SCIENTIFIC COMMITTEE
FIFTEENTH REGULAR SESSION
Pohnpei, Federated States of Micronesia
12-20 August 2019

Alternative Assessment Methods for Oceanic Whitetip Shark

WCPFC-SC15-2019/SA-WP-13

P. Neubauer, Y. Richard and L. Tremblay-Boyer[1]

[1] Dragonfly Data Science, P.O. Box 27535, Wellington, New Zealand 6141

Alternative assessment methods for oceanic white-tip shark

Authors:
Philipp Neubauer
Yvan Richard
Laura Tremblay-Boyer

PO Box 27535, Wellington 6141
New Zealand
dragonfly.co.nz

Cover Notes

To be cited as:

Neubauer, Philipp; Yvan Richard; Laura Tremblay-Boyer (2019). Alternative assessment methods for oceanic white-tip shark, 56 pages. Report prepared for the Western and Central Pacific Fisheries Commission.

EXECUTIVE SUMMARY

The present study evaluates potential alternative assessment methods for sharks, using oceanic whitetip shark (*Carcharhinus longimanus*) as a case study, allowing comparisons with the current age-structured integrated stock assessment of this species. The latter was conducted in parallel to the present study, and used Stock Synthesis 3 (SS3) software.

The most recent previous integrated assessment of oceanic whitetip shark concluded that the stock was overfished and that overfishing was continuing. To minimise ongoing fisheries impacts on this species, a non-retention measure (Conservation and Management Measure CMM 2011-04) was imposed by the Western and Central Pacific Fisheries Commission (WCPFC); however, the non-retention of oceanic whitetip shark also introduced additional uncertainty about the value of indicators such as catch-per-unit-effort (CPUE) for the monitoring of population status. In conjunction with limited data about the efficacy of measure CMM 2011-04 for limiting fishing mortality, the current stock status of this species remains uncertain.

Here, we compared three approaches in conjunction with the current integrated stock assessment of oceanic whitetip shark. These approaches were catch-only simulations, a general spatial risk assessment model, and a Bayesian dynamic surplus production model. We also illustrate the impact of different assumptions on estimates of fishing mortality (F) and risk (F/F_{crash}) to the oceanic whitetip shark stock in the Western and Central Pacific Ocean.

Our findings suggest that catch-only methods are most valuable as a tool to refine Bayesian priors in more sophisticated analyses, as on their own, catch-only methods are dependent on assumptions and provide no relevant management outputs. Nevertheless, we show that by making simple and relatively broad assumptions about the current depletion level, catch alone can constrain initial (unfished and/or starting depletion for the catch time series) population size and productivity and, thereby, serve as a *a priori* constraint on these parameters.

The application of dynamic surplus production models (DSPMs) showed that these model may provide a reasonable tool to rapidly assess shark stocks, either alongside or instead of fully integrated stock assessments. Dynamic surplus production models can be readily applied to sharks: their implementation in widely-available software packages means that they are a cost-effective assessment tool that requires few assumptions. In addition, these models can provide estimates of management-relevant quantities (e.g., stock status, fishing mortality), which have been shown to be robust for sharks. Furthermore, depletion-based catch-only simulations can be used to construct useful priors for Bayesian implementations of these models. Nevertheless, the reliance of DSPMs on a reliable biomass index (e.g., CPUE time series with contrast) and on complete removal estimates (i.e. the availability of a catch series which accurately reflects total catch) limits their application to species for which these time-series data can be derived. This

aspect may exclude the application of DSPMs to species with poor historical identification records such as many shark species.

We also applied a spatial risk assessment (SRA), as this approach only requires recent catch and effort data to estimate fishing mortality, so is less constrained by historical data limitations. Because SRAs generally do not use complete time series of removals, they cannot provide information about stock status. The most commonly employed SRA methods are conceptually similar to fisheries surveys, as they use estimates of gear efficiency to scale observed spatial catch to overall catch via a spatial population density estimate. To derive absolute fishing mortality and risk, however, these methods need to make assumptions about the spatial interaction of the fishing gear with the local population density. This scaling is difficult to establish for longline gear and has a large effect on estimated risk.

For this reason, we suggest that risk assessment methods are employed when 1) no robust time series for catch and CPUE can be derived, and 2) it is possible to make reasonable assumptions about the spatial effect of the fishing gear. Even with these limitations risk assessment methods can be particularly valuable for prioritising assessment and conservation efforts, as they can be readily employed across species in a standardised framework, even for species with limited historical data.

Application of a variety of models to the oceanic whitetip shark stock showed that DSPM, SRAs and SS3 provided similar results, but SRA results were strongly dependent on the assumption of spatial gear effects. All methods suggested that there is a substantial risk that current fishing mortality remains above F_{crash}, the fishing mortality that would lead to extinction in the long term (and by extension, well above F_{lim} and F_{MSM}). The SS3 assessment estimated slightly higher overall fishing mortality and lower productivity and stock status, and therefore provides the most pessimistic view of current fishing mortality and sustainable fishing mortality. All methods suggest that reductions of fishing mortality below likely values in the last year of the assessment (2016; $\approx$ 45% total fishing mortality including haul-back, handling and post-release mortality) would substantially lower existential risks for this stock.

Based on our findings, we suggest that the Scientific Committee considers the following:

- Inferences from different models indicate that oceanic whitetip shark continues to be overfished, and overfishing may still be occurring owing to incidental mortality from fishing, despite non-retention measure CMM 2011-04. Estimated fishing mortality rates for the last year in the assessment (2016) lead to substantial risk that the stock will not persist.
- Spatial risk assessment methods should be employed for species with poor historical records (e.g., poor species identification), but for which recent records are judged reliable. In addition, a standardised methodology based on spatial risk assessment methodology could be

employed to prioritise assessment and conservation efforts.

- Surplus production models can provide a robust cost- and time-effective way to assess shark populations, and provide similar outputs to fully integrated stock assessments such as SS3. Therefore, they may be considered as a rapid assessment tool, either alongside or instead of fully integrated stock assessments, which could be employed for species of high priority.
- Depletion-based catch-only simulations should be considered for constructing priors for DSPMs and to understand the amount of additional information provided by fitting the DSPM.

1. INTRODUCTION

Oceanic whitetip shark (*Carcharhinus longimanus*) occurs globally in tropical and subtropical areas, and frequently features as bycatch in tuna and billfish fisheries worldwide. In the past, this species accounted for a large proportion of total bycatch in Western and Central Pacific Ocean tuna fisheries, suggesting high historical abundance. Nevertheless, precipitous declines have been observed in many abundance indicators of this species worldwide.

In the Western and Central Pacific Ocean, the most recent assessment of oceanic whitetip shark concluded that the stock was overfished (Rice & Harley 2012), even though this conclusion was caveated by uncertainties about input data and biological parameters. Based on this assessment, the Western and Central Pacific Fisheries Commission (WCPFC) adopted Conservation and Management Measure CMM 2011-04, prohibiting retention of oceanic whitetip shark.

An updated catch-per-unit-effort (CPUE) standardisation of catch rates suggested an ongoing decline from the last year of the previous assessment (Rice et al. 2015). In addition, a recent expert review of shark life history in the Western and Central Pacific Ocean further confirmed the considerable uncertainty in the biological parameters for oceanic whitetip shark (Clarke et al. 2015), notably in growth parameters and litter size, which are influential in stock assessments as they relate to stock productivity.

Previous efforts for the reconstruction of catches for this species highlighted low and uneven observer coverage over the temporal and spatial extent of the assessment (Rice 2012, 2018). The adoption of CMM 2011-04 introduced additional uncertainty for the most recent years of the planned 2019 assessment period, as the retention ban could lead to under-estimation of unobserved mortalities given that the requirement to report dead discarded sharks was only agreed by the WCPFC to apply from 2017 onward (WCPFC 2016). Post-release mortality of live captures can also further increase fishing mortality (Common Oceans (ABNJ) Tuna Project 2019). In parallel, Rice (2018) recently reported that CMM 2011-04 had not been adopted comprehensively across fleets, adding to the uncertainty in the predictions of true fishing-induced mortality across fishing fleets. The same report found that the reported catches and discards for oceanic whitetip shark had increased between 2010 and 2015, although this may have been due to increasing attention to accurate reporting of shark catches over time in conjunction with gradual implementation of species-specific reporting requirements.

An updated stock assessment, using an integrated assessment model on the Stock Synthesis 3 (SS3) platform has been conducted for the 15th annual meeting of the WCPFC Scientific Committee (SC 15) to update the stock status and estimates of fishing mortality for oceanic whitetip shark in the Western and Central Pacific Ocean (Tremblay-Boyer et al. 2019). Owing to the above-mentioned uncertainty in data inputs, and corresponding uncertainties about previous assessment outcomes, the Scientific Committee

proposed a project (Project 92) at its 14th annual meeting (SC 14) to compare alternative assessment approaches with the previously applied integrated modelling approach.

A range of models has been applied to sharks in the Western and Pacific Ocean and other regions (Bonfil 2005). While some assessments, such as those for silky shark (Clarke et al. 2018) and oceanic whitetip shark (Rice & Harley 2012) have been based on catch-at-age stock assessment models (Stock Synthesis 3; Methot Jr & Wetzel 2013), other recent assessments have focused on spatial methods, such as risk assessment methods (Neubauer et al. 2018), and hybrid spatial risk-assessment/stock assessment frameworks (Fu et al. 2017, Hoyle et al. 2017). Here, we first describe the general frameworks as they have been applied to shark species in recent assessments for the WCPFC, and compare the different approaches in their assumptions, data requirements and estimated quantities. We then discuss the application of these methods to oceanic whitetip shark, for which we compared a) information provided by catch alone, b) a dynamic Bayesian surplus production model (McAllister & Edwards 2016) and c) a spatial risk-assessment method based on the SAFE method (Sustainability assessment of fishing effects; Zhou et al. 2011) and recently proposed extensions (Zhou et al. 2013, Zhou et al. 2014). These models are compared with data requirements, assumptions and outcomes of the integrated SS3 assessment where possible.

2. METHODS

2.1 Potential assessment methods for sharks

Here, we briefly review methods applied to shark assessments in the Western and Central Pacific Ocean, noting that other methods may provide alternatives (e.g., see in-depth discussion in Bonfil 2005). Where relevant, the latter are referred to here, but the primary focus of the present summary was on methods applied to WCPFC assessments (see Table 1 for the basic data requirements, assumptions and outputs of the different assessment approaches). Included in our review are quantitative assessment methods, representing a Level-3 assessment in the Ecological Risk Assessment for the Effects of Fishing (ERAEF) framework (outlined by Hobday et al. 2011). Qualitative (Level 1) and semi-quantitative (e.g., productivity-susceptibility; Level 2) approaches to (risk-)assessments are not included here; a discussion of the relative data needs and merits of these different levels of (risk-)assessment is provided by Hobday et al. (2011).

2.1.1 Catch-only methods

Most quantitative assessment methods require catch data (except for yield-per-recruit methods; Bonfil 2005). The main difference among methods is how catch data are used (i.e., current catch or time series). In many instances, especially for sharks, catch needs to be scaled from observer data to overall effort (i.e., reported catches often do not provide a reliable estimate

Table 1: Comparison of different approaches to stock assessments reviewed in the present study, showing necessary data inputs, key assumptions, and outputs (CPUE, catch-per-unit-effort, B, biomass; MSY, maximum sustainable yield; F, fishing mortality). Model type distinguishes between models with similar levels of complexity and input requirements. Models and quantities that are used for alternative oceanic whitetip shark assessments in this report are highlighted in bold.

Approach	Model type	Model/package names	Inputs	Assumptions	Outputs
Catch-only methods	Catch abundance	COM-SIR, SSCOM, mPRM	Catch time series	Assumed catch-effort or catch-biomass relationship	Stock status (B/B_{MSY})
	Depletion-based	SSRA, **CMSY**, DBSRA	Catch time series	Current depletion level	**Initial population, productivity**, MSY
Quantitative (Level 3) spatial risk assessment	Spatial contrast	**(e)SAFE**, SEFRA	Current catch and effort	Gear-affected area, species distribution	**Current population, current F**
	Temporal contrast	MIST	Catch and effort time series (sub-area)	Catch known, CPUE reflects abundance, productivity summarised in single rate, catchability transferable across space	Population density, current and past F
Dynamic surplus production (SP) and delay difference (DD)	Dynamic SP	**DBM**, JABBA	Catch and effort (or abundance index) across assessment area	Catch known, CPUE reflects abundance, productivity summarised in single rate	**Stock status, current and past depletion and F**
	DD models		Catch and effort (or abundance index) across assessment area, biological data	Catch known, CPUE reflects abundance, representative biological data (e.g., growth)	Stock status, current and past depletion and F
Integrated (age/size structured) stock assessments		SS3, CASAL, MULTIFAN	Catch and effort (or abundance index) across assessment area, composition data, biological data	Catch known, CPUE reflects abundance, composition data useful to estimate selectivity, recruitment	Stock status, current and past depletion and F

of the number of captured sharks), and total catch in itself is, therefore, uncertain. Some methods attempt to use catch time-series data alone as a proxy for abundance trends or to constrain certain management-relevant parameters, such as maximum sustainable yield (MSY). These methods have not yet been applied to sharks in the Western and Central Pacific Ocean, but were proposed for inclusion in the present review by SC 14 and are, therefore, included here.

Methods that use only catch data usually have to make strong assumptions about how catch trends relate to either abundance directly, or to fishing effort (i.e., effort dynamics). The latter assumption effectively imputes effort to generate catch-per-unit-effort (CPUE) and to estimate relative abundance trends. Both assumptions are strong in the sense that the assumptions about the relationship between catch and abundance or catch and effort directly dictate the estimated abundance trend. This aspect can lead to strongly-biased inferences for any stock alone, even if the methods may perform well on average (Carruthers et al. 2014). Super-ensemble models, which pool estimates from a range of values derived from catch-only methods, can address this shortcoming to some extent (Anderson et al. 2017), but require the fitting of a range of models (and deciding which models to use) and deciding on relative weights.

An alternative approach is to supplement catch data with assumptions about the stock status itself (called "depletion based methods" sensu Carruthers et al. 2014). Although stock status is an assumed rather than an estimated quantity in these methods, it is possible to obtain valuable estimates of other management quantities (e.g., MSY; Martell & Froese 2013, Froese et al. 2017). Stock status assumptions are also used in stochastic stock-reduction-based methods (Walters et al. 2006), such as depletion-based stock reduction analysis (Dick & MacCall 2011).

Depletion-based methods make use of the aspect that catch is often assumed to be a known quantity in stock assessments: it is not data that are fitted to, but rather catch acts as an external constraint on the model. This aspect makes catch particularly useful to *a priori* (in a Bayesian sense) constrain historical and present absolute stock abundance and productivity in assessments of any type. In addition to assumptions about current depletion status, obtaining an *a priori* constraint on absolute population size requires a prior distribution on productivity (e.g., intrinsic population growth rate for a surplus production model, or growth, natural mortality and recruitment for integrated models such as Stock Synthesis 3). With these assumptions, it is possible to evaluate (i.e., by simulating from a population dynamics model) the combinations of population size and productivity that lead to assumed current depletion levels under known catches. The catch-constrained absolute population size and productivity can then be used to derive MSY under a given population dynamics model.

Although depletion-based catch-only simulations can be used to constrain the overall population size, the assumption about current stock status, in conjunction with observed catches, leads to an implicit prior on fishing mortality and, thereby, implies a prior on risk of overfishing. These methods

are, therefore, most valuable in the context of generating informative priors for more sophisticated assessments. For example, we applied a depletion-based method to oceanic whitetip shark, but instead of using the outputs from these simulations as standalone estimates, we used this method to generate prior distributions for a surplus production assessment of this species. This approach recognises that catch and assumptions about stock status are *a priori* constraints on population size and productivity (Froese et al. 2017, Walters et al. 2006), but are most useful if complemented by information that constrains fishing mortality and stock status *a posteriori*. The latter constraint is applied by catch-per-unit-effort data in risk assessments or stock assessments models (Walters et al. 2006, Carruthers et al. 2014, Froese et al. 2017).

2.1.2 Spatial risk assessment methods

Spatial risk assessments (SRAs) augment catch data with spatially resolved effort data and life-history reference points to calculate fishing mortality and sustainability risk (e.g., in the form of the risk-ratio, which sets F relative to a life-history-derived reference point that measures sustainability risk). Risk assessment methods usually only rely on recent catch and, therefore, do not require a complete time series of catch data, nor assumptions about catch being complete or CPUE time series data being unbiased indicators of relative abundance. The lack of temporal contrast to constrain the model is then compensated in one of two ways.

The first way in which a SRA model can derive fishing mortality is via an *a priori* constraint on either the current population size N_{curr} (e.g., Neubauer et al. 2018) or gear efficiency (Q, which is directly related to catchability q, see Appendix subsection A.1; Zhou et al. 2013). For example, for seabirds and other species such as some marine mammals, estimates of N_{curr} are often available from census counts or abundance surveys (e.g., mark-recapture or distance-sampling surveys). A different way to constrain N_{curr} is from sensitive genetic markers through either genetic mark-recapture or population genetics, although the latter often requires strong assumptions or allowing for large uncertainties (Neubauer et al. 2018). In addition, population genetic estimates of population size pertain to a population as a whole (e.g., the Indo-Pacific Ocean population of whaleshark, rather than the Western and Central Pacific Ocean stock of this species), whereas a region-specific assessment (e.g., an assessment of the stock within the Western and Central Pacific Ocean) may only want to consider a proportion of the overall population. In this case, one can consider using the spatial density (relative abundance) of the overall population to estimate the stock that resides within the Western and Central Pacific Ocean. It is then possible to consider the latter as a closed sub-population (stock), and to examine the risk for this sub-population. Alternatively, one can consider the risk posed by fisheries to the sub-population (e.g., within the Western and Central Pacific Ocean) to the total (e.g., Indo-Pacific Ocean) population (e.g., Neubauer et al. 2018), or expand the risk assessment by considering fishing effort from other regions (e.g., as in the porbeagle shark assessment;

Hoyle et al. 2017). Although the former approach of assessing the sub-population only has more direct relevance for management within the Western and Central Pacific Ocean, the latter approach of assessing impact from fishing on the sub-population for the total stock makes fewer biological assumptions about population structure and potential movements within the larger population.

Similarly, methods that assume known gear efficiency, such as the SAFE method (Zhou et al. 2011), can scale directly to a total population size and fishing mortality: this method acts as a fisheries survey in that estimated mean densities for the gear-affected area (the area that is effectively fished by a single unit of effort; e.g., the area that is affected by a single longline set or hook, depending on the unit of effort) are scaled to the overall habitat area (e.g., suitable habitat within the Western and Central Pacific Ocean).

If current population size or gear efficiency cannot be constrained *a priori*, they can, in theory, be estimated together from direct estimation by using contrast in repeated measures of spatial catch and effort, together with a measure of gear-affected area (e.g., by using the N-mixture model used in the eSAFE method; Zhou et al. (2013), Zhou et al. (2014)). The latter has a clear advantage over CPUE trend-based estimation of populations (e.g., in surplus production or catch-at-age models) as it is only necessary to assume that current CPUE in space represents the spatial distribution of the focal species.

The gear-affected area is a key parameter in this framework, as it is a direct factor for scaling estimated densities at the spatial scale of effort units to population size across the total habitat. The gear-affected area may be constrained *a priori* for some gear types (e.g., trawl); however, for longlines, the area that is effectively fished by each set or hook is difficult to estimate; it depends not only on the gear characteristics (e.g., depth), but also on the physical conditions (e.g., currents) and the focal species' sensory and swimming capacity, and its level of attraction to the bait (Zhou et al. 2014, Jordan et al. 2013). While spatial quantitative risk assessments can constrain fishing mortality and risk while using only recent catch and effort data, they do require strong assumptions about the gear efficiency and/or the spatial area that is affected by each unit of effort.

2.1.3 Surplus production models

Dynamic surplus production models (DSPM) are fitted based on state-space equations (McAllister & Edwards 2016, Froese et al. 2017) and do not require equilibrium assumptions that make traditional approaches to surplus production assessments difficult to justify (Bonfil 2005). Examples of packages that implement DSPMs are JABBA ("Just Another Bayesian Biomass Assessment"; Winker et al. 2018) and BDM ("Bayesian biomass dynamics model" Edwards 2017). The DSPMs tend to use an index of abundance (usually CPUE) to constrain the time series of abundance. These models (and other stock assessments) are, therefore, dependent on CPUE being a reliable index of abundance, an assumption that is often violated. In

addition, they rely on either a full time series of catch or a strong assumption about the depletion level at the start of the time series (the initial depletion). Although productivity is usually estimated within DSPMs, it is usually useful to also constrain productivity via an informative prior (Edwards 2017).

