# Accounting for Movement in Spatial Surplus Production Models and Case Studies of Redfish (*Sebastes spp.*, Sebastidae) and Yellowtail Flounder (*Limanda ferruginea*) on the Eastern Grand Banks of Newfoundland

by

© Hoang Nguyen

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science in Fisheries Science (Stock Assessment)

Fisheries and Marine Institute

School of Fisheries

Memorial University of Newfoundland

St. John's Newfoundland and Labrador, May 2025

# Abstract

This thesis explores the integration of spatial modelling and surplus production models (SPMs) for fisheries stock assessment. Typically, SPMs disregard the spatial dynamics of populations. To address this limitation, we propose a novel approach that utilizes the Gaussian random field to capture spatial heterogeneity. This method enhances spatial representation without explicitly parameterizing movement dynamics, offering a more robust framework. The methodology (i.e., the random field model) builds upon existing surplus production models by adapting a triangular grid and employing stochastic process errors to capture spatial variation. Simulations and case studies demonstrate the model's effectiveness in estimating stock biomass dynamics, outperforming non-spatial and movement models. The random field model offers a simplified but robust alternative to the explicit spatial movement model. Applied to the 3LN Redfish stock, the random field model highlights significant spatial heterogeneity and a decline in biomass between 2012-2019. Furthermore, the approach was extended to Yellowtail Flounder in 3LNO Divisions, demonstrating stable biomass distributions with spatial preferences for shallower waters. The findings underscore the importance of spatially explicit models in fisheries stock assessment when sufficient spatial data are available. This study contributes to advancing fisheries stock assessment by providing a scalable and adaptable framework for spatial stock assessment.

# Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Dr. Nan Zheng and Dr. Jin Gao. Dr. Nan Zheng provided exceptional guidance, unwavering support, and invaluable assistance throughout my project. Dr. Jin Gao has been a mentor since the very beginning of my journey, and I am truly grateful for her dedication and encouragement. I could not have reached this point without their expertise and support, for which words cannot fully convey my appreciation.

I extend my heartfelt thanks to Dr. Noel Cadigan for his insightful guidance and comprehensive lectures, which greatly enriched my understanding throughout my study.

To my family and friends: your encouragement, assistance, and belief in me have been my foundation. I would not have made it this far without you, and I am endlessly grateful for your presence in my life.

Finally, I wish to acknowledge the generous funding that made this research possible. Support was provided by the Ocean Frontier Institute through an award from the Canada First Research Excellence Fund. Additional funding was generously provided by the Ocean Frontier Institute and the NSERC Discovery Grant to Dr. Jin Gao [RGPIN-2021–03249], as well as another NSERC Discovery Grant to Nan Zheng [RGPIN-2024-04746].

To all who have supported me on this journey, my sincerest thanks and appreciation.

# **Table of Contents**

Abstra	actii
Ackno	owledgementsiii
List of	f Figuresvi
List of	f Tablesxi
Chapt	er 1. Introduction1
1.	A brief introduction to stock assessment methods1
2.	Why spatial models are important in stock assessment
3.	The benefits of spatial models in stock assessment Error! Bookmark not defined.
4.	Aspects of considerations in spatial stock assessments
5.	Surplus production models
6.	Research motivation and objectives
Chapt	er 2. Accounting for Movement in Spatial Surplus Production Models: A Case Study of
Redfis	sh on the Eastern Grand Banks of Newfoundland12
1.	Introduction
2.	Materials and methods15
3.	Results
4.	Discussion

Grand	Banks of Newfoundland	.56
1.	Introduction	56
2.	Materials and methods	59
3.	Results	63
4.	Discussion	68
Chapt	er 4. Conclusion and future research	.76
1.	Performance of the random field model	76
2.	Insights from case studies and broader applicability	77
3.	Considerations for Application	78
4.	Relevance to spatial management	78
5.	Challenges and future directions	79
Refere	ence	.80
Appen	ıdix1	100

Chapter 3. Applying random field model on Yellowtail Flounder Stock Assessment in the Eastern

# List of Figures

Figure 1. Recorded landings of Redfish in 3LN from 1959 to 2019 with the corresponding total
allowable catch (TAC). Dots represent TAC and solid line represents commercial catches29
Figure 2. Maps of the spatial survey catches in log scale of Redfish biomass (Kg) in the 3LN
Divisions from 2012 to 2019. Bubbles represent survey catches and crosses show survey
locations with zero catch
Figure 3. Maps of commercial fishing effort in log scale of fishing hour by trawlers in the 3LN
Divisions from 2012 to 2019. Bubbles represent commercial fishing data
Figure 4. Simulation results for (A) random field, (B) movement and (C) non-spatial models
including parameter estimator distributions of (1) the average density $\alpha$ , (2) the density
dependence $\beta$ , (3) the standard deviation of spatial variation in productivity $\sigma\omega$ , and (4) the
standard deviation of spatiotemporal process error $\sigma \epsilon$ . The simulation used Tweedie power
parameter $p = 1.5$ . Red lines are the true values of the parameters and blue lines are means of
parameter estimates
Figure 5. Simulation results of correlation between true abundance (i.e., simulated) and estimated
abundance for (1) random field, (2) movement and (3) non-spatial models. The simulation used
Tweedie power parameter $p = 1.5$ . Blue lines are means of correlation estimates
Figure 6. Boxplots of differences in statistical metrics including AIC, BIC, and correlation
between random field and movement models. The simulation used Tweedie power parameter $p =$
1.5. The differences are: AICdifference = AICrandom field – AICmovement;

BICdifference = BICrandom field - BICmovement; and Correlationdifference =	
Correlationrandom field – Correlationmovement	.36

Figure 10. Simulation results of correlation between true abundance (i.e., simulated) and estimated abundance for (1) random field model, (2) movement model and (3) non-spatial model methods. The simulation used Tweedie power parameter p = 1.7. Blue lines are means of between random field and movement models. The simulation used Tweedie power parameter p =

1.7. The differences are: *AICdifference* = *AICrandom field* - *AICmovement*;

Figure 12. Boxplots of differences in statistical metrics including AIC, BIC, and correlation between random field and non-spatial models. The simulation used Tweedie power parameter p = 1.7. The differences are: *AICdifference = AICrandom field - AICnon - spatial*; *BICdifference = BICrandom field - BICnon - spatial*; and *Correlationdifference = Correlationrandom field - Correlationnon - spatial*......40

Figure 15. Model predicted spatial survey indices (i.e., the estimated spatial distribution) of
Redfish in the 3LN Divisions, shown on a logarithmic scale as log (1 + survey index). The color
legend is presented at the bottom of the plot43
Figure 16. Plots of (A) estimated versus observed survey catches for each year and (B) QQ Plots
of the randomized quantile residuals for survey catches. In panel (A), the x axis represents the
estimated survey catches, the y axis represents the observed survey catches, and the red line has a
slope of one through the origin. The bias correction feature in TMB was applied to predict survey
catches
Figure 17. The 3LN Redfish plots of A) observed survey data; B) observed and predicted
commercial landings; C) estimates of stock biomass using ASPIC software; D) estimates of stock
biomass using the random field model. The shading areas in panel B and D show 95% confidence
intervals47
Figure 18. Plots of (A) recorded landings of Yellowtail Flounder (ton) in the 3LNO Divisions;
and (B) total catches of observed survey data recorded for Yellowtail Flounder (ton) in the 3LNO
Divisions
Figure 19. Spatial survey catches in log scale of Yellowtail Flounder (Kg) in the 3LNO Divisions
from 2012 to 2019. Bubbles represent survey catches, and crosses show locations with zero
catch
Figure 20. Spatial commercial fishing effort in log scale of fishing hour by trawlers in the 3LN
Divisions from 2012 to 2019. Bubbles represent commercial fishing data

Figure 21. Maps of population biomass estimates for Yellowtail Flounder in 3LNO divisions,
displayed on a logarithmic scale as log (1 + biomass estimate). The biomass was estimated in Kg.
The color legend is presented at the bottom of the plot
Figure 22. Observed spatial survey catches of Yellowtail Flounder in 3LNO Divisions, shown on
a logarithmic scale as log (1 + survey index). The color legend is presented at the bottom of the
plot
Figure 23. Plot of (A) estimated over observed survey catches for each year and (B) QQ Plots of
the randomized quantile residuals for survey catches. In panel (A), the x axis represents the
estimated survey catches, and the y axis represents the observed survey catches with a 45-degree
line of equality (the red line). The bias correction feature in TMB was applied to evaluate the
estimated survey catches (Thorson and Kristensen, 2016)67
Figure 24. Plot showing (A) observed and predicted commercial catches estimated by the random
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023))
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023))
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023))
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023))
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023))
field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023))

# List of Tables

Table 1. Parameter estimates and standard errors for the random field model fit to the 3LN Redf	ĩsh
data	.43
Table 2. AIC and BIC estimates for using quadratic and not using quadratic dispersion	
relationship approaches	.63
Table 3. Parameter estimates and standard errors when fitting the model to the 3LNO Yellowta	il
Flounder dataset	.68

# **Chapter 1. Introduction**

## 1. A brief introduction to stock assessment methods

In science-based fisheries management, fisheries science provides decision-makers with recommendations on the alternative management strategies derived from scientific studies. Most recommendations are based on quantitative analyses and predictions on how fish stocks respond to alternating management objectives. The process of data collection, investigation, and presentation addressing these objectives is termed "stock assessment" – the backbone of sustainable fisheries management (see Punt and Hilborn, 1997, King, 2007, Haddon, 2011).

The history of fisheries modeling begin with the foundation of Baranov's catch equation (Baranov, 1918), and Russell (1931) introduced the simplest concept of describing the dynamics of fish stock considering four major forces on changes in the stock biomass  $B_t$ . The four forces consist of the growth of individuals  $G_t$ , population recruitments  $R_t$ , the removals due to fishing  $C_t$ , and other reasons  $D_t$  (e.g., discards) with t representing time usually measured in yearly steps:

$$B_{t+1} = B_t + (G_t + R_t) - (C_t + D_t).$$

There are two main types of stock assessment frameworks regarding model structures: 1) the age/size-structured and 2) the age/size-aggregated models (i.e., surplus production framework). The former describes the yearly dynamics for different age/size groups (e.g., Cadigan, 2016, Perreault et al., 2020a, Kumar et al., 2020), while the latter aggregates age/size information using a single production function to describe stock dynamics (e.g., Schaefer, 1954, Beverton and Holt, 1957, Pella and Tomlinson, 1969, Fox, 1970).

The early modeling practices in the stock assessment were deterministic with the assumption of negligible errors in observations. Thus, these deterministic (i.e., non-state-space) methods did not consider any measurement of the uncertainty (i.e., errors). In consideration of parameter estimation in deterministic SPMs, there have been multiple techniques employed, such as ordinary least squares (Uhler, 1980, Jensen, 1984, Prager, 1994), maximum likelihood (Gould and Pollock, 1997, Prager, 2016), and Bayesian inference (Hoenig et al., 1994, Walters and Ludwig, 1994).

Practically, noisy fisheries data are not uncommon. There are two sources of noise in any timeseries data of population dynamics termed process and observation errors. As the name suggested, the observation error results from variation in sampling (i.e., observing), while the process error is typically considered variation in the true population processes (Ahrestani et al., 2013).

The state-space framework, originally introduced by Kalman (1960) and Kalman and Bucy (1961) for engineering purposes, was adopted to fisheries modeling (see Sullivan, 1992, Gudmundsson, 1994) to capture the uncertainty. In principle, state-space fisheries models are built upon the same equations as the previous fisheries modeling framework; however, they include the stochasticity of unobserved errors. In other words, state-space models incorporate process and observation errors.

Two components define a state-space stock assessment model: 1) the process; and 2) the observation models. The process model is referred to how unobserved states of the system under study (e.g., stock dynamics) depend upon preceding states. The observation model describes how observed states (e.g., commercial catch, survey indices) depend on unobserved states (Aeberhard et al., 2018). An estimation algorithm is applied to compute parameter estimates and other statistical components of interest (e.g., residuals). It is assumed that each data source is independent of the others. Parameter estimation methods consist of maximum likelihood estimation (i.e., frequentist approach) and Bayesian approach. It is critical to determine an appropriate likelihood

function (i.e., objective function) for each information component in the analysis. In the frequentist approach, the objective function is based on the joint likelihood function (i.e., the product of component likelihoods – in practice, often using the sum of negative log-likelihoods). Modelers then use an optimizer [e.g., nlminb (R-Core-Team, 2022)] to minimize the negative log-likelihood objective function for parameter estimation. From the Bayesian perspective, the objective function is based on prior distributions.

## 2. Spatial models are important in stock assessment

Since the dawn of animal ecology, the spatial structure has been the fundamental characteristic of organisms' distribution (Elton, 1927). The simplifying assumptions of traditional non-spatial models possibly leads to oversimplified assessments and management advice. These conventional models typically rely on the "unit stock" assumption, treating a fish population as a single, homogenous entity where biological processes and fishing impacts are uniform across its entire range (Cadrin et al., 2004, Kerr et al., 2017). However, this assumption is frequently violated. Fish populations exhibit significant spatial structure driven by habitat preferences, environmental gradients, and critical life history functions (e.g., spawning and nursery areas), leading to varied distributions rather than uniform mixing (Perry and Smith, 1994, Ciannelli et al., 2007). Furthermore, biological parameters like growth and mortality can vary substantially across different regions occupied by the stock (Lorenzen, 2016), and fishing fleets rarely distribute effort randomly, instead targeting specific locations and creating highly uneven fishing mortality patterns (Walters et al., 1998, Branch et al., 2006, Bastardie et al., 2010).

Ignoring this underlying heterogeneity, as non-spatial models do, can lead to significant issues. For example, it can result in biased perceptions of stock status and productivity, potentially masking serious issues like localized depletion even when the overall stock assessment appears positive (Berkeley et al., 2004, Frank et al., 2006). Non-spatial models, which average fishing mortality, fail to capture this dynamic and can be susceptible to the "hyperstability" problem, where catchper-unit-effort (CPUE) indices remain high by targeting remaining aggregations, thus masking true declines in overall abundance until a collapse is potentially imminent (Hilborn and Walters, 1992a). Furthermore, the dynamic nature of fish populations, involving movement through migration, larval dispersal, and ontogenetic shifts, creates complex patterns of connectivity between different areas, influencing recruitment and population resilience in ways that non-spatial models cannot adequately represent (Cowen et al., 2006, Fogarty and Botsford, 2007).

Adopting spatial stock assessment models offers clear advantages by directly addressing these limitations. Primarily, they provide a more realistic representation of fish stock and fishery dynamics by explicitly incorporating spatial heterogeneity in population distribution, movement, biological rates, and fishing pressure (Goethel et al., 2011). This enhanced realism can translate into improved and potentially less biased estimates of model outputs, such as overall biomass, abundance trends, and fishing mortality reference points. Crucially, beyond single stock-wide estimates, spatial models can provide spatially explicit outputs, mapping variations in fish density, productivity, and fishing impact across different areas, offering a much richer understanding of the system's status.

A practical benefit of spatial models lies in their ability to directly inform and evaluate spatial management strategies. As fisheries management increasingly employs spatial tools such as Marine Protected Areas (MPAs), time/area closures, and spatially zoned effort regulations. Spatial assessment models provide the necessary framework to simulate the ecological and socio-economic consequences of these actions that non-spatial models simply cannot do (Pelletier and

Mahévas, 2005, Goethel et al., 2011, Costello et al., 2012). By tracking stock status and fishing pressure at finer spatial scales, these models can detect localized depletion, providing early warnings (Walters, 2003), leading to more nuanced, spatially explicit management advice tailored to local conditions (Apostolaki et al., 2002).

Furthermore, spatial models excel at integrating diverse, spatially referenced data streams that are often underutilized or aggregated in traditional assessments. Data from VMS, detailed logbooks, spatially stratified surveys, tagging experiments, habitat mapping, and environmental monitoring can be synergistically combined within a single analytical framework, leading to more robust and comprehensive assessments (Hintzen et al., 2012, Thorson et al., 2015a, Thorson, 2019). Spatial models provide a powerful platform for investigating the impacts of large-scale environmental changes, particularly climate change, which often manifest spatially through shifts in species distributions, changes in habitat suitability, and altered productivity gradients across a stock's range (Cheung et al., 2010, Hollowed et al., 2013). Ultimately, there is a strong desire and practical need to develop and utilize assessment tools capable of fully integrating spatial information to produce more realistic and reliable analyses (Thorson, 2019, Maunder and Punt, 2013).

## 3. Aspects of considerations in spatial stock assessments

#### Data availability

Explicitly modeling spatial stock assessments requires high-resolution spatial data. This typically includes commercial catch data, survey data, and/or tagging data, all of which need to include precise fishing locations (i.e., longitude and latitude) and span multiple years. However, in real-world fisheries management, such spatial data are not always available. This can be due to various factors, such as management systems in certain fisheries that do not require spatial information or

do not have sufficient capacity to collect spatial data over years, or limited budgets that constrain fisheries research efforts.

#### Spatial stock structure assumptions

In principle, it is essential to define target species, biological and spatial stock structure, as well as management units in implementing stock assessment and fisheries management (Cadrin and Secor, 2009, Cadrin, 2020). However, in practice, the situation can sometimes be messy, for example, Redfish fisheries in the Northwest Atlantic. Their assessment and management are often based upon combined-species data [i.e., *S. mentella* and *S. fasciatus* (Gascon, 2003)]; their biological stock structures are very complex [e.g., compelling evidence of overlapping distribution (Gascon, 2003, Grant, 2004), hybridization, and introgression (Roques et al., 2001)]; there is evidence of mismatching between management and biological units [e.g., movement crossing management units (Gascon, 2003, Benestan et al., 2021)].

