# Cara Rodgveller (Groundfish Biology and Ecology)

### Skipped Spawning

Skipped spawning refers to a phenomenon where fish that are assumed to reproduce annually, elect to skip spawning within a given reproductive cycle. Skipped spawning appears to be a plastic trait in marine teleost fish and appears to occur at a more common rate than we think. Variation in skipped spawning tends to be associated with prevailing environmental conditions, often linked to food availability (poor feeding conditions), coupled with energetic demands and costs associated with reproduction. This has generally been demonstrated in laboratory studies, where fish fed with appropriate diet do not demonstrate signs of skipped spawning, while fish fed with inappropriate diets (i.e., less food) tend to show signs of skipped spawning. As such, the decision to skip-spawn is likely a trade-off between lifetime reproductive potential, natural mortality, and energetic costs associated with spawning activities (e.g., migration, development of oocytes, courting).

In some life-history models, skipped spawning is predicted to be a function of energy availability. However, some unrealistic results also result from such models, wherein skipped spawning also appears to be more prevalent when there is high energy availability, where increased somatic growth in a given year may outweigh the benefits and lifetime reproductive potential of spawning (i.e., grow very large such that your reproductive output in the following years are way higher). Furthermore, it also appears that in certain fish species, the prevalence of skipped spawning declines with age (i.e., Alaska sablefish). From a life-history perspective, this is intuitive because most fish species demonstrate reduced somatic growth at older ages, and the decision to spawn is likely a trade-off between maintenance and spawning at older ages, while at younger ages, the decision to spawn is a trade-off between maintenance, reproduction, and somatic growth.

In marine fishes, there is generally several types of skipped spawning, which broadly categorize into: 1) the resting type, and 2) the abortive type. Within the resting type of skipped spawning, secondary oocytes are not produced within a given spawning year, and primary growth is maintained across this duration. By contrast, the abortive type of skip spawning often results in a rapid reabsorption of secondary oocytes (atresia). Nevertheless, the type of skipped-spawning will often determine how well we can detect if it occurs.

In general, detecting the presence of skipped spawning is best achieved using histological methods, wherein a thick tunica or testicular wall can be indicative of skipped spawning, a lack of evidence for spermatogenesis, mass atresia of oocytes, and lack of evidence for the development of secondary oocytes coupled with thick tunica walls. Other non-histological methods for detecting skipped spawning include: 1) use of 17 estradiol, which controls the synthesis of vitellogenin to aid in the development of oocytes, 2) levels of water soluble proteins, 3) tagging studies to evaluate if tagged fish exhibit known movement behaviors as spawning fish, 4) the detection of mature individuals (use of size-at-maturity) away from spawning grounds during the spawning season, 5) use of otoliths and scale annuli to detect the presence of reduced growth during periods of spawning, and 6) use of stable isotopes to observe if fish traversed into spawning habitats.

### Time-varying Maturity

Similar to the issue of skipped spawning, the prevalence of time-varying maturity is most likely a function of prevailing environmental conditions coupled with food and energy availability, as well as competition from large cohorts. Given that size-at-maturity likely reflects variability in feeding conditions and density-dependent effects, we would *a priori* expect a lower size-at-maturity at declining population sizes due to reduced pressure in competition among individuals. Conversely, we may expect a larger size-at-maturity at large population sizes due to increased competition and reduced food availability, requiring individuals to grow to a larger size prior to fully investing into reproductive activities. ***This phenomenon might be expected because individuals are able to allocate more energy towards both growth and reproduction simultaneously, rather than a continued investment in growth to reach larger sizes, prior to maturing***. Other potential explanations for this mechanism could include individuals prioritizing reproduction rather than growth, in order to have offspring take advantage of favorable conditions. Additionally, time-varying maturity could also be due to genetic of fishery selection, wherein the case of fishery selection, individuals that mature quickly and hence grow quickly may be removed from the population, resulting in a shift towards older/larger maturity-at-size.

For Baltic Sea Cod, time-varying maturity appears to be present and related to SSB and recruitment. ***In particular, a larger size-at-maturity was detected when there was higher SSB and recruitment within the population, indicating potential interactions with food availability and competition in the population.*** Additionally, there appeared to be some metrics correlated with a lower size-at-maturity***, wherein increased reproductive volume and conditions in the Baltic Sea may have facilitated an earlier maturation size.*** From a life-history and biological perspective, these mechanisms are not surprising because during favorable conditions, individuals may be able to allocate more energy towards reproduction, despite being at a smaller size. ***Additionally, time-varying maturity may be governed by mechanisms such as the temperature-size rule, which dictates that younger individuals will grow faster and mature earlier, but reach smaller maximum body sizes***.

### Indices for reproductive potential

#### Stock Assessment Focus

Typically, indices of reproductive potential in stock assessments involves the use of spawning stock biomass as an index. However, this metric may not be appropriate for a variety of reasons. Firstly, I will list off the assumptions that the use of SSB in stock assessment assumes. Namely it assumes:

1. Relative fecundity is constant across ages,
2. SSB and reproductive potential (which was formulated as eggs production in Beverton-Holt’s formulation) are proportional,
   1. Reproductive potential may not be proportional because of difference in egg sizes, allometric scaling of fecundity and size, maternal age effects – i.e., protracted spawning for bet hedging,
3. All eggs are created equal and have an equal chance of surviving.

However, several aspects of fish biology invalidate the assumptions described above. In particular:

1. The above formulation assumes that all individuals contribute to spawning potential on an annual basis, which is invalidated by the presence of skipped spawning,
2. Not all eggs are created equally, wherein maternal age effects through increased provisioning towards eggs (e.g., increased volume of oil globules, larger eggs) can result in eggs produced from older individuals surviving at a higher rate. Similarly, older individuals might produce a larger number of eggs (i.e., relative fecundity is not constant), which may render the use of SSB in stock assessments invalid,
3. Males are surplus and do not necessarily contribute to spawning. However, skewed sex-ratios either due to the environment (increased temperatures) or the biology of the species can result in male-limited systems,
4. Age diversity, age structure and the presence of old individuals are assumed to contribute equally to reproductive potential on a relative basis compared to young individuals. However, this fails to account for hypotheses relating to entrainment, where migratory pathways towards spawning grounds are learned from following old individuals in the population (i.e., breakdown of social and migratory structures, difficulty in finding mates, stock contraction, which may reduce larval advection). It further fails to consider that reproductive potential can be enhanced with increased age diversity, as it can lead to a more robust and larger portfolio of reproductive strategies (i.e., protracted spawning season, overlap in spawning among different age groups), and may better buffer against prevailing environmental conditions,
5. In line with eggs not being created equally, second-time spawners compared to first-time spawners may disproportionately contribute to reproductive potential, presumably because increased experience from second-time spawners may result in higher survival rates for eggs produced.

In all of the cases discussed above, the use of SSB will tend to overestimate the underlying reproductive potential present in the population and optimistic stock status. Several studies have investigated the use of SSB in stock assessment with particular focus on relating alternative metrics of reproductive potential to estimate stock-recruit relationships. Briefly, I will first describe how different studies have formulated alternative metrics for reproductive potential.

1. A study on EBS snow crab (Murphy *et al.*, 2017) ***formulated an egg production index that accounts for: 1) first time-spawners and multi-time spawners, 2) crab size, 3) clutch fullness, 4) abundance, and 5) number of eggs produced.*** Here, they found that SSB and their egg production index was relatively correlated, although the relationship was not great. Furthermore, they found that assuming annual spawning tends to lead to an overestimation of reproductive spawning, where non-annual spawning was estimated using a temperature threshold.
2. A study on Icelandic cod (Marteinsdottir and Thorarinsson, 1998) attempted to improve the stock recruitment relationship by incorporating effects of age diversity. ***Here, they found that incorporating age-diversity (Shannon Weiner Index) in the x-axis of the stock recruitment relationship (e.g., Rec ~ SSB + Age Diversity)*** accounts for additional variation in a stock-recruitment relationship than just accounting for the individual variables themselves. ***The reason that age-diversity might be important in contributing to recruitment is because different age-classes might contribute to unequally to reproduction, more age-diversity may allow for a more protracted spawning season*** (more likely that individuals encounter favorable conditions), expanded spatial and temporal variability which gives rise to a more diverse reproductive portfolio, increased egg quality in some age-classes.
3. Similar to the study described above, a stock-recruitment relationship was reconstructed by using a bioenergetic index of reproductive potential to explain variation in recruitment for Northeast Arctic cod (Marshall *et al.*, 2000). ***Here, they used a metric of total lipid energy, environmental covariates (alongshore wind stress), and mean temperature to explain variation in the stock-recruitment relationship.*** The metric of total lipid energy is useful in this context because it is a useful predictor of egg production and varies as a function of food availability and is also correlated well with SSB. They found that recruitment variation can be explained by a combination of demographic characteristics of spawners (mean length, repeat spawners, total lipid energy), and environmental variables (temperature and wind).
4. Although the following study was not necessarily stock assessment focused, it demonstrates the importance of maternal age effects in Pacific Ocean Perch (Arnold *et al.*, 2018). ***Here, they found that older POP have increased provisioning to offspring, where they are provided with larger oil globule volumes to combat starvation effects and produced a larger number of eggs.*** Additionally, it appears that offspring earlier in the early season were of higher quality, potentially due to a variety of factors such as the benefit of potentially timing good quality offspring to the spring bloom. The other interesting finding was that older POP spawned over a protracted season. In general, they posit that offspring provisioning is a combination of maternal age-effects as well as environmental conditions, resulting in context-dependent maternal effects. ***Thus, this study illustrates the importance of age-diversity in reproductive potential through increased provisioning and protracted spawning seasons.***
5. The following study investigates the impact of skipped-spawning on stock assessment results in Alaska sablefish (Rodgveller *et al.*, 2016). ***Here they adjusted the functional form of maturity by considering skipped spawned fish as immature, which shifted the maturity curve to the right.*** This resulted in a lower overall mature population and an overestimation of SSB, if skipped spawning was not taken into account. This resulted in a lower F40% due to a lower SPR from right shifting the maturity curve.
6. The study described below was more of a review study, but discusses various aspects of the incorrect assumptions of SSB in stock assessment (Witthames and Marshall, 2008). They discuss that Beverton-Holt made two assumptions: 1) that all eggs survive equally, and 2) that SSB is proportional to egg production. However, there are several issues with these assumptions, particularly in the context of stock assessments because SSB is not necessarily proportional to egg production. The issues with SSB in assessments have already been discussed above, but in general, the authors discuss: ***1) the importance of age-diversity, 2) fishery and genetic selection shifting the size-at-age of maturity, 3) maternal age effects, where older females produce larger eggs and may have a protracted spawning season, 4) paternal age effects where older males can fertilize eggs over a longer period, 5) the influence of temperature on growth rates and sex ratios, 6) changes in growth can also impact fecundity, and 7) the issue of skipped spawning.***
7. The following study investigated the impacts of alternative metrics for reproductive potential on fishery reference points and stock recruitment relationships for Atlantic Cod (Murawski *et al.*, 2001). Similar to arguments presented above, they proposed that SSB is an inadequate metric because of the various assumptions it makes. Instead, they introduce some new indices, which directly account for maternal age effects, as well as increase larval quality of certain individuals. ***In particular, they account for: 1) hatch success of first-time and repeat spawners, 2) diameter of eggs produced by certain age classes and subsequent viability of larvae (proportion of larvae surviving to day 10).*** Here, they found that incorporating these indices led to downward estimates of spawning potential because a large proportion of age classes have only spawned once and that it was difficult to maintain a high level of SPR with high F, given that fishing truncates the age structure and these individuals contribute disproportionately to egg production and reproductive potential. Thus, using SSB overestimates the limit reference points. Furthermore, using alternative metrics results in reduced uncertainty in reference points likely because SSB determining Fcrash is poorly determined – i.e., Beverton-Holt SR relationship is less defined, resulting in higher uncertainty. ***They also found that limit reference points were less resilient when using alternative metrics of reproductive potential to estimate a stock recruitment relationship.***
8. In the last study discussed here, they review the importance of integrating reproductive biology into stock assessments, and discusses similar issues as those described above, although with particular focus on American Plaice in Canada (Morgan, 2008). ***In general, they similarly discuss the importance of capturing reproductive potential in the context of: 1) changes in growth, 2) sex-ratios, 3) changes in maturity-at-size or age, 4) alternative metrics of reproductive potential (age 9+ biomass, female SSB) which indicated the need to maintain a higher biomass as their limit reference point, while maintaining a lower F reference point, 5) skipped spawning, and 6) egg and larval viability.***

#### Biological Focus (Determining maturation)

Although the studies discussed above generally refer to reproductive indices and potential within the context of stock assessment models, ***reproductive indices are also useful from a biological perspective, especially in the context of determining whether individuals are mature or immature.***

In general, there are several approaches to determining the maturity and reproductive status of a fish. The easiest method is to use macroscopic methods, wherein we macroscopically examine the gonads to subjectively determine whether a fish is deemed mature or immature. However, macroscopic methods can often be inaccurate, especially because the determination is often not made by experts. Furthermore, the utility of macroscopic methods can depend on when fish are sampled. In particular, it is easy to misclassify fish as immature if sampling takes place early relative to the spawning season, because vitellogenic oocytes may not have recruited (Rodgveller, 2018b) and it may not be obvious for a unexperienced scientist whether an individual was mature or immature. As such, macroscopic methods should always be validated with microscopic histological methods. Another strategy could be to sample during the spawning season. For Alaska sablefish, sampling prior to spawning in December allows for a more accurate determination of maturity status. A comparison of macroscopic, standardized macroscopic, and histological methods showed that macroscopic methods that were unstandardized underestimated the size-at-maturity, when compared to standardized methods and histological methods (Rodgveller, 2018b). Based upon recommendations from Rodgveller 2018b, they suggest that maturity classifications can be made accurately outside of the spawning season if trained scientists are available to make such determinations, or if funding structures are in place to make histological determinations. However, histological methods are required to determine the presence of skipped spawning.

Beyond macroscopic, standardized macroscopic, and histological methods to determining maturity status, alternative approaches are also available, although they are indirect methods. In a study for West Coast sablefish, the following indices were used to infer reproductive status: ***1) hepatosomatic index (HSI), which can be of utility because the liver stores energy content used for reproduction, 2) gonadosomatic index (GSI), which is of utility because the gonads get large during reproduction, 3) levels of 17- estradiol and 11-ketotestoterone for female and male sablefish, which are both used for regulating the synthesis of hepatic vitellogenesis and spermatogenesis for females and males, respectively, and 4) Fulton’s condition index***, which measures length relative to weight – a higher condition presumably relates to a more developed reproductive status. Generally, all of these indices appeared to track well with different reproductive stages for sablefish. In particular, from the time-scale of January to December, all indices were highest in January representing the general onset of the spawning season, gradually declined until August, and then began to increase back up from September to December, potentially relating to processes such as the onset of secondary growth, recruitment of oocytes and vitellogenesis.

A more comprehensive study outside of sablefish was also conducted based on three flatfish species in the northeast, which compared a variety of methods for developing reproductive indices (Wuenschel *et al.*, 2019). ***These metrics included: 1) Fulton’s condition index, 2) Fulton’s condition index but on an allometric scale – because not all fish are cubes, 3) a scale mass index, 4) total bioelectrical conductivity, 5) bioelectric impedance analysis, 6) GSI, 7) HSI, 8) % dry liver weight, 9) % dry muscle weight.*** Here, they found that variation in physiological measures often depended on the life-history of the species. In particular, it was obvious that physiological measures had seasonal cyclic patterns denoting when the onset of spawning occurred for determinate spawners, where the number of oocytes are pre-determined, ***whereas species with indeterminate fecundity and a protracted spawning season had less variation in these measures, and maintained a relatively stable level throughout the course of the year.*** This occurs because indeterminate spawners are usually income breeders, where they will use short term lipid energy as opposed to storing them for later use, resulting in less variability. ***In general, the use of specific indices will depend on the purpose of your study. The use of bioelectric methods was fairly unsuccessful throughout this study. Fulton’s scale allometric relationship is useful for cross-scale comparisons, HSI and GSI are only useful for within species comparisons and for assessing seasonality in reproductive, while % dry weight liver or muscle were good for evaluating the total lipid energy, as well as condition indices. Furthermore, given that these indices vary throughout the season, the collection of annual measures should best be conducted in a standardized manner (i.e., within the same months).***

### Sablefish life history and biology

Sablefish are a long-lived and fast-growing demersal species that are commonly found in depths beyond 300m. They tend to exhibit ontogenetic movements, where young juveniles inhabit inshore regions and bays. As they grow older, they’ll begin migrating towards offshore areas, onto the outer continental shelf and the slope where they spawn at around ages 5 to 6 (Mason *et al.*, 1983; Rodgveller *et al.*, 2016). Most sablefish spawn at depths of up to 300m, near the continental slope. Spawning generally takes place around November to April, where it tends to peak January to mid-February (Mason *et al.*, 1983). This is similar to sablefish in coastal Washington, where they tend to be spawning around November to February, but spawning peaks from December to January (Guzmán *et al.*, 2017). Following the spawning season, larvae catches peak around April (about several weeks), suggesting that there is a gestation period before eggs emerge as larvae into the neuston.

In general, there is a hypothesis that most spawning occurs in the eastern and central Gulf of Alaska. Eggs occupy depths deeper than 400m, where hydrographic conditions are fairly stable, and likely rely on cross-shelf fluxes to advect eggs into onshore regions. Larger females produce more eggs than smaller females (Mason *et al.*, 1983), and this relationship tends to scale in proportion to the length or weight-at-age of an individual, where the relative fecundity of sablefish is age-invariant (Rodgveller *et al.*, 2016). ***While relative fecundity is age-invariant, older individuals tend to produce larger eggs presumably due to maternal age effects***, ***which may confer increased survival rates*** (Rodgveller *et al.*, 2016). However, there is also some conflicting evidence suggesting that relative fecundity actually decreases with age (Rodgveller, 2018a), although this seems unlikely and is most likely due to noisy data.

Sablefish are batch spawners with synchronous oocyte development and exhibit determinate fecundity, where there is a pre-determined number of oocytes that recruit into development. Their reproductive cycle is as follows: 1) during March – April, oocyte development begins, where the onset of secondary growth occurs, 2) from May – July, early vitellogenesis occurs where yolk proteins recruit into oocytes and being their development, resulting in an increased production of steroid hormone levels, 3) from July – September, mid-vitellogenesis occurs, which is indicated by a continued production of steroid hormone levels and an increased diameter of ovarian follicles, 4) from September to November, late-vitellogenesis occurs, 5) from November to February, spawning begins, and 6) February to March denotes individuals that are in post-spawn condition (Guzmán *et al.*, 2017). ***Despite an increase in ovarian follicle diameter throughout the spawning season, it is evident that there are some follicles that do not recruit***, and remain in the perinucleolar stage, where they do not increase in diameter. The above cycle is described for female sablefish. For male sablefish, sperm tends to recruit around April, which quickly develop into spermatozoa stages in July and begin sperminating from November to April. Males tend to exhibit a longer protracted spawning season compared to females, and may reflect a diverse portfolio of reproductive strategies (Guzmán *et al.*, 2017).

Although sablefish are determinate spawners where oocytes recruit as large groups for development, they also exhibit characteristics of skipped spawning. Although there are many types of skipped spawning, sablefish exhibit the “resting” type, where skipped spawning is indicated by a lack of mature oocytes, a thick tunica or ovariance wall, and blood vessels in the lamallae, instead of rapid reabsorption of vitellogenic oocytes. As discussed above, skipped spawning is a plastic trait likely in response to unfavorable environmental conditions and a tradeoff between mortality, reproduction, and somatic growth. This is similarly evidenced in Alaska sablefish (although not evident in WA sablefish; Guzmán *et al.*, 2017), where skipped spawning rates was about 21% in 2011 and was 6% in 2011 (Rodgveller *et al.*, 2016; Rodgveller, 2018b, 2018a). The authors hypothesize that this may be related to a switch in sign in the PDO, where positive PDO events have been related to high recruitment events and presumably favorable conditions for sablefish. Other reasons for skipped spawning could be due to the long migrations that sablefish undertake, and that it may be more beneficial in a given year to elect to skipped spawn as opposed to partitioning energy to undertake a long distance spawning migration (Rodgveller *et al.*, 2016).

Additionally, although there is conflicting evidence with respect to age and the prevalence of skipped spawning, follow up studies have indicated that the prevalence of skipped spawning likely decreases with age. This is most intuitive (as opposed to increasing skipped spawning with age) because individuals at older ages only really require energy for maintenance and reproduction as opposed to maintenance, somatic growth, and reproduction (Rodgveller *et al.*, 2016; Rodgveller, 2018a). ***Given that Rodgveller 2018a found that skipped spawning tends to occur in individuals that have yet to fully mature, this suggests that younger individuals likely prioritize growth rather than reproduction, which may be beneficial for increasing the overall lifetime reproductive potential of an individual***. Some considerations regarding the modelling of skipped spawning include: 1) functional vs. true maturity, 2) whether functional maturity is dome-shaped or asymptotic – where dome-shaped functional maturity can occur if older fish are skip spawning.

Considering sampling methods for Alaska sablefish, maturity is poorly determined using at-sea macroscopic methods, whereas the determination of maturity status can be improved with standardized macrosopcic and histological methods. ***Using macroscopic methods that are unstandardized can result in an underestimation of maturity-at-size or age.*** For sablefish in coastal Washington, the determination of maturity status via macroscopic methods is most likely to be accurate during the early-late fall periods (Guzmán *et al.*, 2017). ***A similar conclusion was drawn for Alaska sablefish where sampling during the summer might not be representative because sablefish might initiate maturity after sampling and be deemed immature.*** Thus, sampling would likely be the most accurate during the last leg of the survey which takes place late August – September, ***because it is easiest to distinguish oocytes that are in the vitellogenic stages vs the perinucleolar stages*** (i.e., oocytes are lkely to be in the developed stages of vitellogenesis; Rodgveller, 2018b). ***However, outside of the survey, sampling during December would lead to the most accurate results,*** because that represents a period prior to spawning, where oocytes would have developed.

# Pete Hulson

## Magnusson Stevens Act and Management Objectives

The MSA is the primary legislature governing fisheries and uses the precautionary approach where the Council incorporates forward looking conservation measures that address differing levels of uncertainty. Further, the Council seeks to make decisions based on sound scientific decision and provide management proactively rather than reactively. FMP plans ***have to*** conform to the 10 National Standards set forth by the MSA. These are:

1. Prevent overfishing and achieve OY on a continuing basis,
2. Conservation and management based on the best available science,
3. Manage individual stocks as a single unit and coordinate with neighboring stocks if necessary,
4. Conservation measures should not be discriminatory or benefit others inequitably,
5. Conservation and management measures should account for variability in catch and the resource,
6. Conservation and management measures should be as efficient as possible,
7. Conservation and management measures should minimize costs,
8. Conservation and management measures should avoid harm and hardship to communities (economic) and provide opportunities for participation in the fishery,
9. Bycatch should be minimized,
10. Safety of human life should be prioritized.

In terms of management objectives for the council, they are broadly grouped into the following categories, which are separate from the MSA, but conform to them:

1. ***To prevent overfishing*** (conservative measures of harvest, evaluation of F40 rule, 2 million mt OY cap, adaptive management for dynamic range of OY),
2. ***Promote sustainable fisheries and communities*** (conservation measures balance both harvest of socio-economic considerations, equitable allocation of resources, and safety at sea),
3. ***Preserve food web*** (development of ecosystem indices, adjust ABC based on ecosystem factors, limit forage fish harvest),
4. ***Reduce bycatch*** (formation of bycatch incentive programs, bycatch limits and research on population status of non-targeted species, reduce economic-related discards, seasonal allocations and gear restrictions, account for bycatch mortality in TAC, reduce waste and maximize retention and utilization),
5. Avoid impacts to sea birds and sea lions,
6. Reduce impacts to habitats (area closures, HAPC, EFH designation and mapping),
7. Promote equitable and efficient fishery resource use,
8. Increase Native Consultation,
9. Improve data quality, monitoring, and enforcement.

## Harvest Specifications

Definitions:

1. Maximum Sustainable Yield – the largest long term catch that can be taken from a stock or complex under the current prevailing environmental and fishery characteristics,
2. Optimum Yield – the amount of fish that provides the greatest benefit to the nation, is prescribed according to MSY with adjustments due to uncertainty, and provides a rebuilding plan to which allows us to reattain MSY if the stock is overfished. For the BSAI, the system wide OY is equal to 1.4 million mt – 2 million mt, while in the GOA, the OY is 116,000 mt – 800,000 mt.

Generally, harvest specifications follow a hierarchy. This hierarchy is as follows: 1) OFL = Maximum Fishing Mortality Threshold, 2) ABC, 3) ACL, which is generally equal to ABC, 4) TAC, which can be adjusted down from the ABC due to uncertainty, and 5) MSST (minimum stock size threshold), below which we are not allowed to fish (i.e., below ½ of B35% for Tier 3 stocks).

TACs are defined in the following manner: 1) define ABC, 2) define a TAC based on some socio-economic and ecological information – the TAC can be below ABC, but not above it, 3) TACs are summed across species to ensure it falls within the range of OY (2 mt), 4) TACs are then apportioned across areas and gear sectors.

## NPFMC Management Measures

1. OY set at 1.4 million mt to 2 million mt, which may not be exceeded,
2. Area closures to reduce bycatch of salmon, crab, herring, crab, etc,\
3. Area closures and transit closures to protect stellar sea lion foraging areas and walrus transit areas,
4. Area closures to protect unique habitats (Bowers Ridge, Alaska Seamounts),
5. Prohibited Species Catch limits and caps, wherein exceeding these limits can lead to area and fishery closures (Chinook, Herring, Crab),
6. EFH designations to understand the life-stages at which certain species are most vulnerable (although not really extensively used in management),
7. Use of prevailing ecosystem information and uncertainty in the management process to reduce ABC to set TACs,
8. Input controls to limit vessel tonnage, horsepower, and size,
9. No directed commercial fishery allowed on forage species

## Tier System and Control Rules for Stock Assessments

* Note that if reliable information for maturity does not exist, a knife-edged maturity assumption can be made.
* For the setting of OFL, these also represent the maximum fishing mortality threshold. The OFL is then applied to the best estimate of biomass, which may or may not be derived from an age-structured model.
* The setting of ABC is flexible, in that it provides an upper bound for harvest. The ABC control rule captures uncertainty in two ways: 1) structured based on the available information, which is qualitatively related to uncertainty, and 2) in the Tier 1 specification, the use of geometric and harmonic means accounts for scientific uncertainty.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Tiers | OFL | ABC | A (B/Bmsy >= 1) | B (on the ramp) | C (B/Bmsy < ) | HCR (Threshold) | Stock Recruit |
| 1 | FOFL = arithmetic mean of Fmsy | FABC = harmonic mean of Fmsy | Fmsy | Fmsy x (B/BMSY - α)/(1 - α) | 0 | Only uses Bmsy reference points | Yes (Age-structured) |
| 2 | FOFL = Fmsy | FABC = Fmsy \* (F40 / F35) | FMSY × (F40% /F35%) | FMSY × (F40% /F35%)× (B/BMSY - α)/(1 - α) | B/BMSY ≤ α | For ABC setting, uses Bmsy and Fx% | Yes (Age-structured) |
| 3 | FOFL = F35 | FABC = F40 | F40% | F40% × (B/B40% - α)/(1 - α) | 0 | Uses both Bx% and Fx% | No (Age-structured) |
| 4 | FOFL = F35 | FABC = F40 | NA | NA | NA | NA | No (Generally Age-structured) |
| 5 | FOFL = M | FABC = 0.75 \* M | NA | NA | NA | NA | No (Generally index method) |
| 6 | FOFL = Avg Catch (1979 – 1995) | FABC = 0.75 \* Avg Catch (1979 – 1995) | NA | NA | NA | NA | No (Catch only methods) |

## Status Determination

**Overfishing Status: *Occurs when catch is too high and the level of catch exceeds the maximum fishing mortality threshold*** (i.e., the OFL). If overfishing did occur, amendments need to be in place to terminate overfishing activities. Note that overfishing and overfished is not the same thing. The key difference is that an overfished stock can be fished in a sustainable manner, while overfishing can occur on a healthy stock that is not overfished.

**Overfished Status:** ***Occurs when stock falls below the minimum stock-size threshold***, which is defined as ½ of Bmsy or B40% depending on the tier of the stock. For tiers 4 – 6, since there is not estimate of Bmsy, overfished status is undefined. Thus, to declare overfished status:

1. if the current year SSB is less than ½ of Bmsy, then the stock is overfished,
2. if it is above, it is not overfished,
3. if it is in between ½ of Bmsy and Bmsy, then we need to do stochastic simulations for 10 years, while fishing at the OFL (F35% or Fmsy) and look at the mean SSB over that 10 year span – if it is below ½ of Bmsy, then the stock is overfished.

Once declared overfished, a rebuilding plan must be put in place with levels of FOFL and Fmsy that will rebuild the stock in a reasonable time frame.

**Approaching Overfished Status**: MSA requires us to determine whether a stock is approaching overfishing. This is done by projecting the current year SSB forward by two years and fishing at ***maxFABC*** to see if we fall below ½ of Bmsy:

1. if we project two years forward and are below ½ Bmsy, then we are approaching an overfished status,
2. if we project two years forward and are above Bmsy, we are not approaching an overfished status,
3. if the projection results in being above ½ Bmsy but below Bmsy, then project it forward 2 years fishing at maxFABC, and then project it an additional 10 yaers forward fishing at OFL (Fmsy or F35%). If we are below Bmsy in the 12th year, then we are approaching an overfished condition.

***For stocks assessed under tiers 4 – 6, it is not possible to determine whether it is overfished or approaching an overfished condition.*** ***Such determinations requires an estimate of B20% and B17.5%. However, it is possible to conclude whether overfishing is occurring, if catches surpass the OFL.***

## AI Pacific Cod

### Biology and Life-History

Pacific cod generally occur at depths from the shoreline to about 500m. Eggs are demersal and are semi-adhesive (eggs tend to hatch around 16 – 28 days), which aids in egg retention, and the larval duration is about 90 days. The survival of eggs and hatching success have been found to be dependent on temperature – recruitment is also highly influenced by temperature. Eggs move quickly to the surface upon hatching and larval stages have fairly good swimming abilities. Larval drift can be great, where larval stages have been found to be transported by currents from the Kenai and Kodiak to Unimak, which suggests some degree of connectivity between the western GOA and Aleutian Islands. Larvae tend to exhibit shoreward movement (i.e., inshore) but this relationship can vary.

Juvenile Pacific cod tend to settle near the seafloor and can also be dependent on nursery habitats (as inferred from Atlantic cod). Habitat use of juveniles is generally in shallower waters from coastal-demersal to shelf-pelagic (0 – 80m). Habitat distribution of juveniles is hypothesized to be as a result of density-dependence, temperatures, and prevalence of demersal predators.

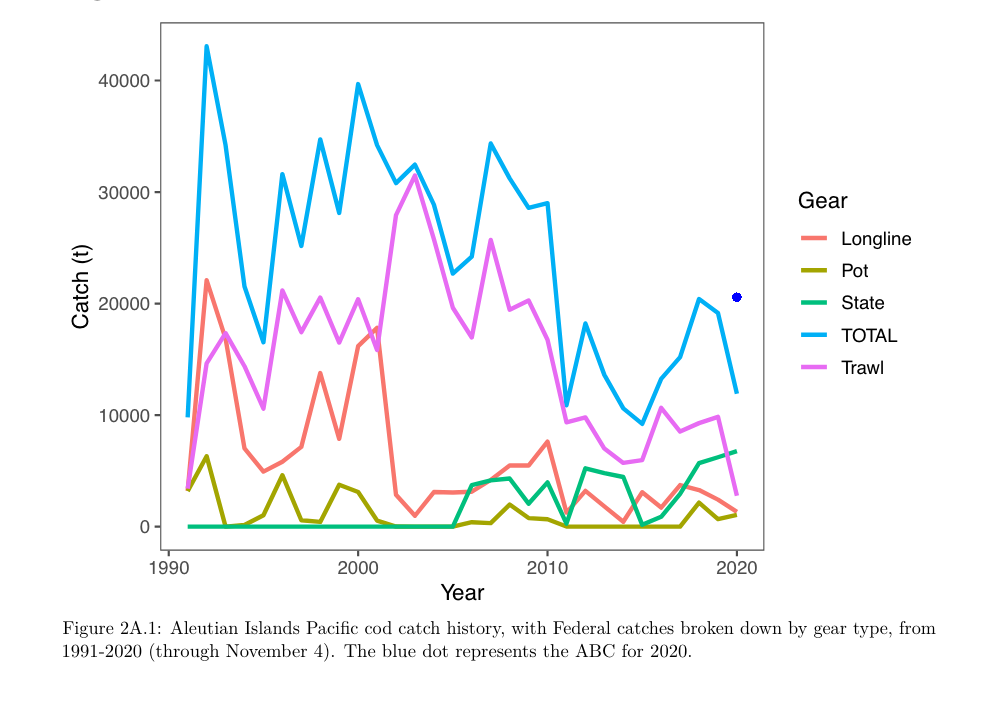
Although juveniles have some flexibility in being somewhat pelagic to demersal, adults are strongly associated with the seafloor and diel vertical migration has been observed. Adults tend to form large spawning aggregations and are annual spawners. Adults tend to spawn around the months of February through April. Spawning aggregations have been found near Zhemchug Canyon, Pribilof Islands, and adjacent islands in the central and western Aleutian Islands. After spawning, they undertake spawning migrations as well as feeding migrations (has been observed to travel about 100nmi to 500nmi).

The diet of Adult Pacific cod in the Aleutian Islands is generally composed of Atka Mackerel, although in recent years, this prey item has declined in proportion, corresponding to declines in Atka Mackerel, which may reflect a lower quality diet item for Pacific cod. However, Pacific cod are a generalist predator and are able to switch between benthic to pelagic prey, vice versa. Furthermore, in the BSAI and GOA, much of their diet is composed of polychaetas, amphipods, and crangonid shrimp. Other common items include, euphausiids, miscellaneous fishes, pollock, offal, yellowfin sole, and crustaceans. Juveniles are more invertebrate feeders, while adults are more piscivorous and can be cannibalistic.

### Fishery Characteristics

The fishery is a multi-gear fishery that is composed predominately of trawl, pot, and hook-and-line gear. The catches among these gears fluctuate quite a bit and the selectivities are likely different, which may suggest the need for time-varying selectivity. ***Nonetheless, most of the catch results from trawl gear, followed by longline gear, and pot gear***. ***The fishery during the feeding season tends to be longline gear, while trawl nets are typically used during the spawning season, which tend to select larger fish.*** Given the migratory patterns there and the different gears used, some thought needs to put into how selectivity should be structured in this respect.

