SCIENTIFIC COMMITTEE
THIRTEENTH REGULAR SESSION

Rarotonga, Cook Islands
9-17 August 2017

STOCK ASSESSMENT OF SWORDFISH (*Xiphias gladius*) IN THE SOUTHWEST PACIFIC OCEAN

WCPFC-SC13-2017/SA-WP-13

Yukio Takeuchi[1],
Graham Pilling and John Hampton

[1] The Pacific Community (SPC), Oceanic Fisheries Program, Nouméa, New Caledonia

Executive summary

This paper presents the 2017 assessment of swordfish (*Xiphias gladius*) in the southwest Pacific Ocean (SWP). The model time period now extends to the end of 2015, adding a further four years of data since the last stock assessment was conducted in 2013. This assessment report describes all the analyses that provided inputs to the assessment, including the development of standardised CPUE analyses.

One of the key recommendations arising from the 2013 assessment of this stock was the use of a sexually-disaggregated stock assessment model, to better account for sexual dimorphism and spatial heterogeneity in sex ratios. For the 2017 assessment, considerable effort was expended in further developing this functionality within MULTIFAN-CL (MFCL), and runs were attempted assuming an explicitly modelled sex-structured population, while allowing most of fishery data to be sex aggregated. As we proceeded, it became evident that some additional important features were necessary in MFCL to conduct an explicitly sex-disaggregated swordfish assessment using primarily sex-aggregated data. It was not possible to undertake the necessary software development in the available time, therefore it was decided that sex-aggregated swordfish population dynamics be modelled for the SWP swordfish assessment in 2017. As for the 2013 assessment, we adopted a two-region model, delineated at 165°E, for the WCPFC area south of the Equator.

The progression from the 2013 reference case model to the 2017 diagnostic case model included the following steps:

1. The 2013 reference case model.
2. The 2013 reference case model with the new MFCL executable.
3. A complete update to the 2013 reference case model – all inputs extended from 2011 to 2015.
4. The previous model with modifications to selectivity constraints, forms, groupings and time-series break in selectivity for the 04_AU_1 fishery.
5. The previous model with updated growth and maturity parameters from Farley *et al*. (2016).
6. The previous model with up-weighted size data, consistent with the approach for WCPO tuna assessments.

In addition to the diagnostic case model, we report the results of one-off sensitivity models to explore the relative impacts of key data and model assumptions for the diagnostic case model on the stock assessment results and conclusions. We also undertook a structural uncertainty analysis (model grid) for consideration in developing management advice where all possible combinations of the most important axes of uncertainty from the one-off models were included. No particular emphasis is placed on the diagnostic case model for the purpose of reporting stock status. Instead it is recommended that management advice is formulated from the results of the structural uncertainty grid.

Across the range of model runs in this assessment, the most influential factor with respect to model output depended upon the management quantity of interest. Depletion was notably influenced by assumptions on diffusion (movement rate) between the two model regions, with increasing levels of movement implying notably more pessimistic results in terms of depletion, and increasing levels of fishing mortality. Fishing mortality relative to F_{MSY} was notably influenced by the assumed value of the stock recruitment relationship (SRR) steepness parameter, with the expected trend of increasing values of steepness leading to more optimistic results and reduced variability in estimates of F_{recent}/F_{MSY}. Depletion estimates were much more robust to the assumed steepness value, and changing that assumption had little impact on the level of variability in estimates.

Based on the results of the model grid, the general conclusions of this assessment are as follows:

1. The grid contains a wide range of models with some variation in estimates of stock status, trends in abundance and reference points. Biomass is estimated to have declined throughout the model period for all models in the grid, but the decline is particularly steep in the last 15 years. Those declines are found in both model regions, but are particularly notable in region 2 (the eastern region).

2. Fishing mortality for juvenile (ages 1-3), maturing (ages 4-6) and adult (ages 7+) swordfish is estimated to have increased since the 1950s. Fishing mortality rate increased notably from the mid 1990s in both model regions, on maturing aged fish in particular (seen in the diagnostic case model), to levels approximately four times that of juveniles and adults.

3. Noting that WCPFC has yet to formally agree a limit reference point for SWP swordfish, we have reported the main stock assessment results in terms of both spawning potential depletion and maximum sustainable yield (MSY)-related reference points. Across the model grid, the terminal spawning potential depletion estimated for all runs, $SB_{latest}/SB_{F=0}$, was above $20\%SB_{F=0}$. The median estimate was 0.35 (range 0.26-0.49). The median ratio of SB_{latest} to SB_{MSY} was 1.61 (range 0.85-4.06, 11% of which were < 1.0).

4. The median estimate of F_{recent}/F_{msy} was 0.86 (range 0.42-1.46), with 23 out of the 72 runs (32%) indicating that $F_{recent}/F_{msy} > 1$. Runs where overfishing was indicated were generally those with a steepness of 0.65 assumed.

5. Unlike in the bigeye and yellowfin assessments, evidence for a strong increase in recent recruitment for swordfish was not found in either the CPUE time series or estimates of recruitment. Variability in the recruitment estimates for swordfish may in part mask any recent trend. We also note that the longline-only nature of the fishery, catching mainly larger, older swordfish, is not strongly informative with regards to recruitment dynamics.

6. The current assessment investigated the impact of a wide range of uncertainties. However, a key axis of uncertainty in the 2013 assessment – growth – has been reduced in the current assessment through the results from Farley *et al*. (2016). Nonetheless, there remains a range of other model assumptions that should be investigated either internally or through directed research. These are noted in the main text, but briefly, include further developments to MFCL to enable the sex-disaggregated assessment of this stock (given the data available), enhancement of sex-separated data collection, investigations into potential stock structure, further analysis of the size data available, and consideration of additional data required to enhance CPUE standardisation given the decline in fishing by key long-term fleets within the SWP.

1. Introduction

Swordfish (*Xiphias gladius*) is one of six species of billfishes commonly reported from commercial longline fisheries within the western and central Pacific Ocean (WCPO) (Molony 2005). Swordfish in southwest Pacific (SWP, WCPFC Area south of the Equator) is an important bycatch species in many domestic and distant water fisheries and has been the focus of recently developing target fisheries in the waters of New Zealand, Australia, and in the high seas of the south Pacific by Spanish flagged longline vessels.

Previous assessments of swordfish in the south Pacific referenced the area 140E–175°W and used the integrated assessment models MULTIFAN-CL (MFCL, Kolody *et al*. 2006) and CASAL (Davies *et al*. (2006). This assessment was updated for SC4 in 2008 (Kolody *et al*. 2008) using MFCL, while a new CASAL-based assessment was also performed for the south-central Pacific alone (175°W-130°W; Davies *et al*. 2008). For the region west of 175°W, the MFCL assessment indicated that overfishing was not occurring and the stock was not in an overfished state. The CASAL assessment attempted for swordfish in the south-central Pacific was unable to determine the stock status due to a range of factors including the shortness and lack of contrast in the Spanish longline CPUE series and the conflict between the CPUE series for the Chinese Taipei fleet and other fleets. Overall it was concluded that the available data did not indicate evidence of significant fishery impacts at that time. Combined assessments of the full area from 140°E to 130°W were unsuccessful.

In 2013 a new assessment was conducted using MFCL, which assumed two model regions delineated at 165°E in the WCPFC Area south of the Equator, based upon the results of electronic tagging programmes. The overwhelming source of uncertainty in that assessment was attributable to the assumptions for the growth, maturity and mortality-at-age schedules. These were taken directly from the 2008 assessment in the absence of new information, and comprised two main schedules, based on estimates provided by scientists from the Pacific Islands Fisheries Science Center (PIFSC) in Hawaii and CSIRO in Australia. These assumptions dominated the uncertainty estimates derived over the key model runs and from the structural uncertainty analysis. It was concluded that the stock was not in an overfished state, with spawning potential at 26 - 60% (range of key model runs) of the level predicted to exist in the absence of fishing. The overfishing status depended upon the assumption regarding growth. $F_{current}/F_{MSY}$ was estimated to be between 0.33 and 1.77 (range of key model runs). Within this range, assuming the PIFSC growth curve produced estimates between 0.40 and 0.70, while assuming the CSIRO growth curve produced estimates between 1.06 and 1.77.

This report details a MFCL assessment of the SWP swordfish stock using a combined two-region "south-western" and "south-central" spatial structure, based upon the results of electronic tagging programmes and comparable to the approach taken in 2013. Substantial new information has been added to the assessment, including an additional four years of data. We used a new growth curve based on Farley *et al*. (2016), which resolved the earlier differences in growth estimated by PIFSC and CSIRO. The model includes five standardised CPUE indices for longline fisheries. Model assumptions for fisheries selectivity and statistical weighting of the model fit to observations have also been updated. Finally, many of the issues examined herein were discussed in detail, and recommendations to the assessment approach made, at the SPC Pre-assessment workshop held in Noumea over 24–27 April, 2017 (Pilling and Brouwer, 2017).

2. Background

Much of the background material in this report repeats that of Kolody *et al*. (2008) and Davies *et al*. (2013) since much of the fisheries and biological information remains relevant.

2.1. Stock structure

Swordfish are one of the most widely distributed pelagic species, distributed globally, and observed from 50°N to 50°S and at all longitudes in the Pacific Ocean. Japanese longline catch rate distributions suggest

three large, relatively high density areas, the North-West, South-West and Eastern Pacific. In contrast, spawning distributions (as inferred from larval surveys, Nishikawa *et al*. 1985, and maturity studies, e.g. Young and Drake 2002, Mejuto *et al*. 2008a) tend to suggest spawning only in tropical and sub-tropical areas, though with conspicuous absence from the Western Pacific equatorial region, and the coastal regions of North and South America. The degree to which individuals migrate and sub-populations mix potentially has important implications for fisheries management, but the effective stock structure is poorly understood. Genetic studies indicate that there is not uniform gene flow among Pacific swordfish populations. Reeb *et al*. (2000) suggest a broad "⊃"-shaped connectivity pattern, such that the SW and NW Pacific populations are the most distinct from each other, with central and eastern populations intermediate between the two. Alvarado Bremer *et al*. (2006) concluded that the SE Pacific population was genetically distinct from the NE and SW. There was additional evidence to suggest that the south-central Pacific represented a population intermediate between the SW and SE, but it was recognized that sample sizes in the south-central region were not sufficient to be conclusive.

In recent years, PSAT and opportunistic conventional tagging programs in the SWP have begun to provide direct information about the movement of individuals (Karen Evans and Chris Wilcox, CSIRO, pers. comm.; Holdsworth *et al*. 2007; Kolody and Davies 2008). Tagging seems to confirm that swordfish undergo directed seasonal migrations between temperate foraging grounds and tropical spawning grounds, but it remains unclear how much site fidelity individuals maintain between these migrations. The large-scale collaboration on swordfish electronic tagging in the South Pacific (Evans *et al*. 2012) has provided information, which is discussed in Section 3.6.

2.2. Biological Characteristics

Swordfish are sexually dimorphic (females grow larger and faster than males) and seem to have different spatial distributions (e.g. Young and Drake 2002; Mejuto *et al*. 2008a). Potential sexual differences in other life history characteristics are largely unknown (e.g. migration patterns, natural mortality, etc.).

There have been a number of studies on swordfish growth rates and maturity in the SWP (e.g. Young and Drake 2002, 2004; DeMartini *et al*. 2000, 2007; Mejuto *et al*. 2008a; Valeiras *et al*. 2008) providing a range of estimates that have contributed to stock assessment uncertainty (Young *et al*. 2008). In response to this uncertainty, WCPFC SC recommended that additional work on age, growth and age validation be undertaken (Project 71). That research and results are fully described in Farley *et al*. (2016). New growth and maturity estimates were developed based on otolith readings, which indicated that swordfish lived longer and grew slower than previously estimated. SC12 endorsed the use of these estimates in future assessments.

2.3. Fisheries

Historically, the majority of swordfish catches represent bycatch from the tuna-target fisheries; a significant amount of recent catches remains a non-target bycatch. Across the two regions within this assessment, catches slowly increased from the early 1970s up until around 2000, caught primarily by Japanese fleets. Thereafter catch by the Japanese fleet has declined due to declining effort in the assessment regions. In contrast, catch levels of other nations increased rapidly from the mid-1990s, as more targeted Australian (south-west Pacific Ocean region) and New Zealand fisheries developed (south central Pacific Ocean region). After the year 2000, increasing catches were also taken by the Spanish and Chinese fleets in the south-central Pacific Ocean (Figure 1). In the current century, the majority of the swordfish catch has been in the tropical waters in the south central Pacific Ocean, particularly following declines in the Australian fishery from the early 2000s (Figure 1).

3. Data compilation

The total catch (in numbers) and size composition data for most fleets were provided from SPC databases. Analyses involving effort standardization of Japanese, Chinese Taipei, Australian and New Zealand fleets were conducted with additional fine-scale data with the cooperation of scientists from specific countries. The following briefly overviews the assessment data used in 2017. Much more detailed summaries and analyses of the catch, effort and size composition data were provided in Campbell (2008), and catch rate standardization analyses are detailed in Campbell *et al*. (2012) and Hoyle *et al*. (2013).

Data used in the swordfish assessment for the SWP consisted of fishery-specific catch and effort data, length-frequency data, weight-frequency data and tag-release-recapture data.

3.1. Spatial stratification

The overall model area for this assessment is consistent with the two-region assessment attempted by Kolody *et al*. (2008), based upon the spatial distribution of catches in the WCPFC region south of the Equator. As in that assessment, the model area is divided into two regions, south-west and south-central. The delineation of these two regions at 165°E followed Davies *et al*. (2013), based on the tagging analysis by Evans *et al*. (2012). The two model regions were each further divided into three sub-areas to aid fishery definition: a northern, central and southern fishery sub-area (Figure 2). In the 2013 assessment the model region was expanded to include a small portion outside of the convention area. This assessment excluded that extension and is based only on the WCPFC area south of the Equator.

3.2. Temporal stratification

Data used in the current assessment cover the period 1952–2015 (Figure 3). As agreed at SC12, the assessment does not include data from the most recent calendar year. This is because these data are only finalized very late and often subject to significant revision post-SC, in particular the longline data on which this assessment greatly depends. Given the seasonal patterns of catch and effort within fleets, the fisheries data were stratified on a quarterly basis (1; Jan–Mar, 2; Apr–Jun, 3; Jul–Sep, 4; Oct–Dec). However, the model dynamics (periodicity of recruitment, movement, etc.) were implemented with an annual time step.

3.3. Definition of fisheries

The fishery sub-areas of the model regions were applied to define the spatial boundaries of the specific fisheries in the SWP. A total of 13 longline fisheries were defined (Table 1), based on sub-area boundaries, nationality and time period. Historically, distant-water fishing nation longline fleets, primarily Japanese, have dominated the catches of swordfish, taken primarily as a bycatch. Recently, catches by Japanese fleets have declined. Instead, other distant-water fishing nations including China, Chinese-Taipei and Korea increased their swordfish catch rapidly. These fleets were grouped together for each fishery sub-area of the two model regions (Fisheries 1-3 and 6-9). The pattern of catches by fishery sub-area for these fisheries (identified by "DW" in the label) is shown in Figure 1.

Since the early 1990s, major longline fisheries have also developed in Australia (one fishery across the three sub-areas in region 1, Fishery 4) and New Zealand (one fishery in sub-areas 2C and 2S in region 2, Fishery 10), as well as the more recently developed EU (Spanish) fleet (one fishery in the three sub-areas in region 2, Fishery 11).

Three additional longline fisheries were defined to account for the other sources of longline effort and catch, one in region 1 (across all three sub-areas, Fishery 6) and two in region 2 (in sub-area 2c and 2s, respectively, Fisheries 12 and 13, Table 1). These other longline fisheries included effort and catches by recently developed longline fisheries of PICTs, and the Australian fleet in region 2 and EU fleet in region 1.

3.4. Catch and effort data

For all fisheries, catch data were expressed as the number of swordfish captured (Table 1) and fishing effort as the number of hooks set. As catches submitted by the EU are in weight, these had to be converted to numbers of fish using average weight data prior to their use in the assessment. Catch and effort data for all fisheries were aggregated within the quarterly time intervals.

Data were supplied in a variety of spatial and temporal resolutions. For example, longline catch and effort data from the distant-water fleets were generally available aggregated by month and 5-degree spatial resolution, while operational-level logsheet data were available for many of the domestic longline fleets. Compared to the 2013 assessment, new operational data sets were available for the Spanish, Chinese Taipei, China, Korea and Japanese fleets.

For each group of fisheries, we describe the patterns in catch and effort, and where appropriate, the standardised CPUE time series developed to inform the assessment.

Distant-water fishing nation longline fisheries (Fisheries 1-3 and 6-9): The distant-water fishery was primarily comprised of vessels from the fleets of Japan, China, Korea and Chinese Taipei. This fleet fished in both regions 1 and 2, primarily in the central and northern fishery sub-areas of these regions (Figure 1). Swordfish catches within the Japanese fleet were highest in sub-area 1C up until the mid-1990s, after which catches subsequently declined. The Korean fleet operated primarily in sub-area 2N from the mid 1970s to the early 2000s, after which catches of this fleet declined in that fishery sub-area. Key Chinese Taipei fleet catches were in sub-areas 2N and 2C from the 1970s, but increased particularly from the late 1990s, catches subsequently declining from the mid-2000s, particularly in sub-area 2C. Chinese fleet catches increased rapidly from 2001 in sub-area 2N in particular, peaking at over 40,000 individuals by the late 2000s.

From these data, two key standardised CPUE series were derived from the Japanese and Chinese Taipei fleets:

- Distant-water fishery CPUE in sub-area 1C (02_DW_1C): Operational catch and effort data from the Japanese fleet for 1952–2015 were supplied by the NRIFSF with detailed information (number of hooks between floats, HBF (mostly after 1975), vessel id (after 1979)). The data offers the longest time series of CPUE trends for swordfish. Fishing effort was standardised using a generalised linear model (GLM) approach. The GLM included the following variables: year/quarter, spatial cell (5° latitude/longitude cell), proportion of moon illumination and HBF. In order to derive stable time series a "core" sets were defined as 1) no. of sets were >25 in 5° latitude/longitude cell and Year/Quarter, 2) Year/Quarter with more than or equal to three 5° latitude/longitude cell satisfying condition 1. Because of the seasonal nature of this fishery a substantial amount of data in the first and fourth quarters were not used for analysis. In addition, declining effort mainly due to the decline of southern bluefin tuna targeted effort from the late 1990s resulted in a substantial amount of data in the second and third quarters of the past decade being excluded. The resulting CPUE indices are presented in Figure 4. For each year/quarter if a standardised CPUE index was available, an index of standardised effort was calculated by dividing the total quarterly catch by the CPUE index derived from the GLM. Estimates of time-variant precision for each standardised index were calculated, with the highest being for those for the early and most recent periods (Figure 4). In sub-area 1C, swordfish CPUE was relatively high between 1970 and the late 1990s, subsequently declined to the mid-2000s and then increased but did not reach the previous high levels.

- Distant-water fishery CPUE in sub-area 2C (07_DW_2C_pre-2001 and 08_DW_2C_post-2001): Data from the Chinese Taipei fleet were available from the late 1960s. In the 2013 assessment, catch and effort data aggregated by 5-degree square and month were used for this fleet. This year, operational catch and effort data for this fleet became available for the first time. Data were supplied by the Overseas Fisheries Council of the Republic of China via the Council of Agriculture. For this data set, vessel id information was available for most of the sets but no hooks-

between-floats (HBF) information was provided. Instead, clustering of proportion of catch by species (albacore, bigeye, yellowfin and swordfish) by year/quarter/vessel was conducted. The results of clustering suggest two major "clusters" probably representing "traditional" albacore targeting and a recently emerging bigeye targeting cluster. There was also a remaining very minor "unknown" targeting cluster. A "core" data set was derived by applying similar criteria as described for DW1C. As a result, most of the data in the first and fourth quarters were not included in the analysis. For the core data set, a CPUE standardisation with Year/Quarter, 5° latitude/longitude cell and clustering of catch proportion by species was conducted. The resulting standardised CPUE indicated an abrupt change of catchability of swordfish around 2001, as reported in the 2013 assessment. Similar to the last assessment, the CPUE time series was split for the period pre- and post-2001. Unique catchability is assumed for each fishery component. The resulting catch rate time series (Figure 4) was relatively stable until the early 2000s (evident in the pre-2001 index), and then declined (post-2001 index).

Australian fisheries (04_AU_1): Following relatively low catches within this fishery in region 1, catches in the central sub-area (1C) increased significantly to over 30,000 individuals per year from the late 1990s to early 2000s, and subsequently declined to below 20,000 individuals by the end of the time series. A small amount of catch was also taken by Australian vessels in sub-area 2C in the early to late 2000s. Those catches were included in fishery 13_Other_2C.

- Australian fishery CPUE in region 1 (04_AU_1): A CPUE index for swordfish caught by vessels operating within the longline sector of the Australian Eastern Tuna and Billfish Fishery (ETBF) has been updated to 2016 (Campbell, 2016.). The time-series of data used for the development of the index begins in mid-1997 when logbooks began collecting information on the gear settings used in the ETBF and continues to the end of 2016. Standardised CPUE indices (with catch being the sum of retained and discarded fish) are calculated for three different size classes of fish (Small, Prime and Large) as well as an index for all sizes classes combined. Retained catch for each trip is apportioned to each size class using size data (individual weight data) collected from processors receiving fish landed in the fishery while discards are apportioned based on observer data. Size data for swordfish have been collected for around 80% of all retained fish. For the all-size-classes-combined index used in this assessment, the CPUE was standardised by fitting both a binomial model to the probability of attaining a catch (with a logit link) and a negative-binomial model to the size of the positive catch (with a log link) and then combining the two results. The main effects in both models consisted of year, quarter, area (7), hooks-per-hook, bait-type, start-time, percentage of hooks with light-sticks, number of hooks-per-kilometer, and set-type (based on the two-way interaction between mainline-length and the distance between floats) as well as several environmental effects (including daily moon-phase, weekly sea-surface temperature, mixed-layer-depth, wind-speed, bathymetry and the monthly value of the Southern-Oscillation-Index) together with two additional effects which account for competition between vessels within each 1-degree square. All effects except moon-phase were fitted as categorical variables with the Year, Quarter and Area effects fitted as either a full 3-way interaction or the sum of both Year*Quarter and Quarter*Area interactions with the latter chosen for the final index. Time-variant precision of the index for the combined size class was obtained from the component of the model fitted to the positive catch only. The index was rescaled to have a mean of 1 over the period 1997-3 to 2016-4. The standardised index displays a steady decline between 1997 and 2003 after which the index increases to 2007 then remains relatively stable through to the end of 2015 before a decrease in 2016. A more complete description of this work was provided to SC8 (see SC8-SA-IP13). The resulting CPUE indices and their precision are presented in Figure 4.

New Zealand domestic fishery (10_NZ_2): Catch from this fishery, in both numbers and tonnes of swordfish, were provided by the Ministry of Primary Industries, New Zealand, for the period 1993–2015. New Zealand catches have occurred in sub-area 2C, increasing from the early 1990s to a peak in the early 2000s. In the 2013 assessment two alternative standardised CPUE time series were developed from 1) New Zealand domestic vessels operating in region 2; and 2) Japanese charter fleet fishing off the west coast of the South Island of New Zealand. In the 2017 assessment, the standardised CPUE estimates were

not updated or used, primarily due to the concern of the effects on the CPUE by recently implemented regulations, as reported at the SPC PAW (Pilling and Brouwer, 2017). In future assessments, these should be revisited.

Spanish fishery (11_EU_2): Operational longline data were provided by the Spanish Oceanographic Institute for the years 2004-2015. Significant catches were taken in region 2, peaking in the late 2000s. The limited catch and effort within region 1 was, in this assessment, incorporated within fishery 05_Other_1 (see below).

- Spanish fishery CPUE in region 2: A standardised CPUE index was derived for the period 2004 to 2015. Notable problems in using these data for deriving a standardised CPUE index include the lack of operational factors available for standardising effort (numbers of hooks were not included in the data provided), and catch was expressed in units of weight (rather than numbers of fish). In addition, the number of vessels operating in the assessment area has declined since 2008, with the main fishing area of this fleet being in Eastern Pacific Ocean. The CPUE index was standardised relative to latitude and longitude, as the main factor influencing proportions of swordfish within catches, similar to the method applied in the 2013 assessment (OFP 2012). The resulting CPUE indices are presented in Figure 4. Catch rates initially declined but recovered somewhat in the most recent years.

Other longline fisheries (05_Other_1, 12_Other_2N and 13_Other_2C): Other longline fleets have also operated within the model region since 1952, aside from the fisheries identified above. These "other" longline fisheries were pooled into the relevant model sub-areas on a quarterly basis. These fisheries included fleets from PICTs (e.g. Fiji, New Caledonia, Papua New Guinea, Tonga, Vanuatu, Cook Islands, French Polynesia, American Samoa and Samoa), plus fleets from distant-water fishing nations other than the key fleets discussed above. All data for these other fisheries were supplied as logsheet data and/or aggregated spatial data, with effort and catches raised as appropriate.

3.5. Size data

Length-frequency and/or weight-frequency data were available from many of the fisheries defined in Table 1, although data were provided in a number of different formats depending on the specific fishery. There was high coverage in particular for the length-frequency data of the Spanish catch and weight-frequency for the Australian and New Zealand catches (Table 2). For the remaining fisheries, temporal coverage of the size-frequency data was somewhat sporadic (Figure 3). For inclusion in the assessment, size data were aggregated by fishery and time strata (year/quarter). Length data were aggregated into 29 10-cm size classes (30–310 cm EFL), which removed the first and second bins from the 2013 assessment. Weight data were aggregated into 31 10-kg intervals (2–312 kg whole weight).

Length data were provided based on two different length measurement methods: eye orbit–fork length (EFL), and lower jaw–fork length (LJFL). A range of weights were supplied including whole weight, Japanese processed weights (gilled, gutted, head and tail left on, bill removed at a point level with the tip of the lower jaw), and gilled, gutted and headed (i.e. trunked) weights. All length measurements were standardised to LJFL from EFL using the following relationship:

$$LJFL = 1.0753 * (EFL + 6.898)$$

(Campbell 2008), and weight measurements were standardised to the equivalent whole (unprocessed) weight. Data from these fisheries were supplied from a combination of regional observer programmes, regional port-sampling programmes, market data and/or from research institutes of distant water fleets.

Australian fisheries (04_AU_1): A large amount of weight data was provided by AFMA for the period 1997–2015 (Table 2, Figure 3). The weight data was originally sourced from the main fish processors receiving swordfish from Australian longline vessels and represents a comprehensive sample of almost the entire catch. Weights were supplied as processed (trunked) weights (i.e. gilled, gutted and head

removed) to the nearest 0.1 kg. To enable comparisons with whole weights, a conversion factor was calculated using processed and whole weight data collected by Australian observers on longline vessels operating in Australian waters. The relationship between the two measures was:

Whole weight (kg) = 1.3717 x (gilled-gutted weight (kg) - 0.5).

Length-frequency data, sourced from the AFMA observer programme, were relatively limited, compared with the weight-frequency data. Several issues were encountered with unrealistic length measurements, for example lengths of lower jaw-fork length (LFL) less than 40 cm, and lengths over 600 cm (several over 1,000 cm). These length records were not included in the model. The data exhibited a mode at relatively low sizes (~110-130 cm, Figure 5), comparable to that sampled from 01_DW_1N from region 1. This corresponded with a mode at approximately 10-30 kg seen in the substantial number of weight-frequency samples from this fishery (Figure 6). The median lengths and weights in the time series of data from this fishery (Figure 7 and Figure 8) showed a general decline until around 2010 and increased after that.

New Zealand fishery (10_NZ_2): Length data were available from observers on board New Zealand longline vessels during 1992–2015. Data were supplied by the Ministry of Primary Industries with lengths measured as LJFL.

Length-frequency data show a mode at 130-150 cm (Figure 5). Substantial weight-frequency samples were available from 2006 to 2011, showing a mode at 10-30 kg, comparable to that seen in 04_AU_1 (Figure 6). Median lengths and weights in the samples from this fishery declined notably across the short time period 2004-2015 (Figure 7 and Figure 8).

Distant water fishing nation fisheries: Within region 1, the following data were available:

- In sub-area 1N (01_DW_1N), length-frequency data were available between 1993 and 2015, while weight-frequency data were available sporadically for 1997 to 2002 (Figure 3). Similar to the 04_AU_1 fishery, the length-frequency from 01_DW_1N had a mode at relatively low size. The limited weight-frequency data available exhibited a mode at relatively higher weights, but these data were excluded from the assessment due to the small sample size.
- Within sub-area 1C (02_DW_1C), length-frequency data were primarily available between 1991 and 1998 with sporadic samples after that time. The data exhibited a mode at larger sizes (150-180 cm) than in the majority of other fisheries (Figure 5).
- Within sub-area 1S (03_DW_1S), length-frequency data were sporadically available from 1991 to 2001.
- Within sub-area 2N (06_DW_2N), length-frequency data were available from 1996 to 2015. From 2006 to 2011 significant amount of length-frequency data became available due to increased length measurements by Chinese-Taipei.

Examining the length-frequency distributions from these fisheries, the modes in the data tended to increase in size from north to south in region 1, with the limited samples available from the distant-water fishing fleet operating in sub-area 1S catches in particular containing samples with a high proportion of relatively large individuals.

Within region 2, the following data were available from the distant-water fleets:

- From fishery 06_DW_2N, samples were available between 1990 and 2015, with increasing coverage toward the end of the time series. Limited weight-frequency samples were sporadically available over the period between 1996 and 2010.
- From fisheries 07_DW_2C_pre-2001 and 08_DW_2C_post-2001, length-frequency samples were available from 1987 to 2011. These were primarily concentrated in the early and later years of this period. No weight-frequency data were available from these fisheries.
- From fishery 09_DW_2S, limited length-frequency samples were available from quarters during the period 1988 to 2006.

Length-frequency samples from the distant-water longline fisheries in region 2 showed comparable distributions in 2N and 2C, with some increase in size moving south. However, the limited samples from the fishery to the south (09_DW_2S) were of notably larger lengths than seen in all other fisheries in this sub-area and generally for the assessment region. Weight-frequency samples from the northern fisheries were comparable to those from the "Other" fisheries in that sub-area. In 06_DW_2N, the median weight within the samples showed a general increase through time over the period 1996-2010 (Figure 8).

Spanish fishery (11_EU_02): A large length-frequency data set was available for the period 2004 to 2015, having a mode around 150-200 cm.

Other fisheries (05_Other_1, 12_Other_2N and 13_Other_2C): Length data were available from fishery 05_Other_1 from 1993, with consistent quarterly samples collected during the period 1998–2015; data from fishery 12_Other_2N were available from 1992, with consistent quarterly sampling from the period 1994–2015; data from fishery 13_Other_2C were available from 1993, with consistent quarterly sampling from the late 1990s. Limited weight-frequency data were also available from fisheries 12_Other_2N and 13_Other_2C, from 1993 to 2015 and 1994 to 2015 respectively.

Length-frequency data from these fisheries displayed modes around 150 cm. For the weight-frequency data, modes were apparent at larger weights than seen in the Australian and New Zealand fisheries (Figure 6). For these fisheries, the median weight in the samples fluctuated over time (Figure 8).

3.6. Tagging data

A summary of the most recent swordfish tagging data was provided by Harley *et al*. (2012).

A large-scale collaboration on swordfish electronic tagging in the South Pacific was described in Evans *et al*. (2012). This programme provided over 50 electronic tag tracks with durations of greater than 30 days. The data indicated that, in combination with long duration conventional recoveries, a division of the stock into south-western and south-central regions (west and east of 175°W), as in the 2008 assessment, was not defensible on biological grounds. Significant differences in behaviour were found between fish tagged in the Tasman Sea and those tagged in the south Pacific Ocean to the east of New Zealand. Movement patterns across the Tasman and Coral Seas suggest limited mixing or the partial overlap of sub-populations that may not mix strongly on the spawning grounds. There appeared to be no mixing between the southern and northern WCPO, nor the WCPO and the eastern Pacific Ocean.

While the electronic tagging data in particular were useful in consideration of the appropriate regional structure for the model, the tagging data from the swordfish fishery were not directly incorporated as a data source in the current assessment model.

4. Model description

As with any model, various structural assumptions have been made in the SWP swordfish model. Such assumptions are always a trade-off to some extent between the need, on the one hand, to keep the parameterization as simple as possible, and on the other, to allow sufficient flexibility so that important characteristics of the fisheries and population are captured in the model. The mathematical specification of structural assumptions is given in Hampton and Fournier (2001). The main structural assumptions and fixed parameter settings used in the swordfish model are discussed below and summarized in Table 3.

As recommended in the last assessment report, sex-disaggregated population dynamics should be a key feature to reduce the uncertainty of assessments on species such as SWP swordfish that have known sexual dimorphism and spatial heterogeneity in sex ratios. However, the fishery data, in particular all the catch and effort data, were aggregated by sex. There were limited amounts of sex-specific length measurements from some fleets (e.g. Australia, New Zealand and Chinese Taipei) and all weight composition data are aggregated by sex.

In the last six months, work has focused on extending MFCL to support a 2 sex structure in the model dynamics, with fitting to either sex-specific or sex-aggregated data as might be available. When, as for SWP swordfish, few sex-specific data are available, a range of model assumptions is necessary to allow stable estimation. During the course of attempting a 2 sex assessment for SWP swordfish this year, we realised that further developments of MFCL are required to support fitting a model with sex-specific dynamics to sex-aggregated fishery data. Because we had insufficient time to implement the required developments to MFCL, we had to revert to a sex-aggregated assessment for 2017. However, because important sexual dimorphism has been observed in terms of SWP swordfish growth and length-weight relationships, and there is evidence of spatial heterogeneity in both sex ratios and size composition (Davies *et al*. 2013), we will continue to develop MFCL so that a sex-specific assessment can be conducted in the future.

As described above, there are differences in the observed size (length and/or weight) structure of the catch among fishery sub-areas of both model regions. These spatial differences were addressed through the method and sub-area specific definitions of the fisheries incorporated in the model and the flexibility to estimate specific size-based selectivity functions for each of the main fisheries within each sub-area. Seasonal and spatial variations in catch rates of swordfish between fisheries are accounted for in the model by estimating fishery-specific catchability parameters incorporating seasonal variation. The principal source of abundance information in the model is derived from the catch and (GLM standardised) effort series for the main fisheries (Figure 4). As described in Section 3.4, there were five standardised effort series from longline fisheries available for the model fitting procedure (02_DW_1C, 04_AU_1, 07_DW_2C_pre-2001, 08_DW_2C_post-2001 and 11_EU_2). The most significant and sustained component of the catch from the model region and throughout the time period is from the distant-water longline fisheries in regions 1 and 2.

