# The sspm R package for spatially-explicit surplus production population models

Valentin Lucet [*1] and **Eric J. Pedersen** [†1,2]

1 Department of Biology, Concordia University, Montreal, CA **2** Department of Biology, Memorial University of Newfoundland, St. John's, CA

Statement of need

Population models are important tools for making management decisions, especially in fisheries, where predictive methods like Surplus Production Models (SPMs) are widely used. Fisheries analysts and managers often lack user-friendly, flexible tools to implement and apply SPMs. In addition, SPMs are rarely spatially explicit and usually cannot account for relevant ecosystem drivers. Therefore, there is a need for tools that implement spatially explicit surplus production models (SSPMs). The Northern Shrimp stock in the Newfoundland and Labrador Shelves is an example of a stock in need of an SSPM that can integrate important spatially-structured ecosystem drivers.

Summary

Population modelling is an exercise of interest within environmental sciences and adjacent fields. Early population models such as the logistic model assumed that while the abundance of a population might change over time, the conditions governing parameters affecting that rate of change, such as the maximum rate of growth or the carrying capacity of the population, stay constant over time (Gotelli, 2008). Modern population models increasingly acknowledge the non-stationary nature of wild populations and work to incorporate environmental fluctuations into dynamic models (Thorson et al., 2015, 2017). Population models designed to answer applied resource management questions, such as fisheries assessment models, increasingly address how the dynamics of stocks vary across space and time.

Resource managers are becoming increasingly interested in how variation in ecosystem factors such as predator abundance and abiotic variables impact the spatiotemporal variability of population parameters, such as productivity (Szuwalski & Hollowed, 2016; Zhang et al., 2021). Further, treating spatially structured stocks as single unstructured stocks can lead to substantially biased estimates of population change (Thorson et al., 2015). However, stock models that explicitly incorporate spatial dynamics and time-varying ecosystem variables are still rare in fisheries science, despite the push for more ecosystem-based management methods in fisheries management (Berkes, 2012; Crowder et al., 2008; Tam et al., 2017).

Surplus production models (SPMs) are one of the classic models used in fisheries and are based on modelling changes in the total biomass of a stock in a given location over time as a function of current stock abundance and fishing pressure (Walters et al., 2008). Classically, SPMs assume single unstructured stocks with purely logistic dynamics (Walters et al., 2008) and, as such, have been of limited use for modelling more complex stocks. They are useful in data-poor contexts where the age structure of the population is not accessible or when age or length structure do not change substantially over time (Prager, 1994; Punt, 2003). SPMs

---

*co-first author
†co-first author

40 typically model spatially constant productivity. They also assume that populations are only
41 affected by past abundance and fishing, which ignores stressors like climate change which
42 affect growth rates independently of fishing pressure.

43 In the context of climate change and shifting ranges, fisheries productivity is likely to be a
44 moving target (Karp et al., 2019), and managers need better methods that account for varying
45 productivity (Szuwalski & Hollowed, 2016). The Northern Shrimp (*Pandalus borealis*) in
46 the Newfoundland and Labrador Shelves, which has undergone several periods of large-scale
47 biomass change in the last two decades, despite a relatively constant harvest regime, is a prime
48 example of a population thought to be affected by environmental conditions (DFO, 2019).
49 These populations currently lack a population model to understand the drivers of this change
50 and to predict how fishing pressure and changing environmental conditions may affect future
51 abundance, which managers are advised to account for.

52 Population models like SPMs usually fall under the two following categories: process-based
53 and statistical models. Process-based models often rely on differential or difference equations
54 and are based on replicating the underlying processes (e.g., predation, recruitment, dispersal)
55 driving population dynamics. Statistical models instead fit a regression model to time series
56 of population abundances, abundance indices, or productivities, with some assumed error
57 distribution for variation around predictions.

58 We have chosen a statistical approach to fitting SPMs. Statistical models allow for estimation
59 of parameter uncertainty and ranges of model predictions and for flexibly incorporating potential
60 ecosystem drivers into models (Plagányi et al., 2014). Statistical models also allow for straight-
61 forward estimation of spatial variation in population parameters such as maximum productivity
62 or density dependence from data, in the absence of theory predicting how these parameters
63 should vary.

