Are assessed stocks a representative sample of fished species in the US?

Philipp Neubauer, Dragonfly Data Science, Wellington, NZ
James T. Thorsen, NOAA Northwest Fisheries Science Center, Seattle
Michael C. Melnychuk, School of Aquatic and Fisheries Science,
University of Washington, Seattle

November 8, 2016

1 Introduction

It is often said "what gets measured gets managed". For over XXXX years (historical citation), fisheries scientists have measured human impacts on ocean populations (finfishes and marine invertebrates) with the goal of balancing both the long-term sustainability of human impacts with maximizing harvest (or value) from fishing. In particular, stock assessments generally involve estimating two measures of human impact: (1) fishing rate, i.e., the instantaneous mortality or annual fraction of the population that is harvested relative to an estimated target level, and (2) the population abundance, i.e., spawning biomass or reproductive output relative to an estimated target level. The value of these two metrics can be plotted on a 2-dimensional "stock status" plot, and fisheries agencies are increasingly committed to maintaining fished populations in the "healthy" quadrant (i.e., fishing rate below a target level, and population abundance above a target level; citation).

Recently, the National Marine Fisheries Service (NMFS, the agency in charge of science supporting fisheries management in the United States) has committed to "end overfishing" for all marine species within regional fisheries management plans (with exceptions granted in a few circumstances; citation). In the US, overfishing is defined as any stock having annual harvest above limit levels. A target (or limit) harvest can in theory be calculated by combining a target harvest rate with a target population abundance. However, the vast majority of overfishing limits are currently estimated using methods that do not individually calculate either harvest rate or population abundance (Berkson and Thorson 2015, Newman Berkson Sautoni 2015). For example, depletion-corrected average catch (DCAC; MacCall 2009) is used to calculate an annual fishing limit for many stocks, but is not used to calculate population abundance. DCAC (and similar methods) can therefore be used to help "end overfishing" but is not otherwise informative about characteristics of the fished population.

However, conservationists and ecologists will often be more interested in estimating population abundance (or abundance relative to equilibrium conditions) than estimating an overfishing limit (citations). Estimating abundance generally requires applying a population model to available harvest data and an index of population depletion (either an index proportional to population abundance, or average size/age data); in these cases, the population model essentially treats the historical fishing as a "depletion experiment" to estimate the fraction of the population that is fished every year. In the following, we call these population models "stock assessments", although we acknowledge that other authors have used the term "stock assessment" more broadly to also include methods for estimating overfishing limits (e.g., DCAC). Although NMFS has estimated overfishing limits for the vast majority of fishes in US fisheries management plans, a much smaller percentage of fished species have a stock assessment (in our sense). This relative dearth of stock assessment models presumably arises because developing a stock assessment is time-consuming and requires extensive financial resources from NMFS and other interested parties (Geremont ref).

We argue that stock assessments (population models for estimating absolute

abundance for fished species) are important for many applied and theoretical questions regarding marine ecosystems. However, there is little previous research regarding which fished species are more or less likely to receive sufficient attention to develop a stock assessment. Understanding which species are more or less likely to be assessed could be useful for the following three reasons (among others):

1. Unassessed stocks may receive less attention from the public or fisheries managers regarding management actions. Therefore, if a given taxon is less likely to be assessed, it might also be less likely to receive rapid or public attention when management changes are warrented.

2. Output from stock assessments has often been used in meta-analysis models to understand ecological characteristics of marine fishes in general (Myers ref, Thorson et al. Fish and Fisheries 2015, other ideas...). Therefore, any systematic bias in which stocks are assessed could also bias our ecological understanding of marine fishes.

3. Stock assessments often require updates periodically (e.g., Pacific hake has been re-assessed annually from X through 2016), and agency resources might be fully expended while assessing a small fraction of stocks. However, it is currently unknown whether the rate of assessing new species for the first time is increasing or decreasing. If this rate is decreasing, this could indicate the need for additional public resources for stock assessment.

We therefore seek to provide the first-ever quantitative analysis of which marine species are likely to have undergone a stock assessment. To do so, we combine two databases representing all fished marine species in the United States: a database of all landed species from 195X to 201X, and a database of management and stock assessment attributes for all US fishes with peer-reviewed stock assessments. For each landed stock in each of four US regions (northeast, southeast, northwest, and southwest), we record the year that it first had a stock assessment; we treat any stock that does not have an assessment by 2015 as a "censored" observation (i.e., it might eventually have a peer-reviewed assessment) We then apply a censored time-to-event model to answer the following questions: (1) what economic or biological characteristics are associated with a high or low annual probability of being assessed for the first time? (2) which US region has assessed stocks relatively quickly or slowly? (3) are there certain taxa (e.g., invertebrates or sharks) that are substantially faster or slower-assessed given these biological and economic attributes? and (4) is the rate of stock assessment accelerating or decelerating over time. We show that price per kilo and and landings are the main drivers of globally increasing stock assessment rates, but some taxa defy these trends and are .

