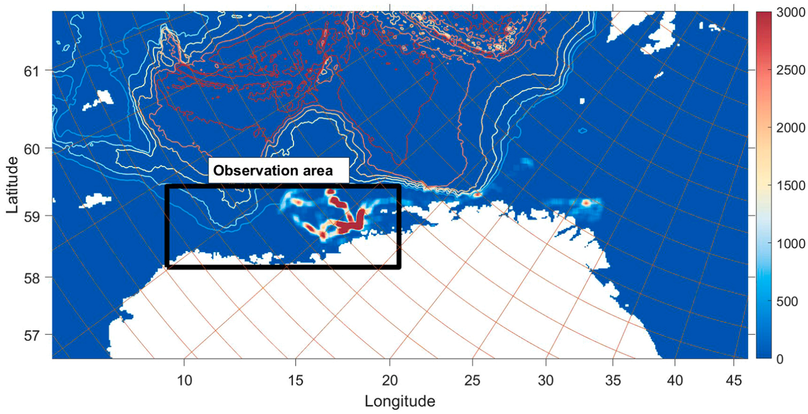


Fig. 2. The area sampled for measurement point during assimilation (black rectangle), where the colored cells represent the true distribution of derived density values (X_T), in kilograms, on day 25 of the simulation. The contours represent the depth in metres. For convenience, the colorbar represents the scale of both X_T and depth. Longitude and latitude ticks extend from the x and y axis, respectively, along the Norwegian coast.

procedure, specifically to avoid extreme correction terms (Table 1).

To investigate the number of observation points needed to make the ensemble converge towards the true state, four separate scenarios were setup to test varying number of observation points, where $S_1 = 100$ points, $S_2 = 200$ points, $S_3 = 400$ points and $S_4 = 800$ points. The points were sampled at equally spaced intervals in the observation area (Fig. 2). To implement a fixed virtual observer system, these same points were sampled on each day of the assimilation period. The output from these scenarios was compared to a control model, which was run in parallel with no assimilation of data.

2.6. Analysis

Quantitative and qualitative measures of performance investigated the capacity to correct the four model scenarios with samples of measurement points and thus, represent the spatiotemporal patterns of the true fish distribution. This is important in the geographical mapping of fish stocks. The quantitative measures used were based on equations from Woillez et al. (2007). The Centre of Gravity (CG) measures the weighted position of the density estimates at a given time. We investigated how this diverged from the control model and converged towards the true CG. Global index of Collocation (GIC) is a measure of the overlap of two separate distributions. It takes into account the CG of the two distributions and the variance around the CG. A value of one is perfect overlap between the two and a value of zero indicates distinct populations. Both CG and GIC were described in terms of latitude and longitude coordinates. They were calculated from the densities of the derived states $\overline{X}$ and X_T, where $\overline{X}$ is the ensemble mean of the model. The derived states were saved once per day during the model simulation, and after the assimilation step.

Table 1
List of model variables and parameters for ensemble simulations.

Name	Description	Unit	Value
State variables			
$\mathbf{P}$	Model position of individuals		
$\mathbf{B}$	Model biomass of individuals		
X	Model derived density states		
$\mathbf{P}_T$	True position of individuals		
$\mathbf{B}_T$	True biomass of individuals		
X_T	True derived density states		
Parameters			
N	Number of ensemble members	–	100
s	Number of individuals	–	10000
m	Number of observations	–	–
n	Number of derived states	–	583420
Δt	Time step	h	4
ϵ_R	Standard deviation in ensemble magnitude	m s^{-1}	0.1
ϵ_Θ	Standard deviation in ensemble angle	°	45
ϵ_B	Standard deviation in ensemble biomass	–	0.002
ω	Biomass reduction for ensemble	–	0.005
μ_B	Mean total initial biomass	kg	5e06
σ_R	Standard deviation in individual magnitude	m s^{-1}	0.01
σ_Θ	Standard deviation in individual angle	°	4.5
σ_B	Standard deviation in individual biomass	–	0.002
σ_O	Observation noise	kg	250
Σ_B	Standard deviation in initial biomass	kg	1e06
α_1	Temporal correlation parameter	–	0.984
α_2	Temporal correlation parameter	–	0.129
ρ	Localization parameter	–	6
c	Localization cut-off	–	15
ψ	Inflation factor	–	1.01

In addition to spatial estimates, the raw error between the density values of the model and true model were analyzed. This ground truth error was taken as the difference root-mean squared difference between density estimates from the model and the true derived density estimates:

$$e_T = \sqrt{\frac{1}{l}\sum\nolimits_{i=1}^{l}\left(\overline{X}_i - X_{Ti}\right)^2} \tag{21}$$

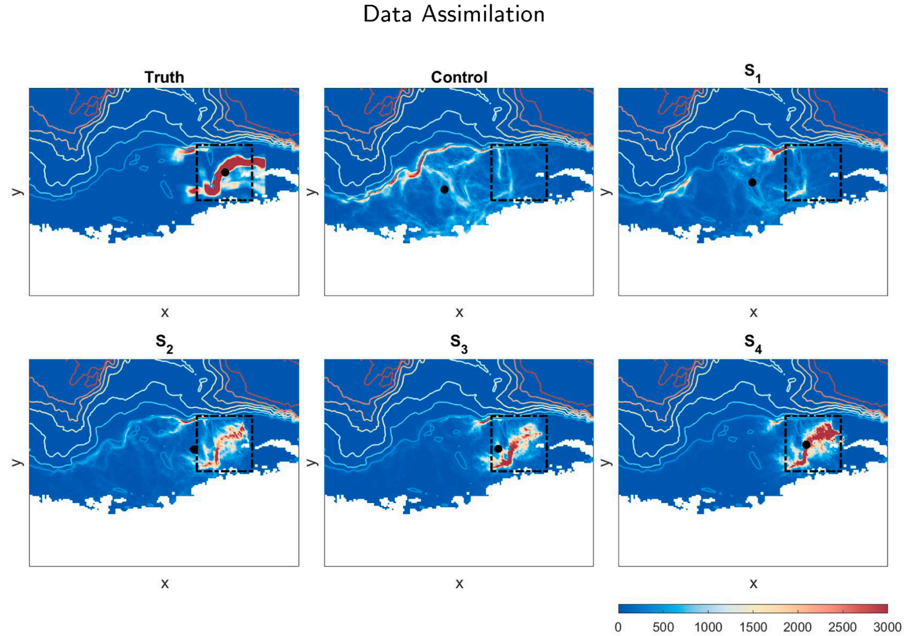
where i is the model coordinate and l is the number of indices within the observation area (Fig. 2).

3. Results

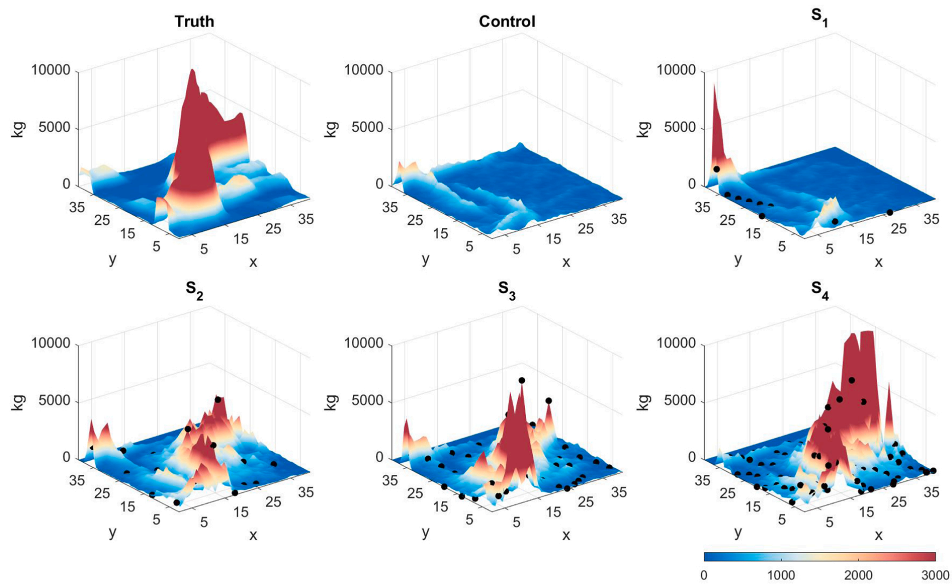
3.1. Qualitative analysis

The ensemble mean of the model ($\overline{X}$) is the best estimate of the model at each time step, and thus in the following results, $\overline{X}$ is the focus for analysis. The modelled migration moves south, with an offshore distribution prior to assimilation. During assimilation, the control simulation continues with this development, while the corrected scenarios develop a more coastal distribution, reflecting the true distribution. Following assimilation, all corrected scenarios deviate from the true distribution, but to a lesser degree than the control distribution.

To visualize the impact of corrections, we plotted derived density maps from two time stamps during the assimilation period, one at day 25 (Fig. 3) and another at day 35 (Fig. 4). The visual comparisons show

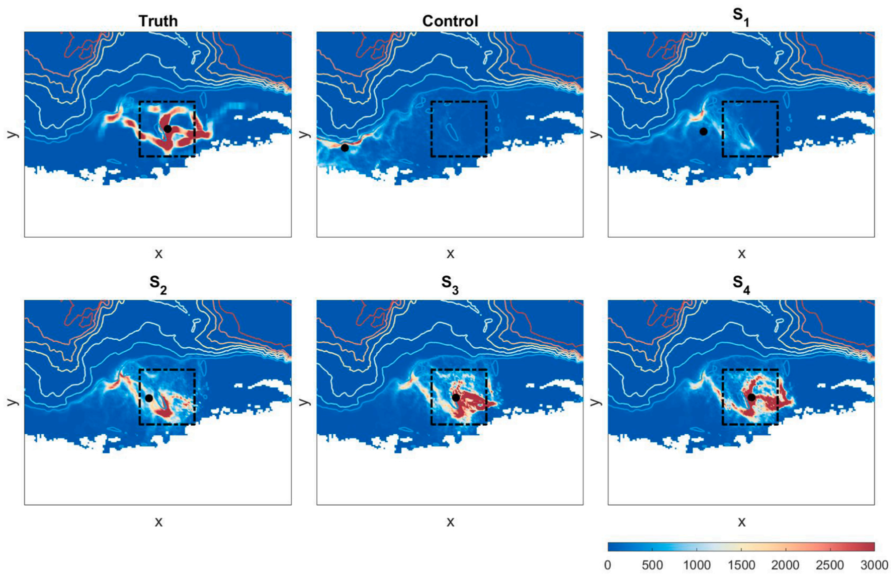


(a) Large scale distribution on day 20.

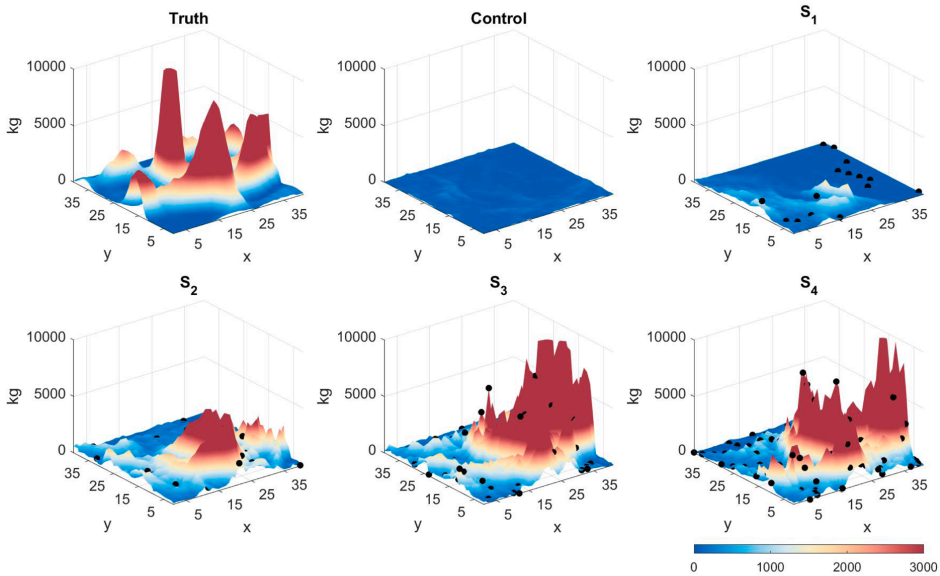


(b) Local distribution on day 20.

Fig. 3. (a) Large scale 2D plot of derived density states of model ($\overline{X}$) and true distribution (X_T) over a selected area of the Norwegian coast on day 20 of the simulation. The density colormap shows values in kg, while the contour lines show depth in metres. The black point shows the centre point (CG). The same colorbar scale is used for both. (b) The local 3D representations of derived states taken from the squared area in (a). Black dots show the location of measurement points. No measurement points were sampled for the control scenario.



(a) Large scale distribution on day 35.



(b) Local distribution on day 35.

Fig. 4. (a) Large scale 2D plot of derived density states of model ($\overline{X}$) and true distribution (X_T) over a selected area of the Norwegian coast on day 35 of the simulation. The density colormap shows values in kg, while the contour lines show depth in metres. The same colorbar scale is used for both. The black point shows the centre point (CG). (b) The local 3D representations of derived states taken from the squared area in (a). Black dots show the location of measurement points. No measurement points were sampled for the control scenario.

large scale distributions (2D plot) concatenated vertically with local distributions (3D surface plot). The large scale model distributions become more similar to the true distribution in the assimilated scenarios. The model CG converges on the true value also. The true distribution tends towards the coast and is concentrated more northerly. Further north (higher on the x axis) there is a clear increase in density values in assimilated scenarios, where measurement values from the true distribution are higher. Further south (lower on the x axis), the density values decrease, as a result of adjustments using zero measurement values.

In any given cell, local densities vary from true values, but on average the ensemble mean approaches the topography of X_T. Location and density of measurement points impact the scale of corrections. The peaks and valleys in the local densities of Fig. 3b and 4b are concentrated in varying locations, related to the position of measurement points. In the case of the control model, the derived density topology is distinct from the true derived topology. With an increasing number of measurement points, the density map starts to resemble the true map. For example, the ridge in S_4 resembles the surface features of the true model (Fig. 4b). In any given cell, the density estimates from $\overline{X}$ may not reflect those from X_T, but on average with increased observation numbers recreates a similar topography.

