NOAA Technical Memorandum NMFS-NE-285

Report of the Index-Based Methods Working Group

US DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, Massachusetts
February 2023

NOAA Technical Memorandum NMFS-NE-285

This series represents a secondary level of scientific publishing. All issues employ thorough internal scientific review; some issues employ external scientific review. Reviews are transparent collegial reviews, not anonymous peer reviews. All issues may be cited in formal scientific communications.

Report of the Index-Based Methods Working Group

Northeast Fisheries Science Center[1]

[1] NOAA Fisheries, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543

US DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, Massachusetts
February 2023

Editorial Notes

Information Quality Act Compliance: In accordance with section 515 of Public Law 106-554, the Northeast Fisheries Science Center (NEFSC) completed both technical and policy reviews for this report. These pre-dissemination reviews are on file at the NEFSC Editorial Office.

Species Names: The NEFSC Editorial Office's policy on the use of species names in all technical communications is generally to follow the American Fisheries Society's lists of scientific and common names for fishes, mollusks, and decapod crustaceans and to follow the Society for Marine Mammalogy's guidance on scientific and common names for marine mammals. Exceptions to this policy occur when there are subsequent compelling revisions in the classifications of species, resulting in changes in the names of species.

Statistical Terms: The NEFSC Editorial Office's policy on the use of statistical terms in all technical communications is generally to follow the International Standards Organization's handbook of statistical methods.

EXECUTIVE SUMMARY

Term of Reference (ToR) 1. Develop methods to create data that if assessed with standard age-based approaches (e.g., Virtual Population Analysis [VPA] or Age-Structured Assessment Program [ASAP]) could exhibit a strong retrospective pattern.

The stock assessment program Woods Hole Assessment Model (WHAM) was used to generate data that exhibited strong retrospective patterns when assessed with an age-based approach. There were 2 alternative sources for the retrospective pattern: unaccounted catch and an unknown increase in natural mortality (M). The magnitudes associated with each source were modified according to the specific scenario to produce a Mohn's rho of approximately 0.5 for spawning stock biomass. This is a strong retrospective pattern and comparable to retrospective patterns seen in some stock assessments in the region. There were 8 combinations of retrospective source, fishing history, and number of fishery selectivity blocks (each factor had 2 levels) that formed the 50-year base period. An index-based assessment was applied every other year during the next 40 years of the feedback period. The data available to the index-based methods (IBMs) contained noise in the observations, as would occur in actual stock assessments. For each scenario, IBM, and control rule combination, 1,000 closed-loop simulations were performed.

ToR 2. Identify a number of IBMs and a range of harvest control rules for use in closed-loop simulation using index-based data resulting from ToR 1.

Thirteen IBMs (12 individual methods and 1 ensemble approach) were selected from the large number of available methods for use in the closed-loop simulations. The selected methods reflect local use, common use elsewhere, or a newly developed approach being proposed for the region. When applied, the IBMs used the M rate from the initial base period to reflect that the change in M was not known in the age-based assessment. There were 2 harvest control rules used. The first method simply applied the catch advice resulting from the IBM directly. The second method reduced the catch advice (multiplied it by 0.75) to approximate a decision that the original catch advice was an overfishing limit and that the acceptable biological catch should be reduced from it. In neither case did the size of the population modify the harvest control rule, as is sometimes done with age-based assessments, because many of the IBMs do not have a way of determining current population abundance relative to a reference point.

Table of Index-based methods (IBMs) with short descriptions. A * in the description indicates the IBM is used in the Ensemble, while a # indicates that the IBM is used in a local assessment.

IBM	Description
AIM	An Index Method *#

IBM	Description
CC-FSPR	Catch Curve F40%SPR *
CC-FM	Catch Curve F=M
DLM	Dynamic Linear Model
Ensemble	Combination of models
ES-FM	Expanded survey biomass F=M
ES-Frecent	Expanded survey biomass recent F*#
ES-FSPR	Expanded survey biomass F40%SPR *#
ES-Fstable	Expanded survey biomass stable F
Islope	common trend based IBM *
Itarget	common level based IBM *
PBS	survey smoother *#
Skate	catch/B driven *#

ToR 3. Identify metrics from the index-based assessment results that could be used in evaluations of trade-offs in performance among harvest control rules and IBMs.

A total of 50 performance metrics were collected from the simulation results. All were based on the true population values that are known because this is a simulation exercise (these values would not be known in application of the IBMs to real data). There were sets of performance metrics for spawning stock biomass, fishing mortality rate, and catch. The metrics were collected for either the short term (the first 6 years of the feedback period) or the long term (the last 20 years of a 40-year feedback period). The metrics included average values compared to the associated reference point, classification of whether an event occurred during the time period (e.g., F goes above F_{MSY}), or the number of years that an event occurred during the time period. Examples of how these performance metrics could be used to score the performance of the IBMs are provided, but these are management decisions and are shown for demonstration purposes only.

ToR 4. Evaluate the combinations of IBMs and control rules using the metrics in ToR 3 to determine candidates for consideration by the Councils or other management authorities.

The IBMs generally formed 2 groups in terms of performance. The first group consists of CC-FSPR, CC-FM, DLM, PBS, ES-Frecent, and Islope, while the second group consists of Skate, AIM, ES-Fstable, ES-FSPR, ES-FM, Ensemble, and Itarget. The IBMs within these groups performed similarly in terms of both the mean values and the type of relationship between SSB/SSB_{MSY} and catch/MSY. The first group generally had lower short-term catch, lower long-term F/F_{MSY}, and higher SSB/SSB_{MSY} than the second group. The second group had a more linear relationship between long-term SSB/SSB_{MSY} and catch/MSY than the first group. IBMs in the first group may be more suited for stocks that are thought to be in good condition, while IBMs in the second group may be more suited for stocks that are thought to need rebuilding.

Multiple lines of evidence pointed to the retrospective source and the catch advice multiplier as the most important factor in determining performance of the IBMs. For some performance metrics, fishing history and interaction terms among factors were also important.

Overall, none of the IBMs considered in these simulations performed better than the rho-adjusted statistical catch-at-age (SCAA) model, so in situations where an SCAA model is rejected due to a strong retrospective pattern, there should not be an expectation that an IBM will perform better than the rejected model.

The reasons for why the IBMs grouped as they did according to the performance metrics is not understood at this time and in need of future research. The results of this study provide a basis for these explorations. Additional questions can also be addressed by the framework developed in this study.

ToR 5. Provide guidance on specific situations that are and are not well-suited for a particular control rule or IBM identified in ToR 4.

For several IBMs, the variability among metrics was primarily due to the factor related to the cause of the retrospective pattern. Thus, if you cannot identify the likely (or dominant) source of a retrospective pattern (catch or M), then using an IBM sensitive to the retrospective source would be risky to the stock and fishery, and this risk could be avoided by using a method robust to this uncertainty. In this regard, the following methods were more robust to retro type: DLM, PBS, ES-Frecent, Islope, and to some extent, the 2 catch curve methods. We anticipate these methods would have more robust performance in situations where a retrospective pattern exists similar to that simulated in this project.

A rebuilt stock (long term $SSB/SSB_{MSY} > 1$) was greatest for the IBMs that were the least sensitive to the source of the retrospective pattern (CC-FSPR and CC-FM, DLM, PBS, ES-Frecent and Islope). DLM, PBS, and CC-FM and CC-FSPR achieved the greatest median SSB/SSB_{MSY}, but catch was lowest for CC-FSPR and CC-FM. PBS, Islope and ES-Frecent had the highest median catch among the methods that achieved rebuilding more than 50% of the time.

Trade-offs in risk (overfishing and rebuilding) and rewards (catch) are inherent in management decisions. Balancing median catch close to MSY in the short term while still maintaining a probability of at least 50% of achieving rebuilding was possible for the ES-Frecent,

PBS, DLM, and Islope, although long-term median catch with these methods was far below MSY. This could indicate that these 4 methods are appropriate short-term models for management advice, but while they are employed, other efforts should be invested to return to an age-based model. As noted in ToR 4, none of the IBMs consistently performed better than the SCAA with rho-adjustment, supporting the recommendation to use age-based approaches despite retrospective patterns.

Caveats for the conclusions relate to the large, but by necessity limited, number of IBMs, scenarios, and simulations that could be conducted. The biological and fishery characteristics were based on local groundfish; other species or fisheries may have different performance for the IBMs. A single source, catch or M, and magnitude, 0.5, of retrospective pattern was considered; multiple, changing, and stronger retrospective patterns may lead to different performance of IBMs. The IBMs were applied every other year in the feedback period to mimic local application, but missing surveys could require 3 years before the next assessment could be conducted. By necessity, the IBMs were applied formulaically within the simulations, so review of data meeting IBM assumptions or poor diagnostics were not evaluated during this study; these results should be considered minimum performance of the IBMs. As with all research, many questions were raised during this study. The framework developed for these simulations is well suited to address many of them immediately or with minor modifications. The Index-Based Methods Working Group (IBMWG) recommends this work be continued to explore the results generated during this study as well as to build on these results to address new questions.

ToR 6. Create guidelines for setting biological reference points for index-based stocks.

This ToR requires additional study. Generally, it is challenging to generate appropriate biological reference points for index-based stocks because they lack production functions that allow examination of trade-offs between catch and future population size. For this reason and due to time constraints, only general guidelines and ideas for future research are provided at this time. The framework developed for this project is well suited for this additional research.

Working Group Process

The working group (Appendix 1) met weekly from March to November via webinar due to travel restrictions associated with the COVID-19 global pandemic. All meetings of the working group were open the public (Appendix 2), with announcements made regarding connection information prior to each meeting through the dedicated website for the working group (https://www.fisheries.noaa.gov/new-england-mid-atlantic/population-assessments/stock-assessment-working-group-index-based-methods-and-control-rules). For each of the 41 meetings (Appendix 3), the chair prepared an agenda and often a set of prompts to allow working group members to contribute thoughts before the meeting. This asynchronous sharing of ideas through Google Docs allowed rapid progress to be made. The working group used a GitHub repository to share code used to run and analyze the simulations (https://github.com/cmlegault/IBMWG). A number of computer resources were used to run the actual simulations. Due to the large size of the resulting files (~300 GB in total), these files could not be stored on the GitHub repository and were instead stored on Google Drive.

INTRODUCTION

In the U.S., age-structured stock assessment models are used when possible to estimate annual stock abundance and fishing mortality rates, as well as management reference points. These models must undergo peer review, where an independent panel of experts determines whether or not results from the model are suitable as the basis for determining stock status and for setting catch advice. There are a number of model diagnostics that are used to evaluate uncertainty and stability of assessment model results, but one that is commonly used and carries substantial weight in the review is the retrospective pattern. A retrospective pattern is a systematic inconsistency among a series of sequential assessment estimates of population size (or other related assessment variables) based on increasing time periods of data used in the model fitting (Mohn 1999). These inconsistencies in assessment estimates are indicative of 1 or more mismatches between model assumptions and patterns in the data used to fit the model. Large or persistent retrospective patterns indicate an instability in model results and may therefore be the basis for a peer review panel determining that model results are not suitable for management purposes (Punt et al. 2020).

Many stock assessments in the Northeast U.S. have a history of strong, positive retrospective patterns in biomass estimates, whereby estimates of biomass are revised downward and estimates of fishing mortality rate are revised upward as new data are added to the model. NOAA Fisheries, the New England Fishery Management Council (NEFMC), the Mid-Atlantic Fishery Management Council (MAFMC), and the Atlantic States Marine Fisheries Commission (ASMFC) manage these stocks, and retrospective issues remain a challenge for managers when setting catch advice and tracking stock status. This problem has been particularly acute for, but not limited to, stocks in the New England groundfish complex (NEFSC 2002a, 2005, 2008, 2015a, 2015b, 2017, 2019; Deroba et al. 2010) managed under NOAA Fisheries and the New England Council's Northeast Multispecies (Groundfish) Fishery Management Plan.

The magnitude of the retrospective pattern is typically measured with a statistic known as Mohn's rho (Mohn 1999). Stock assessments where the rho-adjusted value (terminal year estimate divided by 1 plus Mohn's rho) is outside the 90% confidence interval of the terminal year estimate of spawning stock biomass or fishing mortality rate are classified as strong retrospectives, and the rho-adjusted values are used for status determination and to modify the starting population for projections used to provide catch advice (Brooks and Legault 2016).

There is no formal criteria in the region for rejecting an assessment based on Mohn's rho, but large, positive values of rho (especially those persisting) have played an important role in the rejection of recent age-based assessments for stocks including Atlantic mackerel (Scomber scombrus), Georges Bank Atlantic cod (Gadus morhua), Georges Bank yellowtail flounder (Limanda ferruginea), and witch flounder (Glyptocephalus cynoglossus; Deroba et al. 2010; Legault et al. 2014; NEFSC 2015a, 2015b). In each of these cases, and another where the assessment rejection was not based on the retrospective pattern (black sea bass [Centropristis striatus]; NEFSC 2012), the councils have relied on a variety data-limited approaches for setting catch advice for these stocks (McNamee et al. 2015; NEFSC 2015a, 2015b; Wiedenmann 2015). These approaches have all been ad-hoc, and a recent analysis suggested that some of the data-limited approaches may not be suitable for stocks in the Northeast U.S. with a history of high exploitation rates (Wiedenmann et al. 2019). In addition, large, positive retrospective patterns persist for a number of other stocks in the region (NEFSC 2019), raising concerns that additional stocks may rely on data-limited approaches in the future. Additionally, many of these stocks are in rebuilding plans without a mechanism to track rebuilding progress in the absence of reference points. A recent challenge also emerged in the 2020 stock assessments when index-based methods

(IBMs) in place for some stocks were rejected and alternative IBMs were pursued (i.e., red hake [Urophycis chuss] and Northern windowpane flounder [Scopthalmus aquosus]). Therefore, there is an immediate need to identify suitable data-limited approaches for setting catch advice and stock status determination for stocks with age-based assessments that did not pass review.

Stocks assessments in the region are classified as either management track or research track assessments (see Box 1 for more details on the distinction). Research track assessments can focus on individual stocks, or they can be topic based. Topic-based assessments are meant to provide utility and application in future management track assessments. The Index-Based Methods Working Group (IBMWG) was formed to conduct a topic-based research track assessment to evaluate the suitability of a range of data-limited methods for setting target catches for stocks with assessments with strong retrospective patterns. The Terms of Reference for the Index-Based Methods Research Track Stock Assessment are:

1. Develop methods to create data that if assessed with standard age-based approaches (e.g., Virtual Population Analysis [VPA] or Age-Structured Assessment Program [ASAP]) could exhibit a strong retrospective pattern.
2. Identify a number of IBMs and a range of harvest control rules for use in closed-loop simulation using index-based data resulting from ToR 1.
3. Identify metrics from the index-based assessment results that could be used in evaluations of trade-offs in performance among harvest control rules and IBMs.
4. Evaluate the combinations of IBMs and control rules using the metrics in ToR 3 to determine candidates for consideration by the councils or other management authorities.
5. Provide guidance on specific situations that are and are not well-suited for a particular control rule or IBM identified in ToR 4.
6. Create guidelines for setting biological reference points for index-based stocks.

To address the ToR, the IBMWG developed a management strategy evaluation (MSE) simulation model (e.g., Punt et al. 2016). MSE models are used to evaluate the trade-offs associated with different management options in the face of uncertainty and consist of a series of linked submodels (Figure 0.1). The foundation of MSE is the operating model that controls the population and fishery dynamics in the system. The operating model is run for an initial period of time (called the Base period here) that controls the historical population dynamics and fishing pressure, and it allows for sufficient data to be generated in the observation model to be used in the different assessment/management options being explored in the MSE. After the Base period, a given management approach is applied to set the target catch for the stock, which is then removed from the population. This process is repeated at a fixed interval over a number of years in what is called the Feedback period (Figure 0.2). In many MSEs, catch advice is based on a stock assessment model that estimates current abundance and management reference points. In data-limited cases, or when an assessment model is rejected, the model output is not used, and catch advice is based on the available data required by a given data-limited approach. Details of the operating model are provided in response to ToR 1, and the data-limited methods explored are provided in response to ToR 2 below. Because the Northeast U.S. has a long time series of data from a fishery-independent survey, the IBMWG focused on methods that can utilize this index of abundance, and we broadly refer to these approaches as IBMs.

New Stock Assessment Process in the New England and Mid-Atlantic Region

The Northeast Region Coordinating Council (NRCC) is comprised of leadership from the ASMFC, Greater Atlantic Regional Fisheries Office (GARFO), MAFMC, NEFMC, and the Northeast Fisheries Science Center (NEFSC). The NRCC coordinates science and management activities and resources in the region. A significant responsibility of the NRCC is setting stock assessment priorities and schedules for the region.

Historically, the process to establish stock assessment priorities and schedules was conducted annually on an ad hoc basis and, at times, influenced by outside pressure and demands. In 2018, the NRCC agreed to a new process in determining the region's assessment needs in order to provide for a more strategic and long-term approach with the goal of improving the overall quality of the regions' stock assessments (NRCC 2018). The new process identified 2 types of stock assessments: management track and research track. Management track assessments provide routine and updated advice on a specified schedule to directly inform management actions. Generally, management track assessments utilize the previously peer reviewed stock assessment for a particular species; however, the new process allows for greater flexibility to incorporate new or updated information and approaches to provide for continued improvements. Research track assessments are comprehensive and complex efforts that are typically carried out over several years in order to develop and consider new research. These assessments consider new data sources and new or different modeling approaches. An approved research track assessment can then be applied in future management track assessments and eventually inform management.

Research track assessments may focus on a single species and individual stocks or may examine a topic or issue that could apply to a broad range of species or assessment types. Topic-based research track assessments should demonstrate utility and application to future management track assessments or management actions. The ToRs for a research track assessment should specify what and how outputs and outcomes could apply to future management track assessments. The IBMs research track assessment is the first topic based assessment to be developed, conducted, and peer reviewed through the new NRCC process.

TOR 1. DEVELOP METHODS TO CREATE DATA THAT IF ASSESSED WITH STANDARD AGE-BASED APPROACHES (E.G., VPA OR ASAP) COULD EXHIBIT A STRONG RETROSPECTIVE PATTERN.

Operating Model

We used the Woods Hole Assessment Model (WHAM; Stock and Miller 2021; Miller and Stock 2020; Appendix 4) as the operating model for the simulation study. WHAM is an R package, and the general model is built using the Template Model Builder package (Kristensen et al. 2016). WHAM is a stock assessment model used to estimate parameters from real data, but it can also simulate data, including during a projection period, given a set of parameters and projection specifications. We used WHAM operating models to simulate data with known properties during the base period. Catches during the feedback period were iteratively updated based on an IBM) and harvest control rule that used the simulated observations to make catch advice during the

projection period. We specified the initial base period to be 50 years, labeled 1970-2019, where the age-structured population and catch and index observations were simulated according to user-supplied biological and fishery parameters for each scenario. The following period of 40 years, labeled 2020-2059, normally viewed as a projection period in a WHAM assessment model, is here specified as a feedback period. From the beginning of the feedback period, an IBM was applied to generate catch advice for 2 year blocks, a typical catch specification timeframe for NEFMC and MAFMC managed fisheries. WHAM used these catches, along with the user supplied biological and fishery data, to have the simulated population respond to the IBM.

There were a few assumptions that defined different operating model scenarios. First there were 2 assumptions on history of fully-selected fishing mortality rates:

1. 2.5 x F_{MSY} (based on M=0.2 and terminal selectivity) for the entire base period (ie., always overfishing) or
2. 2.5 x F_{MSY} for first half of base period and F_{MSY} for second half of base period (ie., overfishing the first half and then reduced to F_{MSY} for second half).

Second, there were 2 assumptions about variation in fishery selectivity:

1. constant during the base and feedback period or
2. a change in selectivity after the first half of the base period so that the age at 50% selectivity increased to 5 from approximately 3.7.

Third, we considered 2 types of misspecification:

1. an unknown change in natural mortality (M) or
2. underreported catch.

The biological and fishery characteristics used in the simulations were derived from local groundfish stocks. The high level of fully-selected fishing mortality that we assumed for overfishing ($F>F_{MSY}$) was based on work by Wiedenmann et al. (2019; Figure 1.1) looking at harvest rates relative to fishing mortality reference points for Northeast groundfish stocks. These 2 fishing intensities during the latter half of the base period lead to different starting conditions for each of the simulations.

The 2 selectivity patterns are based on local groundfish fishery selectivity patterns. The changed selectivity pattern when 2 selectivity blocks occur reflects an increase in mesh size of the fishery to avoid younger fish (Table 1.1).

The values we used to adjust M and unreported catch were derived by attempting to achieve a Mohn's rho of approximately 0.5 for spawning stock biomass (SSB) when a statistical catch at age model configuration of WHAM was used to fit the simulated data (Table 1.2). We also fit the same SCAA configuration of WHAM to data without misspecified M or catch to verify that retrospective patterns were not present on average (Figure 1.2). We tried to include changes in survey catchability as an additional source of the retrospective pattern, but preliminary explorations showed we were unable to generate a Mohn's rho as large as 0.5 using this approach. This is an interesting result that deserves further exploration.

For the M misspecification, the true M used to simulate the population and observations changes from 0.2 to 0.32 (for fishing intensity history type 1) or to 0.36 (for fishing intensity

history type 2) linearly between 2000 and 2009 during the base period and remains at the higher level throughout the feedback period. The IBM uses the observations, and those IBMs that require a M rate use the value (0.2) from before any change in M because the change in M is meant to be unknown.

For the catch misspecification, a scalar multiple of the true catch observation is provided to the IBM. The scalar is 5 for fishing intensity history type 1 and both selectivity patterns, 2.25 (for fishing intensity type 2 and a change in selectivity) or 2.5 (for fishing intensity type 2 and constant selectivity) linearly from 2001-2010. Note that this scalar is applied only to the aggregate catch so it affects all catches at age equally. The same scalar is then used to multiply the catch advice from the IBM into a total amount of catch removed from the population. For example, if the scalar is 5 and recent catch removed from the population is 100 tons, then the IBMs would be told that it was actually only 20 tons, and when one of the IBMs then provides catch advice of 25 tons, the simulation model would remove 125 tons from the population.

We used the same set of 1,000 random seeds to initialize simulations for all scenarios. For every random seed, we applied each of the IBMs so the base period is identical for each of them, but the population and observations during the feedback period vary by IBM due to the differences in catch advice provided by each of the IBMs (Figure 1.3). During the review meeting, the question was posed as to whether the retrospective patterns continued during the feedback period. Examination of 2 random number seeds for the 2 sources of retrospective patterns, catch and M, demonstrated that the retrospective patterns for SSB remained strong during the feedback period (Figure 1.4). Another question posed during the review meeting was how often the fishing mortality cap of 2.0 is invoked. Preliminary explorations during the meeting indicated it occurred more frequently than expected, with scenarios using catch as the retrospective source invoking the cap more frequently than the associated scenarios using M as the retrospective source. However, this exploration used a limited number of realizations, and so it should be explored across all realizations before strong conclusions are drawn.

The maximum possible fully selected fishing mortality rate was 2 no matter how high the IBM catch advice was. This limit was maintained even when catch misreporting occurred. This prevented the population from going negative, as could have happened if the catch advice was simply subtracted from the population.

All operating model scenarios assumed autoregressive deviations between the realized recruitment and that expected from a Beverton-Holt stock recruitment curve (Table 1.3). This stochasticity in recruitment is the only source of variation in the population for a given random seed. The initial numbers at age are specified at equilibrium values with total mortality defined by M in the first year, fishery selectivity in the terminal year of the base period, and corresponding F_{MSY}. Therefore, they are the same for all 1,000 seeds of a given scenario, but they differ between scenarios when there are differences in terminal year fishery selectivity. Consistent seeds were used across all the scenarios so differences in performance among IBMs can be attributed to characteristics of the IBMs themselves and not different sequences of observation or process errors; such standardization is considered best practice (Punt et al. 2016).

There were a number of scenarios considered in a full factorial design, as well as a number of additional scenarios denoted as one-offs. The factors included in the full factorial design were: cause of the retrospective pattern, fishing history, fishery selectivity, IBM, and catch multiplier (Table 1.4). This resulted in a total of 2 x 2 x 2 = 8 operating models where 13 IBMs and 2 catch multipliers were applied during the feedback period resulting in 208 scenarios for each of the 1,000 random seeds. Several one-offs were considered (Table 1.5), however due to time limitations, only

the no retrospective pattern and SCAA model analyses could be conducted. The SCAA model was configured and estimated using WHAM, but specified to closely match ASAP (Figure 1.5). The purpose of the extra analyses was to examine specific aspects of the simulations without requiring the full number of runs if they had been part of the full factorial design.

Biological reference points

We used the true values of the population and fishery to define biological reference points and associated metrics to evaluate IBM performance (Table 1.6). The MSY reference points were derived for each year during the base period using the fishery and biological characteristics of that year. This resulted in changes between the first and second halves of the base period due to changes in M and fishery selectivity in appropriate scenarios. Similar to the definition of initial numbers at age, the fully selected fishing mortality during the base period was based on F_{MSY} (defined using M = 0.2) and the selectivity during the terminal years of the base period (Figure 1.6). For example, in scenarios where F history pattern = 2, fully-selected fishing mortality begins at 2.5 times this F_{MSY} and reduces to this F_{MSY} value during the last portion of the base period. Since the changes in the second half of the base period were held at these values during the feedback period, the MSY reference points from the terminal year of the base period, including changes in M, were used for performance metrics in the analyses.

This difference between the reference points used to set the fishing history and those to determine performance metrics led to some counterintuitive starting conditions. Specifically, the large change in M required to achieve the desired retrospective pattern led to the fishing mortality ratio compared to the terminal year reference points being below (in some cases well below) 1 instead of equal to or well above 1 (Figure 1.6). This led to the SSB ratios being above 1 for F history scenarios of overfishing then F_{MSY} in the second half and SSB ratios equal or above 1 for F history scenarios of overfishing throughout (Figure 1.7). In contrast, the reference points associated with catch as the retrospective source had only minor changes due to fishery selectivity changes (Figure 1.8). Had we used the terminal year MSY reference points when setting the fishing mortality rate histories, the overfishing levels would have been so high as to collapse the population. So instead, we used the F histories from the catch retrospective scenarios directly in the M retrospective scenarios and recognize that the terms associated with the F histories do not necessarily apply for the M retrospective scenarios. The changes in MSY reference points associated with M as the retrospective source have consequences for the performance metrics computed for the feedback period and could impact the IBMs that rely on stationarity assumptions when deriving catch advice (such as knowing the M rate or being able to approximate the biomass reference point).

Feedback period procedure

Within the MSE process, the IBMs were conducted in the middle of the calendar year resulting in different years of information being available for different data inputs. Two surveys were conducted each year to produce an index of abundance: the first occurred 0.25 into the year (spring like survey) and the second at 0.75 of the year (fall like survey). The IBM method conducted in year y had the spring like survey in year y available to it as well as the fall like survey from year y-1 and the catch data from year y-1 (e.g., an assessment that occurred in summer 2014 would have the spring survey data from 2014, the fall survey data from 2013, and the catch from 2013). IBM methods that utilized only a single survey were provided with the average of the 2

surveys calculated as the mean of the spring survey in year y and the fall survey in year y-1. As no fall survey took place prior to year one, the first year of mean survey index was the mean of the spring survey in year two and the fall survey in year one approximating the index on January 1st of year two. IBM methods that evaluated the catch over the mean survey index lined up the catch in year y with the survey in year y with survey in year y defined as the mean of the spring survey in year y+1 and the fall survey in year y (e.g. the total catch for the year 2018 would be lined up with the survey index approximated as Jan 1, 2019 [mean of fall 2018 survey and spring 2019 survey]).

Annual aggregate catch and index observations were assumed to be normal after log-transformation. Aggregate catch and index observations were created by the operating model by summing over all ages using the simulated population abundance in numbers, associated selectivity patterns, and weights at age. The associated age composition observations were assumed to be multinomial distributed. Coefficients of variation and effective sample sizes are provided in Table 1.7. During the review meeting, a reviewer asked whether these levels of uncertainty matched local assessments. Fifteen local assessments had the estimated coefficients of variation for the surveys along with effective sample sizes for the fishery and surveys complied (Figures 1.9 and 1.10). The coefficients of variation used in this study for the surveys appear reasonable compared to local assessments, while the effective sample sizes used in this study are a bit higher than found in local assessments. The latter implies the age composition in the simulated data are known more precisely than would be expected from local assessments.

TOR 2. IDENTIFY A NUMBER OF IBMS AND A RANGE OF HARVEST CONTROL RULES FOR USE IN CLOSED-LOOP SIMULATION USING INDEX-BASED DATA RESULTING FROM TOR 1.

A large number of data-limited methods exist for setting catch advice, and they vary widely in complexity, data inputs, and assumptions required (e.g., Carruthers and Hordyk 2018). The northeast U.S. is a data-rich region with a range of information available including the NEFSC spring and fall coast-wide bottom trawl surveys, as well as historical catch data. As a result, the IBMWG focused on a subset of data-limited methods known as IBMs that utilize time series of fishery catch and catch per unit effort (CPUE) from a survey or fishery and omitted methods that require only catch data, snap shots of survey data, or length data (e.g., constant catch methods or length-based methods). The IBMWG also omitted methods that required complete catch histories (from the inception of fishing), assumed an underlying surplus production population dynamics, or required assumptions about relative depletion. Complete catch histories are not available for stocks in the region. Example methods that meet these criteria for omission include Depletion-Corrected Average Catch (DCAC; MacCall 2009), Depletion Based Stock Reduction Analysis (DB-SRA; Dick and MacCall 2011), and the Simple Method for Estimating MSY (SMSY; Martell and Froese 2013). The IBMWG also omitted methods that assume an underlying surplus production model where changes in productivity are driven solely by changes in biomass (e.g., Martell and Froese 2013). Changes in recruitment (Miller et al. 2017; Xu et al. 2018; Tableau et al. 2019) or M (Pershing et al. 2015) for many stocks in the region violate this assumption, and surplus production fits to survey and catch data result in very different estimates of biomass over time compared to age-based assessments for many stocks in the region (Wiedenmann et al. 2019).

Given the long time-series of trawl survey data, the group focused largely on survey IBMs. The relatively short timeline for the IBMWG restricted the total number of methods that could be evaluated to those that have been used or would be considered plausible for the Northeast Shelf. A list of methods currently used in the region is provided in Table 2.1, with 13 IBMs selected by the IBMWG for evaluation. The IBMWG deliberately did not include every approach to limit the overall number of methods explored and to allow for exploration of additional options not currently used but that have been explored or applied elsewhere. Although catch-curve analyses are not currently applied in the region, they were included here since age information is available for most of the stocks and because Wiedenmann et al. (2019) showed they performed well in a retrospective application to groundfish stocks. Due to the availability of age-based information, we did not include length-based methods; this approach also helped to reduce the overall number of methods explored. We also included 2 additional IBMs (Islope and Itarget) not currently used in the region because they have been tested in other applications and shown promise (Geromont and Butterworth 2015a, 2015b; Carruthers et al. 2015; Wiedenmann et al. 2019). An ensemble of models was also considered based on recent findings that improved performance can result from combining the results from multiple models (Anderson et al. 2017; Rosenberg et al. 2017; Spence et al. 2018; Stewart and Hicks 2018). The full range of methods included in this analysis are detailed below with equations provided in Table 2.2. Each method was examined for its ability to produce sustainable catch advice with data that would lead to large retrospective patterns. A comparison table of the data requirements and arbitrary constants used in the IBMs was created by Dr. Paul Rago, chair of the review panel, and presented in Table 2.3 with his kind permission.

Methods that only use the survey index and catch

Plan B Smooth (PBS)

The PBS approach has been used to set catch advice for Georges Bank cod since the rejection of the 2015 age-based assessment (NEFSC 2015, 2017, 2019). The PBS approach combines the spring and fall surveys into an average index, then locally estimated scatterplot smoothing (LOESS) is applied to the average index (with a span = 0.3). The predicted LOESS values in the final 3 years are used in a log-linear regression to estimate the slope, and this slope (transformed back to the linear scale) is used to adjust the most recent 3-year average catch to generate catch advice (Table 2.2).

Islope

The Islope method is similar to the PBS approach in that it uses recent trends in the combined average index (spring and fall) to adjust the recent average catch up or down. This general method was proposed by Geromont and Butterworth (2015a) and has been evaluated in a number of studies (e.g., Geromont and Butterworth 2015a, 2015b; Carruthers et al. 2015; Wiedenmann et al. 2019). A log-linear regression is applied to the final 5 years of the unsmoothed average index, and the slope is used to adjust a multiple (< 1) of the most recent 5 year averaged catch to generate catch advice (Table 2.1). Four formulations of Islope were proposed by Geromont and Butterworth (2015a), and we used the least conservative version here (version 1), although flexibility was incorporated to allow for exploration of the other versions. Key differences between Islope and PBS are the use of an unsmoothed average index, a longer recent time period (5 years vs. 3 years), and a buffered catch (80% of the recent average in version 1; Table 2.2).

Itarget

The Itarget method was proposed by Geromont and Butterworth (2015a) and has been evaluated in a number of studies (Geromont and Butterworth 2014, 2015; Carruthers et al.; Wiedenmann et al. 2019). Instead of using trends in the recent index, catch advice is determined by comparing the most recent 5-year average index to some target based on a reference period. The reference period does not change, and we used the last 25 years of the base period (years 26-50) as our reference period. The target index of abundance is set to some multiple (≥ 1) of the average index value over the reference period, and the catch advice is based on the average catch over the reference period, adjusted up or down based on the recent 5-year average index relative to target index (Figure 2.1; Table 2.2). We combined the spring and fall surveys into an average index to use in the Itarget method. As with Islope, Geromont and Butterworth (2015a) proposed 4 formulations, and we used the least conservative option (version 1; Table 2.2).

Skate method

The skate control rule method was developed by the New England Council's Skate Plan Development Team and endorsed by its Scientific and Statistical Committee to produce catch advice for the skate complex using only a time-series of catch and a survey index for each skate species (Skate FMP Amendment 3). A single time-series of the survey index is calculated as the mean of the fall and spring surveys. The median value of the annual catch divided by the annual index over the entire time series (except the most recent years) provides the relative fishing mortality to produce catch advice. A moving average smoother is applied to the catch and survey index prior to dividing the 2. The relative fishing mortality multiplied by the terminal year of the smoothed survey index is the proposed catch advice. The entire time-series is used, and there is no comparison to a pre-specified reference time period.

An Index Method (AIM)

AIM was developed by Dr. Paul Rago at the Northeast Fisheries Science Center (NEFSC 2002b), and the executable with GUI can be found in the NOAA Fisheries Toolbox (https://nmfs-fish-tools.github.io/AIM/). For the current project, the methods were implemented in R (comparisons were performed to confirm that the NFT program and the R implementation produce the same results; notation herein is consistent with that in the NFT program help files). The AIM model requires 2 inputs—an index of relative stock biomass and total catch in biomass—and seeks to identify a relationship between 2 calculated series: Replacement Ratio (Ψ) and Relative F (F_{REL}). Ψ is calculated as a ratio of the survey index in a given year divided by a weighted average of user-defined survey values; for this application, the numerator of Ψ is the current index year, and the denominator is the average of the previous 5 years of the index. Assuming no density-dependence in the population dynamics, this time series of replacement ratio reflects changes in annual population biomass due to recruitment, growth, and natural and fishing mortality. F_{REL} is the ratio of catch in a given year (y) to an average of recent index values (for this implementation, a 3-year average of the biomass index, centered on y, was used) and in principle reflects the relative magnitude of catches relative to average biomass; the catchability of the index is unknown, and so the scale of F_{REL} is not simply the proportion of biomass loss. The AIM model performs a regression of $\ln(\Psi)$ on $\ln(F_{REL})$ to identify the value of F_{REL} where $\Psi=1$ (i.e. where $\ln(\Psi)=0$, a level of fishing that allows replacement, which we'll refer to as F_{REL*}). Assuming the regression is satisfactory and modeling assumptions are met, then the value of F_{REL*} is a relative fishing rate that allows the population biomass to replace itself. To arrive at an approximation of stable catch advice, F_{REL*} is multiplied by the most recent index value.

Dynamic Linear Model (DLM)

The DLM is a state-space approach requiring an index of relative stock biomass and total catch in biomass to produce an estimated target catch level. For a single or multiple survey indices and a vector of catch data, the method fits a dynamic linear model on the log scale with at least 2 components: 1) a smoothly evolving mean abundance, and 2) a dynamic regression on catch. Because survey indices and catch time series are often strongly correlated, multicollinearity can cause difficulties in model fitting. However, this linear correlation represents redundant statistical information, a reflection of SSB, contained in both time series. Therefore, the average relative exploitation rate is differenced out of the catch time series using a linear regression between the log survey indices and the log catch to create a time series of catch anomalies. Essentially, the catch anomalies values reflect whether more or less catch occurred than one would expect given the survey indices and catch history. Catch is chosen to be modified because it is assumed that the survey indices represent a more reliable measure of stock abundance. If the survey indices and catch time series for a given stock were not correlated, this procedure would only de-mean the catch time series and thus not impact model fitting. To track the mean survey abundances in this simulation experiment, the method incorporates a dynamic linear trend. A dynamic linear trend is estimated by 2 state variables: 1) an initial, static intercept at time 0, and 2) a random walk whose value is added to the intercept variable at each time t. The second state variable is therefore an estimate of the abundance trend in the survey, after the effect of catch is accounted for, at any point in time. This is advantageous in cases where population productivity is nonstationary (e.g.,. M or recruits/spawner change over time) in that forecasts account for the recent trajectory of the population. Once the DLM is fit, the survey indices are projected forward in time with the estimated parameters, and an optimization routine is used to determine the annual catch levels that would result in the reference biomass level on average. For this project, the biomass reference level was set as the 75th percentile of the survey indices, and a 10-year rebuild to the reference level was desired. The first 2 years of catch advice returned by the optimization function is averaged with the catch levels in the 2 years prior to the forecast period in order to smooth out changes in harvest and produce a single value for catch advice. Notably, the quantification of forecast uncertainty and the ability to test different model structures was not employed due to the time constraints of the project. Additional options that could be explored in future work include testing different model structures and using the estimated forecast uncertainty to assess risk in setting different catch advice tailored to the stock of interest. Finally, the DLM structure could also be augmented with additional model components (e.g., other covariates, an autocorrelated component) and/or length or age information.

Swept area biomass methods

Expanded Survey Methods

Estimates of swept area biomass have been used to determine catch advice for a number of stocks in the region. The estimation of swept area biomass was approximated by scaling the simulated spring and fall survey observations to units of biomass using their respective catchabilities. These “scaled-up” biomass indices were then averaged using the spring observation from year y and the fall observation in year y-1, as is common practice. Catch advice was determined as the product of the averaged biomass index from the most recent year and one of 4 target exploitation rates based on: 1) $F_{40\%}$, from a Spawner Per Recruit analysis 2) the exploitation rate associated with stock replacement (F_{REL*} calculated as detailed above for the AIM method),

3) the target Fishing mortality set equal to the assumed M (F=M), and 4) the average of the relative F values, defined as catch divided by the "scaled-up" biomass indices, from the most recent 3 years.

Catch curve methods

The catch curve methods explored by the working group utilized numbers at age of fully-selected fish in each survey to estimate total mortality (Z) using catch curve analysis. An aggregate abundance at age was calculated by summing across the most recent 3 years for a given survey to create a single numbers-at-age vector. The age class with the largest abundance at age in each survey is set to the age at full selection. Ages below the age at full selection are dropped, as is the plus group, and the remaining values are used in a log-linear regression to estimate Z in each survey (Z = the inverse of the slope). An average Z across surveys was calculated and used with the assumed M to calculate an average F (Z = M + F), which is used with the most recent catch estimate to estimate total biomass (Table 2.1). Target catch values were then set using a specified target F, and we explored 2 options (F_{targ} = the assumed M, and $F_{targ} = F_{40\%}$ from a spawners per recruit [SPR] analysis). It is possible for the estimated Z from the catch curve analysis to be quite low, and if $Z \leq M$, it would result in nonsensical estimates of recent biomass because the recent F would be < 0. As a result, we set a constraint such that $Z = \max(Z, M + 0.05)$ so that the estimated recent F would never be < 0.05.

Combination of methods

Ensemble method

An ensemble approach was evaluated that determined catch advice as the median of the advice produced by other IBMs. Only 8 of the 12 IBMs were included in the ensemble. Variants of the catch curve and expanded survey methods were excluded because each variant is not independent of the others and including all of them would skew the performance of the ensemble towards these 2 general methods, effectively weighting these approaches more heavily than other IBMs. Ultimately, 2 variants of the expanded survey method were included because preliminary results suggested that their performance was distinct enough as to negate any concerns about unduly weighting the expanded survey method in the ensemble. The DLM method was excluded because it had relatively long computational times that would have prohibited any evaluation of an ensemble approach in the time available. The IBM methods included in the ensemble were: AIM, catch curve using $F_{40\%}$, expanded survey using $F_{40\%}$, expanded survey using the average relative F values from the 3 most recent years, Islope, Itarget, PBS, and the skate method.

Traditional age-based assessment methods

Statistical Catch-At-Age (SCAA)

An SCAA model and harvest control rule were used in simulations for a subset of operating model scenarios. The SCAA model was configured in the WHAM. The correct CVs and effective sample sizes were assumed for all catch and index data in the SCAA model. No stock-recruit model was assumed, and there were no random effects. M was misspecified at 0.2 for the scenarios where M changed. Mohn's rho was calculated (7 year peels) for abundance at age for all model fits during the feedback period and used to retro-adjust abundance at age for projections (divided by one plus Mohn's rho). Catch advice was determined by specifying F equal to 0.75 of the estimated $F_{40\%}$ (M=0.2).

Application of the methods

Each of the methods produces a single target catch value that was fixed over a 2-year interval. If the methods are being applied in year y, then target catches are set for years y+1 and y + 2 (denoted $C_{targ,y+1:y+2}$). In practice, the timing of setting target catches in the region generally occurs in late summer or early fall in between the spring and fall surveys, and before complete catch data are available. Therefore, in year y, complete catch data are available through year y-1, and survey data are available for the spring survey through year y and for the fall survey through year y-1. In practice, the data-limited methods that have been applied have used an average of the spring and fall index and we followed that approach here. If a method for setting catches uses an average of spring and fall, the average index in year y, $\bar{I}_y$ includes the spring data in year y and the fall data in year y-1:

$$\bar{I}_y = \frac{I_{fall,y-1} + I_{spr,y}}{2}.$$

Control Rules

Most IBMs do not have the ability to estimate a biomass reference point (e.g., B_{MSY}), which made consideration of so-called biomass-based harvest control rules that reduce F or catch in response to estimated changes in relative stock status impossible. Lack of clarity exists, however, on whether the catch advice from IBMs should be treated as an overfishing limit (OFL) or acceptable biological catch (ABC). OFLs are equated to the catch that would result from applying F_{MSY}, whereas an ABC is a catch reduced from the OFL to account for scientific uncertainty. The ABC can be thought of as the actual quota that is implemented. Therefore, each IBM was evaluated using 2 "harvest control rules": 1) the catch advice from a given IBM was applied directly and assumed to serve as a proxy for the catch associated with F_{MSY}, thereby being used directly as the ABC (catch multiplier = 1), and 2) the catch advice from a given IBM was reduced by 25% to account for unspecified scientific uncertainty, thereby making the catch advice from the IBM essentially the OFL and an ad hoc rule applied to create the ABC (catch multiplier = 0.75). Catches were reduced by 25% to approximate an ABC because using the catch associated with $0.75F_{MSY}$ is a common default ABC control rule in the region. The end result is that the quota implemented in the simulations is the catch advice coming from the IBM multiplied by the catch multiplier (1.0 or 0.75).

TOR 3. IDENTIFY METRICS FROM THE INDEX-BASED ASSESSMENT RESULTS THAT COULD BE USED IN EVALUATIONS OF TRADE-OFFS IN PERFORMANCE AMONG HARVEST CONTROL RULES AND IBMS.

The IBMWG identified a broad range of metrics to use in evaluating trade-offs that are typical of similar simulations, such as in management strategy evaluations (Punt 2017). The metrics were chosen because they directly represent management or biological interests, or were of legislative relevance (e.g., overfishing and overfishing metrics related to Magnuson-Stevens Act [MSA]). The metrics fell into 5 categories: 1) catch, 2) variability in catch, 3) biomass of the target species, 4) fishing mortality, and 5) miscellaneous (e.g., ecosystem, legislative). Generally,

each category corresponds to some fundamental objective typical of fisheries management, and trade-offs among categories are likely to be of greatest interest (Tables 3.1, 3.2). Metrics related to catch and variability in catch can also be considered reasonable proxies for economic metrics, such as profit or revenue, and societal metrics, such as stability in production, although the relationship between catch metrics and these societal and economic metrics may not be linear. Ecosystem and multispecies metrics were considered but fell outside the Terms of Reference and time frame available. Thus, all the metrics had a single species focus. Ultimately, not all of the metrics discussed by the IBMWG were used in analyses (see Table 3.2) and may be evaluated in future research.

Some of the metrics within each of the categories are redundant and displayed similar trade-offs. The redundancies were expected, but the IBMWG preferred to maintain a broader list to hopefully capture more metrics of specific interest to varied stakeholders, managers, and legislative matters. Thus, the IBMWG expects that the majority of trade-offs and subsequent decisions will be made using a subset of available metrics.

Performance metrics also varied as to whether they took a short-term (first 6 years of the projection period) vs. a long-term (last 20 years of a 40-year projection period) perspective. The IBMWG felt this critical as trade-offs among short and long-term objectives can be expected, but both are important to consider when comparing strategic performance of control rules.

Performance metrics could also be split into risk-based metrics that used probability statements (e.g., frequency of simulations over which something happens) vs. quantity-based metrics (e.g., the value of SSB relative to a reference point). Risk-based metrics are useful for summarizing and comparing scenarios because a single numeric value bounded by 0 and 1 captures the space of the outcomes. However, these metrics are often more difficult to interpret as they require some acceptable tolerance level for the probability, which can be more challenging for managers to define explicitly. Quantity-based metrics tend to be more easily interpretable and tolerance levels more easy to discuss (but the IBMWG recognizes that explicit transparency about what is "enough" from decision-makers is also commonly a challenge for these metrics, too). However, quantity-based metrics must be reported as summary from simulations, and so choices need to be made when communicating them about how the values and range of values of outcomes across simulations are summarized (e.g., mean/median over simulations vs. median plus/minus some quantile interval vs. some lower/upper quantile that reflects the preferred direction of the metric).

Shorthand definitions for each metric were created for easier graphical displays used in addressing other Terms of Reference. These definitions are below (Table 3.1), but follow the convention that: "l" is long-term, "s" is short-term, "avg" is an average, "is" represents a metric recorded as a probability statement, "n" represents the number of years that a situation occurred during a time period, "less" represents a value being less than some reference point, "ge" represents a value being greater than or equal to some reference point, "gr" represents a value being greater than some reference point, "f" indicates a fishing mortality quantity, "ssb" and "b" are used interchangeably and indicate spawning stock biomass, "catch" is self-evident, "dot" indicates a metric that compares realized fishing mortality to an equilibrium fishing mortality rate that would drive the stock to some given level of biomass (i.e., either 10% B_{MSY} or 50%B_{MSY}), 01 or 1 indicate 10% of B_{MSY}, 05 or 5 indicate 50% B_{MSY}, MSY indicates a maximum sustainable yield reference point (i.e., MSY, F_{MSY}, or B_{MSY}).

During the review, it was noted that many of the metrics should be highly correlated, and a correlation matrix was requested. We computed the correlations across all the simulations

conducted in the base set of analyses and found strong positive and negative correlations among metrics, as expected (Figure 3). These strong correlations indicate that simply using all the metrics as if they were independent results would produce biased inferences. Instead, specific subsets of metrics can be used to address specific questions or else dimension-reducing methods, such as cluster analysis, can be used across all the metrics.

TOR 4. EVALUATE THE COMBINATIONS OF IBMS AND CONTROL RULES USING THE METRICS IN TOR 3 TO DETERMINE CANDIDATES FOR CONSIDERATION BY THE COUNCILS OR OTHER MANAGEMENT AUTHORITIES.

Base and Sensitivity Runs

There were 208 factorial combinations of IBMs and scenarios examined in the base runs, where the catch advice multiplier (1 or 0.75) was considered part of the scenario definition (Table 4.1). Most of these factorial combinations had the full 1,000 simulations produce results (Table 4.2). There were 2 IBMs with 2 scenarios that had individual simulations fail, and the DLM had a much lower number of runs due to time limitations (it took significantly longer to run than all the other IBMs).

There were 2 sets of sensitivity runs. The first applied an SCAA to 4 of the scenarios. The SCAA applied a rho-adjustment to account for the estimated retrospective pattern in each assessment of each simulation. This required considerable computing time, and so only 4 scenarios were examined. The second sensitivity run removed all sources of retrospective pattern for 2 of the scenarios. All the IBMs, except DLM and SCAA due to time constraints, were applied to these scenarios.

There were a total of 230,147 successful simulations produced.

Linear Models

The simulations generated a massive amount of results. Therefore, evaluating the performance of the IBMs using only graphical displays would be both time consuming and subjective. So linear models were conducted to help identify the most important elements of the study design that affected the performance of the IBMs, an approach that has been used in management strategy evaluations (Punt et al. 2008; Fay et al. 2011). The objective for the linear models was to supplement the graphical displays and help focus additional analyses. So while some consideration was given to the validity of the linear models, a thorough evaluation of their assumptions and statistical rigor was not conducted.

A linear model was conducted independently for each of the metrics (ToR 3). Metrics recorded as frequencies or proportions were arcsine square root transformed, while all other metrics were log transformed (Punt et al. 2008; Fay et al. 2011), $tran(metric)$. Explanatory variables included source of the retrospective error, $retro_{type}$, fishing history, $F_{history}$, number of selectivity blocks, $num_{selectivity\ blocks}$, IBM, IBM, multiplier on the catch advice, $catch_{mult}$, and all 2-way interactions:

$$tran(metric) = \mu + retro_{type} + F_{history} + num_{selectivity\ blocks} + IBM + catch_{mult} + two\text{-}way\ interactions$$

where μ is the overall intercept. Results were summarized by creating a table noting which explanatory variables were significant for each metric at the 0.05 level (Sig) or no, (NS). The proportion of times an explanatory variable was significant among metrics grouped by SSB, fishing mortality, or catch, was reported.

The explanatory variables $retro_{type}$, IBM, $catch_{mult}$, and the two-way interactions between $retro_{type}$ and IBM , and $F_{history}$ and IBM were significant for all metrics (Table 4.3). The interaction of $retro_{type}$ and $catch_{mult}$ was also significant for the vast majority of metrics. The interaction of $catch_{mult}$, and IBM was consistently significant for long-term metrics and metrics related to catch. $F_{history}$ was significant for a majority of metrics, most consistently for short-term SSB and catch metrics. The interaction between $F_{history}$ and $retro_{type}$ was significant for a majority of metrics with no discernible difference between long- and short-term metrics. The $num_{selectivity\ blocks}$ variable was significant for the majority of the SSB metrics and all but one of the catch metrics. The remaining explanatory variables were generally significant for less than half of metrics. Given these results, greater emphasis was placed on understanding the effects of $retro_{type}$, IBM, and $catch_{mult}$ because these variables were the most consistently significant relative to the other variables.

Scoring

The mean value of each performance metric for all IBMs was computed across the 1,000 simulations of the 16 scenarios. Since some metrics are better when values are larger and others are better when values are smaller, these mean values cannot easily be combined. So metrics where smaller values were better were multiplied by negative 1 to create a set of values where bigger is better (Table 4.4). This mean and adjustment approach to make bigger values better was also applied to the SCAA scenarios (Table 4.5; more detail about SCAA scenarios provided below).

Scores were generated from these values in 2 ways. The first was simply ranking them, giving the largest value the number of IBMs considered, the smallest value the number 1, and integer values in between (except in the case of ties). The second was to subtract the mean and divide by the standard deviation of each metric across all the IBMs. Both approaches result in scores where bigger values are better for each metric and can be easily summed across selected metrics to determine which IBMs perform best relative to those metrics.

Two examples sets of metrics are provided here to demonstrate that the ordering of the IBMs depends strongly on the metrics selected. The first set of metrics contains the mean ratios of SSB, F, and catch to their respective MSY reference points in the long term (Figure 4.1). The second set of metrics contains the interannual variability in catch across the entire feedback period and the short term mean catch/MSY (Figure 4.2). These sets of metrics produce different orderings of the IBMs. For example, both catch curve (CC) IBMs score highly using the first set of metrics, while these IBMs score poorly using the second set of metrics. A number of additional sets of metrics are provided in Appendix 6 to demonstrate the changes in IBM ordering when different metrics are selected.

To allow users to easily evaluate a large number of sets of metrics, an R Shiny app was developed (see the scorer_app folder in the GitHub repository). This R Shiny app can be copied to a local machine and run to examine the Rank and Resid scores for whatever combination of metrics is desired. The sensitivity runs, denoted as “noretro” and “scaa” in the app, are also available in the R Shiny app. The app was created because specific metrics were not provided as the basis for

determining performance of the IBMs. The app allows users to pick among the 50 metrics to see how the IBMs performed relative to each other.

Ratios of MSY reference points

The ratio of the mean SSB, F, or catch to its respective MSY reference point showed differences among the IBMs by scenarios, with some factors having larger impact than others. Figures 4.3-4.5 show the mean SSB, F, and catch ratios in the long term (i.e., final 20 years), respectively, while Figures 4.6-4.8 show the ratios in the short term (ie., first 6 years). All 6 figures have the IBMs sorted so that the best (largest SSB and catch ratios, smallest F ratios) are at the top based on the mean across all 16 scenarios. The plots show similar patterns for 1 or 2 selectivity blocks in both the short and long term. The plots also generally show similar patterns for fishing histories in the long term, but there are differences in the short term, as expected due to the different starting conditions. The catch multiplier often had the expected effect of reducing catch in the short term but could sometimes result in higher average catch in the long term due to the larger SSB and lower F. The retrospective source had a large impact on the ordering of the IBMs, with groups of IBMs having either high or low performance for either catch or M, but rarely both. One group of IBMs contains the CC-FSPR, CC-FM, DLM, PBS, ES-Frecent, and Islope which performed well in terms of both SSB and F in both the short and long term, while the other group contains Skate, AIM, ES-Fstable, ES-FSPR, ES-FM, Ensemble, and Itarget which performed well in terms of catch in both the short and long term. In the long term, the SSB ratio was above 1 for the M retrospective source for all IBMs, while the catch retrospective source depended on IBM group as to whether it was above 1 or not. The IBM group that performed well for SSB ratios was able to rebuild the stock above SSB_{MSY} on average in the long term, while the other IBM group was not. Thus, if a stock is thought to be in poor condition, the IBMs in the group that performed well in terms of rebuilding would be preferred to the IBMs in the other group.

The distribution of 16 scenarios by IBM or 13 IBMs by scenario can be used to summarize the metrics. For example, the short- and long-term SSB/SSB_{MSY} distributions are shown in Figures 4.9-4.10. Similar figures for all the metrics are available in Appendix 6. These plots show the groupings of IBMs and influence of different scenarios on those groupings but in a more concise way than the 8 panel plots. This allows all the metrics to be presented in Appendix 6.

The distributions of mean values do not express the full range of results, however. When all the simulations are plotted, there is clearly a wide range for each ratio, indicating that performance for a particular series of environmental conditions, expressed through recruitment deviations, can vary widely. For example, Figure 4.11 shows the SSB/SSB_{MSY} and catch/MSY relationship for scenario CF1A (ie., catch retrospective source, F_{MSY} in second half of base period, constant selectivity block, and catch multiplier equal to 1.0) in the long term for the 1,000 simulations. Note the plots for the remaining 15 scenarios, as well as the equivalent short-term plots, are available in Appendix 6. The long- and short-term relationships can also be visualized through bagplots (Rousseeuw et al. 1999). For example, Figure 4.12 shows both the long-term and short-term SSB/SSB_{MSY} and catch/MSY for scenario CF1A. The full set of bagplots are available in Appendix 6.

The same groups of IBMs as noted above display different patterns in the relationship between the SSB and catch ratios in both the plots showing all the simulations and the bagplots. While all IBMs have large ranges for both ratios, Skate, AIM, ES-Fstable, ES-FSPR, ES-FM, Ensemble, and Itarget have a nearly linear relationship while CC-FSPR, CC-FM, DLM, PBS, ES-Frecent, and Islope have a much more diffuse relationship. This pattern by IBM group is consistent

across the different scenarios. These linear or diffuse relationships have implications for the trade-offs among IBMs, with linear relationships having higher certainty of performance but lower population sizes on average. The more diffuse relationships can also result in situations where the population is quite high but the catch is low relative to MSY, meaning the F is quite low.

Examination of the simulation plots in Appendix 6 also demonstrates some of the changes in results by the factors. For example, toggling between Figure A6.39 and A6.40 (A vs. R catch advice multipliers) shows that reducing the catch advice has a big impact on the vertical distribution of the diffuse relationship IBMs (much lower for R than A), while the linear relationship IBMs don't change as much but do appear to move a little to the right and maybe even up. This might occur because the diffuse relationship IBMs with reduced catch multipliers are seeing a population bouncing around an average value, meaning catch advice should be about the same, but the catch advice multiplier of 0.75 keeps reducing it.

Another way to explore the impact of the factors is to make so-called "confetti plots" where the mean value of a metric is shown for each IBM and scenario combination but the points are colored by the factor. For example, Figure 4.12 shows the mean value from the 1,000 simulations for 6 SSB metrics for the 208 combinations of IBM and scenario with the color of the point determined by the retrospective source. Here, the differences are clearly seen between catch and M as the retrospective source for most of the metrics. In contrast, the same plot except the points are colored by the fishing history during the base period shows much more interspersed results (Figure 4.13). The full set of "confetti plots" by metric and factor are provided in Appendix 6.

Risk issues

The average SSB and F relative to their MSY reference points are indicative of the expected status of population under different combinations of IBM and scenario, but other metrics can also be used to examine risk. Specifically, the "_is_" metrics can be used to examine the probability that an event will occur at least once during the period. For example, the average value of the SSB metric l_is_less_05_bmsy from the 1,000 simulations provides the probability that the SSB falls below half SSB_{MSY}, meaning the stock would be declared overfished, at least once during during the last 20 years of the simulation. Similarly, the average value of the F metric l_is_gr_fmsy from the 1,000 simulations provides the probability that the F falls above F_{MSY}, meaning the stock would be declared undergoing overfishing, at least once during the last 20 years of the simulation. The number of times that overfished or overfishing status happened can be found using the associated "_n_" metrics. This allows consideration of how often such an event happened on average. The use of the "avg" metrics of SSB and F relative to their MSY reference points then includes the magnitude of the difference, as well, but not the number of years. Consideration of the metrics together allows for a more complete understanding of the performance of the IBMs across scenarios than using only a single metric. These results can be seen in the R Shiny app as well as through a number of different plots in Appendix 6. The IBMs that have the diffuse relationship between SSB/SSB_{MSY} and catch/MSY performed better than the IBMs that have the linear relationship for these metrics.

Another way to examine risks was provided by the chair of the review panel: odds ratios. The odds ratios suggested were the number of realizations resulting in $F>F_{MSY}$ when $SSB<SSB_{MSY}$ multiplied by the number of realizations when $F<F_{MSY}$ when $SSB>SSB_{MSY}$ divided by the product of the number of realizations in the other 2 quadrants of a Kobe plot. These odd ratios were examined for all the base analyses, and no obvious patterns emerged for either the short-term or

long-term summaries. It was suggested that this approach could be more informative if conducted on annual results instead of short- or long-term aggregates. This is an area for future research.

Catch stability

While overfished and overfishing statuses are regulatory issues, there are other aspects of performance that may be of interest to managers. One commonly mentioned is the stability of catch advice. This was explored in these simulations through the use of the "_iav_" metrics for catch. The interannual variability tries to distinguish between an IBM and scenario combination that has little change from one assessment to the next compared to an IBM and scenario that fluctuates wildly from one assessment to the next, even if they have the same mean value. These results can be seen in the R Shiny app as well as through a number of different plots in Appendix 6. Generally, the IBMs that have the diffuse relationship between SSB/SSB_{MSY} and catch/MSY performed better with lower catch variability than the IBMs that have the linear relationship for this metric. The exceptions to this general rule are the 2 CC methods, which performed poorly according to this metric.

Ensemble

By design, the Ensemble model generally had performance that fell in the middle of the orderings for metrics. It had an equal number of IBMs from the 2 groups (diffuse or linear relationships) of IBMs. This resulted in having an overall performance more similar to the IBMs with the linear relationship because the variability in the diffuse relationship IBMs could offset each other. The Ensemble did perform better than the other linear relationship IBMs in terms of catch stability, as would be expected. There could be benefits to using an Ensemble approach if managers are interested in trying to trade off the benefits from both types of IBMs, although it generally followed the results of the linear relationship IBMs, so the amount of trade-off is limited in these simulation results. The performance of the Ensemble can be seen in the R Shiny app as well as through a number of different plots in Appendix 6.

No Retrospective

The no retrospective sensitivity analysis consists of the scenarios CF1A, CO1A, MF1A, MO1A, NF1A, and NO1A for all the IBMs except DLM. The performance of IBMs did not always improve when there was no source of retrospective error. Some of this was due to the fact that the starting conditions were different from the M retrospective source due to the changing reference points for the latter scenarios. In the long term, the average SSB/SSB_{MSY} and catch/MSY were generally closer to 1.0 than either the catch or M retrospective sources (Figure 4.14). This demonstrates a weakness with the scoring algorithm used in this study, values well above SSB or MSY reference points are scored higher than values close to the reference points. This could be taken into account by developing alternative algorithms for deriving the score, such as mean distance from the reference point with a penalty for being on the bad side of the reference point. This would require additional input from managers about their preferences, and so was not pursued in this study but could be done in future analyses.

Despite the shortcomings of the scoring algorithms, there was some change in the ordering of the IBMs when only the no retrospective scenarios were considered, but generally the same groupings held as were seen in the base analyses. See Appendix 6 for some sample scores using the noretro set and the scorer app to create additional results using other sets of metrics.

The performance of the IBMs when no retrospective source is present can perhaps be most clearly seen in the equivalent of Figures 4.3-4.8, where the points represent the mean values from the 1,000 simulations for each IBM and scenario (see Appendix 6). Note that due to the limited number of scenarios, there are fewer panels in these plots. The long term SSB/SSB_{MSY} for the no retrospective source showed generally good performance among IBMs, although the Skate, AIM, and ES-Fstable methods resulted in a mean value below 0.5 for the fishing history of overfishing throughout the base period. Surprisingly, the long-term F/F_{MSY} mean values were above 1.0 for all the IBMs in the no retrospective source scenarios. This may be due to the averaging across years and the fact that F could go well above F_{MSY} but was limited at 0 in how far below F_{MSY} it could go. Despite the high mean values of F/F_{MSY}, the no retrospective source performed better than the catch retrospective source for nearly all IBMs. The M source performed better than the no retrospective source for F/F_{MSY}, but this is most likely due to the high F_{MSY} values associated with the increased M rate. The long-term catch/MSY for the no retrospective source did not have any of the very low values seen for some of the IBMs in the catch retrospective source, and did generally similar to the M retrospective source despite having much higher MSY values. The 3 short-term plots demonstrate the importance of the starting conditions as the fishing history scenarios were often quite different.

The 1,000 point plots for the no retrospective source scenarios were not that different from the associated catch and M retrospective source. The diffuse patterns tended to be less so, and the linear patterns were moved so that they more closely intersected the (1,1) point. These plots are provided in Appendix 6, along with a large number of plots similar to those from the base scenarios.

SCAA

The SCAA sensitivity analysis used scenarios CF1A, CO1A, MF1A, and MO1A. Note, the no retrospective source scenarios were not included due to time limitations. The SCAA model performed better than all the IBMs when the long-term SSB, F, and catch relative to their MSY reference points was used as the scoring metric (Figure 4.16). While the superior performance of the SCAA model held for many metrics, it did not hold for all of them. For example, the set of metrics containing the interannual variability during the entire feedback period and the short-term catch/MSY had SCAA in the lower half of the IBMs order (Figure 4.17).

The performance of the SCAA model can perhaps be most clearly seen in the equivalent of Figures 4.3-4.8, where the points represent the mean values from the 1,000 simulations for each IBM and scenario (see Appendix 6). In the long term, the SCAA model performed near the top of the ordered list, with no IBM consistently performing better than it. In the short term, the SCAA model's performance varied by the fishing history, with some metrics doing well for one fishing history but not the other, leading to a middling performance across these 3 metrics.

