Spatial Stock Assessment Methods: International Approaches and Advancements

ANALYST GUIDANCE FOR SIMULATION EXPERIMENT

18 July 2021

1. Overview of the analyst process

An overview of the Study Design document and objectives was provided previously and can be found on the GitHub repository. Here, we provide a detailed outline of the expectations of each analyst participating in the simulation experiment at each stage of the study, including information on model parametrization, expected deliverables, and discussion questions to consider. Using the yellowfin tuna (YFT) operating model and/or Antarctic toothfish (TOA) operating model, each analyst will develop, using initially a single representative dataset:

1. A single area, panmictic assessment to then apply it to the 100 one area simulated data sets (i.e., simulation runs) using their chosen assessment application;
2. A 4 area, spatially stratified assessment and apply it to the 100 spatial (4 area) simulated data sets (i.e., simulation runs) using their chosen assessment application;
3. A summary of their decision point model development process, which outlines how analysts arrived at the final spatially stratified stock assessment model parametrization (Study Design **Objective #1**; see Section 4)
4. A summary of results from each model configuration (i.e., 1 area & 4 area models);
5. A brief comparison of the panmictic and spatial assessment outputs and diagnostics (**Objective #2**);
6. Output files to be provided to the workshop/experiment organizers via Box (links and folder structure to be provided).

After the base runs (i.e., 1 and 4 area assessment models) are complete, the analysts are encouraged to develop other model parametrizations or implement alternate data aggregations (e.g. 2 areas or 8 areas for the YFT model) and comparisons across spatial aggregation approaches are very much welcomed. Results can be pushed to a Box

repository (given space limitations on Github) and analysts are encouraged to pursue publication of their results and comparisons across parametrizations (see Section 7.3 for more info on publications and deliverables).

2. Questions to keep in mind

We ask that analysts keep the following questions in mind, while generating panmictic and spatial assessments. These general questions will help guide group discussion sessions during the virtual webinars and aid in developing suggestions for best practices and future needs for spatial models from the simulation experiment. After completing the simulation experiment provide detailed answers to these questions:

- What types of spatial data that are currently collected are necessary to fit when developing spatial stock assessment? What types of novel or unique data could be fit or used to help parametrize spatial models, and, thus, should be more routinely collected and integrated into the next generation of spatial assessment platforms?
- What spatial capabilities did your platform have that were of the most use/importance?
- What spatial capabilities that are not currently incorporated into your platform should be considered in the next generation of spatial assessment platforms?
- What are the primary limitations of current spatial assessment platforms and how might these be overcome in future spatial model platforms?
- What types of diagnostic tools and residual analyses are useful for identifying when spatial processes need to be incorporated into a stock assessment (e.g., based on data fits and patterns observed in panmictic assessments)?
- What do you think are the primary factors limiting wider implementation of operational spatially explicit stock assessments and how might these be addressed in the future?
- How can we develop biological reference points that adequately account for spatial structure?

3. Model parameterization

We would like the analysts to treat this experiment as though they were developing a real world assessment based on observed data and typically available knowledge regarding biological and fishery dynamics. Of course, all analysts will have more information than is typical in a real world assessment development application (e.g., true values of difficult to estimate parameters, such as natural mortality, will be provided), which will reduce uncertainty and produce overly optimistic results. Yet, we believe that, even with a handful of parameters fixed at true values, the resultant model development process will

be extremely useful to better elucidate best practices for building spatial assessments and identifying sub-optimal parametrizations through residual analysis.

The complete model description can be found in the OM description document for each species (YFT and TOA). Below we provide examples of the minimum parameters that we are requesting each analyst to estimate or derive, as well as those that are provided and may either be fixed or estimated. We note that depending on model parametrization (e.g., choice of population structure) some of these parameters may not be estimated or included in a given model.

For each simulated dataset at each assessment aggregation, we ask that the analysts estimate or derive the following parameters.

Parameters to be *Estimated*:

    i. B0: although given in the OM description document, please do not fix at this value in your models, as it is a key scaling parameter that is almost always estimated in an assessment model.
    ii. R0: same as above.
    iii. Recruitment: provide time series estimates of recruitment by area.
    iv. Recruit Apportionment: if apportionment is estimated please provide these estimates for each area and any temporal variation in these values.
    v. Movement: provide estimates of movement rates between areas as well as any temporal variation in these estimates.
    vi. Fishing Mortality: provide time series estimates of F by fleet and area.
    vii. Fishery Selectivity: provide estimates of selectivity for each fleet by age or length along with any spatiotemporal variation that might be estimated in these parameters.

