

A simulation-based comparison of confidence interval coverage, bias, and variance of alternative spatial biomass estimation methods used for the evaluation of Northern Shrimp biomass

John-Philip Williams

A Thesis in the Department of
Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of
Master of Science (Biology)

At Concordia University
Montréal, Québec, Canada

October, 2023

© John-Philip Williams, 2023

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: John-Philip Williams

Entitled: A simulation-based comparison of confidence interval coverage, bias, and variance of alternative spatial biomass estimation methods used for the evaluation of Northern Shrimp biomass

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Pedro Peres-Neto

Chair

_____ Examiner

Dr. Krista D. Baker

_____ Examiner

Dr. Pedro Peres-Neto

_____ External Examiner

Dr. James Grant

_____ supervisor

Dr. Eric Pedersen

Approved by _____

Dr. Robert Weladji, Graduate Program Director

Pascale Sicotte, Dean of Faculty

Date 13/10/2023

Abstract

A simulation-based comparison of confidence interval coverage, bias, and variance of alternative spatial biomass estimation methods used for the evaluation of Northern Shrimp biomass

John-Philip Williams

In fisheries management, reliable estimates of population abundance metrics such as standing biomass are crucial for sustainability and balanced decision-making. The OGive MAPping (OGMAP) method, used by the Department of Fisheries and Oceans (DFO) in Newfoundland and Labrador for biomass estimation, addresses non-normally distributed populations but raises concerns about handling spatial data variations. I conducted a simulation-based comparative analysis comparing OGMAP against Generalized Additive Models (GAMs) and STRATified Programs (STRAP) to answer the following question: “are the uncertainties of the estimates calculated from these different methods reliable?”

Using Northern Shrimp, *Pandalus borealis*, as a reference, I simulated biomass landscapes, exploring parameters like landscape roughness, sampling intensity, and model settings. The analysis consistently showed OGMAP's failure to capture nominal confidence intervals (CIs) compared to alternatives, regardless of the treatment. OGMAP exhibited tighter intervals, raising concerns about overfitting and its inability to reflect the true landscape biomass. However, halving the automatically optimized bandwidths for OGMAP's probability distribution fields significantly improved its realized coverage.

These findings underscore OGMAP's variability, shedding light on its limitations in decision-making by the Department of Fisheries and Oceans. I stress the pivotal role of reliable estimates in fisheries management. Additionally, I suggest that alternative methods, like GAMs, may offer more dependable forecasts given OGMAP's underperformance. This research prompts a review of the fisheries management framework relying on OGMAP, suggesting potential inadequacies in capturing the true uncertainty associated with spatially distributed stocks.

Acknowledgments

I'm grateful to everyone who played a role in bringing this project to completion. I am ever thankful to Dr. Eric Pedersen, my supervisor, for entrusting me with this project during my master's. His continuous support and guidance have expanded my knowledge and helped me produce my best work to this day. I'm also grateful to Dr. Pedro Peres-Neto for introducing me to the world of Biostatistics and sharing his passion for the subject. I'd like to thank Krista Baker for her insight into DFO and her constructive input throughout my project.

To my lab mates, I thank you for your time and input. Your cheerful support made this journey enjoyable despite its challenges and remote working at times... Special thanks to my dear friends Claire, Pierre Olivier, Amely, Gabriel, and Marion for their unwavering support and long hours spent working together. I'm also incredibly grateful to my parents for their consistent encouragement throughout my studies, for supporting my goals, for being understanding of my long working hours and for always cheering me on!

I'd like to thank the Fisheries and Oceans Canada (DFO) for funding this project through the Sustainable Fisheries Science Fund. This enabled me to dedicate myself to my master's project. I'd also like to acknowledge Concordia University for granting me the Concordia Merit Scholarship and the Canadian Digital Research Alliance for providing access to high-performance computing resources. Lastly, I'd like to mention my colleagues at Randolph Dix30, for their understanding and flexibility regarding my academic commitments and constant support when needed.

Once again, I'd like to give a special mention to my supervisor, Eric Pedersen, for his valuable insights and the time he dedicated to working with me on this project. His review, suggestions, and edits have been incredibly valuable in ensuring I produced a quality project. I look forward to collaborating with him should the opportunity arise. Once again, thank you, Eric, for your ongoing support and guidance.

Table of Contents

List of Tables and Figures.....	vi
List of Symbols and Abbreviations	vii
Glossary.....	viii
Introduction	1
Methods.....	5
1. Simulation Design.....	5
i. Original landscape generation.....	7
ii. Generation of landscape variation by adjusting landscape patchiness.....	8
iii. Depth & Temperature Generation.....	8
iv. Biomass Generation.....	9
v. Strata generation.....	10
vi. Data Sampling.....	11
2. Experimental treatments.....	11
3. Statistical approaches.....	12
4. Data analysis.....	14
5. OGMAP bandwidth manipulation experiment	15
Results.....	16
1. Comparing the Effects of Landscape Variation	16
6. Effect of Sampling Intensity:.....	18
7. Comparison with Stock Metrics:.....	18
8. Effect of OGMAP's Bandwidth:	22
Discussion.....	23
References	26

List of Tables and Figures

In-text Tables and Figures

Figure 1 : Comparison of Confidence Intervals with Error Bars: Plots showcasing the estimation of fish biomass over time. Error bars illustrate different confidence intervals—red error bars (left) represent a well-covered interval, while green error bars (middle) indicate an overfitted interval & orange error bars indicate a biased coverage. The blue points represent true biomass values. 4

Figure 2: Framework of the landscape generation and subsequent analysis of OGMAP, STRAP and GAMs. Connections between boxes within the same level represent the order of steps happening for that level; arrows indicate the start of a new process. 5

Figure 3: Sample plot of the resulting raster from the NLMMPD generation algorithm, brighter pixels represent higher values while darker pixels represent lower values; this represents the first step in generating a biomass landscape. 7

Figure 4: Fractal plasma example..... 8

Figure 5: Variation 1-4 (left to right) Landscape examples with Depth (top) of the landscape display from dark blue, shallowest, to yellow, deepest areas, and resulting Biomass (bottom) of the landscape as a function of depth and water temperate where brighter colours indicate increased biomass in metric tons. 10

Figure 6: Strata generation (left), the colors represent individual stratum which are areas representing pixels of similar depth ; Depth map (right) represents the depth of the landscape with lighter colors indicating deeper areas (i.e. yellow is the deepest) 11

Figure 8: Confidence intervals by method A) Confidence intervals (at levels of 90% (blue), 95% (green), and 99% (orange)) for biomass estimates (in units of 1000 tons) using the OGMAP method. Displayed percentages correspond to the realized coverage at the 95% confidence interval. B) Confidence intervals for the GAM biomass estimates. C) Confidence intervals for the STRAP biomass estimates. D) Biomass distribution of the landscape, utilizing colour gradients to represent biomass density across the area. E) Geographical representation of strata based on depth. 16

Figure 9 : Realized coverage of Biomass Estimations; Realized coverage (%) for which the confidence interval of each method was able to capture the true biomass of the landscape at increasing levels of landscape roughness. The GAM model is represented in green, the OGMAP method in blue, and STRAP in red. The dashed line represents the nominal coverage at the set 90%, 95% and 99%, respectively. 17

Figure 10 : Percentage of Successful Biomass Estimations; Percentage of true values, rate of success for which the confidence interval of each method was able to capture the true biomass of the landscape at increasing sampling levels. The GAM model is represented in green, the OGMAP method in blue, and STRAP in red..... 18

Figure 11: Percentage bias of GAM (green), OGMAP (blue), and STRAP (red) methods across ascending variation levels from 1 to 4 (low to high), with error bars denoting the standard deviation (SD) of the percentage bias across different landscapes with the same level of landscape roughness. 19

Figure 12: Variance of each method, OGMAP (blue), GAM (green), and STRAP (red), according to the changing landscape roughness levels from low to high (1-4). The error bars represent the standard deviation (SD) of the variance. 20

Figure 13 : Average width of Confidence Intervals (CI), representing the resulting Confidence Interval Widths across CI levels, Variations 1-4 (low to high roughness), and methods (GAM in green, OGMAP in blue, STRAP in red)..... 21

Figure 14 : Effects of the OGMAP bandwidth on the resulting confidence intervals at levels of 90% (blue), 95% (green), and 99% (orange). The values are ordered by increasing values of model point estimates. The black line represents the absolute true biomass of each Landscape. 22

List of Symbols and Abbreviations

Symbols & abbreviations	Definition
CI	Confidence Interval
CMA	Covariance Matrix Adaptation
GAM	Generalized Additive Model
KS	Kolmogorov–Smirnov test
MAD	Median Absolute Deviation
OGMAP	Ogive MAPping
REML	Restricted maximum likelihood
STRAP	STRatified Analysis Program

Glossary

Term	Definition
<i>Confidence Interval</i>	The confidence interval (CI) is a statistical range derived from a sample statistic, indicating a range of values within which the true population parameter is expected to fall with a defined level of confidence (the <i>nominal coverage</i> of the interval), as long as the assumptions made about the sampling distribution for the population are accurate (Altman, 2005). For example, on average a 95% confidence interval should include the true population parameter in 95 samples out of 100 representative samples.
<i>Bias</i>	The bias of a <i>point estimate</i> from a given estimator is the systematic tendency of the estimated value to deviate in a particular direction from the true value of a parameter or quantity of interest (Bishara & Hittner, 2015). It represents the expected (average) discrepancy between the point estimate and the actual population parameter across multiple samples.
<i>Variance, precision</i>	<p>Variance is a quantitative measure of the spread or dispersion of data points within a dataset around their mean value. When referring to the variance of a <i>point estimate</i> from a given statistical method (estimator), it refers to the expected squared deviation of the point estimate from the expected value of the point estimate across replicated samples (the standard error of the estimate). If one statistical estimator has a higher variance than another, it implies that estimates of the higher variance estimator will be more variable from one sample to the next than the low variance estimator (Wood, 2015).</p> <p>Precision in fisheries measurement involves the level of consistency and exactness in the data collected during stock assessments. A precise estimate of fish stock size indicates that repeated assessments under similar conditions yield similar results. The precision of an estimate is the inverse of the variance of that estimate, and I use both terms as is convenient in the text.</p>
<i>Nominal Coverage</i>	The nominal coverage of a <i>confidence interval</i> procedure denotes the expected fraction of samples for which the CI will include the true population parameter. For instance, when discussing a 95% nominal coverage, it means the method is theoretically projected to capture the population parameter within the calculated confidence interval in 95 samples out of 100. This metric specifically quantifies the anticipated proportion of confidence intervals at a given alpha level, which should include the true population parameter under the presumption of an accurate model (Wackerly et al., 2014).
<i>Realized Coverage</i>	Realized coverage signifies the actual proportion of <i>confidence intervals</i> calculated by a given procedure that encompass the true population parameter (Dodge, 2003). It contrasts with <i>nominal coverage</i> by evaluating the practical performance of a statistical method in real data. For example, if a method claims a 95% nominal coverage, realized coverage checks if, in practice, this method indeed captures the true parameter in about 95 out of 100 samples when applied to observed data. If the true value is captured in only 85 out of 100 samples in a given case, the realized coverage would only be 85%.
<i>Point Estimate</i>	A point estimate is a single numerical value from a statistical model used to estimate an unknown population parameter, such as the true standing stock of fish in a population.

<i>Strata</i>	In the context of fisheries science, a stratum or strata (plural) refers to a grouping of similar characteristics, specifically used to denote continuous areas of the landscape that share a similar depth. Strata are used in <i>random stratified sampling</i> to ensure consistent spatial coverage of a large area across sampling periods.
<i>Random Stratified Sampling</i>	Random stratified sampling involves randomly selecting sampling locations across distinct <i>strata</i> . This method relies on choosing locations within each stratum randomly. The number of samples taken from each stratum is proportional to the area it represents.

Introduction

Estimating biomass and abundance presents a complex challenge, with modelling standing as a potent tool to unveil the intricacies of fisheries management; however, the shadow of uncertainty continues to cast its influence. Despite our reliance on rigorous analyses, the persistent concern of inherent uncertainty persists (Charles, 1998). This is far from a new issue, as numerous endeavours have been undertaken to fathom and tackle the impacts and extent of uncertainty (Charles, 1994, 1998; Hilborn, 1987). It is clear that the focus should shift from point estimates to emphasizing how uncertainty is measured and its profound significance in decision-making (Smith & Addison, 2003). Recognizing the importance of uncertainty assessment has prompted the exploration of techniques such as prediction intervals, which find valuable application in shaping management decisions within fisheries (Berkson et al., 2002).

The dangers of an overly confident comparison of estimates are clear, as such an approach may inadvertently lead to the escalation of fishing rates (Anderson, 2017). Erroneous estimation of biomass can have detrimental repercussions across ecological management and conservation endeavours (Privitera-Johnson & Punt, 2020a). Overestimating current biomass within a stock might lead to inadequate resource protection, while underestimation could trigger overly restrictive measures, impeding the pursuit of sustainable resource utilization. Even the most thorough analyses and conclusions might unintentionally disregard critical components, thereby yielding inaccurate estimates (Shertzer & Prager, 2007). Therefore, a comprehensive understanding of uncertainty and its appropriate incorporation into the decision-making processes is a high-priority item if we wish to protect our fisheries in the long term.

Accurate measures of statistical uncertainty are especially important for managing low-abundance or rapidly declining stocks. The decline in fish stocks is a global concern, encompassing not only the broader fish population but also specific subgroups like groundfish rather than the entire fish population (Palomares et al., 2020). Of particular note is the decline in estimated biomass of Northern Shrimp (*Pandalus borealis*) on the Newfoundland and Labrador (NL) Shelves since 2008-2009 (Pedersen et al., 2022). This raises the question: can the current modeling technique accurately capture the uncertainty of the shrimp population throughout its decline, or might it potentially be unable to faithfully represent the current state of the population? Our understanding of this situation is informed by shrimp trawling surveys, which commenced in 1995 as a direct method of assessment. Answering this question requires an accurate assessment of trends, particularly given the inherent uncertainties associated with such endeavours. The emphasis on the precision and the acknowledgment of potential uncertainties are integral components of discerning the true dynamics at play.