4.1. Population dynamics

4.1.1 Recruitment

Recruitment is the appearance of age-class 1 fish in the population (Fournier *et al*. 1998). Swordfish spawn in the tropical and sub-tropical latitudes (Young & Drake 2002) following seasonal migrations. As for previous assessments, recruitment to the model population was assumed to be annual and occurring in the first quarter.

Recruitment was assumed to be related to spawning potential according to the Beverton-Holt stock-recruitment relationship (SRR). “Spawning potential” in this assessment is defined as the sum over age-classes of the product of numbers-at-age, proportion mature-at-age, weight-at-age (assumed proportional to fecundity) and the sex ratio-at-age for females (assumed to be 0.5 for all age-classes). These assumptions were also used in the 2013 assessment. Deviations in estimated recruitment from the SRR attract a small penalty in the likelihood. In the tropical tuna assessments, we typically keep this penalty as small as possible so that the SRR assumptions do not overly impact the estimates of recruitment (as recommended by the 2011 bigeye assessment review). For this SWP swordfish assessment, with much less substantial data inputs compared to the tropical tuna assessments, a somewhat larger penalty, equivalent to annual recruitments having a CV of 0.5, was required in order to have stable model behaviour.

Typically, fisheries data are very uninformative about SRR parameters and it is generally accepted that the steepness parameter, which controls the shape of the curve at lower stock sizes, is not well estimated in fisheries models. As in previous assessments, we assumed a fixed value of steepness equal to 0.8 for the diagnostic case, with alternative options of 0.65 and 0.95 included in the structural uncertainty grid.

4.1.2 Initial population

The population age structure in the initial time period was assumed to be in equilibrium and determined as a function of the average total mortality during the first five years. This assumption avoids having to treat the initial age structure as independent parameters in the model, which is generally poorly determined.

4.1.3 Growth, maturity and natural mortality parameters

Parameters such as growth rates, maturity schedule, longevity and mortality are important model parameters for MFCL (Fournier *et al*. 1998). While MFCL can estimate many of these parameters, some parameters need to be fixed through time.

Two swordfish growth estimates provided alternative scenarios in the 2013 assessment. These were:

- CSIRO-developed estimates from Australian age samples (Young *et al*. 2008);
- PIFSC-developed estimates from Hawai'i age samples (DeMartini *et al*. 2007).

In each case, the mean of the male and female curve was used as fixed input, with variance on length-at-age inflated to cover both sexes. These two growth hypotheses proved to be a key area of uncertainty in the 2013 assessment.

At SC12 in 2016, new growth and maturity estimates for this stock (Farley *et al*. 2016, SC12-SA-WP12) were presented and endorsed. It resolved the differences between the two alternative growth hypotheses, and provided new results from otolith aging which are expected to better represent the growth of large fish. A sex-aggregated growth curve based on these results was used in the current assessment (Figure 9, Table 3).

SC12-SA-WP12 also presented the new results on swordfish maturity. Those results provided similar maturity-at-age regardless of the choice of growth curve based on either fin rays or otolith aging. The derived swordfish maturity schedule indicated fish attained maturity slightly slower when compared with the maturity schedule applied for the 2013 reference case (Figure 10, Table 3).

Although in the current assessment more reliable age estimation is available, natural mortality-at-age estimates are still highly uncertain. A broad range of M values are assumed in other swordfish assessments worldwide, ranging from 0.2 – 0.5 yr^{-1}. Following the previous two assessments in 2008 and 2013, a similar method (Kolody *et al*. 2008) was applied to update age-specific natural morality vectors for the current assessment. The method relies on two main biological parameters (von Bertalanffy growth curve parameters and age at maturity). In the current assessment, the updated parameters described above were used in the derivation of M vectors. In addition, derivation of mean M-at-age required the anticipated temperature of swordfish habitat. The same two values (14.57 and 22.83 °C) used in previous assessments were used here. This resulted in two M vectors (M1 and M3 in Figure 10). In addition, the method applied in the 2008 and 2013 assessments requires an assumed ratio of average adult M to average juvenile M. We tested two assumptions – adult M > juvenile M (M1 and M3 in Figure 10) and juvenile M > adult M (M2 and M4 in Figure 10). Therefore, 4 alternative vectors of natural mortality-at-age were considered in the current assessment (Table 4).

The assumptions made concerning age and growth in the MFCL model are (i) the lengths-at-age are normally distributed for each age class; (ii) the mean lengths-at-age follow a von Bertalanffy growth curve; and (iii) the standard deviations in length-at-age are a linear function of the mean length-at-age (Fournier *et al*. 1998). The probability distributions of weights-at-age are a deterministic function of the lengths-at-age and a specified weight-length relationship (Table 5).

For any specific model, it is necessary to assume the number of significant age-classes in the exploited population, with the last age-class (20 yr) being defined as a "plus group", i.e. all fish of the designated age and older.

4.1.4 Length and Weight

The parameters for the relationship between LJFL and whole weight were obtained from Davies *et al*. (2005) based upon observer samples for both sexes combined, n = 2835. The parameters for the relationship are provided in Table 3.

4.1.5 Movement

An estimate of swordfish movement rates between the two regions of the current model was developed by Evans *et al*. (2012). They estimated diffusive mixing across the boundary at 165°E (diffusion rate, D = 0.11) as the best estimate of movement between regions at this time.

Evans *et al*. also strongly recommend examining the sensitivity of this assumption, including alternative interpretations at the extremes (i.e. very high and zero mixing), in recognition that this estimate is highly uncertain (and qualitatively wrong if spawning populations really are isolated). Two other values of diffusion were examined in the sensitivity analyses (zero movement, and approximately twice the recommended value). Diffusion rates were translated into quarterly bulk transfer coefficients calculated for model input values based upon a "key" developed by Kolody and Davies (2008, SC4-SA-IP2, see Figure 16), that assumes instantaneous and complete mixing of the population within regions.

4.2. Fishery dynamics

4.2.1 Selectivity

Selectivity is fishery-specific and assumed to be time-invariant and length-based but modelled as age-based (Kleiber et al. 2017). Differences in selectivities among fisheries using the same methods (i.e. longlines) in different fishery sub-areas of the model region may be proxies for spatial structuring of the swordfish population by size. The selectivities-at-age for the longline fisheries were estimated using two forms of parameterization: cubic splines (for all fisheries except 09_DW_2S) and asymptotic forms (only 09_DW_2S). Each selectivity spline function was parameterised with three nodes allowing considerable flexibility in the functional form while minimizing the number of parameters that needed to be estimated. In contrast to the last assessment, the assumption to constrain selectivities to be non-decreasing was removed for all fisheries except for those in the southern sub-regions where larger fish were reported, in order to improve the fitting to size composition data. Common terminal selectivities at age were assumed to start from 19+ years. At the PAW in April 2017, a suggestion was made to allow selectivity time break for fishery 4 (AU_1) to account for a possible selectivity change in the fishery due to the introduction of a different hook type (circle hook). The change from J-hooks to circle hooks occurred rapidly around 2008, therefore 2008 was chosen to be a break point for selectivity, and the sensitivity of model estimates to this assumption examined (Section 5.2).

4.2.2 Catchability

Catchability was assumed to be constant over time for those fisheries where the model was being fitted to a standardised CPUE time series. This was because the CPUE is considered informative of temporal trends in population relative abundance. In this case fishing effort has been standardised to account for systematic trends in catchability associated with temporal and spatial changes in the distribution of fishing effort and changes in gear configuration. While it is considered unlikely that such a statistical approach can account fully for systematic variation in catchability over time, the resulting standardised effort series represent the best available indices of relative abundance for the stock.

Catchability for all other fisheries that lack standardised effort, or having CPUE but not being fitted by the model, was allowed to vary slowly over time (akin to a random walk) using a structural time-series approach. Random walk steps were taken biennially, and the deviations constrained by a prior distribution of mean zero and a variance equivalent to a CV of 0.1. Seasonal variation in catchability was also allowed to explain the strong seasonal variability in CPUE for most of the fisheries.

4.2.3 Effort deviations

Effort deviations, constrained by prior distributions of zero mean and a specified variance, were used to model the random variation in the effort–fishing mortality relationship. For all fisheries, a penalty weight scaled by the square root of the effort was applied to the effort deviations, to reflect the amount of effort and its uncertainty. For the fisheries to which the model was fitted to standardised effort, the time-variant precision estimates were applied multiplicatively to the penalties, i.e. as temporal effort deviate penalties that are higher for more precise effort indices.

4.2.4 Likelihood components

There are three data components that contribute to the log-likelihood function − the total catch data, the weight-frequency data and the length-frequency data. Whereas tagging data informed the assumptions for movement, no tagging data were included in the fitting of this assessment model. The observed total catch data are assumed to be unbiased and relatively precise, with the standard deviation (SD) of residuals on the log scale being 0.002.

The probability distributions for the length- and weight-frequency proportions are assumed to be approximated by robust normal distributions, with the variance determined by the effective sample size (ESS) and the observed frequency proportion. It is necessary to down-scale the ESS in relation to observed or measured sample numbers to recognise (i) that size-frequency samples are not truly random samples; and (ii) that the variance of the samples is affected by a range of process error that is not accounted for in the model. In the diagnostic case, the ESS for the length- and weight-frequency samples for all fisheries except 03_DW_1S and 09_DW_2S were assumed to be 0.05 times the observed sample size with a maximum ESS of 50. For 03_DW_1S and 09_DW_2S, we assumed the ESS was 0.1 times the observed sample size, with a maximum ESS of 100. Greater weight was given to these fisheries because they catch the largest swordfish and it is important for the model estimates to be consistent with this observation.

4.3. Parameter estimation and uncertainty

The parameters of the model were estimated by maximising the log-likelihoods of the data plus the log of the probability density functions of the priors and penalties specified in the model. The maximisation was performed by an efficient optimisation using exact derivatives with respect to the model parameters. Estimation was conducted in a series of phases, the first of which used arbitrary starting values for most parameters. Convergence was judged based on the absolute value of the gradient of the negative log likelihood function with respect to the parameters. If all gradients were less than 0.001, the model was considered to have converged.

The Hessian matrix computed at the mode of the posterior distribution was used to obtain estimates of the covariance matrix, which was used in combination with the Delta method to compute approximate asymptotic confidence intervals for estimated and derived parameters of interest.

Convergence of model runs were further assessed by: i) likelihood profiling in terms of average total biomass size; and ii) multiple runs starting from parameter estimates with slight perturbation ("jittering").

4.4. Stock assessment interpretation methods

Several ancillary analyses are conducted in order to interpret the results of the model for stock assessment purposes. The methods involved are summarized below and the details can be found in Kleiber *et al*. (2017).

4.4.1 Yield analysis

The yield analysis consists of computing equilibrium catch (or yield) and biomass, conditional on a specified basal level of age-specific fishing mortality (F_a) for the entire model domain, a series of fishing mortality multipliers (*fmult*), the natural mortality (M), the mean weight-at-age (w_a) and the SRR parameters. All of these parameters, apart from *fmult*, which is arbitrarily specified over a range of 0–50

in increments of 0.1, are available from the parameter estimates of the model. The maximum yield with respect to *fmult* can easily be determined and is equivalent to the MSY. Similarly the total and adult biomass at MSY can also be determined. The equilibrium yield estimate includes a log-normal bias correction for the assumed distribution of recruitment deviates about the stock-recruitment relationship.

4.4.2 Depletion and fishery impact

Many assessments estimate the ratio of recent to initial biomass as an index of fishery depletion. The problem with this approach is that recruitment may vary considerably throughout the time series, and if either the initial or recent biomass estimates (or both) are "non-representative" because of recruitment variability, then the ratio may not measure fishery depletion, but simply reflect recruitment variability.

We approach this problem by computing biomass time series (at the region level) using the estimated model parameters including the annual estimated recruitments (excluding a spawner stock – recruitment effect), but assuming that fishing mortality was zero. Because both the *real* biomass B_t and the *unexploited* biomass $B_{t_{F=0}}$ incorporate recruitment variability, their ratio at each time step of the analysis $B_t/B_{t_{F=0}}$ can be interpreted as an index of fishery depletion. In the two tuna assessments presented this year (2017), for the calculation of unexploited biomass an adjustment of recruitment was performed to acknowledge the possibility of a reduction in recruitment in an exploited population through the stock recruitment relationship. Due to a technical difficulty specific to the swordfish stock assessment, this adjustment was not applied here and model-estimated recruitments were used instead. This limitation was also applicable to the calculation of unfished spawning potential described in the next section.

4.4.3 Reference points

Historically, MSY-based reference points, in particular the fishing mortality at MSY (F_{MSY}) and the spawning biomass (or spawning potential) at MSY (SB_{MSY}), have been commonly used to define "overfishing" and "overfished" conditions, respectively. Because of the uncertainty in estimating MSY-based reference points, and in particular their sensitivity to assumptions regarding the SRR steepness parameter, WCPFC has decided to apply spawning-potential-depletion-based reference points for the tropical tunas and South Pacific albacore. In particular, $20\%SB_{F=0}$ has been decided as a limit reference point (LRP) for these stocks, where $SB_{F=0}$ is the estimated spawning potential that would have occurred in the absence of fishing.

WCPFC is yet to decide on a LRP for SWP swordfish. Therefore, in this assessment, we report stock status in relation to both MSY-based and depletion-based reference points – $SB_{recent}/SB_{F=0}$, $SB_{latest}/SB_{F=0}$, SB_{recent}/SB_{MSY}, SB_{latest}/SB_{MSY}, and F_{recent}/F_{MSY}, where "recent" refers to the average of 2011-2014 and "latest" to 2015.

4.4.4 Majuro and Kobe plots

For the standard yield analysis (Section 4.4.1), the fishing mortality-at-age, F_a, is determined as the average over some recent period of time (2011–2014 herein). In addition to this approach the MSY-based reference points (F_t/F_{MSY}, and SB_t/SB_{MSY}) and the depletion-based reference point ($SB_t/SB_{F=0[t]}$) were also computed using the average annual F_a from each year included in the model (1952–2014, with no value calculated for the terminal year) by repeating the yield analysis for each year in turn. This enabled temporal trends in the reference point variables to be estimated taking account of the differences in MSY levels under varying historical patterns of age-specific exploitation. This analysis is presented in the form of dynamic Kobe plots and "Majuro plots", which have been presented for all WCPO stock assessments in recent years.

5. Model runs

5.1. Developments from the last assessment

We followed the standard approach for transitioning from the reference case of the last assessment to a "diagnostic case", denoted *DiagCase,* for the current assessment in a number of steps so as to highlight

the impact of the various individual changes incorporated in this assessment. We stress that the *DiagCase* is just one of the models that comprise the overall suite of models that we use to characterise stock status and its uncertainty. The *DiagCase* is used simply to demonstrate the step-wise model development described in this section and as a reference to evaluate the impact of various one-off sensitivities investigated in the assessment. The following steps were undertaken in the development of the 2017 *DiagCase* model:

Step 0: *Ref2013* – This is the "reference case" model for the 2013 swordfish assessment as reported in Davies *et al.* (2013).

Step 1: *Ref2013NewExe* – This model re-runs the 2013 reference case with the latest MFCL executable, which has undergone many changes in the past 4 years. In addition, in this step we also modified the structure of the length-frequency data, to omit the two smallest length bins (10-20 cm and 20-30 cm). Swordfish are highly unlikely to be caught at this size by longline and this change has been made for model stability purposes. We verified that making this change alone resulted in no perceptible change to the assessment results.

Step 2: *Update2015* – In this model, we updated all input data (catch, effort, size) for four additional years of data (2012-2015, inclusive) since the 2013 assessment.

Step 3: *NewSel* – In this model, we made small changes to the constraints on selectivity, as described in section 4.2.1.

Step 4: *NewGrowth* – In this step, we updated the model from step 3 with a composite growth curve (i.e. representing both males and females) and maturity schedule based on otolith and reproductive maturity data presented by Farley *et al.* (2016). Adjustments were also made to the assumed natural mortality-at-age resulting from the changes to growth and maturity.

Step 5: *DiagCase* – Consistent with other assessments, we adopted a size weighting for the *DiagCase* whereby the observed sample size is divided by a factor of 20 and a maximum applied such that the maximum effective sample sizes (ESS) for both length and weight samples is set at 50 fish. In the 2013 assessment, the ESS was set to 10 fish, but we considered that this level of ESS was insufficient to adequately fit the size data.

5.2. Sensitivity analyses

Analyses were undertaken to test the sensitivity of outputs from the *DiagCase* model to changes in: diffusion rate, natural mortality, steepness, size data weighting, the break point year applied for selectivity time blocks applied to 04_AU_1, potential for the existence of a separate sub-stock in the northern part of assessment region, as requested by the 2017 PAW, and effects of the exclusion of Australian length-composition data.

5.2.1 Diffusion rate

In addition to providing a recommended diffusion rate of 0.11, Evans *et al*. (2012) also recommended exploring model sensitivity to this rate due to the uncertainty in the estimate. We used 0.11 in the *DiagCase*, and tested alternatives of 0.0 and 0.25 in the sensitivities. The region-specific quarterly block transfer coefficients were obtained using the method described in Section 4.1.5 using block sizes of 25° and 65° for regions 1 and 2, respectively.

5.2.2 Natural mortality

Four options for the natural-mortality-at-age schedule were explored (section 4.1.3). The M1 vector was used for the *DiagCase*.

5.2.3 Steepness

A fixed value of 0.8 was assumed for the *DiagCase*, and fixed values of 0.65 and 0.95 were tested in the sensitivities.

5.2.4 Size data weighting

Moderate relative weight was assumed for length- and weight-frequency data for the *DiagCase*, as described in Section 4.2.4. A sensitivity was conducted in which the ESS's for all fisheries were reduced to 50% of those used in the *DiagCase*, i.e. to 0.025 of the observed sample sizes (max. 25) for all fisheries except 03_DW_1S and 09_DW_2S, and to 0.05 of the observed sample sizes (max. 50) for those fisheries.

The *DiagCase* and the eight sensitivity runs described above were taken as the key model runs for examining the effects of the primary sources of uncertainty on management reference points in the current assessment (Table 6). In addition to these runs, additional sensitivities, as described below, were also considered in sensitivity runs.

5.2.5 Selectivity time blocks for fishery 4 (04_AU_1)

The *DiagCase* applied selectivity time blocks to the 04_AU_1 fishery to take into account possible selectivity changes due to the introduction of a new hook type regulation around that time. The *DiagCase* model set 2008 as the break point of selectivity time blocks. To explore the uncertainty in the timing of regulation adoption, four alternative breakpoints for the selectivity time blocks (2006, 2007, 2009 and 2010) were applied. In addition, a sensitivity was made removing selectivity time blocks from 04_AUS_1 completely, therefore assuming time-invariant selectivity.

5.2.6 Potential effects of different sub-stock in northern sub region

The *DiagCase* implicitly assumes that swordfish in the assessment region belong to the same (sub) stock. Recent genetic studies suggested multiple swordfish sub stocks in the Pacific Ocean, but to date, there are no quantitative results from genetic studies or tagging data to delineate stock boundaries. However, the potential for the northern part of the current assessment region to represent a different sub-stock of swordfish was raised at the PAW. Therefore, a sensitivity run was performed which, for simplicity, removed data in the two northern sub-regions. However, this sensitivity run has limitations. Although data of fisheries that only occur in the two northern sub-regions (01_DW_1N, 06_DW_2N and 12_Other_2N) were removed from the analysis, the remaining fisheries that operate across the region partly include data from the northern sub-regions.

5.2.7 Removal of the 04_AU_1 length-frequency data

The length data provided for the 04_AU_1 fishery was notably different from that available for the 2013 assessment, and modelling identified a conflict between the length and weight data available for this fishery. To examine the influence of this data set on the outputs of the model, a sensitivity run was performed excluding the length-frequency data from this fishery.

5.3. Structural uncertainty

Stock assessments of pelagic species in the WCPO in recent years have utilised an approach to assess the structural uncertainty in the assessment model by running a "grid" of models to explore the interactions among selected "axes" of uncertainty. The grid contains all combinations of two or more parameter settings or assumptions for each uncertainty axis. The axes are generally selected from those factors explored in the one-off sensitivities with the aim of providing an approximate understanding of variability in model estimates due to assumptions in model structure not accounted for by statistical uncertainty estimated in a single model run, or over a set of one-off sensitivities.

The structural uncertainty grid for the 2017 swordfish assessment was constructed from 4 axes denoted as key sensitivity runs in section 5.2 and Table 6 – steepness (3 settings), size data weighting (2), diffusion rate (3) and natural mortality (4). The final grid thus consisted of 72 models.

6. Results

6.1. Consequences of key model developments

The transition from the 2013 reference case model to the current *DiagCase* is undertaken in five steps, and the transition is displayed in terms of two key stock assessment outputs, the spawning potential (Figure 11a) and spawning potential depletion (Figure 11b). The following step-wise changes are noted:

Step 1: *Ref2013NewExe* – The application of the new MFCL executable to the 2013 assessment, including the re-definition of the length bins, produced an identical result to *Ref2013* (the spawning potential and depletion trajectories in Figure 11a and Figure 11b are overlaid).

Step 2: *Update2015* – In this model, we updated all input data (catch, effort, size) for four additional years of data (2012-2015, inclusive) since the 2013 assessment. This resulted in the spawning potential being substantially down-scaled (Figure 11. Stepwise changes in spawning potential (a, top) and fishing depletion (b, bottom) from the 2014 reference case model through to the 2017 diagnostic case model. For details of each step, see section 6.1.a) but the depletion plots (Figure 11b) show very similar trends and overall scale.

Step 3: *NewSel* – The selectivity changes made at this step included removing some constraints to allow for the possibility of a dome-shaped pattern in some fisheries (while retaining monotonically increasing selectivity with age for one fishery in each model region), grouping selectivity for some fisheries where this was warranted, introducing a time-series break in selectivity for fishery 04_AU_1 and setting selectivity for the youngest age classes to zero in some fisheries to enhance the fit to the size data. These changes, which enhanced the fit to the size data overall, resulted in a steeper decline in spawning potential after 2000 (Figure 11a) and resulted in higher depletion over the same period (Figure 11b).

Step 4: *NewGrowth* – The updated growth and maturity parameters resulted in some relatively small downward changes in spawning potential and depletion.

Step 5: *DiagCase* – In this final step we up-weighted the size data compared to that used in the previous assessment such that the maximum effective sample size for all length and weight samples is 50 fish. This change resulted in moderately higher spawning potential and depletion ratio relative to Step 4.

6.2. Model fit for the diagnostic case model

6.2.1 Catch data

As is always the case for MFCL models, we apply a large penalty to deviations of predicted and observed catches. Therefore, the fit to the catch data is effectively perfect (Figure 12).

6.2.2 Standardised CPUE

The observed and predicted CPUE for those fisheries for which standardised effort was used and constant catchability over years was assumed appears very good (Figure 13). This is also confirmed by the plots of effort deviations for these fisheries (Figure 14) which generally show an even scatter about zero. The effort deviations for the fisheries where time-series changes in catchability were allowed (Figure 15) also show a fairly even distribution about zero, indicating that most of the persistent changes have been captured by the catchability deviation estimates.

6.2.3 Size-frequency data

The overall fits to the length (Figure 5) and weight (Figure 6) frequency data are generally adequate, given the small sample sizes for many fisheries and the conservative weighting applied in the likelihood. Lack of fit to length data is noteworthy for fisheries 02_DW_1C, 07_DW_2C_pre-2001 and 08_DW_2C_post-2001. This is also seen in the time-series plots of the fit to available length data (Figure 6). However, those

for key fisheries appear reasonable, although very recent increases in mean size in 04_AU_1 and 05_Other_1 data are not fully captured (Figure 7). Fits to the weight data are in general good, although recent temporal patterns in the 04_AU_1 data are again not fully captured (Figure 8).

6.3. Model parameter estimates (diagnostic case)

6.3.1 Catchability

The annual catchability (although allowed to vary seasonally) was held constant for the fisheries having standardised CPUE indices to which the model was fitted (02_DW_1C, 04_AU_1, 07_DW_2C_pre-2001, 08_DW_2C_post-2001, 11_EU_2). Catchability was allowed to vary temporally for all other fisheries, with several, e.g. 01_DW_1N, 03_DW_1S, 06_DW_2N, 10_NZ_2 and 12_Other_2N showing strongly increasing trends, in contrast with 05_Other_1, which shows a strongly decreasing trend (Figure 16). These trends generally reflect the nature of the CPUE data available for these fisheries, which may reflect changes in fleet composition, targeting practices or other factors that impact CPUE. Seasonal variability in catchability was apparent in most fisheries, particular those located in the more southerly parts of the region.

6.3.2 Selectivity

Estimated selectivity functions by age-class are displayed in Figure 17. Only two fisheries were constrained to have a monotonically increasing selectivity with age, 03_DW_1S and 09_DW_2S. These fisheries were selected as they capture the largest swordfish in each of the two model regions. All other fisheries were allowed to have a declining right limb in the selectivity curve, if warranted by the data. For most of these fisheries, selectivity was dome-shaped with the maximum selectivity occurring most frequently at age-classes 4 or 5. The exceptions were fisheries 02_DW_1C, for which selectivity peaked at age-class 2 and fisheries 07_DW_2C_pre-2001 and 08_DW_2C_post-2001, where selectivity peaked at age-class 6.

6.3.3 Movement

Movement was fixed in the swordfish model, with two alternative movement patterns ranging from effectively separate stocks in regions 1 and 2 (no movement) to a rapidly mixing population over the entire model domain, explored in the sensitivity grid. These results are described in a later section.

6.3.4 Growth

For this assessment, we fixed growth according to a von Bertalanffy model fitted to length-at-age observations obtained from analyses of otoliths (Farley *et al.* 2016). The growth curve assumed in the assessment is shown in Figure 9, in comparison to that used in the 2013 assessment. The current curve has a similar length at age 20, but a slightly reduced growth coefficient (k) in comparison to that used in 2013.

6.4. Stock assessment results

Some of the results in this section use the *DiagCase* for convenience to show time-series trends or regional comparisons. However, for the critical stock status indicators in the following section, we present the results from the complete structural uncertainty grid so as to fully characterise the uncertainty in the main metrics.

6.4.1 Recruitment

Overall recruitment (Figure 18) shows a large inter-annual variability, but no persistent trend over time. Approximately two-thirds of the recruitment on average is estimated to occur in region 2 (the east) and one-third in region 1. Recruitment estimates have broad confidence intervals indicating substantial uncertainty (Figure 19), with a slight reduction in uncertainty following the mid-1990s, when more size data became available. Annual recruitment estimates are widely scattered about the

SRR **(**

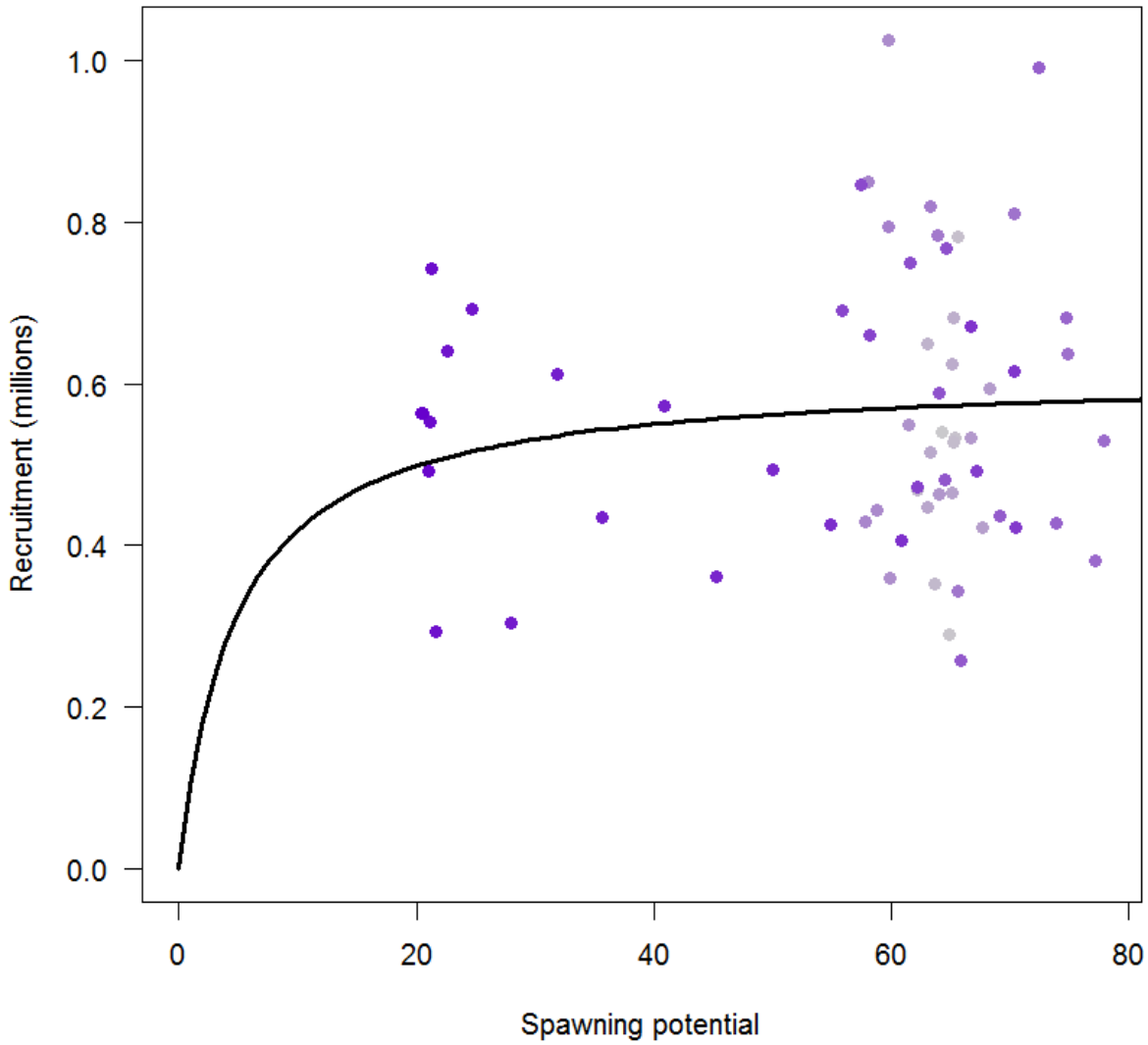

Figure **20**).

6.4.2 Biomass and spawning potential

The annual estimates of spawning potential by region are shown in Figure 21 with the distribution between the two regions being similar to that for recruitment. The main feature of the trend is a strong decline initiated in the late 1990s, and a stabilisation since about 2010. This decline appears to be largely an impact of fishing. There is a high level of uncertainty associated with the spawning potential estimates (Figure 22), which decreases slightly from the late 1990s. This reduced uncertainty may be a result of more size data only being available since the 1990s for the key fisheries. Similar regional patterns and levels of uncertainty are evident for total biomass (Figure 23 and Figure 24).

6.4.3 Fishing mortality

Fishing mortality (exploitation) rates for juvenile (Ages 1-3), maturing (ages 4-6) and adult (ages 7+) swordfish are estimated to have increased over time, notably in the mid 1990s (Figure 25), following the significant increases in catches at that time (see Figure 1). Since that time, fishing mortality rates for maturing age swordfish in particular increased notably until the early 2000s in region 1, and to a lesser extent for the other age groupings, before falling slightly in the mid 2000s to intermediate levels. In region

2, a similar increase is seen in maturing aged swordfish, peaking in the mid 2000s and fluctuating at a relatively high level to the end of the time series. A peak in adult mortality is seen in the early 1970s, which is driven by catches of fishery 9. The reason for this requires further investigation. Combined, the overall fishing mortality rate has been relatively high on the maturing age group, at around four times that on juveniles and adults.

Decadal changes in population numbers-at-age and F-at-age are shown in Figure 26. As the fishery developed through the 1980s, fishing mortality was low and focused on the oldest age-classes. In the last three decades, fishing mortality on sub-adult swordfish aged 3-7 years greatly increased, with a resulting "downstream" depletion of older age-classes.

6.5. Multimodel inference –sensitivity analyses, and structural uncertainty

The following sections refer mainly the key management quantities that are defined in Table 7.

6.5.1 One-off changes

Comparisons of the spawning potential and depletion trajectories for the *DiagCase* and one-off sensitivity runs are provided in Figure 27 and Figure 28. The key reference points for these runs are compared in Table 8 and Table 10, and the likelihood components are provided in Table 9 and Table 11.

Size data weighting [*W2*]

Down-weighting the size data within the model scaled the population upwards moderately, although the relative changes in the time-series were very similar (Figure 27a). The estimates of fisheries depletion are also scaled up slightly compared with the *DiagCase*, with estimates of depletion in the most recent period over 5 percentage points higher (Figure 28a). Estimates of F_{recent}/F_{MSY} and $SB_{latest}/SB_{F=0}$ for the *DiagCase* model run and down-weighted size data run were 0.89 vs 0.72, and 0.31 vs 0.39, respectively (Table 8).

Steepness [*h0.65, h0.95*]

The alternative assumptions of low (h0.65) and high (h0.95) steepness had little impact on spawning potential (Figure 27a) and depletion (Figure 28a and Table 8). However, stronger impacts are seen for MSY-based reference points (e.g. F_{recent}/F_{MSY} varies from 1.17 for h0.65 to 0.59 for h0.95; Table 8).

Natural Mortality [*M2*, *M3*, *M4*]

Assuming alternative functional forms of M-at-age led to very different spawning potential estimates. The M2 assumption led to a small down-scaling of spawning potential, but with a comparable temporal pattern to the *DiagCase* model run (Figure 27a). The trajectory of spawning potential depletion was also comparable to that of the *DiagCase* model run (Figure 28a). Both M3 and M4 scaled spawning potential upwards notably, but again the temporal pattern was comparable to the *DiagCase* model run results. Trajectories of depletion were comparable to the *DiagCase* model run, but diverged in the most recent period in particular to end around 5 percentage points lower. A similar pattern was seen for the MSY- and depletion-based reference points. F_{recent}/F_{MSY} for the *DiagCase* and M2 models were 0.89 and 0.91 respectively, while estimates from the M3 and M4 models were higher at 1.06 and 1.05, respectively (Table 8) while the depletion-based reference points tended to be less sensitive to assumed M-at-age, $SB_{latest}/SB_{F=0}$ ranging from 0.27 to 0.31 (Table 8).