The SCAA model had a near linear relationship between the SSB/SSB_{MSY} and catch/MSY points, with better performance for the M than catch retrospective source (Figure 4.18). The SCAA performed with respect to long-term probability of the stock being overfished or undergoing overfishing compared to the IBMs (Figure 4.19). The full suite of figures for the SCAA sensitivity analysis is available in Appendix 6.

Candidates for consideration

Overall, none of the IBMs considered in these simulations performed better than the rho-adjusted SCAA model. Therefore, in situations where an SCAA model is rejected due to a strong

retrospective pattern, there should not be an expectation that an IBM will perform better than the rejected model. These simulations were by necessity limited in scope, so it is not clear that this will always be the case, especially if the retrospective pattern is much larger than examined in this study.

There were 2 groups of IBMs that performed similarly. In situations where the stock is felt to be in poor condition, CC-FSPR, CC-FM, DLM, PBS, ES-Frecent, and Islope should be candidates for consideration because they had better performance rebuilding an overfished stock. In situations where the stock is felt to be in good condition, Skate, AIM, ES-Fstable, ES-FSPR, ES-FM, Ensemble, and Itarget should be candidates for consideration because they had higher short-term catch.

TOR 5. PROVIDE GUIDANCE ON SPECIFIC SITUATIONS THAT ARE AND ARE NOT WELL-SUITED FOR A PARTICULAR CONTROL RULE OR IBM IDENTIFIED IN TOR 4.

Analysis of variance (ANOVA) for evaluating sensitivity of IBMs to simulation factors

For each iteration of each IBM in the simulation, the performance metrics of SSB, F, and catch relative to their true MSY values from the OM were calculated. Short- and long-term metrics were summarized over a number of years for SSB/SSB_{MSY}, F/F_{MSY}, and catch/MSY. Short-term was defined as the average of each metric over the first 6 years of the projection period while long-term was defined as the average over the last 20 years of the projection. The variability among these metrics was analyzed in an ANOVA to determine which factors in the simulation explained the greatest proportion of total variance by IBM. Simulation factors evaluated were: retro_type (catch or M), Fhist (always overfishing or overfishing followed by $F=F_{MSY}$), n_selblocks (1 or 2 fishery selectivity blocks), catch.mult (multiplier of 1.0 or 0.75 on 2-year catch advice), and time.avg (whether the metric was short- or long-term). The following interactions were also included in the analysis: Fhist:time.avg, retro_type:Fhist, retro_type:catch.mult, retro_type:time.avg, catch.mult:time.avg, and Fhist:catch.mult. Factors explaining a large proportion of total variance are interpreted as something that the metric for a particular IBM is sensitive to, while factors explaining relatively small fractions of the total variance are interpreted as factors that a metric for a particular IBM is relatively insensitive to. This approach is similar to explorations of reference point sensitivity to biological parameters in Brooks et al. (2008).

Before performing the ANOVA, the distributions of the metrics were evaluated for normality. The data were found to be quite skewed, and a square root transformation generally performed better than the natural logarithm transformation in reducing skewness, so all values were square-root transformed (see distributions of data with and without transforms and qq plots of residuals to the fitted linear models in Appendix 6). For the transformed data, all of the above listed factors and 2-way interactions were evaluated in separate ANOVAs for each IBM for each of the 3 metrics. The ANOVA was done separately for each IBM so that differences in performance across methods could be summarized. The value in each table reflects the proportion of total variance explained by that factor, and the final row is the total fraction explained by those factors. “NA” indicates a non-significant factor.

The proportion of total variance explained by each factor is reported in Tables 5.1-5.3 for SSB/SSB_{MSY}, F/F_{MSY}, and catch/MSY. Most factors were significant at the 0.01 or 0.001 level for all 3 metrics, though some of those explained less than 1% of the variability. The 2 factors explaining the most for SSB/SSB_{MSY} and F/F_{MSY} were retro_source and time.avg, while for catch/MSY, time.avg explained the most followed by catch.mult and Fhist.

Sensitivity of IBMs to the different factors tended to fall out in 2 distinct groups for the SSB/SSB_{MSY} and F/F_{MSY} metrics. Specifically, DLM, PBS, ES-Frecent, and Islope were relatively insensitive to the cause of the retrospective (retro_type) and variability in their metrics were primarily explained by whether the value reflected a short- versus long-term average (time.avg). IBMs most sensitive to the cause of the retrospective were Itarget, Skate, AIM, ES-FSPR, ES-FM, ES-FStable, and the Ensemble. The 2 catch curve methods (CC-FSPR and CC-FM) were intermediate to these 2 IBM groupings.

Catch/MSY variability was largely explained by whether the value was a short- or long-term average (time.avg) for all IBMs but Itarget and the 2 catch curves. After the factor time.avg, the same 2 IBM groupings identified for SSB/SSB_{MSY} and F/F_{MSY} held: Islope, ES-Frecent, PBS and DLM had the catch multiplier as the second most important factor, while Itarget, Skate, AIM, ES-FSPR, ES-FM, ES-FStable, and the Ensemble had the interaction retro_type:time.avg and Fhist:time.avg as the second or third most important factor.

Conclusions on suitability of IBMs for different scenarios:

1. For several IBMs, the variability among metrics was primarily due to the factor related to the cause of the retrospective pattern (retro_type). Thus, if you cannot identify the likely (or dominant) source of a retrospective pattern (catch or M), then using an IBM sensitive to the retrospective source would be risky to the stock and fishery, and this risk could be avoided by using a method robust to this uncertainty. In this regard, the following methods were more robust to retro type: DLM, PBS, ES-Frecent, Islope, and to some extent the 2 catch curve methods. We anticipate these methods would have more robust performance in situations where a retrospective pattern exists similar to that simulated in this project.
2. DLM variability in SSB/SSB_{MSY} and catch/MSY is mostly explained by short vs. long term means; this is not surprising since the method aimed to rebuild in 10 years, which is longer than the cut-off for short-term metrics; catch multiplier followed by retro_type were the other main explanatory terms.
3. PBS, ES-Frecent, and Islope had similar results to the DLM IBM. PBS and Islope are similar (use slopes to adjust catch advice), but simpler methodologically, so this result is not surprising.
4. The ES-Frecent IBM generally performed well and differently from the other ES methods. This is due to the survey catchability coefficient canceling out in the application of catch advice. The catch advice is derived as Frecent times the expanded survey. Since Frecent is calculated as catch/expanded survey, the effect of the survey catchability is essentially removed. So this method will generally keep the stock at the current fishing mortality rate, although uncertainty in the survey and catch observations can cause it to change. The other ES methods would be expected to depend strongly on the survey catchability estimate. Future research could explore how the ES methods perform with uncertainty and/or bias in the catchability estimate.

5. Different sensitivity of IBMs to the different simulation factors could be useful for exploring future ensemble composition. Based on the scenarios explored herein, future analyses could consider combining the 2 catch curve methods with the 4 IBMs that were insensitive to the source of retrospective pattern for comparison with the results from the current Ensemble to see if an overall improvement (e.g., consistent rebuilding, higher and less variable catch) is achieved.

Heatmaps for identifying similar performance of median performance metrics

Results from the ANOVAs highlighted which factors explained variability in IBM performance and served as the basis for data subsetting and generation of heatmaps that show groupings of IBMs that exhibit similar performance across the factor groupings. The heatmaps were produced by the function heatmap.2 in the package gplots (Warnes et al. 2020).

As with the ANOVA analyses, X/X_{MSY} metrics for each iteration of each IBM in the simulation were summarized over short- and long-term time periods. At the highest level of aggregation, medians by IBM were calculated across all simulations (13 categories [IBMs] with 16,000 values per IBM if all iterations were conducted and/or converged). More granular analyses calculated medians by IBM and single factor (26 categories with 8,000 values for complete cases) and then medians by IBM and 2-way factors (52 categories with 4,000 values for complete cases).

In the heatmap.2 function, the median values for all 3 metrics are first normalized so that values are on the same scale. Dissimilarities are calculated across the categories and dendrograms reorder the categories. In the heatmaps for this analysis, categories are the factors into which all of the data were subsetted (13 IBMs, or 26 IBMxFactor, or 52 IBMxFactor1xFactor2). In the top left of each heatmap, a key indicates the colors that distinguish where the normalized score falls within the distribution, and a superimposed cyan histogram within that key indicates how the data were distributed over the normalized distribution. The title of each heatmap indicates the subsetting of the data, and the right-side y-axis gives the IBM x Factor category for the median value of the metric in that row. The rows are arranged by the dendrogram on the left-side y-axis, and the color in each cell of the heatmap corresponds to the normalized metric and matches the colors in the key at the top left of the heatmap. Comparing the contrasting colors within a given row shows trade-offs between achieving a higher median SSB/SSB_{MSY} (value >1 indicates rebuilding was achieved) vs. the fraction of MSY that could be caught; similarly, the F/F_{MSY} median indicates whether overfishing occurred more or less than 50% of the time (values >1 indicate overfishing). A table of the median performance metric values was produced for each heatmap, and the row order in the tables matches the right-hand side y-axis labels in the Figures to facilitate easier association between the figure and corresponding table.

The median X/X_{MSY} metric by IBM indicates that >50% of iterations achieved rebuilding ($SSB/SSB_{MSY} > 1$) for CC-FSPR and CC-FM, DLM, PBS, ES-Frecent and Islope, in order of greatest median SSB/SSB_{MSY} (Figure 5.1 and Table 5.4). This is seen on the heatmap as the block in the SSB column with the darkest red shading. These same methods avoided overfishing in >50% of iterations (F/F_{MSY} was 0.31-0.87). The trade-off for rebuilding and avoiding overfishing is seen by the corresponding lighter color blocks in the catch column; however, among the IBMs achieving rebuilding >50% of the time, PBS, ES-Frecent, and Islope achieved median catch that was 65-70% of MSY. By contrast, the Skate, AIM, ES-FStable had median SSB/SSBMSY of 0.53-0.7, while median catch was > MSY.

The heatmap for median performance metrics by IBM and retrospective source indicates that rebuilding was more successful when the retrospective source was M rather than catch (Figure 5.2 and Table 5.5). Scenarios where rebuilding was not achieved at least 50% of the time were all in cases where catch was the retrospective source for the IBMs Skate, ES-FM, ES-FSPR, AIM, ES-Fstable, Ensemble, and Itarget. Several IBMs that achieved at least 50% rebuilding also achieved catch that was greater than MSY: ES-FSPR, ES-FM, ES-FStable, AIM, and Skate. Partitioning the results by retrospective source helped identify that methods that performed poorly at the overall IBM level (AIM and Skate) actually can have good performance for some of the scenarios. This finding is consistent with the ANOVA, which identified that the largest source of variability in metrics for AIM and Skate was the retrospective source.

The factor for retrospective source explained a lot of variability in some IBMs, while in others, the time horizon of metric summary explained the most. The heatmap for median performance by IBM and time horizon not unexpectedly shows that the greatest rebuilding is achieved in the long-term (Figure 5.3 and Table 5.6). However, for the methods PBS, DLM, Islope and ES-Frecent, as well as the 2 catch curves, median rebuilding was greater than 1 even in the short-term (slightly greater for some of these), while long-term median metrics were greater than 2. This factor explained the greatest amount of variability in the metrics for these IBMs. For these same IBMs, catch in the short-term was greater than long-term catch.

When median metrics for IBMs and catch multipliers are analyzed, rebuilding probability was greatest when the catch advice was scaled by 0.75 rather than no scaling; as a result, median catch was far below MSY (Figure 5.4 and Table 5.7).

Median metrics for IBM by Fhistory (1=always overfishing then $F=F_{MSY}$, 2=always overfishing) reveal that higher median SSB/SSBMSY is achieved by Fhistory=1 rather than 2, although for the IBMs PBS, DLM, Islope, and Frecent rebuilding is achieved at least 50% of the time under either Fhistory (Figure 5.5 and Table 5.8). The 2 catch curves were insensitive to this for the SSB metric, but median catch was nearly double when Fhistory=1 rather than 2.

Heatmaps for median metrics by IBM and 2 factors were also examined, and the figures and corresponding tables can be found in Appendix 6. This finer granularity of data subsetting did not alter the bigger picture result that a subset of IBMs consistently outperformed others (PBS, DLM, Islope, ES-Frecent, and the 2 catch curve methods).

Conclusions:

1. Looking at IBM performance across all iterations, rebuilding was greatest for the IBMs that were the least sensitive to source of the retrospective pattern (CC-FSPR and CC-FM, DLM, PBS, ES-Frecent and Islope). DLM, PBS, and CC-FM and CC-FSPR achieved the greatest median SSB/SSB_{MSY}, but catch was lowest for CC-FSPR and CC-FM. PBS, Islope and ES-Frecent had the highest median catch among the methods that achieved rebuilding more than 50% of the time.
2. The IBMs ES-FSPR, ES-FM, Skate, Ensemble, ES-Fstable, AIM, and Itarget on the whole had less success rebuilding, though it depended on the scenario. If the retrospective source was M, some of these achieved median rebuilding.
3. The Ensemble method slightly underperformed at rebuilding, with median SSB/SSB_{MSY} of 0.92, but 90% of MSY was attained for median catch. As the ensemble included several of the IBMs that did not achieve rebuilding, the median in the ensemble was weakened by their inclusion. Future research could look at

optimizing ensemble composition based on the performance in scenarios most similar to that expected in a given assessment setting.

4. Trade-offs in risk (overfishing and rebuilding) and rewards (catch) are inherent in management decisions. Balancing median catch close to MSY in the short-term while still maintaining a probability of at least 50% of achieving rebuilding was possible for the ES-Frecent, PBS, DLM, and Islope, although long-term median catch with these methods was far below MSY. This could indicate that these 4 methods are appropriate short-term models for management advice, but while they are employed, other efforts should be invested to return to an age-based model.

Caveats

Although a large number of IBMs, scenarios, and simulations were conducted, there are always limits to the scope of simulation studies. These caveats should be kept in mind as the results of this study are applied in other situations. The biological and fishery characteristics used in this study were based on local groundfish parameters, and other life histories or fishing histories may show different performance for the IBMs. This framework is well-suited to the application of other biological and fishery characteristics and should be used in the future to do so.

There was only a single source causing the retrospective patterns in the age-based data that remained at the same magnitude throughout the feedback period in these simulations. Recent experience in the region has demonstrated this is not how retrospective patterns typically behave: they either increase or decrease in magnitude over successive assessments. While the framework is currently not formulated to handle such a situation of changes in the forcing function for the retrospective pattern, it could be modified to do so with relative ease. This is an area for future research. The framework is well suited to handle either multiple sources of retrospective patterns or different magnitudes of retrospective patterns (with constant forcing functions). These issues should be explored in the future using the framework.

The simulations applied the IBMs every other year in the feedback period to mimic the common use of IBMs in the region. The IBMs are generally easy to apply with little tuning needed, and so they could be applied every year. However, this would create a burden on the management process that would need to be considered before actual application. Large improvements in performance would need to be demonstrated to justify this additional management cost. The reliance of the IBMs on the surveys means that missing years in the time series can have larger impacts than in age-based models. The impact of missing surveys could be examined through this framework by occasionally using the catch advice for 3 years and delaying an assessment. The framework is well-suited to explore this issue, but it was not considered in this study due to time limitations.

Due to the nature of closed-loop simulations, the IBMs had to be applied formulaically with no ability to modify the IBM formulation to a specific set of observations. This is not how actual index-based assessments are conducted. Whoever is conducting a real index-based assessment will examine diagnostics and look for issues with the particular IBM being applied to the actual data available. This may raise questions about either the IBM suitability or the data themselves. Peer review processes are also used to ensure that assumptions for an IBM are met in actual applications, something that could not be easily considered in this study. The results presented here should be considered minimum performance of the IBMs given the formulaic application of the IBMs without consideration of the data available in any given assessment. As noted during the review meeting, a subset of simulations could be examined to see if removing

realizations with diagnostic problems changes the relative performance of any of the IBMs. This would require detailed examination of each assessment within each realization, and so would be quite time-consuming but could address the formulaic application caveat.

The DLM approach makes no explicit assumptions about the population dynamics, life history, or condition of a stock. It requires at least 1 index of abundance and catch information for use, though it could be employed in a manner similar to PBS if no reliable catch information was available. If missing values are present in the abundance index(ices), the DLM fitting procedure generates a probabilistic imputation in those years based upon all of the available data and thus provides an estimate of what would have been observed with quantified uncertainty. If catch information is included as a covariate and missing values are present, those values will need to be imputed prior to fitting the DLM. The time series length necessary to fit the DLM depends upon the number of available abundance indices and their signal-to-noise ratio(s), but at least 25 years of data should generally be used. Because the DLM "learns" more about the behavior of a stock over time as more data is included in fitting, it can be expected that the model fit achieved will improve over subsequent updates. Due to the constraints of this simulation experiment, a single model structure and weakly informative prior distributions were used. In a management setting, different model components and prior assumptions could be compared to determine the optimal structure for a given stock. Similarly, catch advice in this work was set such that the mean forecasted stock trajectory was on track for a 10-year rebuild (or decline) to the 75th percentile of the observed abundance index. This decision rule was only used for this experiment and is not required by the DLM. The DLM produces a probabilistic forecast of future abundance observations based upon assumed harvest levels (when catch information is included). This forecast and uncertainty can be used by managers to aim toward any desired stock level while assessing the risk of different strategies. While still in development, the DLM structure can also be augmented to include additional covariates (e.g., climate information), length- or age-structure, or multispecies information.

The Skate method was developed using the median catch over biomass ratio assuming that this would provide a reasonable relative fishing intensity. This is appropriate for stocks that have not been undergoing overfishing for half or more of the years considered. It would not be expected to perform well in situations where the stock has been undergoing fishing for the majority of time. In such a situation, a different quantile of the distribution of catch over biomass would be preferred. This aspect was not examined in this study due to time constraints but could be an interesting avenue of future research.

The IBMs that change the catch advice based on recent trends in the surveys (e.g., PBS, Islope, DLM) do not appear well suited to applying a reduction to the catch advice. This is because when the stock rebuilds, the surveys do not change much, yet the catch advice continues to decline due to the reduction (catch multiplier of 0.75 in this study). This can lead to the situation where the stock is well above SSB_{MSY} but the catch advice is well below MSY. This would clearly be a frustrating situation for fishers and managers.

TOR 6. CREATE GUIDELINES FOR SETTING BIOLOGICAL REFERENCE POINTS FOR INDEX-BASED STOCKS.

The IBMs evaluated here used a variety of techniques to derive catch advice. The methods ranged from simple deterministic calculations to fitted state-space models but generally fell into three basic categories:

1) IBMs examined the trend in the time series to determine if catch should be increased or decreased (e.g., PBS);
2) IBMs compared the current survey index with some reference time period of the survey index to set catch advise (e.g., Itarget). These methods often calculate the relative fishing mortality (catch/survey index) to determine what level of fishing would result in the reference index level; and
3) IBMs estimated total biomass (e.g., catch curve) then used a proxy fishing mortality reference level derived from other sources to calculate catch advice.

True MSY reference points are based on the population growth rate of a stock. None of the IBMs estimate the population growth rate and thus cannot provide true MSY reference points. Instead, the IBMs attempt to link the raw survey indices and catch data with thresholds that might serve as reference levels, though they should not be considered proxies for MSY reference points. In addition to the structural challenges associated with IBMs, a full exploration of the results to provide guidance on setting reference points with IBMs was not possible due to time constraints. A review of the results, however, highlighted a number of interesting observations.

The trend IBMs (PBS, Islope) examine the slope of the survey and catch in the most recent years and then modify catch based on the slope of the survey index. The performance of the 2 trend methods with regard to maintaining the population near MSY-based references points differed considerably. PBS consistently ranked as one of the top methods for maintaining SSB above SSBmsy and keeping F below Fmsy in the short and long term. The method tended to produce relatively conservative catch advice, however. Islope tended to perform moderately on these same metrics and was consistently ranked in the middle of the different IBMs. While the 2 methods utilize a similar concept to derive catch advice based on recent trends in catch and survey, the specifics of the methods result in different performance. Islope contains multiple options specifying how conservative the catch advice would be. The IBMWG ran the simulations with the least conservative options suggesting that alternative formulations could have resulted in different outputs.

The survey reference level methods (Itarget, AIM, Skate method, and DLM) utilized a range of techniques to set catch advice in order to realize a target survey level. Some methods, such as AIM and the Skate control rule, have a prespecified system for establishing the reference level and produce catch advice to realize that goal while others were designed to incorporate a user-defined reference level. For example, the IBMWG set the reference period for Itarget as the last 25 years of the base period (years 25-50) and the reference level for DLM to the 75th percentile of the survey index. Both of these methods were then evaluated on their performance based on these decisions, when any time period or any percentile could have been selected as the reference level. Because of the range of functional forms within the methods and the range of potential reference levels, the performance of this group of methods for maintaining SSB and F near MSY reference points varied widely. DLM consistently ranked near the top for SSB and F metrics; however, it was quite conservative and typically resulted in lower catches. Had a different reference level been selected, such as the mean of the survey index for a particular set of years or the 55th percentile of the survey index, both the catch advice and the performance of the DLM method could have been different. This group also contained methods that consistently ranked in the middle of the IBMs and those that tended to rank near the bottom (e.g., AIM, Skate control rule). For the Skate control rule, there is an underlying assumption that the time series of data

captures a period when the stock is in a reasonable condition. If the survey index only captures the stock when it is overfished, the Skate control rule will provide catch advice that maintains it at that overfished level. When setting the reference period for many of these methods, at least qualitative information on the status of the stock is very important. Without some information on an appropriate biomass target level, it will be difficult for many of the methods to produce catch advice to achieve the target.

Two methods estimated total biomass with 2 different techniques (catch curve analysis and the expanded biomass method) and utilized proxy target fishing mortality values to produce catch advice. Despite using the same fishing mortality estimates that could be considered proxy F_{MSY} values (F40% from a spawner per recruit analysis and F=M), the 2 methods performed very differently based on the SSB, F, and catch metrics. The catch curve analysis consistently ranked among the top methods when producing catch advice with F40% or F=M, while the expanded biomass method ranked in the middle and low end of the IBMs. Within a single method, such as within the catch curve analysis, the performance was similar when using F40% or F=M, though not identical. The same was true for the expanded biomass method. The expanded biomass method also evaluated the performance of simply using the mean of the most recent catch to provide catch advice (Frecent). This method tended to perform well, typically in the top half of all methods; however, the starting conditions were likely very important for this method. In many of the scenarios, the IBMs started providing catch advice to a population that was near its MSY reference points. Harvesting at MSY when the biomass is at B_{MSY} should maintain the population. The use of this method when the stock is above or below B_{MSY}, however, would maintain the stock at a low level or forgo potential yield.

Despite a huge volume of simulations and results, the output did not produce consistent guidelines for developing IBM reference levels. A few factors to consider did emerge, however. At least a qualitative knowledge of the status of the stock (e.g., good, bad, ok) is important for setting a reference level. Some examples of qualitative stock indicators include increases in exploitable biomass from surveys, expansion in size or age structure in fishery-dependent and independent data sources, and tracking and monitoring the progress of year classes over time. While there are numerous caveats around why a stock might have been in an ok state at some point in the past (e.g., high productive period, limited harvest), without at least some historical knowledge, a reference level could be inadvertently set that maintains a stock in an overfished condition. The actual functional form of the IBM can be as important as the reference level itself. PBS and Islope both examined the trend of the time series in the most recent years, but their performance was quite different. Similarly, the catch curve analysis and the expanded biomass method both estimated total biomass and used the same fishing mortality rate but produced different results. DLM and Itarget have very different functional forms and were run with different reference levels. It would be interesting to see how they performed if given the same reference level. Future work based on this study could focus on 2 aspects: 1) conducting a similar simulation study with a subset of the IBMs and examine a range of different reference levels and different initial biomass conditions for each method to determine when different reference levels work and 2) further exploration of the results of this and other simulation and real word examples for potential relationships between the survey and catch time series and MSY reference points. As always, all results need to be considered in light of the assumptions, starting conditions, and decisions made by the working group. Different assumptions and starting conditions could produce different results.

CONCLUSIONS AND RECOMMENDATIONS

General conclusions and recommendation

- For stocks that have had an age-based assessment rejected due to a strong retrospective pattern, there is no expectation that an index-based assessment will perform better than a rho-adjusted SCAA analysis.
- The performance of an index-based assessment in a specific situation can be analyzed through the framework developed for this project but requires specific hypotheses about possible sources of the retrospective pattern.
    - The IBMWG recommends this framework be used for all assessments that have changed from age-based to index-based due to retrospective patterns, using biological and fishery settings appropriate for that stock to ensure the selected IBM has a high probability of providing reasonable catch advice.
- The IBMWG recommends future research be conducted to both analyze the results of this study in more detail as well as build on this study to address other questions.
    - Examples of future research on the results of this study include:
        - Filtering the results to have a common value for 1 metric, such as probability of overfishing in the long term, to see how the other metrics compare across IBMs. This is conceptually similar to tuning the IBMs to produce a common risk outcome but is easier to apply across multiple scenarios;
        - Digging into the detailed results to look for reasons why IBMs of similar type did not always perform similarly while IBMs of different types sometimes did; and
        - Exploring the detailed results to look for observed data that could be used to set biological reference points. This is a big task that will probably require additional simulations to fully address.
    - Examples of future research building on this study include:
        - Examining life history and fishery characteristics beyond groundfish;
        - Applying multiple or additional sources to create the retrospective patterns;
        - Applying different magnitudes of retrospective patterns;
        - Using state-space models to see if performance is better than SCAA; and
        - Tuning the framework to specific data for a stock to examine performance of IBMs in a specific situation.

Process-related conclusions and recommendations

- More time needs to be allocated to topic-based research tracks in order to fully address the ToRs.
- COVID-19 travel limitations required weekly meetings with remote collaborations. This may be a useful approach for future topic-based research tracks but may be less useful for stock-specific research tracks. The time-certain weekly meetings helped track the programming and decision making progress necessary for the large simulation study. This may not apply well to stock-specific research tracks.

- GitHub was helpful for coordinating coding among multiple programmers.
    - A training session early on would ensure everyone is able to work together efficiently
- This was a big project that required lots of computing power. The cooperation of network users not involved in the project to free up computing time was greatly appreciated.
- Fast internet speed is an issue for moving large files and large numbers of files
    - Cloud computing would have been helpful
- Google Docs with prompts before a meeting allowed asynchronous contributions which could then be built on during the meeting
- Google Docs are handy for meeting notes but not great for report writing
    - Need to establish a workflow all the way through to final report (including 508 compliance)
    - Would be helpful to be able to use Rmarkdown so we don't have to update tables and figures in report by hand
- Project management software (e.g., Jira) could be useful but would require training

REFERENCES

A'mar ZT, Punt AE, Dorn MW. 2010. Incorporating ecosystem forcing through predation into a management strategy evaluation for the Gulf of Alaska walleye Pollock (Theragra chalcogramma) fishery. Fish Res. 102(1): 98-114.

Anderson SC, Cooper AB, Jensen OP, Minto C, Thorson JT, Walsh JC, Afflerbach J, Dickey-Collas M, Kleisner KM, Longo C, et al. 2017. Improving estimates of population status and trend with superensemble models. Fish Fish. 18(4):732-741.

Brooks EN and Legault CM. 2016. Retrospective forecasting – evaluating performance of stock projections for New England groundfish stocks. Can J Fish Aquat Sci. 73(6):935-950.

Brooks EN, Shertzer KW, Gedamke T, Vaughan DS. 2008. Stock assessment of protogynous fish: evaluating measures of spawning biomass used to estimate biological reference points. Fisheries Bulletin. 106(1):12-23.

Carruthers TR and Hordyk A. 2018. The Data-Limited Methods Toolkit (DLMTool). an R package for informing management of data-limited population. Methods Ecol Evol. 9(12):2388-2395.

Carruthers TR, Punt AE, Walters CJ, MacCall A, McAllister MK, Dick EJ, Cope J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish Res. 153:48-68.

Carruthers TR, Kell LT, Butterworth D, Maunder M, Geromont H, Walters CJ, McAllister M, Hillary R, Levontin P, Kitakado T, et al. 2015. Performance review of simple management procedures. ICES J Mar Sci. 73(2):464-482. Accessible at: doi:10.1093/icesjms/fsv212.

Deroba J, Shepherd G, Grégoire F, Nieland J, Rago P. 2010. Stock assessment of Atlantic mackerel in the Northwest Atlantic for 2010. Transboundary Resources Assessment Committee. Reference Document No. 2010/01. 59 p.

Dick EJ and MacCall AD. 2011. Depletion-Based Stock Reduction Analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fish Res. 110(2):331-341.

Fay G, Punt AE, Smith ADM. 2011. Impacts of spatial uncertainty on performance of age-structure based harvest strategies for blue eye trevalla (Hypheroglyphe antarctica). Fish Res. 110(3):391-407.

Geromont HF and Butterworth DS. 2015a. Generic management procedure for data-poor fisheries: forecasting with few data. ICES J Mar Sci. 72(1):251-261. Accessible at: doi:10.1093/icesjms/fst232.

Geromont HF and Butterworth DS. 2015b. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. ICES J Mar Sci. 72(1):262-274. Accessible at: doi:10.1093/icesjms/fsu017.

Kristensen K, Nielsen A, Berg CW, Skaug H, Bell BM. 2016. TMB: automatic differentiation and Laplace approximation. J Stat Softw. 70(5):1-21. Accessible at: doi:10.18637/jss.v070.i05.

Legault, C.M., Alade, L., Gross, W.E., Stone, H.H. 2014. Stock Assessment of Georges Bank Yellowtail Flounder for 2014. TRAC Ref. Doc. 2014/01. 214 p. Available from http://www.nefsc.noaa.gov/saw/trac/

MacCall AD. 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES J Mar Sci. 66(1):2267-2271.

Martell S and Froese R. 2013. A simple method for estimating MSY from catch and resilience. Fish Fish. 14(4):504-514. Accessible at: doi:10.1111/j.1467-2979.2012.00485.x.

McNamee, J., Fay, G., Cadrin, S. 2015. Data Limited Techniques for Tier 4 Stocks: An alternative approach to setting harvest control rules using closed loop simulations for management strategy evaluation. Final report to the Mid Atlantic Fishery Management Council.

Miller TJ, Hare JA, Alade LA. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern New England yellowtail flounder. Can J Fish Aquat Sci. 73(8):1261-1270. Accessible at: doi:10.1139/cjfas-2015-0339.

Miller TJ and Stock BC. 2020. WHAM: a state-space age-structured assessment model. Version 1.0. NOAA Fisheries. Accessible at: https://timjmiller.github.io/wham/

Mohn R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J Mar Sci. 56(4):473-488. Accessible at: doi:10.1006/jmsc.1999.0481.

[NEFSC] Northeast Fisheries Science Center. 2002a. Assessment of 20 Northeast groundfish stocks through 2001. Report of the Groundfish Assessment Review Meeting (GARM); 2002 Oct 8-11; Woods Hole (MA): US Dept Commer Northeast Fish Sci Cent Ref Doc. 02-16. 544 p. Accessible at: https://repository.library.noaa.gov/view/noaa/5344

[NEFSC] Northeast Fisheries Science Center. 2002b. Re-evaluation of biological reference points for New England groundfish. US Dept Commer Northeast Fish Sci Cent Ref Doc. 02-04; 395 p. Accessible at: https://repository.library.noaa.gov/view/noaa/5316

[NEFSC] Northeast Fisheries Science Center. 2005. Assessment of 19 Northeast groundfish stocks through 2004. Report of the 2005 Groundfish Assessment Review Meeting (GARM); 2005 Aug. 15-19; Woods Hole (MA): US Dept Commer Northeast Fish Sci Cent Ref Doc. 05-13; 499 p. Accessible at: https://repository.library.noaa.gov/view/noaa/5328

[NEFSC] Northeast Fisheries Science Center. 2008. Assessment of 19 Northeast groundfish stocks through 2007. Report of the 3rd Groundfish Assessment Review Meeting (GARM); 2008 Aug. 4-8; Woods Hole (MA): US Dept Commer Northeast Fish Sci Cent Ref Doc. 08-15; 884 p. Accessible at: https://repository.library.noaa.gov/view/noaa/5227

[NEFSC] Northeast Fisheries Science Center. 2012. 53rd Northeast regional Stock Assessment Workshop (SAW) assessment report. US Dept Commer Northeast Fish Sci Cent Ref Doc. 12-05; 559 p. Accessible at: https://repository.library.noaa.gov/view/noaa/4053

[NEFSC] Northeast Fisheries Science Center. 2015a. Stock assessment update of 20 Northeast groundfish stocks through 2014. US Dept Commer Northeast Fish Sci Cent Ref Doc. 15-24; 238 p. https://repository.library.noaa.gov/view/noaa/5293

[NEFSC] Northeast Fisheries Science Center. 2015b. 60th Northeast regional Stock Assessment Workshop (SAW) assessment report. US Dept Commer Northeast Fish Sci Cent Ref Doc. 15-08; 870 p. Accessible at: https://repository.library.noaa.gov/view/noaa/5293

[NEFSC] Northeast Fisheries Science Center. 2017. Operational assessment of 19 Northeast groundfish stocks, updated through 2016. US Dept Commer Northeast Fish Sci Cent Ref Doc. 17-17; 259 p. Accessible at: https://repository.library.noaa.gov/view/noaa/16091

[NEFSC] Northeast Fisheries Science Center. 2019. Operational Assessment of 14 Northeast Groundfish Stocks, Updated Through 2018. US Dept Commer, Northeast Fish Sci Cent Ref. Doc. 22-06. 205 p.

[NRCC] Northeast Region Coordinating Council. 2018. Description of New England and Mid-Atlantic Region Stock Assessment Process [PDF]. 16 p. Accessible at: https://s3.amazonaws.com/nefmc.org/Stock-assessment-process-June2020.pdf

Pershing AJ, Alexander MA, Hernandez CM, Kerr LA, Le Bris A, Mills KE, Nye JA, Record NR, Scannell HA, Scott JD, et al.. 2015. Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. Science. 350(6262):809-812. Accessible at: doi:10.1126/science.aac9819.

Punt AE. 2017. Strategic management decision-making in a complex world: quantifying, understanding, and using trade-offs. ICES J Mar Sci. 74(2):499-510.

Punt AE, Dorn MW, Haltuch MA. 2008. Evaluation of threshold management strategies for groundfish off the US west coast. Fish Res. 94(3):251-266.

Punt AE, Butterworth DS, de Moor CL, De Oliveira JAA, Haddon M. 2016. Management strategy evaluation: best practices. Fish Fish. 17(2):303-334.

Punt A E, Tuck GN, Day J, Canales C M, Cope JM, de Moor CL, De Oliveira JAA, Dickey-Collas M, Elvarsson B, Haltuch MA, et al. 2020. When are model-based stock assessments rejected for use in management and what happens then? Fish Res. 224(105465).

Rosenberg AA, Kleisner KM, Afflerbach J, Anderson SC, Dickey-Collas M, Cooper AB, Fogarty MJ, Fulton EA, Gutiérrez NL, Hyde KJW, et al. 2017. Applying a new ensemble approach to estimating stock status of marine fisheries around the world. Conserv Lett. 11(1):1-9. Accessible at: doi:10.1111/conl.12363.

Rousseeuw PJ, Ruts I, Tukey JW. 1999. The bagplot: a bivariate boxplot. The Am Stat. 53(4):382-387.

Spence MA, Blanchard JL, Rossberg AG, Heath MR, Heymans JJ, Mackinson S, Serpetti N, Speirs DC, Thorpe RB, Blackwell PG. 2018. A general framework for combining ecosystem models. Fish Fish. 19(6):1031-1042.

Stewart IJ and Hicks AC. 2018. Interannual stability from ensemble modelling. Can J Fish Aquat Sci. 75(12):2109-2113.

Stock BC and Miller TJ. 2021. The Woods Hole Assessment Model (WHAM): a general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fish Res. 240:105967.

Tableau A, Collie JS, Bell RJ, Minto C. 2019. Decadal changes in the productivity of New England fish populations. Can J Fish Aquat Sci. 76(9):1528-1540.

Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, et al. 2020. gplots: Various R Programming Tools for Plotting Data. R package version 3.1.0. ScienceOpen. Accessible at: https://CRAN.R-project.org/package=gplots (heatmap.2 function)

Wiedenmann J. 2015. Application of data-poor harvest control rules to Atlantic mackerel. Final report to the Mid-Atlantic Fishery Management Council. Southeast Data, Assessment, and Review. Report No. SEDAR46-RD-09.

Wiedenmann J and Jensen OP. 2018. Uncertainty in stock assessment estimates for New England groundfish and its impact on achieving target harvest rates. Can J Fish Aquat Sci. 75(3):342-356. Accessible at: doi:10.1139/cjfas-2016-0480.

Wiedenmann J, Free CM, Jensen OP. 2019. Evaluating the performance of data-limited methods for setting catch targets through application to data-rich stocks: a case study using Northeast U.S. fish stocks. Fish Res. 209:129-142.

Xu H, Miller TJ, Hameed S, Alade L, Nye JA. 2018. Evaluating the utility of the Gulf Stream Index for predicting recruitment of Southern New England-Mid Atlantic yellowtail flounder. Fish Ocean. 27(1): 85-95.

TABLES

Table 1.1. Biological, index, and fishery characteristics for the age-structured operating model. N=population abundance; M=natural mortality; Fhist=fishing mortality rate history.

Age class	Initial N (x1000) (M=0.2, 1970 selectivity)	Initial N (x1000) (M=0.2, latter selectivity changed)	M (no change, 1970-1994)	M (1995-2059, Fhist =1)	M (1995-2059, Fhist =2)	Maturity	Weight (kg)	Selectivity (no change or 1970-1994)	Selectivity (1995-2059 if changed)
1	5879	6701	0.2	0.32	0.36	0.04	0.15	0.07	0.02
2	4667	5430	0.2	0.32	0.36	0.25	0.5	0.17	0.05
3	3546	4326	0.2	0.32	0.36	0.6	0.9	0.36	0.12
4	2481	3306	0.2	0.32	0.36	0.77	1.4	0.61	0.27
5	1561	2318	0.2	0.32	0.36	0.85	2	0.81	0.5
6	900	1423	0.2	0.32	0.36	0.92	2.6	0.92	0.74
7	494	764	0.2	0.32	0.36	1	3.2	0.97	0.89
8	266	376	0.2	0.32	0.36	1	4.1	0.99	0.96
9	142	178	0.2	0.32	0.36	1	5.9	1	0.99
10+	160	154	0.2	0.32	0.36	1	9	1	1

Table 1.2. Median values of Mohn's rho when fitting Age-Structured Assessment Program (ASAP)-like configuration of Woods Hole Assessment Model (WHAM) to 50 simulated data sets. M=natural mortality; SSB=spawning stock biomass; F=fishing mortality rate; R=recruitment.

Mis-specification	F history	Selectivity Assumption	SSB	F	R
None	1	1	0.02	-0.01	0.08
None	2	1	-0.01	0.02	0.06
None	1	2	0.01	0.00	0.10
None	2	2	0.00	0.02	0.06
Catch	1	1	0.40	-0.28	0.31
Catch	2	1	0.53	-0.39	0.54

Mis-specification	F history	Selectivity Assumption	SSB	F	R
Catch	1	2	0.43	-0.30	0.29
Catch	2	2	0.50	-0.41	0.47
M	1	1	0.54	-0.37	0.65
M	2	1	0.48	-0.38	0.71
M	1	2	0.51	-0.36	0.58
M	2	2	0.45	-0.36	0.66

Table 1.3. Stock recruitment characteristics of the population. R0=unexploited recruitment; AR1=autoregressive function with 1 year lag.

steepness	0.75
R0 (x1000)	10000
AR1 correlation recruitment deviations	0.4
AR1 conditional standard deviation	0.5

Table 1.4. Factors and levels for the full factorial design of scenarios. M=natural mortality; F_{MSY}=fishing mortality rate at maximum sustainable yield; A50=age at 50% selectivity; IBM=index-based method.

Factor	Levels
Cause of the retrospective pattern	Missing catch Increased M
Fishing history	Always overfishing (2.5Fmsy) Overfishing 1970-1994, then Fmsy
Fishery Selectivity	Constant (A50=3.57) Two time blocks: 1970-1994, A50=3.57, 1995-2059, A50 = 5
Index-Based Methods	13 IBMs (see Term of Reference 2 for details)
Catch advice multiplier	1.00 0.75

Table 1.5. Scenario and changes for the one-offs that were planned. Only the "no retrospective pattern" and statistical catch-at-age ("SCAA") scenarios could be conducted due to time limitations.

Scenario	Changes
No retrospective pattern	No change in catch reporting or change in M
SCAA	Use SCAA as IBM
Pelagic biology and fishery	Life history characteristics such as M, maturity, weights at age, fishery selectivity, etc.
Fishing history	Underfishing then overfishing
Fishing history	Always underfishing
Retrospective source	Both missing catch and increased M
Retrospective source	Change in survey catchability
Retrospective source	Fishery selectivity changes (doming)
Retrospective source	Big cohort
Selectivity	Fishery domed
Selectivity	Both fishery and survey domed
Recruitment	No autocorrelation in deviations
Survey observations	Spring and fall in same year
IBM	First Second Derivatives (FSD)
IBM	Expand Survey wrong catchability
IBM	Islope alternative options
IBM	Itarget alternative options
Catch advice cap	Limit amount of inter-assessment change
Biology	Weights at age change over time
Missing data	Survey data not available in some years

Table 1.6. Equilibrium productivity characteristics of the population under alternative scenario assumptions. M=natural mortality; F_{MSY}=fishing mortality rate at maximum sustainable yield; F40=fishing mortality that reduces spawners per recruit to 40% of unexploited value; SSB_{MSY}=spawning stock biomass at maximum sustainable yield; SSB40=spawning stock biomass when fished at F40; MSY=maximum sustainable biomass; Y40=yield (catch in weight) when stock fished at F40; SSB/R(0)=spawning stock biomass per recruit when F=0; SSB/R(F_{MSY})=spawning stock biomass per recruit when F=F_{MSY}; SSB/R(F40)=spawning stock biomass per recruit when F=F40; F(0.5SSB_{MSY})=fishing mortality rate that reduces spawning stock biomass to half of spawning stock biomass at maximum sustainable yield; F(0.1SSB_{MSY})=fishing mortality rate that reduces spawning stock biomass to one tenth of spawning stock biomass at maximum sustainable yield.

Selectivity assumption	1	1	1	2	2	2
M	0.2	0.32	0.36	0.2	0.32	0.36
Fmsy	0.17	0.31	0.38	0.23	0.48	0.67
F40	0.17	0.28	0.34	0.23	0.47	0.60
SSBmsy (mt)	47633	15973	11613	49144	16834	11979
SSB40 (mt)	48636	17122	12838	48636	17122	12838
MSY (mt)	6544	3334	2807	7238	3559	2971
Y40 (mt)	6543	3327	2796	7237	3559	2966
SSB/R (0)	14.08	4.96	3.72	14.07	4.96	3.72
SSB/R (Fmsy)	5.53	1.88	1.37	5.68	1.96	1.41
SSB/R (F40)	5.63	1.98	1.49	5.63	1.98	1.49
F (0.5SSBmsy)	0.32	0.62	0.77	0.48	1.24	1.78
F (0.1SSBmsy)	0.70	1.50	1.89	1.44	4.27	5.78

Table 1.7. Sampling attributes for observations from the 2 survey indices and the fishery. CV=coefficient of variation.

	Index 1	Index 2	Fishery
CV	0.3	0.4	0.1
Multinomial age composition sample size	100	100	200

Table 2.1. Data-limited methods currently used in the New England and Mid-Atlantic regions. GB=Georges Bank; GOM=Gulf of Maine.

Method	Stocks (24 total)
Plan B Smooth (PBS)	3 (GB cod, two monkfish stocks)
An index method (AIM)	2 (two stocks of windowpane flounder; previously used for many more stocks)
Custom model for skate complex	7 skate species (winter, little, thorny, smooth, clearnose, rosette, barndoor)
Stock-specific custom models using the survey index	5 (two silver hake stocks, two squid species, ocean pout)
Expanded survey biomass (different approaches used to set target exploitation fraction)	5 (witch flounder, GB yellowtail flounder, GOM winter flounder, two stocks of red hake)
First second derivative (FSD)	1 (Atlantic halibut)
Depletion corrected average catch (DCAC)	1 (red crab)

Table 2.2. Index-based methods used in this study showing equations and details for each method. C=catch (in weight); targ=target; y=year; F=fishing mortality rate; M=natural mortality; Z=total mortality.

Method	Details
Plan B Smooth (PBS)	$C_{targ,y+1:y+2} = \overline{C}_{3,y}(e^{\lambda})$, Where $\overline{C}_{3,y}$ is the most recent three year average $\overline{C}_{3,y} = \frac{1}{3}\sum_{t=1}^{t=3} C_{y-t}$, and λ is the slope of a log linear regression of a LOESS average index of abundance (spring and fall) with span = 0.3: $\widehat{I}_y = LOESS(\overline{I}_y)$, and $LN(\widehat{I}_y) = b + \lambda y$
Islope (Islope)	$C_{targ,y+1:y+2} = 0.8\overline{C}_{5,y}(1 + 0.4e^{\lambda})$, where $\overline{C}_{5,y}$ is the most recent five-year average catch through year y-1: $\overline{C}_{5,y} = \frac{1}{5}\sum_{t=1}^{t=5} C_{y-t}$,

Method	Details
	and λ is the slope of a log-linear regression of the most recent five years of the averaged index: $LN(\bar{I}_y) = b + \lambda y$.
Itarget (Itarget)	$C_{targ,y+1:y+2} = \left[0.5 C_{ref} \left(\frac{\overline{I}_{5,y} - I_{thresh}}{I_{target} - I_{thresh}} \right)\right]$ $\overline{I}_{5,y} \geq I_{thresh}$ $C_{targ,y+1:y+2} = \left[0.5 C_{ref} \left(\frac{\overline{I}_{5,y}}{I_{thresh}} \right)^2\right]$ $\overline{I}_{5,y} < I_{thresh}$ C_{ref} is the average catch over the reference period (years 26 through 50): $C_{ref} = \frac{1}{25} \sum_{y=26}^{y=50} C_y$, Itarget is 1.5 times the average index over the reference period: $I_{target} = 1.5 \left(\frac{1}{25} \sum_{y=26}^{y=50} \overline{I_y} \right)$, $I_{thresh} = 0.8\ I_{target}$, and $\overline{I}_{5,y}$ is the most recent five year average of the combined spring and fall index: $\overline{I}_{5,y} = \frac{1}{5} \sum_{t=1}^{t=5} \overline{I}_{y-t+1}$.

Method	Details
Skate (Skate)	$C_{targ,y+1:y+2} = F_{rel}\overline{I}_{3,y}$, where $F_{rel} = median\left(\frac{\overline{C}_{3,\mathbf{Y}}}{\overline{I}_{3,\mathbf{Y}}}\right)$ is the median relative fishing mortality rate calculated using a 3 year moving average of the catch and average survey index across all available years (**Y**): $\overline{C}_{3,y} = \frac{1}{3}\sum_{t=1}^{t=3} C_{y-t}$ $\overline{I}_{3,y} = \frac{1}{3}\sum_{t=1}^{t=3} I_{y-t+1}$
An Index Method (AIM)	AIM first calculates the annual relative F, $F_{rel,y} = \frac{C_y}{\frac{1}{3}\sum_{t=1}^{t=3}\overline{I}_{y-t+1}}$ and the annual replacement ratio, $\Psi_y = \frac{\overline{I}_y}{\frac{1}{5}\sum_{t=1}^{t=5}\overline{I}_{y-t}}$. These values are used in a regression: $LN(\Psi_y) = b + \lambda LN(F_{rel,y})$ to determine $F_{rel,*}$, which is the value of $F_{rel,y}$ where the predicted $\Psi = 1$ or $LN(\Psi) = 0$. $F_{rel,*}$ is called either the "stable" or "replacement" F, and is used to calculate the target catch with $C_{targ,y+1:y+2} = \overline{I}_y F_{rel,*}$.
Dynamic Linear Model (DLM)	$SI_t = [1\ 0\ 1]\left[\theta_{int,t}\ \theta_{trend,t}\ \beta_{CA,t}\right]' + \upsilon_t,\ \upsilon_t \sim N(0,V), V \sim IG(a,b)$ $\left[\theta_{int,t}\ \theta_{trend,t}\ \beta_{CA,t}\right] = [1\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 1]\left[\theta_{int,t-1}\ \theta_{trend,t-1}\ \beta_{CA,t-1}\right] + \left[0\ \omega_{1,t}\ \omega_{2,t}\right], \omega_t \sim N(0,W), W \sim IW(\Psi,\upsilon)$ where SI_t is the survey index at time t, θ_{int} is the intercept of the dynamic trend, θ_{trend} is the

Method	Details
	slope of the dynamic trend, β_{CA} is the regression coefficient on catch anomalies, ν is the observation error, ω is the evolutions, V is the observation variance, W is the evolution variance, IG denotes inverse gamma distribution, and IW denotes inverse Wishart distribution.
Expanded survey biomass method 1 $F_{40\%}$ (ES-FSPR)	$C_{targ,y+1:y+2} = B_{\bar{I},y}\mu_{targ}$, where $B_{\bar{I}}$ is the average of estimated fully-selected biomass from each survey: $B_{\bar{I},y} = \frac{1}{2}\left(\frac{I_{spr,y}}{q_{spr}} + \frac{I_{fall,y-1}}{q_{fall}}\right)$, and target exploitation fraction, μ_{targ} is calculated with: $\mu_{targ} = \frac{F_{targ}}{Z_{targ}}\left(1 - e^{-Z_{targ}}\right)$ $F_{targ} = F_{40\%}$ $Z_{targ} = F_{targ} + M$
Expanded survey biomass method 2 $F =$ AIM replacement (ES-Fstable)	Same as the above expanded survey method, but with μ_{targ}being the stable exploitation fraction calculated using the same approach as AIM (see above) to calculate $F_{rel,*}$.
Expanded survey biomass method 3 $F = M$ (ES-FM)	Same as the above expanded survey methods, with the target exploitation rate set to the assumed M: $\mu_{targ} = \frac{F_{targ}}{Z}\left(1 - e^{-Z}\right)$ $F_{targ} = M$ $Z = F_{targ} + M$

Method	Details
Expanded survey biomass method 4 $F =$ recent average (ES-FM)	Same as the above expanded survey methods, with the target exploitation fraction set to the most recent three year average exploitation fraction: $\mu_{targ} = \frac{\sum_{y-2}^{y} \mu_y}{3}$ $\mu_y = \frac{C_{y-1}}{B_{\bar{I},y}}$
Catch curve Method 1 $F_{targ} = F_{40\%}$ (CC-FSPR)	$C_{targ,y+1:y+2} = \frac{F_{targ}}{Z_{avg,y}} B_{cc,y} \left(1 - e^{-Z_{avg,y}}\right)$, where B_{cc} is the estimated biomass: $B_{cc,y} = \frac{C_{y-1}}{\frac{F_{avg,y}}{Z_{avg,y}} \left(1 - e^{-Z_{avg,y}}\right)}$, with $Z_{avg,y} = \frac{Z_{spring,y} + Z_{fall,y-1}}{2}$, $F_{avg,y-1} = Z_{avg,y-1} - M$, and $F_{targ} = F_{40\%}$
Catch curve Method 2 $F_{targ} = M$ (CC-FM)	Same as catch curve method 1 above, but with $F_{targ} = M$
Ensemble Method	Median of catch advice provided by AIM, CCFSPR, ES-Frecent, ES-FSPR, Islope, Itarget, PBS, and Skate methods

Table 2.3. Summary table comparing data requirements and arbitrary constants of the index-based methods. Table kindly provided by Dr. Paul Rago, chair of the review panel. yr=year; F=fishing mortality rate; Z=total mortality; spr=spring; M=natural mortality, pt=point; avg=average; SSB_{MSY}=spawning stock biomass at maximum sustainable yield; SSBthresh=threshold spawning stock biomass set equal to half of SSB_{MSY}.

	Data Requirements				External Parameters		Arbitrary Constants				
Method	Catch	Survey	Survey Age Comp	All Assessment data	Natural Mortality	Survey Catchability	1	2	3	4	5
AIM	X	X					5 yr avg Index for Replacement ratio	3-yr avg for relative F	avg of 2 surveys		
CC-FM	X		X		X		Boundary for Z	Z avg for spr and fall	Selection of n points in regression	Sum of 3 yrs of survey age freq	
CC-FSPR	X		X		X		Boundary for Z	Z avg for spr and fall	Selection of n points in regression	Sum of 3 yrs of survey age freq	
DLM	X	X					Ref level = 75%ile of Survey Indices	10 yr rebuild Period	2 yr forecast averaged with last 2 yr advice		

	Data Requirements				External Parameters		Arbitrary Constants				
Ensemble	X	X	X		X	X	All assumptions in AIM, CCFspr, ES_recent, ES_Fspr, Islope, Itarget, PBS, Skate				Median of catch over all methods
ES-FM	X	X			X	X	F target=M				
ES-Frecent	X	X			X	X	3 yr avg for exploitation				
ES-FSPR	X	X			X	X	all data for SSB/R and 40% SPR crit				
ES-Fstable	X	X				X	Linked to AIM				
Islope	X	X					5 pt regression	0.8 multiplier	0.4 multplier on slope	5 yr avg catch	avg of 2 surveys
Itarget	X	X					25 yr ref period	25 yr target period	Threshold =0.8*Target	5 year avg for avg of spring and fall	target=1.5* avg index in ref period
PBS	X	X					Span=0.3	avg of 2 surveys	3 yr regression of smoothed estimates		

	Data Requirements				External Parameters		Arbitrary Constants				
SCAA	X	X	X	X	X		7 yr peel for Mohn's Rho	All assessment data plus all constraints on model formulation			
Skate	X	X					3 yr avg Catch	3-yr avg Index	75%ile of Survey for SSBmsy, 1/2 of that for SSBthresh	Use of Median for central tendency	

Table 3.1. Performance metrics collected from simulations grouped by category with fundamental objective and shorthand notation described. MSY=maximum sustainable yield; SSB=spawning stock biomass; F=fishing mortality rate.

Fundamental Objective	Category	Metric	Shorthand
Maximize catch relative to MSY	Catch	Mean short-term catch	s_avg_catch
Maximize catch relative to MSY	Catch	Mean long-term catch	l_avg_catch
Maximize catch relative to MSY	Catch	Mean short-term catch/MSY	s_avg_catch_msy
Maximize catch relative to MSY	Catch	Mean long-term catch/MSY	l_avg_catch_msy
Legislative	Catch	Frequency catch > MSY 2 out of 3 years	l_prop_g_msy_2_of_3
Minimize variability in catch	Variability in catch	Short-term standard deviation of catch	s_sd_catch
Minimize variability in catch	Variability in catch	Long-term standard deviation of catch	l_sd_catch
Minimize variability in catch	Variability in catch	Short-term interannual variation of catch	s_iav_catch
Minimize variability in catch	Variability in catch	Long-term interannual variation of catch	l_iav_catch
Minimize variability in catch	Variability in catch	Long-term mean interannual variability in catch (A'mar et al. 2010)	l_iav_catch
Avoid low biomass or overfished state	Biomass of target species	Short-term mean SSB/SSB_{MSY}	s_avg_ssb_ssbmsy
Avoid low biomass or overfished state	Biomass of target species	Long-term mean SSB/SSB_{MSY}	l_avg_ssb_ssbmsy
Avoid low biomass or	Biomass of target	Short-term proportion	s_is_less_01_bmsy

Fundamental Objective	Category	Metric	Shorthand
overfished state	species	of realizations with ≥ 1 year SSB < 0.1 SSB_{MSY}	
Avoid low biomass or overfished state	Biomass of target species	Long-term proportion of realizations with ≥ 1 year SSB < 0.1 SSB_{MSY}	l_is_less_01_bmsy
Avoid low biomass or overfished state	Biomass of target species	Short-term proportion of realizations with ≥ 1 year SSB < 0.5 SSB_{MSY}	s_is_less_05_bmsy
Avoid low biomass or overfished state	Biomass of target species	Long-term proportion of realizations with ≥ 1 year SSB < 0.5 SSB_{MSY}	l_is_less_05_bmsy
Avoid low biomass or overfished state	Biomass of target species	Short-term proportion of realizations with ≥ 1 year SSB ≥ SSB_{MSY}	s_is_ge_bmsy
Avoid low biomass or overfished state	Biomass of target species	Long-term proportion of realizations with ≥ 1 year SSB ≥ SSB_{MSY}	l_is_ge_bmsy
Avoid low biomass or overfished state	Biomass of target species	Short-term mean number of years among realizations with SSB ≥ SSB_{MSY}	s_n_ge_bmsy
Avoid low biomass or overfished state	Biomass of target species	Long-term mean number of years among realizations with SSB > SSB_{MSY}	l_n_ge_bmsy
Avoid low biomass or overfished state	Biomass of target species	Short-term mean number of years among realizations with SSB < 0.1 SSB_{MSY}	s_n_less_01_bmsy
Avoid low biomass or overfished state	Biomass of target species	Long-term mean number of years	l_n_less_01_bmsy

Fundamental Objective	Category	Metric	Shorthand
		among realizations with SSB < 0.1 SSB_{MSY}	
Avoid low biomass or overfished state	Biomass of target species	Short-term mean number of years among realizations with SSB < 0.5 SSB_{MSY}	s_n_less_05_bmsy
Avoid low biomass or overfished state	Biomass of target species	Long-term mean number of years among realizations with SSB < 0.5 SSB_{MSY}	l_n_less_05_bmsy
Avoid overfishing	Fishing mortality	Short-term proportion of realizations with ≥ 1 year F > F_{MSY}	s_is_gr_fmsy
Avoid overfishing	Fishing mortality	Long-term proportion of realizations with ≥ 1 year F > F_{MSY}	l_is_gr_fmsy
Avoid overfishing	Fishing mortality	Short-term proportion of realizations with ≥ 1 year F < F_{MSY}	s_is_less_fmsy
Avoid overfishing	Fishing mortality	Long-term proportion of realizations with ≥ 1 year F < F_{MSY}	l_is_less_fmsy
Avoid overfishing	Fishing mortality	Short-term mean number of years F > F_{MSY}	s_n_gr_fmsy
Avoid overfishing	Fishing mortality	Long-term mean number of years F > F_{MSY}	l_n_gr_fmsy
Avoid overfishing	Fishing mortality	Short-term mean number of years F < F_{MSY}	s_n_less_fmsy
Avoid overfishing	Fishing mortality	Long-term mean	l_n_less_fmsy

Fundamental Objective	Category	Metric	Shorthand
		number of years F < F_{MSY}	
Avoid overfishing	Fishing mortality	Short-term proportion of realizations with ≥ 1 year F > F(0.5 SSB_{MSY})	s_is_gr_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Long-term proportion of realizations with ≥ 1 year F > F(0.5 SSB_{MSY})	l_is_gr_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Short-term proportion of realizations with ≥ 1 year F > F(0.1 SSB_{MSY})	s_is_gr_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Long-term proportion of realizations with ≥ 1 year F > F(0.1 SSB_{MSY})	l_is_gr_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Short-term proportion of realizations with ≥ 1 year F < F(0.5 SSB_{MSY})	s_is_less_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Long-term proportion of realizations with ≥ 1 year F < F(0.5 SSB_{MSY})	l_is_less_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Short-term proportion of realizations with ≥ 1 year F < F(0.1 SSB_{MSY})	s_is_less_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Long-term proportion of realizations with ≥ 1 year F < F(0.1 SSB_{MSY})	l_is_less_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Short-term mean number of years	s_n_gr_f_dot_5_bmsy

Fundamental Objective	Category	Metric	Shorthand
		among realization with F > F(0.5 SSB_{MSY})	
Avoid overfishing	Fishing mortality	Long-term mean number of years among realization with F > F(0.5 SSB_{MSY})	l_n_gr_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Short-term mean number of years among realization with F > F(0.1 SSB_{MSY})	s_n_gr_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Long-term mean number of years among realization with F > F(0.1 SSB_{MSY})	l_n_gr_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Short-term mean number of years among realization with F < F(0.5 SSB_{MSY})	s_n_less_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Long-term mean number of years among realization with F < F(0.5 SSB_{MSY})	l_n_less_f_dot_5_bmsy
Avoid overfishing	Fishing mortality	Short-term mean number of years among realization with F < F(0.1 SSB_{MSY})	s_n_less_f_dot_1_bmsy
Avoid overfishing	Fishing mortality	Long-term mean number of years among realization with F < F(0.1	l_n_less_f_dot_1_bmsy

Fundamental Objective	Category	Metric	Shorthand
		SSB_{MSY})	
Avoid overfishing	Fishing mortality	Short-term mean F/F_{MSY}	s_avg_f_fmsy
Avoid overfishing	Fishing mortality	Long-term mean F/F_{MSY}	l_avg_f_fmsy

Table 3.2. Metrics considered but ultimately excluded from the analysis. SSB=spawning stock biomass; F=fishing mortality rate.

Fundamental Objective	Category	Metric
Avoid low biomass or overfished state	Biomass of target species	Maximum consecutive years among realizations with SSB < threshold
Rebuild a depleted stock	Fishing mortality	F relative to F that would rebuild in 10 years
Rebuild a depleted stock	Fishing mortality	Mean proportion of years among realizations that a method correctly identified rebuilding was necessary
Rebuild a depleted stock	Fishing mortality	Mean number of years among realizations that an overfished stock takes to rebuild
Correctly identify stock status	Miscellaneous (legislative)	Mean proportion of years among realizations that a method correctly identified overfished
Correctly identify stock status	Miscellaneous (legislative)	Mean proportion of years among realizations that a method correctly identified overfishing
Maintain a diverse age structure	Miscellaneous (ecosystem/stock resilience)	Mean age of the population among years and realizations during the feedback period
Maintain a diverse age structure	Miscellaneous (ecosystem/stock resilience)	Mean age of the catch among years and realizations during the feedback period
Use a robust method	Miscellaneous	Degree of retrospective pattern of a method, if applicable
Use a robust method	Miscellaneous	Variation in reference points among years

Table 4.1. Coding scheme used to identify scenarios. There are 4 factors, each with 2 or 3 possible values that combine to make the 4-character code. F_{MSY}=fishing mortality rate at maximum sustainable yield; IBM=index-based method.

Position	Factor	Values
1	Retrospective source	C = catch M = natural mortality N = none
2	Fishing history	F = Fmsy in second half of base period O = overfishing throughout base period
3	Fishery selectivity blocks	1 = constant selectivity 2 = selectivity changes in second half of base period
4	Catch advice multiplier	A = applied as is from IBM R = reduced (multiplied by 0.75) from IBM

Table 4.2. Number of successful simulations conducted for each combination of index-based method (IBM; row) and scenario (column, see Table 4.1 for scenario codes). The factorial combinations are represented by the first 13 IBMs (not including statistical catch-at-age [SCAA]) and the first 16 scenarios (not including NF1A or NO1A). The SCAA row represents the sensitivity run using an SCAA model and was only conducted for 4 of the scenarios. The NF1A and NO1A columns represent sensitivity runs conducted with no retrospective source and were run for all IBMs except the dynamic linear model (DLM) and SCAA due to time limitations.