2 Methods

2.1 Management database

The year of first stock assessment for marine stocks previously assessed in the United States was determined primarily by interviews with regional stock assessment scientists, as supplemented by conducting a literature review of archived assessments. Many types of stock assessments are applied in the US, with varying levels of model complexity and input data requirements. Assessments for any given stock also tend to change over time, typically becoming more complex as warranted by available data. For consistency across populations, we defined a stock assessment as a single-species statistical model of density-dependent population dynamics (e.g., including some combination of individual growth, recruitment, and/or aggregate surplus production), [ADD CAVEAT ABOUT FITTING DATA TO EXCLUDE SRA], which provided time series estimates of some form of abundance (e.g. total biomass, spawning biomass) and/or fishing mortality (or exploitation rate) as well as some type of benchmark for these time series estimates. Benchmarks could include target reference points, reference points based on maximum sustainable yield (MSY) or its proxies, or initial population abundance; ratios of the time series and their corresponding reference points provide a relative index of stock status. Age-structured models, delay-difference models, biomass dynamics models, and surplus production models all qualified as assessment models, and stock-reduction analyses did not qualify as stock assessments under this definition. Exceptions were occasionally made (i.e. definitions of an assessment were more forgiving) for some invertebrate species like squids due to short lifespans and difficulties of aging individuals. We reconize that stock-reduction analyses are often used to estimate overfishing limits for stocks in the absence of a population-dynamics model fitted to data. These SRAs are generally not interpreted as providing a time-series of population-biomass over time (MacCall ref), so they do not provide all the information expected as output from a conventional stock assessment.

Year of first assessment assignments were compared with the NOAA Species Information System (SIS) database to ensure consistency. The SIS database does not contain information about when a stock was first assessed, so comparisons were restricted to recent years, assuming that stocks which qualified as assessed in a previous year continue to qualify as presently assessed. These comparisons generally showed consistency among datasets. The only exceptions... [squids, couple others].

2.2 Defining un-assessed stocks

Not all species landed in US fisheries are assessed using a population-dynamics model, and some are only assessed in a subset of regions. Our management database therefore only included instances of stocks with a stock assessment, and did not contain records of unassessed stocks. Excluding all un-assessed stocks would not allow us to make general inference about the propensity of

stocks to be assessed. We therefore complement the management database with data from NOAA's landings database. For this, we attributed all regional landings from assessed stocks to these stocks [MIKE - I'm talking about the proportions for each stock in the spreadsheet. Can't quite recall how you did this - can you elaborate here?], and then grouped landings that were not attributed to assessed stocks into species/region combinations, which we treated as un-assessed stocks. We distinguish four regions, defined as: Alaska (i.e., the Eastern Bering Sea, Gulf of Alaska, and Aleutian Islands); US West Coast (i.e., the federal waters of Oregon, Washington, and California); Northeast Coast (including the mid-Atlantic Coast); and Southeast Coast (including the South Atlantic Coast and Gulf of Mexico). These regions do not include the portion of US federally-managed waters in the Caribbean or central Pacific, because the management database has not been consistently expanded to these regions [IS THIS CORRECT MIKE?].

This definition of "unassessed stocks" is difficult to apply in the case of highly migratory species, which often span multiple regions or cross national jurisdiction. We therefore expluded all species that are typically assessed by Regional Fisheries Management Organisations. This included all tuna, billfish and oceanic sharks.

2.3 Explanatory variables

Several variables were considered as explanatory factors affecting the year in which a stock was first assessed. Region and habitat typically occupied by the population were each treated as a categorical random effect. Habitat types from FishBase (ref) or SeaLifeBase (ref) were aggregated into six categories: deep sea (¿200m; bathy-pelagic or bathy-demersal); benthic; demersal; benthopelagic; pelagic; and reef-associated. Adjustments were made to the default global species-level classifications for some populations to better reflect local habitat usage.

Maximum body length of the species was also assigned to each population and used as a numerical predictor, drawing from FishBase and SeaLifeBase. Fisheries that developed earlier may be more likely to be assessed earlier than others. The year of fishery development, defined as the year in which landings first reached 25 percent of the maximum landings across the full time series (Sethi et al. 2011? PNAS), provides historical information about the fishery for each population as was used as a numerical predictor. The catch quantity and ex-vessel price of the population together determine landed value of the population; more valuable populations may be more likely to be assessed. We considered maximum annual landings prior to the first assessment and ex-vessel price per kg in the year of first assessment as separate numerical predictors, drawn from the National Marine Fisheries Service landings database. These state-level landings and ex-vessel price data were linked to corresponding biological stock units according to areas of stock distribution as defined in assessments.