64 In this paper, we use a statistical approach to fitting SPMs using Generalized Additive Models
65 (GAMS), estimated using the `mgcv` R package (Wood, 2017) as the backend. We apply this
66 approach to the population of Northern Shrimp of the Newfoundland and Labrador Shelves,
67 leveraging the smoothing properties of GAMs to account for varying productivity across time
68 and space. The resulting model is a spatial SPM (SSPM), implemented via an R package:
69 `sspm`.

70 The R package `sspm` is designed to make SSPMs simpler to estimate and apply to any spatially
71 structured stock. The basic model this package implements was first used to model time-
72 varying production in Newfoundland and Labrador Northern Shrimp stocks (E. J. Pedersen
73 et al., 2022). However, the general modelling approach used here will work for any spatially
74 structured fishery with sufficient data. It includes a range of features to manipulate harvest
75 and biomass data. Those features are organized in a stepwise workflow, whose implementation
76 is described in more detail in Figure 1 and in the next section.

77 Although it was developed in a fisheries context, the package is suited to model spatially-
78 structured population dynamics in general.

Package design

80 The package follows an object oriented design, making use of the S4 class systems to model a
81 stepwise workflow: (Figure 1).

82 The key workflow steps are:

83 ▪ Discretization and aggregation of spatially structured observations into discrete patches,
84 with a range of methods of discretization (random or custom sampling, Voronoi tessella-
85 tion, or Delaunay triangulation).
86   1. Provided boundary data in the form of a shapefile is converted into a `sspm_boundary`
87   object using `spm_as_boundary()` to define the boundary/region of interest.

2. The region within the boundary is discretized into patches with the `spm_discretize()` function, creating a `sspm_discrete_boundary` object.

- Spatiotemporal smoothing of biomass and environmental predictors using GAMs.

3. The `spm_as_dataset()` function turns user-provided data frames of raw observations into `sspm_dataset` objects that explicitly track locations, data types, and aggregation scales for each input. `sspm` recognizes three types of data: **trawl** (i.e., biomass estimates from scientific surveys), **predictors**, and **catch** (i.e., harvest).

4. The `spm_smooth()` function uses GAMs to calculate spatially smoothed yearly estimates of biomass and environmental predictors for each patch from trawl-level data, based on the spatial structure from the `sspm_discrete_boundary` object. The user specifies a GAM formula with custom smooth terms. The output is another `sspm_dataset` object with a `smoothed_data` slot which contains the smoothed predictions for all patches.

- Computation of surplus production based on biomass density and fishing effort.

5. The `spm_aggregate_catch()` function aggregates catch into patches and years and calculates patch-specific productivity for each year as the ratio of estimated biomass density plus catch from the next year divided by estimated biomass density of the current year. The result is returned as a `sspm_dataset`.

6. The `sspm()` function combines productivity and predictor datasets into a single dataset. Additionally, the user may create lagged versions of predictors with `spm_lag()` and split data into testing and training sets for model validation with `spm_split()` at this stage.

- Fitting of SSPMs to productivity estimates with GAMs.

7. The `spm()` function is used to fit a SSPM model to the output of step 6, using a GAM model with custom syntax able to model a range of SSPMs. The output is an `sspm` object.

- Visualization of results, and one-step-ahead projections of biomass for model validation and scenario-based predictions.