It is crucial to recognize the global context of these challenges as fisheries worldwide grapple with similar concerns. The diminishing biomass, whether in NL's once-thriving fish and shrimp stocks or elsewhere, underscores the urgency of sound stock estimation. This, in turn, guides decision-makers to formulate strategies for the sustainability of fishing practices, such steps as setting appropriate catch limits and implementing conservation measures. These actions are essential not only for safeguarding the ongoing viability of fish stocks but also for ensuring the socio-economic well-being of the fishing industry and the communities that depend on them (Kilduff et al., 2009).

Setting Total Allowable Catch (TAC) within fisheries management is intricately tied to uncertainties that influence our understanding of fisheries and their dynamics (Magnusson & Hilborn, 2007; Mangel & Clark, 1983; Privitera-Johnson & Punt, 2020b). The balance between community needs and ecosystem preservation is underscored by the complexity of accurately assessing fish stock status in the presence of inherent uncertainties. Models are employed to guide TAC determination, acting as guardrails for responsible fish harvest (Schrank, 2007). Risk tolerance significantly influences target-setting, as overly

ambitious goals can lead to overfishing, with potential consequences for both communities and ecosystems (Ruckelshaus et al., 2013; Szuwalski & Hollowed, 2016).

The uncertainty that pervades fisheries management emerges from a diverse array of sources, including the complexities of fish dynamics, imprecise stock estimates, and suboptimal harvest quotas (Sethi et al., 2005). This uncertainty encompasses intricate factors such as density-dependent population interactions, environmental variables, management objectives, and institutional frameworks (Charles, 1998). In other words, the challenges in accurately assessing fish stock status and establishing sustainable TACs stem from the complex combination of these different factors. The collapse of the Northern Cod (*Gadus morhua*) population in the 1990s serves as an example of how inadequate management decisions driven by uncertainty can have far-reaching consequences, including the closure of fisheries and economic hardships (Bavington, 2011; Verma, 2019; Walters & Maguire, 1996).

The Northern Cod fishery in NL further underscores the ramifications of inaccurate assessments and overly optimistic models, which can lead to overfishing and significant socio-economic losses (Hutchings & Rangeley, 2011). These instances underscore the imperative of precise stock assessments and the implementation of sustainable management practices within fisheries. The intricate interplay of uncertainty within TAC setting and fisheries management necessitates adaptive strategies to navigate the complexities and mitigate potential risks

Fisheries and Oceans Canada (DFO) has conducted extensive research on the shrimp dynamics of fisheries in the NL Shelves (Fisheries and Oceans Canada, 2019). The Northern Shrimp fishery is of utmost importance to the economic and social well-being of rural coastal communities in the region (Fisheries and Oceans Canada, 2006). Further, Northern Shrimp are also ecologically important in benthic ecosystem of the NL Shelves: they represent a substantial portion of the total benthic biomass and serve as a key resource for several predators (Lilly et al., 2000). A decline in Northern Shrimp stocks has been observed starting in 2008, especially in the more southern Shrimp Fishing Areas (SFA 6 and 7; Fisheries Oceans & Canada, 2017). In 2016, the Northern Shrimp SFA 6 stock shifted from a "Cautious" to a "Critical" state according to the DFO's fish health stock categorizing system (Fisheries and Oceans Canada, 2017). As a result, the DFO implemented shrimp quota cutbacks in SFA 6, and SFA 7 (the most southerly SFA) has been fully closed to shrimp fishing since 2015 (Fisheries and Oceans Canada, 2018). Shrimp biomass in SFA 6 has remained in the Critical Zone despite these cuts in quota (Fisheries and Oceans Canada, 2017). Further, the two northern SFAs, SFA 4 and 5, have also shown declines in estimated biomass from peak levels, but are still both considered to be in the healthy zone (Fisheries and Oceans Canada, 2023). However, all four stocks continue to show strong inter-annual variability in estimated biomass, and it is unclear if this is due to actual population variability or variability in estimated biomass, and there is debate on whether the current biomass estimation methods used for this stock are appropriately incorporating measurement uncertainty into the calculation of CIs (Pedersen et al. 2022).

It is evident that reliable shrimp stock estimates are critical for effective fisheries management and decision-making. As such, since 1995 DFO has monitored the status of shrimp stocks on the NL Shelves through a *stratified random sample* (see glossary for definitions of italicized terms) bottom trawl survey (Chadwick et al., 2007). For each bottom trawl, the total weight of Northern Shrimp caught is measured, and the measurements from each year are used to calculate indices of current stock size. However, a range of alternative spatial statistical models have been used at different points in time to estimate current Northern Shrimp biomass from the trawl survey data (Healey et al., 2019). These models differ in their assumptions about the statistical distribution of shrimp, and how to calculate CIs for current biomass.

The earliest statistical method of estimating shrimp biomass was Stratified Analysis (STRAP), consisting of calculating the summed weighted average biomass across all *strata* (S. J. Smith & G. D. Somerton, 1981). CIs for STRAP are calculated by assuming the biomass distribution across trawls within each stratum was normally distributed. However, this assumption can be problematic for species with long-tailed biomass distributions, with many trawls having zero biomass and a few trawls having very large biomasses (Orr, 2004). This is a particular issue for species like the Northern Shrimp, which aggregates into large groups (Duffy, 2010). Furthermore, STRAP requires every stratum to be sampled in each year to calculate biomass estimates and missing samples can lead to biased or missing biomass estimates. As such, STRAP was replaced with a new statistical model in the early 2000s: Ogive Mapping (OGMAP) (Orr, 2004).

OGMAP was developed as a nonparametric spatially explicit alternative method to STRAP for estimating shrimp biomass (Skanes & Evans, 2005). OGMAP estimates the distribution of biomass at each location in space via kernel regression with a weighting scheme based on distance and depth between locations and survey points (Evans et al., 2000; Geoffrey Evans, Unpublished). OGMAP estimates bandwidths for depth and space using a custom bootstrapping procedure (Evans, Unpublished). This method generates a nonparametric estimate of the probability of getting a trawl of X biomass at a given point based on the distribution of biomass in nearby locations. This model produces estimates for all parts of the distribution of biomass at each location in space: the mean, median, mode, and others. To calculate biomass for a whole region, OGMAP estimates biomass across grid of nodes spread across space, and then sums the biomasses across nodes; it uses bootstrapping to calculate CIs for biomass estimates (Orr, 2004). A simulation study comparing STRAP to OGMAP indicated that OGMAP CIs were smaller (more precise) than STRAP CIs, while still having good *realized coverage* (Skanes & Evans, 2005). However, this study used fixed bandwidths when estimating biomass of simulated data, so did not test how OGMAP's custom bandwidth estimation procedure affected the precision of its estimates.

A third statistical method suggested for estimating biomass for these stocks are Generalized Additive Models or GAMs (Pedersen et al. 2022). GAMs are a semi-parametric statistical model that models the distribution of observations as following from some statistical distribution, with a mean that can vary following a nonlinear function of covariates (such as location in space or depth). This method was suggested for modelling these stocks to account for the substantial spatial variation in the NL Shelf, which ranges in depth from 50 up to 4500 meters (Colbourne et al., 2016). GAMs are highly customizable and are based on a combination of both generalized linear models (GLM) and smoothing methods developed in the 1970s and 1980s (Hastie & Tibshirani, 1986). This method, as the name implies, is based on the addition/sum of many smaller functions. These small functions are called basis functions, and together, they can be combined to form a range of smooth functions in a GAM. Each basis function is multiplied by a coefficient. GAMs use a quadratic penalty on the estimated coefficients, with a penalty estimated using the data, to avoid overfitting the fitted curves to the sample that can occur with a large number of basis functions (Wood, 2017). GAMs are typically estimated via Restricted Maximum Likelihood, based on the complexity of the curve shapes or, in other terms, its wiggleness (Wood, 2017).

GAMs have gained widespread recognition as a versatile and powerful tool for statistical modelling in various fields, encompassing finance, healthcare, and fisheries science (Hastie & Tibshirani, 1986; Rigby & Stasinopoulos, 2005; Szuwalski & Hollowed, 2016; Wood, 2017). Within fisheries science, GAMs offer valuable insights into fish population dynamics, stock assessment, and ecosystem-based management. Likewise, in the context of fisheries, GAMs enable researchers to identify potential factors affecting fish abundance (Pedersen et al. 2022, Lucet and Pedersen 2023), ultimately leading to enhanced fisheries management and conservation efforts. Moreover, the application of GAMs in fisheries science

proves particularly valuable, as they facilitate the modelling of complex interactions within aquatic ecosystems, estimation of fish biomass, and provision of critical information for sustainable fisheries management. Their capacity to capture complex and non-linear relationships between variables, accommodate diverse fisheries data types, and incorporate spatial and temporal variability contributes significantly to the accurate modelling and management of fish populations, aiding in sustainable fisheries practices and conservation efforts (Hastie & Tibshirani, 1986; Rigby & Stasinopoulos, 2005; Szuwalski & Hollowed, 2016; S. N. Wood, 2017)

Pedersen et al. (2022) compared the performance of GAM and OGMAP models for estimating Northern Shrimp biomass in the NL Shelf. Of particular interest was the notable difference in CI widths between the two methods observed in SFA 6, the area with the most rugged bathymetry among all SFAs, where OGMAP exhibited a much tighter (more precise) CI compared to a GAM model of biomass in the same SFA. Interestingly, these tighter intervals played a role in the shift from STRAP to OGMAP in DFO practices (Orr, 2004). It was unclear in Pedersen et al. (2022) which, if either, of the CIs should be trusted for use in management decision making. This range of CIs underscores the nuances that must be considered when analyzing the reliability of these estimated ranges; a CI may not capture the true value in the expected fraction of samples due to bias or overly narrow estimations (Figure 1).

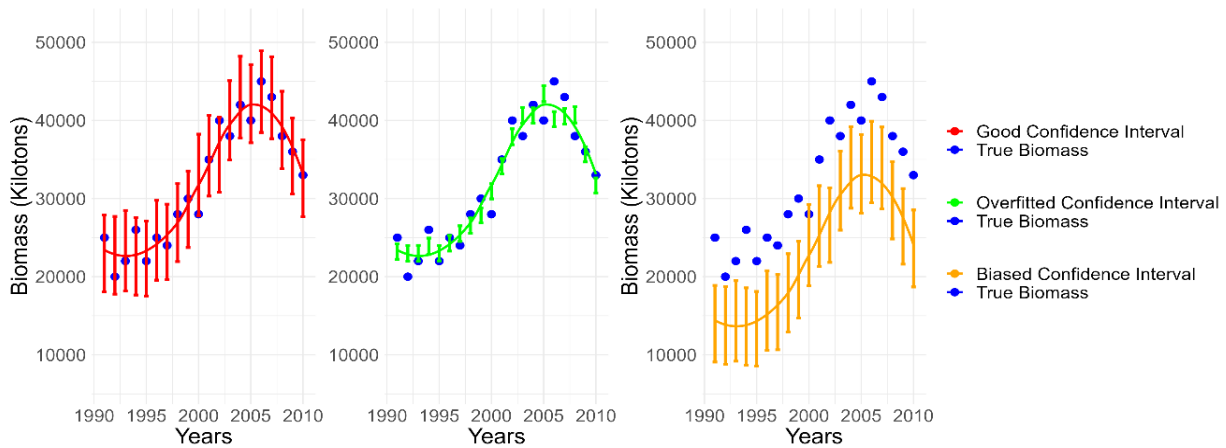


Figure 1 : Comparison of Confidence Intervals with Error Bars: Plots showcasing the estimation of fish biomass over time. Error bars illustrate different confidence intervals—red error bars (left) represent a well-covered interval, while green error bars (middle) indicate an overfitted interval & orange error bars indicate a biased coverage. The blue points represent true biomass values.

This current study is designed to evaluate the cause of this divergence via testing alternative biomass estimation techniques on simulated spatially distributed biomass data, with a specific emphasis on comprehending how various methods for estimating biomass are influenced by different landscapes, particularly within the context of fisheries management and the assessment of spatially distributed species.

This work aims to investigate and further the understanding of models looking at spatially distributed populations, more specifically, the evaluation of methods for estimating spatial measurement error of spatial biomass estimates. I compared the three spatial biomass estimation approaches used in the past for NL Northern Shrimp: OGMAP, GAMs and STRAP. I compared and analyzed the performance of these models to answer the following question: "under what conditions do CIs calculated from these models consistently capture the true stock biomass of the simulated populations?". The statistical performance of

these models was evaluated on a series of spatial biomass simulations. This also involved assessing the reliability of the predicted uncertainty based on the known measurement error of these simulations.

To this end, I performed an analysis of the current OGMAP method for spatial biomass estimation of Northern Shrimp compared with GAM- and STRAP-based approaches. I assessed the performance of the coverage of these models based on prediction error, prediction variance, and residual autocorrelation (Hua et al., 2019; Pedersen, Koen-Alonso, et al., 2020; Potts & Rose, 2018). To understand whether different levels of spatial variation in landscape structure and sampling intensity affected the statistical coverage properties of the different methods, I varied both the spatial structure of simulated landscapes, and the sampling intensity of the simulated stratified random sampling scheme.

Methods

This project involved three main steps: data generation, prediction generation, and data synthesis and analysis (Figure 2). In the data generation phase, I simulated virtual ocean landscapes and spatial distributions of shrimp biomass. Each landscape was stratified into depth strata, mimicking the stratification method used by the DFO trawl survey (Chadwick et al., 2007). I simulated sampling from these landscapes to mimic the data collected through the DFOs NL Shelves bottom trawl survey. The collected data included depth, water temperature, measured biomass, location, and stratum identifier. This artificially generated trawl data were then processed using the selected models to generate predictions. Subsequently, the predictions obtained from the models were analyzed and compared with the true biomass in each simulated landscape to assess the realized CI coverage, CI width, and bias and variance of the point estimates for each method across replicate sample years.