Despite these dependencies on temporally resolved data, their simplicity and relative robustness make DSPMs an appealing framework for shark assessments, where data of biological processes such as growth and recruitment are often sparse (Bonfil 2005). In addition, the direct integration of DSPMs with many theoretical frameworks can be appealing in that productivity parameters can be derived using life-history theory (see below). For these reasons, DSPMs offer a cost-effective alternative to fully-integrated stock assessments, and provide many of the same outputs that can be used for management: being based on a dynamic model, one can easily derive F relative to reference points, and inspect current and/or future stock status under alternative scenarios of fishing mortality.

In the Western and Central Pacific Ocean, DSPMs have not been applied to stock assessments by themselves, but in a hybrid framework in two instances: for porbeagle shark (Hoyle et al. 2017) and bigeye thresher shark (Fu et al. 2017). In both instances, a Bayesian DSPM was fitted to data over a small (calibration) area, and the estimates of catchability were then used to estimate fishing mortality across the larger overall habitat. This approach is conceptually similar to SRAs, but also requires the same assumptions as a surplus production model applied on its own (at least in the calibration sub-area). Nevertheless, it provides additional realism in that it is not necessary to assume that the fishery affects the stock in a homogeneous way across the stock's habitat. Additional biological realism may also be achieved by modelling productivity in more explicit terms (i.e., via delay difference models that represent growth, natural mortality and recruitment, but do not require size or age composition data; Bonfil 2005).

2.1.4 Integrated assessment methods

In contrast to surplus production models, which constrain stock status and fishing mortality by catch and abundance indices alone, integrated assessments such as statistical catch-at-age models (e.g., Stock Synthesis, CASAL) will add further constraints on productivity via composition data (e.g., age-composition data). Other additional data sources can be integrated to further constrain growth, recruitment or movement dynamics (e.g., length-frequency data, growth data, tagging data). The integration of different data sources to constrain processes within the model makes integrated assessment models more biologically realistic and potentially more apt at describing population dynamics and, therefore, at obtaining unbiased estimates of population status and of risk of overfishing. For example, age or size composition data can help resolve dynamics due to transient age and/or size structures and, thereby, lead to reduced bias (e.g., Punt & Szuwalski 2012).

Nevertheless, the added complexity of integrated stock-assessment models and potential conflicts between data sources can also make integrated assessments considerably more difficult to apply and interpret than surplus-production or delay-difference models. For instance, composition data, while being potentially useful, can also be a substantial source of bias if the processes that produce these data are not adequately modelled (Maunder & Piner 2017, Minte-Vera et al. 2017, Francis 2011). For this reason, composition data are often down-weighted *a priori* (i.e., before fitting the assessment model), and often provide little, if any, information about biomass trends, though they may still inform selectivity and recruitment (e.g., Clarke et al. 2018). This aspect is particularly relevant for relatively data-poor species such as most sharks, where composition data are usually sparse and have little weight (or information content) in integrated models (Rice & Harley 2012, Clarke et al. 2018). Similar to its more simple surplus-production and delay-difference counterparts, the temporal contrast in CPUE is often the dominant influence on estimated assessment trends Ludwig and Walters 1985, Maunder and Piner 2017, Minte-Vera et al. 2017, Francis 2011. For both types of assessments (production versus catch-at-age), the assumption that standardised CPUE is an informative indicator of biomass trends is, therefore, a key assumption.

2.2 Application to oceanic whitetip shark

In this section, we applied depletion-based catch-only methods, surplus production models and SRA models to oceanic whitetip shark in the Western and Central Pacific Ocean. To ensure comparable outcomes, we assumed a closed population for all models (as is usual in single-stock assessment models such as surplus production models or single stock Stock Synthesis models), including the risk-assessment model, noting that the population structure of oceanic whitetip shark is poorly resolved, especially in the Pacific Ocean, and that the assumption of a single closed sub-population is a technical convenience rather than a biological reality.

2.2.1 Choosing reference points: estimating R_{max}

A reference point is a benchmark that can be used to measure the status of a species or impact of fishing on a species. Depending on the type of reference points, these may be benchmarks that should be avoided (i.e., a limit reference point) or a target ot be achieved (i.e., a target reference point). Clarke and Hoyle (2014) and Zhou et al. (2018) evaluated methods to derive reference points for elasmobranchs in the Western and Central Pacific Ocean. Both studies proposed a number of potential reference points, and suggested a variety of methods to derive them; however, to date, there are no formally agreed reference points for sharks in the Western and Central Pacific Ocean. Given the small current relative population size of oceanic whitetip shark, we compared current fishing mortality, estimated from different methods to F_{lim} as a tentative limit reference point for sharks, and to F_{crash}, the fishing mortality that would lead to extinction in the long term. If one assumes a simple Schaefer surplus production model,

then $F_{crash} = R_{max}$, the maximum population growth rate (intuitively, a population cannot be sustained if fishing removes more individuals than the population can maximally produce), and $F_{lim} = 0.75R_{max}$. Accordingly, we defined the risk of overfishing as the fishing mortality relative to R_{max} (Zhou et al. 2018). Risk in terms of alternative fishing mortality reference points (e.g., maximum sustainable mortality; MSM) are products of the definition used here (e.g., $F_{MSM} = 0.5F_{crash}$; Zhou et al. (2018)).

Population growth R_{max} was calculated from methods in Pardo et al. (2018) based on the Euler-Lotka equation (see also Zhou et al. 2018), adjusted for survival to age at first maturity (Pardo et al. 2016). Estimating R_{max} serves a dual purpose here: it can act as a reference point for depletion-based catch-only and SRA methods that cannot estimate stock productivity independently, but can also act as a prior for a DSPM for which R_{max} is the productivity parameter.

Life history input values for the Euler-Lotka equation were compiled from ranges and point estimates reported in Clarke et al. (2015), with some adjustments to accommodate recent growth and aging studies (D'Alberto et al. 2017, Joung et al. 2016). Specifically, although a low number of mature individuals were measured, D'Alberto et al. (2017) found a much higher age-at-maturity for oceanic whitetip shark sampled from Papua New Guinea than those found previously in other regions. For each life-history input, values were simulated from a distribution that encompassed reported values in Clarke et al. (2015), D'Alberto et al. (2017), and Joung et al. (2016), with a mode or mean centred on reported means or modes where available (see Figure 1 for the simulated inputs and Figure 2 for the resulting value of R_{max}). When only ranges were reported, the distributions were constructed to encompass those ranges as extreme quantiles (i.e., near the 5th and 95th percentile).

We also integrated over methods to derive natural mortality in the simulation procedure. Specifically, we used methods described in Jensen (1996) (age-at-maturity based), Hewitt and Hoenig (2005) (maximum age based), and Pardo et al. (2016) (expected life-span derived) by simulating R_{max} from the inputs under these mortality assumptions and combining the outputs. This lead to a broader distribution for R_{max} than would be obtained if one considered a single method to estimate natural mortality. The calculated R_{max} included high values compared with values reported in Clarke et al. (2015), but was similar to ranges reported for OCS in Pardo et al. (2016) and point estimates were near identical to those reported from combining methods to estimate F based reference points in Zhou et al. (2018).

2.2.2 Input data

We used predicted total catch in space and time as reconstructed in Tremblay-Boyer and Neubauer (2019). Catch was reconstructed over a time-span of 1995–2016, over a geographic extent from 30° S to 30° N. Catch reconstructions were based on observer data for the Western and Central

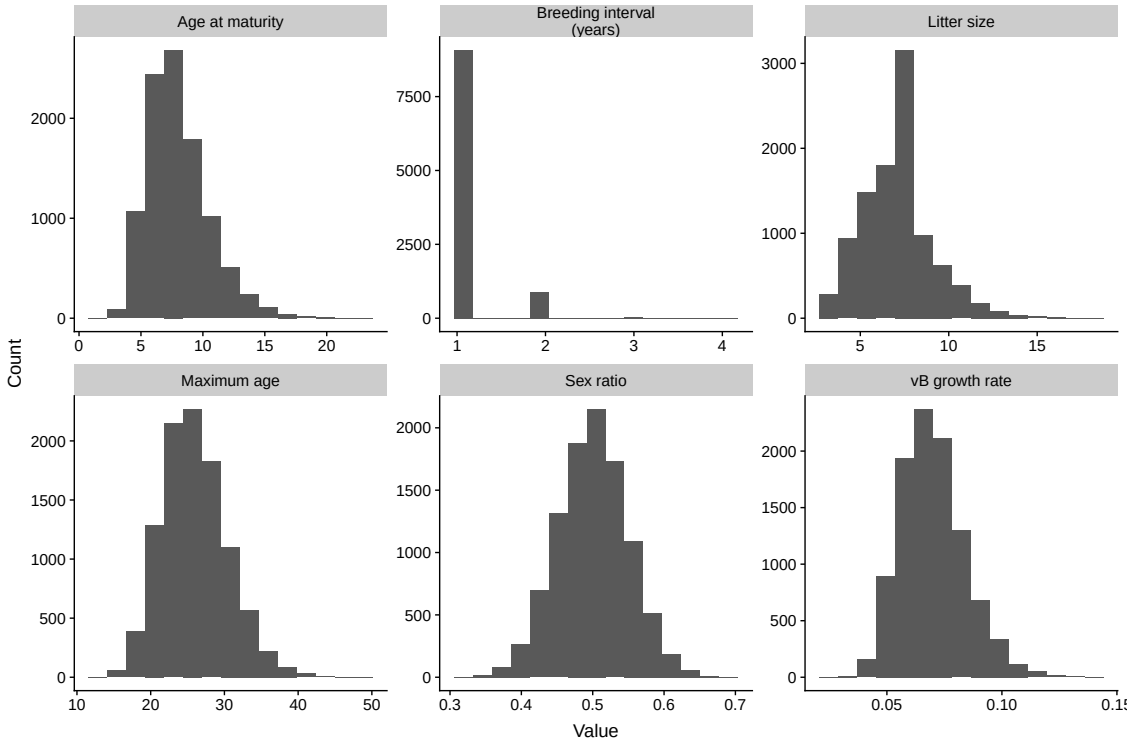

Figure 1: Input values for R_{max} simulations for oceanic whitetip shark, based on parameter values reported in the literature (vB, von Bertalanffy).

Pacific Ocean and used negative-binomial generalised linear mixed models (GLMMs). Briefly – observed interactions were modelled as a function of effort (measured in number of hooks), flag state, targeting and fishing practice (deep vs. shallow sets), and oceanographic variables. The model was then used to extrapolate interactions (captures) from observed effort to total fishing effort. Here, we used the median predicted catch from the best of a set of candidate models as the reference case (the model-predicted trajectories are shown in Figure 3, with spatial predictions shown in Appendix A, Figure B-2).

As retention is prohibited under CMM2011-04, we subtracted live discards from estimated catches, based on the model for reported observer fate codes, with additional assumptions about mortality of discarded sharks. We used three scenarios for discard and post-release mortality (Figure 4):

1. 100% mortality of all retained and discarded sharks, henceforth called "MedianDM100" scenario, where median corresponds to the estimated catch level and DM stands for discard mortality. 100% discard mortality corresponds with assumptions in the previous (2012) stock assessment, providing an upper bound for fishing mortality),

2. 44% total mortality applied on discards (i.e., including haul-back, handling and post-release mortality; henceforth called the "Medi-

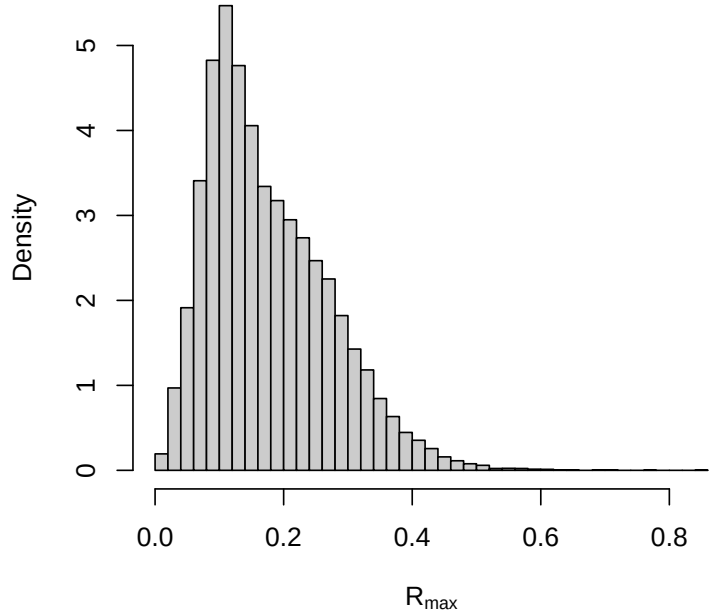

Figure 2: Simulated distribution over R_{max} for oceanic whitetip shark using distributions over input parameters shown in Figure 1.

anDM44"). Discards were taken as median-estimated total discards modelled from observer fate codes. 25% haulback and handling mortality was applied oceanic whitetip shark discards (T. Peatman, cited in Common Oceans (ABNJ) Tuna Project 2019), with additional post-release mortality applied to obtain total mortality corresponding to the lower bound for total mortality found for other species in Common Oceans (ABNJ) Tuna Project (2019). The choice of the lower bound for total mortality for OCS takes into account that oceanic whitetip shark appear to be more frequently released alive than most other species (T. Peatman, cited in Common Oceans (ABNJ) Tuna Project 2019). This scenario was judged the most plausible given available information on post-release mortality of pelagic sharks caught on longlines,

3. 25% mortality applied oceanic whitetip shark discards (T. Peatman, cited in Common Oceans (ABNJ) Tuna Project 2019), and no post-release mortality applied (i.e., to provide a lower bound on fishing mortality, henceforth called "MedianDM25" - all live discards are assumed to survive).

These mortality scenarios adjust recent trends in fishing mortality, but do not vary the absolute catch level across the time series. The latter aspect is inconsequential as assessment models will generally adjust the population

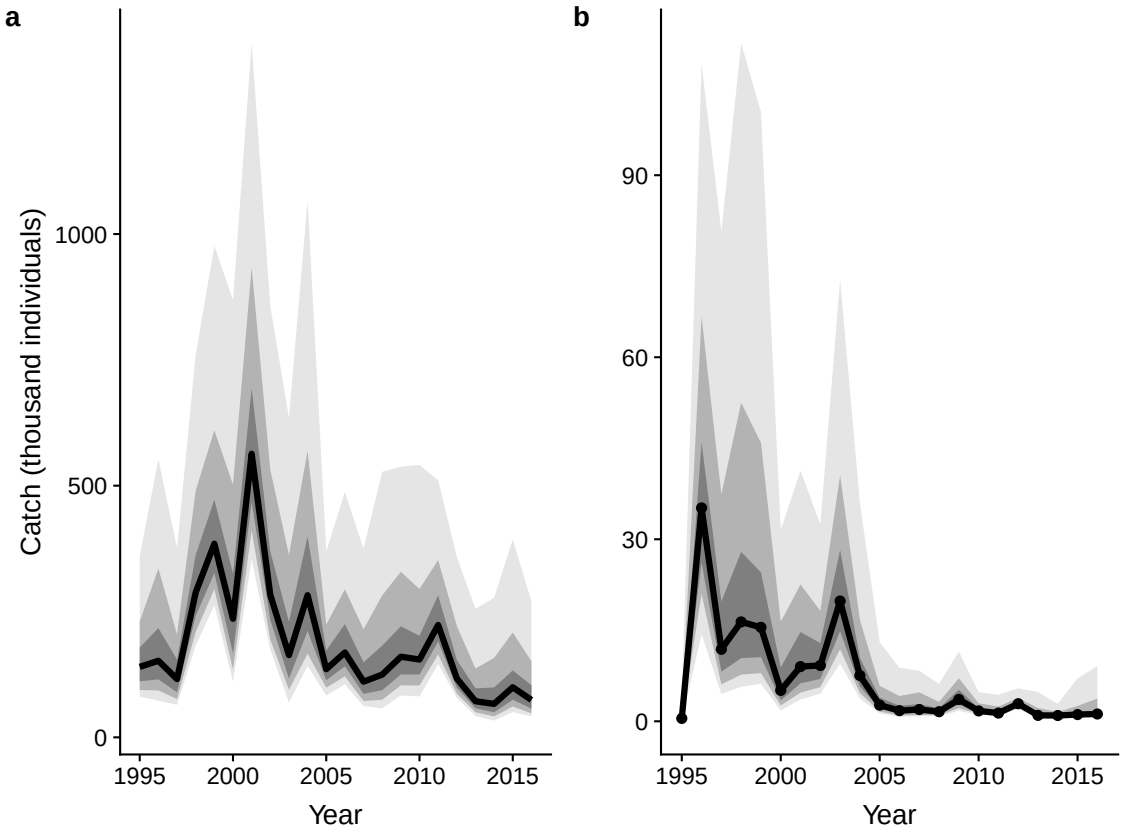

Figure 3: Model-predicted total catch of oceanic whitetip shark (oceanic whitetip shark) by year from (a) longline and (b) purse-seine fishing (grouping associated and un-associated sets. (c) Estimated catch-per-unit-effort (expected oceanic whitetip shark captures per hook).

size to compensate for additional removals (i.e., only changing the trend in catches provides different overfishing-status-relevant information to the model). We, therefore, focused on scenarios with different mortality trends (i.e., the discard mortality scenarios detailed above).

CPUE was analysed with an analogous model (Tremblay-Boyer & Neubauer 2019), but using only a subset of consistently well-observed fisheries (i.e., sub-setting the observer programmes), and modelling on a finer spatial scale with more specific gear covariates (Figure 5).

2.2.3 Applying depletion-based catch-only models: prior constraints on population size and productivity

As mentioned above, catch-only simulations based on prior assumptions about depletion, may be most useful in the context of deriving prior constraints on population size and productivity. Th assumption about the current depletion level can be vague; however, it is likely that catches on the order of magnitude observed for oceanic whitetip shark have had some impact, especially given the life-history traits of shark species. For this reason, current abundance (or biomass) is likely smaller than it was

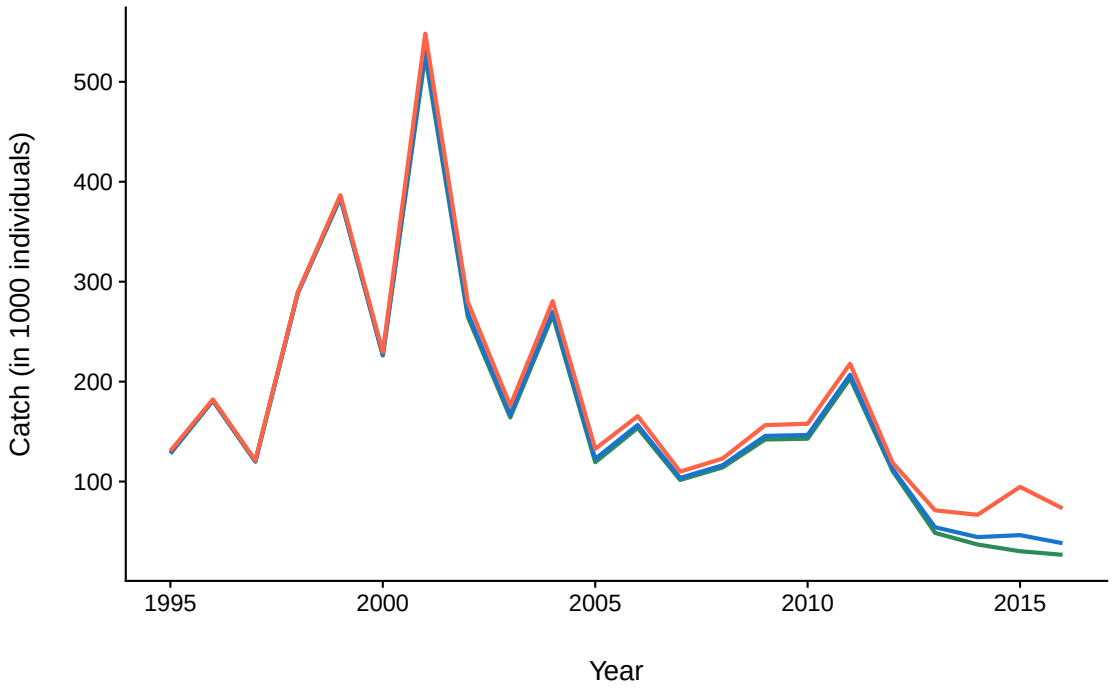

Figure 4: Estimated catch (black line) of oceanic whitetip shark, accounting for estimated discards adjusted for 25% discard mortality (green line), and catch with a total of 44% discard mortality, including haulback, handling and post-release mortality (blue line).

historically. A conservative (i.e., not overly informative) assumption may be that i) current abundance is less than at the start of the catch time-series in 1995 (e.g., , $N_{2016}/N_{1995} < 0.9$) and ii) current abundance is greater than 0 (i.e., based on the observation that catches are not zero).

