Lessons learned for spatial model development from implementation of a tag-integrated assessment for Alaskan sablefish

# Aim

Document our approach and decisions when developing a tag-integrated assessment model for Alaskan Sablefish

# Abstract

Do last

# Introduction

Recent emphasis on spatial stock assessment methods in the literature, with best practice review papers (Punt 2019a, 2019b, Goethel et al 2023), international workshops (space oddity and CAPAM) and simulation studies (Goethel xx) highlighting the pros and pitfalls of spatially explicit models. However, in practice, developing a spatially explicit stock assessment model is complex, requiring identification of primary spatial dynamics and appropriate assumptions. Given the dearth of spatial assessment applications (Punt 2019), documenting and disseminating the model development process can be helpful for highlighting lessons learned to aid future spatial applications.

We documented the development of a spatially explicit tag integrated assessment model for Alaskan Sablefish (*Anoplopoma fimbria*), covering the Gulf of Alaska, Bering Sea, and Aleutian Islands regions (Goethel et al. 2023). This stock provides a unique case study to illustrate how the spatial modeling process can be approached. Alaskan Sablefish represent a highly mobile, data rich population with a dedicated longline survey and over 40 years of tag releases and recaptures. Currently a single, panmictic population is modeled with quotas apportioned to management areas based on area-specific survey biomass (SAFE reference). However, significant spatial heterogeneity exists in the population distribution resulting from multiple hypothesized factors such as age-based habitat preferences and migration patterns.

Many spatial tagging models have been developed for or included the Alaskan Sablefish stock (Hanselman et al 2015, Heifetz and Fujioka 1991, Bracken 1982, Kimura & Shavy 1998) but none have attempted to integrate the tagging data with the assessment data and assessment population structural assumptions. The biggest impediment for previous studies to conduct a fully integrated analysis was computational power. However, with recent advancements in computers and efficient statistical software (TMB, Stan) these models are much more attainable.

# Methods

The first step in the model development was to review the literature with an emphasis on stock structure. Although initially focused on the Alaskan stock, it is important to re-evaluate stock structure assumptions. A recent genetic study by Jasonowicz et al. (2017), found no significant difference between samples off the west coast of US, GOA and Bering Sea (there were no samples off the west coast of Canada in this study), providing evidence for a single panmictic population. A recent morphometric study (M. Kapur et al. 2020), found significant differences in fish traits such as length at age between samples from the west coast of US and GOA. These results were consistent with the findings from Tripp-Valdez et al. (2012). Tagging is another important information source, with extensive tag releases along the entire North Pacific. Tag recovery data provide evidence for a two-stock hypothesis (Kimura and Shavy 1998). With a stock north of Vancouver Island and a stock to the south. We decided to focus on developing a spatial model that had the same north and south spatial extent to the current assessment. We developed a model that could be extended to include Canadian and US west coast fish at a later period.

The next step included defining the number of spatial regions and intra annual time-blocks that we wanted to include. It is important to choose a modelling tool that has flexibility in the number of spatial regions. The model developed for this research was a bespoke TMB model that was general with respect to spatial resolution (see Appendix ?? for model equations). We conducted an exploratory analysis to help inform spatial strata for our tag-integrated model.

## Data and exploratory analysis

Age-length pairs were visually plotted across the six management regions within each sex to inspect for spatial growth differences. Perhaps a more objective approach at identifying growth delineation is described by (Kapur et al 2020). Incorporating spatial varying growth would require tracking numbers by age and length (or length group), which adds another dimension to partition. When tagging is included the partition has the following dimensions . In the five area model

Spatially varying growth in an age-structured model would Due to the considerable computational demand