2.4 Time-to-event model

To assess which factors drive the over-all rate of assessments and the time from first recorded landings to a full stock assessment in US stocks, we applied a time-to-event model. Also called survival models in the medical and ecological literature, these models account for censored data (i.e., species that are landed but not yet assessed) while modeling time-to-assessment within a parametric framework.

The Weibull distribution is often used as a flexible model that has several desireable properties for this type of analysis, and one can easily check whether the Weibull model is appropriate for the data at hand (see Figure A.1). The shape parameter of the Weibull density can be interpreted in terms of the rate of events occurring. A shape greater than one suggests an increasing rate of events, whereas a shape of less than one indicates an decreasing rate. This allows us to directly estimate the change in assessment rates over time.

A further desirable property is that the model can be interpreted in terms of the ratio of event rates as well as time-to-event. For example, one can interpret a model coefficient as decreasing or increasing the likelihood of an event occurring at any particular time relative to the baseline (this is usually called the hazard ratio interpretation). Thus, the rate of events may be modified by a particular variable. The model can also be interpreted in terms of time-to-event parameters, which scale the time to the event by an acceleration factor. For example, in a hyothetical scenario, the median time-to-assessment of a demersal stock may be 0.5 times that of a pelagic stock, suggesting that it takes twice as long for pelagic stocks to get assessed. Such acceleration factors are just transformations of the parmeters obtained for the event rate interpretation - the two interpretation are therefore easily exchangeable in the Weibull model.

We thus model time-to-assessment as Weibull distributed with shape parameter τ and rate λ. The connection between the event rate and the time-to-event interpretations can be made explicit by writing the Weibull density as a function of the product of the probability of the assessment occurring at time t and the probability of the assessment not occurring prior to time t. Thus, $P(T = t)$ depends on the probability of an assessment not having occurred prior to time t, $(A(t) = 1 - P(T \leq t) = 1 - F(t))$, where $F(t) = \exp(-\lambda t^{\tau})$ is the Weibull distribution function, and the rate $r(t)$ with which assessments occurr at time t.

$$T \sim Weibull(\tau, \lambda) \tag{1}$$

$$= A(t) \times r(t) \tag{2}$$

$$= \exp(-\lambda t^{\tau}) \times \lambda \tau t^{\tau - 1} \tag{3}$$

We modeled the scale λ of the Weibull distribution as a function of covariates and categorical variables:

$$log(\lambda_{i,r,h,c,o,f}) = \beta X_i + \alpha_r + \gamma_h + \kappa_c + \omega_o + \zeta_f, \tag{4}$$

where beta is a row-vector of regression coefficients, X_i is a vector of continuous covariates. Continuous covariates were taken as the logarithm of price per kg, maximum landings, and species maximum length, standardised for the analysis. α, γ, κ, ω, and ζ_f were region, habitat, class, order, and family specific random effects. The model was estiamted within a Bayesian framework, using Markov Chain Monte Carlo (MCMC) as implemented in the JAGS package. MCMC was run using three chain of 210 000 iterations each, keeping every 100th iteration, with 10 000 iterations for each chain discarded as burn-in. This provided 6000 samples from the posterior distribution for each parameter.

All random effects were given half-Cauchy priors with a scale of $\Theta = 100$, regression coeficcients had vague normal priors with a precision of $1/sigma^2 = 1e-5$, and τ was estimated using a gamma prior with parameters $a = b = 1e-5$.

3 Results

Across all regions of the US, the number of un-assessed fished populations increased steadily since the 1950s and into the 1990s, while an increasing number of assessments in each region lead to a steadily increasing trend in the proportion of landed stocks that are assessed (Figure 1b).