Based on these simple assumptions, a population model (e.g., a Schaefer surplus production model) was used to simulate population trajectories subject to observed catches from a wide prior on initial abundance and stock productivity (Martell & Froese 2013). Stock trajectories that do not meet the *a priori* constraints on stock status (i.e., $N_{2016}/N_{1995} < 0.9$; $N_{2016} > 0$) are then discarded (filtered). This method relies on a complete catch time series unless initial depletion can be specified. For oceanic whitetip shark, we assumed an initial depletion level (in 1995) between 5% and 80% of the pre-fishing abundance ($0.05 < N_{1995}/K < 0.8$). We then used the sample of filtered trajectories to generate a prior for the the pre-fishing biomass (K), initial depletion (N_{1995}) and productivity (R_{max}) in the surplus production model.

2.2.4 Applying surplus production models to oceanic whitetip shark in the WCPFC

We applied the Schaefer surplus production model implemented in the *bdm* R package (Edwards 2017) to catch and CPUE data for oceanic whitetip shark in the Western and Central Pacific Ocean. We used a classic Schaefer

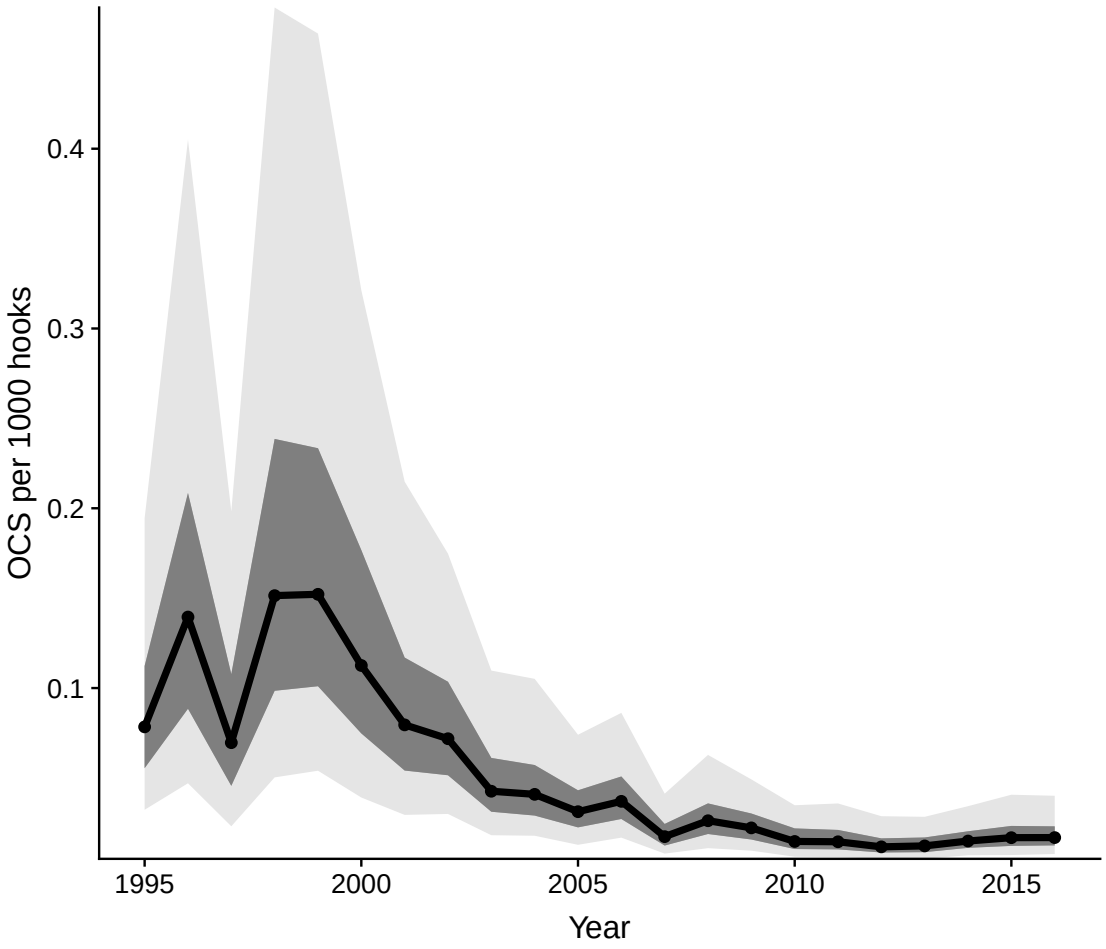

Figure 5: Estimated catch-per-unit-effort (expected oceanic whitetip shark captures per hook).

production model, although other hybrid production functions can be used with this R package. The population dynamics are parametrised in terms of the relative depletion ($x_t = N_t/K$), with relative harvest H_t also expressed in relative terms ($H_t = C_t/K$):

$$x_{t+1} = x_t + g(x_t) - H_t \tag{1}$$
$$g(x_t) = R_{max} x_t \left(1 - x_t\right). \tag{2}$$

Priors for intrinsic population growth R_{max} were derived as a log-normal distribution fitted to simulated, filtered R_{max} values from the depletion-based catch-only simulations. The prior for the carrying capacity K was derived as a log-normal distribution from the prior simulations (outlined above). Importantly, the period for predicted catch starts only in 1995, when catches were relatively high (Figure 3). The latter suggests that fishing mortality prior to 1995 may have been similarly high and, therefore, the initial depletion level is highly uncertain. A prior for initial depletion was

set as log-normal prior based on retained prior-predictive draws. We also included a sensitivity to investigate the sensitivity of conclusions on the assumed initial depletion by roughly halving the prior mean for initial depletion.

CPUE was identical to the index used in the integrated stock assessment (Tremblay-Boyer et al. 2019). CPUE observation error was calculated from the standardised CPUE.

All estimation was done within the bdm package, with Markov Chain Monte Carlo (MCMC) in the underlying Bayesian estimation software Stan (Stan Development Team 2018) used to estimate parameters. We ran the MCMC for 100 000 iteration, discarding the first 10 000 iterations as burn-in, and keeping 100 samples from each of 4 chains. The package allows for efficient set-up of surplus production models, and the sampler was fast, taking about 4 minutes for 100 000 iterations.

2.2.5 Spatial risk-assessment for oceanic whitetip shark in the Western and Central Pacific Ocean

To compare results from the surplus production model to a risk assessment approach, we found that we needed to use model-based estimation of population size: There was no sensible way to constrain either N_{curr} or gear efficiency Q *a priori* for oceanic whitetip shark in the Western and Central Pacific Ocean. A previous study estimated a range of 19 200–191 000 individuals for the historical effective population size (N_e; Ruck 2016); however, these values are not estimates of current census population size, and seem inconsistent with observed and estimated catches over time (estimated at >500 000 for some years). Historical census population size (N_c) may be an order of magnitude (for long-lived species with low natural mortality M, such as whales) to two orders of magnitude (for species with high M such as most fishes) higher than the historical N_e (Dudgeon et al. 2012), and may contain little information about current population size. In the absence of robust information of stock structure and present-day effective population size in the Indo-Pacific, it appears difficult to use genetics to constrain oceanic whitetip shark population size.

An alternative way to specifying N_{curr} or gear efficiency Q was proposed by Zhou et al. (2013): their method extended N-mixture models that are commonly used in terrestrial surveys (Royle & Dorazio 2006, Raftery 1988) to the estimation of N_{curr} in the fished area and Q, based on repeated fishing events on a spatial grid. To apply this method, it is essential to obtain both contrast in Q (i.e., by including different gear types or fleets), and repeated trials in each area (Zhou et al. 2014). Here, we extended this method to simultaneously estimate Q and oceanic whitetip shark density over the total habitat area of oceanic whitetip shark in the WCPO (i.e., assuming a closed population in the WCPO). To derive total current population size and current fishing mortality, population density and gear efficiency are combined with an estimate of the area affected by each unit of effort (the "gear-affected area").

Deriving an estimate of the gear-affected area for longline fisheries is difficult for a number of reasons. Although the efficacy of longlines is known to vary with gear depth and hook spacing, little is known about the spatial area that is effectively fished by the gear (e.g., by a single set or hook Zhou et al. 2014, Jordan et al. 2013). Jordan et al (2013) suggest that olfactory attraction may be on the order of kilometres, but this attraction is likely to depend on currents (advection and diffusion), soak time, and attractiveness of bait. Zhou et al. (2013) used a band of 1 km width as the distance of attraction for a number of elasmobranch species (i.e., 500 m either side of the longline). Here we use a range of potential attraction distances: 500 m, 1 km, 2 km and 10 km width, noting that a distance between 2 km and 10 km corresponds to values suggested in Jordan et al (2013) and appear more likely a priori than short attraction distances given oceanic whitetip shark biology. We will use a value of 2 km as a base-case here, but note that higher values are possible.

As no information of longline characteristics was available, we assumed a hook spacing of 50 m on average. This spacing is potentially slightly larger than the hook spacing for most tuna longlines, but smaller than the spacing for shallow billfish sets (e.g., Bigelow et al. 2006). Hook spacing was multiplied by the width of the affected area (i.e., the attraction distance) per hook to obtain $a = [2.5, 5, 10, 50]ha$ for attraction distances of 500 m, 1 km, 2 km, and 10 km, respectively.

Only observer data from the last three years available for this study (i.e., 2014, 2015 and 2016) were used to fit the N-mixture model, as the model assumes a stationary population. We also restrict ourselves to longline catches since the estimates for purse-seine catches in 2016 are negligible compared to longline catches. Calculations of fishing mortality were derived for the most recent year only (2016), with the same three discard mortality scenarios as for the surplus production model. For internal consistency (and to avoid using the data twice), we predicted catch for the entire fleet from the N-mixture model instead of using the catch-reconstruction model employed by Tremblay-Boyer et al. (2019). The estimated catches were compared to ensure that fishing moralities derived from the N-mixture model were not overly influenced by differences in assumed catches. Additional technical detail on the implementation of the N-mixture model for oceanic whitetip shark is given in Appendix A.2.

3. RESULTS

3.1 Depletion-based catch-only simulations

Catch-only simulations showed that based on estimated removals over time, only a subset of initial values of population size (N_{1995}) and population growth R_{max} were plausible under the assumption that the current stock status is at least 20% lower than it was in 1995 (Figure 6). Especially high values of R_{max} were filtered out as these often lead to scenarios of increasing abundance under known removals. In the absence of any *a priori* constraints on current depletion level, a vague log-normal prior on initial

stock size led to a marginal prior distribution for stock status that implied a nearly unaffected stock and no risk of overfishing (see unfiltered panels of Figure 6). This high *a priori* stock status is an artefact of the scale of N_{1995} and K (i.e., via assumptions about initial depletion level N_{1995}/K), which are both necessarily positive, but without an *a priori*-informed upper bound on the absolute population size. The stock status assumption in conjunction with the catch time-series introduces an upper bound and thereby constrains the priors for initial population size and productivity. Alternative catch scenarios of current mortality (i.e., scenarios of discard mortality) led to similar implied priors for all parameters except the risk ratio (Figure 6, and Appendix A, Figures B-3, B-4). As expected, the value for the risk ratio is highest *a priori* with an assumed 100% discard mortality (the default in the previous assessment) and lowest with an assumed 75% survival (i.e., without post-release mortality).

3.2 Surplus production assessment of oceanic whitetip shark

The joint prior derived from depletion-based catch-only simulations (Figure 7) was applied to the surplus production model, with only small differences in the prior for the three discard mortality scenarios.

The surplus production model estimated all parameters well (Figure 8 and Appendix A, Figure B-5), constraining R_{max} to the lower end of the prior range. The fit to CPUE data was suboptimal for parts of the time series (Figure 9), but this outcome was largely due to high estimated observation error, which provided a relative weak constraint for fitting CPUE. When the observation error was reduced (by a factor of 4, for illustration), the fits were improved (Figure B-6), but did not markedly change any of the parameter estimates and especially derived quantities such as overfishing risk (Figure B-8).

The fits of the surplus production model indicated a steep decline in abundance since the mid-1990s to levels below 10% of unfished abundance in 2016 (0.07; 95% confidence interval [0.03; 0.16] across all discard mortality scenarios; Figures 8 and 9). The depletion estimate was sensitive to assumptions of initial depletion in 1995, with runs with higher initial depletion leading to estimates of lower relative abundance in 2016 (Figure 8).

While the assumption about initial depletion drives differences in other model parameters, it hardly affects the estimate of current fishing mortality relative to F_{crash}. Estimated fishing mortality increased to levels above F_{crash} in the early 2000s and remained critically high until 2012, irrespective of the discard mortality scenario (Figure 9) or initial population status (Figure B-7). Assuming different discard mortality scenarios resulted in different estimates of current fishing mortality and trends: although the estimated stock status in 2016 was near identical across the different scenarios (Figure 8), the highest mortality scenario, assuming 100% discard mortality, estimated a continued high fishing mortality (i.e., $P(F > F_{crash}) \approx 0.97$) and a decline over the most recent years. Both scenarios

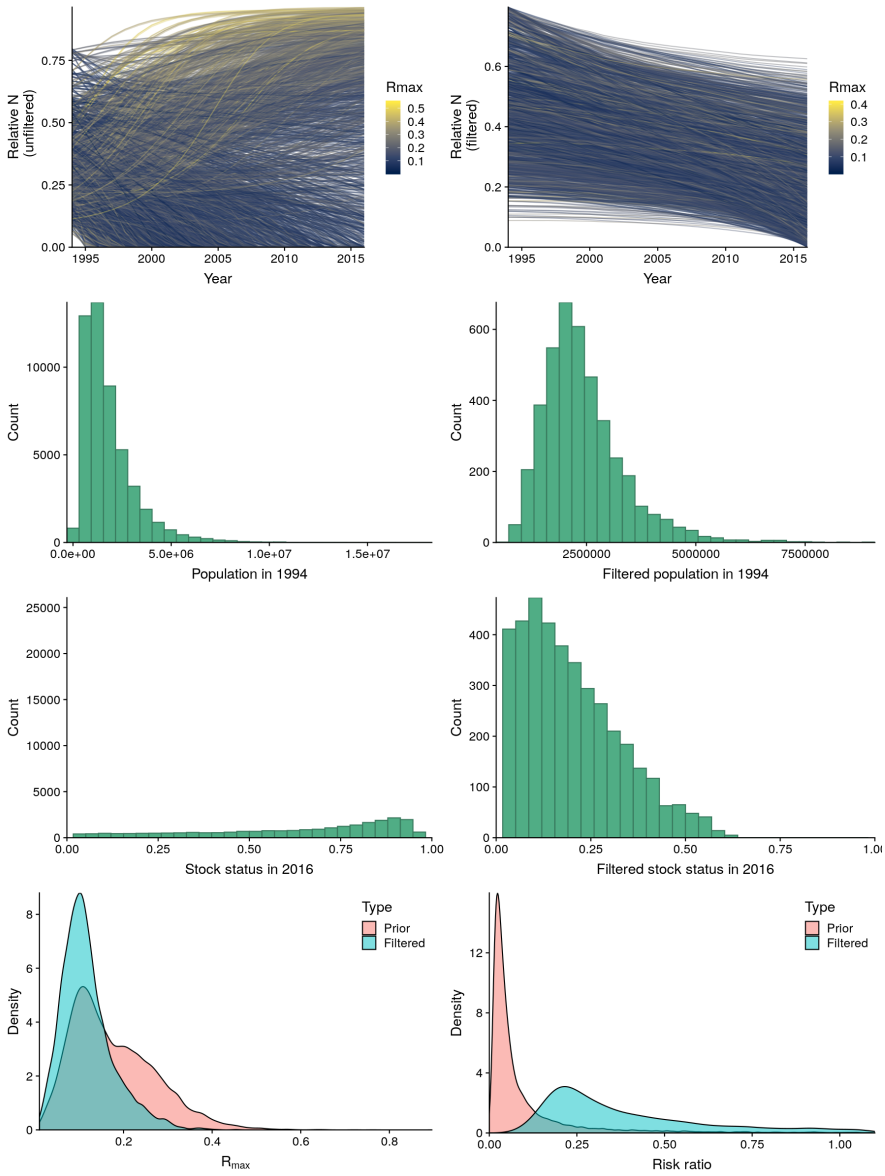

Figure 6: Summary of prior predictive simulations under assumed 44% discard mortality scenario for oceanic whitetip shark. Top row: Simulated population trajectories (in terms of relative abundance N_t/K) coloured by the value of the draw from the prior distribution of R_{max}. For each simulation trajectory, a set of values for carrying capacity, initial depletion, and R_{max} were drawn from their prior distribution, and the median estimated catch from the catch reconstruction was applied, with an assumed 44% total discard mortality. A subset of 1000 trajectories is shown on the left hand side, and a subset of 1000 trajectories from the filtered set (after applying constraints on current depletion (abundance relative to 1994). The corresponding draws from the prior distribution of stock size in 1994 are shown (2nd row) for the original prior and the constrained (filtered) prior. The prior distribution over stock status corresponding to the unconstrained prior (left) and the constrained prior (right) is shown in the third row. The constrained prior can be thought of as a joint Bayesian prior over parameters and current stock status in the simple surplus production model, and therefore implies a constrained prior for R_{max} and overfishing risk (last row; overfishing risk in terms of $F_{curr}/F_{crash} = F_{curr}/R_{max}$).

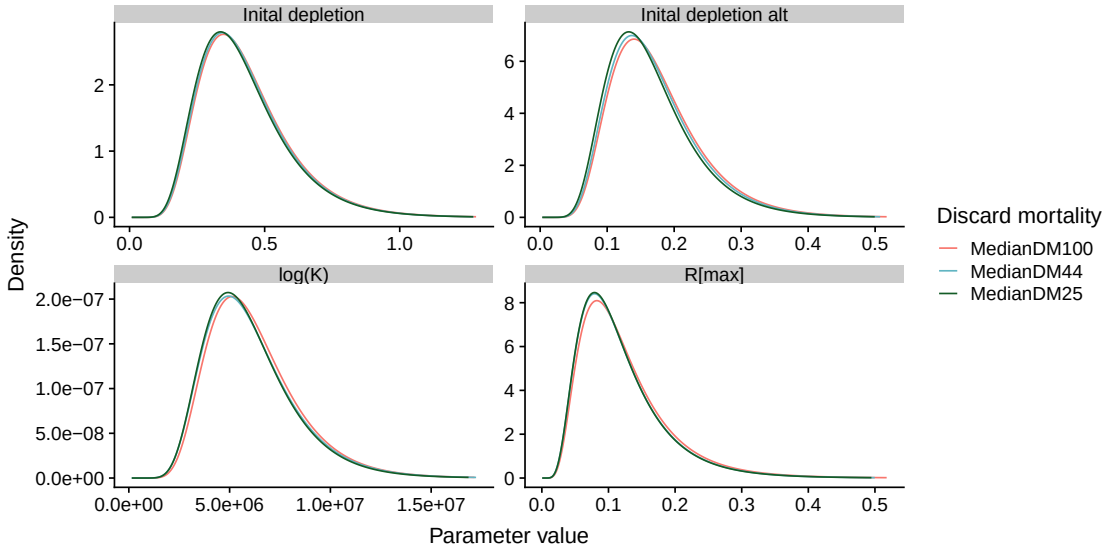

Figure 7: Priors for initial depletion (initial depletion retained from depletion-based catch-only simulations, and alternative (alt) scenario with higher initial depletion), the logarithm of the carrying capacity K, and intrinsic growth (R_{max}) for each discard mortality scenario of oceanic whitetip shark bycatch as applied for the surplus production model of oceanic whitetip shark. Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44).

of discards (25% and 44% total mortality, corresponding to zero and 25% post-release mortality) suggested a relatively flat abundance trend for the last two years, owing to a substantial reduction in fishing mortality to levels near F_{crash} (Figure 9).

Projections using estimated catch and discard mortality scenarios suggested limited to no rebuilding over five years from 2016 (the last assessment year) to 2021 even under reduced mortality (i.e., 0% or 25% post-release mortality), but a high likelihood of a population crash (i.e., the simulated population hit the lower bound of the assessment model set at a relative depletion of 1e-5) under a scenario of 100% discard mortality (Figure 10).

3.3 Spatial risk assessment for oceanic whitetip shark in the Western and Central Pacific Ocean

The N-mixture model converged (Figure B-9) towards an estimate of relatively low gear efficiency over the gear-affected area ($Q \approx$12%; Table B-1). The model fit was considered adequate (Figure B-10); there were no residual patterns with respect to flag or longitude and latitude (Figure B-11). The estimated relative abundance surface indicated a clear preference for equatorial warm-water areas (Figure 11). Grid cells beyond 30°North and 30°South (the limit of data used in the integrated stock assessment ad surplus production model) had a predicted abundance near zero, suggesting that estimates of abundance and risk are comparable.

