# Question 1

In the following sections, differences in data inputs, model structured and assumptions, and how management quantities are estimated will be presented for the three Pacific cod assessments that I reviewed. These include: 1) model 23.1.0.d for the Eastern Bering Sea (EBS; Barbeaux et al. 2023)), 2) model 19.1b for the Gulf of Alaska (GOA; Hulson 2023), and 3) models 23.2 (author preferred) and 13.4 (accepted/used for harvest recommendations) for the Aleutian Islands (AI; Spies et al. 2023).

## Data Inputs

***In the EBS Pacific cod stock assessment***, the earliest data input occurs in 1977, and the model begins in 1976. Model inputs for the EBS assessment include a ***single time-series of fishery catches***, which are summed across the longline, trawl, pot, and other (i.e., jig) fisheries. In addition, ***marginal*** ***size-composition data from the fishery*** are also utilized within the model, which are computed as a weighted average of catch proportions across hauls, vessels, months, gear, and area, only using samples with a minimum of 30 fish sampled for any month, gear, and area combination (i.e., aggregated size-composition data across fisheries). Additionally, input sample sizes (ISS) for fishery size-composition data represent annual hauls, which are then standardized to the mean survey size ISS, such that the mean ISS for both the survey size composition and fishery size compositions are equal. ***A model-based survey abundance index from the Bottom Trawl Survey (BTS)*** is also incorporated in this assessment, which is standardized using a Vector-Autoregressive Spatio-Temporal (VAST) model. The index is constructed with spatial and spatio-temporal fields encapsulating the EBS and the Northern Bering Sea (NBS) surveys following an autoregressive (1) process, and also allows for spatially varying responses with respect to the cold pool in the Bering Sea. ***Marginal age composition from the BTS*** ***(NBS + EBS)*** are also utilized in this assessment model, which are derived and expanded using a VAST model (i.e., uses model-based age composition “data”). Survey ISS for marginal age composition data are based on bootstrap methods (i.e., Hulson and Williams 2024). For a summary of the data sources utilized in model 23.1.0.d, see figure 1.

In the ***GOA Pacific cod stock assessment***, the earliest data input occurs in 1977, and the model begins in 1976. However, contrasting the EBS assessment, ***three separate time-series of fishery catches are utilized***, where the model utilizes separate time series of catches for the longline, trawl, and pot fisheries. ***Marginal size-composition data are also incorporated in the model for all three fisheries*** separately, where the ISS for fishery size-composition data are defined by the number of hauls, with a maximum of 200. In addition, ***conditional-at-age-length (CAAL) data are also utilized in the model and are available for all 3 fisheries*** beginning in 2007, where ISS are based on the number of ages sampled per length-bin, multiplied by 0.14. Two survey indices are utilized in this assessment, both of which are design-based. These survey indices are the ***BTS index and the Longline Survey (LLS) Relative Population Numbers (RPN) index***. For ***both surveys, marginal size-composition data*** are incorporated, and the ISS for these data are set at 100. Additionally, for the BTS, CAAL data are also used in the assessment model, where the ISS is based on the number of age samples per length bin, multiplied by 0.14. Lastly, the GOA model is unique in that ***it incorporates an environmental index representing bottom temperature anomalies*** and is linked to the catchability covariate in the LSS index. For a summary of the data sources utilized in model 19.1b, see figure 2.

Lastly, in the ***AI Pacific cod stock assessment (23.2)***, the earliest data input occurs in 1991, in contrast to the EBS and GOA assessments. Model inputs in the AI assessment include a ***single time-series of catch***, where catches are summed across the longline, trawl, and pot fisheries***. Marginal size-composition data*** are also utilized in the assessment model and are computed as a weighted average of catch proportions across years, areas, gears, and quarters. Here, ISS for marginal size-composition data for the fishery are based on the number of hauls, scaled to the mean survey ISS (i.e., following the same method as the EBS assessment). Similar to the other two assessments, the AI Pacific cod assessment uses a ***design-based abundance index from the BTS***. From the BTS survey, the model also ***utilizes marginal size-composition data and CAAL data***, where the ISS is defined using bootstrap methods developed by Hulson and Williams (2024). For a summary of the data sources utilized in model 23.2, see figure 3. In addition to the author-preferred model (23.2), ***a Tier 5 model (13.0) was also presented in the 2023 AI assessment, which only utilizes the BTS index.***