(a) Number of stocks   (b) Proportion of stocks   (c) Proportion of landings
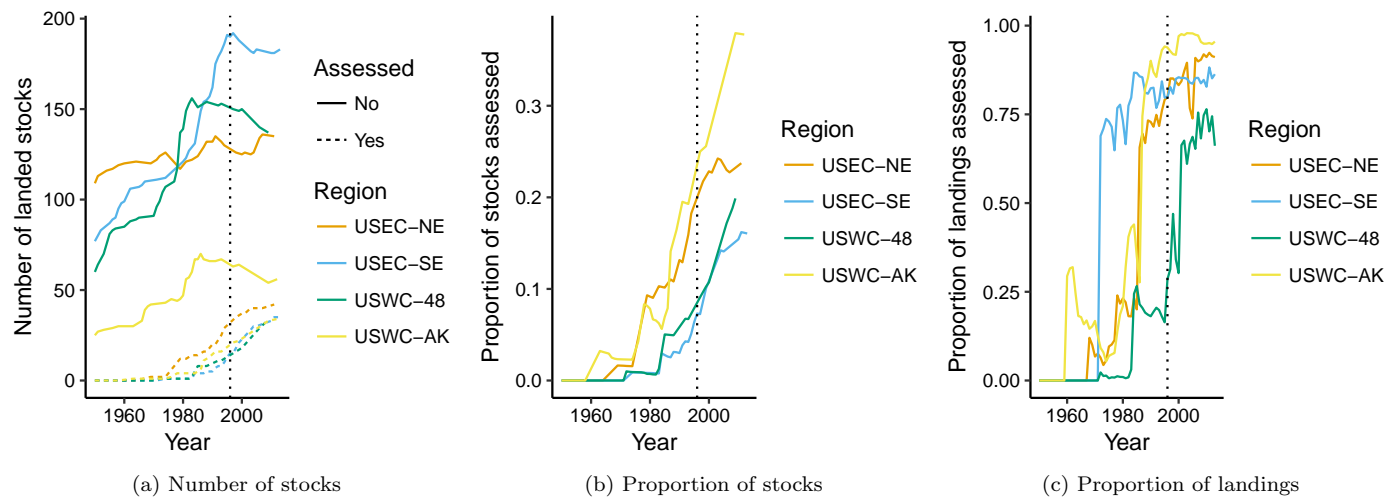
Figure 1: Timeline of a) the number of stocks landed by region and assessment status, b) proportion of landed stocks that are assessed, and c) the proportion of total landings (per year) from assessed stocks. The dotted vertical line marks the enactment of the Sustainable Fisheries Act of 1996
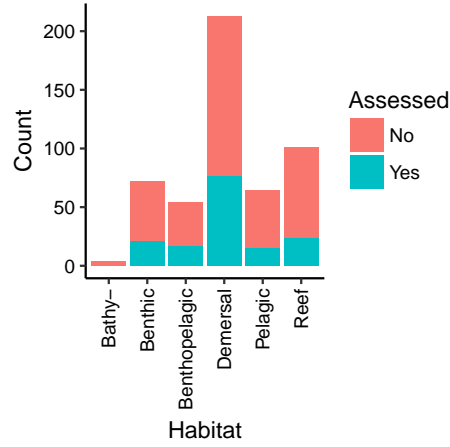
Figure 2: Assessment status at time of last known status (censoring time) by habitat
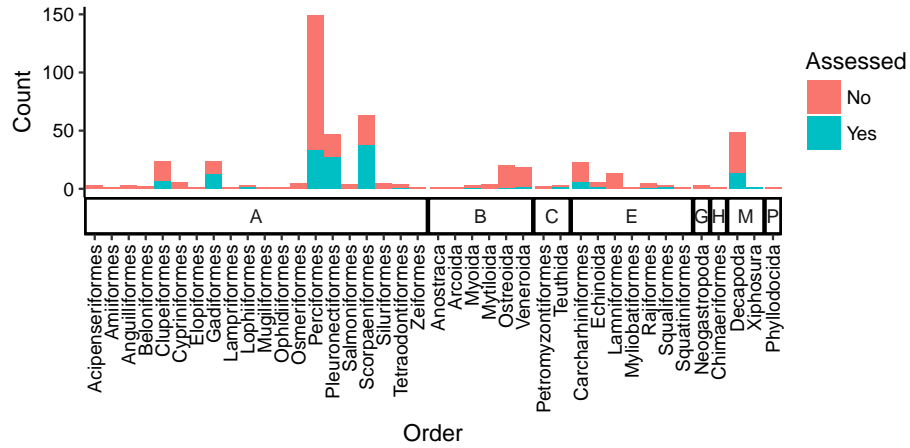

Figure 3: Assessment status at time of last known status (censoring time) by Order and sorted by Class. Classes are abbreviated as A: Actinopterygii, B: Bivalvia, M: Malacostraca, E: Echinoidea, P: Merostomata, G: Elasmobranchii, C: Polychaeta, H: Gastropoda, A: Cephalopoda, B: Cephalaspidomorphi, M: Holocephali, E: Branchiopoda

Table 1: Posterior mean and $P(\theta > 1)$ for model parameters. Parameters can be interpreted as the ratio of rates (rate effect θ, i.e., rates at which stocks with different characteristics are assessed) or as multiplicative acceleration factors (time effect, i.e., $\nu = 0.5$ suggests a stock with these characteristics is assessed twice as fast as the average stock).