8. Plots can be generated with the `plot()` method. Predictions from the fitted model can be obtained using the built-in `predict()` method, including confidence and prediction intervals
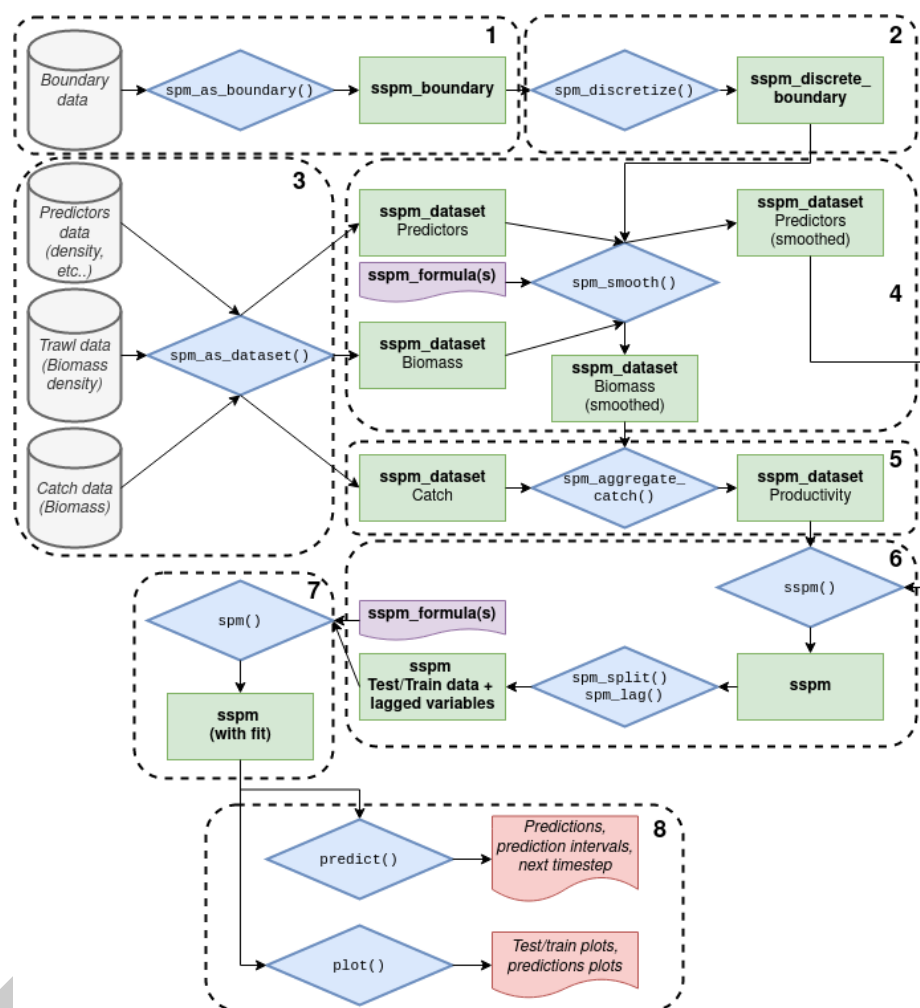
Figure 1: The sspm workflow. Gray cylinders represent raw, unprocessed sources of data. Each blue diamond shape represents a function processing a raw input and validating it, or producing an intermediate package object, represented as a green object. Secondary objects like formulas, which must be created by the user, are represented by a purple document shape. Finally, outputs are represented by a red document shape. The steps of the workflow as described above are denoted by dotted lines and corresponding step number.

Connections to other surplus-production based stock assessment approaches

The `sspm` package uses a model-based, random-effects based approach to estimate the effects of ecosystem drivers on surplus production across space and time. Our approach is conceptually related to the stochastic stock assessment approaches used by the `spict` (M. W. Pedersen & Berg, 2017) and `jabba` (Winker et al., 2018) R packages for surplus production modelling, in that we assume that biomass dynamics can be modelled as effectively a logistic growth model with both process and measurement error. While `sspm` does not currently have the capacity to model biomass dynamics as a continuous-time process, as with `spict`, or incorporate prior parameter information on catchability or biomass dynamics, as in `jabba`, `sspm` can model spatially and temporally varying productivity, which is currently not possible in these models.