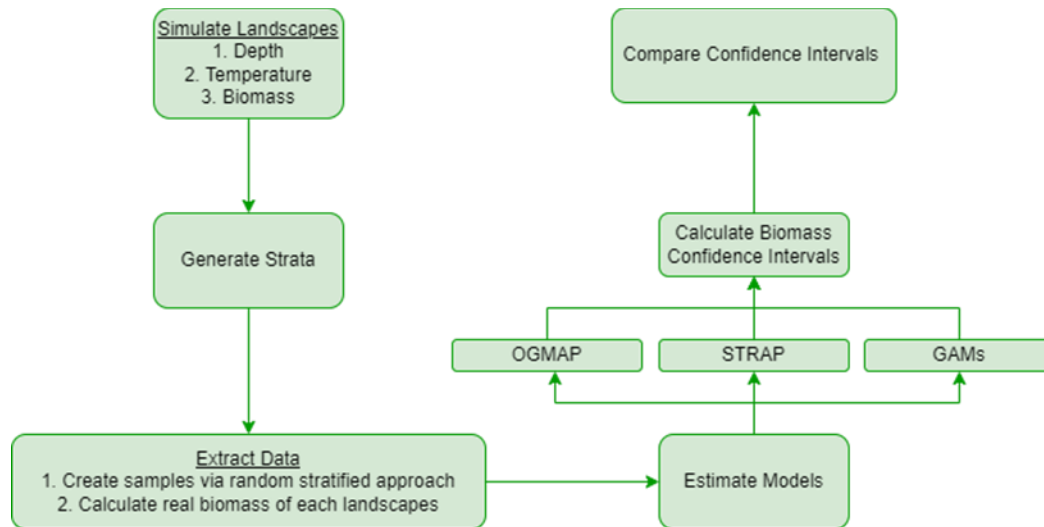


Figure 2: Framework of the landscape generation and subsequent analysis of OGMAP, STRAP and GAMs. Connections between boxes within the same level represent the order of steps happening for that level; arrows indicate the start of a new process.

1. Simulation Design

The study in question was conducted using large-scale simulation data designed to replicate the spatial variation of Northern Shrimp biomass on the NL Shelves, and the trawl-based survey methods used by DFO for the NL multi-species trawl survey (Chadwick et al., 2007). More specifically, I investigated areas that are known to have a great degree of variability in shrimp biomass prediction, as shown by Pedersen et. al. (2022).

The simulations of this work were scripted using R programming language (R Core Team, 2021) Each simulation comprises a raster which contains a pixel grid with assigned values. These rasters have a base size of 500x500 pixels. I simulated values for each pixel for ocean depth, temperature, and shrimp biomass. Here is a summary of the data processing:

1. Simulate an autocorrelated landscape using the NLMR package (Sciaini et al., 2018).
2. Generation of landscape variation by adjusting landscape patchiness.
3. Generation of depth via a median standardization.
4. Generate temperature variation across the landscape based on depth using a historical model.
5. Simulate variation in average shrimp biomass by considering ideal depth and temperature.
6. Divide the landscape into depth-based strata.
7. Sample the landscape and ensure a minimum of three sample trawls per stratum to replicate the DFO methodology. Simulate trawl-level variation in biomass by drawing trawl biomass from a Tweedie distribution (Shono, 2008; Wood, 2023).

i. Original landscape generation

The landscape creation process in this study involves the utilization of the NLMMPD algorithm, which stands for "midpoint displacement neutral landscape model." This algorithm employs a diamond-squares technique to generate synthetic landscapes with a certain degree of autocorrelation, resembling natural environments (Sciaini et al., 2018). The initial stage of landscape generation starts with a grid of points represented by a raster (Figure 3), where each pixel corresponds to a specific location within the landscape, containing information about ocean depth, temperature, and shrimp biomass.

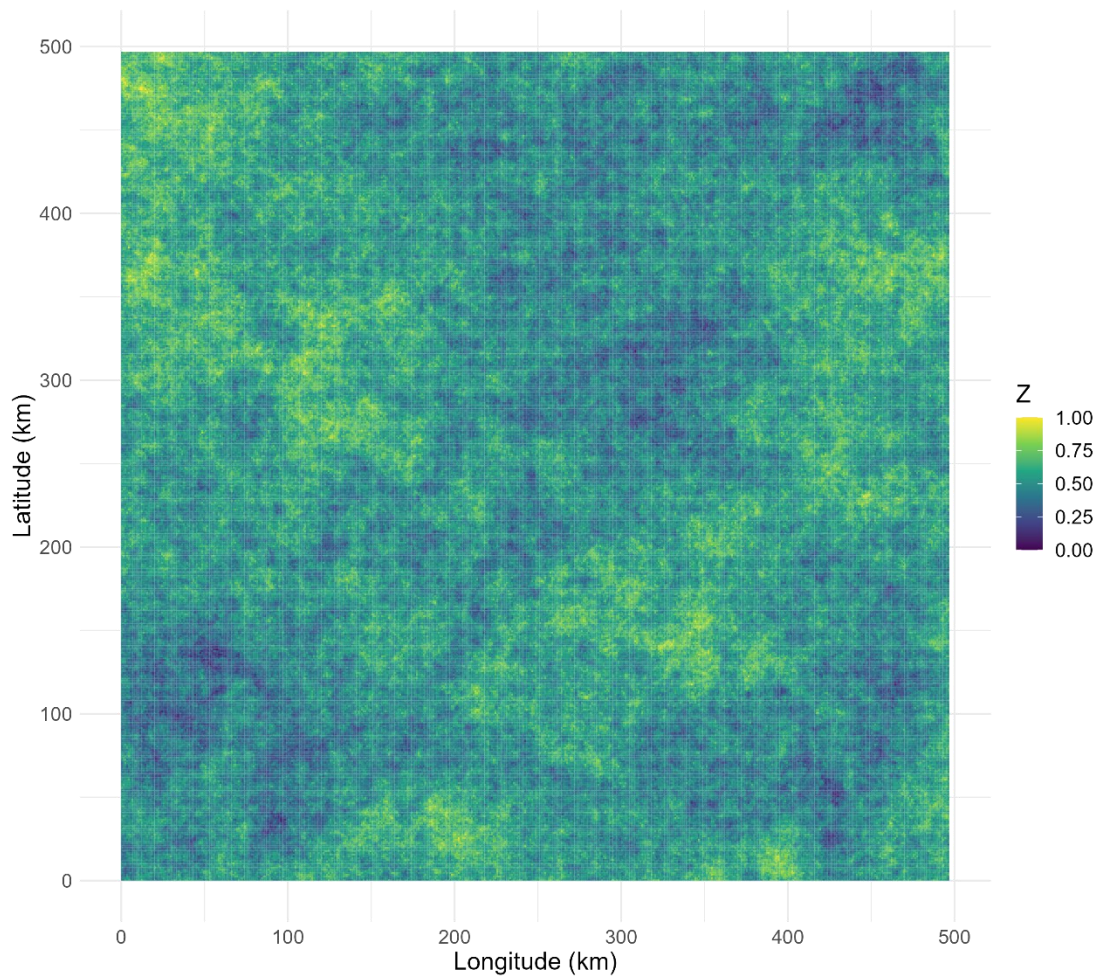


Figure 3: Sample plot of the resulting raster from the NLMMPD generation algorithm, brighter pixels represent higher values while darker pixels represent lower values; this represents the first step in generating a biomass landscape.

Each trial as a whole consisted of a set of runs, each containing a set number of landscapes, which are referred to individually as "sims". These simulations are generated using the ``nlm_mpd`` function from the NLMR R package (Sciaini et al., 2018). The ``nlm_mpd`` applies the NLMMPD technique and is based on a diamond-squares algorithm which generates synthetic landscapes with spatial

autocorrelation by iteratively dividing and averaging values in a grid to create terrain resembling natural environments (Fournier et al., 1982; Miller, 1986). This method is used to create a landscape that varies fractally in space with varying levels of similarity among nearby points (Figure 4). This guarantees that the terrain displays a specific level of autocorrelation, indicating that points close to one another on the landscape tend to share greater similarities compared to more distant points.



Figure 4: Fractal plasma example

ii. Generation of landscape variation by adjusting landscape patchiness.

I employed a patching algorithm to aggregate neighbouring areas. This was achieved by multiplying the newly created landscape with a new classified raster with the following groups: 0, 0.33, 0.66, 1. This raster of clusters was created using an implementation of a naturalistic clustering approach involving nearest-neighbours (Saura & Martínez-Millán, 2000) of the NLRM package. This approach generated a landscape of clusters with an associated value. The shape of which aims to recreate possible patches which could occur in a real-world scenario. As I strive to keep most of the landscape unchanged to preserve its natural correlation, I ensured that 60% of the initial landscape was unaffected by this change. This was achieved using an exponential approach ($\exp(0)=1$), where patches of value 0 were unchanged, and a variation factor was applied to multiply the values of the remaining pixels directly. This landscape was then multiplied with the landscape generated at step 1. This deliberate manipulation was aimed at introducing patches of rough terrain and disrupting the overall smoothness of the Gaussian raster, adding complexity and realism to the landscape.

For these simulations, the goal was to recreate the variability which would be present during a routine multi-species trawl in the Newfoundland and Labrador Shelves region. As a result, it was necessary to have varying degrees of sampling and landscape variability. As the areas covered by these trawls show a high degree of variability, particular attention was given to setting variation in landscape depth/biomass at different levels to replicate this naturally occurring process and how it may affect the performance of each model.

iii. Depth & Temperature Generation

To make sure the depth maps shared a common baseline for comparison, for each simulation, I scaled the output of my initial landscape model so each simulated landscape had a median depth of 200 m and a Median Absolute Deviation (MAD) around the median depth of 100 m. I utilized the MAD as a measure of how spread the depth values were from their median in the landscape generated by the neutral model. This MAD-based scaling factor allowed us to standardize the depth maps effectively and to compare results across simulations with different levels of landscape variation (patchiness). Examples of simulated depth maps are shown in Figure 5 (top panels). To simulate variation in temperature with depth, I used predicted temperature values from a GAM model predicting temperature as a function of depth that I estimated using trawl data in the NL region. The model is based on the relationship between temperature and depth, utilizing the Gaussian family and

adaptive spline smoother (Wood, 2017) with 40 basis functions. I chose this smoother due to the non-linear nature of the data, as the adaptive spline smoother can model functions that change shape at different rates across the function. This model assumes that the temperature at location i follows a Gaussian distribution with a mean average temperature ($\mu_{temp,i}$) that varies smoothly with the square root of depth, and a constant standard deviation (σ_{temp}) across all locations.

$$\mu_{temp,i} = f(\sqrt{depth_i})$$

$$temp_i \sim \text{Gaussian}(\mu_{temp,i}, \sigma_{temp})$$

iv. Biomass Generation

After generating temperature and depth rasters for each landscape, I simulated the spatial distribution of biomass for each replicate year in each landscape by employing specific parameters for depth and temperature within my simulation workflow. I assumed that the mean biomass at a given location was modelled by the product of a Gaussian density function of depth (with a mean of 312.5m and a standard deviation of 200m) and a Gaussian density function of temperature (mean of 2.92°C and standard deviation of 2°C). The mean and standard deviations were chosen so that the peak biomass concentrations across space corresponded with observed patterns of Northern Shrimp abundance conditions in the Newfoundland and Labrador Shelves. By doing so, I obtained biomass estimates for each simulation run (Figure 5, bottom).

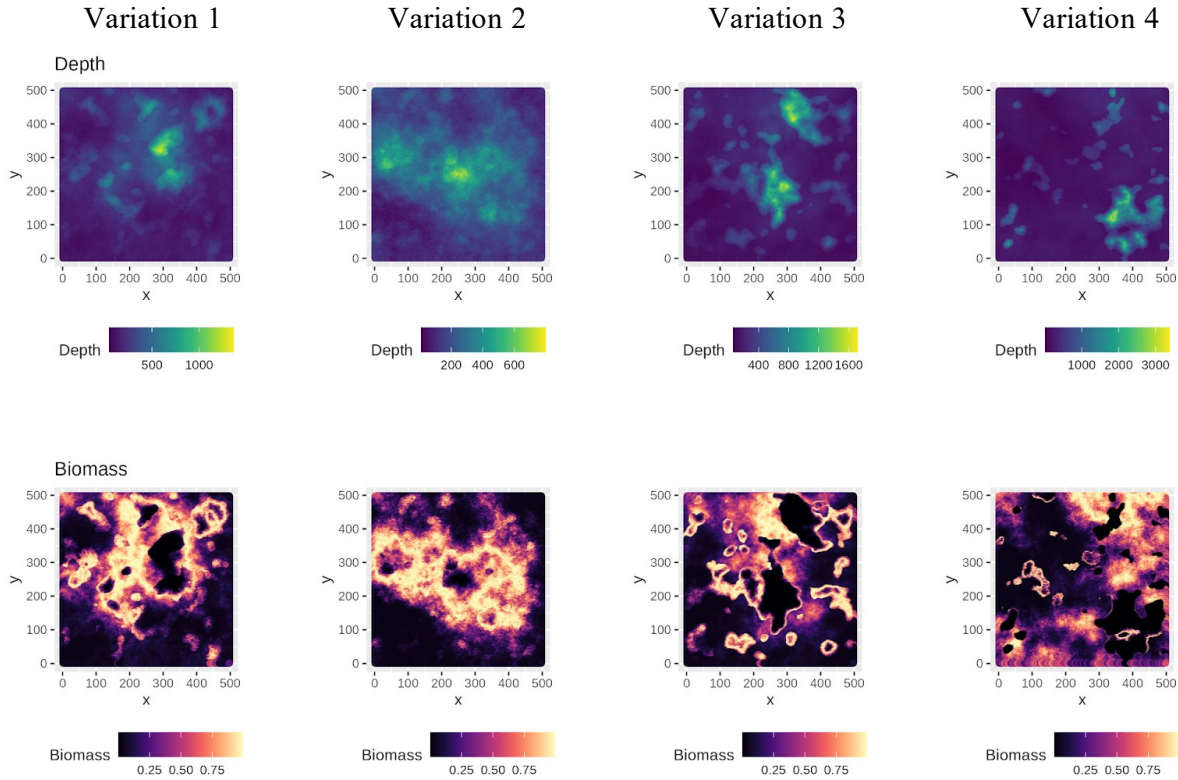


Figure 5: Variation 1-4 (left to right) Landscape examples with Depth (top) of the landscape display from dark blue, shallowest, to yellow, deepest areas, and resulting Biomass (bottom) of the landscape as a function of depth and water temperature where brighter colours indicate increased biomass in metric tons.

v. Strata generation

I divided each simulated landscape into stratas characterized by similar depth levels to mimic the stratified design used by DFO in Newfoundland and Labrador. I implemented a two-step process that ensured both precision and consistency in defining these strata. Initially, I divided each landscape into 16 equal-width depth bins, each representing a distinct range of depth levels. This initial division provided a coarse stratification of the landscape based on depth. To divide large depth bins into smaller strata, I used Voronoi tessellation from the R SSPM package (Lucet & Pedersen, 2023). This technique allowed us to partition each depth bin into smaller, contiguous strata, ensuring that each section had a unique label associated with its continuous depth range. Furthermore, to complete my stratum allocation and enhance the overall coherence of the stratification, I implemented a step to merge small areas into larger ones. This process ensured that my stratas remained ecologically meaningful and manageable for subsequent analysis. The minimum stratum size was set to 200 square kilometres, which is consistent with trawl data from the DFOs NL trawl survey. My stratification process generated a range of 100 to 150 strata for each landscape, each characterized by its specific depth range (Figure 6).