#### Selection of model structure

Modeling fish stock dynamics often begins with choosing either age-structured or age-aggregated (i.e., surplus production) frameworks. Punt (2019) described a general form of the spatially age-structured model as follows:

$$N_{y,a}^{s,A} = \begin{cases} R_y^{s,A} \\ \sum_{s'} \sum_{A'} X_{y-1,a-1}^{s',s,A} N_{y-1,a-1}^{s',A} e^{-Z_{y-1,a-1}^{s',A'}} & \text{if } a = 0 \text{ otherwise} \end{cases}$$

Where  $N_{y,a}^{s,A}$  represents the abundance of fish of age *a* and stock *s* in area *A* at year *y*;  $R_y^{s,A}$  denotes the recruitment (often assumed at age 0) to stock *s* and area *A* at year *y*;  $Z_{y-1,a-1}^{s',A'}$  is the total mortality of age *a* and stock *s* in area *A* in year *y*; and  $X_{y-1,a-1}^{s's,A}$  is the proportion of fish of age *a* and stock *s*' in area *A*' that move to stock *s* and area *A* in year *y* (after mortality in year *y*).

A general form of the age-aggregated framework can be referred to Thorson et al. (2017):

$$B_{t+1} = g[m(B_t e^{(-u_t F_t)})] e^{\varepsilon_t}$$

Where  $F_t$  is a function of fishing mortality in year t,  $B_t e^{(-u_t F_t)}$  describes the product of survival rate  $e^{(-u_t F_t)}$  and density  $B_t$  in year t; m represents a function describing the movement of the net effect of advective and diffusive movements; g is a function approximating the local density in the stock dynamics. The error term is defined by  $e^{\varepsilon_t}$ .

#### Spatial estimation methods

There are different ways to discretise space and approximate movement. The choice of tools for space discretization has varied in the literature, such as using the sf package to generate a set of polygons in the Northern shrimp assessment (Pedersen et al., 2022), the R-INLA package to create a grid of triangles in a study on the big skate (Thorson et al., 2017), snow crab (Olmos et al., 2023), and scallop (McDonald et al., 2021), as well as using a grid of squared cells for Antarctic toothfish (Mormede et al., 2017), and snow crab (Cadigan et al., 2017). The proportion of fish density moving from one spatial cell to another at annual time steps can be computed using a Euler approximation (Thorson et al., 2017) or using probability density functions based on environmental variables and the distance from their previous location (Mormede et al., 2017). Instead of explicitly modelling movement rates, spatial variation parameters can be approximated using a Markov Random Field smoother. For example, Pedersen et al. (2022) modeled Northern Shrimp movement using an explicit advection-diffusion framework, incorporating parameters for directed movement (advection velocities) and random dispersal (diffusion coefficient). A key assumption in their

implementation, common in many such models, is that these movement parameters are constant across space and time.

## 4. Surplus production models

Surplus production models (SPMs) were introduced by Graham (1935) and later developed by multiple quantitative scientists (e.g., Schaefer, 1954, Beverton and Holt, 1957, Pella and Tomlinson, 1969, Fox, 1970). The underlying concept is to use a single production function to describe stock dynamics disregarding the age and length compositions of the population. Accordingly, SPMs are often thought of their simplicities in analytical study; and applicable with only commercial catch and survey index information (Jensen, 1984, Yoshimoto and Clarke, 1993). Thus, SPMs are usually the foundation of developing more complex models (Thorson et al., 2017).

There are three density-independent SPMs commonly studied in fisheries science, including Schaefer (i.e., logistic) (Schaefer, 1954), Pella and Tomlinson (i.e., generalized production) (Pella and Tomlinson, 1969, Fletcher, 1978), and Fox models (Fox, 1970). Of the three, Schaefer SPM is thought of as the most frequently used form (Zhang, 2013, Karim et al., 2020). The model describes the growth of the population size based on a logistic production function, where r represents the stock's intrinsic growth rate, and K represents the carrying capacity as follows:

$$\frac{dB_t}{dt} = rB_t - \frac{r}{K}B_t^2.$$

A study by Pella and Tomlinson (1969), and later restructured by Fletcher (1978), added a shape parameter to the production function governing the skewness of the surplus production curve:

$$\frac{dB_t}{dt} = \gamma \mu \frac{B_t}{K} - \gamma \mu \left(\frac{B_t}{K}\right)^n.$$

Where  $\mu$  is maximum sustainable yield, and  $\gamma$  is a function of n (i.e.,  $\gamma = n^{n/(n-1)}/(n-1)$ ). In the case of Schaefer SPM, n = 2. This model is widely recognized as the generalized form of SPMs (i.e., Pella and Tomlinson SPM) and thought of sometimes more adaptable to the fishing realities (Pella and Tomlinson, 1969, Quinn and Deriso, 1999). However, some studies found that estimating the shape parameter is challenging (Fletcher, 1978, Rivard and Bledsoei, 1978, Hilborn and Walters, 1992a), furthermore the generalized SPM is sensitive to outliers; and thus more likely to generate less precise estimates than the Schaefer model (Prager, 2002, Panhwar et al., 2012).

A study by Fox (1970) presented a modified SPM using the Gompertz function to model the production function, which led to an asymmetrical production curve as follows:

$$\frac{dB_t}{dt} = rB_t(\log(K) - \log(B_t)).$$

In an example on lobster fisheries, it was documented that the Fox model provided more realistic estimates of biological parameters than the Schaefer model (Yoshimoto and Clarke, 1993).

Besides the density-independent models, attention has also been drawn to density-dependent models. Although density-dependent analysis is a very attractive topic in quantitative ecology, spatial modeling in fisheries science has not been thoroughly studied due to the complexity of spatial models. There were studies of applying spatial SPMs to rockfish data (*Sebastes* spp.) (Thorson et al., 2015b) and big skate (*Raja binoculata*) (Thorson et al., 2017). The two examples were developed from the production function derived from the Gompertz equation documented in (Dennis and Taper, 1994):

$$\frac{dB_t}{dt} = B_t \big( \propto + \beta \log(B_t) \big).$$

Where  $\propto$  governs the population density and  $\beta$  represents the degree of the density dependence.

There were also attempts to extend the usage of SPMs, such as incorporating environmental variables into the applications of SPMs (Jensen, 2005, Thiaw et al., 2009, Wang et al., 2018), developing SPMs for multi-species assessments (Mueter and Megrey, 2006, Johnson and Cox, 2021), testing SPMs in the continuous-time functions (Pedersen and Berg, 2017), and using Bayesian statistics for SPMs (Woods and Jonasson, 2017). More importantly, SPMs have been widely applied in formal stock assessments (Ávila de Melo et al., 2020, ABARES, 2021). Pauly et al. (1989) reviewed and stated that SPMs are more realistically applicable to managing tropical fisheries in developing countries where age and length data might not be generally available.

### 5. Research motivation and objectives

Given the broader need to incorporate spatial modeling into stock assessments, along with our observations of the challenges in the current movement model proposed by Thorson et al. (2017), we are strongly motivated to develop a new method to overcome the challenges in computing spatial heterogeneity. The simplicity of SPM further motivated its use, as it provides an ideal framework for building more advanced spatial stock assessment models, where spatial modeling often presents significant complexity.

The primary objective of this study is to develop a methodology — specifically, the random field model — to effectively account for spatial heterogeneity in the assessment of fish population dynamics using a scalable triangular grid on a defined spatial domain. In Chapter Two, I provide a comprehensive description of the model's development and evaluate its performance through

rigorous simulation testing. This chapter also includes an application of the random field model to the 3LN Redfish dataset, demonstrating its practical utility. In Chapter Three, I apply the analysis by refining the understanding of the relationship between expected and observed survey catches through the integration of a quadratic function within the random field framework. Furthermore, I compare the model's output with predicted catches and stock biomass estimates reported in the 2023 Assessment of Yellowtail Flounder in NAFO Divisions 3LNO. Finally, I present a final thoughts and conclusion chapter (i.e., Chapter Four) to summarize the work and my thoughts on the limitations, potential applicability, and considerations for future developments of the proposed methodology.

# Chapter 2. Accounting for Movement in Spatial Surplus Production Models: A Case Study of Redfish on the Eastern Grand Banks of Newfoundland

# 1. Introduction

Fisheries stock assessment relies on quantitative models to provide estimates of population size, evaluate stock status (biomass and harvest rate relative to sustainable reference points) and provide management advice. Although many advanced models exist, the choice of models largely depend on data availability specific to each fish stock. SPMs provide a simple approach to describe stock dynamics and generate benchmarks under the maximum sustainable yield (MSY) framework. These models combine net contributions from growth, recruitment, natural mortality and associated density-dependent processes into a single production function to approximate population dynamics. Punt and Szuwalski (2012) showed that ignoring transient age and size dynamics led to poor estimates of some benchmarks (i.e., E<sub>MSY</sub> and B<sub>MSY</sub>) when using empirical measures of surplus production; however, for data limited stocks where there are not enough age or length frequency data to conduct a fully age- or length-structured stock assessment, surplus production models remain widely used (e.g., Berg et al., 2021, Kimoto et al., 2023, González Herraiz et al., 2023).

Spatial stock management is becoming more important as Canada committed to designate 30% of Canadian oceans as marine protected areas by 2030 (DFO, 2024). Such area-based management tools, including Marine Protected Areas (MPAs), inherently create spatial heterogeneity in fishing mortality and potentially affect ecological processes differently inside versus outside the protected zone, challenging traditional assessment models that assume uniformity across the entire stock area

(Berger et al., 2017). Effectively evaluating the impacts of these protected areas requires spatially explicit assessment models. Beyond addressing specific spatial management tools like MPAs, these models offer broader fundamental benefits. They allow for a more detailed understanding and tracking of ecological processes such as population structure, connectivity, and spatial variations in productivity (Cadrin, 2020). Spatial models also provide a framework to explicitly evaluate and relax the often-violated assumptions of non-spatial models, like spatial homogeneity in population dynamics and fishing impacts (Berger et al., 2017, Cadrin, 2020). Furthermore, they enable assessments to align more closely with the spatial realities of both fish populations and fisheries management strategies. Consequently, detailed management advice from spatial stock assessment models can assist when evaluating the impacts on proposed protected areas. Spatial models can better accommodate survey and catch data with different spatial domains compared (e.g., inshore and offshore surveys or fisheries fleets) to space-aggregated models (Goethel et al., 2011, see Berger et al., 2017).

Spatially explicit models assign distinct quantities to different spatial locations and typically incorporate correlation structures among these locations to enhance statistical inference. Such spatially explicit approaches can track the abundance or density of the population (by age, size, and sex if possible) at points in space and time (see Cadigan et al., 2017, Thorson et al., 2017). Some stocks lack age and length information while high-resolution spatial catch and effort data exist, making them particularly suitable for spatial surplus production models (SSPMs). For example, SSPMs have been developed for a shrimp stock (*Pandalus borealis*) (Pedersen et al., 2022, Lucet and Pedersen, 2023), big skate *Raja binoculata* (Thorson et al., 2017) and Antarctic toothfish (Mormede et al., 2017). An understanding of heterogeneous population distributions has

been recognized as critical for effective fisheries management, highlighting the necessity of developing spatially explicit stock assessment models.

However, in general, explicitly incorporating spatial structure into stock assessment models remains challenging (see Cadrin and Secor, 2009, Cadrin et al., 2020), often because of insufficient spatial information regarding commercial catches and effort in addition to computational challenges. Also, movement parameter estimation has remained challenging and model convergence likely depends on both the life history characteristics of the fish species and the data quality (Punt, 2019, Punt et al., 2020).

The Eastern Grand Banks of Newfoundland (NAFO [Northwest Atlantic Fisheries Organization] Divisions 3L and 3N [3LN]) historically sustained an important commercial fishery targeting Redfish (*Sebastes spp.*, Sebastidae) but experienced a significant decline in landings in 1994 and a fishing moratorium was established in 1998. There have been signs of recovery in recent years and commercial harvesting resumed in 2011 (Rogers et al., 2022).

Redfish are long-lived and very slow-growing. As adults, Redfish are often considered relatively sedentary compared to highly migratory species, typically exhibiting strong associations with specific bathymetric features like slopes and canyons (Gascon, 2003). While large-scale annual migrations are not generally characteristic of adult Redfish, ontogenetic migrations are known to occur, with smaller/younger individuals often inhabiting shallower nursery areas before moving to deeper waters as they (Kenchington, 1991). Connectivity across broader areas likely relies more heavily on the pelagic larval stage, where dispersal is influenced by ocean currents, rather than extensive adult movement (Roques et al., 2001, Benestan et al., 2021). Therefore, while adults may exhibit some localized movement related to feeding or environmental conditions, their dynamics

within a specific area over an annual time step might be less dominated by long-distance directed migration compared to other species.

Due to the challenges in accurately determining their age, ageing data are not available. Therefore, an ASPIC (i.e., A Stock–Production Model Incorporating Covariates) logistic surplus production model (SPM) (Prager, 1994, Prager, 2016) was used to provide management advice by the NAFO during 2008 to 2020; however, this model was rejected during the 2022 assessment because of 1) mismatch between recent observed survey indices and the model biomass estimates and, 2) the use of a fixed MSY approach that resulted in a value of *r* that was considered too high for this species (Rogers et al., 2022). There has been no development of a spatial assessment model for this stock. In this study, we investigate the feasibility of incorporating movement into a SSPM using a spatial random field. Our novel methodology significantly reduces the modelling and computational complexity associated with movement while incorporating spatial heterogeneity in fish stock dynamics. It addresses both natural productivity and human harvesting activities, with a focus on the spatial distribution of the 3LN Redfish stock. We validate our approach through simulations. We then apply the model to survey and commercial catches of Redfish to evaluate the advantages of SSPMs.

# 2. Materials and methods

We begin by summarizing a typical spatial surplus production model introduced by Thorson et al. (2017), highlight its associated issues, and subsequently propose our method to address these challenges. A table of mathematical notations is included in the Appendix for reference.

A general form of an age-aggregated surplus production model for fisheries stock assessment uses a single production function to describe stock dynamics:

$$b_{t+1} = g(b_t e^{(-u_t F_t)}) e^{\varepsilon_t}, \qquad (1)$$

where population abundance (or biomass, as used below) in year t + 1, denoted  $b_{t+1}$ , is a nonlinear function g() of the product of population biomass  $b_t$  and survival rate  $e^{(-u_tF_t)}$  in year t, with the instantaneous fishing mortality rate given by the product of the fishing effort data  $F_t$  and a parameter  $u_t$ . The production function g() incorporates the effects of natural mortality, recruitment, and individual growth rates. The process error term  $e^{\varepsilon_t}$  is used to capture the uncertainty in population dynamics not explained by the production or fishing mortality. Eq. (1) assumes a sequential progression of various effects from fishing mortality to production and then to process error  $\varepsilon_t$ .

#### Spatial modeling of the surplus production process proposed by Thorson et al. (2017)

To extend the above surplus model to accommodate spatial population dynamics, the spatial domain of interest is first divided into a large set of  $n_r$  triangles. As the triangles are sufficiently small, the spatial variation within each triangle can be neglected so that every location within a triangle is homogeneous in terms of fish densities, harvest rates, population growth rates, mortality rates, process errors, etc. Each observation can be classified into one and only one triangle according to its latitude and longitude. The abundance or biomass of fish in a triangle *r* at time *t* is represented as  $b_{r,t}$ , with  $b_t$  being the vector of these abundance values for all triangles at time *t*. Thorson et al. (2017) proposed a spatial population dynamics modeling framework that integrates fish movement and migration, encompassing both random diffusion and directional advection processes:

$$\boldsymbol{b}_{t+1} = g(\boldsymbol{m}(\boldsymbol{b}_t * \exp(-\boldsymbol{u}_t \boldsymbol{F}_t))) * \exp(\boldsymbol{\varepsilon}_t). \quad (2)$$

Here, the function g() is applied elementwise to its vector argument. The function m() symbolizes a movement operation,  $F_t$  is a vector of fishing effort per unit area for all triangles in year t, and  $\varepsilon_t$  represents a vector of process errors for all spatial triangles. Throughout this paper, the symbol '\*' denotes elementwise multiplication for matrices. If there is no '\*' symbol between two matrices, it indicates a standard matrix multiplication. Additionally,  $e^Q$  denotes the matrix exponential of matrix Q, while  $\exp(Q)$  signifies the elementwise exponential. In Eq. (2), the sequence of effects is ordered as fishing mortality, movement, production, and process error. Defining  $p_t =$  $m(b_t * \exp(-u_tF_t))$ , Thorson et al. (2017) applied the Gompertz's production function:

$$g(p_{r,t}) = p_{r,t} e^{\left(\alpha + \omega_r - \beta \log\left(\frac{p_{r,t}}{a_r}\right)\right)}$$
(3)

which ensures the positivity of abundance. Here  $\alpha$  signifies the change in average population abundance on a logarithmic scale,  $\beta$  denotes the logarithmic decline in productivity as population density increases,  $\omega_r$  captures the difference in productivity between triangles (i.e., spatial variance) and  $a_r$  represents the area of the  $r^{th}$  triangle. *a* is a vector of all triangle areas.

The form of Gompertz's production function, introduced by Reddingius (1971) (as cited in Dennis and Taper, 1994), describes the relationship between population abundance or biomass at time

 $t (N_t)$  and at time  $t + l (N_{t+1})$ . The relationship is expressed as:  $N_t e^{(a + b \log(N_t) + \sigma^{Z_t})}$ , where *a* and *b* are constants, and  $\sigma^{Z_t}$  represents a random shock to the population growth. This form was later utilized to develop the Gompertz production model for fish population dynamics (see Thorson et al., 2015b, Thorson et al., 2017).