Diffusion rate [*Mv0*, *Mv0.25*]

The assumption of diffusion rate had a relatively small effect on the spawning potential. Assuming no diffusion between regions 1 and 2 changed the historical temporal pattern of biomass estimates, being slightly more stable than the *DiagCase* model run, while assuming higher rates of diffusion between the two model regions led to a trajectory comparable to that of the *DiagCase* model run. Trajectories of depletion were comparable to the *DiagCase* model run estimates for much of the time period, but diverged in the most recent period. Assuming no diffusion between regions resulted in a much reduced level of terminal depletion, being 10 percentage points higher than in the *DiagCase* model run. Assuming a higher rate of diffusion resulted in a depletion trajectory comparable to the *DiagCase* run. MSY- and depletion-

based reference points for the *DiagCase*, Mv0 and Mv0.25 assumptions were F_{recent}/F_{MSY} values of 0.89, 0.65 and 0.91, and $SB_{latest}/SB_{F=0}$ values of 0.31, 0.48 and 0.31, respectively (Table 8).

Removal of length data from the Australian Fishery [-AUS1L]
Removal of the Australian length composition data from the model scaled spawning potential downward, and resulted in different temporal patterns, particularly within the historical period (Figure 27b). In the most recent period, where considerable weight data were also available, the biomass trajectory was comparable to that of the *DiagCase* model run estimates. In terms of depletion trajectories, while the stock was estimated to be slightly more depleted over much of the historical period, recent estimates were comparable to the *DiagCase* model run estimates. MSY- and depletion-based reference points for the *DiagCase* and this run were F_{recent}/F_{MSY} values of 0.89 vs 0.93, and $SB_{latest}/SB_{F=0}$ values of 0.31 vs 0.32, respectively (Table 10).

Removal of the northern region from the model [-1Nand2N]
Removal of the northern region (model sub-regions 1N and 2N) from the model scaled down the estimate of spawning potential by around 50%, but resulted in comparable temporal dynamics to that from the *DiagCase* model estimates (Figure 27b). The trajectory of depletion indicated slightly higher depletion in the historical period when fishing was more concentrated in the 1C and 2C sub-regions. The depletion trajectory within the more recent time period was comparable to that from the *DiagCase* model run, but recovered above that in the most recent period (where a large increase in fishing is notable within the 2N region). Given the smaller model region, MSY and SB_{MSY} estimates were reduced, the F_{recent}/F_{MSY} estimate lower, and F_{mult} larger. However, the SB-based reference point values were relatively robust to the change. MSY- and depletion-based reference points for the *DiagCase* and this run were F_{recent}/F_{MSY} values of 0.89 vs 0.79, and $SB_{latest}/SB_{F=0}$ values of 0.31 vs 0.32, respectively.

Removal of selectivity change assumption for fishery 4 [-*Selblock*]
Removal of the selectivity time blocks for fishery 4 (the AU1 fleet) scaled the estimates of both spawning potential and depletion down slightly, but the temporal trend was maintained in both cases (Figure 27b and Figure 28b). MSY- and depletion-based reference points for the *DiagCase* and this run were F_{recent}/F_{MSY} values of 0.89 vs 0.92, and $SB_{latest}/SB_{F=0}$ values of 0.31 vs 0.31, respectively (Table 10).

Alternative years in which the selectivity change was assumed in fishery 4 [*Selblk06*, *Selblk07*, *Selblk09*, *Selblk10*]
Shifting the time at which the selectivity time block for fishery 4 was assumed (2006 = *Selblk06*, etc.) tended to scale down biomass estimates slightly in the historical period, and led to slightly different temporal dynamics. Recent dynamics, during the period in which considerable size data were available led to more comparable trends to that of the *DiagCase* model run, with terminal reproductive potential slightly below the *DiagCase* model run (Figure 27b). Temporal depletion trends were comparable to the *DiagCase* model run, but indicated slightly greater terminal depletion (Figure 28b). MSY- and depletion-based reference points for the *DiagCase* compared to *Selblk06*, *Selblk07*, etc. were F_{recent}/F_{MSY} values of 0.89 vs 0.93, 0.94, 0.96 and 0.96, and $SB_{latest}/SB_{F=0}$ values of 0.31 vs 0.31, 0.31, 0.30, and 0.30, respectively.

6.5.2 Structural uncertainty analysis

The results of the structural uncertainty analysis are summarised in several forms – time-series plots of fisheries depletion for all models in the grid (Figure 29), boxplots of F_{recent}/F_{MSY} and $SB_{latest}/SB_{F=0}$ for the different levels of each of the four axes of uncertainty (Figure 30 and Figure 31), Majuro plots showing the estimates of F_{recent}/F_{MSY} and $SB_{latest}/SB_{F=0}$ (and $SB_{recent}/SB_{F=0}$ for comparison) across all models in the grid (Figure 32), Kobe plots showing the estimates of F_{recent}/F_{MSY} and SB_{latest}/SB_{MSY} (and SB_{recent}/SB_{MSY} for comparison) across all models in the grid (Figure 33), and averages and quantiles across the full grid of 72 model runs for all the reference points and other quantities of interest (Table 12) that have also been presented for the *DiagCase* model and one-off sensitivity models.

Many of the results of the structural uncertainty analysis are consistent with the results of previous assessments of swordfish and tuna stocks in the WCPO that used the same uncertainty axes. The general features of the structural uncertainty analysis are:

- The grid contains a range of models that estimate a notable range of estimates of stock status, trends in abundance, and reference points. However, with a key axis of uncertainty – growth – being better understood in this assessment compared to that in 2013, the range of estimates is reduced. In all runs, SB_{latest} was above $20\%SB_{F=0}$. However, in 23 of the 72 runs in the grid (32%), $F_{recent} > F_{MSY}$, and for 18 runs (11%), $SB_{latest} < SB_{MSY}$. These runs were generally those with a steepness of 0.65.
- The most influential axis with respect to model output depended upon the management quantity of interest.
- Depletion was notably influenced by assumptions on diffusion (movement rate) between the two model regions, with increasing levels of movement implying notably more pessimistic results in terms of depletion, and increasing levels of fishing mortality.
- Fishing mortality relative to F_{MSY} was notably influenced by the assumed value of SRR steepness, with the expected trend of increasing values of steepness leading to more optimistic results and reduced variability in estimates of F_{recent}/F_{MSY}. As expected, depletion estimates were much more robust to the assumed steepness value, and changing that assumption had little impact on the level of variability in estimates.
- Down-weighting the influence of the size data tended to lead to more optimistic results in terms of depletion, and led to reduced variability in the depletion estimates, and had a relatively small impact on the level or variability in F_{recent}/F_{MSY} estimates.
- Stock status estimates were relatively robust to the assumptions on natural mortality within the model. Stock status estimates were slightly more pessimistic where the forms M3 and M4 were assumed, while F_{recent}/F_{MSY} estimates were also more variable for those runs.

6.6. Further analyses of stock status

There are several ancillary analyses related to stock status that are typically undertaken on the reference/*DiagCase* model (dynamic Kobe/Majuro analyses, fisheries impacts analyses, etc.). The shift towards relying more on multi-model inference for the 2017 assessment makes it more difficult to present these results over a large number of model runs. In this section, we rely heavily on both fisheries impact analyses from the *DiagCase* and the tabular results of the structural uncertainty grid (Table 12).

6.6.1 Fishery impacts for example models

We measure fishery impact at each time step as the ratio of the estimated spawning potential relative to the spawning potential that would have occurred in the historical absence of fishing. This is a useful quantity to monitor, as it can be computed both at the region level and for the WCPO as a whole. This information is plotted in two ways, firstly as the fished and unfished spawning potential trajectories (Figure 34), and secondly as the depletion ratios themselves (Figure 35).

The *DiagCase* model estimated that declines in spawning potential have occurred in both regions (Figure 34), and that the entire fishery has had a substantial impact on the levels of spawning potential, with current levels being around 30% of what they would have been in the absence of fishing (i.e. a decline of 70%). Key depletion trends are seen from the late 1990s when increases in catch were seen in both regions, although the rate and extent of decline differs among the two regions. Overall depletion in region 1 is estimated to be to around 40% of that in the absence of fishing, and that in region 2 to around 25% (Figure 35).

It is also possible to attribute the fishery impact with respect to depletion levels to specific fishery components, in order to estimate which types of fishing activity have the most impact on spawning

potential. As all key fisheries within the swordfish model are longline fisheries, we plot the impact of these longline fisheries grouped by region (overall impact) and by specific groups of longline fisheries within each region (Figure 36).

Across both model regions, the greatest impact is estimated to have resulted from the distant-water fleets operating in the northern model regions, an impact that increased notably from around 2000 (Figure 36). This is particularly notable in region 2. Impacts of the domestic fleets (e.g. Australia, NZ, Pacific Islands) also increased from the early 1990s to mid-2000s, then appeared to stabilise somewhat. That impact appears primarily in region 1, but also affects region 2 due to movement, in concert with domestic fisheries in that region. Finally, the impact of the distant water-fleets operating in the central region (1C and 2C, noting the latter includes EU fleets) was relatively high prior to the mid 1990s, and increased again in the mid 2000s. The impact from distant-water fleets in the southern sub-region is estimated to have been negligible.

6.6.2 Yield analysis and equilibrium estimates across the grid

The yield analyses conducted in this assessment incorporate the SRR into the equilibrium biomass and yield computations. Importantly, in the *DiagCase* model, the steepness of the SRR was fixed at 0.8 so only the scaling parameter was estimated. Other models in the one-off sensitivity analyses and structural uncertainty analyses assumed steepness values of 0.65 and 0.95.

Across the structural uncertainty grid the equilibrium virgin spawning potential in the absence of fishing ($SB_{F=0}$) was estimated to be between 61,997 and 100,691 mt (Table 12), and the spawning potential that would support the MSY (SB_{MSY}) was estimated to be between 7,251 and 30,400 mt. The ratio of SB_{MSY} to SB_0 was estimated to be between 0.13 and 0.27, while the ratio of SB_{MSY} to $SB_{F=0}$ was estimated to be between 0.12 and 0.30.

A plot of the yield as a function of fishing effort relative to the current effort is shown in Figure 37 for the *DiagCase* model run. Yield is maximized at F_{mult} = 1.12 for an MSY of 8,335 mt per annum for the *DiagCase*, with the grid ranges being 0.68 to 2.38 and 5,898 to 11,380 mt, respectively. Note that the MSY range encompasses the estimates of the average annual catch from the model region in the recent period (C_{latest}).

6.6.3 Dynamic Majuro plots and comparisons with Limit Reference Points

The section summarising the structural uncertainty grid (Section 6.5.2) presents terminal estimates of stock status in the form of Majuro and Kobe plots. Further analyses can estimate the time-series of stock status in the form of Majuro and Kobe plots, the methods of which are presented in Section 4.4.4. The large number of model runs in the structural uncertainty grid precludes undertaking and presenting this process for all runs, however an example from the *DiagCase* model run is presented in Figure 38.

At the start of the assessment period, stock status estimates were close to an $SB/SB_{F=0}$ of one and an F/F_{MSY} approaching zero, but each progressively shifted towards the overfishing and overfished definitions over the remaining period. The *DiagCase* model never reaches $20\%SB_{F=0}$, but reached an F/F_{msy} of one in 2008, and has fluctuated just below that level in recent years. The equivalent dynamic Kobe plot is displayed in Figure 38.

7. Discussion and conclusions

7.1. Development of a sex-specific MULTIFAN-CL assessment model

One of the five key recommendations arising from the 2013 stock assessment for SWP swordfish was the use of a sex-disaggregated stock assessment model, to better account for sexual dimorphism and spatial heterogeneity in sex ratios (Davies *et al*. 2013).

For the 2017 assessment, considerable effort was expended in further developing this functionality within MFCL (Davies *et al*. 2017). This should allow a reliable assessment with explicitly sex-structured population dynamics, if all fishery data including catch and effort, and size samples were reported with sex identification. However, unfortunately catch and effort data were reported in sex-aggregated form, while only some members were able to supply length measurements with sex identified. For this assessment, extensive efforts were made to conduct the assessment using the updated MFCL with an explicitly modelled sex-structured population, while allowing most of fishery data to be sex-aggregated. Several further updates were made in the last six months as reported in Davies *et al*. (2017) to enable these analyses. Nevertheless some of the important features necessary to conduct an explicitly sex-disaggregated swordfish assessment on the data available remain to be finalised. Therefore it was decided that sex-aggregated swordfish population dynamics be modelled for the SWP swordfish assessment in 2017.

7.2. Changes to the previous assessment

The 2017 swordfish assessment introduces a number of changes from the 2013 assessment that have had an influence on the resulting estimates of stock status. Four additional years of data (catch, effort, and size frequencies) were included within the assessment. Within this period there has been some evidence of a stabilisation, and in some cases an increase, in the standardised CPUE indices in the terminal years of the model. This has influenced the stabilisation of biomass in the most recent period.

Other changes made to the model included implementing minor developments to MFCL that have become available since the 2013 assessment. These included developments in the modelling of selectivities, which enhanced the fit to the size data overall, but resulted in a steeper decline in spawning potential, the use of improved growth and maturity estimates for this species, and weighting of the size data within the model consistent with that used for tuna assessments.

7.3. MFCL developments and other modelling considerations

The continued development of MFCL in the recent period has provided a number of new features that have been examined within this assessment, and have improved model estimates. However, further developments specific for swordfish have been noted throughout the report, in particular:

- ***Further development of the sex-disaggregated model***. To better account for sexual dimorphism of swordfish and the spatial heterogeneity in sex ratios, a modelling approach that includes sexual structure is still recommended. Recent developments in MFCL towards this capability will assist in this and with suitable testing should be applied for the next swordfish assessment.
- ***Selectivity***. To conduct a stock assessment with explicitly sex structure population dynamics, modelling selectivity at length rather than age is preferable. The current MFCL development already allows this to be incorporated, although that application could be expanded. The next swordfish assessment should fully explore this feature within the explicitly sex dis-aggregated population dynamics model. As already recommended in the last assessment report fishing mortality may be better estimated via selectivity functions in respect of length rather than age.
- ***Functionality to allow size-frequency data with heterogeneous resolution***. Currently MFCL assumes a uniform resolution (bin size) in size composition data across the whole fishery. This required length data to be coerced into a common length bin (10cm) for the 2017 swordfish assessment, due to the data provided. This is quite a coarse resolution compared to other assessments. Development of functionality to allow the size bin to vary between fisheries would therefore be beneficial for this assessment.

7.4. Additional areas of research

To progress the assessment, further research in the following areas is recommended:

- In the recent decade, one of the major sources of increased fishing pressure on the swordfish stock is from fisheries in the northern sub-regions. One of the sensitivity runs examined herein suggested potentially different stock status if the northern region was excluded from the assessment on the basis that those norther sub-regions represent a different (sub-)stock. Nevertheless, there is currently little known on the stock structure of SWP swordfish. For future quantitative stock assessments, information on stock boundaries should be developed through genetics and/or movement studies.
- After the 2013 assessment, notable advances were made on the knowledge of the growth /maturity/mortality schedules-at-age for SWP swordfish. This assessment used growth curve parameters from an otolith aging study. Validation of otolith-based estimates would be useful to further underpin this work. In turn, spatial heterogeneity in growth and maturity potentially needs to be explored. The priority on studies of spatial heterogeneity should be increased if genetic and/or movement studies provide evidence of the existence of sub-stocks in the assessment region.
- Collection of sex-specific size data is recommended. Currently a limited amount of sex-specific length measurements are available from a limited number of fisheries. If the next assessment is to be conducted with an explicitly sex-specific population dynamics model, comprehensive collection of sex-specific length-frequency data would help reduce the uncertainty.
- The reduced parameter uncertainty for the period where size data are more available has illustrated the importance of size composition data for swordfish. The continued sampling of weight and length information in the key fisheries is strongly recommended. Work to resolve conflicts within these data from within some fisheries, e.g. Australian size data, is encouraged to ascertain their causes, perhaps by reviewing the sampling protocols for weight and length collections.
- Standardised CPUE indices were important for deriving reasonable model estimates in this assessment, in particular for region 2 (07_DW_2C_pre-2001, 08_DW_2C_post-2001 and 11_EU_2). These studies would be enhanced by additional information for the 11_EU_2 catch and effort data to improve standardisation. Specifically, operational factors including the numbers of hooks, numbers of fish landed, use of light sticks, bait type and hook type would be beneficial. This would enable more comprehensive analyses of catch and effort data, particularly variations in targeting and fishing power. In addition, fine-scale analysis of these CPUE data may identify seasonal shifts in areas of relatively high CPUEs which may assist in defining movements of swordfish in the model region. Because of the decline of the Japanese and Chinese-Taipei longline effort in the central sub-regions after 2000, in the next stock assessment, the time series of the key CPUE indices from distant-water fisheries in the central sub-regions (02_DW_1C, 07_DW_2C_pre-2001 and 08_DW_2C_post-2001) used in the assessment may not be extendable due to the reduced spatial/temporal coverage of data. Developments of alternative CPUE indices from Japanese and/or Chinese-Taipei longline in northern sub-regions may be worth consideration, under the assumption that they do not include data from different sub stocks (see above). Alternatively the next assessment should explore additional sources catch and effort data such as that from China and/or Korea. To accomplish this, further examination of the operational data provided from those members will be necessary.

7.5. Main assessment conclusions

The main conclusions of the current assessment below are based upon the total grid of 72 model runs. The general conclusions of this assessment can be summarised as follows:

1. The grid contains a wide range of models with some variation in estimates of stock status, trends in abundance and reference points. Biomass is estimated to have declined throughout the model period for all models in the grid, but the decline is particularly steep in the last 15 years. Those declines are found in both model regions, but are particularly notable in region 2 (the eastern region).

2. Fishing mortality for juvenile (ages 1-3), maturing (ages 4-6) and adult (ages 7+) swordfish is estimated to have increased since the 1950s. Fishing mortality rate increased notably from the mid 1990s in both model regions, on maturing aged fish in particular (seen in the *DiagCase* model), to levels approximately four times that of juveniles and adults.

3. Noting that WCPFC has yet to formally agree a limit reference point for SWP swordfish, we have reported the main stock assessment results in terms of both spawning potential depletion and maximum sustainable yield (MSY)-related reference points. Across the model grid, the terminal spawning potential depletion estimated for all runs, $SB_{latest}/SB_{F=0}$, was above $20\%SB_{F=0}$. The median estimate was 0.35 (range 0.26-0.49). The median ratio of SB_{latest} to SB_{MSY} was 1.61 (range 0.85-4.06, 11% of which were <1.0).
4. The median estimate of F_{recent}/F_{msy} was 0.86 (range 0.42-1.46), with 23 out of the 72 runs (32%) indicating that $F_{recent}/F_{msy} > 1$. Runs where overfishing was indicated were generally those with a steepness of 0.65 assumed.
5. Unlike in the bigeye and yellowfin assessments, evidence for a strong increase in recent recruitment for swordfish was not found in either the CPUE time series or estimates of recruitment. Variability in the recruitment estimates for swordfish may in part mask any recent trend. We also note that the longline-only nature of the fishery, catching mainly larger, older swordfish, is not strongly informative with regards to recruitment dynamics.
6. The current assessment investigated the impact of a wide range of uncertainties. However, a key axis of uncertainty in the 2013 assessment – growth – has been reduced in the current assessment through the results from Farley *et al*. (2016). Nonetheless, there remain a range of other model assumptions that should be investigated either internally or through directed research. These are noted in the main text, but briefly, include further developments to MFCL to enable the sex-disaggregated assessment of this stock (given the data available), enhancement of sex-separated data collection, investigations into potential stock structure, further analysis of the size data available, and consideration of additional data required to enhance CPUE standardisation given the decline in fishing by key long-term fleets within the SWP.

8. Acknowledgements

We are grateful for the assistance provided by Peter Williams (SPC) in the preparation of input data-sets for the assessment. We thank the various fisheries agencies for the provision of the catch, effort and size-frequency data used in this analysis, and in particular to Robert Campbell (CSIRO), for the supply of size and catch-per-unit effort data.

9. References

Alvarado Bremer, J.R. Hinton, M.G. and Greig, T.W. 2006. Evidence of the spatial genetic heterogeneity in Pacific swordfish revealed by the analysis of LDH-A sequences. Bull. Mar. Sci. 79(3): 493-503.

Cadigan, N. G. and Farrell, P. J. 2005. Local influence diagnostics for the retrospective problem in sequential population analysis. ICES Journal of Marine Science: Journal du Conseil, 62(2):256–265.

Cadrin, S. and Vaughan, D. 1997. Retrospective analysis of virtual population estimates for Atlantic menhaden stock assessment. Fishery Bulletin, 95(3):256–265.

Campbell, R. 2008. Data summary pertaining to the catch of swordfish by longline fleets operating in the Southern WCPO. WCPFC-SC4-2008/SA-IP-3.

Campbell, R. 2012. Abundance indices for striped marlin and broadbill swordfish in the south-west Pacific based on standardised CPUE from the Australian longline fleet. WCPFC-SC8-SA-IP-13.

Davies, N., Griggs, L. and Unwin, M. 2005. Information available for developing a stock assessment for New Zealand swordfish (*Xiphias gladius*). Final research report for Ministry of Fisheries Research projects SWO2003-01. NIWA.

Davies, N., Campbell, R. and Kolody, D. 2006. CASAL Stock Assessment for South-West Pacific Broadbill Swordfish 1952-2004. Methods Specialist Working Group paper WCPFC-SC2 ME-WP-4 presented at the 2nd meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission, held 7-16 August, Manila, Philippines.

Davies, N., Bian, R., Kolody, D. and Campbell, R. 2008. CASAL Stock assessment for southwest-Central Pacific broadbill swordfish 1952-2007. WCPFC-SC4-2008/SA-WP-7.

Davies, N., Pilling, G., Harley, S. and Hampton, J. 2013. Stock assessment of swordfish (*Xiphias gladius*) in the southwest Pacific Ocean. WCPFC-SC9-2013/SA-WP-05.

Davies, N., Fournier, D., Takeuchi, Y., Bouyé, F. and Hampton, J. 2017. Developments in the MULTIFAN-CL software 2016-17. WCPFC-SC13-2017/SA-IP-05.

DeMartini E.E., Uchiyama J.H., Williams H.A. 2000. Sexual maturity, sex ratio, and size composition of swordfish, *Xiphias gladius*, caught by the Hawaii-based pelagic longline fishery. *US.Fish.Bull.* 98:489-506.

DeMartini E.E., Uchiyama J.H., Humphreys Jr., R.L., Sampaga, J.D., and Williams, H.A. 2007. Age and growth of swordfish (*Xiphias gladius*) caught by the Hawaii-based pelagic longline fishery. Fish. Bull. 105: 356-367.

Evans, K., Kolody, D., Abascal, F., Holdsworth, J., Maru, P. and Sippel, T. (2012). Spatial dynamics of swordfish in the South Pacific Ocean inferred from tagging data. WCPFC-SC8-2012/SA-IP-05.

Farley, J., Clear, N., Kolody, D., Krusic-Golub, K., Eveson P. and Young, J. (2016). Determination of swordfish growth and maturity relevant to the southwest Pacific stock. WCPFC-SC12-2016/ SA-WP-11.

Fournier, D.A., Hampton, J., and Sibert, J.R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences. 55: 2105–2116.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68: 1124–1138

Graves, J. E. and McDowell, J. R. 2003. Stock structure of the world's istiophorid billfishes: a genetic perspective. Marine and Freshwater Research. 54: 287–298.

Hampton, J. and Fournier, D.A. 2001. A spatially-disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. Marine and Freshwater Research.. 52:937–963.

Harley, S.J. 2011. Preliminary examination of steepness in tunas based on stock assessment results. WCPFC-SC7 SA-IP-08. Pohnpei, Federated States of Micronesia, 9-17 August 2011.

Harley, S., Williams, P., and Hampton, J. 2012. A compendium of fisheries indicators for bigeye, skipjack, yellowfin, and south Pacific albacore tunas and south Pacific swordfish. WCPFC-SC8-2012/SA-WP-02.

Holdsworth, J. C. Sippel, T. J. and P. J. Saul. 2007. An investigation into Swordfish Stock Structure Using Satellite Tag and Release Methods. Western and Central Pacific Fisheries Commission Biology Specialist Working Group Working Paper 3.

Hoyle, S, Davies, N. and Pilling, G. 2012. CPUE standardisation for striped marlin in the Western and Central Pacific Ocean. WPCFC-SC8/SA-IP-09.

Hoyle, S., Davies, N. and Chang, E. 2013. CPUE standardisation for swordfish in the southwestern Pacific Ocean. WCPFC-SC9/SA-IP-03.

Kleiber, P., Fournier, D.A., Hampton, J., Davies, N., Bouyé, F., and Hoyle, S. 2017. MULTIFAN-CL Users' Guide. http://www.multifan-cl.org/userguide.pdf.

Kolody,D. and Davies, N. 2008. Spatial structure in South Pacific Swordfish Stocks and Assessment Models. WCPFC-SC4-2008/SA-IP-2.

Kolody, D., Campbell, R. and Davies, N. 2008. A MULTIFAN-CL stock assessment of south-west Pacific swordfish 1952-2007. WCPFC-SC4-2008/SA-WP-6.

Kolody, D., Campbell, R. and Davies. N. 2006. A Multifan-CL Stock Assessment for South-West Pacific Swordfish 1952-2004. Western and Central Pacific Fisheries Commission Scientific Committee Meeting 2 – Methods Specialist Working Group Working Paper 3.

McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-WP-05, Rarotonga, Cook Islands, 9–17 August 2017.

Molony, B. W. 2005. Summary of the biology, ecology and stock status of billfishes in the WCPFC, with a review of major variables influencing longline fishery performance. Working Paper EB WP–2. 1st Meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission (WCPFC–SC1), Noumea, New Caledonia, 8–19 August 2005.

Nishikawa, Y., Honma, M., Ueyanagi, S. and Kikawa, S. (1985) Average distribution of larvae of oceanic species of scombrid fishes, 1956–1981. Far Seas Fisheries Research Laboratory, Shimizu. S Series 12.

Oceanic Fisheries Programme. 2013. Report from the SPC pre-assessment workshop, Nouméa, April 2013. WCPFC-SC9-2013/SA-IP-01.

Oceanic Fisheries Programme. 2012. Update of progress towards a stock assessment for swordfish in the southern WCPO. WCPFC9-2012-IP10.

Reeb, C.A., Arcangeli, L., Block, B. 2000. Structure and migration corridors in Pacific Ocean populations of the swordfish, Xiphias gladius, as inferred through analysis of mitochondrial DNA. Working paper BBRG-13 presented at the 13th meeting of the Standing Committee on Tuna and Billfish, held 5-12 July 2000, Noumea, New Caledonia.

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. WCPFC-SC13-2017/SAWP-06, Rarotonga, Cook Islands, 9–17 August 2017.

Valeiras, X., J. Mejuto, M. Ruiz. 2008. Age and growth of swordfish (Xiphias Gladius) in the North Pacific. WCPFC-SC4-2008/BI-WP-1.

Williams, P., Harley, S., and Campbell, R. 2011. South Pacific swordfish data available for stock assessments. WCPFC-SC7-2011/ST IP-04.

Young, J. and Drake, A. 2002. Age and growth of broadbill swordsh (*Xiphias gladius*) from Australian waters. FRDC Project 2001/014.

Young, J. and A. Drake. 2004. Age and growth of broadbill swordfish (*Xiphius gladius*) from Australian waters. Final report for project 2001/014, Fisheries Research Development Corporation, Canberra, Australia.

Young, J., Humphreys, R., Uchiyama, J. and Clear, N. 2008. Comparison of maturity and swordish from Hawaiian and Australian waters. WCPFC-SC4-2008/BI-IP-2.

Table 1. Description of the fisheries and summary of information used in the assessment.

Fishery	Sub-area	Label	Method	Flag	Catch	Effort	Years
1	1N	01_DW_1N	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	1952-2015
2	1C	02_DW_1C	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	1953-2015
3	1S	03_DW_1S	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	1962-2015
4	1N, 1C, 1S	04_AU_1	Longline	AU	Number	Hooks	1986-2015
5	1N, 1C, 1S	05_Other_1	Longline	AS, BZ, CK, FM, FJ, PF, GE,GU, IN, ID, KI, MH, NC, NZ, NU, PW, PG, PH, WS, SB, SU, TO, TV, USAS, USMC, USHW, US, VU, VN, EU	Number	Hooks	1983-2015
6	2N	06_DW_2N	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	1952-2015
7	2C	07_DW_2C_pre 2001	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	1954-2000
8	2C	08_DW_2C_post 2001	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	2001-2015
9	2S	09_DW_2S	Longline	CN, CNOS, JPDW, JP, JPOS, KRDW, KR, TWDW, TW, TWOD, TWOS	Number	Hooks	1958-2015
10	2C, 2S	10_NZ_2	Longline	NZ	Number	Hooks	1990-2015
11	2N, 2C, 2S	11_EU_2	Longline	ES,PO	Number	sets	2004-2015
12	2N	12_Other_2N	Longline	AS, AU, BZ, CK, FM, FJ, PF, GE,GU, IN, ID, KI, MH, NC, NU, PW, PG, PH, WS, SB, SU, TO, TV, USAS, USMC, USHW, US, VU, VN	Number	Hooks	1982-2015
13	2C	13_Other_2C	Longline	AS, AU, BZ, CK, FM, FJ, PF, GE,GU, IN, ID, KI, MH, NC, NU, PW, PG, PH, WS, SB, SU, TO, TV, USAS, USMC, USHW, US, VU, VN	Number	Hooks	1982-2015

Table 2. Number of swordfish in length- and weight-frequency samples for each of the defined fisheries.

	Length-frequency	Weight-frequency
01_DW_1N	1,374	0
02_DW_1C	2,368	0
03_DW_1S	171	0
04_AU_1	12,360	391,818
05_Other_1	4,273	0
06_DW_2N	54,156	1,716
07_DW_2C_pre 2001	1,743	0
08_DW_2C_post 2001	11,74	0
09_DW_2S	471	0
10_NZ_2	6472	32,831
11_EU_2	204,562	0
12_Other_2N	10,992	5,699
13_Other_2C	4,205	1,315
Total	**304,271**	**433,379**

Table 3. Biological parameters used in the assessment.

Parameter	Value	Comment	Source
Number of age classes	20	Fixed. Pools all fish 20 years and older together in the oldest age class.	2008 assessment.
Length-weight relationship ($L = aW^b$)	a= 3.879 e-06; b= 3.24	Fixed	Davies *et al.* (2005)
Growth parameters (von Bertalanffy)	Mean length at age 1: 85.75 cm; Mean length at age 20+: 239.20 cm; k: 0.196 $year^{-1}$	Fixed Fixed Fixed	De Martini. (2007)
Mean Natural mortality coefficient	0.31 and 0.256 $year^{-1}$	Fixed	Kolody *et al.* (2008)
Maturity ogive (females)	Age: 1 to 20+ 0.01 0.061 0.03 0.11 0.35 0.71 0.91 0.98 1 1 1 1 1 1 1 1 1 1	Fixed	Kolody *et al.* (2008)
Beverton-Holt stock-recruitment relationship - steepness	0.8 Sensitivity: 0.65, 0.95	Fixed	Kolody *et al.* (2008)

Table 4. Definitions of the four assumed natural mortality vectors.

Input temperature for Pauly's formula	Growth	Maturity at age	Mean natural mortality	AdultM>JuvenileM
High : 22.83 °C	Otolith	100% age 10	0.31	Yes
High : 22.83 °C	Otolith	100% age 10	0.31	No
Low : 14.57 °C	Otolith	100% age 10	0.26	Yes
Low : 14.57 °C	Otolith	100% age 10	0.26	No

Table 5. Main structural assumptions used in the analysis.

Category	Assumption
Observation model for total catch data	Observation errors small, equivalent to a residual SD on the log scale of 0.07.
Observation model for length- and weight-frequency data	Normal probability distribution of frequencies with variance determined by sample size and observed frequency. Effective sample size is assumed to be 0.05 times actual weight-frequency sample size and 0.05 times the actual length-frequency sample size with a maximum effective sample size of 50.
Recruitment	Occurs as discrete events in the first quarter of each year. Annual variation in the proportions of recruitment to each region was estimated. Recruitment is related (CV=0.5) to spawning potential with no lag period via a Beverton-Holt SRR with steepness fixed at 0.8. Alternative, values were 0.65 and 0.95.
Initial population	Equilibrium age structure in the region as a function of the estimated natural mortality.
Age and growth	20 annual age-classes, with the last representing a 20+ age group. A fixed von Bertalanffy growth curve was assumed. Mean weights (W_j) computed internally by estimating the distribution of weight-at-age from the distribution of length-at-age and applying the weight-length relationship. Parameter values are in Table 3.
Selectivity	Constant over time. Fishery 9 had logistic selectivity. For all other fisheries, selectivity was described by splines (with 3 nodes except for fisheries 3, which had 5 nodes). Fishery 3 was constrained to have non-decreasing selectivity. The coefficients for age-classes above age 18 years were constrained to be equal for the non-decreasing selectivities.
Catchability	Seasonal variation for all fisheries. All fisheries, except that for which the CPUE index is fitted, have structural time-series variation, with random steps (catchability deviations) taken every 2 years. Catchability deviations constrained by a prior distribution with a normal mean 0 and SD 0.1.
Fishing effort	Fisheries for which the CPUE index is fitted, the effort deviations are constrained by a temporally-variable penalty weight based upon the index coefficient of variation (constrained to have a mean of 0.2). For other fisheries, variability of effort deviations was constrained by a penalty weight scaled by the square root of the effort.
Natural mortality	Fixed according to the 4 assumed schedules in Table 3 and Figure 10 **Error! Reference source not found.**
Movement	Quarterly and assumed constant at a diffusion rate of 0.11, with sensitivities of 0.0, 0.05 and 0.25.

Table 6. Names and descriptions of the key model runs undertaken for the 2017 swordfish assessment. The diagnostic case is in bold and all other runs are one-off sensitivities to the diagnostic case.