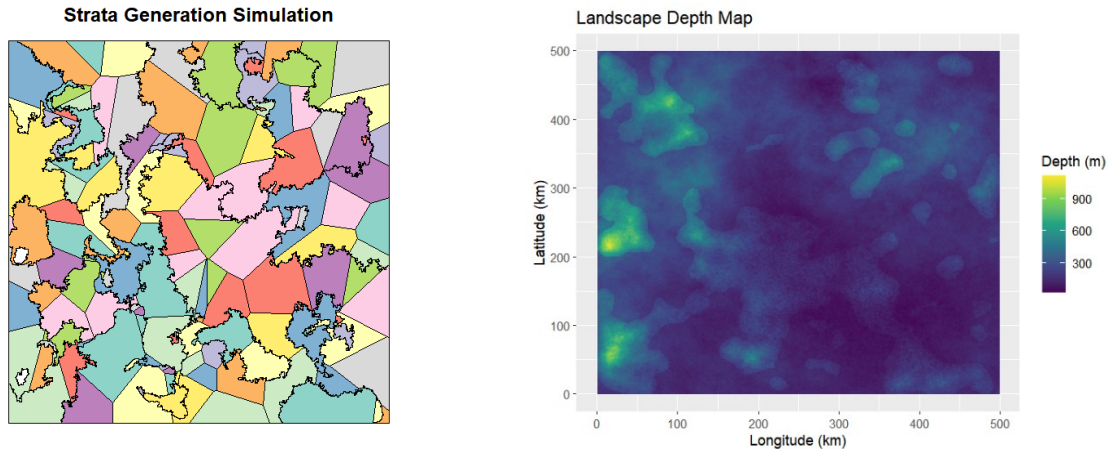


Figure 6: Strata generation (left), the colors represent individual strata which are areas representing pixels of similar depth ; Depth map (right) represents the depth of the landscape with lighter colors indicating deeper areas (i.e. yellow is the deepest)

vi. Data Sampling

I used a stratified random sampling design to simulate trawl samples for each stratum in each year, to mimic the DFO multi-species trawl sampling design (Chadwick et al. 2007). This approach involved the selection of data points from the previously established strata. To maintain consistency through this study, I set a minimum of three samples were drawn from each individual stratum. This decision was twofold in purpose. Firstly, it aligned with the DFO methodology during research trawls, where a minimum of three samples is captured per stratum, ideally. Secondly, it catered to the requirements of the STRAP method for estimating biomass, which relies on a weighted average of mean and standard errors of biomasses from each individual stratum to calculate the estimate and CI for overall biomass. Having three randomly located samples within the same stratum was crucial to obtain accurate estimates of shrimp density, as depth predominantly influences shrimp density, and each stratum exhibits similar depth characteristics across its pixels.

I drew the simulated biomass caught by each trawl from a Tweedie distribution with a mean equal to the simulated biomass density at that location to replicate trawl-to-trawl variation in biomass (Shono, 2008). The Tweedie distribution introduced substantial variability to the trawled biomass values, ranging from high to zero. For this application, I used a shape parameter (P) of 1.76 and a scale parameter (Φ) of 2; this combination ensures a large number of zeros and a moderate spread of values. These values are based on a Tweedie GAM reference model of NL shrimp distribution (unpublished). This variability accurately represented occurrences in research trawls where either zero or very large shrimp trawls were encountered. As a result, this approach generated simulated trawl data that my models could seamlessly process, faithfully recreating the dynamics of a natural trawl scenario.

2. Experimental treatments

The experiment incorporated two treatments: (1) the level of variation, and (2) the sampling intensity, strategically designed to evaluate and enhance the performance of the various models employed in my study. I simulated landscapes with four distinct qualitative levels of variation, ranging from low to high, to investigate how different degrees of landscape variation influence the accuracy and precision of biomass estimates for the statistical models tested (Figure 5). A low level of variation denoted less sharp and smoother variations in depth, while a high level resulted in rapid variation in depth over short spatial

distances. The motivation behind this treatment was to assess how well the models could adapt to landscapes with varying levels of complexity and abrupt depth transitions, reflecting real-world scenarios.

I also varied sampling intensity between the following values: 0.125%, 0.25% and 0.5% of the total landscape sampled. This was done to systematically evaluate how changing sample size affected the statistical performance of my models. This treatment aimed to explore whether any issues identified with the statistical coverage of CIs from my models would improve by increasing the sample sizes.

I generated 150 replicate landscapes for each treatment level, with each landscape having a uniquely simulated bathymetry. For each replicate landscape, I simulated 200 simulated years (simulations) for each landscape. This approach provided a solid foundation for the subsequent analysis and interpretation of the experimental findings.

Furthermore, I reused one of these completed simulations to analyze the influence of OGMAPs bandwidth parameter. Subsequently generating two additional sets of predictions using a narrow and wide bandwidth based on the standard OGMAP bandwidth estimation, so I could compare the coverage properties of OGMAP CIs for different choices of bandwidths.

3. Statistical approaches

For each simulation, I calculated a *point estimate* and the 90%, 95% and 99% CI for the total population biomass using the three methods that have been previously suggested for modelling Northern Shrimp in the NL Shelf stocks: OGMAP, GAM, and STRAP.

OGMAP's approach extends beyond the traditional Ogive method by incorporating a technique for bandwidth optimization. This involved the use of the Covariance Matrix Adaptation evolution strategy (CMA-ES (Hansen & Ostermeier, 1996)), a versatile optimization algorithm implemented through Python (*PyCma*, 2016/2023). The primary objective of this bandwidth optimization process is to fine-tune the selection of bandwidth values, a critical component of probability field estimation. This optimization addresses two aspects: the pursuit of uniformity and independence among the sampled points within the probability field.

Uniformity, a fundamental concept within OGMAP, revolves around achieving a well-balanced distribution of probabilities associated with deviations from Kolmogorov-Smirnov (KS) curves across the $[0, 1]$ range. The key goal is to ensure that each probability value within this range possesses an equal likelihood of occurrence. Optimizing bandwidths significantly enhances the precision and accuracy of the probability field, resulting in more reliable biomass estimates.

Simultaneously, OGMAP places significant emphasis on the concept of "subcauchy" weighting functions. The term "subcauchy" characterizes a specific type of kernel function distinguished by thinner tails compared to the unnormalized Cauchy distribution. These subcauchy kernels play a crucial role in assigning weights to survey nodes based on their proximity to the target node. Notably, the choice of a subcauchy kernel implies that the weights assigned to each survey node depend on their respective distances from the target node. This distribution of weights ensures that survey nodes in closer proximity are assigned higher weights, while those farther away receive lower weights. Consequently, OGMAP's approach captures the local variations in the probability field, leading to more accurate and precise estimates.

Furthermore, OGMAP strongly emphasizes maintaining independence among the sampled probabilities. In this context, "independence curves" represent additional probabilities that are expected to conform to a uniform distribution within the interval $[0, 1]$. These independence curves symbolize probabilities that

exhibit autonomy from one another. Bandwidth selection in OGMAP is designed to select for the widest bandwidths that result in uniform and independent curves (Geoffrey Evans, Unpublished)

After successfully optimizing the bandwidths and generating probability fields that adhere to uniformity and independence criteria, OGMAP proceeds to calculate the total biomass of the landscape. This calculation involves summing up the biomass estimates at each survey node, each weighted according to its proximity and contribution to the target node. By aggregating these localized estimates, OGMAP derives a comprehensive and precise assessment of the total biomass for the entire landscape. OGMAP uses a bootstrap resampling technique to calculate CIs of the biomass at each location for each year. This technique operates by repeatedly resampling from the dataset, creating multiple simulated datasets, each with a different subset of points.

Generalized Additive Models (GAMs), unlike traditional linear models that rely on linear relationships between predictors and the response variable, GAMs are constructed using a collection of basis functions. These basis functions serve as building blocks representing various components of the model, such as individual predictor variables or their interactions. By combining these basis functions, GAMs can flexibly capture complex and non-linear relationships in ecological data.

The fundamental idea behind GAMs is to decompose the relationship between predictors and the response variable into multiple additive components. Each component consists of several nonlinear basis functions of environmental predictors. Basis functions are multiplied by coefficients, and then summed to form the overall model. Finally, GAMs use quadratic penalties on the coefficients of the basis functions to prevent the model from over-fitting the observed data (Wood, 2017). Smoothing functions, such as thin-plate splines, are commonly employed as basis functions in GAMs (Wood, 2003). For this GAM model, I used thin-plate splines as basis functions to capture the relationships between latitude (*lat*), longitude (*long*), and year (*year*) when estimating biomass. As I did not intend on predicting biomass through time, I compounded years in subsets of five years for parallelization and to speed up the modelling process. I specified a tensor product smoother of spatial location and year ("*te*," (Wood, 2006)). I used a tensor product of an isotropic two-dimensional thin-plate spline for space with 27 basis functions and a ten-basis function thin-plate spline smoother for year. This model estimates parameters and smoother penalties using the Restricted Maximum Likelihood (*REML* ;(Wood, 2017)) method and assumes a Tweedie distribution for the response variable (biomass). This enables us to capture complex non-linear relationships and model complex interactions between the predictor and the response variables.

$$\mu_{biomass,i} = f(long_i, lat_i, year_i)$$

$$Biomass_i \sim Tweedie(\mu_{biomass,i}, \phi, p)$$

The GAM model provided a spatially explicit estimate of biomass density for each location in space for each year ($\mu_{biomass,i}$). I translated this into an estimate of biomass for the whole landscape by summing the biomass densities for all locations together into a single value.

To computed confidence intervals (CIs) for the GAM-based biomass estimates, I simulated 1000 replicate coefficient draws from the empirical Bayesian posterior distribution of the fitted model (Wood, 2017). For each posterior draw of each simulation, I calculated the total landscape biomass by summing the estimated biomass density across the landscape. I translated this into confidence intervals for overall biomass by using different quantiles of the posterior distribution, such as 0.1, 0.05, and 0.01, corresponding to 90%, 95%, and 99% CIs, respectively. By analyzing these posterior samples, I obtained robust CIs that provided a comprehensive view of the uncertainty associated with the biomass estimates.

This approach enhances the reliability and robustness of my ecological findings while aligning them with the flexibility and adaptability offered by GAMs.

Finally, the last method used to estimate biomass was STRAP (Stratified Analysis Programs) (S. J. Smith & G. D. Somerton, 1981). For this method, I calculated the point estimate of total biomass (T_{Bio}) as the summation of all average biomasses for each stratum ($\sum_n \overline{bio}_i$), where the average biomass per stratum is determined by the sample mean of biomass densities (kg/km^2) calculated from samples within the same stratum i multiplied by the stratum area ($\overline{bio}_i = \overline{trawl}_i \cdot A_i$). I computed the CI for the STRAP estimate for each simulation from stratum-level estimates of the standard error of the biomass, which is based on the sum of the squared stratum area multiplied by the squared within-stratum sample standard error of the biomass density (S_i^2) ($\sum_n S_i^2 = A_i^2 \cdot SE(bio)_i^2$). The square root of which is equal to the standard error (SE) of the biomass ($B_{SE} = \sqrt{\sum_n S_j^2}$). I then multiplied this standard error by the critical values (z) corresponding to 90%, 95% and 99% CI, according to the Z-distribution, to obtain the respective CIs in combination with the estimate biomass point estimate. ($CI = T_{bio} \mp Z * B_{SE}$)

4. Data analysis

In this study, I evaluated the biomass uncertainty present in spatially distributed populations using one-to-one comparisons of the models. This involved the analysis of prediction data from the models with absolute landscape values initially generated. I used three different statistical methods to estimate the total landscape biomass and CIs for this biomass for each replicate simulation: OGMAP (Orr, 2004), GAM (Hastie & Tibshirani, 1986), and STRAP (S. J. Smith & G. D. Somerton, 1981). These models were built using the sampled data and were subsequently used to generate predictions for each point in the 500 by 500 pixels landscapes, the same as the original data generation. These predictions were then compared to the initial landscapes to estimate the models' performance.

I used the realized coverage of CIs from each method to evaluate the reliability of each model's uncertainty estimates. CIs provide upper and lower bounds of my model's biomass predictions. Since I have access to the actual biomass values from simulations, I can assess how often the CI captures the true value (i.e. the *realized coverage* of the CI). A well-performing CI should contain the true value in a fraction of simulations equal to its nominal coverage probability (i.e., a 95% interval should contain the true value in 95% of simulations). If the CI consistently falls outside the true value, it prompts us to investigate the landscape features responsible for such deviations and refine the models accordingly.

To evaluate the performance of each method for accurately assessing biomass, I calculated two additional statistical metrics of performance: the *bias* and *variance* of estimated values across simulations. Bias represents the average difference between the estimated and true biomass for each model. I used percentage bias as the bias measure, which was calculated using the model point estimates and the original absolute true biomass of the simulated landscape. The bias was obtained from the mean of the difference between the model point estimate (bio_p) and the true value (bio_t) divided by the true value, then multiplied by 100% ($BIAS\%_{bio} = \frac{bio_p}{bio_t} \cdot 100$). Looking at biomass bias allows to gauge the reliability of my models in predicting biomass uncertainty in different habitats where aquatic populations exist. Biomass bias can indicate if different models consistently make similar kinds of errors of over- or under-estimation of population size. Recognizing these patterns of error is essential since I want to rectify these mistakes to enhance the accuracy of the model predictions.