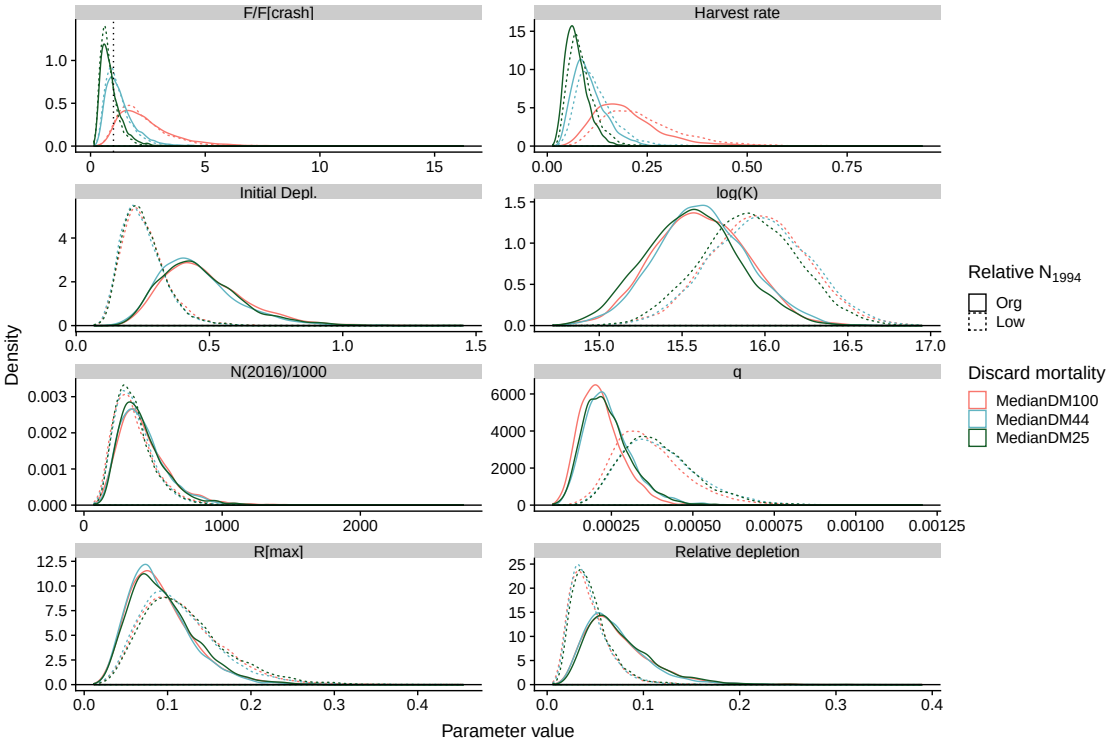

Figure 8: Marginal posterior densities for derived parameters (harvest rate, risk ratio of overfishing F/F_{crash}) and selected estimated parameters (carrying capacity K, catchability q, intrinsic population growth R_{max} and relative depletion) for different discard mortality scenarios of oceanic whitetip shark bycatch, and different assumptions of initial depletion level (low vs high initial depletion; see Figure 7). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). Vertical line for F/F_{crash} corresponds with fishing mortality of oceanic whitetip shark that leads to extinction (i.e., $F/F_{crash} = 1$).

Estimated catch from the N-mixture model corresponded closely with catch from the catch-reconstruction model for 016 (Figure 12). The estimated total abundance was markedly affected by the assumption about the size of the gear-affected area, which indicated a strong dependency of the risk of overfishing on the assumed spatial extent of the gear-affected area (Figure 12). For a large (0.5 km^2 or 50 ha) gear-affected area, all mortality scenarios on discards led to substantial risk of $F > F_{crash}$. For an assumed area of 0.1 km^2 (or 10 ha), any post-release mortality led to a substantial risk that $F > F_{crash}$. Even for a smaller assumed spatial extent of the gear-affected area (0.05 km^2 or 5 ha), there was some risk of $F > F_{crash}$.

3.4 Comparing methods

Only fishing mortality rates can be meaningfully compared among SRA, DSPM, and SS3 models: the SRA does not give information about the current depletion status. We note nevertheless, that inferred stock trajectories and depletion levels were very similar between SS3 and surplus

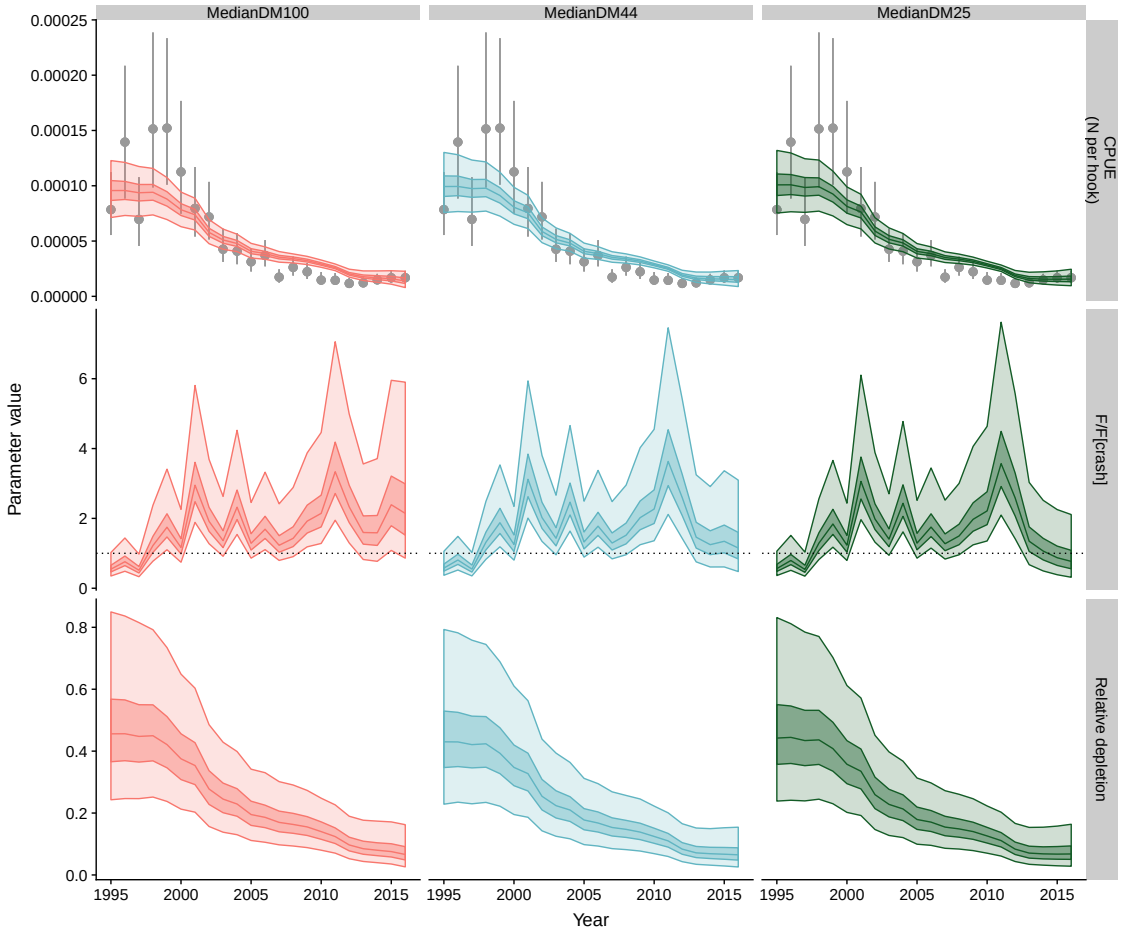

Figure 9: Fitting of catch-per-unit-effort (CPUE) data by the surplus production model of oceanic whitetip shark over time (dark shading, inter-quartile; light shading, 95% confidence interval). Top row: Predicted CPUE with input CPUE (points) and observation error (inter-quartile range) from the surplus production model. Middle row: Time series of risk ratio of overfishing F/F_{crash} estimated in the surplus production model. Bottom row: Estimated relative depletion (relative to unfished abundance K). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). (Note that the stock was not unfished in the first year of the time-series.)

production assessments (see Tremblay-Boyer et al. 2019, for estimated biomass trends and depletion levels). In addition, we can compare estimated abundance in 2016 from the DSPM and the SRA. For the SRA, estimated posterior median population size for intermediate (and potentially most plausible) gear-affected areas of 5 and 10 ha were 483 000 (CI[294 000,1 040 000]) and 241 000 (CI[147 000,520 000]), respectively. The DSPM output falls within the range of median SRA outcomes, with an estimate of around 363 000 remaining OCS in 2016 (CI[164 000,816 000]; Figure 8).

All three assessment types suggested that the most likely scenario of current fishing mortality (approximately 45% total discard mortality) still carries a high risk of inducing fishing mortalities at or above F_{crash} (and by extension,

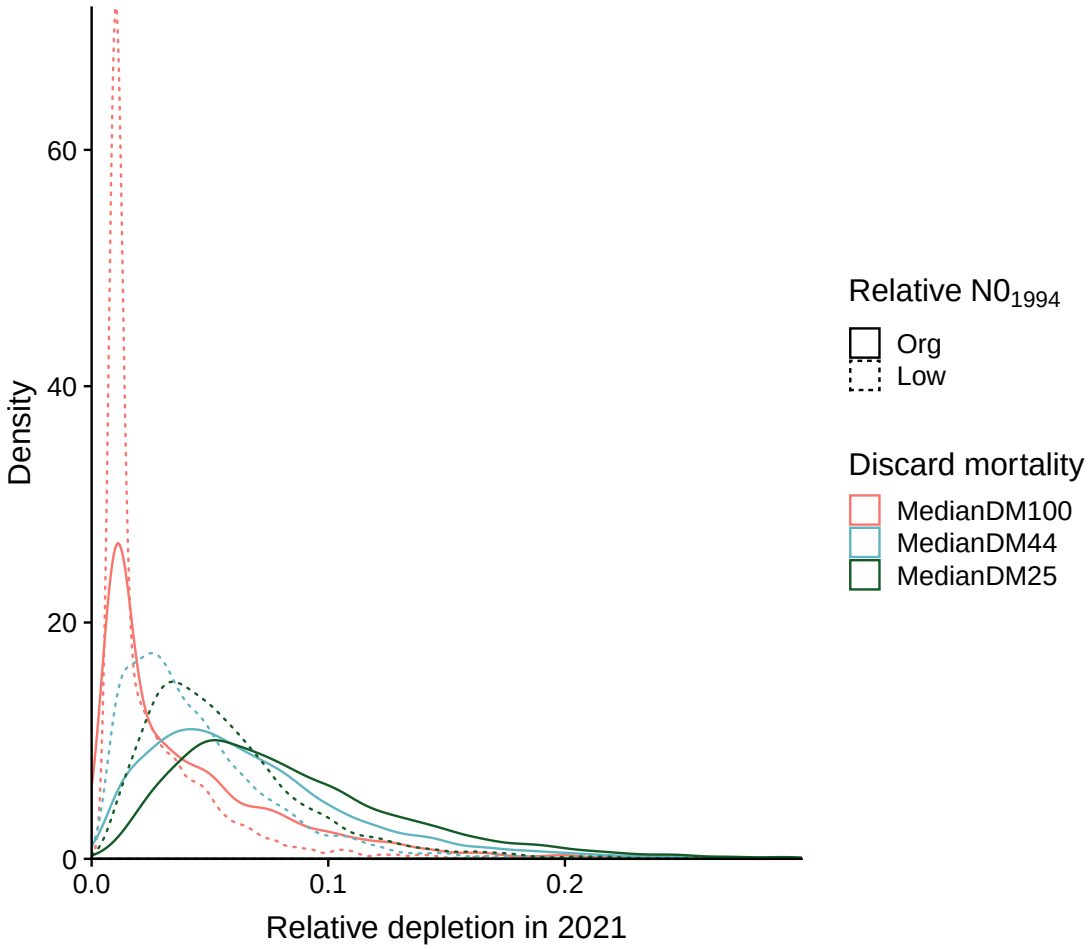

Figure 10: Marginal posterior densities of projected oceanic whitetip shark abundance in 2021 from the Bayesian surplus production model under different bycatch discard mortality scenarios. Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44).

well above F_{MSM} and F_{lim}; Table 2, Figure 13). The integrated stock assessment suggested consistently higher fishing mortality than both the DSPM and the SRA. The latter two approaches estimated that the scenario that did not consider post-release mortality had some chance that $F < F_{crash}$, whereas the SS3 assessment suggested that all three fishing mortality scenarios lead to $F_{2016} \geq R_{max}$ when one considers the estimated R_{max} from the DSPM as a suitable value for F_{crash}.

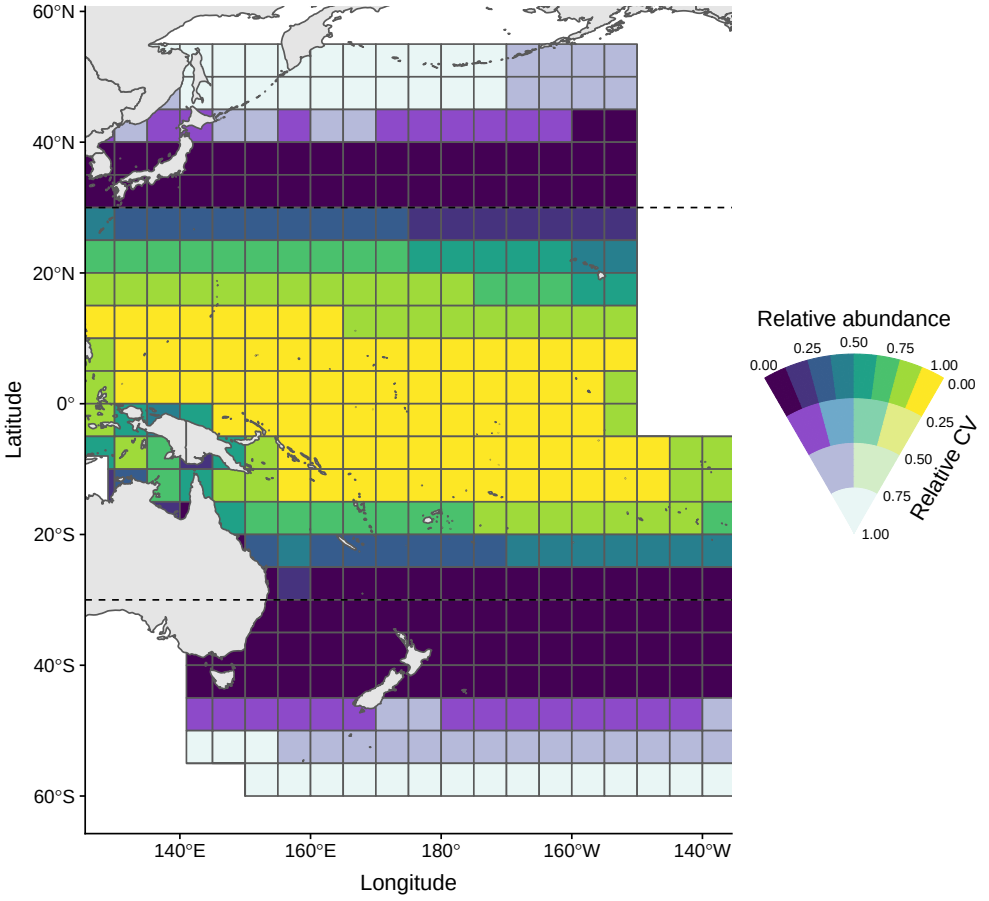

Figure 11: Estimated relative abundance surface determined by sea-surface temperature by the Bayesian N-mixture model for oceanic whitetip shark in the Western and Central Pacific Ocean (CV, coefficient of variation). The limit of the data used for the integrated stock assessment (30°North to 30°South) is shown for reference.

4. DISCUSSION

We applied a surplus production and spatial risk assessment framework to provide alternative perspectives on the status of oceanic whitetip shark in the Western and Central Pacific Ocean. Although age-structured stock assessment models are often regarded as the "gold standard" for stock assessments, it can be argued that for information-poor fisheries "[t]here is no single "best" model that should be used for fisheries stock assessments" (Bonfil 2005). Applying a range of models fit to available data provides an opportunity to detect inconsistencies, coincidences and patterns. Such a multi-model assessment approach also includes a comparison of outcomes across different analyses to critically assess the conclusions, allowing for the improvement of data and future assessments.

We showed that estimated catches alone, in conjunction with a prior belief that observed catches have affected the stock status of oceanic whitetip shark, can constrain the initial population size and R_{max} to some extent, but these simulations cannot provide useful management information as

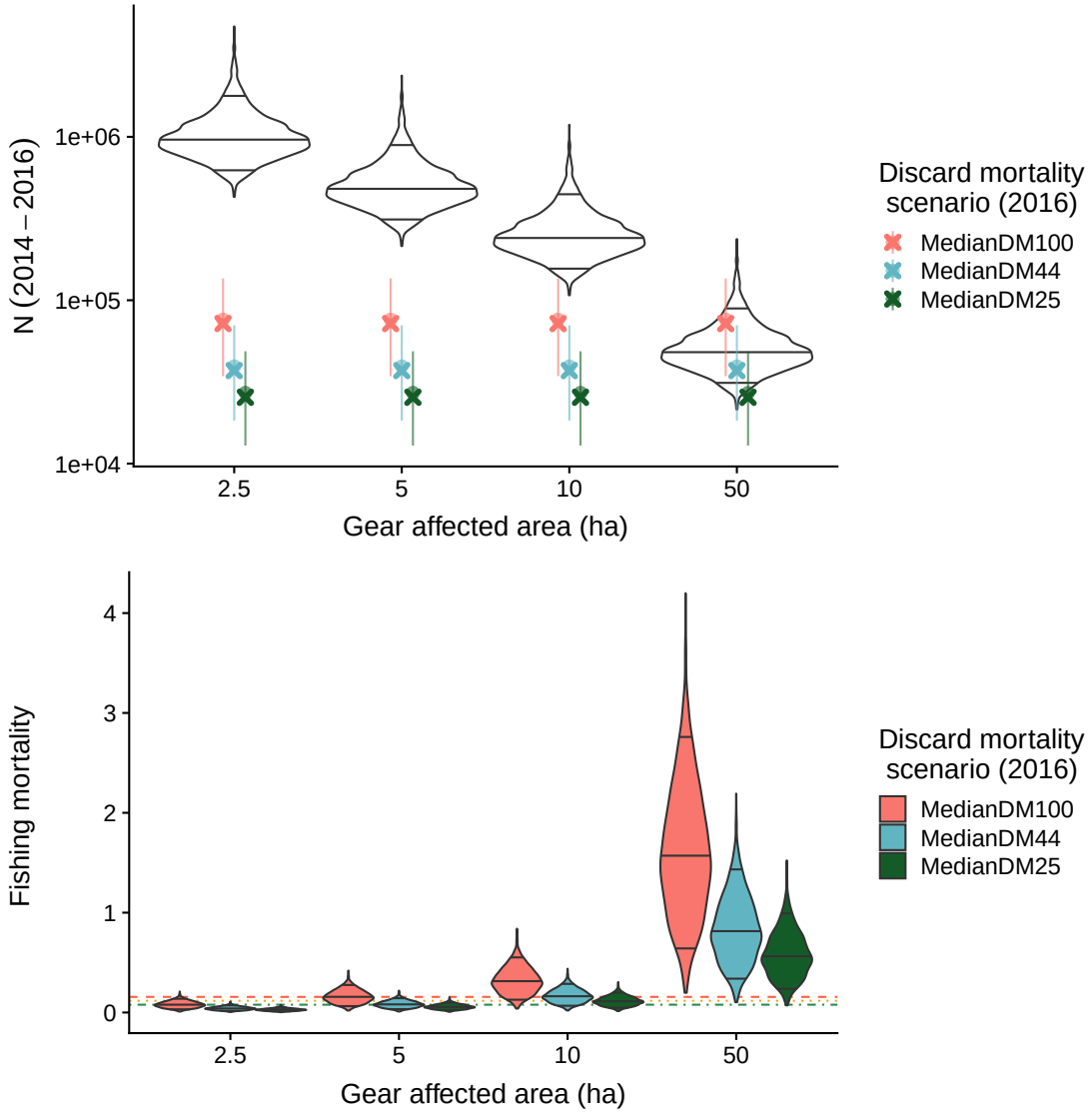

Figure 12: Estimated oceanic whitetip shark abundance and estimated fishing mortality in the Western and Central Pacific Ocean, under three different bycatch discard mortality scenarios. Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). Top panel: Marginal posterior distribution of estimated shark abundance for four different assumed sizes of gear-affected area (violin plots with 2.5%, 50%, 97.5% posterior quantiles) and estimated discard mortalities (median and 95% confidence interval). Bottom panel: Fishing mortality (violin plots with 2.5%, 50%, 97.5% posterior quantiles) for four different assumed sizes of gear-affected area and the three discard mortality scenarios. Also shown is the median value (red dashed horizontal line) for the *a priori* derived $R_{max} = F_{crash}$, as well as $0.75R_{max} = F_{lim}$ (orange dotted horizontal line) and $0.5R_{max} = F_{MSM}$ (dark green dashed-dotted horizontal line).