In Eq. (3), the process error  $\boldsymbol{\varepsilon}_t$  and the spatial variation  $\boldsymbol{\omega}$ , a vector of all the  $\omega_r$ 's, are both assumed to follow a Gaussian field with a mean of zero and a Matérn covariance structure. For example, the Matérn covariance between two components of  $\boldsymbol{\varepsilon}_t$  at  $s_1$  and  $s_2$  is given by

$$Cov(\boldsymbol{\varepsilon}_t(s_1), \boldsymbol{\varepsilon}_t(s_2)) = \frac{\tau_{\varepsilon}^{-2}}{2^{\nu-1}\Gamma(\nu)} (\kappa_{\varepsilon}|s_1 - s_2|)^{\nu} K_{\nu}(\kappa_{\varepsilon}|s_1 - s_2|).$$
(4)

Here  $\Gamma(v)$  is the Gamma function of the smoothness parameter v,  $K_v()$  is the modified Bessel function of the second kind with order v, and  $|s_1 - s_2|$  represents the distance between two spatial points. The parameter  $\tau_{\varepsilon}$  accounts for the pointwise variance of  $\varepsilon_t$ , and  $\kappa_{\varepsilon}$  represents the geographical range of correlations (Babyn et al., 2021). The same concept is applied to  $\omega$ . v is poorly identified in typical applications (Lindgren et al., 2011), so it is commonly set at 1 (Thorson et al., 2017).

A Gaussian field possesses a dense covariance matrix, whose factorization cost is  $O(n^3)$ , where *n* denotes the matrix size. This renders the approach impractical for dealing with large-scale problems. Lindgren et al. (2011) developed an effective methodology to approximate a Gaussian Matérn field using a Gaussian Markov random field (GMRF) based on a triangular spatial grid. This approximation provides the advantages of modelling as a Gaussian field while maintaining the computational efficiency of a GMRF. Thorson et al. (2017) provides R code, based on the R-INLA package (Lindgren and Rue, 2015b), to create the set of spatial triangles with data points positioned near the triangle centers. With R-INLA, users can specify lower bounds for triangle angles and edges, thereby regulating the size and shape of the triangles within the grid design. R-INLA also provides the corresponding design matrices required to construct the precision matrix for the GMRF approximation of the Gaussian field.

Thorson et al. (2017) specified the initial abundance as a deviation from the equilibrium abundance, accounting for spatial fishing mortalities equal to those in the first year and movement,

$$\boldsymbol{b}_0 = m\left(\boldsymbol{a} * \exp\left(\mu_0 + \frac{\omega}{\beta}\right) * \exp\left(-u_1 \boldsymbol{F}_1\right)\right), \ (5)$$

Where  $\mu_0$  is a model parameter to estimate, representing the average logarithm of the equilibrium biomass density across all spatial locations. It establishes the baseline level for the initial biomass distribution ( $\boldsymbol{b}_0$ ), which is then modified by spatial productivity variations ( $\boldsymbol{\omega}$ ), estimated fishing mortality in the first year ( $\boldsymbol{F}_1$ ), and movement effects (m()).

#### Difficulties with the spatial surplus production model

The primary challenges with the spatial surplus production model outlined above lie in the modeling and implementation of movement.

**Modeling Challenges**: Developing a detailed model for diffusion and advection (i.e. *m*() in Eq. 2) with manageable computation costs poses a significant challenge. To address this, Thorson et al. (2017) adopted a simplified approach, assuming fish movement follows a first-order homogeneous differential equation,

$$\frac{\partial \boldsymbol{b}}{\partial t} = \boldsymbol{N}\boldsymbol{b}$$
 (6)

with the solution  $\mathbf{b}_{t+1} = e^N \mathbf{b}_t = M \mathbf{b}_t$ . N is a matrix of instantaneous movement rates; employing the Metzler matrix, i.e., matrix with nonnegative off-diagonal elements (Kemp and Kimura, 1978), for N ensures that M is nonnegative, thereby guaranteeing  $\mathbf{b}_{t+1}$  remains nonnegative if  $\mathbf{b}_t$  is nonnegative. For this formulation,  $\mathbf{p}_t = \mathbf{M}(\mathbf{b}_t * \exp(-u_t \mathbf{F}_t))$  in Eq. (3). Because the spatial domain is discretised into triangles, Thorson et al. (2017) further simplified the instantaneous movement matrix N by restricting it to traverse only neighbouring triangles. This movement is governed by four estimated velocity parameters representing constant rates along the cardinal directions: east (e), south (s), west (w) and north (n). The overall matrix N is constructed as the sum  $\mathbf{N} = v_e N_e + v_s N_s + v_w N_w + v_n N_n$ , where  $N_e$ ,  $N_s$ ,  $N_w$  and  $N_n$  are metrices defining the structure of movement in each respective direction with corresponding velocity parameters  $v_e$ ,  $v_s$ ,  $v_w$ ,  $v_n$ .

Despite all these simplifications, this approach still results in a complex formulation and coding. Furthermore, fully capturing the intricate dynamics of fish diffusion and advection remains a challenge for current methods, although future developments may better account for these dynamics than is currently feasible.

**Implementation Challenges**: Although N is a Metzler matrix involving only neighbouring movement, the yearly movement matrix  $M = e^N$  is not sparse unless N is an identity matrix, resulting in impractical computational costs for spatiotemporal modelling. Thorson et al. (2017) addressed this issue by employing the Euler approximation  $(I + N/n_{time-step})^{n_{time-step}} \approx e^N$ , where  $n_{time-step}$  represents the number of time-steps in the approximation. This approach presents a dilemma: ensuring sparsity in M via the Euler approximation requires  $n_{time-step}$  to be sufficiently small, but a low  $n_{time-step}$  can compromise the accuracy of the movement matrix approximation.

**Parameter Estimation Challenges**: Preliminary simulation analyses conducted for this study reveal challenges in accurately estimating the four velocity parameters governing movement along the four 2D directions. This estimation also impacts the identifiability of other parameters. For instance, both Thorson et al. (2017) and our simulations demonstrate that, in the presence of movement parameters, the spatial correlation range parameter  $\kappa$  cannot be separately estimated for

 $\varepsilon_t$  and  $\omega$  due to model convergence issues, requiring the same  $\kappa$  for both Gaussian fields. This is functionally important because  $\kappa$  determines the spatial scale of correlation, and forcing the scale of persistent productivity patterns ( $\omega$ ) to match the scale of annual process errors ( $\varepsilon_t$ ) is a restrictive assumption that limits the model's ability to capture potentially different spatial dynamics for these different processes.

Similarly, when utilizing the Tweedie distribution for survey catch data, the Tweedie power parameter *p* becomes unidentifiable in the presence of movement parameters. This parameter is critically important because it fundamentally defines the statistical relationship between the expected survey catch and the variance of those catches, a relationship formally expressed as  $Var(Y) = \sigma^2 * \mu^p$ . This parameter is crucial as it defines the variance structure in the survey catch distribution. The failure to identify these crucial parameters, alongside the movement parameters, significantly hampers the practical application of this approach.

#### Accounting for movement by Gaussian random fields

In the spatial surplus production model (Eq. 2), the spatial variation in fish abundance arises from both the movement function, m(), and the process error,  $\varepsilon_t$ . This coexistence can lead to confounding, posing challenges in simultaneously estimating both factors and resulting in the inability to identify certain parameters. To address this issue, we propose an alternative approach by not explicitly modelling movement. Instead, we accommodate movements using a more flexible process error with a distinct correlation range from  $\omega$ . While this approach does not directly compute fish movements, our primary focus is on the spatial distribution of fish, considering movement as nuisance effects. We refer to this approach as the random field model, distinguishing it from Eq. (2), which we term the movement model. The random field model is defined for an individual triangle r as

$$b_{r,t+1} = g(b_{r,t}e^{(-u_tF_{r,t})})e^{\varepsilon_{r,t}},$$
 (7)

where the production function g() retains its definition from Eq. (3), despite a modified argument.  $\varepsilon_t$  and  $\omega$  also maintain their previous definitions and distributional assumptions, but now have distinct distributional parameters for estimation, a modification facilitated by the exclusion of movement from the model, as previously explained. For the initial abundance, we continue to utilize the equilibrium abundance with deviations (Eq. 5), albeit without any subsequent movement,

$$\boldsymbol{b}_0 = \boldsymbol{a} * \exp\left(\mu_0 + \frac{\omega}{\beta}\right) * \exp(-u_1 \boldsymbol{F}_1). \quad (8)$$

We generate data with the movement model and fit it using both the movement model and the random field model. If the random field model accurately estimates fish density and production parameters, demonstrating comparable or superior performance to the movement model, then it will be considered as an effective spatiotemporal surplus production model that is adept at accounting for complex fish movements.

#### Connecting the spatial surplus production model to harvest and stock index components.

Estimation of surplus production model parameters needs to integrate models of fishery harvest and an index of relative stock abundance (typically derived from scientific surveys or fishery catchper-unit-effort):

$$H_{r,t} = b_{r,t} \left( 1 - e^{(-u_t F_{r,t})} \right)$$
(9)  
$$H_t = e^{\delta_t} \sum_{r=1}^{n_r} H_{r,t}$$
(10)  
$$\mu_{r,t} = q_w \left( \frac{b_{r,t}}{a_r} \right)$$
(11)

where the commercial harvest  $H_{r,t}$  in the *r*th triangle and year *t* is a product of triangular stock abundance (or biomass) and fishing mortality rate (i.e.,  $1 - e^{(-u_t F_{r,t})}$ ), Eq. (9),  $H_t$  is the total commercial harvest in year *t*,  $e^{\delta_t}$  represents the measurement error in the harvest data, Eq. (10), *w* is the area swept for a given tow,  $\mu_{r,t}$  is the expected catch of this sample in triangle *r* and year *t*, and  $b_{r,t}/a_r$  is the population density for triangle *r* in year *t*, and *q* represents the catchability coefficient Eq. (11). Of these equations, Eq. 9 and 10 relate to commercial catches while Eq. 11 relate to survey catches.

Our investigations suggest that the Tweedie distribution is robust to the heavy zero-inflations observed in the survey indices of some species such as Redfish in NAFO Divisions 3LN. The Tweedie distribution can also naturally include a multiplicative structure on its mean (Foster and Bravington, 2013). In ecological applications, a multiplicative structure for the expected value is often appropriate, where the expected response variable is proportional to the product of the covariate effects, reflecting a multiplicative rather than an additive relationship (see Foster and Bravington, 2013). In fisheries studies, a multiplicative structure on the expected value appropriately reflects the proportional increase in catches with an increase in sampling effort such as a trawl's swept-area and sampling duration. Therefore, we model the survey catches  $I_{observed}$  using the Tweedie distribution with mean  $\mu_{r,t}$  defined by Eq. (11), power parameter p ranging from 1 to 2 for zero-inflated nonnegative continuous data, and positive dispersion parameter  $\varphi$  (i.e.,  $I_{observed} \sim Tweedie(\mu, p, \varphi)$ ). The variance of a Tweedie distribution follows the power law  $variance = \varphi \mu^p$  (Jørgensen, 1997).

Parameter estimation

The joint likelihood of the data and random effects, which is conditional on the fixed-effect parameters  $\theta$  defining the model structure, can be written as

$$L_{joint}(\boldsymbol{\theta}, \boldsymbol{D}, \boldsymbol{\Psi}) = p_{\boldsymbol{\theta}}(\boldsymbol{D}|\boldsymbol{\Psi})p_{\boldsymbol{\theta}}(\boldsymbol{\Psi}) \qquad (12)$$

where D represents the observed data (i.e., commercial catch, survey indices),  $\Psi$  is the vector of random effects, and  $\theta$  is the vector of fixed-effect parameters. We compute the negative log-marginal likelihood and its gradients using the Template Model Builder (TMB) package in R (R-Core-Team, 2022), which assess the integration over  $\Psi$  using Laplace approximation (Kristensen et al., 2016). Based on these TMB outputs, we obtain the maximum marginal likelihood estimates for  $\theta$  by minimizing the negative log-marginal likelihood through the nlminb() function in R.

#### Simulation study

In this simulation study, we employ sampling sites and production parameters closely aligned with 3LN Redfish data and corresponding analyses presented in the next section, ensuring a close resemblance to real-world fisheries sampling and dynamics. The spatial mesh, consisting of 168 triangles, is optimally determined based on the 3LN Redfish sampling sites using the R-INLA package. Following Thorson et al. (2017), the triangular mesh facilitates decomposing instantaneous movements among adjacent triangles into directional components  $N = v_e N_e + v_s N_s + v_w N_w + v_n N_n$ , where  $N_e, N_s, N_w, N_n$  are movement matrices for east (*e*), south (*s*), west (*w*) and north (*n*) directions, respectively, with corresponding velocity parameters  $v_e, v_s, v_w, v_n$  set to 1. The net movement entering or leaving a triangle edge is calculated as the sum of the projections of the four cardinal movements (i.e., east, south, west, and north) along the direction perpendicular to that edge , as described in Thorson et al. (2017). The yearly movement matrix is  $M = e^N$ . Starting with the initial abundance modeled using Eq. (5), we apply a spatial surplus production model Eq. (3) to simulate 68 years of population dynamics. This extended simulation period allows

the modeled population dynamics to stabilize and move beyond potentially arbitrary initial conditions, ensuring the system reaches a state more representative of its long-term behavior under the specified parameters and stochastic processes (Thorson et al., 2017). Following this stabilization period, we use the simulated data from only the final 8 years to approximate natural abundance distributions in the final 8 years.

The production parameters in model Eq. (3) are set to  $\beta = 0.89$  and  $\alpha = 1.19$  that are based on fitting the 3LN Redfish data. Likewise, the correlation range parameters  $\kappa$  of Gaussian random fields are derived from this fitting. We define that the spatial variation across triangles in productivity is characterized by a standard deviation (i.e.,  $\sigma_{\omega}$ ) of 1 and the standard deviation of spatiotemporal process errors (i.e.,  $\sigma_{\varepsilon}$ ) also takes the value of 1. Fishing efforts  $F_t$  for the 168 spatial triangles over this period are determined by the outer product of a spatial vector (168 elements) and a temporal vector (68 elements). The spatial vector is generated using the geoR package (Ribeiro and Diggle, 2001) as the exponential of a Gaussian random field with a mean of 1 and Matérn covariance (standard deviation: 0.5, range parameter:  $\kappa_F = 1$ ), while the temporal vector is generated elementwise from a lognormal distribution with a mean of 0.2 and a standard deviation of its logarithm set to 0.5. The spatial variation  $\omega$  is generated as a Gaussian random field with a mean of -0.5 and Matérn covariance (marginal standard deviation: 1, range parameter:  $\kappa_F = 1$ ). The mean of -0.5 is chosen to maintain an exponential mean of 1. The process errors  $\varepsilon_t$ are independently generated for each year as Gaussian random fields with a mean of 0 and Matérn covariance (marginal standard deviation: 1, range parameter:  $\kappa_F = 1$ ).

The spatial abundances over the last 8 years are subsequently used to generate commercial harvest data and survey catch data using Eqs. (9)-(11), with a consistent yearly sample size of 388, mirroring the average sample size of 3LN Redfish survey data. The errors  $\delta_t$  in harvest data follow
an independent normal distribution with a mean of 0 and a standard deviation of 0.2, which is close to the fitted value for the Redfish data. Of the 168 triangles, we implemented random selection with replacement to distribute simulated survey data to the triangles. The mean survey catch is calculated using Eq. (11), where q=1 and w=1. The survey catches follow a Tweedie distribution with dispersion parameter  $\varphi = 3$  and power parameter p=1.5 and 1.7. Both movement and random field models are fitted to the harvest and catch data. As the movement model cannot estimate the power parameter, p is set to 1.7 during fitting. When fitting the random field model, p is estimated alongside other model parameters. This simulation is iterated 100 times to obtain sampling distributions for the estimators.

To demonstrate the necessity of spatial modeling, we also fitted the simulated spatiotemporal data using a non-spatial SPM (i.e., not account for spatial variation among triangles) (1), using the Gompertz's production function  $g(b_t) = b_t \exp(\alpha - \beta \log(b_t))$ ; that is, we assumed that fish density is uniform across the 3LN region to fit the spatial individual survey catches, as was done for SSPMs using variable spatial distributions. Note that non-spatial SPMs are typically fitted to annual aggregated observations. However, this approach requires a long time series to reveal production contrasts, including time-periods of higher catches that cause the stock to decline and time-periods of lower catches that allow the stock to grow. Also, our purpose here was to investigate the improvements in fit provided by our spatial model and we cannot directly compare spatial and non-spatial model fits to spatial and aggregated data respectively using Akaike Information Criterion (AIC) (Akaike, 1981) and Bayesian Information Criterion (BIC) (Gideon, 1978). In contrast, our method accommodates spatial data over fewer years due to its larger sample size and spatial contrast in harvest rates and stock production that provide enough information to estimate the model parameters. Also, our approach uses the same data as the other methods in this paper, ensuring that AIC and BIC are comparable for these models.

To compare the performances of random field, movement and non-spatial models, we computed the difference in AIC (Akaike, 1981) and BIC between the two methods in a pair. We also compare the methods based on the correlation between the true (i.e., simulated) and the estimated population densities to investigate which method can better reconstruct the stock density, which is the primary focus of the modeling efforts. The Pearson correlation coefficient was calculated to compare the true and estimated values. We generated histograms of the parameter estimates to examine parameter estimates to their true values.

## Case study application

We applied the model to assess the spatiotemporal dynamics of 3LN Redfish. The available data include annual commercial landings from 1959 - 2019, commercial fishing effort information from 2012 - 2022, and spatiotemporal survey catches between 1983 - 2019, which include Canadian Spring and Fall survey data. In this study, we combined the information from both surveys for the analysis. We only selected information from 2012 - 2019 - 2019 in our analysis for consistency regarding time intervals among different data sources.

Commercial catches sourced from NAFO Data and Statistics (<u>www.nafo.int/Data</u>) exhibit an overall increase in landings since 2012 (Fig. 1). Survey catch information consists of 3098 records between 2012 and 2019, obtained through bottom trawl surveys in the 3LN Divisions by Fisheries and Oceans Canada (DFO). Each record contains catch year, location (i.e., longitude and latitude) and the total number and weight of Redfish caught per tow. In this study, we utilized survey catch

data that had been pre-standardized to a uniform swept area per tow across all catches. The survey catch dataset is highly zero-inflated (Fig. 2), particularly in shallow areas.