Parameters to be *Derived*:

    i. SSB: provide time series estimates of SSB by area.
    ii. Abundance-at-age: provide time series estimates of abundance-at-age by area.
    iii. Catch: provide a time series predicted total catch by fleet and area.
    iv. CPUE: provide a time series of predicted CPUE for the longline fleet by area and any associated parameters (e.g., catchability).
    v. Length Compositions: provide a time series of predicted length compositions by fleet and area.
    vi. Tag Recaptures: provide a time series of predicted tag recaptures by release cohort, recapture year, recapture area, and recapture age (if tagging data is fit in the model).

The following parameters are provided in the OM description documents and can either be fixed at the true value or directly estimated:

a. Steepness: the steepness of the stock-recruit curve can be fixed at the true value (for the YFT OM steepness is fixed at 0.8).
b. Natural Mortality: the natural mortality rate can be fixed at the true value (for the YFT OM, natural mortality varies by age, but not area; see Fig. 3d the YFT OM Description document).
c. Tag Retention: initial tag retention can be estimated or fixed (for the YFT OM, tag retention was assumed to be 90%).
d. Tag Reporting Rate: the proportion of tags reported (for the YFT OM, 90% tag reporting was assumed in the purse seine fishery and 0% reporting was assumed for all other fisheries).
e. Chronic Tag Loss: chronic tag loss can be estimated or fixed (for the YFT OM, chronic tag loss was assumed to 0.91% per quarter).

4. Analyst documentation of model development (Objective #1 of Study Design

4.1 Decision Point Analysis

The focal point of the experiment is a comparison among analysts' approaches to develop, parametrize, and assess goodness of fit for spatial stock assessment models. Towards this end, we request that each analyst describe and document, the parametrization of their spatial model, the assumptions implicit or explicit in their model platform, and how these influence the decisions made during the modeling process. The following questions provide a general guideline of major decision points to be documented (loosely based on Punt, 2019), but are not all-inclusive:

1. How did you arrive at your final spatial model parametrization? Did you first develop a panmictic assessment and look for residual patterns indicative of critical spatial processes to be included? Did you simply use the OM description? Did you use any particular diagnostic analyses to develop and refine parametrization (see sub-section 4.2 for more questions related to diagnostics)?
2. How did you identify which spatial processes to incorporate? Were these mainly guided by the platform being used or pre-existing knowledge (i.e., OM descriptions) or both?
3. How did you model fishery and demographic processes?

    a. Were selectivity functions directly estimated for each fleet or shared across areas/fleets?
    b. How was recruitment modeled (i.e., what population structure was assumed and how was area specific recruitment estimates handled)?
    c. Was natural mortality fixed or estimated? Was age, area, or temporal variation accounted for in natural mortality?

    d. How was movement modeled? What influenced the final parametrization of movement?

4. Which spatial dynamics could not be incorporated in your model due to platform or data limitations?
5. Were there any data limitations that impeded model development or estimation (e.g., either sample size limitations in length compositions or simply data that was not simulated that could have been useful for guiding parametrization or estimating specific parameters)?
6. Was tagging data fit in the model? Why or why not? How were tagging parameters dealt with (e.g., fixed or estimated? time-invariant or time-varying?)?

4.2 Model Diagnostics

As residual analysis and model diagnostics are critical components of the assessment process, which indicate whether a given model is adequately fitting observed data, we are particularly interested in the diagnostics that analysts use to determine goodness of fit in a spatial context. In particular, did particular residual patterns from panmictic assessments guide parametrizations of resultant spatial assessments? We ask that you consider and document responses to the following questions throughout the model development and parametrization stages:

1. Are the typical residual analyses used in panmictic assessments adequate for spatial models?
2. Can violation of spatial assumptions or ignorance of spatial structure (i.e., when implementing a panmictic assessment) be identified from commonly observed residual patterns? Was residual analysis from your fit of the panmictic assessment model used to identify key spatial processes to be incorporated or did it inform your spatial parametrization in any way?
3. Can we use the results of a non-spatial model to identify what spatial processes need to be included in a spatial model (e.g., are there diagnostics that can help infer that major assumptions are violated)? If so, please elaborate.
4. Are alternate diagnostics needed for spatial models (e.g., by area or sub-populations)? Did you or have you in the past developed such diagnostics? If so, please elaborate.

Please fill out the accompanying google form for decision point analysis and model diagnostics.