Given the different data sources described above, some key differences emerge. In particular, the ISS weighting scheme for fishery and survey marginal size-composition data and/or conditional-the AI and EBS assessments coincide with each other, where fishery size-compositions are scaled to the mean of survey size-composition ISS, and survey size-composition ISS and CAAL are derived using bootstrap methods from Hulson and Williams, 2024. By contrast, the GOA assessment weights samples either by the number of hauls with a maximum of 200 for fishery size-composition data, sets it at a constant of 100 for survey size-composition data, or sets it at the number of ages sampled per length bin multiplied by 0.14 for CAAL data. Additionally, another key difference across these models is that the GOA and AI assessments utilize CAAL data, while the EBS assessment uses marginal age-composition data. Further, the EBS and AI assessments aggregate catch and combine size-frequency data across the different fisheries, whereas the GOA assessment treats them separately. Across these assessments, all of them utilize abundance indices derived from the BTS, although there are some variations with respect to how they are treated. In particular, the EBS assessment uses a model-based method for the survey index, whereas the AI and GOA assessment utilize design-based indices. The GOA assessment is also the only one to incorporate abundance indices derived from the LSS, which is uniquely linked to an environmental covariate (bottom temperature anomalies) to explain changes in catchability.

## Model Structure and Assumptions

Most models discussed in this section are conducted using the Stock Synthesis 3 (SS3) platform (except for model 13.0 for the AI assessment, which is conducted using the REMA framework), where quantities are estimated using penalized maximum likelihood methods. All models conducted under the SS3 platform are single-sex and area, age-structured models that integrate a variety of data sources (which are discussed above). I will first begin by detailing the model structure of the EBS assessment, followed by the GOA, and the AI assessment.

In the ***EBS Pacific cod stock assessment, a single fishery fleet and survey fleet*** is modelled. There are several parameters that are assumed to vary annually, which are further described below. Given this, these parameters need to be constrained (as they would not be identifiable otherwise) by some variance term. These variance terms are iteratively tuned using the methods are Methot and Taylor (2011). Mean recruitment is assumed in this model, with lognormal deviations and a regime shift parameter (multiplicative) to describe recruitment prior to1977 (i.e., in 1976). The fishery fleet aggregates catches and size-composition data (described in data inputs) across 4 different gear types and assumes length-based double-normal selectivity, where certain parameters are fixed, such that it mimics the dynamics of logistic selectivity. ***Fishery selectivity here is modelled with two time-blocks*** (1977 – 1989, 1990 – 2023). A single survey is also assumed in this model, which is associated with marginal age-composition data, marginal size-composition data, and an abundance index, with a single catchability covariate (i.e., time-invariant) associated with the abundance index. Similar to the fishery, ***survey selectivity is assumed to be length-based and follows the formulation of a double-normal***, although with certain parameters fixed, such that the functional form mimics logistic selectivity. A notable difference between the fishery and survey here is that ***survey selectivity assumes annual additive deviations on the parameter describing the ascending-width*** of the double normal parametrization. A common theme with all 3 models is that growth is estimated internally within the model***. Growth in the EBS model uniquely assumes a Richards growth function***, which is described by 4 parameters where the ***Lmin and rho parameters are assumed to vary annually*** to capture changes in growth dynamics. Within the model, abundance indices assume a lognormal distribution, catches are assumed known, while compositional data (i.e., age and size) assume a multinomial distribution, where the ISS describing the weight assigned to these sample sizes are ***re-weighted using the methods of Francis, 2011*** to reflect effective sample sizes (ESS). Parameters and inputs that are estimated outside of the assessment model include: 1) ageing error/variability relationship, 2) the length-weight relationship, 3) length-based maturity, which is based on histological samples from Stark, 2007, and 4) ***natural mortality, which is assumed to be time-invariant and known within the model*** (0. 3866), which is derived using recently developed methods (phylogenetic structural equation models) from Thorson, 2024. Lastly, management quantities are estimated using traditional egg production methods (i.e., SPR40%) where values of F40% and B40% (SPR0% multiplied by mean recruitment) are derived and input into a sloping threshold harvest control rule (HCR) via Tier 3 methods, given that no stock recruitment relationship is assumed. Catches to the state are then further apportioned by multiplying the defined ABC from the EBS assessment by 0.12.