Run name	Description
DiagCase	**Steepness = 0.8; Movement diffusion rate = 0.11;Natural mortality M1; Fit to 02_DW_1C, 04_AU_1, 07_DW_2C pre-2001, 08_DW_2C_post-2001, 11_EU_2 CPUE indices; size data relative weight is 03_DW_1S/09_DW_2S: n/10, Other: n/20.**
M2	M2 natural mortality option
M3	M3 natural mortality option
M4	M4 natural mortality option
h0.65	Assume stock-recruitment relationship steepness = 0.65
h_0.95	Assume stock-recruitment relationship steepness = 0.95
mv_0	Assume no movement among regions
mv_0.25	Assume a movement diffusion rate = 0.25
DownWeight	Low size data relative weight is 03_DW_1S/09_DW_2S: n/20, Other: n/40.
Remove1N/2N	Removed fisheries in sub regions 1N and 2N (01_DW_1N, 06_DW_2N, and 12_Other_2N)
-Selblock	Removed selectivity time blocks for Fishery 4 (04_AU_1)
Selblock06	Break point of selectivity time blocks is shifted to 2006
Selblock07	Break point of selectivity time blocks is shifted to 2007
Selblock09	Break point of selectivity time blocks is shifted to 2009
Selblock10	Break point of selectivity time blocks is shifted to 2010.
-AUS1L	Removed length composition data for Fishery 4 (04_AU_1)

Table 7. Description of symbols used in describing the stock assessment results and yield analysis.

Symbol	Description
C_{latest}	Catch in the last year of the assessment (2015)
F_{recent}	Average fishing mortality-at-age for a recent period (2011–2014)
F_{MSY}	Fishing mortality-at-age producing the maximum sustainable yield (MSY)
MSY	Equilibrium yield at F_{MSY}
$F_{recent}/F_{\mathrm{MSY}}$	Average fishing mortality-at-age for a recent period (2011–2014) relative to F_{MSY}
SB_0	Equilibrium unexploited spawning potential
SB_{latest}	Spawning potential in the latest time period (2015)
SB_{recent}	Spawning potential for a recent period (2011–2014)
$SB_{F=0}$	Average spawning potential predicted to occur in the absence of fishing for the period 2005–2014
SB_{MSY}	Spawning potential that will produce the maximum sustainable yield (MSY)
$SB_{latest}/SB_{F=0}$	Spawning potential in the latest time period (2015) relative to the average spawning potential predicted to occur in the absence of fishing for the period 2005–2014
$SB_{latest}/SB_{\mathrm{MSY}}$	Spawning potential in the latest time period (2015) relative to that which will produce the maximum sustainable yield (MSY)
$SB_{recent}/SB_{F=0}$	Spawning potential in for a recent period (2011–2014) relative to the average spawning potential predicted to occur in the absence of fishing for the period 2005–2014
$20\%SB_{F=0}$	WCPFC adopted limit reference point – 20% of spawning potential in the absence of fishing average over years $t-10$ to $t-1$ (2005–2014)

Table 8. Important reference points and model results summarized by quantities from One-off sensitivity runs.

	Diag.Case	M2	M3	M4	Mv0	Mv0.25	DownWeight	h0.65	h0.95
C_{latest}	9329.79	9461.68	9342.80	9343.06	9930.90	9339.07	9663.57	9335.54	9322.57
MSY	8335.00	8177.00	7508.00	7490.00	8945.00	8302.00	9380.00	7105.00	10080.00
$Y_{F_{recent}}$	8271.00	8134.00	7492.00	7475.00	8087.00	8255.00	8856.00	6946.00	9128.00
f_{mult}	1.12	1.10	0.95	0.95	1.55	1.10	1.38	0.85	1.70
F_{MSY}	0.16	0.14	0.15	0.14	0.16	0.16	0.16	0.12	0.23
$F_{recent}/F_{\mathrm{MSY}}$	0.89	0.91	1.06	1.05	0.65	0.91	0.72	1.17	0.59
SB_{MSY}	15750.00	15320.00	19100.00	19660.00	16630.00	15670.00	17630.00	21890.00	8889.00
SB_0	73200.00	74340.00	87500.00	92930.00	78800.00	72840.00	82860.00	78680.00	70150.00
$SB_{F=0}$	64904.22	66963.44	79840.11	85219.44	74841.00	63692.00	74434.56	67041.56	63783.67
SB_{latest}/SB_0	0.28	0.27	0.25	0.25	0.45	0.27	0.35	0.28	0.28
$SB_{latest}/SB_{F=0}$	0.31	0.30	0.27	0.27	0.48	0.31	0.39	0.33	0.31
$SB_{latest}/SB_{\mathrm{MSY}}$	1.30	1.33	1.13	1.18	2.14	1.27	1.63	1.01	2.20
$SB_{recent}/SB_{F=0}$	0.33	0.33	0.30	0.31	0.45	0.33	0.40	0.35	0.33
$SB_{recent}/SB_{\mathrm{MSY}}$	1.18	1.15	0.91	0.91	1.76	1.16	1.55	0.77	2.22

Table 9. Data component and stock recruitment relationship negative log likelihoods from One-off sensitivity runs and the diagnostic case model.

Component	Diag.Case	M2	M3	M4	Mv0	Mv0.25	DownWeight	h0.65	h0.95
Beverton Holt	11.7603	11.5296	12.2025	12.0428	8.85862	12.2781	9.9831	12.1682	11.6204
Effort devs	1105.08	1104.86	1101.71	1104.20	1044.60	1117.54	1078.90	1103.33	1105.34
Catch devs	97.28	95.976	98.9038	99.05	86.18	99.47	92.28	97.89	97.24
Length comps	-16943.32	-16936.60	-16946.18	-16945.24	-16936.95	-16942.54	-10675.63	-16942.54	-16943.62
Wgt comps	-10359.64	-10357.54	-10362.45	-10358.63	-10361.85	-10358.84	-7707.20	-10359.62	-10359.54
Total	-26061.02	-25753.68	-25768.03	-25760.75	-25821.25	-25748.10	-16871.78	-25760.70	-25761.04

Table 10. Important reference points and model results summarized by quantities from other sensitivity runs.

	Diag.Case	-AUS1L	-1Nand2N	-Selblock	Selblk06	Selblk07	Selblk09	Selblk10
C_{latest}	9329.79	9560.48	4150.40	9401.95	9368.56	9381.78	9369.99	9361.82
MSY	8335.00	7547.00	4526.00	8085.00	7981.00	7944.00	7935.00	7924.00
$Y_{F_{recent}}$	8271.00	7525.00	4387.00	8050.00	7955.00	7923.00	7928.00	7916.00
f_{mult}	1.12	1.07	1.27	1.09	1.08	1.07	1.04	1.04
F_{MSY}	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
$F_{recent}/F_{\text{MSY}}$	0.89	0.93	0.79	0.92	0.93	0.94	0.96	0.96
SB_{MSY}	15750.00	14220.00	8301.00	15250.00	15030.00	14960.00	14890.00	14890.00
SB_0	73200.00	66180.00	38830.00	70920.00	70120.00	69760.00	69440.00	69410.00
$SB_{F=0}$	64904.22	61605.33	36022.70	63615.22	62656.00	62332.67	61479.44	61446.67
SB_{latest}/SB_0	0.28	0.30	0.30	0.27	0.27	0.27	0.27	0.27
$SB_{latest}/SB_{F=0}$	0.31	0.32	0.32	0.31	0.31	0.31	0.30	0.30
$SB_{latest}/SB_{\text{MSY}}$	1.30	1.40	1.40	1.28	1.28	1.28	1.24	1.24
$SB_{recent}/SB_{F=0}$	0.33	0.32	0.33	0.33	0.32	0.32	0.31	0.31
$SB_{recent}/SB_{\text{MSY}}$	1.18	1.11	1.39	1.14	1.12	1.11	1.06	1.06

Table 11. Data component and stock recruitment relationship negative log likelihoods from other sensitivity runs and the diagnostic case model.

Component	Diag.Case	-AUS1L	Remove1N/2N	-Selblock	Selblk06	Selblk07	Selblk09	Selblk10
Beverton Holt	11.76	9.47	15.80	11.48	11.19	10.50	11.20	10.83
Effort devs	1105.08	1144.67	862.22	1109.10	1139.251	1080.92	1135.45	1136.84
Catch devs	97.28	104.59	3.14	97.37	99.96	98.99	99.83	99.83
Length comps	-16943.32	-14623.80	-10801.59	-16942.55	-16974.12	-16974.38	-16975.58	-16976.79
Wgt comps	-10359.64	-10356.45	-7374.45	-10355.55	-10355.10	-10356.31	-10357.79	-10356.65
Total	-26061.02	-23701.53	-17222.35	-25752.84	-26055.43	-26087.55	-26062.56	-26062.53

Table 12. Summaries of important reference points and model results summarized by quantities across all models in the structural uncertainties.

	Mean	Median	Min	25	75	Max
C_{latest}	9748.41	9884.85	9318.12	9395.00	10088.20	10287.80
MSY	8176.83	7914.00	5898.00	6947.50	9190.00	11380.00
$Y_{F_{recent}}$	7622.21	7752.00	5223.00	6880.25	8522.50	9538.00
f_{mult}	1.30	1.16	0.68	0.94	1.62	2.38
F_{MSY}	0.16	0.14	0.10	0.12	0.20	0.23
$F_{recent}/F_{\text{MSY}}$	0.86	0.86	0.42	0.62	1.06	1.46
SB_{MSY}	17275.74	17705.00	7251.00	10935.00	21760.00	30440.00
SB_0	84167.78	84075.00	57070.00	75517.50	92425.00	111000.00
$SB_{F=0}$	78615.14	78301.44	61996.89	69430.06	86581.64	100691.44
SB_{latest}/SB_0	0.33	0.32	0.24	0.27	0.40	0.46
$SB_{latest}/SB_{F=0}$	0.35	0.35	0.26	0.30	0.40	0.49
$SB_{latest}/SB_{\text{MSY}}$	1.86	1.61	0.85	1.25	2.20	4.06
$SB_{recent}/SB_{F=0}$	0.37	0.36	0.29	0.33	0.40	0.46
$SB_{recent}/SB_{\text{MSY}}$	1.52	1.23	0.46	0.91	2.13	3.38

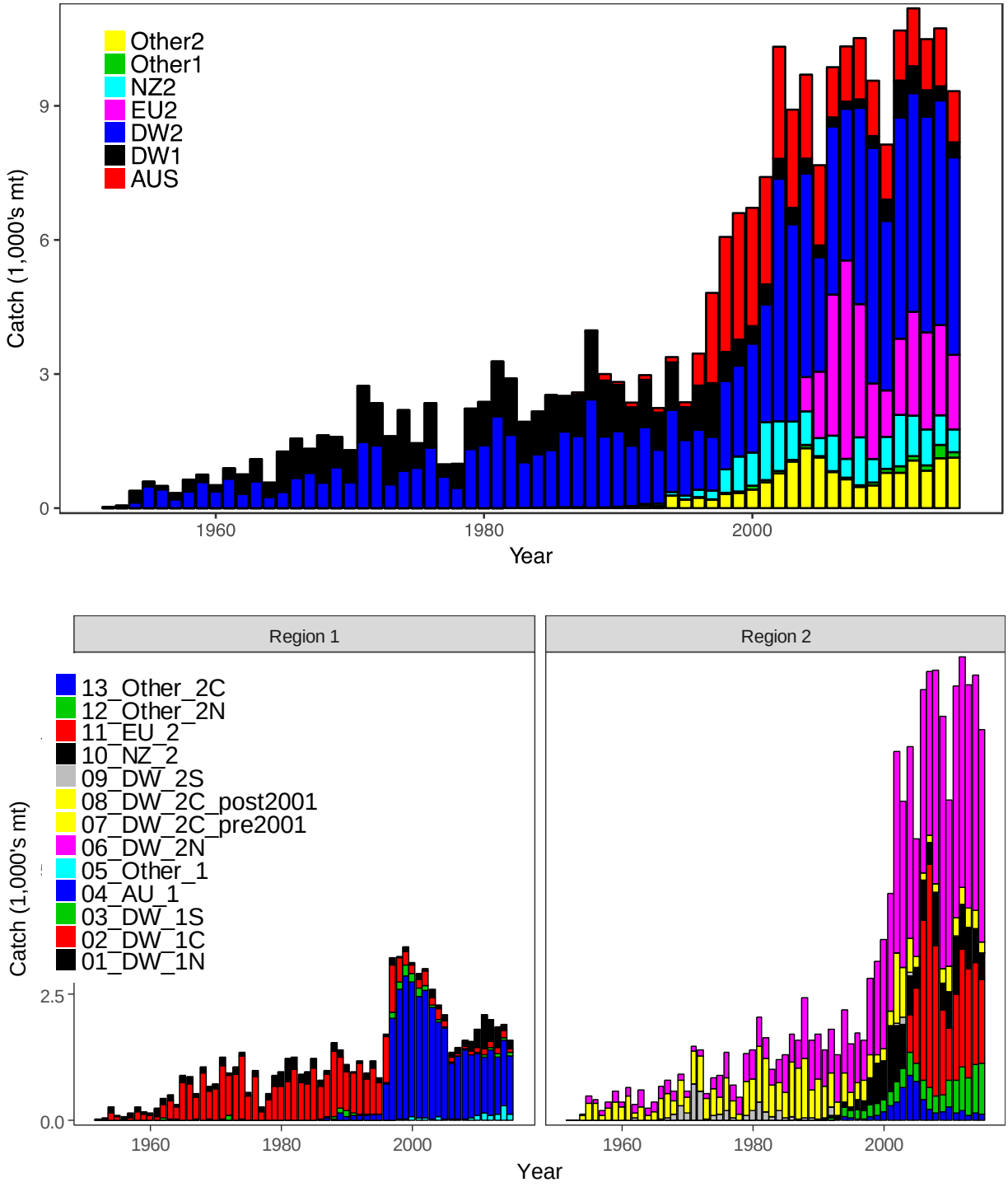

Figure 1. Total swordfish catches in weight (top), those by region (bottom) grouped by major longline-method fisheries in the model regions, 1952–2011. In the upper panel: DW1 - distant water fleet region 1; AUS – Australian region 1; Other1 - Other fisheries region 1; DW2 - distant water fleet region 2; NZ2 - New Zealand region 2; EU2 – EU (Spanish) region 2; Other2 - other fisheries region 2.

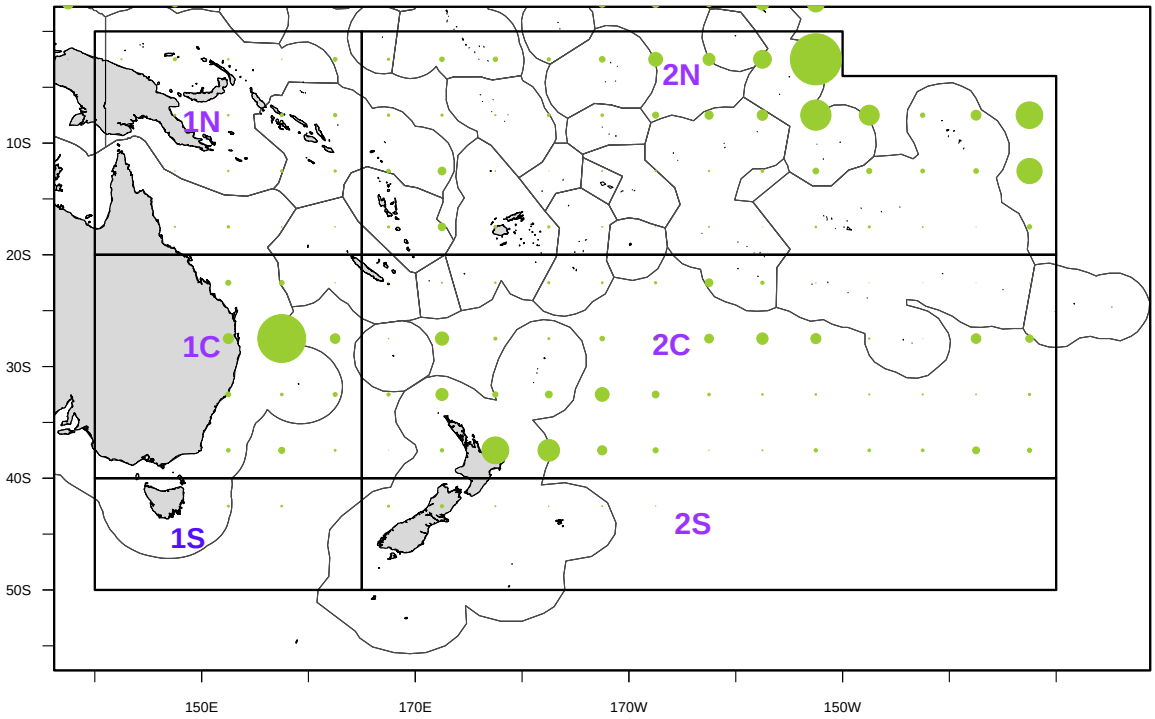

Figure 2. Catches of swordfish (numbers) in the southwest Pacific, 2006–2015. Source: raised catch estimates available from the SPC. The black lines represent the boundaries of the assessment regions 1 and 2 (outer lines) for swordfish in the southwest Pacific Ocean, and the six fishery sub-areas within those regions.

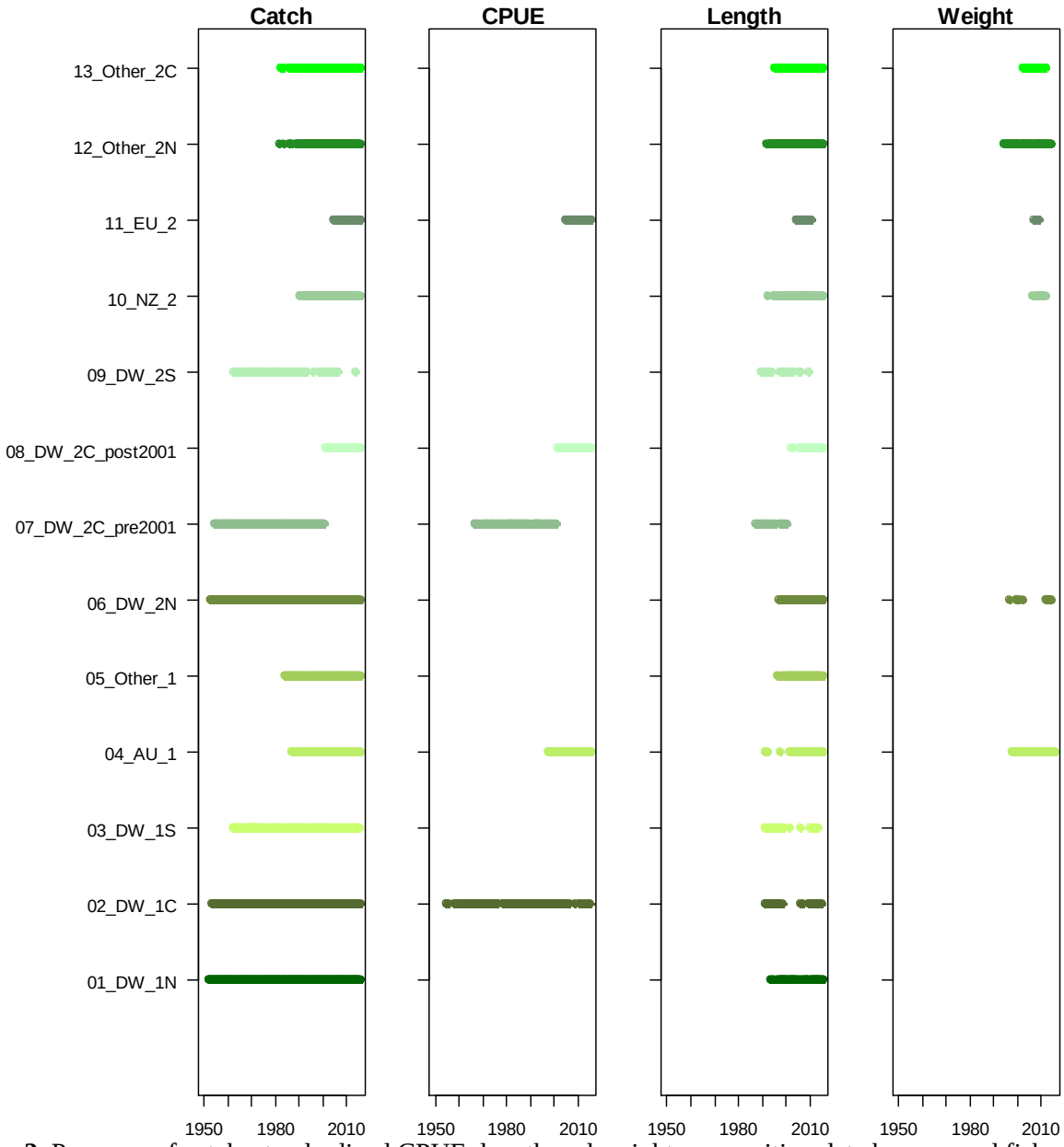

Figure 3. Presence of catch, standardized CPUE, length and weight composition data by year and fishery for the diagnostic case model over the full assessment period.

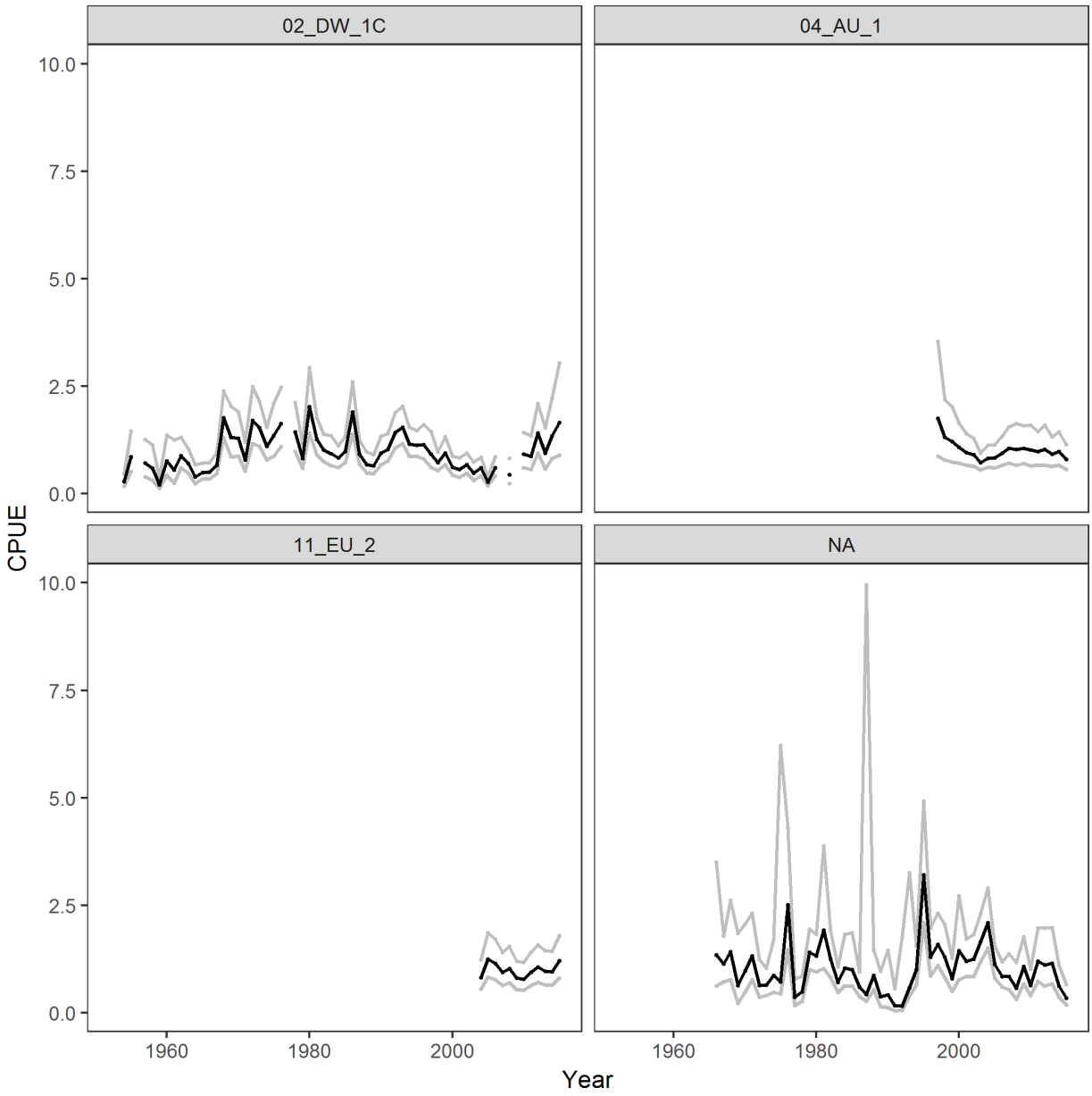

Figure 4. Standardized catch-per-unit-effort (CPUE) indices used for diagnostic case model. The light grey lines represent the 95% confidence intervals derived from effort deviation penalties used in the diagnostic case model.

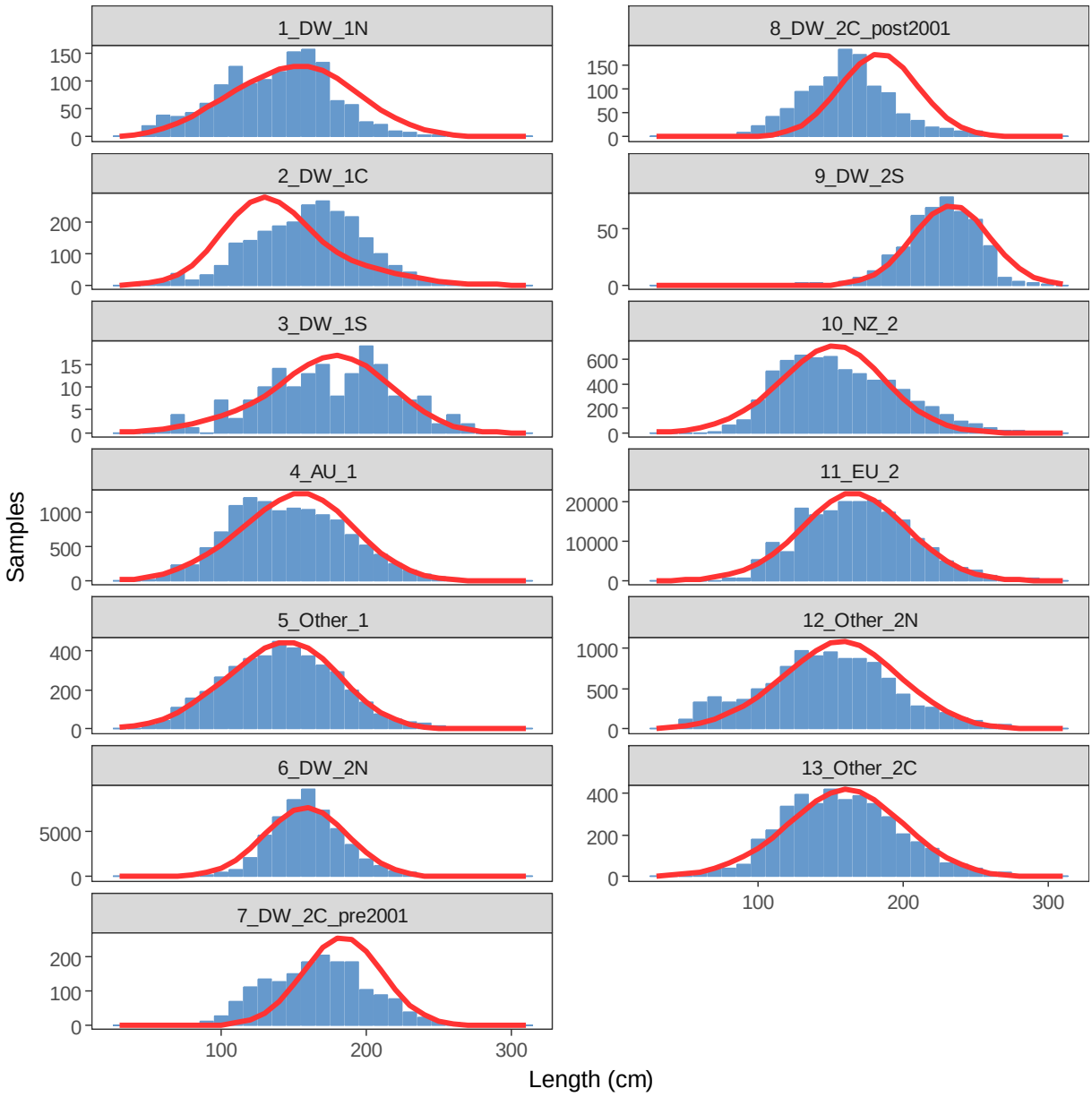

Figure 5. Composite (all time periods combined) observed (blue histograms) and predicted (red line) length compositions for all fisheries for the DiagCase.

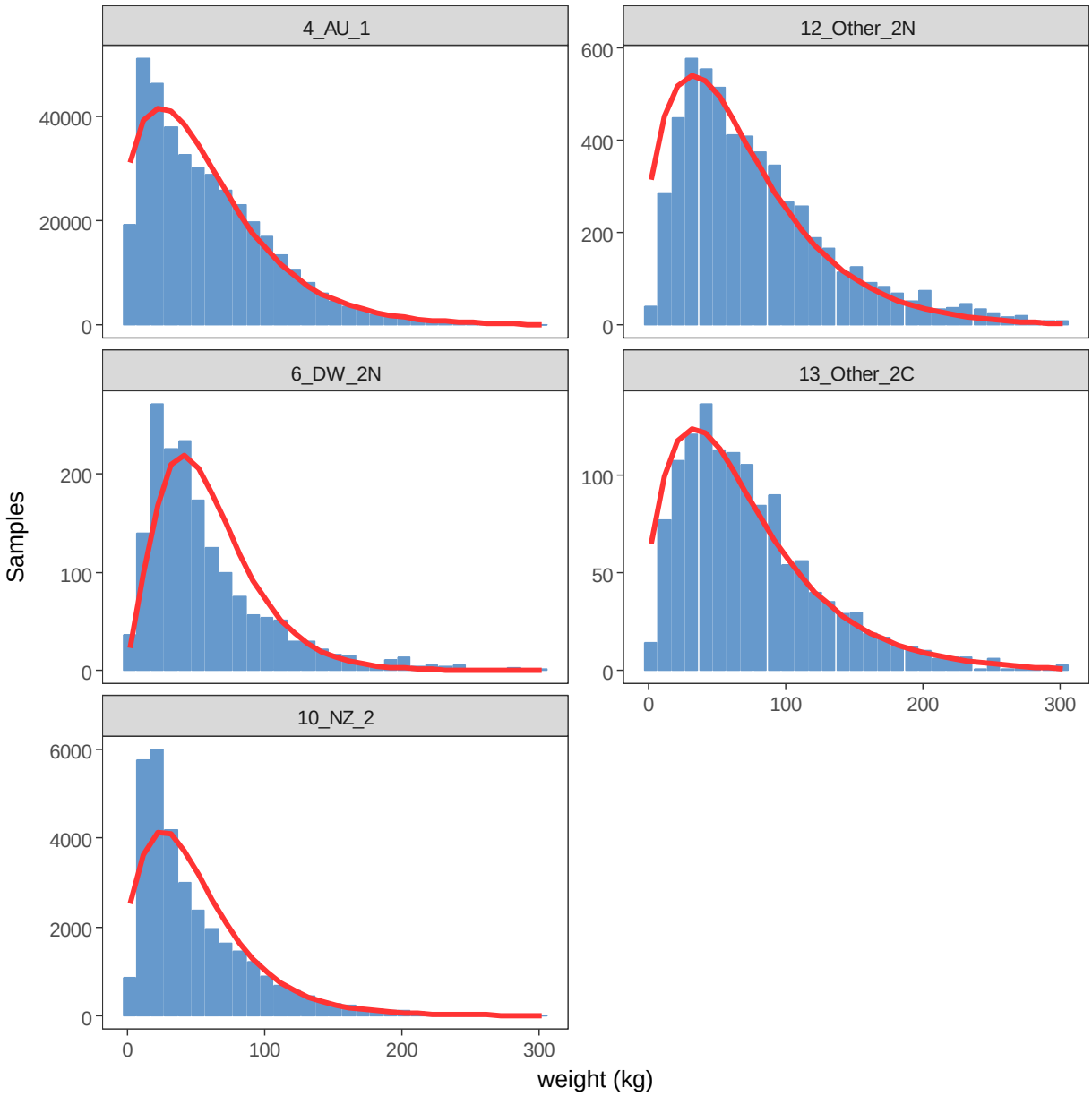

Figure 6. Composite (all time periods combined) observed (blue histograms) and predicted (red line) weight compositions for all fisheries for the DiagCase.