3.2. Quantitative analysis

The time series of CG of the ensemble mean for three scenarios was compared to the true CG. The CG is calculated in both latitude and longitude axes (Fig. 5). During the assimilation period there is convergence of CG towards the true point. The standard deviation across the ensemble reduces during the assimilation period and is sharply reduced with a higher density of observations. This sharp reduction is pronounced on the first call to the assimilation function. A large number of instances of the model are heavily penalized at this point. The standard deviation increases rapidly post-assimilation. There is faster convergence on latitude, reflecting the greater difference in latitude points, which was the main axis of variation for the simulation period. The inflation factor (ψ) is partly responsible for maintaining the standard deviation across the ensemble. The CG and standard deviation identical in all scenarios prior to assimilation. With a low density of observations, there are relatively weak corrections and convergence on the truth. In all cases there is divergence from the true CG post-assimilation. However, with a higher density of observations, there is less divergence. This is clear when we compare S_3 to S_1.

In Fig. 6, we compare each scenario to the true and control CG (Fig. 6a) and overlap (Fig. 6b). Before the assimilation period, the model and true distribution diverge and there is less overlap. The non-assimilated control model continues to diverge from the truth during the assimilation period. There is immediate divergence from the control on day 18 and convergence to the true CG for all scenarios. This is reflected in the overlap, which approaches a value of one over time.

The ground truth error (e_T) evaluates the raw error in the observation area between the model derived density values and the true derived density estimates from Equation (21). The error increases initially as the initial distribution of the truth and ensemble diverge in spatial characteristics. This pattern continues for the control model, until it eventually plateaus. The e_T is generally reduced from S_1 to S_4, with an initial sharp reduction, followed by a gradual decline in errors, with some irregularities. The e_T remains lower than the control for some days post-assimilation, until it eventually converges to a similar value at the end of the simulation.

4. Discussion

In this article, we have presented a novel general method for

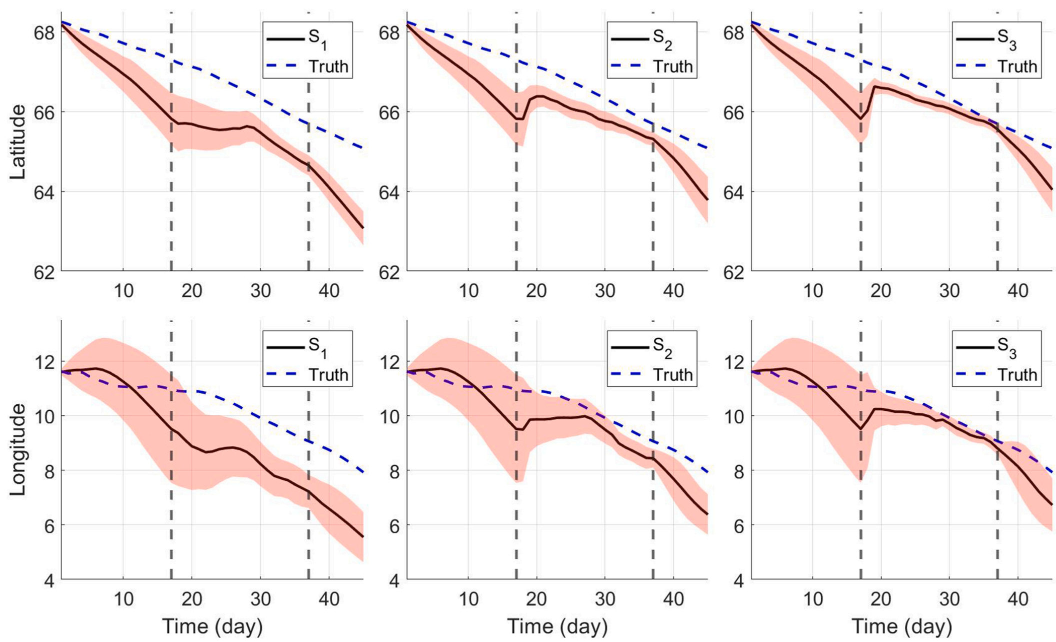
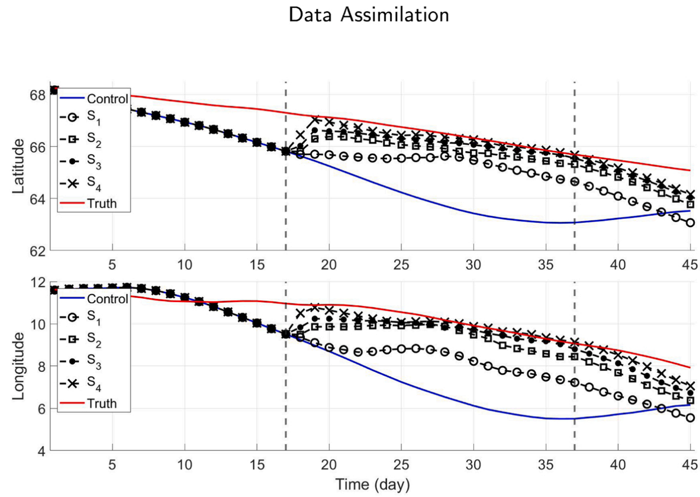
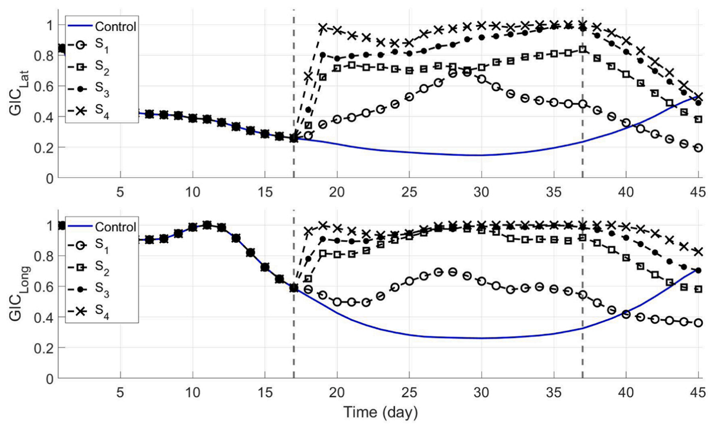


Fig. 5. Time series of Centre of Gravity (CG) in terms of latitude (first row) and longitude (second row) values during the simulation period. The CG of the true derived states (X_T) is shown with the dotted blue line in each panel, while the CG of the ensemble mean of the model derived states ($\overline{X}$) is shown with the black line, with each column representing a separate scenario. The vertical dotted grey lines represent the boundaries of the assimilation period.



(a) Centre points for model and true distribution.



(b) Overlap between model and true distribution.

Fig. 6. (a) Time series of centre points (CG) for the four scenarios, true distribution and control model in terms of latitude (top panel) and longitude (bottom panel). The vertical dotted grey lines represent the boundaries of the assimilation period. (b) Time series of overlap (GIC) between the model and true distributions for the four scenarios and control model in terms of latitude (top panel) and longitude (bottom panel). The vertical dotted grey lines represent the boundaries of the assimilation period.

assimilating data with an IBM operating in a high dimensional system. Assimilating data sources with a population of unique, discrete individuals is challenging, given observation sources like catch data, which do not preserve individual identity. Our suggestion is the use of derived states, which map individuals onto a discrete grid, with each grid cell expressing total densities of individuals. These derived states can then be assimilated with observation data, using an ensemble approach, to calculate a posterior density grid. Derived states can be remapped to the IBM states, without excessive manipulation of the model structure. Such a method is particularly useful for spatially and temporally explicit predictions of fish distributions. In the setup tested here, we compared scenarios for a bounded time period, where observations were available at frequent discrete intervals. The prior- and post-assimilation periods assumed no access to observations. In scenarios with access to many measurement points, the large scale and local density field converge on the true distribution. Importantly, we have shown how the assimilated scenarios outperform the non-assimilated control scenario in spatiotemporal predictions during the assimilation period. Performance is also superior for the time stamps directly succeeding assimilation. Towards the end of the time series, the model estimates eventually diverge from the true distribution and converge on the control case. Future work on incorporating fisheries dependent data can improve predictions and validate this method with real data.

4.1. Making the IBM compatible with the EnKF

The EnKF was chosen given the highly non-linear nature of the system modelled. Additionally, the EnKF shifts values in the model, rather than reinitializing model components. This prevents degeneracy of the model structure since each IBM instance is altered with minimal manipulation during assimilation. The IBM states were perturbed with Gaussian errors, but upon simulation the distribution of the ensemble of derived states becomes non-Gaussian. However, while the EnKF implicitly assumes a Gaussian state-space, it provides good approximate solutions in cases where systems violate this assumption (Katzfuss et al., 2016).

In the real system, the observations will be sampled from an underlying non-negative concentration field (fish per unit area), and furthermore the field will have a bias towards values of zero in locations outside of the distribution of the migrating fish at any time. This has two consequences. First, the assumption of gaussian measurement noise is a poor fit to real world observations outside of the area covered by the migrating fish. Second, perturbation of observations using gaussian distributed random values will lead to a high number of negative values in those same areas. For these reasons, observations were treated as deterministic in this study. To compensate for the lower ensemble spread resulting from this choice, an ensemble inflation factor was applied (Evensen, 2009). To relax the need for adding observation errors, a square root EnKF variant could be considered (for example: Bishop et al., 2001). In future studies using real world observations, the statistics of the sampling process should be investigated in detail for the actual observations made, and the assimilation process customized accordingly. One approach could be to use approximate Bayesian inference along the lines proposed by Eidsvik et al. (2008).

An innovation of our method is the use of derived states that convert from particles to a field of density values. This allows corrections of density values rather than unique individual values, for which we don't have measurement data to describe. This would require, for example, large-scale tagging studies or time-sensitive acoustic back-scattering data, which are not fully developed as of now. Also, a spatial density field is easier to interpret and compare with observation data sources, relative to a cloud of particles at large spatial scales. When mapping the posterior state back to individual states, there were two manipulations. Firstly, the negative X^a values were removed to omit negative biomass values. Secondly, individuals that had zero biomass values (post-assimilation), were moved into positions with positive X^a values, until either none remained to be moved or all positive X^a values were assigned, in a process similar to the randomized redistribution described in Cocucci et al. (2022). This prevented loss of information during assimilation, without heavily intruding on the mechanics of the IBM directly.

The parameter values in assimilation were calibrated to ensure corrections were applied without extreme effects. The inflation factor kept spread around the ensemble, preventing excessive convergence of model on the observations, given observations were treated as deterministic. Localization was used to limit impacts of observations spatially and the choice of localization distance affects the corrections of cells between measurement points. Random perturbations on model states generated variance in the evolution of the migration scenarios. Balance between observation noise and model perturbations determined the overall scale of the corrections. One must note that assimilation is an approximate method of estimation, and operates under the assumption of uncertainty in model states and parameters. More persistent effects of the data assimilation can be achieved by also estimating model parameters in the data assimilation process, and for the present system the average swimming speed is a natural choice. Using parameter estimation, one would not only update the model state, but also attempt to tune the model to better match the real system at a fundamental level.

4.2. Impacts of measurements on the fish distribution

We used the twin model experiment (Fig. 1) to generate virtual observations, gauging the impact of corrections on the model IBM. The twin model was designed to configure a hypothetical shift in the spatial distribution of the fish relative to the prior assumption of our model. Inter-annual shifts in distribution are common in many migratory fish species, as captured often in surveys. For example, the herring spawning migration usually ends with masses of individuals spawning around Møre, but often, spawning occurs further north (Slotte and Fiksen, 2000). Our intention was not to explicate those reasons, but to gain insight into how assimilation of real-time data may modify the distribution to reflect a hypothetical disagreement between modelled and true distributions. In reality, fish distributions are highly uncertain in real-time as we have access to sparse observations, such as catch data. The twin model experiment design is useful as we are omniscient of the underlying true distribution and can easily analyse the impact of measurements.

Qualitatively, the large scale and local spatiotemporal distributions increasingly resemble the true distribution with denser clusters of observations (Fig. 3 and 4). Quantitatively, the centre points and overlap of the model converge on the true indices during assimilation to an increasing degree with more observations (Fig. 6a and b. Additionally, the deviations between ensemble instances is reduced with measurements, meaning the estimates are of higher certainty (Fig. 5). Finally, the ground truth error between the model and true derived density states is reduced with observations, showing, with access to more measurements, the model becomes more predictive in an absolute sense (Fig. 7).

At any one location, corrections are highly sensitive to placement of observation points. For example, at the coordinate (5,38) in (Fig. 3b) there are high density values in scenarios S_1, S_2 and S_3, while this peak is absent in S_4. This is related to the position of measurement points at this step of the analysis and the previous position of measurement points. However, on average, the denser the observations, the more the features reflect the true spatial distribution. The overall topography of S_4 resembles the true distribution more closely at the large and local scale (Fig. 3a and 3b).

4.3. Opportunities for model implementation

Today, there is much interest in utilizing fisher's knowledge, as it is considered part of the best available information for research studies. This is complementary to research survey data, which much work has relied on until now. Utilizing spatially explicit data, such as position and speed from vessel monitoring systems, we can improve our understanding of the state of the fishery in real-time. The estimation approach presented in this article is intended to be coupled with such data sources and thus, facilitate real-time monitoring of fish stocks. This has potential applications in fisheries management, marine planning and tracking of migrations. We note that this method is suggested to support decisions in these areas alongside complementary sources of information. Explicit decisions in fisheries systems are complex and require human deliberation and intervention. Thus, our model offers increased situational awareness without explicitly directing the decision-making process. Decision-making is the responsibility of the end user.

The method also has theoretical value for tuning parameters and

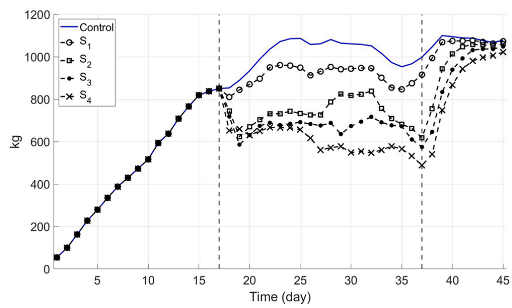


Fig. 7. The ground truth error (e_T) between the model and true distribution for the simulation period in kg. The vertical dotted grey lines represent the boundaries of the assimilation period.

improving models of fish dynamics. We have shown that applying corrections to model estimates improve prior predictions and with enough coverage, model estimates converge on the true spatial distribution. Further work will attempt to validate this method with real fisheries observation data. Furthermore, we wish to improve predictions when observations are not available, for example during time windows with little access to measurements.

CRediT authorship contribution statement

Cian Kelly: Conceptualization, Methodology, Writing, Simulation. **Finn Are Michelsen**: Writing – review & editing. **Morten Omholt Alver**: Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

No data was used for the research described in the article.

Acknowledgements

We greatly acknowledge the support, input and feedback from the project participants: NTNU, SINTEF, NAIS and UiB. The work is part of the FishGuider project, which is funded by the project participants and the Norwegian Research Council (project number 296321).