Unlike the EBS and AI assessments, the ***GOA Pacific cod assessment*** separately models its fishery fleets, where there are a total of 3 fisheries modelled (pot, trawl, and longline) and 2 surveys that are incorporated (BTS and LLS). Similar to the EBS assessment, mean recruitment is assumed in this model, with lognormal deviations and a regime shift parameter (multiplicative) to describe recruitment prior to1977 (i.e., in 1976). However, the variance parameter for recruitment here is not tuned, but is instead fixed at a value of 0.44. Instead of aggregating the catches and size-composition for the fishery, 3 fishery fleets are modelled separately, each with differing assumptions regarding time-variation in selectivity, as well as the underlying functional form (although all fisheries assume length-based selectivity). In particular, some interesting parameterizations regarding time-variation emerge. ***For the pot fishery***, a double normal parameterization is assumed for selectivity with certain parameters fixed (I did not have time to check to see which parameters they were), but allows for ***dome-shaped dynamics*** for larger individuals and is estimated in ***2 time blocks*** (1997 – 2012, and 2013 – 2022). ***For the trawl fishery***, a double normal parameterization is also assumed for selectivity, although certain parameters were fixed to allow the selectivity curve to mimic a logistic function. Furthermore, it appears ***that time-varying selectivity is modelled as a combination of annual deviations*** (periods prior to 1990; a total of 39 deviation parameters) constrained by a sigma of 0.2 ***and 4 time blocks*** (1990 – 2004, 2005 – 2006, 2007 – 2016, 2017 – 2023). ***Selectivity for the longline fishery is likely parameterized similarly.*** Here, the assessment details that annual deviations in longline fishery selectivity occurs for periods prior to 1990, with a total of 24 deviations parameters, of which are constrained by a sigma of 0.2 estimated. However, in Table 2.12, I note that there appears to be 15 additional fixed parameters estimated, which are most likely time blocks, but were not necessarily explicitly discussed, although I could have missed it, given limited review time. Similar to fishery selectivity, selectivity for the two survey fleets incorporated also follow the parameterization of a double normal function, although with some parameters fixed. ***For the BTS, 3 time-blocks are implemented*** (1990 – 1995, 1996 – 2006, and 2007 – 2002) and certain parameters are fixed such that ***selectivity mimics a logistic shape.*** ***For the LLS, selectivity is assumed to be time-invariant, with a single parameter fixed*** (i.e., 5 estimated), such that LLS selectivity reflects dome-shaped dynamics for larger individuals. Additionally, both surveys a fit to abundance indices, and thus catchability parameters need to be estimated. ***For the BTS, a single time-invariant catchability parameter is freely estimated. For the LLS, a single mean catchability parameter is estimated, coupled with an estimated “slope” parameter*** and the bottom temperature anomaly index to allow for multiplicative deviations in catchability for the LLS. This was done given that the minimum depth of the LLS was 150m, and that during warm years, cod would become more available to the LLS (presumably due to thermal refugia). Similar to both the EBS and AI assessments, growth is estimated internally. However, for the GOA assessment, ***growth is assumed to follow a von Bertalanffy function and is time-invariant.*** Furthermore, ***priors based upon a non-linear least squares approach estimated outside the assessment are imposed on von Bertalanffy parameters estimated within the assessment to constrain model estimates***, to help achieve convergence and stability. A key structural difference between the three Pacific cod models involves how natural mortality is estimated. In the EBS model, this is assumed time-invariant and is fixed based on methods from phylogenetic structural equation models. For the AI assessment, two time-blocks are assumed, and natural mortality is freely estimated (3 parameters; mean parameter and two deviation parameters). However, in the GOA assessment, natural mortality is estimated with a lognormal prior of -0.81 and a standard deviation of 0.41. Similar to the AI assessment, two time blocks are assumed (2014 – 2016, and all other years) although only two parameters for natural mortality are estimated here. Assumptions for likelihoods here are fairly similar across models, where compositional data (CAAL and size) assume a multinomial distribution, while indices assume a lognormal distribution. ***However, ISS in the GOA Pacific cod model do not appear to be iteratively re-weighted using tuning methods*** (i.e., Francis or Mcallister-Ianelli). For parameters that are estimated outside the assessment, these include: 1) ageing error/variability relationship, 2) the length-weight relationship, and 3) length-based maturity, which is based on histological samples from Stark, 2007, which are fairly similar to those described in the EBS assessment. Lastly, the estimation of management quantities here are similarly based off of SPR rates by using F40% and B40% as target reference points, coupled with a sloping threshold HCR. In contrast to the other assessments that only utilize a single fishery-fleet, multiple fleets here are incorporated, and thus, the reference selectivity used for SPR calculations were a weighted average of the fishing mortality rates across the three fleets from 2018 – 2023. Catches were then further apportioned by estimating the relative biomass in each area using the REMA framework (based on trawl survey estimates).