the overfishing risk is indirectly defined *a priori*. Contrast in CPUE or spatial catch and effort are required to constrain current population size, stock status and/or the risk of overfishing *a posteriori*. We did not apply catch-only frameworks that infer stock status and/or the risk of overfishing

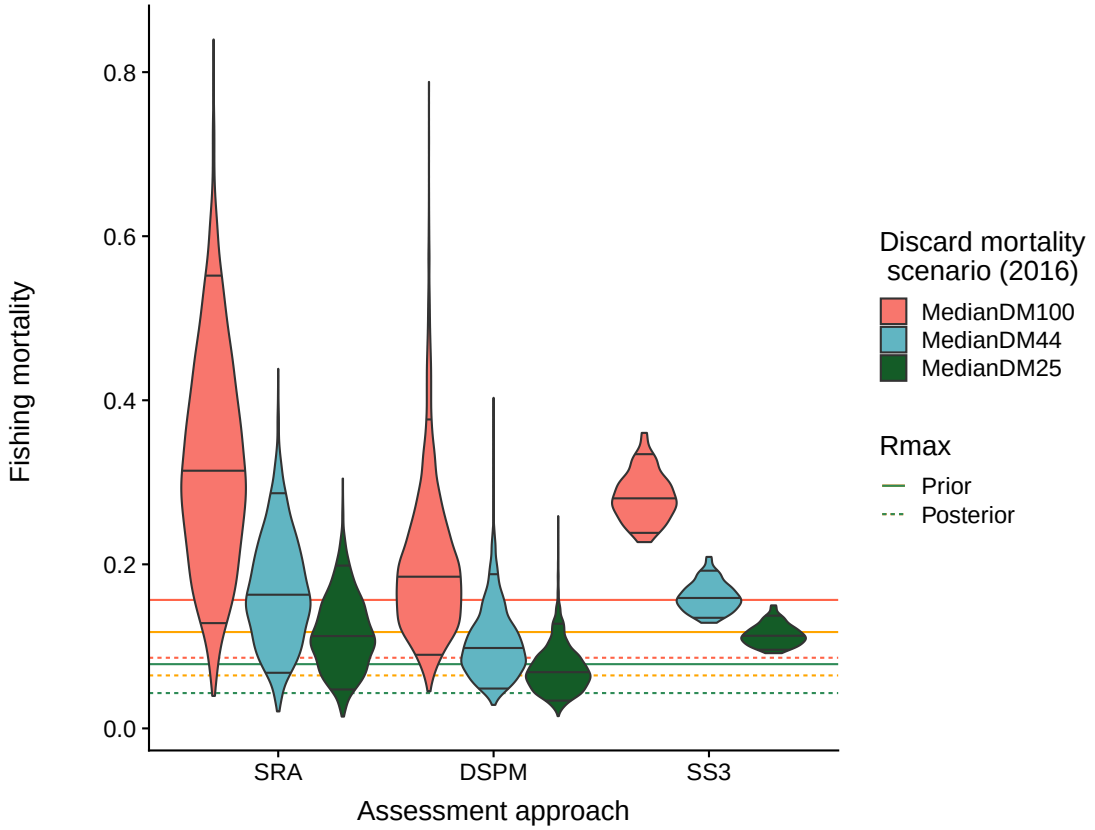

Figure 13: Fishing mortality estimated for oceanic whitetip shark (violin plots; plotted with identical widths for visibility (i.e., the area of the surface does not integrate to 1), with 2.5%, 50%, 97.5% posterior quantiles) for three different assessment methods (spatial risk assessment (SRA) using a 10 ha gear affected area corresponding to a 2km attraction range, or 1 km either side of baited hooks), dynamic surplus production model (DSPM), and stock synthesis (SS3)). Solid horizontal lines show the simulated prior median value for $F_{crash} = R_{max}$ [red] ($F_{lim} = 0.75R_{max}$ [orange], $F_{MSM} = 0.5R_{max}$ [green]) and posterior median estimates of the same quantities from the DSPM (dotted lines). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). The distribution for fishing mortality values for the spatial risk assessment (SRA) and the dynamic surplus production model (DSPM) reflects parameter uncertainty, whereas the distribution for the SS3 fishing mortality reflects model grid-runs, which use maximum likelihood inference and do not have measures of estimation (parameter) uncertainty attached. The size of the inter-quantile ranges is therefore not comparable between SS3 and the other model types.

risk from trends in catch alone as these inferred trends tend to be strongly confounded with the effects of (or assumptions about) changing effort, targeting or avoidance measures (Anderson et al. 2017). Although multi-model inference can integrate over these assumptions to derive a consensus estimate of stock status from catch alone (Anderson et al. 2017), this integration entails fitting a large number of different models, and will still depend on the specific models and their weight in such an ensemble approach.

Table 2: Probability (risk) that fishing mortality exceeds F_{crash}, F_{lim}, and F_{MSM} for the spatial risk assessment (SRA; using a 10 ha gear affected area corresponding to a 2km attraction range, or 1 km either side of baited hooks), the dynamic surplus production model (DSPM), and Stock Synthesis 3. Risk is shown in terms of *a priori* derived R_{max} values (Prior) and in terms of estimated R_{max} from the surplus production model (Posterior). The risk for fishing mortality values for the SRA and the DSPM reflects parameter uncertainty, whereas the distribution for the SS3 fishing mortality reflects model grid-runs, which use maximum likelihood inference and do not have measures of estimation (parameter) uncertainty attached. Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44).

Model	Mortality	P($F > F_{crash}$)		P($F > F_{lim}$)		P($F > F_{MSM}$)	
		Prior	Posterior	Prior	Posterior	Prior	Posterior
SRA	MedianDM100	0.96	0.98	0.98	1.00	1.00	1.00
	MedianDM44	0.74	0.90	0.88	0.96	0.97	0.99
	MedianDM25	0.42	0.73	0.69	0.87	0.89	0.97
DSPM	MedianDM100	0.84	0.97	0.95	0.99	0.99	1.00
	MedianDM44	0.29	0.62	0.57	0.84	0.88	0.98
	MedianDM25	0.07	0.27	0.21	0.56	0.63	0.87
SS3	MedianDM100	1.00	1.00	1.00	1.00	1.00	1.00
	MedianDM44	1.00	1.00	1.00	1.00	1.00	1.00
	MedianDM25	0.27	1.00	1.00	1.00	1.00	1.00

The surplus production model can be thought of as a simplified assessment (relative to age- or size-structured assessments such as SS3). Although the biological assumptions in surplus production models may be considered overly simple, they also serve to make fewer assumptions. For sharks, large uncertainties often remain about fundamental biological parameters such as age-at-first-reproduction, theoretical maximum age, fecundity or growth rates. The single productivity parameter R_{max} that is used in DSPMs can be reasonably constrained from theory and by using depletion-based catch-only simulations, and no additional assumptions about productivity are needed. Biological information and uncertainty can still considered when deriving a Bayesian prior R_{max}, and is therefore still integral to a DSPM assessment. The relative performance of either the surplus production or the age-structure (e.g., SS3) model will depend on trends in CPUE and the influence of transient age and size structures (e.g., Ludwig & Walters 1985, Punt & Szuwalski 2012). This choice is difficult to make *a priori*, and producing (formal or informal) multi-model inference for stocks with sparse compositional data could, therefore, be considered the most robust approach (Bonfil 2005). However, given cost-constraints, DSPMs may provide useful information of stock dynamics and over-fishing risk that may be sufficient in many cases.

The surplus production model inference showed a substantial decline in oceanic whitetip shark numbers from the mid 1990s until the most recent years in the assessment (the minimum year depends on the assumed mortality scenario, see Figure 9). The estimated stock status below 10% of carrying capacity, and the unsustainable estimated fishing mortality over most of the assessment period suggest that the extent of reductions in fishing

mortality through non-retention and minimisation of discard mortality are the key uncertainties about the persistence of the stock into the future. Under scenarios of high discard mortality, the stock is unlikely to be able to persist.

The stock status estimate from the surplus production assessment was slightly higher than that from the integrated stock assessment (7% in the DSPM vs. 3% in SS3), the latter estimating a relative abundance under 5% of un-fished abundance. There are a number of potential causes for this discrepancy: For example, assumptions about initial depletion/fishing mortality or selective removal of mature individuals in the integrated model can lead to differences in the estimate stock status. The surplus production model cannot account for selective removal of large, mature individuals, and therefore represents total abundance rather than spawning stock abundance. As such, the surplus production model may underestimate the degree of recruitment over-fishing, and over-estimate productivity at low population size.

Both the integrated SS3 assessment and the DSPM use standardized longline CPUE to drive abundance trends in the models. CPUE was highly uncertain and variable for early years, when observer coverage was sparse, and may have been additionally impacted by conservation measures such as CMM-2011-04, which potentially lead to a decline in reported numbers of hooked sharks if the latter are cut free before being landed on board. Additionally, the ban on shark lines and wire traces brought about by CMM-2014-05 may have led to reduced shark catch-rates, meaning that recent CPUE may be under-estimated.

The SRA does not rely on a complete time-series of CPUE, and could therefore be considered a more robust alternative to estimate recent fishing mortality. Despite key differences in data inputs and methods, results about the magnitude of 2016 fishing mortality from the SRA were similar to those from the DSPM and SS3. We aligned the spatial risk assessment modelling approach as closely as possible with the other assessment methodologies to highlight the key assumptions and impacts. Estimated catch from this method and the catch-reconstruction model used for the other assessment methods applied to oceanic whitetip shark correspond closely. Estimates of the risk of overfishing are, therefore, only linked to different assumptions. The spatial risk assessment method used here can be considered a model-based estimation of a fisheries survey, and scaling up to total abundance and risk of overfishing risk is dependent on assumptions about the spatial impact of the gear. This assumption is necessary for any SRA in which population size cannot be constrained *a priori*.

The SRA allowed us to derive a current population size, and estimates of F_{2016} on the basis of estimated catch relative to the estimated population. For assumptions of large gear-affected areas (i.e., $\geq$0.1 km^2 or 10 ha, our base assumption), we obtained estimates of current F that corresponded with conclusions of the surplus production assessment: with 100% assumed discard mortality, there was a high risk that $F_{2016} > F_{crash}$, whereas the estimate for F_{2016} was considerably lower for scenarios of higher discard

survival. Nevertheless, at the most likely mortality scenario, which includes direct mortality and post-release mortality, a non-negligible risk remained that $F_{2016} > F_{crash}$ (Table B-1). There is, therefore, a non-zero risk that the population will not persist under this mortality scenario.

In combination, the DSPM and SRA approaches highlight the need to further minimize fishing mortality to sustain the oceanic whitetip shark population in the Western and Central Pacific Ocean. Current fishing mortality estimates for this species are highly uncertain owing to high uncertainty about discard mortality (Common Oceans (ABNJ) Tuna Project 2019), nevertheless, even the most optimistic scenario (no post-release mortality) carries a high risk for OCS for all models but the DSPM, and the most realistic scenario (44% total mortality) carries substantial risk across all models, especially when one considers the posterior distribution of R_{max} from the DSPM as a proxy for F_{crash} (in that case $\mathrm{P}(F > F_{crash}) \geq 0.62$ across all models; Table 2). Understanding discard mortality (i.e., the sum of haul-back, handling and post-release mortalities) for oceanic whitetip shark therefore needs to be a priority to ensure that measures aimed at minimising fishing mortality of oceanic whitetip shark are effective and contribute to rebuilding the population from current low levels.

Based on the method comparison and results of this study, a possible framework for assessing shark and other bycatch species is illustrated in Figure 14. The comparison of assessment approaches showed that when data are sufficient (i.e., a time series of removals and CPUE exist and judged a sufficiently reliable indicator of abundance), a surplus production model could be considered a cost-effective method to rapidly and frequently assess bycatch species such as pelagic sharks. By carefully constructing priors for the assessment, one can integrate biological and domain knowledge, yet avoid having to run a large amount of model-uncertainty scenarios that are often necessary in integrated assessments such as SS3 (i.e., grid runs in the integrated assessments). As a result, one can focus on key uncertainties such as mortality scenarios and initial depletion level. More time-intensive and costly integrated assessments such as the SS3 assessment could then be used for priority species, where prioritization could include cost, stock status, and overfishing risk. SRA approaches can help with prioritisation as the relatively simple SRA framework can be applied across a range of species, even those with poor historical time-series data. If the gear-affected area can be quantified, the SRA can be used to get an estimate of absolute F, however, SRA methods can still be used in a semi-quantitative (i.e.,relative) way to prioritise assessment and conservation efforts.

ACKNOWLEDGEMENTS

Many thanks to Dr. Shelley Clarke for valuable pointers over the course of this project. The authors also thank Sebastian Pardo for kindly sharing his R-code for estimating uncertainty in r_{max}. Funding for this work was provided through the Global Environmental Facility (GEF) funded United Nations (UN) Food and Agriculture Organization (FAO) Common Oceans (ABNJ)

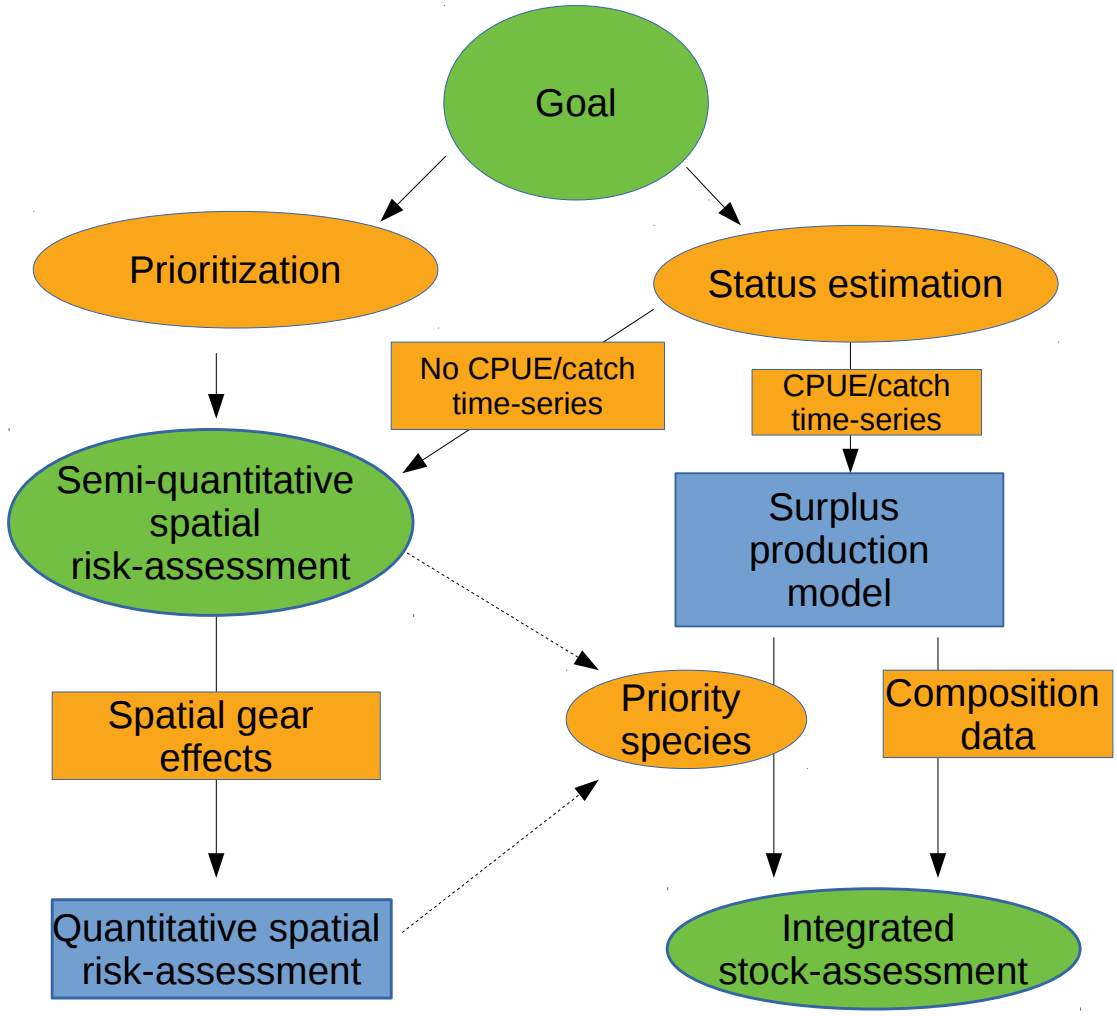

Figure 14: Possible framework for assessing shark and other bycatch species, with management considerations in orange ovals and data constraints in orange squares: a semi-quantitative spatial risk assessments could be used to prioritise species in terms of over-fishing risk, and help designate species for which additional research and/or full quantitative stock assessments are needed. If reasonable assumptions can be made about the spatial interaction of fishing gear and bycatch species, then the relative risk framework can be extended to estimate absolute fishing mortality and over-fishing risk. Surplus production models can be used to cost-effectively assess shark species on a more regular basis, provided a realistic time-series for catch and CPUE can be constructed. For species of high conservation concern and sufficient data, full integrated stock assessments can be attempted.

Tuna Project executed through the Western and Central Pacific Fisheries Commission (WCPFC).

5. REFERENCES

Anderson, S. C.; Cooper, A. B.; Jensen, O. P.; Minto, C.; Thorson, J. T.; Walsh, J. C.; Afflerbach, J.; Dickey-Collas, M.; Kleisner, K. M.; Longo, C., et al. (2017). Improving estimates of population status and trend with superensemble models. *Fish and Fisheries*, *18*(4), 732–741.

Bigelow, K.; Musyl, M. K.; Poisson, F., & Kleiber, P. (2006). Pelagic longline gear depth and shoaling. *Fisheries research*, *77*(2), 173–183.

Bonfil, R. (2005). Fishery stock assessment models and their application to sharks. In R. Bonfil & J. A. Musick (Eds.), (p. 154). Rome: FAO.

Carruthers, T. R.; Punt, A. E.; Walters, C. J.; MacCall, A.; McAllister, M. K.; Dick, E. J., & Cope, J. (2014). Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research*, *153*, 48–68.

Clarke, S.; Coelho, R.; Francis, M.; Kai, M.; Kohin, S.; Liu, K.; Simpendorfer, C.; Tovar-Availa, J.; Rigby, C., & Smart, J. (2015). *Report of the Pacific Shark Life History Expert Panel Workshop*. WCPFC-SC11-2015/EB-IP-13. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Eleventh Regular Session, 5–13 August 2015, Pohnpei, Federated States of Micronesia.

Clarke, S. & Hoyle, S. (2014). *Development of limit reference points for elasmobranchs*.

Clarke, S.; Langley, A.; Lennert-Cody, C.; Aires-da-Silva, A., & Maunder, M. (2018). *Stock assessment of silky sharks in the Western and Central Pacific Ocean*. WCPFC-SC14-2018/SA-WP-08. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Fourteenth Regular Session, August 8–16, Busan, Korea.

Common Oceans (ABNJ) Tuna Project (2019). Joint analysis of shark post-release mortality tag results. *Western and Central Pacific Fisheries Commission*.

D'Alberto, B. M.; Chin, A.; Smart, J. J.; Baje, L.; White, W. T., & Simpfendorfer, C. A. (2017). Age, growth and maturity of oceanic whitetip shark (*Carcharhinus longimanus*) from Papua New Guinea. *Marine and Freshwater Research*, *68*(6), 1118–1129.

Dick, E. & MacCall, A. D. (2011). Depletion-based stock reduction analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. *Fisheries Research*, *110*(2), 331–341.

Dudgeon, C. L.; Blower, D. C.; Broderick, D.; Giles, J. L.; Holmes, B. J.; Kashiwagi, T.; Krück, N. C.; Morgan, J. A. T.; Tillett, B. J., & Ovenden, J. R. (2012, April). A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology*, *80*(5), 1789–1843. doi:10.1111/j.1095-8649.2012.03265.x

Edwards, C. T. T. (2017). *Bdm: Bayesian biomass dynamics model*. R package version 0.0.0.9022.