IBMlab	CF1A	CF1R	CF2A	CF2R	CO1A	CO1R	CO2A	CO2R	MF1A	MF1R	MF2A	MF2R	MO1A	MO1R	MO2A	MO2R	NF1A	NO1A
AIM	999	1000	1000	1000	999	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	999	999
CC-FM	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
CC-FSPR	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
DLM	734	535	735	535	733	535	735	535	735	535	734	535	733	535	735	535	NA	NA
Ensemble	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-FM	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-Frecent	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-FSPR	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
ES-Fstable	1000	1000	1000	1000	999	1000	999	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	999
Islope	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Itarget	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
PBS	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Skate	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
SCAA	1000	NA	NA	NA	1000	NA	NA	NA	1000	NA	NA	NA	1000	NA	NA	NA	NA	NA

Table 4.3. Results of linear model analysis. “Sig” indicates a given explanatory variable was significant ($p<0.05$) for a given metric, while “NS” indicates not significant. Rows labeled “Fraction_Sig” display the fraction of metrics for which an explanatory variable was significant, and this value was calculated for 3 separate groups of metrics: 1) those related to spawning stock biomass (SSB), 2) those related to fishing mortality rate (F), and 3) those related to catch.

	retro_type	Fhist	n_selblocks	IBMlab	catch.mult	retro_type:Fhist	retro_type:n_selblocks	retro_type:IBMlab	retro_type:catch.mult	Fhist:n_selblocks	Fhist:IBMlab	Fhist:catch.mult	n_selblocks:IBMlab	n_selblocks:catch.mult	IBMlab:catch.mult
l_avg_ssb_ssbmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	Sig
l_is_ge_bmsy	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_is_less_01_bmsy	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	Sig
l_is_less_05_bmsy	Sig	NS	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_n_ge_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	NS	NS
l_n_less_01_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_n_less_05_bmsy	Sig	NS	NS	Sig	Sig	NS	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
s_avg_ssb_ssbmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	NS	NS	NS
s_is_ge_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	Sig	Sig	NS	Sig
s_is_less_01_bmsy	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	NS
s_is_less_05_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	NS	NS	NS	NS
s_n_ge_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	NS	NS
s_n_less_01_bmsy	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	NS	NS	NS	NS
s_n_less_05_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	Sig	Sig	NS	NS	NS
Fraction_Sig	1	0.86	0.86	1	1	0.64	0.64	1	0.79	0.36	1	0.29	0.64	0.21	0.5

	retro_type	Fhist	n_selblocks	IBMlab	catch.mult	retro_type:Fhist	retro_type:n_selblocks	retro_type:IBMlab	retro_type:catch.mult	Fhist:n_selblocks	Fhist:IBMlab	Fhist:catch.mult	n_selblocks:IBMlab	n_selblocks:catch.mult	IBMlab:catch.mult
l_avg_f_fmsy	Sig	Sig	Sig	Sig	Sig	NS	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_is_gr_f_dot_1_bmsy	Sig	NS	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	Sig
l_is_gr_f_dot_5_bmsy	Sig	NS	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_is_gr_fmsy	Sig	NS	Sig	Sig	Sig	NS	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_is_less_f_dot_1_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	Sig	Sig
l_is_less_f_dot_5_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_is_less_fmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_n_gr_f_dot_1_bmsy	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	NS	Sig	Sig	Sig
l_n_gr_f_dot_5_bmsy	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_n_gr_fmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_n_less_f_dot_1_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	NS
l_n_less_f_dot_5_bmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	NS
l_n_less_fmsy	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	NS
s_avg_f_fmsy	Sig	NS	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	NS
s_is_gr_f_dot_1_bmsy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS
s_is_gr_f_dot_5_bmsy	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	NS	NS	Sig	NS	Sig	NS	Sig
s_is_gr_fmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	NS
s_is_less_f_dot_1_bmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	Sig	Sig	NS	NS	NS	NS
s_is_less_f_dot_5_bmsy	Sig	NS	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	Sig	NS	NS	NS
s_is_less_fmsy	Sig	Sig	NS	Sig	Sig	Sig	Sig	Sig	NS	NS	Sig	Sig	NS	NS	Sig
s_n_gr_f_dot_1_bmsy	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	NS	Sig	Sig	NS
s_n_gr_f_dot_5_bmsy	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
s_n_gr_fmsy	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig
s_n_less_f_dot_1_bmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	NS	NS	NS
s_n_less_f_dot_5_bmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	NS	NS	Sig	NS	NS	NS	NS
s_n_less_fmsy	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	NS	NS	NS
Fraction_Sig	1	0.77	0.42	1	1	0.69	0.27	1	0.81	0.04	1	0.15	0.73	0.19	0.54

	retro_type	Fhist	n_selblocks	IBMlab	catch.mult	retro_type:Fhist	retro_type:n_selblocks	retro_type:IBMlab	retro_type:catch.mult	Fhist:n_selblocks	Fhist:IBMlab	Fhist:catch.mult	n_selblocks:IBMlab	n_selblocks:catch.mult	IBMlab:catch.mult
a_iav_catch	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	NS	Sig	NS	Sig	NS	Sig
l_avg_catch	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_avg_catch_msy	Sig	NS	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_iav_catch	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	NS	Sig	NS	Sig	Sig	Sig
l_prop_g_msy_2_of_3	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
l_sd_catch	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig	NS	Sig
s_avg_catch	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig
s_avg_catch_msy	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	Sig
s_iav_catch	Sig	Sig	NS	Sig	Sig	Sig	NS	Sig	NS	NS	Sig	NS	NS	NS	Sig
s_sd_catch	Sig	Sig	Sig	Sig	Sig	Sig	NS	Sig	NS	NS	Sig	NS	NS	NS	Sig
Fraction_Sig	1	0.9	0.9	1	1	0.9	0.6	1	0.6	0.2	1	NA	0.8	0.1	1

Table 4.4. Modified mean value of metric across all scenarios in base set of simulations. The metrics with negative values have had the actual mean multiplied by -1 so that the values closer to 0 are better than the values further from 0. The means are calculated across all successful simulations of all scenarios for each IBM.

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	Skate
catch_a_iav_catch	-0.633	-0.698	-0.683	-0.488	-0.417	-0.682	-0.326	-0.651	-0.635	-0.169	-0.712	-0.196	-0.745
catch_l_avg_catch	4056	3066	3103	1716	3492	3301	2568	3337	4088	2658	4802	2362	2807
catch_l_avg_catch_msy	0.903	0.532	0.536	0.361	0.778	0.793	0.534	0.798	0.907	0.557	0.979	0.485	0.712
catch_l_iav_catch	-0.401	-0.597	-0.592	-0.530	-0.290	-0.349	-0.288	-0.346	-0.404	-0.120	-0.635	-0.178	-0.313
catch_l_prop_g_msy_2_of_3	-0.247	-0.085	-0.087	-0.047	-0.240	-0.226	-0.116	-0.226	-0.247	-0.144	-0.242	-0.105	-0.212
catch_l_sd_catch	-2767	-2509	-2507	-1214	-1991	-1965	-873	-1990	-2815	-634	-5249	-562	-1687
catch_s_avg_catch	7590	3869	3872	4694	6702	7132	4932	7147	7576	5160	6246	4688	7916
catch_s_avg_catch_msy	1.481	0.738	0.735	0.957	1.267	1.330	1.005	1.330	1.479	1.056	1.115	0.956	1.561
catch_s_iav_catch	-0.428	-0.266	-0.262	-0.171	-0.211	-0.389	-0.186	-0.381	-0.431	-0.089	-0.300	-0.105	-0.425
catch_s_sd_catch	-5190	-1795	-1757	-1001	-2391	-4799	-1031	-4707	-5195	-614	-3578	-616	-5479
f_l_avg_f_fmsy	-2.628	-1.072	-0.972	-0.640	-3.227	-3.844	-1.833	-3.757	-2.610	-2.598	-2.248	-1.320	-4.864
f_l_is_gr_f_dot_1_bmsy	-0.479	-0.216	-0.195	-0.125	-0.461	-0.503	-0.224	-0.488	-0.480	-0.324	-0.553	-0.157	-0.602
f_l_is_gr_f_dot_5_bmsy	-0.707	-0.285	-0.287	-0.208	-0.633	-0.630	-0.334	-0.630	-0.712	-0.454	-0.755	-0.235	-0.825
f_l_is_gr_fmsy	-0.906	-0.358	-0.360	-0.274	-0.773	-0.863	-0.411	-0.857	-0.908	-0.520	-0.852	-0.314	-0.945

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	Skate
f_l_is_less_f_dot_1_bmsy	0.991	0.993	0.998	0.989	0.901	0.865	0.926	0.869	0.991	0.876	1.000	0.953	0.793
f_l_is_less_f_dot_5_bmsy	0.952	0.989	0.994	0.987	0.847	0.751	0.889	0.758	0.955	0.829	0.999	0.929	0.668
f_l_is_less_fmsy	0.805	0.982	0.985	0.981	0.784	0.622	0.841	0.628	0.813	0.765	0.993	0.894	0.544
f_l_n_gr_f_dot_1_bmsy	-3.399	-1.629	-1.402	-0.840	-5.043	-6.369	-2.942	-6.282	-3.383	-4.475	-3.552	-1.990	-8.131
f_l_n_gr_f_dot_5_bmsy	-7.037	-2.294	-2.220	-1.420	-7.125	-8.798	-4.277	-8.764	-6.945	-6.185	-5.526	-2.885	-11.198
f_l_n_gr_fmsy	-11.445	-3.131	-3.102	-2.172	-9.616	-11.442	-5.598	-11.420	-11.329	-7.529	-7.461	-3.989	-14.113
f_l_n_less_f_dot_1_bmsy	16.601	18.371	18.598	19.160	14.957	13.631	17.058	13.718	16.617	15.525	16.448	18.010	11.869
f_l_n_less_f_dot_5_bmsy	12.963	17.706	17.780	18.580	12.875	11.203	15.724	11.237	13.055	13.815	14.474	17.116	8.802
f_l_n_less_fmsy	8.555	16.869	16.898	17.828	10.384	8.558	14.402	8.580	8.671	12.471	12.539	16.011	5.887
f_s_avg_f_fmsy	-4.585	-1.391	-1.358	-1.328	-2.760	-4.145	-1.470	-4.072	-4.581	-1.772	-2.945	-1.275	-5.574
f_s_is_gr_f_dot_1_bmsy	-0.463	-0.120	-0.116	-0.091	-0.280	-0.410	-0.099	-0.405	-0.465	-0.139	-0.301	-0.072	-0.534
f_s_is_gr_f_dot_5_bmsy	-0.615	-0.210	-0.208	-0.248	-0.466	-0.520	-0.273	-0.523	-0.617	-0.323	-0.394	-0.224	-0.699
f_s_is_gr_fmsy	-0.818	-0.365	-0.364	-0.494	-0.664	-0.646	-0.556	-0.650	-0.815	-0.591	-0.526	-0.493	-0.885
f_s_is_less_f_dot_1_bmsy	0.711	0.951	0.956	0.990	0.933	0.729	0.981	0.740	0.710	0.966	0.850	0.991	0.564
f_s_is_less_f_dot_5_bmsy	0.571	0.910	0.911	0.913	0.722	0.578	0.901	0.581	0.573	0.859	0.704	0.918	0.486

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	Skate
f_s_is_less_fmsy	0.439	0.850	0.850	0.765	0.540	0.500	0.734	0.502	0.443	0.647	0.626	0.755	0.309
f_s_n_gr_f_dot_1_bmsy	-2.342	-0.513	-0.490	-0.300	-1.123	-2.105	-0.348	-2.078	-2.339	-0.529	-1.457	-0.239	-2.948
f_s_n_gr_f_dot_5_bmsy	-3.131	-0.883	-0.879	-0.955	-2.229	-2.826	-1.071	-2.824	-3.133	-1.384	-2.061	-0.876	-3.669
f_s_n_gr_fmsy	-4.156	-1.494	-1.487	-2.192	-3.381	-3.400	-2.443	-3.416	-4.140	-2.825	-2.678	-2.208	-4.770
f_s_n_less_f_dot_1_bmsy	3.658	5.487	5.510	5.700	4.877	3.895	5.652	3.922	3.661	5.471	4.543	5.761	3.052
f_s_n_less_f_dot_5_bmsy	2.869	5.117	5.121	5.045	3.771	3.174	4.929	3.176	2.867	4.616	3.939	5.124	2.331
f_s_n_less_fmsy	1.844	4.506	4.513	3.808	2.619	2.601	3.557	2.584	1.860	3.176	3.322	3.792	1.230
ssb_l_avg_ssb_ssbmsy	0.937	2.757	2.778	2.659	1.201	0.975	2.142	0.986	0.945	1.822	1.236	2.385	0.730
ssb_l_is_ge_bmsy	0.601	0.946	0.954	0.942	0.675	0.549	0.814	0.550	0.607	0.738	0.847	0.866	0.488
ssb_l_is_less_01_bmsy	-0.366	-0.143	-0.114	-0.102	-0.394	-0.433	-0.204	-0.433	-0.365	-0.295	-0.320	-0.149	-0.518
ssb_l_is_less_05_bmsy	-0.714	-0.278	-0.274	-0.241	-0.648	-0.640	-0.337	-0.639	-0.715	-0.460	-0.726	-0.251	-0.815
ssb_l_n_ge_bmsy	7.021	16.423	16.469	16.361	8.839	7.585	13.912	7.607	7.054	11.894	9.293	15.320	5.408
ssb_l_n_less_01_bmsy	-3.513	-1.244	-0.888	-0.923	-5.060	-6.114	-2.685	-5.954	-3.447	-3.965	-1.796	-1.935	-7.595
ssb_l_n_less_05_bmsy	-9.193	-2.536	-2.384	-2.265	-8.446	-9.720	-4.563	-9.709	-9.111	-6.451	-6.882	-3.307	-11.832
ssb_s_avg_ssb_ssbmsy	0.936	1.410	1.411	1.303	1.142	1.040	1.288	1.042	0.938	1.259	1.190	1.312	0.860

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	Skate
ssb_s_is_ge_bmsy	0.604	0.738	0.738	0.693	0.639	0.627	0.675	0.628	0.607	0.662	0.678	0.684	0.583
ssb_s_is_less_01_bmsy	-0.176	-0.044	-0.038	-0.027	-0.080	-0.167	-0.037	-0.151	-0.176	-0.051	-0.095	-0.028	-0.231
ssb_s_is_less_05_bmsy	-0.560	-0.303	-0.303	-0.282	-0.433	-0.505	-0.284	-0.509	-0.562	-0.315	-0.462	-0.266	-0.623
ssb_s_n_ge_bmsy	2.260	3.466	3.468	3.335	2.779	2.542	3.285	2.534	2.263	3.169	2.878	3.357	2.006
ssb_s_n_less_01_bmsy	-0.458	-0.115	-0.098	-0.065	-0.185	-0.420	-0.099	-0.368	-0.462	-0.138	-0.214	-0.072	-0.651
ssb_s_n_less_05_bmsy	-2.474	-1.139	-1.132	-1.126	-1.706	-2.297	-1.163	-2.288	-2.469	-1.291	-1.816	-1.093	-2.798

Table 4.5. Modified mean value of metric across all scenarios in base set of simulations. The metrics with negative values have had the actual mean multiplied by -1 so that the values closer to 0 are better than the values further from 0. The means are calculated across all successful simulations of all scenarios for each index-based method (IBM). The scenarios are limited to the ones examined in the statistical catch-at-age sensitivity analyses (see Table 4.2).

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	SCAA	Skate
catch_a_iav_catch	-0.807	-0.798	-0.722	-0.507	-0.524	-0.914	-0.389	-0.841	-0.801	-0.285	-0.836	-0.237	-0.453	-1.035
catch_l_avg_catch	3137	3487	2894	2480	2272	2400	3141	2620	3174	2684	4296	3169	4387	1879
catch_l_avg_catch_msy	0.793	0.665	0.522	0.553	0.622	0.687	0.697	0.715	0.797	0.601	0.929	0.689	0.904	0.580
catch_l_iav_catch	-0.447	-0.630	-0.593	-0.525	-0.310	-0.388	-0.314	-0.378	-0.446	-0.213	-0.727	-0.208	-0.276	-0.351
catch_l_prop_g_msy_2_of_3	-0.214	-0.122	-0.084	-0.085	-0.204	-0.203	-0.217	-0.215	-0.216	-0.241	-0.212	-0.206	-0.251	-0.178
catch_l_sd_catch	-2377	-3132	-2409	-1577	-1302	-1545	-1117	-1643	-2419	-876	-5330	-716	-2741	-1270
catch_s_avg_catch	7404	4252	3758	5295	6800	7069	5410	6774	7403	5611	5925	5145	5396	7610
catch_s_avg_catch_msy	1.536	0.852	0.747	1.132	1.361	1.401	1.155	1.313	1.534	1.204	1.103	1.101	0.994	1.587
catch_s_iav_catch	-0.482	-0.274	-0.265	-0.144	-0.230	-0.453	-0.199	-0.423	-0.487	-0.115	-0.321	-0.115	-0.351	-0.492
catch_s_sd_catch	-5668	-2039	-1804	-972	-2651	-5514	-1179	-5151	-5694	-776	-3895	-723	-4570	-6058
f_l_avg_f_fmsy	-3.196	-2.019	-1.165	-1.491	-4.749	-4.675	-4.309	-4.154	-3.157	-5.512	-2.354	-3.220	-1.323	-6.218
f_l_is_gr_f_dot_l_bmsy	-0.635	-0.387	-0.232	-0.316	-0.612	-0.579	-0.575	-0.521	-0.635	-0.750	-0.676	-0.412	-0.317	-0.793
f_l_is_gr_f_dot_5_bmsy	-0.837	-0.455	-0.290	-0.407	-0.729	-0.745	-0.664	-0.637	-0.843	-0.799	-0.767	-0.489	-0.454	-0.911
f_l_is_gr_fmsy	-0.985	-0.540	-0.364	-0.513	-0.878	-0.969	-0.783	-0.913	-0.982	-0.869	-0.856	-0.623	-0.511	-0.990
f_l_is_less_f_dot_l_bmsy	0.976	0.975	0.994	0.956	0.744	0.786	0.767	0.840	0.979	0.662	1.000	0.838	1.000	0.604
f_l_is_less_f_dot_5_bmsy	0.926	0.965	0.986	0.952	0.694	0.708	0.704	0.758	0.931	0.584	1.000	0.794	1.000	0.555
f_l_is_less_fmsy	0.789	0.948	0.978	0.942	0.658	0.601	0.622	0.648	0.800	0.488	0.999	0.725	0.994	0.513

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	SCAA	Skate
f_l_n_gr_f_dot_1_bmsy	-4.859	-3.339	-1.771	-2.352	-8.048	-7.619	-8.030	-6.712	-4.762	-10.975	-4.029	-5.632	-1.334	-10.855
f_l_n_gr_f_dot_5_bmsy	-8.342	-4.174	-2.348	-3.193	-9.411	-9.704	-9.531	-8.719	-8.232	-12.324	-5.315	-6.833	-2.994	-12.502
f_l_n_gr_fmsy	-12.908	-5.276	-3.139	-4.435	-11.469	-12.690	-11.560	-11.461	-12.743	-13.994	-6.975	-8.672	-5.848	-15.012
f_l_n_less_f_dot_1_bmsy	15.141	16.661	18.229	17.648	11.952	12.381	11.970	13.288	15.238	9.025	15.972	14.369	18.666	9.145
f_l_n_less_f_dot_5_bmsy	11.658	15.826	17.652	16.807	10.589	10.297	10.470	11.281	11.768	7.677	14.686	13.167	17.006	7.499
f_l_n_less_fmsy	7.092	14.724	16.862	15.565	8.531	7.311	8.441	8.540	7.257	6.006	13.025	11.329	14.152	4.989
f_s_avg_f_fmsy	-5.439	-1.835	-1.516	-1.857	-3.360	-5.139	-2.056	-4.599	-5.441	-2.433	-3.264	-1.764	-3.154	-6.552
f_s_is_gr_f_dot_1_bmsy	-0.547	-0.172	-0.139	-0.192	-0.359	-0.486	-0.203	-0.436	-0.553	-0.262	-0.313	-0.157	-0.253	-0.637
f_s_is_gr_f_dot_5_bmsy	-0.696	-0.272	-0.216	-0.386	-0.555	-0.563	-0.413	-0.524	-0.695	-0.465	-0.396	-0.355	-0.285	-0.751
f_s_is_gr_fmsy	-0.899	-0.446	-0.375	-0.639	-0.753	-0.746	-0.711	-0.647	-0.896	-0.732	-0.528	-0.647	-0.471	-0.943
f_s_is_less_f_dot_1_bmsy	0.649	0.928	0.944	0.980	0.916	0.652	0.959	0.705	0.651	0.922	0.826	0.980	0.791	0.508
f_s_is_less_f_dot_5_bmsy	0.552	0.875	0.905	0.850	0.654	0.525	0.835	0.560	0.557	0.785	0.700	0.858	0.771	0.475
f_s_is_less_fmsy	0.370	0.803	0.846	0.616	0.463	0.470	0.598	0.498	0.373	0.498	0.626	0.617	0.697	0.245
f_s_n_gr_f_dot_1_bmsy	-2.736	-0.744	-0.594	-0.623	-1.467	-2.560	-0.729	-2.250	-2.736	-1.038	-1.557	-0.533	-1.427	-3.379
f_s_n_gr_f_dot_5_bmsy	-3.424	-1.163	-0.904	-1.561	-2.687	-3.085	-1.717	-2.903	-3.426	-2.033	-2.068	-1.455	-1.556	-3.885
f_s_n_gr_fmsy	-4.636	-1.884	-1.527	-3.063	-3.877	-3.790	-3.333	-3.378	-4.607	-3.701	-2.678	-3.095	-2.304	-5.114
f_s_n_less_f_dot_1_bmsy	3.264	5.256	5.406	5.377	4.533	3.440	5.272	3.750	3.264	4.962	4.443	5.468	4.573	2.621
f_s_n_less_f_dot_5_bmsy	2.576	4.838	5.097	4.439	3.313	2.915	4.283	3.098	2.574	3.967	3.933	4.546	4.444	2.116
f_s_n_less_fmsy	1.364	4.116	4.473	2.937	2.123	2.210	2.667	2.622	1.393	2.299	3.322	2.906	3.696	0.886
ssb_l_avg_ssb_ssbmsy	0.731	2.296	2.756	1.855	0.930	0.803	1.160	0.933	0.740	0.877	1.165	1.507	1.621	0.558

metric	AIM	CC-FM	CC-FSPR	DLM	Ensemble	ES-FM	ES-Frecent	ES-FSPR	ES-Fstable	Islope	Itarget	PBS	SCAA	Skate
ssb_l_is_ge_bmsy	0.520	0.875	0.927	0.847	0.548	0.504	0.587	0.527	0.524	0.473	0.825	0.686	0.776	0.437
ssb_l_is_less_01_bmsy	-0.526	-0.320	-0.195	-0.300	-0.559	-0.523	-0.504	-0.499	-0.526	-0.654	-0.479	-0.378	-0.248	-0.640
ssb_l_is_less_05_bmsy	-0.828	-0.458	-0.289	-0.489	-0.743	-0.735	-0.640	-0.644	-0.832	-0.790	-0.791	-0.494	-0.443	-0.905
ssb_l_n_ge_bmsy	5.448	13.892	16.137	12.548	6.956	6.394	8.335	7.424	5.493	5.995	8.737	10.692	11.742	4.207
ssb_l_n_less_01_bmsy	-6.325	-3.179	-1.800	-2.944	-8.642	-8.723	-7.182	-7.894	-6.184	-9.444	-2.962	-5.335	-1.329	-10.416
ssb_l_n_less_05_bmsy	-11.278	-4.879	-2.890	-5.331	-10.962	-11.029	-9.749	-10.161	-11.214	-12.443	-7.778	-7.324	-5.072	-13.437
ssb_s_avg_ssb_ssbmsy	0.841	1.307	1.360	1.186	1.043	0.931	1.168	0.992	0.842	1.136	1.143	1.195	1.168	0.775
ssb_s_is_ge_bmsy	0.576	0.698	0.720	0.637	0.602	0.599	0.623	0.612	0.578	0.612	0.663	0.632	0.669	0.557
ssb_s_is_less_01_bmsy	-0.322	-0.087	-0.069	-0.060	-0.169	-0.326	-0.083	-0.286	-0.320	-0.105	-0.178	-0.060	-0.183	-0.405
ssb_s_is_less_05_bmsy	-0.617	-0.355	-0.328	-0.352	-0.495	-0.549	-0.355	-0.524	-0.625	-0.394	-0.488	-0.335	-0.461	-0.683
ssb_s_n_ge_bmsy	2.041	3.202	3.353	3.025	2.517	2.297	2.948	2.453	2.042	2.822	2.778	3.036	2.829	1.839
ssb_s_n_less_01_bmsy	-0.891	-0.229	-0.182	-0.140	-0.399	-0.857	-0.221	-0.723	-0.885	-0.288	-0.424	-0.155	-0.459	-1.198
ssb_s_n_less_05_bmsy	-2.774	-1.397	-1.267	-1.445	-1.999	-2.622	-1.499	-2.476	-2.780	-1.646	-1.994	-1.408	-2.051	-3.114

Table 5.1. Analysis of variance (ANOVA) results with fraction of total Sums of Squares explained by factor for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) by index-based method (IBM). Bold values indicate >10% explained, italics indicate 5-10% explained, roman font indicates 1-5% explained. "—" indicates less than 1% explained, and "NA" indicates not significant.

Factor	Islope*	Itarget*	Skate*	AIM*	ES-Frecent*	ES-FSPR*	ES-FM	ES-Fstable	CC-FSPR*	CC-FM	PBS*	DLM	Ensemble*
retro_type	0.05	**0.54**	**0.50**	**0.50**	0.06	**0.58**	**0.58**	**0.50**	**0.13**	**0.13**	0.05	0.05	**0.43**
Fhist	0.02	0.01	0.01	0.02	0.02	--	--	0.02	--	0.01	0.02	0.03	0.01
n_selblocks	--	**NA**	--	--	--	--	--	--	0.01	--	--	--	--
catch.mult	**0.10**	--	0.02	0.02	**0.12**	0.01	0.01	0.02	0.02	0.02	0.09	0.08	0.03

Factor	Islope*	Itarget*	Skate*	AIM*	ES-Frecent*	ES-FSPR*	ES-FM	ES-Fstable	CC-FSPR*	CC-FM	PBS*	DLM	Ensemble*
time.avg	0.05	--	--	**NA**	0.10	--	--	**NA**	**0.21**	**0.20**	**0.15**	**0.20**	--
Fhist:time.avg	0.02	--	--	--	0.01	0.01	0.01	--	0.04	0.05	0.01	0.01	0.01
retro_type:Fhist	--	--	--	--	--	--	--	--	**NA**	--	--	--	--
retro_type:catch.mult	**NA**	--	0.01	0.01	**NA**	--	--	0.01	--	--	**NA**	**NA**	--
retro_type:time.avg	--	**NA**	--	--	--	0.01	0.01	--	--	--	--	--	0.02
catch.mult:time.avg	0.07	--	--	0.01	0.08	--	--	0.01	0.01	0.01	0.06	0.06	0.01
Fhist:catch.mult	--	--	--	**NA**	--	--	--	**NA**	--	--	**NA**	--	**NA**
SS.frac.explained	0.32	0.56	0.56	0.56	0.40	0.61	0.62	0.56	0.42	0.42	0.39	0.43	0.51

Table 5.2. Analysis of variance (ANOVA) results with fraction of total Sums of Squares explained by factor for fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) by index-based method (IBM). Bold values indicate >10% explained, italics indicate 5-10% explained, roman font indicates 1-5% explained. "—"" indicates less than 1% explained, and "NA" indicates not significant.

Factor	Islope*	Itarget*	Skate*	AIM*	ES-Frecent*	ES-FSPR*	ES-FM	ES-Fstable	CC-FSPR*	CC-FM	PBS*	DLM	Ensemble*
retro_type	0.10	**0.43**	**0.81**	**0.54**	0.08	**0.69**	**0.69**	**0.53**	**0.13**	**0.13**	0.07	0.06	**0.45**
Fhist	0.02	0.04	--	--	0.02	0.09	0.09	--	0.04	0.05	0.02	0.04	--
n_selblocks	0.03	--	0.03	--	0.02	--	0.01	--	--	0.01	0.01	--	0.01
catch.mult	**0.13**	--	0.01	0.02	**0.15**	0.01	0.01	0.02	0.03	0.03	**0.12**	0.08	0.04

Factor	Islope*	Itarget*	Skate*	AIM*	ES-Frecent*	ES-FSPR*	ES-FM	ES-Fstable	CC-FSPR*	CC-FM	PBS*	DLM	Ensemble*
time.avg	0.02	0.01	--	0.08	--	--	--	0.09	--	--	--	0.04	--
Fhist:time.avg	0.02	0.04	--	--	0.01	--	--	--	--	**NA**	0.01	0.02	--
retro_type:Fhist	--	0.05	--	--	--	0.07	0.07	--	0.01	0.02	--	0.02	0.01
retro_type:catch.mult	0.02	--	--	--	0.03	--	--	--	0.01	0.01	0.03	0.02	0.01
retro_type:time.avg	--	0.04	--	0.06	--	--	--	0.06	--	**NA**	--	--	--
catch.mult:time.avg	0.04	--	--	--	0.04	--	--	--	--	--	0.03	--	--
Fhist:catch.mult	--	--	--	--	--	--	--	--	--	--	--	--	**NA**
SS.frac.explained	0.391	0.618	0.869	0.732	0.369	0.874	0.881	0.723	0.252	0.259	0.3	0.29	0.542

Table 5.3. Analysis of variance (ANOVA) results with fraction of total Sums of Squares explained by factor for catch/maximum sustainable yield (MSY) by index-based method (IBM). Bold values indicate >10% explained, italics indicate 5-10% explained, roman font indicates 1-5% explained. "—"" indicates less than 1% explained, and "NA" indicates not significant.

Factor	Islope*	Itarget*	Skate*	AIM*	ES-Frecent*	ES-FSPR*	ES-FM	ES-Fstable	CC-FSPR*	CC-FM	PBS*	DLM	Ensemble*
retro_type	0.02	0.02	0.06	0.02	0.01	0.01	0.01	0.02	0.05	0.06	0.01	0.01	--
Fhist	0.02	0.08	0.04	0.05	--	0.04	0.04	0.05	**0.12**	**0.13**	--	--	0.03
n_selblocks	--	--	--	--	--	0.01	--	--	0.04	--	--	--	--
catch.mult	0.07	--	--	--	**0.13**	--	--	--	0.04	0.04	**0.15**	**0.11**	--
time.avg	**0.20**	--	**0.24**	**0.13**	**0.19**	**0.11**	**0.12**	**0.13**	0.03	0.03	**0.21**	**0.21**	**0.11**

Factor	Islope*	Itarget*	Skate*	AIM*	ES-Frecent*	ES-FSPR*	ES-FM	ES-Fstable	CC-FSPR*	CC-FM	PBS*	DLM	Ensemble*
Fhist:time.avg	--	0.05	0.05	0.04	--	0.10	0.10	0.04	0.02	0.02	**NA**	--	0.06
retro_type:Fhist	0.01	0.04	--	0.01	0.01	--	--	0.01	--	--	0.01	--	0.02
retro_type:catch.mult	--	--	--	--	--	--	--	--	--	--	--	--	--
retro_type:time.avg	**NA**	0.06	0.07	0.04	**NA**	**0.14**	**0.15**	0.04	0.01	0.01	--	**NA**	0.07
catch.mult:time.avg	--	--	--	--	--	--	--	--	--	--	0.01	--	0.01
Fhist:catch.mult	--	**NA**	--	--	--	--	--	--	**NA**	--	--	--	--
SS.frac.explained	0.325	0.269	0.469	0.307	0.364	0.425	0.429	0.302	0.316	0.296	0.41	0.348	0.313

Table 5.4. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM). Table rows correspond to the ordering in the heatmap in Figure 5.1.

IBM	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-Fstable	1.98	1.03	0.70
AIM	1.98	1.04	0.69
ES-FM	1.75	0.93	0.75
ES-FSPR	1.71	0.93	0.74
Ensemble	1.39	0.90	0.92
Itarget	1.26	0.87	0.97
Islope	0.87	0.70	1.23
PBS	0.55	0.65	1.56
DLM	0.49	0.58	1.60
ES-Frecent	0.65	0.67	1.42
CC-FSPR	0.31	0.47	1.79
CC-FM	0.31	0.47	1.78
Skate	3.78	1.03	0.54

Table 5.5. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and source of retrospective pattern (C=catch, M=natural mortality). Table rows correspond to the ordering in the heatmap in Figure 5.2.

IBM x Retro	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-FM-M	1.04	1.58	0.62
ES-FSPR-M	1.04	1.58	0.62
Ensemble-M	0.96	1.67	0.58
Islope-M	0.79	1.58	0.63
PBS-M	0.72	1.90	0.41
DLM-M	0.66	1.94	0.39
ES-Frecent-M	0.76	1.78	0.49
Itarget-M	0.77	1.74	0.42
CC-FM-M	0.39	2.38	0.17
CC-FSPR-M	0.39	2.36	0.17
CC-FM-Catch	0.59	1.15	0.67
CC-FSPR-Catch	0.59	1.16	0.65
DLM-Catch	0.50	1.18	0.70
PBS-Catch	0.58	1.18	0.78
ES-Frecent-Catch	0.59	1.04	0.94
Islope-Catch	0.61	0.87	1.20
AIM-M	1.15	1.38	0.88
ES-Fstable-M	1.14	1.38	0.88

IBM x Retro	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
Skate-M	1.21	1.21	1.29
ES-FM-Catch	0.74	0.25	8.01
ES-FSPR-Catch	0.74	0.25	7.89
Skate-Catch	0.67	0.18	8.73
AIM-Catch	0.87	0.31	5.76
ES-Fstable-Catch	0.88	0.32	5.68
Ensemble-Catch	0.78	0.42	4.63
Itarget-Catch	0.98	0.51	4.11

Table 5.6. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and time horizon (short- or long-term summary, "S" or "L"). Table rows correspond to the ordering in the heatmap in Figure 5.3.

IBM x Time	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-Fstable-L	0.71	0.84	1.69
AIM-L	0.70	0.83	1.69
ES-FSPR-L	0.70	0.75	1.71
ES-FM-L	0.71	0.75	1.82
ES-FM-S	0.76	1.13	1.71
ES-FSPR-S	0.77	1.13	1.71
Itarget-L	1.02	0.89	1.63
Ensemble-L	0.93	0.73	1.50
Ensemble-S	0.91	1.06	1.30
PBS-S	1.14	0.86	0.78
DLM-S	1.13	0.83	0.79
ES-Frecent-S	1.12	0.90	0.85
Itarget-S	0.94	0.83	0.79
Islope-S	1.08	0.96	0.96
CC-FM-S	1.19	0.55	0.45
CC-FSPR-S	1.19	0.55	0.44
Islope-L	1.66	0.43	0.55
CC-FM-L	2.66	0.33	0.14
CC-FSPR-L	2.66	0.34	0.13
PBS-L	2.34	0.30	0.19
DLM-L	2.45	0.16	0.11
ES-Frecent-L	2.10	0.35	0.28
ES-Fstable-S	0.69	1.28	2.61
AIM-S	0.69	1.29	2.58
Skate-L	0.42	0.62	3.10

IBM x Time	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
Skate-S	0.60	1.39	4.31

Table 5.7. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and catch advice multiplier (1 or 0.75). Table rows correspond to the ordering in the heatmap in Figure 5.4.

IBM x Catch Multiplier	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-Fstable-0.75	1.42	1.01	0.81
AIM-0.75	1.41	1.01	0.80
ES-FM-0.75	1.34	0.90	0.84
ES-FSPR-0.75	1.30	0.90	0.83
Itarget-1	1.48	0.91	0.91
ES-Frecent-1	1.26	0.92	1.01
Ensemble-1	1.93	0.95	0.78
Islope-1	1.81	0.92	0.85
PBS-1	0.91	0.87	1.20
Ensemble-0.75	0.93	0.86	1.08
Itarget-0.75	1.04	0.84	1.03
DLM-1	0.74	0.76	1.32
CC-FM-0.75	0.21	0.36	2.02
CC-FSPR-0.75	0.21	0.36	2.01
PBS-0.75	0.25	0.37	2.05
ES-Frecent-0.75	0.29	0.44	1.97
DLM-0.75	0.20	0.27	2.18
CC-FM-1	0.47	0.60	1.57
CC-FSPR-1	0.46	0.59	1.58
Islope-0.75	0.42	0.55	1.72
ES-Fstable-1	2.59	1.07	0.61
AIM-1	2.59	1.07	0.60
ES-FM-1	2.31	0.97	0.67
ES-FSPR-1	2.16	0.97	0.67
Skate-0.75	3.06	1.00	0.62
Skate-1	4.88	1.06	0.48

Table 5.8. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and fishing history (1=overfishing then F=F_{MSY}, 2=always overfishing). Table rows correspond to the ordering in the heatmap in Figure 5.5.

IBM x Fhistory	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
CC-FSPR-1	0.42	0.69	1.73

IBM x Fhistory	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
CC-FM-1	0.43	0.70	1.71
PBS-1	0.46	0.72	1.81
ES-Frecent-1	0.53	0.75	1.69
DLM-1	0.39	0.62	1.86
Islope-1	0.67	0.77	1.48
CC-FM-2	0.23	0.33	1.89
CC-FSPR-2	0.23	0.33	1.87
PBS-2	0.75	0.58	1.22
DLM-2	0.70	0.53	1.26
ES-Frecent-2	0.94	0.61	1.08
Islope-2	1.21	0.63	0.91
Itarget-2	1.08	0.68	0.88
Ensemble-2	1.43	0.82	0.80
ES-FSPR-2	1.44	0.81	0.75
ES-FM-2	1.52	0.81	0.76
Itarget-1	1.63	1.09	1.07
Ensemble-1	1.33	1.00	1.05
ES-Fstable-1	1.86	1.20	0.83
AIM-1	1.86	1.20	0.82
AIM-2	2.06	0.90	0.60
ES-Fstable-2	2.05	0.90	0.60
Skate-1	4.24	1.22	0.61
ES-FM-1	3.76	1.12	0.72
ES-FSPR-1	3.26	1.11	0.73
Skate-2	3.45	0.89	0.47

FIGURES

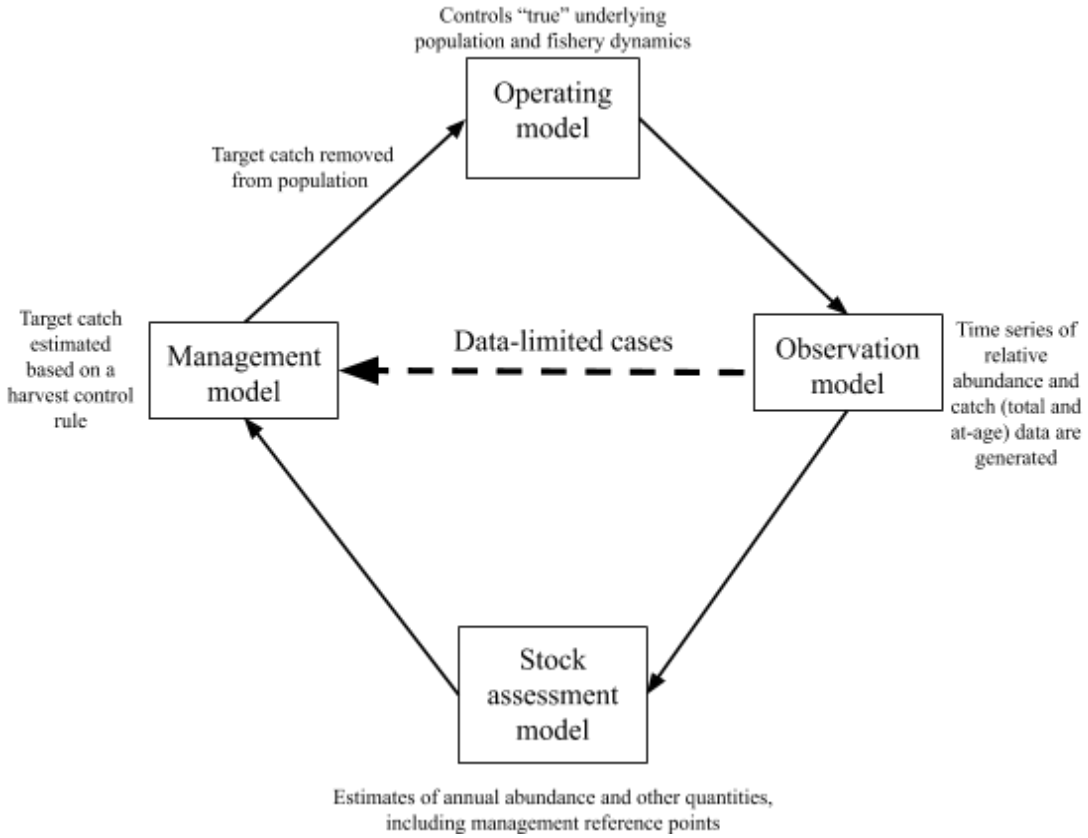


Figure 0.1. Overview of the closed-loop simulation approach used in this study.

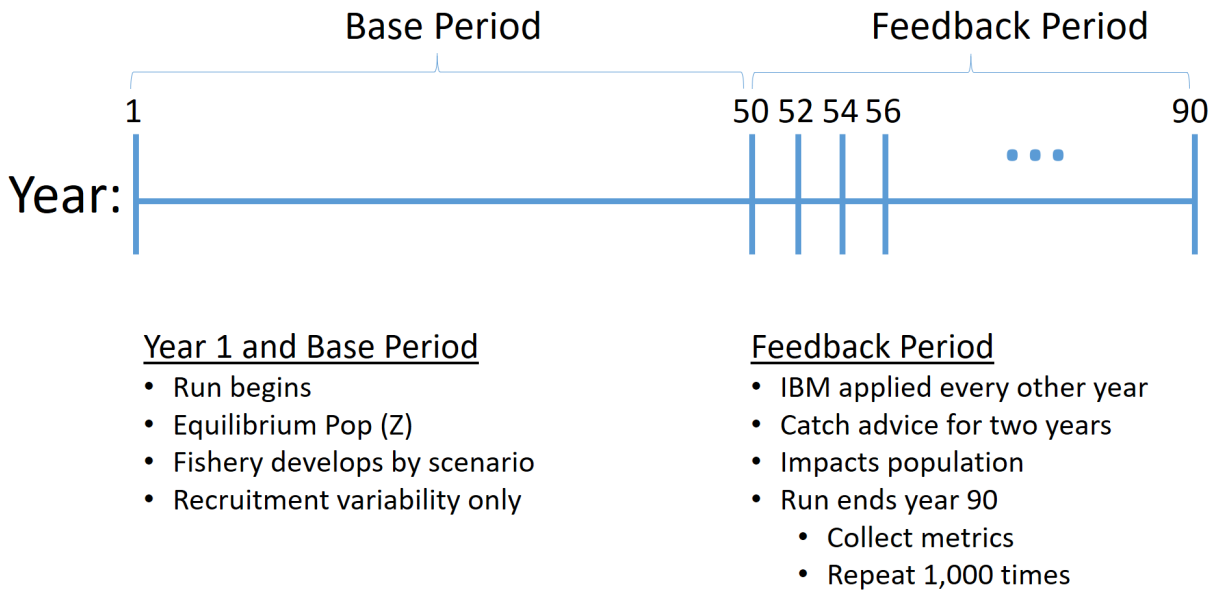


Figure 0.2. Overview of the structure of the simulations used in this study.

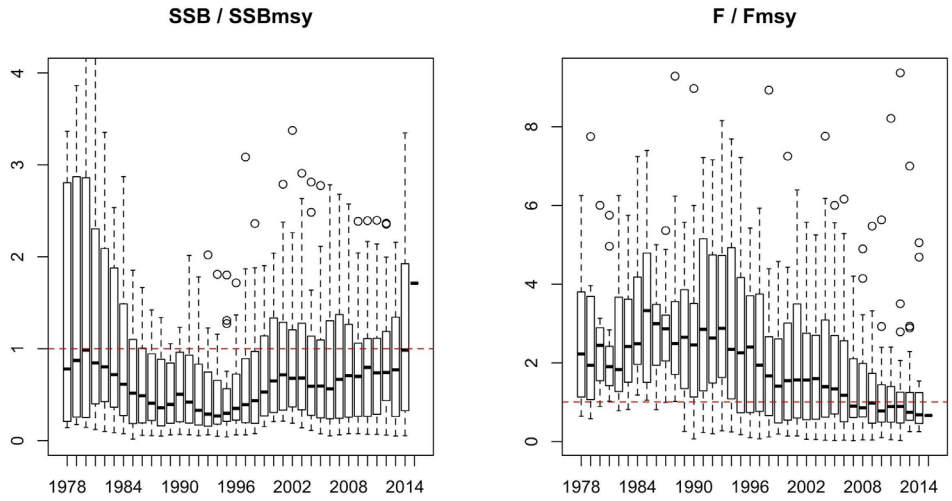


Figure 1.1. Ratios of spawning stock biomass (SSB) and fishing mortality rate (F) to associated maximum sustainable yield (MSY) reference points for 19 stocks in the Northeast U.S. (15 managed by the New England Fishery Management Council and 5 by the Mid-Atlantic Fishery Management Council).

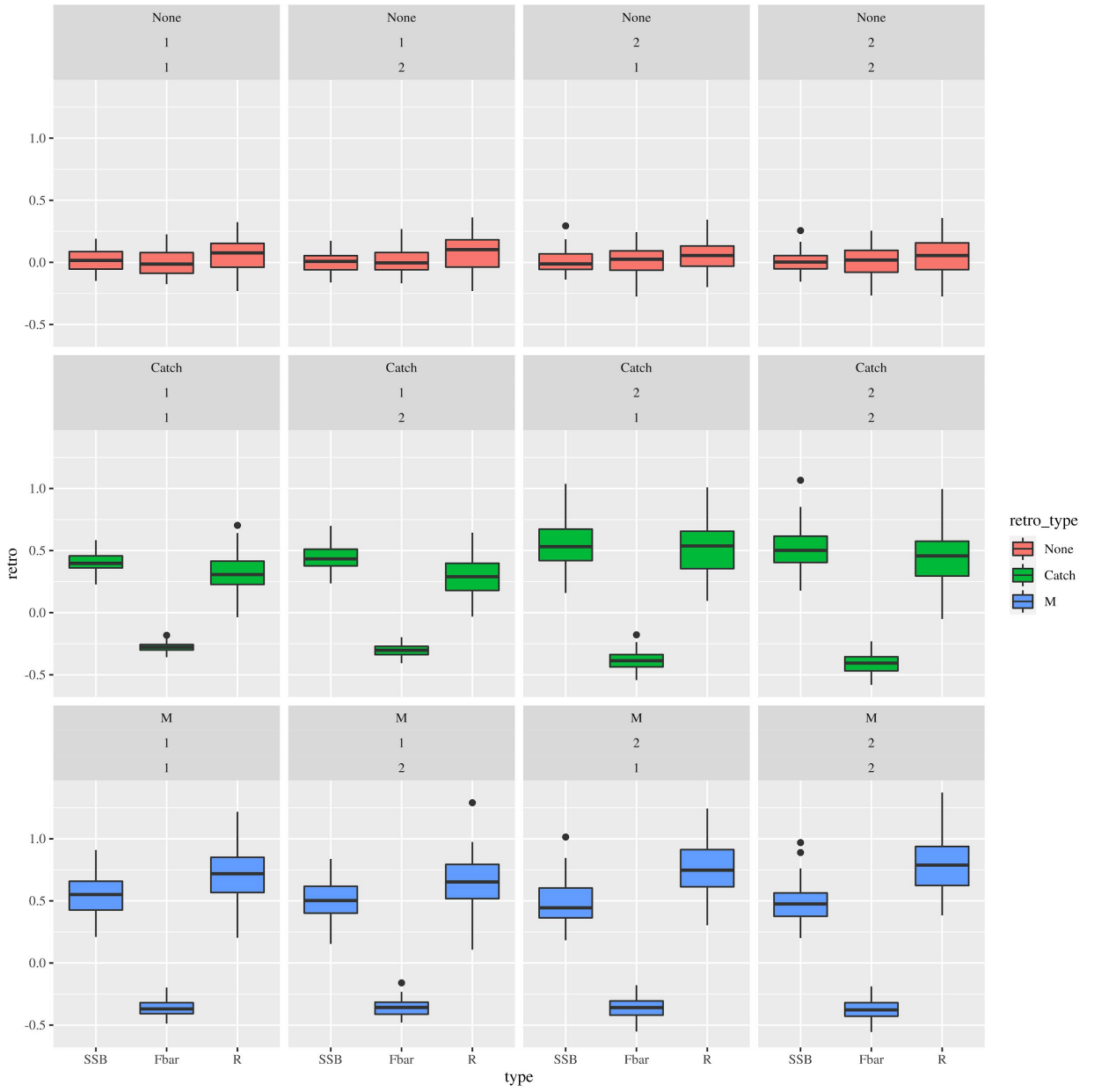


Figure 1.2. Distribution of 50 Mohn's rho values for no misspecification (top row), catch misspecification (middle row), and natural mortality misspecification (bottom row) for spawning stock biomass (SSB), fishing mortality rate (Fbar), and recruitment (R). Panels are defined by retrospective type (none, catch, natural mortality), fishing intensity (1=always overfishing, 2= overfishing reducing to maximum sustainable yield [Fmsy]), and number of selectivity blocks (1 or 2).

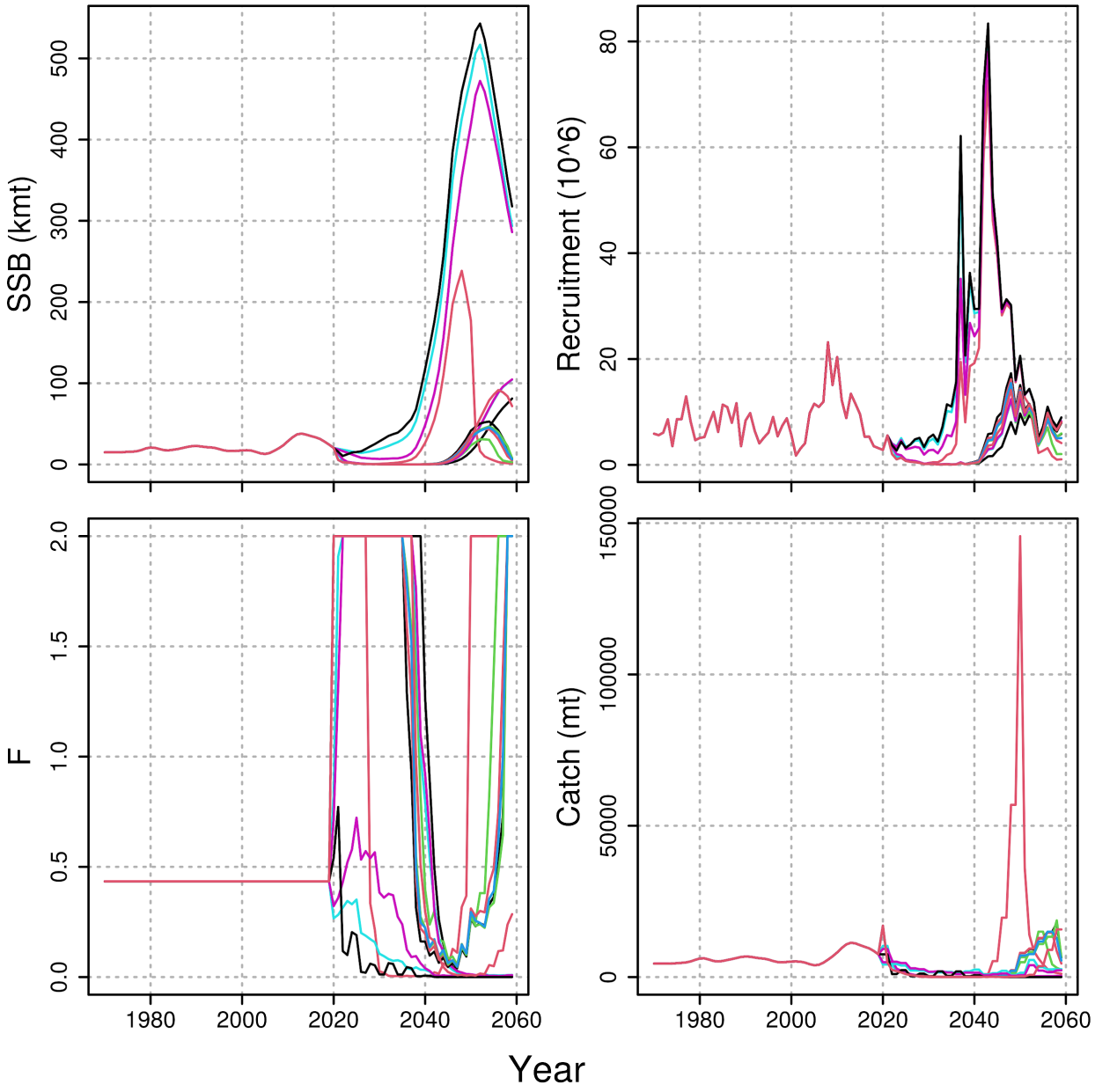


Figure 1.3. Example of variation in spawning stock biomass (SSB), recruitment, fully-selected fishing mortality, and catch trajectories during feedback period with alternative index-based methods (IBMs) and harvest control rules.

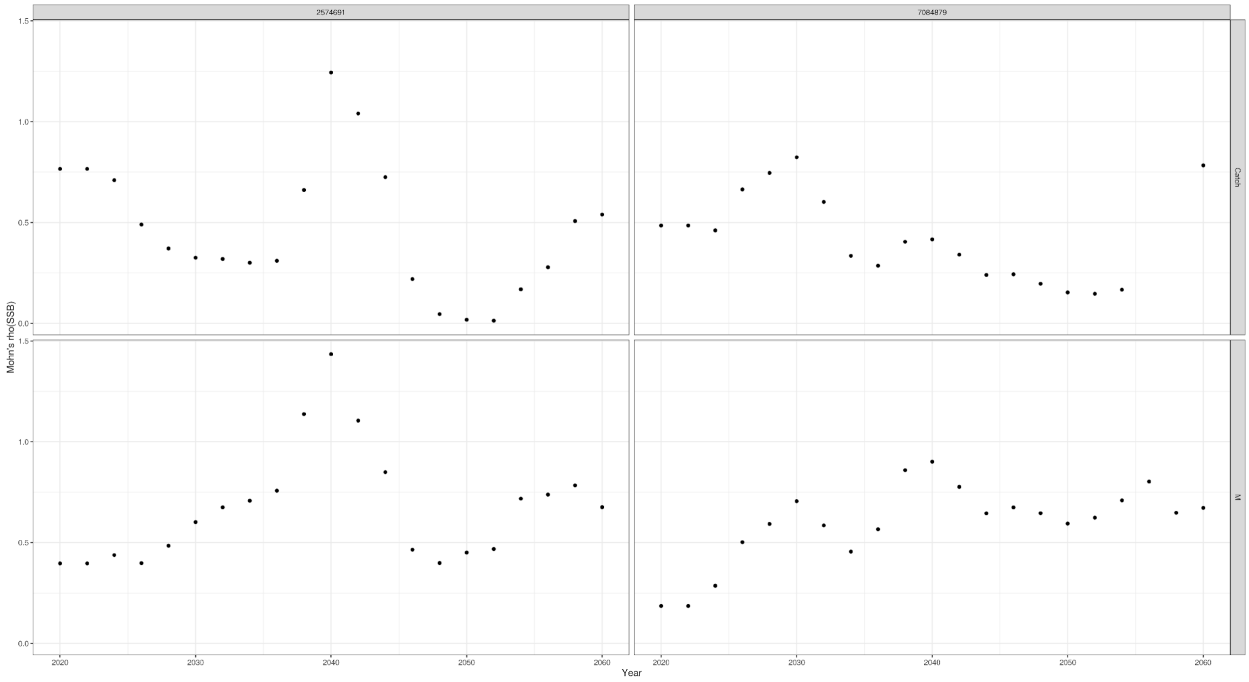


Figure 1.4. Mohn's rho for spawning stock biomass (SSB) during the feedback period using 2 different random number seeds (columns) and either catch (top row) or natural mortality (bottom row) as the source of the retrospective pattern.

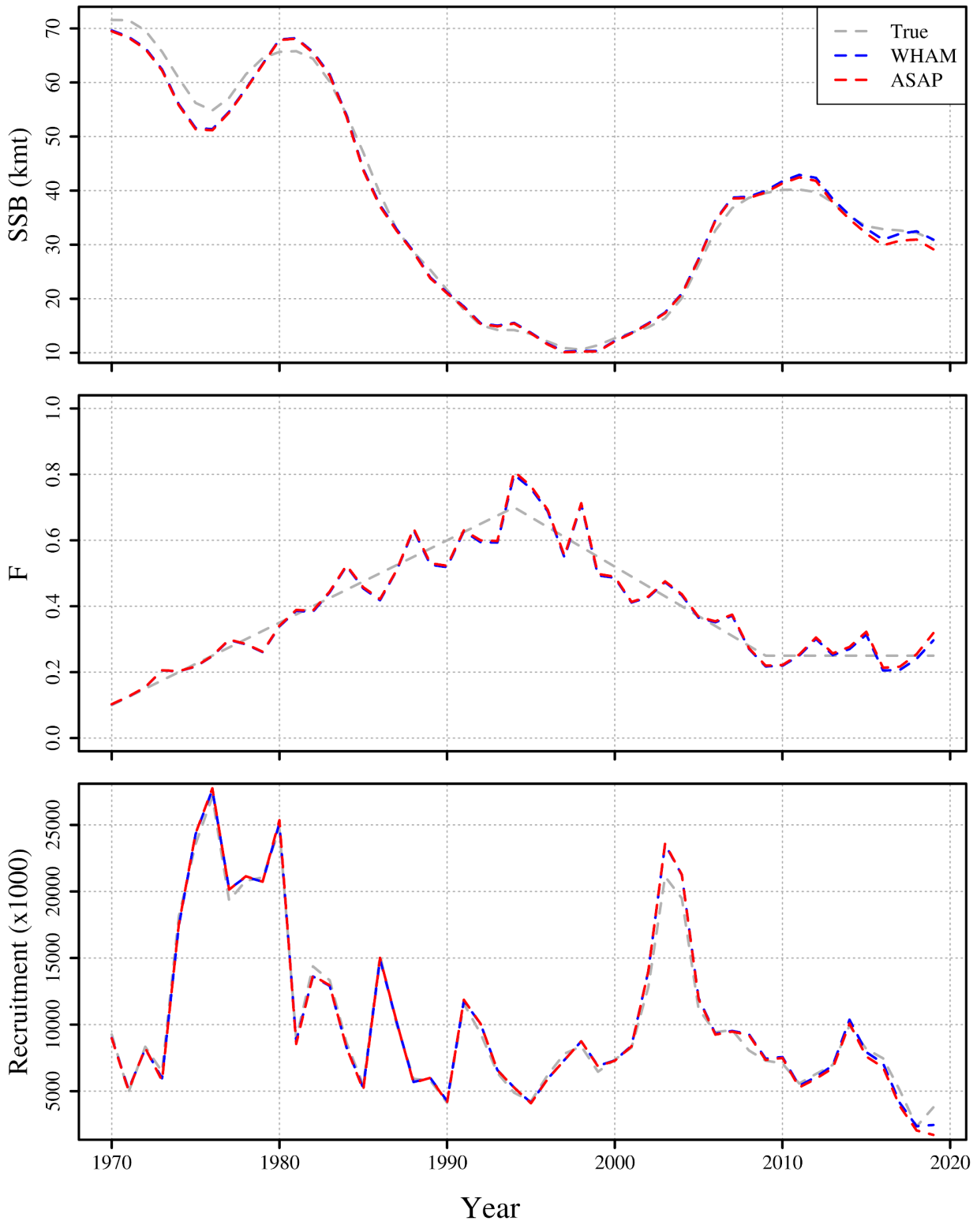


Figure 1.5. Comparison of estimates of spawning stock biomass (SSB), fishing mortality rate (F), and recruitment from comparable configurations of statistical catch-at-age (SCAA) models in the Age-Structured Assessment Program (ASAP) and the Woods Hole Assessment Model (WHAM).

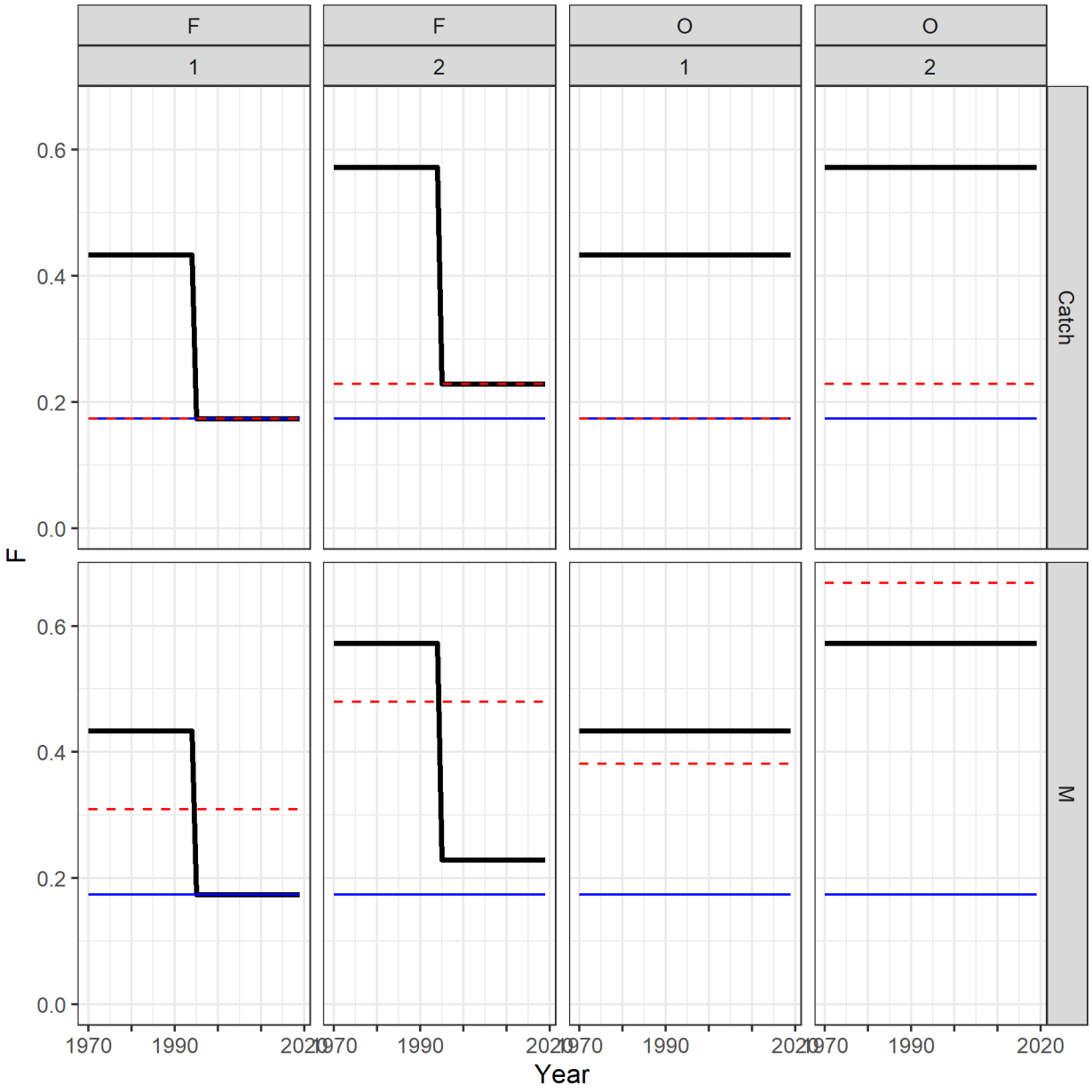


Figure 1.6. Fishing mortality rate (F) during the base period (black line) for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The horizontal blue line denotes the F at maximum sustainable yield (Fmsy) computed from conditions in the first year of the base period (1970) while the red dashed horizontal line denotes the Fmsy computed from conditions in the terminal year in the base period (2019).

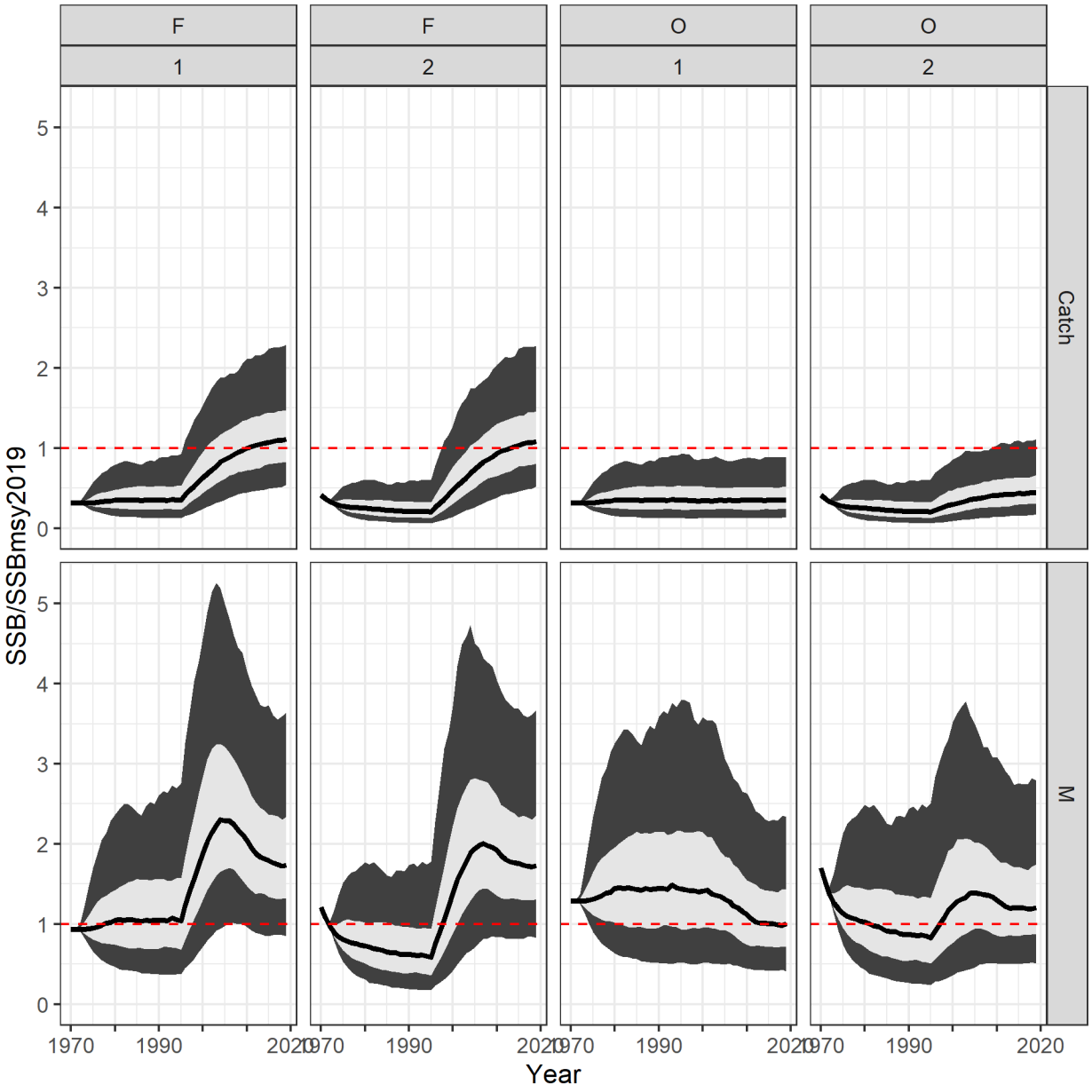


Figure 1.7. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) computed from the conditions in the terminal year in the base period (2019) with the median value shown as the black line, the interquartile range shown in light gray, and the 95% range shown in dark gray for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]).