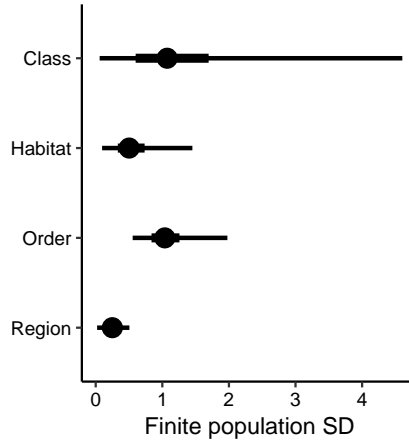| Parameter	Rate effect (θ)	Time effect (ν)	$P(\theta > 1)$
Length	1.06	1.00	0.52
Maximum landings	1.78	0.79	0.99
Mean price per kg	5.39	0.49	1.00

Figure 4: Comparison of finite population standard deviation (i.e., variance attributed to each variable) for random effects in the Weibull survival model. The circle shows the posterior median, with thick bars showing the inter-quartile range of the posterior and the thin line is the 95% confidence interval
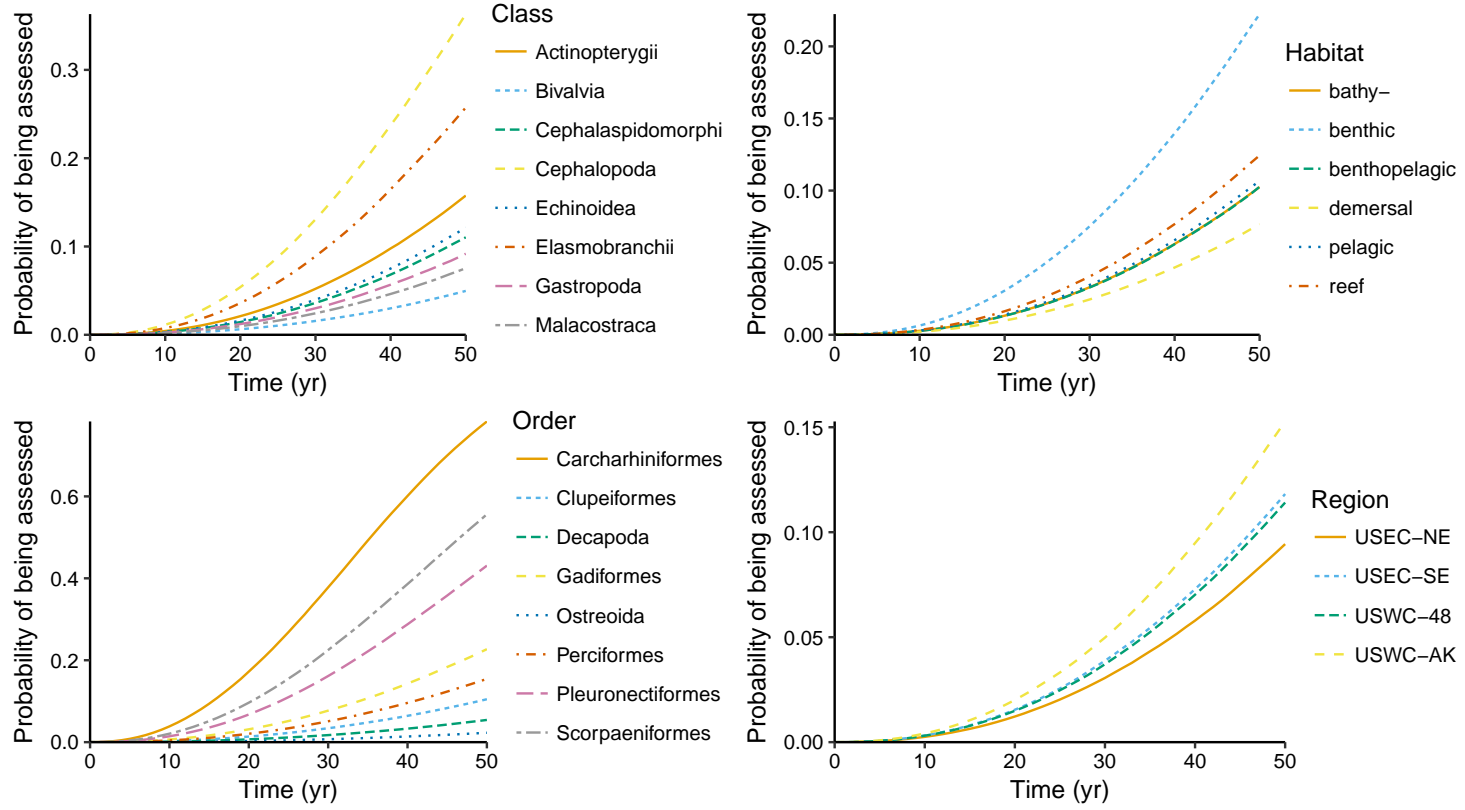
Figure 5: Marginal probability of a stock in category k being assessed as a function of time $(P(T_k \leq t) = F_k(t) = \exp(-\lambda_k t^\tau))$, for stocks of various taxonomic orders, class, regions and habitats. For taxonomic variables, only the eight levels with the most stocks represented in our dataset are shown. Marginal probabilities were evaluated at the mean of (centered) continuous covariates.