Furthermore, given that the AI region is managed as three separate regions (central, western, and eastern), the overall fishing mortality rate in each of these sub-areas differ quite a bit (East: 61%, Central: 26%, and West: 13% in 2020). These differences in fishing effort are likely a function of: 1) distance from port, and 2) due to Stellar sea lion protective regulations. Considering that differences in fishing mortality can lead to dome-shaped population selectivity patterns, despite the fishery demonstrating logistic gear selectivity (i.e., due to differential survival in each area), it may be prudent to investigate the use of dome-shaped selectivity in the assessment. Additionally, fishery length composition data suggest that larger fish are caught in the winter (spawning aggregations) relative to the summer months. Furthermore, most of the composition data results from trawl and hook-and-line fisheries, with limited samples collected for the pot fishery.



### 2020 AI Pacific Cod

#### SSC and Plan Team Comments

1. SSC suggests exploring averaging multiple surveys (not sure if there is data for that) as well as using a VAST model for the purpose of apportionment, but not updates were made and no VAST models were explored due to a lack of a survey,
2. The SSC also recommends fitting the maturity curve inside the assessment,
3. The SSC appreciates the efforts in the exploration of age-structured methods and recommends further explorations. They recommend using *M­*-prior methods from Jason Cope to explore the estimate of natural mortality used in the assessment. This was not done in the current assessment cycle. The age-structured model was not updated due to a lack of survey data.

#### Assessment Structure

The 2020 AI Pacific Cod stock assessment is defined as a Tier 5 stock ***(REMA model).*** As such, the FOFL is defined as *M* (0.34, which uses Jensen’s age-at-maturity estimator) while the FABC is defined as 0.75*M* (0.255), and the resulting catch advice would be the biomass estimated, multiplied by these quantities. Thus, the 2019 AI Pacific Cod assessment uses a simple random effects model, wherein the index of the NMFS Bottom Trawl Survey in the AI region is used (triennial early on and biennial in recent periods):

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 using a state-space random walk model.

Considering that this is a tier 5 stock, an estimate of *M* is needed to provide management advice, which is multiplied with to provide an estimate of catch advice (either *M* or 0.75*M*). This value was updated several times in accordance with natural mortality assessments from the EBS Pacific cod assessment, but this practice is no longer done given concerns with the EBS *M* estimate not necessarily being equal to that of the AI. Thus, I believe the authors here have continued their use of Jensen’s age-at-maturity estimator for *M*, which is currently specified at 0.34.

The assessment estimates fairly wide confidence intervals for 2020, ***which they attributed to a lack of survey data since 2018*** (i.e., no survey data in 2019 and 2020 because of COVID).

#### Harvest Apportionment

For AI Pacific Cod, there are several issues that need to be addressed with respect to apportionment of harvest. These include: 1) apportionment of harvest from the state, and 2) apportionment of harvest given Stellar Sea Lion protection measures.

Prior to 2014, apportionment of BSAI Pacific Cod for the state was done by multiplying 3% of the TAC from the assessment. Following 2014, when the AI Pacific Cod stock was managed separately from EBS, apportionment to the state was done by multiplying about 27% - 39% of the AI Pacific Cod to the state. The percentage increases by 4% if the catch reaches 90% of the guideline harvest level from the state in the previous year, but may not exceed 39% (6804 t). Further, the federal + state harvest may not exceed the ABC.

Prior to 2014, there was a regulation that prevented the harvest of Pacific Cod in Area 543 (in the AI region). However, there was an amendment in 2015, that now puts a harvest limit in Area 543, instead of fully restricting fishing. There, an apportionment of the TAC is made by:

where is the harvest limit in area 543 and is calculated by subtracting the TAC by the guideline harvest level from the state, and then multiplying the proportion of biomass in area 543 relative to the total biomass of the assessed stock (). This can be calculated several ways including: 1) using the average survey raw proportions, 2) the most recent survey raw proportions, 3) using the average survey estimate from Area 543 and dividing that by the average survey estimate from the entire area, and 4) using the most recent survey estimate from 543 and dividing that by the survey estimate in the most recent year. The last approach is what was used in this assessment.

### 2021 AI Pacific Cod

#### SSC and Plan Team Comments

1. SSC suggests exploring averaging multiple surveys as well as using a VAST model for the purpose of apportionment, but not updates were made and no VAST models again this year,
2. The SSC appreciates the efforts in the exploration of age-structured methods and recommends further explorations. They also recommend fitting the maturity curve inside the assessment but this was not done due to potential confounding with ageing error – which I do not think is a great excuse – fitting the maturity curve in the assessment would allow for the propagation of ageing error in the maturity curve,
3. Several growth models were investigated, and it was eventually decided based on AIC and parsimony that the von Bertalanffy growth model would fit the best,
4. The SSC requested the author to look into the use of Jason Cope’s method for determining *M*. This was done in this assessment cycle and the estimate came out to 0.36, although a point estimate of 0.4 was used in the assessment model this cycle,
5. The Plan Team wanted the author to bring forward both model runs of maturity (observer and Stark 2007). The teams recommended the observer data because it has more samples and is more representative than the Stark (2007) estimates, but wanted histology verification,
6. The authors requested guidance on data-weighting from the SSC, but no data-weighting exercises were really attempted this assessment cycle (survey ages weighted by hauls, and fishery was set to standardize at a mean of 20 for ISS),
7. Results from 3 age-structured assessments were brought forward this cycle, and the SSC recommended not bringing forward one of the models that dropped fishery length data. These assessments differed in 1) the estimate of M (0.34 vs. 0.4) and 2) the maturity curves used.
8. Despite the *M­-*prior methods suggesting a point estimate of 0.36, the SSC and authors recommended a value of 0.4, to balance the tradeoff in the likelihood profile indicated by the fishery and survey (0.3 vs. 0.8). The estimate of 0.4 follows the general mode of *M* used in this assessment in previous years. However, there is no firm justification as to why the value of 0.4 was really chosen.

#### Assessment Structure

##### Data

The 2021 assessment brings forward 3 age-structured assessment models:

1. ***old maturity curve + M = 0.34,***
2. ***new maturity curve + M = 0.34,***
3. ***new maturity curve + M = 0.4.***

and one tier 5 assessment model (rolled over from 2020). For the tier 5 model, it uses what was described in the 2020 AI Pacific Cod assessment model – BTS biomass estimates using a random-walk smoother.

The age-structured assessment model uses the following datasets:

1. Fishery catch (1991 – 2021; more during spawning season (winter), larger fish)
2. Fishery size compositions (1991 – 2021)
3. Biomass index from BTS survey (biennial and triennial, most recent = 2018; more during summer months, smaller fish),
4. Age composition from BTS survey (biennial and triennial, most recent = 2018)

In the age-structured assessment model, there are a decent number of age-composition samples available for use in modelling (500 – 1000 samples per survey year). Length-composition data are not used for the survey because age-data are available (although conditional age-at-length might be used). The survey biomass index is a design-based index that is expanded to the strata and summed. The NMFS LL survey is also conducted in this same region, which I believe to be biennial. I believe that the author should make efforts to incorporate these data into the age-structured model and investigate the quality and quantity of length and age-composition data from the LL survey.

For fishery data, gears and statistical areas from AI are combined and modelled together. Catch data are summed. For length composition data, most of these data come from longline and trawl fisheries. The assessment document indicates that length compositions are combined by weighting by the relative catch in each statistical area. Depending on how well each gear, season, and area overlap, this might not be an appropriate approach, and the length frequency data should instead be weighted by relative catch across seasons, gears, and areas given that there appears to be differences in the sizes of fish caught depending on these characteristics (larger fish during winter months, which tend to be trawl gears). However, given that most of the fishing takes place in the winter months, season weighting might not be necessary. Nonetheless, I think gear \* area might be an appropriate relative weighting scheme for these data.

##### Model Structure

The model used is an age-structured model with a single-sex and a 1:1 sex-ratio (10 ages, 10+ is the plus group). The survey and fishery both assume logistic age-based selectivity. A growth function following von Bertalanffy dynamics is estimated outside the assessment, which is used to construct an age-length matrix as well as compute weight-at-age. An ageing error matrix is used in the assessment as well. Recruitment and fishing mortality parameters are time-varying, while all other parameters are constant. Survey catchability and selectivity are estimated within the assessment, while maturity is estimated outside of the assessment. Natural mortality is fixed inside the assessment (either as 0.34 or 0.4) and fishery length frequencies are weighted by the relative catch in each statistical area.

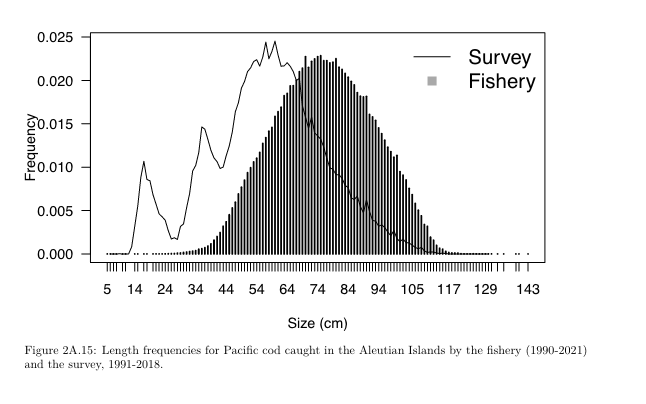
Data-weighting methods were attempted to weight survey age-composition and length frequencies but led to unreasonably high likelihood weights – this resulted in decreased survey catchability and biomass estimates. They ended up weighting the samples using the number of hauls in each year for survey age compositions. Different weighting values resulted in expected changes in fits to data sources. Increasing weights to length-frequencies led to better fits to these data, but poorer fits to survey indices. However, this tended to result in poor convergence criteria. Thus, they ended up weighting the length frequency data as the number of lengths sampled to retain annual sampling variability, but weighted so that the mean of the weights were 20, so as not to overwhelm the model by forcing fits to length data.

A new maturity curve was also implemented in this model, because the old maturity curves was based off the EBS and had limited samples. The new maturity curve was similar but had a slightly lower age-at-50% maturity (4 years old). The maturity curve was estimated outside the assessment using length-based maturity, and was then converted to age-based.

Length-at-age and associated age-length keys were also constructed using a von Bertalanffy growth model, instead of more highly parameterized and flexible models as they all performed similarly. The age-length transition matrix was constructed by simulating values conditioned on the mean length-at-age and its associated CV/variance. Length-at-age was then converted to weight-at-age using an allometric relationship.

With respect to natural mortality, a value of 0.34 was used based on age-based maturity estimators of *M* initially. The estimate of natural mortality differ quite a bit from the GOA and the BSAI stocks (0.47 vs. 0.3ish). ***Likelihood profiles of M showed strong data conflicts – fishery length data indicated a value of 0.3 would be best, while survey ages, biomass, and the recruitment penalty suggested a value of 0.8 which is very unreasonable (but this could also be due to large fish moving out of the surveyed area – i.e., not seeing any large fish there which would result in M being estimated fairly high, while the fishery sees large fish, which is why you see that data conflict there; the conflict could also be due to the seasonal and spawning migrations as well as the fishery and survey operating in different times).*** Following this, a value of *M* = 0.4 was used instead to compromise between these data conflicts, which also matched some of the modes of *M* used in this assessment. 0.4 was considered a good starting point given estimates of 0.47 from GOA and 0.348 from the EBS assessment. The *M*-prior methods from Jason Cope suggested a value of 0.36. Given all of the above, *M* was set at 0.4 in this assessment.

Catchability estimates were estimated inside the model for the survey, but the fishery was set at 1. In general, the availability of cod and hence its catchability is fairly high for a majority of surveys (fairly high). Selectivity for both the fishery and survey were estimated as logistic age-based. ***One could make an argument for a more flexible selectivity form for the fishery given the combined fleet structure, although I realize its informed by primarily length-composition data. However, it might be worthwhile to look at length-based selectivity given the reliance on length-data here, although it may not work as expected given that Pacific cod can demonstrate dramatic fluctuations in size-at-age throughout the season and annually.*** The justification of similar selectivities between the survey and fishery is that the length distributions match up fairly well at those initial ages. Furthermore, looking at the data, dome-shaped selectivity does not appear to be warranted – likely because they do not reside in untrawlable habitats and larger cod do not appear to leave the region entirely.



However, based on the figure above, it seems like the survey catches individuals that are much smaller than the fishery. This could potentially be due to a mismatch between survey and fishery timing, and I do think there is some potentially for dome-shaped selectivity in the survey, simply given the seasonal distribution of Pacific cod.

##### Model Results

Regarding the use of diagnostics, they examine goodness of fits statistics and residuals for the biomass index and composition data. They also use likelihood profiles for *M* to examine potential data conflicts and also use retrospective diagnostics. However, given the model instability to data-weighting, I think an additional jitter analysis should be conducted to see how much results change with incremental changes in initial values. Additionally, while a profile of *M* was conducted, I think a profile of *q* for the survey would also be appropriate to understand the values governing this parameter and whether composition data conflict with this estimate. A conflict in *q* could also be due to mis-specified.

Due to the fixing of M = 0.4, which is a higher value than all the other models presented, it resulted in a much higher initial biomass and scale of the population, leading to an ABC that was 2x the model with M = 0.34. This is presumably due to the fact that a higher value of M requires the population to be larger to sustain the level of catches as well as the relative decline in the indices observed. Whether these quotas and the resulting estimate of M is realistic remains debatable.

Retrospective analysis indicated that SSB was consistently positive biased (Mohn’s rho = 0.15). They use Hurtado-Ferro’s paper to justify that the retrospective inconsistencies were not significant – constitutes best available science at that time. To conclude, the authors recommend this stock to use tier 5 designations for harvest specifications, but for the SSC to consider upgrading the assessment to tier 3 designation. A key problem in this assessment is due to a lack of survey data in recent years.

##### Summary on Problems in Assessment

Below is a summary of what I think is going wrong with the age-structured assessment at this point:

1. The estimate of *M* used does not have good justification. Furthermore, there seems to be data conflicts of what *M* should be based on likelihood profiles between fishery data and survey data,
2. The authors might consider separating out the fishery fleets, potentially due to differences in selectivities and changes in the ratio of catch taken by a given gear,
3. Given that the fleets operate at very distinct times, it may be worthwhile to disaggregate some of them (pot and longline = aggregated, trawl = separate),
4. Furthermore, while the authors suggest that there is no evidence for dome-shaped selectivity in the survey, this seems a bit dubious because of the life-history of Pacific cod (seasonal and spawning migrations), which can lead to dome-shaped selectivity (O’Boyle *et al.*, 2016). Looking at the age-and length-composition data for the survey and comparing them against the fishery, the survey selects more younger fish with a strong peak in those middle age-classes, while the fishery selects a more uniform spread of age-classes. This may suggest a potential for dome-shaped selectivity and should be explored,
5. There is no biomass index available for the new age-structured assessment, leading to increased uncertainty,
6. Additionally, the biomass-based index is not well fit in the assessment,
7. Despite the authors suggesting the retrospective bias is not significant, I tend to disagree. The retrospective pattern could potentially be a result of aggregating all the fleets, but treating them as time-invariant – there seems to be shifts in the catch ratio among fleets,
8. The uncertainty in ageing from the maturity curve is not propagated into the assessment,
9. Data-weighting methods were unsuccessful – further methods (e.g., Francis instead of McAllister and Ianelli can be explored),
10. Lastly, the interpretation of fishery performance and the use of CPUE in this assessment is incorrect – you cannot just aggregate all those gears with different units of effort and interpret fishery performance from that.
11. The estimate of survey catchability was at 0.83 and 0.7 corresponding to models with M = 0.34, and M = 0.4, which seems fairly reasonable given external studies and *a priori* knowledge of trawls and Pacific cod habitat requirements. However, a difference in 0.13 catchability is potentially large, and this estimate might be revised, by estimating M with a prior as opposed to fixing it (likely because of correlation with *M*).

##### Recommendations on Model Structure

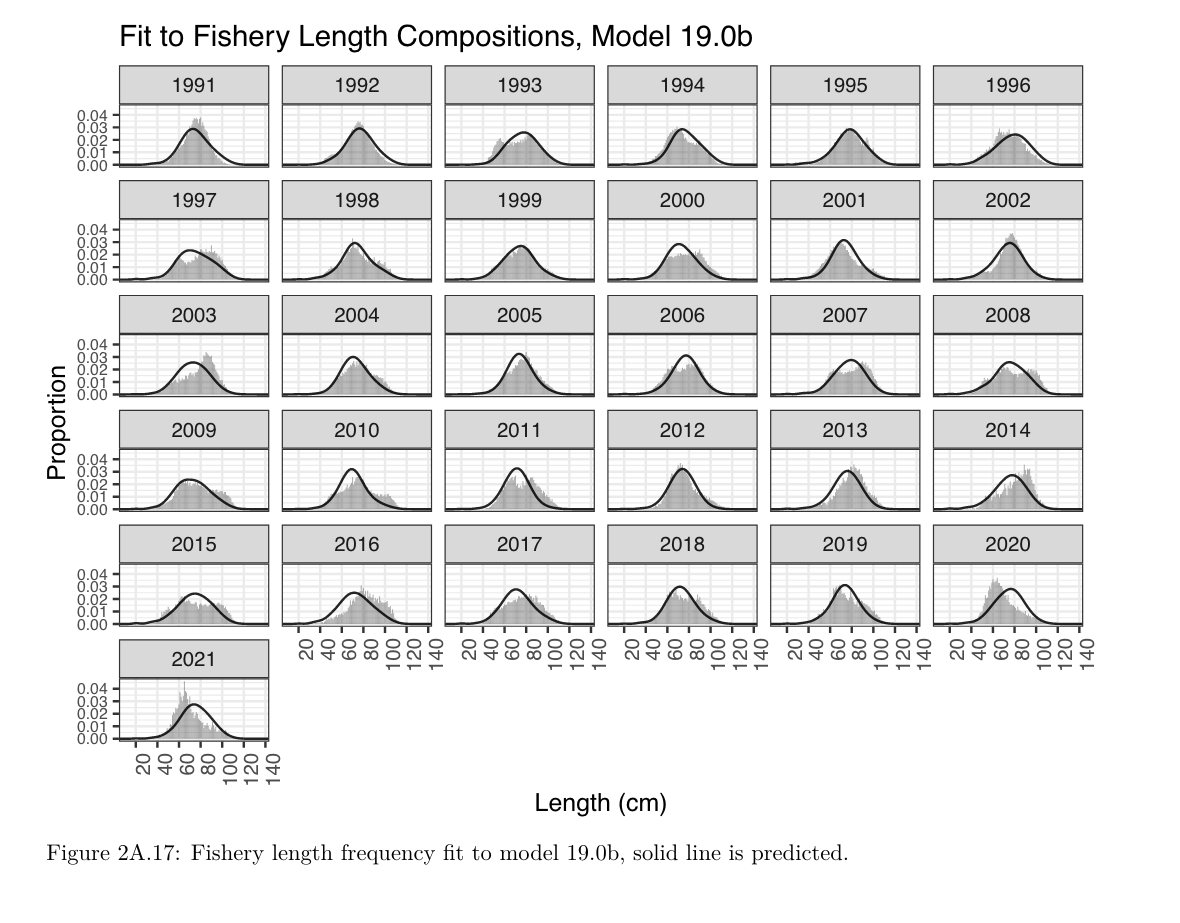
Given that the survey and fishery operate at very distinct times, I think the estimate of *M* might be different in terms of the likelihood profiles between these different data sources because the survey isn’t seeing as many old fish, while the fishery explicitly targets old fish. Furthermore, the difference in survey and fishery timing might resulting in some confounding in the estimate of *M* because of the seasonal migrations of feeding and spawning (i.e., large Pacific Cod move out of the area to feed), suggesting the potential need to construct a spatial model to examine potential movement rates, and better reconcile the estimates of *M*. However, looking at the likelihood profiles, it seems like the survey data suggest that the profile for *M* is really flat beyond 0.3, and the likelihoods do seem to dip at 0.3. I think 0.3ish is a good estimate, but definitely would be interesting to look at some of the spatial structuring in this respect to see if the descending limb of the likelihood profiles for these datasets bound back up. It is also quite surprising that fishery length composition data contain so much information about the estimate of *M*. Nonetheless, the limited information from the survey could/is most likely due to a mixture of factors confounding with each other: 1) fishery operates in the winter and catches larger fish, 2) the survey operates in the summer, capturing smaller fish, where there is limited information about older age classes and their probability of survival. Using a prior to estimate *M* would potentially be more useful and allows for the propagation of uncertainty, as opposed to fixing it *a priori.* It may also help resolve some of the large differences in estimated survey catchability.

A graph with different colored lines

Description automatically generated

Spatial structure aside, the fishery also appears to operate at very distinct times. In particular, the trawl fishery seems to primarily operate during the months of Feb – Mar, while the longline and pot fisheries operate year-round. Considering these seasonal movements, I think that the pot gear and hook-and-line gear should be combined into its own fleet, while modelling the trawl fishery as its separate fleet and following *pulse-*fishing dynamics potentially. This would be a fruitful avenue to explore. This is potentially supported by some of the bimodal length frequencies that are observed in these data and some misfits to fishery data (long shoulders could be because of combining data from gears with distinct selectivities).

A graph of different gear types

Description automatically generated with medium confidence

### 2022 AI Pacific Cod

#### SSC and Plan Team Comments

1. The SSB recommends changing the point estimate of *M* to 0.36 as in Jason Cope’s method. However, the assessment opted to estimate *M* inside the assessment and without a prior,
2. Maturity was requested to be fit within the assessment model, but was opted to be estimated outside the model in this cycle,
3. The SSC recommended using the LL survey in the age-structured models, where this was used in the assessment during this cycle,
4. The plan team recommended some data-weighting exercise and sigmaR being tuned, to which was employed in this assessment cycle (Francis and iterative tuning of sigmaR)

#### Assessment Structure

##### Data

In the 2022 AI Pacific Cod assessment, two age-structured assessments are brought forward ***(all data without LL survey and combined fishery, another with all data including LL survey but disaggregated fisheries)***. A tier 5 biomass-based index assessment is also brought forward. The age-structured assessment models uses the following datasets:

1. Fishery catch (1991 – 2022),
2. Fishery size compositions (1991 – 2021)
3. Age composition data from the fishery (2020 – 2021),
4. Biomass index from BTS survey (biennial and triennial, most recent = 2022; more during summer months, smaller fish),
5. Age composition from BTS survey (biennial and triennial, most recent = 2018),
6. Abundance index from LL Survey (biennial, most recent = 2022),
7. Size composition data from LL Survey (biennial, most recent = 2022)

##### Model Structure

As noted above, two age-structured models are presented. These are:

1. Model 22.0, where the fisheries are combined and the length frequency data are weighted by catch, season, gear, and area, and also uses the BTS Trawl Survey (164 parameters),
2. Model 22.1, where fisheries are disaggregated and uses both the BTS Trawl Survey and the LL Survey (274 parameters).

In these models, they are:

1. single sex, with growth estimated *within* the model,
2. an ageing-error matrix is incorporated,
3. fishing mortality, recruitment, selectivity, and catchability are estimated within the model,
4. natural mortality is also estimated *within* the model (presumably with a prior, although it is unclear whether or not a prior was utilized),
5. and the trawl survey has a prior of 1 for catchability. It appears that the standard deviation of the catchability prior is at 0.01, and thus is essentially fixed in the assessment. However, there is no good justification for the catchability prior with respect to the variance of the penalty/prior. Another approach that might be explored is to analytically solve for catchability,
6. additionally, the maturity curve is estimated outside the model using observer data,
7. these models were weighted using Francis-reweighting. However, no weights from Francis-reweighting are provided, which makes it impossible to know which data sources are being up or down-weighted.
8. For selectivity, all fisheries and surveys were fit using a double normal. For most fisheries and surveys, the double normal descending limb was fixed at its upper bound to resemble logistic selectivity. However, for the LL survey, the descending limb was estimated and was dome-shaped due to a lower proportion of older fish caught. Thus, the fisheries and BTS used logistic selectivity, while the LL survey used double normal selectivity.
9. In terms of the recruitment standard deviation, this was tuned iteratively, following methods of Thompson et al. (2008).

##### Model Results

The combined fleet model (22.0) fit the trawl survey biomass estimate much better than the disaggregated fleet model (22.1), but both models generally fell within the confidence intervals. I believe that there is model instability and confounding caused by the number of parameters estimated by the disaggregated fleet model (22.1), resulting in poor survey index fits as well as unreasonable estimates of catchability when estimated with the bounds of the prior relaxed. Nonetheless, the fits to the primary source of information about population trends and scale do not necessarily fit well.

A graph with numbers and lines

Description automatically generated

A graph of a number of people

Description automatically generated with medium confidence

Likelihood profiles were constructed and profiled across R0. In general, for model 22.0, the likelihood profile for the fishery agreed with the MLE, while for the survey, it showed higher likelihoods at larger values. For model 22.0, the estimate of R0 was in between the fishery and survey data. In terms of model 22.1, the trawl fishery suggested a much higher R0 value relative the longline fishery and the trawl survey. The pot fishery and LL survey showed an increasing likelihood of R0 as it increased, and no local minima were detected. The fact that likelihood profiles show conflicting results for R0 is likely to suggest that there is some mis-specification for selectivity. This is likely/potentially the case potentially attributed to the time-invariant selectivity estimated, although the use of the R0 profile has sometimes been shown to be inconsistent (Lee *et al.*, 2014). Additionally, it could be due to selectivity for some of the fisheries being dome-shaped rather than asymptotic, which can result in the scaling parameter/absolute abundance being conflicted between those data sources.

A screenshot of a graph

Description automatically generated

Both models estimated somewhat reasonable levels of natural mortality (0.35 and 0.36), which are consistent with other methods and assessments that estimate *M* for Pacific cod. However, it remains unclear whether a prior was used (most likely not), and whether use of a prior for natural mortality would help stabilize some of these results.

In terms of retrospective analysis, both models showed significant retrospective inconsistency when peeling back 10 years, which were about 0.25 to 0.31. The authors attribute this to a lack of survey data in 2020, it is also potentially likely that the model may be mis-specifying processes in which time-variation may occur. In particular, these include natural mortality (given extreme heatwaves), and time-varying selectivity, given the recent provision of harvest limits in the western AI and potentially shifts in fishing effort, as well as changes in catch ratios across gears. Overall, none of these retrospective values were within acceptable bounds of those defined by Hurtado-Ferro et al, 2015.

Overall, model 22.1 led to a poor fit to the trawl survey biomass relative to model 22.0. This misfit could be due to conflicting signals from the LL survey and the trawl survey given that it may be unrepresentative of AI Pacific cod population trends, resulting in model tension in attempting to fit to both data sources simultaneously. Additionally, likelihood profiles were particularly conflicted in this model, potentially due to certain data components suggesting larger stock sizes, which could be attributed to differences in fishery and survey timing. By contrast, model 22.0 performed reasonably with smooth likelihood profiles about R0, however, a fairly strong retrospective pattern was detected, likely attributed to a lack of a survey for multiple years, as well as some processes not being modelled (time-variation).

In terms of the time-series trends, the stock has experienced a continued decrease from 1992, although in recent years there are several indications of increases in the population. ***Total biomass estimates from 22.0 predict smaller increases, whereas the scale of increase from model 22.1 are much larger (2-fold of bottom trawl survey). Furthermore, model 22.0 shows less optimistic stock trajectories as model 22.1 (22.0 is at about B20, while 22.1 is a decent amount above that).*** Recruitment estimates also seem to be higher for 22.1, although both models follow similar trends. ***Interestingly, despite all the surveys showing downward trends, the population trajectories from the age-structured assessment models trend up.*** However, the assessment notes that the LL survey may not be representative because: 1) it only targets half of the Aleutian Islands, and 2) variable sampling due to gear loss and Pacific cod are a non-target species. Looking at some of the trends between the surveys, they track each other fairly nicely, despite the difference in spatial scale – although the LL survey shows signs of increase in 2015 while the BTS survey shows stability – which could potentially be due to differences in selectivity (asymptotic vs. domed). However, it is unclear what these catchabilities for the LL survey are estimated at from the document. Knowing if the catchability was much lower for the survey would provide indication of what is driving the difference in scale between these two models.

***Comparing model 22.0 and 22.1, the projected SSB is almost two-times more than for model 22.1 in the terminal period. The recommended harvest levels are also almost two-times more for model 22.1***. ***It is apparent that there is a scaling issue present in these models, given difficulties in estimating survey catchability as well as conflicts in the R0 parameters.*** ***In particular, there is almost a 20% difference in R0 between the two models when exponentiated onto a normal scale. However, it is unclear whether the scale issue is due to: 1) disaggregating the fleets, or 2) the use of the LL survey, and would be understood if the author used incremental changes instead of skipping certain developments of the model.***

The difference in scale also starts to begin in year 2009, wherein during this period, there was a lack of BTS data, and the LLS showed an increasing trend in its index. This is similarly seen in recruitment estimates, while model 22.1 had slightly larger recruitment estimates prior to 2009, they significantly diverged from 22.0 after 2009. As noted before, during this period, data from the BTS survey was lacking and the only available survey data that were available were from the LL survey. Furthermore, the LL survey index trends up and this tracks with when this divergence in biomass also begins trending up. ***This gives credence to the fact that the LLS might not be a representative index and is likely the main culprit behind driving the population biomass to higher levels, relative to model 22.0.*** Other issues are also plausible, such as the incorrect estimation of certain selectivity curves, given the number of data partitions done for a fleet-disaggregated approach. However, it would have been better to conduct a stepwise model building exercise, adding complexity and data sources incrementally to understand model behavior.

A graph of a graph showing the growth of a stock biomass

Description automatically generated with medium confidence

***A graph with red and blue lines

Description automatically generated***

Given sparse data available for the pot fleet, I do not think there are sufficient data to do a fully fleet-disaggregated approach, which may be leading to model instability. It would have been helpful to provide the standard deviations of selectivity parameters for these fleets and the potential exploration of a combined hook-and-line and pot fleet model would be warranted, where the trawl fleet is separate. Furthermore, looking at some of the selectivity profiles, the longline and pot fleet have similar selectivities, suggesting that it is likely appropriate to combine these two fleets together (blue and green).This approach is also likely appropriate because the pot fishery has consistently low catches and is fairly stable over time (i.e., fleet structure is not changing much). By contrast, the trawl fishery and the hook-and-line fishery fluctuate quite a bit in terms of their catch ratios. Assumptions about selectivity forms aside, this suggests that under a fleet-aggregated approach, selectivity is likely time-varying, which is problematic when considering that the selectivity for the trawl fishery is preferential towards older fish (i.e., selectivity is time-varying).

A graph of different colored lines

Description automatically generated

A graph of different colored lines

Description automatically generated

##### Summary on Problems in Assessment

Below is a summary of what I think is going wrong with the age-structured assessment at this point and what was also pointed out by the authors:

1. The biomass index from the BTS survey is not well fit for both models. This can be due the strong prior placed on the index, or incorrect selectivity assumptions on the survey(i.e., it could be dome-shaped not necessarily be logistic),
2. The BTS survey size-composition data does not appear to be fit well for model 22.1. Furthermore, no plots of model fits to CAAL data are provided for either models, making it hard to determine model adequacy,

A graph of a graph of a graph

Description automatically generated with medium confidence

1. In model 22.1, the LL survey is also not well fit, which is likely due to the survey being unrepresentative of Pacific cod dynamics,
2. A large retrospective pattern was detected in these models, which is potentially due to: 1) not modelling time-varying *M* during the heatwave period, and 2) selectivity for the fisheries are varying,
3. For trawl survey catchability, the model with multiple fleets (22.1, 274 parameters) led to a very low estimate of catchability when estimated freely, which is inconsistent with how trawls interact with cod. Thus, a strong prior with a mean of 1 and a sd of 0.01 was imposed. This might be happening due to the estimation of *M* within the model, which does not appear to be constrained by a prior. An analytical method for solving for trawl survey catchability could potentially shed light on the plausibility of *q* = 1.
4. It is unclear how well estimated the selectivities for model 22.1 are. Given the lack of composition data for some of these fleets and the reliance on only length composition data (all fleets rely solely on length composition data), some of these parameters may not be well estimated, and might be confounding with other parameters (*M* and *q*). Associated uncertainty for these parameters should be provided,
5. For the point above, give the lack of data for the pot fleet and the similarities in estimated selectivities between the two gears, these should potentially be combined into a fixed-gear fleet, while the trawl fleet should be separate given the season in which the gears operate in, as well as the relatively larger size of fish the trawl fleet captures,
6. While the assessment authors suggest that dome-shaped selectivity is unreasonable in the BTS survey, the size-compositions suggest (and the lack of fit for model 22.1, also compare the average composition data between trawl survey and trawl fishery) that dome-shaped selectivity options should be explored, especially when considering the life-history of the species (seasonal migrations can lead to availability differences resulting in dome-shaped patterns),
7. Lastly, the likelihood profiles for model 22.1 suggested substantial data conflicts, with the trawl fishery suggesting a much higher scale compared to all other data sources. This could be due to a variety of factors, including mis-specified selectivities, natural mortality (i.e., might be time-varying), and just general model stability and sparse data for some fisheries.

##### Recommendations on Model Structure

Firstly, it would be incredibly beneficial to have an exercise that sequentially adds new data into the model, to understand the influence of particular data components, surveys and fisheries. In particular, it would be helpful to understand the influence of the LLS, in the absence of a fleet-disaggregated model. ***Next, it somewhat worries me and seems dubious that M is so well estimated with low uncertainty even without a prior imposed on it.*** Given this point, I was surprised to see no attempt some form of a jitter analysis to see how much *M* changes as you change the starting values. It seems dubious that length-composition data are so influential in determining these natural mortality estimates and a likelihood profile on *M* should be done (based on last year’s model – 2021 assessment). My next point is in regard to the strong prior placed on survey catchability – which is essentially fixed at 1. It is possible that this strong prior placed on catchability might result in a fairly well estimated *M*, given these two factors are confounded, and might be interplaying with some of the lack of fit to the indices. I believe that some attempt to estimate *M* with a prior instead of freely estimated should be explored. In addition, it is unclear from the assessment document what the value of catchability for the LL survey is and how it is parameterized. If it still has a strong prior of 1, that seems unreasonable for a LL survey, especially given that the survey does not directly target Pacific cod and because it is a LL survey, where the units of catchability do not necessarily translate on a scale of 0 to 1. Clarification needs to be provided here. The assessment notes that dome-shaped selectivity is not reasonable for the BTS survey. However, given these seasonal movements, I would actually argue that it might be worthwhile to attempt and explore (O’Boyle *et al.*, 2016).

Another key issue with the use of the LL survey is that it only covers half of the area occupied by the stock, which may be a driver of the lack of fit for the index. A potential solution could be to model a single aggregate survey fleet by standardizing the indices between the BTS and LL survey, while modelling and controlling for spatial-temporal effects. Although there is little literature on this specific topic, it may be a fruitful avenue to explore if we want to incorporate all possible data sources into the analysis. However, I recognize that there might be some difficulties in implementing such an approach given the large spatial imbalance, the need to aggregate composition data.