Lastly, the ***AI Pacific cod stock assessment (model 23.2)*** begins in the year 1990, in contrast to both the GOA and EBS, due to a lack of data prior to 1990. Similar to the previous models, this model assumes mean recruitment, although the deviation parameter here is fixed at 0.636, instead of being tuned (EBS) or fixed at 0.44 (GOA). Similar to the EBS model, the AI assessment assumes a single fishery fleet and survey fleet (BTS). In the author preferred model, ***length-based selectivity is assumed for both the fishery and survey and follow the double normal parameterization***, with certain parameters fixed, such that they ***mimic dynamics of logistic selectivity***. However, in the AI model, no time-blocks or time-variation in selectivity are allowed (at least in the author preferred model). Furthermore, with respect to the BTS index, a strong lognormal prior is imposed on the catchability parameter (mean = 0, sd = 0.01; essentially fixed), where in both the GOA and EBS models, catchability parameters were freely estimated. A key distinction in the author-preferred AI model is the parameterization of natural mortality and growth. In particular, two time-blocks are estimated for natural mortality to account for potentially increased natural mortality as a result of the marine heatwave, which was similarly experienced by GOA Pacific cod. Here, the time blocks correspond to 1991 – 2015 and 2016 – terminal year. However, instead of 2 parameters being used to describe two time-blocks, 3 parameters are estimated in the AI assessment (1 mean/base parameter with 2 deviation parameters). In addition to natural mortality, growth, which followed a von Bertalanffy growth function was also allowed to vary. Here, two time blocks were also estimated, and the kappa parameter (growth rate parameter) was allowed to vary from 1991 – 2003, and 2004 – terminal year. However, similar to natural mortality, 3 parameters are estimated here to describe 2 time blocks. Likelihoods are structured most similarly to the EBS assessment, where composition data (size and CAAL) are multinomially distributed with Francis re-weighting, while indices are lognormally distributed. For parameters that are estimated outside the assessment, these include: 1) the maturity-at-length relationship, which is derived using observed samples from the AI (and not from Stark 2007), and 2) the ageing error/variability matrix. Lastly, estimation of management quantities for model 23.2 (author-preferred), utilize SPR methods. Here, estimates for F40% and B40% are calculated as target reference points, where levels of fishing mortality are adjusted using the sloping threshold HCR. Catch allocations are then apportioned across areas using the relative trawl survey biomass data.

However, the author-preferred age-structured model was used for management advice this year. Thus, model 13.0 was utilized instead, which models the BTS index as a simple random walk process, utilizing a state-space framework:

where is the observed index of abundance (trawl survey biomass), are lognormally distributed deviates constrained by the variance observed from the trawl survey, and are treated as random effects and represent the true unobserved trawl survey biomass, which follows a random walk process:

where are deviations from a random walk, which are constrained by a process error variance term (. In the assessment, only one process error variance term is estimated as a fixed-effect, and are estimated as latent unobserved random effects. Essentially, this model smooths over the observed survey biomass as a state-space random walk model. Given that SPR methods are not possible when using these more data-limited methods, management advice here is derived using a hybrid approach. In particular, natural mortality is used as a proxy for Fmsy, where values of natural mortality are derived from the estimates from the age-structured stock assessment. ABC is then defined by multiplying the modelled trawl survey biomass by the estimated natural mortality rate and 0.75 (i.e., ABC = ), which is then apportioned across areas using the relative survey biomass across areas.

## Suggestions

Several aspects of these three models could be better aligned, although there are certain limitations to data quality and availability that would preclude such suggestions. Herein, I will attempt to make practical suggestions for each assessment, while briefly synthesizing some key differences among models.

Firstly, the Pacific cod fishery consists of several fishery fleets (typically 3). However, there are some differences across these models with respect to how fleets are structured. In particular, the EBS and AI models assume a single fishery fleet, while the GOA model treats fishery gears as separate fleets. However, the number of fleets modelled will likely depend on the available data and quality. Thus, I believe that it is justified to retain some differences in this respect. In particular, the AI model has incredibly sparse data to justify the modelling of 3 fishery fleets, as was observed in the 2022 assessment, and it is likely more appropriate to model a single fleet to aid in model convergence and stability. However, I will note that the EBS model combines 4 fishery gears and assumes 2 time-blocks that extend for a fairly long period. Upon inspecting the catch time-series, it appears that the proportion of catch resulting from these different gears likely varies over time (i.e., their Figure 2.5) and that selectivity likely differs across these gears. For the GOA Pacific cod assessment, 3 separate fishery fleets are modelled, although I found the parameterization of time-variation to be quite odd. Firstly, it is my belief that the whole point of modelling separate fleets (aside from the potential of gear allocations or better model diagnostics) is to allow for each fishery fleet to have selectivities that remain fairly time-invariant. However, in the trawl and longline fisheries, multiple time blocks coupled with annual deviations in selectivity were utilized to represent time-varying selectivity. It is unclear why the estimation of annual deviations in selectivity was reverted to vary as time-blocks. It is also unclear to me the benefit of modelling separate fleets with time-varying dynamics, when instead, gears could be combined into a single-fleet and allowed to vary continuously, which likely reduces the number of fixed effects parameters estimated. As such, a potential consensus suggestion would be: 1) only disaggregate gears as fleets if selectivity can be assumed to remain relatively time-invariant, and 2) if fleets are treated as aggregated, selectivity differences among gears are suspected, and the catch proportions vary over time, it will be of merit to allow for selectivity dynamics to vary over time.