Commercial fishing effort information was extracted from an open-source Application Programming Interface (API) published by Global Fishing Watch (GFW, <u>globalfishingwatch.org</u>). GFW's API provides a comprehensive view of the world's fishing efforts, tracking vessels' movement, identity, and activities. By utilizing advanced satellite technology and machine learning algorithms, the platform takes into account both the automatic identification system (AIS) and vessel monitoring systems (VMS), which combine global positioning with a transmitter to regularly broadcast vessel locations. Furthermore, GFW integrates data from over 30 public vessel registries and various satellite imagery sources to create its open-source API. Fisheries scientists have been using data from GFW's platform for their publications since 2016 (e.g., White et al., 2017, Sala et al., 2021, Carneiro et al., 2022) (for a list of publications, see <u>globalfishingwatch.org/publications/</u>).

We obtained fishing effort by commercial trawler operations using the gfwr package published by GFW for the R environment (see <u>github.com/GlobalFishingWatch/gfwr</u>). Each record contains the fishing location (i.e., longitude and latitude) and fishing effort measured in fishing hours (Fig. 3). We then converted the data to a matrix of  $n_r$  rows (i.e.,  $n_r$  triangles) and  $n_t$  columns (i.e.,  $n_t$  years). Each matrix element results from dividing the sum of fishing effort by the area of the corresponding triangle for the given year *t* and triangle *r*. The matrix elements serve as the fishing effort per unit area  $F_{r,t}$  in the movement model Eq. (3) and random field model Eq. (7) to fit the data.

In a preliminary study, the movement model (Eqs. 3 and 5) was applied to the 3LN Redfish data to inform the setup of the simulation studies. However, only results from the random field model are presented in the results section due to its superior performance observed in the simulations.

To demonstrate the goodness of fit, we evaluated the randomized quantile residual (Dunn and Smyth, 1996) for each observed survey catch, a suitable model validation device for non-normal observations. Following this, we generated a Q-Q plot using the qqnorm() function in R to examine the fit of the random field model to the Redfish data. We also compared the estimated and observed survey catches. The conventional "plug-in" estimators are biased, as explained by Thorson and Kristensen (2016). To address this issue, we applied the bias correction feature in TMB, which utilizes the "epsilon" estimator proposed by Thorson and Kristensen (2016) to obtain bias-corrected estimates of survey catches.



Figure 1. Recorded landings of Redfish in 3LN from 1959 to 2019 with the corresponding total allowable catch (TAC). Dots represent TAC and solid line represents commercial catches.



Figure 2. Maps of the spatial survey catches in log scale of Redfish biomass (Kg) in the 3LN Divisions from 2012 to 2019. Bubbles represent survey catches and crosses show survey locations with zero catch.



Figure 3. Maps of commercial fishing effort in log scale of fishing hour by trawlers in the 3LN Divisions from 2012 to 2019. Bubbles represent commercial fishing data.

# 3. Results

# Simulation testing

Simulation testing with p=1.5

Our inspection of parameter estimation revealed that, in the simulation setup with p = 1.5, the parameters  $\alpha$ ,  $\beta$ ,  $\sigma_{\omega}$  and  $\sigma_{\varepsilon}$  are estimated reasonably well using movement and random field models (Fig. 4). The random field model clearly outperformed the movement model in estimating  $\alpha$  and  $\sigma_{\omega}$ . The random field model slightly overestimates  $\sigma_{\varepsilon}$  due to the incorporation of fish

movement by  $\boldsymbol{\varepsilon}_t$ . Quantitative summaries of estimator performance, including relative bias and coefficient of variation, are provided in Appendix Tables S1 and S2.

The correlation between the simulated and estimated population abundance across triangles and years is generally high for both spatial modeling methods with the means of correlation at 0.78 and 0.76 for random field and movement models, respectively, suggesting accurate estimation of the spatiotemporal distribution of population density through each approach (Fig. 5, panel 1 and 2). In contrast, the non-spatial model presents a much greater degree of discrepancy in the correlation between simulated and estimated population abundance with a mean of correlation coefficient at 0.54. In some cases, its estimates are opposite to the simulated data trend (with negative correlation between simulated and estimated population abundance) (Fig. 5, panel 3). This outcome is expected, as the non-spatial model assumes negligible spatial variation in population dynamics, a condition that does not hold for the simulated data.

The random field model results in better estimates with 56% higher correlations (Fig. 6, panel 3) and 54% lower AIC and BIC, scores compared with those of the movement model (Fig. 6, panels 1 and 2). Similarly, the random field model outperforms the non-spatial method with 74% higher correlations and 65% lower AIC and BIC scores (Fig. 7).

Both the random field model and the non-spatial model yield unbiased estimates of the total population biomass (Fig. 8, panel 1 and 3). However, the standard error of the random field model estimator is significantly smaller than that of the non-spatial model. In contrast, the total population biomass estimate derived from the movement model shows a slight bias and exhibits a larger standard error compared to the random field model (Fig. 8, panel 2).

# Simulation testing with p=1.7

In the alternative simulation, we raised the Tweedie power parameter p to 1.7 and estimated it in the random field and non-spatial model approaches, while setting p to 1.5 in the movement model approach due to challenges in parameter identification. The simulation results reveal significant positive biases in  $\alpha$  estimation with the movement model approach, whereas the random field model yields more accurate estimates of  $\alpha$  (Fig. 9, panels A1 and B1). Similarly, the random field model outperforms the movement model in  $\sigma_{\omega}$  estimation (Fig. 9, panels A3 and B3). While the random field model still overestimates  $\sigma_{\varepsilon}$  to accommodate movement as previously explained, the movement model completely missed the true value of  $\sigma_{\varepsilon}$  (Fig. 9, panels A4 and B4), contrasting starkly with its previously decent estimation of  $\sigma_{\varepsilon}$  in Fig. 4, panel B4. This discrepancy suggests that this parameter is highly sensitive to the inaccurate observational model used in the movement model approach.

When comparing the correlation between simulation and estimates of population abundance, the random field model performs at least as well as the movement model. Means of correlation estimates are 0.70 and 0.68 for random field and movement models respectively (Fig. 10). Conversely, the non-spatial method exhibits the poorest performance (0.52), with its estimates occasionally deviating significantly from the simulated abundance (Fig. 10).

In 57% and 61% of simulations, the random field model achieves lower AIC and BIC scores compared to the movement model, respectively (Fig. 11). 55% of the correlations between simulated and estimated population abundances are higher for the random field model compared to the movement model (Fig. 11).

In 67% of the simulations, the random field achieved lower AIC and BIC scores compared to the non-spatial model (Fig. 12). Additionally, the correlation between simulated and estimated population abundance is higher in 74% of the simulations for the random field model compared to the non-spatial model (Fig. 12), indicating that the random field model can reconstruct the population abundance better than the non-spatial model.

Both the random field model and the non-spatial model provide unbiased estimates of total population biomass (or abundance); however, the random field model achieves a significantly smaller standard error (Fig. 13, panel 1 and 3). In contrast, the movement model produces a biased estimate with a larger standard error compared to the random field model (Fig. 13, panel 2).

Our simulations demonstrated that the movement model method frequently encountered convergence issues. In contrast, the random field model approach had no issues with model convergence regardless of different simulation settings.



Figure 4. Simulation results for (A) random field, (B) movement and (C) non-spatial models including parameter estimator distributions of (1) the average density  $\alpha$ , (2) the density dependence  $\beta$ , (3) the standard deviation of spatial variation in productivity  $\sigma_{\omega}$ , and (4) the standard deviation of spatiotemporal process error  $\sigma_{\varepsilon}$ . The simulation used Tweedie power parameter p = 1.5. Red lines are the true values of the parameters and blue lines are means of parameter estimates.



Figure 5. Simulation results of correlation between true abundance (i.e., simulated) and estimated abundance for (1) random field, (2) movement and (3) non-spatial models. The simulation used Tweedie power parameter p = 1.5. Blue lines are means of correlation estimates.



Figure 6. Boxplots of differences in statistical metrics including AIC, BIC, and correlation between random field and movement models. The simulation used Tweedie power parameter p =1.5. The differences are:  $AIC_{difference} = AIC_{random field} - AIC_{movement}$ ;  $BIC_{difference} =$  $BIC_{random field} - BIC_{movement}$ ; and  $Correlation_{difference} = Correlation_{random field} Correlation_{movement}$ .



Figure 7. Boxplots of differences in statistical metrics including AIC, BIC, and correlation between random field and non-spatial models. The simulation used Tweedie power parameter p = 1.5. The differences are:  $AIC_{difference} = AIC_{random field} - AIC_{non-spatial}$ ;  $BIC_{difference} = BIC_{random field} - BIC_{non-spatial}$ ; and  $Correlation_{difference} =$  $Correlation_{random field} - Correlation_{non-spatial}$ .



Figure 8. Mean of differences between predicted versus generated annual biomasses of 1) random field model; 2) movement model; and 3) nonspatial model. The simulation used Tweedie power parameter p = 1.5. The difference is  $Biomass_{difference} = Biomass_{generated} - Biomass_{predicted}$ . The shading areas show 95% confidence intervals.



Figure 9. Simulation results for (A) random field, (B) movement and (C) non-spatial models, including parameter estimator distributions of (1) the average density  $\alpha$ , (2) the density dependence  $\beta$ , (3) the standard deviation of spatial variation in productivity  $\sigma_{\omega}$ , and (4) the standard deviation of spatiotemporal process error  $\sigma_{\varepsilon}$ . The simulation used Tweedie power parameter p = 1.7. Red lines are the true values of the parameters and blue lines are means of parameter estimates. In the movement model, p is fixed, whereas, in the random field and non-spatial models, p is estimated.



Figure 10. Simulation results of correlation between true abundance (i.e., simulated) and estimated abundance for (1) random field model, (2) movement model and (3) non-spatial model methods. The simulation used Tweedie power parameter p = 1.7. Blue lines are means of correlation estimates. In the movement model, p is fixed, whereas, in the random field and non-spatial models, p is estimated.



Figure 11. Boxplots of differences in statistical metrics including AIC, BIC, and correlation between random field and movement models. The simulation used Tweedie power parameter p =1.7. The differences are:  $AIC_{difference} = AIC_{random field} - AIC_{movement}$ ;  $BIC_{difference} =$  $BIC_{random field} - BIC_{movement}$ ; and  $Correlation_{difference} = Correlation_{random field} -$ 

*Correlation<sub>movement</sub>*. In the movement model, p is fixed, whereas, in the random field model, p is estimated.



Figure 12. Boxplots of differences in statistical metrics including AIC, BIC, and correlation between random field and non-spatial models. The simulation used Tweedie power parameter p = 1.7. The differences are:  $AIC_{difference} = AIC_{random field} - AIC_{non-spatial}$ ;  $BIC_{difference} = BIC_{random field} - BIC_{non-spatial}$ ; and  $Correlation_{difference} =$  $Correlation_{random field} - Correlation_{non-spatial}$ .



Figure 13. Mean of differences between predicted versus generated annual biomasses of 1) random field model; 2) movement model; and 3) nonspatial model. The simulation used Tweedie power parameter p = 1.7. The difference is  $Biomass_{difference} = Biomass_{generated} - Biomass_{predicted}$ . The shading area shows 95% confidence intervals. In the movement model, p is fixed, whereas, in the random field and non-spatial models, p is estimated.

## Case study application

The estimated spatial distribution of Redfish biomass within the 3LN Division, as shown in Fig. 14, exhibits heterogeneity. A peak in estimated biomass occurs at the southern part of Division 3N. Redfish populations predominantly aggregate in deepwater zones, specifically within the bathymetric range of 200 to 500 meters. Conversely, in the shallower littoral and nearshore environments, Redfish biomass is much lower, with some areas exhibiting negligible or non-existent populations. In the areas of high concentration, fluctuations in Redfish biomass are evident from 2012 to 2019. The estimated spatiotemporal distribution of biomass (Fig. 14) demonstrates a high agreement with the predicted survey catch data (Fig. 15) over the period extending from 2012 to 2019.

In addition, the alignment between model-estimated and observed catches (Fig. 16, panel A), along with the straight-line pattern in the normal Q-Q plot, indicates that the model fits the survey data very well (Fig. 16, panel B). The parameter estimates and their corresponding standard errors are presented in Table 1. The large difference between  $K_{\varepsilon}$  and  $K_{\omega}$  highlights the advantage of the random field model, as the movement model is constrained to treat these parameters as

equal. In addition, the Tweedie power parameter *p* takes on a value (1.77) that is not easily anticipated and is best estimated, which is also beyond the movement model.



Figure 14. Maps of population biomass estimates for Redfish in the 3LN Divisions, displayed on a logarithmic scale as log (1 + biomass estimate). The color legend is presented at the bottom of the plot.



Figure 15. Model predicted spatial survey indices (i.e., the estimated spatial distribution) of Redfish in the 3LN Divisions, shown on a logarithmic scale as  $\log (1 + \text{survey index})$ . The color legend is presented at the bottom of the plot.

Table 1. Parameter estimates and standard errors for the random field model fit to the

3LN Redfish data.

Parameters	Estimates	Standard Error
β	0.43	0.31

α	0.97	2.08
Kε	0.53	0.30
K <sub>ω</sub>	7.76	0.36
$\sigma_{arepsilon}$	1.43	0.24
$\sigma_{\omega}$	3.32	0.39
$\mu_0$	15.54	4.73
q	0.95	6.42
φ	9.19	0.031
p	1.77	1.15
$\sigma_{catch}$	0.25	0.75





Figure 16. Plots of (A) estimated versus observed survey catches for each year and (B) QQ Plots of the randomized quantile residuals for survey catches. In panel (A), the x axis represents the estimated survey catches, the y axis represents the observed survey catches, and the red line has a slope of one through the origin. The bias correction feature in TMB was applied to predict survey catches.

# 4. Discussion

## Non-spatial SPM vs. spatial SPM

In our simulation testing, we applied non-homogeneous distributions for catches and survey indices across the spatial area (i.e., the triangles) to closely mimic a real-world fisheries scenario. Our simulation study revealed that the non-spatial model struggled to accurately estimate the parameter  $\alpha$ , exhibiting substantial bias and large standard errors, in the presence of spatial variation (Figs. 4 and 9), as observed in the Redfish distribution. Separately, previous research has shown that conventional (non-spatial) Gompertz models can also yield biased estimates of density dependence ( $\beta$ ) when spatial density variations are present (Thorson et al., 2015b). Together, these findings highlight that ignoring spatial structure can lead to mischaracterizations of multiple key population dynamic parameters. The non-spatial model also produced significantly wider confidence intervals for total population biomass compared to the random field model (Figs. 8 and 13), indicating its lower estimation precision. Therefore, we suggest applying spatial surplus production models (SSPMs), when possible, in data limited fish stock assessments such as our case study, where the time-series was short (only 8 years), age data were not available, and only total annual catches for the entire area were available. It will usually not be possible to fit an aggregated SPM to such limited data, but we could fit a spatial SPM because of the increase in spatial sample size and associated information on how spatial variations in fishing effort and harvest rates affect stock production.

Analysis of case study for the 3LN Redfish





Figure 17. The 3LN Redfish plots of A) observed survey data; B) observed and predicted commercial landings; C) estimates of stock biomass using ASPIC software; D) estimates of stock biomass using the random field model. The shading areas in panel B and D show 95% confidence intervals.

From 2008 to 2020, the assessment of 3LN Redfish utilized a logistic surplus production model (SPM) executed through a non-spatial ASPIC software suite (Prager, 1994, Prager, 2016). There is a clear discrepancy between the observed survey biomass and the biomass estimates from the ASPIC model (Fig. 17, panels A and C). Moreover, Rogers et al. (2022) reported that Redfish in the size range of 15 to 20 cm have consistently been lower than the long-term mean since the mid-2010s, as observed in both Spring and Autumn Canadian 3LN surveys and the EU-Spain 3L and 3N survey series. These signals indicate issues with a reduction in the stock biomass. This potential decline, coupled with the trend of increasing fishing pressures recently (Fig. 17, panel B), raises significant concerns about the sustainability of the population over the long term. Conversely, the random field model indicated that stock biomass declined almost 50% between 2012 and 2019 (Fig. 17, panel D). This trend corresponds with the patterns seen in the available survey indices (Rogers et al., 2022) and correlates with rising fishing pressures (Fig. 17, panel B).

SSPMs offer the significant benefit of depicting the stock's spatial structure. Although spatial survey data alone can map population distribution across space, incorporating spatially explicit commercial data, as described in our method, i.e., Eq. (7), allows the estimation of "absolute" stock abundance. This integration makes the random field model particularly robust by enhancing its ability to reconstruct stock structure accurately. In addition to that, the normal QQ-plot confirms the goodness of fit when applying the model to the 3LN Redfish survey data.

#### Factors considered in spatial stock assessment models

To extend a non-spatial model to explicitly handle spatial population dynamics, it is essential for modelers to conduct space discretization. In this study, we utilized the R-INLA package (Lindgren and Rue, 2015a, Roger et al., 2015) to discretize the spatial domain of interest into a set of triangles. This technique was also utilized to implement spatial division in earlier studies on spatial fisheries models (e.g., Thorson et al., 2017, Thorson, 2022, Olmos et al., 2023). A noticeable benefit of the technique is that it provides adjustable spatial resolution, enabling the precise modeling of spatial variation with a customizable number of triangles to match the biologically relevant scales. It also can accommodate irregular boundaries, often found in marine environments, through the strategic use of triangles, making it particularly suited for areas with complex coastlines. However, in practice, modelers are required to consider trade-offs between spatial resolution and computational costs, as higher resolution increases the costs.