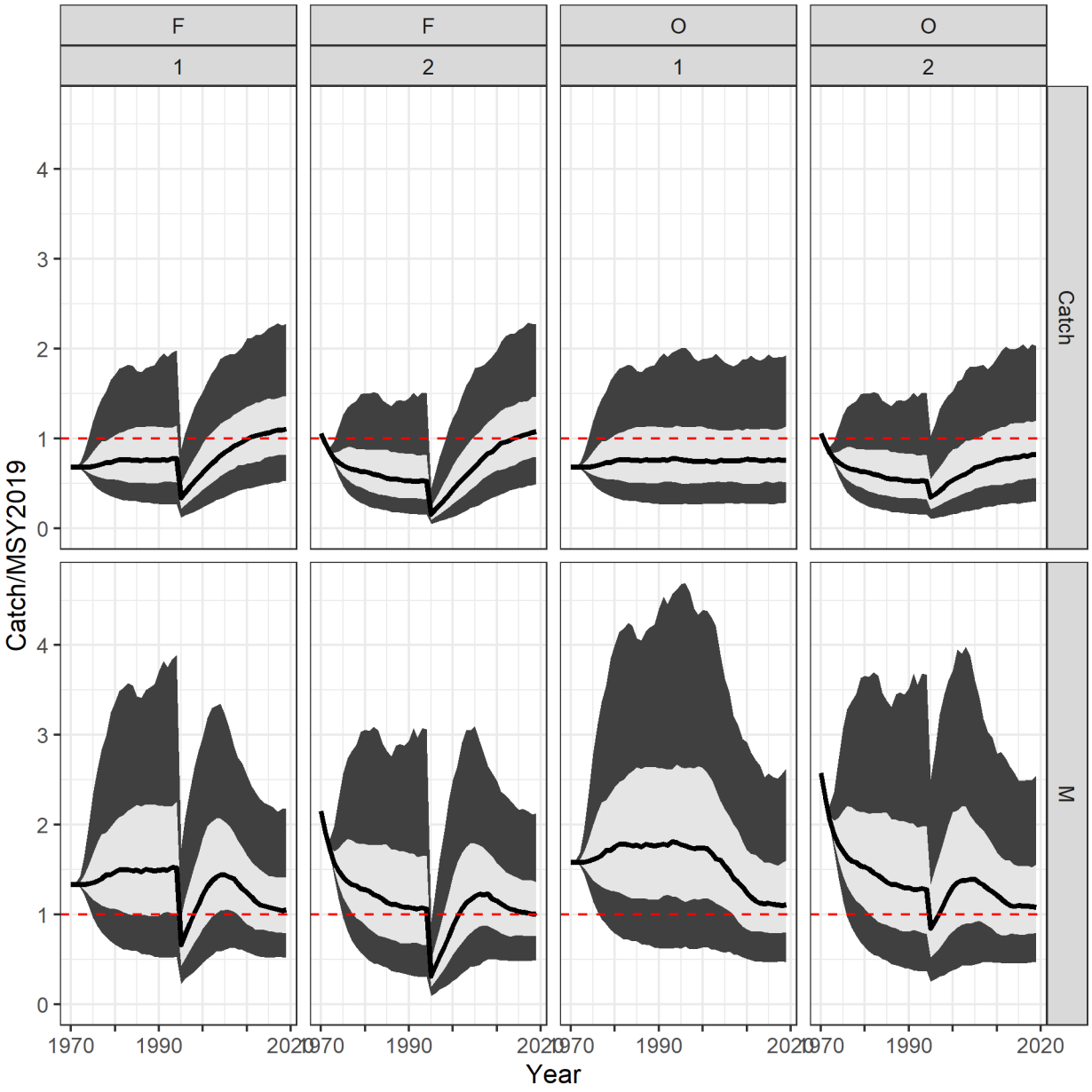


Figure 1.8. Catch relative to maximum sustainable yield (MSY) computed from the conditions in the terminal year in the base period (2019) with the median value shown as the black line, the interquartile range shown in light gray, and the 95% range shown in dark gray for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]).

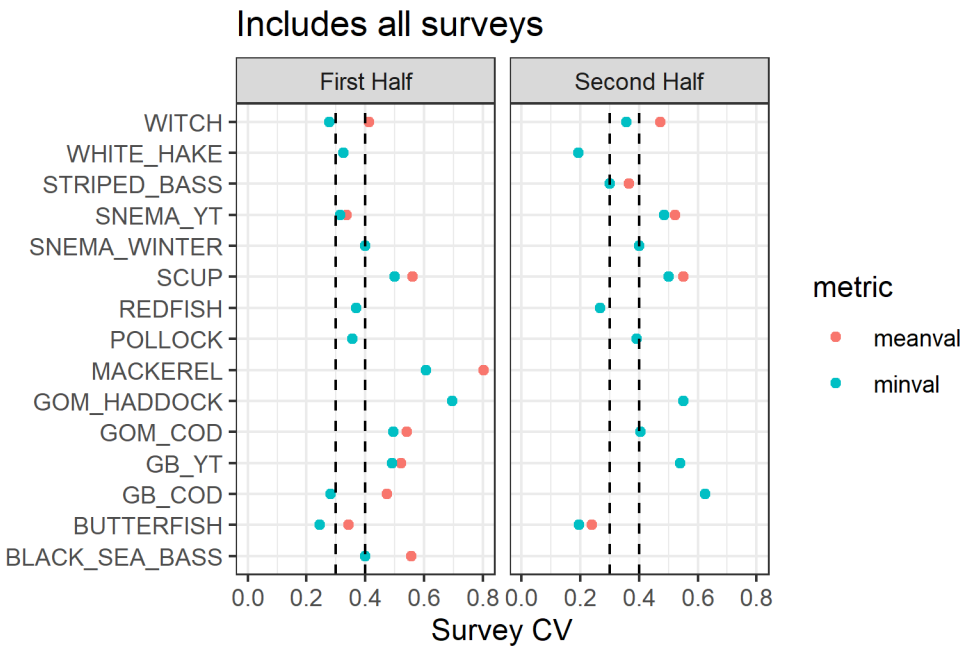


Figure 1.9. Coefficient of variation (CV) for the surveys used in 15 local stock assessments separated by time of the survey (Jan-June = first half, July-Dec = second half). The metrics shown are the mean value (red) and minimum value (blue) of the survey CVs that occurred in a given half year for that stock assessment. The vertical lines show the CVs assumed in the simulation modeling.

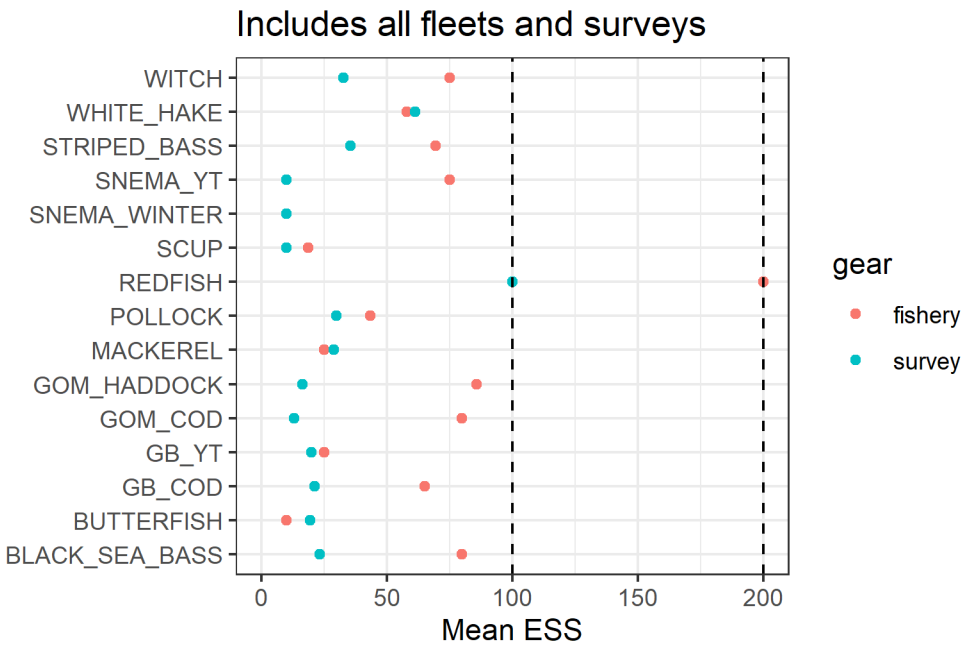


Figure 1.10. Mean effective sample size (ESS) for the fishery and surveys from 15 local assessments. The vertical lines show the ESS values assumed in the simulation modeling.

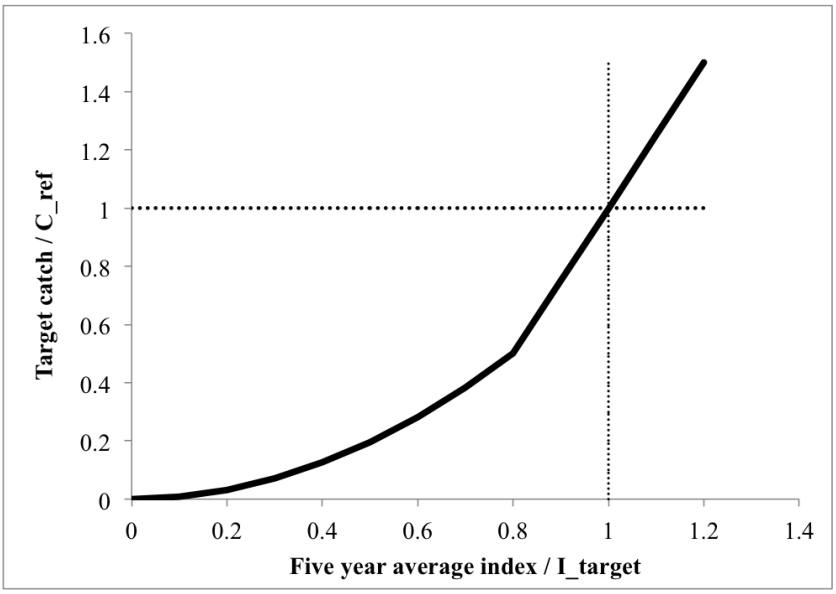


Figure 2.1. Relationship between relative target catch and relative index used in Itarget method to determine catch advice based on recent index observations.

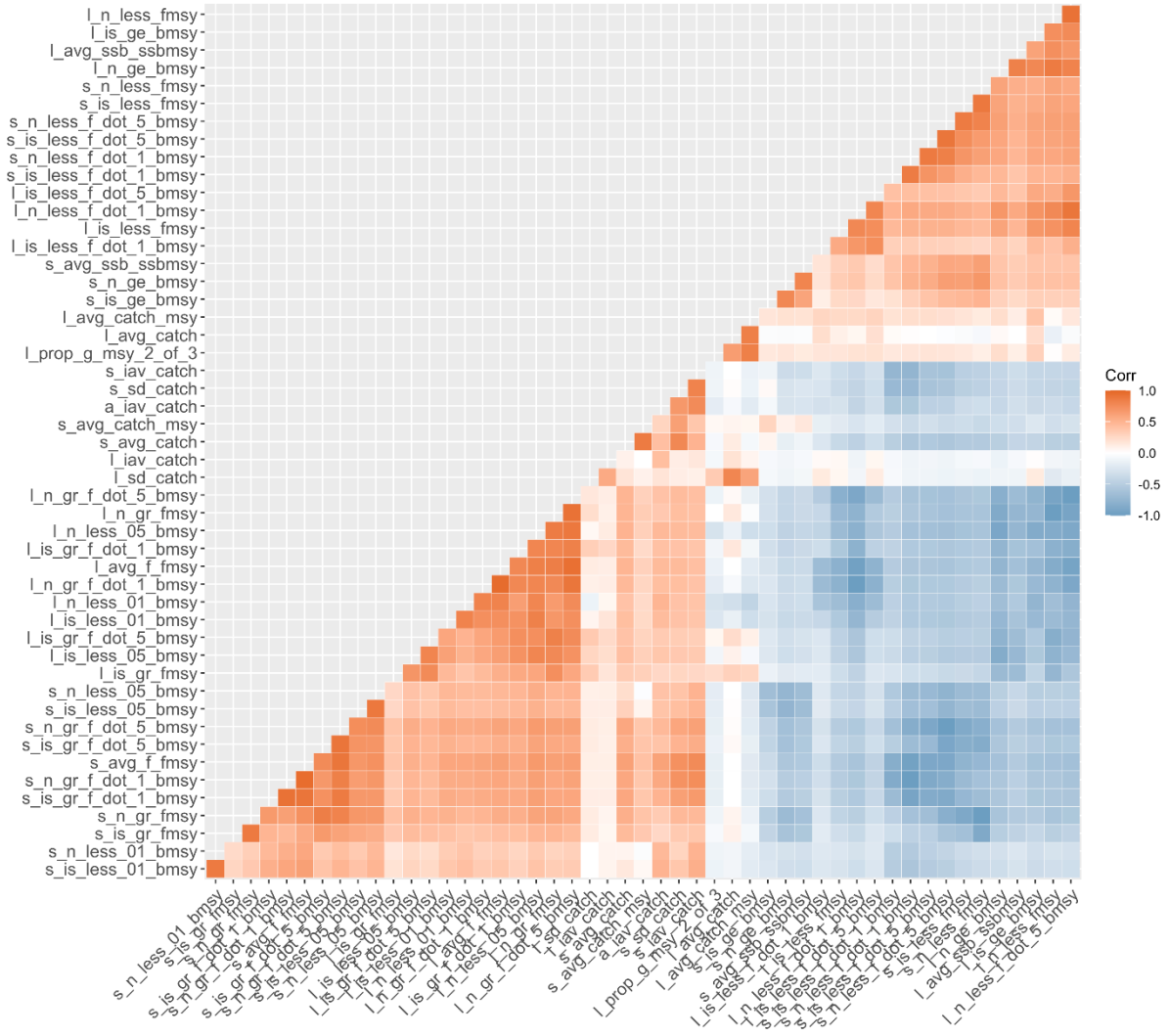


Figure 3.1. Correlation plot for all 50 metrics used in this study from the base analyses. The ordering of the metrics in the axes uses hierarchical clustering to create regions of strong positive and negative correlation.

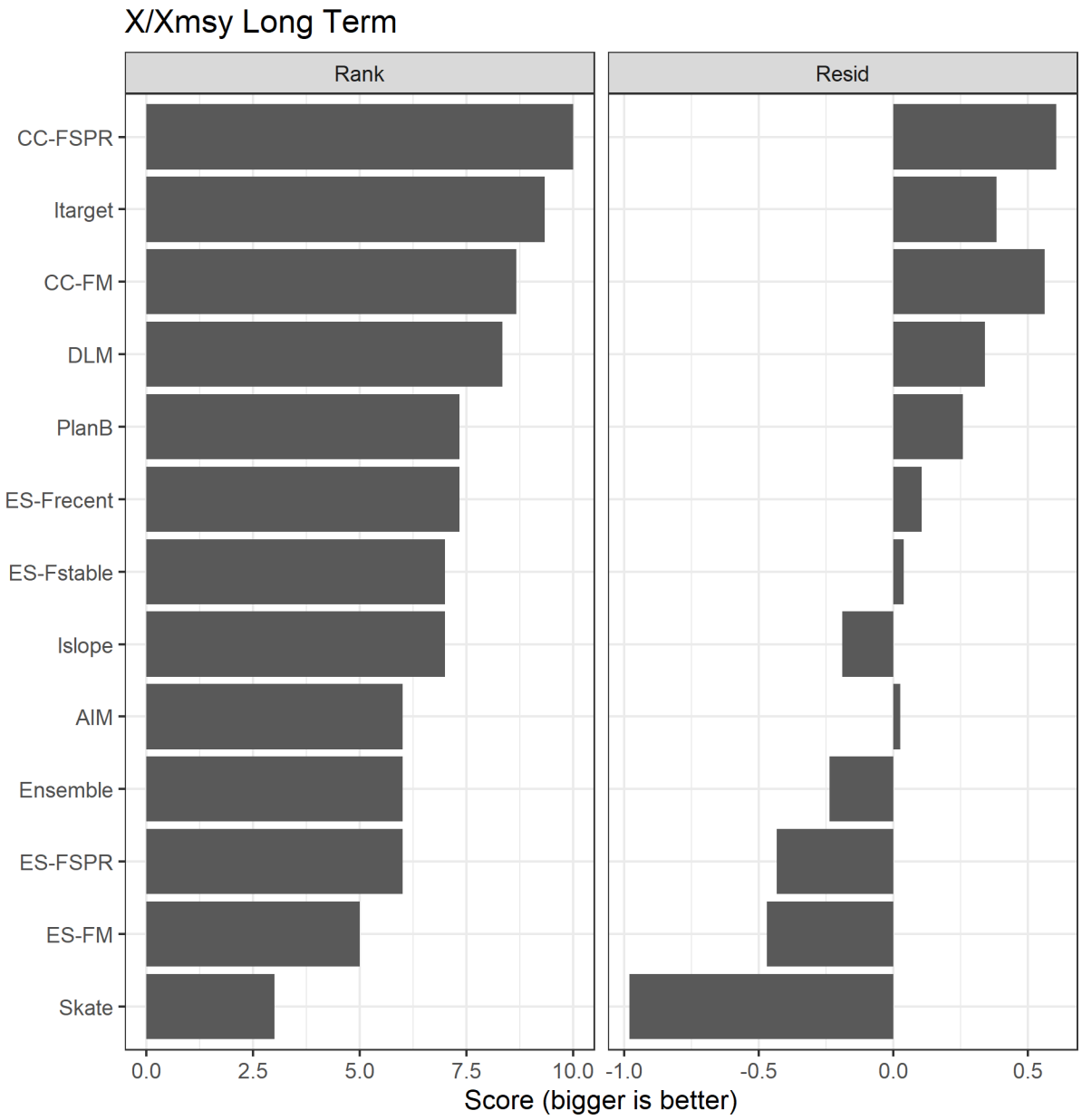


Figure 4.1. Two sets of scores, Rank and Resid, for the base analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/Xmsy) in the long term.

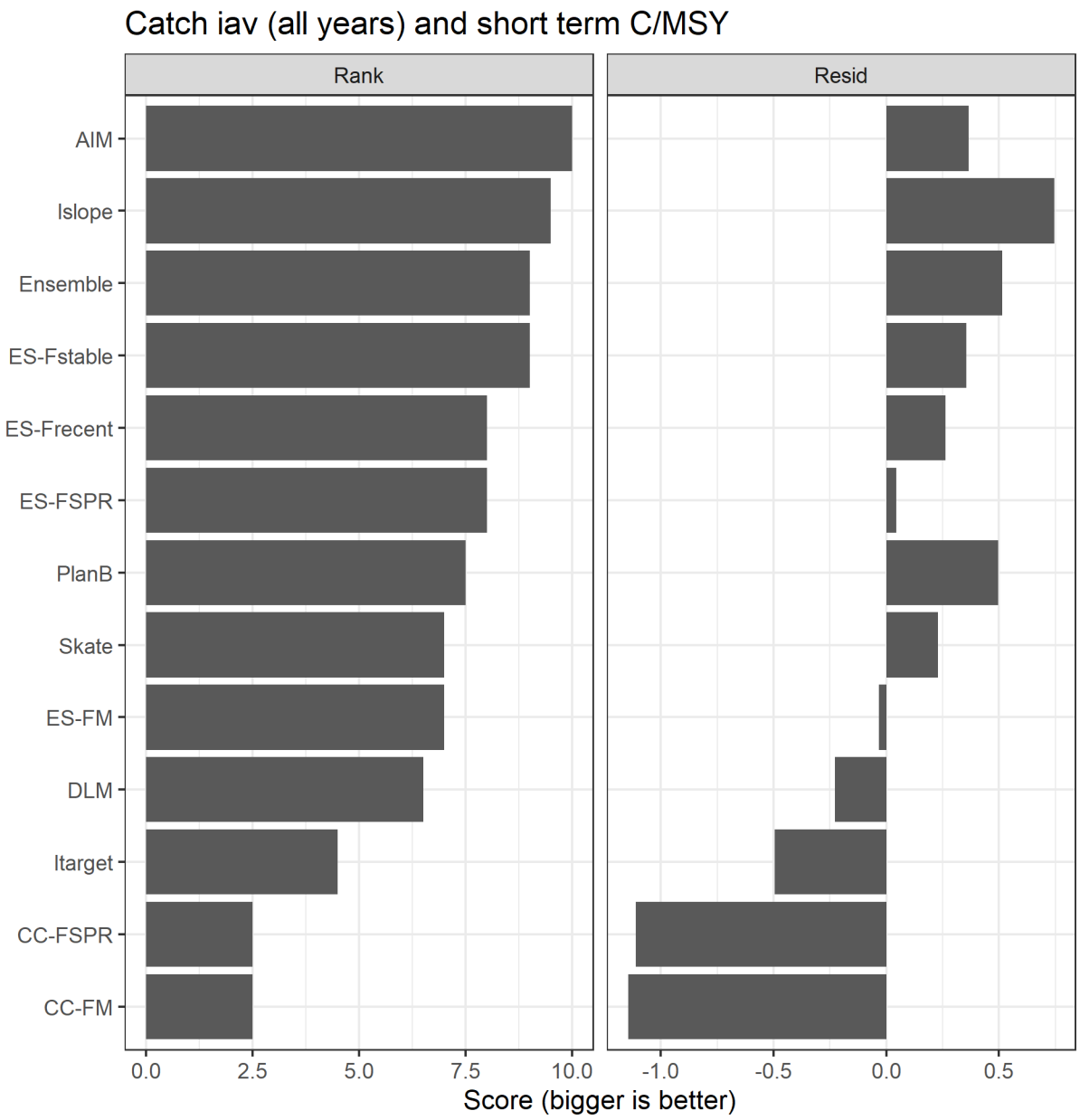


Figure 4.2. Two sets of scores, Rank and Resid, for the base analyses for the 2 metrics of interannual variability in catch over the entire feedback period and the short-term mean catch/maximum sustainable yield (MSY).

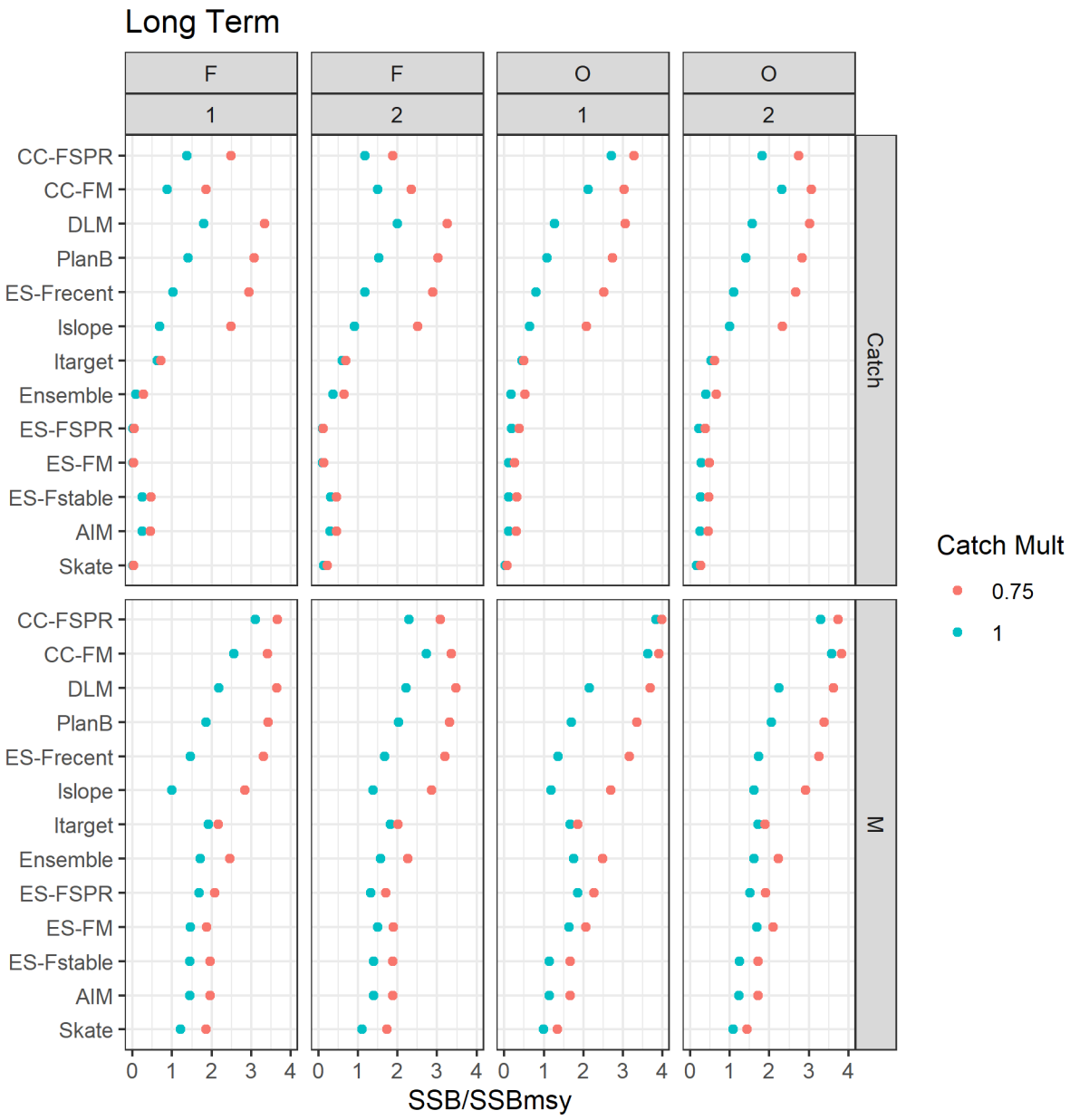


Figure 4.3. Mean values of spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

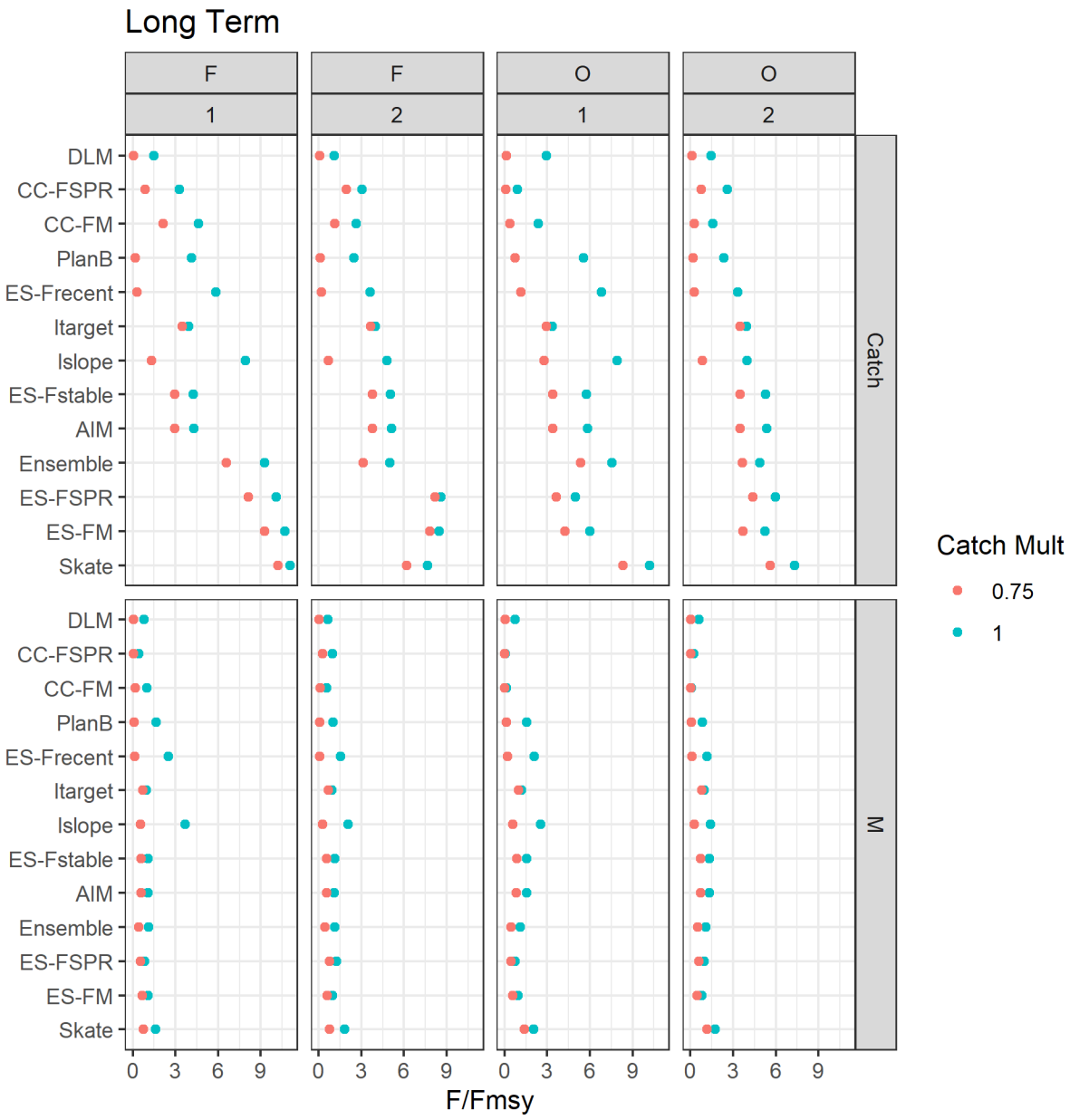


Figure 4.4. Mean values of fishing mortality rate (F) relative to the F at maximum sustainable yield (Fmsy) by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that lower values are at the top.

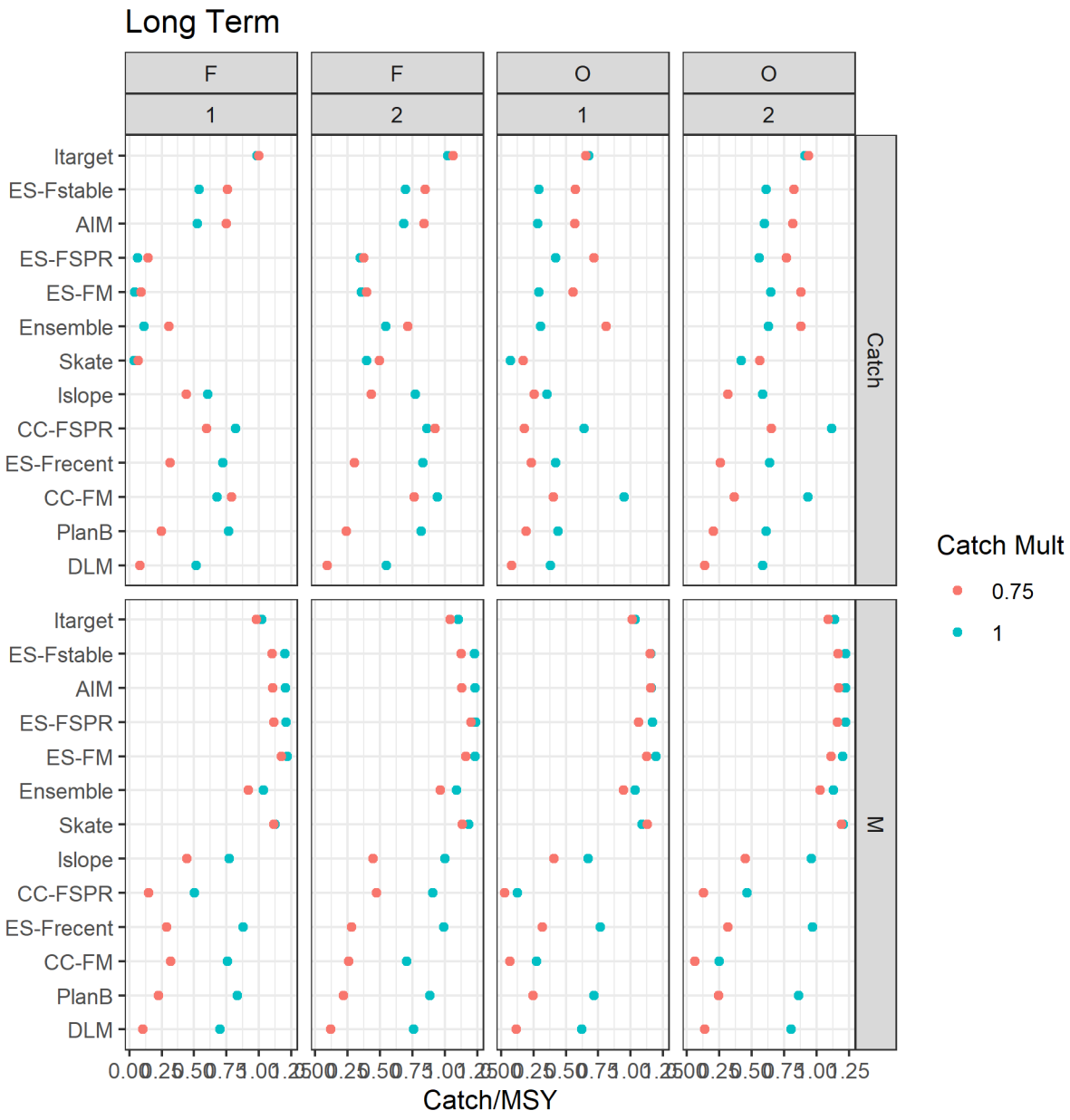


Figure 4.5. Mean values of catch relative to the maximum sustainable yield (MSY) by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

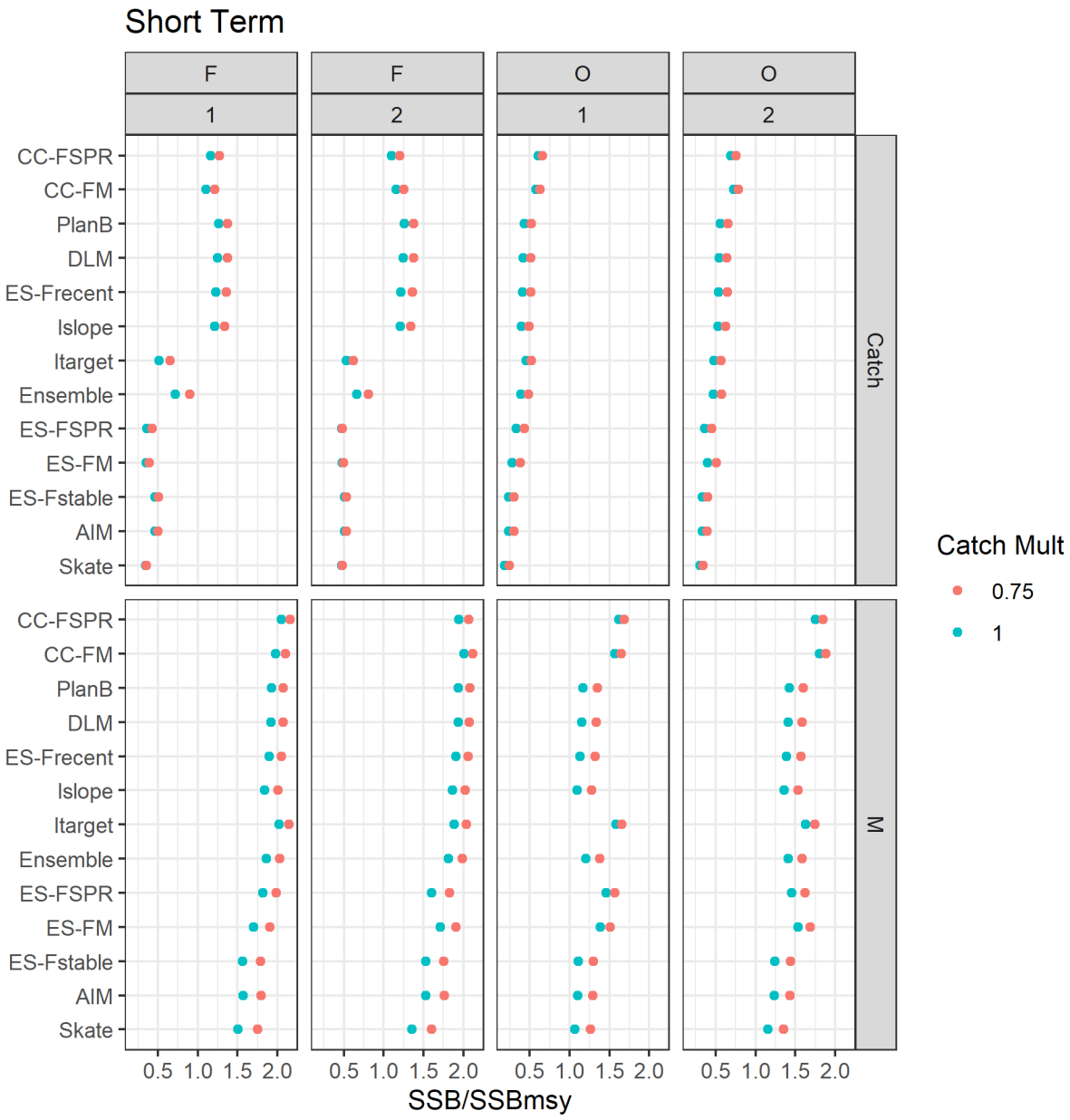


Figure 4.6. Mean values of spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

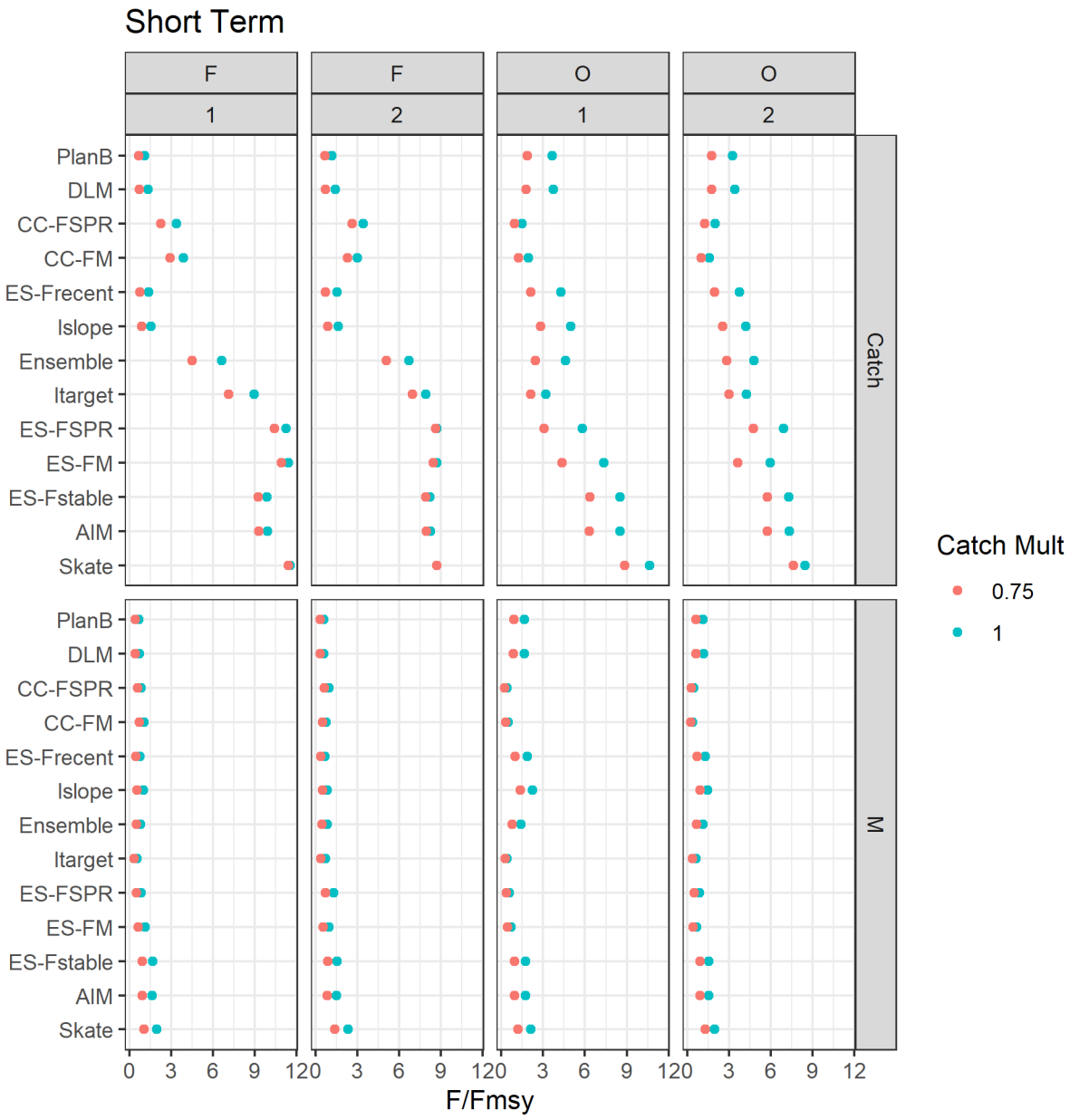


Figure 4.7. Mean values of fishing mortality rate (F) relative to the F at maximum sustainable yield (Fmsy) by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that lower values are at the top.

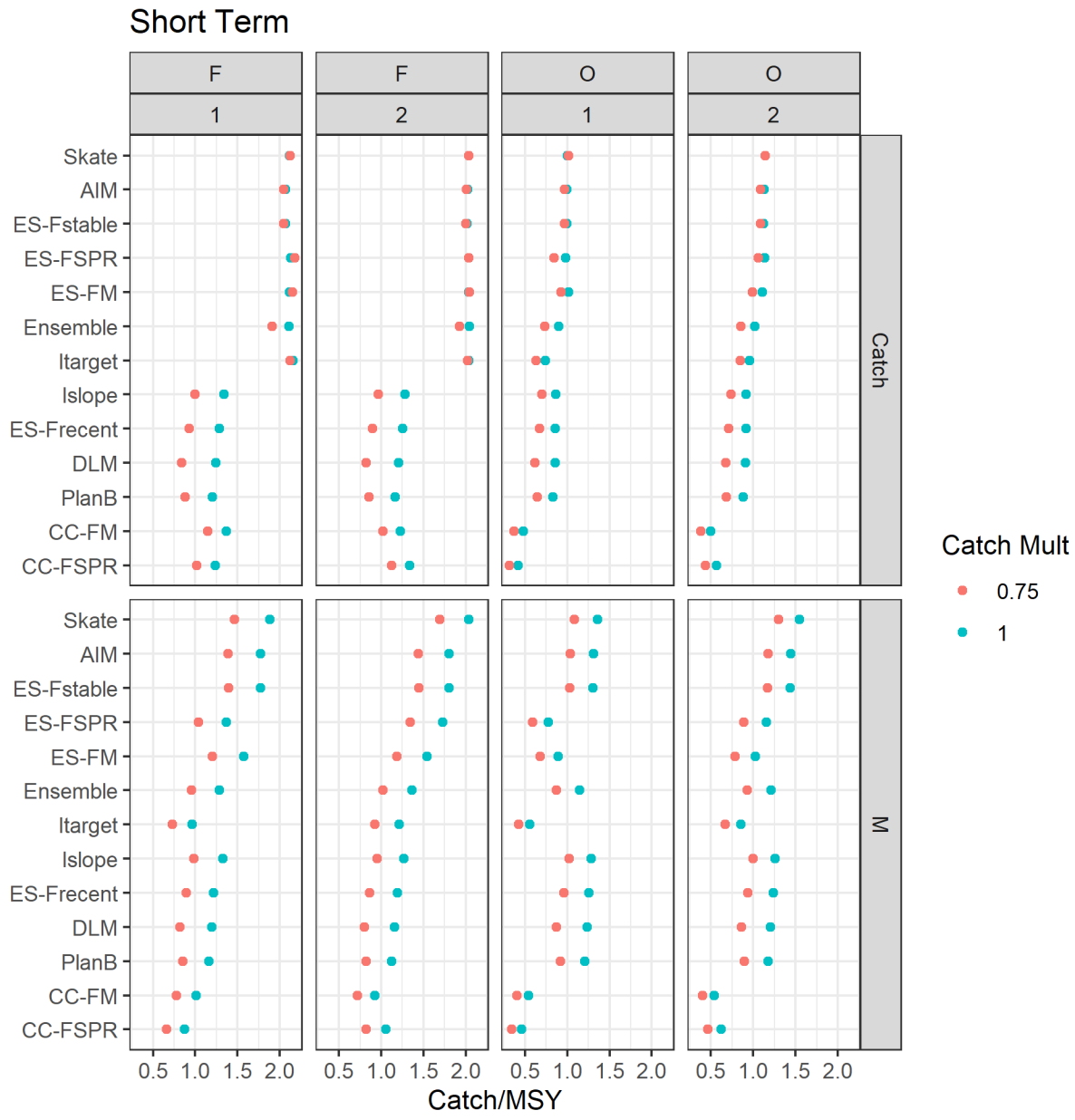


Figure 4.8. Mean values of catch relative to the maximum sustainable yield (MSY) by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of F history (F = Fmsy in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

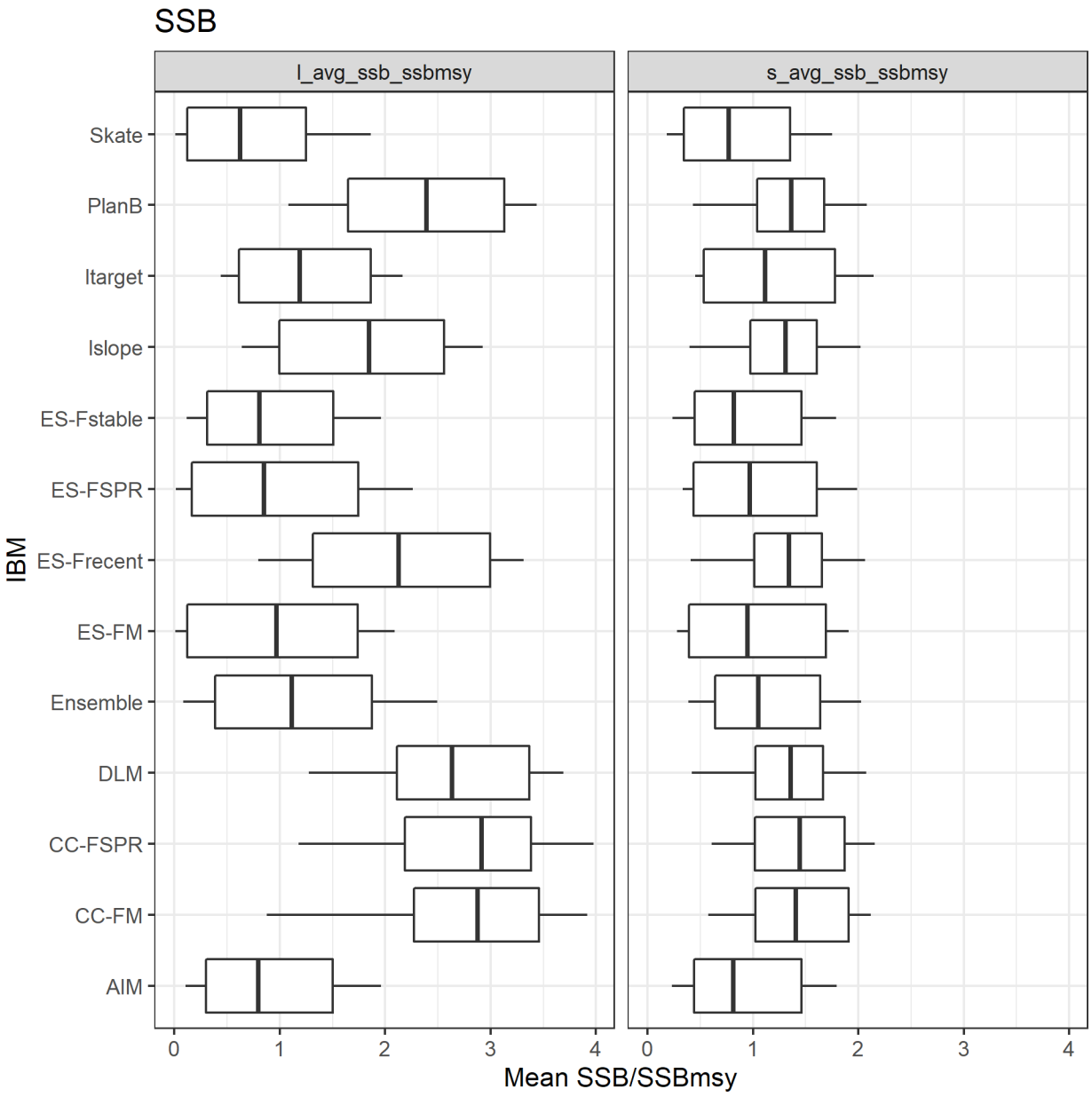


Figure 4.9. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) by index-based method (IBM) where the boxplot shows the distribution of the 16 scenarios in both the long and short term.

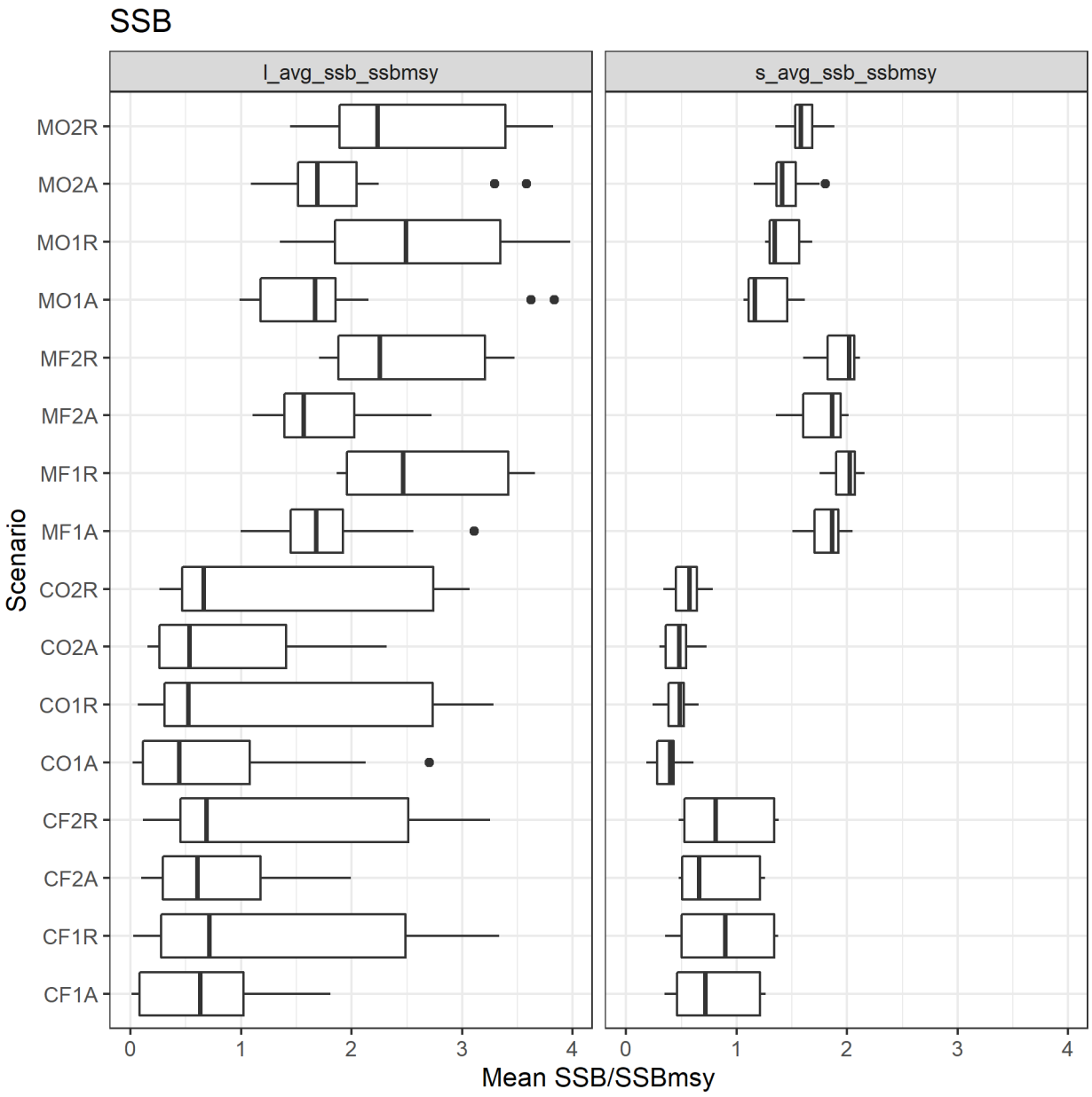


Figure 4.10. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) by scenario where the boxplot shows the distribution of the 13 index-based methods (IBMs) in both the long and short term.

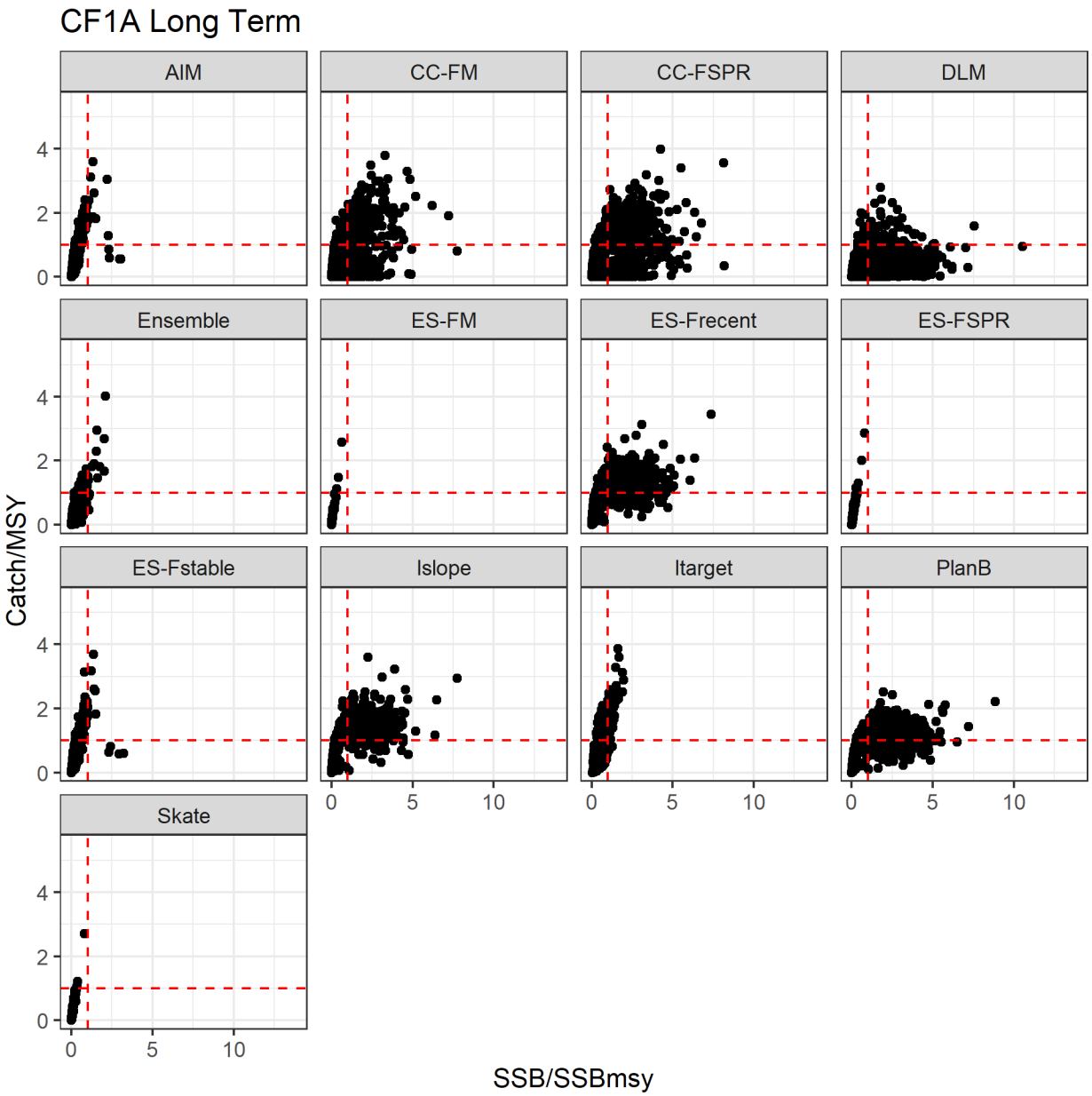


Figure 4.11. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) and catch relative to MSY for scenario CF1A in the long term with each dot representing 1 of the 1,000 simulations.

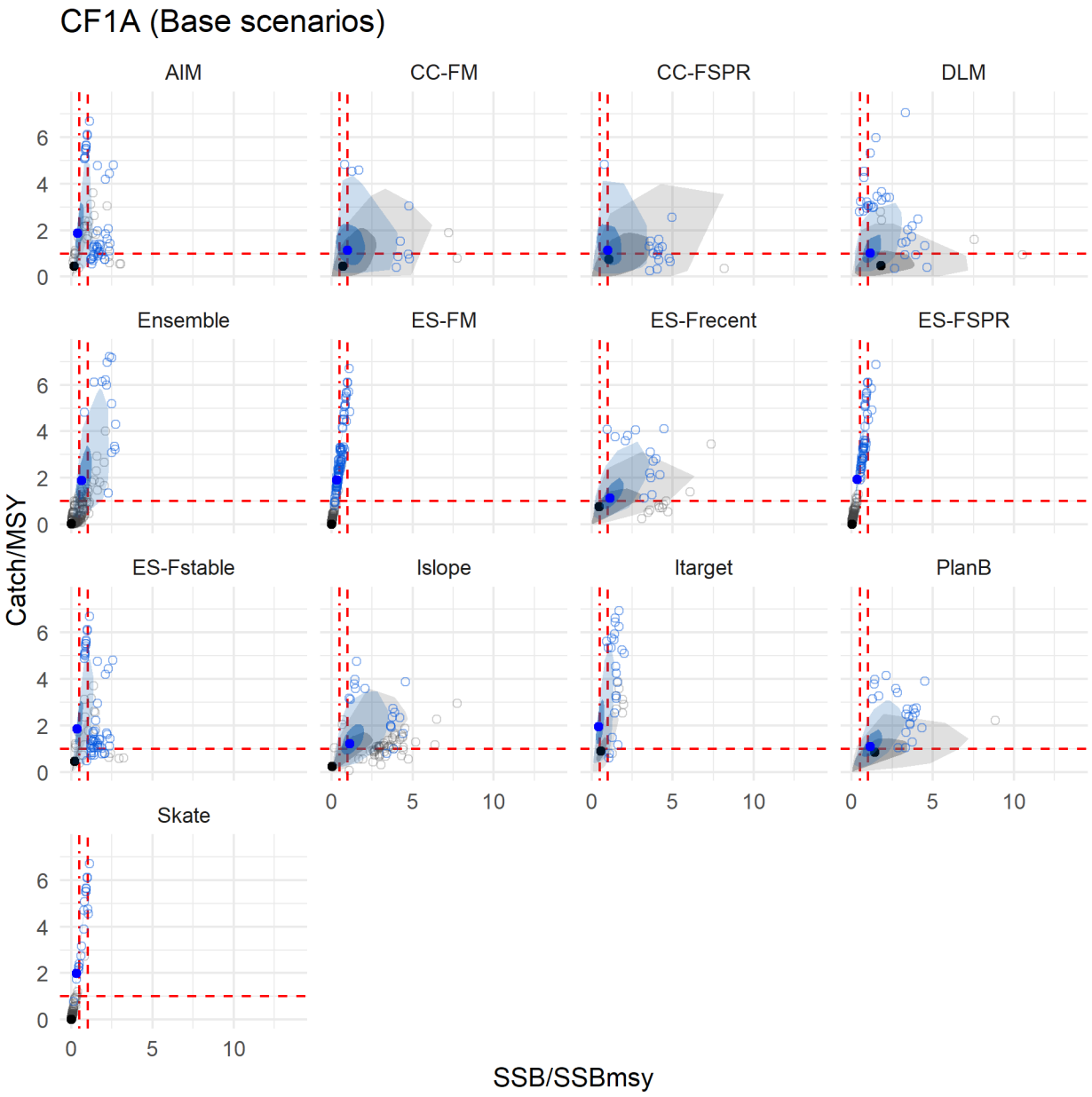


Figure 4.12. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSBmsy) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

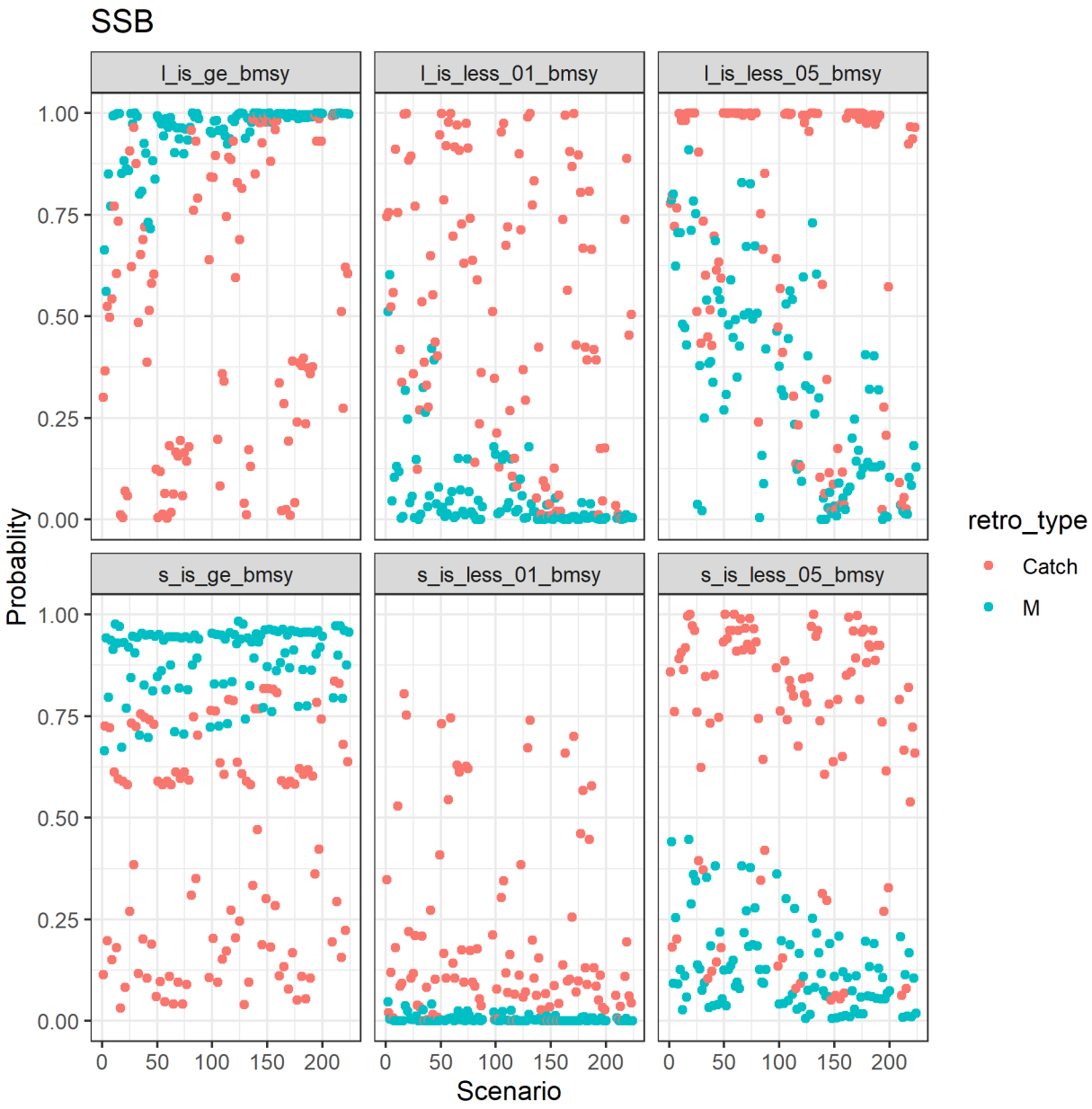


Figure 4.13. Mean value for 6 spawning stock biomass (SSB) metrics from the 1,000 simulations across the 208 combinations of index-based methods (IBMs) and scenarios with the color of the dots denoting the retrospective source.