References

Alver, M.O., Michelsen, F.A., 2015.Data assimilation with SINMOD. Technical Report. SINTEF Fisheries and Aquaculture.
Baker, R.E., Peña, J.M., Jayamohan, J., Jérusalem, A., 2018. Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol. Lett. 14, 20170660 https://doi.org/10.1098/rsbl.2017.0660.
Barbaro, A., Einarsson, B., Birnir, B., Sigurðsson, S., Valdimarsson, H., Pálsson, O., Sveinbjörnsson, S., Sigurðsson, T., 2009. Modelling and simulations of the migration of pelagic fish. ICES J. Mar. Sci. 66, 826–838. https://doi.org/10.1093/icesjms/fsp067.
Bez, N., Walker, E., Gaertner, D., Rivoirard, J., Gaspar, P., 2011. Fishing activity of tuna purse seiners estimated from vessel monitoring system (VMS) data. Can. J. Fish. Aquat. Sci. 68, 1998–2010. https://doi.org/10.1139/f2011-114.
Bishop, C.H., Etherton, B.J., Majumdar, S.J., 2001. Adaptive sampling with the ensemble transform Kalman Filter. Part I: theoretical aspects. Mon. Weather Rev. 129, 420–436. https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.
Buchwald, A.G., Adams, J., Bortz, D.M., Carlton, E.J., 2020. Infectious disease transmission models to predict, evaluate, and improve understanding of COVID-19 trajectory and interventions. Ann. ATS 17, 1204–1206. https://doi.org/10.1513/AnnalsATS.202005-501PS.
Burgers, G., Jan van Leeuwen, P., Evensen, G., 1998. Analysis scheme in the ensemble Kalman filter. Mon. Wea. Rev. 126, 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.
Chu, D., 2011. Technology evolution and advances in fisheries acoustics. J. Mar. Sci. Technol. 19 https://doi.org/10.51400/2709-6998.2188.
Cocucci, T.J., Pulido, M., Aparicio, J.P., Ruíz, J., Simoy, M.I., Rosa, S., 2022. Inference in epidemiological agent-based models using ensemble-based data assimilation. PLOS ONE 17, e0264892. https://doi.org/10.1371/journal.pone.0264892.
DeAngelis, D.L., Grimm, V., 2014. Individual-based models in ecology after four decades. F1000Prime Rep. 6 https://doi.org/10.12703/P6-39. ⟨https://facultyopinions.com/prime/reports/b/6/39/⟩.
Eidsvik, J., Martino, S., Rue, H., 2008. Approximate Bayesian Inference in Spatial Generalized Linear Mixed Models. Scand. J. Stat. https://doi.org/10.1111/j.1467-9469.2008.00621.x.
Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. 99, 10143. https://doi.org/10.1029/94JC00572.
Evensen, G., 2009. Data Assimilation: The Ensemble Kalman Filter. Springer.
Fu, W., She, J., Zhuang, S., 2011. Application of an ensemble optimal interpolation in a north/baltic sea model: Assimilating temperature and salinity profiles. Ocean Model. 40, 227–245. https://doi.org/10.1016/j.ocemod.2011.09.004.
Grimm, V., Railsback, S.F., 2005. Individual-based Modeling and Ecology. stu - student edition ed. Princeton University Press, pp. 3–21 (pp.). ⟨http://www.jstor.org/stable/j.ctt5hhnk8.5⟩ (pp.).
Houtekamer, P., Mitchell, H.L., 2005. Ensemble Kalman filtering. Q. J. R. Meteorol. Soc. 131, 3269–3289. https://doi.org/10.1256/qj.05.135.
Houtekamer, P.L., Mitchell, H.L., 2001. A sequential ensemble kalman filter for atmospheric data assimilation. Mon. Weather Rev. 129, 123–137. https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.
Katzfuss, M., Stroud, J.R., Wikle, C.K., 2016. Understanding the ensemble Kalman Filter. Am. Stat. 70, 350–357. https://doi.org/10.1080/00031305.2016.1141709.
Kelly, C., Michelsen, F.A., Kolding, J., Alver, M.O., 2022. Tuning and development of an individual-based Model of the herring spawning migration. Front. Mar. Sci. 8, 754476 https://doi.org/10.3389/fmars.2021.754476.
Keppenne, C.L., Rienecker, M.M., Jacob, J.P., Kovach, R., 2008. Error covariance modeling in the GMAO ocean ensemble Kalman filter. Mon. Weather Rev. 136, 2964–2982. https://doi.org/10.1175/2007MWR2243.1.
Kieu, L.M., Malleson, N., Heppenstall, A., 2020. Dealing with uncertainty in agent-based models for short-term predictions. R. Soc. Open Sci. 7, 191074 https://doi.org/10.1098/rsos.191074.
Koopman, J.S., Lynch, J.W., 1999. Individual causal models and population system models in epidemiology. Am. J. Public Health 89, 1170–1174. https://doi.org/10.2105/AJPH.89.8.1170.
Mandel, J., 2006.Efficient implementation of the ensemble kalman filter. Center for Computational Mathematics Reports.
Niu, S., Luo, Y., Dietze, M.C., Keenan, T.F., Shi, Z., Li, J., Iii, F.S.C., 2014. The role of data assimilation in predictive ecology. art65 Ecosphere 5. https://doi.org/10.1890/ES13-00273.1.
Simon, E., Bertino, L., 2009. Application of the gaussian anamorphosis to assimilation in a 3-d coupled physical-ecosystem model of the north atlantic with the enkf: a twin experiment. Ocean Sci. 5, 495–510. https://doi.org/10.5194/os-5-495-2009. ⟨https://os.copernicus.org/articles/5/495/2009/⟩.
Slagstad, D., McClimans, T.A., 2005. Modeling the ecosystem dynamics of the barents sea including the marginal ice zone: I. physical and chemical oceanography. J. Mar. Syst. 58, 1–18. https://doi.org/10.1016/j.jmarsys.2005.05.005. ⟨https://www.sciencedirect.com/science/article/pii/S0924796305001296⟩.
Slotte, A., Fiksen, Ø., 2000. State-dependent spawning migration in Norwegian spring-spawning herring. J. Fish. Biol. 56, 138–162. https://doi.org/10.1111/j.1095-8649.2000.tb02091.x.
de Souza, E.N., Boerder, K., Matwin, S., Worm, B., 2016. Improving fishing pattern detection from satellite AIS using data mining and machine learning. PLoS ONE 11, e0158248. https://doi.org/10.1371/journal.pone.0158248.
Tu, C.Y., Tseng, Y.H., Chiu, T.S., Shen, M.L., Hsieh, C.H., 2012. Using coupled fish behavior-hydrodynamic model to investigate spawning migration of Japanese anchovy, Engraulis japonicus, from the East China Sea to Taiwan: Spawning migration model of Japanese anchovy. Fish. Oceanogr. 21, 255–268. https://doi.org/10.1111/j.1365-2419.2012.00619.x.
Ward, J.A., Evans, A.J., Malleson, N.S., 2016. Dynamic calibration of agent-based models using data assimilation. R. Soc. Open Sci. 3, 150703 https://doi.org/10.1098/rsos.150703.
Woillez, M., Poulard, J.C., Rivoirard, J., Petitgas, P., Bez, N., 2007. Indices for capturing spatial patterns and their evolution in time, with application to European hake (Merluccius merluccius) in the Bay of Biscay. ICES J. Mar. Sci. 64, 537–550. https://doi.org/10.1093/icesjms/fsm025.

Paper 3

Kelly, C., Michelsen, F. A., and Alver, M. O. (In prep). Model-based estimation of fish catch potential using assimilation of synthetic measurements from a neural network driven by fishing vessel activity data.

This paper is not yet published and is therefore not included.

Paper 4

Kelly, C., Michelsen, F. A., Reite, K. J., Kolding, J., Varpe, Ø., Berset, A. P., and Alver, M. O. (2022). Capturing big fisheries data: Integrating fishers' knowledge in a web-based decision support tool. Frontiers in Marine Science 9, 1051879. doi:10.3389/fmars.2022.1051879

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 12 December 2022
DOI 10.3389/fmars.2022.1051879

Check for updates

OPEN ACCESS

EDITED BY
Nathalie A. Steins,
Wageningen Marine Research
IJmuiden, Netherlands

REVIEWED BY
Julia Calderwood,
Marine Institute, Ireland
Martin Pastoors,
Pelagic Freezer-trawler Association,
Netherlands

*CORRESPONDENCE
Cian Kelly
cian.kelly@ntnu.no

SPECIALTY SECTION
This article was submitted to
Marine Fisheries, Aquaculture and
Living Resources,
a section of the journal
Frontiers in Marine Science

RECEIVED 23 September 2022
ACCEPTED 21 November 2022
PUBLISHED 12 December 2022

CITATION
Kelly C, Michelsen FA, Reite KJ, Kolding J, Varpe Ø, Berset AP and Alver MO (2022) Capturing big fisheries data: Integrating fishers' knowledge in a web-based decision support tool.
Front. Mar. Sci. 9:1051879.
doi: 10.3389/fmars.2022.1051879

Capturing big fisheries data: Integrating fishers' knowledge in a web-based decision support tool

Cian Kelly[1*], Finn Are Michelsen[2], Karl Johan Reite[2], Jeppe Kolding[3], Øystein Varpe[3,4], Are Prytz Berset[5] and Morten Omholt Alver[1]

[1]Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway, [2]Fisheries and new Biomarine Industry, SINTEF Ocean, Trondheim, Norway, [3]Department of Biological Sciences, University of Bergen, Bergen, Norway, [4]Norwegian Institute for Nature Research (NINA), Bergen, Norway, [5]GAGN AS consulting, Ålesund, Norway

There is increasing interest in utilizing fishers' knowledge to better understand the marine environment, given the spatial extent and temporal resolution of fishing vessel operations. Furthermore, fishers' knowledge is part of the best available information needed for sustainable harvesting of stocks, marine spatial planning and large-scale monitoring of fishing activity. However, there are difficulties with integrating such information into advisory processes. Data is often not systematically collected in a structured manner and there are issues around sharing of information within the industry, and between industry and research partners. Decision support systems for fishing planning and routing can integrate relevant information in a systematic way, which both incentivizes vessels to share information beneficial to their operations and capture time sensitive big datasets for marine research. The project Fishguider has been developing such a web-based decision support tool since 2019, together with partners in the Norwegian fishing fleet. The objectives of the project are twofold: 1) To provide a tool which provides relevant model and observation data to skippers, thus supporting sustainable fishing activity. 2) To foster bidirectional information flow between research and fishing activity by transfer of salient knowledge (both experiential and data-driven), thereby supporting knowledge creation for research and advisory processes. Here we provide a conceptual framework of the tool, along with current status and developments, while outlining specific challenges faced. We also present experiential input from fishers' regarding what they consider important sources of information when actively fishing, and how this has guided the development of the tool. We also explore potential benefits of utilizing such experiential knowledge generally. Moreover, we detail how such collaborations between industry and research may rapidly produce extensive, structured datasets for research and input into management of stocks. Ultimately, we suggest that such decision support services will motivate fishing vessels to collect and share data, while the available data will foster increased research, improving the decision support tool itself and consequently knowledge of the oceans, its fish stocks and fishing activities.

KEYWORDS

decision support, interface, knowledge, experience, observations, model

1 Introduction

There is a global movement towards better understanding and utilization of data and experience of fishers in order to inform research activity and management decisions (Johannes et al., 2008; Stephenson et al., 2016; Dyrset et al., 2022). This is due to an increasing awareness that it is advantageous to consider fishers' knowledge, as the quantity of information available to modern fleets is vast given the temporal and spatial extent of global fisheries operations, which is estimated at four times the spatial extent of agriculture (Kroodsma et al., 2018). Such knowledge includes the experiences of fishers themselves and information processing systems onboard, and is considered part of the best available information (Stephenson et al., 2016). This information can be used in stock assessment, marine spatial planning and mapping species abundance and distribution (Holm and Soma, 2016). Modern applications to real-time monitoring of vessel tracks can screen for illegal fishing activity and map the global footprint of effort (de Souza et al., 2016; Kroodsma et al., 2018). Remote sensing of environmental variables may be a cost-effective method of supporting fishing activities (Santos, 2000). The recent paper from Jones et al. (2022) demonstrates how high resolution data from the US reference fleet has contributed to abundance indices for several stocks, while footprints of fishing vessels can inform planning of offshore wind projects. A similar Norwegian reference fleet program found that gathering species and age composition data from fishing vessels is a cost-effective method of sampling and producing CPUE time series for cod, haddock and redfish (Hjelle et al., 2021).

Given the multitude benefits of using fishers' knowledge to inform policy, it begs the question why it's underutilized? For catch data, there is the issue of bias in samples for density estimates, as catch logs exclusively record instances of fishing activity, neglecting areas not targeted by fishers, which biases predictions of species distributions (Karp et al., 2022). Also, given the unsystematic way much of fishers' knowledge is handled, it is often neglected (Hind, 2015). This means that although the quantity of information is high, the quality is highly variable and potentially skewed. It's challenging to filter from individual knowledge claims to scientific input that is legitimate and salient for decision-making (Brattland, 2013; Röckmann et al., 2015). Regardless, there is the charge that biologists don't take fishers ecological knowledge seriously, where such information can avert collapses of spawning stocks (Johannes et al., 2008).

In addition to the benefits to decision makers of incorporating fishers' knowledge, there are increasingly clear incentives for fishers to contribute in meaningful ways. The historical trajectory of the Norwegian fishing industry has been to long-term sustainable harvesting. For example, advances in fish finding equipment, with the uptake of echosounders and sonar, has improved vertical and horizontal profiling of fish and led to more offshore and targeted exploitation of stocks (Nakken, 2008; Gordon and Hannesson, 2015). Advances in mechanical winches for trawling gear reduced the labour involved in hauling nets, and introduction of non-rotting synthetic fibres made nets pressure resistant, increasing catch efficiency (Hamre and Nakken, 1971; Jennings et al., 2001). A modern purse seiner makes particularly effective use of the listed advances, and is relatively fuel efficient, using approximately 0.1kg of fuel per kilo of fish (Schau et al., 2009). In addition to technological developments, structural changes to the fleet, from introduction of tradeable quotas, decommissioning schemes and general movement of labour away from the industry, have reduced overcapacity and increased operating margins (Standal and Asche, 2018; Fisheries Directorate, 2021). However, such technological advances are a double-edged sword. The cumulative impact of technological innovation, especially mechanical hauling, led to increased catch rates and the collapse of the North-East Atlantic herring stocks in the 1970s (Fiksen and Slotte, 2002; Gordon and Hannesson, 2015; Standal and Asche, 2018). Therefore, prudent management of stocks is essential alongside such developments.

