Summing the parts: Improving population estimates using a state-space multispecies production model

Paul M. Regular1\*, Mariano Koen-Alonso1, M. Joanne Morgan1, Pierre Pepin1, Rick M. Rideout1

1 Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, 80 East White Hills, St. John’s, Newfoundland and Labrador, A1C 5X1, Canada  
  
\*Corresponding author  
E-mail: [Paul.Regular@dfo-mpo.gc.ca](mailto:Paul.Regular@dfo-mpo.gc.ca)  
  
2023-03-22

**Abstract:** …  
  
**Keywords:** …

# Introduction

Paragraph ideas:

* Carrying capacity, the concept that a environment can only accommodate a finite number of specific species, is a foundational concept in applied population ecology (e.g., use of MSY in fisheries management).
* Usefulness of this concept has been called into question (e.g., changing K in variable environments). Standard formulations also tend to focus on single-species and ignore species interactions.
* We propose an extension of single-species production model that assumes that inter and intra-specific competition contributes to density-dependent effects within ecosystem production units.
* Apply the model to commercially important demersal fish species off the east coast of Canada.

# Methods

## Model formulation

Trends in fish populations have frequently been described using state-space production models of the form

where is biomass at the start of year , is the catch through year , is production as a function of biomass, is an index of relative abundance in year from survey , is the time-invariant catchability coefficient for survey index , is process error, and is observation error. Statistical challenges aside, the most difficult aspect of this model to parameterize is the production function as it needs to capture changes caused by growth, recruitment, and natural mortality. Schaefer ([1954](#ref-schaefer1954)) proposed a solution by applying the logistic equation to describe self-limiting growth,

where is the maximum per-capita rate of change and is the carrying capacity. That is, a populations’ intrinsic ability to grow () is limited by the size of the current population relative to the maximum biomass the system can support (). While this formulation offers an elegant description of single-species population dynamics, it assumes that density-dependent effects are solely caused by intraspecific competition and ignores the potential effects of other species inhabiting the same ecological area. We present an extension of equation (3) that attempts to account for intra and interspecific competition by assuming that density-dependent effects are incurred when the total biomass of multiple species, represented by , exceeds the capacity of the system,

While intrinsic rates of growth may vary across species, this formulation implies that the growth of all species is ultimately limited by the finite amount of energy in a region (i.e., as the population in the system increases towards , year-over-year growth of all species slows). Combining equations (1), (2), and (4), our model becomes

The inclusion of multiple species in the model permits the estimation of covariance. While covarying changes in observed populations may be described using observation errors, we assume that most covariance stems from population processes. We therefore assume that observation errors are independent and normal distribution such that , where the standard deviation parameter represents species and survey specific levels of observation error. A more flexible error structure is used to describe the process errors as ecological processes may contribute to species or temporal dependencies. Beyond the global limiting effect of , species interactions may elicit positive or negative population responses resulting from direct or indirect associations. We therefore apply the multivariate normal distribution to account for the possibility that process errors are not independent across species. Deviations from expected production may also display temporal dependence if the factors contributing to the process errors change slowly through time. Such inertia may cause positive or negative process errors to persist across several years. A first-order autoregressive (AR1) process was therefore applied to account for temporal dependence. Both sources of dependence are modeled using a multivariate AR1 process where

and

The degree of temporal correlation is controlled by , where values between 0 to 1 represent low to high correlation, and species-to-species correlations are described by , where values between -1 to 1 represent negative to positive correlation. This is a flexible structure that allows for the testing of alternate hypotheses that process errors are independent through time or across species (i.e, or ). The possibility that process errors are similarly correlated across all species may also be tested by estimating only one parameter. Finally, the magnitude of the process error deviations is controlled by the species-specific standard deviation parameters, .