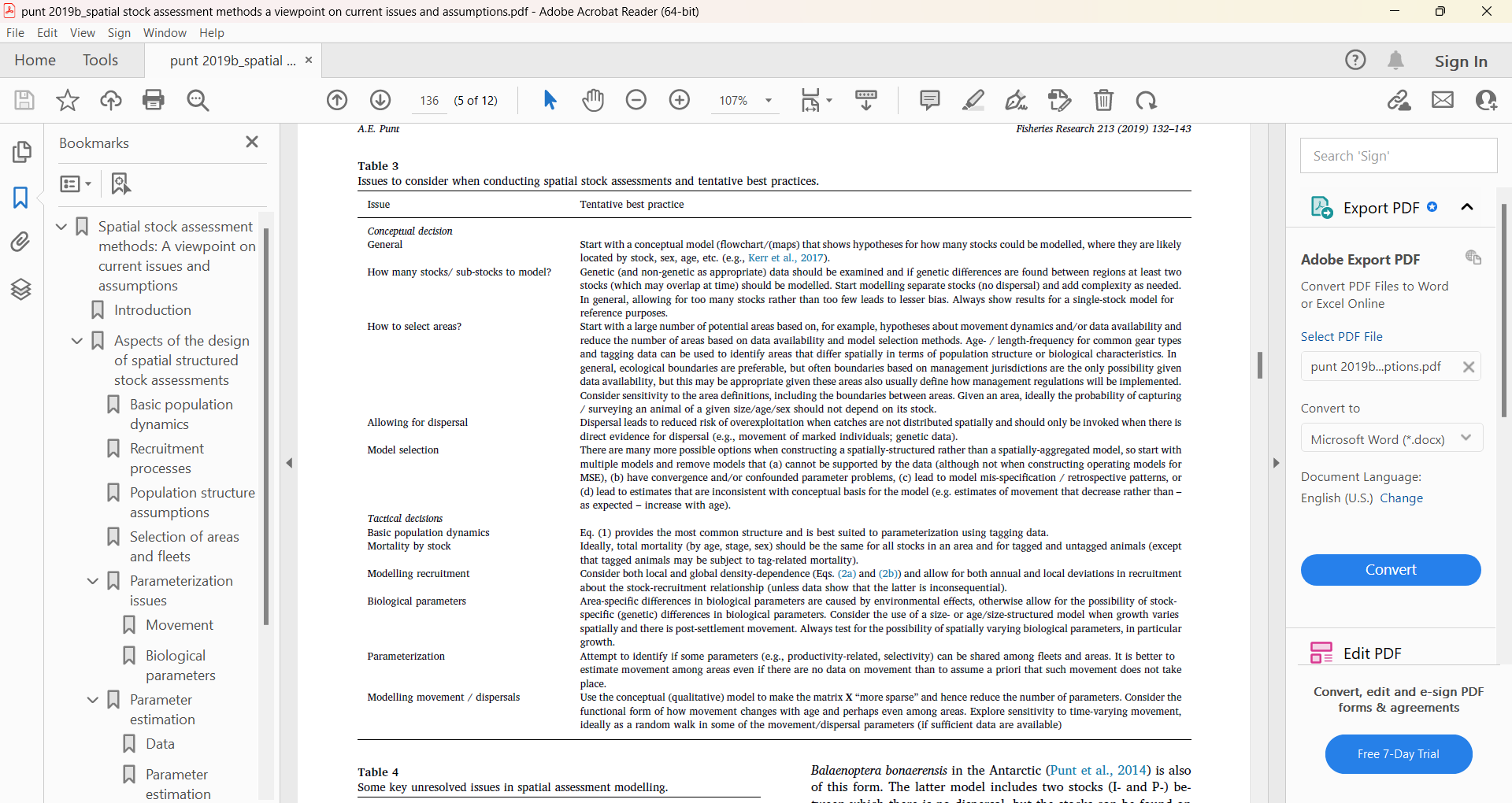
* + What exploratory analysis were conducted to help inform spatial/temporal structure in the model
    - Natural mortality
    - Growth – spatial-temporal variation, we made a decision that it had to be substantial before it was considered for inclusion into the model. The reason for this is the computational cost of including spatially varying growth into a spatial model. To properly include spatially varying growth we would need to add a length dimension to the partition which, if tagging is included in the partition, has substantial computation efficiency implications. The current dimension of the partition is n\_ages x n\_regions x n\_sexes x n\_tagged\_years. When n\_tagged\_years = 10 (as is explore in the model) this equates to 30 x 5 x 2 x 10 = 3000 numbers at age that the model tracks in a given year. Increasing this would seriously impact the model run time.
    - Fishery structure – gear type, spatial distribution of catch over time.
    - Stock regions/production regions/movement – developed a model that can evaluate multiple spatial resolutions and develop flexible input code to estimate age/length frequencies at multiple spatial resolutions. Use clustering algorithms to help delineate spatial structures. The default spatial regions were based on the six current management regions due to convenience of data analysis and historical data inclusion (historic data often could not be disaggregated at a finer spatial resolution). Then with more recent length data we did an exploratory analysis using length composition data and the approach described by (??). This indicated two splits one partway through the aluetian island region and the other to the eastern boundary of the central gulf region (Figure??). The first break was thought to be of little use given the lack of catch that is taken west of the Aleutian islands. It was also hypothesized that this was split was driven by early data which had larger fish. The second split corresponds to the biological split described by (Kapur 2022). This split would result in cutting the central gulf region approximately 70:30. Simulations were needed to confirm if this was a large enough divide to justify an explicit boundary, we subjectively chose to keep the central and eastern gulf management boundary due to the assumptions in historical data inputs that would be required to create that new regional definition.
    - Tagging data inclusion – Generally tag data is length-based due to the cost of ageing. External age-length keys are used to convert length at releases to age at releases for use in age-structured model. This was possible due to the survey releasing the tagged fish which resulted in a representative age-length key for conversions. It is possible to input the tag releases and recoveries as a length based process and use the model’s growth transition matrix to convert to lengths to ages for the age-structured model. This has been done in CASAL stock assessments (toothfish and snapper model), but there is a smearing of information which is discussed in Appendix A which is why we recommend against this method. 