Francis, R. I. C. C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(6), 1124–1138. doi:10.1139/f2011-025

Froese, R.; Demirel, N.; Coro, G.; Kleisner, K. M., & Winker, H. (2017). Estimating fisheries reference points from catch and resilience. *Fish and Fisheries*, *18*(3), 506–526.

Fu, D.; Roux, M.-J.; Clarke, S.; Francis, M.; Dunn, A., & Hoyle, S. (2017). Pacific-wide sustainability risk assessment of bigeye thresher shark (alopias superciliosus). *WCPFC-SC13-2017/SA-WP-11 (rev 2)*.

Hewitt, D. A. & Hoenig, J. M. (2005). Comparison of two approaches for estimating natural mortality based on longevity. *Fishery Bulletin*, *103*(2), 433.

Hobday, A.; Smith, A.; Stobutzki, I.; Bulman, C.; Daley, R.; Dambacher, J.; Deng, R.; Dowdney, J.; Fuller, M.; Furlani, D.; Griffiths, S.; Johnson, D.; Kenyon, R.; Knuckey, I.; Ling, S.; Pitcher, R.; Sainsbury, K.; Sporcic, M.; Smith, T.; Turnbull, C.; Walker, T.; Wayte, S.; Webb, H.; Williams, A.; Wise, B., & Zhou, S. (2011). Ecological risk assessment for the effects of fishing. *Fisheries Research*, *108*, 372–384.

Hoyle, S. D.; Edwards, C. T. T.; Roux, M.-J.; Clarke, S., & Francis, M. (2017). Southern hemisphere porbeagle shark stock status assessment. *WCPFC-SC13-2017/SA-WP-12 (rev. 1)*.

Jensen, A. (1996). Beverton and holt life history invariants result from optimal trade-off of reproduction and survival. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*(4), 820–822.

Jordan, L. K.; Mandelman, J. W.; McComb, D. M.; Fordham, S. V.; Carlson, J. K., & Werner, T. B. (2013). Linking sensory biology and fisheries bycatch reduction in elasmobranch fishes: A review with new directions for research. *Conservation Physiology*, *1*(1), cot002.

Joung, S.-J.; Chen, N.-F.; Hsu, H.-H., & Liu, K.-M. (2016). Estimates of life history parameters of the oceanic whitetip shark, *Carcharhinus longimanus*, in the western North Pacific Ocean. *Marine Biology Research*, *12*(7), 758–768.

Ludwig, D. & Walters, C. J. (1985). Are age-structured models appropriate for catch-effort data? *Canadian Journal of Fisheries and Aquatic Sciences*, *42*(6), 1066–1072.

Martell, S. & Froese, R. (2013). A simple method for estimating msy from catch and resilience. *Fish and Fisheries*, *14*(4), 504–514.

Maunder, M. N. & Piner, K. R. (2017). Dealing with data conflicts in statistical inference of population assessment models that integrate information from multiple diverse data sets. *Fisheries Research*. Data conflict and weighting, likelihood functions, and process error, *192*, 16–27. doi:10.1016/j.fishres.2016.04.022

McAllister, M. & Edwards, C. (2016). Applications of a bayesian surplus production model to new zealand fish stocks. *New Zealand Fisheries Assessment Report*, *52*, 79.

Methot Jr, R. & Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, *142*, 86–99.

Minte-Vera, C. V.; Maunder, M. N.; Aires-da-Silva, A. M.; Satoh, K., & Uosaki, K. (2017). Get the biology right, or use size-composition data at your own risk. *Fisheries Research*. Data conflict and weighting, likelihood functions, and process error, *192*, 114–125. doi:10.1016/j.fishres.2017.01.014

Neubauer, P.; Richard, Y., & Clarke, S. (2018). *Risk to the Indo-Pacific Ocean whale shark population from interactions with Pacific Ocean purse-seine*

fisheries. WCPFC-SC14-2018/SA-WP-12 (rev.1), Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Fourteenth Regular Session 8–16 August 2018, Busan, Republic of Korea.

Pardo, S. A.; Kindsvater, H. K.; Reynolds, J. D., & Dulvy, N. K. (2016, August). Maximum intrinsic rate of population increase in sharks, rays, and chimaeras: The importance of survival to maturity. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(8), 1159–1163. doi:10.1139/cjfas-2016-0069

Pardo, S. A.; Cooper, A. B.; Reynolds, J. D., & Dulvy, N. K. (2018, January 5). Quantifying the known unknowns: Estimating maximum intrinsic rate of population increase in the face of uncertainty. *ICES Journal of Marine Science*. doi:10.1093/icesjms/fsx220

Punt, A. E. & Szuwalski, C. (2012). How well can fmsy and bmsy be estimated using empirical measures of surplus production? *Fisheries Research*, *134*, 113–124.

Raftery, A. E. (1988). Inference for the binomial n parameter: A hierarchical bayes approach. *Biometrika*, *75*(2), 223–228.

Rice, J. (2012). *Alternate catch estimates for silky and oceanic whitetip sharks in Western and Central Pacific Ocean*, WCPFC-SC8/SA-IP-12. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Eighth Regular Session, 7–15 August 2012, Busan, Korea.

Rice, J. (2018). *Report for Project 78: Analysis of observer and logbook data pertaining to key shark species in the Western and Central Pacific Ocean*. WCPFC-SC14/EB-WP-02. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Fourteenth Regular Session, 8–16 August 2018, Busan, Korea.

Rice, J. & Harley, S. (2012). *Stock assessment of oceanic whitetip sharks in the western and central Pacific Ocean*. WCPFC-SC8-2012/SA-WP-06 Rev 1. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Eighth Regular Session, 7–15 August 2012, Busan, Korea.

Rice, J.; Tremblay-Boyer, L.; Scott, R.; Hare, S., & Tidd, A. (2015). *Analysis of stock status and related indicators for key shark species of the Western Central Pacific Fisheries Commission*. WCPFC-SC11/EB-WP-04. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Eleventh Regular Session, 5–13 August 2015, Pohnpei, Federated States of Micronesia.

Royle, J. A. & Dorazio, R. M. (2006). Hierarchical models of animal abundance and occurrence. *Journal of Agricultural, Biological, and Environmental Statistics*, *11*(3), 249–263.

Ruck, C. L. (2016). *Global genetic connectivity and diversity in a shark of high conservation concern, the oceanic whitetip,* Carcharhinus longimanus. Unpublished M.Sc. thesis, Nova Southeastern University, Fort Lauderdale, United States.

Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.17.3. Retrieved from http://mc-stan.org/

Tremblay-Boyer, L.; Carvalho, F.; Neubauer, P., & Pilling, G. (2019). *Stock assessment for oceanic whitetip shark in the western and central pacific ocean*. WCPFC-SC15/SA-WP-06.

Tremblay-Boyer, L. & Neubauer, P. (2019). *Data inputs to the stock assessment for oceanic whitetip shark in the Western and Central Pacific Ocean*. WCPFC-SC15/SA-IP-XX. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Fifteenth Regular Session, 12–20 August 2018, Pohnpei, Federated States of Micronesia.

Walters, C. J.; Martell, S. J., & Korman, J. (2006). A stochastic approach to stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, *63*(1), 212–223.

WCPFC (2016). *WCPFC13 Summary Report Attachment G: Scientific data to be provided to the commission*.

Winker, H.; Carvalho, F., & Kapur, M. (2018). Jabba: Just another bayesian biomass assessment. *Fisheries Research*, *204*, 275–288.

Zhou, S.; Daley, R.; Fuller, M.; Bulman, C.; Hobday, A.; Courtney, T.; Ryan, P., & Ferrel, D. (2013). Era extension to assess cumulative effects of fishing on species. *Final Report on FRDC Project No. 2011/029*.

Zhou, S.; Deng, R.; Hoyle, S., & Dunn, M. (2018). *Identifying appropriate reference points for elasmobranchs within the WCPFC*. WCPFC-SC14-2018/MI-WP-07. Report to the Western and Central Pacific Fisheries Commission Scientific Committee. Fourteenth Regular Session, 8–16 August 2018, Busan, Korea.

Zhou, S.; Klaer, N. L.; Daley, R. M.; Zhu, Z.; Fuller, M., & Smith, A. D. (2014). Modelling multiple fishing gear efficiencies and abundance for aggregated populations using fishery or survey data. *ICES Journal of Marine Science*, *71*(9), 2436–2447.

Zhou, S.; Smith, A. D., & Fuller, M. (2011). Quantitative ecological risk assessment for fishing effects on diverse data-poor non-target species in a multi-sector and multi-gear fishery. *Fisheries Research*, *112*(3), 168–178.

APPENDIX A: Supplementary technical detail

A.1 Relationship between surplus production and risk assessments.

One can show that the relationship between catchability q as used in a simple stock assessment and gear efficiency Q as used in a spatial risk-assessment or survey is simply:

$$q = \frac{aQ}{A} \tag{A-1}$$

with a the area affected by the gear (e.g., area swept by trawl gear, or attraction area for baited longline or hook-an-line), and A the over-all area of the assessment. To derive this, note that in a typical assessment, catch (C)-per-unit effort (E) is assumed to be proportional to abundance, such that:

$$\frac{C}{E} = qN \tag{A-2}$$

In the survey method or spatial risk assessment, the catch in the gear-affected area a is simply the size of the area, multiplied by the gear efficiency Q (i.e., the proportion of the local density that is removed per unit effort) and the density of the focal species (N/A)

$$\frac{C}{E} = \frac{aQN}{A} \tag{A-3}$$

Equating equations A-2 and A-3 gives the above stated relationship between q and Q.

A.2 Statistical detail for the N-mixture model used in the spatial risk assessment for OCS

The N-mixture model can be written as:

$$n_{i,k} \sim P(h_{i,k} a D_i), \tag{A-4}$$
$$C_{i,k} \sim B(n_{i,k}, Q_k), \tag{A-5}$$

where $n_{i,k}$ is the abundance of oceanic whitetip shark in the gear-affected area, $h_{i,k}$ is the number of hooks in area i for fleet k, and a, is the affected area per hook. The parameter $D_i = N_i/A_i$ is the density (numbers per grid area) of oceanic whitetip shark in grid i. The model can be considered a generative model: the number of sharks available per spatial unit of effort (i.e., the area affected by a unit of effort, e.g., a single hook) for each fleet in

each grid cell i is random and follows a Poisson (P) distribution with mean $D_i^a = aD_i = aN_i/A_i$. The catch for each fleet k in grid cell i is a binomial (B) draw given a specific fleet's gear efficiency Q_k.

For consistency, we defined fleets in terms of the same variables that were included in the catch-reconstruction model that was used for both the current project and the integrated assessment ((Tremblay-Boyer et al. 2019)). Specifically, fleets were characterised by vessel flag, set type (deep versus shallow sets), and target cluster (i.e., k is defined by the unique combinations of these variables). Although different set types may affect a, we subsumed this effect into the gear-efficiency parameter Q. We modelled Q as:

$$\mathsf{log}\left(\frac{Q}{1-Q}\right) = \beta_Q X + \omega_{\text{flag}} + \zeta_{\text{flag-year}} \tag{A-6}$$

The parameter X is a design matrix with contrasts for (predicted) deep and shallow sets and for targeting cluster (as per catch reconstruction in Tremblay-Boyer et al. 2019), with coefficients β_Q estimated for both effects. Flag (ω_{flag}) and flag-year ($\zeta_{\text{flag-year}}$) were treated as random effects, which capture differences in fishing operations (flag) and potentially changing practices due to implementation of conservation measures by fleets (flag-year effect). All random effects had informed half-normal priors with a standard deviation of 2 for the random effect standard deviation (on the logit scale).

The spatial density of oceanic whitetip sharks was modelled as $\mathsf{log}\, D_i^a = \mu_D + f(\text{SST}_i)$, where μ_D is the mean density per unit of effort and $f(\text{SST}_i)$ is a smooth (spline) function of sea-surface temperature in grid i. The overall model differs from the catch-reconstruction model in that the abundance and capture process are explicitly separated. This separation allowed extrapolation of shark density beyond fished grid cells.

For the estimation, it is possible to explicitly integrate over the latent abundance process (the Poisson component of the model above), yielding a Poisson model $P(h_{i,k} D_i^a Q_k)$ (e.g., Raftery 1988). The expected number of individuals N_i in grid i is then $a^{-1} A_i D_i^a$. To account for the overdispersion of captures, we used a three-parameter version of the negative-binomial (NB) distribution (see Tremblay-Boyer et al. 2019 for detail) instead of the Poisson distribution. The NB model is commonly interpreted in terms of overdispersion (e.g., aggregation of individuals) relative to a Poisson distribution (randomly-distributed individuals in space). The NB model used here has two additional parameters (compared to the poisson): an overdispersion parameter (ϕ) and an exponent ν that describes how overdispersion varies with the mean of the distribution.

The definition of current abundance $N_{curr} = \sum_i N_i = \sum_i a^{-1} A_i D_i^a$ implies that it is possible to estimate the current abundance N_{curr} in principle. Nevertheless, this estimate is reliant on an estimate of the gear-affected area. For this reason, the estimation of N_{curr} has the same limitation as the scaling of survey estimates of biomass to a total biomass estimate. Although the estimation of N_{curr} provides a method for not requiring knowledge of gear

efficiency Q and the distribution of the focal species *a priori*, this method does not remove all unknowns from the estimation of N.)

The difficulty to disentangle Q and N in N -mixtures is well known, and Bayesian priors can help place constraints on these parameters (Raftery 1988). For instance, it can be assumed that the likelihood of catching an oceanic whitetip shark (given that it is present in the gear-affected area) is probably >1%; i.e., more than 1 in 100 sharks will be hooked given they are present in the area, but this value depends on what is considered the spatial extent of the gear-affected area. Given that some sharks may be satiated or otherwise not attracted to the bait, the gear efficiency is probably well below 1. We used a prior on the logit scale for the mean efficiency that was centred on 0 (i.e., $Q = 0.5$), with a standard deviation of 1. We then formulated our prior for mean density per hook area D_i^a, so that the product of Q and D_i^a was within a reasonable range: it would appear *a priori* implausible that the mean density is such that over plausible values of Q we would observe one oceanic whitetip shark per 10 hooks on average, or one oceanic whitetip shark for every ten million hooks (given known catches and total number of hooks used in the Western and Central Pacific Ocean). Therefore, we constrained the prior for $\log D_i^a$ as a normal distribution that leads to a joint prior over Q and D_i^a that remains within the *a priori* constraints (see Figure A-1 for the joint prior).

The stan model used for inference is shown below in full. The model was run using MCMC for 5000 iterations after discarding 1000 iterations as burn-in.

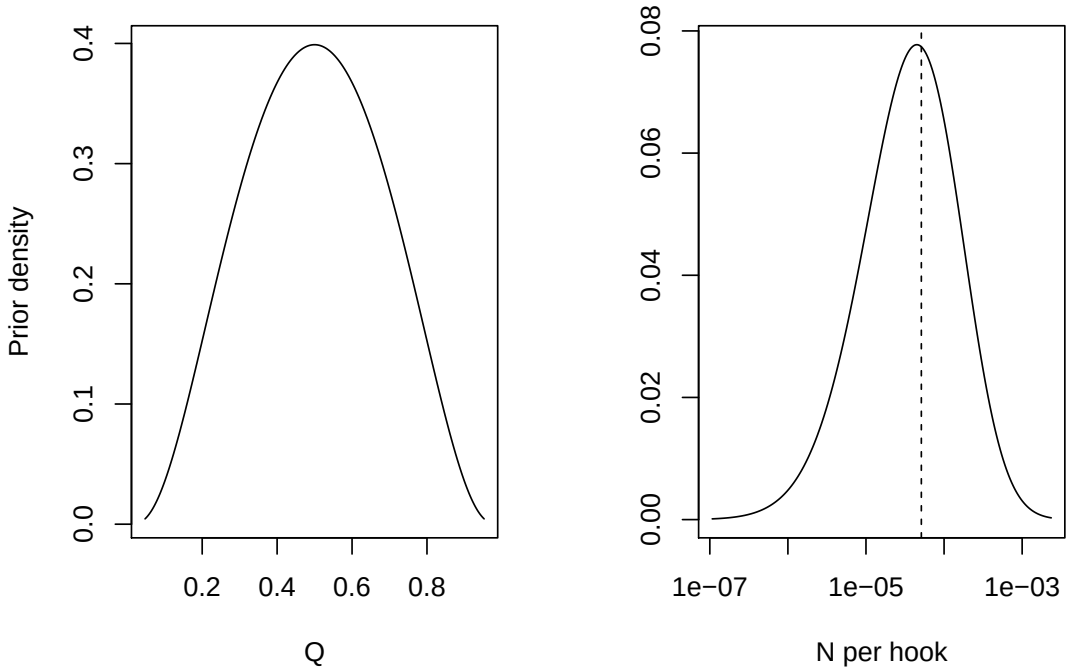

Figure A-1: Priors to constrain gear efficiency Q and mean density per hook area D_i^a. Left panel: Prior over Q. Right panel: Expectation for the number of oceanic whitetip shark caught per longline hook implied by the joint prior over Q and D_i^a. (The observed mean number of oceanic whitetip shark for the dataset is plotted as the vertical line for reference.)

```
// modified from code generated with brms 2.8.0
functions {

    real neg_binomial_2_trials_lpmf(int y, real mu,
       real phi, int T) {
        return neg_binomial_2_lpmf(y | mu * T, phi * T)
           ;
    }
    int neg_binomial_2_trials_rng(real mu, real phi,
       int T) {
        return neg_binomial_2_rng(mu * T, phi * T);
    }
    real neg_binomial_2_nu_lpmf(int y, real mu, real
       phi, real nu, int T) {
        return neg_binomial_2_lpmf(y | mu * T, phi * T
           * (mu^nu));
    }
    int neg_binomial_2_nu_rng(real mu, real phi, real
       nu, int T) {
        return neg_binomial_2_rng(mu * T, phi * T * (mu
```

```
            ^nu));
    }

}
data {
  int<lower=1> N;  // number of observations
  int<lower=1> N_grid;  // number of grid cells
  int<lower=1> N_pred;  // number of pred grid cells
  int pred_grid[N_pred]; //prediction grid
  int pred_hooks[N_pred]; //hooks for catch prediction
  int<lower=1> n_sc;  // number of gear-affected area
     scenarios
  real hook_area[n_sc]; //gear affected area
  vector[N_grid] Area; // area of each grid cell
  int Y[N];  // response variable - inference
  int trials[N];  // number of hooks - inference
  int<lower=1> K;  // number of population-level
     effects
  matrix[N, K] X;  // population-level design matrix
  matrix[N_pred, K] Xp; //effects for prediction (
     design matrix)
  // data for smooth terms
  int Ks;
  matrix[N, Ks] Xs;
  matrix[N_grid, Ks] Xss;
  // data of smooth s(SST,k=3)
  int nb_1;  // number of bases
  int knots_1[nb_1];
  matrix[N, knots_1[1]] Zs_1_1;
  matrix[N_grid, knots_1[1]] Zss_1_1;
  // data for group-level effects of ID 1 (flag)
  int<lower=1> N_1;
  int<lower=1> M_1;
  int<lower=1> J_1[N];
  int<lower=1> Jp_1[N_pred];
   // data for group-level effects of ID 2 (flag-year)
  int<lower=1> N_2;
  int<lower=1> M_2;
  int<lower=1> J_2[N];
  int<lower=1> Jp_2[N_pred];
  matrix[N_1,3] fate; // mortality scenarios, predicted
      discards by flag * mort
}
transformed data {
  int Kc = K - 1;
  matrix[N, K - 1] Xc;  // centered version of X
  matrix[N_pred, K - 1] Xcp;  // centered version of X
  vector[K - 1] means_X;  // column means of X before
     centering
```

```
  for (i in 2:K) {
    means_X[i - 1] = mean(X[, i]);
    Xc[, i - 1] = X[, i] - means_X[i - 1];
    Xcp[, i - 1] = Xp[, i] - means_X[i - 1]; // for
       predictions
  }
}
parameters {
  vector[Kc] b;  // population-level effects
  real temp_Intercept;  // temporary intercept
  real b_Intercept_mu;  // temporary intercept
  // parameters for smooth terms
  vector[Ks] bs;
  real<lower=0> phi;
  real<lower=0> nu;
  // parameters of smooth s(SST,k=3)
  vector[knots_1[1]] zs_1_1;
  real<lower=0> sds_1_1;
  vector<lower=0>[M_1] sd_1;  // group-level standard
     deviations
  vector[N_1] z_1[M_1];  // unscaled group-level
     effects
  vector<lower=0>[M_1] sd_2;  // group-level standard
     deviations
  vector[N_2] z_2[M_2];  // unscaled group-level
     effects

}
transformed parameters {
  vector[knots_1[1]] s_1_1 = sds_1_1 * zs_1_1;
  // group-level effects
  vector[N_1] r_1_1 = (sd_1[1] * (z_1[1]));
  vector[N_2] r_2_1 = (sd_2[1] * (z_2[1]));
  // group-level effects
  vector[N] mu = b_Intercept_mu + Xs * bs + Zs_1_1 *
     s_1_1; //expected population density
  vector[N] p  = temp_Intercept + Xc * b; //gear
     efficiency
  for (n in 1:N) {
    p[n] += r_1_1[J_1[n]]+r_2_1[J_1[n]]; // add random
       effects for gear
    p[n] = inv_logit(p[n]);
  }
}
model {

  // priors including all constants
  target += normal_lpdf(b | 0, 3);
  target += normal_lpdf(b_Intercept_mu | -9, 2);
```

```
  target += normal_lpdf(temp_Intercept | 0, 1);
  target += normal_lpdf(bs | 0, 3)
    - 1 * normal_lccdf(0 | 0, 3);
  target += normal_lpdf(zs_1_1 | 0, 1);
  target += student_t_lpdf(sds_1_1 | 3, 0, 20)
    - 1 * student_t_lccdf(0 | 3, 0, 20);
  target += student_t_lpdf(sd_1 | 3, 0, 2)
    - 1 * student_t_lccdf(0 | 3, 0, 2);
  target += student_t_lpdf(sd_2 | 3, 0, 2)
    - 1 * student_t_lccdf(0 | 3, 0, 2);
  target += normal_lpdf(z_1[1] | 0, 1);
  target += normal_lpdf(z_2[1] | 0, 1);

  // likelihood including all constants
  for (n in 1:N)  Y[n] ~ neg_binomial_2_nu(exp(mu[n]) *
      p[n],1/phi^2,nu, trials[n]);

}
generated quantities {
  // actual population-level intercept
  real b_Intercept_p = temp_Intercept - dot_product(
     means_X, b);
  vector[N] pY;
  vector[N_grid] mu_p;
  matrix[N_grid,n_sc] N_p;
  matrix[N_grid,n_sc] FS;
  vector[N_pred] p_pred;
  matrix[N_grid,3] pred_catch;
  vector[N_grid] hooks;

  // prediction for psoterior predictive diagnostics
  for (n in 1:N) pY[n] = neg_binomial_2_nu_rng(exp(mu[n
     ]) * p[n],1/phi^2,nu,trials[n]);

  // predictions on over-all grid
  mu_p = b_Intercept_mu + Xss * bs + Zss_1_1 * s_1_1;
     // predicted density
  p_pred  = temp_Intercept + Xcp * b; //
  pred_catch = rep_matrix(0,N_grid,3); //predicted
     catch
  hooks = rep_vector(0,N_grid);
  for (n in 1:N_pred) {
    int draw;
    p_pred[n] += r_1_1[Jp_1[n]] + r_2_1[Jp_2[n]];
    p_pred[n] = inv_logit(p_pred[n]);
    draw = neg_binomial_2_nu_rng(p_pred[n] * exp(mu_p[
       pred_grid[n]]), 1/phi^2, nu, pred_hooks[n]);
    hooks[pred_grid[n]] += pred_hooks[n];
    for(f in 1:3)  pred_catch[pred_grid[n],f] += draw*
```

```
        fate[Jp_1[n],f]; //apply mortality scenarios
    }

    for (sc in 1:n_sc) {
      N_p[,sc] = exp(mu_p) .* (Area/hook_area[sc]); //
         expected density over grid
      FS[,sc] = (inv_logit(b_Intercept_p) * exp(mu_p)
         .* hooks) ./ N_p[,sc];// spatial F = expected
         catch over expected numbers in each grid
    }
}
```