***With respect to fleet structure in model 22.1, given the lack of data for the pot fishery and similarities between pot and longline gears (selectivity and seasons), exploration should be conducted to combine these two gears, while modelling the trawl fishery as a separate fleet.*** ***Perhaps the longline and pot gear fleets should incorporate some form of dome-shaped selectivity as it appears that the larger age/size-classes are less selected relative to the trawl fishery.*** Furthermore, if a fleet-aggregated approach is clear, attempts should be made to allow for time-varying selectivity – this seems likely give that the selectivity among the three primary fleets differ, and the catch-ratios between the hook-and-line and trawl fleet fluctuate quite a bit.

My next point is in regard to the use of conditional-age-at-length data, despite this method generally thought to be reserved for data-rich assessments, where this stock in my opinion, should be considered relatively data poor to data moderate. Explorations should be conducted into fitting marginal age-compositions and length-compositions instead of conditional-age-at-length, given that the use of CAAL can lead to biases in growth and management quantities if these data are not representative of the age-structure of the population – no model fits nor justification was provided in terms of the use of these data (Lee *et al.*, 2019), although I will admit I am not an expert on CAAL; I just know that there are some nuances for these data to be of utility (Lee *et al.*, 2024). In particular, there are some issues pertaining to unmodelled processes and using CAAL data, where misspecification of growth variation, natural mortality, or the observation model can distort the CAAL and lead to estimates of incorrect population age structure.

### 2023 AI Pacific Cod

#### SSC and Plan Team Comments

1. The SSC recommends getting more data from the pot fleet to support a fleet-disaggregated approach, given that data from the pot fleet are sparse, but it makes up a good chunk of the harvest,
2. If fleet-disaggregated approaches are pursued, they recommend attempting dome-shaped selectivity for the HAL fleet. However, we could potentially also combine the HAL and pot fleets, and model that with dome-shaped selectivity, while maintaining logistic selectivity for the trawl fleet,
3. If the fleet-aggregated approach is pursued, they recommend using time-varying fishery selectivity to address retrospective behavior,
4. The SSC recommends a hybrid approach in setting TACs, where the M estimate is from the age-structured model, but the TAC setting uses Tier 5 methods,
5. The Plan Team also recommends bringing forward the growth varying model, as well as a growth + time-varying selectivity model, however, three alternative models were brought forward instead (detailed below).

#### Assessment Structure

##### Data

In the 2023 AI Pacific Cod assessment, three age-structured assessments were brought forward. The general model structure for this assessment combines fisheries which uses fishery length composition data and catch data, while only one survey fleet is used (BTS survey with CAAL data). ***The three assessment models are:***

1. ***three growth time-blocks to account for shifts in growth and warmer temperatures, which were determined using Kapur et al. (2020) and another defined a priori,***
2. ***selectivity time-blocks (n = 5) to accommodate shifts in fishery targeting practices, but there was no justification for when and why,***
3. ***two time-blocks on growth and two time-blocks on natural mortality to accommodate warmer temperatures.***
4. A tier 5 biomass-based index assessment is also brought forward.

The age-structured assessment models use the following datasets:

1. Catches from the combined fishery (1991 – 2023),
2. Abundance indices from the BTS survey (1991 – 2022),
3. Length composition data from the combined fishery (1991 – 2022),
4. Length composition data from the BTS survey (1991 – 2022),
5. CAAL data from the BTS survey (1991 – 2022)

Note that in the current assessment year, no trawl survey was conducted. To summarize, the model uses CAAL data as well as independent length composition samples for the survey, length composition from the fishery, an index from the survey, and catch data from the fishery. A fair number of ages are sampled from the survey, almost surpassing sablefish in some cases.

##### Model Structure

As noted above, three age-structured models are presented. These are:

1. Model 23.0, where three growth time-blocks (blocks are estimated based on *k* parameter) are incorporated due to a documented change in growth using Maia’s method for detecting shifts in growth,
2. Model 23.1, five fishery selectivity time-blocks are incorporated on the ascending width of a double-normal in addition toe three growth blocks,
3. Model 23.2, two time-blocks on natural mortality and two growth time-blocks due to changes in regimes (marine heatwave).

In these models, they are:

1. single sex,
2. with growth estimated *within* the model (model 23.0 and 23.1 has 3 time blocks on growth, model 23.2 has 2 time blocks on growth and natural mortality),
3. an ageing-error matrix is incorporated,
4. fishing mortality, recruitment, selectivity, and catchability are estimated within the model (model 23.1 has 5 time blocks on selectivity),
5. natural mortality is also estimated *within* the model (***without a prior***; model 23.3 has 2 time blocks on *M*),
6. and the trawl survey has a prior of 0 for catchability in log space. It appears that the standard deviation of the catchability prior is at 0.01, and thus is essentially fixed in the assessment.
7. Additionally, the maturity curve is estimated outside the model using observer data,
8. these models were weighted using Francis-reweighting. Furthermore, input sample sizes were developed using a bootstrap method to reflect annual variability in sampling using Pete and Ben’s method.

For selectivity, all fisheries and surveys were fit using a double normal. A single combined fishery fleet was modeled and it its generally appropriate how the length composition were weighted and expanded and used in the model. Similar to the previous assessment, Francis-reweighting was used. However, the authors use AIC as a metric to compared models, which is invalidated by Francis-reweighting because of the different likelihood weights.

For growth, which was examined outside the assessment initially, a shift in growth in 2002 was detected. The author indicates that a shift in growth occurred in two time-blocks, but does not provide data or figures to support this (i.e., the first derivative plots in Maia’s paper). Furthermore, the authors indicate that there was a shift in growth by using length-frequency data pre and post 2004 (2004 because lack of survey data in 2003), and shows that there were differences in the length frequencies. ***However, this could have entirely been due to a change in the age-structure of the population and not necessarily attributed to growth.*** ***For example, M could have changed during this period, resulting in less old and larger fish observed***.

For models 23.0 and 23.1, an additional growth block in 2017 was included (i.e., pre-2003, 2003 – 2016, and 2017 blocks) to attempt to better fit a decline in survey biomass attributed to changes in temperature regimes, inducing slower growth. However, declines in survey biomass could be due to below average recruitment, increased mortality, and a variety of other factors; not necessarily because of a change in growth. Furthermore, it does not appear that either of these model fit the BTS biomass estimate that well, particularly in the later period. This later time block while implemented, was not supported by a data-driven method, and I don’t see why it was attempted here. A graph with numbers and lines

Description automatically generated

They justify this by stating that it was attempted because it improved fits and retrospective performance. On the latter part of this statement, I disagree that it improved retrospective behavior. However, retrospective behavior does not appear to be much improved, and has a weird dip, likely due to the time blocking structure imposed, that was not adequately dealt with.

A graph with colorful lines and numbers

Description automatically generated

Natural mortality was suggested to change in 2015 due to a thermal lag from marine heatwaves in 2013-2014, which appears appropriate and justified from other Pacific cod assessments. Incorporating a change in natural mortality appeared to allow for much better fits to the survey biomass index as well as much improved retrospective behavior, which seems appropriate. As such, incorporating time-varying M adjusts the models expectations to explain recent declines in BTS biomass, attributed to the heatwave period.

A graph with numbers and lines

Description automatically generated

For BTS survey catchability, two options are presented in the document. In general, catchability is estimated in log space with a strong prior of 0 and a small standard deviation of 0.01. The first approach used a prior to estimate catchability, while the second approach was a sensitivity test to examine the influence of the prior, wherein catchability was analytically calculated. The author posits that small differences were detected between these two options, however, the virgin SSB and natural mortality were fairly different. This might be due to the inherent correlation in *M* and *q*, and the removal of the prior led to this difference (prior is very informative).

While the author indicates that selectivity for model 23.1 was estimated with five time-blocks, no justification was given as to why these blocks were used. Lastly, natural mortality (model 23.2) was freely estimated as two time-blocks in the model. ***Note that in this assessment, time-blocked parameters for growth and natural mortality have an extra estimated parameter. In particular, if two time-blocks were implemented, three parameters were estimated. The correct and more appropriate method would be to estimate two separate parameters, instead of three. It appears that the current parameterization uses a mean parameter and estimates additional deviations about this mean.***

##### Model Results

In general, all three models fit the survey biomass and length frequencies well. However, it appears that model 23.2 fits the biomass index the best. ***The author uses AIC to compare models and says that it is appropriate given that the same datasets are used. However, this is incorrect because Francis-reweighting is used to weight the likelihoods, and thus, your likelihood components are going to be different, and AIC is not a valid tool to use in this respect***. However, if Francis-reweighting is not used (it doesn’t really seem like it and no weights are provided), then this approach is appropriate.

Likelihood profiles about R0 did not seem to indicate anything jarring and was a compromise between fishery and survey data.

A graph of different colored lines

Description automatically generated

***Retrospectives were best for model 23.2, however, there are several issues with the retrospective analysis, especially when time-blocks are used.*** In particular, the retrospective peels prior to the implementation of time-blocks are not comparable as they constitute entirely different models. Nonetheless, I still believe that it provides insight with respect to how the assessment would have performed if those parametrizations were used during that particular period. For all 3 age-structured models, the retrospective patterns were within reasonable bounds. Furthermore, I believe that some of the sharp drops observed in the retrospective patterns are potentially attributed to blocks being implemented in a terminal year, where there is insufficient data to inform a parameter estimate (i.e., model parameters are not identifiable given insufficient information).

At present, the stock does not appear to be in good condition, which could be due to adverse ecosystem conditions and impaired recruitment/survival. SSB has declined to about B20%, which would indicate a fishery closure is likely necessary. ***This is one of the detriments of using Tier 5 methods, is that it is not possible to tell whether a stock is in overfished condition. Perhaps an alternative would be to use the start of the time-series and assume that it is in unfished condition, take 20% of that value and use it as our limit reference point, while applying potentially a knife-edged or threshold control rule.***

Lastly, the assessment did their harvest projections using the most recent growth block but average natural mortality. I believe they presented more projections in their slides, but these are not present in the document. Nonetheless, ***none*** of the projections presented indicated that the stock was subject to overfishing or was in an overfished condition. ***However, note that there is potential for the stock to actually be below B20% based on the confidence intervals, and there is concern for fishery closures there.*** Based on ecosystem considerations, it appears that the stock may be experiencing better conditions in the recent period (lower summer temps (but still above average), higher winter temps, more Pacific cod in seabird diet data), but signals are variable based on other indicators (temps that are still above average indicate higher bioenergetic demand, etc, lower sea lion abundance).

##### Summary on Problems in Assessment

Several problems are identified in this assessment:

1. Unnecessary parameters are estimated for growth and natural mortality time blocks. Furthermore, there is not enough sufficient justification for some of the time-blocks for the fishery as well as some of the growth time blocks,
2. The growth time-block parameters are estimated very similarly. There is no reason why these should be treated as separate parameters. Additionally, a figure of length-frequencies is presented to justify a growth block. I do not see how that is appropriate given that this could be due to a shift in age-structure, not necessarily growth. However, it does potentially help justify a shift in *M*, given the abundance of younger fish observed, relative to older fish (i.e., individuals are dying more quickly and are unable to reach older ages),
3. Additionally, the use of such an informative prior on *q* is somewhat concerning. I think that the reason the analytical *q* method was so different in terms of *M* and SSB estimated was because of the strong prior imposed on *q*, and the strong correlation between the two parameters,
4. Next, the retrospective patterns are not necessarily comparable within a given model, considering the use of time-blocks changes the model structure, and does not result in a apples-to-apples comparison, although I do believe that it provides insights as to how the assessment would have performed given its retrospective assessment structure. Furthermore, some of the sharp declines observed in the retrospective analysis could potentially be due to the use of time blocks during the terminal year of an assessment, where these parameters are not identifiable with the available data,
5. Additionally, the use of AIC to compare models in this assessment is invalid. This is because the likelihood weights applied are different when Francis-reweighting is used. Although the document indicates that Francis-reweighting is used, no likelihood weights are provided and the likelihood components are very similar. Clarification is needed there,
6. There does appear to be a shift in selectivity within the fishery although figures of these blocks are not provided. Given the dynamic nature of fleet structure in this stock, time-varying selectivity really should be considered in conjunction with mortality. Furthermore, given some of the size-compositions relative to the trawl fishery for longline and pot, I think it would be prudent to investigate dome-shaped selectivity for the fixed-gear fleet***. I have the same sentiment for the bottom trawl survey, where dome-shaped selectivity might be considered, which may help explain the need for such a strong prior on q and previous unreasonable estimates (model 19.0b in 2021, M = 0.4, with observer maturity curve) of q when it was freely estimated. It has been documented that mis-specified selectivity forms can lead to unreasonable estimates of catchability, given that these two variables are related (Cadrin et al., 2016).***

##### Recommendations on Model Structure

Some recommendations I have on model structure are detailed below:

1. In general, I think model 23.2 is a good starting point for next years assessment, although I would make the following changes – remove the time-block on growth given the similarities in kappa. I think the *M* time-block is well justified in this case, and does appear to help model performance significantly. Furthermore, I would remove the estimation of the unnecessary parameters for these time blocks,
2. I would consider aggregating the longline and pot fleets, while separating out the trawl fleet for fleet structure. Given the pot fleet has limited data, this approach would allow us to increase the information content available to use a semi-fleet disaggregated approach. This is also most likely appropriate because the longline and pot fleets operate at very similar times of the year and previous years assessments where a fleet-specific model was employed, has shown that longline and pot selectivity are estimated fairly similar (composition data also indicate this),
3. In a similar vein, investigations should be attempted to allow for dome-shaped selectivity in the fixed-gears because of differences in size-composition relative to the trawl fleet (see Figure 10 in the 2022 assessment),
4. Similarly in the survey, there does seem to be potential for dome-shaped selectivity especially when we consider that seasonal movements can lead to dome-shaped selectivity patterns – domed selectivity can manifest because of availability differences (Cadrin *et al.*, 2016; O’Boyle *et al.*, 2016). The reason that such a strong catchability prior is needed may be due to a mis-specified survey selectivity form and may help explain why models brought forward in 2021 estimated lower catchabilities for the survey. However, I do note that the length-composition data are generally well-fit, although the average length compositions for the survey towards those older size-bins are descending much faster than the fishery. Fits to CAAL data are not provided and ghost fits to age-composition are not provided, making it difficult to evaluate the adequacy of logistic selectivity for the survey (best practice is to use OSA residuals, so I refrained from interpreting the Pearson residual plots),
5. Further expanding on point 4, while the authors indicate that there is not direct evidence for dome-shaped selectivity in the survey using Weinberg et al. 2016, this study investigate such dynamics by increasing vessel tow speed and escape from nets. However, this study did not look into the temporal aspect of availability, especially considering the seasonal migrations Pacific cod undertake (i.e., stock assessment selectivity is both contact and availability),
6. Next, I recommend imposing priors on estimating natural mortality. It seems dubious that it is so well estimated especially given that composition data are not directly available after a small amount of exploitation from the fishery – which has been hypothesized to allow for adequate estimation of natural mortality (Lee *et al.*, 2011). The reason that *M* is so well estimated could be due to the strong prior imposed on catchability, and might be supported by the fact that the estimates for *M* change a decent bit when *q* is analytically determined,
7. I further recommend likelihood profiles on *M* to understand how it changes and how different data sources inform this estimate – it could be that the likelihood profile is super shallow and that the estimate of *M* is actually stuck at a local-minima. A jitter analysis would be incredibly helpful.

## Harvest Control Rules

## 

Harvest control rules (HCRs) represent a set of pre-defined rules and agreements in which tactics to regulate the fishery are set based upon the available data (i.e., stock status), and are aimed to achieve a set of explicit fishery management objectives and trade-offs among these objectives. The rules governing these tactics can increase, decrease, or keep constant the amount of catch, effort, or fishing mortality allowed in a given fishery, and allow the prescription of advice even when our present state of nature deviates from our targets. A key benefit of the development of an HCR is that it shifts the reliance away from developing complex models to set catch advice towards a decision-making process that better addresses trade-offs in conservation, maximizing catch, and minimizing catch variability, that are not easily captured within complex models. Thus, it shifts from model-based decisions towards rule-based decisions that are more transparent, and further anchors decision-makers to a pre-defined set of rules and mitigate the influence of politics in this process. Often, HCRs take on different forms to address various fishery management objectives, while balancing the complexity to ensure catch limits are easily conveyed to stakeholders clearly. While the use of HCRs is not a panacea for fisheries management, it does make setting fishery limits (catch, F, effort, etc.) more explicit and relieves political tensions and influence on the decision-making process, given that decision-makers are confined to make decisions within a constrained set of rules.

Most HCRs contain several reference points. These include: 1) limit reference points, which represent limits that should never be reached, and if they are approached, drastic management actions should be undertaken (i.e., fishery closures), 2) threshold reference points, which represent reference points that serve as early warning signs to reduce the probability of exceeding a target or limit reference point due to observation or management uncertainty and management measures are adjusted when a fishery in within this threshold to ensure that the probability of reaching a limit or surpassing a target reference point is reduced with increasing uncertainty – these reference points are commonly used in systems that are high variable, species at the edge of their geographic ranges, or are unresilient (there is misuse of this term, as it can refer to limit reference points, however, it can also refer to buffers between OFL and ABC, as well as maximum fishing mortality and minimum stock size thresholds), and 3) target reference points, which represent reference points that are desirable and should be maintained over the long term. For stocks where stock-recruitment relationships are estimable, some common limit reference points include ½ of Bmsy, arithmetic mean of Fmsy, while target reference points for these stocks are Bmsy and the harmonic mean of Fmsy. For stocks where such relationships are not estimable, SPR reference points are commonly used, where limit reference points can include ½ of B35% and F35%, while target reference points represent B40% and F40%. For stocks where there are limited data, alternative reference points can include the use of *M* as a limit reference point, with x*M* representing the target reference point, where x is a scalar that is less than 1.

In general, HCRs are constructed to meet several common fishery objectives, which include: 1) maximizing catches over some time horizon, 2) minimizing the variability in catches, 3) avoidance of extreme harvest scenarios, whether this be due to fishery closures or rapid increases in catch, 4) the duration for which a stock remains below some undesirable level (i.e., time needed for the stock to be rebuilt from a low level), 5) maximizing profits, and 6) ensuring adequate prey base for predators (e.g., pollock and Pacific cod). A key challenge in accommodating all of these objectives into a HCR, is that these are mostly competing objectives, and often the challenge is finding a trade-off among all of these objectives. From a stakeholder perspective, it is ideal for HCRs to both maximize catches in the long term, while maintaining minimal variability in catches, while from a conservation perspective, it is ideal to maintain the stock size above some pre-defined level/limit in order to avoid a high probability of stock collapse and recruitment overfishing. Furthermore, from a stakeholder perspective, it is ideal that HCRs are not complex and are transparent, where mechanisms underlying changes in annual catch limits can be easily conveyed through the HCR.

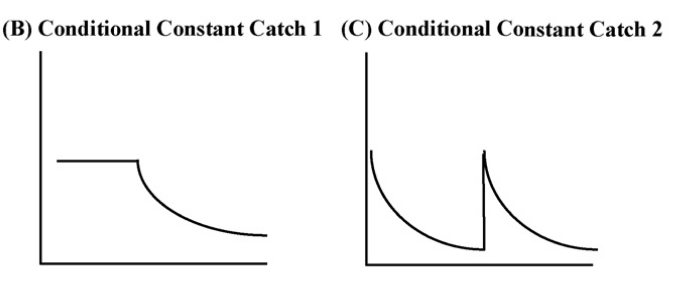
There are several forms of HCRs, which range from less complex to more complex, each with its own individual trade-offs and aim to address differing fishery objectives. HCRs typically come in the form of 1) constant catch, 2) constant harvest rate (or F), 3) constant escapement, 4) threshold control rules, and 5) empirical control rules, where harvest is set based on some biomass index (also high-density areas). As discussed above, the form and shape of the HCR will depend on the objectives of the fishery, but also the life-history of the species. Furthermore, there are generally trade-offs with each of these HCRs, whether it be in the form of increased fishery closures, increased variability in catch, increased rebuilding time, stock collapse, or reduced yield. As noted above, each HCR form serves in own purpose and address trade-offs different. Below, I’ll list off some benefits and trade-offs of each HCR, the different objectives they meet and don’t meet.

**Constant Catch:** The constant catch rule is the simplest and represents a data-poor/limited method, where catch limits are constant, irrespective of the stock size. These limits can be set by multiplying the average historical catch by some proportion, or by setting the catch limit at MSY. There are several benefits to the constant catch rule, which includes: 1) there isn’t a need for regular stock assessments, because catch limits are constant, 2) it provides stability in catches, because it is constant and allows for better planning in the future. However, it performs poorly in the following aspects: 1) catches and profits are seldom maximized because as stock sizes increase (foregone yield), your catch limits still remain the same, 2) in a similar vein, there is a risk of driving the stock to extinction because constant catch rules exhibit depensatory mechanisms, and exert a disproportionately high F rate on the stock, at low stock sizes, and will result in recruitment overfishing, 3) in a stochastic environment, even if catch levels are set at the MSY level, there is potential for driving the stock to extinction if the stochasticity is autocorrelated. Lastly, the constant catch rule only has a limit reference point (the catch limit).

**A line graph with a line and a line

Description automatically generated with medium confidence**

**Conditional Constant Catch:** Conditional constant catch (CCC) rules are similar to that of constant catches but include some alterations in the rule to allow for adjustments depending on stock size (i.e., accounts for the problem of disproportionate harvest rates at low stock sizes). In the case of conditional constant catch, constant catch is taken unless if removing that amount would exceed some limit reference point. If it does exceed this threshold reference point, then it reverts to a constant F strategy. This control rule has the benefit of maintaining: 1) stable catches across the time period, and 2) avoids high fishing mortality rates at low stock sizes. Thus, this control rule minimizes the risk of over-exploitation. However, the CCC rule suffers the same detriment as constant catch, where there is potential for foregone yield.



**Constant F/Harvest Rate:** Constant F or harvest rate rules essentially maintain the same level of F or harvest rate irrespective of stock size. However, given that this is a harvest rate, catches are adjusted in accordance to stock size, where catches are in proportion to the present stock sizes, and thus catches are adjusted downward or upward accordingly. This is a relatively simple rule, where a constant harvest rate or catch is set. For data-limited stocks, proxies for the limit F reference point could use some proportion of an estimate of *M*. In general, constant F rules are useful for: 1) reducing variability in catch because you are always able to fish, 2) maximizing catch, because you are not only able to fish at low stock sizes but are also able to increase harvest at high stock sizes. However, it does poorly with respect to minimizing the risk of overexploitation and increase rebuilding times, because catch is still allowed even when the stock biomass is at low levels. Furthermore, constant F/harvest rate rules can be problematic especially if there are errors in the biomass estimates, because higher but incorrect levels of removals in previous years cannot be adjusted in subsequent years and instead remain constant.

**A graph with a line

Description automatically generated with medium confidence**

**Constant Escapement:** Constant escapement rules are typically used for salmonid stocks. In general, this is a special case of a threshold control rule, where no catch is allowed if you are below your escapement goal (i.e., Smsy from a stock recruit analysis, which is the limit reference point), and catches are allowed to increase infinitely when the escapement limit/goal is met. In general, this rule allows: 1) maximizing profits because you can harvest a bunch at high stock sizes, and 2) it minimizes overexploitation because there is a limit reference point that prevents fishing. However, there are some detriments to this rule, namely: 1) higher variability in yields because of risk of fishery closures, as well as the ability to harvest at an infinite rate at high stock sizes, 2) it is highly susceptible to issues relating to autocorrelated recruitment, because the constant escapement goal is predicated on Smsy (can change depending on the autocorrelated nature of the parameter), which requires parameter estimates from a stock recruitment relationship and can lead to unnecessary fishery closures, 3) assumes that profits are linearly related to harvest, which it is not (i.e., allowing harvest rate to approach infinity at high stock sizes), and 4) requires an identification of the origin of the stock and management on an independent stock level, which can be complicated in a mixed-stock fishery. ***There appear to be some studies proposing that constant F rules outperform constant escapement rules, likely because they are less susceptible to autocorrelated errors and prevent unnecessary fishery closures (constant escapement rules are highly sensitive to the estimate level of maximum recruitment and when it is autocorrelated, it can exert imperfect control on spawning biomass and expected recruitment),*** and do not allow harvest rates to increase infinitely, which minimizes variation in yields as well, and can produce similar levels of catch.

A graph of a mass or abundance

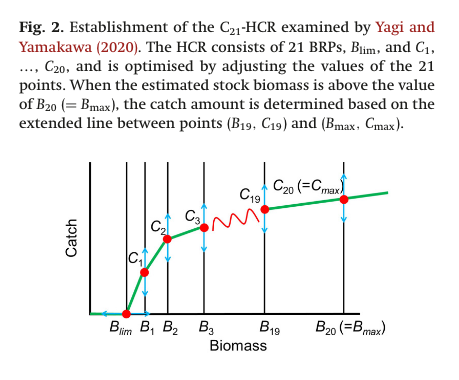
Description automatically generated

**Threshold:** There are a variety of ways to parameterize a threshold control rule. These control rules involve the use of a biomass limit reference point, where if biomass levels are below this limit, all fishing is halted. Furthermore, it also involves a target reference point for both fishing mortality and biomass, where the target is to maintain B/Bmsy and the target F is the F rate that produces Bmsy. These are often associated with more data-rich methods, where these reference points are derived via model-based methods. Some benefits of this method are that: 1) it can maintain a good level of catch (allows for more catch at high stock sizes), 2) it prevents overexploitation, 3) reduces the amount of time needed for rebuilding, 4) is relatively stable in terms of the levels of catches it prescribes, 5) it is not knife-edged generally, which allows for some buffer of estimation error and allows for some level of fishing even if you over or underestimate your biomass limits, and thus reduces catch variability in this respect, as well as buffers and aids in helping with recovery time, and 6) it reduces F with decreasing stock size, which likely maintains the biomass at a reasonable sustainable level, as opposed to driving it to depletion. However, there are some trade-offs with this HCR, which include: 1) difficult to parameterize given the need for limit reference points and target reference points, 2) there is some potential for fishery closures which may increase variability in yield, 3) under a scenario with a stock-recruit relationship, the benefits of maintaining biomass at some level do not really materialize, because there is not compensatory or depensatory mechanism, and it may be better to just harvest at a constant F rate.

A diagram of a graph

Description automatically generated

**Flexible:** Lastly, there are some flexible reference points that have solutions that yield a Pareto-optimal front and simultaneously address issues pertaining to variability in catches, maximizing catches over the long term, and a risk-averse function (i.e., risk-averse in the context of minimum biomass limits) (Yagi and Yamakawa, 2023), and can contain a variety of reference points within the HCR. While this flexible HCR can be used to meet various fishery stakeholder objectives simultaneously, they are quite complex to construct (requires numerical optimization to optimize various utility functions) and from a stakeholder perspective, it can be difficult to convey how the HCR works, and why certain catches were prescribed in a given year. However, they do allow for weights applied to the utility functions which increased stakeholder involvement in the management process, which can aid in the communication aspect of this HCR. Furthermore, they also allow for time-varying capabilities and are easily generalizable to the current paradigms of assessment because they include Bmsy, F, and catch reference points. Generally, these HCRs indicate that under no observation error, it is best to do a constant escapement goal, but as uncertainty increases, it requires a dampening of the linear increase in constant escapement to adjust for uncertainty. Furthermore, to maintain a higher level of biomass, the HCR needs to be right shifted, to maintain a higher biomass limit reference point, while to have lower variability in catches, the HCR needs to have a lower slope at high biomass levels (avoid harvest a ton at high stock sizes).



A group of graphs and diagrams

Description automatically generated

There are several factors that can influence the performance of HCRs, which can range from non-stationarity, autocorrelated recruitment errors, the presence of a stock-recruitment relationship, uncertainty and errors in estimates of stock status, implementation and management error, and the life-history of the species.

With respect to non-stationarity in management reference points and in the context of threshold control rules, there can be unintended consequences, where shifting reference points towards lower productivity states can result in high fishing mortality rates. In particular, studies have illustrated using a surplus production model where the carrying capacity of the stock (and hence MSY) decreases, the use of a threshold control rule will shift the both the carrying capacity and MSY reference point to the left, and result in more favorable estimates of stock status and allow for higher fishing than would be allowed if assuming the status-quo (Szuwalski *et al.*, 2023). There still lacks consensus on how non-stationarity should be dealt with within the management framework, and irrespective of how an HCR is configured, non-stationarity in reference points due to selectivity, natural mortality, recruitment, etc, will likely result in more variable yield and could perform poorly if there are large errors in these non-stationary processes. On a related note, non-stationarity can be accounted for in a HCR by including an environmental index, as is done in Pacific sardine. However, these relationships can break down (as has been shown in the sardine stock) and can similarly have unintended consequences for harvest limits. Under climate change scenarios and non-stationarity, resiliency can be built into HCRs and management structures by 1) implementing ecosystem caps, as is done in Alaska, 2) replacement of constant F rules with threshold rules to reduce rebuilding times, 3) develop empirical HCRs to back up model-based HCRs in cases where model-based methods fail, 5) better incorporate uncertainty buffers into the management process and resultant HCRs.

Another key factor that contributes to the performance of HCRs is the presence of autocorrelated recruitment errors. In particular, the performance of constant F rules outperforms a constant escapement rule, when the asymptotic parameter of a Beverton-Holt is autocorrelated because that parameter determines the maximum recruitment of the stock, and hence the optimal escapement reference point. As such, when there is autocorrelation, there can be imperfect control on levels of spawning biomass and realized recruitment, which can result in constant escapement rules performing poorly (this is assuming that low and high productivity states are not accounted for, and constant escapement is assumed, i.e., fishery closures when there should not be, and openings when they should not be). By contrast, a constant F rule is not dependent on these estimates of escapement, and thus, harvest is allowed irrespective of the escapement goal (and the fishery does not close) and is adjusted in proportion to the level of stock size so as to avoid exerting high F at low stock sizes.

In the presence of a stock recruitment relationship, some of these HCRs can similarly perform differently as well. In particular, in the extreme case where a stock-recruitment relationship does not exist (i.e., constant mean recruitment), there is no benefit to maintain a stock size at Bmsy or B40%, because recruitment is independent of these values. As such, it may be better to just harvest at a constant rate under such circumstances. Thus, without a stock recruitment relationship, yield can potentially be more variable due to the potential for fishery closures, despite the lack of a stock-recruitment relationship. This is similar in the context of a constant escapement strategy in the absence of a stock recruitment relationship, where it may be more beneficial to allow for a constant F, because it avoids unnecessary fishery closures (and hence reduced variability in yield), and allowing for more escapement does not produce more recruits.

In a similar vein, the life history of a species can also impact the performance of a specific HCR. In particular, constant F rules can perform poorly when a species has low steepness and is not resilient, because you allow harvest at low stock sizes, when recruitment is potentially severely impaired. By contrast, a threshold control rule for stocks with low steepness can be of utility, given that it prevents fishing at low stock sizes, and also ramps down fishing mortality as stock status is < 1, which buffers against some degree of uncertainty, and may potentially promote faster rebuilding times.

In addition to the form of the HCR, different limit and target reference points in the context of a threshold rule can also impact their relative performance in managing a fishery and avoiding overexploitation. In particular Punt et al (2010) showed that for a stock with high steepness values, their limit reference point can be much lower than those typically defined by management (e.g., 0.25 \* B0), ***indicating that the minimum stock size threshold should be life-history dependent***. Similarly, target reference points for stocks with high steepness can be lower such that Bmsy/B0 is smaller than 0.4 \* B0. This happens for stocks with high steepness because they are able to sustain high levels of catch and can be maintained at relatively lower biomass levels to produce the maximum amount of biomass in the system. ***Interestingly, for stocks with high steepness, maintaining a stock size above the threshold and at an MSY level of 0.4B0 is an inconsistent goal, because the stock Bmsy/B0 is actually < 0.25B0.*** In a similar vein, stocks with low steepness might need to have higher stock size thresholds, because of strong depensatory relationships.

For stocks with high recruitment variability, there is still probability of being in an overfished state and being below some threshold level, simply given the variability in the system, irrespective of the level of F. Thus, it may be more appropriate to redefine the minimum stock size threshold by relating the distribution of spawning biomass under no fishing to some probability of dropping below the overfished threshold by x%, under a no fishing condition (as opposed to a fixed percentage of 0.25B0). Additionally, setting lower F-rates can also help counteract the effects of high recruitment variability slightly.

***A 60-20 control rule tended to result in a high probability of being above the MSST relative to other threshold levels, likely because the level of F prescribed for the stock is generally lower, and it provides similar levels of catch.*** Furthermore, under low steepness values, it tends to provide the highest catches, because there is less fishery closures. This control rule performed best particularly when steepness was low, however, under high steepness levels, constant F control rules performed just as well, because stocks with high steepness tend to have low probability of overfishing. Expanding upon the utility of a constant F rule, this seems to outperform threshold rules when steepness values are high because it results in less variability in catch, likely because the probability of being overfished is fairly independent of the HCR, as the high steepness values counteract that.

Lastly, a factor determining the status of an HCR involves uncertainty and errors in estimates of stock status. Specifically, uncertainty in biomass estimates results in decreased yield, increased variability in yield, and increased probability of population collapse. Constant escapement rules are quite sensitive to errors in biomass and escapement because the fishery could close in an instant and there is essentially no room for error. Although fishery closures do not happen for constant F rules, incorrect estimates of biomass could erroneously lead to higher removals in certain years that are not subsequently adjusted for in following years, which could lead to a higher probability of stock collapse. Lastly, there is generally mixed information about how threshold control rules perform under imperfect information. Theoretically, they likely outperform constant F rules, because incorrect estimates of biomass can be adjusted downward in subsequent years if you are on the linear ramp and provides some buffer against uncertainty and errors in previous years. However, it likely results in more variable catches due to the presence of a minimum stock size threshold (although note that the purpose of the ramp is to avoid dropping down to the threshold level and rebuild the stock to a target quicker). Furthermore, they likely perform better as discussed by some studies above, because it better hedges against the probability of overfishing, compared to constant F rules.

The use of uncertainty and buffers in HCRs can be done in several ways and can be useful in addressing scientific uncertainty, environmental conditions, and unmodelled processes in an assessment model. The easiest way is to set target and limit reference points (i.e., target = F40%, limit. = F35%), and catches are not allowed to exceed the limit reference point. Here, the difference between the target and limit reference point referents an uncertainty buffer. This is analogous to: OFL > ABC > ACL > ACT. Other ways to incorporate scientific uncertainty and buffers is the Pstar approach, where the ABC is calculated based on the OFL posterior (e.g., F35%), where it is the xth percentile of the OFL posterior distribution, representing x% probability of overfishing. Additionally, we could also specify that ABC is x% of the OFL to provide some sort of buffer (akin to the first approach discussed). While the use of uncertainty buffers reduces overfishing risk and variability in catch, it does come at the cost of foregone yield, and the degree of the buffer necessary may depend on data quality as well as prevailing conditions.