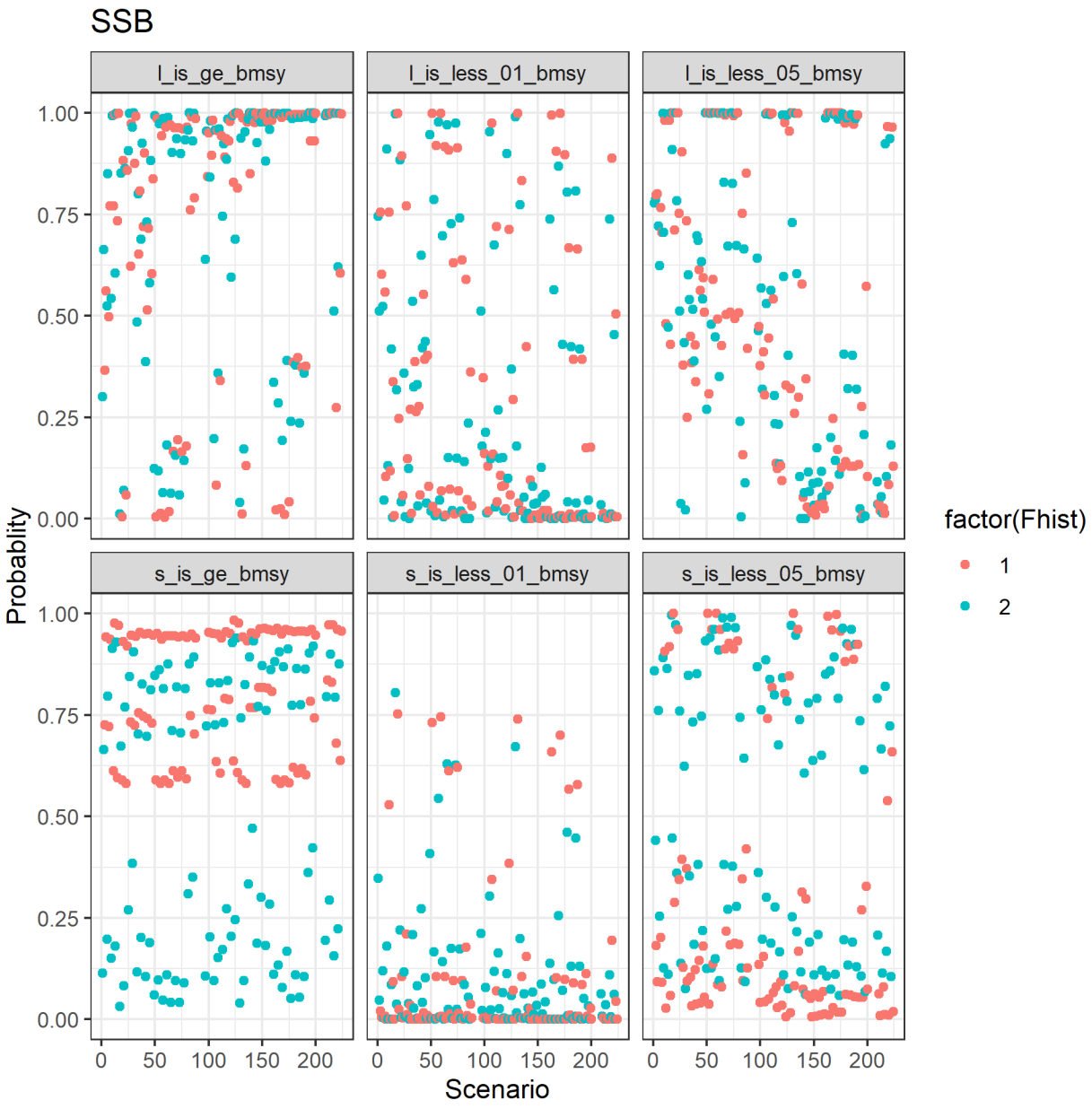


Figure 4.14. Mean value for 6 spawning stock biomass (SSB) metrics from the 1,000 simulations across the 208 combinations of index-based methods (IBMs) and scenarios with the color of the dots denoting the fishing history during the base period.

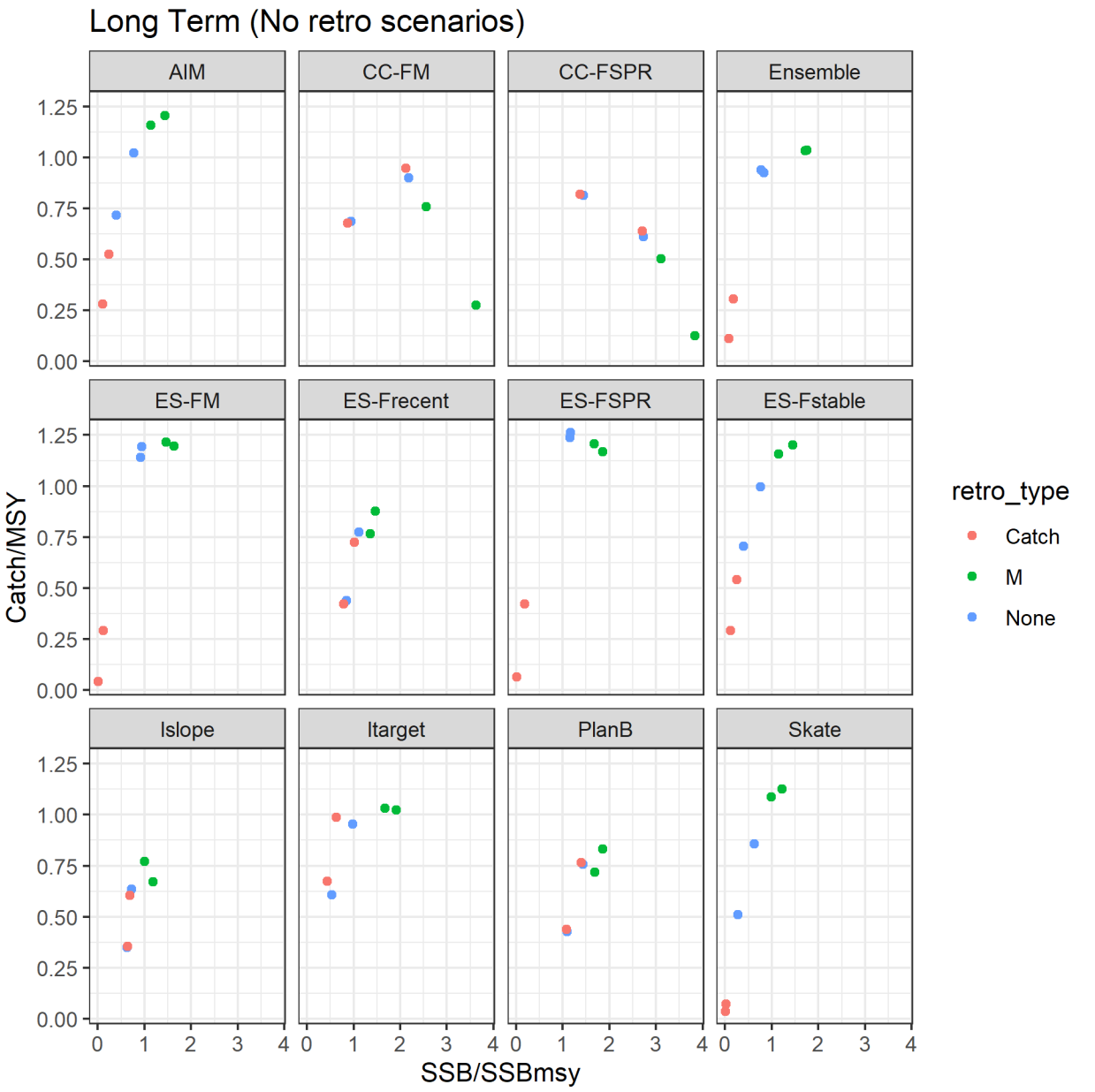


Figure 4.15. Mean spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) and catch relative to MSY for 3 sources of retrospective pattern (catch, natural mortality [M], and none) and 2 fishing histories in the long term based on 1,000 simulations.

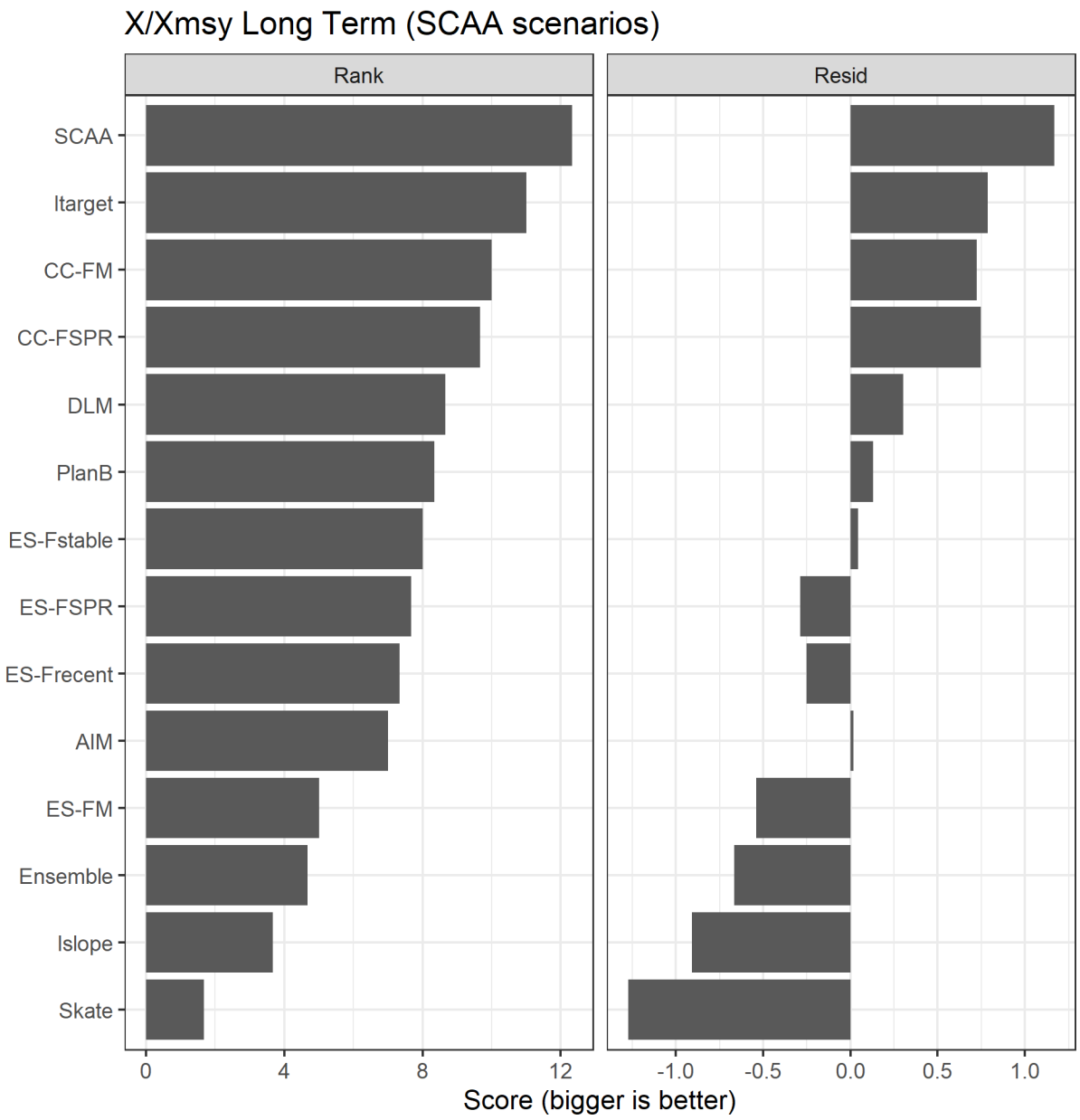


Figure 4.16. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/Xmsy) in the long term.

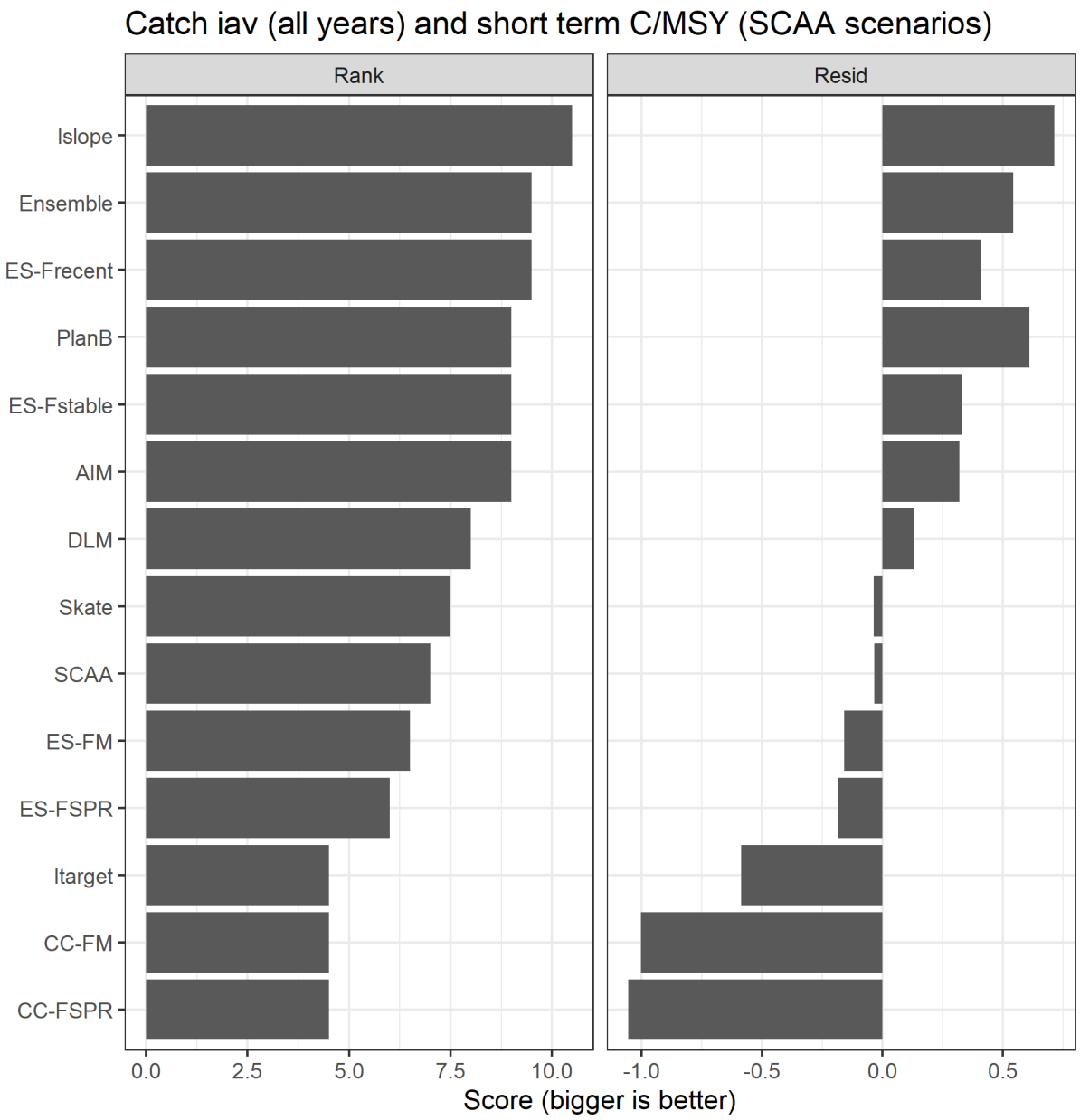


Figure 4.17. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the 2 metrics of interannual variability in catch over the entire feedback period and the short term mean catch/maximum sustainable yield (MSY).

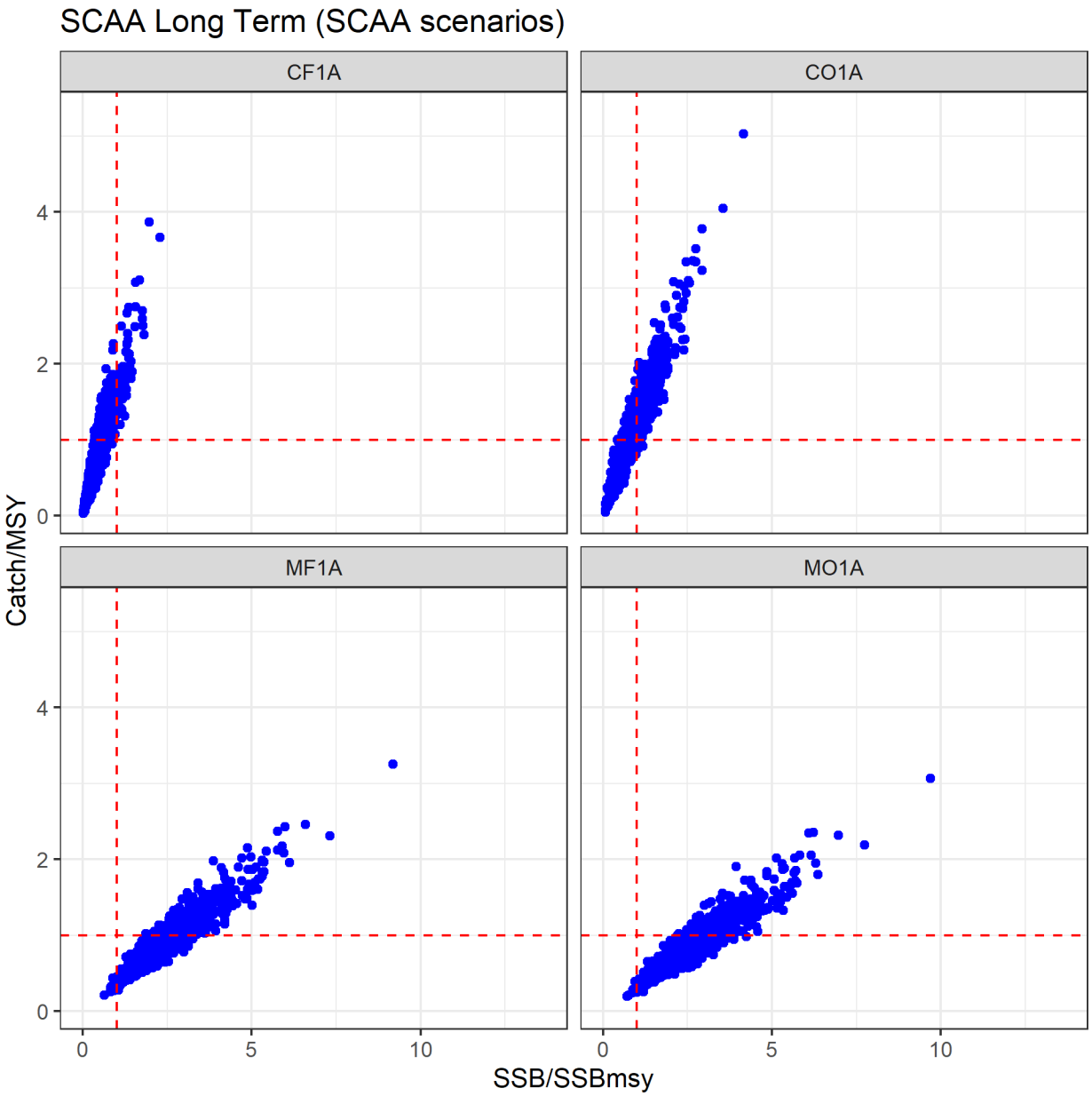


Figure 4.18. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSBmsy) and catch relative to MSY for the statistical catch-at-age (SCAA) model applied to 4 scenarios in the long term with each dot representing 1 of the 1,000 simulations.

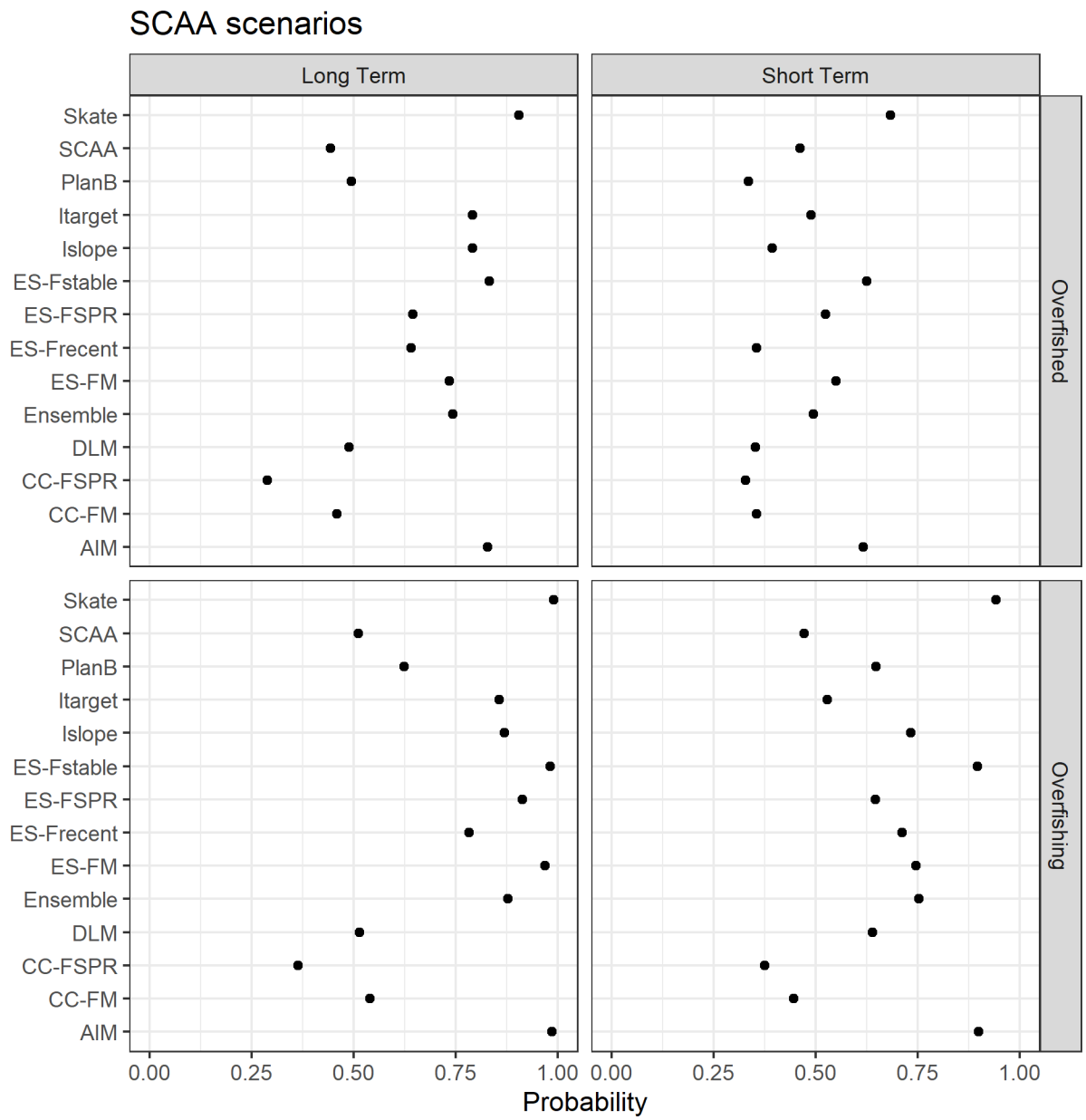


Figure 4.19. Probability stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all statistical catch-at-age (SCAA) analyses.

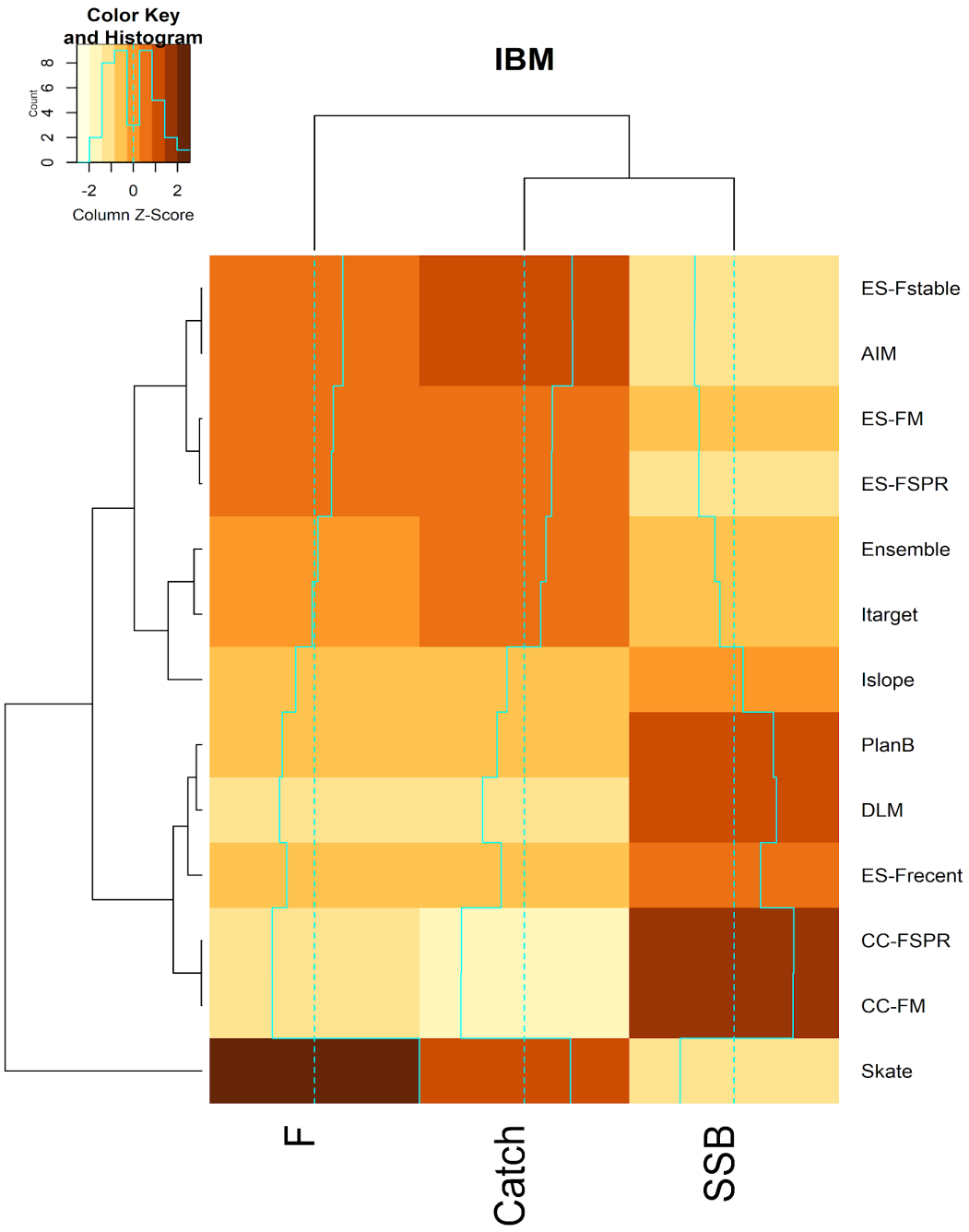


Figure 5.1. Heatmap of the median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate the position of that cell relative to the mean. Values for each cell in the heatmap can be found in Table 5.4 (rows in the same order as the heatmap).

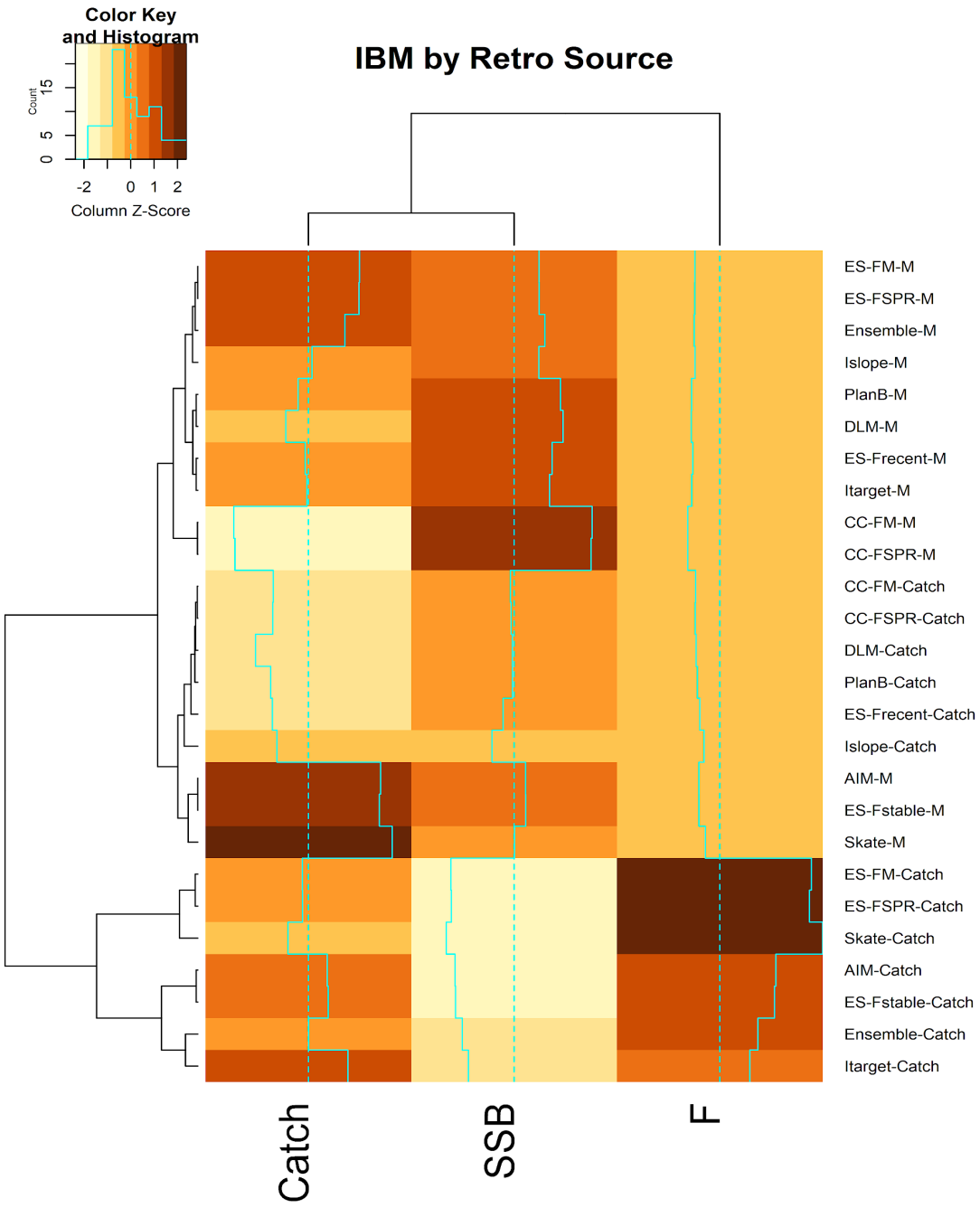


Figure 5.2. Heatmap of the median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY},) fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and source of retrospective pattern (C=catch, M=natural mortality). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate position of that cell relative to the mean. Values for each cell in the heatmap can be found in Table 5.5 (rows in the same order as the heatmap).

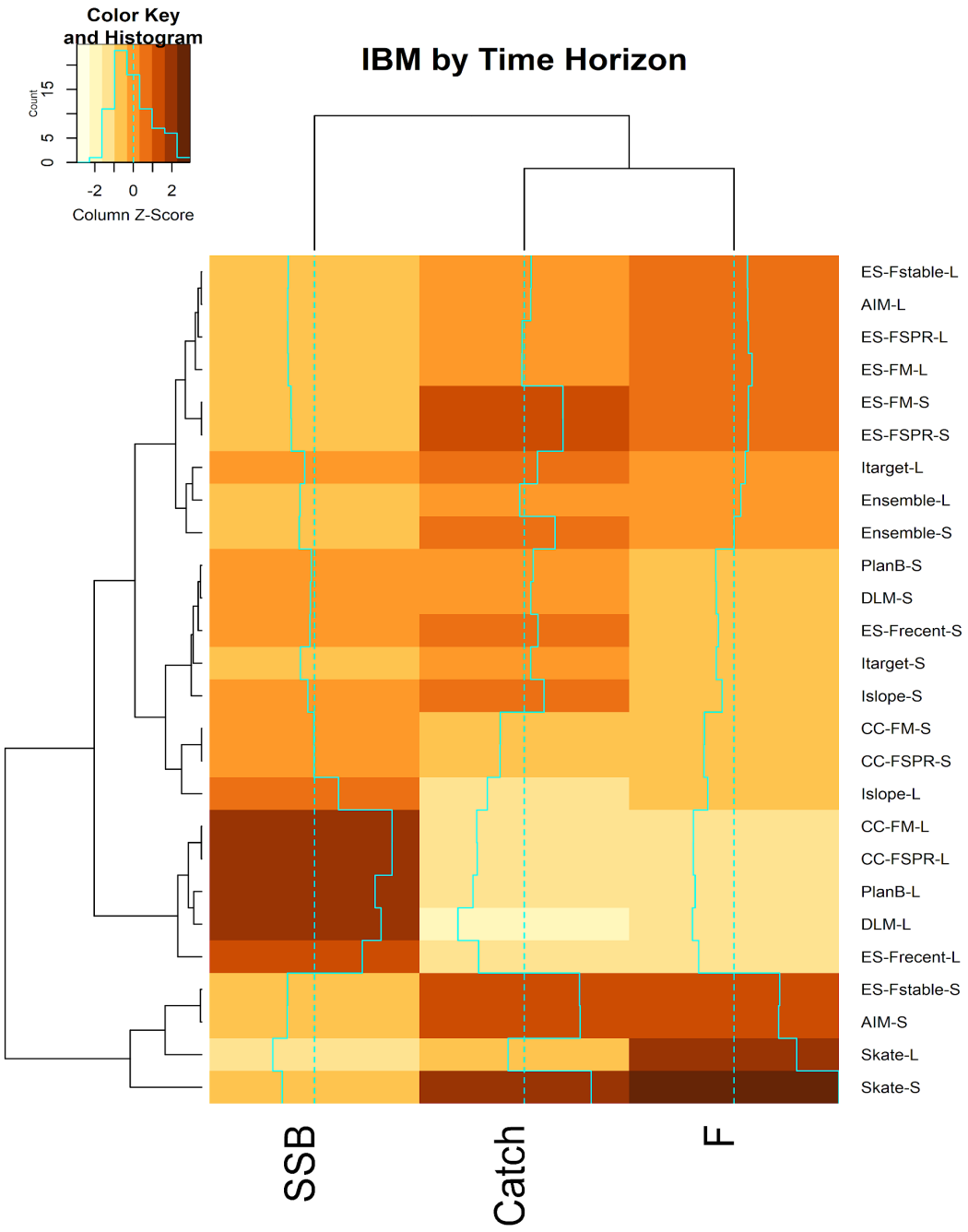


Figure 5.3. Heatmap of the median values for SSB/SSB_{MSY}, F/F_{MSY}, and Catch/MSY by IBM and time horizon (short- or long-term summary, 'S' or 'L'). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate the position of that cell relative to the mean. Values for each cell in the heatmap can be found in Table 5.6 (rows in the same order as the heatmap).

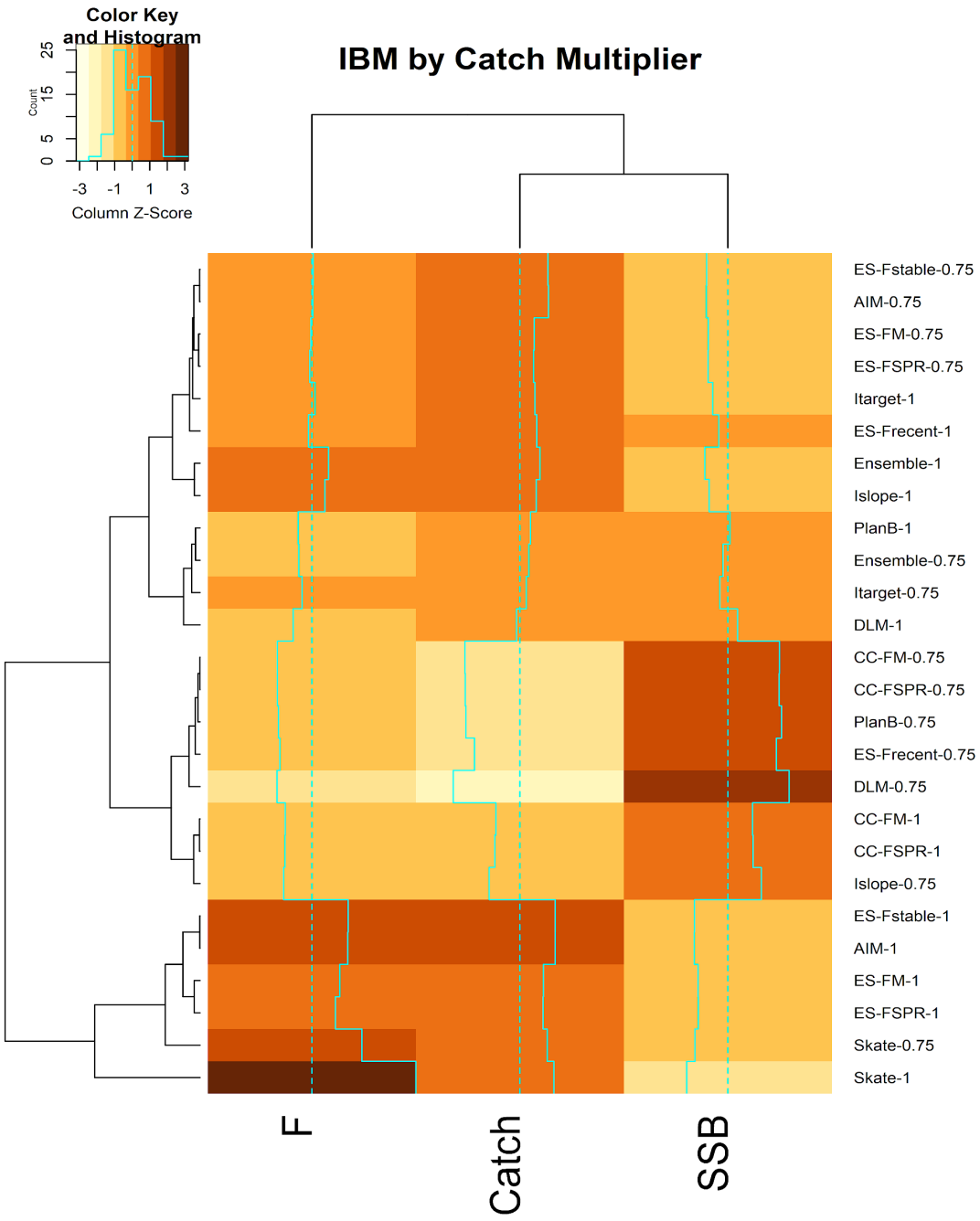


Figure 5.4. Heatmap of the median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and catch advice multiplier (1 or 0.75). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate the position of that cell relative to the mean. Values for each cell in the heatmap can be found in Table 5.7 (rows in the same order as the heatmap).

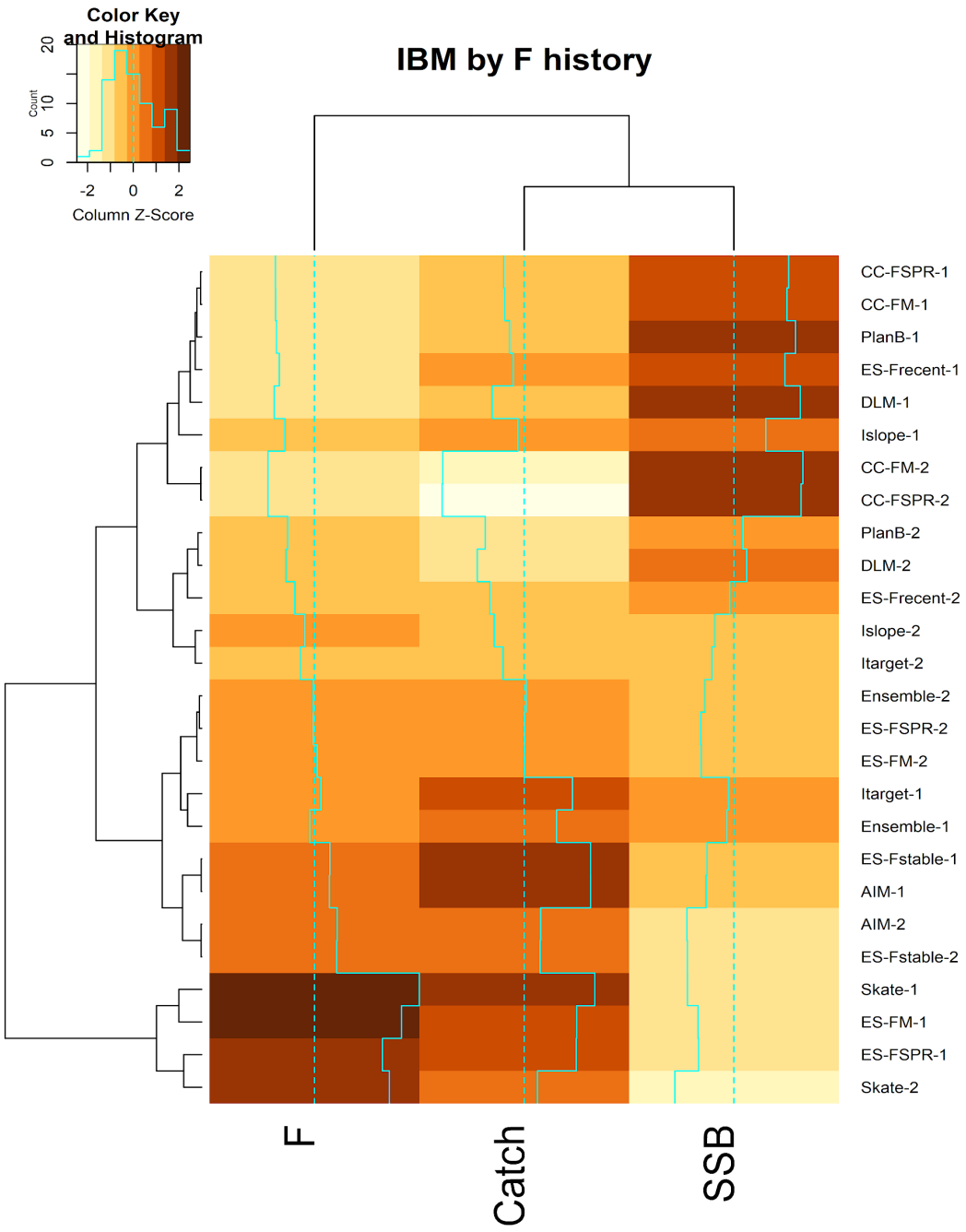


Figure 5.5. Heatmap of the median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM) and fishing history (1=overfishing then F=F_{MSY}, 2=always overfishing). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate the position of that cell relative to the mean. Values for each cell in the heatmap can be found in Table 5.8 (rows in the same order as the heatmap).

APPENDICES

Appendix 1 List of Working Group Members

Name	Affiliation
Andy Jones	NEFSC, Cooperative Research Branch
Brandon Muffley	Mid-Atlantic Fishery Management Council
Chris Legault (chair)	NEFSC, Population Dynamics Branch
Gavin Fay	University of Massachusetts Dartmouth, School for Marine Science and Technology
Jamie Cournane	New England Fishery Management Council
John Wiedenmann	Rutgers University
Jon Deroba	NEFSC, Population Dynamics Branch
Rich Bell	The Nature Conservancy
Tim Miller	NEFSC, Population Dynamics Branch

NEFSC = Northeast Fisheries Science Center

Appendix 2 List of Meeting Participants

(attended at least one meeting)

Brian Linton, NEFSC, Population Dynamics Branch
Brooke Wright, NEFSC, Cooperative Research Branch
Charles Perretti, NEFSC, Population Dynamics Branch
Deb Lambert, NOAA Fisheries, Office of Sustainable Fisheries
Dvora Hart, NEFSC, Population Dynamics Branch
Gary Shepherd, NEFSC, Population Dynamics Branch
Joe Langan, University of Rhode Island
Karen Greene, NOAA Fisheries, Office of Sustainable Fisheries
Kathy Sosebee, NEFSC, Population Dynamics Branch
Larry Alade, NEFSC, Population Dynamics Branch
Liz Brooks, NEFSC, Population Dynamics Branch
Mark Grant, GARFO, Sustainable Fisheries Division
Mark Terceiro, NEFSC, Population Dynamics Branch
Michele Traver, NEFSC, Assessment Process Lead

Paul Nitschke, NEFSC, Population Dynamics Branch
Paul Rago, Mid-Atlantic Fishery Management Council's Scientific and Statistical Committee
Susan Wigley, NEFSC, Population Dynamics Branch
Toni Chute, NEFSC, Population Dynamics Branch

Appendix 3 Meeting Summaries

On March 26, 2020, the IBMWG and other interested parties met via a webinar to discuss ways to approach the simulations needed to address the Terms of Reference. They also discussed tools for collaborating on this work, the limited timeline for the work, potential simulation software and modeling approaches, and aspects of the study design.

The group considered a wide range of possible factors to examine in the simulations, including:

- sources of retrospective patterns;
- number and type of IBMs (in addition to mean catch per tow);
- number of biological and fishery history patterns;
- changes in productivity over time;
- selectivity differences between the survey and fishery;
- multiple indices with varying spatial coverage;
- use of ensemble models; and
- estimation of biological reference points and rebuilding strategies.

Clearly, the working group does not have sufficient time to explore all combinations of these factors so will need to prioritize different aspects.

On April 2, 2020, the IBMWG and other interested parties met via a webinar to discuss IBMs for consideration in simulation testing. Working group members presented a review of currently used IBMs in the region and overviews of recent publications on the topic. The group discussed pros and cons of different methods but did not make any final selections.

On April 9, 2020, the IBMWG and other interested parties met via webinar to discuss causes of retrospective patterns that could be included in the simulations. Prior to this discussion, a Working Group member presented the IBM proposed for the red hakes as a follow up from the previous meeting. The group discussed a range of possible causes of retrospective patterns and priorities for including them as part of the operating model for the simulations but did not make any final decisions.

On April 16, 2020, the IBMWG and other interested parties met via webinar to discuss desired features of the operating model to be used in the simulation studies. The group focused on a number of decision points regarding aspects of the model with emphasis on those necessary to meet the Terms of Reference.

On April 23, 2020, the IBMWG and other interested parties met via webinar to discuss whether to condition the operating model on specific stocks in the region or to develop a number of situations that could apply more broadly. The group was unanimous in favor of the latter approach but requested confirmation that such an approach would allow leads for stocks in a management track to change the index-based approach based on the outcome of this working group.

On April 30, 2020, the IBMWG and other interested parties met via webinar to discuss how to present results of the work to both technical and management audiences. The chair first

presented an R Shiny app that demonstrated a number of features desired in the simulation studies. The discussions then focused on how to summarize the large amount of information that results from closed-loop simulations in meaningful ways for different audiences.

The May meetings were all scheduled for 2 hours to allow additional deliberations, if a topic required them.

On May 7, 2020, the IBMWG and other interested parties met via webinar to discuss parameter values and default settings for the operating model to be used in the simulations. The group relied on a number of presentations provided by IBMWG members regarding data from stocks in the region in order to make the default settings as realistic as possible without exactly matching any one particular stock. The group agreed to continue development of code used to demonstrate results from preliminary settings.

On May 14, 2020, the IBMWG and other interested parties met via webinar to discuss a proposed workflow for developing the simulator for this project. The group agreed on developing a new R package and writing an R script using the functions in this package would allow meeting the Terms of Reference. The group discussed a preliminary list of functions and a conceptual model of how the R script would work.

On May 21, 2020, the IBMWG and other interested parties met via webinar to take stock of R code readily available for us in this project. A member of the working group discussed the possibility of using a new stock assessment model (Woods Hole Assessment Model [WHAM]) simulation capability as the operating model. Further work is needed to determine whether the feedback portion of the simulations can be completed using this approach. A number of sources of R code were identified for use by the group.

On May 28, 2020, the IBMWG and other interested parties met via webinar to discuss which IBMs and control rule combinations should be examined in the simulations. A member of the working group provided an example using WHAM as the operating model for the simulations including the feedback period. The group decided to use WHAM as the operating model and noted that a small helper function is needed to initialize the base period.

On June 4, 2020, the IBMWG and other interested parties met via webinar to discuss metrics to collect from the simulations to allow comparisons of trade-offs among IBMs and control rules. The working group first agreed to some terminology before discussing the metrics. The metrics agreed to focus on catch, biomass, and fishing mortality rate both in the short and long term of the simulation feedback period. Two approaches to storing the simulation results were presented.

On June 11, 2020, the IBMWG and other interested parties met via webinar to compare and contrast 2 approaches to saving simulation results. The working group selected the tibble approach, which allows saving all the information from each run in a list of lists format. Saving these files in compressed RDS format makes storage of this large amount of information reasonable. Subgroups were formed to work on some specific coding aspects remaining before the simulations can begin.

On June 18, 2020, the IBMWG and other interested parties met via webinar to discuss progress being made by multiple subgroups on coding. The code now allows retrospective patterns to be generated by missing catch, increased M, or changing survey q. The group agreed to set up the input files to use a single base approach with modifications from it using the list feature in R with default values. The subgroup developing the IBM functions presented a draft list of features with emphasis on each function receiving the same list of data and producing an output list with catch advice as its first element. How to combine the 2 indices was tabled until next week.

On June 25, 2020, the IBMWG and other interested parties met via webinar to discuss coding updates from the multiple subgroups. Default settings for WHAM as the operating model were described. A function to create the default input settings has been created. Local stocks have been examined for reasonable weight at age estimates. Coding of the index-based approaches has begun with a draft document describing the object containing all the information needed across all approaches drafted. Default settings for the feedback period were agreed. How to combine the 2 indices was again deferred.

On July 2, 2020, the IBMWG and other interested parties met via webinar to discuss additional coding updates from the subgroups. A description of the elements provided with the simulated data from a WHAM operating model was provided. Progress continues on the functions to create new inputs from the default settings as well as the functions to compute the catch advice from the different approaches. The working group decided to use as the default, when needed, an average of the fall survey in year t with the spring survey in year t+1. Sensitivity runs will be conducted to explore any differences when the spring and fall survey in the same year are averaged.

The IBMWG did not meet on July 9, 2020, or July 16, 2020, due to multiple working group members having conflicts with those dates.

On July 23, 2020, the IBMWG and other interested parties met via webinar to discuss additional coding updates from the subgroups. A Workflow proposal has been added to the Wiki on the GitHub site describing how the functions can be used to create new scenarios for the simulations. The IBMs functions subgroup provided a list of questions prior to the meeting that were addressed by the group. The functions are nearly complete and will be ready for testing soon.

On July 30, 2020, the IBMWG and other interested parties met via webinar to discuss a range of topics. A working group member presented a management strategy evaluation (MSE) Shiny app that allowed comparison of results across multiple scenarios as a possible teaching tool for readers of our report. Subgroups updated progress. The design matrix for the scenarios to be run was discussed and modified. Turning the R functions into a package and a strawman outline of the report were tabled until next week.

On August 6, 2020, the IBMWG and other interested parties met via webinar to get updates from the sub-groups and discuss a draft outline of the report. A new subgroup was formed to join together the work produced so far and produce a simple set of outputs. Another subgroup was formed to begin drafting the introduction section of the report. A potential outlet for a primary literature publication based on this work was identified.

On August 13, 2020, the IBMWG and other interested parties met via webinar to get updates from the subgroups and discuss creation of an R package for this work. All the subgroups are making good progress. A number of short-term items were identified for each subgroup. The group will try to create an R package of functions used in this work to have another product, in addition to the report, that can be used in the future but will need to determine if there is sufficient time to do so.

On August 20, 2020, the IBMWG and other interested parties met via webinar to update progress by all the subgroups. Excellent progress continues, with a number of issues resolved during the week. A demonstration of 4 scenarios with 10 realizations was completed, and the results were posted to the GitHub site to allow thinking about summary tables and figures. The working group agreed to pursue the Frontiers in Marine Science special issue on data-limited research in stock assessment as an outlet for this work.

On August 27, 2020, the IBMWG and other interested parties met via webinar to discuss a punch list of remaining items in the simulation code that need to be addressed before the runs can

commence. Many of the items were identified as completed, and others were assigned to specific individuals. The working group also began discussions about plots that could be used in the final report to summarize results based on some demonstration plots created with fake data.

On September 3, 2020, the IBMWG and other interested parties met via webinar to discuss the few remaining items that need to be addressed before the simulations can begin. Many issues identified last week were completed. Some new issues arose during the discussions that were assigned to meeting participants with the goal of having the code fully functional by next week.

On September 10, 2020, the IBMWG and other interested parties met via webinar to review the remaining coding issues and plan for starting the simulation runs. A new file structure for the code and settings was agreed for the final versions of the code, with only a few minor tweaks remaining. Approaches for writing the document were discussed and agreed to, as well.

On September 17, 2020, the IBMWG and other interested parties met via webinar to discuss the few remaining issues on the code To Do list. The final code was moved to the new directory structure and final issues identified. Working group members and others volunteered to draft sections of the report.

On September 24, 2020, the IBMWG and other interested parties met via webinar to discuss some issues that arose in the simulation code. A number of these issues were quickly resolved and the others tasked with short deadlines. The working group also reviewed a presentation on how to run the simulations and updated on progress with report writing.

On October 1, 2020, the IBMWG and other interested parties met via webinar to discuss approaches that could be used to summarize the results of the simulations. The large number of factorial combinations and many realizations per scenario means that a lot of results will be produced. The working group examined a number of summary approaches that could provide both filtering of poor performing methods as well as means of comparing large numbers of scenarios.

On October 8, 2020, the IBMWG and other interested parties met via webinar to discuss preliminary results from the simulations. Demonstration tables and plots were examined for potential use in the report and to increase understanding of IBM performance. A problem in one factor of the simulations was discovered during this examination, requiring code changes.

On October 15, 2020, the IBMWG and other interested parties met via webinar to examine additional summary plots and tables with more simulations. A number of suggestions were made to improve the plots, and additional plots were suggested. The problem identified last week was addressed and simulations continue to be run. Draft text continues to be added to the report sections that do not depend on the results.

On October 22, 2020, the IBMWG and other interested parties met via webinar to decide on the IBMs to be included in the ensemble method. Preliminary results of a linear modeling exercise to determine significance of factors and interactions across performance metrics and a scoring approach to rank the IBMs were examined to inform the decision. Both pragmatic consideration (amount of time to conduct simulations) and balance among model types informed the decision regarding which IBMs to include in the ensemble model.

On October 29, 2020, the IBMWG and other interested parties met via webinar to discuss results to date. The metrics to include in scoring of the IBMs were defined, with a layered approach using few metrics to begin and inclusion of additional metrics to answer more detailed questions. The one-off simulations to conduct were also determined, acknowledging the limited time remaining before the report is due.

On November 5, 2020, the IBMWG and other interested parties met via webinar to discuss preliminary results and conclusions to include in the report. The final data set for the factorial

combinations of IBMs and scenarios is complete. The results from a statistical catch-at-age model used in place of an IBM are also available. Soon to be added are the last results from the no retrospective source scenario.

On November 12, 2020, the IBMWG and other interested parties met via webinar to discuss a Shiny app developed to allow users to compare results for IBMs across different metrics. The group also discussed topics to highlight in the report, focusing on ToR 4. It was noted that a lot has been accomplished and the deadline for the report is coming fast, so writing is important.

On November 19, 2020, the IBMWG and other interested parties met via webinar to discuss the few remaining items to be decided in the draft report to be submitted for review. The group edited the executive summary to ensure consensus on the statements made there and highlighted where finishing touches were needed in the main report. The chair thanked the working group and everyone who contributed to the project for all their hard work.

Appendix 4 General Overview of WHAM

WHAM (Miller and Stock 2020) is a R package that employs a TMB-based generalization and extension of the age-structured models of Miller et al. (2016) and Miller and Hyun (2018). Stock and Miller (2021) give a full description and simulation study of the statistical behavior of the model under alternative configurations. There are several similarities to the Age-Structured Assessment Program (ASAP; Legault and Restreppo 1998; Miller and Legault 2015). WHAM can be configured to fit statistical catch-at-age models nearly identically to ASAP, and there is functionality in the WHAM package to migrate ASAP input files to R inputs needed for fitting WHAM models. WHAM uses many of the same types of data inputs, such as empirical weight-at-age, so existing assessments in the U.S. Northeast can be easily replicated and tested against models with state-space and environmental effects in a single framework.

WHAM is different from ASAP and other statistical catch-at-age (SCAA) models through the implementation of random effects on 3 processes—interannual transitions in numbers at age, M, and selectivity—and allowing effects of environmental covariates on recruitment and M. These random effects make WHAM a state-space model, and models are fitted by maximizing the marginal log-likelihood with respect to the fixed effects parameters that also include the variance parameters associated with the random effects distributions. WHAM also includes state-space models for environmental covariates where the true values are represented by random effects that may be observed with error. There are several alternatives for how to model the correlation of the different random effects in WHAM.

As in ASAP and many other SCAA assessment frameworks, WHAM assumes separability in the fishing mortality rate by age and year. Three parametric forms are available (logistic, double-logistic, and decreasing-logistic), as well as a non-parametric option to estimate each age-specific selectivity parameter. To allow for temporal changes in selectivity as in ASAP, WHAM can estimate selectivity in user-specified time blocks and as a time series process with random effects within each block as mentioned above.

There are numerous ways to configure environmental covariate models and effects on recruitment and M, but these are not employed in the operating model we used for our simulation studies. See Stock and Miller (2021) and vignettes at https://timjmiller.github.io/wham/ for more details.

Like ASAP, there are observation likelihood components for aggregate catch and abundance index for each fleet and index, and associated age composition. The log-aggregate catch log-aggregate index observations are assumed to have a normal distribution where bias correction

terms associated with log-normal variables can be used. WHAM includes several options for the catch and index age compositions including multinomial (default), Dirichlet, Dirichlet-multinomial, logistic normal, and a zero-one inflated logistic normal.

Projections are possible within the WHAM model and the number of projection years equates to the feedback period described in ToR 1. The population is projected for a specified number of years. Fishing mortality during the projections are based on a fully-selected fishing mortality that can be user-specified, the average over specified model years, or that associated with a percentage of spawning potential ratio. The age-specific values of fishing mortality are obtained by using selectivity at age averaged over a specified set of model years. Reference points during the projection period also use values of maturity, weight, and M at age averaged over specified years. For models with random effects on numbers at age, M, or environmental covariates, the default is to continue the stochastic process into the projection years.

References for Appendix 4

Legault CM and Restrepo VR. 1998. A flexible forward age-structured assessment program [PDF]. ICCAT Working Document SCRS/98/58. 15 p.

Miller TJ and Hyun S-Y. 2018. Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. C J Fish Aquat Sci. 75(5):691-703. Accessible at: doi:10.1139/cjfas-2017-0035

Miller TJ, Hare JA, Alade LA. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to Southern New England yellowtail flounder. Can J Fish Aquat Sci. 73(8):1261-1270. Accessible at: doi: 10.1139/cjfas-2015-0339

Miller TJ and Legault CM. 2015. Technical details for ASAP version 4.0. US Dept Commer Northeast Fish Sci Cent Ref Doc. 15-17. 136 p.

Miller TJ and Stock BC. 2020. WHAM: a state-space age-structured assessment model. Version 1.0. NOAA Fisheries. Accessible at: https://timjmiller.github.io/wham/

Stock BC and Miller TJ. 2021. The Woods Hole Assessment Model (WHAM): a general state-space assessment framework that incorporates time- and age-varying processes via random effects and links to environmental covariates. Fish Res. 240:105967.

APPENDIX 5 DESCRIPTION OF WORKFLOW

To conduct some simulations, you first need to copy the files from the GitHub repository to your computer. The file structure is assumed within the calls, so you need to have the R code in the code directory, the input and tracking files in the settings directory, and have available an output directory to save the results. This is most easily accomplished by cloning the repository.

You'll next want to install the latest version of wham using devtools::install_github("timjmiller/wham", dependencies=TRUE). Note WHAM has a lot of package dependencies and you'll need to install all of them. This may take a while.

Once you have the files on your computer and the latest version of WHAM installed, you're ready to start running simulations. In R, you need to open the do_mse.R file and run the code there. This will prepare you to be able to conduct the actual simulations. Once you've run the code in do_mse.R, you are ready to conduct some simulations. To do so, you simply need to provide 3 arguments to the do_mse function:

- nsim = the number of simulations you would like to run (an integer greater than 0)
- user = your initials (just for tracking purposes)
- write_to_google = TRUE or FALSE to determine whether to automatically write the output to the Google Drive; if you select FALSE, you need to do so by hand

Note that it is best not to run the call to do_mse() from do_mse.R itself.

@gavinfay has created a run_script.R in the root folder of the repo that can be used to run simulations from the command line.

There isn't a need to change the script for the user and simulation numbers, but this is currently hardwired to set write_to_google = FALSE.

This can be run at the command line by typing something like:

$ Rscript run_script.R 10 GF

This will run 10 simulations for user "GF".

Under the Hood

So how did we get to the point where a single function call can conduct a number of simulations and keep us from repeating or missing runs? Teamwork! There are a series of codes that were run to set up the scenarios and control files for these simulations.

- rseed_list created the 1000 random seeds used for the 1000 realizations in each scenario
- design_matrix created the 224 factorial scenarios (and will eventually also create all the one-offs)
- setup_scenarios created the 3 rds files in the settings directory that keep track of all the scenarios that are needed and which have already been run

WHAM is the R package used as the operating model to simulate the data according to each scenario's specifications, a lot of prior work was needed using this package to define the simulations to achieve the desired Mohn's rho for spawning stock biomass (SSB) for each retrospective source and other settings

There are a number of files sourced by do_mse.R that allow easy definition of each scenario and calculation of desired metrics:

- base_input defines the all the basic biological and fishery characteristics used in the simulations
- change_input allows easy access to all the scenario combinations
- IBM_options sets the defaults for all the IBMs
- performance_metrics collects the summary statistics of interest (catch, SSB, and fishing mortality rate [F] mainly) that will be used to compare the different IBMs
- summarize_results is used after the runs are completed

- wham_mse_functions contains the code for all the IBMs
- wham_retro_functions contains the code needed to create the different sources of retrospective patterns

File structure for the GitHub repository

This section describes how to find files or place new files on the GitHub repository (https://github.com/cmlegault/IBMWG). If you clone the project, you should see these folders and files on your computer.