Spatial discretization could also be done using a grid of squared cells (see Cadigan et al., 2017). The grid cell approach can provide a solution to handle high-resolution data and provide flexibility in scenario testing. An alternative technique involves employing Voronoi Tessellations, as described in (Pedersen et al., 2022), where a random selection of locations from previous trawl surveys within the study areas was used to calculate a Voronoi Tessellation. Each polygon

encompasses all points nearer to its generating point than any other. This ensures that each polygon represents an area of the sea floor closest to its corresponding trawl survey location. This approach was defined to satisfy specific criteria prioritized for Northern shrimp assessment. Cadigan et al. (2022) developed a method to implement GMRFs using DFO strata as spatial grid units, which are irregular in size and shape. Since DFO strata are designed based on ocean floor depth, an important environmental factor for groundfish like Redfish, this approach inherently incorporates depth information, making it well-suited for modeling groundfish. However, the irregular grid shapes complicate movement modeling, as this requires estimating velocities perpendicular to grid edges. In contrast, our random field approach is better suited for spatial modeling on irregular grids. As future work, we will implement an SSPM using DFO strata as spatial grids, incorporating the GMRF methodology proposed by Cadigan et al. (2022).

An essential consideration is addressing spatial variation in the modeling process. Generally, modelers establish a mechanism to compute the variation between a spatial unit (i.e., a square cell or a triangle) and its neighboring units across a temporal step (e.g., yearly). The computation often accounts for spatiotemporal variations in process errors (i.e., model uncertainties) and biological processes, such as abundance, biomass, and recruitment dynamics, with the integration of age or size classes when applicable (Punt, 2019). One approach involves directly modeling movements using movement matrices and estimating movement parameters, as demonstrated for big skate by (Thorson et al., 2017). Another approach employs GMRF to estimate spatiotemporal variation in process error and biological functions such as our random field model for Redfish, studies on snow crab (*Chionoecetes opilio*) (Olmos et al., 2023), and scallop in the Bay of Fundy (McDonald et al., 2021). Within the scope of this study, we further discuss the advantages of employing the random

field model compared to using the movement model approach when tackling spatiotemporal variation.

# The advantages of random field model

First, both SSPMs, the random field and movement models, can apply production contrasts across space to fit the SPMs to a relatively short time series of data (i.e., eight years), which can be beyond the conventional approach of using aggregated spatial data.

In comparing the two spatial models tested in this study (i.e., the random field and the movement models), our simulation study shows that employing the movement model does not provide a reliable estimation of movement rate parameters, which is consistent with a finding presented by Thorson et al. (2017). Attempts to apply the explicit movement model to the 3LN Redfish dataset encountered challenges (e.g., convergence issues, unreliable parameter estimates), reaffirming the limitations previously observed in simulations regarding its ability to accurately compute movement rate parameters in this context. In the context of SSPMs (Eq. 2), the spatial heterogeneity is attributed to the confluence of the movement function, (m()), and the stochastic process error  $(\boldsymbol{\varepsilon}_t)$ . We suspect this simultaneous occurrence to have a confounding effect, thereby hindering the accurate identification and differentiation of those parameters. Another dimension of the challenges in parameter estimation involves the use of the Euler approximation (i.e.,  $(1 + N/n)^n \approx e^N$ ) where a sufficiently small value of n is chosen to ensure sparsity in the movement matrix M (i.e.,  $M = e^{N}$  as the solution of Eq. 6). In this context, the parameter n is instrumental in modelling the annual movement of fish across adjacent triangles. Employing a small value for *n* might not be justifiable for species characterized by active swimming behaviours, potentially leading to unreliable estimation of the movement parameters. Note that these difficulties pertain only to the simplified movement model. Incorporating more realistic modeling of fish movements, such as diffusion and advection, can introduce significantly more complex challenges.

Also, we note that spatial heterogeneity is a fundamental aspect of marine ecosystems and is expected to manifest in both the underlying productivity potential and the annual process errors within fish stock dynamics. Spatial heterogeneity in productivity (i.e.,  $\omega_r$ ) arises from environmental features and biological interactions. Factors such as depth, substrate type, bottom temperature gradients, oceanographic features (e.g., fronts), localized food availability, or spatially varying predator fields can create regions with consistently higher or lower potential for growth, survival, and recruitment, leading to spatial structure in average productivity. Both process error and spatial productivity differences are assumed in this modeling framework to follow a Gaussian random field, thereby being influenced by the geographical range of correlations (i.e.,  $\kappa$ ) and pointwise variance (i.e.,  $\tau$ ).

Our investigation uncovered an additional limitation of the movement model approach: its inability to estimate the geographical range of correlations for process error (i.e.,  $\kappa_{\varepsilon}$ ) and spatial productivity variance (i.e.,  $\kappa_{\omega}$ ). Functionally, this constraint means the model must assume that the characteristic spatial difference of productivity is identical to the spatial difference of year-to-year random variations. However, this forced equivalence is unlikely to reflect ecological reality, as baseline productivity patterns and annual stochastic deviations represent distinct processes often driven by factors operating at different spatial extents. To address this problem, we assumed the geographical range of correlations for these two parameters was equal (i.e.,  $\kappa_{\varepsilon} = \kappa_{\omega}$ ). In contrast, the random field model is capable of separately and effectively estimating the geographical correlation ranges for each of these processes (i.e.,  $\kappa_{\varepsilon} \neq \kappa_{\omega}$ ). By circumventing the challenges of explicitly modeling movements, our approach introduces flexibility and simplifies the parameter estimation process. This strategy effectively concentrates on addressing the spatial heterogeneity of the fish stock dynamics.

Our study faced convergence challenges with the movement model, similar to those reported by Thorson et al. (2017). Previously identified computational challenges are likely contributing to the observed convergence issues. Additionally, even though setting the instantaneous movement rate N in Eq. (6) as a Metzler matrix ensures positive population abundances after movements, Euler approximation of the movement matrix often results in negative abundances. Consequently, the movement model has to frequently call the 'posfun' function (Fournier et al., 2012) to maintain non-negative population abundances. While essential, this approach may result in slower computational performance and potentially exacerbate model convergence problems. In contrast, the random field model does not encounter these issues and shows no convergence problems when fitting either simulated or empirical Redfish data. This indicates its superiority in terms of convergence and robustness. Simulation testing revealed that the random field model achieves significantly higher rates of convergence (p=1.5 convergence rate = 97%, and p=1.7 convergence rate = 95%) compared to the movement model (p=1.5 convergence rate = 51%, and p=1.7 convergence rate = 49%). These findings have been included in the Appendix (Fig. S3).

The simulation results consistently favored the random field model across multiple metrics. The higher correlations between estimated and true abundance achieved by this model (Figs. 5 and 10) are indicative of better estimates regarding the reproduction of relative spatio-temporal patterns. Concurrently, the lower AIC and BIC scores consistently obtained for the random field model (Figs. 6, 7, 11, 12) are indicative that this model would likely be chosen over the movement and non-spatial alternatives in practice based on these standard model selection criteria assessing relative fit to the observation data. Furthermore, direct comparisons confirmed the random field model

produced unbiased or less biased estimates of total annual biomass with higher precision than the alternatives (Figs. 8 and 13).

#### Limitations for implementing spatial stock assessment models

The unavailability of detailed spatial information is a major constraint when applying spatial stock assessment models. While spatial survey data is usually available, commercial catches often lack sufficient or accurate information regarding harvesting locations. Although some fisheries require reporting fishing locations in logbooks, for some fisheries it is rare to have catches distinguished by species and their corresponding harvesting locations for each haul. Typically, commercial logbooks include spatial information, the total catch and the gear type used. It is noted that even when detailed spatial locations are recorded in logbooks, confidentiality regulations often restrict access to this data for research purposes.

In certain fisheries, such as Northern Shrimp, the catch predominantly consists of a single target species, making bycatch species negligible (Pedersen et al., 2022). In these situations, using SSPMs that incorporate commercial catches could yield much more accurate estimates. Another consideration is the availability of supporting information such as age, size, and environmental indicators. In fisheries where such data are available, explicit spatial models can uncover spatial recruitment and natural dynamics (McDonald et al., 2023), investigate recruitment and maturity, and suggest a connection between environmental conditions (e.g., cold pool) and size class abundance (Olmos et al., 2023).

A limitation associated with this GFW dataset is that it represents the effort of commercial trawlers within the specified area and timeframe and could not be filtered to isolate effort specifically targeting Redfish. Consequently, the fishing effort used in this analysis likely includes effort directed at other species. In addition, the data might include operations from vessels of varying sizes and efficiencies, yet a 'fishing hour' is treated uniformly in this analysis regardless of vessel characteristics (e.g., size, horsepower, gear configuration). It is also based on AIS/VMS signals indicating fishing behaviour inferred by GFW algorithms, which may not perfectly capture all fishing operations or misclassify some activity. We acknowledge this data constraint, and incorporating more refined, species-specific effort data, potentially from logbooks or other sources if they become available, represents an important area for improvement in future research.

The simulation testing conducted in this study incorporated spatial heterogeneity through Gaussian random fields for productivity and process errors, this generated random spatial structure according to the specified correlation range. A valuable area for future research would be to explicitly test the model's performance under scenarios with more deterministic or persistent spatial structures. For example, simulating populations where large areas consistently have near-zero density (due to unsuitable habitat) or where specific locations consistently exhibit high densities (e.g., predictable aggregation sites), reflecting patterns observed for species like Redfish which associate strongly with specific bathymetric features, was not explicitly part of this study's simulation design. Assessing how well the random field model captures biomass trends and avoids bias when confronted with such strong, persistent spatial structuring, potentially driven by underlying habitat covariates not explicitly included in the current model formulation, would further clarify the model's robustness and limitations for species with highly defined spatial distributions.

In our study, the random field model demonstrated good performance in simulation studies, suggesting its high potential for applications in fish stock assessments, particularly when sufficient spatial data are available. With the advantages mentioned above, our random field model approach

along with multiple environmental predictors (e.g., predator and prey densities) may produce a more comprehensive understanding of the ecosystem dynamics.

# Chapter 3. Applying random field model on Yellowtail Flounder Stock Assessment in the Eastern Grand Banks of Newfoundland

# 1. Introduction

Yellowtail Flounder (*Limanda ferruginea*, Storer, 1839), a flatfish species native to the eastern North American coast, inhabits the waters from Newfoundland and Labrador, Canada to Chesapeake Bay, USA. There are American stocks (NOAA, 2024) and the NAFO Grand Bank stock (NAFO, 2024). Their distribution has remained relatively stable since the mid-80s, with a significant portion of the population concentrated north of 45 degrees latitude (Dwyer et al., 2003, Maddock Parsons et al., 2023).

The species transitions from a pelagic larval stage to a benthic juvenile and adult phase (Murua and Saborido-Rey, 2003). Yellowtail Flounder are batch spawners, with peak spawning occurring from April to June on the Southeast Shoal of the Grand Bank, a region abundant in juveniles (Ollerhead et al., 2009). The species is characterized by its flat, ovate shape, upturned mouth, and distinctive yellow tail fin. The fish are adept at camouflage, adjusting their coloration to match the ocean floor. They mature at about 4–5 years and they usually grow to 38 – 40 centimeters in length and 0.5 – 0.6 kilograms in weight. Some individuals can reach up to 12 years of age, though most do not live past 10 years (Bowering and Brodie, 1991, Walsh, 1992). Yellowtail Flounder feed on invertebrates such as polychaete worms and crustaceans and are preyed upon by predators like cod and hake (NOAA, 2024).

The Yellowtail Flounder stock in NAFO Divisions 3LNO has been managed under Total Allowable Catch (TAC) regulations since 1973 (Maddock Parsons et al., 2023). Catches have fluctuated greatly since the 1960s, with a peak of 39,000 tons in 1972 and recent catches ranging from 3,100 to 14,800 tons. Catches exceeded quotas in the late 1980s and early 2000s but have generally been below quotas since then. Industry factors and difficulty locating fish due to potentially changing water temperatures are thought to be behind recent lower catches. Most catches are from Division 3N and are taken using otter trawls. From 2000 to 2015, stock assessments were conducted using ASPIC (Prager, 1994), a surplus production model, incorporating covariance catch and survey indices (Maddock Parsons et al., 2018). Concerns were raised about the ASPIC model's lack of sensitivity; therefore, a Bayesian model was adopted for 3LNO Yellowtail Flounder, replacing the previous ASPIC-based model since 2018 (Maddock Parsons et al., 2018). The model employs the Schaefer equation, characterized by two key parameters: r (the intrinsic rate of population growth) and K (the carrying capacity). Stochasticity in population dynamics is integrated into the model by accounting for process error. The unobserved stock biomass is modeled through an equation that establishes a relationship between the research survey indices and the catchability parameter (q), incorporating observation error (Maddock Parsons et al., 2023). Ocean Choice International (OCI), the primary fishing company in this region, records Yellowtail Flounder catches in onboard logbooks, including discards at sea, as well as on landing tickets (Knapman et al., 2020).

Several studies have attempted to incorporate spatially explicit analyses into stock assessment models for Yellowtail Flounder. Goethel et al. (2014) conducted a simulation analysis using tagrecapture data to evaluate the performance of a spatially explicit, tag-integrated model. Building on this, Goethel et al. (2015) expanded the work by applying the model to assess three interconnected Yellowtail Flounder stocks in New England. Robertson et al. (2021) applied the Vector Autoregressive Spatiotemporal (VAST) model to survey data collected by DFO to analyze spatial distributions of Yellowtail Flounder in Grand Banks off Newfoundland. Commercial fishing information has not been utilized to investigate the spatiotemporal dynamics of Yellowtail Flounder, particularly for the 3LNO stock.

Fisheries surveys employing trawling are a fundamental tool for collecting data essential to the evaluation and management of fish stocks across global marine ecosystems (Gunderson, 1993). However, these surveys are subject to measurement errors due to uncertainties in the efficiency of the trawl at each tow. Furthermore, variability in the swept area/volume arises from differences in trawl wing spread, current speed and direction, herding, and other operational factors. Addressing these discrepancies between observed and expected survey catches is vital for enhancing the precision and reliability of stock assessments. In this context, considerable research has focused on improving the standardization of survey methodologies. Nonetheless, measurement error still exists in survey catches as indices of local population size at a tow site. Cadigan and Chen (2011) found that the variance of the measurement error was a quadratic function of the mean local stock density. He used this relationship to refine confidence intervals for the average population survey catch. Building on these results, we were motivated to further investigate the relationship between expected and observed survey catches when using the random field model developed in the previous chapter and applying this approach to a new case study: Yellowtail Flounder in the Eastern Grand Banks of Newfoundland.

In this study, we aim to 1) investigate the application of the random field model to assess the Yellowtail Flounder stock in the 3LNO Divisions of the Northwest Atlantic; and 2) refine the understanding of the relationship between expected and observed survey catches using the

quadratic function, building on previous methodologies to improve the accuracy of stock assessments. In addition, we compare our results using the random field model with predicted catches and stock biomass estimates presented in the 2023 Assessment of Yellowtail Flounder in NAFO Divisions 3LNO by Maddock Parsons et al. (2023) using the Bayesian SPM model.

# 2. Materials and methods

We applied the random field model described in Chapter 2 to the assessment data ofr Yellowtail Flounder (*Limanda ferruginea*) stock in NAFO Divisions 3L, 3N and 3O (3LNO, the Eastern Grand Bank of Newfoundland). In Chapter 2, we applied a Tweedie distribution to model the survey catches with mean  $\mu_{r,t}$  defined in Eq. 11. While the Tweedie distribution has proven effective, advancing its application to zero-inflated data may benefit from extending its dispersion relationship. Cadigan and Chen (2011) found that a quadratic dispersion relationship was appropriate for fisheries trawl survey catches. Therefore, when using the Tweedie distribution to model the survey catches (i.e.,  $I_{observed} \sim Tweedie(\mu, p, \varphi)$ ), we were motivated to utilize a quadratic function to account for the parameter  $\varphi$  as follows

$$\hat{\varphi} = \frac{a * \hat{\mu}_{r,t} + b * \hat{\mu}_{r,t}^{2}}{\hat{\mu}_{r,t}^{p}}, \quad (12)$$

where a and b are quadratic variance parameters,  $\hat{\mu}_{r,t}$  is the expected survey catch in triangle *r* and time *t*, *p* is the power parameter of the Tweedie distribution ranging from 1 to 2. In Eq. 12,  $\varphi$  is the positive dispersion parameter which is a function of  $\mu_{r,t}$ .

We compared model performance between using the quadratic dispersion relationship and not using this relationship in calculating the parameter  $\varphi$  when modelling the survey catches as a Tweedie distribution. We computed AIC (Akaike, 1983) and BIC (Gideon, 1978) scores to examine the goodness of fit between the two approaches.

The spatial random field surplus production model (SSPM) was used to investigate the spatiotemporal dynamics of Yellowtail Flounder in 3LNO. The data consisted of annual commercial landings, spatiotemporal survey catches and commercial fishing effort data between 2012–2019.

Commercial landings data, sourced from NAFO Data and Statistics (<u>www.nafo.int/Data</u>), contains 8 records between 2012 to 2019. The catch increases sharply from 2012 to 2013 but then declined until 2015. After 2015 landings increased steadily and then substantially increased in 2019, reaching a high of about 12,000 tons (Fig. 18, panel A).

Survey catch dataset comprises 4,217 records, combining Spring and Fall surveys, collected from bottom trawl surveys conducted by DFO in the 3LNO Divisions between 2012 and 2019. Each record includes details such as the year, geographic location (i.e., latitude and longitude), and the total weight of Yellowtail Flounder. The 2014 Fall and 2015 Spring did not cover the entire stock area, resulting in their estimates not being considered representative. There are Spring and Fall Canadian surveys.