APPENDIX B: Supplementary figures and tables

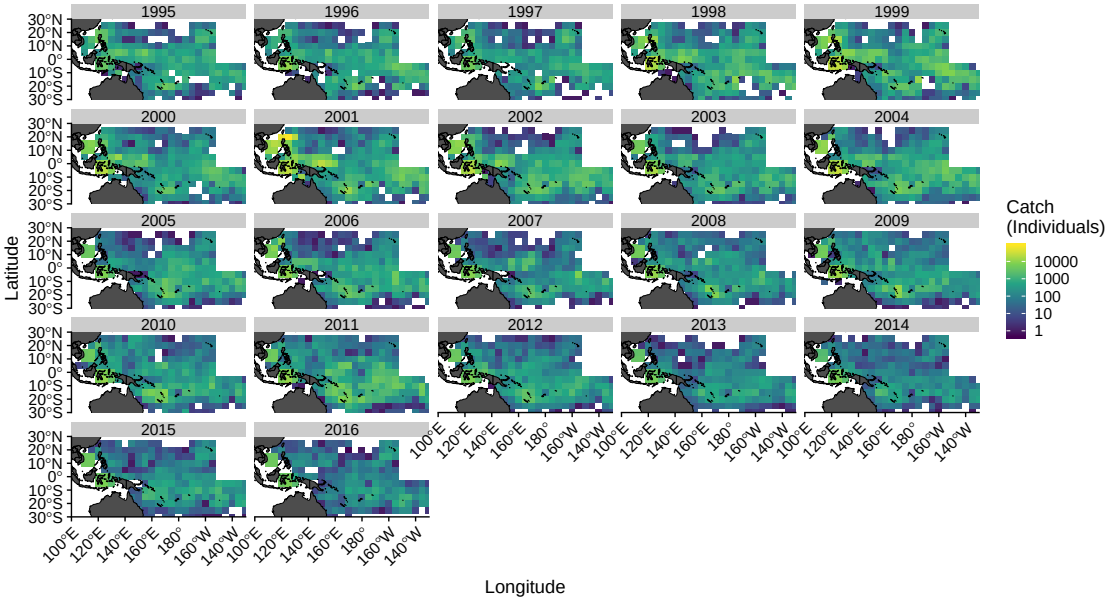

Figure B-2: Model-predicted longline catch of oceanic whitetip shark by year for the Western and Central Pacific Ocean, with the boundary of the Western and Central Pacific Fishery Commission indicated by the red line.

Table B-1: Marginal posterior estimates for selected parameters of the N-mixture model applied to oceanic whitetip shark in the Western and Central Pacific Ocean. Parameters included shark density D = exp(μ_D); gear efficiency Q= intercept component of β_Q (back-transformed); sd, standard deviation of random effects parameters.

Parameter	Mean	SD	2.5%	50%	97.5%	Rhat
D	4.33e-05	3.77e-01	2.17e-05	4.26e-05	9.58e-05	9.99e-01
Q	1.41e-01	4.84e-01	5.58e-02	1.43e-01	3.01e-01	9.99e-01
phi	6.03e-01	2.50e-01	2.47e-01	5.62e-01	1.18e+00	1.00e+00
nu	1.20e+00	8.19e-02	1.05e+00	1.20e+00	1.36e+00	1.00e+00
sd(Flag)	7.74e-01	4.44e-01	3.63e-02	7.98e-01	1.68e+00	1.00e+00
sd(Flag-year)	7.95e-01	4.27e-01	5.44e-02	8.09e-01	1.62e+00	1.00e+00

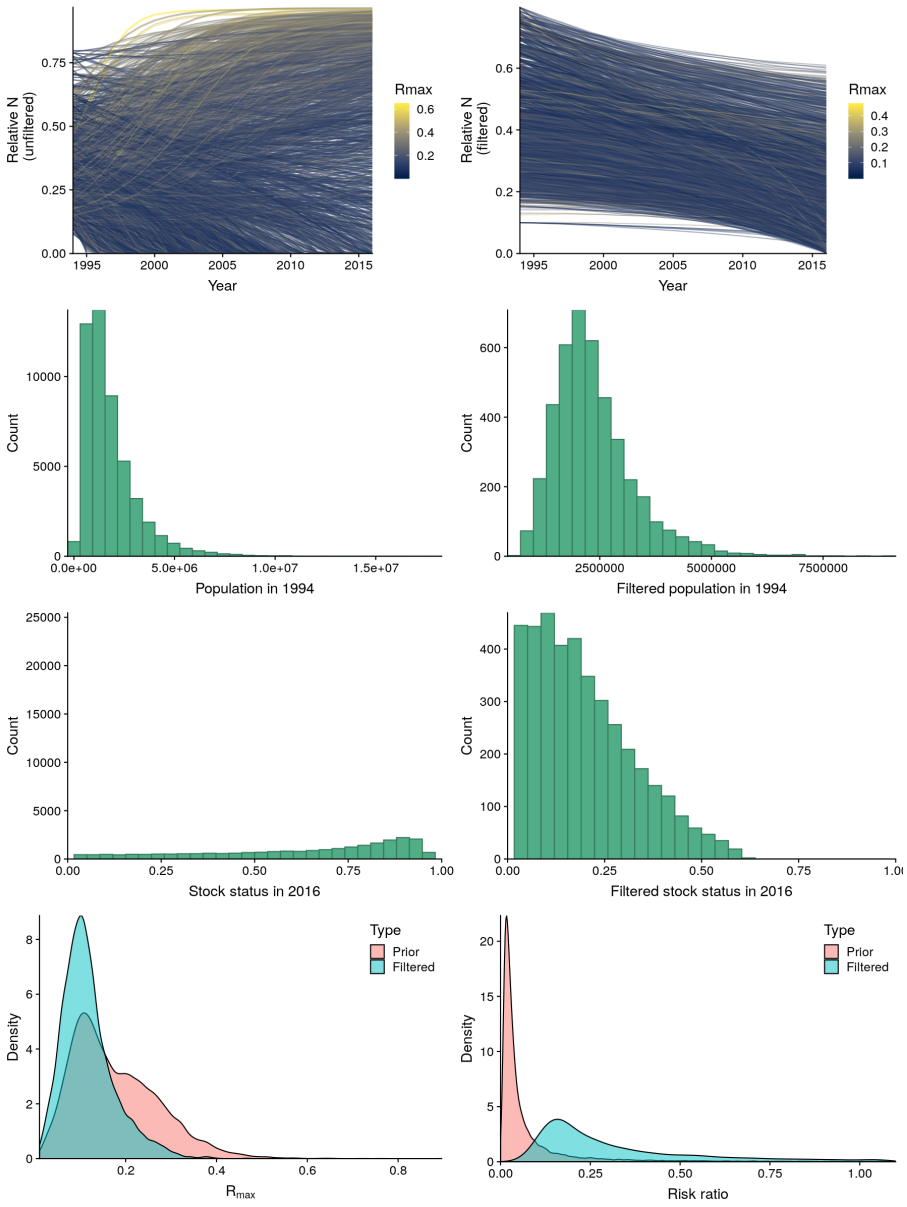

Figure B-3: Summary of prior predictive simulations under assumed 75% live-discards for oceanic white-tip shark, and no post-release mortality. Top row: Simulated population trajectories (in terms of relative abundance N_t/K) coloured by the value of the draw from the prior distribution of R_{max}. For each simulation trajectory, a set of values for carrying capacity, initial depletion, and R_max were drawn from their prior distribution, and estimated catch with estimated discards were applied with no discard mortality. A subset of 1000 trajectories is shown on the left hand side, and a subset of 1000 trajectories from the filtered set (after applying constraints on current depletion (abundance relative to 1994). The corresponding draws from the prior distribution of stock size in 1994 are shown (2nd row) for the original prior and the constrained (filtered) prior. The prior distribution over stock status corresponding to the unconstrained prior (left) and the constrained prior (right) is shown in the third row. The constrained prior can be thought of as a joint Bayesian prior over parameters and current stock status in the simple surplus production model, and therefore implies a constrained prior for R_{max} and overfishing risk (last row; overfishing risk in terms of $F_{curr}/F_{crash} = F_{curr}/R_{max}$).

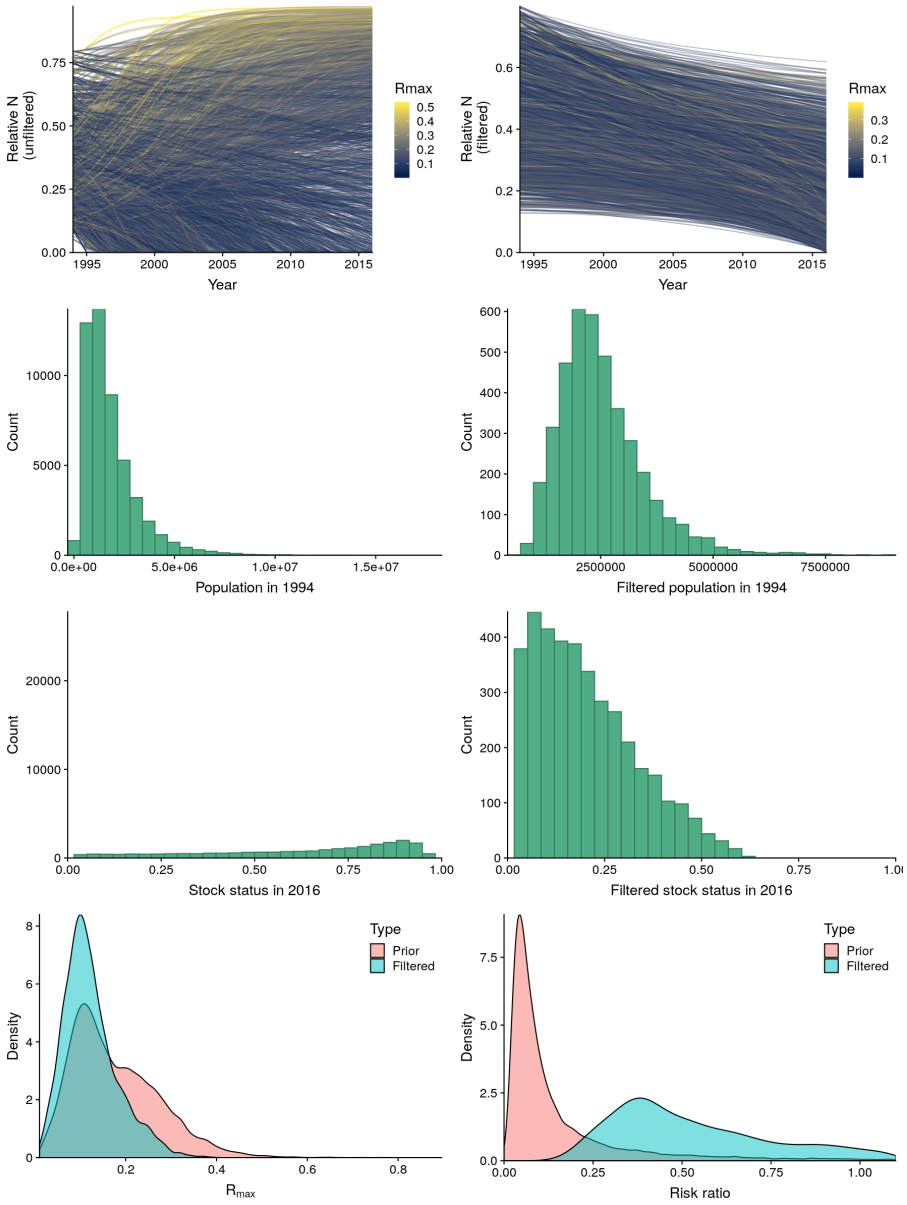

Figure B-4: Summary of prior predictive simulations under assumed 100% discard mortality for oceanic white-tip shark. Top row: Simulated population trajectories (in terms of relative abundance N_t/K) coloured by the value of the draw from the prior distribution of R_{max}. For each simulation trajectory, a set of values for carrying capacity, initial depletion, and R_max were drawn from their prior distribution, and estimated catch from the catch reconstruction was applied with 100% mortality (i.e., no survival for discarded sharks). A subset of 1000 trajectories is shown on the left hand side, and a subset of 1000 trajectories from the filtered set (after applying constraints on current depletion (abundance relative to 1994). The corresponding draws from the prior distribution of stock size in 1994 are shown (2nd row) for the original prior and the constrained (filtered) prior. The prior distribution over stock status corresponding to the unconstrained prior (left) and the constrained prior (right) is shown in the third row. The constrained prior can be thought of as a joint Bayesian prior over parameters and current stock status in the simple surplus production model, and therefore implies a constrained prior for R_{max} and overfishing risk (last row; overfishing risk in terms of $F_{curr}/F_{crash} = F_{curr}/R_{max}$).

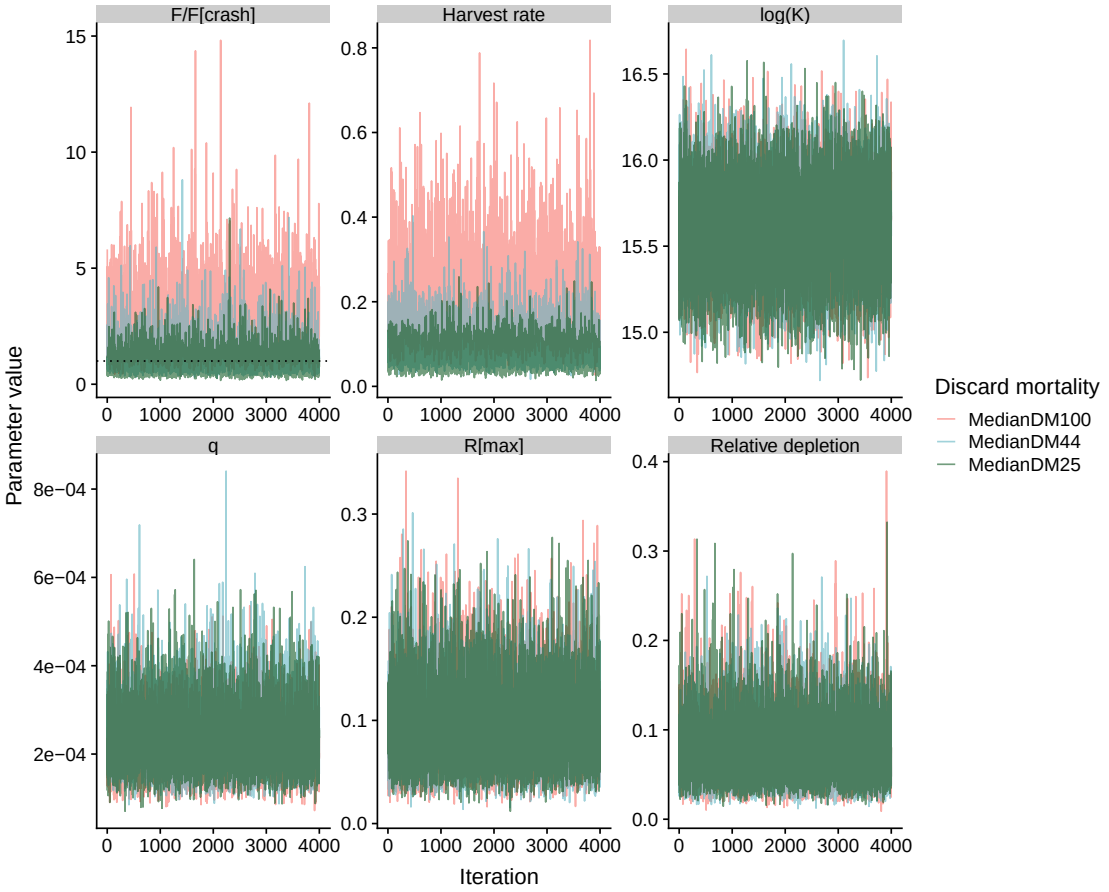

Figure B-5: Markov Chain Monte Carlo traceplots for derived parameters (harvest rate and risk ratio of overfishing (F/F_{crash}) and selected estimated parameters in the surplus production model of oceanic whitetip shark (log carrying capacity K, catchability q, intrinsic population growth R_{max} and relative depletion). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). The horizontal line for F/F_{crash} corresponds with $F/F_{crash} = 1$, the fishing mortality that leads to extinction.