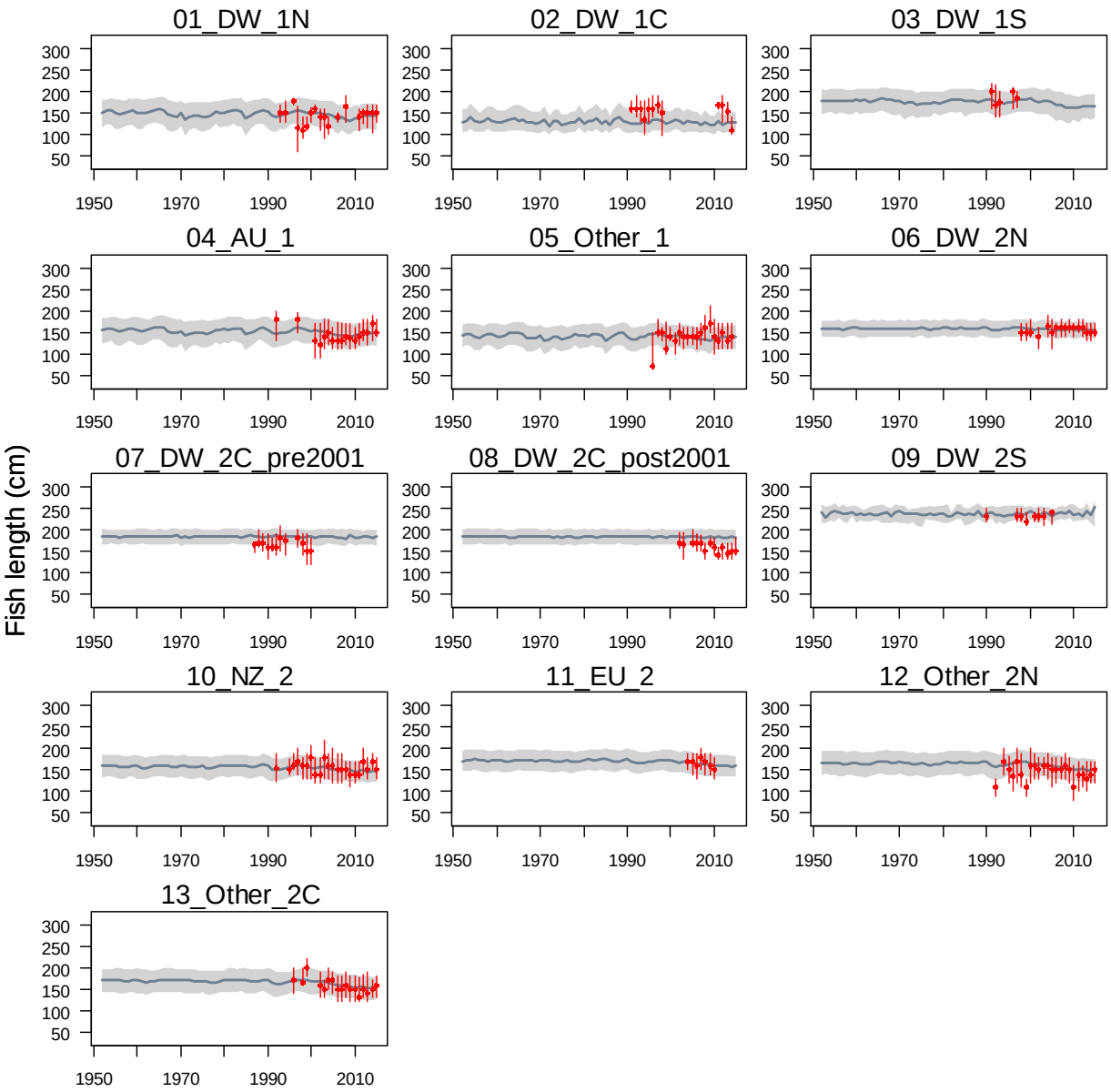

Figure 7. A comparison of the observed (red points) and predicted (grey line) median fish length (LJFL cm) for all fisheries with length samples for the diagnostic case model. The uncertainty intervals (grey shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are aggregated by year and only length sample with a minimum of 30 fish per year are plotted. For all fisheries for the DiagCase.

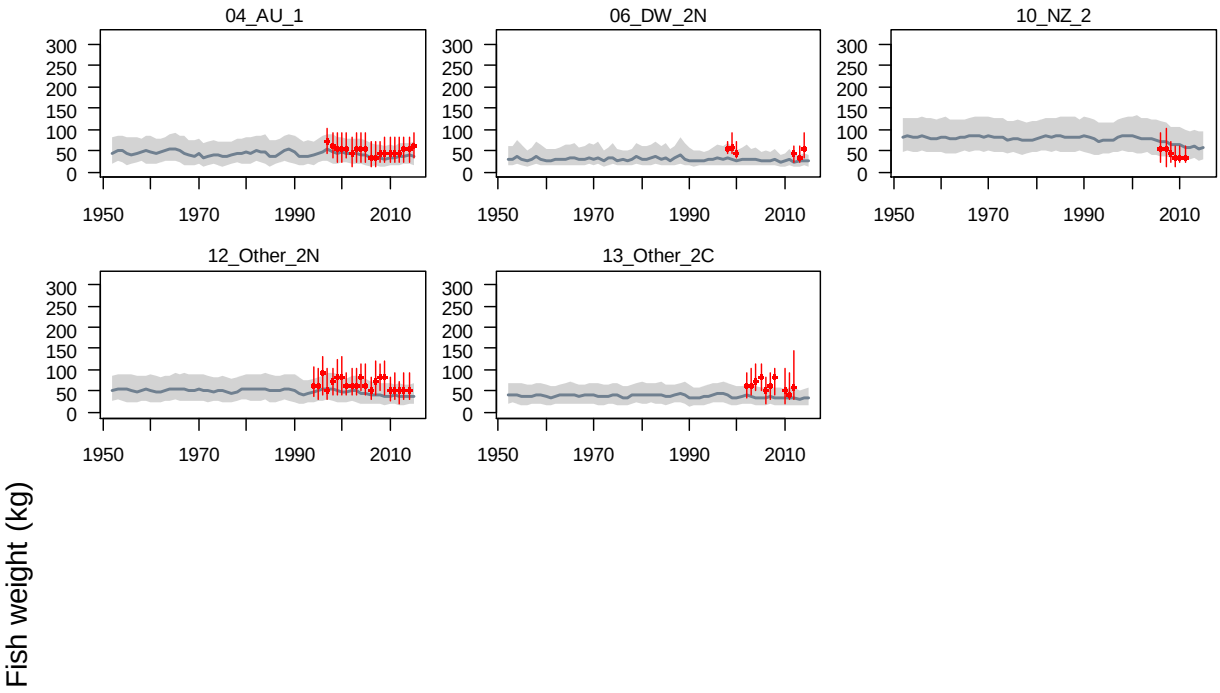

Figure 8. A comparison of the observed (red points) and predicted (grey line) median fish length (kg) for all fisheries with length samples for the diagnostic case model. The uncertainty intervals (grey shading) represent the values encompassed by the 25% and 75% quantiles. Sampling data are aggregated by year and only length sample with a minimum of 30 fish per year are plotted. For all fisheries for the DiagCase.

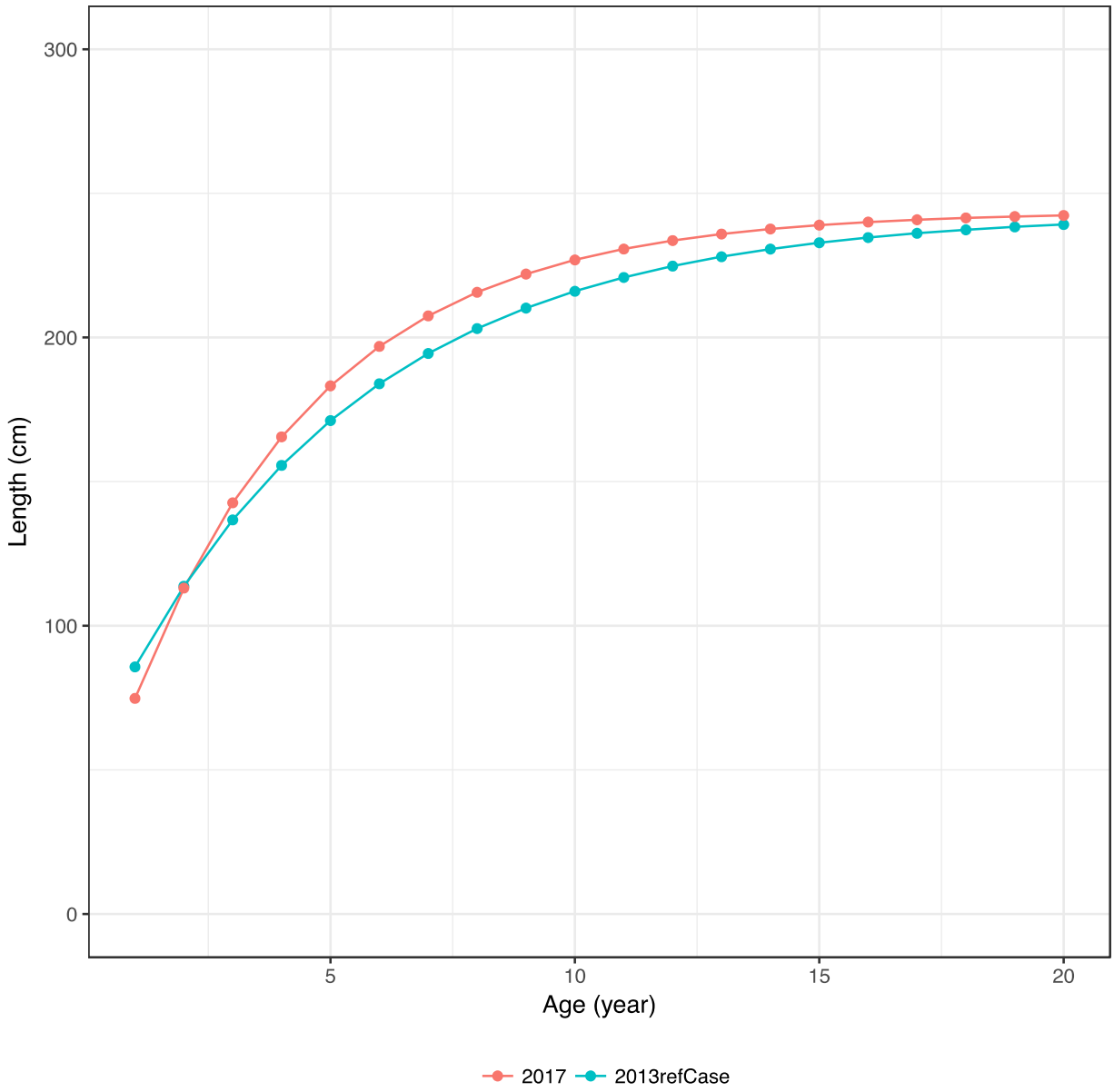

Figure 9. Length-at-age (in years) as assumed in this assessment (orange). Length-at-age (in years) as assumed in the 2013 ref.case (green).

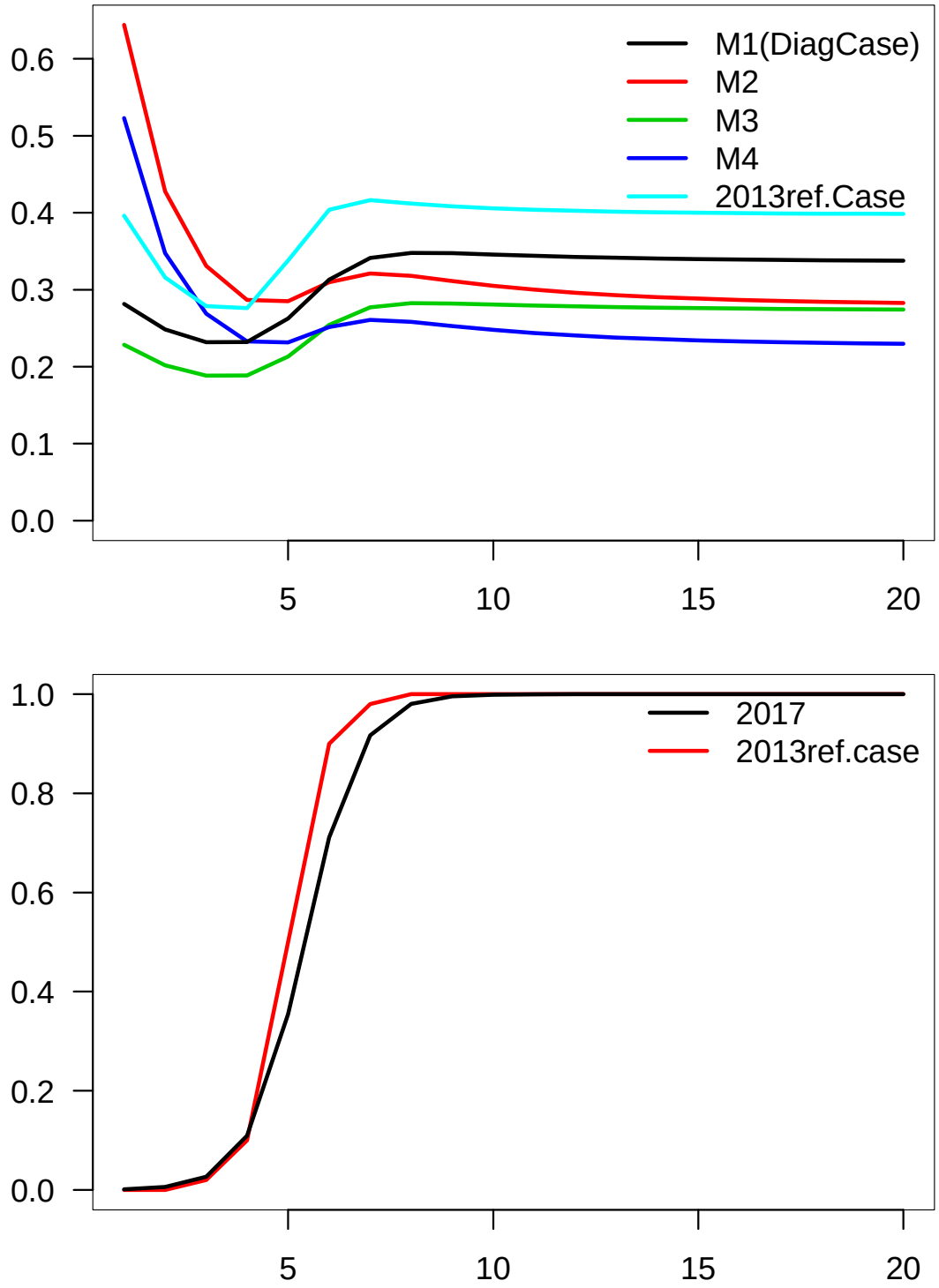

Figure 10. Top: Annual natural mortality-at-age as assumed in this stock assessment. Black line represents M1 (applied for diagnostic case). Red, green and blue lines represent M2, M3 and M4 respectively (applied for sensitivity and grid), Light blue line represents the natural mortality at age applied for 2013 ref.case. Bottom: maturity at age assumed in the 2017 assessment (black) and in the 2013 ref.case (red).

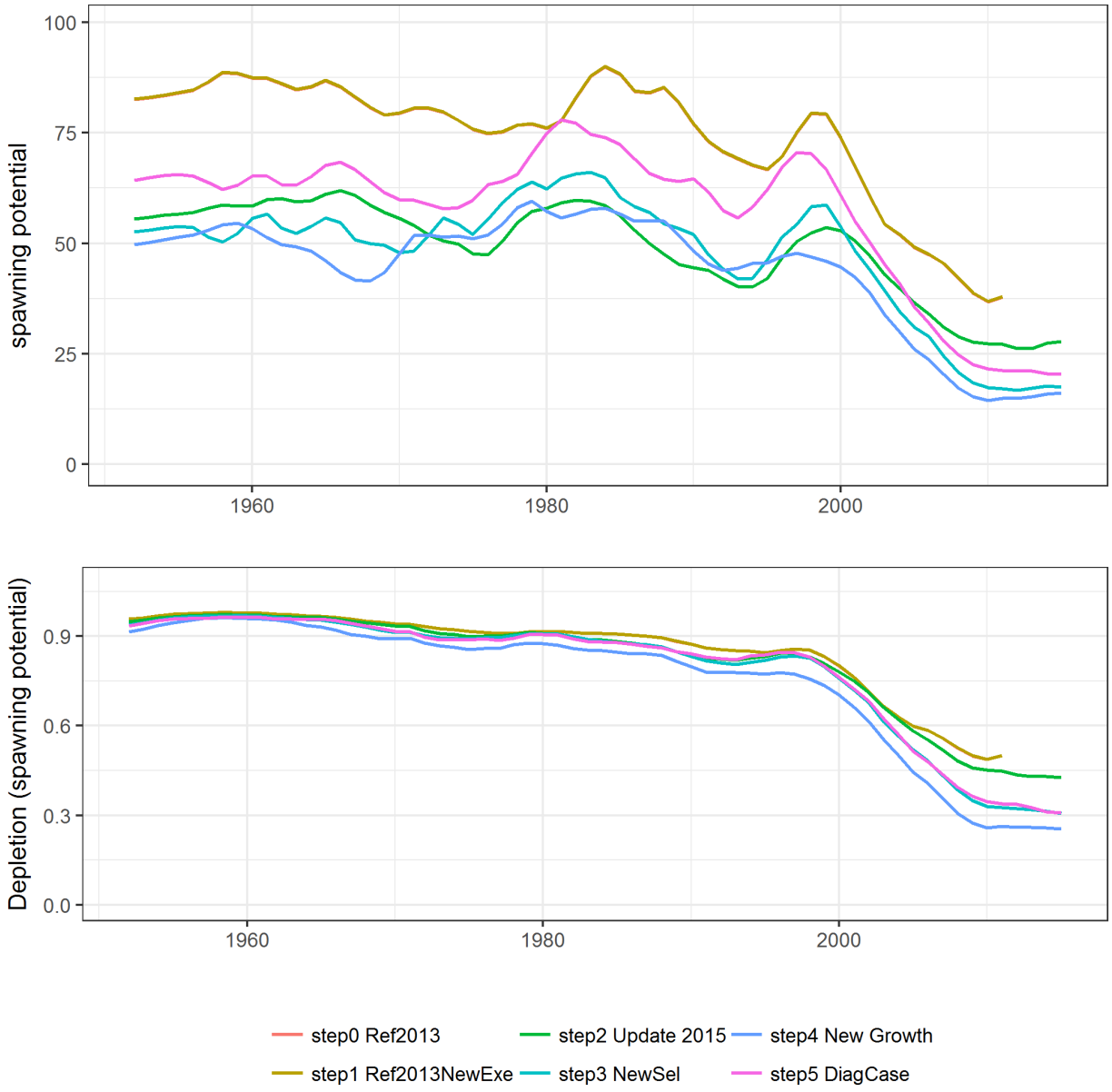

Figure 11. Stepwise changes in spawning potential (a, top) and fishing depletion (b, bottom) from the 2014 reference case model through to the 2017 diagnostic case model. For details of each step, see section 6.1.

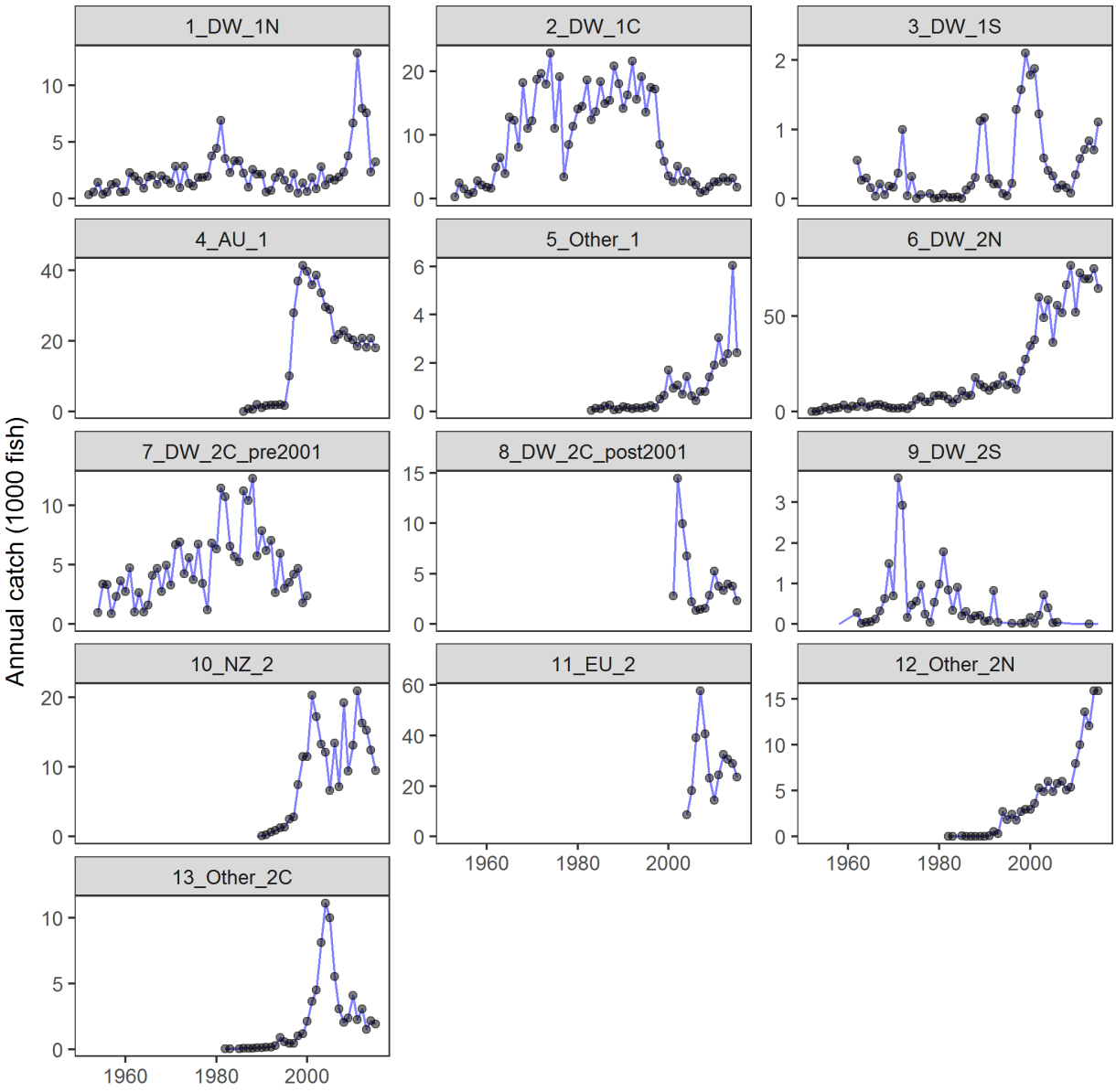

Figure 12. Observed and predicted catch for the diagnostic case model.

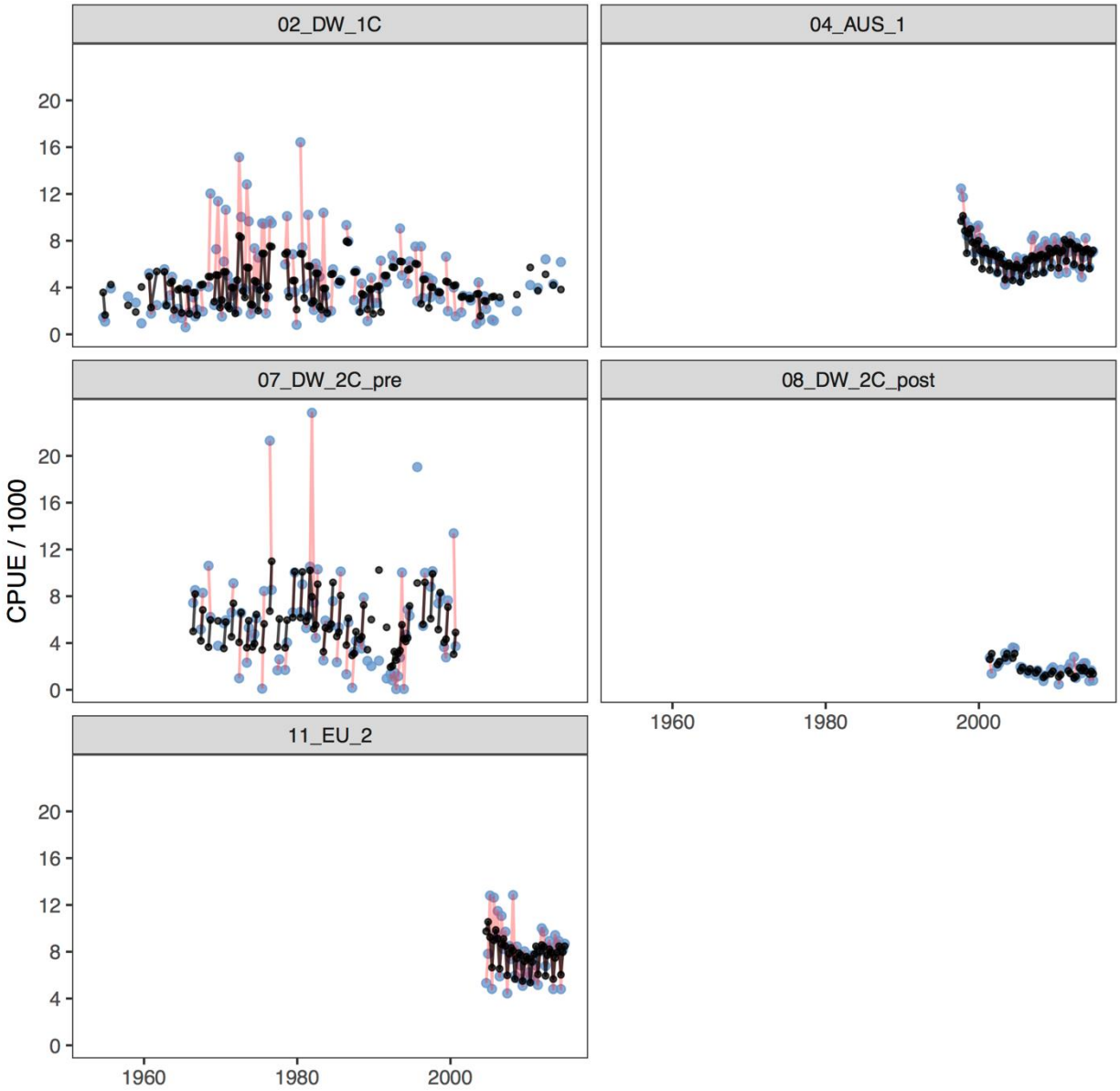

Figure 13. Observed (blue points with red lines) and predicted cpue (black points and lines) CPUE for the 5 fisheries which received standardized CPUE indices in the diagnostic case model.

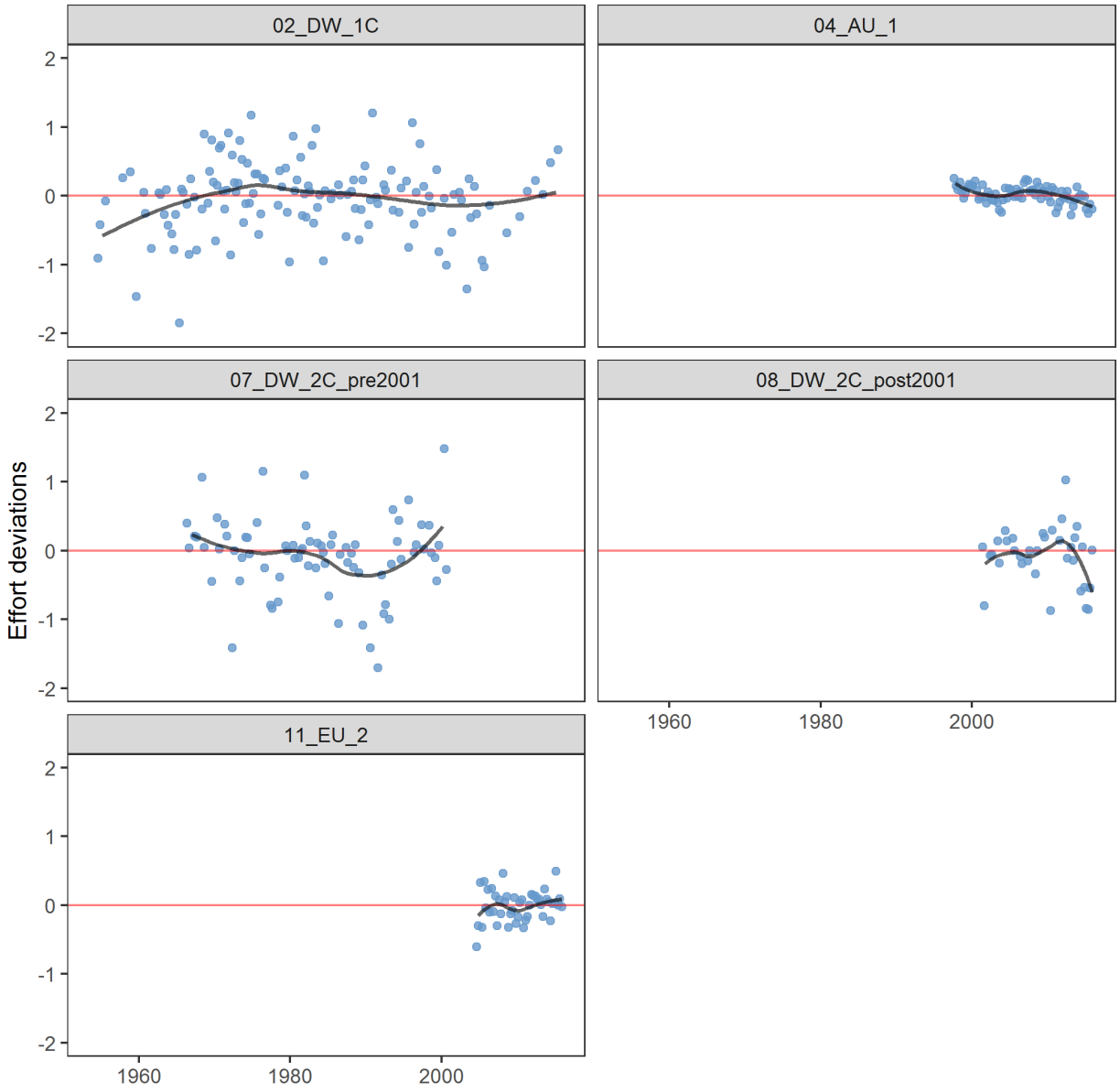

Figure 14. Effort deviations by time period for each of the fisheries receiving standardized CPUE indices in the diagnostic case model. The dark line represents a lowess smoothed fit to the effort deviations.

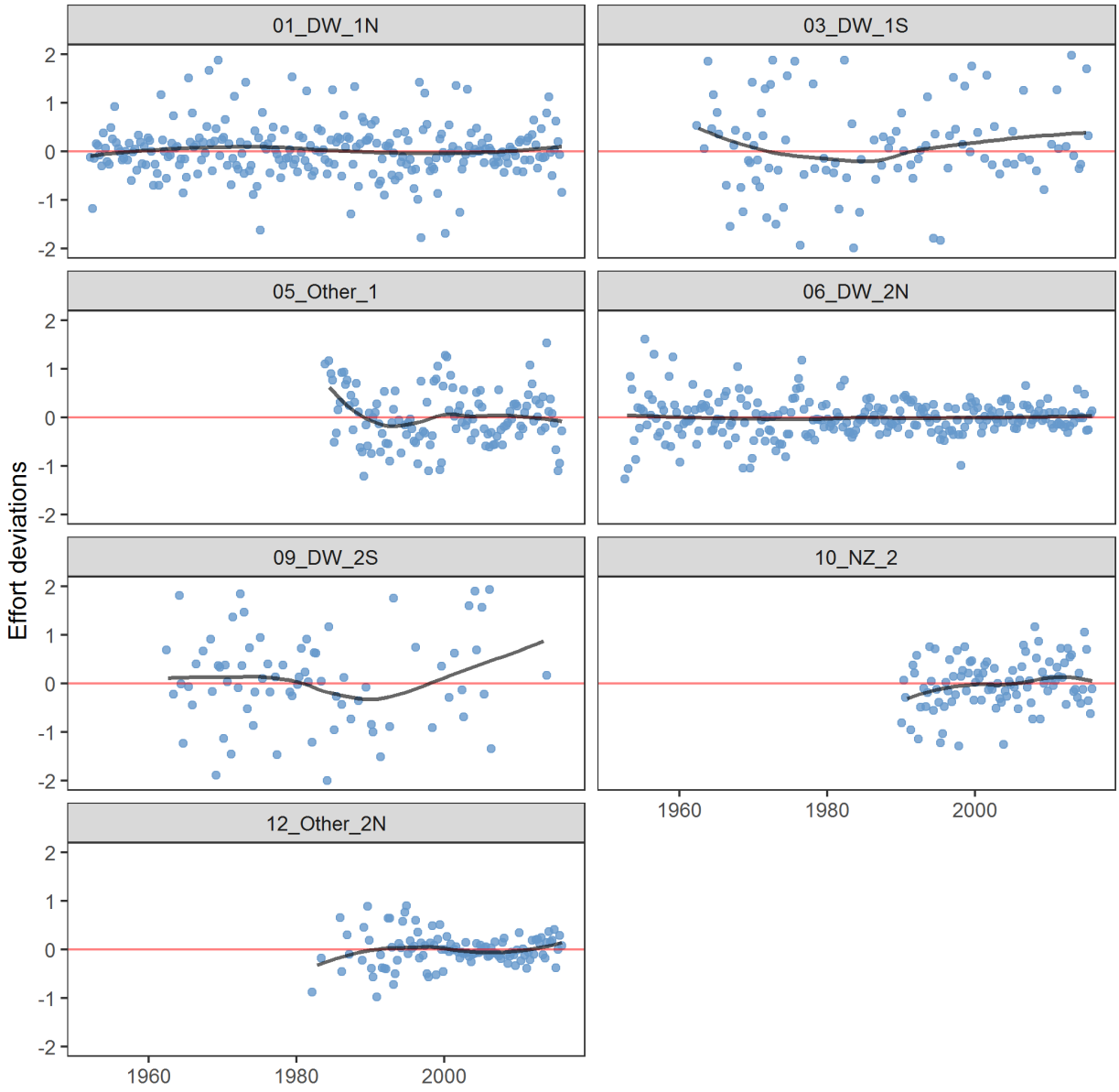

Figure 15. Effort deviations by time period for each of the fisheries that did not receive standardized CPUE indices in the diagnostic case model. The dark line represents a lowess smoothed fit to the effort deviations.