A group of white rectangular objects with black text

Description automatically generated with medium confidence

Lastly, I will discuss some general issues with HCRs. Throughout this section, I have discussed that HCRs will generally depend on the life-history of the stock as well as management objectives. As such, a good HCR attempts to reach an adequate trade-off between a multitude of conflicting objectives ranging from maximizing catch, minimizing catch variability, and reducing the risk of overfishing and reaching an overfished status. It will also be able to conform to the current management system and is conformed to and can be readily enforced – having a HCR that is not enforced is the same as not having any management. In addition, a good HCR will anchor decision-makers to a pre-defined set of rules such that they are unable to easily deviate from it, and reduces the influence of politics in influencing decision-making. However, this can potentially be a determinant and may require substantial effort to revisit/revise an HCR when it is needed. A good HCR should also be thoroughly tested using MSE, to ensure that it adequately meets the competing objectives of the fishery. HCRs should also explicitly acknowledge the scientific, environmental, and model uncertainty in advice, and allow for buffers. Furthermore, HCRs should allow for a reduction in catch when there is increased scientific uncertainty, or if a stock is below some target value, to allow for the stock to quickly rebuild to a target (i.e., threshold).

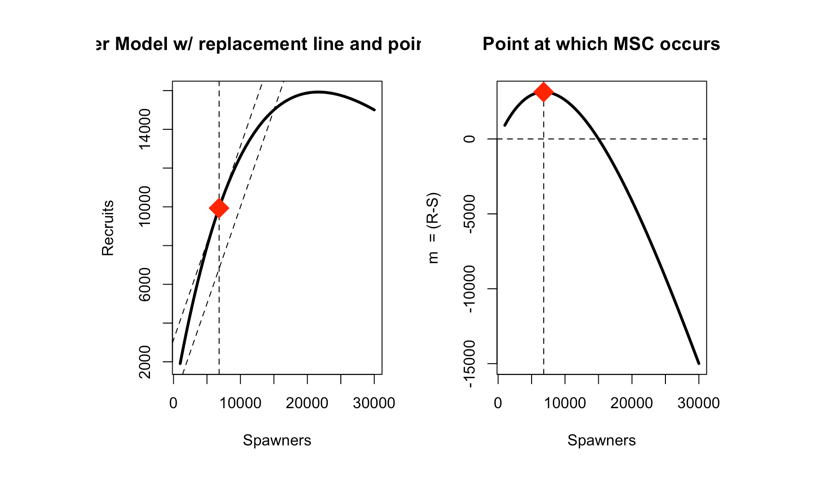
Finally, a good HCR should have meta rules. This can include: 1) ensuring that catch limits don’t exceed 10% from the previous year, 2) conditions and timing of a review of a strategy, 3) specifications and rules for exceptional circumstances. Provisions can also be made to adjust the HCR. However, the number of provisions and meta-rules should not be result in difficulties in decision-making or less efficient performance of HCRs in achieving objectives. Provisions are generally made to protect stakeholder’s interests or to address uncertainty and shocks. While provisions can be helpful, they often make the HCR less effective, and can be mutually exclusive which can make resolving conflicting provisions difficult and subject to political pressure, which gets away from the benefit of an HCR. As such, provisions can be seen as allowing decision makers to provide scientific advice, while diminishing the impact of science on the outcome of the policy. A key example of having a large number of provisions is Northeast Arctic: 1) TAC in a given year is a 3 year average of the current fishing mortality regime, 2) catches cannot change by more than 10%, 3) if the HCR implies an F below 0.3, catch should be set at F = 0.3, and 4) if the stock is below a limit reference point, then no restrictions are needed for the 10% variability clause.

# Dan Goethel

## Reference Points

Management reference points are generally model-based and are used for setting catch limits within fisheries management. These reference points are based upon equilibrium theory, wherein, if we were to fish at a particular reference point in the long term, then on average, we could be able to maintain the biomass at X, while allowing us to remove X fish from the system. In practice, these reference points are commonly integrated into our harvest control rules, as opposed to being directly used, which allows us to implicitly deal with uncertainty in the reference point, particularly at lower stock sizes. Generally, there are three types of models to derive reference points, which include: 1) spawner-recruit, 2) dynamic pool (per-recruit), and 3) production models (from a surplus production model).

Reference points derived from spawner-recruit analysis allows for the estimation of the number of spawners that are required in order to produce a certain number of recruits. However, these reference points are often reserved only for semelparous species because there is no age-structure incorporated in these analyses and spawners are only assumed to spawn once, thus, their contribution to recruitment can be easily enumerated. By contrast, stocks that are iteroparous can contribute to recruitment more than once, and it becomes difficult to understand the contribution of a given spawner to recruitment over its lifetime. The principles of spawner recruitment revolve around the fact that productivity of your stock declines as your spawners increase, that is to say, your R/S declines as stock size gets large, reflecting density-dependent compensation mechanisms. Depending on the stock-recruitment relationship assumed, the origin of density dependence can vary. In particular, assuming a Beverton-Holt stock recruitment relationship assumes that all density-dependence results from the early life-history, such that these processes act upon juveniles, which include can include competition or egg survival. By contrast, assuming a Ricker stock recruitment relationship assumes that increasing number of spawners can result in 0 recruitment, implying that density-dependence acts upon the adult population through cannibalism or competition for spawning grounds. For these types of models, finding the maximum harvest at equilibrium is done by taking the derivative of the function, and solving it to find a slope of 1, which corresponds to the slope of the replacement line and represents the largest vertical difference between the replacement line and the number of spawners (which is our MSY point).



A lot of fishery systems utilize per-recruit based reference points, which come in two flavors: 1) Yield Per-Recruit (YPR), and 2) Spawning Biomass Per-Recruit (SPR). For YPR reference points, these seek to minimize the probability of growth overfishing and maximize catches on a per-recruit basis, resulting in our Fmax reference point. It considers weight-at-age, natural mortality, fishery selectivity, and attempts to trade-off between these values in order to maximize catch. A yield curve can be produced from this method, where the yield curve is generally dependent on selectivity. In particular, if selectivity is more left shifted, the yield curve will generally be more dome-shaped, whereas if it is more right shifted, the yield curve at higher fishing mortality should generally be flatter. Changing values of natural mortality and weight-at-age will generally change the scale of the yield curve, but the shape will remain fairly invariant. In order to maximize catches here (and in general), F needs be infinite when the cohort biomass is maximized (i.e., at the critical age). In addition to Fmax reference points that can be derived from a YPR analysis, practitioners commonly use F0.1 as a proxy for Fmax. In particular, this is done because Fmax is often larger than Fmsy, and fishing at Fmax can potentially result in overfishing. The F0.1 reference point is more conservative and is below Fmax, although the 0.1 is arbitrarily defined. Nonetheless, the F0.1 reference point represents where the YPR curve is at 10% of the slope of the YPR curve at its origin. While YPR analysis accounts for growth overfishing, it fails to account for potential for recruitment overfishing, which leads into why SPR reference points are more used nowadays.

SPR reference points are often used in contemporary management systems in the absence of a stock recruitment relationship, because it better accounts for factors relating to recruitment overfishing, whereas YPR reference points fail to do so. In particular, SPR reference points require inputs of fishery selectivity, natural mortality, maturity, and weight-at-age. SPR analyses are done as follows: 1) an unfished SPR is calculated representing 100% SPR, 2) a fished SPR is calculated, 3) the fished/unfished SPR is calculated to represent some percentage, 4) a level of F is found to try to keep the SPR at that percentage. In general, 35% and 40% are typically used as and represents a proxy for MSY quantities, in cases where a stock recruitment relationship is not estimable. Bx% reference points here are then usually calculated by multiplying the Bx% reference point by average recruitment, as a proxy for equilibrium recruitment. These reference points (and most reference points) are generally sensitive to the levels of M, where higher values of M would result in a higher F40%, due to the need to remove individuals more quickly to satisfy the constraints of x%. However, while this rule prevents recruitment overfishing, there are some alternative considerations that may need to be considered. This includes economic considerations as well as life history considerations. In particular, fishing at Fx% essentially penalizes the stock for being at a high biomass, and drives the stock back down to Bx% , which may or may not be appropriate depending on the stock and the fishery. Nonetheless, these reference points are generally more preferable than YPR reference points because they account for reproductive potential.

Lastly, surplus production models can derive management reference points which implicitly assume that fishing the biomass at some intermediate level will allow it to produce maximum latent productivity and accounts for density-dependent effects (i.e., high biomass reduces maximum latent productivity). Different flavors of deriving MSY reference points here result from differences in the production model. A Schaeffer-Gordon model assumes that the MSY point is carrying capacity/2, the Fox model assumes that MSY is 0.37 \* carrying capacity, while a Pella-Tomlinson estimates where this point should be. However, most of these models do not account for age-structure, selectivity effects, assume equilibrium conditions, and do not necessarily account for recruitment (although modifications can be made). Production models to derive reference points outside the context of a surplus production model, involves the combination of YPR, SPR, and stock-recruitment. Here, the combination of these three models allows for the derivation of Fmsy and Bmsy quantities. In particular, a given level of F corresponds to a given level of SPR. For that given level of F, equilibrium recruitment can be found by inverting the SPR (S/R => R/S), which represents the slope of the stock recruitment relationship. Where the slope intersects the stock recruitment curve then represents equilibrium recruitment. With the value of equilibrium recruitment, YPR can be multiplied by equilibrium recruitment to obtain equilibrium yield, representing our MSY quantity. This type of reference point is able to account for recruitment overfishing through the use of a stock recruitment relationship and seeks to maximize yield while considering recruitment overfishing. In general, a higher steepness values will allow for a higher Fmsy quantity in these cases, while lower steepness values will result in lower Fmsy values, given that your equilibrium recruitment will be lower.

As briefly discussed above, there are several model-based reference points that can be used within a typical fisheries management framework. These can include SPR reference points, which can be configured to be maintained at a higher or lower level, depending on the life-history of the stock. For low resilience stocks, these percentages would likely be higher, such as to avoid recruitment overfishing. Commonly, 48% is used as a proxy for maximum economic yield, where these reference points typically allow lower levels of fishing than MSY-based reference points, while maintaining similar levels of yield. Other reference points include Fmax from a YPR analysis, which corresponds to the F where YPR is maximized, and F0.1, which corresponds to the F at 0.1 \* the origin of the YPR curve. From the combination of a stock recruit, YPR, and SPR analysis, Fmsy and Bmsy based reference points can also be derived which represents catch and biomass levels that maintain maximum latent productivity.

In addition to model-based reference points, empirical reference points are sometimes used, particularly in more data poor cases. Often, *M* is used as a proxy for Fmsy (typically as a limit reference point), although it is well recognized that *M* typically overshoots the Fmsy reference point and some proportion of *M* is used as the Fmsy proxy instead. In particular, the NPFMC uses 0.75 \* *M* as the ABC, while life-history studies have found that Fmsy tends to correspond to 0.85 \* *M* (Zhou *et al.*, 2012). Although this requires an estimate of *M*, this can be derived via empirical estimators (maximum age, maturity, etc.). Other empirical reference points can include the use of length-based indicators (e.g., mean length, proportion of large fish), which I will discuss in a much later section (when summarizing Curry’s papers). Additionally, empirical reference points can be derived using average catches (although these don’t necessarily represent MSY), and using indices derived from surveys, as well as surveys themselves as reference points.

With particular focus on spatial aspects of reference points, I will first discuss why spatial considerations need to be accounted for. In particular, most traditional reference points assume that vital rates are homogenous across areas, which they seldom area. In particular, growth, natural mortality, recruitment, movement, and fishery (selectivity and effort) processes are likely non-homogenous across the unit stock area. When such spatial differences are not adequately accounted for, biases can result in reference points and localized depletion can be a key concern in such systems. This can result in certain areas being disproportionately harvest at a higher rate or certain areas being fished in an unsustainable manner, given that the derived reference points do not accurately represent the true reference point, had those values been derived from a spatial perspective.

Some key issues with non-spatial reference points are:

1. productivity values for sub-areas are likely to be overestimated or underestimated, which can result in their overharvest or underharvest,
2. movement dynamics (e.g., natal homing, source-sink) and population structure are not accounted for, which can result in certain areas being harvested inappropriately,
3. recruitment processes might be heterogenous, and applying non-spatial reference points can disproportionately harvest areas with lower recruitment,
4. differences in selectivity in certain areas are present, which can result in reference points in certain areas differing (i.e., F reference points will generally be lower if selectivity removes young individuals in certain areas),
5. unequal harvest pressure in certain areas (e.g., marine reserve areas, or distance to port, or productive regions).
   1. In the context of a marine reserve, a panmictic assessment is going to underestimate the productivity of the stock because the marine reserve area is not reflected in the assessment via its data, additionally, biomass is underestimated because the estimated age-structure of the model is more truncated in a panmictic population, then would be expected if marine reserves were accounted for. Thus, estimating a reference point here in this case would indicate that your stock is likely less productive than it actually is, resulting in management advice that is a bit more conservative (McGilliard *et al.*, 2015).
   2. In the context of productive regions, when it is assessed under a single area model, but fishing effort is focused more on the productive region, your system wide MSY harvest rate is going to be partitioned out less to the less fished area, and more to the harder fished area. Given the different F’s, your harvest rate MSY is going to allocated incorrectly than it actually would be if spatial reference points were used, such that the more productive area is harvested harder than the true productive\_umsy value, which is why you end up depleting the more productive stock.

There are several flavors of spatial reference points.

1. In terms of empirical reference points, one could use highest-density areas (HDA) from a survey (Reuchlin-Hugenholtz *et al.*, 2016), which is calculated as the proportion of tows in a given year that falls within the highest x% quantile of density (e.g., 5kg per area) over its entire time series. For example, if our 5% quantile is 5kg per tow across the entire time series, if we only get 5 tows that are larger than or equal to this value, then our HDA value is 0.05. The x% quantile cutoff is a bit arbitrary, although a high quantile could swamp the signal of what a high density is, while a low quantile could result in overly conservative advice, and a lot of 0s.
   1. Thus, in years where there are lots of high-density areas, the HDA value will be high and the stock is considered to be in good status. By contrast, if it is low, the stock is considered to be in poor status.
   2. For many stocks, HDA has a concave relationship with SSB, where below some threshold SSB value, HDA declines more quickly than a unit decline of SSB. This point can be considered as a limit reference point, and can foreshadow disproportionate declines in SSB, if fishing is not controlled in an adequate manner.
      1. Similarly, 70 – 80% of the maximum observed HDA could also be considered a target reference point.
   3. For many stocks, 30% of the maximum HDA observed corresponds to when 0.5 \* SSB\_MSY is. With respect to the determined threshold, this is usually 20 – 30% of the maximum observed HDA, such that if a stocks HDA declines by 70 – 80% of its maximum, there is potential for SSB to drop precipitously.
      1. For a lot of stocks, the threshold HDA reference point is higher than 0.5 \* SSB\_MSY, while for others, it is lower. Thus, if we were to use this in a real world scenario, choosing a x% quantile that results in a threshold point that is larger than 0.5 \* SSB\_MSY would be considered the precautionary approach.

A diagram of a curve

Description automatically generated

Some of the key advantages to these reference points include a lack of assumptions regarding natural mortality, spatial structure, selectivity, and simply relies on a survey to provide us with this information. These high-density areas could be of utility given that it protects localized regions of high densities that may disproportionately contribute to spawning, protects spatial structure in the context of maintain a decent number of high-density areas, and implicitly protects sub-populations that are highly concentrated/productive. With respect to the calculation of spatial reference points that are non-empirical, this can be done by running a model to equilibrium (Goethel and Berger, 2017; Bosley *et al.*, 2019), and then searching across a combination of F values across areas and fleets that maximized the system-wide yield. Other approaches include using an iterative search algorithm and calculating equilibrium recruitment for an area-specific stock recruitment relationship (Kapur *et al.*, 2021), as well as estimating area-specific reference points, which implicitly ignore movement dynamics.

There are some implications of difference population structures and movement on how spatial reference points behave, as well as the consequences of mis-specifying population structure. In particular, there are considerations regarding whether meta-population or spatial heterogeneity stock structure exists, source-sink dynamics, movement dynamics, catch allocations, impacts of selectivity and maturity, and connectivity. Briefly, it is assumed that metapopulation structure tends to arise from environmental factors or a combination of genetic factors (i.e., localized density dependence), while spatial heterogeneity arises from differences in fishing effort allocation (e.g., marine reserves) or differences in how recruitment is allocated to different regions due to prevailing currents. In the context of the impacts of population structure on spatial reference points:

1. System wide harvest rates are generally fairly invariant of the underlying population structure as well as movements, although it sometimes depends on if there is metapopulation structure coupled with source-sink dynamics (Goethel and Berger, 2017; Bosley *et al.*, 2019).
   1. In particular, in the case of source-sink dynamics, the source population generally needs be protected and thus, has a lower harvest rate at MSY, while maintaining a higher harvest rate at MSY for the sink population, in order to maximize system-wide yield. Additionally, source populations tend to have lower SSBMSY because it continually loses SSB to the sink population and hence, results in lower recruitment (Goethel and Berger, 2017).
   2. Additionally, SSBMSY is generally more dependent on adult movement in source sink dynamics, rather than larval dispersal, because more biomass is lost via movement as adults, rather than larvae.
   3. Under scenarios with metapopulation structure, bidirectional movement, and different productivity, estimates of SSBMSY and reference points are similar to that of source-sink dynamics, where areas with high productivity should have a lower harvest rate, and areas of low productivity should have a high harvest rate, given that it receives contributions from the highly productive population. However, for these scenarios, SSBMSY was lower for the productive region, because more fish were moving out of the productive region, and hence, contributed less to SSB.
   4. Similarly, when there is directional ontogenetic movements, nursery areas should not be fished, while adult areas should have high fishing pressure. The no fishing in the nursery area ensures that individuals grow old enough to allow for maximizing yield (Bosley *et al.*, 2019).
2. Under scenarios of natal homing, system-wide SSBMSY is generally lower and harvest rates are lower, because these individuals occur in mixed areas, and we want to ensure that they contribute to SSB at some point, and not remove all of them at once.
3. Comparing metapopulation structure and spatial heterogeneity structure, where the latter does area-wide recruitment with recruitment apportionment, while the former does area-specific stock-recruitment relationships, differences in associated harvest rates and area-specific depletion can arise, likely because having area-specific stock recruitment relationships especially under ontogenetic movements can potentially result in higher equilibrium recruitment in adult areas.
4. Additionally, both maturity and selectivity play a role in mediating optimal harvest rates from a spatial context, and these interactions can be counterintuitive.
5. In stocks with high residency, this is distinct stocks and thus can potentially be treated as such without much consequence to reference points. This is akin to scenarios without any movement, because fish are more inclined to stay in a given area.
6. Lastly, cases where movement is present and where movement is not present can lead to some distinct differences in yield isopleths. In particular, in cases where movement is not present, the yield that maximizes the catch for both systems has a narrower range. By contrast, in areas where there are high movement rates and exchange between two stocks, the yield that maximizes system wide values encompasses a broader range of values, because populations are not independent and there is some allowance for demographic rescue effects (i.e., transfer of individuals across areas) as well as allows for dispersal of fishing effort over the model domain that causes area-specific harvest strategies (e.g., harvesting one region has impacts on another region for spatial stocks, whereas independent harvesting manifests for non-spatial stocks).
7. In many population structures, different area-specific harvest rates are required to maximize system-wide yield. Thus, assuming the same harvest rate for each area is not appropriate, if spatial structure is present.

In general, getting the spatial movement dynamics in movement and biological factors is more important than getting the population structure (i.e., heterogeneity vs. metapopulation) right, in terms of having reference points that perform adequately. Source-sink dynamics as well as stocks with ontogenetic movements can lead to nursery areas and source areas experiencing 0 fishing mortality effort, which is inconsistent with how management and fisheries operate, and alternative strategies need to be devised. Furthermore, in most of these movement and population structure scenarios, system-wide depletion often tends to be fairly invariant and insensitive to assumptions made regarding population structure and movement, although area-specific depletion can manifest into localized depletion.

In cases where spatial reference points are unavailable, due to data limitations, etc, management systems may be constrained by the use of a single system wide MSY value. There are several methods in which this system-wide value can be allocated on a spatial scale, and the goal is to partition this system-wide value to represent spatial reference points and to avoid localized depletion. However, this often will depend on how the allocation is being done and the susceptibility to errors in the allocation method. Several methods are possible including using a survey index (some average or some weighting scheme to minimize the impacts of error), using a recruitment index, or equal allocation among areas. However, Bosley et al. 2019 showed that the utility of a survey index in allocating catch will depend greatly on movement dynamics as well as the error in an index. Using a recruitment index with high variability led to poor performance. Furthermore, when there was not movement among regions, the survey allocation worked well. However, under scenarios with high movement rates, survey allocation methods and equal allocation worked equally well, likely because biomass was equally spread across the two areas and the index tracks this aspect***. Noting some of the spatial harvest strategies in the context of ontogenetic movements and not allowing fishing in nursery*** ***areas, some of these allocation methods can be inconsistent with these strategies, because indices do not necessarily take into account the age-structure of the population, and thus will allow fishing in regions with any biomass, irrespective of their age-structure.***

### Spatial Modelling

Spatial dynamics in the marine environment are ubiquitous, and rarely are fish species stationary in space and across time. However, a common assumption that is often made within stock assessment modelling is the unit stock assumption, wherein individuals within an assessed area are assumed to have homogenous vital rates and demographic structure (age, size), a closed population is assumed, homogenous fishery dynamics, and that individuals are uniformly mixed across the model domain. However, this assumption is seldom met. Presently, there are several approaches to confront spatio-temporal dynamics in stock assessments, which include: 1) CPUE standardization across space to remove the effect of spatial dynamics, 2) areas-as-fleets (spatially implicit approaches), 3) box-transfer models (spatially explicit but stratified approaches), and 4) spatio-temporal models. Such approaches are appropriate for accounting for spatially correlated residuals, aids in the detection of localized depletion, increases biological realism (spatially varying biological and fishery parameters, as well as how recruitment is apportioned across space), allows for explicit accounting of spatial features such as marine reserves and spawning grounds.

A core assumption in many contemporary spatial stock assessment models following the box-transfer paradigm is that movement occurs instantaneously and that the transfer/movement coefficients are dependent on the size/area and shape of the population area. The diffusion model here also assumes that as fish move into another area, they became instantaneously well-mixed, with all fish having the same probability of movement and residence. However, in reality, this assumption is rarely met because individuals that are at the boundary of the population area are more likely to move – although these models often perform well in adequately representing movement dynamics. Furthermore, an important assumption is that movement rates are a first-order Markov process, such that movements are a random process and that future states only depend on the current state, and not previous states. While movements on a small time-scale are not necessarily a random process, if we were to model these processes on an annual time-step, they would represent a collection of different movement paths, and the assumption of random diffusion is likely to be met. Additionally, while the assumption of Markovian movement can be appropriate for many stocks, these assumptions are not valid for stocks that exhibit movement patterns such as natal homing and feeding migrations, and modifications can be made to account for that (advection terms). Lastly, as discussed below and by Beverton and Holt 1957, it is best that population areas were divided into strata as small as possible.

When conducting a spatial assessment, there are certain assumptions and decisions that are required:

1. Assumptions regarding population structure,
   1. To start off a spatial model, a key decision point that needs to be made is that stock structure of the population. In particular, this involves knowledge whether there is exchange of fish across areas, if there are genetic or phenotypic differences across areas, if fish exhibit natal homing dynamics, whether certain areas are more productive than others, oceanographic features, etc. In this case, a full suite of stock ID methods should be employed to discern the most complex spatial structure plausible, which should then be consolidated based on data availability.
   2. There are various types of population structures that can be represented which are governed by the scale of density-dependence, degree of reproductive mixing, connectivity among areas, as well as spatial variation in demographic and vital rates. These different population structures can be modelled in a variety of ways, which include: 1) single stock with movement among areas, 2) single stock with no movement post settlement, 3) multi-stock multi-area movement, but without dispersal among stocks, such that stocks are genetically distinct, but mix together during feeding, 4) multi-stock multi-area movement, with dispersal among areas. Note that differences in demographics due to area effects are implicitly driven by environmental conditions, while stock effects are generally driven by a combination of genetics and environment. Nonetheless, these structures can be broadly categorized into the following:
      1. Panmictic, where vital rates are assumed homogenous, with high movement rates within the modelled domain,
      2. Spatial Heterogeneity, where vital rates can be spatially varying, but reproduction occurs on a global scale. Such a scenario often occurs when there is substantial larval drift and represents post-settlement dynamics. Under these conditions, if the population is well-mixed and exhibit high movement rates, similar population trajectories can be achieved with a panmictic population, given these high movement rates,
      3. Meta-population, where vital rates can also be spatially varying, but reproduction occurs on a local scale. This scenario occurs when there is localized reproduction with limited dispersal,
2. Assumptions regarding the number of fleets to model and the number of areas to model,
   1. The number of areas to model should generally be dictated by *a priori* hypotheses about what stock structure looks like, which should then be consolidated down based on the availability of data. More generally, having too many areas or fleets is less bad than having too few areas or fleets, although the trade-off there is increased imprecision in model estimates.
   2. Often, spatial areas and fleets match the scale in which data are collected and can lead to biased advice if management boundaries are not coherent with the spatial population structure. As such, areas and fleets should be selected such that demographic rates, population trends, and composition data within a given area should be as homogenous as possible. This can be done using binary regression trees. Thus, the availability aspect is adequately modelled and differences in selectivity or composition data are mainly due to the effects of contact selectivity.
   3. Some assumptions can be made to reduce the number of areas or fleets to model and to increase model tractability. In particular, fleets-as-areas models as a spatially implicit model can also be conducted. However, note that this is a crude approximation of spatial dynamics, and differences in composition data are attributed only to selectivity effects, and the model essentially assumes a uniformly distributed stock. Additionally, we might assume shared selectivities and catchabilities among areas, or share information across areas (share information to weak stocks). However, if this is the case, dome-shaped selectivity dynamics might be considered, given differences in availability across areas, as well as differences in fishing effort across areas,
3. Assumptions regarding the temporal structure of the model,
   1. The temporal structure of the model can greatly depend on the life-history characteristics of the stock, as well as the characteristics of the fishery. In particular, decisions are required for natal homing as well as feeding migrations (also growth) regarding seasonal time-steps (i.e., where and when mixed-stock catches occur). By contrast, another decision that could be made is to assume an annual time-step, although these considerations will likely be dictated by data availability and model parsimony,
4. The parameterization of movement dynamics,
   1. Most studies have shown that correctly parametrizing movement dynamics and getting these correct, is more important than getting the population structure correct. In particular, having movement dynamics that are too inflexible can result in biases in both spatial reference points as well as estimates of biomass. In general, movement parameters can be inferred through changes in age-composition across areas, although the use of tagging data can substantially aid in model convergence, stability, and reduce parameter correlation when they are available.
   2. Several methods are available to model movement dynamics, which include flexible approaches – either as continuous random walks or time-blocks, anchor parameter estimates to some functional form (i.e., older ages have higher movement rates, vice versa), linear interpolations across ages via a gravity model, use of habitat preference functions or covariate to force movement, assume no movement post-settlement (i.e., model a recruitment apportionment factor and assume no movement in adults), or the use of random-walks to force movement. Utilizing an approach that is minimally complex and flexible will likely provide the most robust results, particularly when movement dynamics are uncertain.
   3. Generally, the estimation of movement is possible without the use of tag data and is reliant on good age data (Hulson *et al.*, 2013; Bosley *et al.*, 2022). The use of tag data comes with its own sets of assumptions and difficulties which are discussed further below. However, with the use of tag data, parameter and model estimates are generally more stable and confounding is less of an issue. In particular, without the use of tag data, it is difficult for a spatially explicit model to infer whether changes in abundance are due to increased recruitment in one area or due to movement into a given region. Furthermore, the use of tag data has been seen to improve uncertainty estimates of fishing mortality, precision in abundance estimates, and catchability, because additional information is being provided to inform these parameters through the tagging sub-model (Hulson *et al.*, 2013).
5. The degree of variability in biological parameters across space,
   1. With respect to how variability in biological parameters should be modelled across time and space, these demographic rates should be modelled, because it is difficult to approximate these with other processes, and failing to adequate represent these spatially, when they vary, can lead to biased parameter estimates.
   2. Some key considerations when modelling vital rates in spatial stock assessments are the possibility of shrinkage in size or un-maturing for individuals that move. Shrinkage in size can occur due to differences in biological parameters in an area, while un-maturing can result if the age-at-maturity in one area is smaller than another area.
      1. With respect to shrinkage in size, this can easily be dealt with using a size-structured model, where there is some probability of growing to a given size as you move areas (i.e., use of growth increments). However, this is more difficult for age-structured models, and requires tracking where individuals moved from, in order to correctly calculate growth rates. Another newer approach could be to base growth increments on the difference between the current size and the area’s asymptotic size to do the transition.
      2. With respect to un-maturing, this can easily be dealt with by separating out the stock into mature and immature components, similar to many crustacean assessments.
6. Tractability and practicality in parameter estimation, parsimonious model structures,
   1. The last key decision in spatial stock assessments is how to ensure tractability and parsimony in parameter estimation. Often this is dependent on the available data and the scale in which data were collected at.
      1. If jurisdictional boundaries do not necessarily align with modelled areas, some assumptions and modifications to the available data need to be made, which can include the use of likelihood functions that are fit to aggregated composition or indices, while fitting to spatially-explicitly composition or indices, particularly if data are sparse,
      2. With respect to the use of tag data, there are several nuisance parameters that need to be estimated, which are often confounded with recruitment and natural mortality. These parameters include tag mortality, tag shedding, tag reporting rates, and tag-mixing. In terms of dealing with tag-mixing, some common assumptions that are made include using some cutoff with respect to the time-at-liberty for individuals. Tag-mixing assumptions are required and important, because you want your tagged population to experience/be representative of the rest of the population. Tag mixing can also be dealt with by assuming unmixed individuals experience a higher F during some initial period, or using a likelihood function that accounts for overdispersion (negative binomial or Dirichlet instead of multinomial).
      3. Additionally, some approaches to dealing with tag-shedding and tag mortality would be to simultaneously conduct double-tag studies and using these external estimates to reduce the burden of estimation for an assessment model. Tag reporting rates are a bit more difficult to deal with and fixing it at the wrong value can result in biases on par with ignoring movement all together.
      4. Outside of parameters related to tagging, if certain demographic parameters are not suspected to exhibit strong spatial variation, it is also possible to share information across stocks, by assuming that these parameters are constant across space, which may help increase the precision of model estimates.
      5. As discussed above, minimally complex but flexible movement parameterizations are required for robust model results. While there are flexible approaches that can encapsulate a wide-variety of movement patterns, the use of habitat preference functions, covariates, and functional forms in movement rates can help constrain movement parameter estimation and increase model stability.
      6. Lastly, a simplification to a spatially explicit approach could include the use of a spatially-implicit fleets-as-areas approach. This removes the need to estimate movement parameters and treats each area as a fishery fleet. A similar approach would be to model areas as separate stocks, assuming limited to no movement between areas – this allows for an assessment of stock-specific abundances and can potentially be robust, as it avoids smearing estimates across an entire region. This case can be of utility particularly if there is no tag data available, or that estimates for movement are unrealistic, in the absence of tag data. Lastly, another approach to dealing with spatial structure in stock assessments as well as movement dynamics that have seemed to work is: 1) re-weighting sample sizes, and 2) allow for flexible time-varying selectivity. These approaches likely work because in the first approach, it deals with unmodelled processes and reduces the influence of fitting composition data to other data sources such as abundance indices. In the second approach, aggregating spatial dynamics and catches generate complex dome-shaped patterns that vary over time, which can be implicitly captured with time-varying selectivity approaches.

As discussed above, there is potential for misaligning the biological population structure with management boundaries, given the scale at which data were collected at. However, there can be severe consequences in both estimates of biomass and associated reference points if these boundaries are misaligned with biological boundaries because it causes “demographic leakage” such that the incorrect processes are being averaged. This is particularly of concern because management boundaries are often defined based on political decisions, ease of data collection, or based on ecological theory, but are treated as static. As such, movements outside of these areas can erode the assumed population structure, and it can appear that the stock is depleted, even if it just moved. As shown by a study from Berger et al. (2020), misaligning population structure can lead to severe biases, which are further exacerbated by movement dynamics as well as recruitment dynamics, when movement was unidirectional (i.e., movement outside assessed areas). This implies that it is important to monitor the assessment and management boundaries, and to be adaptive to changes in distribution. The impact of spatial heterogeneity in growth and fishery processes were generally less influential on the degree of boundary incoherence, although still had non-negligible levels of bias. In addition to biases in biomass dynamics, biases in management advice can also occur, where some areas end up getting fished harder than others, resulting in under or overharvest. While system-wide results are generally unbiased if boundaries are incorrectly specified, there is a need to have thoughtful and pragmatic catch allocation schemes to avoid the potential of localized depletion.

In addition to the importance of correct aligning management boundaries, there are similarly implications for misaligning/mis-specifying the true underlying population structure (i.e., assume meta-population when they are one unit stock) (Hulson *et al.*, 2013). Similar to themes discussed above, incorrectly assuming a panmictic population when there is spatial structure generally allows for an adequate estimation of system-wide values, although the estimation of area-specific values results in bias and may manifest as localized depletion. Furthermore, mismatching the underlying population structure generally performed well, only if the movement dynamic were flexible enough or adequately represented. Furthermore, it is generally you are better off using a panmictic population if you get movement dynamics incorrect – getting the movement dynamics incorrect is more detrimental than getting the spatial structure of the stock incorrect.