Another difference between the three models is how growth is modelled. In the EBS model, a Richard’s growth function is assumed. By contrast, in both the AI and GOA models, a von Bertalanffy growth function is utilized. The former parameterization requires the estimation of additional parameters, and there does not appear to be a strong biological reasoning why growth relationships for the same species would differ such that alternative growth functions would be required (i.e., in what way does the Richard’s growth function better describe growth patterns/in what way does the VBGF fail to adequately describe growth patterns). Thus, I suggest a more thorough inspection of growth dynamics across these three regions, to understand whether a consensus growth function can be utilized or whether growth functions differ across regions for the same species. Furthermore, growth dynamics not only differ in their functional form, but also whether they vary overtime. Considering the sensitivity of Pacific cod growth to the prevailing environment, it is likely that growth varies over time (likely continuously), and adequately modelling these dynamics are necessary to achieve appropriate management advice. However, across these three assessment models, time-variation in growth is modelled fairly distinctly. In the EBS, the Lmin and rho parameter in the Richard’s growth function is allowed to vary continuously. In the AI, the kappa parameter in the von Bertalanffy growth function varies as a function of time-blocks (although there appears to be little difference across blocks). Lastly, growth in the GOA model is constrained by priors, and is assumed to be time-invariant. As a consensus, I believe that growth dynamics across these models should be allowed to vary, given the sensitivity of Pacific cod to the prevailing environment. Preferably variation in growth would be modelled to vary continuously, if supported by the available data and underlying model parameterization. When growth dynamics are assumed to vary continuously, I believe that deviations should not be placed on parameters themselves, but rather on the entire functional form. This approach is not only more parsimonious if multiple parameters are varying overtime, but circumvents the need to subjectively specify which parameters in a particular growth function varies. If such methods are not possible, objective model-based methods to detect breakpoints in growth are warranted (i.e., Kapur et al., 2021).

Another suggestion for harmonization is to ensure that the data-weighting and methods for determining ISS are the same across models. These parameterizations are the same between the EBS and AI assessments, where survey ISS is determined using a bootstrap method, while fishery ISS are scaled in accordance to the mean survey ISS. However, this is not consistent for the GOA assessment, and I recommend the GOA assessment mimic the methods of the EBS and AI. Additionally, ISS from the EBS and AI are re-weighted using Francis re-weighting – the GOA model does not appear to be iteratively re-weighted (not described in the assessment so assumed it’s not used), and I also recommend making this aspect of the GOA more consistent with the other Pacific cod assessments.

Finally, I do believe that the AI Pacific cod model should move towards an age-structured assessment, as it represents best available science, and because the Tier 5 methods that are currently used are relatively conservative and may result in suboptimal harvest strategies. I believe that with certain refinements, the author-preferred model (23.0) is close to being operational. In particular, I would remove the time-block on growth given the similarities in kappa. I also believe that the *natural mortality* time-block is well justified in this case, and does appear to help model performance significantly, but would recommend removing the estimation of the unnecessary parameters for these time blocks, as well as imposing priors on natural mortality. Lastly, given the gear-aggregated nature of this model (and in line with my first recommendation), I would recommend further explorations of time-varying selectivity with model 23.0, given that selectivity (trawl selects larger fish, pot and longline selects smaller fish and operate at different times) and catch proportions are likely to differ.

# Figures

A graph with blue dots

Description automatically generated

Figure 1. Summary of data sources utilized by model 23.1.0.d in the EBS Pacific cod stock assessment.

A chart of different colored circles

Description automatically generated

Figure 2. Summary of data sources utilized by model 19.1b in the GOA Pacific cod stock assessment.

A chart of different types of fish

Description automatically generated

Figure 3. Summary of data sources utilized by model 23.2 in the AI Pacific cod stock assessment.