- **code**: how to run the simulations as well as all the functions needed to do so
- **demonstrations**: contains a number of sub-folders with code from working group members to demonstrate specific features of local data or simulation approaches
- **output**: where tibbles of runs will be stored (note: due to the large size of the files, they are not stored on GitHub; instead, they are stored in a Google Drive folder (approximately 300 GB total)
- **plots**: all the nice figures summarizing either individual runs or comparing scenarios
- **results**: files containing performance metrics from all runs separated by base, scaa, and noretro situations (note: the raw base files are cleaned to remove duplicate runs. Two database files for each set show the filename and path as well as the scenario and simulations run to allow back tracking of specific simulations)
- **settings**: files used to define different scenarios to be run; contains biological, fishery, and uncertainty parameters that create the base period of a simulation
- **tables_figs**: just what the name suggests, uses final cleaned data

APPENDIX 6 ADDITIONAL ANALYSES AND RESULTS

Please note: Table A6.6 below provides a lookup table for types of figures in this appendix, an example from the main text, and the set of analyses.

Table A6.1. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), time horizon (short- or long-term summary, "S" or "L"), and catch advice multiplier (1 or 0.75). Table rows correspond to the ordering in the heatmap in Figure A6.446.

IBM x Catch Multiplier	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-FM-S-0.75	1.28	1.03	0.82
ES-FSPR-S-0.75	1.26	1.03	0.82
ES-Fstable-L-0.75	1.18	0.88	0.88
AIM-L-0.75	1.19	0.88	0.87
Islope-S-1	1.27	1.09	1.00
PBS-S-1	1.00	0.99	1.08
DLM-S-1	1.01	0.97	1.08
Ensemble-S-0.75	0.99	0.93	1.00
ES-Frecent-S-1	1.11	1.04	1.05
Itarget-S-1	1.01	0.93	0.87
ES-FM-L-0.75	1.48	0.78	0.87
ES-FSPR-L-0.75	1.34	0.79	0.85
Itarget-L-0.75	1.46	0.89	1.08
ES-Fstable-S-0.75	1.78	1.19	0.75
AIM-S-0.75	1.74	1.19	0.75
Ensemble-S-1	1.81	1.19	0.83
Itarget-L-1	1.79	0.90	0.95
ES-Frecent-L-1	1.69	0.78	0.94
AIM-L-1	2.04	0.78	0.54
ES-Fstable-L-1	2.02	0.79	0.55
ES-FSPR-L-1	1.94	0.72	0.56
ES-FM-L-1	2.15	0.71	0.57
Ensemble-L-1	2.05	0.66	0.69
CC-FM-S-1	0.54	0.64	1.14
CC-FSPR-S-1	0.54	0.64	1.15
DLM-S-0.75	0.57	0.66	1.20
ES-Frecent-S-0.75	0.66	0.77	1.19
PBS-S-0.75	0.62	0.73	1.21
Itarget-S-0.75	0.64	0.74	1.00
Ensemble-L-0.75	0.84	0.78	1.22
Islope-S-0.75	0.75	0.84	1.15
PBS-L-1	0.76	0.76	1.42
CC-FM-S-0.75	0.37	0.48	1.24
CC-FSPR-S-0.75	0.37	0.48	1.24
DLM-L-1	0.41	0.51	1.75
PBS-L-0.75	0.06	0.21	2.99

IBM x Catch Multiplier	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
CC-FM-L-0.75	0.06	0.21	2.98
CC-FSPR-L-0.75	0.06	0.20	2.99
ES-Frecent-L-0.75	0.08	0.27	2.86
DLM-L-0.75	0.02	0.06	3.27
CC-FM-L-1	0.33	0.54	2.27
CC-FSPR-L-1	0.31	0.54	2.28
Islope-L-0.75	0.15	0.39	2.53
ES-FM-S-1	2.46	1.24	0.71
ES-FSPR-S-1	2.31	1.23	0.71
Islope-L-1	3.00	0.64	0.57
Skate-L-0.75	2.54	0.68	0.56
ES-Fstable-S-1	3.29	1.37	0.64
AIM-S-1	3.25	1.38	0.64
Skate-S-0.75	3.34	1.31	0.64
Skate-S-1	5.22	1.48	0.56
Skate-L-1	3.93	0.56	0.33

Table A6.2. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), catch advice multiplier (1 or 0.75), and fishing history (1=overfishing then $F=F_{MSY}$, 2=always overfishing). Table rows correspond to the ordering in the heatmap in Figure A6.447.

IBM x Fhistory x Catch Multiplier	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-FM-2-0.75	1.22	0.79	0.85
Itarget-2-1	1.25	0.72	0.83
PBS-2-1	1.33	0.79	0.88
ES-Fstable-2-0.75	1.51	0.90	0.71
AIM-2-0.75	1.51	0.90	0.71
Itarget-1-0.75	1.33	1.05	1.13
Islope-1-1	1.38	1.02	1.07
ES-Fstable-1-0.75	1.29	1.15	0.94
AIM-1-0.75	1.29	1.16	0.93
ES-FSPR-2-0.75	1.11	0.80	0.85
Ensemble-2-0.75	1.02	0.79	0.94
DLM-2-1	1.11	0.71	1.01
Itarget-2-0.75	0.92	0.65	0.93
ES-Frecent-1-1	0.93	1.00	1.26
Ensemble-1-0.75	0.81	0.94	1.21
ES-FM-2-1	1.94	0.84	0.67

IBM x Fhistory x Catch Multiplier	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
Ensemble-2-1	1.90	0.85	0.67
ES-Frecent-2-1	1.78	0.83	0.73
ES-FSPR-2-1	1.76	0.83	0.67
Itarget-1-1	1.94	1.14	0.99
Ensemble-1-1	1.99	1.08	0.89
Islope-2-1	2.24	0.81	0.63
AIM-2-1	2.67	0.91	0.50
ES-Fstable-2-1	2.65	0.91	0.51
Skate-2-0.75	2.72	0.88	0.54
ES-FM-1-0.75	2.80	1.05	0.80
ES-Fstable-1-1	2.51	1.25	0.73
AIM-1-1	2.46	1.25	0.72
ES-FSPR-1-0.75	2.41	1.05	0.81
Skate-1-0.75	3.38	1.17	0.70
CC-FM-2-1	0.34	0.43	1.68
CC-FSPR-2-1	0.33	0.43	1.67
ES-Frecent-2-0.75	0.39	0.42	1.67
PBS-2-0.75	0.33	0.35	1.78
CC-FSPR-1-0.75	0.28	0.57	1.97
CC-FM-1-0.75	0.28	0.56	1.96
Islope-1-0.75	0.36	0.60	1.93
ES-Frecent-1-0.75	0.25	0.46	2.17
PBS-1-0.75	0.21	0.39	2.24
DLM-1-0.75	0.17	0.29	2.37
CC-FM-2-0.75	0.15	0.24	2.11
CC-FSPR-2-0.75	0.16	0.24	2.09
DLM-2-0.75	0.27	0.26	1.93
CC-FM-1-1	0.69	0.84	1.49
CC-FSPR-1-1	0.68	0.82	1.51
PBS-1-1	0.72	0.94	1.44
DLM-1-1	0.57	0.81	1.57
Islope-2-0.75	0.58	0.50	1.38
ES-FM-1-1	5.11	1.19	0.67
Skate-1-1	5.09	1.29	0.55
ES-FSPR-1-1	4.14	1.18	0.68
Skate-2-1	4.46	0.90	0.40

Table A6.3. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), fishing history (1=overfishing then F=F_{MSY}, 2=always overfishing), and time horizon (short- or long-term summary, "S" or "L"). Table rows correspond to the ordering in the heatmap in Figure A6.448.

IBM x Fhistory x Time	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-FM-2-S	1.41	0.80	0.73
ES-FSPR-2-S	1.40	0.80	0.72
Ensemble-2-S	1.36	0.84	0.71
ES-Frecent-2-S	1.34	0.84	0.72
ES-Fstable-1-L	1.45	0.87	0.78
AIM-1-L	1.46	0.86	0.77
ES-FSPR-2-L	1.46	0.83	0.81
DLM-2-S	1.25	0.79	0.73
PBS-2-S	1.21	0.79	0.75
Ensemble-1-L	1.46	0.66	0.90
ES-Fstable-2-L	1.85	0.81	0.65
AIM-2-L	1.84	0.80	0.64
ES-FM-2-L	1.63	0.83	0.82
Islope-2-S	1.56	0.87	0.68
Itarget-2-L	1.61	0.85	0.93
Ensemble-2-L	1.53	0.79	0.97
Itarget-1-L	1.64	0.94	1.08
Itarget-1-S	1.58	1.38	1.05
Ensemble-1-S	1.21	1.38	1.13
PBS-1-S	0.58	0.90	1.47
DLM-1-S	0.56	0.88	1.45
ES-Frecent-1-S	0.62	0.95	1.44
CC-FM-1-S	0.51	0.85	1.41
CC-FSPR-1-S	0.51	0.84	1.42
Islope-1-S	0.70	1.02	1.40
Islope-1-L	0.57	0.48	1.69
Islope-2-L	0.53	0.39	1.62
CC-FM-2-S	0.41	0.39	0.96
CC-FSPR-2-S	0.40	0.39	0.96
Itarget-2-S	0.60	0.49	0.84
CC-FSPR-1-L	0.29	0.51	2.22
CC-FM-1-L	0.30	0.50	2.21
ES-Frecent-1-L	0.27	0.37	2.19
ES-Frecent-2-L	0.30	0.33	2.01
PBS-1-L	0.18	0.32	2.41
PBS-2-L	0.19	0.28	2.28

IBM x Fhistory x Time	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
DLM-2-L	0.12	0.16	2.35
DLM-1-L	0.10	0.17	2.55
CC-FM-2-L	0.06	0.22	3.05
CC-FSPR-2-L	0.06	0.21	3.05
Skate-1-L	3.38	0.61	0.41
ES-FM-1-L	3.20	0.63	0.47
Skate-2-L	2.95	0.63	0.42
AIM-2-S	2.40	1.00	0.57
ES-Fstable-2-S	2.39	0.99	0.57
ES-FSPR-1-L	2.41	0.63	0.45
ES-Fstable-1-S	3.20	1.64	0.87
AIM-1-S	3.17	1.65	0.86
ES-FSPR-1-S	3.97	1.56	0.82
Skate-2-S	4.03	1.06	0.49
ES-FM-1-S	4.86	1.56	0.82
Skate-1-S	6.46	1.75	0.70

Table A6.4. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), source of retrospective pattern (C=catch, M=natural mortality), and fishing history (1=overfishing then $F=F_{MSY}$, 2=always overfishing). Table rows correspond to the ordering in the heatmap in Figure A6.449.

IBM x Retro x Fhistory	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-FM-M-2	0.91	1.56	0.54
ES-FSPR-M-2	0.91	1.56	0.54
Ensemble-M-2	0.95	1.51	0.65
ES-Frecent-M-2	0.75	1.50	0.63
PBS-Catch-1	0.72	1.48	0.62
PBS-M-2	0.71	1.63	0.54
DLM-M-2	0.67	1.70	0.49
Itarget-M-2	0.69	1.60	0.43
DLM-Catch-1	0.59	1.52	0.51
Ensemble-M-1	0.98	1.80	0.53
Islope-M-1	0.80	1.76	0.52
ES-Frecent-M-1	0.76	1.98	0.40
Itarget-M-1	0.84	1.88	0.41
CC-FM-Catch-2	0.41	1.12	0.51
CC-FSPR-Catch-2	0.40	1.10	0.50
PBS-M-1	0.73	2.11	0.35
DLM-M-1	0.65	2.11	0.33

IBM x Retro x Fhistory	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
CC-FM-M-1	0.57	2.31	0.23
CC-FSPR-M-1	0.56	2.29	0.22
CC-FM-M-2	0.26	2.46	0.11
CC-FSPR-M-2	0.27	2.44	0.12
AIM-M-2	1.09	1.21	1.01
ES-Fstable-M-2	1.08	1.22	0.99
CC-FM-Catch-1	0.89	1.16	1.04
CC-FSPR-Catch-1	0.88	1.19	0.94
Islope-M-2	0.79	1.35	0.81
ES-Frecent-Catch-1	0.72	1.37	0.71
Islope-Catch-1	0.75	1.20	0.86
ES-Fstable-M-1	1.21	1.53	0.78
AIM-M-1	1.22	1.53	0.77
ES-FM-M-1	1.16	1.60	0.70
ES-FSPR-M-1	1.16	1.61	0.70
Skate-M-1	1.30	1.37	1.09
Skate-M-2	1.12	1.06	1.44
PBS-Catch-2	0.47	0.67	1.22
DLM-Catch-2	0.43	0.69	1.16
ES-Frecent-Catch-2	0.48	0.59	1.45
Islope-Catch-2	0.49	0.49	1.93
AIM-Catch-2	0.70	0.25	5.43
ES-Fstable-Catch-2	0.71	0.25	5.36
Itarget-Catch-1	1.35	0.56	5.15
ES-FSPR-Catch-2	0.71	0.28	4.73
ES-FM-Catch-2	0.71	0.28	4.61
AIM-Catch-1	1.17	0.37	6.27
ES-Fstable-Catch-1	1.18	0.38	6.17
Ensemble-Catch-1	1.06	0.48	5.94
Ensemble-Catch-2	0.66	0.37	3.64
Itarget-Catch-2	0.68	0.46	2.84
ES-FM-Catch-1	0.93	0.21	8.73
ES-FSPR-Catch-1	0.93	0.21	8.73
Skate-Catch-1	0.99	0.23	8.73
Skate-Catch-2	0.58	0.15	8.73

Table A6.5. Median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), source of retrospective pattern (C=catch, M=natural mortality), and time horizon (short- or long-term summary, "S" or "L"). Table rows correspond to the ordering in the heatmap in Figure A6.450.

IBM x Retro x Time	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
ES-FM-M-S	0.97	1.50	0.57
ES-FSPR-M-S	0.96	1.50	0.57
Ensemble-M-S	0.97	1.47	0.61
ES-Frecent-M-S	0.96	1.48	0.59
PBS-M-S	0.91	1.50	0.55
DLM-M-S	0.89	1.49	0.54
ES-FM-M-L	1.09	1.65	0.67
ES-FSPR-M-L	1.09	1.66	0.66
Itarget-M-L	0.96	1.79	0.76
Ensemble-M-L	0.96	1.88	0.54
CC-FM-Catch-L	0.60	2.01	0.44
CC-FSPR-Catch-L	0.59	2.04	0.41
Islope-M-L	0.51	1.88	0.46
ES-Frecent-Catch-L	0.30	1.90	0.34
CC-FM-M-S	0.53	1.68	0.26
CC-FSPR-M-S	0.53	1.68	0.26
Itarget-M-S	0.53	1.67	0.26
Islope-Catch-L	0.35	1.37	0.71
ES-Fstable-M-L	1.08	1.44	0.84
AIM-M-L	1.09	1.43	0.83
Islope-M-S	1.03	1.43	0.69
AIM-M-S	1.23	1.32	0.93
ES-Fstable-M-S	1.22	1.32	0.92
Skate-M-S	1.38	1.21	1.26
Skate-M-L	1.07	1.21	1.31
PBS-Catch-S	0.79	0.79	1.10
DLM-Catch-S	0.78	0.77	1.14
ES-Frecent-Catch-S	0.83	0.77	1.20
Islope-Catch-S	0.87	0.75	1.33
CC-FM-Catch-S	0.59	0.75	0.76
CC-FSPR-Catch-S	0.58	0.75	0.76
ES-Frecent-M-L	0.42	2.29	0.24
PBS-Catch-L	0.27	2.15	0.22
DLM-Catch-L	0.11	2.27	0.10
PBS-M-L	0.34	2.54	0.16
DLM-M-L	0.25	2.61	0.12

IBM x Retro x Time	F/F_{MSY}	Catch/MSY	SSB/SSB_{MSY}
CC-FM-M-L	0.21	3.24	0.05
CC-FSPR-M-L	0.20	3.24	0.05
ES-FSPR-Catch-S	1.38	0.37	8.73
ES-FM-Catch-S	1.38	0.36	8.73
AIM-Catch-S	1.37	0.35	8.73
ES-Fstable-Catch-S	1.36	0.35	8.73
Skate-Catch-S	1.41	0.30	8.73
ES-FM-Catch-L	0.25	0.09	7.23
ES-FSPR-Catch-L	0.28	0.10	7.03
Skate-Catch-L	0.10	0.03	8.27
Ensemble-Catch-S	1.23	0.52	3.85
Itarget-Catch-L	0.84	0.55	3.90
AIM-Catch-L	0.55	0.26	4.17
ES-Fstable-Catch-L	0.56	0.27	4.13
Ensemble-Catch-L	0.29	0.22	5.07
Itarget-Catch-S	1.30	0.47	6.28

Table A6.6. Lookup table for figures in Appendix 6 by analysis set (Base, No retro, or statistical catch-at-age [SCAA]) by figure type with an example from the main report for context.

Figure Type	Report Example	Base	No retro	SCAA
Number of sims	N/A	1	1	1
Scores	4.1	2-14	188-200	269-281
Boxplots	4.9	15-32	201-218	282-299
Trade off (means)	4.15	33-38	219-224	300-305
1,000 points	4.11	39-96	225-260	306-341
Bagplots	4.12	97-125	N/A	342-359
Scenario panel sorted	4.3	126-131	261-266	360-365
Status	4.19	132-133	267-268	366-367
Confetti	4.13	134-187	N/A	N/A
ANOVA plots	N/A	368-445	N/A	N/A
Heatmaps	5.1	446-450	N/A	N/A

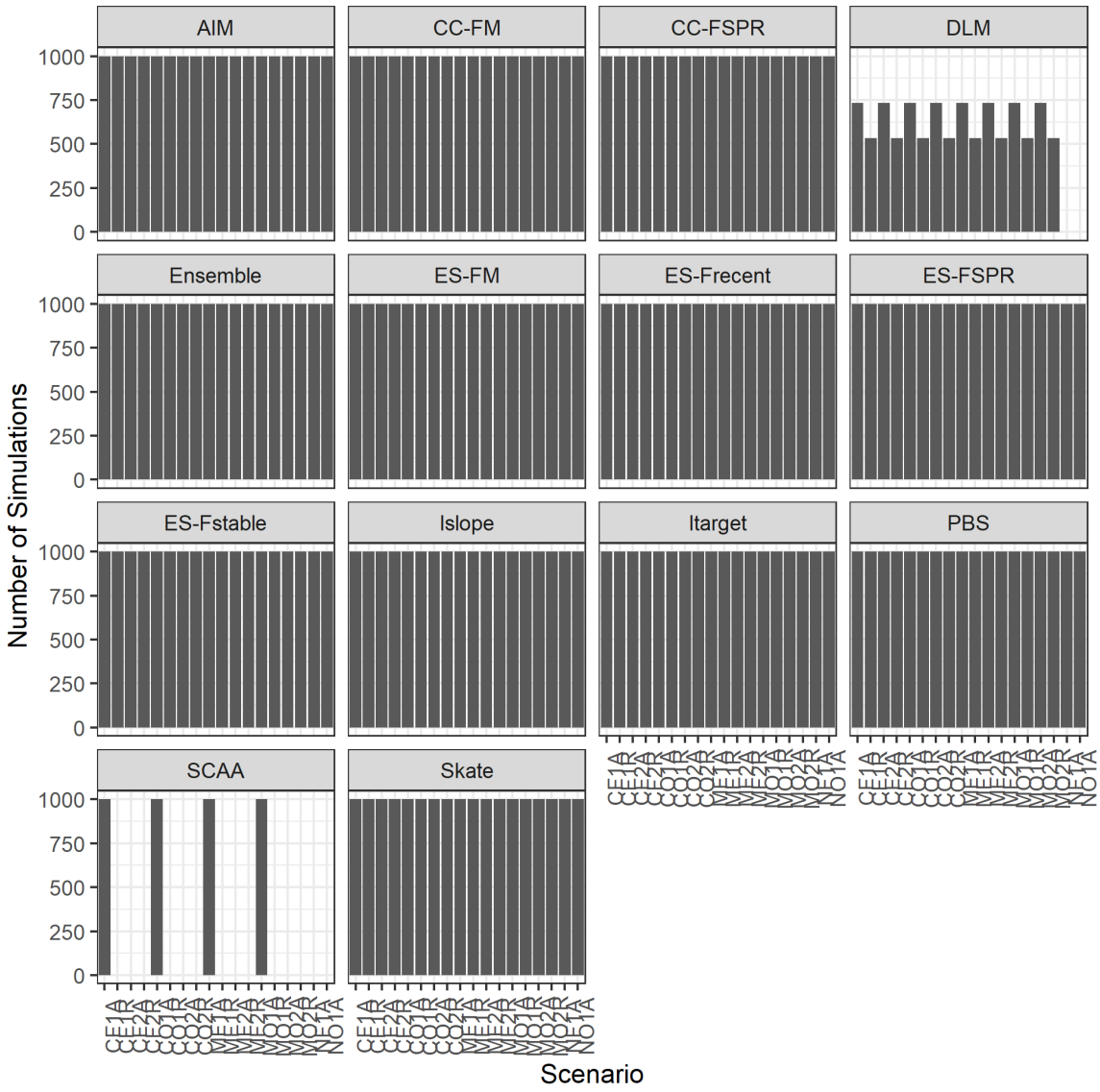


Figure A6.1. Number of successful simulations by scenario.

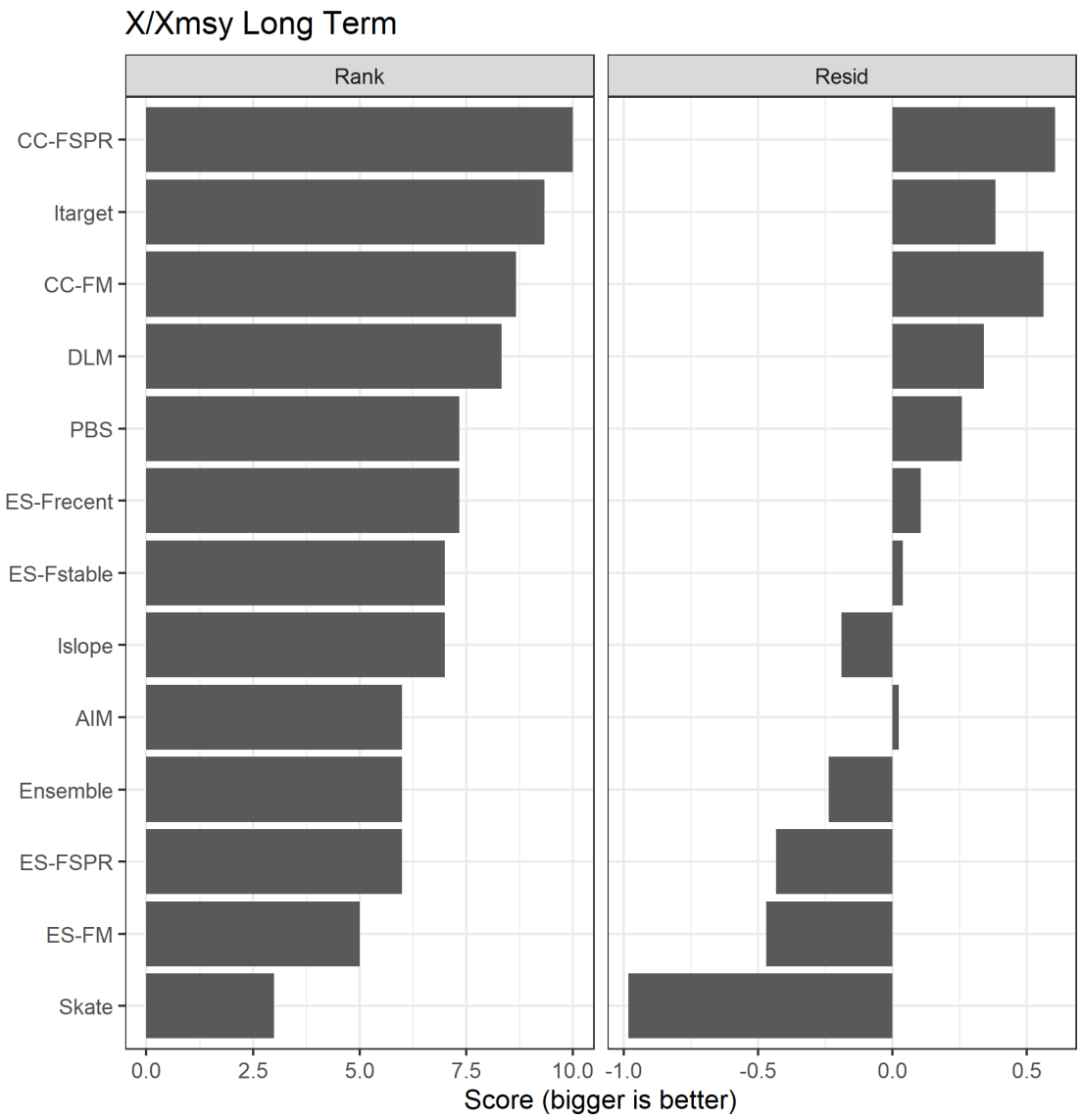


Figure A6.2. Two sets of scores, Rank and Resid, for the base analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in the long term.

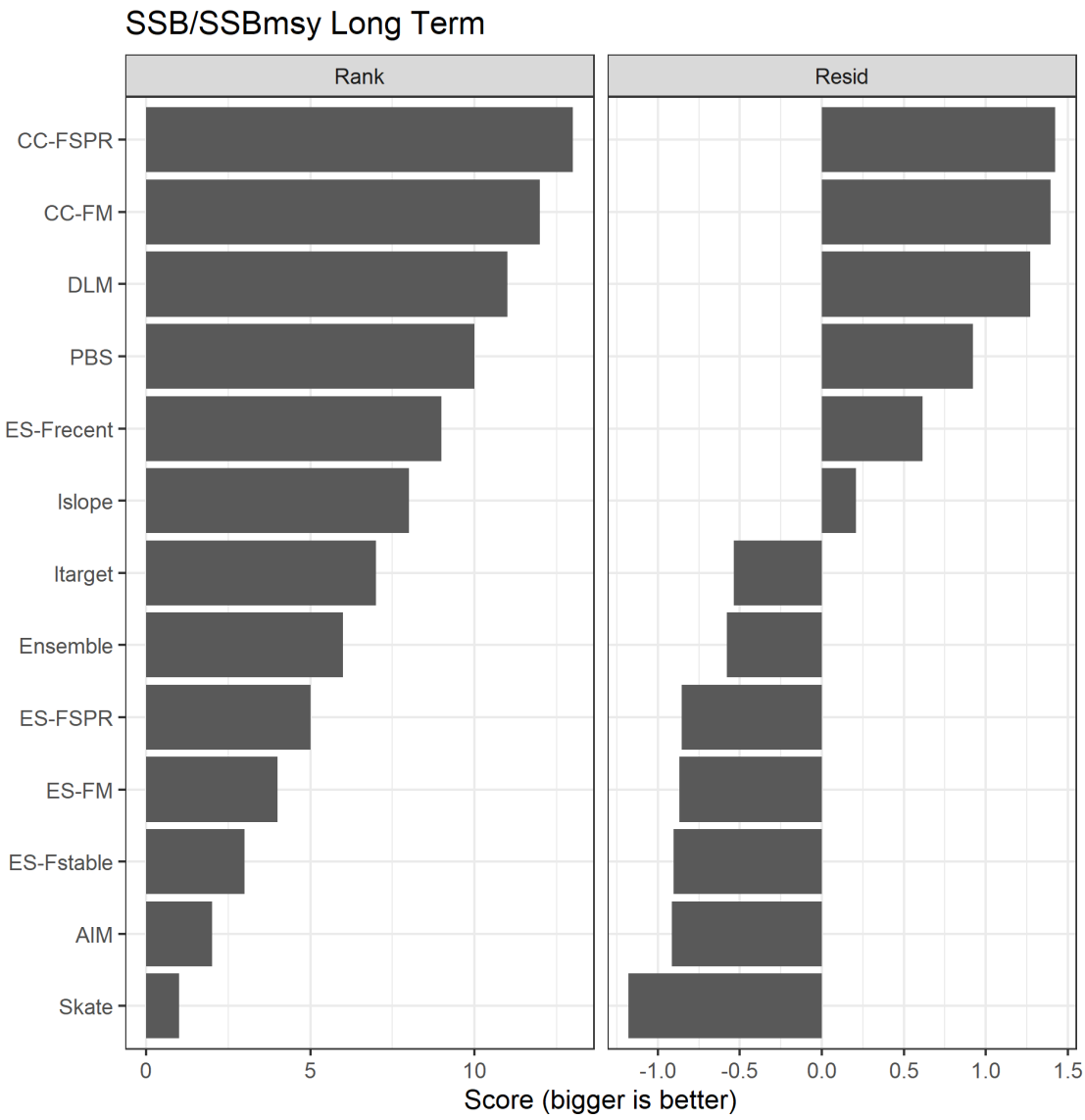


Figure A6.3. Two sets of scores, Rank and Resid, for the base analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in the long term.

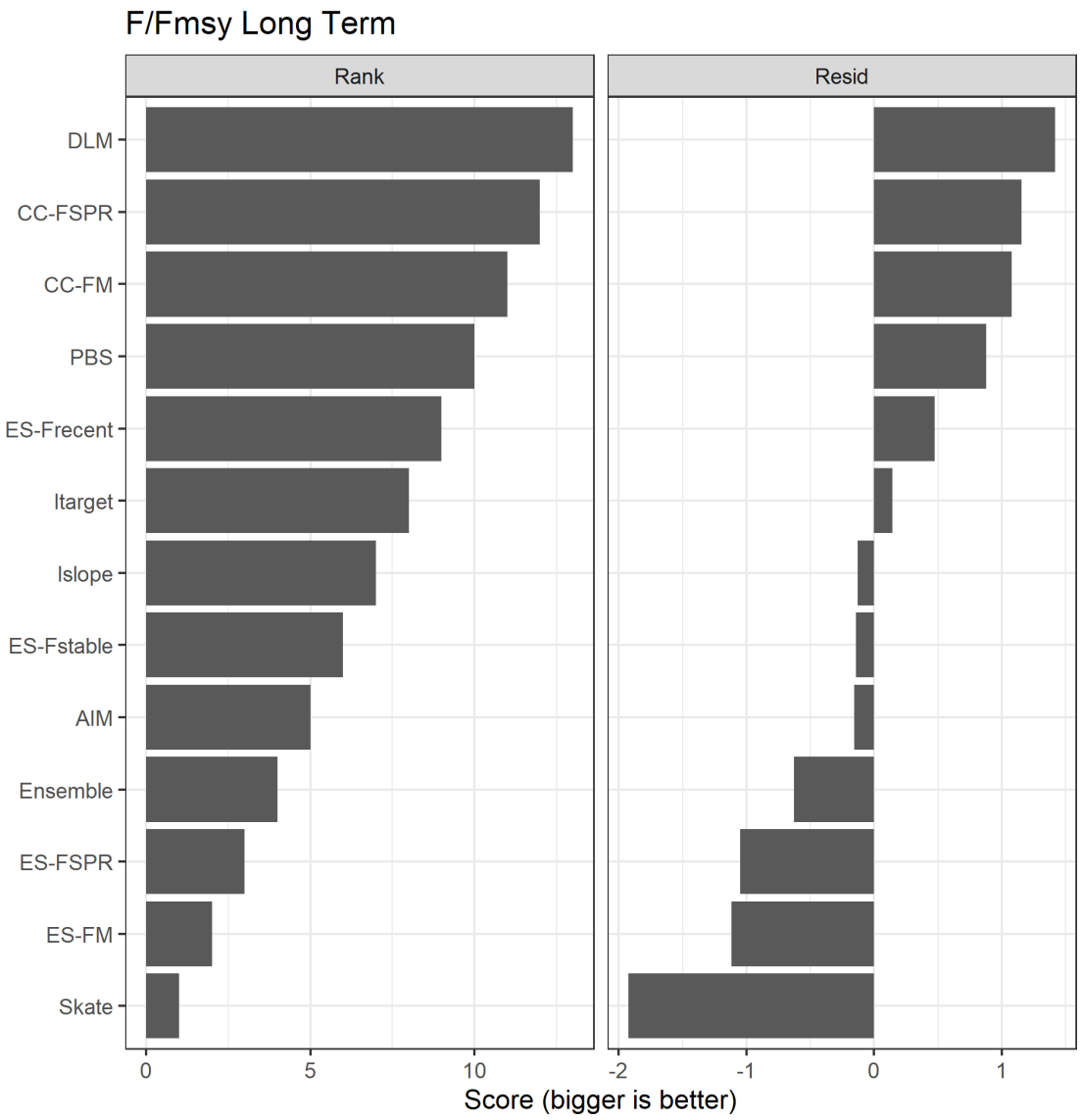


Figure A6.4. Two sets of scores, Rank and Resid, for the base analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in the long term.

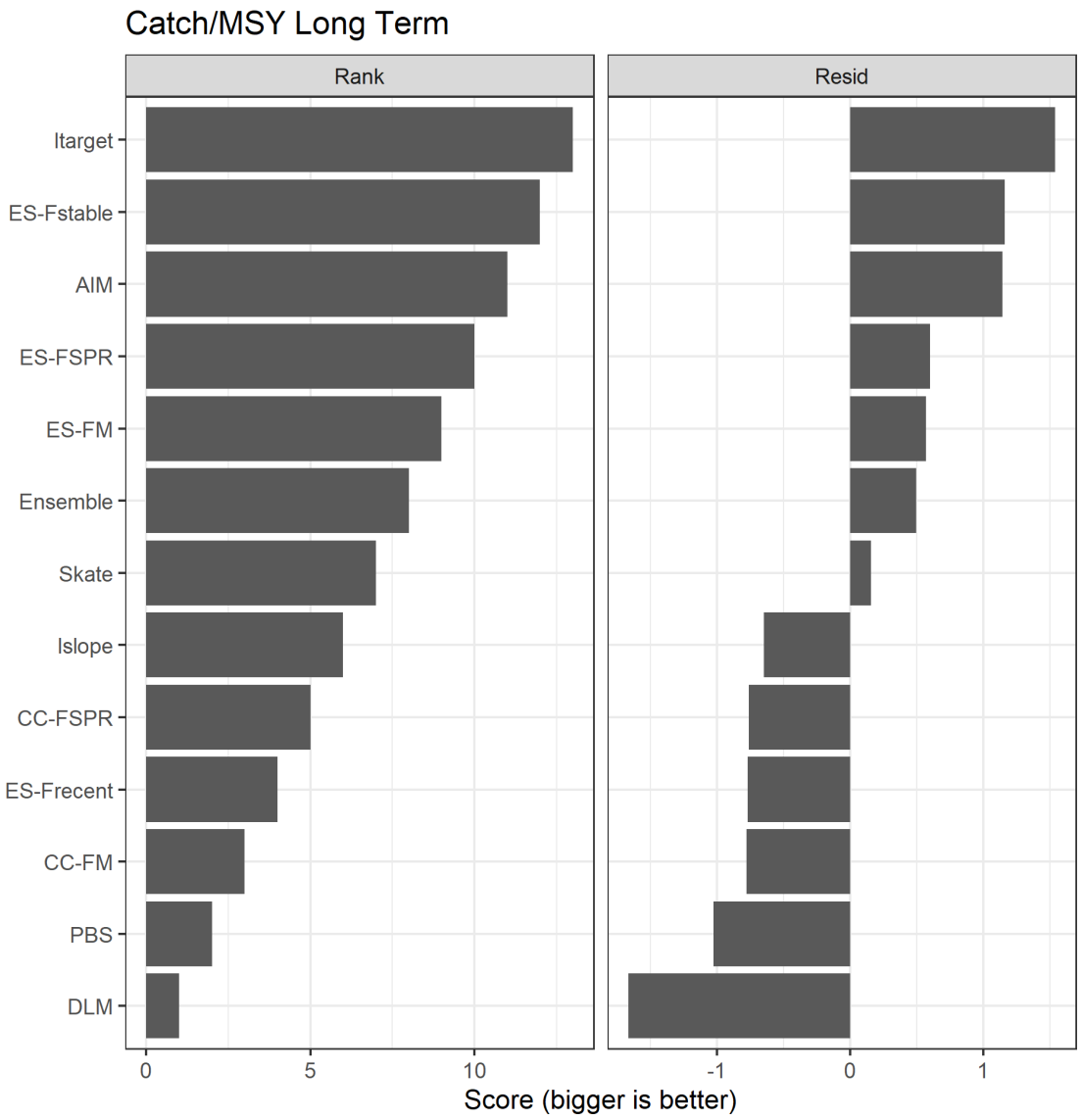


Figure A6.5. Two sets of scores, Rank and Resid, for the base analyses for the catch relative to maximum sustainable yield (MSY) in the long term.

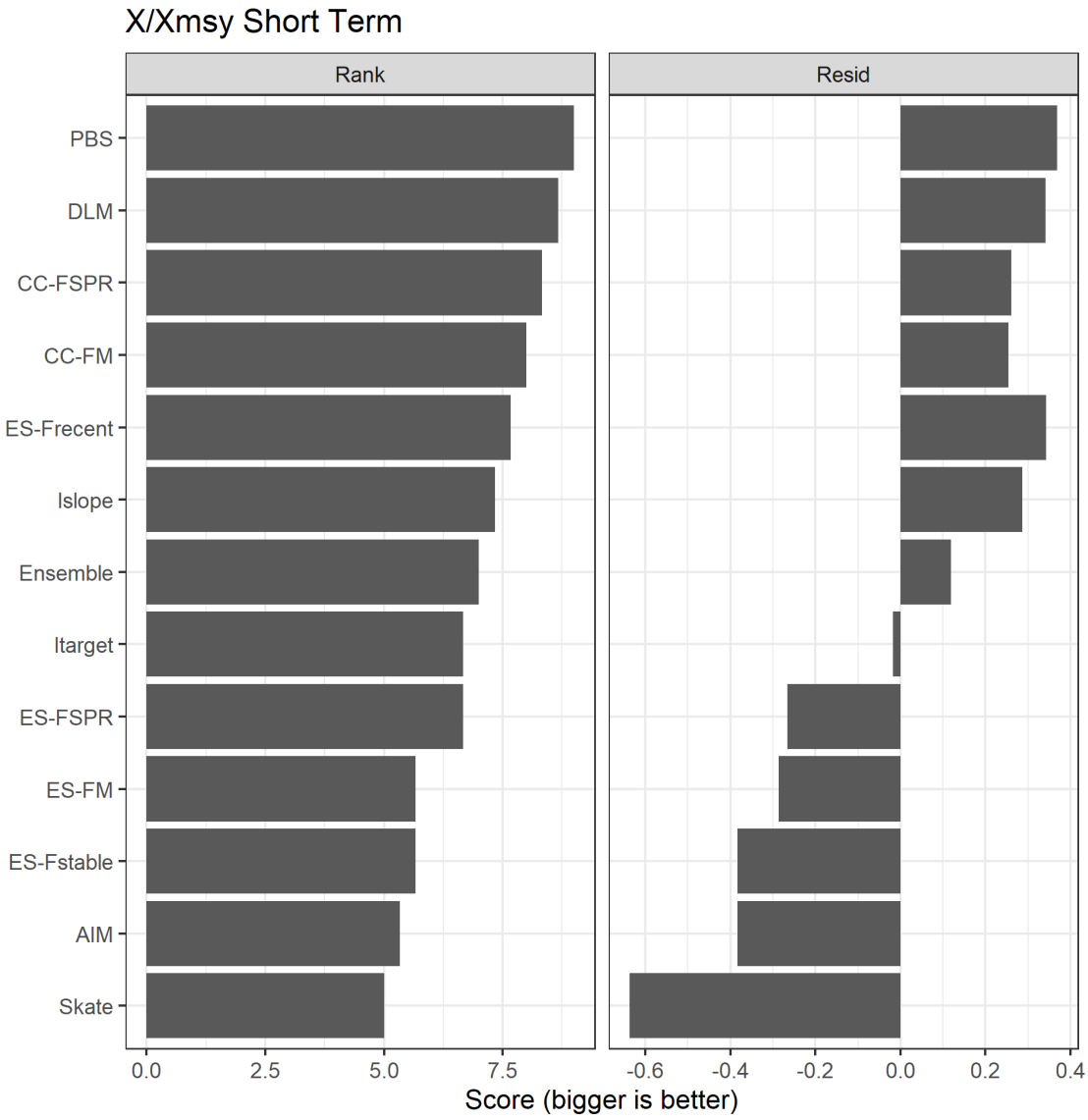


Figure A6.6. Two sets of scores, Rank and Resid, for the base analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in the short term.

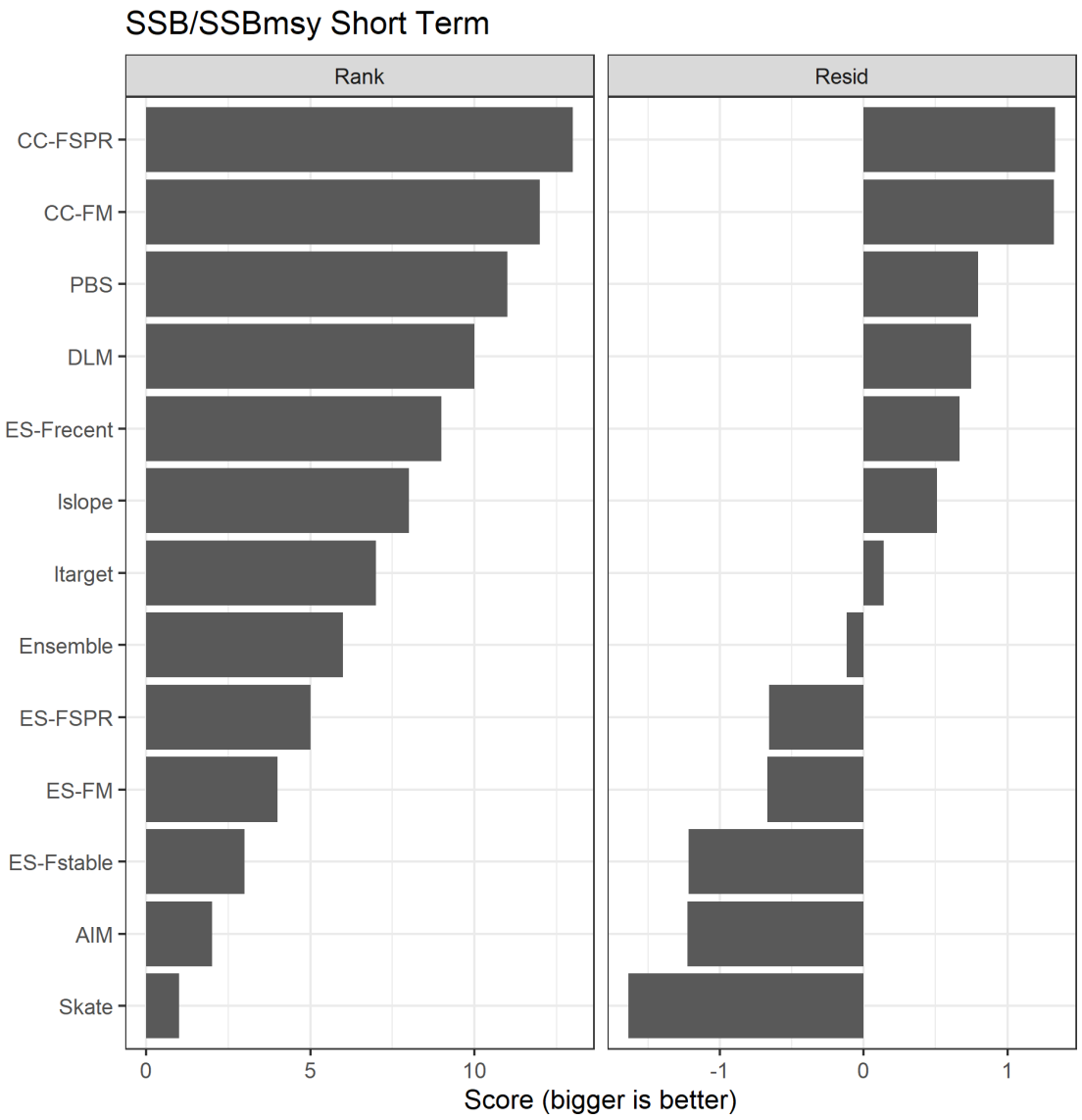


Figure A6.7. Two sets of scores, Rank and Resid, for the base analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in the short term.

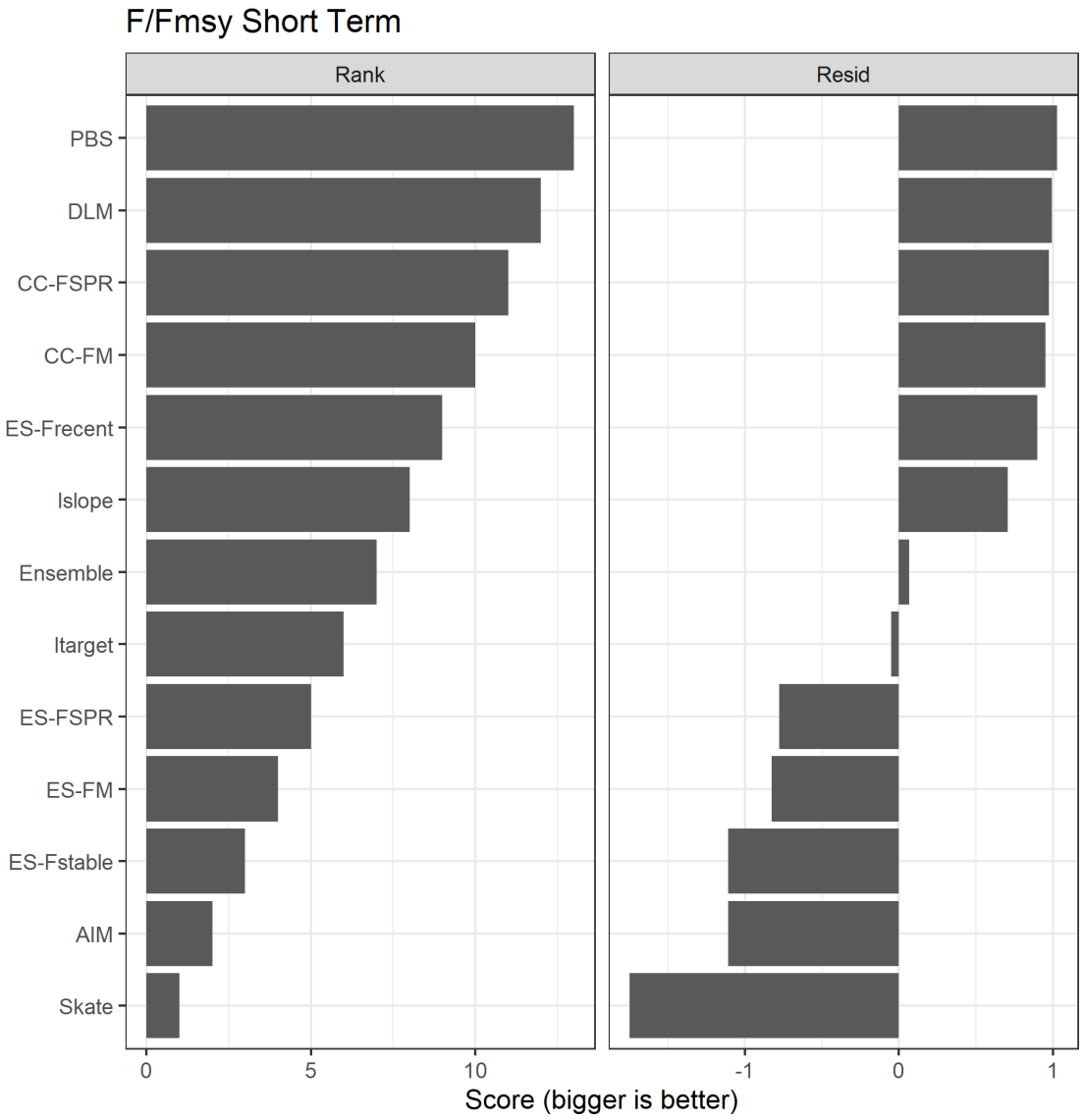


Figure A6.8. Two sets of scores, Rank and Resid, for the base analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in the short term.

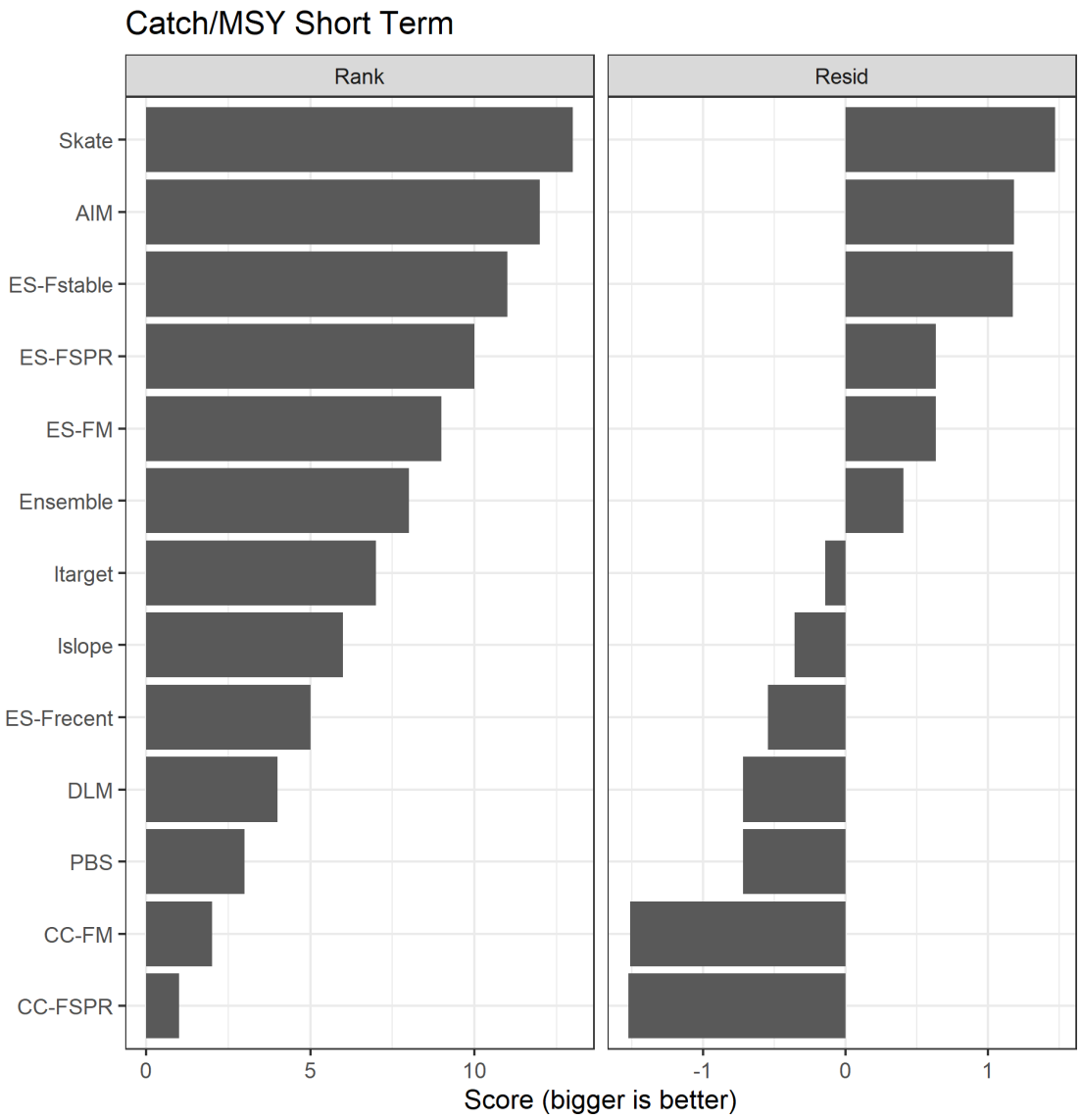


Figure A6.9. Two sets of scores, Rank and Resid, for the base analyses for the catch relative to maximum sustainable yield (MSY) in the short term.

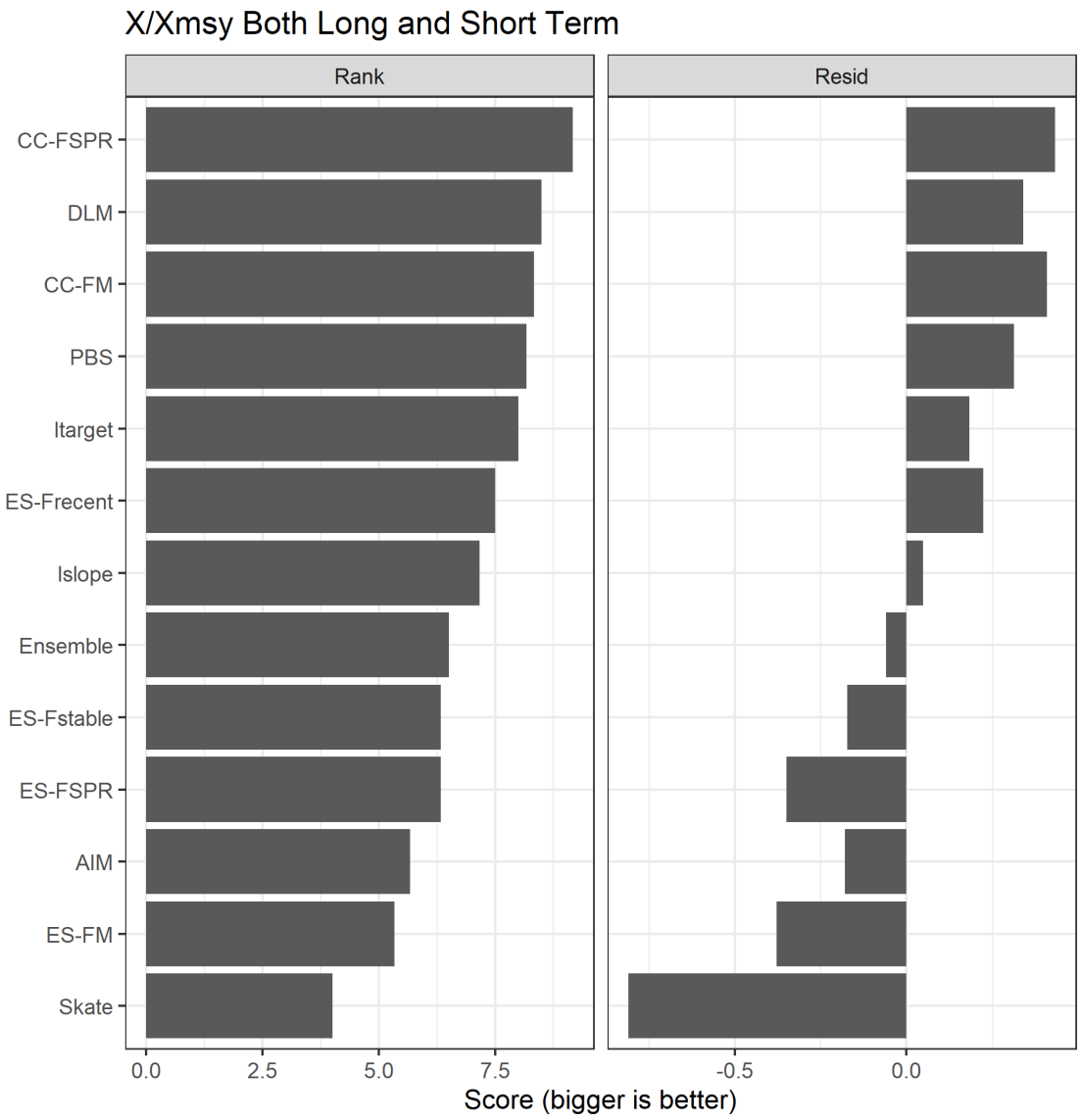


Figure A6.10. Two sets of scores, Rank and Resid, for the base analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in both the long and short term.

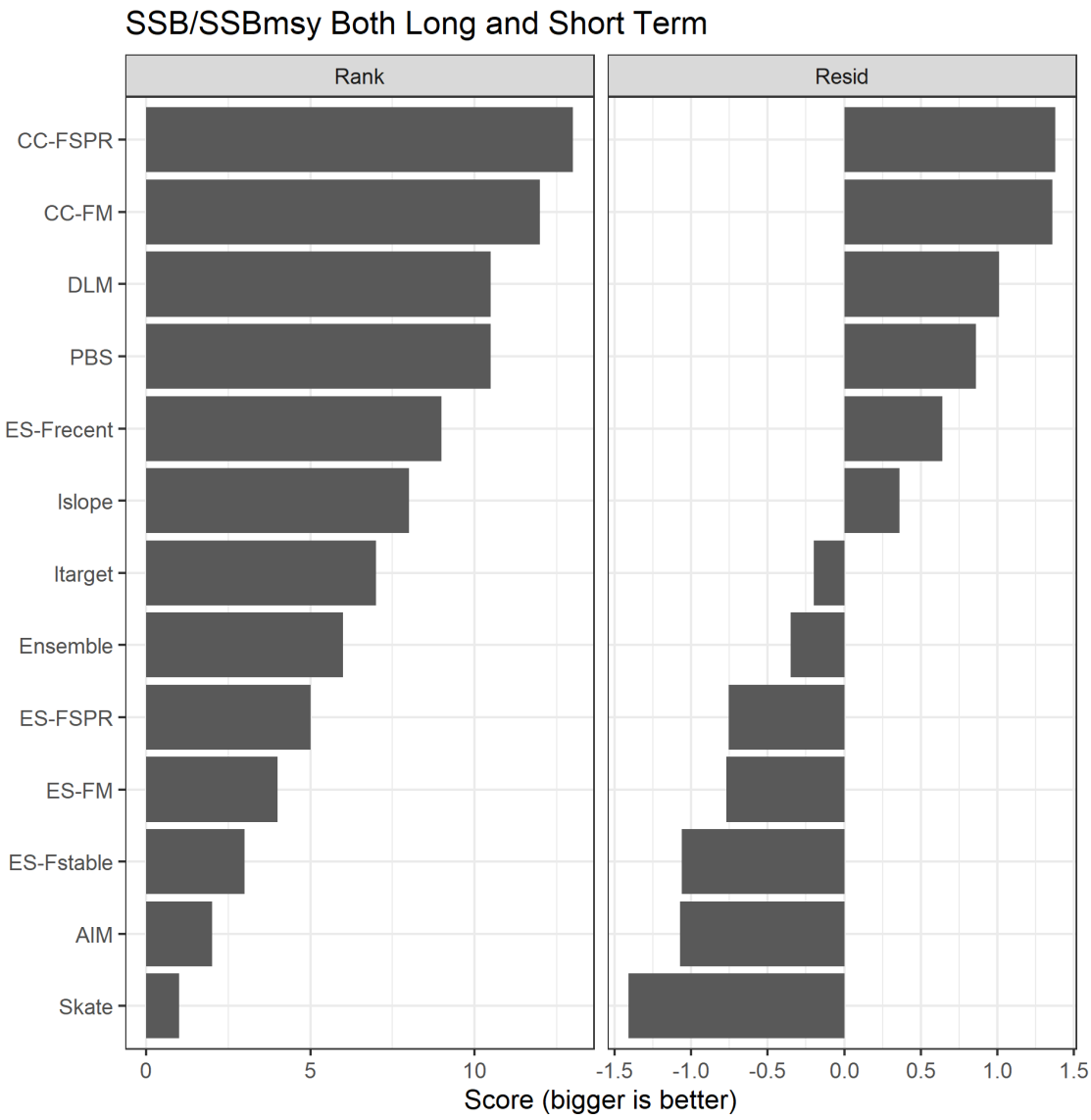


Figure A6.11. Two sets of scores, Rank and Resid, for the base analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in both the long and short term.

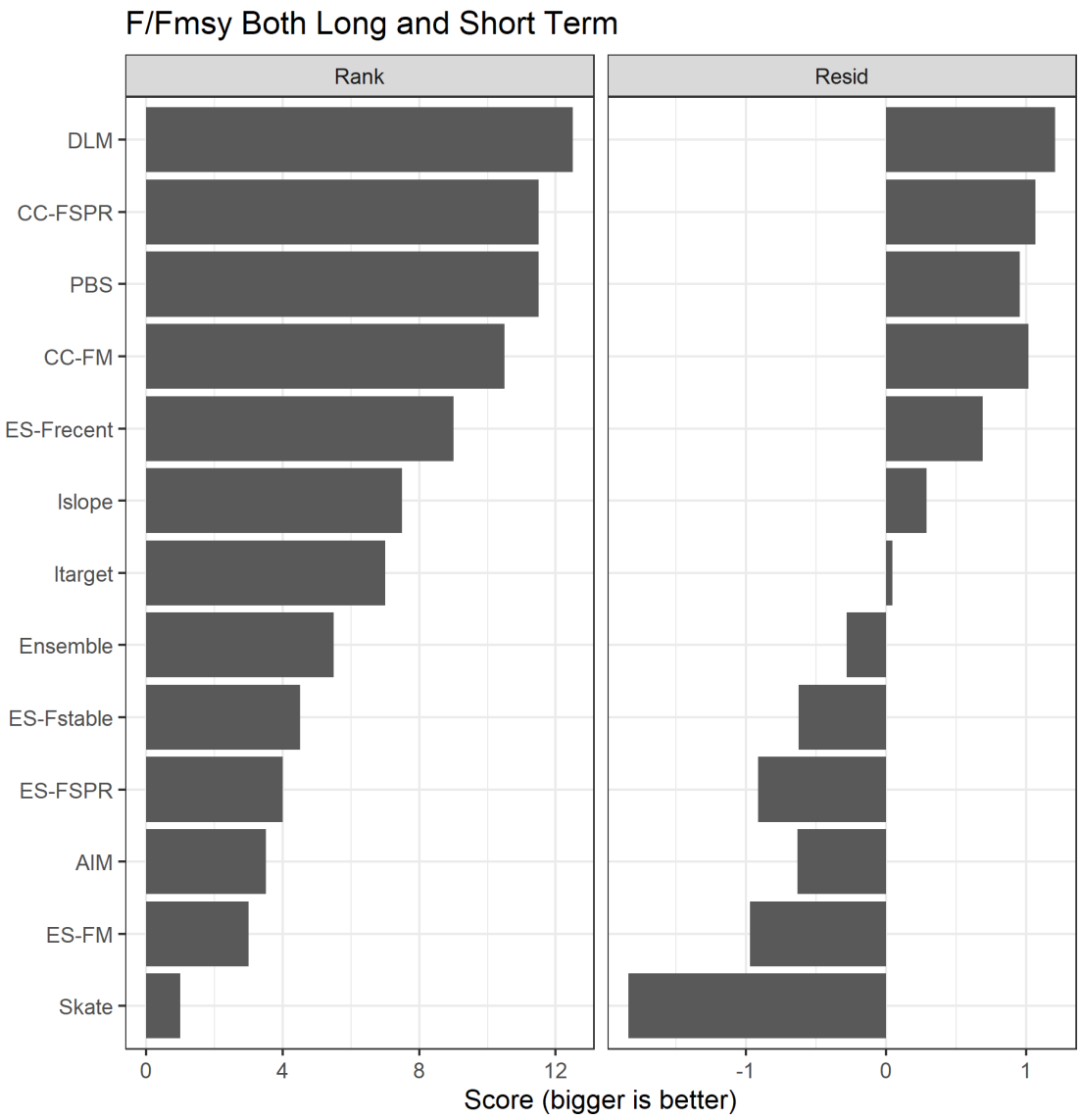


Figure A6.12. Two sets of scores, Rank and Resid, for the base analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in both the long and short term.