As discussed above, it appears to more important to get movement dynamics right, than it is to get the spatial structure of the stock correct, although there are some considerations with respect to sampling designs (i.e., 1 year vs 5 year tag programs) and the type of movement patterns (i.e., ontogenetic vs directional vs combined) This is observed in other studies (Goethel *et al.*, 2021), where allowing more minimally complex but flexible movement parameterizations (e.g., time-blocks for age-and year in two year increments) were robust to the underlying true movement dynamics (i.e., ontogenetic, climate movement, etc), but allowing for overly simple movement dynamics lead to severely biased estimates of biomass. Similarly, incorrectly specifying the type of movement can be worse than assuming no movement at all (i.e., assuming age-based when its yearly). This was also found in another study, where assuming ontogenetic and climate-forced movements when movement dynamics were actually random led to fairly large errors in biomass for Eastern Bering Sea walleye pollock, and that assuming a panmictic population model can perform just as well if movement dynamics are mis-specified (Hulson *et al.*, 2013). Interestingly, Hulson et al. (2013) posit that misspecification in movement patterns can potentially be detected through misfits to regional abundance indices. In particular, if movement is mis-specified under climate forcing scenarios, assessment models will often severely underestimate the biomass index in that region, which can potentially be inferred through fits to the index. Furthermore, some other considerations need to be accounted for such as when ontogenetic movements were present, ignoring these movement dynamics led to the most bias, presumably because there are not a lot of processes that can account for ontogenetic movements (i.e., recruitment cannot really account for changes in age-classes across areas). In a similar vein, when no movement parameters are estimated, the model compensates by increasing recruitment in one area, relative to another area. This will likely predict rebuilding of a stock and a depletion in another stock, when in reality, the stock is actually not changing and just shifting its distribution and may result in incorrect management advice. Additionally, when unidirectional movement dynamics are present, estimating complex movement parametrizations are generally less beneficial, presumably because a unidirectional movement represents a one-way trip, and tag data become more informative for one population rather than the other – under such circumstances, it may be better to disproportionately have higher tagging effort for the less informative population.

Although the use of spatial stock assessment models is necessary to estimate unbiased estimates of population quantities, data limitations and complex parametrizations can preclude their use in management systems, as well as institutional inertia. The use of single-area assessments or spatially implicit assessments can still achieve management goals if coupled with a robust harvest strategy, despite potential biases in population biomass. One particular example that is a common occurrence includes the presence of no-take marine reserves, which displaces fishing effort and creates spatial heterogeneity in the population (i.e., age-structure in the reserve is much larger). In circumstances where there is limited post-settlement movement between marine reserves and fished areas, it is common to use spatially implicit models such as the FAA approach to assess stocks (because movement dynamics can generally be ignored), which can sometimes lead to biases. Using a spatially aggregated approach coupled with a B20:B35:B48 rule, management goals were able to be achieved in the presence of spatial variation in growth and closures, although there was higher catch variability. FAA models also achieve similar results, although they were slightly better, with less catch variability, likely because biomass estimates, and the age-structures were better represented. In general, with assessments that are structured incorrectly, particularly in a spatial context, adding more data does not necessarily lead to more precise estimates, because the spatial scale in which these data are collected are incoherent with the assumed model structure. In particular, spatial heterogeneity in growth and differences in the scale of abundance will result in a non-proportional exploitable biomass index, which can increase assessment error, when data assumptions are not met.

In situations where management goals are not achieved when area-closures are implemented, these were often driven by the choice of area closures by: 1) closing highly productive area, rapidly reducing catch and redistributing fishing effort to other areas of potentially much lower productivity, and 2) increased spatial heterogeneity in the underlying population dynamics, resulting in exacerbated errors in the assessment method. This suggests that it can be important/beneficial to employing a spatially implicit approach as an approximation. However, there can also be some detriments. Even in conditions without movement, selectivity estimates can get severely biased if recruitment apportionment is not accounted for in a spatial model – which a spatially implicit model fails to capture, because selectivity needs to get distorted to match the zone-specific age structure. Given that selectivity drives a lot of the values for reference points, this can result in the incorrect provision of management advice. Nonetheless, it is possible to achieve adequate management outcomes, even when spatial heterogeneity is ignored within an assessment, as long as a robust harvest control rule is utilized.

Despite all of the work that has been done in spatial assessment modelling, there are still many advances and work to be done. In particular, recent developments have allowed for the combination of multiple tag types (floy tags, SAT tags) and point count data (survey and fishery) to estimate fine-scale movement rates as well as abundance on a spatio-temporal scale (Cao *et al.*, 2020; Thorson *et al.*, 2021; Olmos *et al.*, 2023). Here, these methods seek to integrate advection processes represented as vector fields from larval dispersal and currents, taxis processes represented as habitat preference functions towards particular habitats, and diffusion processes resulting from random movement. However, these methods require fine-scale resolution data, and are highly reliant on the use of tag data to provide realistic estimates of movement as well as precise survey data. In addition to fine-scale resolution spatio-temporal models, other advances can potentially enhance the utility of spatial stock assessments, which includes the integration of novel and diverse datasets. In particular, recent advances in genetics has allowed for the development of close-kin mark recapture (CKMR) techniques that enable us to estimate quantities of absolute abundance as well as understand demographic structure and connectivity among populations. However, that are still some limitations that need to be addressed, including: 1) lack of age-samples if these are not collected, because relying on an age-length key that has high variability can make assigning ages difficult, which the method relies upon, in order to ensure parents were mature when the birth of an offspring occurred, 2) similarly related, the birth year and the offspring and estimating the number of individuals in a given year is heavily reliant on age data, and can lead to imprecise and potentially biased estimates if ages are uncertain, 3) sex-markers are needed to understand and differentiate between males and females to match up individuals to their parents, 4) on a related note, estimating CKMR within an assessment can potentially help inform some of the sex-ratios in the population and resultant estimates from CKMR, 5) sampling needs to be done such that there is complete mixing – i.e., sampling independent populations, if there is not, biased estimates of abundance can result (Trenkel *et al.*, 2022).

Other types of data that can be and should be integrated into assessment models include electronic tag data, PSTAT tag data, etc, which can be done with recent advances using advection-diffusion-taxis models from Thorson et al. 20201. Additionally, tag data provide a means to:

1. estimate movement or abundance via a Brownie tag attrition model or a Peterson Abundance estimator,
2. detect spawning and foraging grounds, as well as stock structure,
3. estimate natural mortality,
4. enable the disentangling of changes in abundance, recruitment, or movement,
5. can aid in identification of larval dispersal through the use of nano-tags for larvae,
6. estimate catchabilities and availability from satellite tagged fish, and its interactions with harvest,
7. understanding of use of marine areas and their utility, as well as the utility of dynamic area closures.

However, some key challenges with using some of the more unique tag data types include: 1) data representativeness of electronic tag data, and how sampling designs can be maximized to gain inferences both on a population scale as well as on an ecological scale, and 2) combining Langraigian movements with Eulerian movements.

Lastly, I want to briefly synthesize the utility of the fleets-as-areas approach, when its use is most appropriate, detriments to it, and some of the key assumptions it makes. Fundamentally, a FAA approach assumes a single homogenous population and all individuals are well-mixed. Here, fleets are treated as areas and captures spatial variation in population age or size structure. Thus, it approximates spatial distribution in ages or sizes due to movement or differences in exploitation rates or biological differences through selectivity. However, this approach does not address why variation in age-composition arises; in particular, it does not acknowledge spatial variation due to differences in larval dispersal, differences in natural mortality in these regions, or movement etc., and instead, implicitly acknowledge these through selectivity. The FAA approach is known to be biased when faced with spatial heterogeneity as it fails to encapsulate the factors mentioned above. ***However, when confronted with the following, FAA approaches can be appropriate: 1) limited understanding of spatial structure, 2) little demographic variation, 3) no concern of local depletion, 4) limited ability to estimate movement rates, 5) adequate managements strategies, and 6) differences in fishing mortality among areas.*** ***In particular, they can be feasible particularly when there are no-take marine reserves where there are differences in fishing mortality across areas, but stock ID supports a single-stock assumption (e.g., limited demographic variation), or if there are seasonal migrations that temporarily reduce the availability of the stock.*** Additionally, single-area assessments can also be useful in accounting for/adequately representing system-wide values, where time-and age-varying selectivity approaches can be employed to approximate changes in availability, as well as down-weighting composition data in an fleets-as-areas approach to reduce its influence in the estimation process (Hulson *et al.*, 2013; Lee *et al.*, 2017).

## Stock Identification Methods

Prior to conducting a full stock assessment, one of the first steps is defining the underlying stock structure of the population, which is also commonly referred to as stock identification. The goal of stock ID is to identify self-sustaining components of the population and to align the effects of fishing with a biological unit, allowing us to audit the effects of fishing against renewals. In practice, a unit stock has a flexible definition and is often characterized by the goals of management and the extent to which management can be enforced. In particular, a unit stock can be characterized as technical interactions, ecological and biological attributes, fishery attributes, and commercial value. As such, a unit stock should not be defined simply by its biological attributes, but should also consider the attributes of fishery dynamics, given that a unit stock can experience differential fishing mortality rates across space, which can alter their demographic structure, resulting in spatial heterogeneity.

***Stock structure and the unit stock assumption arises from the following interacting factors, including the dispersal of juveniles and adults, geographic variation in traits (phenotypic or genotypic), and the boundaries it encapsulates (i.e., where it travels within a given year).*** These factors generally encapsulate the impacts of oceanographic features, fish behavior (e.g., natal homing), and local adaption to environment.  ***Defining a unit stock implicitly assumes three things: 1) there is no movement inside or outside of the defined area, 2) individuals mix homogenously within the area (because we want to be able to infer population trends from one area to the next – i.e., they experience similar conditions), and 3) vital rates are homogenous within the defined area (mortality).*** Additionally, the unit stock assumption assumes that the production of young fish results from the amount of old fish in that population. As such, if there is some degree of immigration or emigration, a stock-recruitment function will inadequately predict incoming recruits, because the number of spawners are not correctly accounted for (and mean recruitment might be a better representation). ***Thus, a unit stock reacts independently to exploitation given that the processes of growth and mortality outweigh movement dynamics, where the unit stock can be exploited independently, without have its effects propagate into neighboring areas.***

There are several consequences of incorrectly specifying stock structure and treating political boundaries as biological units. In particular, the effects of fishing can be incorrectly audited to a given area, which can bias management advice, and may result in overharvest of the stock, as well as underutilization. Importantly, cases of lack of stock recovery has been attributed to incorrect definitions of stock structure and boundaries. In cases where we are assessing stocks that shift or are actively mixing with other stocks, a less productive stock is going to experience a larger relative fishing effort, when they are prosecuted in areas where they are actively mixing. With respect to incorrectly specifying stock structure and not reviewing stock boundaries, fish species can routinely shift their distributions, potentially in response to climate change. Given that, if individuals shift outside of their stock boundary, assessments will infer an increase in abundance in the area it shifts towards, while inferring a decrease in abundance in areas it shifts away from, even when abundance is stationary, which can lead to radical management advice that is likely to be inappropriate. In a more complicated sense, when fish species experience natal homing dynamics, the ability to audit the effects of fishing towards these spawning populations becomes more difficult, and certain spawning populations may get depleted, despite the overall population remaining fairly stable (i.e., masked localized depletion). As such, the identification of these spawning contingents and stock structure is important to avoid localized depletion. In general, the main concern with incorrect delineation of stock structure is: 1) localized depletion, 2) overharvest, and 3) underutilization, because we are not correctly auditing the effects of fishing against the population, whether it be due to movement from outside the stock area, or movement into the stock area. From a stock assessment perspective, incorrectly defining stock boundaries and stock structure can result in data that are noisier than would be if stock structure was correctly defined, wherein year-class strength (number of spawners to produce recruits are incorrectly characterized), natural mortality, and selectivity can be confounded by movement dynamics if they are not appropriately modelled, which may lead to biased stock perceptions and incorrect provision of management advice.

For many fisheries worldwide, stock structure is generally misaligned with the management boundary, and as discussed above, this can have severe consequences for the assessment of a fish stock. ***One key example is the collapse of the Atlantic cod***, where a misalignment between the stock and management boundary played a central role in driving its collapse. In particular, a misleading index of abundance from the inshore fishery led to the perception of a healthy stock status, despite offshore populations being more declined, and a shift in fish distribution towards southern regions. The misalignment and aggregated assessment of offshore and inshore regions led to the misperception of healthy stock status and overestimates of productivity, resulting in increased harvest even when the stock was declining and shifting its distribution. ***Another example of the consequences of ignoring stock structure and spawning populations is from Atlantic herring***, where small populations were progressively eroded in size and were ignored in assessments because they were deemed too small to be significant contributors to the population. However, since the recognition that these populations may be important, conservation measures have been placed upon these populations, which has helped rebuild some herring stocks***. Similarly, in many crustacean stocks in Alaska***, failing to monitor expansion to new fishing rounds led to the sequential depletion of fishery resources from inshore areas to offshore areas. Failing to separately consider the expansion of fishing grounds to new areas resulted in an inability to detect declines in previously fished areas, resulting in a misperception of healthy stocks tatus because the new expanded areas have higher catch rates. Thus, this suggests that stock structure does not simply encapsulate spatial variation in the fish species themselves, but should also fully consider fishery dynamics.

Despite the potential for overfishing and underutilization as well as the depletion of important spawning contingents, regular monitoring and advances in stock identification can help realign management boundaries to better represent biological units, which can aid in stock recovery. Such approaches can be employed, even when spatial stock assessments are not feasible. Generally, there are couple of approaches to using information about stock-structure in the management process. This includes:

1. status-quo management, where the unit stock assumption is used, despite there being stock structure,
   1. ***Atlantic bluefin tuna*** are a prime example of where despite there being evidence of stock and spatial structure, as well as mixing, a status-quo approach is utilized. Note that the historical development of the stock was initially a single aggregate stock between eastern and western areas, where they were later separated out to be assessed as independent stocks (status-quo). However, there is some degree of mixing, and the eastern area can contribute to the western area up to 63% during their feeding months, and assessing these stocks as independent units will lead to an incorrect attribution of removals to these respective areas, which may bias management results. In these contexts where stock mixing only occurs in certain months, mixed-stock composition analysis is a helpful tool to better allocate removals back to the correct region, as well as monitoring SSB in those respective regions, as well as using a spatial model to estimate stock mixing. However, this approach is better than assessing them as a single aggregated stock, because localized depletion will likely be of concern, and there are differences in productivity between these two areas.
2. weakest-link management, where the weakest-link of the stock component is protected,
   1. In the case of weakest-link management, localized depletion is less of a concern because the weak component is protected from fishing, however, it can lead to situations where yield is underutilized. Conversely, it can also result in unnecessary fishing effort being placed upon populations that are more productive. ***In an example of herring***, a stock collapse led to the protection of the weakest component of the stock, with only 11% of the system-wide TAC allocated to this region. That particular area ended up recovering and being a substantial contributor to the overall population productivity.
3. the use of temporal and spatial closures, where spatial and temporal closures are implemented to protect spawning contingents,
   1. ***For Atlantic cod*** that exhibit spring and winter spawning complexes that are genetically distinct but mix with each other during the feeding migrations, seasonal and temporal closures are implemented to protect spawning contingents, coinciding with when and where these complexes spawn.
4. use of stock composition data, where identifying the population of origin for different stocks that mix in the same area allow for attributing fishery removals to that specific population of origin,
   1. For ***eastern and western Baltic Sea cod,*** which are genetically distinct but mix together, the use of otolith shape analysis to determine the degree of stock mixing resulted in management finding out that the eastern stock was contributing more to the western stock during their mixing period (i.e., more productive stock contributing to the less productive stock) and allowed for parsing out catches to eastern and western populations. If these analyses were no conducted, it would appear that the western stock was growing in size, despite it actually remaining the same, where in reality, it was the eastern stock contributing to increase biomass in the western area and would lead to erroneous allowance of increased fishing pressure in the western region. Stock mixing is a complicated subject and failing to account for these processes can result in a misperception of stock status.
5. redefinition of management boundaries, where separate management units are defined to approximately align with biological units, while maintaining and leveraging existing sampling designs,
   1. Lastly, an example of a redefinition of management boundaries is ***Sand eel,*** where it was historically managed as 4 stocks, 3 stocks, and finally managed as 7 stocks. In this situation, Sand eel exhibit highly sedentary life-histories and are fairly reproductively isolated and allowing for area-specific TACs as opposed to lumping them all together will allow for utilization of weaker-stocks and reduced probability of foregone yield. However, if they were lumped into coarser units, there is a higher probability of depletion of weak stocks.
   2. Another example of a redefinition of management boundaries is ***red fish*,** where two distinct stocks were initially present, but a new stock was discovered (deep sea, pelagic, and deep sea pelagic). However, a key complication is that stock structure is not necessarily longitudinal or latitudinal, but was rather structured by depth, making adapting sampling designs more difficult. As such, to reconcile both biological and management complexity, stock boundaries were realigned to coincide with where stocks were commonly exploited, while maintain existing sampling designs. Thus, stock structure does not necessarily need to be defined directly with biological units and compromises will sometimes need to be made.

Importantly, these case-studies demonstrate that spatial stock assessments and all movement dynamics need not be incorporated into the management process, because robust harvest control rules and management procedures can help combat some of the problems with a single area stock assessment model. However, the above case-studies do highlight the need to continually review hypotheses regarding stock structure and to simulate test management procedures under various stock structure hypotheses, even when data are limited to conduct a spatial stock assessment.

While adequately representing spatial and stock structure is necessary for appropriate management advice, other considerations with respect to stock and spatial structure are also important to account for within fisheries. In particular, maintaining spatial and stock structure as well as sustainable levels of biomass should both be equally important within fisheries management. Namely, stock, and spatial structure generally arises from local adaption to certain environments (density-dependence to expand or contract, environmental differences driving different demographic rates), differences in fish behavior (homing due to imprinting, learned migrations), and oceanographic features (trade-off between dispersal and isolation: larval and migratory connectivity, isolation due to genetic stock structure and natal homing, physical barriers). The erosion of stock and spatial structure can potentially be a prelude to fishery collapse, which can result from fishing effort eroding population structure as well as responses to climate change. The maintenance of spatial stock structure is important for fish species because spatial heterogeneity can contribute to population resiliency by contributing to an increasingly diversified portfolio of life-history strategies. ***In the context of the member vagrant hypothesis,*** where individuals home to their natal origin to exploit favorable oceanographic features to retain larvae, while maintaining the same migration and seasonal pathways that have previously been favorable, a loss of these natal homing dynamics (spatial stock structure) can lead to an increased number of vagrants that deviate from these traditional pathways, which may experience lower survival rates. On a similar note, the ***entrainment theory*** posits that older members of the population lead juveniles into spawning grounds and allow migration pathways to be learned, which may help with survival and propagation of offspring. However, erosion of these spawning areas and older individuals from the population can subsequently erode spatial structure of the population, and may lead to increased population instability. These theories all contribute to a diversity of reproductive strategies within fish populations: imprinting and natal homing as well as larval retention capitalize upon long term stability via the member vagrant hypothesis, while straying dynamics allows for colonization of new habitats and may manifest as demographic rescue effects, as well as buffers against prevailing environmental conditions. As such, maintaining a diverse reproductive portfolio through a maintenance of spatial stock structure are linked, and contribute population stability.

As noted above, the erosion of spatial stock structure can be a prelude to fishery collapses, because these spatial stock structures contribute to heterogeneity within the population that act as a portfolio to buffer against adverse events that may occur. Spatial structure can be eroded because of fishery effects (i.e., fishing out all old fish, resulting in reductions in entrainment, and loss of known spawning habitats), human factors such as the implementation of dams but also the introduction of new species, and environmental effects such as climate change. With respect to fishery effects, overexploitation of a sub-stock can result in an immediate erosion of spatial structure, while in panmictic populations, removals of localized hotspots may result in an inability to recolonize these areas. As noted, a truncation of age-structure can result in an erosion of migratory pathways. Additionally, the depletion of certain components of the stock, can also lead to mismatches with environmental conditions that have historically been exploited by that specific component (i.e., removal of spring spawning stock that takes advantage of a spring bloom). Lastly, erosion of spatial structure can also result if oceanographic pathways break down as a result of climate change, altering the transport of larvae and eggs, and their subsequent survival. Some of the consequences of an erosion of spatial structure includes: 1) changes in trophic and food web (destabilization of predator-prey dynamics) linkages, 2) reduced genetic variability and local adaptation to certain spatial regions, thereby reducing the potential for portfolio effects, 3) reduction of the potential for demographic rescue effects, because there is a lack of individuals to exchange individuals with neighboring populations, 4) longer recovery times (related to the third point), and 5) reduced probability of learned migratory pathways. In general, erosion of spatial structure is more detrimental for long-lived species, because recovery of these dynamics can take a longer time, relative to short-lived species that can rebound quickly (i.e., herring).

There are several examples can case-studies of the consequences of eroding spatial stock structure and the implications of localized depletion. These include:

1. Pollock in Alaska, where a combination of environmental and fishery effects has eroded populations structure. In particular, overfishing of the Aleutian Basin stock has impeded its recovery, where a complete erosion of the stock has impeded the ability of neighboring stocks to promote demographic rescue effects. Additionally, in areas egg distributions and larvae distributions have contracted near Shelikof as well as in the Bering Sea, relative to historical distributions, which are attributed to reductions in spawning biomass and associated suitable habitats,
2. Scotian Shelf Cod, where following overfishing, the stock experienced an extreme contraction in both its distribution as well as the temporal timing of spawning (resulting in less buffering of YOY against prevailing conditions), and has also reduced recruitment synchrony among other regions, suggesting that the erosion of spatial stock structure resulted in reduced larval connectivity,
3. Norwegian Spring Spawning Herring, where at low abundances, their migratory pathways and feeding areas substantially changed, relative to historical knowledge.

Both erosion of spatial stock structure, identification of stock structure, and correctly aligning management boundaries with the biological unit require methods to identify how such boundaries should be aligned, as well as detect changes in distributions. There are a variety of methods that can be used to identify stock structure, all of which can tell us a different story. However, by combining all of these different methods, you can be fuller picture of what the underlying hypotheses for stock structure should be. Many stock ID methods rely on: 1) tagging and movement, 2) parasites from certain regions, 3) otolith shapes – where 2 and 3 encapsulate phenotypic and natural markers, 4) oceanographic features impacting larval dispersal, as well as IBM modelling, 5) life-history and differences in demographics as well as differences in fishery characteristics (size and age structure of different regions, differences in abundance trends), 6) stable isotopes, and 7) genomic approaches such as SNPs and microsatellites and resulting boundary analyses.

As such, good practice for stock ID is to reconcile differences among all of these approaches and utilize all available information to construct hypotheses. Additionally, stock ID methods should ensure that stock structure and boundaries re defined as closely to the biological population unit as possible, although some deviations can be made (stock structure should in general be defined to encapsulate the entire population such that no immigration or emigration occurs). It is important to stray away from the status-quo of using political boundaries to delineate stock structure. Various scenarios of stock structure and plausible hypotheses should also be constructed, as opposed to just one hypothesis to rule them all. Importantly, stock structure and its associated boundaries should not be static, but instead should be dynamic and continually reviewed and reconciled with new information. Populations that are isolated should be treated as independent, while populations that exhibit strong metapopulation structure should be assessed as its unit stock, while incorporating spatial structure within the assessment and management framework. However, populations that exhibit limited post larvae movement can potentially be assessed as a separate stock, if data are limited although there is value in sharing information among these stocks. Similarly, if stock structure cannot be delineated due to metapopulation structure, spatially implicit approaches can be used to implicitly account for spatial variation. Furthermore, even if stock structure is recognized, data limitations can impede the implementation of a spatially explicit assessment. Under such circumstances, fishing mortality should be set lower than the maximum allowable level to buffer against such uncertainty. Importantly, management procedures and assessment methods should be thoroughly simulation tested against hypotheses constructed about stock structure to ensure that they can provide robust management advice that achieves management objectives.

## Pacific Cod Stock Structure

Pacific cod are a wide-ranging species that reside from Japan to California and occur at depths from the shoreline to about 500m, although they tend to prefer depths of about 260m. Pacific cod can exhibit high movement rates and undertake both spawning and feeding migrations, where they exhibit characteristics of natal homing. Spawning tends to take place around January – April, where they have been observed to move quickly away from spawning grounds following April, into feeding grounds. However, some tagging studies have shown that a partial migration strategy is also plausible in Pacific cod, where certain components of the population may remain in spawning areas and may not undertake feeding migrations (although the representativeness of these studies are questionable; Bryan *et al.*, 2021). Nonetheless, while it is well known that Pacific cod exhibit site fidelity to spawning grounds, it is unclear whether they also exhibit site fidelity towards feeding grounds. Spawning tends to take place at depths of 200m, ***where eggs are semi-adhesive and remain near the bottom for about a month.*** Furthermore, their larval duration can be up to 90 days. ***As such, there is likely a fairly long period of imprinting for eggs and larvae from which these individuals were spawned in – individual based models have shown that the early stages of Pacific cod likely do not disperse far from their natal areas and are often retained in their natal areas, attributed to the semi-adhesive nature of their eggs, which limits larval dispersal*** (Hinckley *et al.*, 2019).

In addition to their general life-history with respect to spawning, movements, depth preferences, etc, differences in prey availability may also serve to drive stock structure, wherein individuals are locally adapted to feed on specific prey items, unique to their respective regions. Pacific cod are generalist predators that feed on a variety of prey items. For Pacific cod residing in the Bering Sea region, individuals are typically found feeding on walleye pollock and crab, whereas in the Aleutian Islands, Pacific cod are typically found to be feeding on Atka mackerel, sculpins, squid, and myctophids (Spies, 2012; Spies *et al.*, 2022). Thus, individuals residing in these areas may be locally adapted to feed and take advantage of the available species and may exhibit patterns of natal homing in order to remain locally adapted to these regions. In particular, this local adaption (foraging, predator avoidance, orientation, and social behavior) could be driven by a transitional zone in Samalga Pass, which divides warm nutrient poor water from the ACC and the cold nutrient rich Alaska stream to the west, where there are differences in water clarity driven by differences in levels of chlorophyll and other organic matter from the ACC (Spies *et al.*, 2022).

Currently, there are several methods that have been commonly used for identifying and delineating stock structure in Pacific cod. Most methods that have been commonly employed to infer stock structure has involved the use of genetics (Cunningham *et al.*, 2009; Spies, 2012; Drinan *et al.*, 2018; Spies *et al.*, 2021, 2022), although some work has also inferred potential differences in stock structure through the use of individual based models coupled with hydrodynamic modelling (Hinckley *et al.*, 2019), as well as through the use of movement and tagging analyses (Shimada and Kimura, 1994; Bryan *et al.*, 2021). Genetic studies typically use neutral genetic markers to detect differences in population structure, although some population assignment studies have also coupled typically genetic markers with outlier loci to accurately assign population origin and detect stock structure. Such approaches are useful because population assignment can be difficult for stocks with typical IBD patterns (Drinan *et al.*, 2018). Although the use of natural markers such as parasite composition and otolith stable isotopes to delineate stock structure for Pacific cod in the west coast (British Columbia to Washington) has previously been done, such methods have seldom been explored for Pacific cod in Alaska (references in Cunningham *et al.*, 2009).

With respect to movement dynamics of Pacific cod, it is well understood that feeding and spawning migrations are undertaken as discussed above, and that they can exhibit high movement rates, as well as move across long distances. Furthermore, some limited pop-up satellite tagging studies have found that a limited number of Pacific cod individuals can swim across deep passes in the Aleutian Islands, which have previously been speculated to be a barrier to movement, given their depth preference of about 250m, although this may be the exception rather than the rule, given the limited number of tags these conclusions were inferred from (Bryan *et al.*, 2021). Nonetheless, it does suggest that Pacific cod are capable of moving through deep passes, although the energetic demands for doing so, may not outweigh the benefits. Based on a fairly extensive floy tag study, Shimada and Kimura 1994 found that Pacific cod tend to exhibit movement dynamics that are more random during the spring and summer months, which are more reflective of opportunistic foraging behavior, while in the winter months, movement patterns were more directed towards both spawning and feeding grounds. Furthermore, Pacific cod appear to demonstrate seasonal migration patterns where individuals move more towards the Unimak region in the winter for spawning, move into inner shelf primarily during the summer for feeding, as well as the outer shelf (seasonal cross shelf movements), and move into the outer shelf during the fall, prior to returning to Unimak and the Pribilofs (Shimada and Kimura, 1994). Furthermore, fish that are tagged in the eastern Bering Sea appear to be able to move into the western GOA, where some released individuals in the EBS are caught near Shumigan Islands. Similarly, fish tagged in the western Aleutian Islands have also been recaptured in the Bering Sea, suggesting some degree of connectivity between the EBS, WGOA, and AI regions (Shimada and Kimura, 1994; Bryan *et al.*, 2021). Given the sensitivity of Pacific cod to water temperatures, seasonal movements are likely somewhat attributed to changes in water temperature. In particular, seasonal movements away from the inner and outer shelf during winter months are likely due to Pacific cod avoiding cooler water temperatures from the formation of sea ice.

Although it is possible for stock structure to arise due to life-history, movement dynamics, and oceanographic features, it is also possible that part of the current stock structure observed in Pacific cod are attributed in part to historical glaciation events, in relation to founder effects. Specifically, during the last glacial maxima period, global sea levels dropped to about 120m, which likely forced marine species to seek refugia in the south, which was free of sea-ice, and in the northern areas. These glacial events likely drove some form of stock structure, which are reflected in contemporary stock structure, where there is detectable genetic divergence between the Northwest (Japan, Russia, and Korea) and Northeast Pacific stocks (Washington, GOA, and BSAI) of Pacific cod (Canino *et al.*, 2010) and further differences within the northwest and northeast Pacific cod stocks. Such divergence was inferred based upon differences in haplotype frequencies between the northeast and northwest stocks, as well as discontinues in neutral genetic markers that are not under selection and is likely attributed to the presence of two glacial refugia: 1) northern offshore areas that were ice-free, which northwest Pacific cod stocks likely colonized, and 2) southern refugia that were ice-free, which northeast Pacific cod stocks likely colonized. The colonization of the northwest likely occurred first, followed by the colonization to the northeast, and then into the Georgia basin. In the Northwest Pacific, Canino et al 2010 detected differences between the Japanese and Korean stocks, where the subdivision between the two stocks was proposed to be due to an ice-age cooling period, resulting in a formation of a sill in between these two regions (Cunningham *et al.*, 2009; Drinan *et al.*, 2018), limiting gene flow. This is also partially observed in differences in ZP3 haplotypes (Spies *et al.*, 2021). In the Northeast Pacific, subdivisions between the greater Northeast Pacific and stocks residing in the Georgia Basin (Washington) was proposed to be due to a staggered period in between colonizing the GOA and the Georgia Basin. In particular, the colonization of the GOA and BSAI occurred first as a single panmictic population, because glacial recession towards the north receded faster than the glaciers in the south. This is supported by an isolation-by-distance pattern in the coastal northeast Pacific cod stocks. Furthermore, distinctions between the northeast Pacific stocks and the Georgia basin stocks are likely due to long-term allopatric speciation, where the presence of a spatial barrier likely restricts movements and gene flow between these two stocks.

As described above, genetic differentiation between the northeast and northwest Pacific cod stocks were likely due to the structuring of glacial refugia between these two regions, which eventually resulted in stock structuring. However, the northeast Pacific, which is better studied and is the focus of my studies, further exhibits more fine-scale genetic differentiation in the form of isolation-by-distance patterns, which is supported by studies of the use of microsatellites to differentiate stock structure (Cunningham *et al.*, 2009). IBD patterns are common in many marine species and indicate that as geographic distance between samples gets further and further, there is greater genetic divergence, where the dispersal of individuals in a continuous distribution is limited (i.e., through natal homing or larval retention). The strong IBD pattern found in Pacific cod is strongly supported by several studies and has resulted in the identification of several unique stocks in the northeast Pacific. However, this stock also has some departures from IBD, particularly for localized inshore components of the stock (fjord environments), likely a result of spatial and geographic barriers for dispersal.

Firstly, there is strong stock structure between coastal North Pacific and Alaska stocks, and the Georgia Basin (Puget Sound + Strait of Georgia), which were colonized after the recolonization of the GOA (Cunningham *et al.*, 2009; Drinan *et al.*, 2018). With respect to stock structuring between the stocks in the Georgia Basin, this is likely attributed to physical-geographical barriers (the presence of narrows and sills) that limit both adult and larval dispersal between complex fjord systems and the west coast of America (Cunningham *et al.*, 2009). Furthermore, there is also stock structure between stocks located in coastal Washington and BC (i.e., Hecate strait, Strait of Juan De Fuca, Coastal WA), with those in Alaska (Drinan *et al.*, 2018; Spies *et al.*, 2021). Such structuring is likely attributed to a combination of environmental factors as well as life-history (egg and larval retention and natal homing). Based on the limited ability for population assignment studies to correctly assign the coastal west coast stocks (combining WA and BC), it is likely that there is some degree of mixing between these regions, and that they are their own separate stock (Drinan *et al.*, 2018)

With respect to stocks that are located in Alaska, there is stock structure between the Bering Sea, Aleutian Islands, Western Gulf of Alaska, and Eastern Gulf of Alaska. There is likely a combination of environmental, life-history, and oceanographic features at play. In particular, there is a well-known transitional zone at Samalga Pass that exhibits differences in both productivity and species composition that likely acts as a barrier (either geographic, life-history, feeding, adaptation) between the eastern Bering Sea and the Aleutian Islands (Amutka Pass also acts as a potential barrier; Spies, 2012) and species may be locally adapted to these differences for both spawning and feeding purposes. Furthermore, oceanographic circulation and dynamics in the Aleutian Islands are way more complex (volcanic islands with substantial water exchange) than those in the Bering Sea (single continental shelf), which could limit both larval and adult dispersal. Specifically, the presence of deep trenches and passes along the Aleutian Islands may limit the dispersal of adults, given their preference for depths of about 260m. Strong tidal currents are also present in these regions, which can act to entrain and retain both larvae and eggs within the AI region. This stock structure is further supported by studies that have used RADSeq for population assignment as well as SNPs, where they were able to adequately assign individuals from the AI and EBS back to their population of origin (Drinan *et al.*, 2018). Additionally, sea-ice dynamics in the Bering Sea could also be acting selectivity on the genome of Pacific cod, which may further drive local adaption and stock structure in these stocks. In particular, the Bering Sea is characterized by the cold pool and sea-ice dynamics, both of which are absent in the Aleutian Islands; these selective forces may be acting upon the ZP3 gene which regulates glycoproteins and antifreeze proteins for the embryo (Spies *et al.*, 2021). There is some evidence of differences in haplotype frequencies for the ZP3 gene between the Bering Sea and Aleutian Islands regions. Consequently, local adaptations to currents, passes, velocities and light levels likely result in stock structure in the Aleutian Islands, while sea-ice dynamics and the cold pool likely led to stock structure in the Bering Sea.

On a related note to local adaptations, islands of differentiation and the loci where differentiation occurs also provides information on differences in population stock structure as well as the potential for local adaptation. For Pacific cod, the use of Pool-Seq detected several islands of differentiation between the EBS, AI, and Washington stocks, which were detected arounds regions associated with vision-related genes (Spies *et al.*, 2022). Differentiation was found to be correlated/related to differences in salinity, velocity, chlorophyll, followed by temperature. Nonetheless, it is likely that differentiation due to vision-related genes could be due to differences in light penetration, which is lower in the EBS than in the AI, and suggests that these vision-related genes combined with the ZP3 gene may act to select for certain individuals, resulting in stock structure between these regions (i.e., darker in EBS because of sea-ice, and the need to be locally adapted). These islands of differentiation were also found between the EBS and Washington stocks, suggesting that they are fairly important across the range of Pacific cod.