The `sspm` package can be viewed as a spatiotemporal Model of Intermediate Complexity [a 'MICE-in-space' model; Thorson et al. (2019)] that can incorporate effects of other species

and ecosystem drivers as well as changes in fishing pressure on stock status. Our approach is closely connected to approaches used by other modern model-based spatial abundance estimation software, such as the VAST R package (Thorson et al., 2019) and the sdmTMB R package (Anderson et al., 2022). Our method shares the same approach as both VAST and sdmTMB of using spatially explicit models to estimate local biomass density (Figure 1 steps 1,2, and 4), then aggregating up from those models to predict aggregate stock-level metrics such as total biomass and productivity (Figure 1 steps 8). The multiplicative surplus production model used by sspm is also conceptually similar to the vector-autoregressive model for biomass changes used by these two packages, as both VAST and sdmTMB can model local temporal changes as autoregressive processes on the link-scale of a generalized linear model. The sspm package cannot, however, model the dynamics of multiple species simultaneously; multi-species modelling would require generating a separate surplus production model (Figure 1 steps 5 and 6) for each species of interest.

One major difference between the sspm package and other model-based spatiotemporal modelling packages is its special-purpose nature. The default spatial_smooth function uses a computationally simpler (although somewhat less flexible) Intrinsic Conditional Autoregressive (ICAR) model (Rue & Held, 2005) for modelling spatial variation in covariates and biomass, as compared to the more complex spatial random effects possible with VAST and sdmTMB. This has the advantage of computational speed and less user knowledge of how to set up complex spatial grids, although it is less flexible. This means that sspm should be easier to adapt to novel fisheries than more complex packages that require more user modelling knowledge. Further, it is possible to specify alternative spatial smoothers than the ICAR model in the spatial_smooth function via the bs= argument, although this functionality has not been well-tested and should be considered experimental.

The other benefit of sspm, relative to other modelling packages, is the ability to model productivity rates directly (Figure 1 steps 5 and 6), rather than implicitly via an auto-regressive processes as used in VAST or sdmTMB. This means that it is possible in sspm to model nonlinear relationships between environmental covariates and productivity, or to easily include factors such as time-lagged effects of predictors on productivity in a given year. This approach does, however, sacrifice the ability to propagate measurement error into uncertainty about rates of change. One of the future directions for development of this package is to include variance propagation methods into the surplus production modelling step.

Acknowledgements

References

Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). *sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields* (p. 2022.03.24.485545). bioRxiv. https://doi.org/10.1101/2022.03.24.485545

Berkes, F. (2012). Implementing ecosystem-based management: Evolution or revolution? *Fish and Fisheries*, *13*(4), 465–476. https://doi.org/10.1111/j.1467-2979.2011.00452.x

Crowder, L. B., Hazen, E. L., Avissar, N., Bjorkland, R., Latanich, C., & Ogburn, M. B. (2008). The impacts of fisheries on marine ecosystems and the transition to ecosystem-based

179   management. *Annual Review of Ecology, Evolution, and Systematics*, *39*(1), 259–278.
180   https://doi.org/10.1146/annurev.ecolsys.39.110707.173406

181 DFO. (2019). *An assessment of Northern Shrimp (Pandalus borealis) in Shrimp Fishing Areas*
182   *4–6 and of Striped Shrimp (Pandalus montagui) in shrimp fishing area 4 in 2018* (Technical
183   Report No. 2019/027; p. 23). Canadian Science Advisory Secretariat (CSAS).

184 Gotelli, N. J. (2008). *A Primer of Ecology* (Fourth Edition). Oxford University Press.
185   ISBN: 978-0-87893-318-1

186 Karp, M. A., Peterson, J. O., Lynch, P. D., Griffis, R. B., Adams, C. F., Arnold, W. S.,
187   Barnett, L. A. K., deReynier, Y., DiCosimo, J., Fenske, K. H., Gaichas, S. K., Hollowed, A.,
188   Holsman, K., Karnauskas, M., Kobayashi, D., Leising, A., Manderson, J. P., McClure, M.,
189   Morrison, W. E., … Link, J. S. (2019). Accounting for shifting distributions and changing
190   productivity in the development of scientific advice for fishery management. *ICES Journal*
191   *of Marine Science*, *76*, 1305–1315. https://doi.org/10.1093/icesjms/fsz048