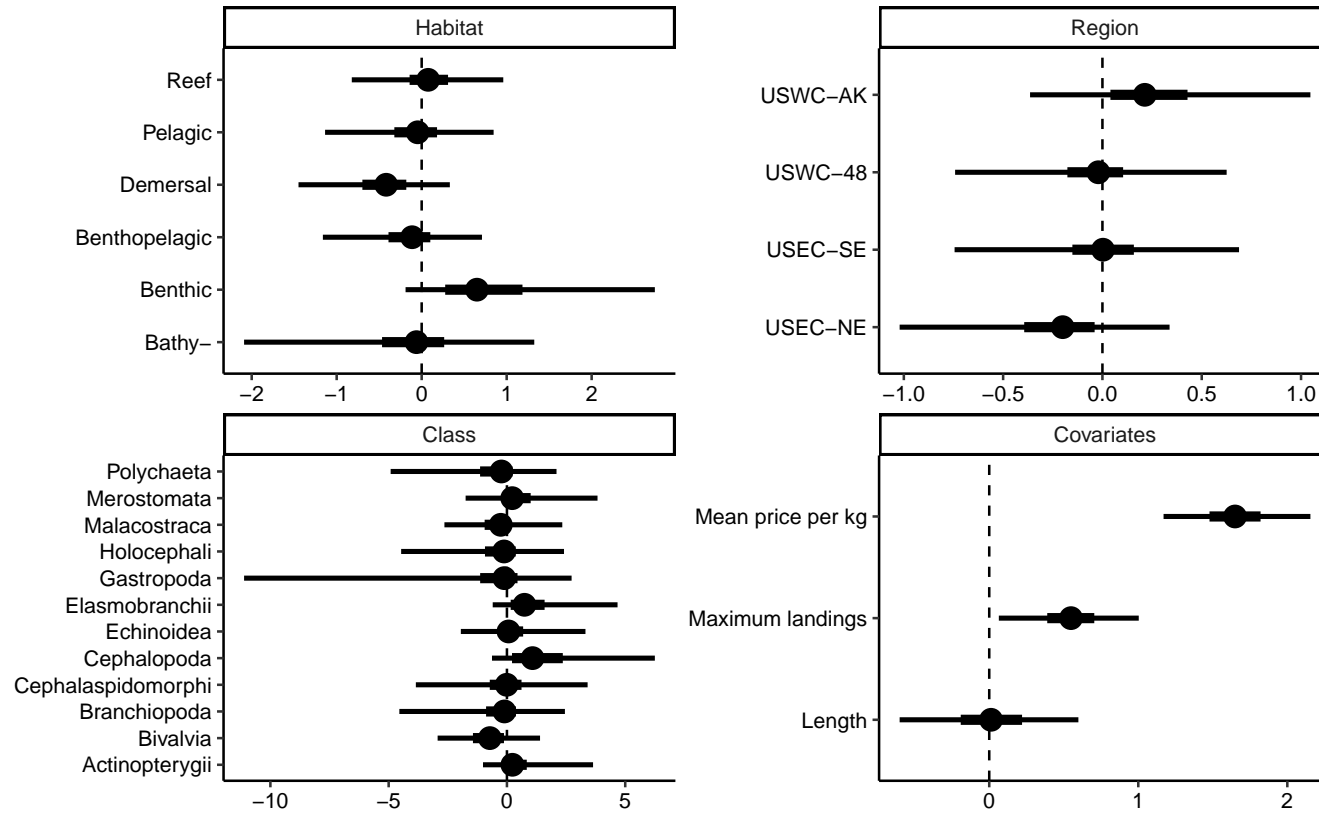
Figure 6: Summaries of estimated posterior distributions for a) continous covariates in the model, b) habitat random effects, c) regional random effects, and d) taxonomic class random effects. The circle shows the posterior median, with thick bars showing the inter-quartile range of the posterior and the thin line is the 95% confidence interval.
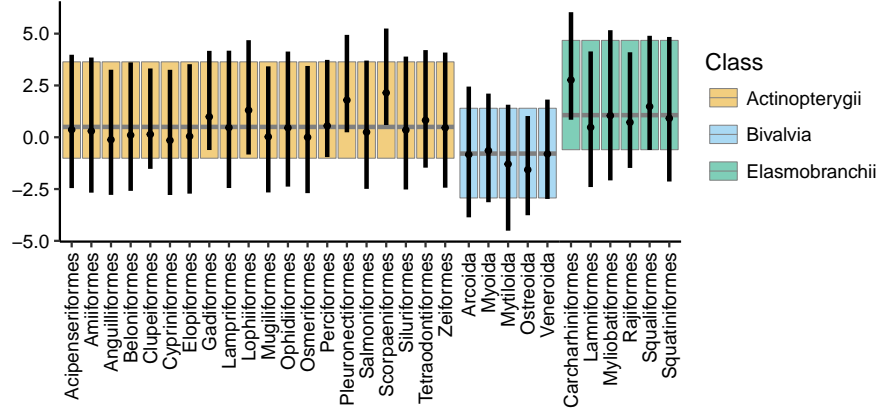
Figure 7: Summaries of estimated posterior distributions for order within class, showing a summary (grey line: posterior mean, coloured box: 95% confidence) of class effects, and order effects relative to class effects (points: posterior mean, black line: 95% confidence), for classes with multiple orders in the dataset.
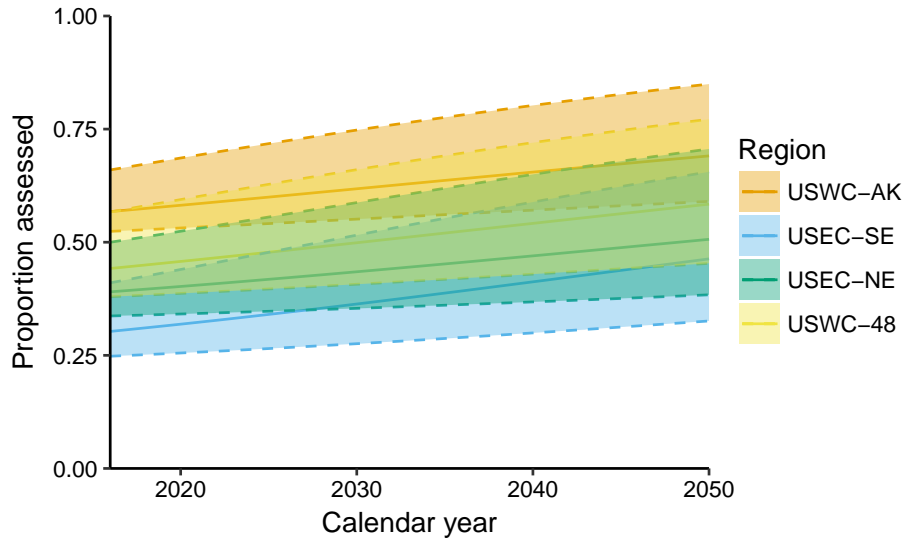