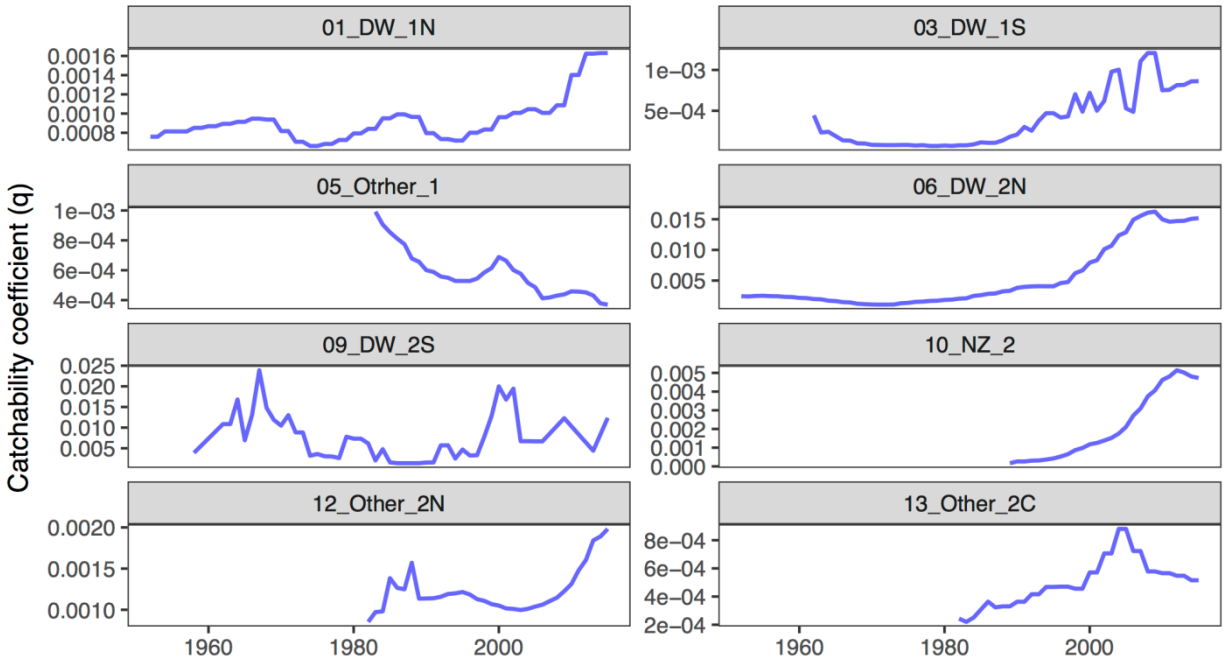

Figure 16. Estimated time series of catchability for those fisheries assumed to have random walk in those parameters. Values shown are the annual means which removes seasonal variability.

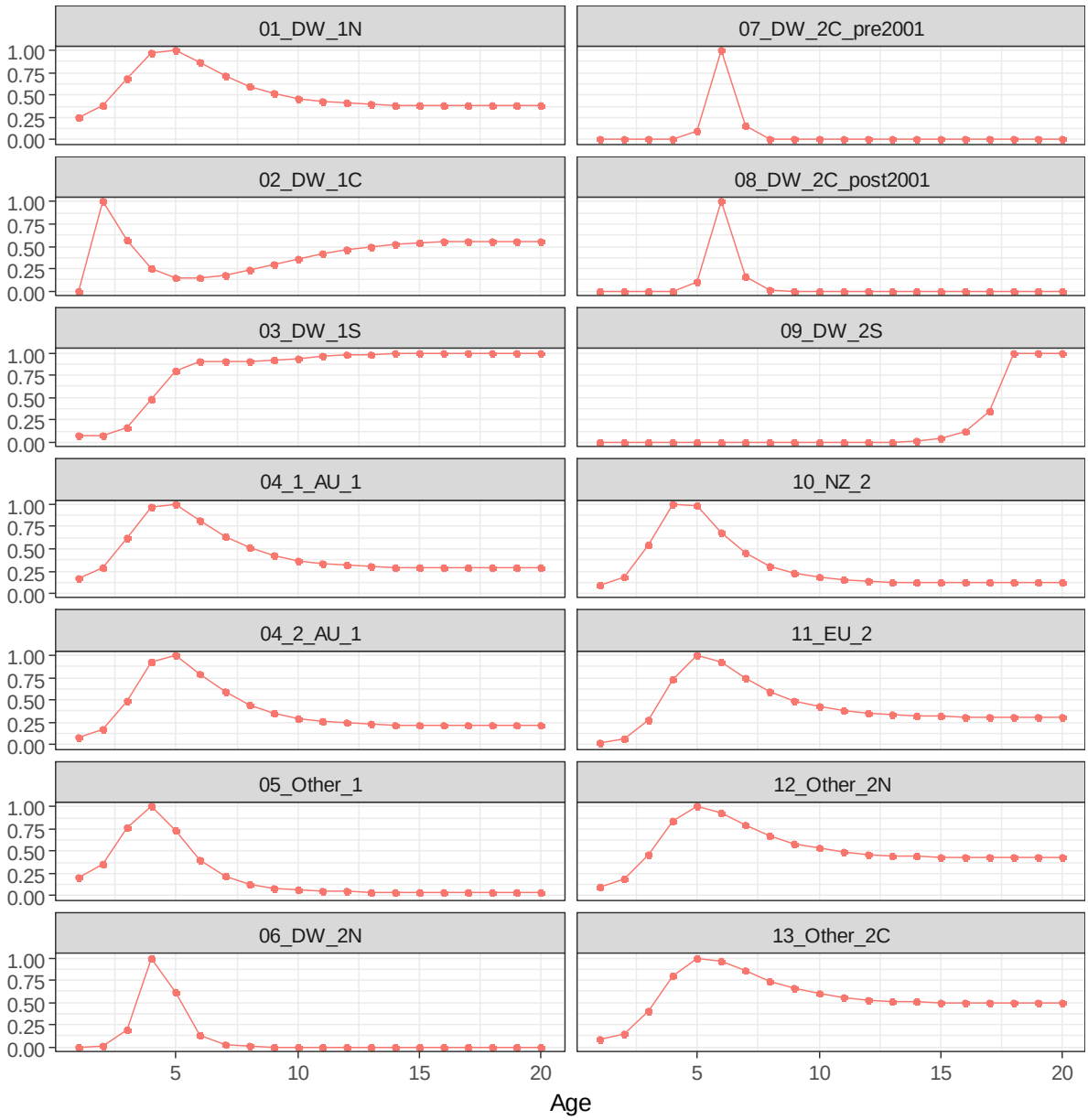

Figure 17. Estimated age-specific selectivity by fishery for the diagnostic case model. 04_1_AU_1 and 04_2_AU_1 refer to 04_AU_1 prior to 2008 and after 2008 respectively.

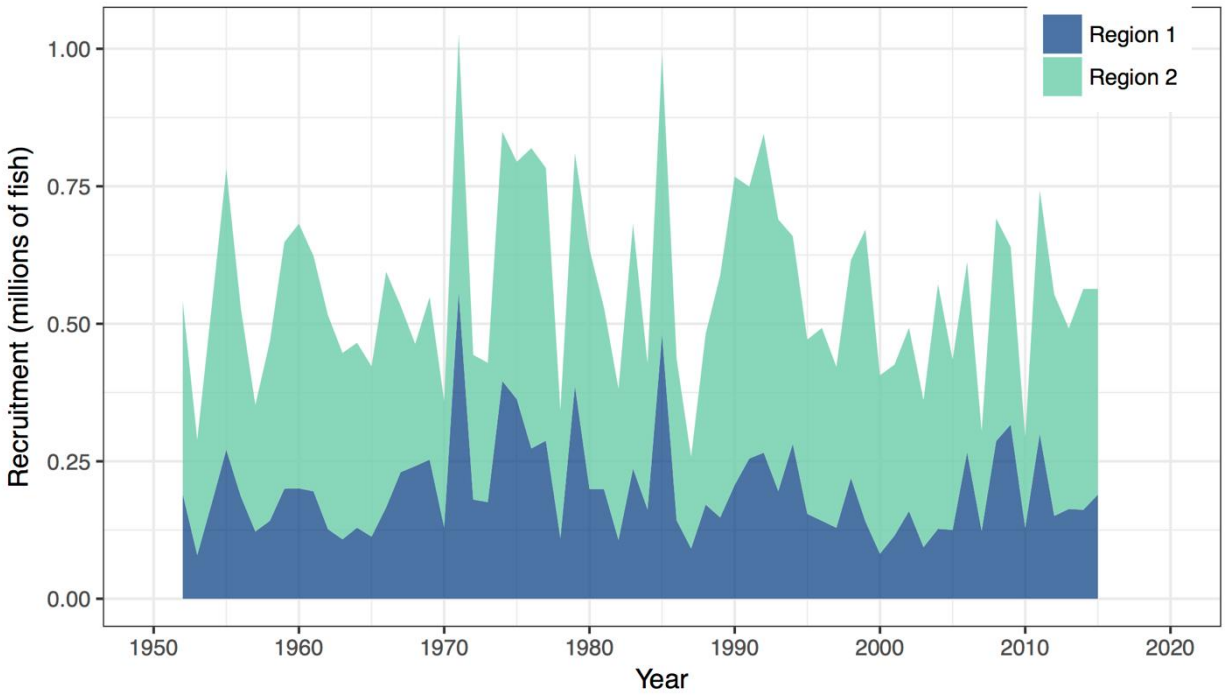

Figure 18. Estimated recruitment by year by model region for the diagnostic case model.

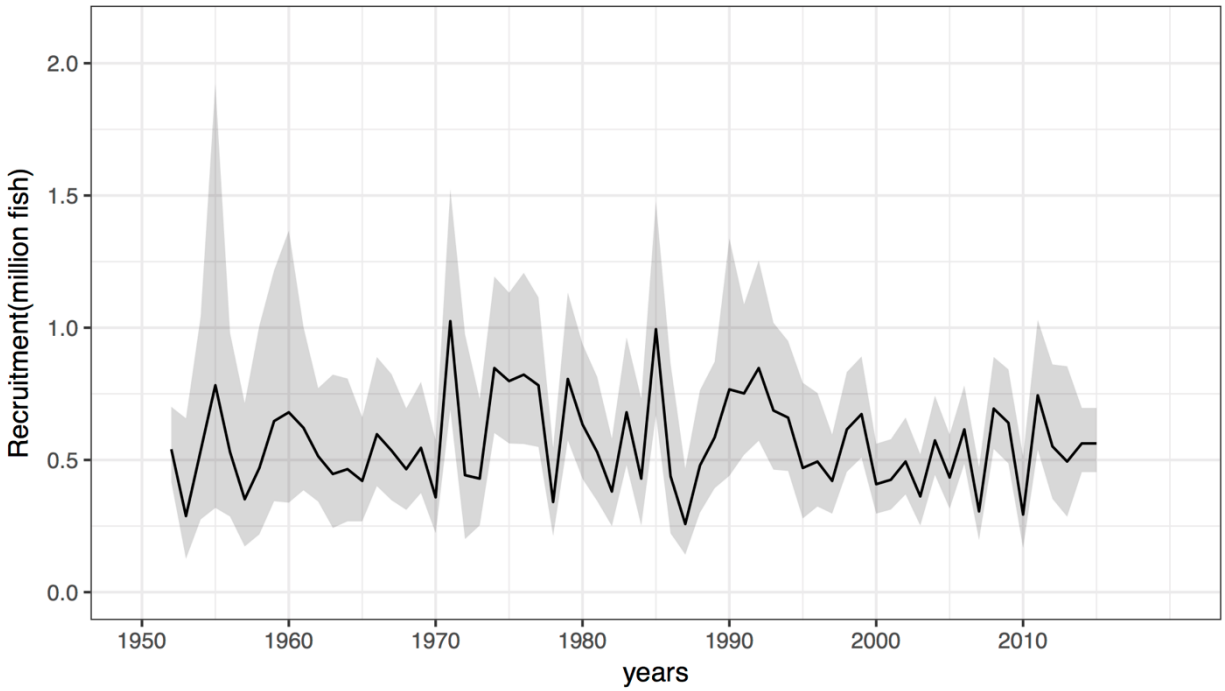

Figure 19. Estimated overall recruitment by year (black line) with 95% asymptotic confidence limits (shaded area) for the diagnostic case model.

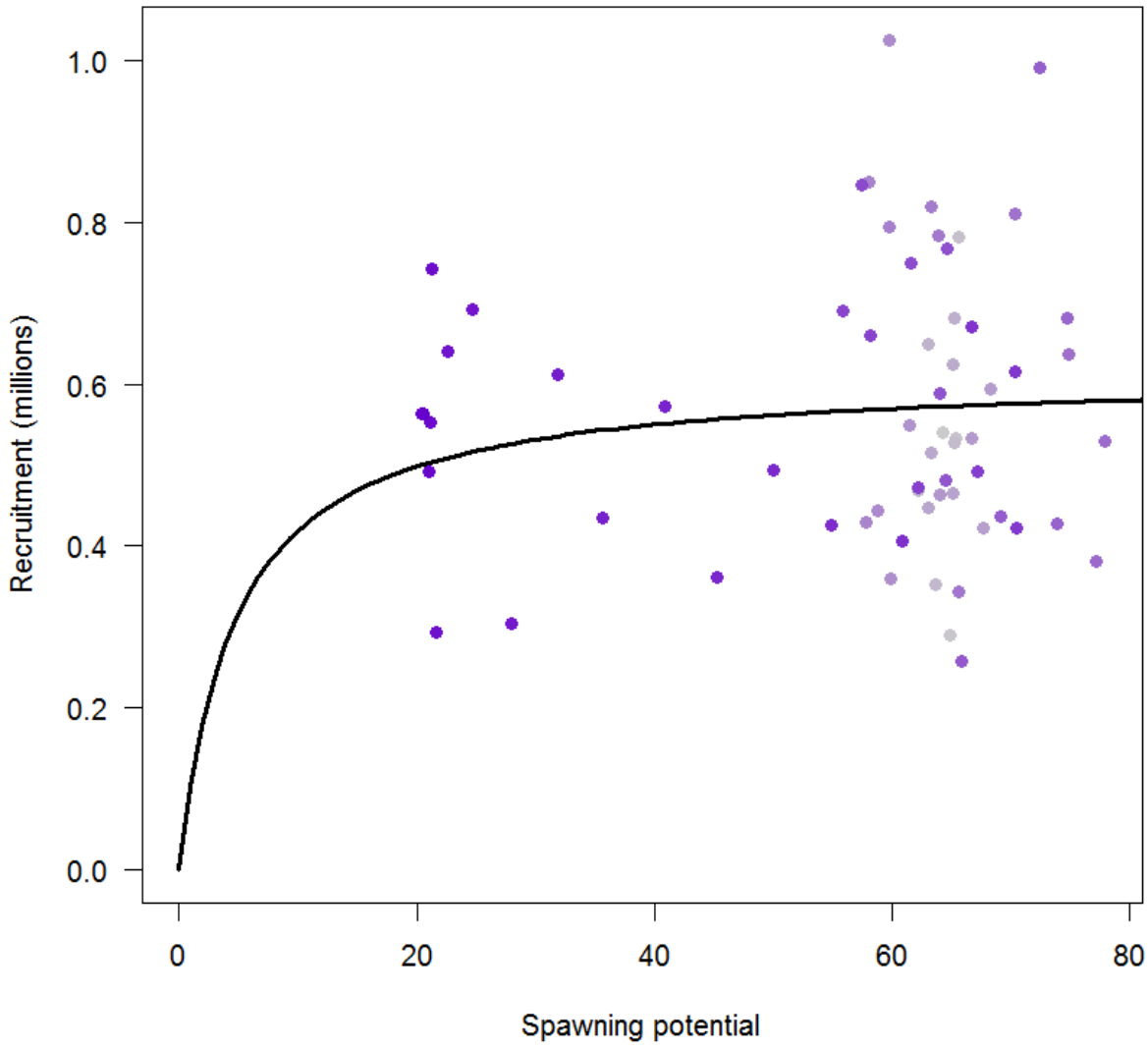

Figure 20. Estimated relationship between recruitment and spawning potential for the diagnostic case model. Points with lighter colours represent earlier years. Points with darker colours represent more recent years.

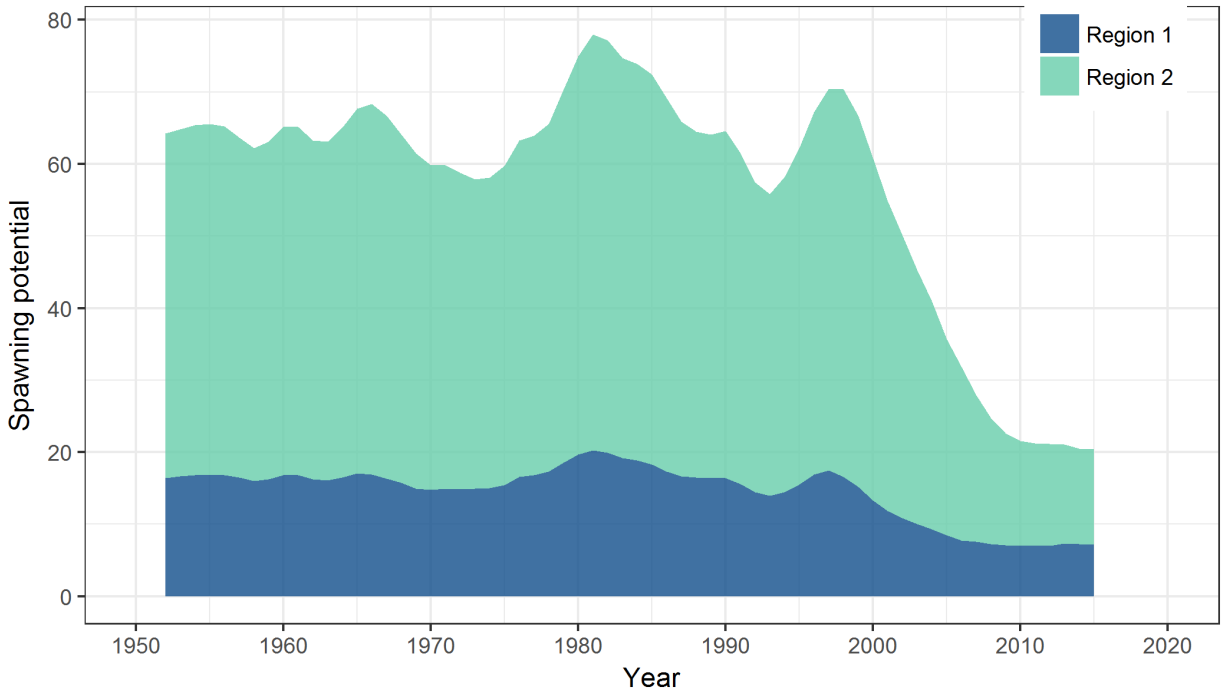

Figure 21. Estimated spawning potential by year by model region for the diagnostic case model.

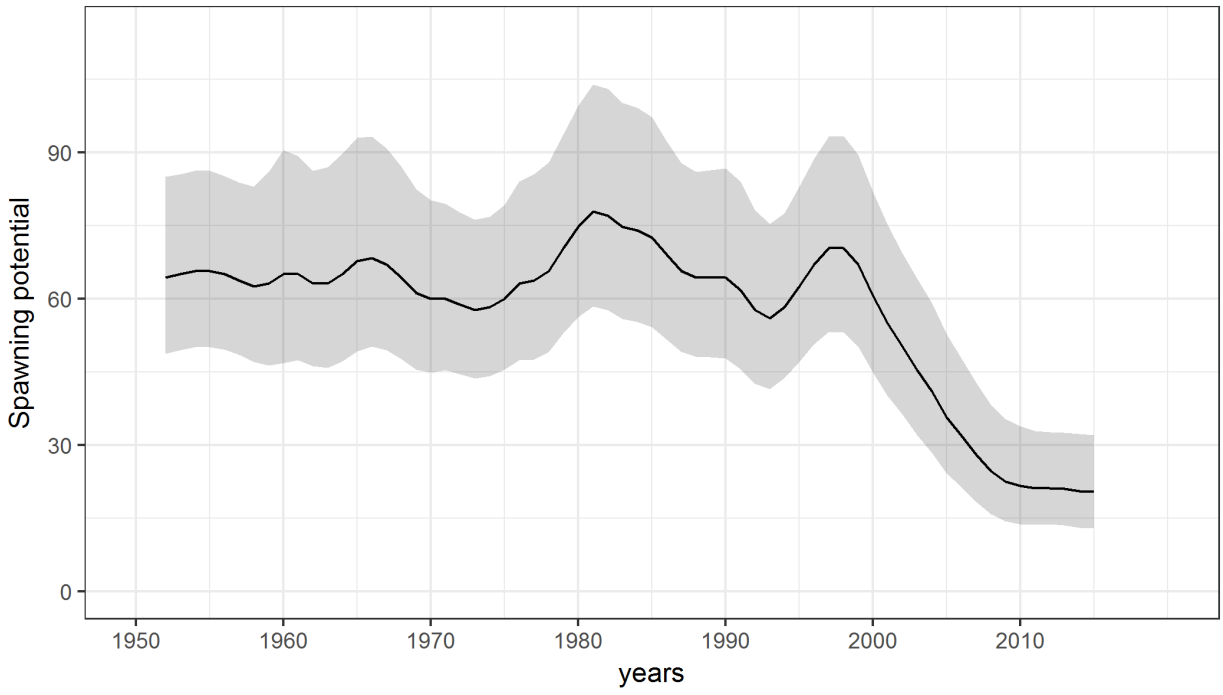

Figure 22. Estimated overall spawning potential by year (black line) with 95% asymptotic confidence limits (shaded area) for the diagnostic case model.

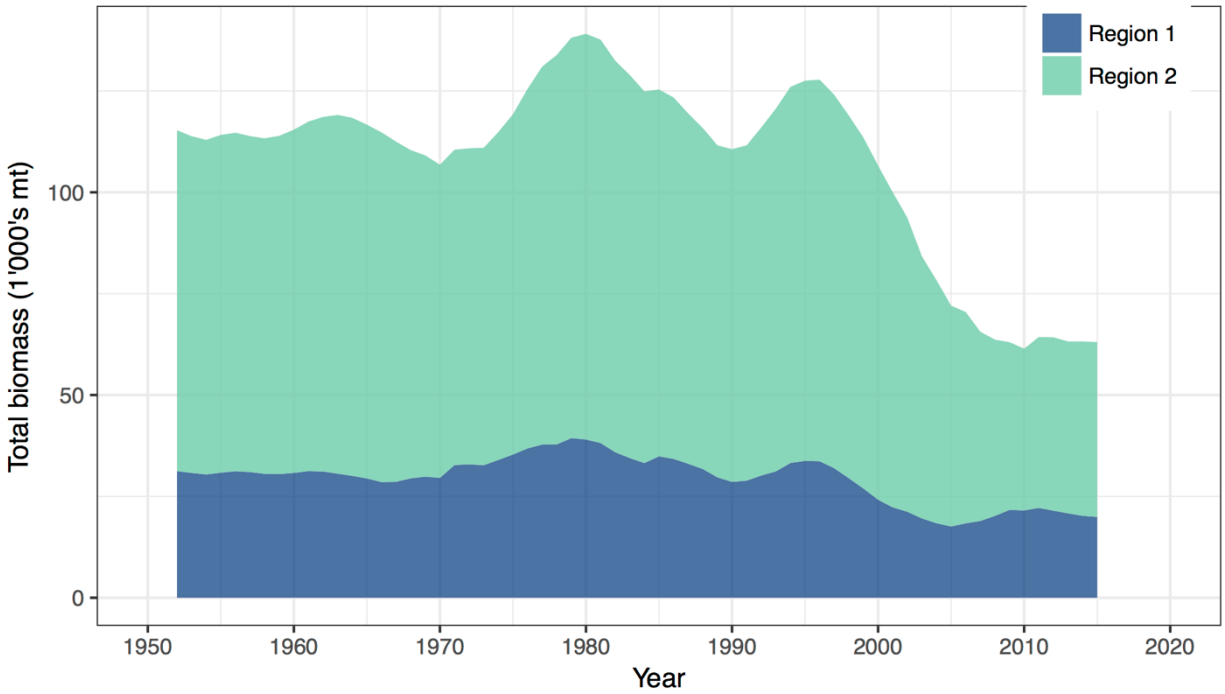

Figure 23. Estimated total biomass by year by model region for the diagnostic case model

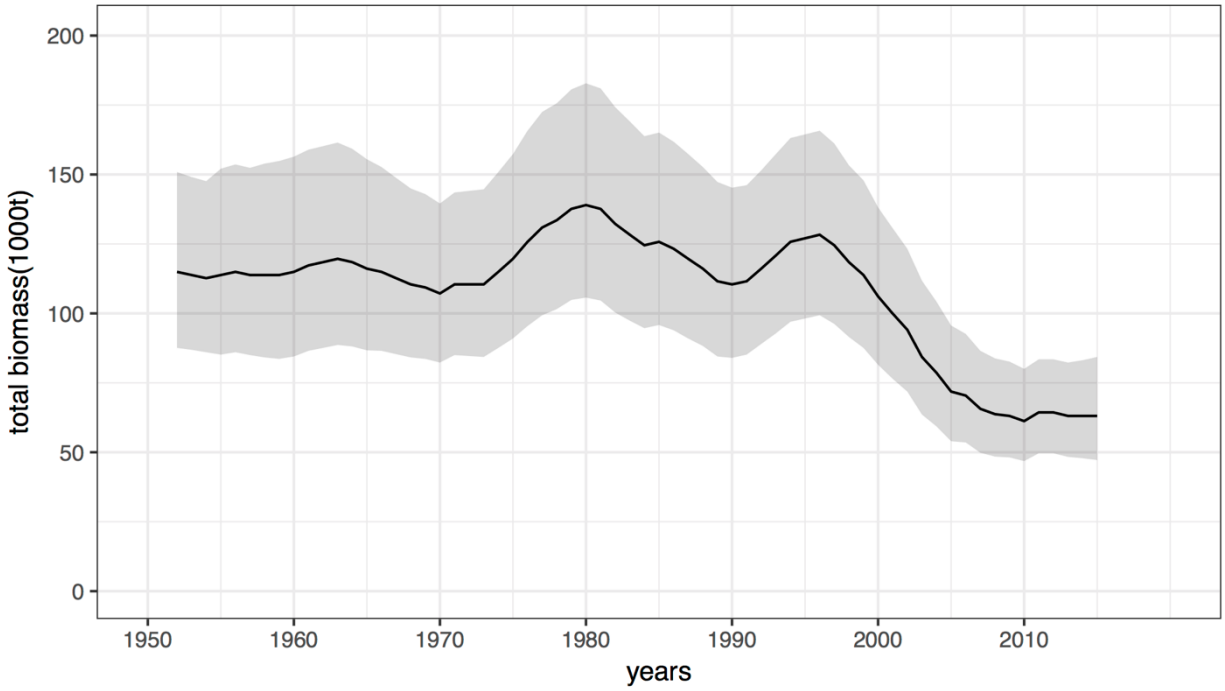

Figure 24. Estimated total biomass by year (black line) with 95% asymptotic confidence limits (shaded area) for the diagnostic case model.

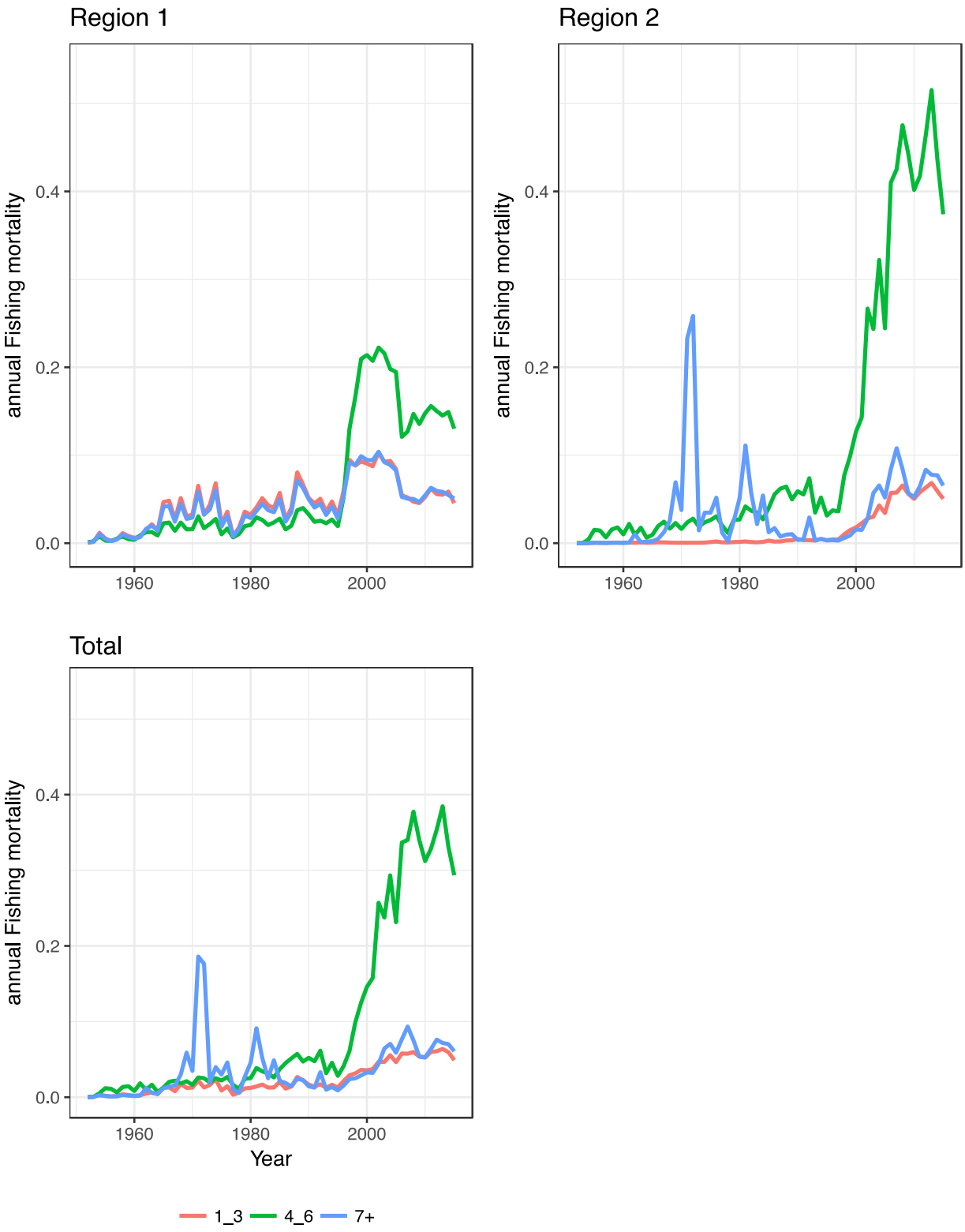

Figure 25. Estimated annual average fishing mortality at age by age groups (red; ages 1-3, green; ages 4-6 and blue; ages 7 and older), over time.

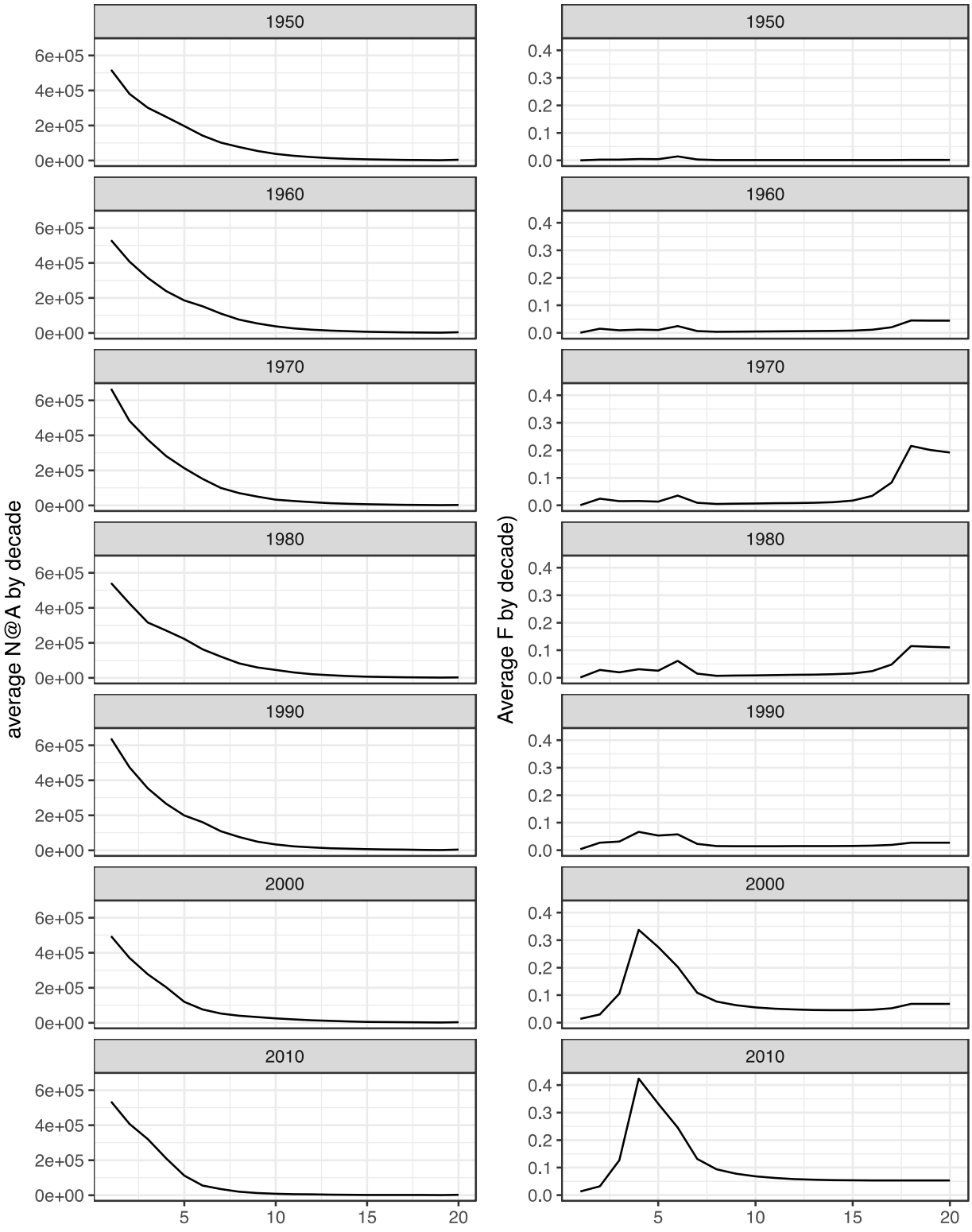

Figure 26. Estimated proportion of the population at age (left panels) and fishing mortality at age (right panels), at decadal intervals, for the diagnostic case model.