Such modernization of the industry means vessels spend long periods at sea with advanced equipment such as echo-sounders and sonars, covering vast geographical areas, and thus, have access to large quantities of information. To utilize such information effectively, collaboration between researchers and fishers is important. Increased knowledge of the environment fishers operate within can contribute to achieving long-term objectives. In this work, the first objective is to supply fishers with information that reduces time spent searching for fishing grounds, while simultaneously reducing fuel use of vessels. The second objective is to build a system that automatically captures and stores data gathered while vessels are at sea.

Decision support systems (DSS) are tools that can integrate knowledge sources to achieve these objectives. Formulations of DSS include: manufacturing DSS that help deliver products and services to customers, clinical DSS used to improve healthcare delivery using clinical knowledge and patient information, and organizational DSS used to inform decisions on complex activities within a large organization (e.g. governmental body), through integration of knowledge such as norms and roles in the organization (Jacob and Pirkul, 1992; Sala et al., 2019; Sutton et al., 2020).

In the maritime context, the major application of DSS tools have been in the shipping industry, aimed mainly towards optimizing speed and routes of vessels and avoiding collisions between vessels (Lazarowska, 2014; Bal Beşikçi et al., 2016; Lee et al., 2018). As described in Gilman et al. (2022), forms of shipping DSS can be applied to support fishing route optimization. In this article we will refer to such computer based tools in the context of supporting stakeholder decisions in the fishing industry specifically. In this context, DSS that have been applied to support management decisions in spatial allocation of effort and bycatch management (Truong et al., 2005; Granado et al., 2021). Moreover, they have been used to provide model estimates on presence and size of fishing banks directly to fishers, thus reducing time and fuel spent on fishing operations (Iglesias et al., 2007). There are a wide range of actors who may benefit from such tools, from managers to ship owners and skippers. As researchers, it's important that research knowledge is integrated with the needs of industry to facilitate uptake of tools Röckmann et al. (2015). In this way, DSS can provide a vital link between research and the fishing industry, where two way information transfer can garner interest in results of research as directed towards their operation, while at the same time encouraging more engagement between parties.

The Fishguider project began in 2019 as a science-industry research collaboration aimed at both reducing fuel use and search time of the Norwegian fishing fleet and fostering two-way information transfer between fishers and researchers. Importantly, this was an industry directed project, where an umbrella organization of motivated fishing companies was founded to partially fund work activities, under the name of the North Atlantic Institute for Sustainable Fishing (NAIS). In consultation between NAIS and researchers, a DSS tool was conceived of as an appropriate method to co-create knowledge necessary to achieve long-term objectives of industry. Such co-creation of knowledge between research and industry is an effective way of building mutual trust between researchers and fishers (Holm and Soma, 2016). Additionally, the delivery of such a software solution is well placed for systematically capturing and sharing data between participants, and supporting management decisions through production of salient and legitimate knowledge. A key component of the project is the participation of fishers in the pilot program to ascertain the feasibility of the DSS tool. There is evidence suggesting that participation can increase in science-industry collaborations if results are perceived to be positive for industry (Calderwood et al., 2021).

In this article, we present the conceptual framework for the DSS tool being developed as part of Fishguider and it's current status, reflecting on similarities to other DSS tools mentioned above. The capacity to systematically capture and share information through a user interface is explored and we discuss how data-driven input and experiential knowledge inform the development of this interface. In addition, a questionnaire is presented, detailing fishers' experiences of which factors are most relevant when considering when and where to fish. Finally, we consider challenges in interpreting, capturing and sharing knowledge through this project.

2 Literature on DSS tools in the fisheries context

DSS tools are described as computer-based programs that integrate diverse information sources in order to support complex decision-making processes (Truong et al., 2005; Bal Beşikçi et al., 2016; Granado et al., 2021; Gilman et al., 2022). In a DSS, computer output virtually represents the real fisheries system, reducing uncertainties that constrain decision making (Truong et al., 2005). Decisions that require support systems usually address problems where there are competing interest groups, such as fishing effort allocation. Therefore, human participation and intervention are essential in the process (Bal Beşikçi et al., 2016; Gilman et al., 2022). In this way, DSS plays a supporting role in decision-making, rather than an executive role. Regardless, there are a multitude of areas where they can give insight, as shown in Table 1. The two broad applications are within fisheries management and industry-related optimization. A diverse range of inputs are used, from data-driven input such as remote sensing and vessel speed to knowledge based input from interdisciplinary collaboration and stakeholder engagements.

Fishers face many practical issues when searching for fishing grounds, such as uncertainties in weather conditions, quality and location of fish, and prices and costs being variable. In the face of these issues, they must make concrete decisions on how to organise fishing activities. The scales of fishing activity decisions can be separated based on duration into three categories: strategic, tactical and operational decisions. Strategic decisions (weeks to months to years) refers to long-term planning of location and timing of fishing based on expectations of both the market and fishing possibilities (Reite et al., 2021). Tactical decisions (hours to days) are decisions about which fishing grounds to visit, the number of grounds to visit and where and when to return to port to land catches (Granado et al., 2021). Long-term tactical decisions may involve, for example, planning of whether to target herring or mackerel based on market prices (Reite et al., 2021). Short-term tactical decisions include motion planning of fishing vessels and controlling position and course of vessels relative to schools of fish (Haugen and Imsland, 2019; Haugen and Kyllingstad, 2021; Kyllingstad et al., 2021). Operational decisions (near real-time) involve immediate control of the vessel, such as speed and heading of fishing vessels between waypoints defined through tactical decisions Granado et al. (2021). Assuming waypoints are clearly defined, operational decisions can be informed through routing optimization, which has been addressed using DSS tools in the shipping industry to reduce fuel consumption (Bal Beşikçi et al., 2016; Granado et al., 2021). However, defining strategic and tactical decisions is a complex task for fishing vessels searching for fish, given the uncertainties in stock distribution and abundance at these scales and therefore, the Fishguider DSS tool is designed to support these decisions.

DSS tools are designed with of a number of interconnected components. Fundamentally, they require high quality data sources, where data can be obtained from remote sensing of environmental variables such as sea surface temperature, weather archive data, information systems on board such as positional data, as well as manual input from ship operators (Iglesias et al., 2007; Bal Beşikçi et al., 2016; Lee et al., 2018). Data can also be gathered from national or global databases, such as historical catch data, where the data is directly relevant to fishers

TABLE 1 A selection of literature sorted chronologically on decision support in fisheries and shipping, describing the input used and the area of application.

Article	Input	Application
Lane and Stephenson (1998)	Interdisciplinary knowledge	Co-management of fisheries
Truong et al. (2005)	Fisheries-dependent data	Optimize fishing schedules
Koutroumanidis et al. (2006)	Time series modelling of fisheries landings	Fisheries management
Iglesias et al. (2007)	Remote sensing	Prediction of fishing banks
Carrick and Ostendorf (2007)	Spatial information and survey data	Economically sustainable fishing activity
Jarre et al. (2008)	Knowledge-based logical system	Ecosystems approach to fisheries management
Vinu Chandran et al. (2009)	Remote sensing	Identify potential fishing grounds
Azadivar et al. (2009)	Systems approach- optimization of schedules	Spatial management of stocks
Dowling et al. (2016)	Questionnaire and stock assessment	Management strategy evaluation
Hobday et al. (2016)	Dynamic ocean modelling	Fishing activity
Bal Beşikçi et al. (2016)	Vessel speed	Reducing fuel consumption of ships
Reite et al. (2017)	Vessel operation and energy system	Reducing fuel consumption of ships
Lee et al. (2018)	Vessel speed	Reducing fuel consumption of ships
Macher et al. (2018)	Stakeholder engagement	Management Strategy evaluation
Skjong et al. (2019)	Combining onboard sensors and mathematical models	Generic decision support
Granado et al. (2021)	Vessel speed and heading	Fishing route optimization
Macher et al. (2021)	Transdisciplinary partnerships	Ecosystem based management in fisheries
Reite et al. (2021)	Oceanographic simulations, catch data analyses	Prediction of fishing grounds
Gilman et al. (2022)	Categorization of mitigation	Bycatch management

operations and can improve their situational awareness. This data is uploaded to a database, where information is compiled and can be queried directly by the user. There is also typically a model solver which takes input and produces estimates of relevant information. Often the problems are complex and require pattern detection through machine learning and data mining algorithms, where artificial neural networks have been particularly effective (Bal Beşikçi et al., 2016).

This information is mapped to a user interface, where the user (fisher or manager) may query databases directly (Bal Beşikçi et al., 2016). User interfaces are typically tuned to the experience and requirements of the user. Information is often displayed in interactive layers which compile the most salient knowledge for decision making. For example, (Granado et al., 2021) describes decision layers developed for fishers to display routes based on an optimization algorithm which allows for interaction with the user. In addition, explicit costs associated with decisions may be displayed, such as in management decisions where there are multiple conflicting objectives such as safety and economic viability (Gilman et al., 2022).

3 Case study: The fishguider DSS tool

3.1 Description

The Fishguider DSS tool was requested by fishing companies working together in an umbrella organization called North Atlantic Institute for Sustainable Fishing (NAIS), who spend much time and fuel searching for fishing grounds, while lacking systemized knowledge to assist in making informed decisions on where and when to fish. The system desired should aid in communication of information between fishing vessels and allow them to both contribute and ascertain relevant information to minimize uncertainties when operating. Importantly, the fishers involved are motivated to collaborate with researchers and understand the ecosystem they operate within. The interested parties wish to build a knowledge base to ensure present and future sustainable harvesting. Specifically, a DSS system may aid in handling decisions made in light of the complexities of climate change, the potential shifts in distributions of fish stocks and instabilities in fuel prices. Improving the situational awareness through knowledge co-creation will help the fishers meet these demands, specifically aiding with strategic and tactical decision making.

The DSS is currently designed as a proof of concept which can be refined and scaled for industrial use. The scaling of the system relies in part on connecting more vessels to the project. Therefore a pilot programme of vessels is underway, where they are now utilizing the system during the fishing season. Participants are from a variety of fisheries, targeting both demersal and pelagic species, with different gears, quotas and sizes of vessels. At the time of writing, there are 19 vessels involved with the pilot project. Of those 6 are classed as coastal vessels, 3 large coastal, 7 ocean-going trawlers and 3 ocean-going purse seiner. In addition, 16 of these vessels are above

21m in length. These classes determine the quotas and areas where the vessel operates. For example, ocean-going vessels cannot operate within fjords without special permission. For pelagic species, the fishers are most active from October to December when herring overwinter near the coast and in Northern fjords and then again January to March during the spawning migration and spawning for the herring, while mackerel are mainly targeted during their wintering cycle in southern Norway from September to December when the market prices are highest, although there is inter-annual variability (Varpe et al., 2005; Nøttestad et al., 2016; Ølmheim, 2021; Reite et al., 2021). The following sections describe the DSS tool according to its data sources, model-based inputs and the user interface (Figure 1).

3.2 Knowledge sources for DSS

3.2.1 Experiential

In an effort to build a tool that is useful for the fishers, a survey in the form of a questionnaire was designed and 13 of the skippers in NAIS responded. The questionnaire was conducted by phone in 2020 in Norwegian and answers were translated into English. An online or paper-based solution were not possible due to logistic challenges with communication. The skippers surveyed are the most actively involved in the project. They target both pelagic and demersal species, but we learned from project meetings that they perceive the most immediate use of the tool in targeting herring and mackerel. Therefore, the questionnaire focused on these two species.

There were two categories of questions asked. The first related to the importance of a variety of factors in deciding when and where the new fishing season should begin (Figure 2). This set of questions corresponded to strategic decisions. The second related to to the importance of factors during the season (Figure 3). This set of questions corresponded to tactical decisions. The survey was designed to gauge the information fishers in NAIS consider important, regardless of availability, in making decisions to choose fishing grounds. Questions were chosen based on wide-ranging project meetings between researchers and active fishers in NAIS. Fishers expressed the importance of a full ecosystem understanding in decision-making, from plankton to whales, and therefore, questions of this nature were included.

Respondents were asked to rate the importance of items from both categories on an evaluative rating scale from 1 to 6, 6 being the highest value. Items were categorized based on their importance to fishers now and their potential importance in the future. The questionnaires displayed are the results for questions related to the targeting of herring. The mean values for the 13 respondents are displayed in the horizontal barplots (Figure 2 and 3). Given the sample surveyed, we don't assume this is completely representative of the fishing industry as a whole, especially given the number of large vessels involved. Additionally, social factors such as business structures and working rhythm may influence strategic and tactical decisions (Schadeberg et al., 2021). Nevertheless, the survey offered relevant input to the design of the support tool in order to make it

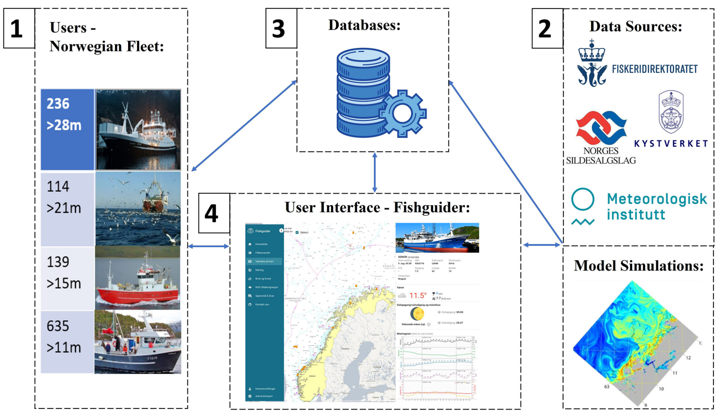


FIGURE 1
Conceptual model of the Fishguider tool: 1) The Norwegian Fleet of vessels over 11m in length who may contribute information, both from experiential knowledge and from information systems onboard vessels (such as satellite and acoustic data). 2) Fishers can access external information, such as meteorological forecasts, real-time auction prices and relevant model output. 3) The data sources are collected in databases developed in conjunction with the project. 4) The final user interface is a web portal that displays relevant layers to the skipper. The design of the interface is largely driven by the requests of participating fishers.

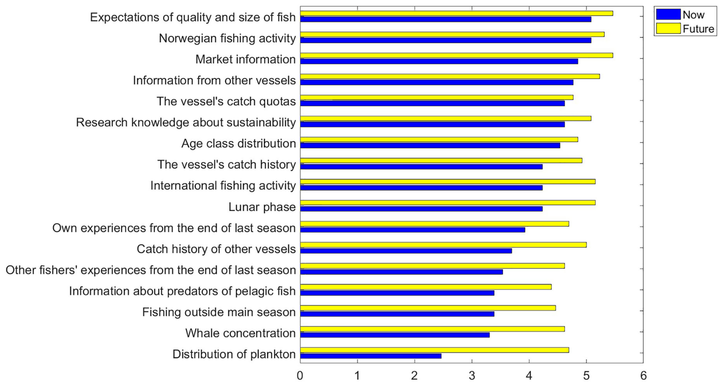


FIGURE 2
Response to Question: How important are the following factors when deciding when and where the new fishing season should begin? The blue bars indicate how important they are now, while the yellow bars signal the importance of better information in the future. The bars display the mean value (N = 13).

relevant for industry implementation, which was our main objective. A table of questionnaire responses for both herring and mackerel can be found in the appendix, with additional informal commentary from respondents included (Appendix A).