The variance of an estimate is a metric used to provide information about the variability and spread of the model's prediction. I calculated the variance of each model estimate, in kg^2 , from the mean of the squared

residual (bio_{res}^2) of the models ($bio_{var} = \overline{bio_{res}^2}$), where the squared residual is obtained from the difference between the model point estimate (bio_p) and the true value (bio_t) ($bio_{res}^2 = (bio_p - bio_t)^2$). This metric informs us on how consistent the data is; a low variance would indicate that the data is more consistent and less influenced by the initial sample, while a high variance would be associated with the opposite and greater sensitivity. In conclusion, these metrics serve as a tool for assessing why the CIs of the model are not consistently capturing the true value expected.

5. OGMAP bandwidth manipulation experiment

Finally, the last treatment I explored how biomass estimates were affected by the way the OGMAP model processes its kernel regression. By default, OGMAP automatically optimizes its bandwidths to find the most relevant weights for its probability distribution. However, as highlighted in a prior OGMAP investigation (Skanes, 2005), these bandwidths have yet to be explored in depth. Therefore, for my initial exploration of the effects of OGMAP bandwidth, I settled on a test experiment. This experiment aimed to observe the effects of the OGMAP bandwidth on the resulting CIs. I hypothesized that the observed low coverage of the OGMAP intervals could be attributed to OGMAP underestimating the appropriate bandwidth, consequently leading to the under-smoothing of local biomass estimates. This under-smoothing phenomenon can make the model overly responsive to minor local fluctuations in the data, resulting in narrower CIs than the nominal coverage. I delved into the bandwidth parameter to further evaluate the impact of OGMAP's performance and its influence on CIs. I achieved this by calculating OGMAP estimates and CIs three times on the same set of simulations: once using bandwidths estimated using OGMAP's default bandwidth estimation method, once with bandwidths equal to half the default estimates, and once with bandwidths twice as wide as the default estimates. This experiment aimed to understand how changes in the bandwidth of the kernel regression process can affect the model's predictions.

Results

1. Comparing the Effects of Landscape Variation

All three methods produced similar point estimates when applied to different simulated years from a single landscape (Figure 8A-C); Figure 8D shows the average biomass map for the simulated landscape and Figure 8E shows the strata used for sampling for this example. However, the three methods showed a great deal of variation in CI widths: CIs from GAM (Figure 8B) and STRAP (Figure 8C) models had realized coverages close to the nominal 95% confidence level. In comparison, the realized coverage of nominal 95% CIs calculated by OGMAP was only 59.5% across simulations from this landscape (Figure 8A).

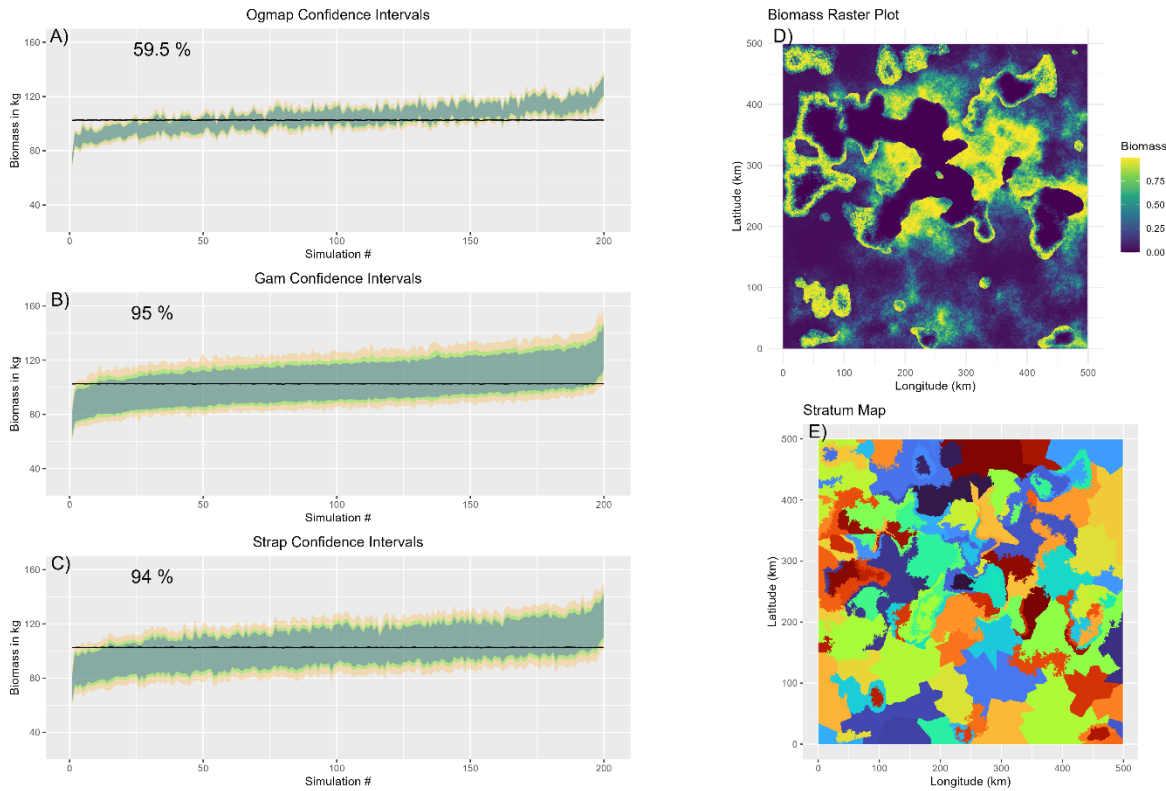
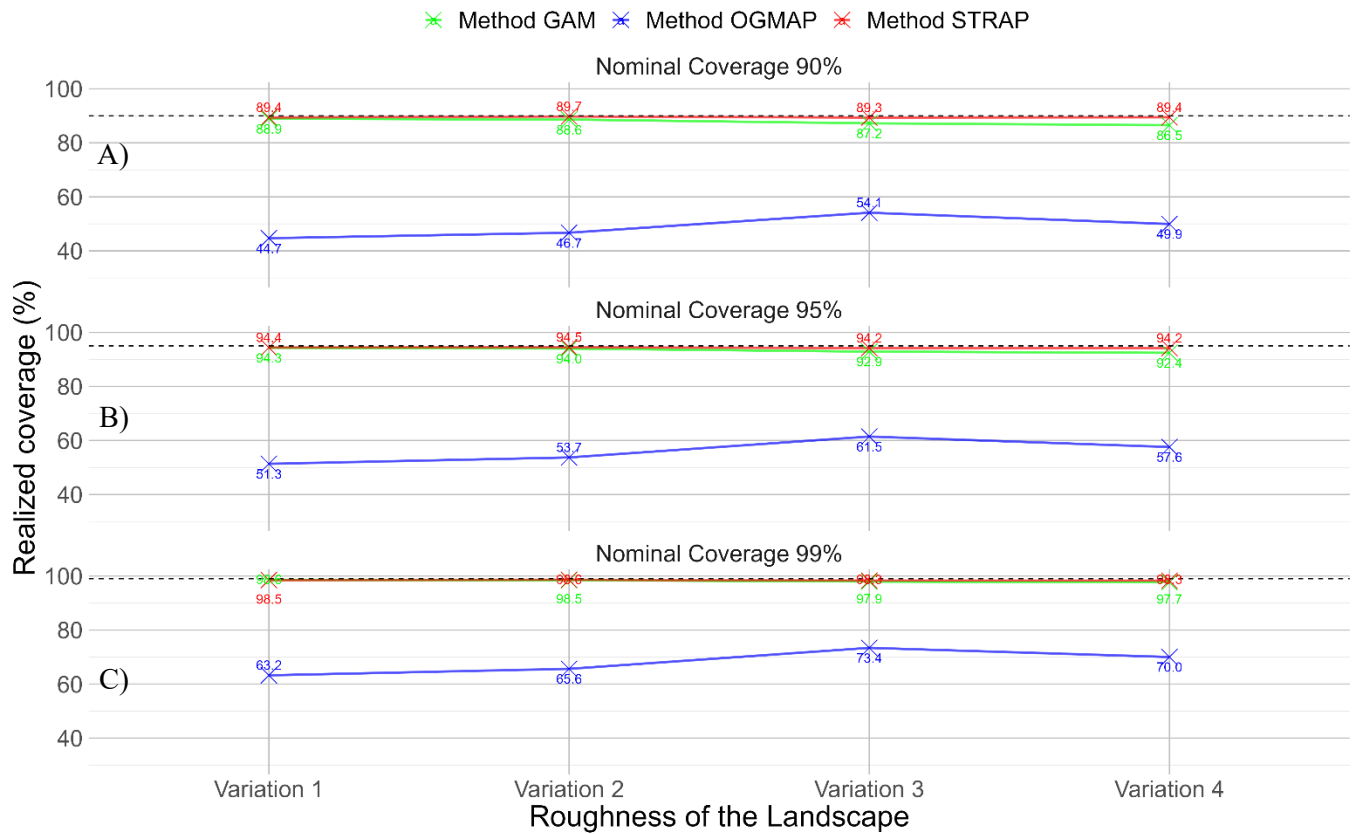


Figure 7: Confidence intervals by method

A) Confidence intervals (at levels of 90% (blue), 95% (green), and 99% (orange)) for biomass estimates (in units of 1000 tons) using the OGMAP method. Displayed percentages correspond to the realized coverage at the 95% confidence interval. B) Confidence intervals for the GAM biomass estimates. C) Confidence intervals for the STRAP biomass estimates. D) Biomass distribution of the landscape, utilizing colour gradients to represent biomass density across the area. E) Geographical representation of strata based on depth.

Figure 9 illustrates how the realized coverage of CIs from each method differed across varying levels of landscape variation. OGMAP coverage ranged from 51.3% to 61.45% at the 95% nominal coverage level (Figure 9B). While OGMAP showed a relative increase in realized coverage as the landscape roughness is increased, its overall ability to capture the true biomass of the landscape is low at all confidence levels, peaking at 70% in the best-case scenario at the 99% nominal coverage level (Figure 9). In contrast, the GAM and STRAP methods display minor CI fluctuations, remained within 2% below the nominal 95%

24 coverage. The same trends were observed at the 90% and 99% CI, with the GAM and STRAP methods
 25 showing nominal coverage in contrast to OGMAP (Figure 9A-9C).

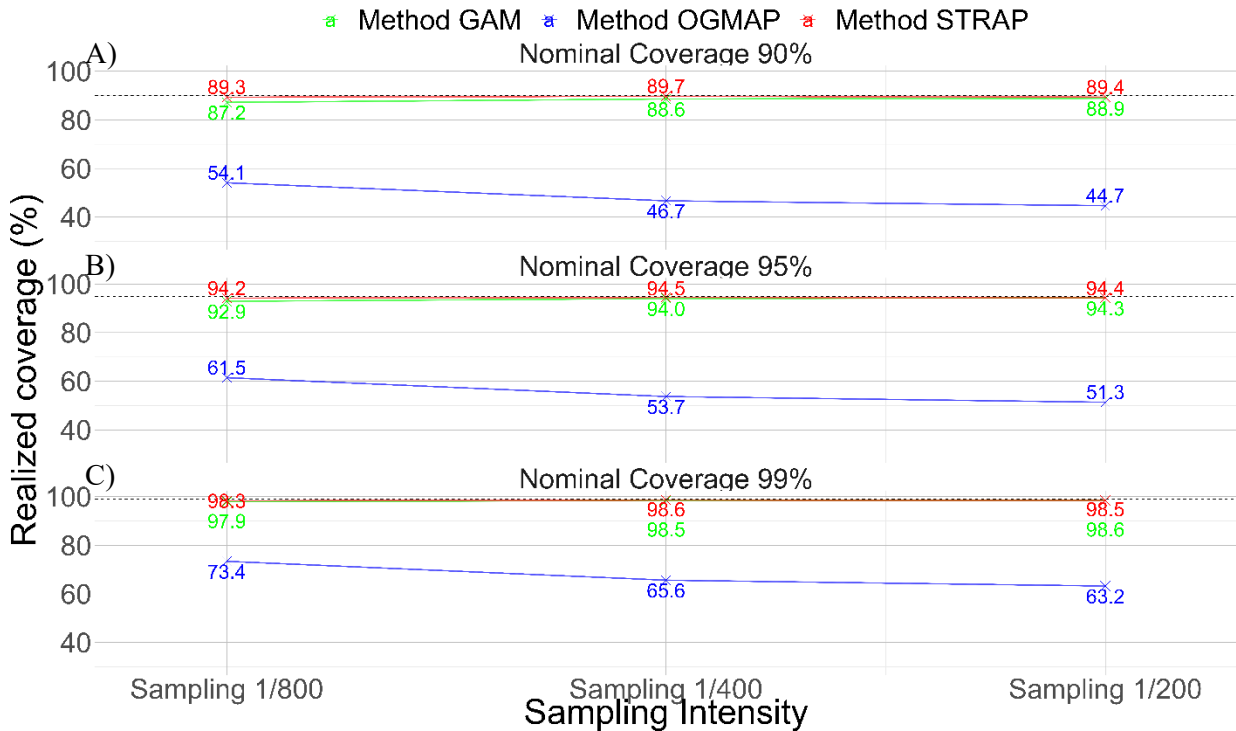


26
 27 *Figure 8 : Realized coverage of Biomass Estimations; Realized coverage (%) for which the confidence interval of each method*
 28 *was able to capture the true biomass of the landscape at increasing levels of landscape roughness. The GAM model is*
 29 *represented in green, the OGMAP method in blue, and STRAP in red. The dashed line represents the nominal coverage at the set*
 30 *90%, 95% and 99%, respectively.*

31

32 **6. Effect of Sampling Intensity:**

33 CIs from both STRAP and GAM confidence intervals maintained realized coverage very close to the
 34 nominal coverage regardless of the sampling intensity (the total number of trawls per simulation) used
 35 (Figure 10). This held true for 90% (Figure 10A), 95% (Figure 10B) and 99% CIs (Figure 10C). However
 36 the realized coverage of OGMAP intervals declined with increasing sampling intensity. At the 95%
 37 nominal coverage level, this resulted in a shift from a realized coverage of 56.7% to 43.5% as I increased
 38 the number of samples (Figure 10B). The performance of OGMAP lagged behind the other investigated
 39 methods (Figure 8). Increasing the sample size generally enhanced the reliability of CIs by providing a
 40 more comprehensive representation of the data distribution.