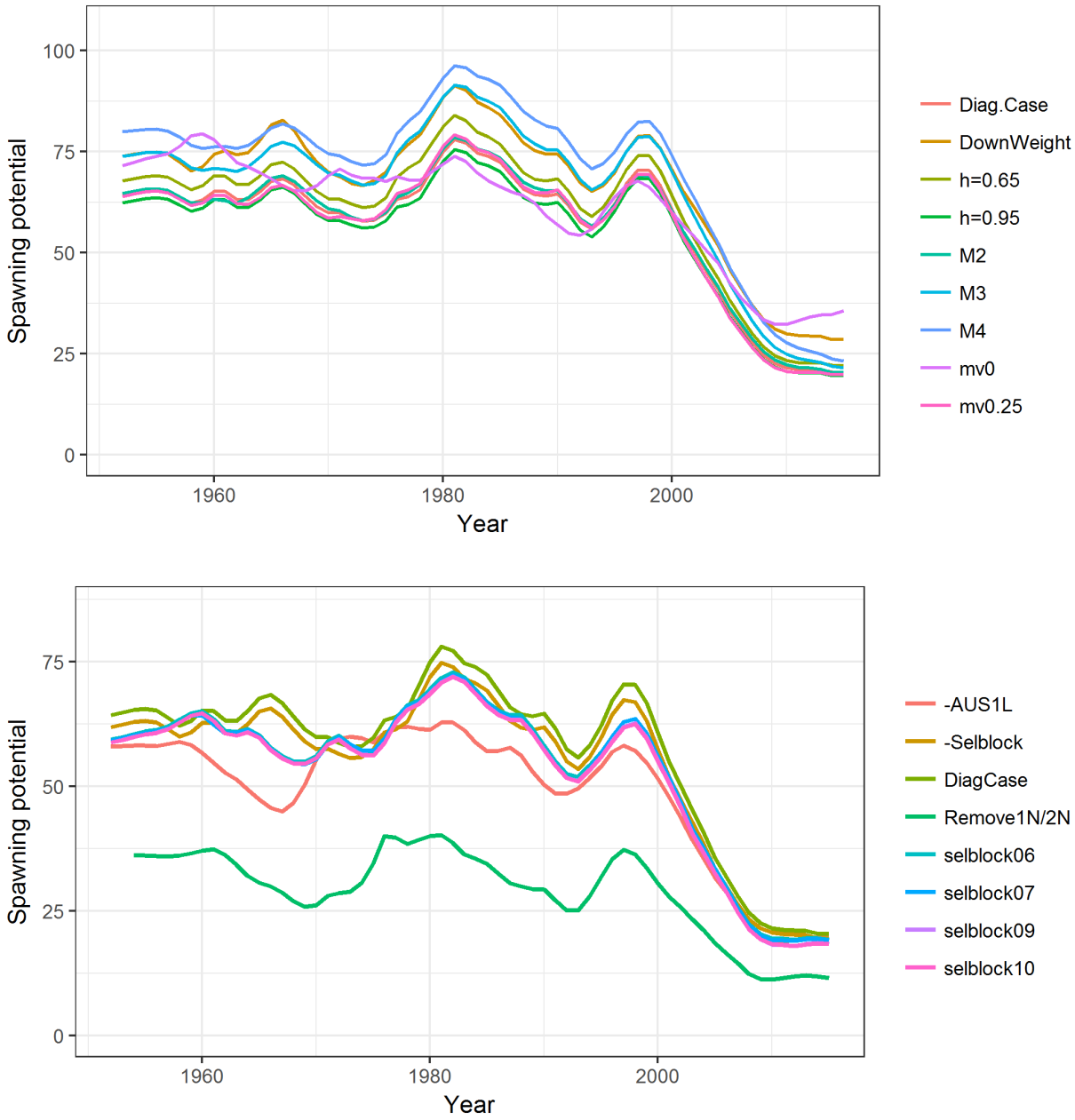

Figure 27. Estimated spawning potential for each of the one-off sensitivity models (a, top) and other sensitivity models (b, bottom) investigated in the assessment.

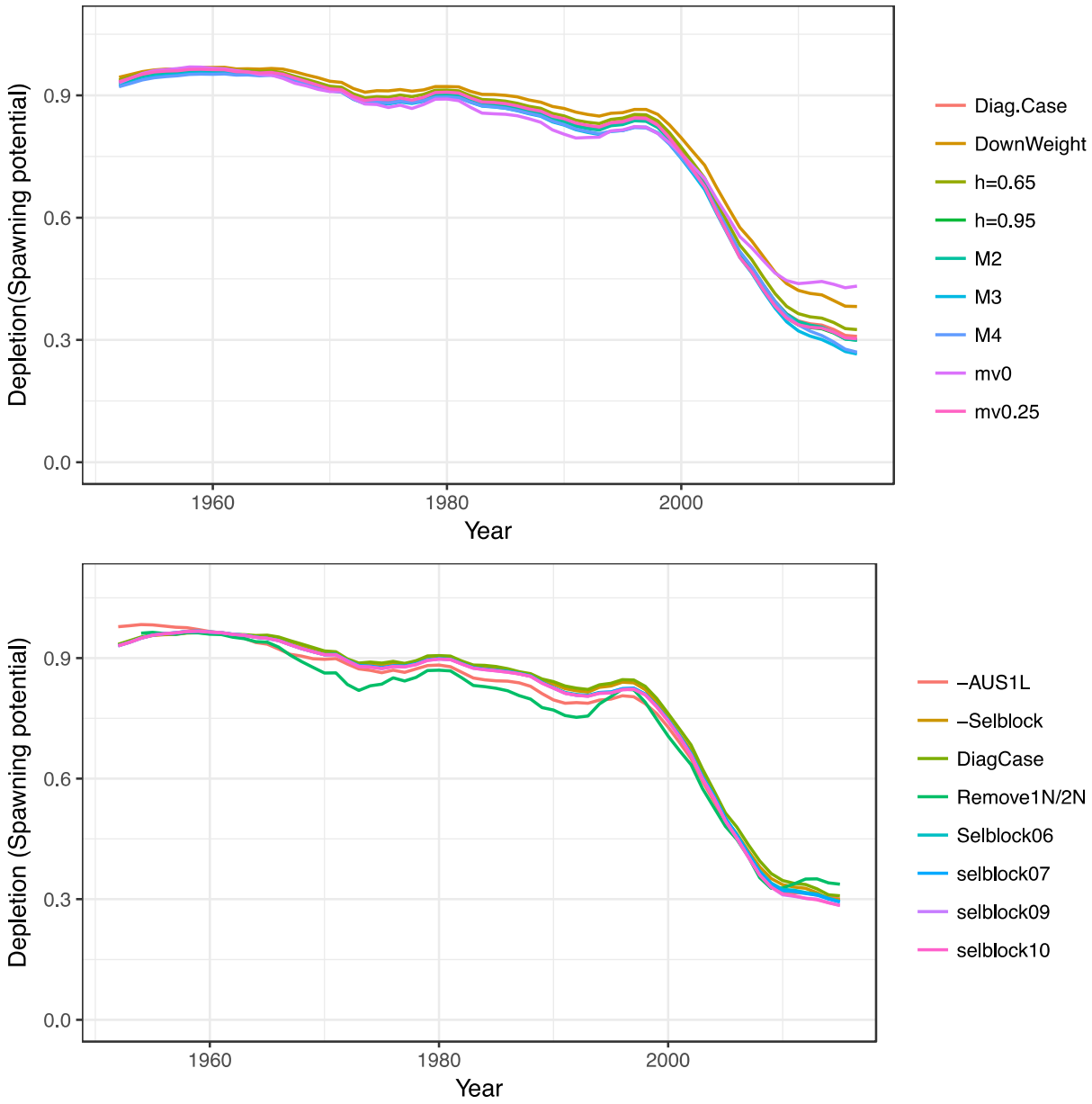

Figure 28. Estimated fishing depletion (of spawning potential) for each of the one-off sensitivity models (top) and other sensitivity models (bottom) investigated in the assessment.

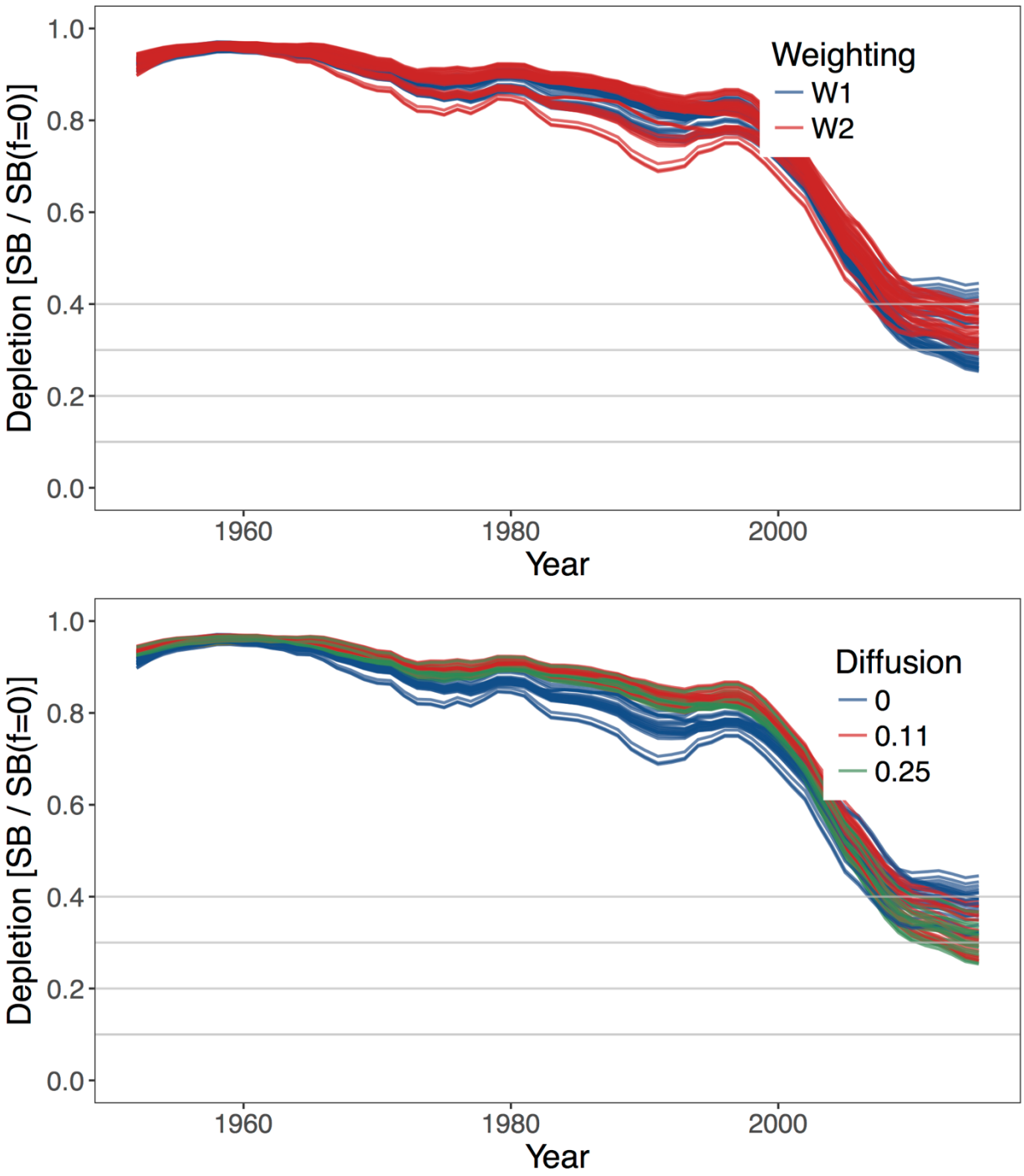

Figure 29. Plots showing the trajectory of fishing depletion (of spawning potential) for the model runs included in the structural uncertainty grid by size weighting axis (top) and diffusion (bottom).

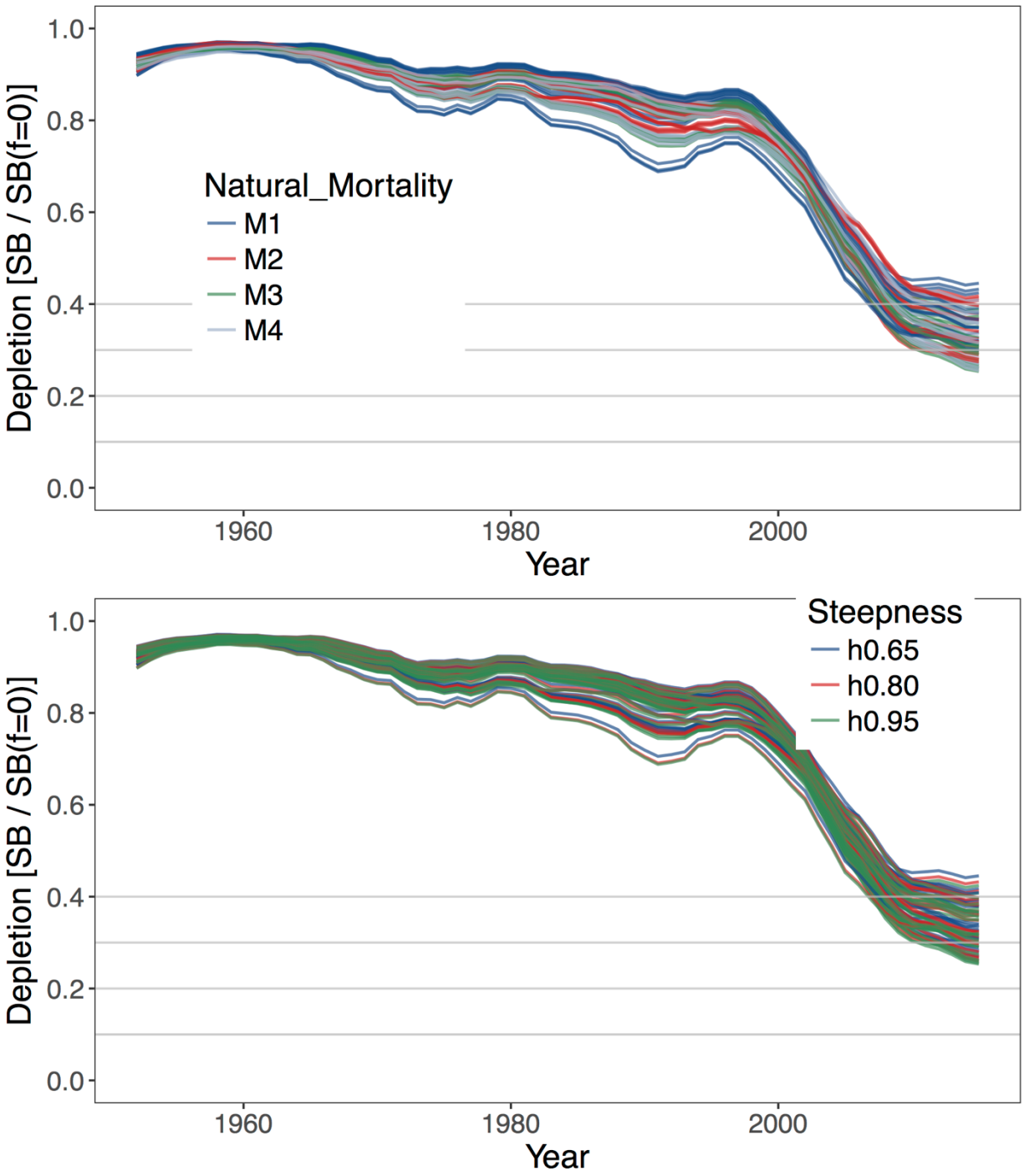

Figure 29 (cont). Plots showing the trajectory of fishing depletion (of spawning potential) for the model runs included in the structural uncertainty grid by natural mortality (top) and steepness (bottom).

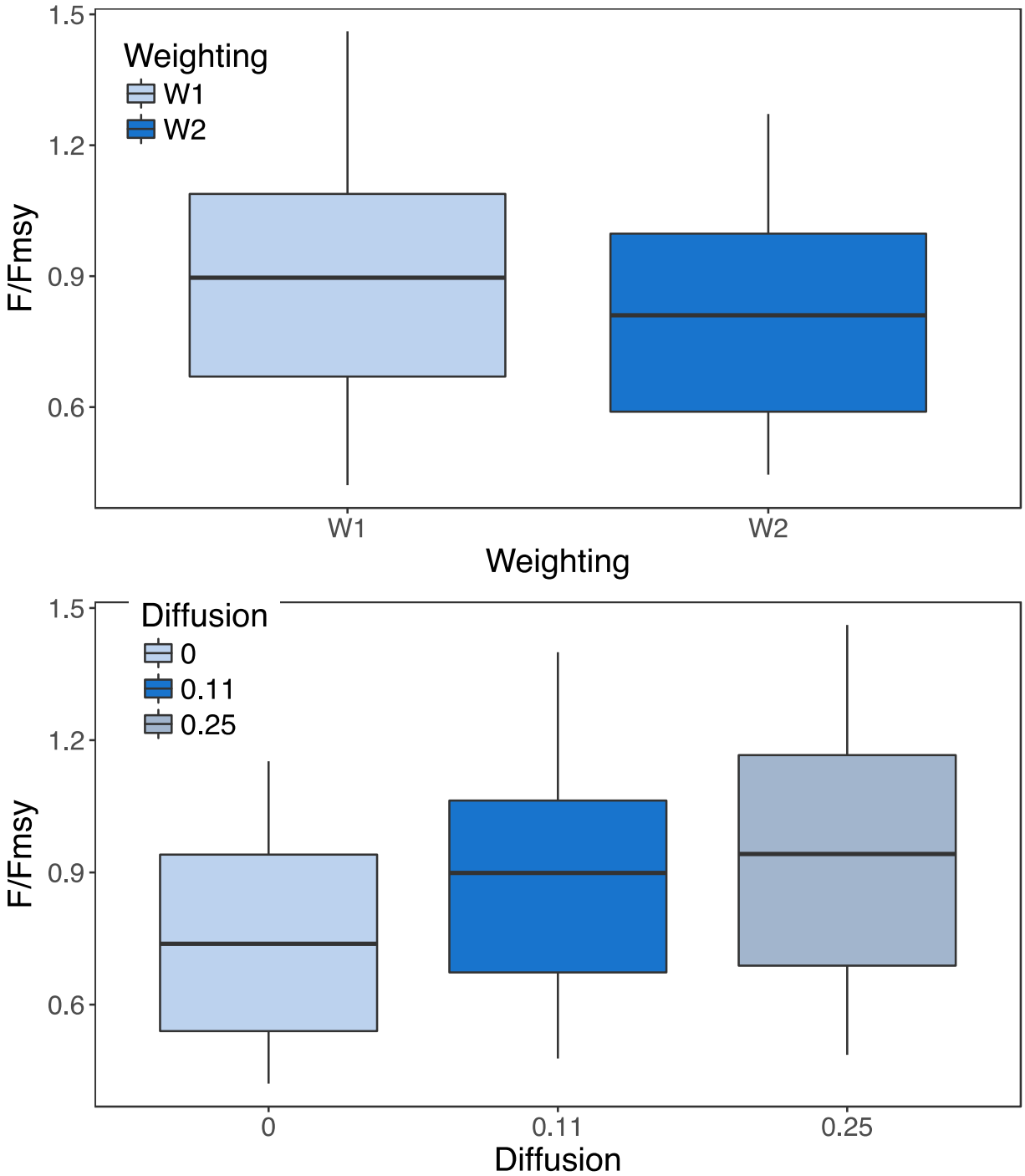

Figure 30. Boxplots summarizing the results of the structural uncertainty grid with respect to the fishing mortality reference points F_{recent}/F_{MSY} by size weighting axis (top) and diffusion (bottom).

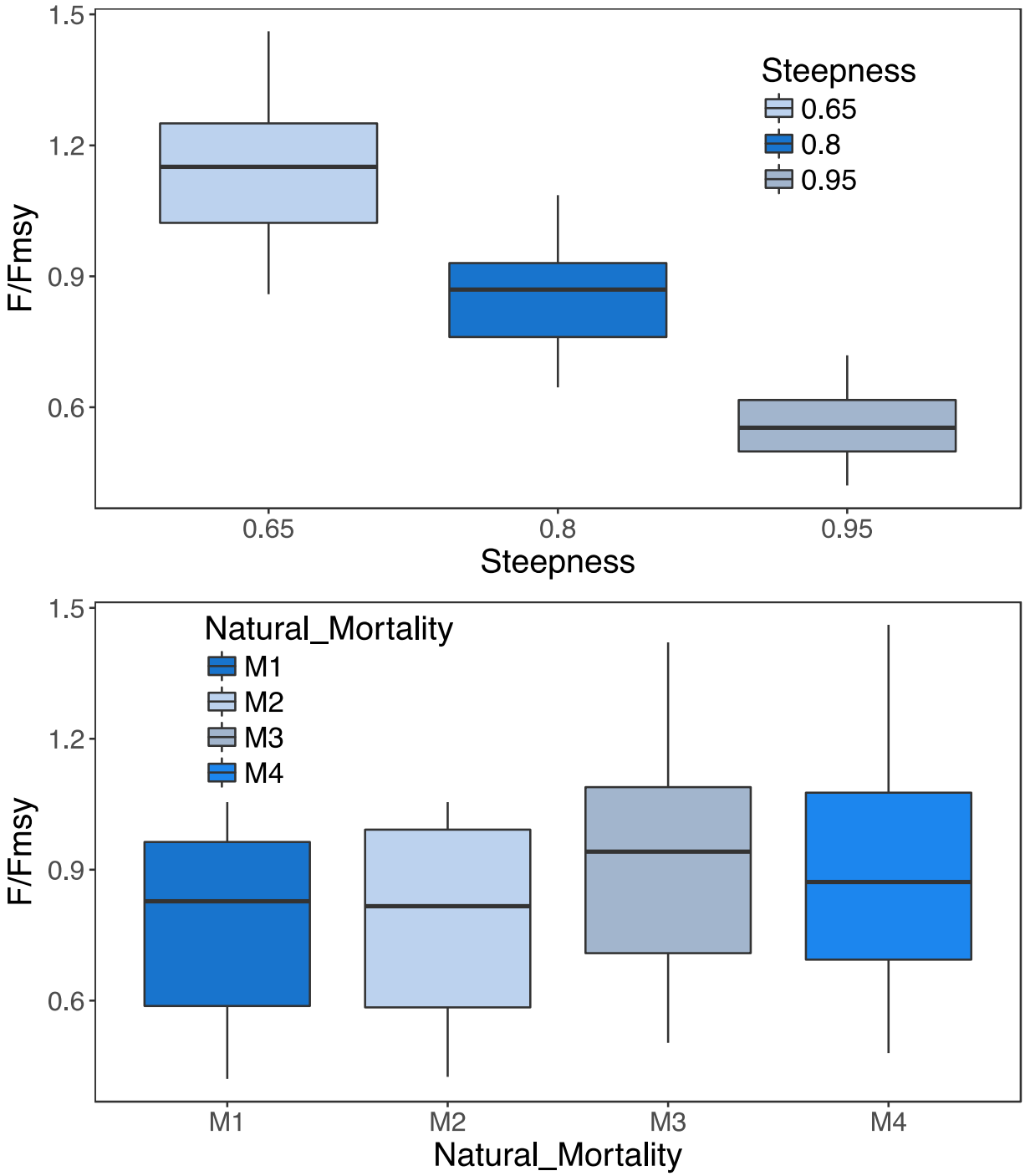

Figure 30 (cont). Boxplots summarizing the results of the structural uncertainty grid with respect to the fishing mortality reference points F_{recent}/F_{MSY} by Steepness (top) and by natural mortality (bottom).

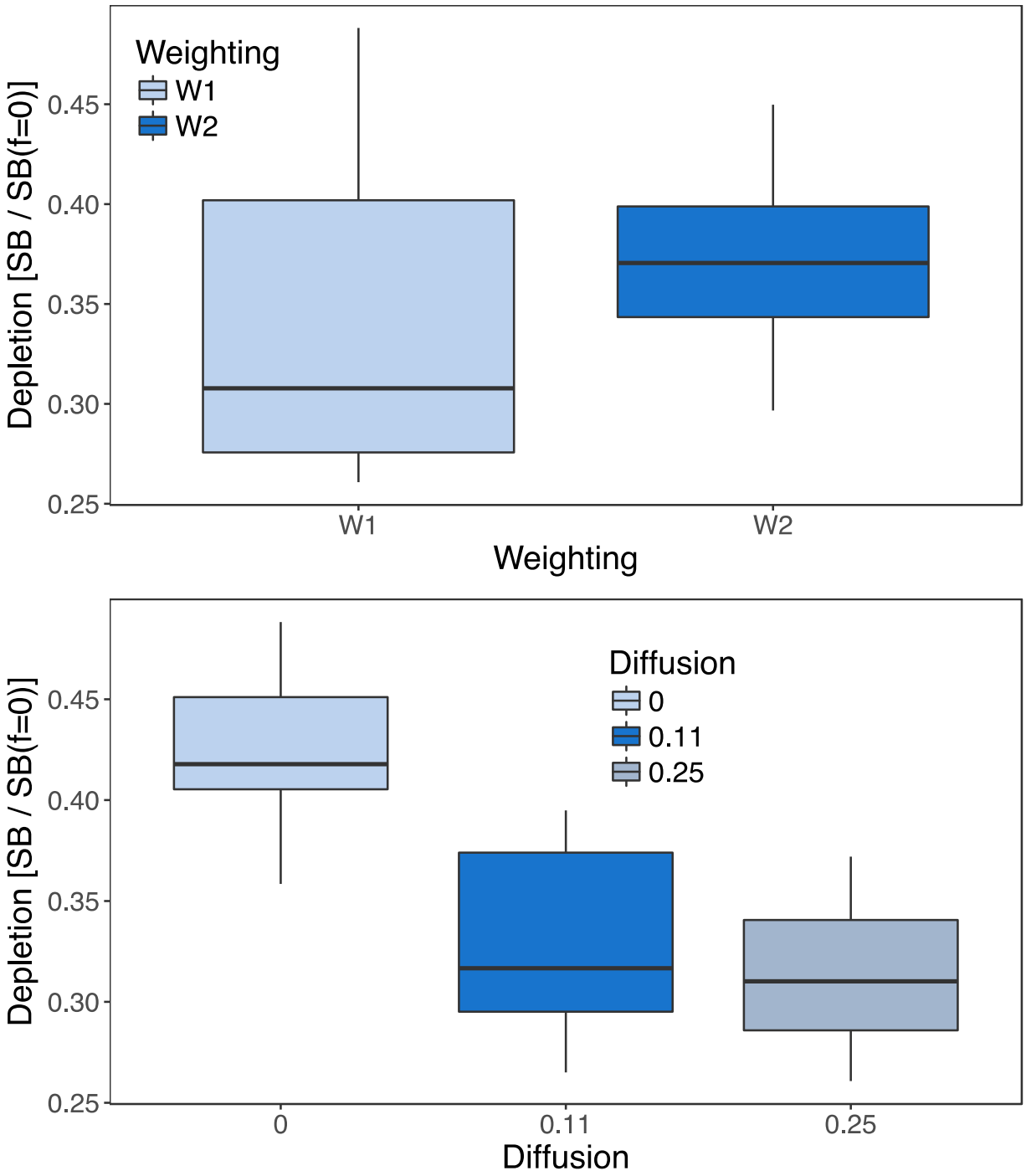

Figure 31. Boxplots summarizing the results of the structural uncertainty grid with respect to the spawning potential reference points by size weighting axis (top) and diffusion (bottom).

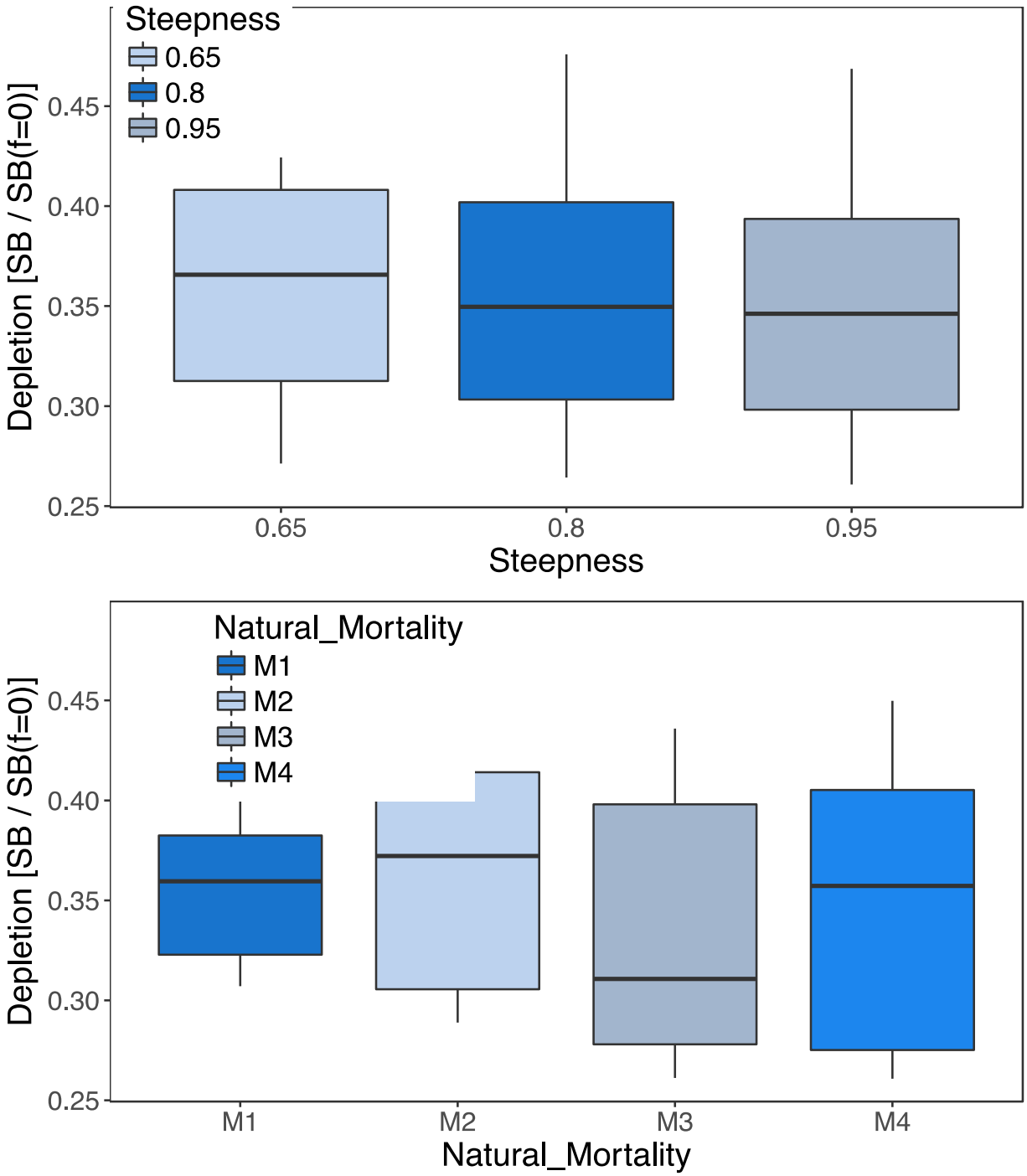

Figure 31 (cont). Boxplots summarizing the results of the structural uncertainty grid with respect to the spawning potential reference points by steepness (top) and by natural mortality (bottom).

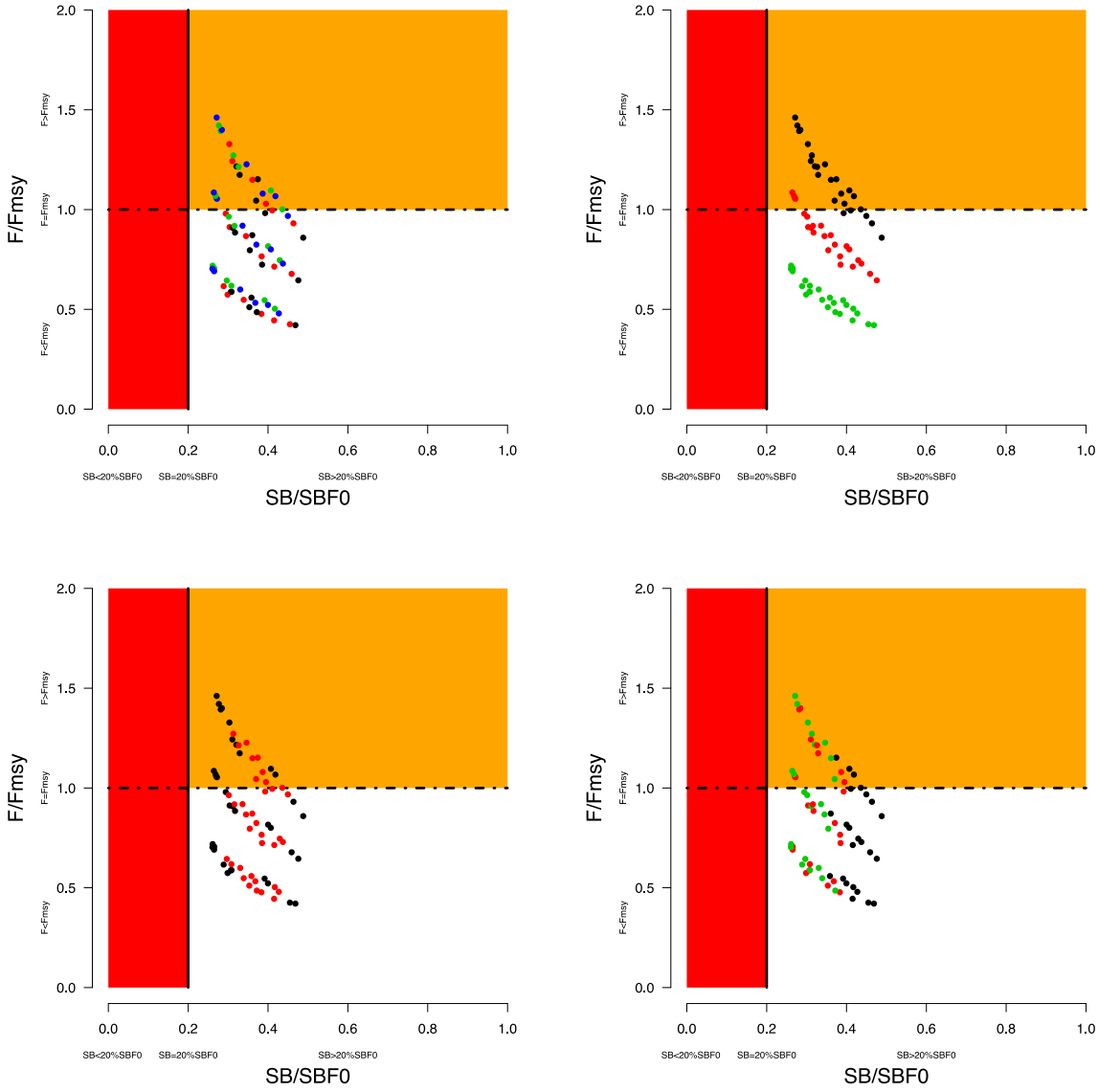

Figure 32. Majuro plots summarizing the results for each of the models in the structural uncertainty grid by each axis in grid. From top left, Natural mortality (top left), Steepness (top right), Weighting (bottom left), Diffusion (bottom right). The plots represent status in terms of spawning potential depletion and fishing mortality. The red zone represents spawning potential (SB_{latest}) levels lower than 20%$SB_{F=0}$. The orange zone is for fishing mortality greater than F_{MSY} (F_{MSY} is marked with the black line). The points represent $SB_{latest}/SB_{F=0}$ for each model run. Colors of circle in each panel indicates for Natural mortality (M1(black), M2(red), M3(green) and M4(blue))), for Steepness (0.65(black), 0.80(red), 0.95(green)), Weighting (W1(black) and W2(red)), Diffusion (mv0(black), mv11(red), mv25(green)).