Generally, practical considerations such as the vessel's quota, catch history and Norwegian fishing activity are important now and are considered important in the future in strategic decision making (Figure 2). Ecological information such as whale concentration, plankton forecasts and information about predators are not strategically utilized now, but attaining such information is perceived as useful in the future.

Similarly, when asked what factors are important in tactical decision making, plankton forecasts and distribution of seabirds and whales are not utilized now, but such information may be valuable in the future (Figure 3). It must be noted that the perspective of fishers on the data they use today is likely based on their ongoing assessment of the quality of data available, while the question of future utility is made under the assumption that high quality data may be readily available. In real-time fishing activity, communication with other vessels, market forecasts and weather forecasts are seen as the most important factors to consider.

The questionnaire, complimented by meetings with fishers, has informed the development of the web portal over the past two years. Many of the information sources fishers deem important are publicly available and a major part of the work is compiling these in one place. Currently, communication between fishers is being facilitated through messaging options in the portal, weather forecasts are attained from the meteorological institute, such as wind speeds and swell at the vessels' location, and oceanographic data (particularly ocean currents), plankton and fish distribution data from model simulations are included. In addition, based on project meetings, it was discovered that fishers deemed the lunar phase an indicator of the timing of the initiation of herring spawning migrations. This factor was thus included in the questionnaire, and has been integrated into the support tool (Figure 4). Finally, the catch history of vessels, market information such as auction prices and vessel quotas, and the trajectories of individual vessels are now being connected to the portal. In the next sections, we explore the major knowledge sources available for the DSS tool.

3.2.2 Data-driven

In addition to the fishers'experiences as obtained from the questionnaire, data is being gathered from several sources. The catch and activity reporting (ERS) and Vessel Monitoring System (VMS) are electronic reporting systems for fisheries data provided by the Norwegian Directorate of Fisheries (https://www.fiskeridir.no/English/Fisheries/Electronic-Reporting-Systems). Whereas ERS data includes vessel positions for fishing activities such as 'in catch operation', 'pumping' and 'steaming' to and from harbour, VMS data includes more detailed position data for all types of vessels with a length of 15 meters and above, logged at minimum one hour sampling frequency. The ERS logs replaced physical logs of catches in 2005 where there is a principle of reporting all Norwegian fishing activity, with widespread adoption (https://www.fiskeridir.no/English/Fisheries/Electronic-Reporting-Systems).

Automatic identification systems data (AIS) are, like VMS data, detailed position data for all types of vessels. There are many sources available such as Marine Traffic (https://www.

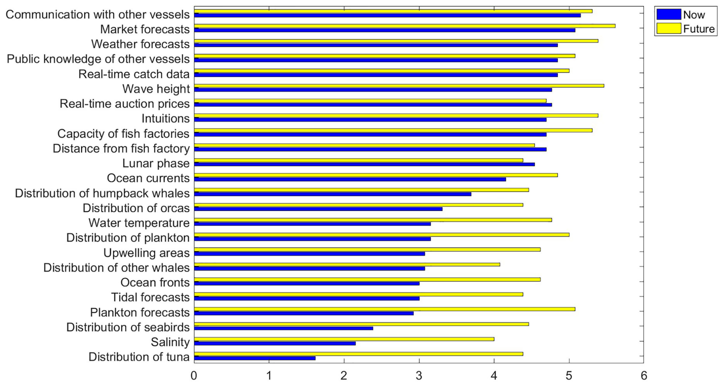


FIGURE 3
Response to Question: How important are the following factors for choosing a fishing spot during the season? The blue bars indicate how important they are now, while the yellow bars signal the importance of better information in the future. The bars display the mean value (N=13).

marinetraffic.com/en/ais/) and The Norwegian Coastal Administration (https://www.kystverket.no/en/navigation-and-monitoring/ais/access-to-ais-data/), which provides AIS data in real time, either as raw data or online traffic information displayed in charts. AIS data is primarily used by coastal administration to avoid shipping collisions and locating a given vessel quickly in an emergency situation. Given the time sensitivity needed to avoid collisions or respond to emergencies, data is transmitted approximately every 10 seconds. ERS, VMS and AIS data are complementary data for monitoring fishing vessel movements which are being integrated in the DSS tool.

The Norwegian Fishers' Sales Organization for Pelagic Fish (or *Norges Sildesalslag* in Norwegian: https://www.sildelaget.no/) is a fisher-owned sales organization that trades fish through an electronic auction. Fresh catches are offered to buyers while vessels are at sea, after the catch is registered over phone, and a commission price on the value of each catch is payed by the fisher (0.65 percent of each catch). Real-time auction prices are highly relevant to direct decisions on fishing, as reflected in the questionnaire responses, and will be integrated in the DSS tool (Figure 3).

Finally, vessels in the Norwegian fleet continuously gather observations using sonar and echosounder, but this data is usually discarded. A future version of Fishguider is expected to collect, aggregate and make decision support based on a fleet supplying such observations, and this work has begun.

3.2.3 Model simulations

Model simulations of conditions alongshore and offshore the Norwegian coast are currently being integrated into the DSS tool. Ocean model estimates of sea surface temperature, current and salinity are loaded from a model called SINMOD (Slagstad and McClimans, 2005). The output from this model has a 4km resolution and is centered on the Norwegian Sea. The model has a time resolution of 10 seconds. An eulerian model of the copepod species Calanus finmarchicus has been coupled to the SINMOD model, where plankton distributions are mainly driven by atmospheric fields including wind, air temperature and precipitation, river discharge, and bottom topography (bathymetry) (Wassmann et al., 2006). This species is a key prey item for many pelagic stocks in the Norwegian Sea. In addition, a model of the spawning migration of herring is coupled to SINMOD, where information on current, temperature and bathymetry are used to drive the fish motion towards their spawning areas (Kelly et al., 2022). All model ouputs are loaded to the web portal in near real-time.

Minimizing the gap between the true system and the model estimates depends on integrating as many vessels into the project. Capturing of data by vessels included in the project can strengthen input to these models, which improves their predictive capacity. An Ensemble Kalman Filter setup has been designed to allow assimilation of observation data into the migration model (Kelly et al., submitted). In the long-term this can develop larger datasets for studying effects of climate change, understanding life cycles and migrations of fish, and providing input into stock assessment.

3.3 Databases

Both national and international databases are being integrated into the web portal. FishGuider is currently being

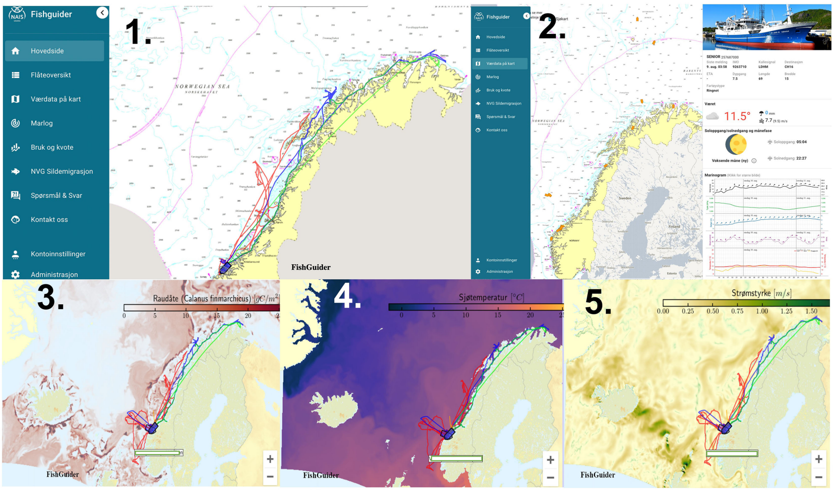


FIGURE 4
A selection of output layers in the Fishguider portal, with Norwegian text: 1) Homepage with tabs for various layers centered on the Norwegian Sea. The red, blue and green lines are the tracks of individual vessels based on GPS coordinates. 2) Weather data and forecast from the position of one of the NAIS vessels based on meterological institute data. 3) Modelled *Calanus finmarchicus* distribution and abundance in grams of carbon per meter squared 4) Sea surface temperature output on a single day in degrees Celsius. 5) Horizontal components of current velocities in meters per second.

integrated with FiskInfo (https://fhf-prod.azurewebsites.net), Kystverkets NAIS service (https://nais.kystverket.no/), BarentsWatch (https://www.barentswatch.no/en), Marine Traffic (https://www.marinetraffic.com/en/ais/), Ocean Resource Watch (https://resourcewatch.org/dashboards/ocean-watch) and other complementary information tools.

3.4 User interface

The user interface of Fishguider web portal provides layers of information tailored to the needs of the fisher. As mentioned, it has been curated according to the experiences of fishers, considering important features when planning fishing operations and in real-time (Figures 2, 3). The web portal has a fleet overview tab, which details the vessels involved in the project and their specifications. There are also messaging possibilities and contact points for the fishers, if they have any difficulties with usability of the tool. The tracks that are displayed in the interface are based on GPS transmitters installed onboard, which are being trialed (Figure 4). This interface has facilitated information flow between fishers and researchers, where fishers now have access to spatiotemporal data on current, temperature, nitrate, plankton and herring from research-based models developed, while researchers have access to observations from vessels which provide input to the models (Figure 4). This input can allow improvement of the accuracy of the model predictions, while also correcting errors in model output.

4 Takeaways

4.1 Knowledge co-creation

Such collaborations between industry and research may rapidly produce extensive, structured datasets for research and input into management of stocks. Involving enough fishers and/or vessels improves collaboration and will give more access to quality information. In general, the more vessels involved, the better. The vessels involved are representative of a subset of the coastal and oceanic fleet in Norway. The project results are presented at project meetings and industry conferences, such as Norfishing and The Midsund Conference (*Midsundkonferansen*). In this way, both participants and industry at large can provide input and feedback on the design of the tool. Additionally, as fishing companies are partly funding the project, key results are communicated to these larger audiences. Fishers seem interested to participate and share data on condition that the platform will yield useful input in guiding operations.

Sharing of information between researchers and fishers is key to achieving this. Given that the fishers themselves are interested in this work, they have been quite open to sharing vessel data. Due to competition between fishers there is a potential scepticism in sharing information, but this issue has become less important the last decade, as individual vessel quotas are the main limiting factor, less vessels participate and there is more transparency because of open data sources (AIS, VMS, ERS). By limiting the spread of information to those who contribute, this should not be a big issue for this project in the future. Furthermore, engagement with the DSS tool will develop the user experience and friendliness of the application, which in turn will encourage more participants to join the project.

The questionnaire results give insight into what fishers deem important factors in strategic and tactical decision making. However, the small sample size and evaluative scale used means we cannot gauge the prioritization of factors by fishers. Further work should consider ranking factors and matching available knowledge based on this. It is important to avoid the inclusion of all desired sources in the DSS tool at the cost of adequate user experience.

4.2 Research-based inputs

Collaboration between fishers and scientists has provided direct results that are salient for decisions regarding fishing activity. For example, fishing routes can be optimized to meet strategic, tactical and operational decisions (Granado et al., 2021). Spatially and temporally explicit maps of fish distribution are particularly useful for planning operations, and can be obtained through analysing remote sensing data (Iglesias et al., 2007). In our work, a migration model has been implemented to estimate the development of the herring spawning migration (Kelly et al., 2022). Lifting the modelling of fish migration, implementation and visualisation of the model to a level that gives the fishers useful additional information and promotes more active engagement with the tool.

Coordinating the various ideas and requirements from the diverse set of fishers is challenging, as there can be variability in the problems they face, depending on the target stock, vessel size and fuel consumption. Additionally, when asked about the utility of various factors in the future, almost all were considered useful in some way, especially research output which is not capitalized upon today (Figure 2 and Figure 3). Therefore, continuous dialogue and soliciting of feedback from fishers is central to qualifying the true importance of information for decision making. Understanding the behaviour of fishing vessels themselves is also important and progress has been made on categorizing activities automatically based on position, speed and heading of vessels (de Souza et al., 2016).

4.3 Advisory processes

Finally, DSS tools can contribute to advisory processes by reducing uncertainties involved in executive decision making. For example, offshore wind farms are planned along the coast of Norway, and the potential conflicts with industry may be anticipated and captured through understanding the movements of fishing vessels. Firstly, the formalized knowledge of fishers is relevant input into decision-making on management of stocks throughout the season. Fine grain information about individual vessels can improve CPUE indices, an important input for stock assessments (Campbell, 2004). Secondly, the legality of fishing activity can be monitored through automatic detection of vessel activities (Arasteh et al., 2020). Automatic monitoring of activity from data they can contribute, may be more desirable and less invasive than physical monitoring through observers or drones. Thirdly, collaboration between researchers, fishers and managers can improve decision-making on these issues. Of crucial importance is that the knowledge base is considered legitimate to decision-makers (Röckmann et al., 2015).

5 Conclusion

The Fishguider project has developed a functional pilot of a DSS tool which is being used for testing and development of the interface, databases and models, while simultaneously helping connect more vessels to the project. Currently, only small number of companies are involved, but the entire Norwegian fleet of fishing vessels are seen as potential participants. Fishguider was setup to primarily facilitate environmentally sustainable fishing activity by reducing search time and fuel consumption of fishing vessels. As the project has evolved, fuel prices have risen, and concerns about climate change have grown, making DSS tools like this one even more crucial. The knowledge being created should therefore be central to fishing activity, marine research and management going forward.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

CK: Wrote the manuscript and created figures and tables. FM: Provided details about model and observations inputs. KR: Provided information about the fishing industry and decision support literature. JK: Proofread the manuscript and provided additional references. ØV: Proofread the manuscript, provided feedback and added discussion points. AB. Provided details and images of the Fishguider tool. MA: Gave detailed feedback. All authors were involved editing a shared version of the manuscript in overleaf. All authors contributed to the article and approved the submitted version.

Funding

The work is part of the FishGuider project, which is funded by the project participants and the Norwegian Research Council (project number 296321).

Acknowledgments

We greatly acknowledge the support, input and feedback from the project participants: NTNU, SINTEF, NAIS and UiB.