Figure 8: Projected proportion of stocks assessed by region and calendar year, based on assessment probabilities of stocks within each region over the projected year range.

13

A   A: Model fit
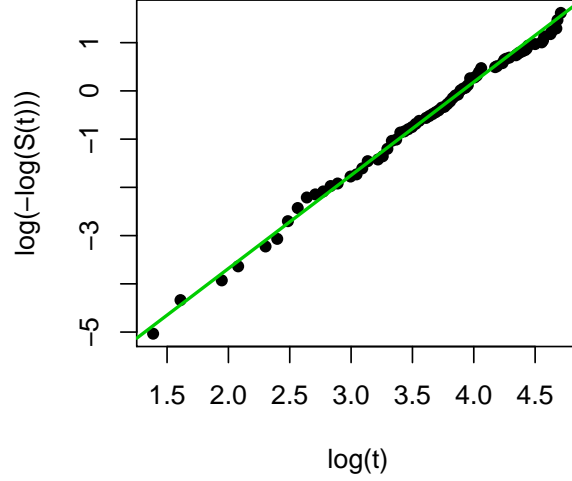
Figure A.1: Appropriateness of the Weibull event-time model for the time-to-assessment dataset. If the Weibull applies, the time from fishery development to assessment should fall on a line with slope p (the Weibull shape parameter) between $log(-log(\hat{S}(t)))$, where $\hat{S}(t)$ is the non-parametric Kaplan-Meyer estimate of survival at time t, and the log of t. Here, p evaluates to 1.93, suggesting an increasing assessment rate with increasing time t.

Figure A.2: Model fit of the Weibull survival model, based on Cox-Snell residuals calculated at the posterior mean of the linear predictor. For a perfect fit all data points (solid points) would lie on the y=x line.
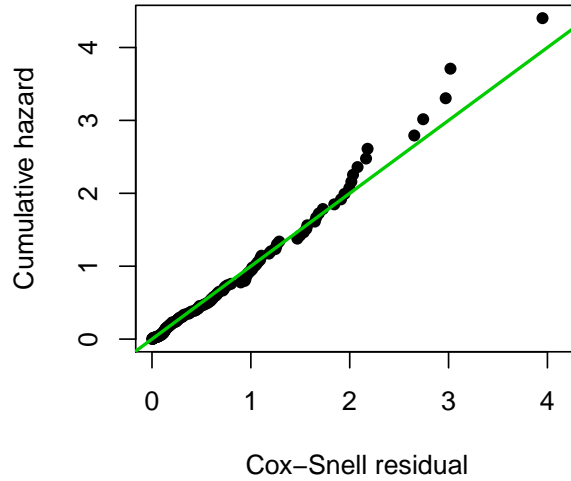
Table 2: Parameter estimates and probability $P(\theta_k > 0)$, the probability that stocks in this category k have a higher than average rate of assessment. Parameters can be interpreted as the ratio of rates between a stock of category k and an average stock (rate effect θ) or as multiplicative acceleration factors (time effect, e.g., $\nu = 0.5$ suggests a stock with these characteristics is assessed twice as fast as the average stock).