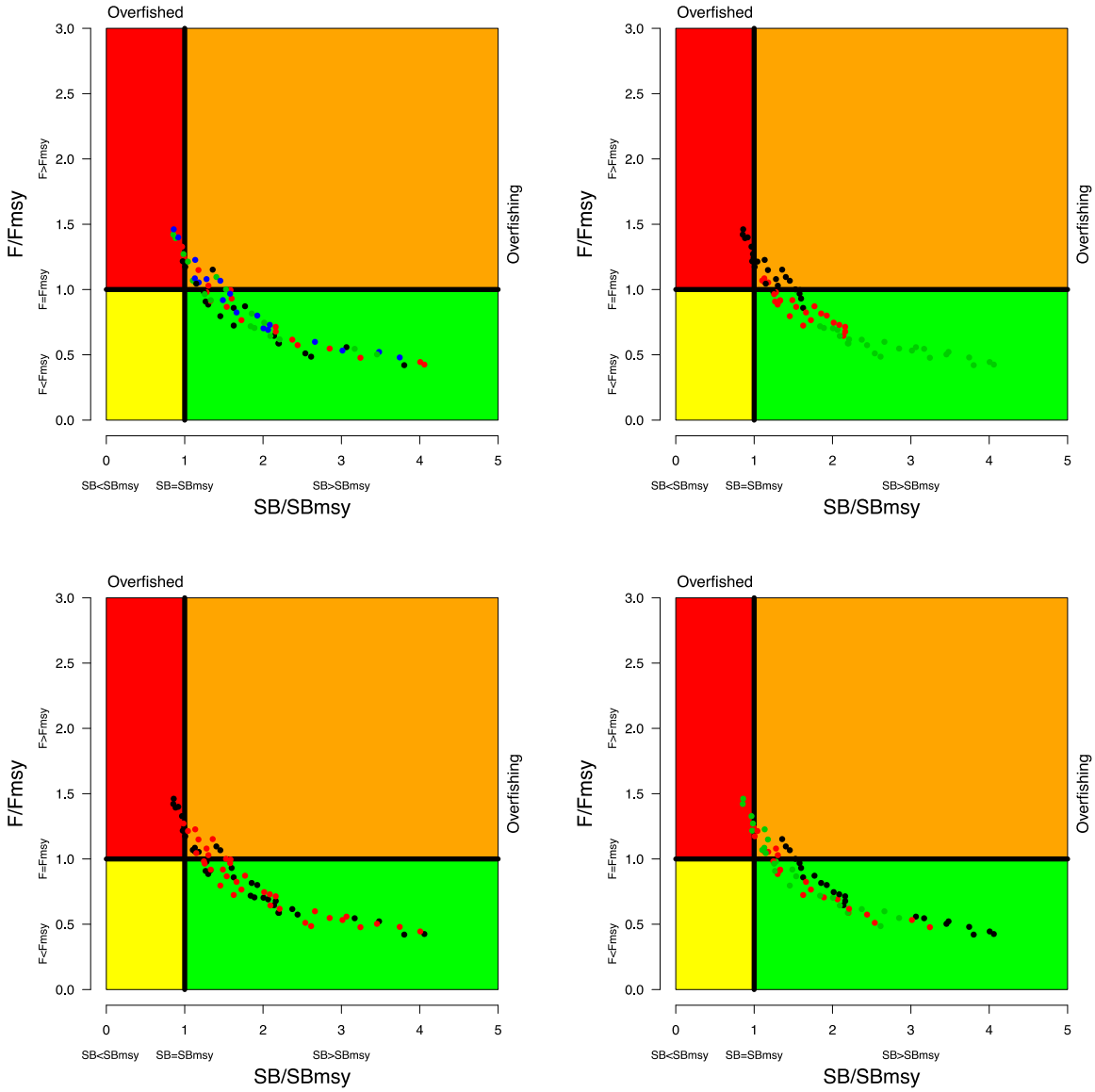

Figure 33. Kobe plots summarizing the results for each of the models in the structural uncertainty grid by each axis in grid. From top left, Natural mortality (top left), Steepness (top right), Weighting (bottom left), Diffusion (bottom right). The points represent $SB_{latest}/SB_{F=0}$ for each model run. Colors of circle in each panel indicates for Natural mortality (M1(black), M2(red), M3(green) and M4(blue)), for Steepness (0.65(black), 0.80(red), 0.95(green)), Weighting (W1(black) and W2(red)), Diffusion (mv0(black), mv11(red), mv25(green)).

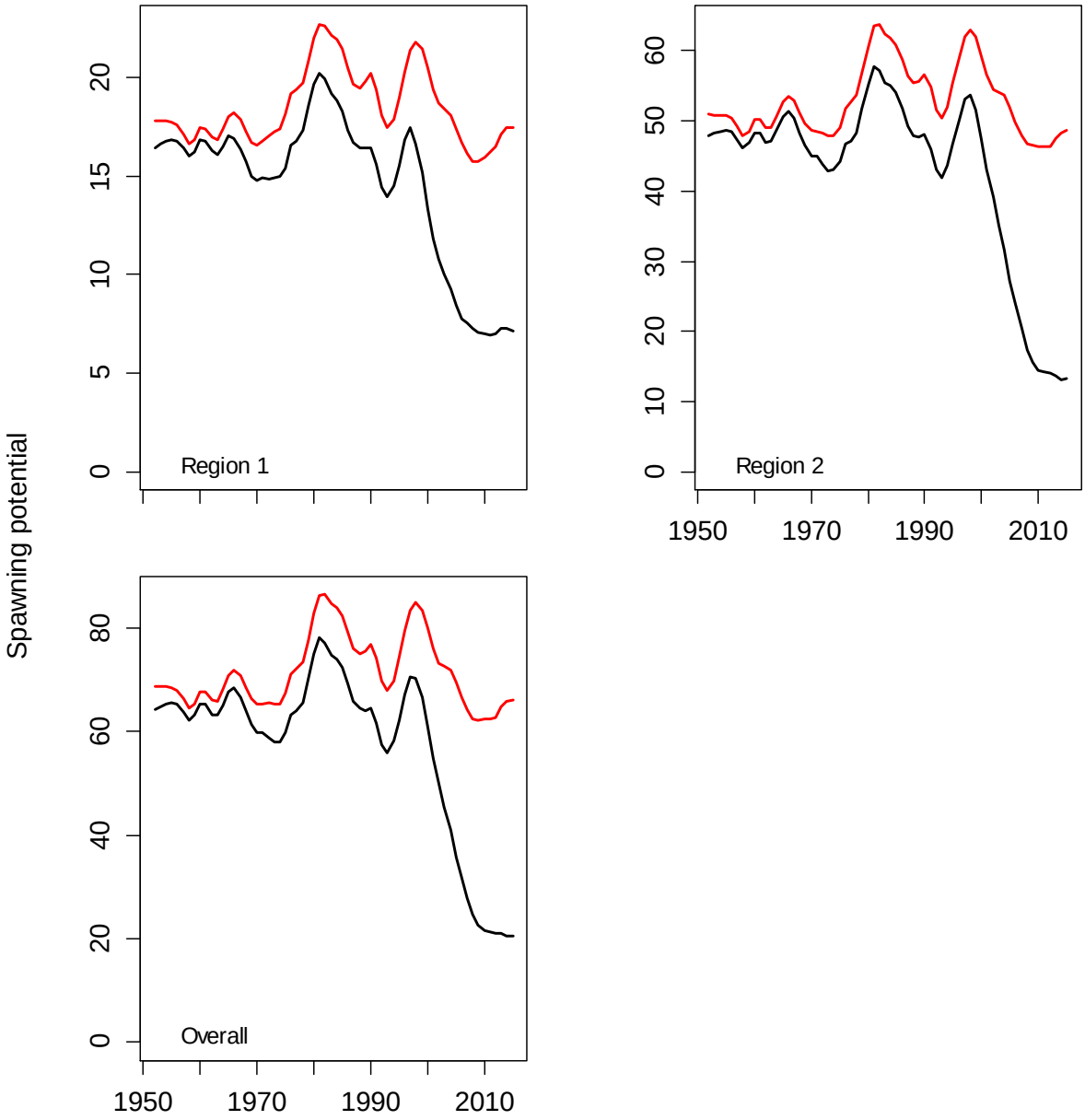

Figure 34. Comparison of the estimated annual spawning potential trajectories (black lines) with those trajectories would have occurred in the absence of fishing (upper red lines)

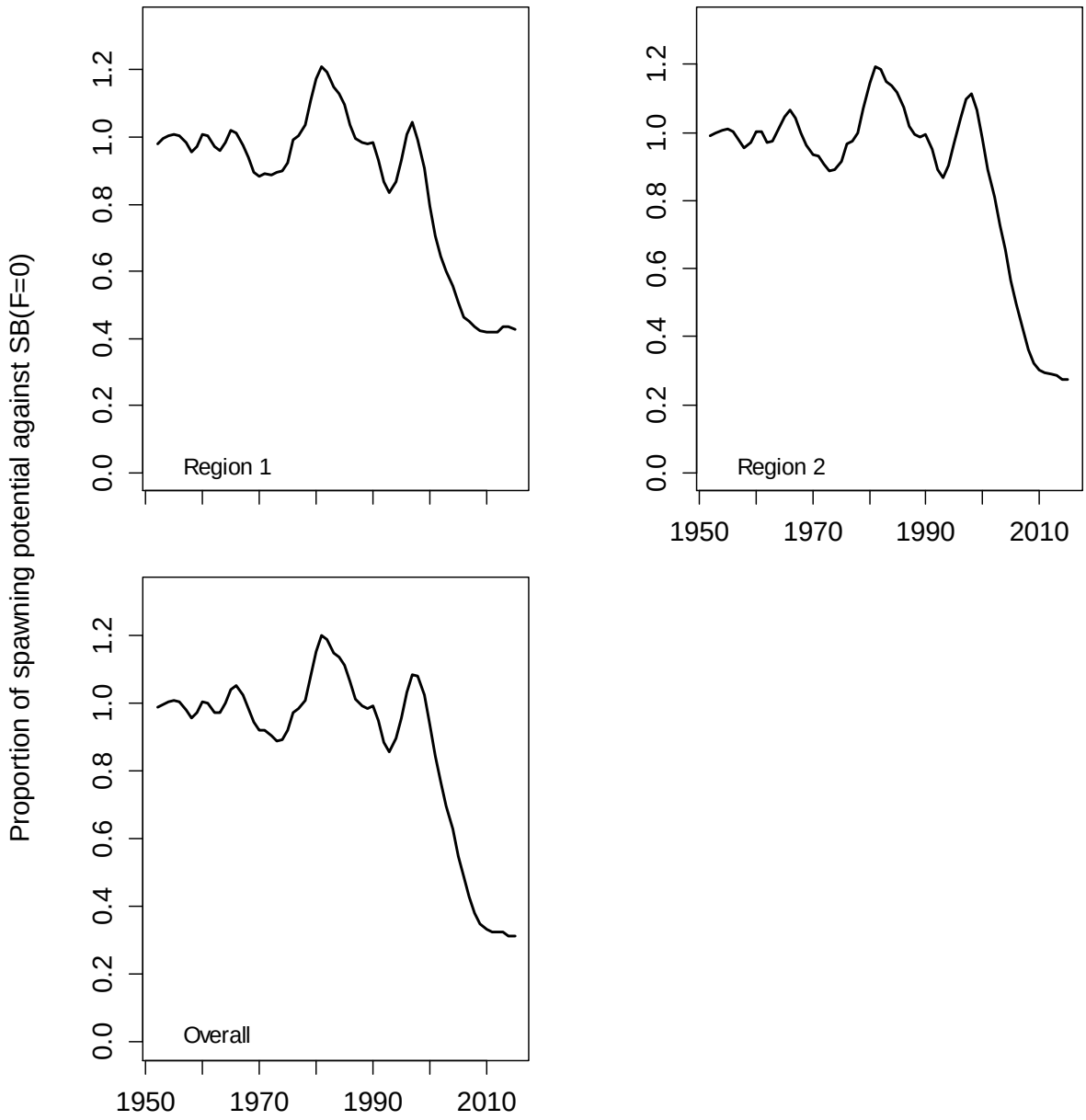

Figure 35. Ratio of exploited to unexploited spawning potential, $SB_{latest}/SB_{F=0}$ for each region and overall for the diagnostic case model.

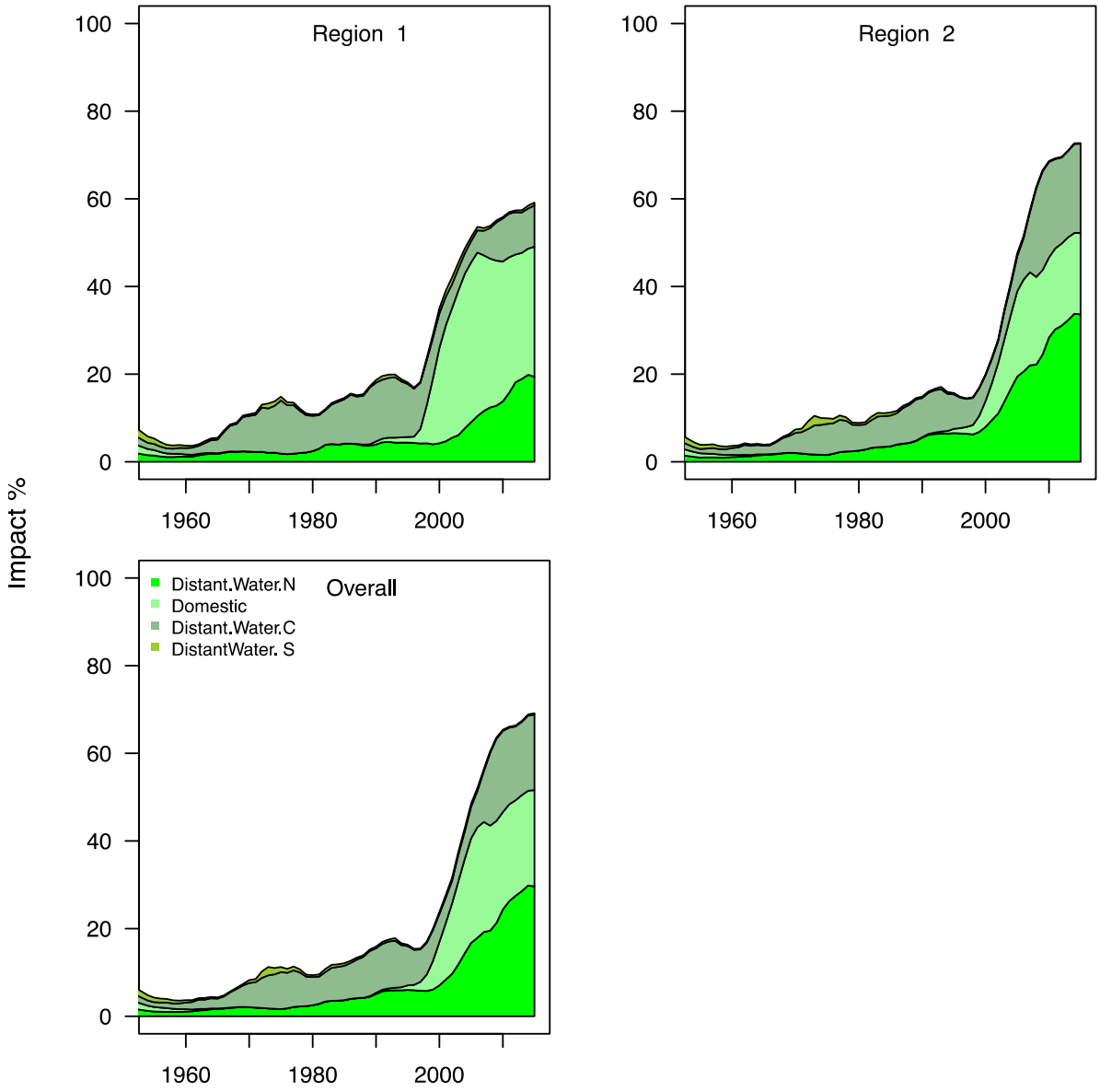

Figure 36. Estimates of reduction in spawning potential due to fishing (fishery impact=1-$SB_{latest}/SB_{F=0}$) by region, and overall regions, attributed to distant water fisheries in northern, central and southern sub-regions and domestic fisheries.

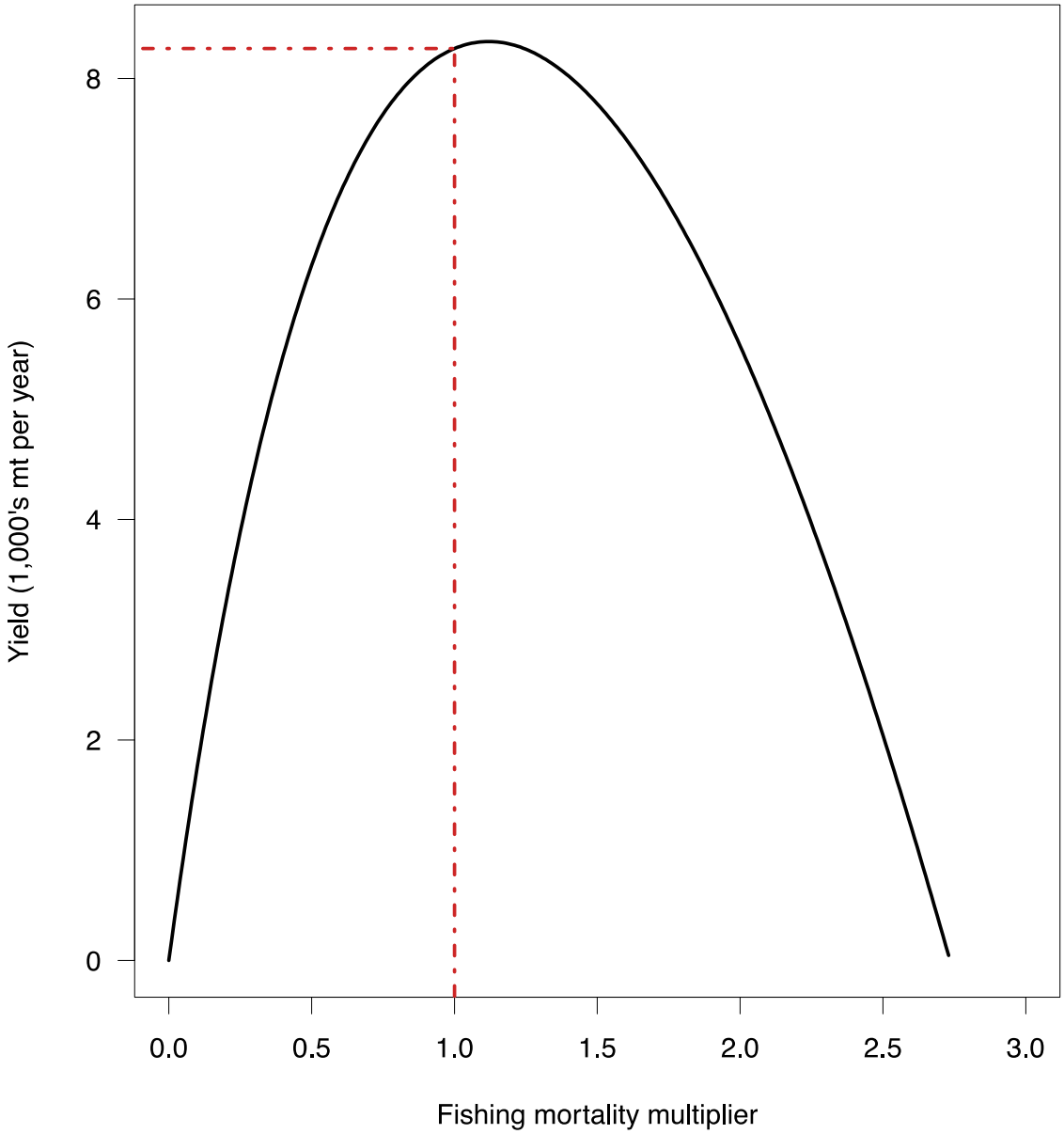

Figure 37. Estimated yield as a function of fishing mortality multiplier for the diagnostic case model.

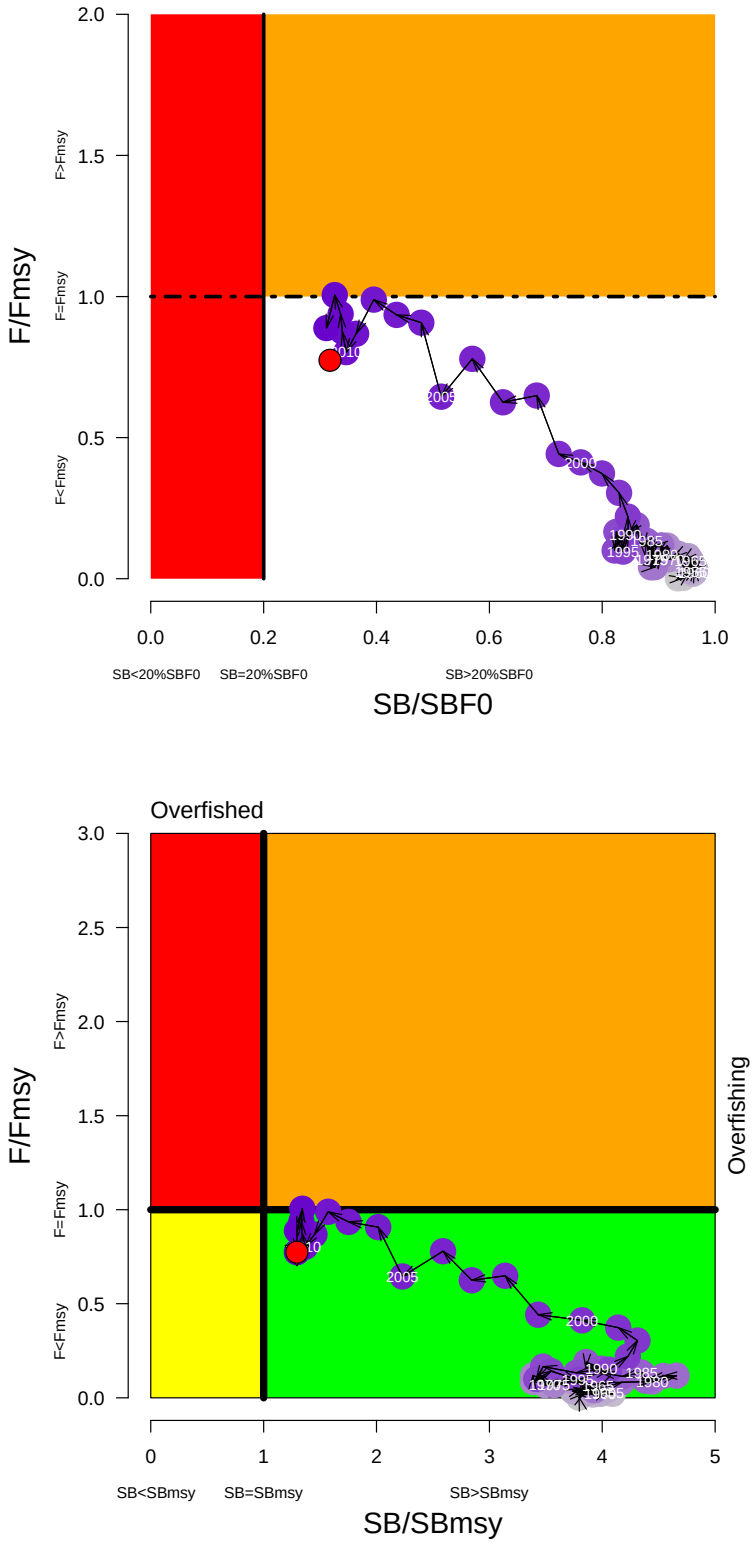

Figure 38. Majuro plot (a) for the diagnostic case model representing stock status in terms of spawning potential depletion and fishing mortality. The red zone represents spawning potential levels lower than 20% of $SB_{recent}/SB_{F=0}$ which is marked with the solid black line. The orange region is for fishing mortality greater than F_{MSY} (marked with the black dashed line). The green pink circle is $SB_{recent}/SB_{F=0}$ which are both detailed in **Table 7**. The equivalent Kobe plot is provided for comparison for the DiagCase.

10. Appendix

10.1. Likelihood profile

The approach for calculating a likelihood profile of the derived parameter, mean total biomass over the assessment period (to represent scale of the stock size) is outlined in Section 4.2.4. The profile was constructed by sequentially moving from the MLE in either direction while progressively penalising the mean total biomass at increasingly high and low values until it was determined that the minimum value had been reached for all data components. The profile reflects the loss of fit over all the data, i.e. the overall objective function value, and the individual data components, caused by changing the population size from that of the maximum likelihood estimated value. The change in likelihood relative to the maximum likelihood estimate is shown for the total likelihood (black line) and the individual data components (coloured lines) in Figure-Appendix 1 and displays significant declines in the parameter moves further away from the maximum value of the diagnostic case model, although the curves for the individual components display different values of support for the mean total biomass.

10.2. Retrospective analyses

Retrospective analyses involve rerunning the selected model by consecutively removing successive years of data to estimate model bias (Cadrin and Vaughan, 1997; Cadigan and Farrell, 2005). A series of four additional models were fitted starting with the full data-set (through 2015), followed by models with the retrospective removal of all input data for the years 2015–2012 sequentially. The models are named below by the final year of data included (e.g., 2011–2015). A comparison of the spawning potential, recruitment and depletion trajectories are shown in Figure-Appendix 2.

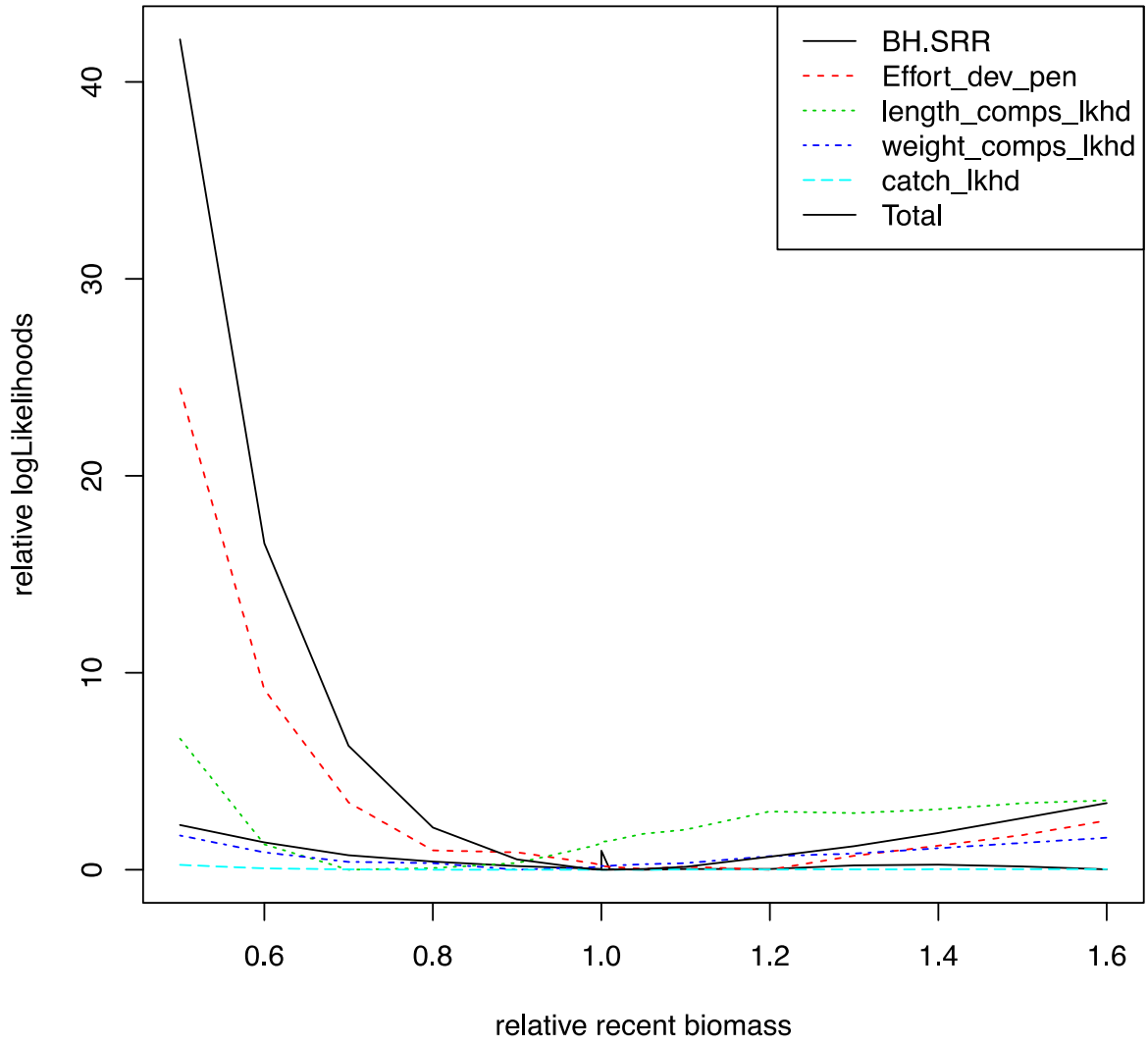

Figure-Appendix 1 Changes in the total, and individual data component log-likelihood with respect to relative mean total biomass in recent 4 years (2011-2014) period to the diagnostic case model.

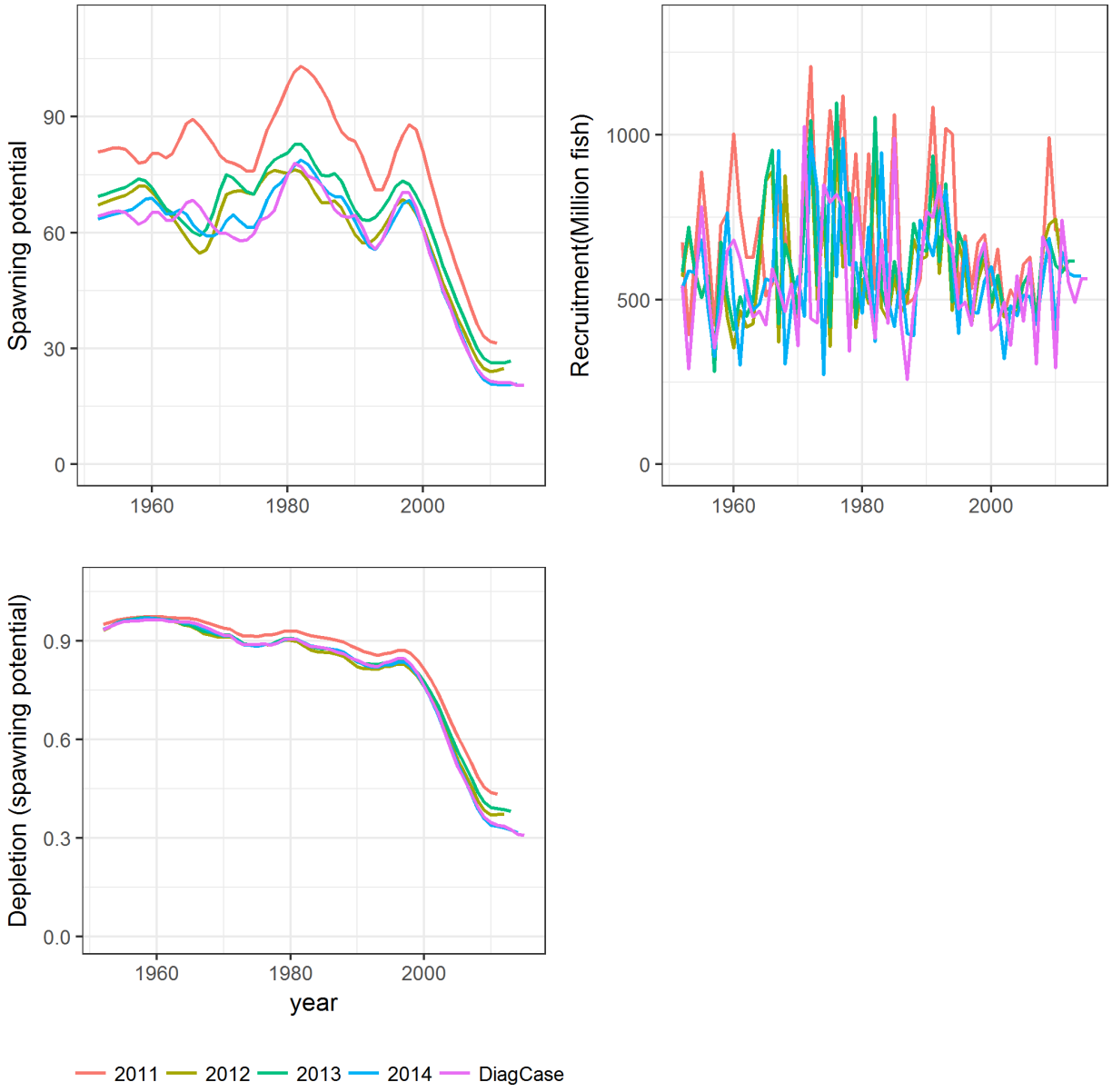

Figure-Appendix 2 Estimated spawning potential, recruitment and fishery depletion ($SB/SB_{F=0}$) for each of retrospective runs.