Minor extensions of the formulation also permit the fitting of covariates which may describe an underlying linear effect. Two options were implemented, one that affects the process errors by substituting in equation (5) with , and another that affects the carrying capacity by substituting in equation (5) with , where parameters capture linear effects of covariates included in design matrices, . The idea is that some factors may affect positive or negative changes in the populations while others may affect change in the total carrying capacity of the system. A covariate option for intrinsic rates of increase was not implemented as one goal of this model is to obtain estimates of this vital rate, which is not expected to change rapidly as it is shaped by natural selection ([Hutchings, 2021](#ref-hutchings2021)). The formulation was also modified to fit the single-species Schaefer production function by dropping the summation of biomass in equation (5) and estimating species-specific carrying capacities, (i.e., apply equation (3) indexed by species).

## Statistical framework

This model was implemented using template model builder (TMB; [Kristensen et al., 2016](#ref-kristensen2015)), which is package for R ([R Core Team, 2021](#ref-R)) that enables the fitting of complex nonlinear random effects such as the latent variable in state-space production models (equation (1)). Such variables are not directly measured but are inferred indirectly via observed values. Data fitting is accomplished using a combination of Laplace approximation and automatic differentiation to evaluate the joint likelihood ([Kristensen et al., 2016](#ref-kristensen2015)). Like the production model described by Pedersen et al. ([2017](#ref-pedersen2017)), both frequentist and Bayesian inference of model parameters are possible. In development, we found that estimation was generally more successful when vaguely informative priors are specified as parameters were, in some cases, not identifiable when unconstrained.

## Case study

### Data

The multispecies production model described above requires two basic inputs for one or more species in a region: a time-series of catch ( in equation (5)), and an index of population size ( in equation (6)). The Northwest Atlantic Fisheries Organization (NAFO) and Fisheries and Oceans Canada (DFO) have been collecting and curating such information for multiple fish populations along the shelves of Newfoundland and Labrador (NL) since the 1970s. The communities inhabiting these shelves can be divided into several regions with distinct productivity [i.e. ecosystem production units; Pepin et al. ([2014](#ref-pepin2014))]. For our case study, we tallied catch data and calculated survey indices of multiple demersal fish populations from three regions: 1) the Northeast NL Shelf (NAFO divisions 2J3K), 2) the Grand Bank (NAFO divisions 3LNO), and 3) Southern NL (NAFO sub-division 3Ps; Figure 1).

Catch data were extracted from the STATLANT 21A database maintained by NAFO (<https://www.nafo.int/Data/STATLANT-21A>, accessed 2022-01-21) and aggregated by region, species, and year. Survey indices were derived from the standardized, stratified random bottom-trawl surveys conducted each spring and fall by DFO; this is perhaps the largest fisheries-independent survey conducted in the world, which aims to cover more than 500,000 km2 annually (roughly the size of Sweden or the Yukon, Canada) to depth up to 1500 m. Since the inception of this program in 1971, survey platforms and protocols have undergone a series of changes that affect the continuity of the data collected in each region and season. A Yankee then Engel otter trawl, with nets designed to catch large demersal fish, were used between 1971 to 1994. In 1995, both survey vessels and gear were changed. Since then, a Campelen shrimp trawl has been used and, because of its smaller mesh size, a broader range of species and size groups have been captured ([Chadwick et al., 2007](#ref-chadwick2007)). Within each era of the survey (Yankee, Engel, or Campelen) and for each season and region, samples were limited to strata that were covered most years (> 80%) and to species found across more than 10% of these core strata. This often resulted in the exclusion of strata >750 m as these areas have been inconsistently covered by the survey. Stratified analyses ([Smith & Somerton, 1981](#ref-smith1981)) were then conducted on the remaining species to obtain indices of total biomass. To minimize bias introduced by inconsistent survey coverage, indices from years where more than 20% of the biomass was likely missed, inferred from time averaged percent occupancy within strata, were excluded from our analysis [*sensu* NAFO guidelines; page 10, NAFO ([2019](#ref-nafo2019))]. Finally, species were ranked by cumulative commercial catch and limited to the the seven most commonly caught species, or species group when catch was not consistently distinguished, within each region. On the Northeast NL Shelf, species include Redfish spp. (*Sebastes fasciatus* or *S. mentella*), Wolffish spp. (*Anarhichas lupus* or *A. minor*), Witch Flounder (*Glyptocephalus cynoglossus*), American Plaice (*Hippoglossoides platessoides*), Greenland Halibut (*Reinhardtius hippoglossoides*), Atlantic Cod (*Gadus morhua*), and Skate spp. (*Amblyraja radiata* or *Malacoraja senta*). On the Grand Bank, species include Redfish spp., Yellowtail Flounder (*Limanda ferruginea*), American Plaice, Greenland Halibut, Haddock (*Melanogrammus aeglefinus*), and Atlantic Cod. Finally, along Sourthern NL, species include Redfish spp., Witch Flounder, American Plaice, White hake (*Urophycis tenuis*), Haddock, Atlantic Cod, and Skate spp.