Figure : Punt 2019

* + Key concerns/uncertainties from data exploration
    - Migration – age-based
    - Spatial resolution
    - Bzero-non-starting conditions
* Initial model setup

# Results

## 1A

* + Start year? 1960 more stable
  + Include tag data?
  + Comp likelihood?
  + Time-blocks for selectivities? Trawl 1, Fixed 2, survey 3. The fixed gear selectivity time-block differed from the assessment in that we had two time-blocks which was start\_year-2014, and 2015:2021. The assessment assumes a time-block prior to the 1994 IFQ change. The earlest composition data is length data starting in year 1990. This left 4 length comps to estimate a selectivity, which was difficult to estimate. For this reason we constrained the selectivity to be the same over the IFQ change for model stability.
  + Based on explorations and decision points
  + Mis-fits and diagnostics that were used to reduce dimensionality and help identify model-misspecification
  + Problematic parameters – overdispersion parameters for compositional likelihoods, were sometimes running to bounds.
  + Including tag-data caused a conflict with other data sets resulting in a worse fit to comp data and index data. The inclusion of tagged data for a single area model required further investigation on why the misfit was caused. Single area models presented here did not include Tagging data. Attempts were made to remedy this conflict by using the negative binomial distribution and estimating the overdispersion parameter and estimated time-varying tag-reporting rates which resulted in unstable models.
* Self-test
* Did this tell us anything? Early age-devs need to be constrained to produce mean unbiased estimates of SSB and depletion for 1960 model. This is because early LF data are not very informative with respect to early recruitment deviations. Use this to justify 1977 vs 1960 model?
* 1977 estimated mean unbiased SSB when data simulated from 1960 model, suggesting it should be used instead of 1960 model. Caveat B0 was biased, so shouldn’t use depletion based reference points.

## 3A

Initial model assumed regional average recruitment with global recruitment deviations estimated and applied for each region no tagging data and no movement among regions.

Model development

* The next step we freed up the recruitment assumption and estimated regional recruitment deviations
* We then included the tagging data (assumed to be Poisson distributed) and estimated movement rates

Self-test with candidate models

## 5A

# Simulations

How well does the single area model perform with survey based catch allocation versus a full five area tag-integrated model?

Using a conditioned five area model estimated from real data we used it as an OM to simulate data. We then ran conducted a closed loop MSE to compare the 1A model with survey based catch estimates

How was data aggregated from OM to 1A EM? In the real-world application, you would generate a single area model input for index and compositional using all the raw data points. Survey abundances were accumulated across all regions and the average uncertainty used. Survey compositional were abundance weighted. And fishery compositions were catch weighted when aggregated.

Where, is the abundance input for single area model and is the regional abundance value from the five-area operating model. The composition data for the survey were aggregated across regions with an abundance weighting following.

The composition data from the fishery were aggregated across regions with a catch weighting following.

Reference points

The NPFMC has established harvest control rules based on a tier system (NPFMC 2018a, 2018b). The NPFMC tier 3 harvest control rule is applied to the biomass-based reference point to produce an ABC limit for each year, and the realized total allowable catch was assumed to equal the ABC. Under tier 3, ABC is calculated with an equal to , which is the fishing mortality rate that would lead to . is the biomass where spawners produced per recruit is at 40% of unfished biomass. If the spawning biomass/ ratio falls below 1.0 (EQ below), the harvest control rule reduces fishing mortality *F* below the projected ,

ABC was apportioned annually to management areas using a 5-yr exponentially weighted moving average of fishery and survey abundance indices, for each year *y* and area *m*. The survey index had double the weight (=2) of the fishery index (=1). This was the method accepted by the NPFMC for apportioning sablefish ABC for 2000-2013.

Where, is the exponential smoothing coefficients where . Catch was split between each gear type based on recent historical allocations.

# Appendix

* + Supplementary data exploration plots
  + Spatial model equations
  + Inputs
    - How were compositional data generated for input? Take from Bookdown
    - Indices? How did we deal with different countries etc.
    - Catch for 1960 Models