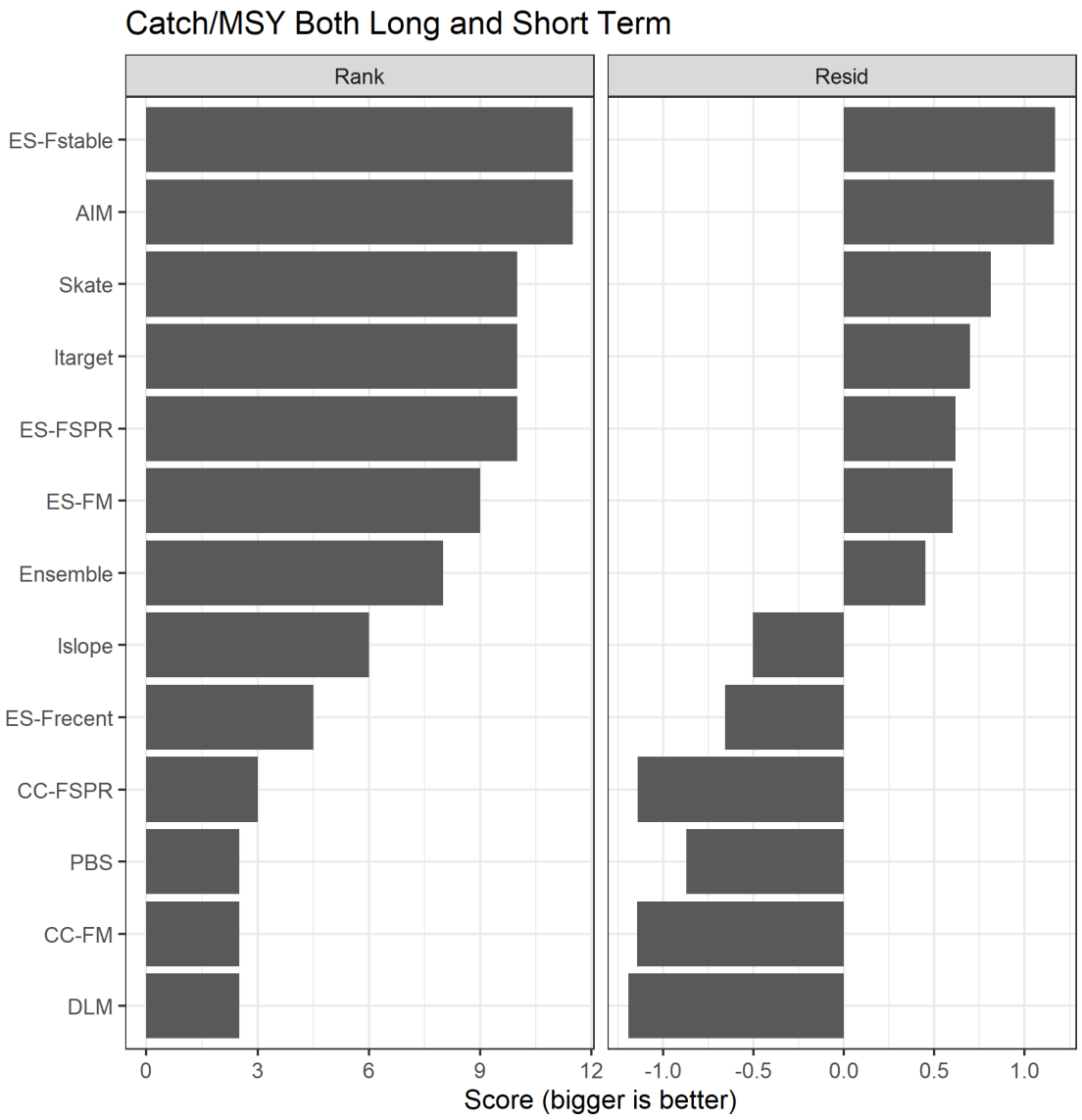


Figure A6.13. Two sets of scores, Rank and Resid, for the base analyses for the catch relative to maximum sustainable yield (MSY) in both the long and short term.

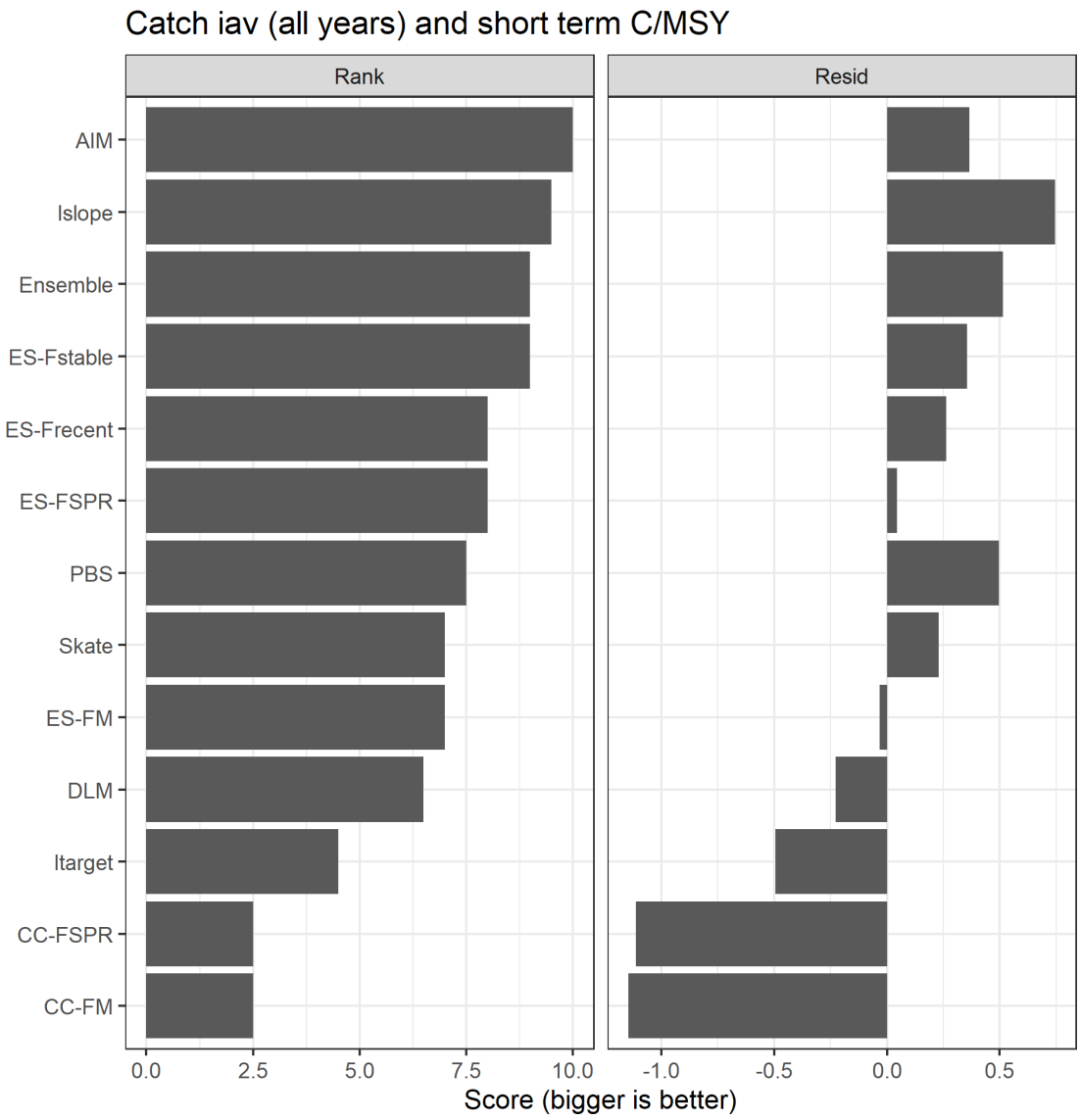


Figure A6.14. Two sets of scores, Rank and Resid, for the base analyses for the 2 metrics of interannual variability in catch over the entire feedback period and the short-term mean catch/maximum sustainable yield (MSY).

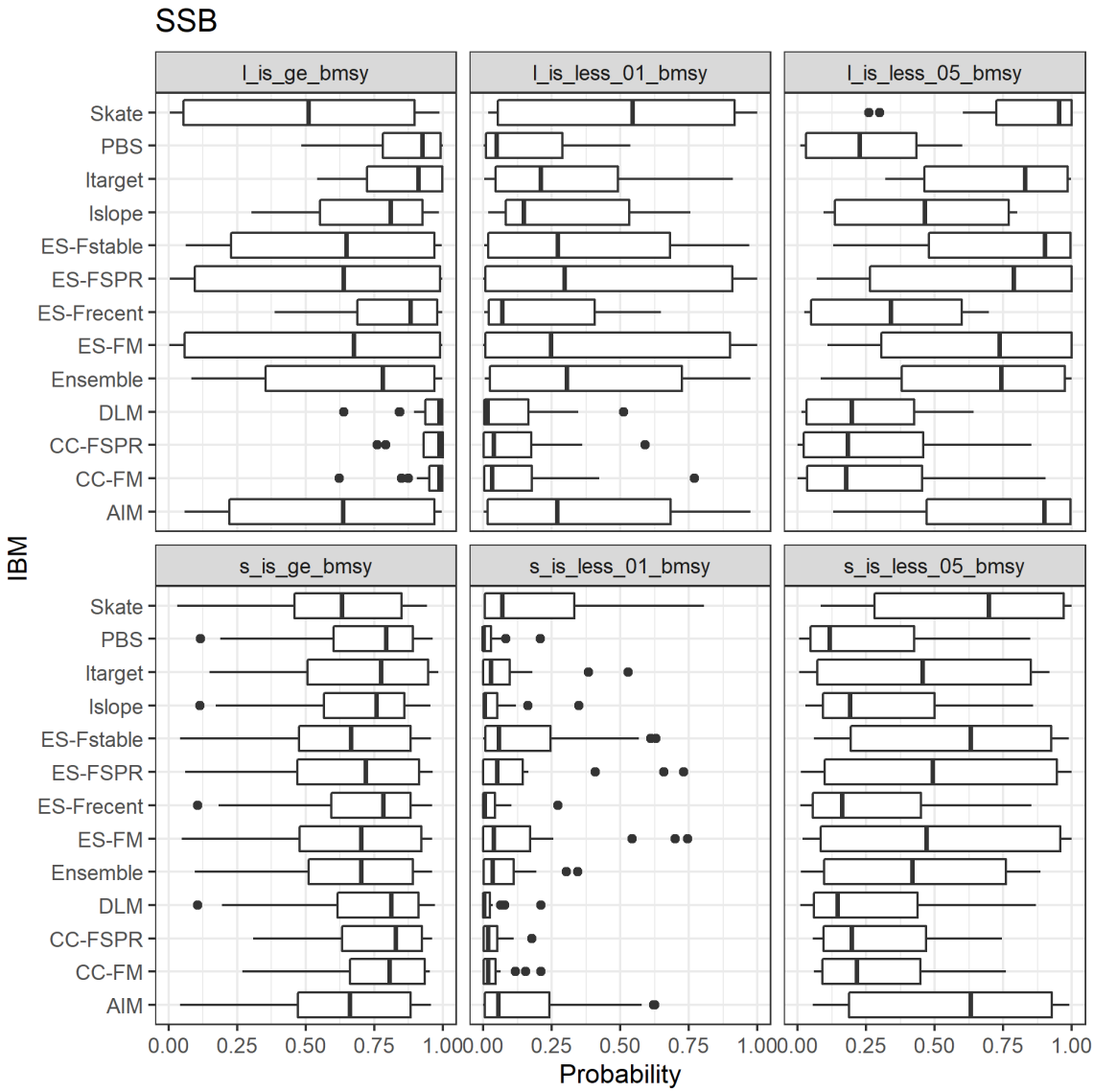


Figure A6.15. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the base analyses by index-based method (IBM).

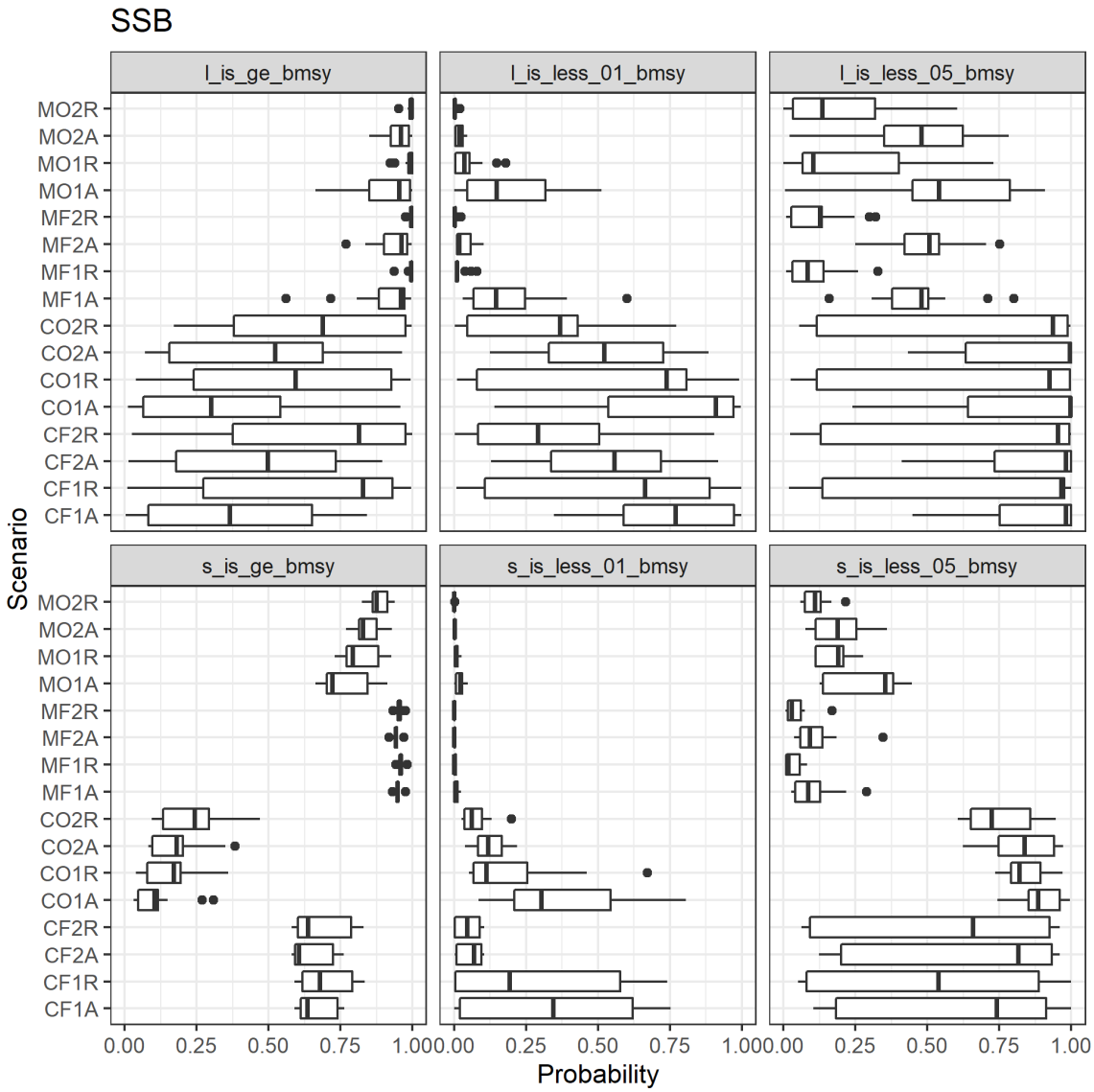


Figure A6.16. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the base analyses by scenario.

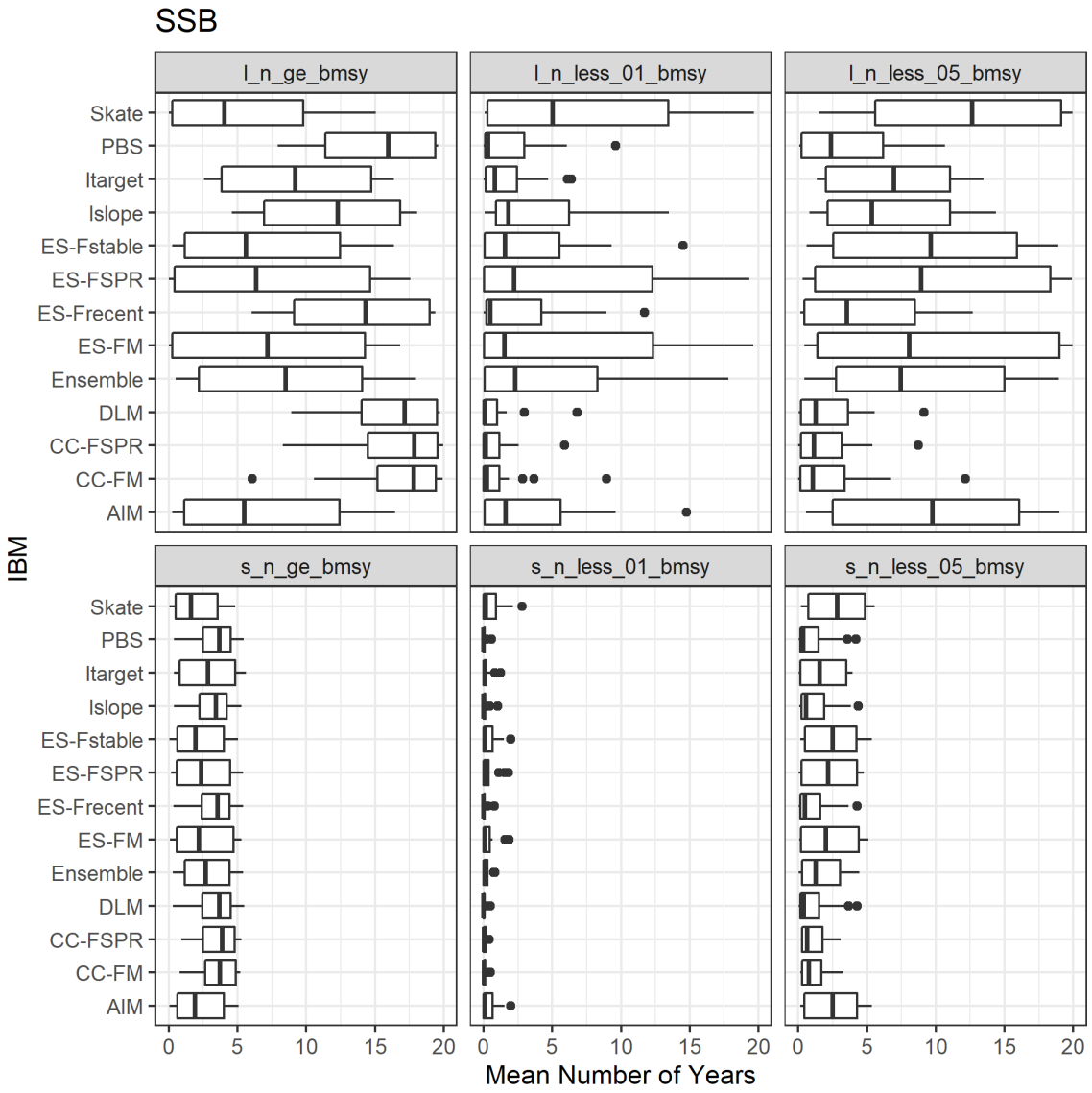


Figure A6.17. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the base analyses by index-based method (IBM).

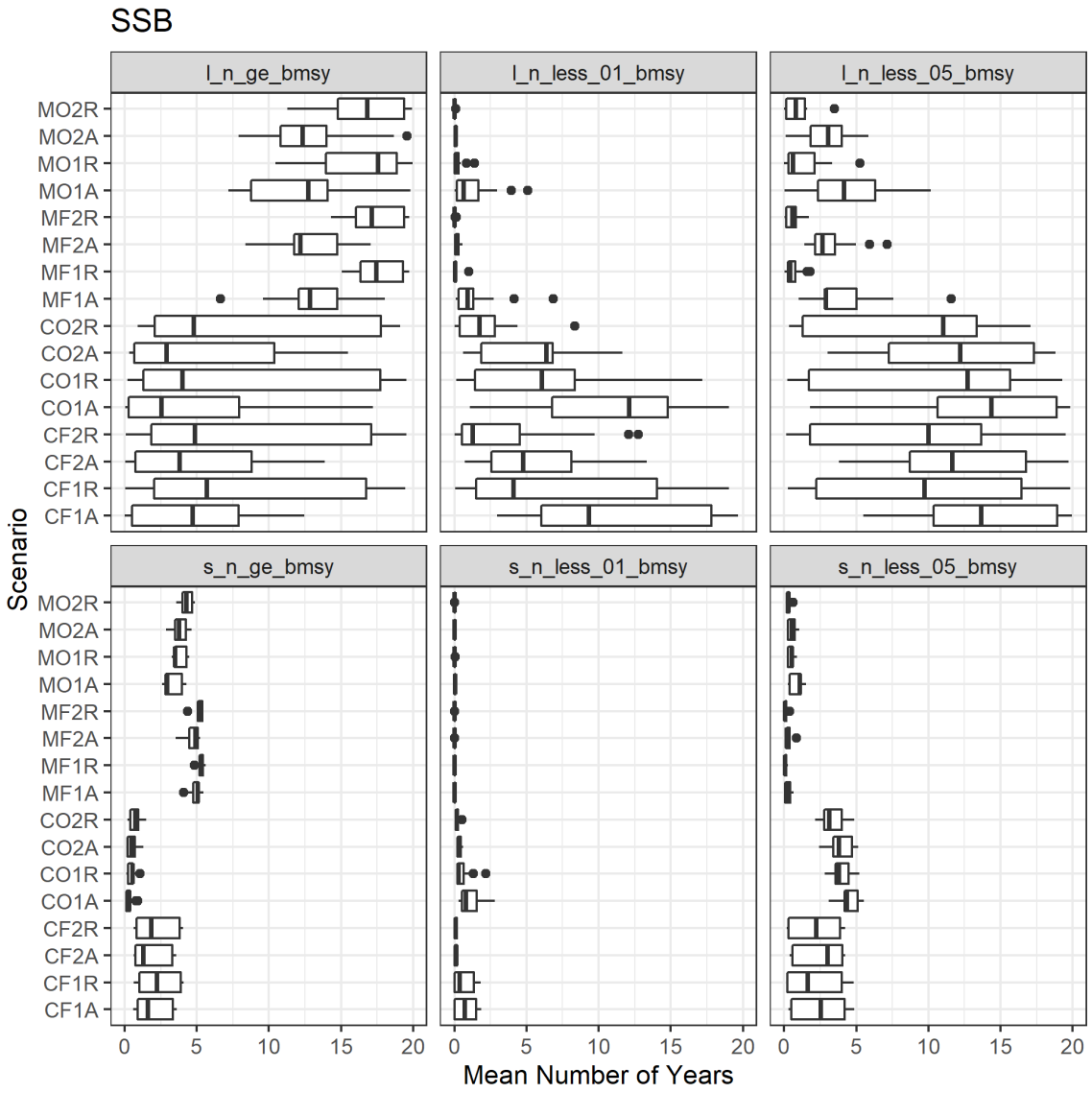


Figure A6.18. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the base analyses by scenario.

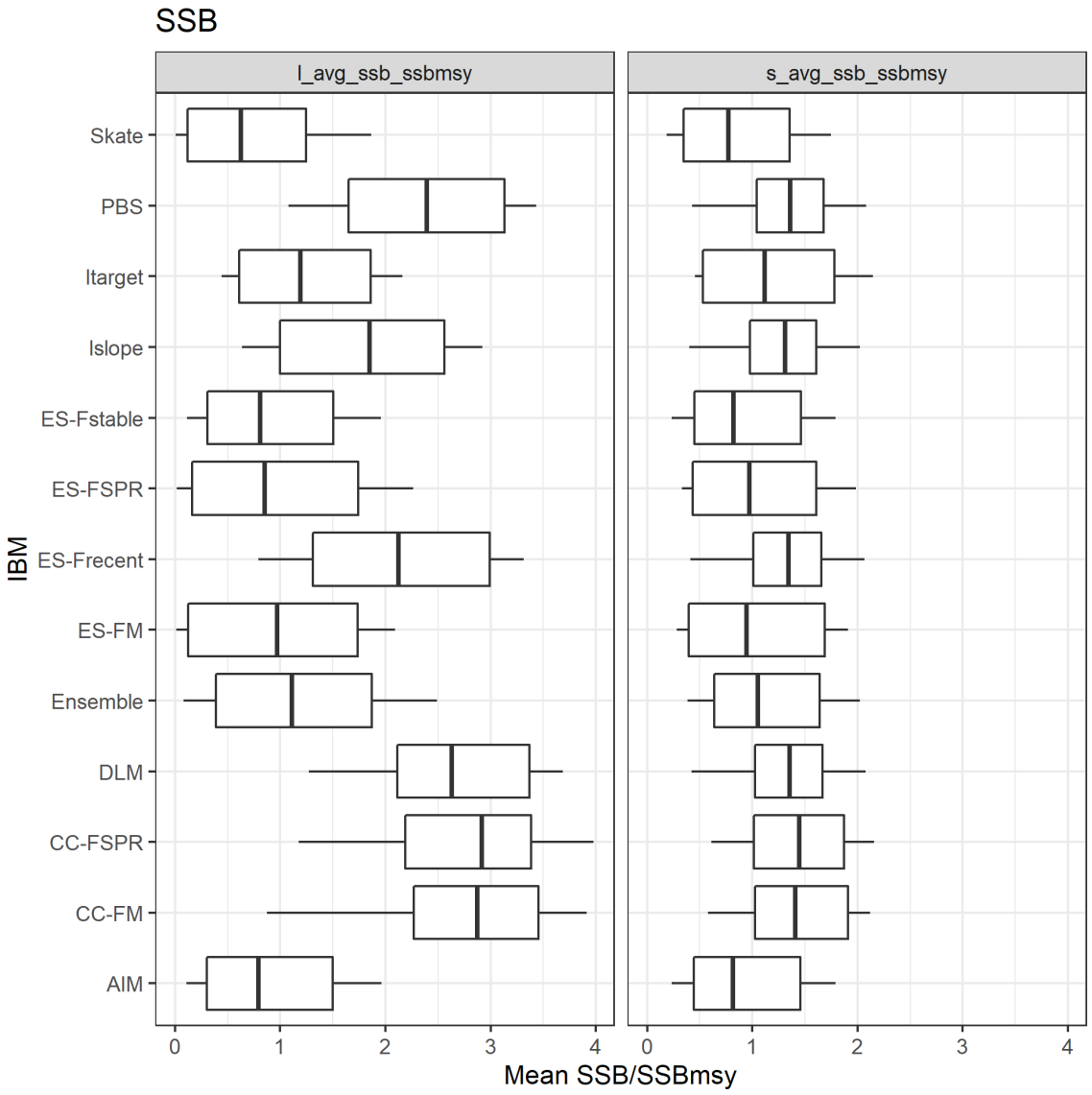


Figure A6.19. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the base analyses by index-based method (IBM).

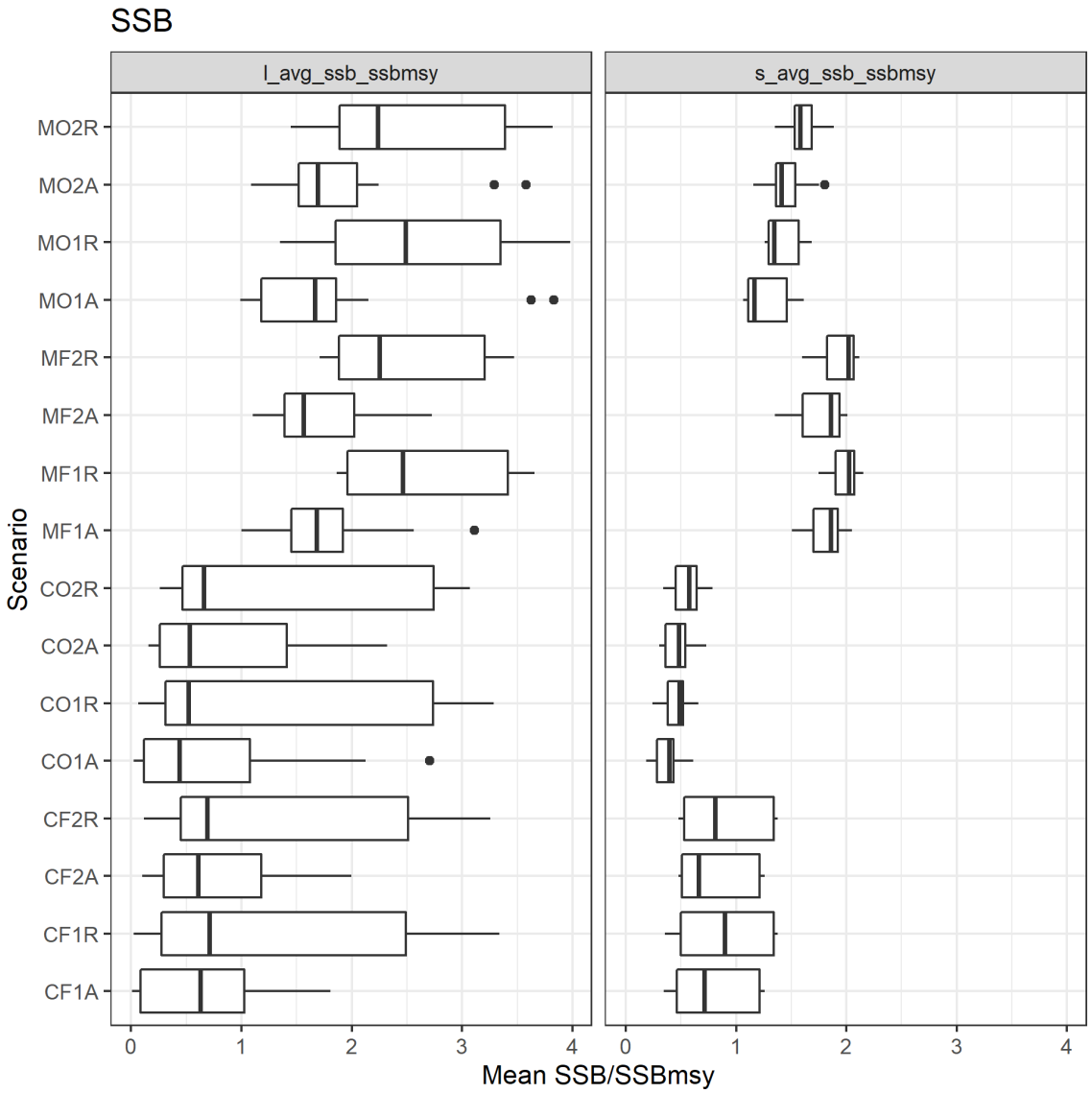


Figure A6.20. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the base analyses by scenario.

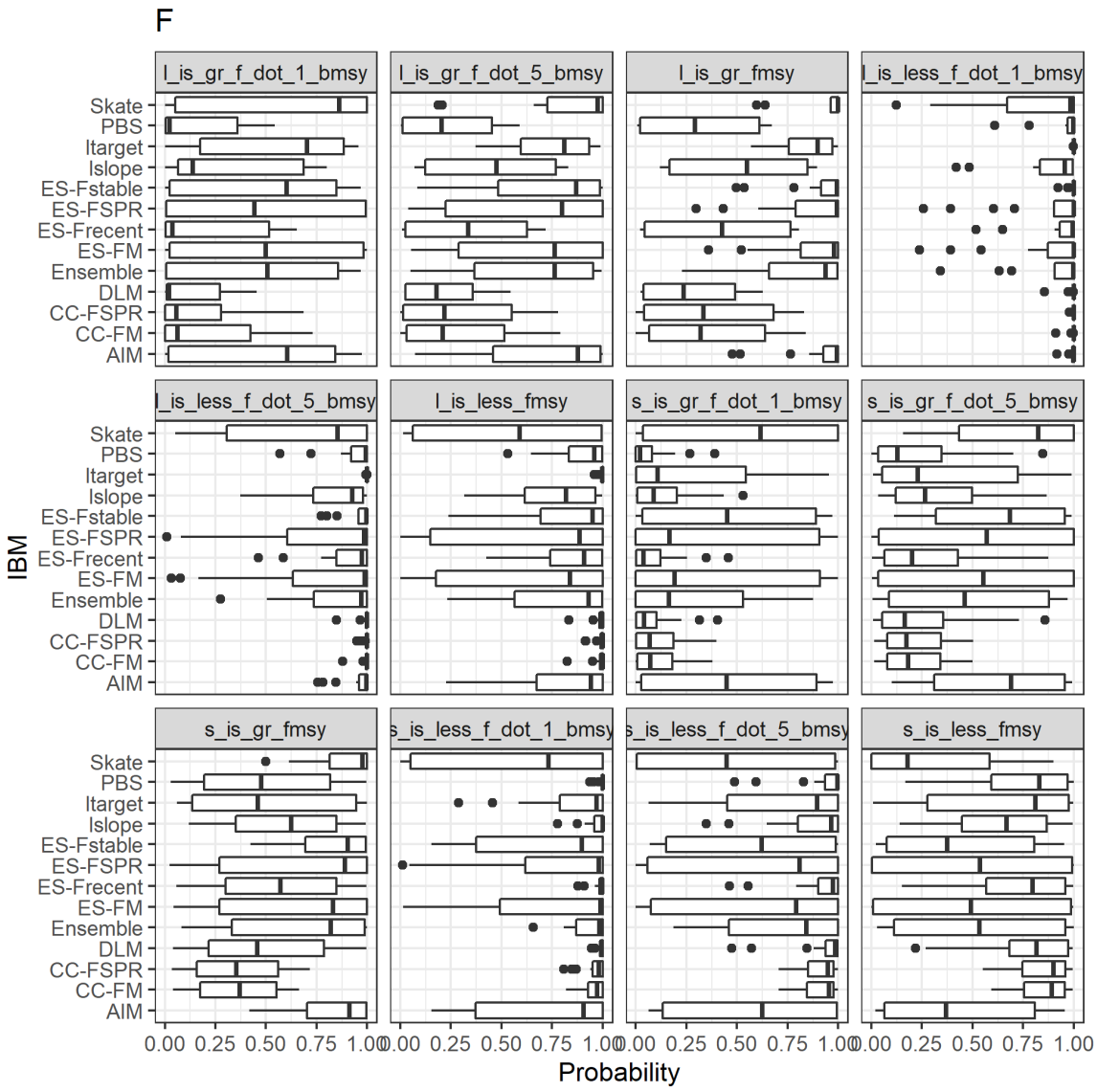


Figure A6.21. Boxplot of the mean values for the fishing mortality rate (F) metrics in the base analyses by index-based method (IBM).

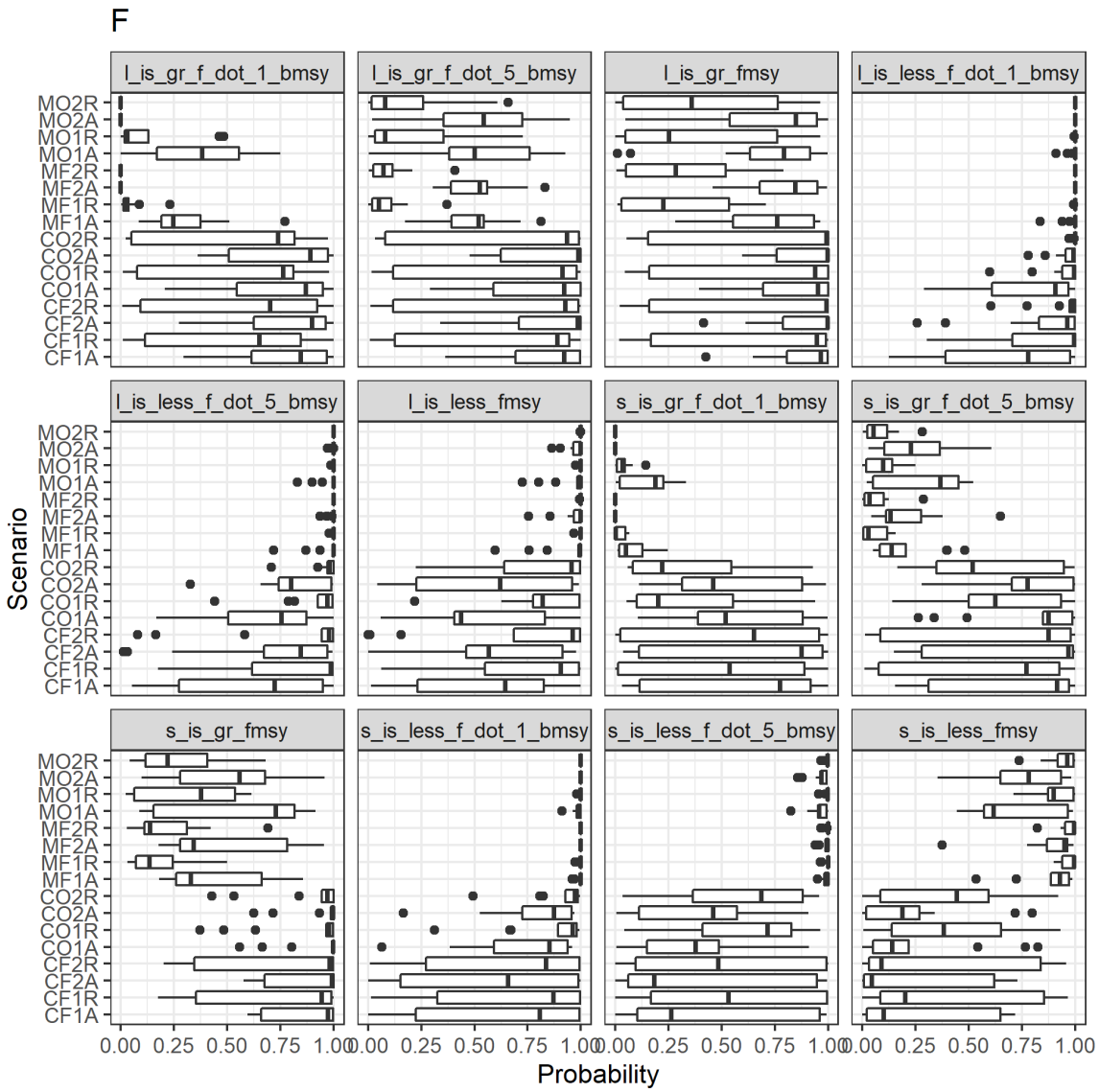


Figure A6.22. Boxplot of the mean values for the fishing mortality rate (F) metrics in the base analyses by scenario.

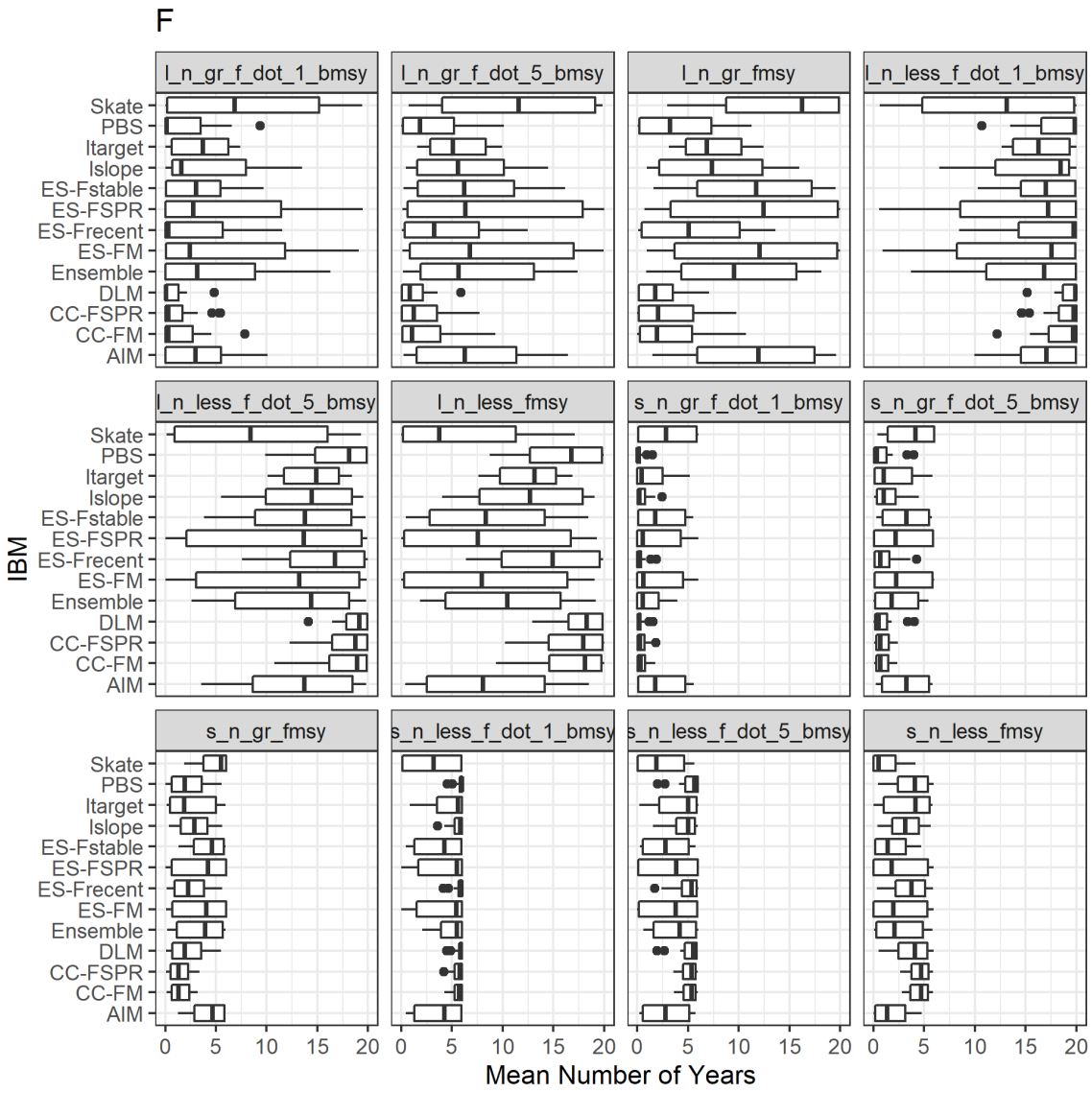


Figure A6.23. Boxplot of the mean values for the fishing mortality rate (F) metrics in the base analyses by index-based method (IBM).

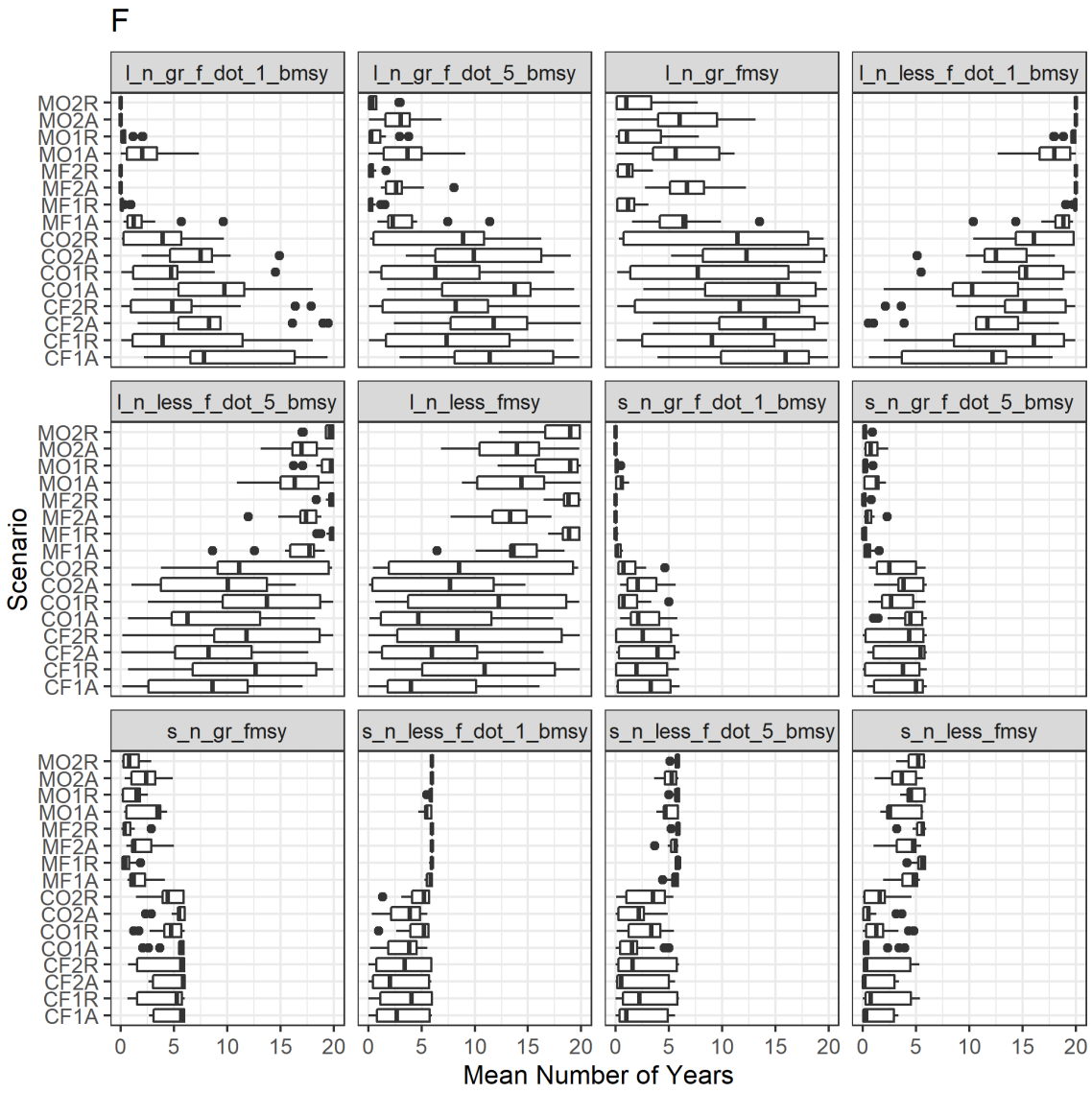


Figure A6.24. Boxplot of the mean values for the fishing mortality rate (F) metrics in the base analyses by scenario.

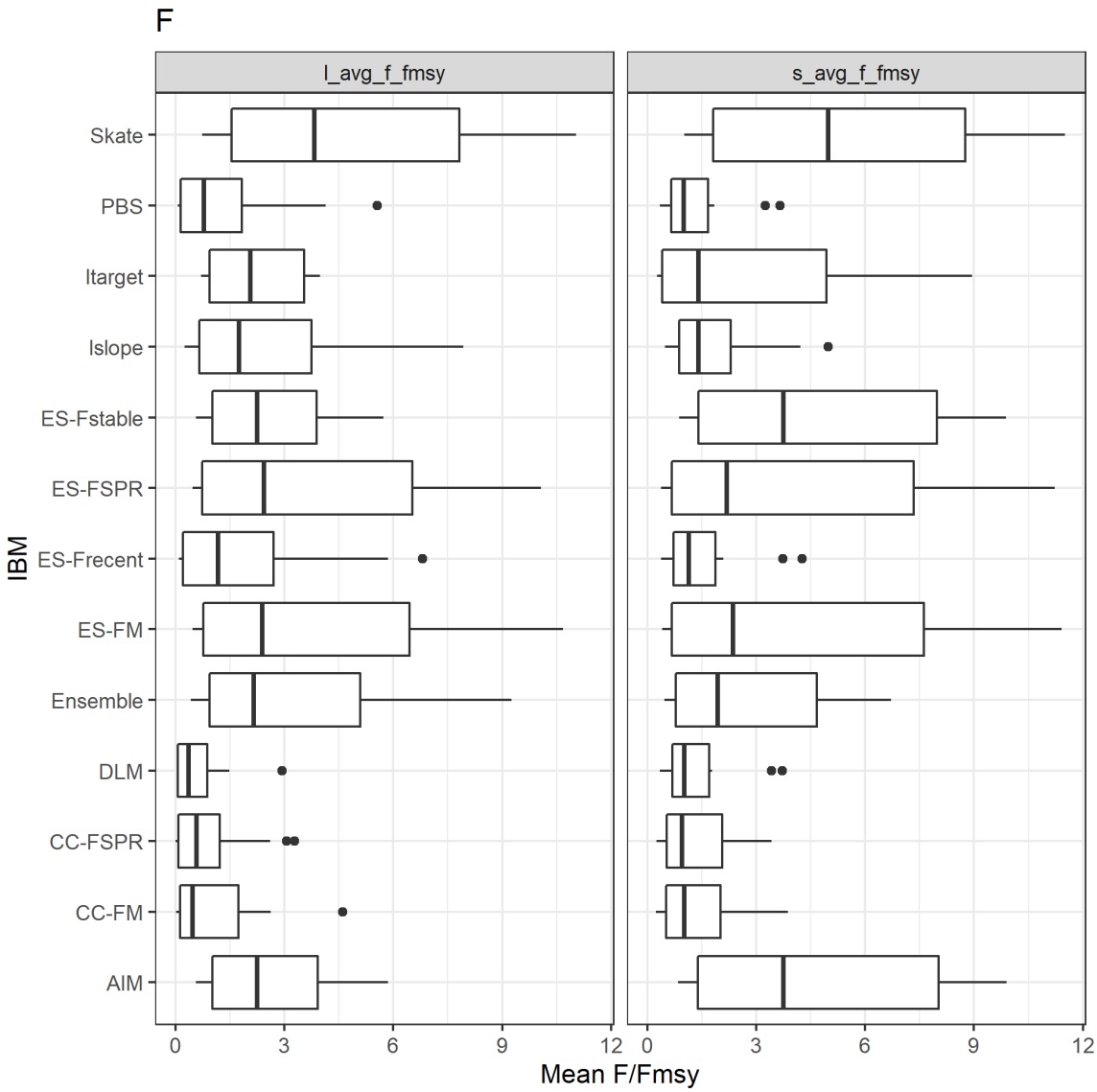


Figure A6.25. Boxplot of the mean values for the fishing mortality rate (F) metrics in the base analyses by index-based method (IBM).

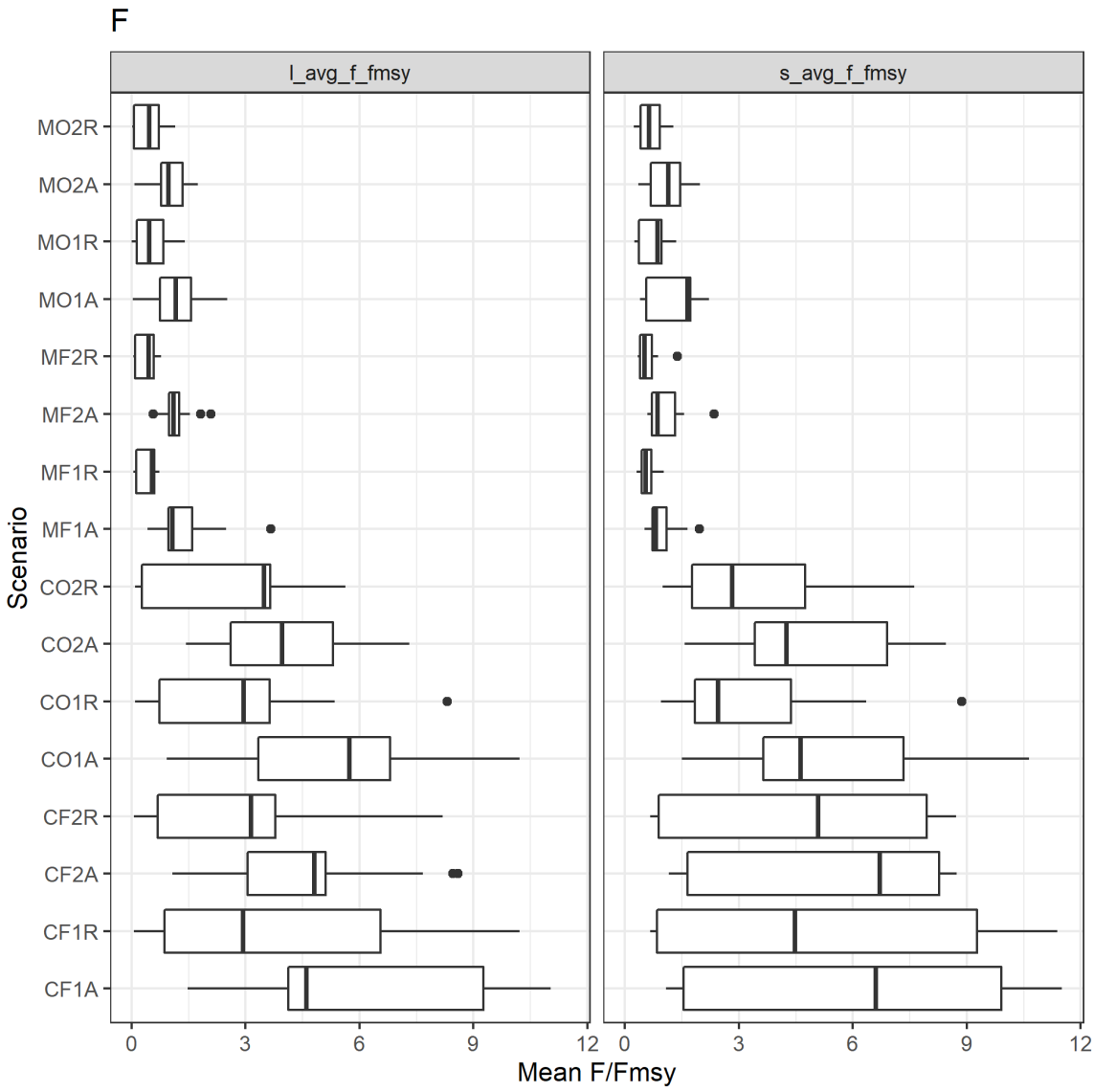


Figure A6.26. Boxplot of the mean values for the fishing mortality rate (F) metrics in the base analyses by scenario.

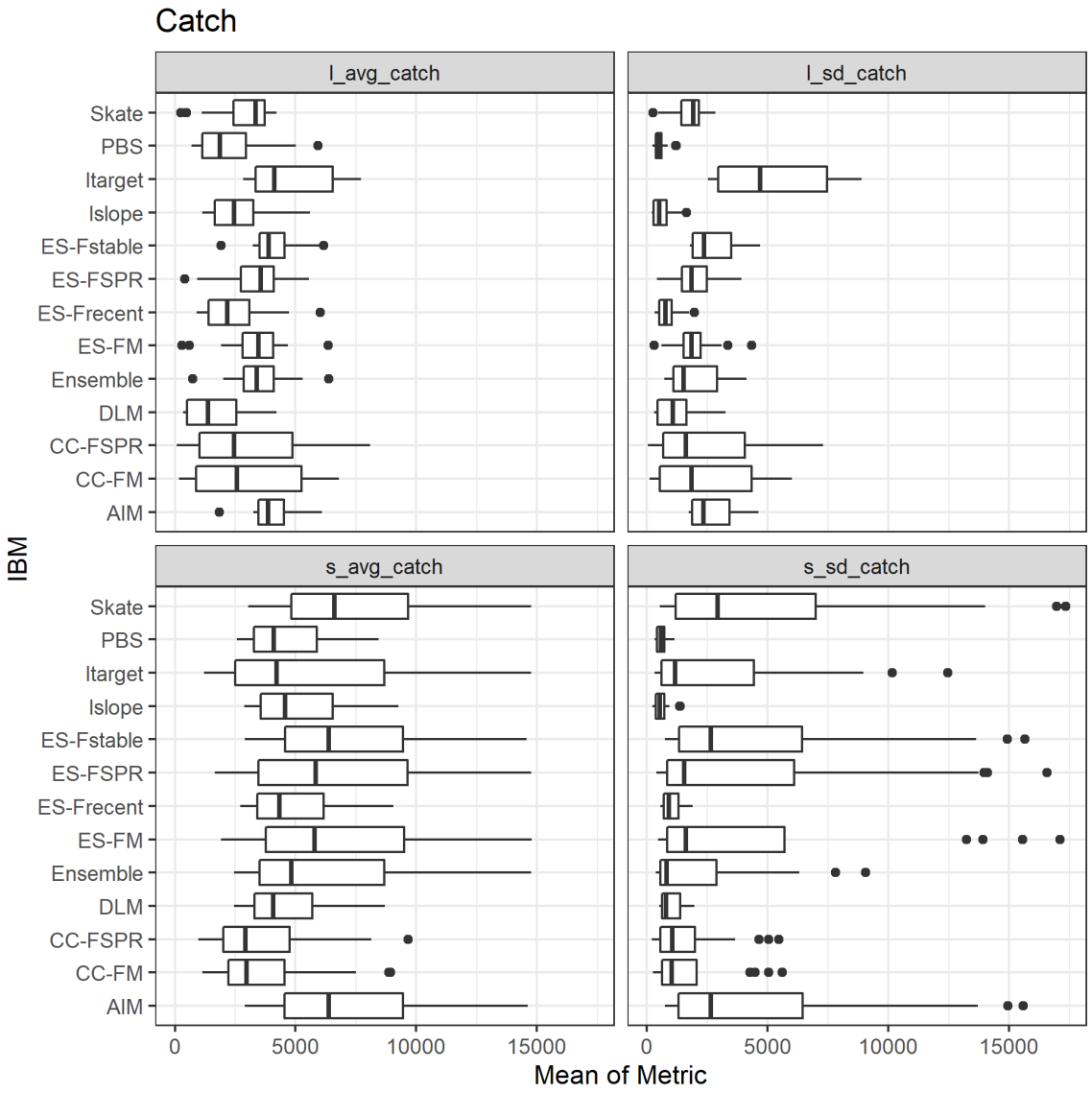


Figure A6.27. Boxplot of the mean values for the catch metrics in the base analyses by index-based method (IBM).

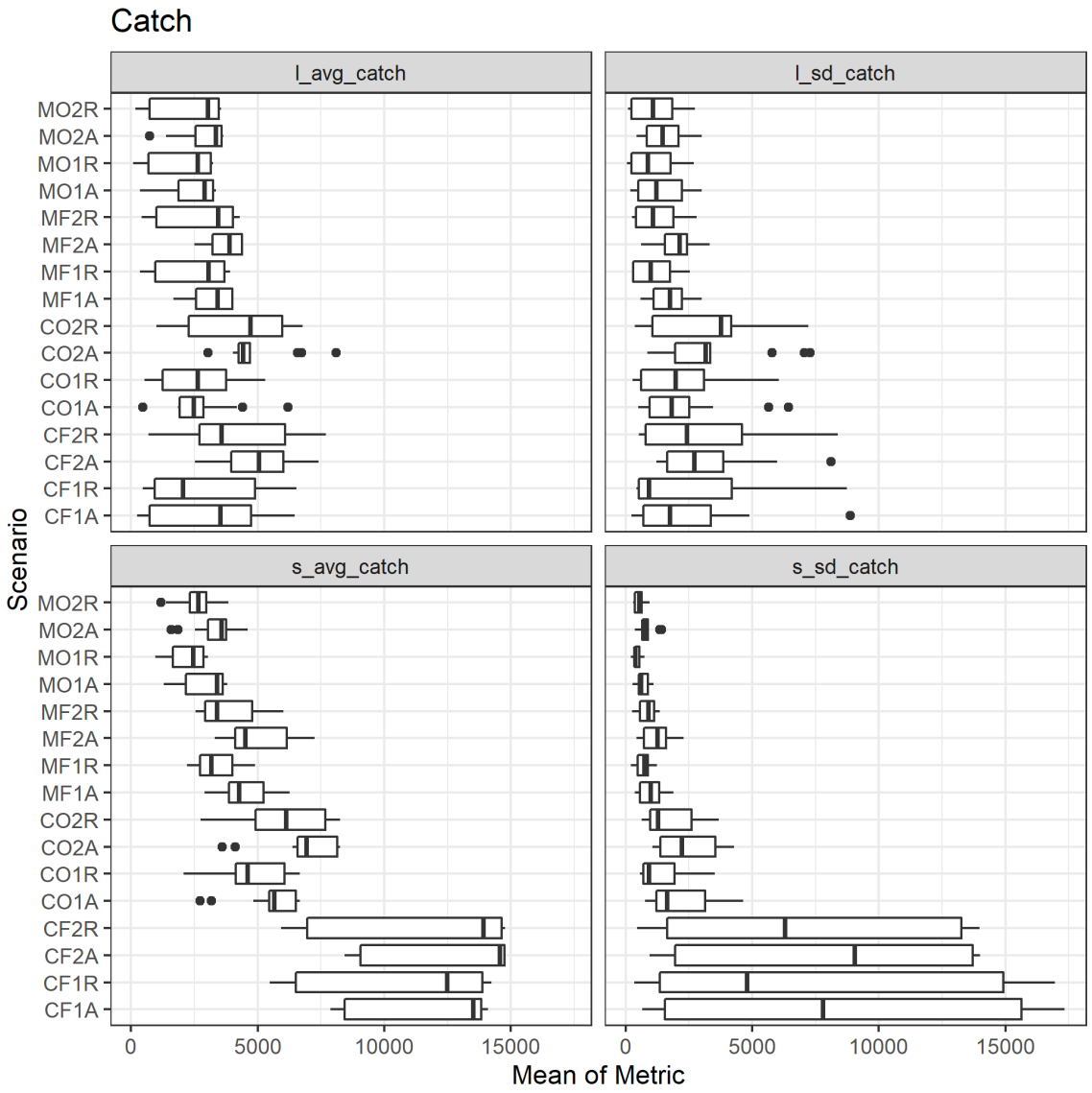


Figure A6.28. Boxplot of the mean values for the catch metrics in the base analyses by scenario.

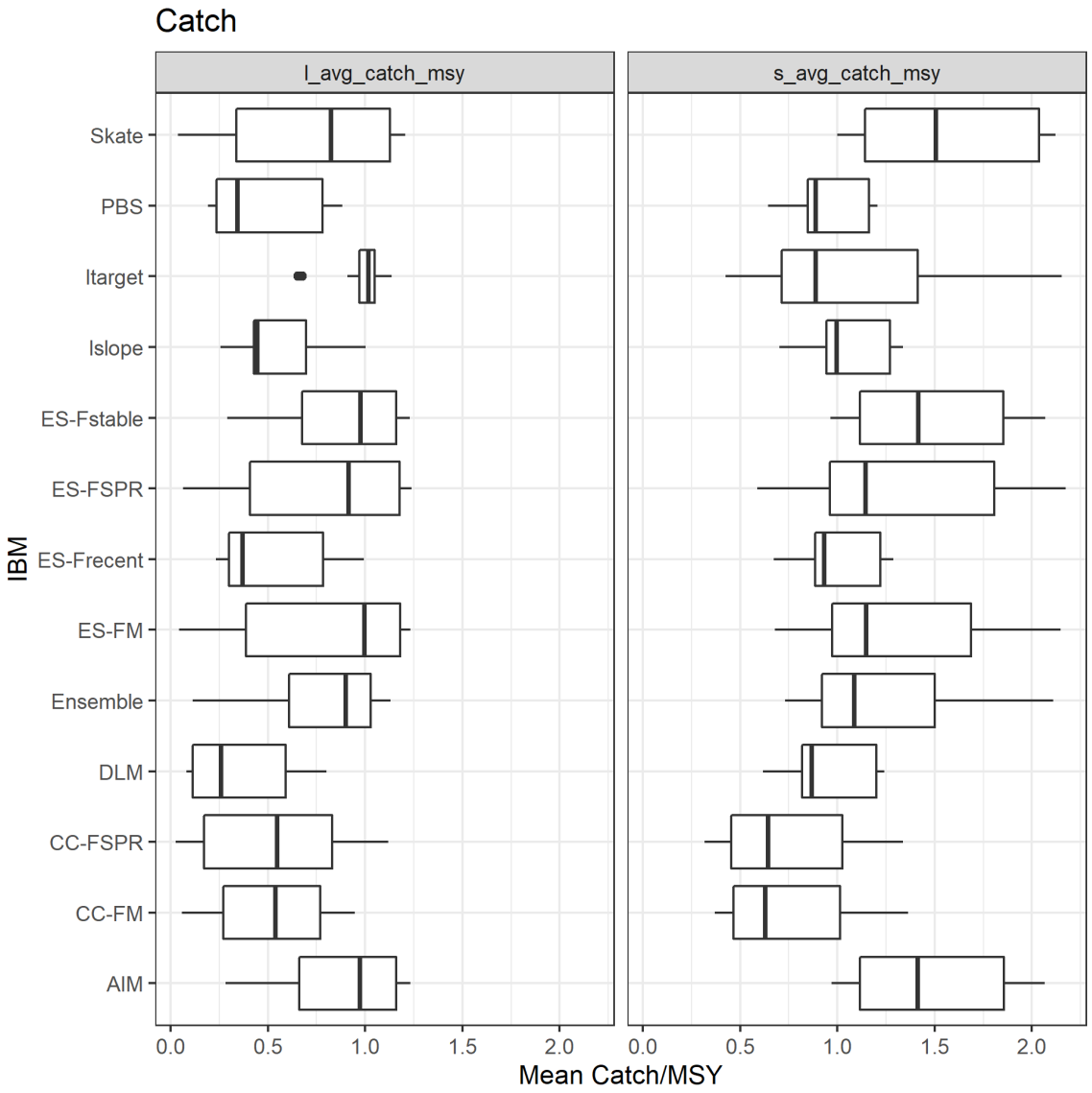


Figure A6.29. Boxplot of the mean values for the catch metrics in the base analyses by index-based method (IBM).