While there is well established stock structure between the Aleutian Islands and the Bering Sea, some studies have suggested that there is even more fine-scale stock structure within each of these larger regions, where in the Bering Sea, Unimak Pass and the Pribilof Islands may be separate stocks, while in the Aleutian Islands, western Aleutian Islands (west of Amchitka Pass) and eastern Aleutian Islands may also be separate stocks, although results were somewhat variable and suffers low sample sizes (Spies, 2012). For the Bering Sea, sub-structuring of the stock between Unimak and the Pribilofs can potentially be attributed to: 1) currents, where different current transport systems could disperse larvae elsewhere, 2) difference in spawn timing, attributed to sea-ice dynamics (later spawning in the Pribilofs), and 3) natal homing. For the Aleutian Islands, sub-structuring the stock between the western and eastern Aleutian Islands can be attributed to: 1) differences in depth, 2) differences in velocity, and 3) differences in currents, driving stock structure (Spies, 2012; Spies *et al.*, 2020). However, the sub-structuring for stocks in the Bering Sea and Aleutian Islands requires further research, although it is quite apparent that the EBS and AI are separate stocks (Spies, 2012).

Within the Bering Sea region, recent changes in the locations in which Pacific cod are caught have resulted in some speculation of sub-stock structure in the Bering Sea. However, studies have shown that individuals caught in the Bering Sea are a single stock, without any stock-structure between the eastern and northern Bering Sea. Their recent occurrence in these regions are likely due to changes in distribution in response to climate warming, rather than a genetically distinct stock between the EBS and NBS (Spies *et al.*, 2020). This is evidenced by the fact that individuals caught in the NBS had loci that clustered from various locations in the EBS and also because individuals caught in the NBS had higher genetic diversity than those individually considered in the EBS, because it was a combination of 3 areas – if the NBS was established due to founder effects, you would expect reduced genetic variability, because they represent a smaller group of individuals straying from their population of origin. Thus, their recent expansion into the NBS is likely due to a feeding migration. While that is true, if spawning habitats in the NBS are no longer covered by sea-ice, it is possible that Pacific cod could move into and colonize these regions, which could be detrimental to the trawl fishery, because they are not allowed to trawl in the NBS.

Outside of the EBS and AI regions, there is likely stock structure between the EGOA and WGOA. In particular, population assignment studies have found that differences between stocks located off of Kodiak+Shumigan Island (WGOA) and Prince William Sound (EGOA and state fishery) (Drinan *et al.*, 2018). Outside of population assignment studies that have used outlier loci for assignment purposes, studies looking at the ZP3 outlier loci in the context of directional selection have similarly found stock structure and differences in haplotype frequencies for this loci that may perhaps explain stock structure between the EGOA, WGOA, EBS, AI, and the lower 48 (Spies *et al.*, 2020, 2021). Specifically, the ZP3 loci is related to processes with reproduction and sexual development, such as the production of antifreeze proteins and proteins to protect the embryo, which is central for reproductive isolation in many stocks. Based on differences in haplotype frequencies of the ZP3 gene, Spies et al 2021 found that haplotypes in the Kodiak and Shumigan Islands (WGOA) were very different from those from the PWS (EGOA), which resembled more of the Eastern Pacific stocks (PWS – Washington Coast) and showed almost no shared haplotypes. By contrast, some haplotype frequencies were shared between the WGOA, BS, and AI regions, suggesting that some degree of connectivity and mixing may occur, as well as similar selective processes acting upon these populations. In particular, these selective processes may be due to the narrower thermal tolerance of eggs for individuals located west of Kodiak Island; following a marine heatwave, EGOA stock remained fairly stable but low, but the CGOA and WGOA portion of the stock continued to decline, potentially attributed to their ZP3 gene being more adapted to cooler and narrower temperatures (3-6 C), and continuing adverse temperature conditions that they were unable to adapt to (Spies *et al.*, 2021). Such differences between these two regions are likely due to differences in oceanographic dynamics between the EGOA and WGOA. Specifically, IBM larval connectivity studies have shown that the PWS (fish from upstream of PWS to southeast AK) and Kodiak/Shumigan Island regions serve as collection regions, where collection in the PWS regions from SEAK are likely due to a narrow continental shelf in SEAK and the EGOA, and impingement of the Alaska current as well as period oceanic eddies that occur in that region (i.e., Yakutat, Sitka, Haida eddies) (Hinckley *et al.*, 2019), as well as a transition in this region from the wide and slow moving Alaska Current into the narrow and fast Alaska Stream.

For stocks like Pacific cod that exhibit high IBD patterns, as well as natal homing dynamics, it is imperative that management fully considers the use of genetics in delineating stock structure, and managing separate stock units, coupled with research-oriented spatial models that have flexible parameterizations of movement dynamics. For Pacific cod that undertake such long ranging seasonal and spawning migrations with the potential of feeding migrations to intermix with other stocks, it is important to be able to assign catches or attribute catches correctly to certain areas, or to be able to model spatial models on a finer time-step. However, such approaches are not always feasible given limitations in the available data. In general, if stocks exhibit low migratory rates between areas, it is best to manage them as separate stocks, although as noted above, data limitations may limit the management of separate stocks, particularly if there are strong but small areas of stock structure. ***In cases where individual area management is not possible, catch-cascading can be a valuable alternative where a combined assessment is conducted, but is apportioned based upon the relative biomass in each area. Additionally, managing high fishing effort areas separate from low fishing effort areas can also potentially be useful in reducing stock overexploitation***, likely because if you have high fishing mortality in one region, but low fishing in another region, but are managed as a single unit, the regions with lower F will disproportionately contribute to the biomass in regions with higher F, and you will think you are able to harvest those areas harder than you should (i.e., localized depletion and the issue of applying a higher fishing mortality rate to regions with lower biomass). By contrast, splitting out to high and low F areas will allow you to adjust F according to the population size in the high F region and adequately audit the effects of fishing to those areas. This is particularly of importance if stocks exhibit limited movement (Spies *et al.*, 2015).

Under an IBD pattern, it is difficult to figure out how management should be structured because high gene flow can result in a difficult determination of population structure. When there is population structure that is not addressed or separately modelled, the areas closest to the port result in a high probability of overfishing due to an increased F, while those away from port did not suffer in terms of catches, likely because they were only lightly fished (Spies and Punt, 2015). Furthermore, ignoring genetic stock structure and managing as a single population can result in overfishing in certain areas that inhibit any sort of recovery. By contrast, if stock structure is detected after x number of years, separately managing those areas can help facilitate some form of population recovery as well as increase harvester revenue, because of higher stock status and higher population sizes. For a species like Pacific cod that exhibit high IBD with low dispersal, the benefits of demographic rescue effects are negligible, and thus, the benefits of demographic rescue effects cannot be seen if one area is severely depleted and area-specific biomasses or some proxy of that will need to be considered in order to avoid localized depletion (i.e., demographic rescue effects are only helpful in high dispersal rate species).

# Franz Mueter

## General GOA Oceanography

The Gulf of Alaska is a complex system dominated by strong storms that vary on both monthly and decadal time scales, with a variety of complex oceanographic features that drive the productivity of the system. The GOA makes up about 13% of the US’s continental shelf (370,000 km2), with the eastern part of the GOA continental shelf being narrower (i.e., from southeast to PWS) and the western part of the GOA (i.e., Koidak to Shumigan Islands) being wider. The complexity of the GOA is generally encapsulated by numerous coastal mountains that have complex interactions with the atmosphere, deep shelves, troughs and banks that mediate water velocity, flow, and direction, coupled with a complex coastline. Such complexity has been shaped by both past and present tectonic uplift as well as glacial processes. While the GOA is extremely productive, it is somewhat paradoxical given how the currents move, such that the system is a downwelling dominated system, resulting in nutrient poor (nitrate poor) freshwater water getting down-welled, with limited upwelling through the movement of currents. The downwelling processes are caused by downwelling favorable winds that blow parallel to shore, resulting in more nutrients being transported to deeper waters, with less nutrient rich cold water being upwelled, except for when downwelling winds relax during the summer. However, like many systems, its productivity is highly dependent on the spring phytoplankton bloom and subsequent zooplankton biomass, which is likely influenced by climate variability and atmospheric induced variations in the system. These spring blooms are important components of the GOA food web because they result in subsequent increases in zooplankton biomass (calanoid copepods and euphausiids), which help with production in higher trophic levels.

The GOA is characterized by several dominant circulation and current systems. Currents are cyclonic in the northern hemipshere, and thus, operate in a counterclockwise fashion. Firstly, the Alaska Coastal Current is a prominent feature along the inner shelf and continental slope and is a freshwater dominated current that is mediated predominately by wind driven processes, as well as freshwater discharge, and is driven by highly seasonal dynamics (in the winter, runoff decreases, reduces salinity at depth with vertical mixing and downwelling; in the summer, salinity in the bottom increases due to upwelling and on-shelf flux of slope water) . It starts from the shelf of British Columbia and bifurcates around the Kennedy Stevenson Entrance (tip of Kodiak Island) where a majority of it goes down Shelikof Strait and into Unimak and Samalga Pass, with some of it joining up with the Alaska Stream as it bifurcates. It is a nutrient poor current due to its reliance on freshwater discharge and is bounded from offshore currents due to a salinity gradient. Generally, the current is confined along the inner shelf and continental shelf due to downwelling favorable winds (freshwater would pool at the source if no downwelling winds were present), and retains a lot of the freshwater discharge along the coast, which can result in a baroclinic structure in the ACC (non-parallel water density and pressure gradients). The baroclinic structure seems to be exemplified when there is high precipitation and runoff, while transport of the current is greatest in the winter due to wind-driven processes (higher winds equate to increased transport in the winter). The ACC is less prominent in the fall likely due to the influence of freshwater discharge. For a majority of species, the ACC is the dominant mode of transport for dispersal. The ACC is characterized by both organized and unorganized features depending on where the current is located (i.e., Seward Line is unorganized), which can help aid in across shelf transport of both larvae as well as nutrients.

In addition to the ACC, the Alaska Current is also a dominant feature within the GOA system (it is the eastern boundary current), which develops by bifurcating from the North Pacific Current (from the West Wind Drift) and is a wide slow-moving current. This current is generally mediated by wind stress curl and advects warmer waters from lower latitudes into the northern regions. Furthermore, the current is characterized by an abundance of eddies as well as meanders. It generally follows the isobaths via bathymetric steering. At the head of the GOA (about the head of Kodiak and tail of Kenai), the Alaska Current turns into a narrow swift-moving current called the Alaska Stream along the continental slope that travels down through the Aleutian Islands before it rejoins the North Pacific Current. The sharp change between the Alaska Current and Alaska stream is characterized by a sharp change in water velocity as well as thermohaline gradients (temperature and salinity). Similar to the Alaska Current, eddies can also develop on the Alaska Stream and are not uncommon – in particular, eddies commonly form near Kodiak Island, which can help with cross-shelf fluxes and transport of nutrients to the surface. Given the dominance of both the Alaska Current and Alaska Stream along the eastern and western GOA, respectively, flow along the shelf break is more organized than flow along the inner shelf and slope, which is dominated by the Alaska Coastal Current.

In terms of the thermohaline characteristics of the GOA, the shelf maintains fairly low temperatures at about 7C although temperature gradients are pretty weak. As such, given the high freshwater discharge from glaciers and the ACC (adds iron and silica acid into the ACC), thermohaline gradients in the GOA are predominately driven by freshwater/salinity gradients in mediating both horizontal and vertical density gradients. As expected, freshwater discharge tends to be highest in the summer, which tends to result in high levels of stratification due to the difference in densities and salinities between freshwater and saltwater during this period. In addition to the freshwater discharge resulting from inland freshwater inputs (i.e., glaciers), freshwater inputs enter into the GOA through the presence of coastal mountains where cyclonic storms linger and spin down near the coastal mountains (typically during fall and spring) and intensify winds, resulting in increased and enhanced precipitation near coastal GOA, eventually driving some of the baroclinic structure seen in the ACC. These storms and coastal runoff results in downwelling favorable winds as well. Given that the Alaska Current and Stream bring in water from the North Pacific Basin that are generally colder and more nutrient rich, the outer shelf, inner slope, and shelf break are generally dominated by saltier nutrient rich waters, while the inner shelf, which is generally more mediated by the ACC is dominated by fresher nutrient poor warm waters (although high in iron and silica acid due to glacial processes and mountains). The mid-shelf characteristics tend to be an intermix between nutrient rich waters and nutrient poor waters.

As noted above, processes governing stratification are largely due to gradients in salinity as a result of freshwater discharge and precipitation. In particular, spring water temperatures are fairly uniform, and thus, gradients are formed through salinity differences. Stratification tends to form near freshwater discharge areas which then get propagated to offshore areas through cross-shelf transport/circulation and vertical mixing. However, salinity is not the only thing that causes stratification. In particular, in the absence of freshwater inputs, increased solar radiation as well as decreasing wind speeds can similarly lead to stratification in the middle and out shelf regions. During the winter, the increased frequency of cooling and storms reduces stratification and increases the mixed layer depth, which can entrain nutrients to the surface layers. This then entrains phytoplankton to the bottom with increased cloud cover, which limits phytoplankton production. During the spring, changes in temperature and salinity result in stratification (i.e., solar heating and vertical wind-mixing), as well as a weakening of winds, which when coupled with increased radiation and reduced cloud cover, can lead to the onset of a phytoplankton bloom. However, note that during the spring bloom, stratification from the ACC and other regions are generally weak, because there is still limited runoff and precipitation during this period. Furthermore, given the highly seasonal nature of the GOA, variability in the timing of the spring bloom can be linked to cloud cover and the intensity of stratification. During the summer, stratification is likely strong given increased solar heating as well as freshwater discharge, however, at this point, nutrients have been depleted from the spring bloom, but are characterized generally with relatively high chlorophyll levels. An important component of the spring bloom is subsequent production of both primary and secondary producers, where if zooplankton arrive prior to the spring bloom, growth will be slow and mortality will be high, which can have impacts on higher trophic levels.

In terms of zooplankton and phytoplankton dynamics, it has shown that phytoplankton blooms can vary greatly over small spatial scales, and largest differences appear to be in the mid-shelf regions, suggesting that it is a boundary likely attributed to differences in the availability of iron between the outer and inner-shelf. In the ACC, phytoplankton blooms in the spring may occur earlier than other regions because phytoplankton are entrained in the surface layer of the mixed-layer depth and are exposed to higher solar irradiance relative to plankton on the shelf, although this might not be a generality. In general, the following dynamics take place: 1) phytoplankton blooms occurs, 2) microzooplankton graze on phytoplankton, and 3) macrozooplankton feed on microzooplankton. Zooplankton that develop in offshore areas can be advected via onshore surface Ekman transport onto the inner-shelf that serve as a prey base for juvenile fish residing in these areas and can be transported throughout the GOA via the ACC.

Obviously, the production of phytoplankton and zooplankton in this system will be dependent on the dynamics of nutrients. In general, nutrient rich salty waters flow onshore when downwelling winds relax, which typically occurs during the summer. In addition to weak upwelling that occurs during the summer, other processes can also result in upwelling and increase the input of nutrients into the euphotic zone. Exchanges via the edges of eddies, surface onshore Ekman transport (convergence of nitrate rich basin water and iron rich freshwater from the coast), canyons intersecting with shelf breaks that transport/advect nutrients to the surface, as well as topographically induced upwelling are all processes that can upwell nutrients to the surface and contribute to the overall productivity of the GOA. Additionally, nutrients from the surface layer can be supplied from more productive offshore regions to inner-shelf regions via winds and cross-shelf transport mechanisms. Specifically, exchange between coastal waters with high chlorophyll levels, with central GOA waters rich in macro-nutrients likely occur through swirls, eddies, and meanders. Other mechanisms for nutrient transport in the GOA include tidal modulation near the Kennedy-Stevenson entrance (also probably southeast of Cook Inlet) to vertically mix the water column to bring nutrients to the surface, as well as estuarine inflow in Shelikof Strait (intrusion of ACC slope water with Shelikof Sea Valley), which can entrain nutrients in the upper layers. Additionally, winds may also deposit iron throughout the land, and strong currents may resuspend sediments and bring iron to the surface. Barrier jets may also enhance the strength of along shore winds, which can lead to positive wind stress curl and result in local upwelling.

With respect to differences in productivity across the EGOA, CGOA, and WGOA, it appears that the WGOA supports a higher biomass of individuals. Such differences in productivity between these regions can be due to a variety of factors and tends to suggest that the WGOA supports higher availability of food resources and productivity in that region. In particular, primary productivity measured through differences in color are higher in the WGOA, which could be attributed to a higher frequency of upwelling, as well as interactions with the Alaska Stream with submarine canyons and strong tidal mixing and amplification. Furthermore, given that the Alaska Current turns into the Alaska Stream in this region, which is more narrow and fast moving, these stronger alongshore currents, advection, and nutrient inputs from major river sources (e.g., Kenai) may further contribute to the productivity of WGOA. Lastly, the shelf area in the EGOA is narrower compared to the CGOA and WGOA (C and WGOA could result in more habitat to support benthic productivity as well, increased sunlight exposure), which is believed to result in reduced nutrient enrichment.

Although there are many localized features within the GOA that govern the productivity and characteristics of this region, other large-scale dynamics such as atmospheric forcing and wind-drive processes can similarly exert a large influence on the productivity of the system. As discussed above, the GOA is a system dominated by cyclonic storms, which lead to increased precipitation and drive the baroclinic structure in coastal regions. During fall through spring, storms are cyclonic and facilities downwelling along the coast and increases the mixed layer depth during this period, but results in increased upwelling in the central GOA. From summer through fall, winds are more variable although they sometimes weaken. Anticyclonic winds are also sometimes observed during this period which promotes upwelling along the coast.

Regional atmospheric variability is predominately driven by the Pacific North American in the winter, which relates to the strength and position of the Aleutian Low, where during an intense Aleutian Low, storms are generally stronger, while during a weak Aleutian Low, storms are weaker and result in calmer conditions (less storms). On the scale of 3 – 7 years, the ENSO is a key contributor to climate variability, where El Nino events tend to lead to positive temperature anomalies and downwelling favorable winds (warmer and less saline water), while La Nina events tend to lead to negative temperature anomalies and upwelling favorable winds (cooler more saline water). On the decadal scale, the PDO is a dominant contributor to variability within the GOA and is the leading mode of North Pacific sea surface temperature variability, which is further coupled with changes in sea level pressure. In general, parts of the PDO resemble the ENSO, where positive PDO is generally fairly similar to a El Nino, whereas a negative PDO is similar to a La Nina event. During the summer, positive PDO events tend to lead to higher water temperatures and increased runoff, while low runoff years tend to be associated with cooler temperatures and a negative PDO phase. Variability during the winter is caused by the Aleutian Low (dominant SLP features in the Aleutians) and the Pacific North American (air pressure patterns with north and south dipoles). By contrast, variability during the summer is thought to be a function of cloud cover, solar radiation, frequency of upwelling events and storms.

As noted above, the PDO is a dominant driver of climate and oceanographic variability within the GOA. The PDO is a special case of the PDV (PDV = El Nino patterns + TPDV – variability in the North Pacific of SST anomalies), which operates on a longer time scale. The PDO contrasts the ENSO because it occurs more on time scales of 20 – 30 (also 50 – 70) years, and is characterized by a horseshoe shape, while the ENSO is more on the scale of 3 – 7 years, characterized by a band near the equator. The PDO is made up of an unrotated EOF calculated with temperature anomalies relative to the global average. During warm PDO phases, cooler temperatures result in the central North Pacific, and warmer temperatures result along the coast (warmer and wetter in coastal AK), while during cold PDO phases, warmer temperatures result in the central North Pacific followed by cooler temperatures along the coast (cooler and drier in coastal AK). Additionally, during warm PDO phases, sea-level pressure is generally lower over the North Pacific, resulting in stronger counterclockwise winds. In general, the formation of the PDO is dependent on the Aleutian Low. In particular, during a strong intense Aleutian Low, counterclockwise winds increase in strength. In the eastern side of the US west coast, stronger winds result in warm and humid air being advected along the west coast of America into the PNW and Alaska. Furthermore, there is some relationship with the Kuroshio current, where the persistence of a strong Aleutian Low can result in atmospheric forcing, propagating this slow moving warm water current via Rossby Waves into the North Pacific region, manifesting as warm temperature anomalies (this relationship can form teleconnections and excite ENSO precursors). In addition, the presence of El Nino events can also result in positive PDO trends. In particular, the development of an El Nino event can manifest through NPO forcing, which results in lower SLP in the southern hemisphere, weakening trade winds, reduced evaporation, while increasing SST. As SST gets warmer, a positive feedback results and leads to continued weakening of the trade winds, thereby activating meridional modes, which propagate warmer SST into the equator and forms an El Nino event. Once an El Nino event develops, atmospheric bridges through Rossby Waves can form teleconnections to the extratropic regions and inject extra variance into the North Pacific Basin through the Aleutian Low, which then drive the following year’s PDO pattern (positive PDO).

In terms of production of groundfish and species in response to different PDO phases, warmer PDO appears to be associated with higher productivity with increased recruitment, as well as changes in species assemblages, where the inverse is posited for PDO effects in the lower 48 (i.e., lower productivity during warm PDO events). Increased productivity followed warm PDO events for some stocks can be due to a variety of factors. However, I will primarily discuss the role of eddies here (circular currents of waters than move) because some work has discussed the role of ENSO and warm PDO events in altering and destabilizing the Alaska Current, resulting in the formation of more eddies. Additionally, the development of eddies has been shown to be related to changes in interannual surface winds, downwelling favorable coastal winds, which are similarly associated with positive PDO, El Nino and intense Aleutian Low events. Eddies typically develop near southern and eastern GOA and stay near the shelf break regions and can influence cross shelf exchange by trapping coastal water in their interior while interacting with waters along the shelf and slope. Their formation can differ depending on their location but in general, their formation is a result of baroclinic instability in currents and interactions with the bathymetric along the continental slope, as well as the agglomeration of smaller eddies caused by the discharge of warm freshwater. This interaction can aid in distributing phytoplankton, zooplankton, and fish larvae from the slope to the shelf (i.e., cross-shelf exchange) as well as along eastern GOA and can help distribute individuals into more favorable feeding areas. By contrast, in the absence of meso-scale eddies, it is likely that shelf-spawned individuals will be advected into unfavorable feeding habitats. In general, there are two main eddy formation regions, which include Yakutat and Sitka, where both eddies will eventually end up interacting with the Alaska Stream and can break off into two separate eddies over time (ends up following the Alaska Stream and interacting with water in the Aleutian Islands). There are also some more variable eddy formation regions in the EGOA, which are typically measured/found in Kodiak, and often form on the shelf break of the EGOA (attributed to EGOA because of similarities in water properties of the eddy to EGOA). Given that the formation of eddies takes place at fairly distinct regions, the water properties of these eddies are also fairly distinct and contain a distinct freshwater core, indicating their origin. This is important because eddies will contain anomalous heat, salinity, nitrate, and nutrients from their formation region down the Alaska Stream, and can provide a mechanism for cross-shelf transport of nutrients as well as larval dispersal. Furthermore, eddies can transport long filaments of chlorophyll from the shelf into the gulf, which can provide for rich foraging areas as well as increase nutrient concentrations along the edges of the eddy through remineralization of nutrients from phytoplankton within the eddy.

With respect to specific eddy formation regions and its biological characteristics, the characteristics of Haida and Sitka eddies are more similar, likely attributed to their proximity. In addition, Haida eddies tend to have higher species diversity likely because it represents the southernmost eddy. In general, the oceanographic features of eddies group into three categories: 1) outside of eddies, 2) eddy edges, and 3) eddy centers, suggesting that the water properties in these regions are distinct from each other, and cluster together (depth, temperature, and salinity).

Additionally, species richness tends to be highest in the eddy center, likely due to their initial formation where water parcels get sequester and larvae and eggs get entrained. Species richness tends to decline as distance gets further from the eddy center. Furthermore, the species composition along the shelf-break greatly determines the composition of the eddy, likely because that is where water gets sequestered and where the eddies primarily form. Thus, species assemblages as well as nutrient composition of a given eddy will depend on the characteristics of the region in which they were formed in. Comparing newly formed and old eddies, species assemblages differ between surrounding slope and basin water, which can be due to ichthyoplankton developing swimming capabilities and exiting eddies, as well as the ability for eddies to still carry and transport anomalous heat, nutrients, and species through continual interaction with surrounding waters across its lifespan. As such, species assemblages along a given eddy will likely differ as the eddy ages, either from the development of swimming capabilities from larvae, or entrainment of nutrients and species within the eddy as it ages.

The transport pathways of fish from the slope to the shelf is extremely important, given that shelf waters are usually more productive and contain more suitable feeding habitats. ***In the realm of eggs, larvae, and spawning behavior, some of the most critical decision points is: 1) ensuring that eggs survive immediately after spawning, and 2) anticipating survival in later life-stages such that larvae and juveniles are well positioned to survive after hatching.*** For slope and offshore spawning individuals, larvae face a challenge in that suitable habitat for nursery and feeding are seldom availability in directly adjacent waters. ***Thus, species often need to spawn in a particular manner, season, location (proximity to nursery grounds), conditions to ensure that larvae can take advantage of the underlying oceanographic processes to reach their nursery grounds.*** There are several ways larvae can transport themselves onto the shelf. For example, larvae can utilize selective tidal stream transport, where they utilize environmental cues to control their vertical distribution and take advantage of currents or countercurrents to flow into their preferred destinations. Other strategies include local adaptations to regional current dynamics (taking advantage of currents and maintain an extended larval duration), leveraging the predictability of seasonal events (spring bloom), and locating conspecifics that are located in regions with high food abundance, and low predator abundance. There are several methods in which larvae can take advantage of the oceanographic features in the GOA to facilitate cross-shelf transport, which include:

1. The narrow shelf of GOA, with a myriad of features due to glacial formation, and can lead to tidal amplifications, local patterns of both upwelling and dowelling, as well as geostrophic flow,
2. Bathymetric heterogeneity in the GOA can allow for the formation of eddies as well as other transport mechanisms for cross-shelf fluxes,
3. The formation of basin wide eddies as well as meanders in the current can propagate along the shelf edge and facilitate exchange between basin and shelf waters,
4. Weak upwelling with a relaxation of winds in the summer can facilitate upwelling as well as transport of nutrients and larvae,
5. Downwelling favorable winds to facilitate transport of larvae from surface to inshore,
6. Additionally, freshwater inputs into the ACC can drive baroclinic gradients and structure, which can in turn lead to the formation of eddies, ultimately aiding in transport of larvae. This baroclinic structure can be driven by differences in wind strength as well as temperature gradients, which can be related to variations in the Aleutian Low,
7. Interactions with troughs, submarine canyons, and banks can result in tidally generated internal waves, shelf-slope eddies, and current meanders that can also bring oceanic water up to the shelf,
8. Lastly, other mechanisms such as the development of swimming capabilities at a young age can importantly facilitate cross-shelf transport of larvae into inshore regions.

Some groundfish (flatfish to be specific) in the GOA take advantage of these different features. Firstly, larvae and eggs that are spawned offshore and, on the slope, generally need to move from the slope onto the shelf, get from the outer shelf to the mid-shelf, and then get from mid-shelf to inner-shelf regions. Within the GOA, this pathway can be facilitated by: 1) relaxation of downwelling winds in the summer, 2) formation of both large-scale basin eddies and meso-scale eddies along the shelf break, 3) bathymetric steering with submarine canyons that bring water up, well as strong tidal mixing and amplification, 4) episodic downwelling transporting waters from the surface inshore. Given that much of these features can be seasonal, it is imperative for spawning and the larval period to be timed in accordance to these events to facilitate onshore transport. For Rex Sole, ***who spawn on the slope and occupy a large area, spawning grounds tend to be more located on the exit region of submarine canyons, which may allow larvae to take advantage of deep water circulation patterns that transport slope water up to the shelf and surface***. Furthermore, ***spawning near canyons can allow eggs and larvae to take advantage of the complex bathymetry present in those regions***. In particular, such complexities can amplify tidal signals that ***generate vertical instabilities and allow for cross-isobath movement of water***. ***Interactions with the Alaska stream can also result in water flowing up canyons and associated transport mechanisms, facilitating cross-shelf transport.*** Such interactions typically occur in the winter, which also coincide with when some species spawn on the slope (Halibut and Arrowtooth flounder), and larvae may be transported to the coastal edge of the Alaska Stream and get transported onto the shelf. ***In addition to cross-shelf transport mechanisms (i.e., once transport onto the shelf has been successful), along-shelf transport mechanisms can also be important***, which often occur as a result of bathymetric steering. As such, the transport of larvae can follow distinct bathymetric isobaths that closely follow the topography of the GOA. For species that spawn later into the year (Dover and Rex Sole), they likely better exploit seasonal changes in transport to move from inshore regions to nursery areas. In particular, Dover and Rex Sole occur more in the mixed layer and may benefit from downwelling favorable winds in the autumn to advect them inshore. ***Lastly, both spawning and larval behavior will impact the dispersal of juveniles.*** In particular, species that spawn at deeper depths (i.e., Halibut and Arrowtooth) are more likely to ***retain eggs in deeper waters and are less likely to be impacted by currents, and thus, rely more on canyons to allow larvae to rapid rise and get pushed inshore***. By contrast, species that spawn shallower (Rex and Dover sole) ***have eggs rapidly float to the surface and are more impacted by currents, thus, they are more widespread and generally have longer planktonic periods to allow for facultative settling in certain areas.*** Given the complexity and heterogeneity of habitats in the GOA, recruitment processes and use of currents and oceanographic features in these regions are likely to be different from the relatively homogenous EBS – both passive and active transport mechanisms are likely to play a role (passive = slope to shelf, active = shelf to inshore due to complexity).

The distribution of groundfish and fish species are mediated by interactions with oceanographic features such as eddies and currents, as well as temperature tolerances. As such, changes in temperature regimes can reflect distributional shifts in longitude, latitude, and depth, and such changes are unlikely to be one-directional. Furthermore, interactions with temperature preferences for certain age- or size-classes can also lead to multi-dimensional responses to changes in temperature. From an oceanographic feature perspective, across the coast of the GOA and the west coast of America, mean bottom temperatures tend to increase from WGOA to southern US, as well as mean depth. Furthermore, temperature anomalies in the GOA tend to be larger relative to southern regions, likely because the GOA is a downwelling dominant system, with limited upwelling of cold nutrient rich water to buffer against warming waters (southern US has more upwelling), resulting in larger temperature fluctuations. Furthermore, shallower depths in the WGOA tend to be cooler than intermediate depths. This can potentially be due to stronger vertical mixing in upstream regions (Kennedy Entrance to Shelikof Strait) being transported downstream into the WGOA, resulting in differences in temperature and water properties. By contrast, other regions of the GOA and the US West Coast tend to have decreasing temperatures with depth. Such differences in both depth and temperature preferences may indicate that species might not respond uniformly to changes in temperatures, and instead, may respond in multi-dimensional ways. For example, warm years in certain regions may not reflect warm years in others. Additionally, given that shallower depths in WGOA can potentially be cooler than more intermediate depths, species may actually move towards shallower waters, as opposed to deeper, to find thermal refugia. Furthermore, species may not necessarily shift latitudinally, because thermal refugia may exist in deeper depths. As such, fish species will generally move towards cooler waters during warm years, but this can be in the form of longitudinal shifts, latitudinal shifts, or shifts in depth distributions. Considering that temperature anomalies are less punctuated in the US west coast, these fish species tend to move less during warm events, given buffering from upwelling, while species in the GOA tend to move more in search of thermal refugia. In the EGOA, fish species have also been observed to move southeast, potentially due to higher primary productivity in those regions during warm years. In addition, fish movements in response to temperature are likely species-specific – for species that inhabit deep waters (i.e., sablefish), movements are unlikely to be great during warm years in search of thermal refugia, because bottom temperatures are fairly unaffected by warming conditions. By contrast, species that occupy shallower depths will show more variable response, given that shallower layers tend to warm up more quickly. Given that fish responses differ based on the sub-regional scale, and exhibit differential responses to warm temperatures, there are likely different thermal ecoregions with respect to how fish movement responds to warming temperatures (i.e., changes in depth, latitudinal or longitudinal shifts). Oceanographic features such as gullies, troughs, canyons, depth, and topography will all mediate the response of fish species to warmer temperatures.

Using Pacific Halibut as an example, I will illustrate how larval dispersal, movement dynamics, fish behavior, and oceanographic features govern the stock-structure of Pacific Halibut. Pacific Halibut are wide-ranging and range from Norton Sound to Cape Mendocino – they are also found in both Russian and Japanese waters. There are generally distinct populations from the Aleutian Islands, GOA, and Bering Sea, with potential sub-stock structuring along Amchitka Pass for the AI region. The AI region is separated from the GOA through Samalga Pass, which is a known transition zone that divides nutrient poor water from the ACC and the cold nutrient rich water from the Alaska Stream to the west. Additionally, there is stock structure between the GOA, the western US, and BC (breakpoint near Haida Gwaii), which has been attributed to differences in salinity, temperature, pH gradients, parasite composition, habitat quality, and prey availability (all of which are driven by oceanographic features). For Pacific Halibut, the pelagic larval stage is primary mediated by oceanographic currents. Spawning of eggs for Pacific Halibut should in theory be in regions where survival and transport is of larvae is maximized. Spawning takes place in depths of about 300 – 600m on the outer shelf, with eggs inhabiting 100 – 400m, where they are less influenced by wind and surface currents. Larvae are typically found in depths of about 150 – 200m, and first feed at about 55 days post hatch, moving inshore to shallow waters at about 3 months. Juveniles tend to occupy the inshore regions, mid-juveniles begin to slowly move offshore, and late juveniles will occupy offshore regions.

In order for eggs and larvae to get from the outer-shelf and slope onto the inner shelf, several factors are involved. Firstly, Pacific Halibut larvae can exhibit high degrees of movement, and in particular, vertical movement, where larvae can take advantage of tidal mixing and vertical instability from interactions with canyons (via interactions with the Alaska Stream) to get transported to the inner shelf. Furthermore, relaxation of downwelling winds, sporadic downwelling events pushing larvae from the surface to inshore regions, as well as eddies and bathymetric steering can facilitate both slope-shelf and alongshore transport and allow larvae to move inshore. In general, transport of eggs and larvae are presumed to be primarily influenced by deep water circulation patterns. Interactions with currents such as the Alaska Coastal Current, Alaska Current, and Alaska Stream can also facilitate alongshore transport, and can be further enhanced by the development of both mesoscale and basin eddies, which can be strengthened through intense Aleutian Lows and positive PDO events (i.e., higher connectivity during these conditions). Through the Alaska Current and Alaska Stream, the westward drift of eggs and larvae in the EGOA likely get retained in the CGOA near the PWS region, prior to changing over to the Alaska Stream, while those spawned in the CGOA likely end up in WGOA and the BSAI region. These differences in transport can be related to various factors including: 1) the Alaska Current turning into the Alaska Stream, and 2) a narrower continental shelf in the eastern GOA versus the western GOA. Thus, eggs in the CGOA have high connectivity with WGOA, while those in the EGOA have connectivity with CGOA, but also in southern regions as well. Beyond the GOA, eggs that are spawned near the Pribilof canyons are likely to remain in the Bering Sea, likely due to how those currents move and are somewhat confined in that system, although some larval transport towards the western Pacific (i.e., Russia and Japan) likely occurs. It is also possible that eggs in the GOA can be transported into the GOA through the ACC around the Samlga Pass and Unimak Pass regions. Additionally, eggs that are spawned earlier in the year will likely allow for greater distances travelled, relative to later period, which could potentially be beneficial in larval survival. Given that the larval duration of Pacific Halibut is fairly long (6 – 8 months), it is possible that this is an advantage for promoting broad larval dispersal via passive transport of currents.