41
 42 *Figure 9 : Percentage of Successful Biomass Estimations; Percentage of true values, rate of success for which the confidence*
 43 *interval of each method was able to capture the true biomass of the landscape at increasing sampling levels. The GAM model is*
 44 *represented in green, the OGMAP method in blue, and STRAP in red.*

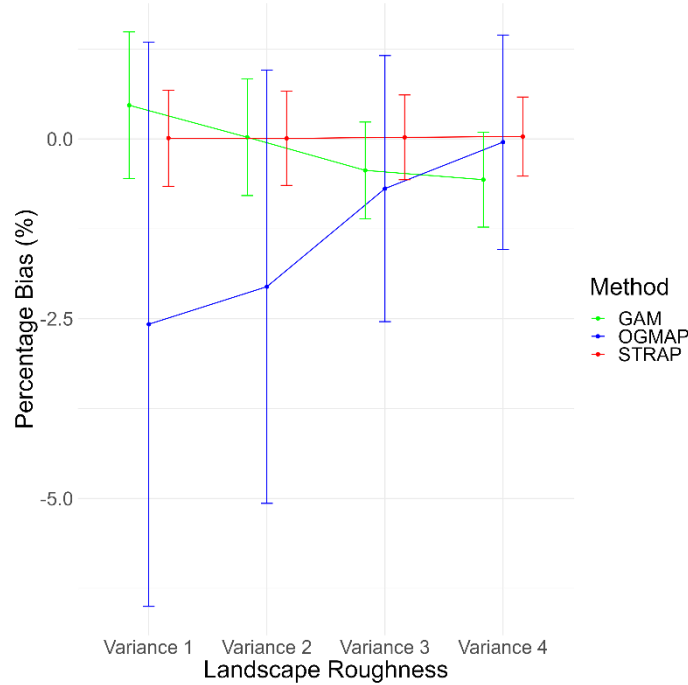
45
 46 **7. Comparison with Stock Metrics:**

47 Point estimates for all three methods showed very little estimate bias as a percentage of true biomass,
 48 regardless of the amount of landscape variation (Figure 11). STRAP estimates had a maximum bias of
 49 0.05% of the true biomass. GAMs tend to yield predominantly unbiased estimates when averaged across
 50 the dataset; with a maximum of 0.45% and a low of -0.79% with a mean of -0.21%, indicating that GAM
 51 estimates are on average only 0.21% lower than the true value of the landscape. GAMs also tend to
 52 become slightly more biased towards underestimating as the landscape roughness increases. OGMAP
 53 estimates were consistently biased downward relative to the true biomass; however, this bias was very
 54 small relative to the absolute biomass, with bias values ranging from -0.05% to -2.56%. The magnitude of

55 the bias of OGMAP decreased as landscape roughness increased. (Figure 11), indicating that it is unlikely
56 that the poor coverage of the OGMAP intervals shown in Figure 10 was due to estimation bias.

57 I then aimed to assess the proximity of the estimates to the mean by examining the mean of the squared
58 residuals of the mean estimates for each method. As illustrated above, both STRAP and GAM display
59 remarkably similar levels of variance. This similarity suggests that the mean estimates from both methods
60 are consistently close to the true mean of the landscape. Moreover, this closeness gradually improves as
61 the level of variation increases, indicating more accurate estimates (Figure 11).

62

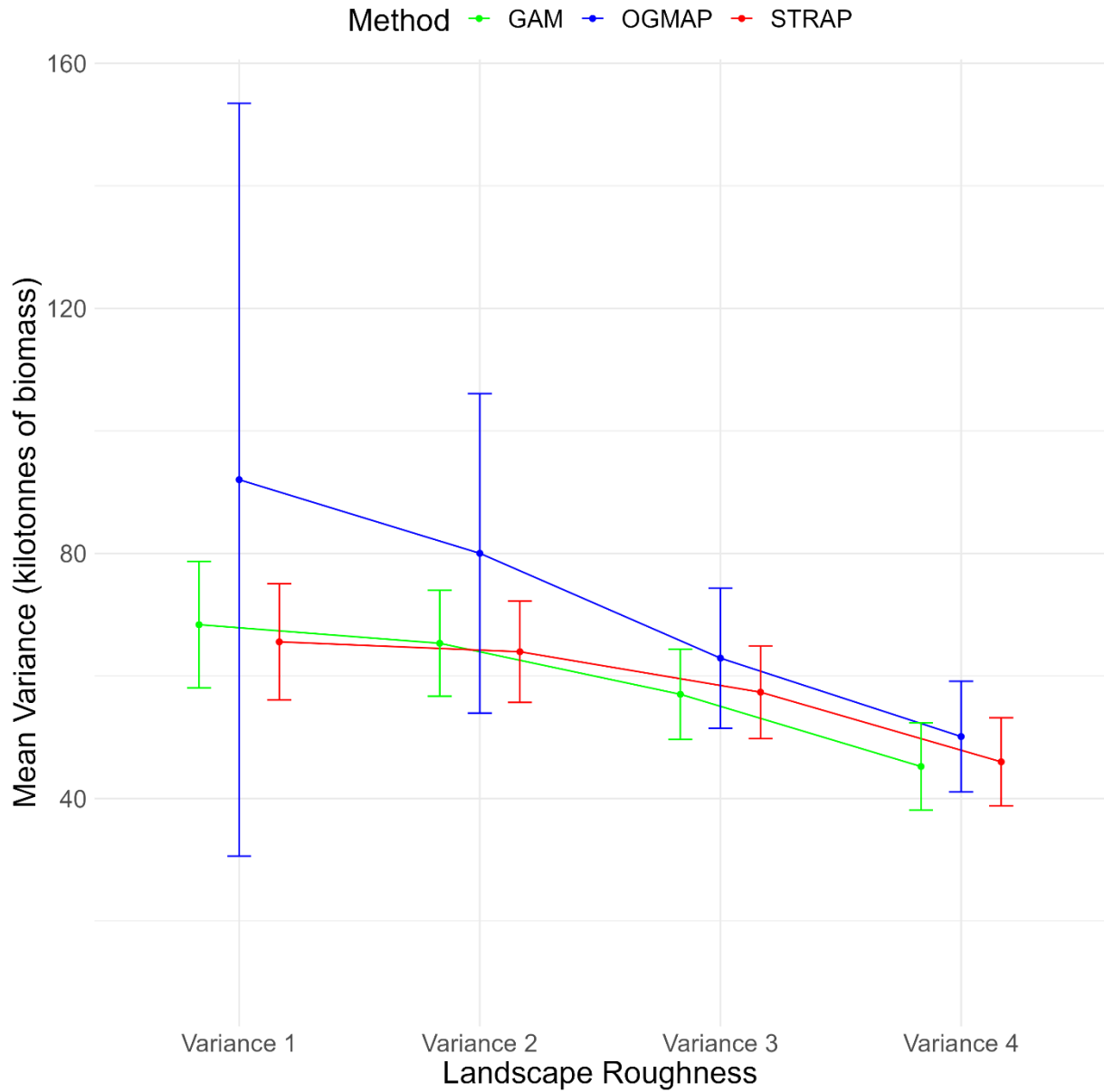


63

64 *Figure 10: Percentage bias of GAM (green), OGMAP (blue), and STRAP (red) methods across ascending variation levels from 1*
65 *to 4 (low to high), with error bars denoting the standard deviation (SD) of the percentage bias across different landscapes with*
66 *the same level of landscape roughness.*

67

68 A comparable trend emerges when analyzing OGMAP. However, it's important to note that OGMAP
69 exhibits a wider spectrum of variance outcomes. This suggests that OGMAP's estimates tend to deviate
70 further from the true mean compared to either GAM or STRAP, especially in scenarios with higher levels
71 of variation (Figure 11).

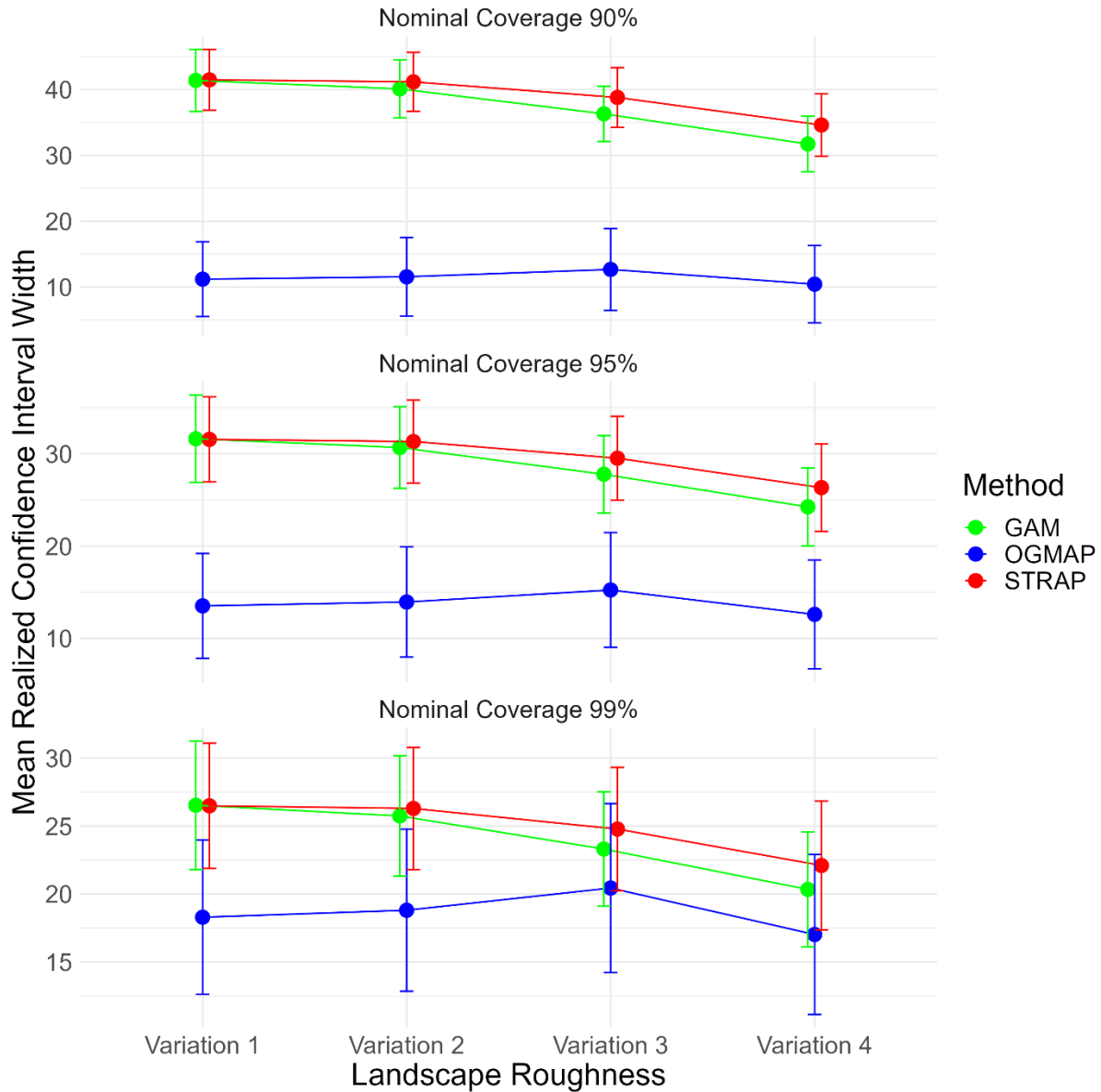


72

73 *Figure 11: Variance of each method, OGMAP (blue), GAM (green), and STRAP (red), according to the changing landscape*
 74 *roughness levels from low to high (1-4). The error bars represent the standard deviation (SD) of the variance.*

75

76 The width of CIs gradually contracts with increased landscape variation for all three methods (Figure 13).
 77 Upon examining the 99% CI, I find that the narrowest intervals consistently occur at the fourth variation
 78 level, characterized by rough terrain.



79

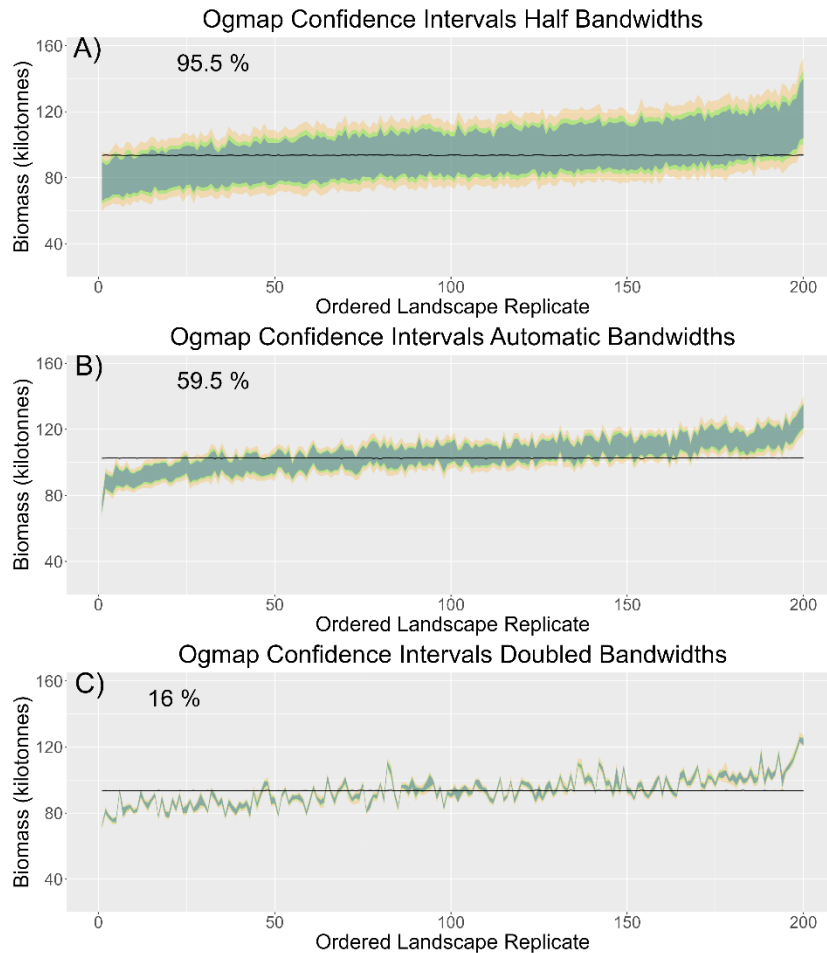
80 *Figure 12 : Average width of Confidence Intervals (CI), representing the resulting Confidence Interval Widths across CI levels,*
 81 *Variations 1-4 (low to high roughness), and methods (GAM in green, OGMAP in blue, STRAP in red).*

82 However, OGMAP distinguishes itself by consistently exhibiting the same trend as the other methods
 83 while also boasting significantly tighter CIs across all levels of variation. This distinctive behaviour
 84 underscores how the OGMAP method estimates respond to variations in landscape characteristics, where
 85 extreme values exert a significant influence on the resulting CIs. This influence is evident in the narrower
 86 CIs observed in rougher landscapes at the same variation level.