### Priors

For simplicity, all priors were normally distributed and, in most cases, upper and lower inflection points (; ~68% of the total area under the curve) of each normal prior were defined using values in log or logit space. Upper and lower values were based on previous research or knowledge to impose fairly generic and vaguely informative priors.

#### *Intrinsic growth rate, , and carrying capacity,*

To capture a broad range intrinsic growth rates ([Thorson, 2020](#ref-thorson2020)), 0.01 and 1 were chosen as the lower, , and upper, , values; this translate to a normal prior with a mean of -1.15 and a standard deviation of 1.15 on the log-scale. The log-scale prior was informed by total levels of catch and and ,

where and are the lower and upper values. On the lower end, this prior imposes the assumption that the fishery is unlikely to have caught more fish in a single year than the system is capable of supporting and, on the upper end, it assumes that the maximum observed catch represents a portion of the true carrying capacity of the system (*sensu* [Froese et al., 2017](#ref-froese2017)). The division by the lower and upper values for also accounts for the potential range of productivity. Note that the maximum time-series catches are species specific when estimating species-specific carrying capacities.

#### *Starting biomass,*

Like , catch was used to constrain the plausible range of the biomass at the beginning of the time series, , which we will denote as to simplify notation. Specifically,

where and are the lower and upper values that are log transformed to define the normal prior for . Again, adjusting for the potential range of productivity from to , this assumes that the fishery did not catch all of the biomass in the first year and it assumes that the catch represents a portion of the biomass in the first year. A flaw with the upper value for this prior is that a lack of market demand or conservation concerns may contradict the assumption that landings are coarsely proportional to stock size. However, the upper range chosen was considered reasonable for this case study as there was an active fishery for each species at the start of the time series.

#### *Process error variance, , temporal correlation, , and species-to-species correlation,*

Considerable process variability may arise from variable recruitment, natural mortality, and/or growth. For instance, species in the Scorpaenidae family, which includes the *Sebastes* genus, have notoriously variable recruitment which frequently results in “spasmodic” stock dynamics ([Cadigan et al., 2022](#ref-cadigan2022); [Licandeo et al., 2020](#ref-licandeo2020)). Also, there is evidence that heightened and variable natural mortality contributed to the collapse and slow recovery of Atlantic cod along the Northeast NL Shelf ([Cadigan, 2015](#ref-cadigan2015); [Paul M. Regular et al., 2021](#ref-regular2021)). To account for a wide range of possible standard deviations, , a vague prior with 0.01 and 1 as the lower and upper values were chosen. In log-space, this translates to a normal distribution with a mean of -2.3 and and sd of 2.3.

There is little information to inform the potential level of temporal or species-to-species dependence present in the focal systems; vague priors were therefore defined for and . For , 0.1 and 0.9 were logit transformed [] to define the upper and lower inflection points for a normal prior for the logit of . For , the logit transformation was shifted [] to capture negative and positive lower (-0.9) and upper (0.9) values. When these parameters are estimated, these priors give most credence to moderate temporal correlation (0.5) and no species-to-species correlation (0) but still allow the possibility of high levels of correlation (0.9).