Conflict of interest

Author AB is employed by GAGN AS consulting.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.1051879/full#supplementary-material

References

Arasteh, S., Tayebi, M. A., Zohrevand, Z., Glässer, U., Shahir, A. Y., Saeedi, P., et al. (2020). "Fishing vessels activity detection from longitudinal AIS data," in *Proceedings of the 28th international conference on advances in geographic information systems* (Seattle WA USA: ACM), 347–356. doi: 10.1145/3397536.3422267

Azadivar, F., Truong, T., and Jiao, Y. (2009). A decision support system for fisheries management using operations research and systems science approach. *Expert Syst. Appl.* 36, 2971–2978. doi: 10.1016/j.eswa.2008.01.080

Bal Beşikçi, E., Arslan, O., Turan, O., and Ölçer, A. (2016). An artificial neural network based decision support system for energy efficient ship operations. *Comput. Operations. Res.* 66, 393–401. doi: 10.1016/j.cor.2015.04.004

Brattland, C. (2013). Proving fishers right. effects of the integration of experience-based knowledge in ecosystem-based management. *Acta Borealia.* 30, 39–59. doi: 10.1080/08003831.2013.768053

Calderwood, J., Marshall, C. T., Haflinger, K., Alfaro-Shigueto, J., Mangel, J. C., and Reid, D. G. (2021). An evaluation of information sharing schemes to identify what motivates fishers to share catch information. *ICES. J. Mar. Sci.*, 1–22. doi: 10.1093/icesjms/fsab252

Campbell, R. A. (2004). CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. *Fish. Res.* 70, 209–227. doi: 10.1016/j.fishres.2004.08.026

Carrick, N., and Ostendorf, B. (2007). Development of a spatial decision support system (DSS) for the Spencer gulf penaeid prawn fishery, south Australia. *Environ. Model. Software.* 22, 137–148. doi: 10.1016/j.envsoft.2005.07.025

de Souza, E. N., Boerder, K., Matwin, S., and Worm, B. (2016). Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PloS One* 11, e0158248. doi: 10.1371/journal.pone.0158248

Dowling, N., Wilson, J., Rudd, M., Babcock, E., Caillaux, M., Cope, J., et al. (2016). "FishPath: A decision support system for assessing and managing data- and capacity- limited fisheries," in *Assessing and managing data-limited fish stocks*. Eds. T. Quinn II, J. Armstrong, M. Baker, J. Heifetz and D. Witherell (Alaska, United States: Alaska Sea Grant, University of Alaska Fairbansk). doi: 10.4027/amdlfs.2016.03

Dyrset, G., Margaryan, L., and Stensland, S. (2022). Local knowledge, social identity and conflicts around traditional marine salmon fisheries – a case from mid-Norway. *Fish. Manage. Ecol.* 29, 131–142. doi: 10.1111/fme.12522

Fiksen, Ø., and Slotte, A. (2002). Stock-environment recruitment models for norwegian spring spawning herring (*clupea harengus*). *Can. J. Fish. Aquat. Sci.* 59, 211–217. doi: 10.1139/f02-002

Fisheries Directorate (2021). *Profitability survey of the Norwegian fishing fleet 2019* (Norway: Fisheries Directorate).

Gilman, E., Hall, M., Booth, H., Gupta, T., Chaloupka, M., Fennell, H., et al. (2022). A decision support tool for integrated fisheries bycatch management. *Rev. Fish. Biol. Fish.* 32, 441–472. doi: 10.1007/s11160-021-09693-5

Gordon, D. V., and Hannesson, R. (2015). The Norwegian winter herring fishery: A story of technological progress and stock collapse. *Land. Econ.* 91, 362–385. doi: 10.3368/le.91.2.362

Granado, I., Hernando, L., Galparsoro, I., Gabiña, G., Groba, C., Prellezo, R., et al. (2021). Towards a framework for fishing route optimization decision support systems: Review of the state-of-the-art and challenges. *J. Cleaner. Prod.* 320, 128661. doi: 10.1016/j.jclepro.2021.128661

Hamre, J., and Nakken, O. (1971). "Technological aspects of the modern norwegian purse seine fisheries," in *Modern fishing gear of the world*, vol. 3. Ed. H. Kristjonsson (Oxford: Fishing News (Books) Ltd).

Haugen, J., and Imsland, L. (2019). Optimization-based motion planning for trawling. *J. Mar. Sci. Technol.* 24, 984–995. doi: 10.1007/s00773-018-0600-0

Haugen, J., and Kyllingstad, L. T. (2021). *Course planning for purse seiners* (Norway: SINTEF). Available at: https://hdl.handle.net/11250/2991650.

Hind, E. J. (2015). A review of the past, the present, and the future of fishers' knowledge research: a challenge to established fisheries science. *ICES. J. Mar. Sci.* 72, 341–358. doi: 10.1093/icesjms/fsu169

Hjelle, H., Gundersen, S., Nedreaas, K., Vølstad, J. H., and Kolding, J. (2021). "Fleet composition, fishing effort and contributions to science," in *The coastal reference fleet 2007-2019* (Norway: Institute of Marine Research).

Hobday, A. J., Spillman, C. M., Paige Eveson, J., and Hartog, ,. J. R. (2016). Seasonal forecasting for decision support in marine fisheries and aquaculture. *Fish. Oceanography.* 25, 45–56. doi: 10.1111/fog.12083

Holm, P., and Soma, K. (2016). Fishers' information in governance — a matter of trust. *Curr. Opin. Environ. Sustainabil.* 18, 115–121. doi: 10.1016/j.cosust.2015.12.005

Iglesias, A., Dafonte, C., Arcay, B., and Cotos, J. (2007). Integration of remote sensing techniques and connectionist models for decision support in fishing catches. *Environ. Model. Software.* 22, 862–870. doi: 10.1016/j.envsoft.2006.05.017

Jacob, V. S., and Pirkul, H. (1992). Organizational decision support systems. *Int. J. Man-Machine. Stud.* 36, 817–832. doi: 10.1016/0020-7373(92)90074-U

Jarre, A., Paterson, B., Moloney, C. L., Miller, D. C., Field, J. G., and Starfield, ,. A. M. (2008). Knowledge-based systems as decision support tools in an ecosystem approach to fisheries: Comparing a fuzzy-logic and a rule-based approach. *Prog. Oceanography.* 79, 390–400. doi: 10.1016/j.pocean.2008.10.010

Jennings, S., Kaiser, M. J., and Reynolds, ,. J. D. (2001). *Marine fisheries ecology* (Oxford ; Malden, MA, USA: Blackwell Science).

Johannes, R., Freeman, M., and Hamilton, R. (2008). Ignore fishers' knowledge and miss the boat. *Fish. Fish.* 1, 257–271. doi: 10.1111/j.1467-2979.2000.00019.x

Jones, A. W., Burchard, K. A., Mercer, A. M., Hoey, J. J., Morin, M. D., Gianesin, G. L., et al. (2022). Learning from the study fleet: Maintenance of a Large-scale reference fleet for northeast U.S. fisheries. *Front. Mar. Sci.* 9. doi: 10.3389/fmars.2022.869560

Karp, M. A., Brodie, S., Smith, J. A., Richerson, K., Selden, R. L., Liu, O. R., et al. (2022). Projecting species distributions using fishery-dependent data. *Fish. Fish.* 1–22 doi: 10.1111/faf.12711

Kelly, C., Michelsen, F. A., Kolding, J., and Alver, M. O. (2022). Tuning and development of an individual-based model of the herring spawning migration. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.754476

Koutroumanidis, T., Iliadis, L., and Sylaios, G. K. (2006). Time-series modeling of fishery landings using ARIMA models and fuzzy expected intervals software. *Environ. Model. Software.* 21, 1711–1721. doi: 10.1016/j.envsoft.2005.09.001

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., et al. (2018). Tracking the global footprint of fisheries. *Science* 359, 904–908. doi: 10.1126/science.aao5646

Kyllingstad, L. T., Haugen, J., Tenningen, M., and Breen, M. (2021). *Catch control in seine fisheries targeting pelagic species* (Norway: SINTEF). Available at: https://hdl.handle.net/11250/2991646.

Lane, D. E., and Stephenson, R. L. (1998). Fisheries co-management: Organization, process and decision support. *J. Northwest. Atlantic. Fishery. Sci.* 23, 251–265. doi: 10.2960/J.v23.a16

Lazarowska, A. (2014). Ant colony optimization based navigational decision support system. *Proc. Comput. Sci.* 35, 1013–1022. doi: 10.1016/j.procs.2014.08.187

Lee, H., Aydin, N., Choi, Y., Lekhavat, S., and Irani, Z. (2018). A decision support system for vessel speed decision in maritime logistics using weather archive big data. *Comput. Operations. Res.* 98, 330–342. doi: 10.1016/j.cor.2017.06.005

Ølmheim, O. (2021). *Economic and biological figures from Norwegian fisheries – 2020* (Norway: Fishery Directorate).

Macher, C., Bertignac, M., Guyader, O., Frangoudes, K., Frésard, M., Le Grand, C., et al. (2018). The role of technical protocols and partnership engagement in developing a decision support framework for fisheries management. *J. Environ. Manage.* 223, 503–516. doi: 10.1016/j.jenvman.2018.06.063

Macher, C., Steins, N. A., Ballesteros, M., Kraan, M., Frangoudes, K., Bailly, D., et al. (2021). Towards transdisciplinary decision-support processes in fisheries: experiences and recommendations from a multidisciplinary collective of researchers. *Aquat. Living. Resour.* 34, 13. doi: 10.1051/alr/2021010

Nøttestad, L., Utne, K. R., Óskarsson, G. J., Jónsson, S. T., Jacobsen, J. A., Tangen, Ø., et al. (2016). Quantifying changes in abundance, biomass, and spatial distribution of northeast Atlantic mackerel (scomber scombrus) in the Nordic seas from 2007 to 2014. *ICES. J. Mar. Sci.: J. du. Conseil.* 73, 359–373. doi: 10.1093/icesjms/fsv218

O. Nakken (Ed.) (2008). *Norwegian Spring-spawning herring & northeast Arctic cod: 100 years of research management* (Trondheim: Tapir Academic Press). OCLC: ocn262434637.

Reite, K.-J., Haugen, J., Michelsen, F., and Aarsæther, K. (2021). Sustainable and added value small pelagics fisheries pilots. In: Södergård C, Mildorf T, Habyarimana E, Berre A. J., Fernandes J. A. and Zinke-Wehlmann C (eds) Big data in bioeconomy (Cham: Springer). doi: 10.1007/978-3-030-71069-9_30

Reite, K. J., Ladstein, J., and Haugen, J. (2017). "Data-driven real-time decision support and its application to hybrid propulsion systems," in *International conference on offshore mechanics and Arctic engineering*, vol. 57748. (Trondheim, Norway: American Society of Mechanical Engineers), V07BT06A024.

Röckmann, C., van Leeuwen, J., Goldsborough, D., Kraan, M., and Piet, G. (2015). The interaction triangle as a tool for understanding stakeholder interactions in marine ecosystem based management. *Mar. Policy* 52, 155–162. doi: 10.1016/j.marpol.2014.10.019

Sala, R., Pezzotta, G., Pirola, F., and Huang, G. Q. (2019). Decision-support system-based service delivery in the product-service system context: Literature review and gap analysis. *Proc. CIRP.* 83, 126–131. doi: 10.1016/j.procir.2019.03.140

Santos, A. P. (2000). Fisheries oceanography using satellite and airborne remote sensing methods: a review. *Fish. Res.* 49, 1–20. doi: 10.1016/S0165-7836(00)00201-0

Schadeberg, A., Kraan, M., and Hamon, K. G. (2021). Beyond métiers: social factors influence fisher behaviour. *ICES. J. Mar. Sci.* 78, 1530–1541. doi: 10.1093/icesjms/fsab050

Schau, E. M., Ellingsen, H., Endal, A., and Aanondsen, S. A. (2009). Energy consumption in the Norwegian fisheries. *J. Cleaner. Prod.* 17, 325–334. doi: 10.1016/j.jclepro.2008.08.015

Skjong, S., Kyllingstad, L. T., Reite, K.-J., Haugen, J., Ladstein, J., and Aarsæther, K. G. (2019). "Generic on-board decision support system framework for marine operations," in *Ocean engineering*, vol. Volume 7A. (Glasgow, Scotland, UK: American Society of Mechanical Engineers), V07AT06A032. doi: 10.1115/OMAE2019-95146

Slagstad, D., and McClimans, T. A. (2005). Modeling the ecosystem dynamics of the barents sea including the marginal ice zone: I. physical and chemical oceanography. *J. Mar. Syst.* 58, 1–18. doi: 10.1016/j.jmarsys.2005.05.005

Standal, D., and Asche, F. (2018). Hesitant reforms: The Norwegian approach towards ITQ's. *Mar. Policy* 88, 58–63. doi: 10.1016/j.marpol.2017.11.005

Stephenson, R. L., Paul, S., Pastoors, M. A., Kraan, M., Holm, P., Wiber, M., et al. (2016). Integrating fishers' knowledge research in science and management. *ICES. J. Mar. Sci.* 73, 1459–1465. doi: 10.1093/icesjms/fsw025

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., and Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digital. Med.* 3, 17. doi: 10.1038/s41746-020-0221-y

Truong, T., Rothschild, B., and Azadivar, F. (2005). "Decision support system for fisheries management," in *Proceedings of the winter simulation conferenc* (Orlando, FL. USA: IEEE), 2107–2111. doi: 10.1109/WSC.2005.1574494

Varpe, Ø., Fiksen, Ø., and Slotte, A. (2005). Meta-ecosystems and biological energy transport from ocean to coast: the ecological importance of herring migration. *Oecologia* 146, 443–451. doi: 10.1007/s00442-005-0219-9

Vinu Chandran, R., Jeyaram, A., Jayaraman, V., Manoj, S., Rajitha, K., and Mukherjee, C. K. (2009). Prioritization of satellite-derived potential fishery grounds: an analytical hierarchical approach-based model using spatial and non-spatial data. *Int. J. Remote Sens.* 30, 4479–4491. doi: 10.1080/01431160802577980

Wassmann, P., Slagstad, D., Riser, C. W., and Reigstad, M. (2006). Modelling the ecosystem dynamics of the barents sea including the marginal ice zone: II. carbon flux and interannual variability. *J. Mar. Syst.* 59, 1–24. doi: 10.1016/j.jmarsys.2005.05.006

Bibliography

Adibi, P., Pranovi, F., Raffaetà, A., Russo, E., Silvestri, C., Simeoni, M., et al. (2020). Predicting Fishing Effort and Catch Using Semantic Trajectories and Machine Learning. In *Multiple-Aspect Analysis of Semantic Trajectories*, eds. K. Tserpes, C. Renso, and S. Matwin (Cham: Springer International Publishing), 83–99