For the purpose of this analysis, which focuses on demonstrating the application of the random field model to estimate overall annual biomass trends and surplus production parameters consistent with previous applications, the data from both the Spring and Fall surveys were pooled annually. Operationally, this means that all survey observations from a given year (Fall and Spring), were used collectively. This approach leverages the increased spatial coverage and data density. Treating the Spring and Fall surveys as separate indices with potentially different catchability coefficients

(i.e., q) might provide a more nuanced understanding of seasonal dynamics, such a detailed seasonal model was considered beyond the scope of this specific application (Fig. 18, panel B).

The spatial distribution of observed survey catches from 2012 to 2019 reveals a significant concentration of fish in the central and southern regions of the spatial domain (Fig. 19). All survey catches have been standardized to a common swept area calculated as the product of the mean distance tow (0.8 nautical miles, equivalent to 1.4816 km) and the standard wing spread of the trawl door (55.25 feet, equivalent to 0.1685 km).

Commercial fishing effort data for the spatial domain of the 3LNO divisions were obtained via the GFW API, as detailed in Chapter 2. Each record contains the information of fishing time (year) geographic coordinates (longitude and latitude) representing the fishing location, and associated fishing effort quantified in fishing hours (Fig. 20).

To assess the goodness of fit, we calculated the randomized quantile residual (Dunn and Smyth, 1996). Subsequently, we created a Q-Q plot using the qqnorm() function in R to evaluate how well the random field model fits the Yellowtail Flounder dataset.


Figure 18. Plots of (A) recorded landings of Yellowtail Flounder (ton) in the 3LNO Divisions; and (B) total catches of observed survey data recorded for Yellowtail Flounder (ton) in the 3LNO Divisions.



Figure 19. Spatial survey catches in log scale of Yellowtail Flounder (Kg) in the 3LNO Divisions from 2012 to 2019. Bubbles represent survey catches, and crosses show locations with zero catch.



Figure 20. Spatial commercial fishing effort in log scale of fishing hour by trawlers in the 3LN Divisions from 2012 to 2019. Bubbles represent commercial fishing data.

# 3. Results

Table 2. AIC and BIC estimates for using quadratic and not using quadratic

dispersion relationship approaches.

Approach	AIC	BIC
Quadratic dispersion relationship	17321.30	17435.55
Non-quadratic dispersion relationship	17669.77	17777.67

I find that the quadratic dispersion relationship outperforms the alternative non-quadratic one based upon AIC and BIC score (Table 2). From this line onward, only results from the quadratic dispersion relationship approach are presented.

Over the eight years between 2012 – 2019, Yellowtail Flounder's spatial biomass distributions appear to remain concentrated in the central and southeastern portions of the 3LNO region. These regions consistently show greater biomass estimates. Slight interannual variations are evident, with some shifts in the intensity and extent of these high-biomass locations from year to year (Fig. 21). Spatial distributions of observed survey data align with that of estimated stock biomass from 2012 to 2019. While the DFO scientific surveys cover the entire 3LNO area, positive catches of Yellowtail Flounder primarily occurred in the central and southern areas (Divisions 3N and 3O) of the spatial domain (Fig. 22). Additionally, the alignment between the model-estimated and observed catches (Fig. 23 panel A), combined with the linear pattern observed in the normal Q-Q plot (Fig. 23 panel B), demonstrates a strong fit of the model to the survey catch data.



Figure 21. Maps of population biomass estimates for Yellowtail Flounder in 3LNO divisions, displayed on a logarithmic scale as log (1 + biomass estimate). The biomass was estimated in Kg. The color legend is presented at the bottom of the plot.



- 500 m - 200 m 0.0 1.0 2.0 3.0 4.0 5.0 6.0

Figure 22. Observed spatial survey catches of Yellowtail Flounder in 3LNO Divisions, shown on a logarithmic scale as  $\log (1 + \text{survey index})$ . The color legend is presented at the bottom of the plot.



Figure 23. Plot of (A) estimated over observed survey catches for each year and (B) QQ Plots of the randomized quantile residuals for survey catches. In panel (A), the x axis represents the estimated survey catches, and the y axis represents the observed survey catches with a 45-degree

line of equality (the red line). The bias correction feature in TMB was applied to evaluate the estimated survey catches (Thorson and Kristensen, 2016).

Table 3. Parameter estimates and standard errors when fitting the model to the 3LNO Yellowtail

Parameters	Estimates	Standard Error
β	0.31	0.67
α	0.88	2.03
Kε	0.23	0.44
Κ <sub>ω</sub>	0.70	0.49
$\sigma_{arepsilon}$	0.19	0.24
$\sigma_{\omega}$	1.53	0.80
μ <sub>0</sub>	1.24	6.52
q	0.95	1.10
а	72.26	0.07
b	0.0055	12.26
р	3.43	0.048
$\sigma_{catch}$	0.25	0.70

Flounder dataset.

## 4. Discussion

Spatial dynamics of 3LNO Yellowtail Flounder

A strong agreement between biomass estimates and observed survey catches can be seen between 2012 and 2019. While the core biomass regions are stable, subtle shifts in the intensity of biomass concentration are slightly noticeable between years (Fig. 21 and 22). Considering the depth contours (i.e., -500 m and -200 m), the concentration of high-density regions in shallower areas suggests that Yellowtail Flounder may prefer shallower waters (Fig. 21). This preference is reinforced in the distribution map of observed survey catches (Fig. 22) and in agreement with findings by earlier studies (see Walsh, 1992, Pereira et al., 2012). In fisheries assessment, survey catches are often considered as the proportional representation of the true stock dynamics; therefore, the results demonstrate that the random field model can effectively reconstruct the spatial dynamics of the stock.

The DFO's 2023 assessment reveals that predicted annual commercial catches consistently exceed observed catches during 2012 – 2019 (Fig. 24 panel B). This discrepancy suggests potential underreporting in commercial landings data, a common challenge in fisheries management (e.g., Bousquet et al., 2010, Cadigan and Marshall, 2016, Van Beveren et al., 2017).

Additionally, the pattern could indicate that the assessment model, which assumes constant natural mortality, may be underestimating the true natural mortality rate (M) for Yellowtail Flounder. Time-varying natural mortality, potentially linked to environmental conditions or predation as observed in other regional groundfish stocks like American Plaice (Robertson et al., 2025) and Atlantic Cod (Regular et al., 2022, Cadigan et al., 2024), would lead to lower actual biomass available for catching than predicted by a model assuming lower, constant M.

Addressing potential underreporting remains a valid avenue for model refinement; future research could explore treating reported landings as lower bounds using a censored likelihood approach (Perreault et al., 2020b). Simultaneously, investigating potential time-varying natural mortality for

Yellowtail Flounder, perhaps linked to environmental covariates or condition indices similar to the approach used for plaice and cod, also warrants consideration.

Biomass estimates from the random field model and reported results from DFO in 2023 (Maddock Parsons et al., 2023) both indicate a general decline in biomass over the eight years from 2012 to 2019. However, they differ in the magnitude of the decline subsequent recovery patterns, and the confidence intervals produced by the random field model are narrower overall than that of the DFO's assessment (Fig. 25, panels A and B). The differences may stem from variations in modeling approaches and discrepancies in the data inputs used for the two computations. It is noted that the DFO assessment incorporated data from two sources, Canadian and Spanish surveys, while the random field model used only Canadian survey information. Observed survey catches align with this decline trend (Fig. 18 panel B). Decline trends can also be observed from estimates of biomass for Spring and Fall DFO surveys. Additionally, stock biomass estimates from the random field model (Fig. 25, panel B) more closely match the biomass calculated for Spring and Fall surveys (Fig. 26, panel A and B) than those estimated by SPM in the DFO assessment report (Fig. 25, panel A).



Figure 24. Plot showing (A) observed and predicted commercial catches estimated by the random field model and (B) observed and predicted commercial catches from DFO's 2023 assessment (data source: Maddock Parsons et al. (2023)).



Figure 25. Plot of (A) stock biomass estimates by DFO stock assessment report (data source: Maddock Parsons et al. (2023)); (B) stock biomass estimates using the random field model.



Figure 26. Plot of (A) estimates of biomass for Spring surveys and (B) estimates of biomass for Fall surveys of Yellowtail Flounder in 3LNO Divisions. Red dots depict data point when surveys did not cover the entire the stock area, resulting in unrepresentative estimates (data source: Maddock Parsons et al. (2023)).

#### Suitability of the random field model for different life history strategies

SPMs, in general, provide a framework for stock assessment by aggregating age and size structures into a single biomass dynamic. They model the net outcome of recruitment, growth, and natural mortality. The practical application of SPMs generally requires a time series of total catches and a corresponding index of relative abundance, often derived from scientific surveys or fishery catchper-unit-effort (CPUE). The effectiveness of traditional, non-spatial SPMs relies on the presence of sufficient contrast in the time series data – ideally showing periods where fishing pressure has reduced the stock and periods of subsequent recovery. This contrast is necessary to reliably estimate the stock's productivity parameters (Hilborn and Walters, 1992b). This requirement can pose challenges for species with slow dynamics (e.g., long-lived species) if the available time series is short or lacks significant biomass fluctuations, as the "surplus" produced each year is small relative to the total biomass. For short-lived species with high productivity and potentially large biomass fluctuations, contrast might be achievable more quickly.

SSPMs extend the traditional framework by explicitly incorporating a spatial dimension, tracking biomass density and associated processes across discrete geographical locations (e.g., grid cells or polygons or triangles). This requires spatially explicit input data, such as geo-referenced survey catches and spatially resolved fishing effort or catch data (i.e., logbooks). A fundamental advantage of SSPMs, demonstrated in this thesis, is their ability to utilize spatial contrast in addition to temporal contrast. By analyzing how density changes differently in areas experiencing different levels of fishing mortality and productivity, SSPMs can gain information about population dynamics even when the temporal fluctuation is limited.

The spatial random field model appears particularly well-suited for longer-lived, slower-growing species that exhibit significant spatial structure either in their own distribution (e.g., due to habitat preferences) or in the fishery targeting them. For species, like the Redfish, Yellowtail Flounder analyzed here or the Big Skate in (Thorson et al., 2017), the lack of strong temporal contrast over assessment periods often hinders traditional SPMs. SSPMs can overcome this by extracting information from spatial contrasts in density and fishing pressure. The handling of movement in the random field model may also be appropriate if movement is relatively localized or diffusive over the annual time step, rather than involving large-scale directed migrations that significantly redistribute biomass across the entire assessed area annually. It is important to clarify that 'movement' in this modeling context refers not to the tracking of individual fish paths, but rather to the net change in the spatial distribution of biomass density across the model grid (i.e., among triangles) between time steps (annually, in this study).

For short-lived species with highly variable dynamics often driven by recruitment pulses, SSPMs could potentially track spatial aspects of large fluctuations (e.g., shifting distributions, localized blooms/depletions). However, high recruitment variability might be difficult to capture solely with the process error term (Myers, 2001). Very rapid or large-scale directed movement (if present) might challenge the random field model.

#### Other aspects

Commercial fishing effort data for the 3LNO Divisions were obtained via the GFW API, as detailed previously (Chapter 2). It is important to acknowledge that this dataset represents the total effort of commercial trawlers operating in the area and could not be filtered to isolate effort specifically targeting Yellowtail Flounder. Incorporating species-specific effort data, should it become available, would be a valuable improvement for future assessments.

Goethel et al. (2015) developed a tag-integrated model as an approach for Yellowtail Flounder spatial stock assessment. Models based on tag-recapture data are often limited by the low recapture rates, such as only 7% of tagged releases being recaptured. While such data can validate mortality estimates, it offers minimal improvements in estimating movement rates (Goethel et al., 2015). In contrast, the random field model demonstrates its superior capability in reconstructing the spatial dynamics of fish stocks over time.

Robertson et al. (2021) utilized VAST model to reveal that Yellowtail Flounder's spatial distribution recovered as water temperatures exceeded previous levels. In addition, it is evident that this species spatial dynamics are strongly correlated with temperature changes by Miller et al. (2016). These findings, combined with the robustness of the random field model, present a compelling case for future research to apply this model to Yellowtail Flounder by integrating

environmental indicators such as bottom temperature, predator and prey densities. A potential approach can be using logbook data (Knapman et al., 2020) and modeling a linear relationship between biomass and environmental factors as presented by Pedersen et al. (2022), such research could provide deeper insights into the species' stock dynamics and enhance understanding of environmental influences on spatial patterns.

A key advantage of integrating such covariates within this spatial random field framework is the potential to model their influence not just temporally, but also spatially, allowing for the investigation of how factors like temperature might exert different effects across various parts of the stock's range, representing a significant advance over traditional non-spatial approaches. Furthermore, future refinements could also involve treating the Spring and Fall surveys as distinct indices within the model; separating these surveys would allow for a more explicit investigation of seasonal dynamics and how they might interact with spatially and temporally varying environmental factors.

## **Chapter 4. Conclusion and future research**

## 1. Performance of the random field model

This study presents a novel development of spatial surplus production models using a random field approach to account for spatial heterogeneity in fisheries stock assessments. The methodologies and findings represent advancements in spatial modeling, particularly in addressing the challenges inherent in movement models. By incorporating the Gaussian random field framework, the study overcomes computational and parameter estimation complexities, offering a robust alternative for assessing the spatiotemporal dynamics of fish populations.

The innovation lies in replacing explicit movement modeling with a flexible random field approach. This method simplifies the modeling of spatial fish stock dynamics by treating the spatial movement as part of the stochastic process error. The random field model estimates spatial variations in productivity and process errors with distinct correlation parameters, reducing the heavy confounding effects observed in the movement model. The movement model often faces convergence issues due to computational demands and parameter identifiability, which arise from the need for matrix exponentiation (or approximations like Euler's method) and the potential for confounding between movement and process error parameters. The random field model exhibited reliable parameter estimation and high convergence rates across all tested scenarios.

The method of using triangular spatial grids by employing the R-INLA package optimally balances spatial resolution and computational cost, providing a scalable framework adaptable to various fisheries contexts. Simulation results demonstrate the superiority of the random field model over both non-spatial and movement models. The model consistently produces higher correlations between simulated and estimated population densities, alongside improved AIC and BIC scores. These outcomes validate the random field model's ability to reconstruct spatial fish population distributions accurately.

### 2. Insights from case studies and broader applicability

#### Redfish in NAFO 3LN divisions

The model confirmed significant spatial heterogeneity, with biomass concentrated in deeper waters (200-500m). Crucially, it revealed a concerning decline in estimated biomass between 2012 and 2019, aligning with survey trends and increasing fishing pressure reported elsewhere. This finding contrasts with previous non-spatial assessments (ASPIC) which failed to capture this decline, underscoring the management need for spatially explicit assessments for this stock complex. The results suggest that management might consider monitoring spatial fishing patterns.

#### Yellowtail Flounder in NAFO 3LNO divisions

The assessment indicated relatively stable spatial biomass distributions from 2012-2019, primarily concentrated in the central and southeastern regions, consistent with the species' known preference for shallower waters (<200m). The model effectively captured these dynamics, aligning well with survey indices. Refining the model using a quadratic dispersion relationship for survey catches improved the fit, demonstrating the model's flexibility. While the overall biomass showed a slight decline, spatial stability suggests management can continue focusing on important areas, but incorporating environmental factors known to influence Yellowtail distribution (like temperature) could further enhance assessment accuracy.

Applicability to other stocks

The random field approach holds strong potential for application to other regional fish stocks, particularly where spatially resolved data—such as survey or commercial effort and catch data—are available, and where spatial structure is evident. This is especially relevant for species known to exhibit aggregation behavior or distinct habitat preferences, such as Greenland Halibut and Atlantic Cod, where spatial dynamics are well documented.

### 3. Considerations for Application

Data Availability: The model requires, at a minimum, spatially explicit survey and commercial catch or effort data (ideally from fishing logbooks). When logbook data are unavailable, alternative sources such as Global Fishing Watch data, as used in this study, can be effective substitutes. The quality, spatial resolution, and time-series length of these datasets are critical to the model's performance. Spatial contrast makes it possible to apply surplus production models to shorter time series than previously expected in spatially aggregated surplus production models.

Population Biology: The model used in this study accounts for spatial heterogeneity in population dynamics on an annual time step. This may present limitations for species exhibiting highly directed, large-scale migrations or distinct seasonal movement patterns.

### 4. Relevance to spatial management

Mapping hotspots: The spatially explicit biomass estimates (e.g., Fig. 14 and 21) directly identify areas of high density, potential refugia, or areas experiencing decline, guiding spatial conservation and management efforts.

Evaluating spatial closures: The framework can be used in a simulation capacity to evaluate the potential consequences of closing specific areas to fishing.

Informing effort allocation: Understanding the spatial distribution of biomass and productivity can help inform decisions about spatially varying quotas or effort controls, potentially directing effort away from vulnerable components of the stock.

### 5. Challenges and future directions

Commercial catch data: The reliance on publicly available GFW effort data, while innovative, is less precise compared to detailed, haul-by-haul spatial data from commercial logbooks. Access to and incorporation of such data, where possible, would significantly improve accuracy.

Underreported catches: The discrepancy noted for Yellowtail Flounder between model-predicted and observed landings highlights the challenge of underreporting. Future work could implement censored likelihood approaches, treating reported landings as lower bounds, to estimate true removals more reliably.

Environmental drivers: While the current model captures spatial heterogeneity, explicitly incorporating key environmental covariates (e.g., bottom temperature, depth, habitat type, prey/predator fields) known to influence fish distribution and productivity could further enhance realism and explanatory power.

Concluding Statement: This thesis contributes a valuable advancement in spatial stock assessment by providing a robust, scalable, and reliable model framework. It effectively addresses key limitations associated with explicit movement modeling, particularly parameter identifiability and computational burden. The successful applications to Redfish and Yellowtail Flounder demonstrate potential to generate spatially explicit insights relevant to sustainable management.