#### *Catchability, , and observation error variance,*

The catchability of the spring and fall surveys conducted by DFO likely changed given the shift in gear from a Yankee to Engel to Campelen trawl as well as shifts in coverage of the strata in each region. Survey catchability is therefore indexed by season and gear, , and species, . Moreover, the lower inflection point for the priors were informed by the average survey coverage by gear and season. Survey coverage within each region was computed by dividing the average spatial coverage of strata across years by the of the total area of all strata (i.e., average area covered / area of survey domain). Survey coverage was then multiplied by 0.2 for deep-water species (Greenland Halibut, Atlantic Halibut, Witch Flounder, Redfish spp., White Hake, Silver Hake, and Monkfish) and 0.5 for the remainder. The upper inflection point was set to 1 as it is unlikely, but still possible, that the survey indices represent overestimates of the true population size. The lower range is widened, especially for deep-water species, to account for gear selectivity issues [e.g., escapement under the footgear; Walsh ([1992](#ref-walsh1992))] and availability issues (e.g., portion of the stock in deeper water than covered by the survey).

A prior for observation error variance was informed by design-based estimates survey variance associated with annual biomass estimates ([Smith & Somerton, 1981](#ref-smith1981)). These estimates were used to calculate the coefficient of variation (CV) for each survey, year, and species; specifically,

where represents the design-based indices of biomass used in equation (2). These CVs were log transformed and survey, , and species, , specific means and standard deviations were used to inform the prior for in log-space. Rather than treat the estimates of survey CV as a perfect indicator of total observation error, the prior was widened by multiplying the standard deviation of log CV by two to account for observation variance introduced by distributional shifts outside of the survey domain. Another two times multiplier was applied for deep-water species as they have more scope for shifting outside the survey domain.

### Model selection

It is unknown whether density-dependence of the demersal fish community inhabiting the NL shelves is primarily driven by intra or interspecific competition, or the degree to which species interactions affect population dyanmics. It is, however, well known that the community has been fished for more than 500 years and this history was punctuated by a collapse of several stocks, most notably cod, in the early 1990s ([Lear, 1998](#ref-lear1998)). This collapse was not isolated to commercial stocks, as the biomass of several non-commercial species collapsed at the same time. This collapse was followed by a reorganization of the community, which implies that the system experienced a regime shift ([Pedersen et al., 2017](#ref-pedersen2017)). Given this context, and the structure of the multispecies surplus production model described above, a series of hypotheses were tested:

**Full:** For this hypothesis, it is posited that population dynamics within regions are governed by an overarching carrying capacity that affects all focal species as the aggregate biomass approaches system limits (i.e., inter-specific density-dependent effects are assumed). For this hypothesis and all others, annual reported landings of each species are subtracted. Residual variations not explained by intrinsic growth, density-dependent effects or landings are described by process errors that are 1) correlated across species and said correlations are assumed to be unstructured, meaning relationships can be of differing strengths — positive, neutral, or negative — for each species-to-species pair; and 2) assumed to be temporally correlated, following an AR1 structure. Finally, a shift covariate was applied to the carrying capacity parameter to enable the estimation of different system limits before and after the community-wide collapse to assess support for a regime shift in each systems’ capacity for the focal demersal species.

**Just shift:** Model structure is the same as the **full** model, except the process errors are assumed to be independent and identically distributed random variables (*iid*) to assess the hypothesis that temporal and species correlations are explained by the shift.

**Just correlation:** Model structure is the same as the **full** model, except population dynamics are assumed to be affected by a common and time-invariant carrying capacity. The shift covariate was not applied.

**Shared correlation:** Model structure is the same as the **just correlation** model, except species by species correlations in the process errors are assumed to be the same across all pairs. This structure implies that there is a common but unknown environmental variable affecting the population dynamics of all species. The shift covariate was not applied.