192 Pedersen, E. J., Skanes, K., Le Corre, N., Koen-Alonso, M., & Baker, K. (2022). *A new*
193   *spatial ecosystem-based surplus production model for SFA 4-6 Northern Shrimp* [Canadian
194   Science Advisory Secretariat (CSAS) Research Document]. https://www.dfo-mpo.gc.ca/
195   csas-sccs/Publications/ResDocs-DocRech/2022/2022_062-eng.html

196 Pedersen, M. W., & Berg, C. W. (2017). A stochastic surplus production model in continuous
197   time. *Fish and Fisheries*, *18*(2), 226–243. https://doi.org/10.1111/faf.12174

198 Plagányi, É. E., Punt, A. E., Hillary, R., Morello, E. B., Thébaud, O., Hutton, T., Pillans, R.
199   D., Thorson, J. T., Fulton, E. A., Smith, A. D. M., Smith, F., Bayliss, P., Haywood, M.,
200   Lyne, V., & Rothlisberg, P. C. (2014). Multispecies fisheries management and conservation:
201   Tactical applications using models of intermediate complexity. *Fish and Fisheries*, *15*(1),
202   1–22. https://doi.org/10.1111/j.1467-2979.2012.00488.x

203 Prager, M. H. (1994). A suite of extensions to a nonequilibrium surplus-production. *Fishery*
204   *Bulletin*, *92*, 374–389.

205 Punt, A. E. (2003). Extending production models to include process error in the population
206   dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, *60*(10), 1217–1228.
207   https://doi.org/10.1139/f03-105

208 Rue, H., & Held, L. (2005). *Gaussian Markov random fields: Theory and applications*. CRC
209   press.

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

213 Tam, J. C., Link, J. S., Rossberg, A. G., Rogers, S. I., Levin, P. S., Rochet, M.-J., Bundy, A.,
214   Belgrano, A., Libralato, S., Tomczak, M., Wolfshaar, K. van de, Pranovi, F., Gorokhova,
215   E., Large, S. I., Niquil, N., Greenstreet, S. P. R., Druon, J.-N., Lesutiene, J., Johansen, M.,
216   … Rindorf, A. (2017). Towards ecosystem-based management: Identifying operational food-
217   web indicators for marine ecosystems. *ICES Journal of Marine Science*, *74*(7), 2040–2052.
218   https://doi.org/10.1093/icesjms/fsw230

219 Thorson, J. T., Adams, G., & Holsman, K. (2019). Spatio-temporal models of intermediate
220   complexity for ecosystem assessments: A new tool for spatial fisheries management. *Fish*
221   *and Fisheries*, *20*(6), 1083–1099. https://doi.org/10.1111/faf.12398

222 Thorson, J. T., Jannot, J., & Somers, K. (2017). Using spatio-temporal models of population
223   growth and movement to monitor overlap between human impacts and fish populations.
224   *Journal of Applied Ecology*, *54*(2), 577–587. https://doi.org/10.1111/1365-2664.12664

225 Thorson, J. T., Skaug, H. J., Kristensen, K., Shelton, A. O., Ward, E. J., Harms, J. H., &
226   Benante, J. A. (2015). The importance of spatial models for estimating the strength of

227 density dependence. *Ecology*, *96*(5), 1202–1212. https://doi.org/10.1890/14-0739.1

228 Walters, C. J., Hilborn, R., & Christensen, V. (2008). Surplus production dynamics in declining
229 and recovering fish populations. *Canadian Journal of Fisheries and Aquatic Sciences*,
230 *65*(11), 2536–2551. https://doi.org/10.1139/F08-170

231 Winker, H., Carvalho, F., & Kapur, M. (2018). JABBA: Just Another Bayesian Biomass
232 Assessment. *Fisheries Research*, *204*, 275–288. https://doi.org/10.1016/j.fishres.2018.03.
233 010

234 Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, 2nd Edition* (2nd
235 ed.). CRC Press. ISBN: 978-1-4987-2837-9

236 Zhang, F., Reid, K. B., & Nudds, T. D. (2021). The longer the better? Trade-offs in
237 fisheries stock assessment in dynamic ecosystems. *Fish and Fisheries*, *22*(4), 789–797.
238 https://doi.org/10.1111/faf.12550