Alver, M. O., Broch, O. J., Melle, W., Bagøien, E., and Slagstad, D. (2016). Validation of an Eulerian population model for the marine copepod Calanus finmarchicus in the Norwegian Sea. *Journal of Marine Systems* 160, 81–93. doi: 10.1016/j.jmarsys.2016.04.004

Alver, M. O. and Michelsen, F. A. (2015). *Data assimilation with SINMOD*. Tech. rep., SINTEF Fisheries and Aquaculture

Azadivar, F., Truong, T., and Jiao, Y. (2009). A decision support system for fisheries management using operations research and systems science approach. *Expert Systems with Applications* 36, 2971–2978. doi:10.1016/j.eswa.2008.01.080

Baker, R. E., Peña, J.-M., Jayamohan, J., and Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biol. Lett.* 14, 20170660. doi:10.1098/rsbl.2017.0660

Bal Beşikçi, E., Arslan, O., Turan, O., and Ölçer, A. (2016). An artificial neural network based decision support system for energy efficient ship operations. *Computers & Operations Research* 66, 393–401. doi:10.1016/j.cor.2015.04.004

Balchen, J. (1981). Recent Progress in the Control of Fish Behavior. *IFAC Proceedings Volumes* 14, 3627–3632. doi:10.1016/S1474-6670(17)64013-8

Balchen, J. G. (1980). Modeling and Identification of Marine Ecological Systems with Applications in Management of Fish Resources and Planning of Fisheries Operations. *Modeling, Identification and Control: A Norwegian Research Bulletin* 1, 67–68. doi:10.4173/mic.1980.2.1

Balchen, J. G. (2000). Thirty Years of Research on the Application of Cybernetic Methods in Fisheries and Aquaculture Technology. *Modeling, Identification and Control: A Norwegian Research Bulletin* 21, 3–64. doi:10.4173/mic.2000.1.1

Barbaro, A., Einarsson, B., Birnir, B., Sigurdsson, S., Valdimarsson, H., Pálsson, O., et al. (2009). Modelling and simulations of the migration of pelagic fish. *ICES Journal of Marine Science* 66, 826–838. doi:10.1093/icesjms/fsp067

Bauer, S. and Klaassen, M. (2013). Mechanistic models of animal migration behaviour - their diversity, structure and use. *J Anim Ecol* 82, 498–508. doi: 10.1111/1365-2656.12054

Beaugrand, G., Reid, P. C., Ibañez, F., Lindley, J. A., and Edwards, M. (2002). Reorganization of North Atlantic Marine Copepod Biodiversity and Climate. *Science* 296, 1692–1694. doi:10.1126/science.1071329

Beveridge, M. C. M., Thilsted, S. H., Phillips, M. J., Metian, M., Troell, M., and Hall, S. J. (2013). Meeting the food and nutrition needs of the poor: the role of fish and the opportunities and challenges emerging from the rise of aquaculture [a]. *Journal of Fish Biology* 83, 1067–1084. doi:10.1111/jfb.12187

Bez, N., Walker, E., Gaertner, D., Rivoirard, J., and Gaspar, P. (2011). Fishing activity of tuna purse seiners estimated from vessel monitoring system (VMS) data. *Can. J. Fish. Aquat. Sci.* 68, 1998–2010. doi:10.1139/f2011-114

Blaxter, J. H. S. (1985). The Herring: A Successful Species? *Canadian Journal of Fisheries and Aquatic Sciences* 42, s21–s30. doi:10.1139/f85-259

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences* 99, 7280–7287. doi:10.1073/pnas.082080899

Boyd, R., Sibly, R., Hyder, K., Walker, N., Thorpe, R., and Roy, S. (2020). Simulating the summer feeding distribution of Northeast Atlantic mackerel with a mechanistic individual-based model. *Progress in Oceanography* 183, 102299. doi:10.1016/j.pocean.2020.102299

Breivik, M. and Sand, G. (2009). Jens Glad Balchen: A Norwegian Pioneer in Engineering Cybernetics. *Modeling, Identification and Control: A Norwegian Research Bulletin* 30, 101–125. doi:10.4173/mic.2009.3.2

Bruge, A., Alvarez, P., Fontán, A., Cotano, U., and Chust, G. (2016). Thermal Niche Tracking and Future Distribution of Atlantic Mackerel Spawning in Response to Ocean Warming. *Frontiers in Marine Science* 3. doi:10.3389/fmars.2016.00086

Béné, C., Barange, M., Subasinghe, R., Pinstrup-Andersen, P., Merino, G., Hemre, G.-I., et al. (2015). Feeding 9 billion by 2050 – Putting fish back on the menu. *Food Security* 7, 261–274. doi:10.1007/s12571-015-0427-z

Cheung, W. W. L., Watson, R., and Pauly, D. (2013). Signature of ocean warming in global fisheries catch. *Nature* 497, 365–368. doi:10.1038/nature12156

Cocucci, T. J., Pulido, M., Aparicio, J. P., Ruíz, J., Simoy, M. I., and Rosa, S. (2022). Inference in epidemiological agent-based models using ensemble-based data assimilation. *PLOS ONE* 17, e0264892. doi:10.1371/journal.pone.0264892

de Souza, E. N., Boerder, K., Matwin, S., and Worm, B. (2016). Improving Fishing Pattern Detection from Satellite AIS Using Data Mining and Machine Learning. *PLoS ONE* 11, e0158248. doi:10.1371/journal.pone.0158248

DeAngelis, D. L. and Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000Prime Reports* 6. doi:10.12703/P6-39

DeAngelis, D. L. and Mooij, W. M. (2005). Individual-Based Modeling of Ecological and Evolutionary Processes. *Annual Review of Ecology, Evolution, and Systematics* 36, 147–168. doi:10.1146/annurev.ecolsys.36.102003.152644

Dellacasa, A. (1987). Refrigerated transport by sea. *International Journal of Refrigeration* 10, 349–352. doi:10.1016/0140-7007(87)90121-6

Deshpande, P. C., Brattebø, H., and Fet, A. M. (2019). A method to extract fishers' knowledge (FK) to generate evidence for sustainable management of fishing gears. *MethodsX* 6, 1044–1053. doi:10.1016/j.mex.2019.05.008

Dowling, N., Wilson, J., Rudd, M., Babcock, E., Caillaux, M., Cope, J., et al. (2016). FishPath: A Decision Support System for Assessing and Managing Data- and Capacity- Limited Fisheries. In *Assessing and Managing Data-Limited Fish Stocks*, eds. T. Quinn II, J. Armstrong, M. Baker, J. Heifetz, and D. Witherell (Alaska Sea Grant, University of Alaska Fairbansk). doi:10.4027/amdlfs.2016.03

Dragesund, O. (1970). Factors influencing year-class strength of norwegian spring spawning herring (clupea harengus l.). *Fiskeridirektoratets Skrifter Serie Havundersøkelser* 15, 381–450

Dyrset, G., Margaryan, L., and Stensland, S. (2022). Local knowledge, social identity and conflicts around traditional marine salmon fisheries – A case from Mid-Norway. *Fisheries Management and Ecology* 29, 131–142. doi:10.1111/fme.12522

Edwards, M. and Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430, 881–884. doi:10.1038/nature02808

Elde, S., Kvalvik, I., Nøstvold, H., Rødbotten, R., Dalmannsdottir, S., Halland, H., et al. (2018). *The Arctic as a Food Producing Region. Phase 1: Current status in five Arctic countries*. Tech. rep., Nofima AS

Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter* (Springer)

Fernö, A., Pitcher, T. J., Melle, W., Nøttestad, L., Mackinson, S., Hollingworth, C., et al. (1998). The challenge of the herring in the Norwegian sea: Making optimal collective spatial decisions. *Sarsia* 83, 149–167. doi:10.1080/00364827.1998.10413679

Fiksen, Ø. and Slotte, A. (2002). Stock-environment recruitment models for norwegian spring spawning herring (*Clupea harengus*). *Canadian Journal of Fisheries and Aquatic Sciences* 59, 211–217. doi:10.1139/f02-002

Fisheries Directorate (2021). *Profitability survey of the Norwegian fishing fleet 2019*. Tech. rep., Fishery Directorate, Norway

Fu, W., She, J., and Zhuang, S. (2011). Application of an ensemble optimal interpolation in a north/baltic sea model: Assimilating temperature and salinity profiles. *Ocean Modelling* 40, 227–245. doi:https://doi.org/10.1016/j.ocemod.2011.09.004

Gallagher, C. A., Chudzinska, M., Larsen-Gray, A., Pollock, C. J., Sells, S. N., White, P. J. C., et al. (2021). From theory to practice in pattern-oriented modelling: identifying and using empirical patterns in predictive models. *Biological Reviews* 96, 1868–1888. doi:10.1111/brv.12729

Gardner, M. (1970). Mathematical games. *Scientific American* 223, 120–123

Giske, J., Huse, G., and Fiksen, O. (1998). Modelling spatial dynamics of fish. *Reviews in Fish Biology and Fisheries* 8, 57–91. doi:10.1023/A:1008864517488

Gordon, D. V. and Hannesson, R. (2015). The Norwegian Winter Herring Fishery: A Story of Technological Progress and Stock Collapse. *Land Economics* 91, 362–385. Publisher: [Board of Regents of the University of Wisconsin System, University of Wisconsin Press]

Granado, I., Hernando, L., Galparsoro, I., Gabiña, G., Groba, C., Prellezo, R., et al. (2021). Towards a framework for fishing route optimization decision support systems: Review of the state-of-the-art and challenges. *Journal of Cleaner Production* 320, 128661. doi:10.1016/j.jclepro.2021.128661

Grimm, V. (1999). Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling* 115, 129–148. doi:10.1016/S0304-3800(98)00188-4

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198, 115–126. doi:https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V. and Railsback, S. F. (2005). *Individual-based Modeling and Ecology* (Princeton University Press). Stu - student edition edn., 3–21

Gullestad, P., Abotnes, A. M., Bakke, G., Skern-Mauritzen, M., Nedreaas, K., and Søvik, G. (2017). Towards ecosystem-based fisheries management in Norway – Practical tools for keeping track of relevant issues and prioritising management efforts. *Marine Policy* 77, 104–110. doi:10.1016/j.marpol.2016.11.032

Gullestad, P., Aglen, A., Bjordal, A., Blom, G., Johansen, S., Krog, J., et al. (2014). Changing attitudes 1970–2012: evolution of the Norwegian management framework to prevent overfishing and to secure long-term sustainability. *ICES Journal of Marine Science* 71, 173–182. doi:10.1093/icesjms/fst094

Gullestad, P., Blom, G., Bakke, G., and Bogstad, B. (2015). The "Discard Ban Package": Experiences in efforts to improve the exploitation patterns in Norwegian fisheries. *Marine Policy* 54, 1–9. doi:10.1016/j.marpol.2014.09.025

Gvoždík, L. (2018). Just what is the thermal niche? *Oikos* 127, 1701–1710. doi:10.1111/oik.05563

Hagen-Zanker, A. (2009). An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation. *International Journal of Geographical Information Science* 23, 61–73. doi:10.1080/13658810802570317

Hamre, J. and Nakken, O. (1971). Technological aspects of the modern norwegian purse seine fisheries. In *Modern Fishing Gear of the World, Vol. 3 (ed. H. Kristjonsson).* (Fishing News (Books) Ltd. Oxford.)

Handegard, N. O., Algrøy, T., Eikvil, L., Hammersland, H., Tenningen, M., and Ona, E. (2021). Smart fisheries in norway: Partnership between science, technology, and the fishing sector

Hannesson, R. (2007). Growth accounting in a fishery. *Journal of Environmental Economics and Management* 53, 364–376. doi:10.1016/j.jeem.2006.11.001

Hannesson, R., Salvanes, K. G., and Squires, D. (2010). Technological Change and the Tragedy of the Commons: The Lofoten Fishery over 130 Years. *Land Economics* 86, 746–765. Publisher: [Board of Regents of the University of Wisconsin System, University of Wisconsin Press]

Hatlebrekke, H. H., Gundersen, S., Nedreaas, K. H., Vølstad, J. H., and Kolding, J. (2021). The coastal reference fleet 2007-2019-fleet composition, fishing effort and contributions to science. *Rapport fra havforskningen*

Haugen, J. and Imsland, L. (2019). Optimization-based motion planning for trawling. *Journal of Marine Science and Technology* 24, 984–995. doi:10.1007/s00773-018-0600-0

Hilborn, R. and Walters, C. J. (2015). *Quantitative fisheries stock assessment: choice, dynamics and uncertainty* (Dordrecht: Springer), digital duplication of the orig. ed. 1992 edn.

Hind, E. J. (2015). A review of the past, the present, and the future of fishers' knowledge research: a challenge to established fisheries science. *ICES Journal of Marine Science* 72, 341–358. doi:10.1093/icesjms/fsu169

Holland, J. H. (2006). Studying complex adaptive systems. *Journal of systems science and complexity* 19, 1–8

Houtekamer, P. and Mitchell, H. L. (2005). Ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.* 131, 3269–3289. doi:10.1256/qj.05.135

Houtekamer, P. L. and Mitchell, H. L. (2001). A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review* 129, 123 – 137. doi:10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2

Huse, G., Railsback, S., and Feronö, A. (2002). Modelling changes in migration pattern of herring: collective behaviour and numerical domination. *Journal of Fish Biology* 60, 571–582. doi:10.1111/j.1095-8649.2002.tb01685.x

Huston, M., DeAngelis, D., and Post, W. (1988). New Computer Models Unify Ecological Theory. *BioScience* 38, 682–691. doi:10.2307/1310870

Iglesias, A., Dafonte, C., Arcay, B., and Cotos, J. (2007). Integration of remote sensing techniques and connectionist models for decision support in fishing catches. *Environmental Modelling & Software* 22, 862–870. doi: 10.1016/j.envsoft.2006.05.017

Jacob, V. S. and Pirkul, H. (1992). Organizational decision support systems. *International Journal of Man-Machine Studies* 36, 817–832. doi:https://doi.org/10.1016/0020-7373(92)90074-U

Jennings, S., Kaiser, M. J., and Reynolds, J. D. (2001). *Marine fisheries ecology* (Oxford ; Malden, MA, USA: Blackwell Science)

Jones, A. W., Burchard, K. A., Mercer, A. M., Hoey, J. J., Morin, M. D., Gianesin, G. L., et al. (2022). Learning From the Study Fleet: Maintenance of a Large-Scale Reference Fleet for Northeast U.S. Fisheries. *Frontiers in Marine Science* 9, 869560. doi:10.3389/fmars.2022.869560

Karp, M. A., Brodie, S., Smith, J. A., Richerson, K., Selden, R. L., Liu, O. R., et al. (2022). Projecting species distributions using fishery-dependent data. *Fish and Fisheries* , faf.12711doi:10.1111/faf.12711