**Just species correlation:** Model structure is the same as the **shared correlation** model, except the process errors at each time step are assumed to be *iid*. This structure implies that there is a common but unknown environmental process affecting all species, but the process is noisy with no temporal dependence. The shift covariate was not applied.

**Just temporal correlation:** Model structure is the same as the **shared correlation** model, except species by species correlations are assumed to be *iid*. This structure implies that environmental processes affect each species differently, however, there may be carry-over effects from one year to the next. The shift covariate was not applied.

**No correlation:** Model structure is the same as the **shared correlation** model, except both temporal and species correlations are assumed to be *iid*. That is, population dynamics are thought to be affected by a common time-invariant carrying capacity and residual variations not explained by intrinsic growth, inter-specific density-dependent effects or landings are noisy and independent across time and species. The shift covariate was not applied.

**Single-species:** Model structure is the same as the **no correlation** model, except population dynamics are assumed to be governed by species-specific carrying capacities (i.e., intra-specific density-dependent effects are assumed). This formulation is analogous to standard state-space Schaefer production models. The shift covariate was not applied.

The predictive ability of each of these models was tested using two cross-validation approaches: 1) leave-one-out cross-validation (LOO-CV), and 2) hindcast cross-validation (Hindcast-CV). LOO-CV is a form of exhaustive cross-validation where the model is repeatedly conditioned on a training set missing one observation until the number of model folds equal the number of observations in the data. The missing observations are predicted at each fold, permitting assessments of the models’ ability to predict the actual value that was left-out at each fold. The hindcast-CV approach is similar, however it focuses on the models’ ability to predict the future. Under this approach, the model is repeatedly conditioned on a training set missing observations from the terminal year such that each fold excludes an increasing number of years worth of data from the tail of the time series ([Kell et al., 2016](#ref-kell2016)). We folded back 20 years and, for each fold, predicted survey indices were compared to the observed survey indices (e.g., observed indices from 2020 were compared with predicted survey indices for 2020 from the model conditioned on data from 1978-2019). For both approaches, we denote predicted survey indices as and the left out observations , where represents all unique combinations of years, species, and survey indices present in the left out data. LOO-CV and Hindcast-CV prediction error scores for each model for each region were calculated by taking the mean squared error,

These scores were also averaged across methods (LOO-CV, Hindcast-CV) and region (Northeast NL Shelf, Grand Bank, Southern NL) to obtain an overall score of predictive ability of each model. A well-fitting model will result in predicted values that are close to the excluded values and, therefore, result in lower scores. Ultimately, these scores enable a model-comparison approach to hypothesis testing.

# Results

Both cross-validation metrics (LOO-CV and Hindcast-CV scores) indicate that most multispecies production model formulations outperform a single-species production model when applied to seven species within three ecosystem production units (Northeast NL Shelf, Grand Bank, and Southern NL) off the east coast of Canada (Figure 2). Focusing on overall scores, the performance of the single-species production model was similar to a multispecies formulation that assumes there is no correlation in the process errors across species or time. There tend to be more notable decreases in the scores as temporal and species correlations are introduced, indicating an improvement in the predictive ability of these models. The “shared correlation” and “just correlation” formulations, in particular, tended to receive the lowest scores, and dropping the species and temporal correlations in lieu of a shift covariate (“just shift” formulation) resulted in a deterioration of predictive ability. Scores were improved when temporal and species-to-species correlations were introduced along with the shift covariate (“full” formulation); however, the fit of the “full” model tended to be poorer than the “just correlation” formulation, which further indicates that the “shift” covariate degraded the predictive ability of the model. Subsequent plots focus on the best fitting formulation, “just correlation”, to demonstrate model predictions.