5. Reporting model outputs

We request that raw assessment outputs be provided at each specific platform's Box repository (to be provided to each analyst/group of analysts) for each of the 100 assess-

ment runs for each of the assessment aggregations (1 area and 4 areas). Results should include, but are not limited to:

i. R0, B0.
ii. Recruitment apportionment (if estimated).
iii. Movement rates among regions.
iv. Fishery selectivity by fleet and area.
v. Fishing mortality time series by fleet and area.
vi. Tag reporting rate (if estimated, by fleet and area).
vii. SSB time series by area and summed across all areas.
viii. Recruitment time series by area and summed across all areas
ix. Abundance-at-age time series by area.
x. All predicted data values (i.e., catch, CPUE, length compositions, and tag recaptures).
xi. Spatial Reference Points (if possible in your platform).

Additionally, we request that the analyst upload all raw report files from each assessment run (200 total) to the Box directory under the appropriate spatial aggregation (further details on *Box* to be provided).

6. Communication during experiment

Questions can be directed to Caren Barcelo (caren.barcelo@noaa.gov) or other workshop organizers via email or by way of Github Team Discussions. The Github discussions will serve as a location for Q&A with organizers as well as to offer a platform for analysts to discuss amongst themselves details related to the experiment. Organizers will also reach out, likely in September 2021, to coordinate best times for each analyst (or team of analysts) to give presentations on their findings.

7. Simulation experiment deliverables

7.1 Virtual presentations of results and comparisons

Analysts are asked to give a 90 minute virtual webinar (approximately 45min seminar with 45min for questions and discussion with analysts using other platforms). These will be public seminars and are intended to elicit feedback and discussions on the decision point analysis, model diagnostics, and within-model comparisons.

We will provide further details regarding guidance on webinar and presentation formats. But, generally, we hope analysts will walk through their model development approach including providing responses to the questions outlined in the decision point analysis section, explore model diagnostics, provide population trends and stock status, and

compare results across different model runs (e.g., between the panmictic and spatially stratified models). For the within model comparisons (i.e., panmictic vs. spatially stratified assessments; **Objective #2** of Study Design, we request that you provide summary statements and figures addressing the following questions:

a. How did model estimates compare?
b. Did panmictic models give similar estimates of total biomass?
c. Was overall stock status similar?
d. Was there evidence to suggest that tracking regional dynamics could be a useful management tool? Why or why not?
e. How did uncertainty compare among models, and what were the main differences among model parameters?

7.2 Organizers synthesize data

Organizers will compile all information from each assessment platform and synthesize differences and similarities among model types and across panmictic v. spatial assessment results (**Objective #3** of Study Design. The organizers will conclude the virtual webinars with a summary presentation discussing these aggregate results. The results will be made available to each analyst at the conclusion of the study and all will be encouraged to participate in a collaborative manuscript(s) resulting from the experiment (Section 7.3).

7.3 Manuscript development

We aim to develop results from the YFT simulation experiment into a collaborative manuscript on best practices for spatial assessment model development. We will use information from each analyst's description of their decision points when developing models, diagnostics used, and a comparison of spatial vs panmictic assessment results for the YFT case study. We invite all those who are interested/have time and who actively participated in the simulation experiment with their assessment model/platforms to be co-authors on the resulting manuscript led by the workshop organizers.

Analysts are also encouraged to develop their own collaborative manuscripts based on the results of their work within the broader simulation experiment design. We envision that many interesting results will be uncovered through this collaborative experiment, which can be useful for advancing the state of knowledge regarding performance and development of both spatially explicit and spatially aggregated stock assessment models. Manuscripts might include comparing across a variety of model parametrizations or working with other analyst/platform groups to do further cross model comparisons. The workshop organizers will provide true values of important population parameters from each SPM run to analysts wishing to develop summary statistics or figures of model

fit and bias for manuscripts once final model results are submitted to the organizers. Pending interest a special issue in a journal (TBD) may be organized.

7.4 Spatial assessment methods in-person workshop 2022

We also plan to have an in-person workshop in 2022 (dates and locations TBD) as a forum for further discussion and collaboration on topics of spatial assessment development and emerging issues. Details will be provided as the experiment progresses and the travel restrictions due to the COVID pandemic are further loosened.

Thank you for your participation! Please reach out to any of the workshop organizers with questions or concerns.

References

Punt, A.E. (2019). Spatial stock assessment methods: a viewpoint on current issues and assumptions. Fisheries Research. 213: 132-134. DOI: 10.1016/j.fishres.2019.01.014.