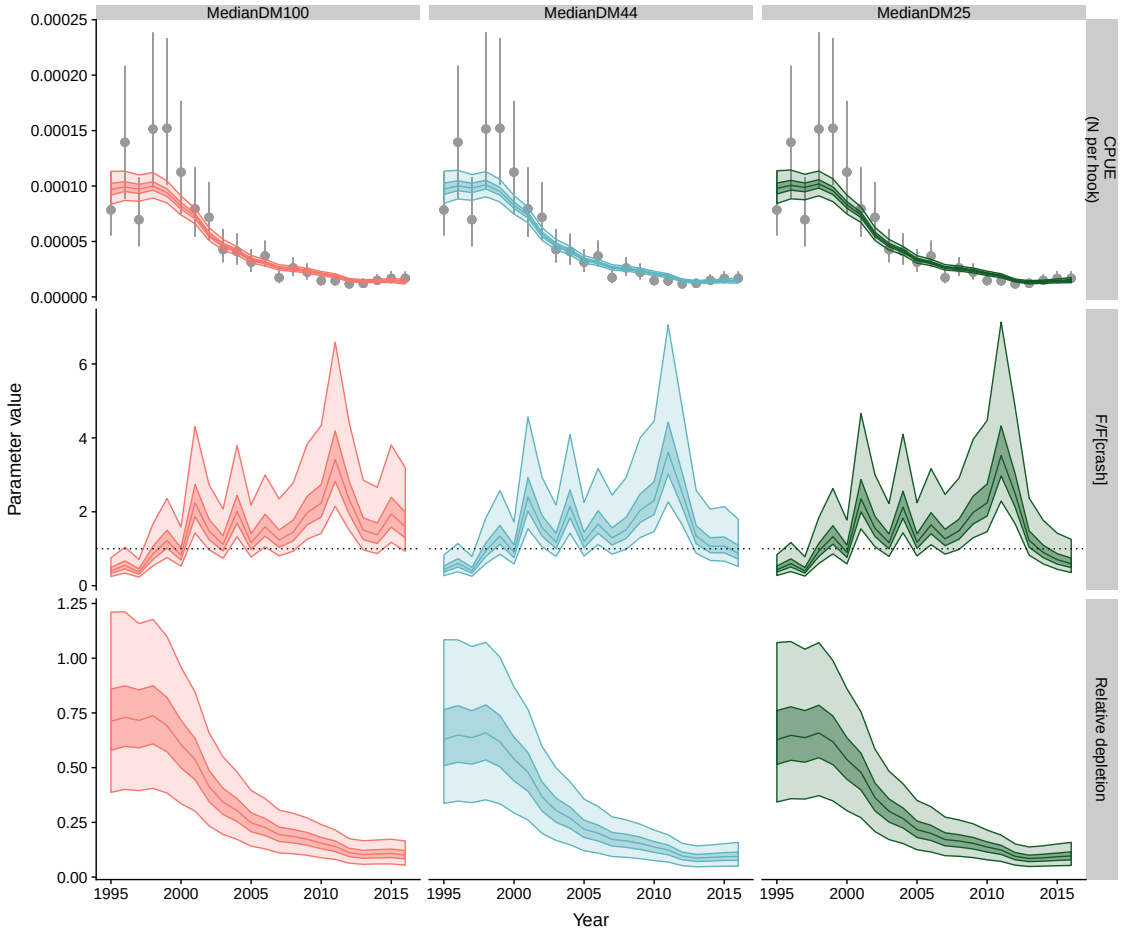

Figure B-6: Fitting of catch-per-unit-effort (CPUE) data using the surplus production model for oceanic whitetip shark with reduced CPUE observation error (original observation error plotted; dark shading, inter-quartile; light shading, 95% confidence interval). Top row: Predicted CPUE and inter-quartile from the surplus production model with reduced CPUE observation error. Middle row: Time series of risk ratio of overfishing F/F_{crash} estimated in the surplus production mode. Bottom row: Estimated relative depletion (relative to unfished abundance K). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). (Note that the stock was not unfished in the first year of the time-series.)

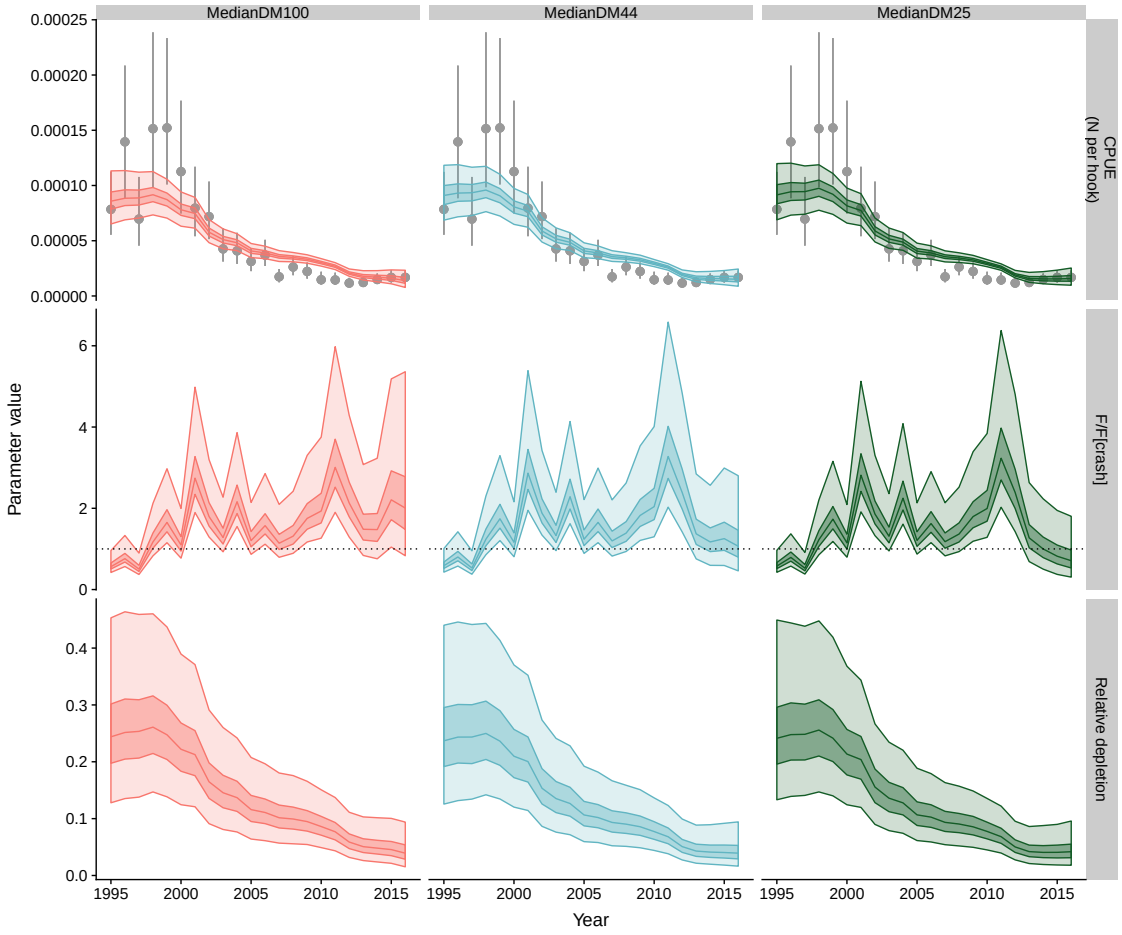

Figure B-7: Fitting of catch-per-unit-effort (CPUE) data using the surplus production model for oceanic whitetip shark with a reduced prior for the initial population (median CPUE and observation error plotted; dark shading, inter-quartile; light shading, 95% confidence interval). Top row: Predicted CPUE and inter-quartile from the surplus production model. Middle row: Time series of risk ratio of overfishing F/F_{crash} estimated in the surplus production model. Bottom row: Estimated relative depletion (relative to unfished abundance K). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). (Note that the stock was not unfished in the first year of the time-series.)

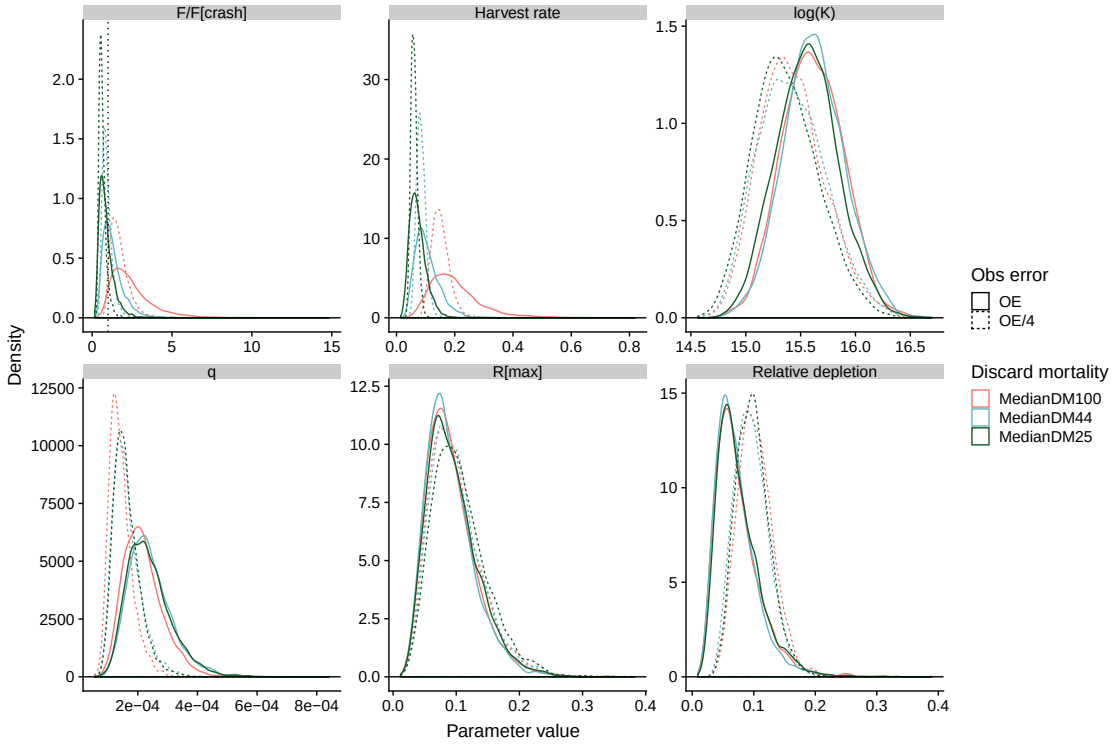

Figure B-8: Marginal posterior densities for selected estimated parameters (carrying capacity K, catchability q, intrinsic population growth R_{max} and relative depletion) and derived parameters (harvest rate and risk ratio of overfishing (F/F_{crash}) in the surplus production model for models run with the estimated observation error (OE; solid lines) and observation error set to a quarter of the original observation error (OE/4, dashed lines). Discard mortality scenarios were 100% discard mortality (MedianDM100), median estimated live discards (75% alive) with no post-release mortality (MedianDM25), and assuming 44% total discard mortality (MedianDM44). The vertical line for F/F_{crash} corresponds with $F/F_{crash} = 1$, the fishing mortality that leads to extinction.

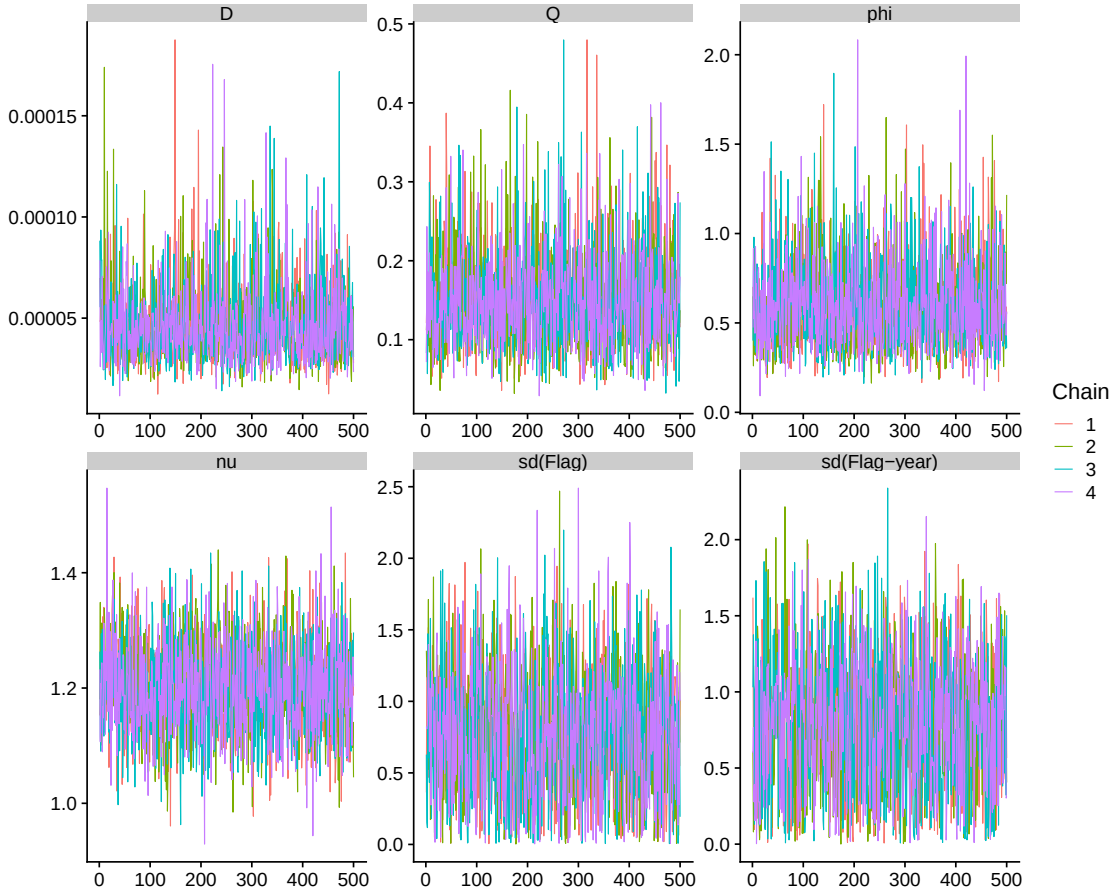

Figure B-9: Markov Chain Monte Carlo traceplots for selected parameters of the N-mixture model of oceanic whitetip shark (shark density D = exp(μ_D), gear efficiency Q= intercept component of β_Q (back-transformed)).

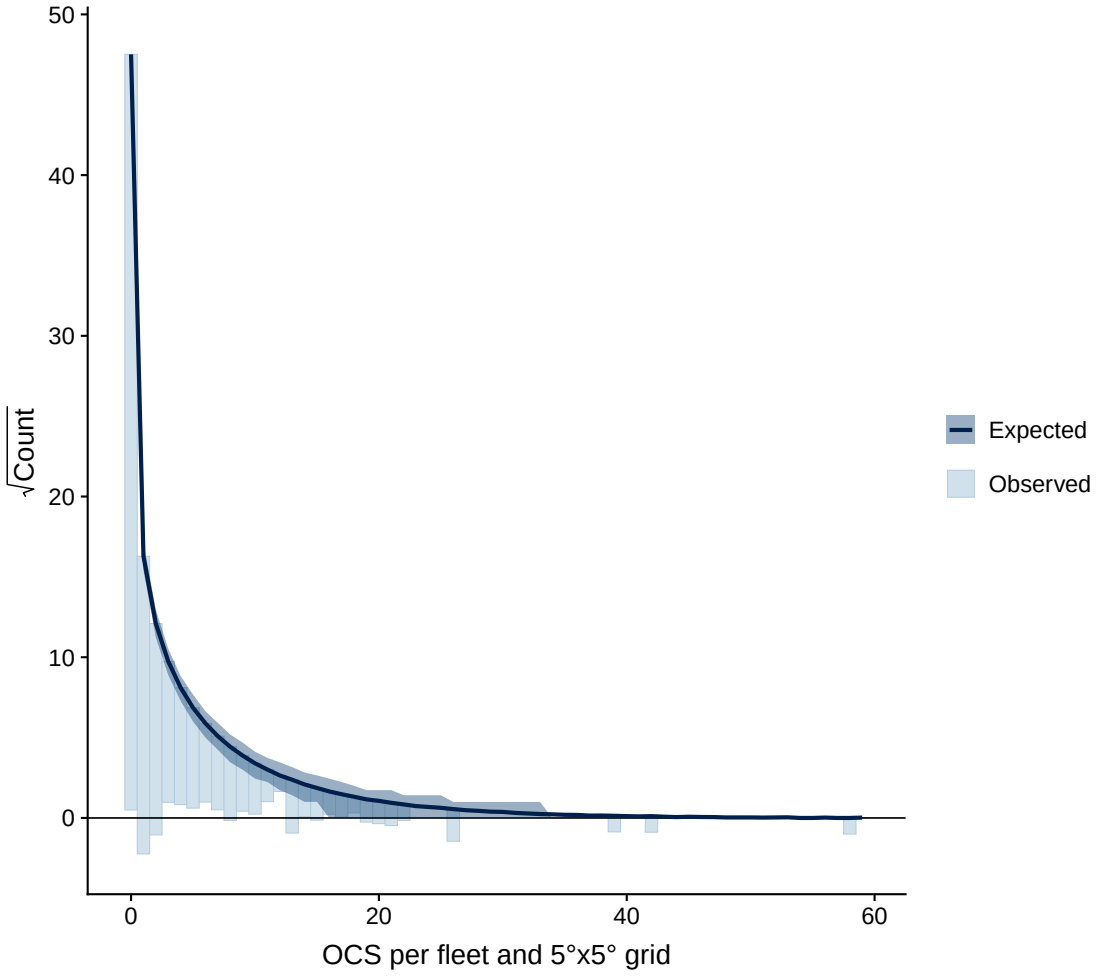

Figure B-10: Rootogram to assess model fit of the N-mixture model for oceanic whitetip shark (OCS) in the Western and Central Pacific Ocean. Expected counts are shown as a line (on square-root scale), with observed counts "hanging" from the line, indicating over- and under-prediction based on the distance of the bar from the y = 0 line.

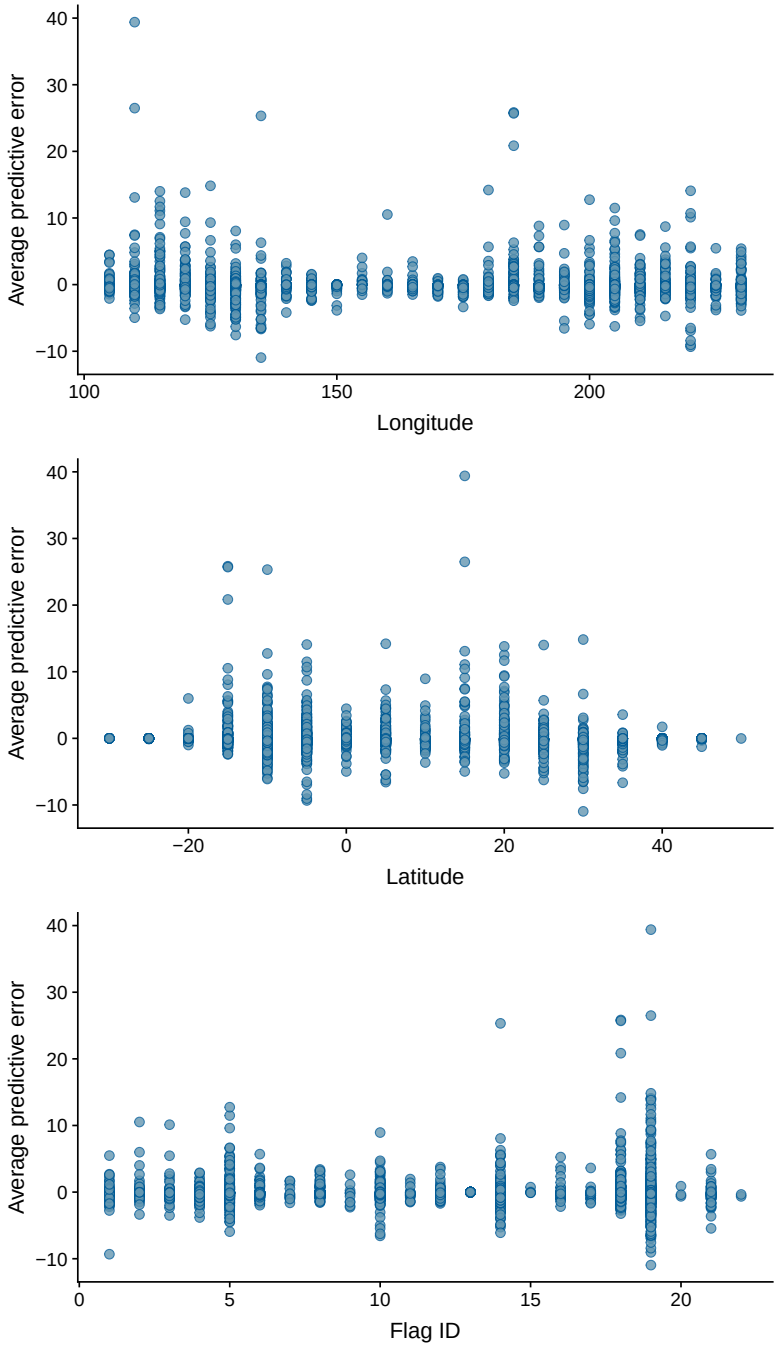

Figure B-11: Residuals from the N-mixture model for longitude, latitude and fishing fleet flag identification (ID).