As individuals grow, juveniles will primarily occur near the west of Kodiak, with mid-juveniles being more wide-ranging and dispersed, while for the BS, juveniles stay in that region and occur on the shelf near the Pribilofs, Bristol Bay, and Alaska Peninsula. Juveniles tend to exhibit an east to west movement in the GOA, while adults tend to exhibit a west to east movement, with individuals spawning near CGOA and EGOA. Thus, annual movement rates are much higher in juveniles than adults – there is also some evidence that adults will continue their migration towards southern regions. Additionally, adults tend to undertake summer feeding migrations with fairly high stie fidelity (April to September) and migrate offshore to spawn in the winter (September to April).

Using Alaska sablefish as an example here, I will detail how the combination of life-history behavior, oceanographic currents, and prevailing environmental conditions govern the dispersal of sablefish and contribute to successful recruitment. Starting with their initial life-stages, eggs are spawned on the slope at about 300 – 800m and get retained at depths of about 400m and deeper, remaining somewhat buoyant. Spawning tends to occur during January – March typically hypothesized to be concentrated around the eastern and central GOA, where larvae hatch from eggs at about 12 - 30 days (several weeks) (~6mm in size, fairly big) and only exhibit spontaneous movement after 20 days of hatching during which they rapidly ascend to the neuston, with an ability to control their vertical movements. This suggests that during that 20-day period post-hatch, sablefish larvae are likely somewhat planktonic and depend on the connectivity of prevailing currents to reach inshore regions. However, given that sablefish can develop fairly large pectoral fins during the larval and juvenile life stage, they are likely able to maintain their position in the water column as well as have relatively good swimming capabilities (beneficial for feeding and predator avoidance), and are likely able to exhibit some degree of directed movements when being carried by currents. As larvae develop into juveniles, they tend to be found in inshore coastal bays and inlets (a lot of juveniles found in Saint John Baptist Bay in Sitka) and can exhibit high movement rates. However, during high recruitment events, juveniles will also inhabit the shelf regions suggesting that they can utilize a variety of habitats. During the egg to larval stage, eggs and larvae tend to be onshore from the months of April – May, generally coinciding with the spring bloom in in offshore regions, while juveniles tend to be inshore during the months of June – August, coinciding with a potential summer bloom. The thermal tolerance of sablefish tends to show that they are sensitive to cold waters (lethal at temperatures of 2C). Furthermore, survival of larvae is highly correlated with an abundance of food, where large year classes have been associated with increases in copepod abundance and increased water temperatures. Primary food items for juvenile sablefish include: 1) euphausiids, 2) copepods, 3) pelagic tunicates, 4) seldom larval fish, and 5) amphipods. As individuals age, they begin to exhibit ontogenetic movements towards deeper offshore regions. While some previous studies have suggested that sablefish exhibit an east-west migration pattern, some recent studies have more variable movement patterns (29% moving west from CGOA, and 39% moving east from CGOA). Furthermore, these recent studies have suggested that movement is correlated with the abundance of old individuals, where movement rates are lower with high SSB.

Coupling oceanographic features with larval transport, studies have found that when spawning occurs uniformly along the slope of the GOA, a large proportion of eggs are transported outside of the model domain – i.e., outside of WGOA, which could enter into the Aleutians and Bering Sea. Thus, retention of individuals alongshore is fairly low. In the same study, individuals settling in the GOA are most likely to have resulted from the EGOA (southeast; Sitka, Cross Sound were likely source locations), which tend to settle in the CGOA near PWS and Icy Bay. By contrast, individuals tended to be transported outside of the GOA, when spawned along the slope of PWS to West Shumigan Islands. ***Increased settlement success in the EGOA could be due to the narrower shelf, making it easier for cross-shelf transport mechanisms to transport larvae onto the shelf.*** Connectivity patterns of larvae tend to exhibit a westward direction, likely following the ACC and the Alaska Current, and thus, larvae and eggs are transported in an east-west direction. ***In general, it appears that increased total connectivity, increased spring offshore and summer onshore primary productivity, and annual cross-shelf flow are related to recruitment strength.*** In terms of the influence of oceanographic features on recruitment strength and transport:

1. Onshore advection from slope to shelf (cross-shelf flow) can be facilitated by storms generated during strong Aleutian Lows, which tend to coincide with a positive PDO phase,
2. Onshore advection can also be related to the development of meanders and eddies that result from baroclinic instability (increased precipitation) as well as instability in the Alaska Current, which has been proposed to be related to positive PDO and El Nino events,
3. Given that IBM studies assuming uniform spawning and the absence of horizontal movement have found that a lot of eggs get transported outside of the model domain, and because most juveniles have been observed in SJBB, it is likely that spawning is concentrated near the eastern and central GOA, with the ability for directed horizontal movements towards coastal inlets and bays like SJBB,
4. Given the influence of the ACC, Alaska Stream, and Alaska Current, it is likely that individuals spawning in WGOA and CGOA are unlikely to settle in the domain of the GOA, and rather are likely to bet transported into the BSAI region, assuming no horizontal movements occur. However, if directed horizontal movements occur, settlement in the WGOA is likely, as well as the utilization/leveraging mechanisms related to canyon transport through tidal amplification and mixing,
5. The overall total connectivity of the system is an important factor for governing recruitment dynamics. IBM studies and fine-scale oceanographic models show that positive PDO and strong Aleutian Low events can increase connectivity through stronger winds, the instability of the Alaska Current developing eddies and meanders, which could allow for enhanced transport towards inshore regions. Coupled with increased spring and summer primary productivity (which was correlated with recruitment strength in an IBM study), the transport of nutrients and food likely increase, which all contribute to the success of recruitment. Thus, spawning in the winter and hatch timing during the spring likely allow sablefish to take advantage of offshore spring primary productivity, and summer onshore primary productivity to facilitate survival and recruitment success.

My current hypothesis for the recruitment success of sablefish coupled with oceanographic features is as follows:

1. A strengthening of the Aleutian Low (generally associated with positive PDO events) increases winter cool air from the Arctic that enhances productivity in the central North Pacific basin by decreasing SST and increasing turbulence in this region resulting in enhanced upwelling and vertical mixing of cold nutrient rich waters. The intense Aleutian Low also increases winds, and hence, wind-driven transport of nutrients from offshore regions to nearshore, thereby increasing primary productivity,
2. The productivity in the central North Pacific basin is advected onto the GOA shelf around April – May through the Alaska Current, which can increase productivity along the Alaska coast and in the shelf regions,
3. During an intense Aleutian Low, the ACC and AC can strengthen (which can enhance GOA circulation and overall connectivity from slopes to inshore) and interact with downwelling favorable winds, which can increase the activity and development of anticyclonic eddies, enhancing productivity and serving as a transport mechanism onto the shelf and other regions,
4. The development of eddies and meanders enhances nutrient production and can result in interactions between iron-rich coastal waters and nitrate-rich offshore waters – tends to result in more summer blooms,
5. In general, it is believed that mean recruitment is governed by spawners in the CGOA and EGOA, with smaller contributions in the west. Increased activity of anticyclonic eddies can in general, increase productivity and growth of sablefish that get entrained there. Furthermore, survival of western populations may better take advantage of increased eddy activity, and may be swept into the Bering Sea,
6. Given that sablefish can develop fairly large pectoral fins, they should have fairly strong swimming capabilities either in the horizontal or vertical direction, and can traverse into inshore areas, as well as take advantage of productivity within the eddy,
7. At the same time, an intense Aleutian Low will carry moist warm air from the south along the Alaska coast, increasing precipitation, water temperatures along the coast, and downwelling, as well as result in increases in wind and associated offshore to neashore transport of nutrients,
8. The conditions above can result in earlier water column stability, stratification, and more intense but short-lived spring blooms (blooms happen typically in May and June when larvae abundance peaks). Intense and short-lived spring blooms can result because nutrients get depleted more quickly, and increased water stability can reduce mixing of nutrient-rich waters, which can lead to underutilization of phytoplankton and reduced zooplankton biomass. However, some studies suggest that intense and strong stratification can be beneficial for phytoplankton and zooplankton growth. Furthermore, Royer et al. 2001 suggests that increased freshwater discharge could increase onshore transport of nutrient-rich subsurface waters,
9. The intensity and duration of phytoplankton blooms aside, it could be possible that the increased productivity advected from the central Pacific basin in combination with increased discharge and precipitation, which has been proposed to enhance onshore transport of nutrient-rich subsurface waters, and the increased development and activity of mesoscale eddies, allow for a maintenance of enhanced primary and secondary productivity during the spring and the summer. In fact, recruitment strength has been related to increases in phytoplankton (proxied as chlorophyll levels) during the summer for sablefish (Yasumiishi *et al.*, 2015). Late summer blooms can increase the prey base for sablefish, and potentially increase fish energy stores and contribute to increased overwinter survival.
10. Additionally, the increased activity of eddies and meanders as a result of intense Aleutian Lows can bring nutrients to the surface, which can be exploited by doliolids (tunicata; gelatinous zooplankton) and has been observed to result in blooms during these events (Pinchuk *et al.*, 2021), and in the diets of juvenile sablefish (Sigler *et al.*, 2001; Pinchuk *et al.*, 2021).
11. Lastly, warmer SST conditions also allow for increased metabolism and growth, and are preferable for sablefish, given their intolerance for cold water. In particular, studies have found that recruitment is related to warmer temperatures.

# Curry (Data Limited Methods + Size-Structured Models)

## Data Limited Methods

### a4a approach (Jardim et al 2014)

The a4a approach is proposes a method that constructs an age-structured stock assessment model in the form of a linear model. However, this approach is not a data-limited method, given that it requires age-composition data and all the other components that are commonly used in a typical age-structured model. However, this method is proposed to be more intuitive for non-assessment practitioners attempting to conduct a stock assessment and seeks to enhance participation in the field of stock assessments, as well as conduct an increased number of stock assessments, even for stocks that are not commercially important, because there are still many stocks worldwide that are left unassessed, to provide information on ecosystem status. Thus, this approach bridges the gap between complex assessments and linear models that are typically used within the ecological field. In the a4a approach, sub-models such as selectivity can be parameterized as a linear model (i.e., F ~ age + year) or a generalized additive model to allow for more flexibility and correlations across ages and years (i.e., F ~ te(age,year)). ***The a4a approach requires data inputs for catch, composition data, and indices of abundance.***

### LIME (Rudd and Thorson 2017)

In many data-limited fisheries, catch and length-composition data are the easiest sources of information that are collected. Although methods are developed to utilize this information to assess the status of a stock, many of these methods assume equilibrium conditions. These methods include length-based SPR as well as mean-length mortality estimators developed by Beverton and Holt, and Ehrhardt and Ault, although these methods are treated as independent and do not allow for information sharing and are only applied to one year of data at a time. Further, the equilibrium assumption is often violated because recruitment and fishing mortality is seldom in equilibrium. Even if it appears that mean length is constant over time, your fishery may not be in equilibrium because increasing fishing mortality will lead to decreasing length, while a constant fishing mortality with a recruitment pulse will also lead to decreasing length. If these two processes interact with each other, it is possible that mean length stays constant and cancels each other out.

The LIME method is essentially an age-structured model that is only fit to length-composition data and catch data. ***It requires at least one year of length data, a catch time series, and life history information (natural mortality, age-length relationship, and maturity). However, it is also able to accommodate additional information such as more years of length data as well as index data.*** ***It estimates fishing mortality parameters, recruitment deviations as random effects, DM weighting parameters, and length-based selectivity.*** Conversions from length to age to fit length-composition data is done using an age-length transition matrix derived from the length-at-age relationship. ***The LIME method is also able to represent LB-SPR if only one year of length-data and no catch data are available. This is done by setting recruitment to 1 and not estimating recruitment and fishing mortality deviations.***

In general, LIME is able to correctly estimate SPR with only one year of data (but is imprecise), where estimates improve both in accuracy and precision when incorporated with abundance index data as well as additional years of length data. Furthermore, LIME outperforms LB-SPR in most scenarios, particularly when data were simulated using LIME and when equilibrium assumptions were violated. However, under equilibrium assumptions, LB-SPR performs slightly better with higher precision. ***However, this seems to be limited to short-lived and medium-lived species. In general, values are best estimated for short-lived species, followed by medium-lived, then long-lived species.*** Presumably, this occurs because recruitment deviations are estimated from length-data, which is easier for the model to differentiate when species are short-lived, given that there is less overlap in length-at-age data, where for medium and long-lived species, the increasing number of ages can result in increased overlap between length-at-age as well as variability at-age, degrading these recruitment estimates (***easier to track cohorts with short-lived species***).

Additionally, when length data are collected on monthly time steps, LB-SPR performs better when pooling these data together, compared to LIME. This occurs particularly for short-lived species (not for long-lived fish), ***likely because growth occurs rapidly – modelling temporal dynamics in monthly time-steps for LIME mostly resolves this issue.*** ***Thus, this suggests that for length-based models and short-lived species, it is important to consider the temporal structure of the model given the potential for rapid growth within a given year.*** With respect to data collection programs, increasing lengths of up to 100 samples for short-lived species is beneficial, while more lengths (up to 500) are required for medium to long-lived species to better resolve model estimates, likely due to increased variability in length-composition data and the length-at-age relationship. ***Furthermore, increasing the time-series of length data reduces biases, but increases imprecision because the uncertainty for the age-length relationship is further propagated through the model.*** ***In general, incorporating additional data types results in better model estimates***, rather than increasing the amount of data collected for a given data type for LIME. ***Across the board, LIME performs better than LB-SPR under non-equilibrium conditions, but not for equilibrium conditions. LIME is also more flexible because it can accommodate multiple data types and scan mimic/perform just as well compared to LB-SPR and allows for information sharing in previous years to help estimate recruitment and other processes, whereas LB-SPR treats each year of data independently.***

### Catch Curve Stock Reduction Analysis (Thorson and Cope 2015)

***Typical stock reduction analysis*** ***combines time series of catch data with a guess of what depletion is at the end of the time series to reconstruct abundance and exploitation rates***. ***It also utilizes priors for steepness and natural mortality and information on weight-at-age – recruitment is deterministic given that there are no data to inform it.*** This is done through the catch equation coupled with a population dynamics model (i.e., Bt = Bt-1 exp(-(F+M)) + Rec). The traditional model can be extended to relax the assumption of uniform selectivity as well as incorporate a production function to the biomass dynamics model, however, it still requires a guess of depletion at the end of the time series. However, this model requires an assumption of final depletion, and estimates can be quite sensitive to this assumption.

For some fisheries, there is also a year of ***age-composition data available, which can help support the use of catch-curve methods***. Catch curve methods use age-composition data to estimate total mortality of a species. Commonly, a single year of age-composition data is used, but traditional catch curves require the tracking of a cohort, instead of a synthetic cohort. This is necessary because differences in recruitment can alter the perception of what total mortality is. It also requires total mortality to be constant beyond a certain age. Catch curves are quite simple and are simply a regression with log Catch ~ Intercept -Z \* Age, where the intercept is starting abundance multiplied by some scalar. However, both catch-curves and SRA are flawed because they violate assumptions and require quantities that are unknown to be known. ***Note that catch-curves do not provide information on stock-status***. Catch-Curve Stock Reduction Analysis (CC-SRA) aims to combine these two methods to avoid the need for violation of assumptions and the sensitivity to final depletion.

CC-SRA aims to combine both catch-curve and stock-reduction analysis into a single model. Here, parameters are identifiable given the inclusion of age-composition in the final year. ***The model is a simple age-structured model with fishery selectivity parameters estimated, Beverton-Holt stock-recruitment, and recruitment in the first year occurring from an unfished by non-equilibrium state. Recruitment is penalized towards the mean Beverton-Holt function, and thus, deviations are allowed. Lastly, priors on M and steepness are required here, as well as information on weight-at-ae and maturity-at-age for SSB.*** Age-composition in the final year is fit using a multinomial, which is based on the catch-at-age in the final year. ***The inclusion of age-composition in the final year replaces the catch-curve, and thus, allows an estimation of fishing mortality in the final year, which relaxes the need to specify a depletion value.***

Comparing SRA and CC-SRA models, CC-SRA does better because no assumptions are required about the final-year depletion since this is estimated through the final year F with age-composition data. Thus, under circumstances where depletion is not at 40%, which is a common assumption for SRA, CC-SRA performs better, although with much larger confidence intervals. As noted above, catch-curves are unable to provide estimates of stock status and assume that mortality is constant – although the mean fishing mortality is generally adequately estimated. ***Under situations with high recruitment variability, confidence intervals understandably get much wider, because recruitment and selectivity are much more difficult to estimate. Furthermore, for species with fast-life histories (increased natural mortality), estimation tends to get more difficult because there are more 0s present in the composition data.*** In general, CC-SRA performs better than catch curves and traditional SRA because it avoids the need to specify a final depletion level and there is no need for assumptions regarding constant recruitment and mortality (informed by the time series of catches). Similar to the theme of many data-limited methods, CC-SRA has a heavy reliance on prior information on life-history characteristics.

### Length-based SPR (Hordyk et al. 2015)

***Length-based SPR is a method that allows for the estimation of stock-status through estimation of SPR simply using length-data. It is applied independently to a given year and utilizes equilibrium assumptions.*** It also requires inputs and prior knowledge on ratios of Beverton-Holt life-history invariants ***(M/k)*** and assumes that natural mortality can be calculated through the max-age estimator (i.e., M = -log(0.01) / max age), ***Linf***, ***maturity data***, ***CV of length-at-age***. Importantly, the underlying theory is that:

1. Under equilibrium conditions, two stocks that have the same M/k and F/M ratio will have the same length composition. Thus, the observed length-composition data under unfished conditions is determined only from M/k and F/M,
2. The von Bertalanffy growth function when standardized by Linf, can be represented through the ratio of M and k, and the expected length composition of the catch can be calculated through selectivity and interactions with mortality and k,
3. All species with the same M/k and F/M values will have same SPR under equilibrium conditions, regardless of the absolute values of M and k.

Going through the steps of how the model is calculated and works:

1. Everything is relative and only one year of length-composition data are required,
2. Specify a value for M/k, which allows you to calculate a standardized von Bertalanffy curve,
3. Specify a value for the CV of length-at-age, which allows the derivation of a standardized age-length transition matrix,
4. The estimation and specification of M can be derived through a max-age estimator,
5. The total mortality in a given year can then be calculated through M, F/M, and selectivity estimates (latter two are estimated),
6. Numbers-at-age/length can be calculated through an exponential mortality model and an age-length transition key,
7. Length-composition data are used to estimate F/M and selectivity parameters (logistic) using a multinomial likelihood. ***Note that length-composition data are standardized by Linf***. This is needed because the equation for a standardized von Bertalanffy is standardized by Linf in order to leverage the M/k ratio to determine SPR rates from length-composition (or some related metric Lmax),
8. The use of selectivity, natural mortality, and estimates of fishing mortality can then be plugged into our standard egg production method to calculate yearly SPR.

In general, the LB-SPR is unbiased when all values are specified correctly and there is no variability in recruitment. However, if M/k is mis-specified, F/M will be biased – if M/k is low, F/M is high, vice versa. This occurs because a low M/k implies that M is lower, which goes into the denominator of F/M. Additionally, the model is quite sensitive to the values of Linf, likely because it standardizes the length-composition data and fits to that. A nicety of this model is that it is fairly insensitive to values of CV for LAA. Under scenarios of recruitment variability, values generally are median unbiased (SPR, F/M and selectivity), although they become increasingly imprecise. The imprecision can lead to over or underestimates of up to 100%. ***With high recruitment variability, it is likely that LB-SPR does poorly because length-composition can become bimodal and makes it impossible to fit.*** One solution is to model length-composition and apply LB-SPR in finer time-steps or to aggregate these data over coarser timescales such that they are more unimodal.

***Furthermore, when dome-shaped selectivity is present, where LB-SPR assumes logistic selectivity, SPR is underestimated because it assumes that all individuals are removed from the population due to fish, and hence why they are not observed, resulting in a high F/M ratio. Species that have low M/k ratios will be particularly sensitive to these scenarios because they will have length-composition data more stacks towards the asymptotic size, and the model will assume that their absence is due to extremely high fishing rates, resulting in severely positively biased values of F/M – the converse is true where with high M/k, you will be less affected by dome-shaped selectivity because most individuals have disappeared towards the older ages.*** Compared to catch-curve methods, LB-SPR provide similar estimates, although they are a bit more conservative. LB-SPR works well for species with M/k ratios > 0.53 because there will be a length distribution that is a bit righter skewed, allowing for more information about what fishing mortality would have been (i.e., more information on fishing mortality if older fish are depleted from the population, whereas less information on fishing mortality if all fish quickly grow and you only see the large fish in the population – fishing is unlikely to have a large impact on the length composition data).

Thus, LB-SPR will generally only ***do well under low-recruitment variability stocks (less bimodality in length data) as well as those with high M/k ratios*** which allow for more information to be conferred about fishing mortality rates. It also ***does poorly under domed selectivity*** because it overestimates F/M to compensate for the fact that no old fish are observed. Lastly, this model compares an equilibrium length composition to the observed length composition to estimate F/M and selectivity***. If your population is not at equilibrium, large biases will manifest.***

### Cook (2013) Assessments with only survey data

***In cases where survey data are available but catch data are either unreliable (i.e., misreporting) or unavailable, assessments in a relative sense can still be conducted***. The use of catch data when misreporting occurs can distort the population scale in the absolute sense, and there may be a benefit to omitting catch data – although I note that there are methods to estimate misreported catch using censored likelihoods. Note that this is in a relative sense because information is lacking regarding the absolute scale of the population. ***In such circumstances, a general age-structured model can still be utilized, although quantities are now expressed on a relative scale*** (i.e., ***survey indices can be used to estimate relative declines and increases*** ***and hence, relative fishing mortality and recruitment***). However, these models often run into issues with confounding in natural mortality, selectivity, and catchability, because there is not enough information to estimate these values (i.e., lack of catch data, composition data, etc.) and the use of Bayesian Models coupled with informative priors can be helpful in these cases. Such a model is illustrated by Cook (2013), where they only fit to several survey abundance indices (abundance-at-age data) as well as composition data from the survey. However, the model is constrained by priors on: 1) strong priors age-specific natural mortality, 2) strong priors on commercial fishery selectivity based on minimum landing sizes, 3) moderate priors on initial fishing mortality, 3) weak priors on initial abundance and recruitment, 4) weak priors on survey selectivity parameters, 5) weak priors on survey index variance, and 6) weak uniform priors on survey catchability.

Comparing this model that encapsulates age-structured dynamics with a traditional age-structured model fit using misreported catch data, estimates of fishing mortality and SSB are similar, but tend to diverge during periods where misreporting begins. On a relative scale, estimates of removals are also fairly similar, except that during certain periods, removals are higher on the relative scale for the survey only model, likely encapsulating catch misreporting. Selectivities are also estimable for both fishery and surveys through the use of informative priors, giving credence and support for this method and tends to fit the proportion-at-age data well. With the incorporation of additional surveys, uncertainty in both recruitment and SSB tends to decrease as well. As such, when comparing this model with traditional age-structured models, it is generally able to reconstruct selectivities through the use of informative priors, relative SSB, fishing mortality rates, and relative recruitment. While such a model does not estimate quantities on an absolute scale, it is able to do so on a relative scale and provides information on stock status, which can in turn be used to adjust fishing effort (i.e., limit fishing) and to monitor the stock of interest. Further, it serves as a good diagnostic when catch data is suspected to be misreported, as well as ensuring that trends on a relative scale are similar between two competing models. However, note that informative survey data are necessary for this method to work well (i.e., good contrast in data to provide estimates of fishing mortality, recruitment).

### Cope et al (2015) Simple Stock Synthesis (Age-Structured Models constrained with priors on life-history and terminal year species-specific stock status)

***For models that require some specification of relative stock status, it may be beneficial to develop a stock-specific stock status prior***, which may help better inform the estimation process and result in more robust model estimates, given the sensitivity of many stock-reduction analysis methods to this stock status value (a common assumption is a prior centered at 0.4, which does not perform well is mis-specified). This can be done by developing a prior during a period earlier than the terminal year, as well as using some qualitative life-history information, where there is some information on stock status, as opposed to setting stocks tatus in the terminal year as a prior – this way, you lose out on information on what stock status should be and define your status *a priori.*

Thus, to develop an informative relative stock status prior:

1. You need a measure that is related to stock status, and
2. Determine a period in which stock status should be based on prior to the terminal year (needs to be a period in which consistent fishing/vulnerability occurs),
3. A measure that can be related to stock status are indices derived from a productivity-susceptibility analysis, which uses qualitative information about a species life-history and fishery characteristics to define the vulnerability and susceptibility to overfishing in a given year (ranks species productivity based on maximum age, mortality, maturity, recruitment and susceptibility based on biomass of spawners, catchability geographic concertation, and vertical overlap – vulnerability index is calculated via Euclidean Distance of all these ranks),
4. These PSA metrics can be linked to stock assessments to formulate a relationship between the vulnerability of the stock and its stock status,
5. ***Using the formulated predictive relationship, the stock-status prior can be predicted by developing a vulnerability score for a given stock. This vulnerability index can potentially be recalculated annually, based on annual rankings of the aforementioned qualitative attributes to derive annual priors for depletion*** (Patrick *et al.*, 2009)***.***

***This developed prior can then be applied to models that combine catch and index data, or models that only use catch data to estimate stock status using the simple stock synthesis framework, where they are age-structured assessment models but with priors placed on natural mortality, steepness, and relative stock status.*** Additionally, in these data-reduced models, selectivity is set equal to maturity, and general information is also needed for weight-at-age and other life history parameters that are typically used in an age-structured assessment model. This then estimates initial abundance, catchability for the model that incorporates and index. Recruitment is treated as deterministic, given that there is not age-composition data available.

In general, the relationship between relative stock status and vulnerability is that higher vulnerability tends to result in lower stock status. Furthermore, these stock status priors developed using vulnerability scores tend to differ from priors that are centered around 0.4, with some stocks that are fairly diffuse, suggesting that they can range from highly depleted to less depleted, potentially due to the stock’s inherent recruitment variability. When comparing models using more informed stock status priors, they are generally estimating values that are closer to stock status estimated by a full assessment method. ***In particular, performance tends to mimic a full age-structured assessment better for stocks that have high vulnerability because their stock status is more likely to be away from 0.4, whereas those with low vulnerability are more likely to be closer to 0.4.*** When comparing different methods, those that incorporated both catch and index data outperformed those with only catch data, although when index data are variable, they perform similarly. Thus, stocks that have priors developed by relating vulnerability scores to stocks status through a predictive relationship can outperform the traditional prior centered at 0.4, particularly if the stock has high vulnerability, where it is less likely that their terminal year stock status is at 0.4.

### Froese et al (2016) Catch-MSY (cMSY) and Bayesian Surplus Production Model (BSM) (Models that use catch and resilience information)

Several newer approaches have been developed that use informative priors derived from theory to constrain estimation to estimate biomass and reference points, while using only catch data or both catch and index data. ***This includes Catch-MSY (cMSY) that is a stock reduction analysis model with a production function via a Schaffer curve, where estimation is constrained by informative priors on r, k, and depletion estimates, while using only catch data. BSM is also a production model, but utilizes information from both catch and index data, and thus, has some information on depletion, but still relies on informative priors derived using resilience.*** ***A key issue with many production models that are not constrained by informative priors is that stock productivity tends to be overestimated at low stock size***, because: 1) there is not enough contrast in the data, and 2) r and k are correlated such that increases in r (high productivity) leads to decrease in k (low stock size).

Typical production models utilize catch and indices or abundance information to estimate productivity. ***Conversely, CMSY uses productivity and catch information to estimate abundance.*** Here, priors for productivity (r and k) are derived by finding the most viable pairs. Viable pairs of r and k are found by running the model such that corresponding biomass trajectories do not result in negative catches and biomass. In previous iterations, the viable pairs were derived using the geometric mean of those viable pairs***. By contrast, CMSY uses the tip of the triangle (k~r) distribution (75th percentile) such that the viable pair contains the definition of r – the maximum intrinsic rate of increase is used (because many r,k combinations can result in viable pairs).*** The theory behind this is that at a large population, you can sustain vaiorus catch levels and k but at a smaller population only certain values of r-k will allow you to produce similar biomass trajectories, and thus, the tip of that triangle represents the most plausible pairs of parameters. ***The initial r ranges here are derived from resilience estimates from FishBase***. ***By contrast, k ranges are derived using three assumptions:*** 1) unexploited stock size is larger than the maximum catch, 2) when the stock is depleted, max catch is a larger fraction than k, and 3) Fmsy depends on the productivity of the stock and thus, maximum catch depends on the productivity of the stock (max catch divided by lower and upper bounds of r form lower and upper bounds of k). For the CMSY method, given that there is no information on depletion, priors on starting and ending depletion are required as well.

In comparison to the CMSY model, the ***BSM utilizes information from both catch and index data, and thus, does not require an estimate of initial and terminal year depletion (includes both process and observation error).*** Priors for r and k are derived from the uniform ranges defined by resilience in FishBase and the k ranges from the CMSY method were also transformed into prior densities. Additionally, priors for catchability are needed here given the use of an index of abundance, which are derived using the relationship between equilibrium yield and biomass. Additionally, the BSM model includes a linear decrease in surplus production to represent depensation at low stock sizes.

When both CMSY and BSM are compared in the real world and in simulation testing, prior ranges generally encapsulate the true value. Additionally, most parameters for these methods are contained within the 95% confidence intervals (MSY, r, q, k, final depletion, final biomass). Outside of simulation testing, when both methods are applied to real stocks, they generally encapsulate the true values of those stocks and correspond well. Outside of using resilience as priors for r and k, this can also be done using r = 2Fmsy = 2M = 3K. However, for the CMSY method, it can be a bit more sensitive to the priors and ranges set for initial and terminal year biomass – if the time series is long, incorrect ranges for initial biomass can be reconciled but not for terminal year biomass. Nonetheless, it is surprising that CMSY performs that well, utilizing only catch and qualitative information (resilience and r-k pairs).

### Then et al (2015) Length-based Mortality Estimators

One of the earliest data limited methods involves the use of length composition data to estimate total mortality of a given fish stock. ***This method requires knowledge of a mean length of captured fish that are fully recruited, von Bertalanffy growth parameters, and the length-at-first capture (and thus assumes knife-edged selectivity).*** Furthermore, like many length-based methods, it assumes equilibrium conditions, where fishing mortality, natural mortality, recruitment are assumed to be constant, and no individual variation is present in growth (i.e., CV of growth = 0%). Some methods have extended this method to allow for variable mortality, and recruitment by incorporating an index of year class strength. ***Furthermore, this method calculates mean length by integrating ages to infinity, and thus assumes that there is an infinite exploitable lifespan and that there is not upper age truncation, but there is a lower age truncation.*** This method is known as ***the Beverton-Holt (BH) mortality estimator***.

However, there are some detriments to this method, aside from the restrictive equilibrium assumption***. In particular, some have argued that some fish species have a finite exploitable life span, particularly for species that have high growth rates and high mortality (i.e., time does not***. Thus, species will have non-constant mortality – i.e., fish at the tail end of the distribution may not experience constant Z, and thus, both lower and upper truncations are necessary for the estimator to work well (i.e., – age-at-first capture where Z is assumed constant; lower truncation, and – upper truncation). ***As such, some alterations to this mortality estimator includes both upper and lower truncations by integrating ages from some minimum length to some maximum length.*** This results in what we know as the Ehrhardt and Ault method (EA) where Z is solved using numerical optimization. Similar to the BH estimator, it requires estimates of von Bertalanffy parameters, length composition data, and estimates of where the upper and lower truncation occurs. As such, it estimates knife-edged selectivity on both sides of the spectrum. ***There are two particular ways in which the truncation length can be done, which can implicitly allow for and account for variability in growth. In particular, this involves:***

1. defining mean length with the upper and lower truncation limits or,
2. defining mean length without an upper length truncation – but the calculation of Z involves the use of the upper length truncation.

Furthermore, one of the key requirements for defining the upper length for the EA estimator is that . In general, this upper length truncation should be selected so that it is not to small such that you exclude ages that experience constant Z and large enough so that not too many samples are discarded.

Comparisons of these two methods across various levels of Z, k, CV of the length-at-age relationship, show that the ***BH estimator consistently overestimates*** Z irrespective of what Z and k are. By contrast, ***the EA method was a bit more variable***, with biases ranging from -80% to 140%. Negative biases tend to emerge with over truncation (truncating smaller sizes) and positive biases with under truncation (truncating larger sizes). ***This likely occurs because with over truncation, you are seeing more large fish than you expect in equilibrium, so Z needs to be reduced, while for under truncation, you are seeing fewer large fish than you expected in equilibrium, so Z needs to be increased and therefore overestimated.*** Furthermore, biases reached similar levels to the BH estimator when under truncation was almost to Linf – essentially mimics the BH estimator. Furthermore, for the EA estimator truncating at the correct length does not result in unbiased estimates. It is also preferable to calculate mean lengths using the upper length truncation limit defined, as opposed to defining mean lengths to Linf. The BH method also appears to be more robust to violations in variability in growth, relative to the EA method. ***Given that the performance of BH is consistent, the EA method does not result in unbiased estimates when correctly defined, and that the BH is fairly conservative, this estimator should be used instead of the EA method (also sensitive to high Z and mode of variability in growth). In practice, it would be good to compare the estimates from both methods – the BH should generally perform better, while acknowledging that Z is overestimated.***

## Cross Comparison Studies (Length-based and Catch-based methods)

### Chong et al (2019) Comparison of Length-Based Methods

As discussed above, there are many different length-based methods that are data-limited that can be used in stock assessments, although studies are required to understand which of these methods are most robust under violations to assumptions and different scenarios. This study compares the following methods:

1. Thompson and Bell Length-Based YPR, which calculates stock status assuming logistic selectivity and maturity as well as equilibrium, and requires:
   1. length data,
   2. VBGF parameters,
   3. LW parameters,
   4. selectivity which is derived using length-based catch curves,
   5. an estimate of natural mortality,
   6. maturity data,
   7. width of selectivity and maturity
   8. an estimate of total mortality from length-based catch curves
2. Length-Based SPR, which assumes equilibrium and logistic selectivity and requires,
   1. length data,
   2. VBGF parameters (Linf, CV, and k),
   3. Ratio of M/K,
   4. Length-weight parameters
   5. Maturity data,
3. LIME, which does not assume equilibrium and is an age-and length-structured model. It assumes logistic selectivity and requires:
   1. length data,
   2. VBGF parameters,
   3. LW parameters,
   4. maturity data,
   5. priors for steepness and natural mortality,
4. LBRA, which uses length composition data to estimate Z and assumes equilibrium conditions and knife-edged selectivity. It requires:
   1. Upper and Lower length truncation,
   2. Length data,
   3. VBGF parameters,
   4. Maturity data, LW data, and selectivity from length-based catch curves are used to estimate SPR.