## Reference

- ABARES 2021. Fishery status reports 2021. Australian Bureau of Agricultural and Resource Economics and Sciences (ABARES).
- AEBERHARD, W. H., MILLS FLEMMING, J. & NIELSEN, A. 2018. Review of State-Space Models for Fisheries Science. *Annual Review of Statistics and Its Application*, 5, 215-235.
- AHRESTANI, F. S., HEBBLEWHITE, M. & POST, E. 2013. The importance of observation versus process error in analyses of global ungulate populations. *Scientific Reports*, 3, 3125.
- AKAIKE, H. 1981. Likelihood of a model and information criteria. *Journal of Econometrics*, 16, 3-14.
- AKAIKE, H. 1983. Information measures and model selection. *Bulletin of the International Statistical Institute*. Bulletin of the International Statistical Institute.
- APOSTOLAKI, P., MILNER-GULLAND, E. J., MCALLISTER, M. K. & KIRKWOOD, G. P. 2002. Modelling the effects of establishing a marine reserve for mobile fish species. *Canadian journal of fisheries and aquatic sciences*, 59, 405-415.
- ÁVILA DE MELO, A. M., NUNO, B., ALPOIM, R., GONZÁLEZ TRONCOSO, D.,

GONZÁLEZ, F. & POCHTAR, M. 2020. Scientific council meeting – June 2020: The status of redfish (*S. mentella* and *S. fasciatus*) in Divisions 3LN and two medium term scenarios (when recruitment is low, Risk Based Management Strategy or common sense?). Northwest Atlantic Fisheries Organization.

- BABYN, J., VARKEY, D., REGULAR, P., INGS, D. & MILLS FLEMMING, J. 2021. A gaussian field approach to generating spatial age length keys. *Fisheries research*, 240, 105956.
- BARANOV, F. I. 1918. On the question of the biological basis of fisheries, Nauchnye Issled, Ikhriologicheskii Ins.
- BASTARDIE, F., NIELSEN, J. R., ULRICH, C., EGEKVIST, J. & DEGEL, H. 2010. Detailed mapping of fishing effort and landings by coupling fishing logbooks with satelliterecorded vessel geo-location. *Fisheries research*, 106, 41-53.
- BENESTAN, L. M., ROUGEMONT, Q., SENAY, C., NORMANDEAU, E., PARENT, E.,
  RIDEOUT, R., BERNATCHEZ, L., LAMBERT, Y., AUDET, C. & PARENT, G. J.
  2021. Population genomics and history of speciation reveal fishery management gaps in
  two related redfish species (*Sebastes mentella* and *S. fasciatus*). *Evolutionary Applications*, 14, 588-606.
- BERG, C., COLEMAN, P., COOPER, A., HANSEN, H. Ø., HASLOB, H., HERRARIZ, I. G., KOKKALIS, A., MILDENBERGER, T., MOURA, T. & PENNINO, M. G. 2021.
  Benchmark Workshop on the development of MSY advice for category 3 stocks using Surplus Production Model in Continuous Time; SPiCT (WKMSYSPiCT).

BERGER, A. M., GOETHEL, D. R., LYNCH, P. D., QUINN, T., II, MORMEDE, S.,
 MCKENZIE, J. & DUNN, A. 2017. Space oddity: The mission for spatial integration.
 *Canadian journal of fisheries and aquatic sciences*, 74, 1698-1716.

- BERKELEY, S. A., HIXON, M. A., LARSON, R. J. & LOVE, M. S. 2004. Fisheries Sustainability via Protection of Age Structure and Spatial Distribution of Fish Populations. *Fisheries (Bethesda)*, 29, 23-32.
- BEVERTON, R. J. H. & HOLT, S. J. 1957. On the dynamics of exploited fish populations. Fish. Invest. *London Ser*, 2, 1-533.
- BOUSQUET, N., CADIGAN, N., DUCHESNE, T., RIVEST, L.-P. & QUINN, T. 2010.
   Detecting and correcting underreported catches in fish stock assessment: trial of a new method. *Canadian journal of fisheries and aquatic sciences*, 67, 1247-1261.
- BOWERING, W. R. & BRODIE, W. B. 1991. Distribution of commercial flatfishes in the Newfoundland-Labrador region of the Canadian Northwest Atlantic and changes in certain biological parameters since exploitation. *Netherlands journal of sea research*, 27, 407-422.
- BRANCH, T. A., HILBORN, R., HAYNIE, A. C., FAY, G., FLYNN, L., GRIFFITHS, J.,
  MARSHALL, K. N., RANDALL, J. K., SCHEUERELL, J. M., WARD, E. J. &
  YOUNG, M. 2006. Fleet dynamics and fishermen behavior: lessons for fisheries
  managers. *Canadian journal of fisheries and aquatic sciences*, 63, 1647-1668.
- CADIGAN, N., ROBERTSON, M. D., NIRMALKANNA, K. & ZHENG, N. 2022. The complex relationship between weight and length of Atlantic cod off the south coast of Newfoundland. *Canadian journal of fisheries and aquatic sciences*, 79, 1798-1819.
- CADIGAN, N. G. 2016. A state-space stock assessment model for northern cod, including underreported catches and variable natural mortality rates. *Canadian journal of fisheries and aquatic sciences*, 73, 296-308.

- CADIGAN, N. G. & CHEN, Y. 2011. Confidence intervals for trawlable abundance from stratified-random bottom trawl surveys. *Canadian journal of fisheries and aquatic sciences*, 68, 781-794.
- CADIGAN, N. G. & MARSHALL, C. T. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. *Canadian journal of fisheries and aquatic sciences*, 73, 296-308.
- CADIGAN, N. G., WADE, E. & NIELSEN, A. 2017. A spatiotemporal model for snow crab (*Chionoecetes opilio*) stock size in the southern Gulf of St. Lawrence. *Canadian Journal* of Fisheries & Aquatic Sciences, 74, 1808-1820.
- CADIGAN, N. G., WEERASEKERA, S. J. W. W. M. M. P., REGULAR, P. M. & RIDEOUT, R.
  M. 2024. Time varying M with starvation mortality in a state-space stock assessment
  model: Part 2: Atlantic cod (Gadus morhua) on the southern Grand Bank of
  Newfoundland. *Fisheries research*, 280, 107174.
- CADRIN, S. X. 2020. Defining spatial structure for fishery stock assessment. *Fisheries research*, 221, 105397.
- CADRIN, S. X., FRIEDLAND, K. D. & WALDMAN, J. R. 2004. *Stock identification methods: applications in fishery science*, Burlington, MA, Elsevier Academic Press.
- CADRIN, S. X., MAUNDER, M. N. & PUNT, A. E. 2020. Spatial Structure: Theory, estimation and application in stock assessment models. *Fisheries research*, 229, 105608.
- CADRIN, S. X. & SECOR, D. H. 2009. Accounting for Spatial Population Structure in Stock Assessment: Past, Present, and Future. *In:* BEAMISH, R. J. & ROTHSCHILD, B. J.

(eds.) *The Future of Fisheries Science in North America*. Dordrecht: Springer Netherlands.

- CARNEIRO, A. P. B., CLARK, B. L., PEARMAIN, E. J., CLAVELLE, T., WOOD, A. G. & PHILLIPS, R. A. 2022. Fine-scale associations between wandering albatrosses and fisheries in the southwest Atlantic Ocean. *Biological Conservation*, 276, 109796.
- CHEUNG, W. W. L., LAM, V. W. Y., SARMIENTO, J. L., KEARNEY, K., WATSON, R. E.
  G., ZELLER, D. & PAULY, D. 2010. Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global change biology*, 16, 24-35.
- CIANNELLI, L., DINGSØR, G. E., BOGSTAD, B., OTTERSEN, G., CHAN, K.-S., GJØSÆTER, H., STIANSEN, J. E. & STENSETH, N. C. 2007. Spatial anatomy of species survival: effects of predation and climate-driven environmental variability. *Ecology (Durham)*, 88, 635-646.
- COSTELLO, C., OVANDO, D., HILBORN, R., GAINES, S. D., DESCHENES, O. & LESTER, S. E. 2012. Status and Solutions for the World's Unassessed Fisheries. *Science (American Association for the Advancement of Science)*, 338, 517-520.
- COWEN, R. K., PARIS, C. B. & SRINIVASAN, A. 2006. Scaling of Connectivity in Marine Populations. Science (American Association for the Advancement of Science), 311, 522-527.
- DENNIS, B. & TAPER, M. L. 1994. Density Dependence in Time Series Observations of Natural Populations: Estimation and Testing. *Ecological monographs*, 64, 205-224.

- DFO. 2024. Canada's marine protected and conserved areas. [Online]. Available: <u>https://www.dfo-mpo.gc.ca/oceans/conservation/areas-zones/index-eng.html</u> [Accessed December 15 2024].
- DUNN, P. K. & SMYTH, G. K. 1996. Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics*, 5, 236-244.
- DWYER, K. S., WALSH, S. J. & CAMPANA, S. E. 2003. Age determination, validation and growth of Grand Bank yellowtail flounder (Limanda ferruginea). *ICES Journal of Marine Science*, 60, 1123-1138.

ELTON, C. S. 1927. Animal ecology, London, Sidgwick & Jackson, ltd.

- FLETCHER, I. F. On the restructuring of the Pella-Tomlinson system. 1978. Fishery Bulletin, 515-521.
- FOGARTY, M. J. & BOTSFORD, L. W. 2007. Population Connectivity and Spatial Management of Marine Fisheries. *Oceanography (Washington, D.C.)*, 20, 12-12.
- FOSTER, S. D. & BRAVINGTON, M. V. 2013. A Poisson–Gamma model for analysis of ecological non-negative continuous data. *Environmental and Ecological Statistics*, 20, 533-552.
- FOURNIER, D. A., SKAUG, H. J., ANCHETA, J., IANELLI, J., MAGNUSSON, A., MAUNDER, M. N., NIELSEN, A. & SIBERT, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization methods & software*, 27, 233-249.

- FOX, W. 1970. An exponential surplus-yield model for optimizing exploited fish populations. *Transactions of The American Fisheries Society - TRANS AMER FISH SOC*, 99, 80-88.
- FRANK, K. T., PETRIE, B., SHACKELL, N. L. & CHOI, J. S. 2006. Reconciling differences in trophic control in midalatitude marine ecosystems. *Ecology letters*, 9, 1096-1105.
- GASCON, D. 2003. Redfish multidisciplinary research zonal program (1995-1998): final report.
- GIDEON, S. 1978. Estimating the Dimension of a Model. The Annals of Statistics, 6, 461-464.
- GOETHEL, D. R., LEGAULT, C. M. & CADRIN, S. X. 2014. Testing the performance of a spatially explicit tag-integrated stock assessment model of yellowtail flounder (Limanda ferruginea) through simulation analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 72, 582-601.
- GOETHEL, D. R., LEGAULT, C. M. & CADRIN, S. X. 2015. Demonstration of a spatially explicit, tag-integrated stock assessment model with application to three interconnected stocks of yellowtail flounder off of New England. *ICES Journal of Marine Science*, 72, 164-177.
- GOETHEL, D. R., QUINN, T. J. & CADRIN, S. X. 2011. Incorporating Spatial Structure in Stock Assessment: Movement Modeling in Marine Fish Population Dynamics. *Reviews in fisheries science*, 19, 119-136.

GONZÁLEZ HERRAIZ, I., VILA, Y., CARDINALE, M., BERG, C. W., WINKER, H.,
AZEVEDO, M., MILDENBERGER, T. K., KOKKALIS, A., VÁZQUEZ VILAMEA, A.
A. & MORLÁN, R. 2023. First Maximum Sustainable Yield advice for the Nephrops
norvegicus stocks of the Northwest Iberian coast using stochastic Surplus Production
model in Continuous Time (SPiCT). *Frontiers in marine science*, 10, 1062078.

- GOULD, W. R. & POLLOCK, K. H. 1997. Catch-effort maximum likelihood estimation of important population parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 54, 890-897.
- GRAHAM, M. 1935. Modern Theory of Exploiting a Fishery, and Application to North Sea Trawling. *ICES Journal of Marine Science*, 10, 264-274.

GRANT, S. M. 2004. Habitat requirements of marine finfish species occurring

- in the newfoundland region. LGL Limited environmental research associates, P.O. Box 13248, Stn. A, St. John's, NF A1B 4A5.
- GUDMUNDSSON, G. 1994. Time Series Analysis of Catch-At-Age Observations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43, 117-126.
- GUNDERSON, D. R. 1993. Surveys of fisheries resources, New York, NY : John Wiley & Sons, Inc.
- HADDON, M. 2011. *Modelling and quantitative methods in fisheries*, Boca Raton, Boca Raton : CRC Press.
- HILBORN, R. & WALTERS, C. J. 1992a. *Quantitative Fisheries Stock Assessment Choice, Dynamics and Uncertainty,* New York, NY, Springer US.
- HILBORN, R. & WALTERS, C. J. 1992b. *Quantitative fisheries stock assessment: choice, dynamics, and uncertainty,* Boston, MA, Springer.
- HINTZEN, N. T., BASTARDIE, F., BEARE, D., PIET, G. J., ULRICH, C., DEPORTE, N., EGEKVIST, J. & DEGEL, H. 2012. VMStools: Open-source software for the processing,

analysis and visualisation of fisheries logbook and VMS data. *Fisheries research*, 115, 31-43.

- HOENIG, J. M., WARREN, W. G. & STOCKER, M. 1994. Bayesian and Related Approaches to Fitting Surplus Production Models. *Canadian Journal of Fisheries and Aquatic Sciences*, 51, 1823-1831.
- HOLLOWED, A. B., BARANGE, M., BEAMISH, R. J., BRANDER, K., COCHRANE, K., DRINKWATER, K., EMAN, M. G. G., HARE, J. A., HOLT, J., ITO, S.-I., KIM, S., KING, J. R., LOENG, H., MACKENZIE, B. R., MUETER, F. J., OKEY, T. A., PECK, M. A., RADCHENKO, V. I., RICE, J. C., SCHIRRIPA, M. J., YATSU, A. & YAMANAKA, Y. 2013. Projected impacts of climate change on marine fish and fisheries. *ICES journal of marine science*, 70, 1023-1037.
- JENSEN, A. L. 1984. Logistic Surplus-Production Model with Explicit Terms for Growth, Mortality, and Recruitment. *Transactions of the American Fisheries Society*, 113, 617-626.
- JENSEN, A. L. 2005. Harvest in a fluctuating environment and conservative harvest for the Fox surplus production model. *Ecological Modelling*, 182, 1-9.
- JOHNSON, S. D. N. & COX, S. P. 2021. Hierarchical surplus production stock assessment models improve management performance in multi-species, spatially-replicated fisheries. *Fisheries Research*, 238, 105885.
- JØRGENSEN, B. 1997. The theory of dispersion models, London : Chapman and Hall.
- KALMAN, R. E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82, 35-45.

- KALMAN, R. E. & BUCY, R. S. 1961. New Results in Linear Filtering and Prediction Theory. *Journal of fluids engineering*, 83, 95-108.
- KARIM, E., QUN, L., HASAN, S. J., ALI, M. Z., HOQ, M. E. & MAHMUD, M. Y. 2020.
  Maximum sustainable yield estimates of marine captured shrimp fishery of the Bay of Bengal, Bangladesh by using surplus production model. *Thalassas: An International Journal of Marine Sciences*, 36, 471-480.
- KEMP, M. C. & KIMURA, Y. 1978. Introduction to Mathematical Economics, New York, Springer. pp 102–114.
- KENCHINGTON, T. J. 1991. Vertical Distribution and Movements of Larval Redfishes (Sebastes spp.) in the Southern Gulf of St. Lawrence. Journal of Northwest Atlantic fishery science, 11, 43-49.
- KERR, L. A., HINTZEN, N. T., CADRIN, S. X., CLAUSEN, L. W., DICKEY-COLLAS, M., GOETHEL, D. R., HATFIELD, E. M. C., KRITZER, J. P. & NASH, R. D. M. 2017.
  Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES journal of marine science*, 74, 1708-1722.
- KIMOTO, A., ORTIZ, M. & TAYLOR, N. 2023. REVIEW OF STOCHASTIC SURPLUS PRODUCTION MODEL IN CONTINUOUS TIME (SPICT) METHODOLOGY FOR THE ICCAT SOFTWARE CATALOGUE. *Collect. Vol. Sci. Pap. ICCAT*, 80, 80-84.
- KING, M. G. 2007. *Fisheries biology, assessment, and management*, Oxford ; Ames, Iowa : Blackwell Pub., c2007. 2nd ed.

- KNAPMAN, P., COOK, R. & BLYTH-SKYRME, R. 2020. OCI Grand Bank Yellowtail Flounder Trawl: Public Comment Draft Report June 2020.
- KRISTENSEN, K., NIELSEN, A., BERG, C. W., SKAUG, H. & BELL, B. M. 2016. TMB: Automatic Differentiation and Laplace Approximation. *Journal of statistical software*, 70, 1-21.
- KUMAR, R., CADIGAN, N. G., ZHENG, N., VARKEY, D. A. & MORGAN, M. J. 2020. A state-space spatial survey-based stock assessment (SSURBA) model to inform spatial variation in relative stock trends. *Canadian journal of fisheries and aquatic sciences*, 77, 1638-1658.
- LINDGREN, F. & RUE, H. 2015a. Bayesian Spatial Modelling with R-INLA. *Journal of Statistical Software*, 63, 1 - 25.
- LINDGREN, F. & RUE, H. 2015b. Bayesian Spatial Modelling with R INLA. *Journal of statistical software*, 63, 1-25.
- LINDGREN, F., RUE, H., LINDSTRÖM, J., KENT, J. T., DIGGLE, P. J., ILLIAN, J. B., SIMPSON, D. P., GNEITING, T., SCHEUERER, M., FURRER, R., FURRER, E., NYCHKA, D., FEARNHEAD, P., CHALLENOR, P., ANDRIANAKIS, Y., STEPHENSON, G., MØLLER, J., HU, X., SIMPSON, D., BOLIN, D., BROWN, P. E., CAMELETTI, M., MARTINO, S., COOLEY, D., HOETING, J. A., CRUJEIRAS, R. M., PRIETO, A., FERREIRA, M. A. R., FUGLSTAD, G.-A., GELMAN, A., GUTTORP, P., DAS, B., HAALAND, B., NOTT, D. J., HASLETT, J., JOSHI, C., GARRETA, V., HÖHLE, M., IPPOLITI, L., MARTIN, R. J., BHANSALI, R. J., JANDHYALA, V. K., FOTOPOULOS, S. B., JUN, M., KONGSGÅRD, H. W., LASINIO, G. J., POLLICE, A.,

LEE, C., LEE, W. T., KAUFMAN, C. G., LI, B., GENTON, M. G., LINDGREN, G., MARDIA, K. V., MATEU, J., MONDAL, D., MÜLLER, W. G., WALDL, H., OTTAVI, A., PAPASPILIOPOULOS, O., PORCU, E., SAEZ, M., SCHMIDT, A. M., STEIN, A., SWITZER, P., TURKMAN, K., WIKLE, C. K. & HOOTEN, M. B. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach [with Discussion]. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 73, 423-498.