The multispecies production model with unstructured species-to-species correlation and AR1 temporal correlation (“just correlation”) offered an explanation of the trends in survey indices of focal species across three ecosystem production units with little signs of systematic bias (see residual plots included in model dashboards; Supplement 1). Predicted indices track observed values and, by estimating catchability parameters by species and survey, indices from temporally fragmented surveys are stitched together and their trends are used to inform a continuous underlying trend in biomass (Figure 3). The earlier Yankee and Engel eras of the Canadian surveys tended to receive lower catchability estimates than the Campelen era survey; indices since 1996 therefore tend to be closer to the underlying estimates of biomass from the model.

Isolating residual changes in biomass not explained by reported fisheries landings or the production function (equation (4)) reveals substantive subtractions in the early 1990s across all three ecosystem production units (Figure 4). These residual subtractions exceed the absolute scale of landings taken in the late 1980s from the Northeast NL Shelf and Grand Bank ecosystem production units. The scale of changes derived from the production function and other processes are of a scale similar to landings reported in the Southern NL ecosystem production unit. Focusing on estimates of biomass, total biomass exceeded the carrying capacity estimated for the Northeast NL Shelf through the 1980s and, in the early 1990s, all focal species displayed abrupt declines. Comparing community composition in the early 1980s to the 2010s, there are no clear shifts in the relative biomass of the focal species in the Northeast NL Shelf. Elsewhere, total biomass exceeded or approached system carrying capacity in the late 1980s, after which all species declined. Estimates of total biomass have gradually increased on the Grand Bank and off Southern NL since the mid 1990s, largely due to increasing Redfish spp. biomass estimates.

Further inspection of the process errors reveals common patterns across all focal species across three ecosystem production units (Figure 5). Like the exponentiated and unstandardized process errors (“residual change”) shown in Figure 4, the standardized process errors shown in Figure 5 highlight substantive subtractions in the early 1990s, representing time-series lows for 19 out of 21 populations. The standardized values also reveal parallel and periodic increases and decreases within each region. The species-to-species correlations in the process errors estimated by the “just correlation” model indicate that only 10 out of 63 pairs were negatively correlated. Also note that the “shared correlation” model estimates of a common species-to-species correlation parameter were 0.79 (95% CI: 0.70, 0.86), 0.68 (95% CI: 0.58, 0.76), and 0.81 (95% CI: 0.66, 0.90) for the Northeast NL Shelf, Grand Bank, and Southern NL ecosystem production units, respectively. This model received similar cross-validation scores as the “just correlation” model which, taken together, further supports the inference that the process errors are primarily positively correlated across species within each production unit.

# Discussion

Main points in jot note form:

* Rather than assuming that populations are primarily regulated by species-specific carrying capacities, our multispecies production model assumes that both intra and interspecific competition stunts growth as total biomass approaches the environment’s maximal load.
* The structure of our model resolves an issue with applying the carrying capacity concept as it accounts for species interactions.
* Our case study of the population dynamics of commercially important demersal fish stocks off the east coast of Canada demonstrates the utility of concurrently modeling the production of multiple sympatric species.
* Account for correlated dynamics across species appears to be especially important and we provide empirical evidence of synchronous changes in the focal demersal fish community. These results imply effects from common bottom-up drivers.
* Populations across all three regions declined in the early 1990s. The prevailing hypothesis is that fisheries were the primary driver of these declines; however, results indicate that substantial subtractions from factors not related to *reported fishing activity* are required to explain the declines. A portion of these losses may be attributed to illegal fishing activity, but it seems unlikely that the industry had the capacity to covertly exceed typical annual catches in the late 1980s. For instance, annual catches in the late 1980s across the Northeast NL Shelf and the Grand Bank totaled ~450 kt while residual losses estimated by the model in the early 1990s was ~1000 kt. The illegal fishery would have had to double its efforts to explain the declines. It follows that the decline must at least in part, if not primarily, be due to an unknown environmental driver. (See also Pedersen).
* It is unclear whether the collapse represents a shift in the carrying capacity of the system; the model which assumes the systems capacity changed was not well supported relative to the other production model formulations. That said, there were clear shifts in the communities in the region. Interspecific competition may have contributed to these shifts, however shifting energy pathways may also be at play. Observed restructuring of the communities may be akin to the **paradox of plankton** where the continuous interaction of ecological and environmental factors give rise to “oscillations and chaos, with a continuous wax and wane of species within the community” (Scheffer et al. 2003; <https://doi.org/10.1023/A:1024404804748>)
* While our multispecies production model accounts for more ecological interactions than standard single-species production models, it is still an imperfect abstraction of nature. Results indicate that total biomass in Northeast NL Shelf production unit was above the carrying capacity of that region in the early 1990s. This result is at odds with historic records that suggest populations such as Northern cod were at substantially higher levels in the 1970s and earlier. However, it is possible that the 1970s represents a productive period when the system may have exceeded the carrying capacity.
* Results from the Grand Bank and Southern NL also indicate that the demersal fish community is currently dominated by redfish; this result may be an artifact of unrealistically low survey catchability estimates or the model’s inability to properly account for year effects. Observation errors are assumed to be lognormally distributed, however, extreme catch events / black swan events in space (cite Anderson) can introduce ‘year effects’ that may be better accounted assuming the a distribution with heavier tails, such as the t-distribution.
* Wrap up: there is growing need to apply an ecosystem approach to fisheries management, however, the data requirements of multispecies assessments models built to support ecosystem-informed decisions are often prohibitive. There are currently few data-limited approaches. We document a production model that can estimate the population dynamics of multiple stocks using commonly available landings and survey index data. The methods presented here may therefore represent a relatively tractable approach for informing management decisions for multiple sympatric and data-limited species.

# Acknowledgements

We thank the many ship’s crew and research staff involved in leading the surveys and collecting the data used in this study. These surveys were supported by Fisheries and Oceans Canada (DFO). This work has benefited from valuable feedback from numerous colleagues, including …

# References

Cadigan, N. G. (2015). A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(2), 296–308.

Cadigan, N. G., Duplisea, D. E., Senay, C., Parent, G. J., Winger, P. D., Linton, B., & Kristinsson, K. (2022). Northwest atlantic redfish science priorities for managing an enigmatic species complex. *Canadian Journal of Fisheries and Aquatic Sciences*, *ja*.

Chadwick, E., Brodie, W., Colbourne, E., Clark, D., Gascon, D., & Hurlbut, T. (2007). History of annual multi-species trawl surveys on the Atlantic coast of Canada. *Atlantic Zonal Monitoring Program Bulletin*, *6*, 25–42.

Froese, R., Demirel, N., Coro, G., Kleisner, K. M., & Winker, H. (2017). Estimating fisheries reference points from catch and resilience. *Fish and Fisheries*, *18*(3), 506–526.

Hutchings, J. A. (2021). *A primer of life histories: Ecology, evolution, and application*. Oxford University Press.

Kell, L. T., Kimoto, A., & Kitakado, T. (2016). Evaluation of the prediction skill of stock assessment using hindcasting. *Fisheries Research*, *183*, 119–127.

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., & Bell, B. (2016). TMB: Automatic differentiation and laplace approximation. *Journal of Statistical Software*, *70*(5), 1–21. <https://doi.org/10.18637/jss.v070.i05>

Lear, W. H. (1998). History of fisheries in the northwest atlantic: The 500-year perspective. *Journal of Northwest Atlantic Fishery Science*, *23*.

Licandeo, R., Duplisea, D. E., Senay, C., Marentette, J. R., & McAllister, M. K. (2020). Management strategies for spasmodic stocks: A canadian atlantic redfish fishery case study. *Canadian Journal of Fisheries and Aquatic Sciences*, *77*(4), 684–702.