Effect	Category	Rate effect (θ)	Time effect (ν)	$P(\theta > 1)$
Covariates	Length	1.01	0.99	0.52
Covariates	Maximum landings	1.73	0.79	0.99
Covariates	Mean price per kg	5.20	0.49	1.00
Region	USEC-NE	0.82	1.09	0.18
Region	USEC-SE	1.00	1.00	0.51
Region	USWC-48	0.98	1.01	0.45
Region	USWC-AK	1.24	0.91	0.82
Habitat	Bathy-	0.94	1.03	0.44
Habitat	Benthic	1.92	0.75	0.92
Habitat	Benthopelagic	0.89	1.05	0.36

Table 2 – continued from previous page

Effect	Category	Rate effect (θ)	Time effect (ν)	$P(\theta > 1)$
Habitat	Demersal	0.66	1.20	0.11
Habitat	Pelagic	0.95	1.02	0.45
Habitat	Reef	1.08	0.97	0.60
Class	Actinopterygii	1.26	0.90	0.66
Class	Bivalvia	0.49	1.37	0.18
Class	Branchiopoda	0.90	1.05	0.43
Class	Cephalaspidomorphi	0.99	1.00	0.49
Class	Cephalopoda	2.96	0.62	0.85
Class	Echinoidea	1.07	0.97	0.55
Class	Elasmobranchii	2.10	0.72	0.84
Class	Gastropoda	0.89	1.05	0.43
Class	Holocephali	0.88	1.06	0.41
Class	Malacostraca	0.77	1.12	0.35
Class	Merostomata	1.25	0.91	0.64
Class	Polychaeta	0.79	1.11	0.37
Order	Acipenseriformes	0.90	1.05	0.46
Order	Amiiformes	0.87	1.06	0.44
Order	Anguilliformes	0.60	1.25	0.26
Order	Anostraca	0.90	1.05	0.45
Order	Arcoida	0.98	1.01	0.49
Order	Beloniformes	0.72	1.15	0.35
Order	Carcharhiniformes	5.12	0.49	0.99
Order	Chimaeriformes	0.83	1.08	0.43
Order	Clupeiformes	0.69	1.18	0.25
Order	Cypriniformes	0.57	1.28	0.25
Order	Decapoda	0.68	1.19	0.31
Order	Echinoida	1.10	0.96	0.55
Order	Elopiformes	0.69	1.18	0.33
Order	Gadiformes	1.55	0.82	0.81
Order	Lamniformes	0.60	1.25	0.27
Order	Lampriformes	0.98	1.01	0.49
Order	Lophiiformes	2.16	0.71	0.84
Order	Mugiliformes	0.68	1.18	0.32
Order	Myliobatiformes	0.97	1.01	0.49
Order	Myoida	1.15	0.94	0.57
Order	Mytiloida	0.70	1.17	0.35
Order	Neogastropoda	0.92	1.04	0.46
Order	Ophidiiformes	1.00	1.00	0.50
Order	Osmeriformes	0.65	1.21	0.31
Order	Ostreoida	0.48	1.39	0.17
Order	Perciformes	1.01	1.00	0.51

Table 2 – continued from previous page

Effect	Category	Rate effect (θ)	Time effect (ν)	$P(\theta > 1)$
Order	Petromyzontiformes	0.95	1.02	0.48
Order	Phyllodocida	0.79	1.11	0.40
Order	Pleuronectiformes	3.43	0.58	1.00
Order	Rajiformes	0.70	1.17	0.32
Order	Salmoniformes	0.81	1.10	0.41
Order	Scorpaeniformes	4.95	0.50	1.00
Order	Siluriformes	0.90	1.05	0.46
Order	Squaliformes	1.45	0.85	0.68
Order	Squatiniformes	0.89	1.06	0.45
Order	Tetraodontiformes	1.36	0.87	0.65
Order	Teuthida	2.68	0.65	0.83
Order	Veneroida	0.97	1.01	0.49
Order	Xiphosura	1.34	0.88	0.63
Order	Zeiformes	0.98	1.01	0.49