87 **8. Effect of OGMAP's Bandwidth:**

88 When I increased the bandwidths for a single landscape relative to OGMAP's estimated bandwidths, I
89 essentially incorporated a broader range of depth and spatial positions into the biomass estimates for any
90 given location. However, this led to narrower CIs and subsequently lower realized coverage compared to
91 the bandwidths estimated by OGMAP (Figure 14A). This phenomenon suggests that the larger
92 bandwidths are considering a wider array of values, which, paradoxically, results in under-smoothing, as
93 evidenced by the narrower CIs.

94



95

96 *Figure 13 : Effects of the OGMAP bandwidth on the resulting confidence intervals at levels of 90% (blue), 95% (green), and*
97 *99% (orange). The values are ordered by increasing values of model point estimates. The black line represents the absolute true*
98 *biomass of each Landscape.*

99 On the other hand, when I decreased the bandwidths relative to OGMAP's estimates (Figure 14C), I
100 obtained a much smoother estimate (than Figure 14B), resembling the coverage pattern observed in
101 GAMs (Figure 8B). This reduction in bandwidth meant that fewer values were taken into account,
102 assigning higher weights to individual measurements across the estimated field. Consequently, this led to
103 less smoothing and wider total confidence intervals.

104 As a result, GAMs displayed more consistent and reliable interval coverage across different sample sizes,
105 showcasing their robustness in handling increased sampling intensity.

106 **Discussion**

107

108 In this research, I examined the impact of landscape characteristics on the performance of the chosen
109 modelling techniques with the aim of enhancing the reliability of shrimp biomass estimates which is a key
110 metric for their effective management. I identified the role played by the smoothness or roughness of
111 depth transitions in influencing the performance of the models. I observed a wide spectrum of CI
112 behaviours among methods, revealing disparities in their ability to accurately estimate shrimp quantities.
113 OGMAP's realized coverage consistently fell below the nominal coverage threshold across all
114 simulations, whereas CIs calculated for GAM and STRAP models consistently achieved realized
115 coverage levels close to their nominal rates.

116 The STRAP model had the lowest bias across simulations, while the analysis of GAMs showed weak
117 bias, which decreased at higher variation levels. In contrast, OGMAP consistently underestimated the true
118 biomass despite showing a reduction in bias as variation levels were increased (Figure 11). However,
119 none of the three models showed a high degree of overall bias, with average bias below 2% regardless of
120 the method. The examination of variance revealed comparable decreasing trends in both STRAP and
121 GAM, while OGMAP exhibited a similar pattern but with initially higher variance (Figure 12). OGMAP
122 consistently demonstrated narrower CIs across variation levels compared to the other two methods
123 (Figure 12). As such, the poor nominal coverage of OGMAP intervals was likely not due to systematic
124 biases in estimated biomass, but instead because OGMAP CIs were overly narrow (i.e. falsely precise).

125 In my simulations the widths of OGMAP's CI were narrower on average in landscapes with greater
126 spatial variation, resulting in reduced realized coverage. This behaviour contrasted with the stable CI
127 fluctuations observed in both GAM and STRAP methods when variation levels were adjusted.
128 Furthermore, my exploration extended to sampling density, where I found that higher sampling intensity
129 corresponded to even lower levels of nominal coverage for OGMAP CIs.

130 In my comparative analysis of biomass estimation methods, one notable observation was the relative
131 underperformance of OGMAP in contrast to other techniques employed in my study. While OGMAP
132 displayed tighter CIs compared to all other methods, which could be a sign of more accurate results, it
133 turned out that OGMAP consistently lagged behind alternative methods when assessing the accuracy of
134 biomass estimates. Specifically, the realized coverage of OGMAP consistently fell behind that of both
135 GAMs and STRAP. This unexpected underperformance raised questions regarding the suitability of
136 OGMAP for this specific research objective. Consequently, I delved deeper into the method's underlying
137 assumptions and limitations, shedding light on the necessity of a nuanced approach in selecting the most
138 appropriate methodology for biomass estimation in this study. These findings highlight the importance of
139 method selection and how the choice of approach can significantly impact the outcomes of biomass
140 estimation studies (Patterson et al., 2001).

141 The bandwidth parameter for OGMAP emerged as the most important factor that influences the biomass
142 estimate outcome. This parameter dictates the scale at which data are considered and weighted during
143 estimation. Opting for a smaller bandwidth enhances spatial resolution by assigning greater significance
144 to localized data points, making it adept at capturing fine-scale spatial variations. Importantly, it addresses
145 the challenge of dealing with extreme data points, such as unusually high catches, by downplaying their
146 impact through proximity-based weighting. Furthermore, the bandwidth parameter has a direct bearing on
147 the width of CIs around biomass estimates, with smaller bandwidths resulting in larger intervals. Halving
148 the bandwidths notably improved OGMAP's performance, increasing its coverage from 51.3% to over
149 92% at the 95% confidence level. This highlights the pivotal role of bandwidth settings in OGMAP,

150 influencing estimates and CIs. This investigation underscored the critical role of bandwidth settings in
151 OGMAP, emphasizing their influence on probability distribution estimates and CIs.

152 One of the key insights I obtained from this analysis was that larger bandwidth settings led to the
153 generation of overly narrower CIs, due to under-smoothing, resulting in overfitting the observed data. I
154 also noted using narrower bandwidths than those estimated by OGMAP had the opposite effect, causing
155 an increase in the width of the CI during model fitting. As such, the low realized coverage of OGMAP is
156 likely due to the method that OGMAP currently uses for bandwidth estimation overestimating the
157 bandwidth.

158 This highlights that the procedure OGMAP currently uses for bandwidth estimation is over-estimating the
159 width of bandwidths for biomass, resulting in over-smoothing of biomass, and thus generating falsely low
160 estimates of spatial biomass variability. In the context of fisheries science, where accurate and reliable
161 estimates are crucial for effective management and decision-making, accurate estimates of model
162 uncertainty are particularly important. Unrealistically narrow CIs may result in overconfidence in the
163 precision of estimates, potentially leading to suboptimal management decisions (e.g. setting TACs higher
164 or lower than they should be given the true biomass). My findings suggest that optimizing OGMAP's
165 bandwidth selection process, especially in scenarios with higher sampling intensity, holds the potential to
166 enhance its interval coverage and overall performance.

167 OGMAPs consistent underperformance compared to GAM and STRAP reflects the broader concern
168 highlighted in the introduction regarding uncertainty in fisheries management. These disparities amongst
169 the techniques underscore the ongoing challenge of accurately estimating shrimp quantities, emphasizing
170 the pressing need to address uncertainties in stock assessments (Charles, 1998).

171 In my analysis of bias, I found distinctions that shed light on the reliability and performance of each
172 approach. STRAP emerged as a particularly dependable method, exhibiting minimal bias across each
173 research scenarios. What sets STRAP apart is its ability to consistently maintain either a positive or
174 neutral bias, effectively steering clear of under/overestimation pitfalls. This reliability factor is a
175 significant asset, ensuring that STRAP's estimates are generally unbiased. While STRAP is unbiased I
176 need to note that STRAP has a major limitation which does not affect OGMAP or GAM. STRAP requires
177 samples covering the entire studied area, so missing trawls have severe repercussions on this technique
178 leading to loss of spatial information and thus inaccurate estimates. In contrast, when I examined the
179 performance of GAMs, I observed a predominantly unbiased pattern when averaging their estimates
180 across the dataset. This suggests that, on average, GAMs provide estimates that are relatively close to the
181 actual values, making them a reasonable choice for biomass estimation across a range of spatial
182 conditions. Additionally, the incorporation of additional environmental variables into GAM models
183 present promising avenues for advancing our understanding of spatial modeling in ecology and
184 contributing to more accurate and reliable estimation of ecological parameters.

185 Ensuring that methods used to produce biomass estimates produce reliable measures of statistical
186 uncertainty such as CIs is crucial for effective fisheries management, as it lays the foundation for
187 establishing proper guidelines that promote the conservation and sustainability of managed stocks
188 (Martinet et al., 2016). The OGMAP method, as previously mentioned, currently serves as the standard
189 approach employed by DFO for estimating Northern Shrimp. OGMAP was developed as a response to
190 dealing with non-normally distributed data, offering a potential solution to the challenges posed by such
191 data. It also showed much tighter CIs compared to STRAP, which was considered to be one benefit of the
192 approach when first introduced (Orr, 2004; Skanes and Evans, 2005). However, as underlined, the effect

193 and range of bandwidths had yet to be thoroughly explored at that time. This work indicates that OGMAP
194 CIs, as currently calculated, may be too narrow to have the advertised (nominal) coverage rates.

195 The importance of having reliable estimates is a well-established concept in ecological research
196 (Vallecillo et al., 2021). It is a fundamental aspect that allows us to gain a deeper understanding of the
197 populations and interactions within the landscape, enhancing the credibility of the results obtained from
198 the models. This emphasis on reliability is not a recent development but has been a consistent theme in
199 fisheries management and ecological studies. Recent research (Benoît & Allard, 2020; Jacques Allard &
200 Hugues P. Benoît, 2019) has emphasized the importance of dependable estimates in fisheries
201 management.

202 Reliable estimates are essential for the conservation and sustainability of managed stocks (Charles, 1998;
203 Sethi et al., 2005). This simulation based approach provides valuable insights into estimating the
204 confidence of crucial parameters, particularly biomass. Importantly, I demonstrated that OGMAP
205 consistently yields unreliable measures of the precision of biomass estimate, prompting a reconsideration
206 of the stock assessment framework employed by the DFO, where OGMAP serves as a key model for
207 decision-making.

208 This issue becomes especially critical when considering the Precautionary Approach, where the reference
209 point serves as the threshold determining whether the fishery should be closed (Fisheries and Oceans
210 Canada, 2006). Overly narrow CIs may inadvertently allow the fishing of species below the reference
211 point, violating principles of sustainable fisheries management. These findings underscore the importance
212 of robust estimation methods to ensure that decisions align with the conservation goals outlined in the
213 Precautionary Approach. By offering a robust assessment of my studied models, I contribute essential
214 information for making informed TAC decisions and aligning those TAC limits with the ecological
215 conditions of the target species.

216 Given these substantial implications, I strongly recommend that the Department of Fisheries and Oceans
217 (DFO) review the use of OGMAP in light of these findings and consider alternative model-based spatial
218 stock assessment approaches for biomass estimation. By reevaluating the model's appropriateness and
219 considering alternative methods, DFO can enhance the reliability of its stock assessments and ensure
220 effective fisheries management. This research provides a foundation for this reconsideration, emphasizing
221 the need for precision and reliability in estimating ecological parameters used for making informed
222 decision- in the dynamic field of fisheries management.

223 **References**

- 224 Altman, D. G. (2005). Why we need confidence intervals. *World Journal of Surgery*, 29(5), 554–556.
225 <https://doi.org/10.1007/s00268-005-7911-0>
- 226 Anderson, R. P. (2017). When and how should biotic interactions be considered in models of species
227 niches and distributions? *Journal of Biogeography*, 44(1), 8–17. <https://doi.org/10.1111/jbi.12825>
- 228 Bavington, D. (2011). *Managed Annihilation: An Unnatural History of the Newfoundland Cod Collapse*.
229 UBC Press.
- 230 Benoît, H. P., & Allard, J. (2020). The dependability of fishery monitoring programs: Harmonising the
231 quality of estimates with the risks to the conservation of aquatic populations. *DFO*.
- 232 Berkson, J. M., Kline, L. L., Orth, D. J., & American Fisheries Society (Eds.). (2002). *Incorporating*
233 *uncertainty into fishery models: Proceedings of the Workshop Incorporating Uncertainty into*
234 *Fishery Models held at Jekyll Island, Georgia, USA, 28-29 June 1999*. American Fisheries
235 Society.
- 236 Bishara, A. J., & Hittner, J. B. (2015). Reducing Bias and Error in the Correlation Coefficient Due to
237 Nonnormality. *Educational and Psychological Measurement*, 75(5), 785–804.
238 <https://doi.org/10.1177/0013164414557639>
- 239 Chadwick, E., Brodie, W., Colbourne, E., Clark, D., Gascon, D., & Hurlbut, T. (2007). History of Annual
240 Multi-Species Trawl Surveys on the Atlantic Coast of Canada. *AZMP Bull.*, 6.
- 241 Charles, A. T. (1994). Towards sustainability: The fishery experience. *Ecological Economics*, 11(3), 201–
242 211. [https://doi.org/10.1016/0921-8009\(94\)90201-1](https://doi.org/10.1016/0921-8009(94)90201-1)
- 243 Charles, A. T. (1998). Living with uncertainty in fisheries: Analytical methods, management priorities
244 and the Canadian groundfishery experience. *Fisheries Research*, 37(1), 37–50.
245 [https://doi.org/10.1016/S0165-7836\(98\)00125-8](https://doi.org/10.1016/S0165-7836(98)00125-8)
- 246 Colbourne, E., Holden, J., Senciall, D., Bailey, W., Snook, S., & Higdon, J. (2016). *Physical*
247 *Oceanographic Conditions on the Newfoundland and Labrador Shelf during 2015*.
- 248 Dodge, Y. (2003). *The Oxford Dictionary of Statistical Terms*. Oxford University Press.