NAFO. (2019). Report of the Scientific Council, 31 May – 13 June 2019, Halifax, Canada. *NAFO SCS Doc*, *19/20*. <https://www.nafo.int/Portals/0/PDFs/sc/2019/scs19-20.pdf>

Paul M. Regular, Buren, A. D., Dwyer, K. S., Cadigan, N. G., Gregory, R. S., Koen-Alonso, M., Rideout, R. M., Robertson, G. J., Robertson, M. D., Stenson, G. B., Wheeland, L. J., & Zhang, F. (2021). Indexing starvation mortality to assess its role in the population regulation of northern cod. *Fisheries Research*, *247*, 106180. https://doi.org/<https://doi.org/10.1016/j.fishres.2021.106180>

Pedersen, E. J., Thompson, P. L., Ball, R. A., Fortin, M.-J., Gouhier, T. C., Link, H., Moritz, C., Nenzen, H., Stanley, R. R., Taranu, Z. E., et al. (2017). Signatures of the collapse and incipient recovery of an overexploited marine ecosystem. *Royal Society Open Science*, *4*(7), 170215.

Pepin, P., Higdon, J., Koen-Alonso, M., Fogarty, M., & Ollerhead, N. (2014). Application of ecoregion analysis to the identification of Ecosystem Production Units (EPUs) in the NAFO Convention Area. *NAFO Sci. Counc. Res. Doc*, *14*, 069.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Schaefer, M. B. (1954). Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. *Inter-American Tropical Tuna Commission Bulletin*, *1*(2), 23–56.

Smith, S., & Somerton, G. (1981). *STRAP: A User-Oriented Computer Analysis System for Groundfish Research Trawl Survey Data* (p. 66). Canadian Technical Report of Fisheries; Aquatic Sciences No. 1030.

Thorson, J. T. (2020). Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. *Fish and Fisheries*, *21*(2), 237–251.

Walsh, S. J. (1992). Size-dependent selection at the footgear of a groundfish survey trawl. *North American Journal of Fisheries Management*, *12*(3), 625–633.

# Figures



Fig 1: Map of Newfoundland and Labrador (NL) case study area showing the Northeast NL Shelf (purple), Grand Bank (green), and Southern NL (yellow) ecosystem production units.

Fig 2: LOO-CV and Hindcast-CV scores from a series of production model formulations of increasing complexity, from a single-species production model (“Single-species”) to a multispecies production model with a covariate effect and species and temporal correlations (“Full”), fit to seven species in three ecosystem production units (Northeast NL Shelf, Grand Bank, and Southern NL) off the east coast of Canada. Overall scores for each model across score type and ecosystem production unit are shown on the rightmost facet.

Fig 3: Trends in survey indices (dots) along with predicted indices (dotted line) and model estimates of biomass from the best fitting multispecies production model. Uncertainty in the biomass estimates are represented by 95% confidence intervals (shaded region). Results from three ecosystem production units (Northeast NL Shelf, Grand Bank, and Southern NL) off the east coast of Canada are presented. Though data from seven species were used in each region, the species facet is limited to four for simplicity. Differences in biomass estimates from the indices stem from differing survey catchability estimates. Note differences in scale across facets.

Fig 4: Trends in biomass, expected production, reported landings, and residual change from the best fitting multispecies production model across three ecosystem production units (Northeast NL Shelf, Grand Bank, and Southern NL) off the east coast of Canada. Species-specific trends from a subset of species are shown along with totals including seven species. Expected production represents expected changes from intrinsic growth and density-dependence (i.e., change from the production equation) and residual change represent process errors not explained by the production equation or reported landings. Note difference in scale across facets.

Fig 5: Trends (upper facet) and species-to-species correlations (lower facet) in standardized process errors from the best fitting multispecies production model applied to seven species within three ecosystem production units (Northeast NL Shelf, Grand Bank, and Southern NL) off the east coast of Canada. Process errors represent changes not explained by the production equation or reported landings. Points in the correlation plots are coloured and scaled by the level of species-to-species correlation, .