Katzfuss, M., Stroud, J. R., and Wikle, C. K. (2016). Understanding the Ensemble Kalman Filter. *The American Statistician* 70, 350–357. doi:10.1080/00031305.2016.1141709

Keppenne, C. L., Rienecker, M. M., Jacob, J. P., and Kovach, R. (2008). Error Covariance Modeling in the GMAO Ocean Ensemble Kalman Filter. *Monthly Weather Review* 136, 2964–2982. doi:10.1175/2007MWR2243.1

Kieu, L.-M., Malleson, N., and Heppenstall, A. (2020). Dealing with uncertainty in agent-based models for short-term predictions. *R. Soc. open sci.* 7, 191074. doi:10.1098/rsos.191074

Kooijman, B. (2009). *Dynamic Energy Budget Theory for Metabolic Organisation* (Cambridge University Press), 3 edn. doi:10.1017/CBO9780511805400

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., et al. (2018). Tracking the global footprint of fisheries. *Science* 359, 904–908. doi:10.1126/science.aao5646

Lane, D. E. and Stephenson, R. L. (1998). Fisheries co-management: Organization, process and decision support. *Journal of Northwest Atlantic Fishery Science* 23

Langbehn, T. J. and Varpe, Ø. (2017). Sea-ice loss boosts visual search: fish foraging and changing pelagic interactions in polar oceans. *Global Change Biology* 23, 5318–5330. doi:10.1111/gcb.13797

Langård, L., Fatnes, O., Johannessen, A., Skaret, G., Axelsen, B., Nøttestad, L., et al. (2014). State-dependent spatial and intra-school dynamics in pre-spawning herring Clupea harengus in a semi-enclosed ecosystem. *Mar. Ecol. Prog. Ser.* 501, 251–263. doi:10.3354/meps10718

Lee, H., Aydin, N., Choi, Y., Lekhavat, S., and Irani, Z. (2018). A decision support system for vessel speed decision in maritime logistics using weather archive big data. *Computers & Operations Research* 98, 330–342. doi:10.1016/j.cor.2017.06.005

Mandel, J. (2006). Efficient implementation of the ensemble kalman filter. *Center for Computational Mathematics Reports*

Maunder, M. N. and Punt, A. E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries Research* 70, 141–159. doi:10.1016/j.fishres.2004.08.002

Maunder, M. N., Sibert, J. R., Fonteneau, A., Hampton, J., Kleiber, P., and Harley, S. J. (2006). Interpreting catch per unit effort data to assess the status of individual stocks and communities. *ICES Journal of Marine Science* 63, 1373–1385. doi:10.1016/j.icesjms.2006.05.008

May Petry, L., Soares, A., Bogorny, V., Brandoli, B., and Matwin, S. (2020). Challenges in vessel behavior and anomaly detection: From classical machine learning to deep learning. In *Advances in Artificial Intelligence*, eds. C. Goutte and X. Zhu (Cham: Springer International Publishing), 401–407

McLane, A. J., Semeniuk, C., McDermid, G. J., and Marceau, D. J. (2011). The role of agent-based models in wildlife ecology and management. *Ecological Modelling* 222, 1544–1556. doi:https://doi.org/10.1016/j.ecolmodel.2011.01.020

Moral, P. D. and Doucet, A. (2014). Particle methods: An introduction with applications. *ESAIM: Proceedings* 44, 1–46. doi:10.1051/proc/201444001

Nakken, O. (2008). *Norwegian spring-spawning herring & Northeast Arctic cod: 100 years of research management* (Trondheim: Tapir Academic Press). OCLC: ocn262434637

Natale, F., Gibin, M., Alessandrini, A., Vespe, M., and Paulrud, A. (2015). Mapping Fishing Effort through AIS Data. *PLOS ONE* 10, e0130746. doi: 10.1371/journal.pone.0130746

Niu, S., Luo, Y., Dietze, M. C., Keenan, T. F., Shi, Z., Li, J., et al. (2014). The role of data assimilation in predictive ecology. *Ecosphere* 5, art65. doi:10.1890/ES13-00273.1

Nøttestad, L. (1998). Extensive gas bubble release in Norwegian spring-spawning herring (Clupea harengus) during predator avoidance. *ICES Journal of Marine Science* 55, 1133–1140. doi:10.1006/jmsc.1998.0416

Parker, R. W. R., Blanchard, J. L., Gardner, C., Green, B. S., Hartmann, K., Tyedmers, P. H., et al. (2018). Fuel use and greenhouse gas emissions of world fisheries. *Nature Climate Change* 8, 333–337. doi:10.1038/s41558-018-0117-x

Parker, R. W. R. and Tyedmers, P. H. (2015). Fuel consumption of global fishing fleets: current understanding and knowledge gaps. *Fish and Fisheries* 16, 684–696. doi:10.1111/faf.12087

Pelletier, N., André, J., Charef, A., Damalas, D., Green, B., Parker, R., et al. (2014). Energy prices and seafood security. *Global Environmental Change* 24, 30–41. doi:10.1016/j.gloenvcha.2013.11.014

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. doi:10.1093/icesjms/fsw062. _eprint: https://academic.oup.com/icesjms/article-pdf/73/9/2302/31231202/fsw062.pdf

Perry, A. L., Low, P. J., Ellis, J. R., and Reynolds, J. D. (2005). Climate Change and Distribution Shifts in Marine Fishes. *Science* 308, 1912–1915. doi:10.1126/science.1111322

Perry, G. L. and Enright, N. J. (2006). Spatial modelling of vegetation change in dynamic landscapes: a review of methods and applications. *Progress in Physical Geography: Earth and Environment* 30, 47–72. doi:10.1191/0309133306pp469ra

Pe'er, G. and Kramer-Schadt, S. (2008). Incorporating the perceptual range of animals into connectivity models. *Ecological Modelling* 213, 73–85. doi:https://doi.org/10.1016/j.ecolmodel.2007.11.020

Pinsky, M. L., Eikeset, A. M., McCauley, D. J., Payne, J. L., and Sunday, J. M. (2019). Greater vulnerability to warming of marine versus terrestrial ectotherms. *Nature* 569, 108–111. doi:10.1038/s41586-019-1132-4

Pinsky, M. L., Selden, R. L., and Kitchel, Z. J. (2020). Climate-Driven Shifts in Marine Species Ranges: Scaling from Organisms to Communities. *Annual Review of Marine Science* 12, 153–179. doi:10.1146/annurev-marine-010419-010916

Politikos, D. V., Huret, M., and Petitgas, P. (2015). A coupled movement and bioenergetics model to explore the spawning migration of anchovy in the Bay of Biscay. *Ecological Modelling* 313, 212–222. doi:10.1016/j.ecolmodel.2015.06.036

Pörtner, H.-O., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegría, A., et al. (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* (Cambridge University Press)

Railsback, S. F. (2001). Concepts from complex adaptive systems as a framework for individual-based modelling. *Ecological Modelling* 139, 47–62. doi:https://doi.org/10.1016/S0304-3800(01)00228-9

Reed, M. and Balchen, J. G. (1982). A Multidimensional Continuum Model of Fish Population Dynamics and Behaviour: Application to the Barents Sea Capelin (Mallotus Villosus). *Modeling, Identification and Control: A Norwegian Research Bulletin* 3, 65–109. doi:10.4173/mic.1982.2.1

Reite, K.-J., Fernandes, J. A., Uriondo, Z., and Quincoces, I. (2021a). The Potential of Big Data for Improving Pelagic Fisheries Sustainability. In *Big Data in Bioeconomy*, eds. C. Södergård, T. Mildorf, E. Habyarimana, A. J. Berre, J. A. Fernandes, and C. Zinke-Wehlmann (Cham: Springer International Publishing). 371–376. doi:10.1007/978-3-030-71069-9_28

Reite, K.-J., Haugen, J., Michelsen, F., and Aarsæther, K. (2021b). Sustainable and added value small pelagics fisheries pilots. *Big Data in Bioeconomy: Results from the European DataBio Project* , 389

Röckmann, C., van Leeuwen, J., Goldsborough, D., Kraan, M., and Piet, G. (2015). The interaction triangle as a tool for understanding stakeholder interactions in marine ecosystem based management. *Marine Policy* 52, 155–162. doi:10.1016/j.marpol.2014.10.019

Sakov, P. and Oke, P. R. (2008). A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. *Tellus A: Dynamic Meteorology and Oceanography* 60, 361. doi:10.1111/j.1600-0870.2007.00299.x

Sala, R., Pezzotta, G., Pirola, F., and Huang, G. Q. (2019). Decision-Support System-based Service Delivery in the Product-Service System Context: Literature Review and Gap Analysis. *Procedia CIRP* 83, 126–131. doi:10.1016/j.procir.2019.03.140

Schau, E. M., Ellingsen, H., Endal, A., and Aanondsen, S. A. (2009). Energy consumption in the Norwegian fisheries. *Journal of Cleaner Production* 17, 325–334. doi:10.1016/j.jclepro.2008.08.015

Slagstad, D. and McClimans, T. A. (2005). Modeling the ecosystem dynamics of the barents sea including the marginal ice zone: I. physical and chemical oceanography. *Journal of Marine Systems* 58, 1–18. doi:https://doi.org/10.1016/j.jmarsys.2005.05.005

Slotte, A. and Fiksen, Ø. (2000). State-dependent spawning migration in Norwegian spring-spawning herring. *Journal of Fish Biology* 56, 138–162. doi:10.1111/j.1095-8649.2000.tb02091.x

Slotte, A., Salthaug, A., Utne, K., Ona, E., Vatnehol, S., and Pena, H. (2016). *Distribution and abundance of Norwegian spring spawning herring during the spawning season in 2016*. Tech. rep., Institute of Marine Research (IMR), Bergen, Norway

Sousa, T., Domingos, T., Poggiale, J.-C., and Kooijman, S. A. L. M. (2010). Dynamic energy budget theory restores coherence in biology. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 3413–3428. doi:10.1098/rstb.2010.0166

Squires, D. (1994). Sources of growth in marine fishing industries. *Marine Policy* 18, 5–18. doi:10.1016/0308-597X(94)90085-X

Standal, D. and Asche, F. (2018). Hesitant reforms: The Norwegian approach towards ITQ's. *Marine Policy* 88, 58–63. doi:10.1016/j.marpol.2017.11.005

Stephenson, R. L., Paul, S., Pastoors, M. A., Kraan, M., Holm, P., Wiber, M., et al. (2016). Integrating fishers' knowledge research in science and management. *ICES Journal of Marine Science* 73, 1459–1465. doi:10.1093/icesjms/fsw025

Stillman, R. A., Railsback, S. F., Giske, J., Berger, U., and Grimm, V. (2015). Making Predictions in a Changing World: The Benefits of Individual-Based Ecology. *BioScience* 65, 140–150. doi:10.1093/biosci/biu192

Stoeva, G., Tsiouvalas, A., Humpert, M., Raspotnik, A., Mordal, M. H., Fleener, C., et al. (2022). *Blue Fisheries & Aquaculture. Fisheries and Aquaculture in Alaska and North Norway*. Tech. rep., Nord University

Strogatz, S. H. (2001). Exploring complex networks. *Nature* 410, 268–276. doi:10.1038/35065725

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., and Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digital Medicine* 3, 17. doi:10.1038/s41746-020-0221-y

Swaminathan, M. S. (2009). Norman E. Borlaug (1914–2009). *Nature* 461, 894–894. doi:10.1038/461894a

Tamario, C., Sunde, J., Petersson, E., Tibblin, P., and Forsman, A. (2019). Ecological and Evolutionary Consequences of Environmental Change and Management Actions for Migrating Fish. *Front. Ecol. Evol.* 7, 271. doi:10.3389/fevo.2019.00271

Totland, A. and Johnsen, E. (2022). Kayak Drone – a silent acoustic unmanned surface vehicle for marine research. *Frontiers in Marine Science* 9, 986752. doi:10.3389/fmars.2022.986752

Touzeau, S., Lindroos, M., Kaitala, V., and Ylikarjula, J. (2000). Economic and biological risk analysis of the norwegian spring-spawning herring fishery. *Annals of Operations Research* 94, 197–217. doi:10.1023/A:1018973217951

Truong, T., Rothschild, B., and Azadivar, F. (2005). Decision Support System for Fisheries Management. In *Proceedings of the Winter Simulation Conference, 2005.* (Orlando, FL. USA: IEEE), 2107–2111. doi:10.1109/WSC.2005.1574494

Tu, C.-Y., Tseng, Y.-H., Chiu, T.-S., Shen, M.-L., and Hsieh, C.-H. (2012). Using coupled fish behavior-hydrodynamic model to investigate spawning migration of Japanese anchovy, Engraulis japonicus, from the East China Sea to Taiwan: Spawning migration model of Japanese anchovy. *Fisheries Oceanography* 21, 255–268. doi:10.1111/j.1365-2419.2012.00619.x

Varpe, Ø., Daase, M., and Kristiansen, T. (2015). A fish-eye view on the new Arctic lightscape. *ICES Journal of Marine Science* 72, 2532–2538. doi:10.1093/icesjms/fsv129

Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W. L., Edgar, G. J., Mouillot, D., et al. (2022). A quantitative review of abundance-based species distribution models. *Ecography* 2022, ecog.05694. doi:10.1111/ecog.05694

Walker, E. and Bez, N. (2010). A pioneer validation of a state-space model of vessel trajectories (VMS) with observers' data. *Ecological Modelling* 221, 2008–2017. doi:10.1016/j.ecolmodel.2010.05.007

Ward, J. A., Evans, A. J., and Malleson, N. S. (2016). Dynamic calibration of agent-based models using data assimilation. *R. Soc. open sci.* 3, 150703. doi:10.1098/rsos.150703

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442. doi:10.1038/30918

Wiegand, T., Jeltsch, F., Hanski, I., and Grimm, V. (2003). Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. *Oikos* 100, 209–222. doi:10.1034/j.1600-0706.2003.12027.x

Williams, R. J. and Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature* 404, 180–183. doi:10.1038/35004572

Ziegler, F., Winther, U., Hognes, E. S., Emanuelsson, A., Sund, V., and Ellingsen, H. (2013). The Carbon Footprint of Norwegian Seafood Products on the Global Seafood Market. *Journal of Industrial Ecology* 17, 103–116. doi:10.1111/j.1530-9290.2012.00485.x

Ølmheim, O. (2021). *Economic and Biological Figures from Norwegian fisheries – 2020*. Tech. rep., Fishery Directorate, Norway

ISBN 978-82-326-6763-5 (trykt utg.)
ISBN 978-82-326-6907-3 (elektr. utg.)
ISSN 1503-8181 (trykt utg.)
ISSN 2703-8084 (online ver.)

NTNU
Kunnskap for en bedre verden