Comparing the methods above, several scenarios were simulated which include, fishing at 40% SPR, overexploited, unexploited, recruitment variability, equilibrium, autocorrelated recruitment, and different life histories are compared. In general, they found that TB and LB-SPR performs pretty well (bias < 30%) but are imprecise. The TB method performs best for short-lived species and when overexploitation occurs. LB-SPR, while robust, was imprecise when underexploited, autocorrelated recruitment and under recruitment variability – i.e., does best under equilibrium conditions. LIME does best for long-lived species and medium-lived species but performs best when under or overexploited. LBRA only does well for short-lived species and tends to overestimate F/Fmsy because Z is overestimated given the need for an age-truncation specification.

Thus, to generalized, all methods tend to do poorly for short-lived species (particularly LBSPR and LIME), and those that performed the best for these life-histories were the TB method and LBRA. Short-lived species likely perform poorly because it requires time-steps that are finer than annual time steps to capture fishery and growth dynamics (inference on mortality is typically on the upper tails of length data, and it is hard separate the effects of fishing, recruitment, and natural mortality if the length data are left skewed). Medium-lived species tend to be easiest to estimate because there is more information about how the cohort persists across time, and the CV in LAA is not large enough to make it difficult to parse out the effects of recruitment. Additionally, LIME did not perform well in this simulation test, because it relies heavily on more than one year of monthly length frequency data – only having that one year makes it hard to disentangle the effects of fishing, mortality, and recruitment. ***They recommend using the TB method with selectivity derived using length-based catch curves as an initial start because it was robust to uncertainties and non-equilibrium conditions. This would be a robust start because it only needs 1 year of data. Additionally, LBSPR with only one year of data should be used as exploration as well, given that it is fairly robust. LIME should be used when you have more than one year of data, while LBRA should be used as a last resort, noting that it is conservative and overestimates total mortality in most cases.***

### Pons et al (2020) Comparison of Length-Based and Catch-Based Methods

Catch and length data are typically the easiest to collect. These two different methods use either catch data or length data, but often do not utilize both data sources. Cross comparison studies to assess which type of method/data is more robust is difficult because they output different quantities. In this study, they compare the performance of a variety of catch-based and length-based methods using estimates of exploitation rate. ***For catch-based methods, these are able to estimate stock biomass, MSY, while for length-based models, these are only able to estimate exploitation rates and SPR.*** Note that because these methods are conditional on depletion, they can be sensitive to how the stock may have been exploited in the past, as well as the guess/prior for depletion, whereas for length-based models, they are pretty sensitive to life-history because differences in growth can greatly mediate how length composition data arise. ***An age-structured simulation model is used to assess the robustness of each method in estimating exploitation rates investigating factors of depletion, life history (recruitment variability), and different harvest scenarios.***

For catch-based methods this includes the use of:

1. Catch-MSY, which is a stock reduction analysis with Schaeffer dynamics. r,k viable pairs are selected based on resilience from FishBase using the geometric mean of viable pairs. It also requires priors on termainl and inital stock depletion,
2. cMSY, which is also a stock reduction analysis, and is the same as Catch-MSY, but instead, uses the 75th percentile of the tip of the triangle of r,k viable pairs to correspond with the definition of r – the maximum intrinsic rate of increase
3. DBSRA, which is also a stock reduction analysis but formulated as a delay-difference production dynamics, and only incorporates spawning/recruitment from mature individuals. It requires priors for depletion, Fmsy/M, and Bmsy/B0,
4. Simple Stock Synthesis, which is an age-structured model using only catch data, but requires priors on almost every population dynamics process.

For length-based methods, this includes the use of:

1. LB-SPR, which assumes equilibrium, requires inputs of M/k, maturity, growth parameters, to estimate F/M and selectivity,
2. LIME, which does not assume equilibrium to estimate recruitment and F, assuming time-invariant logistic selectivity,
3. LBB, which is a length-based Bayesian method that uses priors on life-history parameters and only requires length data.

Comparing these different methods, both SSS and cMSY perform best for catch-based methods. However, SSS did appear to underestimate exploitation rates by 30%, while cMSY performed best for medium to short-lived species. However, Catch-MSY, which uses the geometric mean of viable r,k pairs to calculate biomass does the most poorly. Catch-based methods are more sensitive to lightly depleted stocks and for long-lived species.Length-based methods on the other hand are least biased for medium-lived species and ***LBSPR was most biased***, ***while LIME was most robust.*** ***Furthermore, they provide better estimates of exploitation compared to catch-based methods*** – particularly, LIME which has stronger data requirements. However, for long-lived species, length-based methods do pretty poorly.

Across catch-based and length-based methods, SSS and CMSY do best for short-lived species, and LIME was the most robust under these scenarios. All methods perform worse for long-lived species. ***The poor performance of length-based models for long-lived species and also short-lived species is likely because: 1) variability in the LAA relationship to parse out recruitment, and 2) because of extremely fast growth rates, which could make estimation more difficult given that there is less information in the composition data.*** Furthermore, for stocks that are ***strongly depleted, estimation performance of LB-SPR is going to perform poorly because there is less information to infer mortality from the RHS of the length frequency data (i.e., all those old fish are gone).***  ***For catch-based models, the better performance for strongly depleted stocks could be due to increased contrast in the catch time series to inform parameter estimates (increased number of parameter rejection draws to produce the same biomass trajectories). By contrast, it does poorly for long-lived stocks because there may not have been enough time to observe the full range of dynamics.*** In terms of catch-based methods, SSS perhaps does better compared to production models, because it is more realistic and tracks numbers-at-age instead of biomass, although with large requirements for life-history parameters.

***In data limited situations, it may be more preferable to have biased but precise estimates than unbiased and imprecise estimates***. This will likely be the case, because if assessments are conducted every year, declining trends can be captured by a consistently biased, but precise method, and allow for some form of management action, whereas such trends cannot be captured by inconsistent methods. Furthermore, instead of taking a central tendency of some value as a point estimate for management, we could also use percentiles to reflect the risk tolerance of management. Lastly, SSS and CMSY perform best when catch data are available, and CMSY has less requirements for priors. When length-data are the only available thing, LIME will do best, particularly if recruitment variability is a big issue, although other studies have found that TB and LBSPR can perform well (LB-SPR does not perform well in this study).

## Solutions for managing unassessed stocks and data-limited fisheries

### Berkson and Thorson (2015) Better ways for setting catch limits

In the US as well as other parts of the world, there are many stocks that are unassessed, or are considered data-poor, where often, catch data or length data are the only available data sources. Despite that the available data in many stocks do not support the estimation of MSY, many laws worldwide require us to set catch limits for most stocks, even for stocks that have insufficient data (i.e., the Magnuson Stevens Act). ***Setting MSY and catch limits requires knowledge on: 1) the production function, 2) growth, 3) compensatory recruitment, 4) natural mortality, and 5) fishery selectivity.*** Setting MSY limits for stocks that only have catch data is particularly problematic because catch data do not necessarily provide information on production, growth, recruitment, natural mortality, and selectivity, and can often be confounded by a variety of factors. ***Setting MSY using catch-based methods is thus fairly dubious and can result in highly imprecise MSY estimates***. In the US, 52% of stocks use catch-only methods and scalar approaches are typically used the most differently, with a lot of different variations in how risk levels are set, the value of the scalar, and the reference period that is used. In the PFMC, DB-SRA and DCAC are commonly used instead of scalar approaches. Furthermore, data-poor methods in one region do not necessarily translate to the same category in other regions, resulting in substantial variability in terms of how these methods are applied across regions.

Currently, there are several methods to setting MSY limits using data-limited methods (catch only methods):

1. ***The Restrepo approach***, which defines a reference period of catch to use (a period in which catch is stable – not increasing or decreasing), and some scalar/statistic is applied to this reference period of average catches to define MSY. However, as noted above, such methods do not necessarily have information about production, mortality, selectivity, recruitment, and growth, and setting catch limits using such approaches appears dubious. ***Furthermore, this approach assumes that the fishery is at a sustainable equilibrium of catch, when these limits are set using a given period***,
2. The ***Only Reliable Catch Stocks (ORCS) method***, which is similar to the Restrepo approach, but allows for the incorporation of risk levels and defines scalars using qualitative information about stock status, fishery characteristics species life-history and susceptibility, through a Table of Attributes (i.e., includes more ecological information),
3. ***Depletion-Corrected Average Catch***, which is similar to an average catch approach, although it corrects for depletion in the reference period and back-calculates what a “pristine unfished sustainable catch” should have been. ***As such it requires information of catch (not necessarily the full time series), natural mortality, some ratio of Fmsy/M, and some information on how much the stock could have been depleted during the period in which catch is summarized***,
4. ***Depletion-based Stock Reduction Analysis***, which requires the ***entire catch-history of a stock, some idea of the depletion-status, which is used as a prior, and natural mortality***. The model has a production function (i.e., incorporates growth, recruitment, and mortality) and assumes uniform selectivity to back-calculate catches from the depletion level defined by the prior distribution,
5. The use of meta-analysis to steal life-history information from the rich and share with the poor.

As noted above, it is dubious to set MSY limits based on catch information alone, because catches do not necessarily reflect production, mortality, growth, selectivity, and recruitment processes, all of which are needed to determine MSY levels. Instead of using such reference points, more robust management strategies that are more model-free and robust on average should be developed, which can be simulation tested via an MSE. In particular, strategies that provide high average catches with low probability of depletion should be preferred. ***Model-free empirical approaches can also be use***d, such as using survey indices, average length of catches (i.e., length-based indicators), density of fish inside or outside of MPAs, proportion of highest density areas. ***Thus, management systems should attempt to move away from using catches to define MSY references because they do not necessarily provide information on production, natural mortality, and recruitment.***

### Costello et al (2012) Status and Solutions for Unassessed Fisheries

Many stocks worldwide are left unassessed. Common methods for understanding the status of unassessed stocks worldwide include: 1) looking at the fraction of declined catches, 2) mean trophic levels of catch, 3) applying indicators from data rich fisheries to data poor fisheries. However, many of these methods are flawed because of nuanced interpretations of catch data, stock status, and the meaning of trophic levels of catches, which fail to incorporate economic decisions with respect to why catches change or why mean trophic levels of catch decline, as well as life-history theory and resultant species susceptibility. ***One alternative approach to circumvent the aforementioned issues is to use a multivariate regression approach (panel regression) on stocks that are assessed, with information on species and harvest history (time harvested) to predict (i.e., B/Bmsy ~ Species + Time + Catches) the stock status of unassessed fisheries with some information on catches. However, it assumes that such relationships applied to assessed stocks also apply to unassessed stocks.*** Nonetheless, this approach found that for stocks that have formal assessment methods, 63% have levels that are below MSY, but most of them have lowered exploitation rates to allow for recovery and with fairly low depletion rates (0.94). Similarly, for stocks that are unassessed, 64% of these stocks have biomasses that are below Bmsy, but with a larger percentage of collapsed stocks (0.18) and with increased depletion (0.64). In general, stocks that represent coastal miscellaneous fish and shark species have worse status, relative to larger stocks. Rebuilding unassessed stocks would allow for an increase in 56% biomass in the ocean, while increasing yields to up to 8 – 40%. ***Some solutions for these unassed stocks include: 1) territorial user right fisheries, 2) fishery cooperatives, 3) co-management approaches, and 4) use of data-limited stock assessment methods.***

### Froese (2004) Length-based indicators for overfishing

As mentioned in Berkson and Thorson 2015, the use of catch data to set MSY limits is a bit dubious because it does not reflect the production function and assumes the reference period represents a period of stable equilibrium catches. They proposed using more empirical model-free approaches that can result in a low probability of depletion and high average catches. ***Froese 2004 proposes several model free approaches that utilize only length-data***. In particular, he discusses the concepts of Malthusian overfishing, which occurs due to poverty and a lack of alternatives and convenience overfishing, which occurs because management agencies set quotes irrespective of stock status due to political conflicts, in part due to lack of participation in fisheries as a result of the complexity of many assessment methods. Three length-based indicators are proposed, which includes:

1. ***Letting all individuals spawn at least once (recruitment overfishing).*** This indicator measures the percentage of fish that are captured that are fully mature. Ideally, our length composition from our catches would reflect the abundance of large fish that have already been fully mature,
2. ***Letting them grow (growth overfishing).*** This indicator aims to harvest fish within 10% of the optimal length – determine from a YPR analysis to find the critical length. Ideally most of our fish are within this optimal length,
3. ***Conserve mega-spawners (recruitment overfishing).*** This indicator aims to protect old fish (entrainment, fecundity, higher survival of eggs, better fitness and genetics) and measures the percentage of fish harvested that are larger than the optimal length plus 10%. This indicator can be used to either protect the harvest of large fish, make sure that the catch has at least 30 – 40% of catch from large fish, indicating a healthy length structure, and that anything below 20% is a cause for concern.

These indicators are useful because mean length of the population is reflective of fishing mortality. Furthermore, these indicators serve to protect against recruitment overfishing (letting them all spawn at least once and protect large spawners) and growth overfishing (letting them grow). If most spawners can reproduce at least once, overfishing has been found to be less of a problem. These strategies can have the following goals, which include minimum size limits to let them grow, and maximum size limits to conserve mega spawners. ***Thus, these methods shift the focus away from model-based methods and regulating effort, and instead monitor the length-structure of the population to regulate fishing. They are based on the assumption that overfishing alters the length-structure of the population.*** ***These methods can be readily used as monitoring tools, as well as tools to generally assess stock status, to supplement traditional stock assessments, and could be incorporated into HCRs either as target or limit reference points as well as used as scalars.***

### Cope and Punt (2009) Comparison of length-based indicators for overfishing

As discussed by Froese (2004), there are several ways in which we can construct indicators to reflect stock status and the potential for overfishing. These indicators include: 1) ***Pmat***, which represents the catch length composition containing only of fully mature individuals (spawned at least once), 2) ***Popt***, which reflects the capture of optimally sized individuals that allow for us to maximize yield by harvesting only individuals that have maximized growth, and 3) ***Pmega***, which reflects the catch of mega-spawners, and ideally this would either be at 0% or range from 30-40% to ensure a protracted population. These indicators assume that overfishing will be reflected through changes in length-composition, and while correct in theory, there are some issues with these indicators as illustrated by Cope and Punt 2009.

Using a simulation study, they constructed an age-structured population and derived the aforementioned indicators (*Pmat –* sum of proportion of fish above L50%, *Popt* – sum of proportion of fish above the optimal size where cohort growth is maximized, *Pmega* – sum of proportion of fish that are larger 1.1 times larger than *Popt*). The use of these indicators relies on information on catch length composition, length-at-maturity, and length-at-optimal size. The length-at-optimal size can be derived through a per recruit analysis (where cohort weight is maximized when fishing is at 0), and Lmat can be derived using the ratio of Lmat/Lopt (across a range of 0.65, 0.8, and 0.9). ***Using these metrics, they investigated how they changed in response to different selectivity curves (corresponding to ideal selectivity scenarios and unideal selectivity scenarios – selection of small fish and not large fish), different steepness values, different life-history traits, different fishing mortality rate values (0, 0.3, 0.6, 1.0, 1.5, F40, and F25).***

***These metrics were often not sensitive to differences in steepness or fishing mortality rates*** (you can have huge fishing mortality rate, but your length-based indicator will still be at about 1, if selectivity is knife-edged for Lmat or Lopt), but were sensitive to the values specified for Lmat/Lopt, because they define where those cutoff points are. ***However, they are dependent on the selectivity forms*** – in particular, Pmat and Popt are much lower if selectivity selects for young individuals relative to old, which reflects undesirable fishery selectivity curves. Additionally, higher fishing mortality rates tend to be allowed is selectivities target older fish, when compared to targeting younger fish, so as to avoid growth overfishing***. Importantly, the length-based indicators are not sensitive to stock status***, except for when steepness is high. This can occur if a stock is fished really hard (i.e., high F) but has knife-edged selectivity towards old mature individuals. ***Similarly, if the stock is fished to the point where there are no more old fish left in the population, Pmega can be at 0, which indicates a healthy population according to Froese.*** As such, the metrics above area really sensitive to the underlying selectivity pattern.

Another approach would be to take the sum of all three indicators, which allows for some information on selectivity to be conveyed***. If the sum of the three metrics (Popt + Pmat + Pmega) is below 1, it suggests that the selectivity selects a lot of young immature fish, wheras if it is above 1, it suggests that the selectivity more follows Froese’s recommendations of selecting large fish. Further, if the metric is < 1 and Popt and Pmega is 0, this suggests that the fishery is still capturing a lot of immature fish.*** However, these indicators in general are not sensitive enough to indicate whether a stock is at the target or limit reference point. Thus, these indicators should be considered as trigger reference points where reaching some value will result in some sort of management action. ***Their study ends by providing a decision tree, for values of the combined indicator metric, where certain values will indicate whether it is above or below some reference point/stock status value.***

Thus, their study indicates that size-based indicators when used alone, may not really reflect stock status and should be avoided, as it could lead to misperceptions. Furthermore, length-data can often be somewhat variable, which could lead to increased imprecision in these metrics and increased misperception relative to stock status. By combining all three length-based indicators here, practitioners can make some sense of the selectivity attributes of the fishery, and using the decision-tree designed can provide some rough measure of what stock status is. However, these metrics should not be used as limit or target reference points given their insensitivity to stock status (SSB/B40, or SSB/B20), and instead, should be used as trigger reference points. Furthermore, these indicators are not sensitive to life-history and are fairly irresponsive to changes in fishing mortality. Thus, combining all three indicators together is more well-suited for the purpose of assessing whether we are above or below some reference point, and can be used as a trigger for management measures, as opposed to as a target or limit.

Below, I’ll provide some examples as to when and why these individual indicators may not work well:

1. When *Popt* or *Pmat* are less than 1, sustainable fishing can still occur if fishing mortality rates are maintained at lower levels and older individuals are allowed to persist in the population,
2. In cases where *Popt* is at 1, a high fishing mortality coupled with knife-edged selectivity for *Lopt* can result in such circumstances. Thus, even if the stock is overfished, you can still have an “ideal” value of *Popt*,
3. Similarly, *Pmega* can be maintained at levels of 0 if, for example, no old fish exist due to high fishing rates. Conversely, *Pmega* can be maintained at levels of 30 – 40% if knife edged selectivity occurs for this size-ranges,

Given that, the use of *Pobj* which combines all three metrics will likely fare better in a management context. *Pobj* can be used in the following manner:

1. If *Pobj < 1*, then we can incorporate it as a trigger reference point in an HCR, where we restrict catches more heavily because it reflects that the fishery is taking primarily immature fish, and reducing the fishing mortality or catches will allow us to sustainably fish the stock,
2. Given that values for *Pmega* can remain at 0 if no fishing occurs if all the old fish are removed, *Pobj* could be forced to be constrained heavily at 2 (such that *Pmega* + *Popt* sum to 1 and that old fish are actually present in the population) with *Pmat* constrained to be close to 1, just to make sure there are some old fish left in the population and that individuals are spawning at least once 50% of the time.

Nonetheless, this study suggests that these length-based indicators may not best reflective stock status given their insensitivity and the use of alternative metrics + approaches to understand stock status in attempt to manage the fishery is likely warranted (i.e., using *Pobj* coupled with the decision tree as trigger reference points).

### Froese et al (2016) Revisting Biological Limits

For many fish stocks, the following indicators can be used to describe whether fishing is occurring within safe limits:

1. Fmsy is less than M, where M is used as an upper limit for fishing for many fish stocks. Note that fishing FMSY = M is problematic because the stock will essentially experience 2x of mortality and you would reduce the life expectancy and the reproductive potential of the stock by half,
2. The fishery selection should occur such that individuals have spawned at least once (i.e., selecting only full mature individuals),
3. The stock size is within SSBlim, which is the limit of SSB below which recruitment is impaired (i.e., recruitment overfishing)

In general, SSBlim and Fmsy are derived using traditional stock recruitment analysis. However, these stock recruitment functions are often criticized because productivity tends to be highest at low stock size and lowest at high stock sizes. ***Instead, a segmented regression approach might be more appropriate in modelling recruitment***, because it assumes a linear increase in recruitment at low stock sizes, such that productivity is constant, after which, productivity declines at higher stock sizes (i.e., modelled as a constant rate and then declines, as opposed to an exponential decline). The point at which it transitions from a slope to constant recruitment is where SSBlim is defined. However, these approaches can sometimes produce unrealistic SSBlim estimates because this point has not been observed. Thus some alternative reference points can be developed by:

1. Defining the lowest level of recruitment observed as SSBlim,
2. The lowest level of recruitment observed times 1.4 is a precautionary buffer reference point, where if we are within this limit, we are fishing in a safe limit,
3. And 2 times the precautionary limit is where SSBMSY occurs.

The proxy reference points described above are derived from production model theory, where assuming that SSBpa = 0.2B0, then 2SSBpa = SSBMSY because in production models, SSBMSY = 0.5B0. There are several ways to conduct a stock recruitment analysis using a segmented regression approach, to result in better estiamtes of SSBlim and to constrain the estimation process when there is not enough contrast observed in the data:

1. A traditional segmented approach where recruitment is constant above SSBlim, and below SSBlim, which is estimated, recruitment is sloped,
2. A rule based segmented regression approach, where a breakpoint in the midpoint of the observed biomass is defined – the geometric mean of recruits is taken above this breakpoint. Then, any SSB that produces a recruitment level larger than the geometric mean of recruitment is above SSBlim, and anything below is below SSBlim. A precautionary reference point is then obtained by multiplying the SSBlim reference here by 1.4, where below this point, 95% of recruitment is below the recruitment related SSBlim reference point defined by the geometric mean,
3. The third approach uses a rule-based hockey stick segmented regression approach, but with Bayesian priors using the rules defined above, with the variability defined by the observed variability in the stock itself.

When these methods are applied to a bunch of ICES stocks, almost half of the stocks had SSBlim reference points that were lower than those derived with these hockey-stick shaped recruitment curves and most stocks were below SSBMSY. Furthermore, most stocks had FMSY that were larger than M and selectivity was way earlier than the age-at-first maturity. However, there are some issues with these simpler approaches for deriving reference points – in particular, if your stock has never been above SSBlim or below SSBlim, your estimates of reference points are going to be biased because you have never seen these data before. For example, if the stock has never been constant, then your highest recruitment will be estimated as SSBlim (it may underestimate SSBlim). If the stock has never been declining, then the lowest recruitment point is going to be estimated as SSBlim, even if it may sustain even lower SSBlim (it may overestimate SSBlim). Thus, if there is insufficient contrast in the data, it is not possible to estimate SSBlim reliable, particularly if there is not enough information to constrain the estimation process, suggesting that the use of priors can help result in more reliable reference points estimates.

## Size-Structured Models

### Size-transition matrices

#### Sideek et al (2016) Compare methods for estimating size-transition matrix

Growth and size-transition matrices are fundamental in the use of size-structured models. ***In general, growth and size-transition matrices are distinguished, where a growth transition matrix models molt increments while incorporating molt probabilities, while a size-transition matric directly models changes in size.*** To estimate these matrices, recapture data are used, and in this study, a multinomial likelihood is used to fit recapture data by conditioning on selectivity, observed release size, the size-transition matrix, to predict size-at-recaptures. These matrices can be represented with a variety of functional forms, which include the gamma function, a student’s t function, a normal function, etc. For size-transition matrices, the ranges for the integral can directly use sizes. ***However, in a molt increment growth transition matrix model, the ranges for the integral need to be standardized a scale parameter (at least in the case of a gamma). Molt probabilities tend to be estimated using a logistic function (declining logistic) with molt probability declining with size, while for the growth increment model, increments can be estimated through a linear mode – linear increment from pre-molt size to post-molt size (although other options are available), which is basically a von Bertalanffy model (Growth Increment = a + b\*length\_mids).***

In this study, they compare the performance of estimating either a growth or size-transition matrix within a size-structured stock assessment. ***Estimating a size-transition matrix within the model allows for accounting of selectivity effects but also allows for the propagation of uncertainty in the estimation process. Furthermore, it also allows the use of auxiliary information from size-composition data to inform the matrix, although a drawback is parameter confounding. On a related note, if tag data are not separated by molt and non-molt stages, there is also increased potential for confounding.*** The size-structured model uses catch, discards, retained, and a CPUE index, as well as tag-recapture and length-composition data. Comparing the two methods, there are minimal differences in model estimates and fits. However, there are some differences with respect to the values estimated, parameter correlations, and AIC values. ***With respect to parameter correlations, the growth transition matrix had higher correlations, likely because of an increased number of parameters estimated, and likely confounds more with selectivity and catchability, although this method had lower AIC values. Furthermore, this approach is more biologically meaningful (i.e., molt probability + growth increment).*** In terms of values in the transition matrix, the diagonal values for the growth transition matrix were larger, because it directly incorporates the probability of not molting, and the probability of molting but not growing.

***Thus, the benefits of estimating growth transition matrices inside a model include: 1) leveraging size-composition data to inform the matrix, 2) propagation of uncertainty, 3) allows for accounting of selectivity effects. However, it can lead to increased confounding with selectivity and catchability, because those factors can impact your recaptured lengths.*** Furthermore, there is likely no meaningful difference between modelling growth increments of size-transition matrices, although the latter is more biological meaningful, but also increases parameter confounding with selectivity and catchability due to the modelling of molt probabilities as well as growth increments.

#### Punt et al (2009) Compare methods for estimating size-transition matrix

Generally, methods that estimate size-transition matrices incorporate the same error term for the whole population. However, it is sometimes important to account for individual variability, particularly because observations are not independent, and because incorporating such process error may lead to better estimates:

1. Individual variability in von Bertlanffy Linf or k parameters to model the growth increment, and the size-transition matrix is estimated analytically via MCMC, because there is no closed-form solution. These methods assumes that Linf and k are independent of size class,
2. Variability in the “age” at release, which can be estimated as a random variable. This is important to account for because that can impact how much an individual grows – i.e., older grows less, and thus, accounting for age-at-release is important. Similarly, this requires the size-transition matrix to be estimated analytically via MCMC, because there is no closed-form solution,
3. Estimating the size-transition matrix directly with two variance terms to govern the probability of transitioning from one size-class to the next,
4. Estimating the size-transition matrix directly, but with differences in how the variance of the growth increment is modeled. In particular, this can include:
   1. Varies in proportion to the increment,
   2. A power function to the increment, and
   3. The square root of the sum of the proportion growth increment with some constant.

These different methods discussed above are compared using a simulation study that varies Linf, k, both Linf and k. Age-structured dynamics are simulated with an exponential mortality model, with logistic selectivity. Additionally, different sample sizes, measurement error, stochastic growth in increments, different distributions for Linf and k variability, uniform selectivity, and spawning 20 weeks prior to the start of tagging are also factors that were investigated.

***The most robust method across simulations was when Linf varied as well as age-at-release as a random variable***. Furthermore, different ways of treating growth increments in terms of the variance does not matter as much. Under the different scenarios (stochastic growth, sample sizes, etc.) allowing Linf and random age-at-release, as well as, random k and age-at-release both did well. ***Nonetheless, there does not appear to be a method that performs best across all scenarios***. ***Additionally, incorrect estimates of Linf or k do not impede good performance, because these parameters are negatively correlated, and multiple combinations can still result in a similar matrix being estimated***. For example, as shown here, Linf and a varying was most robust when comparing equilibrium lengths, mean lengths, and the entries of the matrix, but Linf was biased. When tagging occurred at 20 weeks old instead of when they were first spawned, all methods perform worse because larger tagged individuals are faster growing individuals, violating the assumption that Linf and k are independent of size class. In real world cases, allowing for k and age-at-release to be random can lead to very unrealistic estimates given parameter confounding and correlation. Lastly, when selectivity was uniform, transition matrices and population estimates tended to be better, likely because selectivity mediates the recapture component of the model and failing toa account for it will lead to bias – thus, estimating it inside a model will be the most appropriate in most cases. In general, allowing for individual variability when it is present will perform best, and ignoring it will lead to biased population means and transition matrices.

### Integrated models

#### Zheng et al (1994) Size-Structured Model for Red King Crab

A fundamental part of size-structured assessments is that there is a size-transition matrix model that govern how individuals progress and grow through the population. To its core, a size-structured stock assessment is essentially a stage-structured assessment model with a projection matrix. In this paper, they illustrate the use of a size-structured assessment for BBRKC primarily using data from a trawl survey. They use an abundance index from the survey, catch data, as well as length composition from the survey.

***The following features and potential issues are included in this model:***

1. Abundance by sex, shell length, and shell condition (old shell vs. new shell). Old shell and new shell need to be modelled because old shell individuals may have a higher probability of molting than new crab. Additionally, note that molt probability decreases with age, and for males, they can molt multiple times per year, while for females, they molt annually,
2. The model estimates selectivity, recruitment, initial age-at-abundance, catchability, molt probability, parameters for the growth-transition matrix, and length-specific natural mortality.
   1. Natural mortality is bowl shaped to account for increased mortality during early life-history and senesce at older ages,
      1. Related to modelling natural mortality and growth in size-structured assessments is potential confounding between these variables because growth and natural mortality parameters can covary.
      2. For example, if M is high then Linf will generally be low, given that fish are not able to survive to those older ages. Given that Linf and k are negatively correlated, this then results in a higher value of k. Also higher k tends to be associated with higher M (fast pace of life).
      3. Similarly, natural morality is confounded with selectivity (in both size and age-structured contexts) because there is generally less fish available at older ages, and so it’s hard to parse out whether this is due to a lack of fish at older ages or a high M at older ages.
   2. Recruitment is estimated annually and is partitioned out to different length classes using a gamma function. Thus, if the mean length of recruitment is known, only one parameter is required to be estimated,
   3. Note that it is difficult to estimate a stock recruitment function within a size-structured model, because it is difficult to know what year of SSB contributes to which age of recruitment since no age information are available (i.e., no clue when crabs of different sizes actually enter the population), but also because temperature changes can cause growth differences, which in turn, can impact the distribution of sizes at recruitment.
      1. On a similarly related note to recruitment, parameter confounding can be more severe in size-structured assessments because when there is variation in size-at-age, that can make it harder to discern cohorts in the size-composition data. Although this can potentially be better resolved with tagging and recapture data.
   4. In terms of issues with selectivity estimates, with a low recruitment to the first length-class, that can be compensated by a low selectivity to the first length-class as well, which is something else to keep an eye out for. If recruitment is low for the first size-class, a selectivity of 0 or 1 for that first size-class can fit those data just as well.

Given the above, a stock-recruitment relationship (Ricker) is conducted outside the model with some assumption regarding which years of SSB contribute to recruitment in a given year. This was done to assess whether there was autocorrelation in recruitment as well as the presence of a stock recruitment relationship. ***Furthermore, different metrics for stock biomass are used which include: 1) female mature abundance, 2) the product of female and male abundance, and 3) assuming that male’s mate with an average of three females when larger than 160 mm, and only mates with one female if larger than 120mm. If female abundance is less than reproductive potential, that mature female biomass is used instead.*** Recruitment sizes was defined at 97mm for females and 105mm for males, and used the lagged 7 years SSB metric.

In general, allowing for time-varying natural mortality in this model tends to improve fits to survey abundance and better matches the steep decline observed in the survey around the 80s. Length-frequency data also appeared to be better fit with time-varying natural mortality. A stock recruitment relationship fit outside the model indicated that there was some compensation and was supported by a ricker, with some dome-shaped dynamics (overcompensation) at high stock sizes. Furthermore, the stock recruitment was better fit with autocorrelation. Natural mortality is also more likely to be bowl shaped and likely varies across time for RKC. However, in this model***, a catchability of 1 needed to be assumed to avoid parameter confounding with M – M and q are negatively correlated*** – lower catchability with higher mortality because less individuals left in the population.

#### Zheng et al (1996) Size-Structured Model for KRKC and BBRKC

This study uses the same model as the one constructed above, and includes information on commercial catches, effort, the mean and variance of growth increments, natural mortality, and length frequency data. ***It then compares this length-based model only using catch data from the fishery, as well as length frequencies from the catches to survey abundance estimates to see how well they correspond to each other*** (i.e., a more data limited case). Time-varying molt probabilities is incorporated in this model as well, with old and new shell crab sharing the same parameters. Parameter estimation minimizes catch length-frequencies, effort data, and catch data. Additionally, varying the weighting to effort data, catch data, and length-frequency data was also attempted to evaluate its impact. Various values for fixing M at different values and varying M at known values across three different periods were also attempted.

In general, they found that abundance estimates and fit to data often depended on the values of natural mortality and weighting factors – i.e., sensitive to data weighting. As expected, higher mortality values will lead to high initial abundances, because you need to have a higher scale to see the same rate of decline with a higher mortality rate. Furthermore, higher weights to fishing effort leads to better fits to the CPUE data, but lower and intermediate weights were generally preferable. A value of natural mortality at 0.4 tends to correspond better with the absolute indices of abundance from the survey***. Despite the sensitives to the different weighting factors and natural mortality, they were generally able to reproduce the relative trends and dynamics following the survey abundance quite well, suggesting that even if the absolute scale is incorrect, these data limited length-structured models can be quite accurate in reproducing relative trends. Thus, it may prove to be of use for data-limited management contexts. However, if we wanted to manage under a fixed exploitation strategy, it would be good to understand the value of M, to avoid over or under-exploiting the stock (i.e., being able to estimate reference points).*** Furthermore, the model is quite sensitive to estimates of mortality as well as weighting factors – time-varying natural mortality does not really seem to help model fits either. Thus, like age-structured models, these models should have sensitivity tests for M as well as some refinement to data weighting methods. While increasing weights for effort data leads to different solutions, the use of effort data can be a bit sketchy in assessments because it assumes F = qE, and effort can change due to changes in catchability, etc, which can distort trends estimated in abundance. Next, for a more crab-specific case, changes in population size may need to be linked to catchability give their aggregating behavior and contraction behavior – at low stock sizes, catchability could be low, while at high stock sizes, it could be high. The use for time-varying catchability may be valuable in these cases.