249 Duffy, J. E. (2010). Crustacean Social Evolution. In M. D. Breed & J. Moore (Eds.), *Encyclopedia of*
250 *Animal Behavior* (pp. 421–429). Academic Press. [https://doi.org/10.1016/B978-0-08-045337-](https://doi.org/10.1016/B978-0-08-045337-8.00350-8)
251 [8.00350-8](https://doi.org/10.1016/B978-0-08-045337-8.00350-8)

252 Evans, G. T., Parsons, D. G., Veitch, P. J., & Orr, D. C. (2000). A Local-influence Method of Estimating
253 Biomass from Trawl Surveys, with Monte Carlo Confidence Intervals. *Journal of Northwest*
254 *Atlantic Fishery Science*, 27, 133–138. <https://doi.org/10.2960/J.v27.a12>

255 Fisheries and Oceans Canada. (2006). *A harvest strategy compliant with the precautionary approach*
256 (Can. Sci. Advis. Sec. Sci. Advis. Rep. 2006/023). [https://waves-vagues.dfo-mpo.gc.ca/library-](https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/322464.pdf)
257 [bibliotheque/322464.pdf](https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/322464.pdf)

258 Fisheries and Oceans Canada. (2017). *An assessment of Northern Shrimp (Pandalus borealis) in Shrimp*
259 *Fishing Areas 4-6 in 2017* (Can. Sci. Advis. Sec. Sci. Advis. Rep. 2018/018).

260 Fisheries and Oceans Canada. (2018). *Northern shrimp and striped shrimp – Shrimpfishing areas 0, 1, 4-*
261 *7, the Eastern and Western Assessment Zones and North Atlantic Fisheries Organization (NAFO)*
262 *Division 3M* (Integrated Fisheries Management Plans 2018). [https://www.dfo-](https://www.dfo-mpo.gc.ca/fisheries-peches/ifmp-gmp/shrimp-crevette/shrimp-crevette-2018-002-eng.html)
263 [mpo.gc.ca/fisheries-peches/ifmp-gmp/shrimp-crevette/shrimp-crevette-2018-002-eng.html](https://www.dfo-mpo.gc.ca/fisheries-peches/ifmp-gmp/shrimp-crevette/shrimp-crevette-2018-002-eng.html)

264 Fisheries and Oceans Canada. (2019). *An Assessment of Northern Shrimp (Pandalus borealis) in Shrimp*
265 *Fishing Areas 4-6 in 2019* (Can. Sci. Advis. Sec. Sci. Advis. Rep. 2021/010).

266 Fisheries and Oceans Canada. (2023). *Assessment of Northern Shrimp (Pandalus borealis) in Shrimp*
267 *Fishing Areas 4-6 in 2021* (Can. Sci. Advis. Sec. Sci. Advis. Rep. 2023/038).

268 Fournier, A., Fussell, D., & Carpenter, L. (1982). Computer rendering of stochastic models.
269 *Communications of the ACM*, 25(6), 371–384. <https://doi.org/10.1145/358523.358553>

270 Geoffrey Evans. (Unpublished). *Ogive mapping*.

271 Hansen, N., & Ostermeier, A. (1996). *Adapting arbitrary normal mutation distributions in evolution*
272 *strategies: The covariance matrix adaptation*. 312–317.
273 <https://doi.org/10.1109/ICEC.1996.542381>

274 Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, 1(3), 297–310.
275 <https://doi.org/10.1214/ss/1177013604>

276 Healey, B., McKenzie, C. H., & Boyd, S. (2019). *Proceedings of the Regional Peer Review for the*
277 *Assessment of Northern Shrimp in Shrimp Fishing Areas (SFAs) 4, 5 and 6.*

278 Hilborn, R. (1987). Living with Uncertainty in Resource Management. *North American Journal of*
279 *Fisheries Management*, 7(1), 1–5. <https://doi.org/10.1577/1548->
280 [8659\(1987\)7<1:LWUIRM>2.0.CO;2](https://doi.org/10.1577/1548-8659(1987)7<1:LWUIRM>2.0.CO;2)

281 Hua, C., Zhu, Q., Shi, Y., & Liu, Y. (2019). Comparative analysis of CPUE standardization of Chinese
282 Pacific saury (*Cololabis saira*) fishery based on GLM and GAM. *Acta Oceanologica Sinica*,
283 38(10), 100–110. <https://doi.org/10.1007/s13131-019-1486-3>

284 Hutchings, J. A., & Rangeley, R. W. (2011). Correlates of recovery for Canadian Atlantic cod (*Gadus*
285 *morhua*). *Canadian Journal of Zoology*, 89(5), 386–400. <https://doi.org/10.1139/z11-022>

286 Jacques Allard & Hugues P. Benoît. (2019). Unified framework for the statistical assessment of fishery
287 monitoring programs. *DFO*.

288 Kilduff, P., Carmichael, J., Latour, R., & Berger, T. L. (2009). Guide to Fisheries Science and Stock
289 Assessments. *Atlantic States Marine Fisheries Commission*, 74.

290 Lilly, G. R., Parsons, D. G., & Kulka, D. W. (2000). Was the Increase in Shrimp Biomass on the
291 Northeast Newfoundland Shelf a Consequence of a Release in Predation Pressure from Cod?
292 *Journal of Northwest Atlantic Fishery Science*, 27, 45–61. <https://doi.org/10.2960/J.v27.a5>

293 Lucet, V., & Pedersen, E. J. (2023). The sspm R package for spatially-explicit surplus production
294 population models. *Journal of Open Source Software*, 8(86), 4724.
295 <https://doi.org/10.21105/joss.04724>

296 Magnusson, A., & Hilborn, R. (2007). What makes fisheries data informative? *Fish and Fisheries*, 8(4),
297 337–358. <https://doi.org/10.1111/j.1467-2979.2007.00258.x>

298 Mangel, M., & Clark, C. W. (1983). Uncertainty, search, and information in fisheries. *ICES Journal of*
299 *Marine Science*, 41(1), 93–103. <https://doi.org/10.1093/icesjms/41.1.93>

300 Martinet, V., Peña-Torres, J., De Lara, M., & Ramírez C., H. (2016). Risk and Sustainability: Assessing
301 Fishery Management Strategies. *Environmental and Resource Economics*, 64(4), 683–707.
302 <https://doi.org/10.1007/s10640-015-9894-0>

303 Miller, G. S. P. (1986). The definition and rendering of terrain maps. *ACM SIGGRAPH Computer*
304 *Graphics*, 20(4), 39–48. <https://doi.org/10.1145/15886.15890>

305 Orr, D. C. (2004). OGIVE MAPping (ogmap) as an Alternative Means of Estimating Indices and Setting
306 TACs. *DFO*, 20.

307 Palomares, M. L. D., Froese, R., Derrick, B., Meeuwig, J. J., Noël, S.-L., Tsui, G., Woroniak, J., Zeller,
308 D., & Pauly, D. (2020). Fishery biomass trends of exploited fish populations in marine
309 ecoregions, climatic zones and ocean basins. *Estuarine, Coastal and Shelf Science*, 243, 106896.
310 <https://doi.org/10.1016/j.ecss.2020.106896>

311 Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V.,
312 Skagen, D. W., & Stefánsson, G. (2001). Estimating uncertainty in fish stock assessment and
313 forecasting. *Fish and Fisheries*, 2(2), 125–157. <https://doi.org/10.1046/j.1467-2960.2001.00042.x>

314 Pedersen, E. J., Koen-Alonso, M., & Tunney, T. D. (2020). Detecting regime shifts in communities using
315 estimated rates of change. *ICES Journal of Marine Science*, 77(4), 1546–1555.
316 <https://doi.org/10.1093/icesjms/fsaa056>

317 Pedersen, E. J., Skanes, K., Corre, N. L., Alonso, M. K., & Baker, K. D. (2022). A New Spatial
318 Ecosystem-Based Surplus Production Model for Northern Shrimp in Shrimp Fishing Areas 4 to 6.
319 *DFO*.

320 Potts, S. E., & Rose, K. A. (2018). Evaluation of GLM and GAM for estimating population indices from
321 fishery independent surveys. *Fisheries Research*, 208, 167–178.
322 <https://doi.org/10.1016/j.fishres.2018.07.016>

323 Privitera-Johnson, K. M., & Punt, A. E. (2020a). A review of approaches to quantifying uncertainty in
324 fisheries stock assessments. *Fisheries Research*, 226, 105503.
325 <https://doi.org/10.1016/j.fishres.2020.105503>

326 Privitera-Johnson, K. M., & Punt, A. E. (2020b). Leveraging scientific uncertainty in fisheries
327 management for estimating among-assessment variation in overfishing limits. *ICES Journal of*
328 *Marine Science*, 77(2), 515–526. <https://doi.org/10.1093/icesjms/fsz237>

329 *Pycma*. (2023). [Python]. CMA-ES. <https://github.com/CMA-ES/pycma> (Original work published 2016)

330 R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for
331 Statistical Computing. <https://www.R-project.org/>

332 Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized Additive Models for Location, Scale and
333 Shape. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 54(3), 507–554.
334 <https://doi.org/10.1111/j.1467-9876.2005.00510.x>

335 Ruckelshaus, M., Doney, S. C., Galindo, H. M., Barry, J. P., Chan, F., Duffy, J. E., English, C. A.,
336 Gaines, S. D., Grebmeier, J. M., Hollowed, A. B., Knowlton, N., Polovina, J., Rabalais, N. N.,
337 Sydeman, W. J., & Talley, L. D. (2013). Securing ocean benefits for society in the face of climate
338 change. *Marine Policy*, 40, 154–159. <https://doi.org/10.1016/j.marpol.2013.01.009>

339 S. J. Smith & G. D. Somerton. (1981). *STRAP: A User-Oriented Computer Analysis System for*
340 *Groundfish Research Trawl Survey Data*. DFO.

341 Saura, S., & Martínez-Millán, J. (2000). Landscape patterns simulation with a modified random clusters
342 method. *Landscape Ecology*, 15(7), 661–678. <https://doi.org/10.1023/A:1008107902848>

343 Schrank, W. E. (2007). Is there any hope for fisheries management? *Marine Policy*, 31(3), 299–307.
344 <https://doi.org/10.1016/j.marpol.2006.09.001>

345 Sciaini, M., Fritsch, M., Scherer, C., & Simpkins, C. E. (2018). NLMR and landscapetools: An integrated
346 environment for simulating and modifying neutral landscape models in R. *Methods in Ecology*
347 *and Evolution*, 9(11), 2240–2248. <https://doi.org/10.1111/2041-210X.13076>

348 Sethi, G., Costello, C., Fisher, A., Hanemann, M., & Karp, L. (2005). Fishery management under multiple
349 uncertainty. *Journal of Environmental Economics and Management*, 50(2), 300–318.
350 <https://doi.org/10.1016/j.jeem.2004.11.005>

351 Shertzer, K. W., & Prager, M. H. (2007). Delay in fishery management: Diminished yield, longer
352 rebuilding, and increased probability of stock collapse. *ICES Journal of Marine Science*, 64(1),
353 149–159. <https://doi.org/10.1093/icesjms/fsl005>

354 Shono, H. (2008). Application of the Tweedie distribution to zero-catch data in CPUE analysis. *Fisheries*
355 *Research*, 93(1), 154–162. <https://doi.org/10.1016/j.fishres.2008.03.006>

356 Skanes, K., & Evans, G. (2005). *Testing OGMAP as a Tool for Estimating Biomass and Abundance of*
357 *Northern Shrimp* (p. 3).

358 Smith, M. T., & Addison, J. T. (2003). Methods for stock assessment of crustacean fisheries. *Fisheries*
359 *Research*, 65(1), 231–256. <https://doi.org/10.1016/j.fishres.2003.09.017>

360 Szuwalski, C. S., & Hollowed, A. B. (2016). Climate change and non-stationary population processes in
361 fisheries management. *ICES Journal of Marine Science*, 73(5), 1297–1305.
362 <https://doi.org/10.1093/icesjms/fsv229>

363 Vallecillo, D., Gauthier-Clerc, M., Guillemain, M., Vittecoq, M., Vandewalle, P., Roche, B., &
364 Champagnon, J. (2021). Reliability of animal counts and implications for the interpretation of
365 trends. *Ecology and Evolution*, 11(5), 2249–2260. <https://doi.org/10.1002/ece3.7191>

366 Verma, J. T. (2019). *Cod Collapse: The Rise and Fall of Newfoundland's Saltwater Cowboys*. Nimbus
367 Publishing Limited.

368 Wackerly, D., Mendenhall, W., & Scheaffer, R. L. (2014). *Mathematical Statistics with Applications*.
369 Cengage Learning.

370 Walters, C., & Maguire, J.-J. (1996). Lessons for stock assessment from the northern cod collapse.
371 *Reviews in Fish Biology and Fisheries*, 6(2). <https://doi.org/10.1007/BF00182340>

372 Wood. (2003). Thin Plate Regression Splines. *Journal of the Royal Statistical Society Series B: Statistical*
373 *Methodology*, 65(1), 95–114. <https://doi.org/10.1111/1467-9868.00374>

374 Wood. (2006). Low-rank scale-invariant tensor product smooths for generalized additive mixed models.
375 *Biometrics*, 62(4), 1025–1036. <https://doi.org/10.1111/j.1541-0420.2006.00574.x>

376 Wood. (2015). *Core Statistics*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107741973>

- 377 Wood. (2017). *Generalized Additive Models: An Introduction with R, Second Edition*. CRC Press.
- 378 Wood. (2023). *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation (1.9-0)*
- 379 [Computer software]. <https://cran.r-project.org/web/packages/mgcv/index.html>
- 380