- LORENZEN, K. 2016. Toward a new paradigm for growth modeling in fisheries stock assessments: Embracing plasticity and its consequences. *Fisheries research*, 180, 4-22.
- LUCET, V. & PEDERSEN, E. J. 2023. The sspm R package for spatially-explicit surplus production population models. *Journal of open source software*, 8, 4724.
- MADDOCK PARSONS, D., MORGAN, M. J. & ROGERS, R. 2018. Assessment of Yellowtail Flounder in NAFO Divisions 3LNO using a new Stock Production Model in Bayesian Framework. *NAFO SCR Doc. 18/038, Serial No.N6828.*
- MADDOCK PARSONS, D., SKANES KATHERINE, R. & RIDEOUT RICK, M. 2023. 2023 Assessment of Yellowtail Flounder in NAFO Divisions 3LNO.
- MAUNDER, M. N. & PUNT, A. E. 2013. A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142, 61-74.
- MCDONALD, R. R., KEITH, D. M., SAMEOTO, J. A., HUTCHINGS, J. A. & FLEMMING, J.
   M. 2021. Explicit incorporation of spatial variability in a biomass dynamics assessment model. *ICES Journal of Marine Science*, 78, 3265-3280.

- MCDONALD, R. R., KEITH, D. M., SAMEOTO, J. A. & MILLS FLEMMING, J. 2023. Integrating habitat features into spatio-temporal biomass dynamics models for a better understanding of stock productivity: a case study of sea scallop in the Bay of Fundy. *ICES journal of marine science*, 80, 1710-1726.
- MILLER, T. J., HARE, J. A. & ALADE, L. A. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern New England yellowtail flounder. *Canadian Journal of Fisheries and Aquatic Sciences*, 73, 1261-1270.
- MORMEDE, S., DUNN, A., PARKER, S. & HANCHET, S. 2017. Using spatial population models to investigate the potential effects of the Ross Sea region Marine Protected Area on the Antarctic toothfish population. *Fisheries Research*, 190, 164-174.
- MUETER, F. J. & MEGREY, B. A. 2006. Using multi-species surplus production models to estimate ecosystem-level maximum sustainable yields. *Fisheries Research*, 81, 189-201.
- MURUA, H. & SABORIDO-REY, F. 2003. Female Reproductive Strategies of Marine Fish Species of the North Atlantic. *Journal of Northwest Atlantic fishery science*, 33, 23-31.
- MYERS, R. A. 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. *ICES journal of marine science*, 58, 937-951.
- NAFO 2024. Yellowtail Flounder (Limanda ferruginea) in Divisions 3L, 3N and 3O. *NAFO SCS Doc. 18-19, Serial No. N6849.* Halifax, Nova Scotia.
- NOAA. 2024. Yellowtail Flounder [Online]. NOAA Fisheries. Available: https://www.fisheries.noaa.gov/species/yellowtail-flounder [Accessed October, 24 2024].

OLLERHEAD, L. M. N., CANADA. DEPARTMENT OF, F., OCEANS, CANADA.

DEPARTMENT OF, F. & OCEANS, S. 2009. *Mapping spawning times and locations for ten commercially important fish species found on the Grand Banks of Newfoundland,* Ottawa, Ont, Majesty the Queen in Right of Canada.

- OLMOS, M., CAO, J., THORSON, J. T., PUNT, A. E., MONNAHAN, C. C., ALGLAVE, B. & SZUWALSKI, C. 2023. A step towards the integration of spatial dynamics in population dynamics models: Eastern Bering Sea snow crab as a case study. *Ecological Modelling*, 485, 110484.
- PANHWAR, S. K., LIU, Q., AMIR, S. A. & KALHORO, M. A. 2012. Performance comparison between logistic and generalized surplus-production models applied to the Sillago sihama fishery in Pakistan. *Journal of Ocean University of China*, 11, 401-407.
- PEDERSEN, E. J., SKANES, K. R., NICOLAS LE, C., KOEN-ALONSO, M. & BAKER, K. D. 2022. A New Spatial Ecosystem-Based Surplus Production Model for Northern Shrimp in Shrimp Fishing Areas 4 to 6.
- PEDERSEN, M. W. & BERG, C. W. 2017. A stochastic surplus production model in continuous time. *Fish and Fisheries*, 18, 226-243.
- PELLA, J. & TOMLINSON, P. K. 1969. A generalized stock production model = Un modelo generalizado de la producción del stock, La Jolla, Calif, Inter-American Tropical Tuna Commission.
- PELLETIER, D. & MAHÉVAS, S. 2005. Spatially explicit fisheries simulation models for policy evaluation. *Fish and fisheries (Oxford, England),* 6, 307-349.

- PEREIRA, J., SCHULTZ, E. & AUSTER, P. 2012. Geospatial analysis of habitat use in yellowtail flounder Limanda ferruginea on Georges Bank. *Marine Ecology Progress Series*, 468, 279-290.
- PERREAULT, A., WHEELAND, L., MORGAN, M. & CADIGAN, N. 2020a. A state-space stock assessment model for American plaice on the Grand Bank of Newfoundland.
- PERREAULT, A. M. J., WHEELAND, L. J., MORGAN, M. J. & CADIGAN, N. G. 2020b. A state-space stock assessment model for American plaice on the Grand Bank of Newfoundland. *Journal of Northwest Atlantic fishery science*, 51, 45-104.
- PERRY, R. I. & SMITH, S. J. 1994. Identifying Habitat Associations of Marine Fishes Using Survey Data: An Application to the Northwest Atlantic. *Canadian journal of fisheries and aquatic sciences*, 51, 589-602.
- PRAGER, M. 1994. A Suite of Extensions to a Nonequilibrium Surplus-Production Model. *Fishery Bulletin*, 92, 374-389.
- PRAGER, M. H. 2002. Comparison of logistic and generalized surplus-production models applied to swordfish, Xiphias gladius, in the north Atlantic Ocean. *Fisheries Research*, 58, 41-57.
- PRAGER, M. H. 2016. User's Guide for ASPIC Suite, version 7: A Stock–Production Model Incorporating Covariates and auxiliary programs [Online]. Available: <u>http://www.mhprager.com/aspic7-guide.pdf</u> [Accessed].
- PUNT, A. E. 2019. Spatial stock assessment methods: A viewpoint on current issues and assumptions. *Fisheries research*, 213, 132-143.

- PUNT, A. E., DUNN, A., ELVARSSON, B. Þ., HAMPTON, J., HOYLE, S. D., MAUNDER, M. N., METHOT, R. D. & NIELSEN, A. 2020. Essential features of the next-generation integrated fisheries stock assessment package: A perspective. *Fisheries research*, 229, 105617.
- PUNT, A. E. & HILBORN, R. A. Y. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. *Reviews in fish biology and fisheries*, 7, 35-63.
- PUNT, A. E. & SZUWALSKI, C. 2012. How well can FMSY and BMSY be estimated using empirical measures of surplus production? *Fisheries Research*, 134-136, 113-124.
- QUINN, T. J. & DERISO, R. B. 1999. *Quantitative fish dynamics*, New York, Oxford University Press.
- R-CORE-TEAM 2022. R: A language and environment for statistical computing.
- REGULAR, P. M., BUREN, A. D., DWYER, K. S., CADIGAN, N. G., GREGORY, R. S.,
  KOEN-ALONSO, M., RIDEOUT, R. M., ROBERTSON, G. J., ROBERTSON, M. D.,
  STENSON, G. B., WHEELAND, L. J. & ZHANG, F. 2022. Indexing starvation mortality
  to assess its role in the population regulation of Northern cod. *Fisheries research*, 247, 106180.
- RIVARD, D. & BLEDSOEI, L. J. Parameter estimation for the Pella-tomlinson stock production model under nonequilibrium conditions. 1978.
- ROBERTSON, M. D., CADIGAN, N. G., NIRMALKANNA, K. & WHEELAND, L. J. 2025. Impact of starvation-induced mortality on the collapse and lack of recovery of American plaice on the Newfoundland Grand Banks. *Canadian journal of fisheries and aquatic sciences*, 82, 1-14.

- ROBERTSON, M. D., GAO, J., REGULAR, P. M., MORGAN, M. J. & ZHANG, F. 2021. Lagged recovery of fish spatial distributions following a cold-water perturbation. *Scientific Reports*, 11, 9513.
- ROGER, B., VIRGILIO, G.-R. & HÅVARD, R. 2015. Spatial Data Analysis with R-INLA with Some Extensions. *Journal of statistical software*, 63, 1-31.
- ROGERS, B., ANDREA, M. J. P., SIMPSON, M. & VARKEY, D. A. 2022. Assessment of 3LN redfish using the ASPIC model in 2022 (Sebastes mentella and S. fasciatus).
- ROQUES, S., SÉVIGNY, J.-M. & BERNATCHEZ, L. 2001. Evidence for broadscale introgressive hybridization between two redfish (genus *Sebastes*) in the North-west Atlantic: a rare marine example. *Molecular ecology*, 10, 149-165.
- RUSSELL, E. S. 1931. Some theoretical Considerations on the "Overfishing" Problem. *ICES* Journal of Marine Science, 6, 3-20.
- SALA, E., MAYORGA, J., BRADLEY, D., CABRAL, R. B., ATWOOD, T. B., AUBER, A., CHEUNG, W., COSTELLO, C., FERRETTI, F., FRIEDLANDER, A. M., GAINES, S.
  D., GARILAO, C., GOODELL, W., HALPERN, B. S., HINSON, A., KASCHNER, K., KESNER-REYES, K., LEPRIEUR, F., MCGOWAN, J., MORGAN, L. E., MOUILLOT,
  D., PALACIOS-ABRANTES, J., POSSINGHAM, H. P., RECHBERGER, K. D., WORM, B. & LUBCHENCO, J. 2021. Protecting the global ocean for biodiversity, food and climate. *Nature*, 592, 397-402.
- SCHAEFER, M. B. 1954. Some Aspects of the Dynamics of Populations Important to the Management of the Commercial Marine Fisheries, Inter-American Tropical Tuna Commission.

- SULLIVAN, P. J. 1992. A Kalman Filter Approach to Catch-at-Length Analysis. *Biometrics*, 48, 237-257.
- THIAW, M., GASCUEL, D., JOUFFRE, D. & THIAW, O. T. 2009. A surplus production model including environmental effects: Application to the Senegalese white shrimp stocks. *Progress in Oceanography*, 83, 351-360.
- THORSON, J. T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research*, 210, 143-161.
- THORSON, J. T. 2022. VAST model structure and user interface [Online]. James-Thorson-NOAA/VAST. Available: <u>https://github.com/James-Thorson-</u>

NOAA/VAST/blob/main/manual/VAST\_model\_structure.pdf [Accessed Jan 10th 2022].

- THORSON, J. T., JANNOT, J. & SOMERS, K. 2017. Using spatio-temporal models of population growth and movement to monitor overlap between human impacts and fish populations. *Journal of Applied Ecology*, 54, 577-587.
- THORSON, J. T. & KRISTENSEN, K. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research*, 175, 66-74.
- THORSON, J. T., SHELTON, A. O., WARD, E. J. & SKAUG, H. J. 2015a. Geostatistical deltageneralized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science*, 72, 1297-1310.
- THORSON, J. T., SKAUG, H. J., KRISTENSEN, K., SHELTON, A. O., WARD, E. J., HARMS, J. H. & BENANTE, J. A. 2015b. The importance of spatial models for estimating the strength of density dependence. *Ecology (Durham)*, 96, 1202-1212.
- UHLER, R. S. 1980. Least Squares Regression Estimates of the Schaefer Production Model: Some Monte Carlo Simulation Results. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 1284-1294.
- VAN BEVEREN, E., DUPLISEA, D., CASTONGUAY, M., DONIOL-VALCROZE, T., PLOURDE, S. & CADIGAN, N. 2017. How catch underreporting can bias stock assessment of and advice for northwest Atlantic mackerel and a possible resolution using censored catch. *Fisheries research*, 194, 146-154.
- WALSH, S. J. 1992. Factors influencing distribution of juvenile yellowtail flounder (Limanda ferruginea) on the grand bank of Newfoundland. *Netherlands Journal of Sea Research*, 29, 193-203.
- WALTERS, C. 2003. Folly and fantasy in the analysis of spatial catch rate data. *Canadian journal of fisheries and aquatic sciences,* 60, 1433-1436.
- WALTERS, C. & LUDWIG, D. 1994. Calculation of Bayes Posterior Probability Distributions for Key Population Parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 51, 713-722.
- WALTERS, C., PAULY, D., PITCHER, T. J. & HART, P. J. B. 1998. Designing fisheries management systems that do not depend upon accurate stock assessment.

- WANG, J., CHEN, X., STAPLES, K. W. & CHEN, Y. 2018. A stock assessment for *Illex argentinus* in Southwest Atlantic using an environmentally dependent surplus production model. *Acta Oceanologica Sinica*, 37, 94-101.
- WHITE, T. D., CARLISLE, A. B., KROODSMA, D. A., BLOCK, B. A., CASAGRANDI, R., DE LEO, G. A., GATTO, M., MICHELI, F. & MCCAULEY, D. J. 2017. Assessing the effectiveness of a large marine protected area for reef shark conservation. *Biological Conservation*, 207, 64-71.
- WOODS, P. & JONASSON, J. P. 2017. Bayesian hierarchical surplus production model of the common whelk Buccinum undatum in Icelandic waters. *Fisheries Research*, 194, 117-128.
- YOSHIMOTO, S. S. & CLARKE, R. P. 1993. Comparing Dynamic Versions of the Schaefer and Fox Production Models and Their Application to Lobster Fisheries. *Canadian journal of fisheries and aquatic sciences*, 50, 181-189.
- ZHANG, Z. 2013. Evaluation of logistic surplus production model through simulations. *Fisheries Research*, 140, 36-45.

## Appendix

List of mathematical notations

Parameters	Description		
β	The density dependence	Fixed effect	
α	The average density	Fixed effect	
Kε	The geographical range of correlations for process error	Fixed effect	
K <sub>w</sub>	The geographical range of correlations for spatial productivity	Fixed effect	
$\delta_{arepsilon}$	The standard deviation of spatiotemporal process errors	Fixed effect	
$\delta_{\omega}$	The standard deviation of spatial productivity	Fixed effect	
$\mu_0$	The initial stock biomass	Fixed effect	
q	The catchability coefficient	Fixed effect	
φ	The positive dispersion parameter of the Tweedie distribution	Fixed effect	
р	The power parameter of the Tweedie distribution	Fixed effect	
$\sigma_{catch}$	The standard deviation of catch	Fixed effect	
b <sub>r,t</sub>	Matrix of population biomass at spatial triangle $r^{th}$ in time $t$	Random effect	
$\omega_g$	Represents the spatial variation in productivity	Random effect	



Figure S1. Simulation results for (A) random field, (B) non-spatial models including parameter estimator distributions of the Tweedie power parameter using (1) p = 1.5 and (2) p = 1.7. Red lines are the true values of the parameters and blue lines are means of parameter estimates.



Figure S2. Plot of estimated versus observed survey catches for each year (in log scale). The x axis represents the estimated survey catches, the y axis represents the observed survey catches, and the red line has a slope of one through the origin. The bias correction feature in TMB was applied to predict survey catches.



Figure S3. Plot of convergence rate comparison between the simulation of random field and movement models using Tweedie power parameter: (1) p=1.5 and (2) p=1.7. The convergence rate for random field and movement models are 97% and 51% for p=1.5; 95% and 49% for p=1.7, respectively.

	Random field		Movement		Non-spatial	
Metrics Parameter	Relative Bias	Coefficient of Variation	Relative Bias	Coefficient of Variation	Relative Bias	Coefficient of Variation
α	2.42%	0.17	12.64%	0.18	23.44%	0.56
β	0.5%	0.14	1.96%	0.21	23.19%	0.49
$\sigma_{\omega}$	2.72%	0.19	7.80%	0.21		
$\sigma_{\varepsilon}$	12.05%	0.09	19.12%	0.12		

Table S1. Simulation results for parameter estimation performance (p = 1.5)

	Random field		Movement		Non-spatial	
Metrics Parameter	Relative Bias	Coefficient of Variation	Relative Bias	Coefficient of Variation	Relative Bias	Coefficient of Variation
α	4.25%	0.18	10.51%	0.26	6.56%	0.68
β	1.68%	0.16	2.67%	0.21	6.88%	0.54
$\sigma_{\omega}$	3.18%	0.21	8.91%	0.19		
$\sigma_{\varepsilon}$	14.14%	0.11	35.05%	0.08		

Table S2. Simulation results for parameter estimation performance (p = 1.7)

The metrics calculation methods

 $Relative Bias = \frac{Mean(Parameter \ estimates) - Parameter \ true \ value}{Parameter \ true \ value}$ 

Coefficient of Variation =  $\frac{SD(Parameter estimate)}{Mean(Parameter estimate)}$