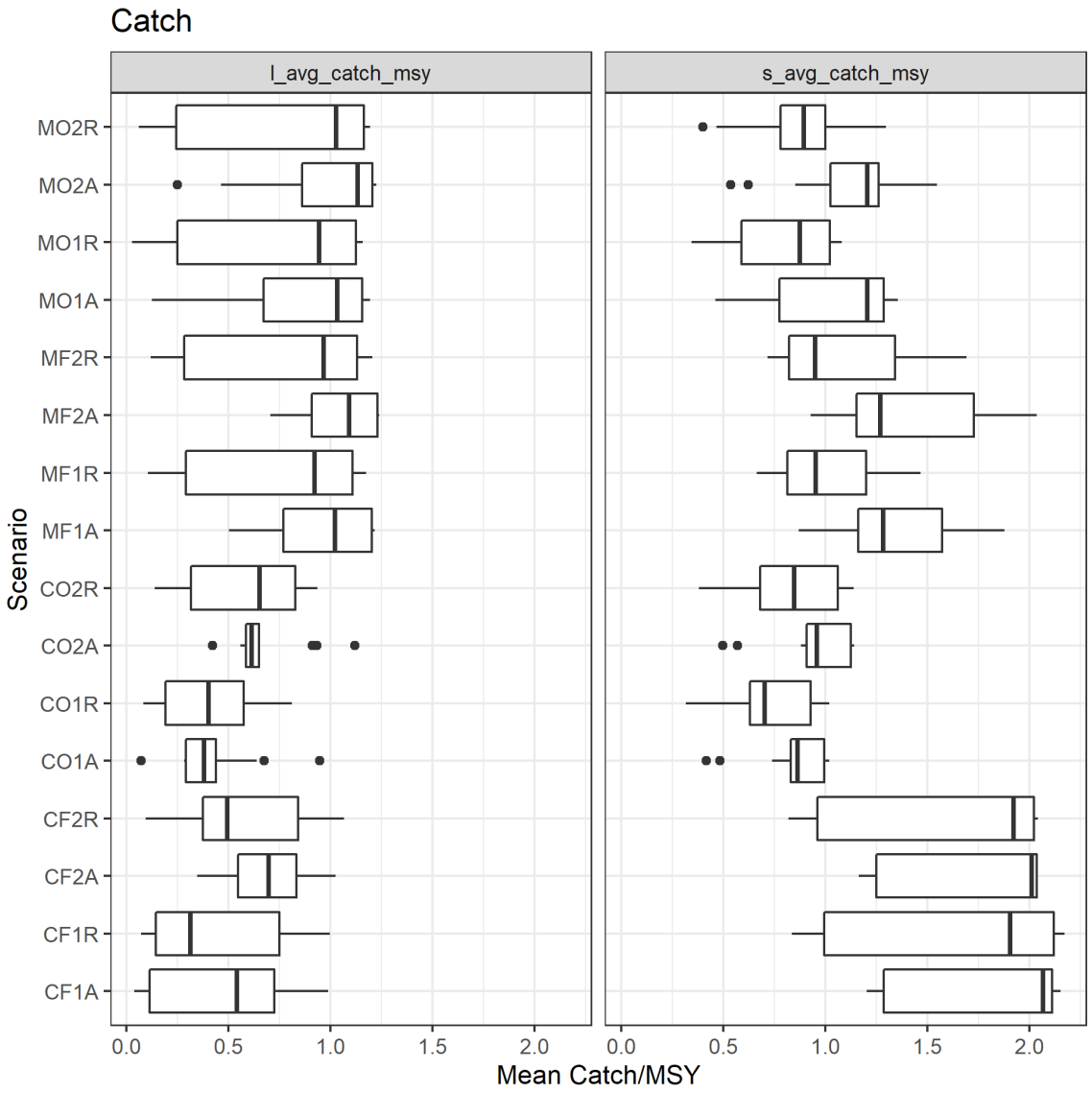


Figure A6.30. Boxplot of the mean values for the catch metrics in the base analyses by scenario.

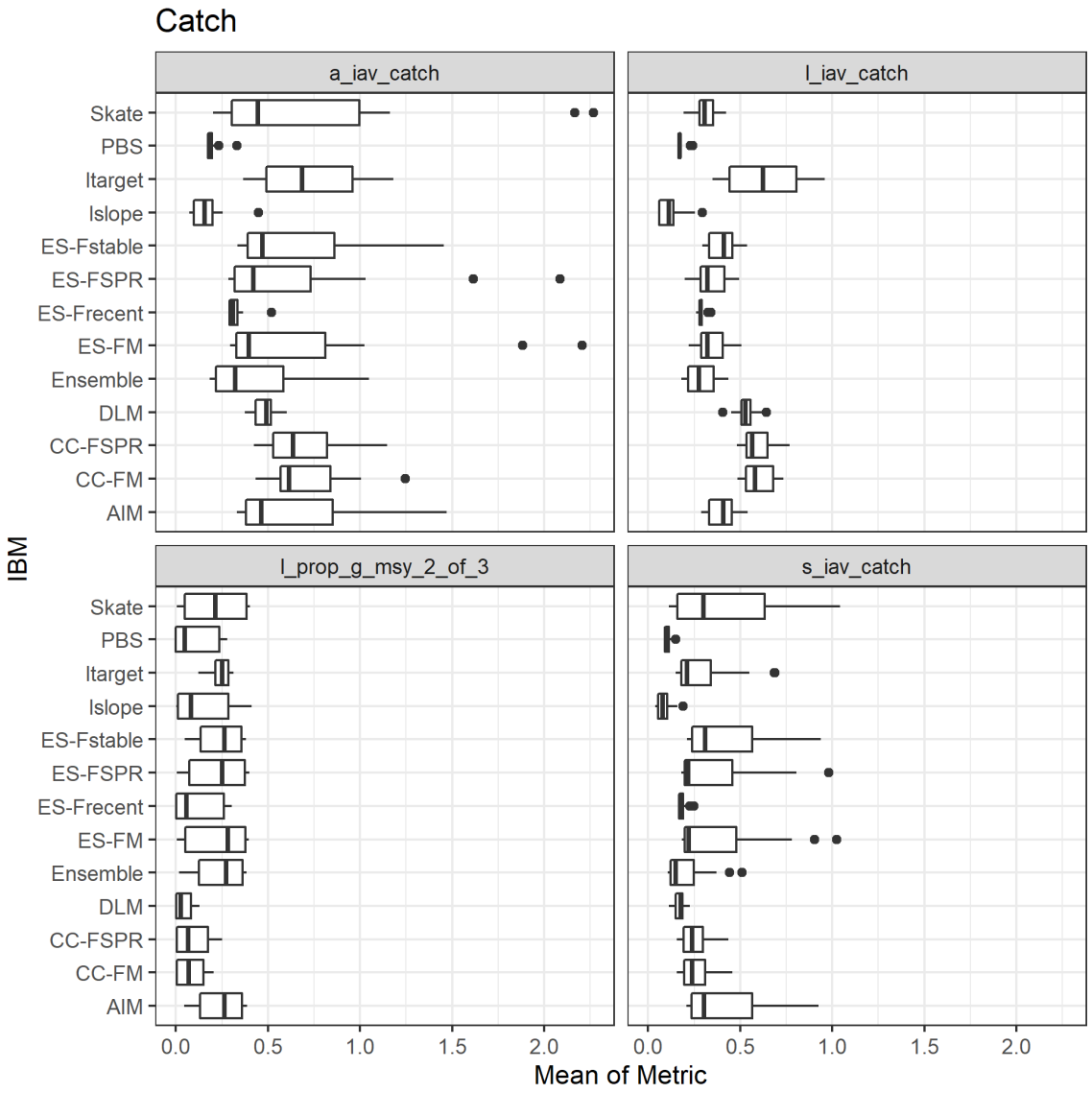


Figure A6.31. Boxplot of the mean values for the catch metrics in the base analyses by index-based method (IBM).

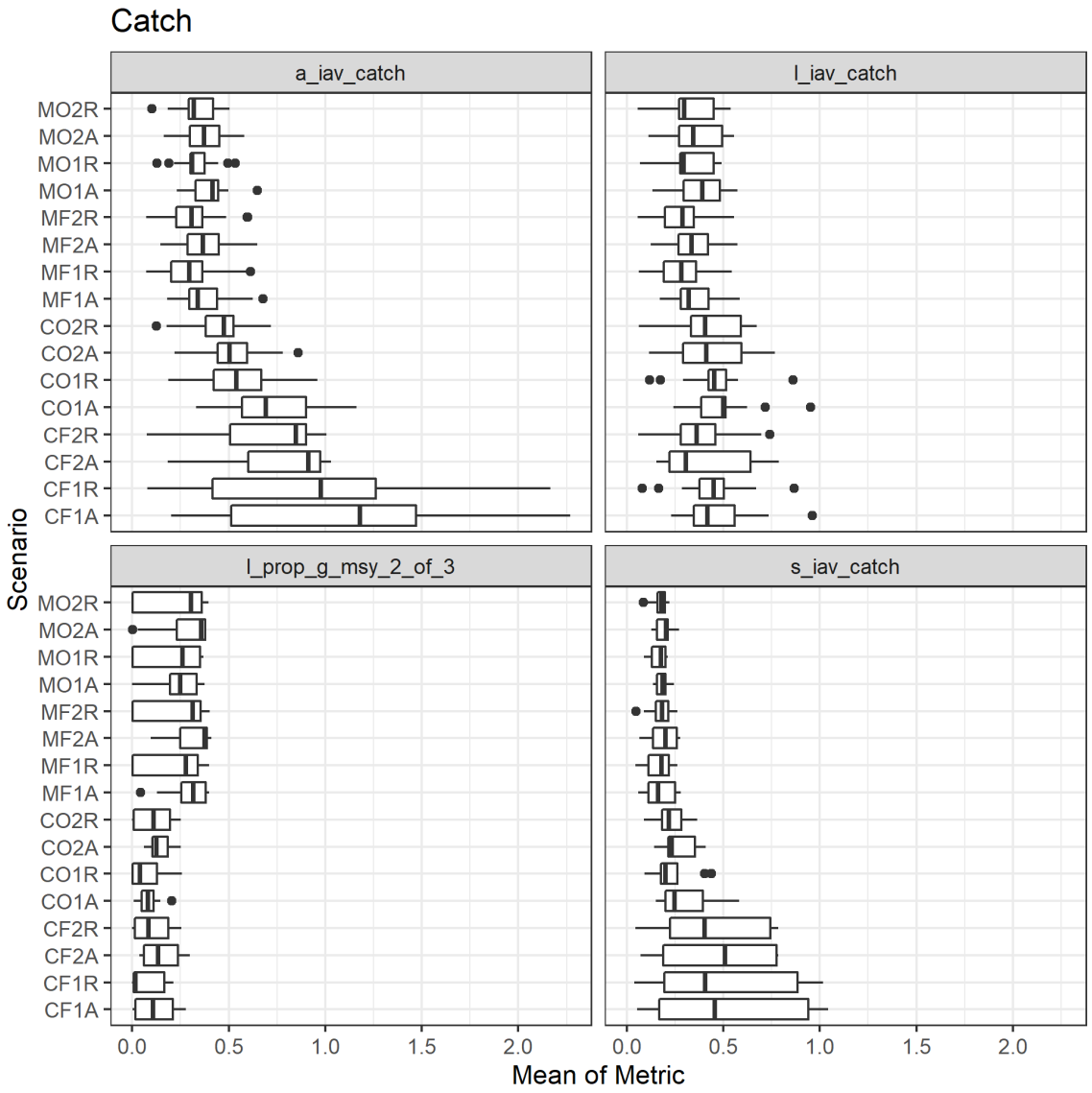


Figure A6.32. Boxplot of the mean values for the catch metrics in the base analyses by scenario.

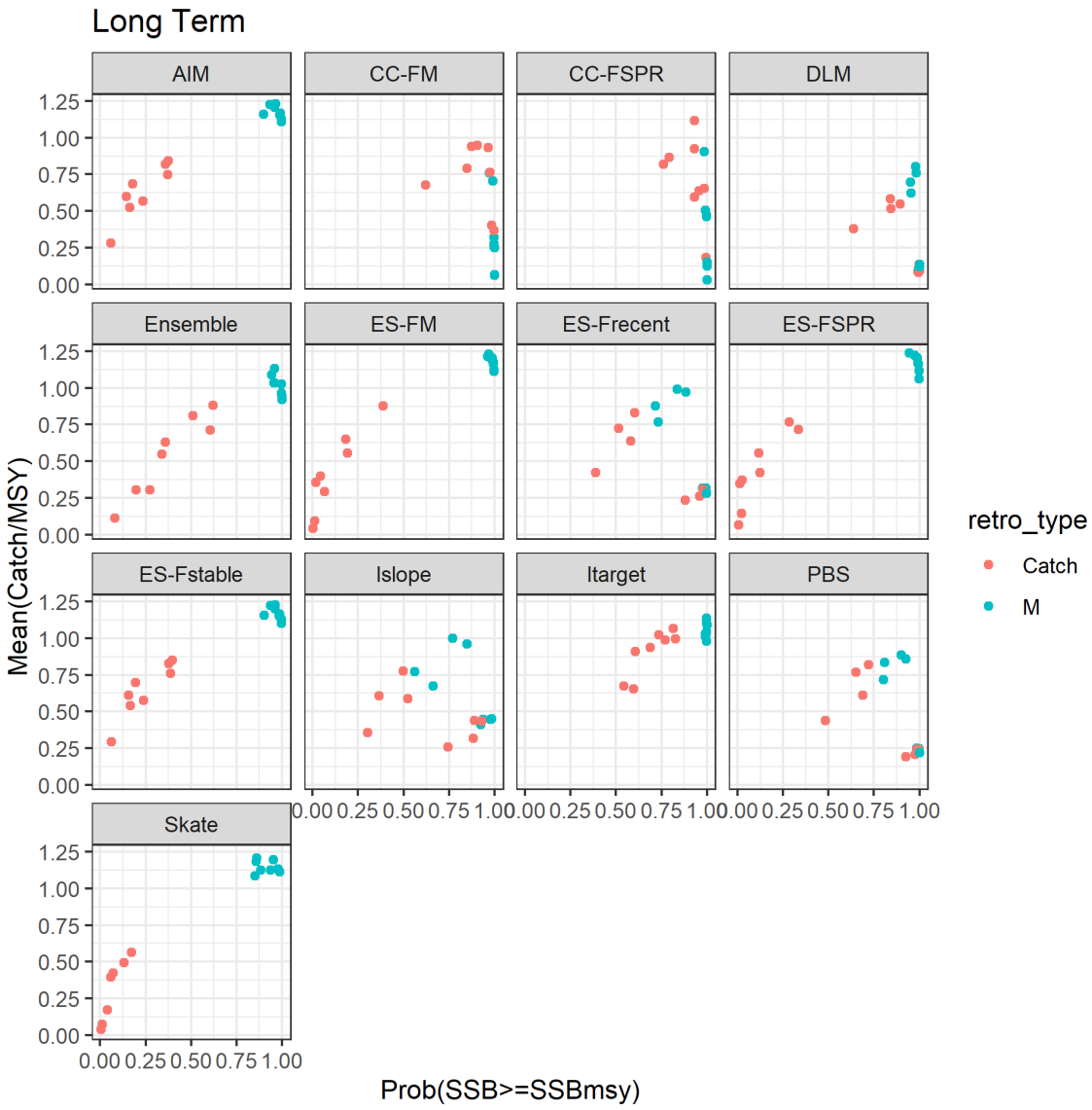


Figure A6.33. Trade off plot by index-based method (IBM) for the base analyses between the probability of a rebuilt stock and the mean catch relative to maximum sustainable yield (MSY) in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

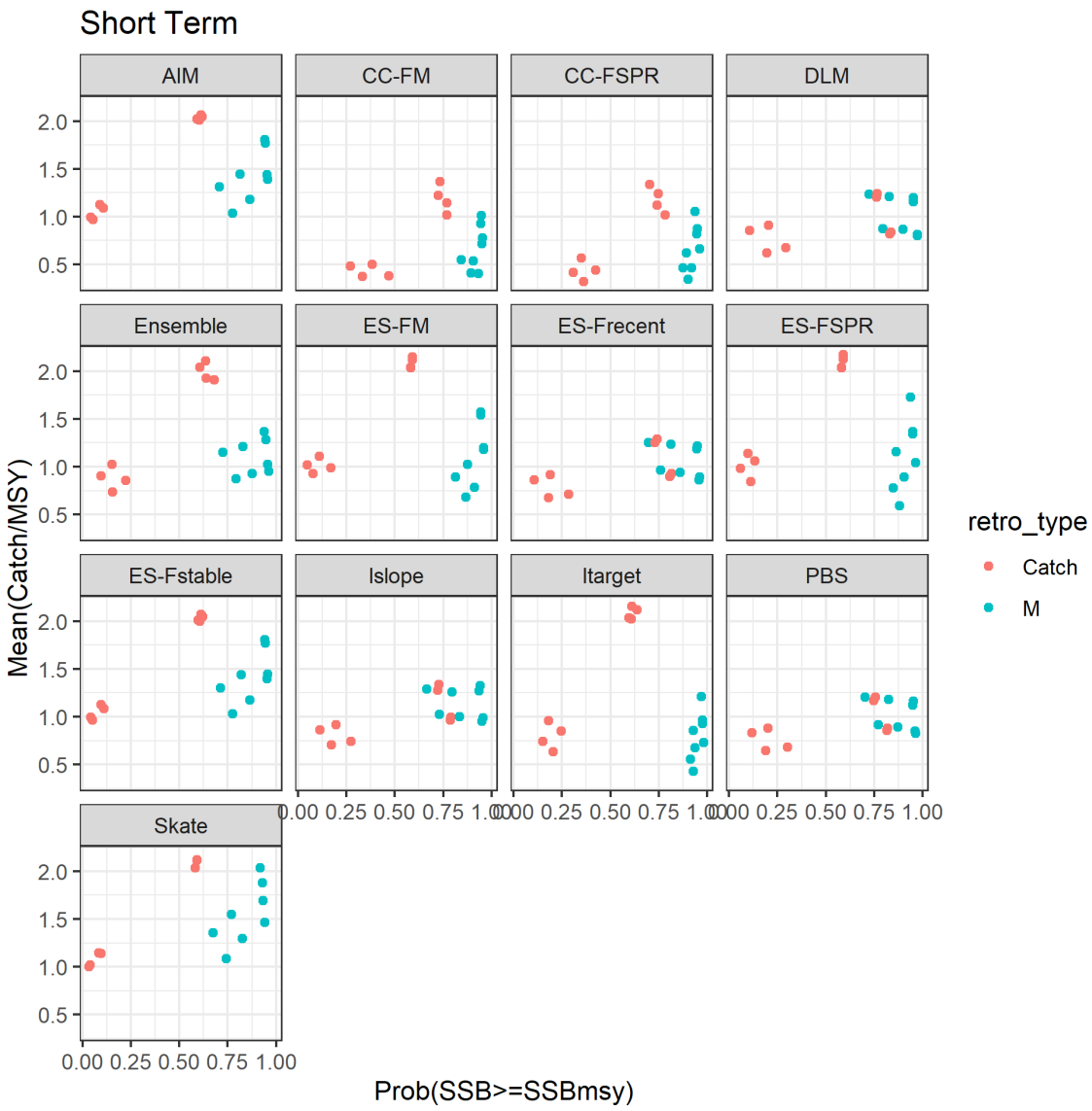


Figure A6.34. Trade off plot by index-based method (IBM) for the base analyses between the probability of a rebuilt stock and the mean catch relative to maximum sustainable yield (MSY) in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

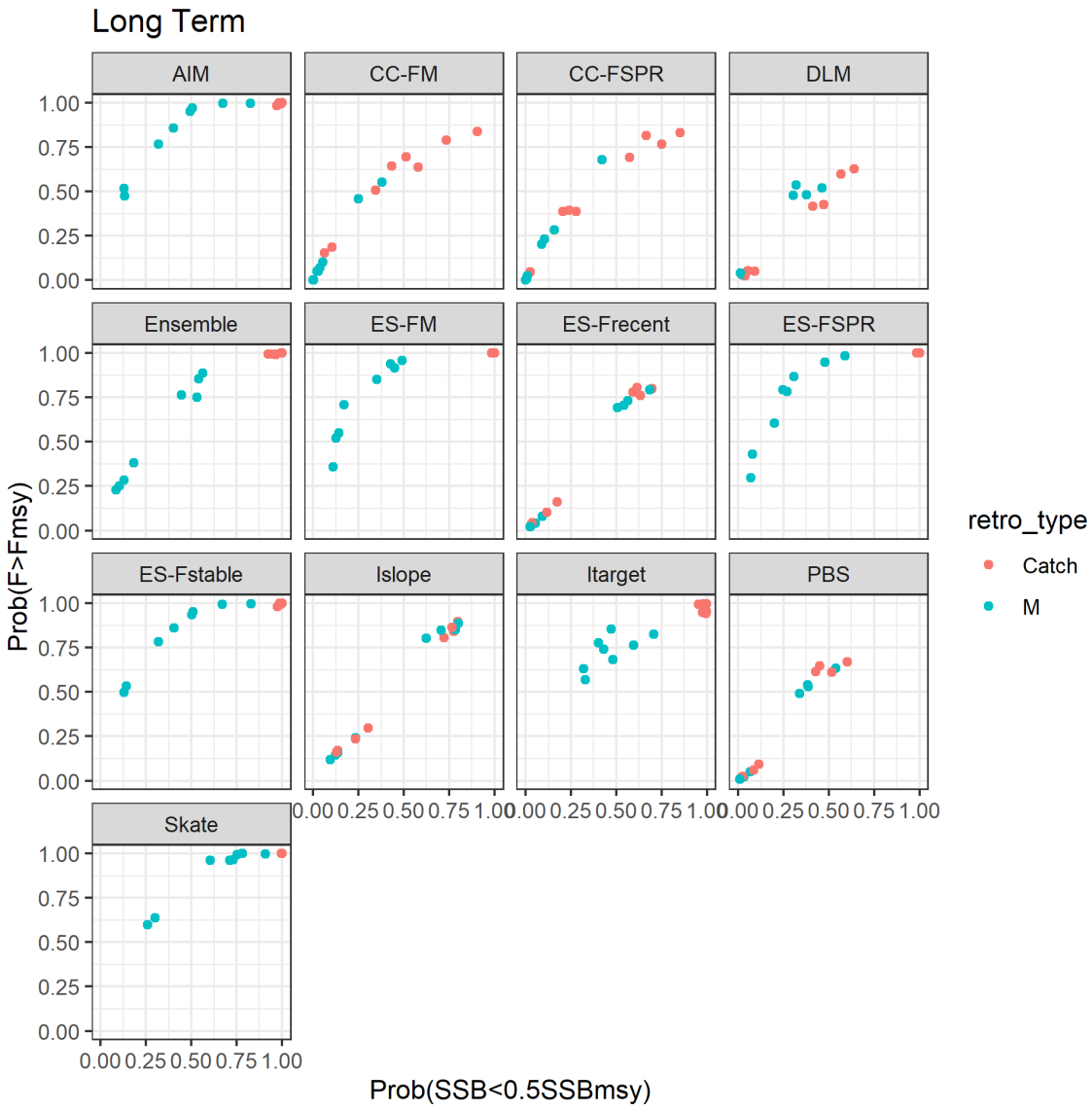


Figure A6.35. Trade off plot by index-based method (IBM) for the base analyses between the probability the stock is overfished and the probability that overfishing is occurring in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

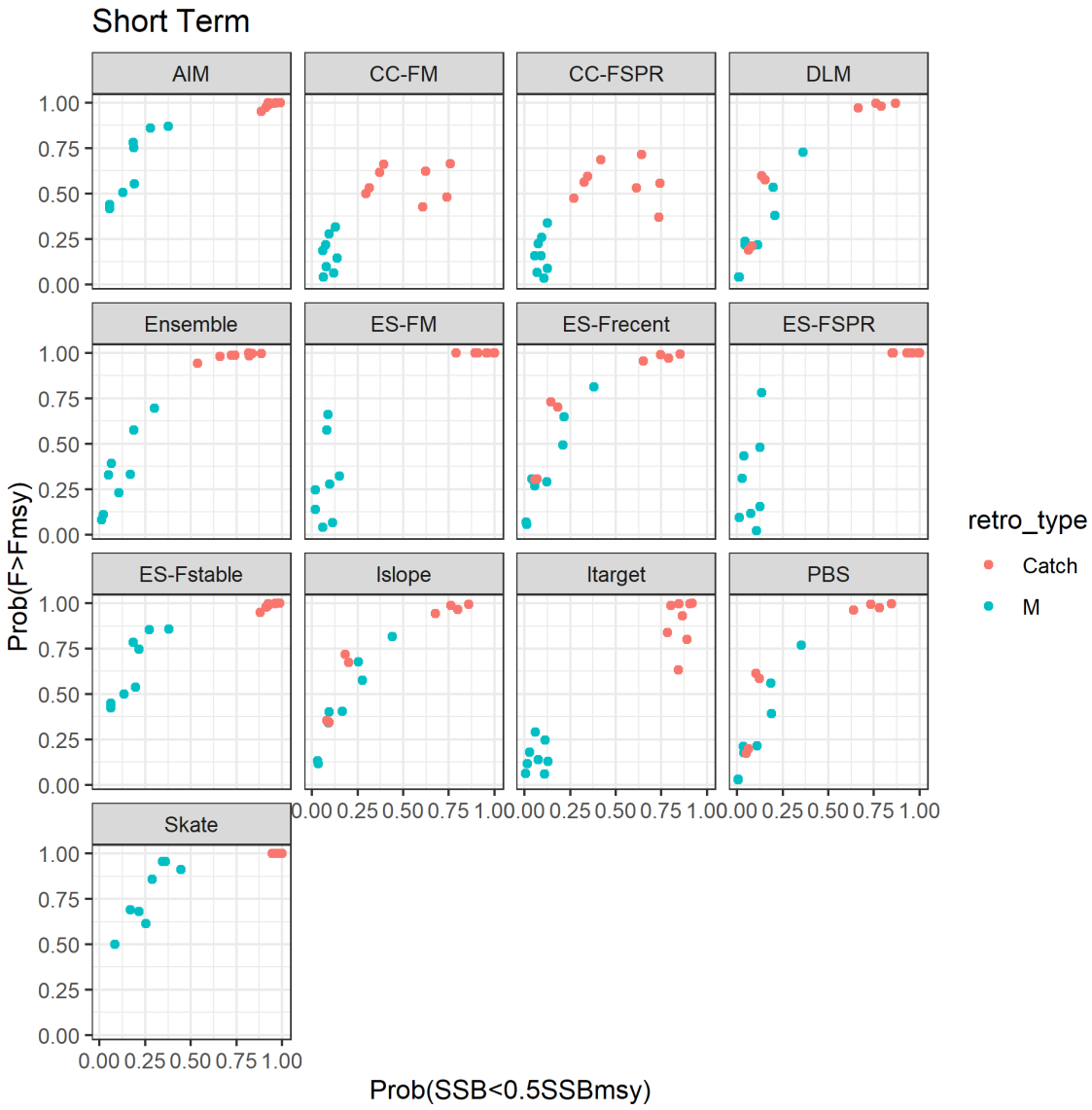


Figure A6.36. Trade off plot by index-based method (IBM) for the base analyses between the probability the stock is overfished and the probability that overfishing is occurring in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

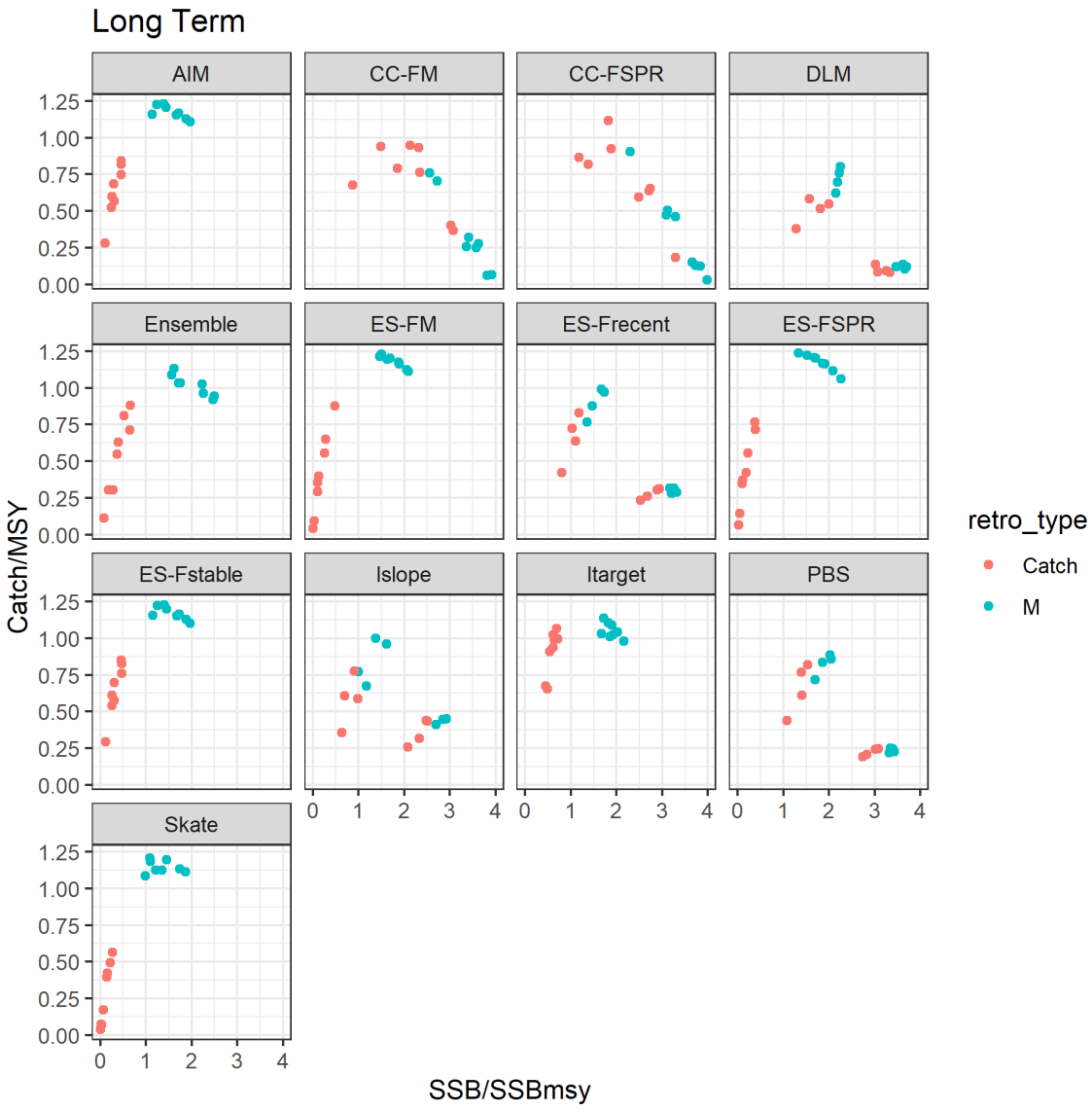


Figure A6.37. Trade off plot by index-based method (IBM) for the base analyses between the average spawning stock biomass (SSB) and catch relative to their maximum sustainable yield (MSY) reference points in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

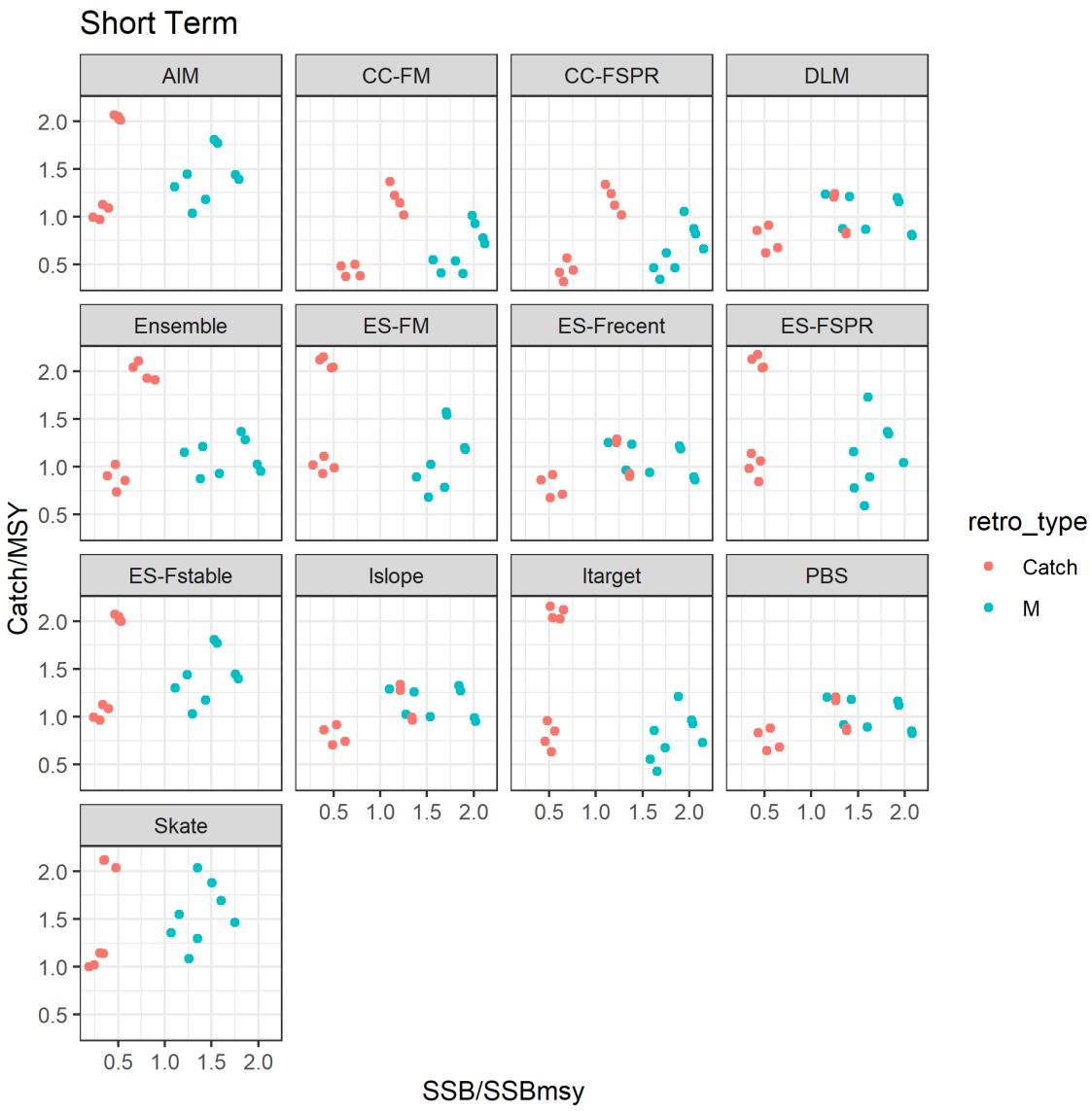


Figure A6.38. Trade off plot by index-based method (IBM) for the base analyses between the average spawning stock biomass (SSB) and catch relative to their maximum sustainable yield (MSY) reference points in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

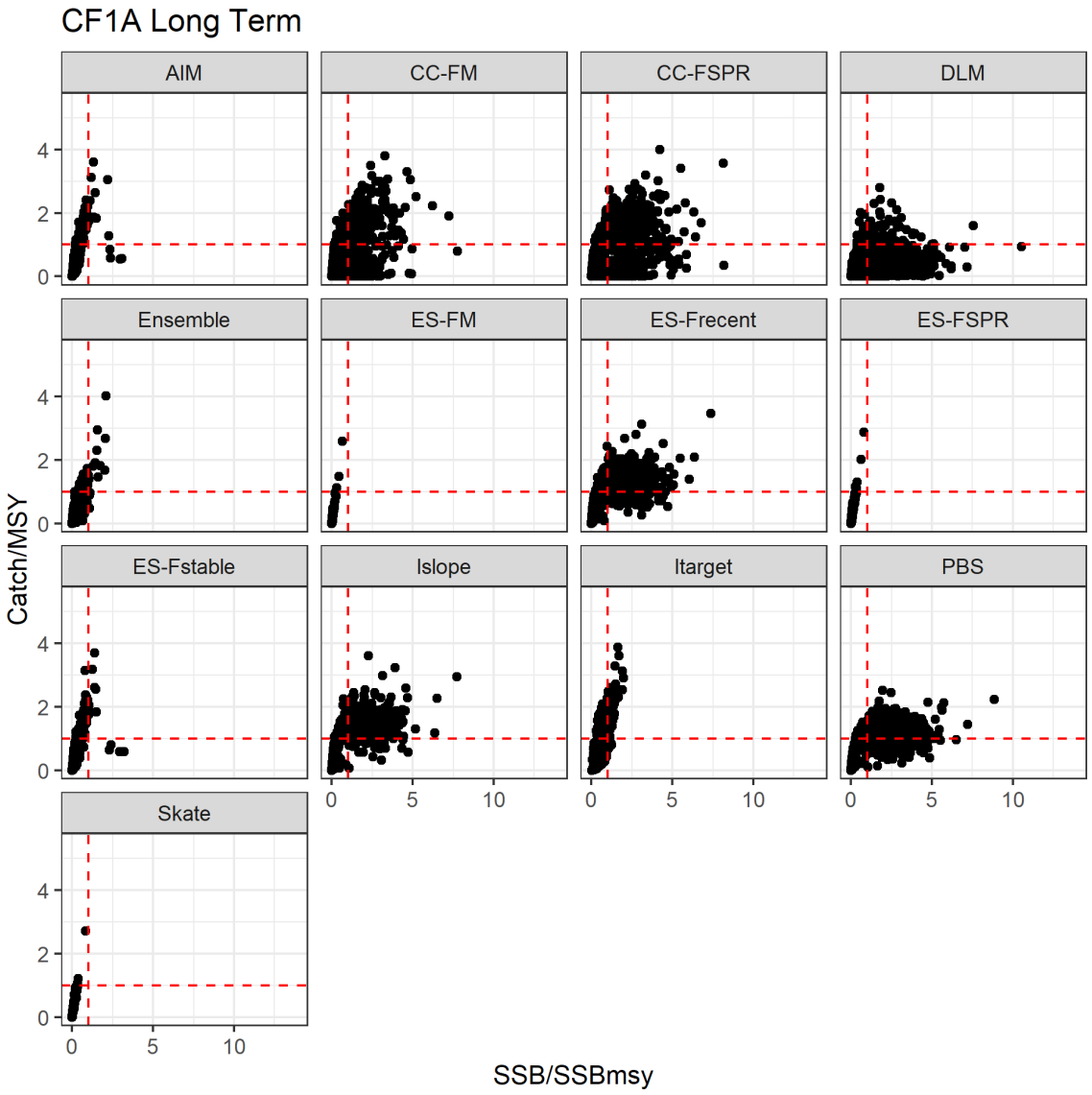


Figure A6.39. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

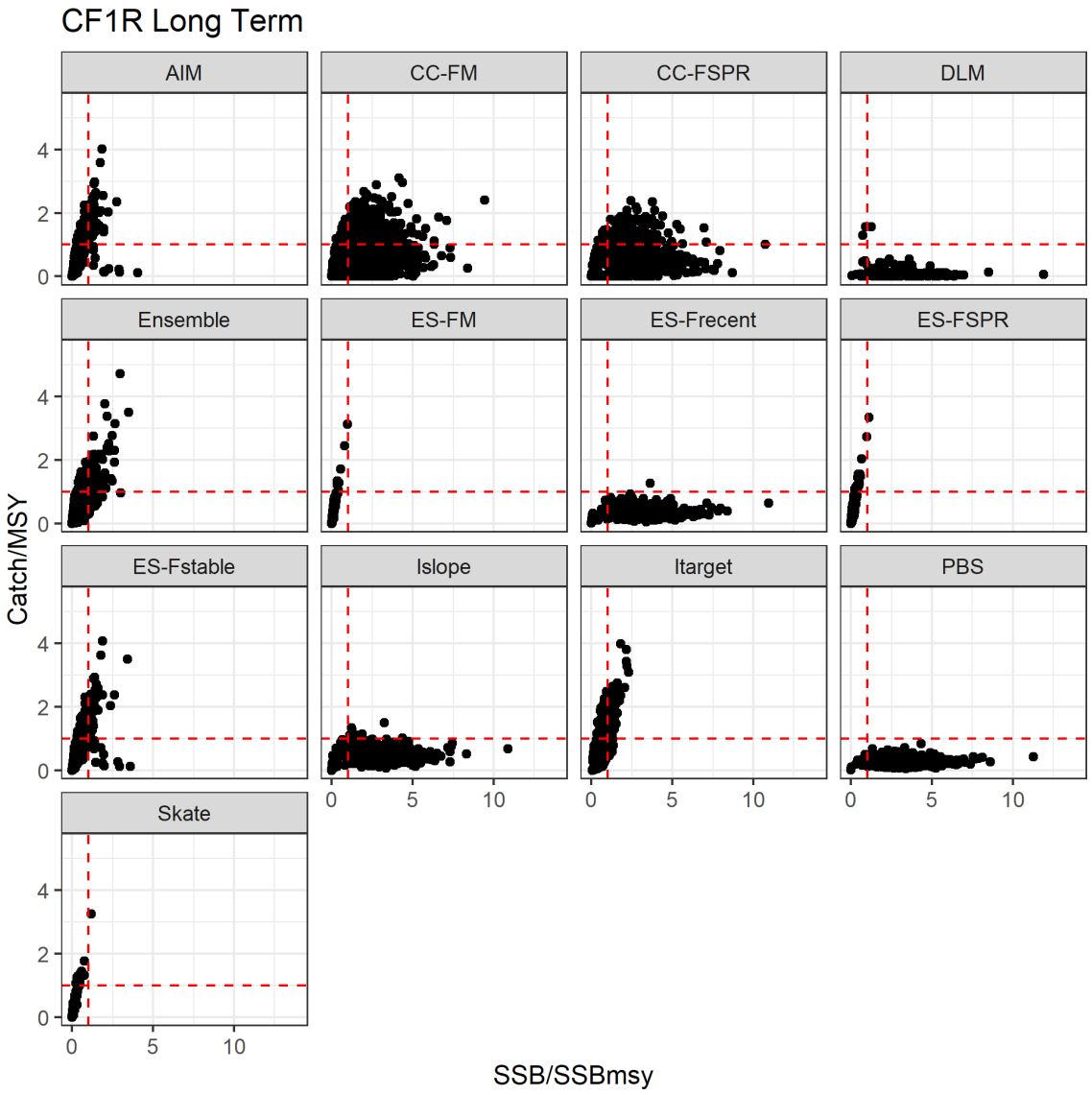


Figure A6.40. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

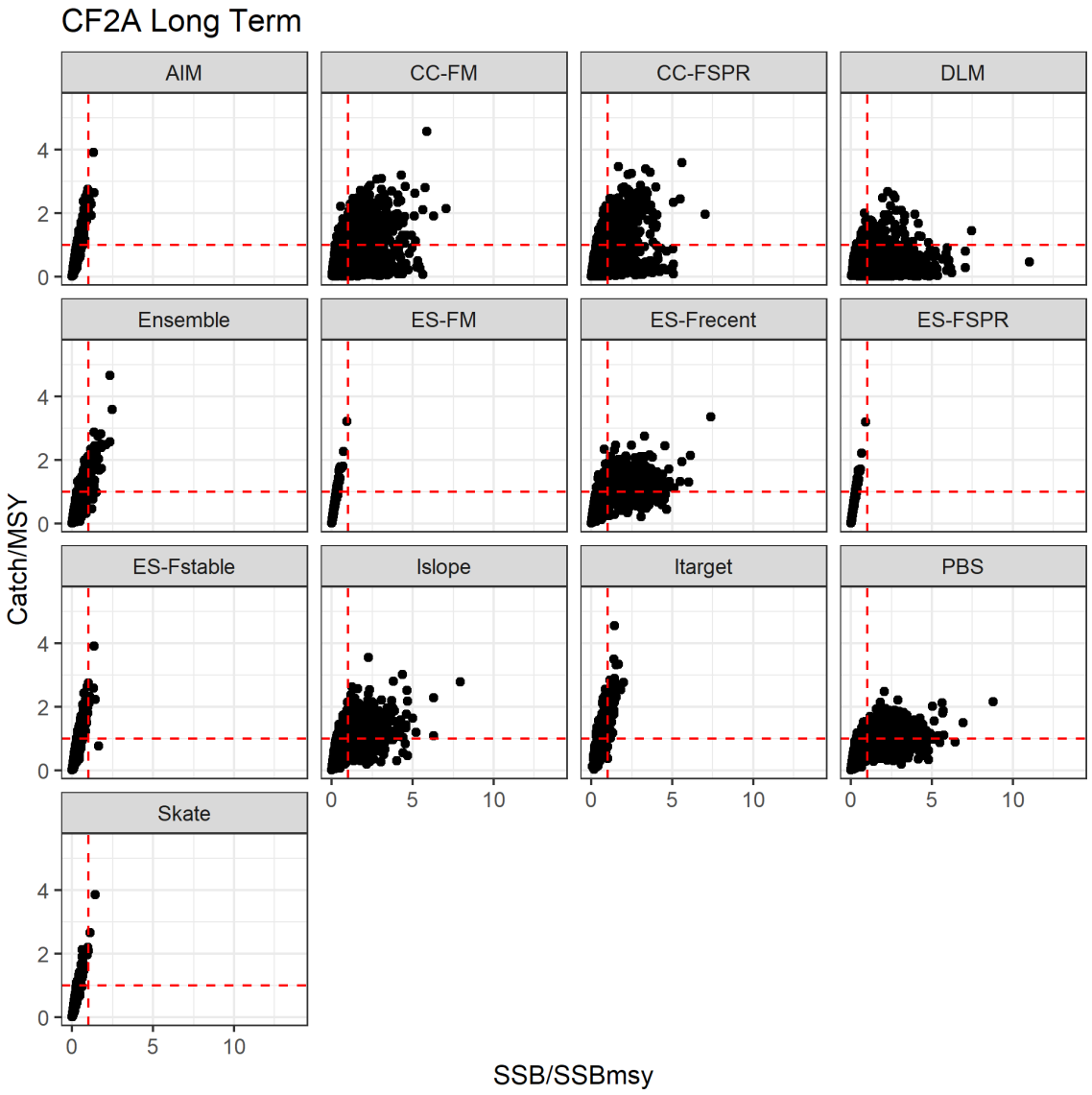


Figure A6.41. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

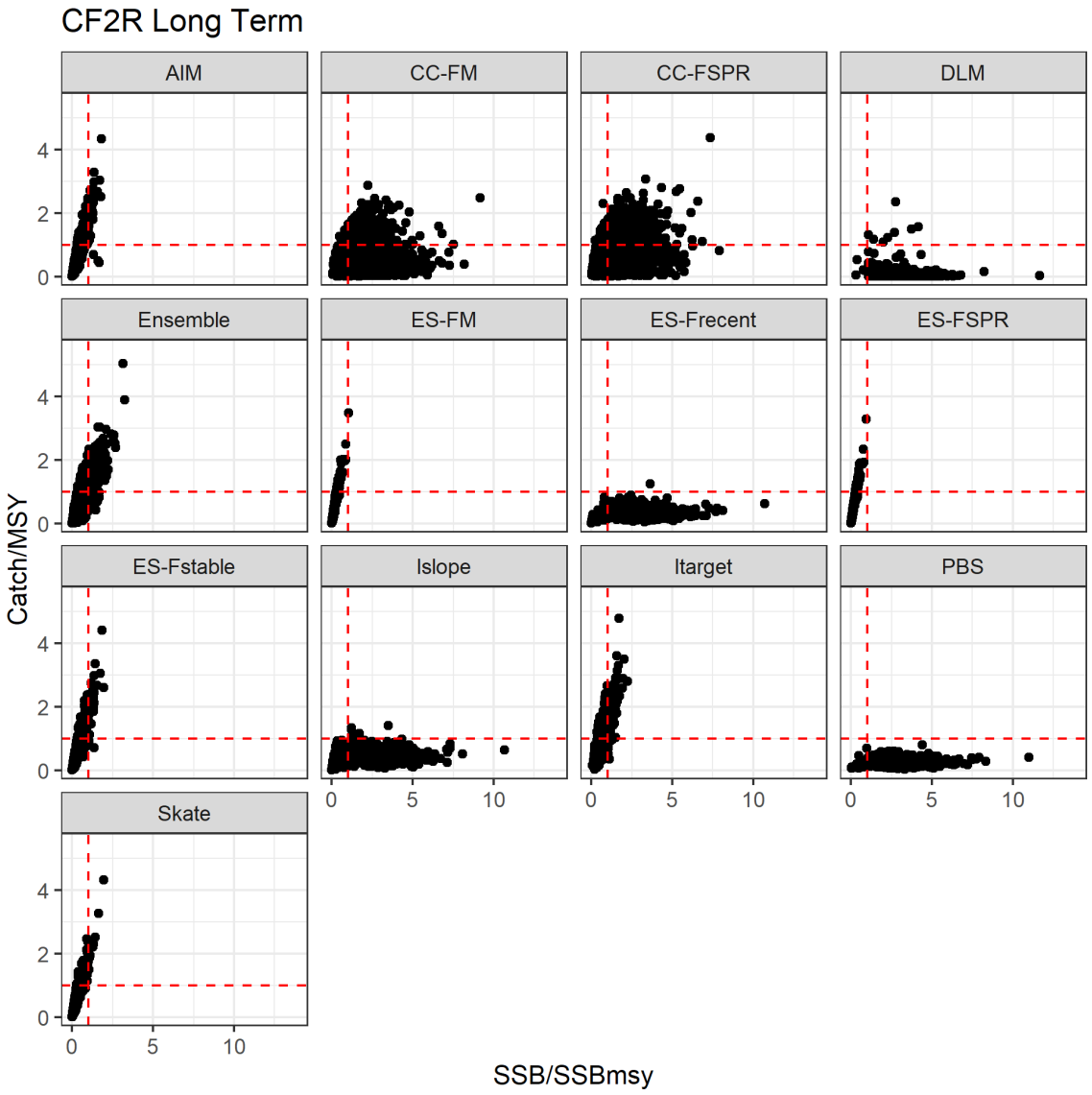


Figure A6.42. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

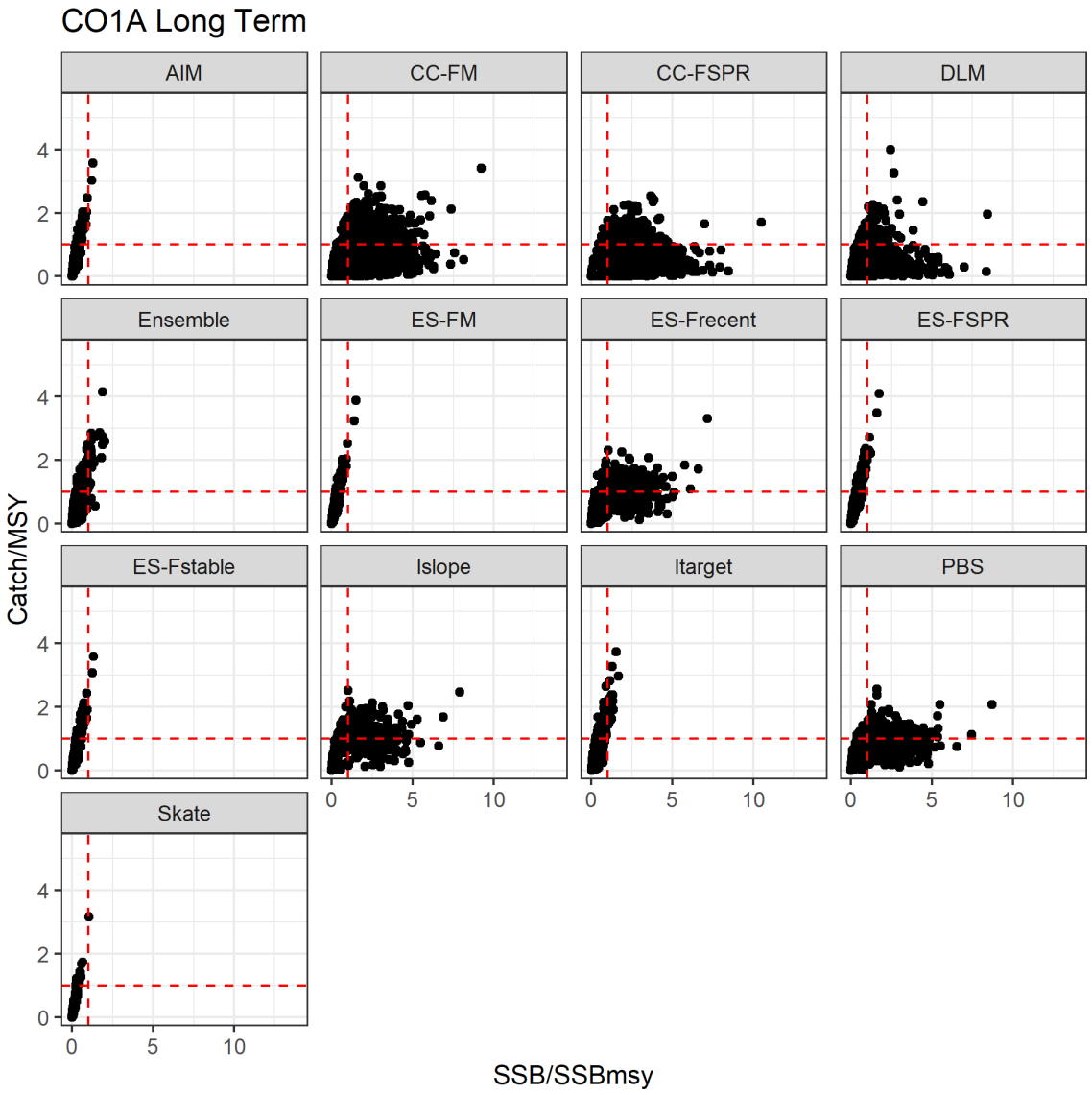


Figure A6.43. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

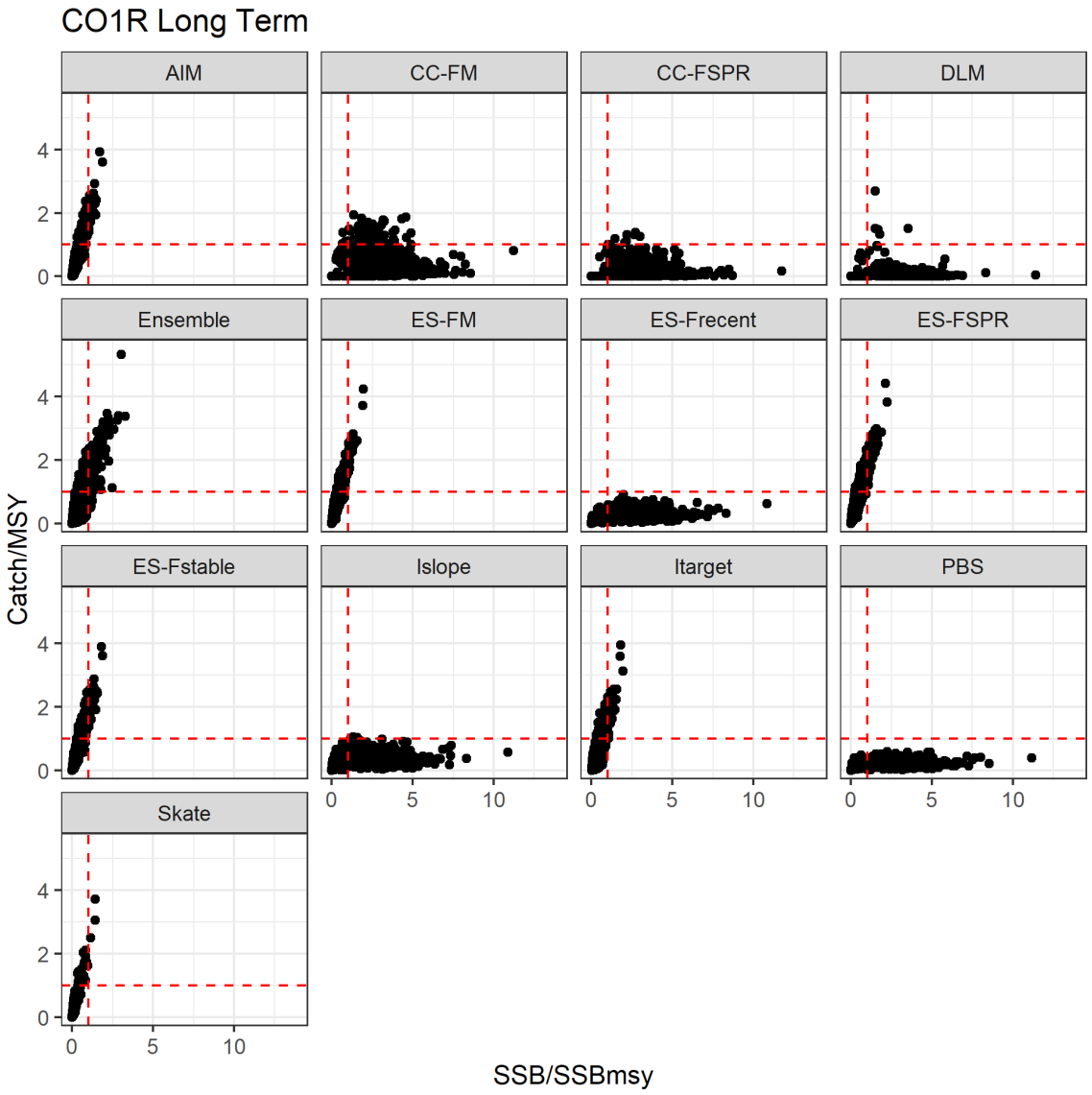


Figure A6.44. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

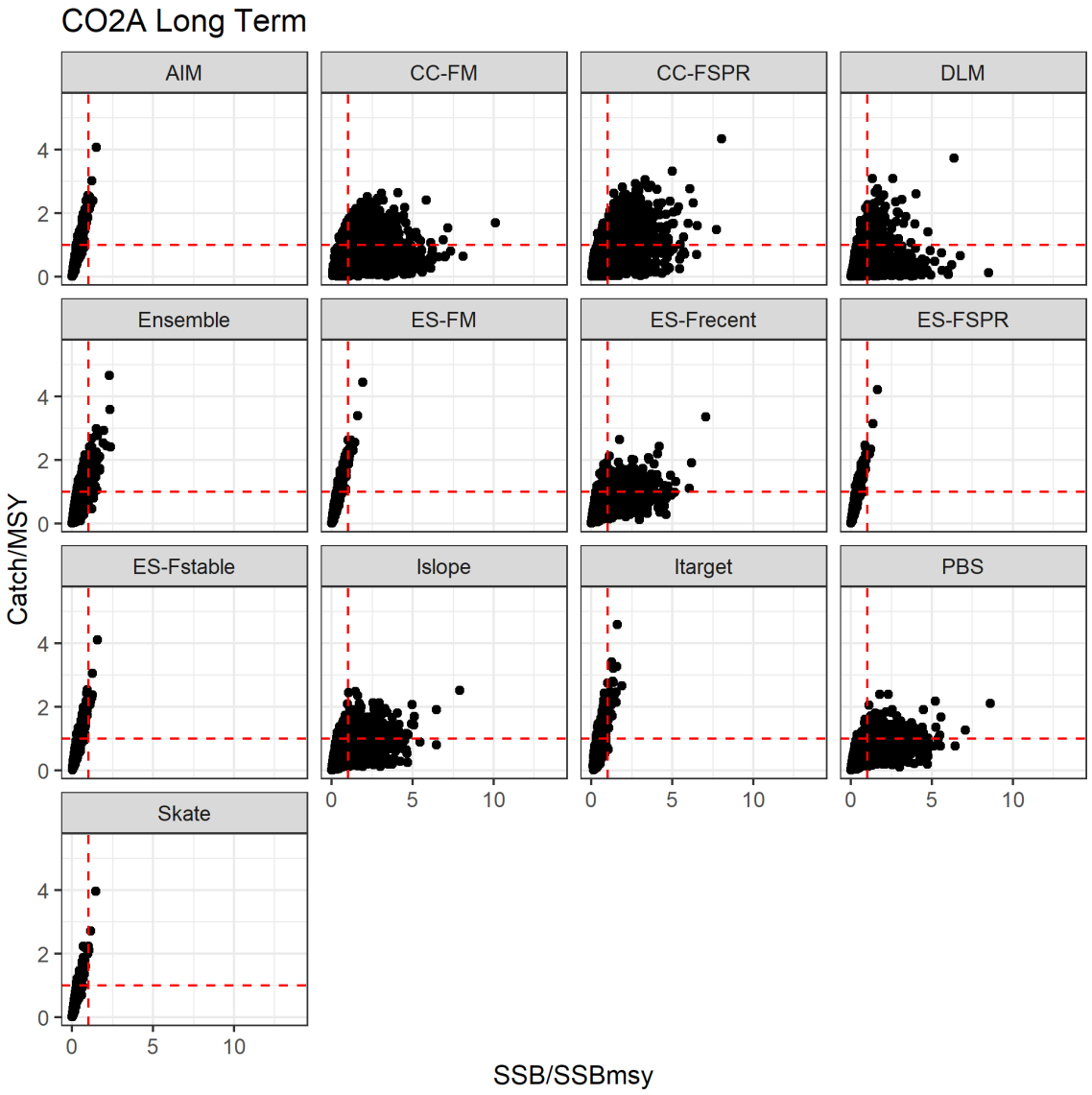


Figure A6.45. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

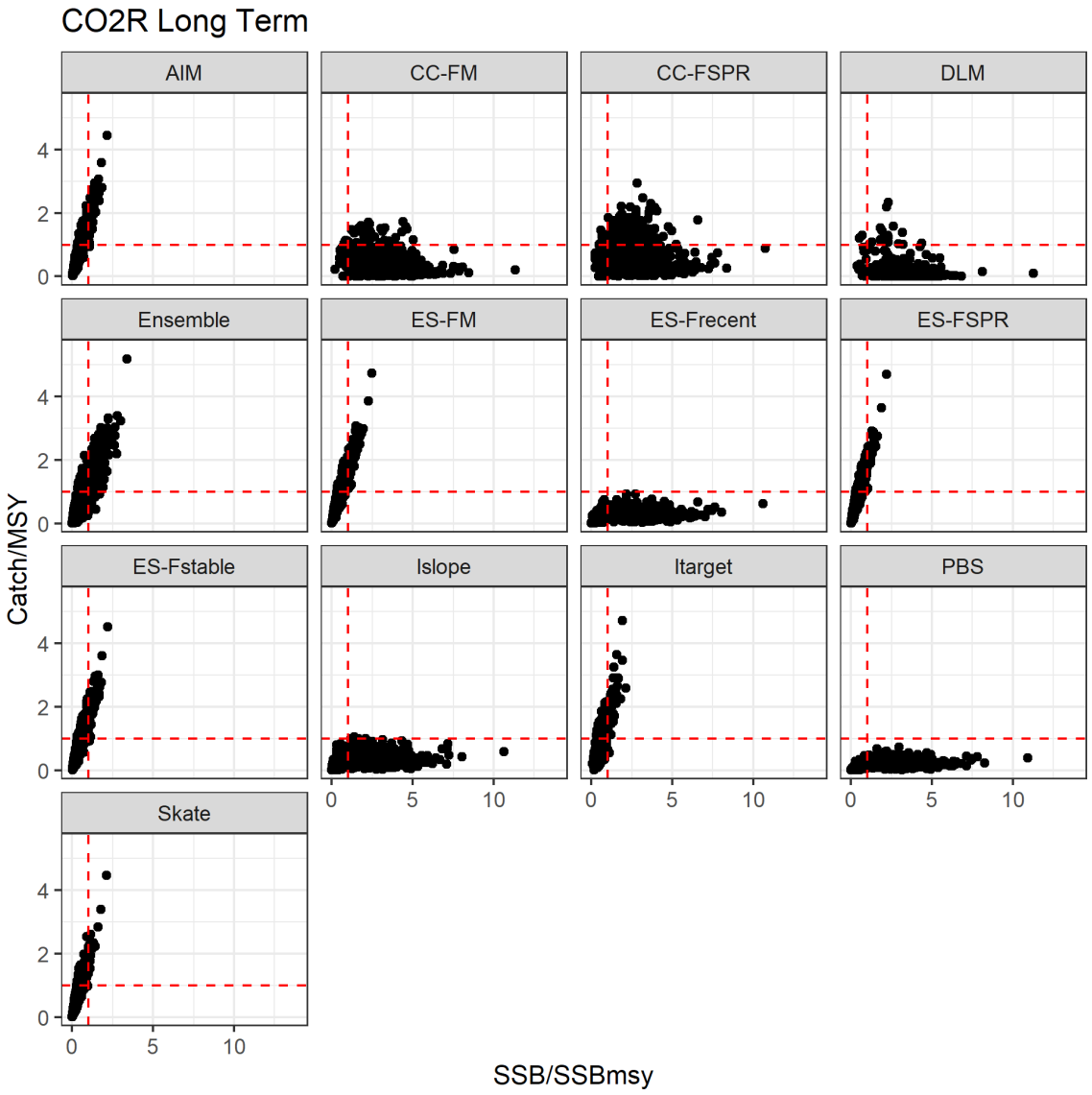


Figure A6.46. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

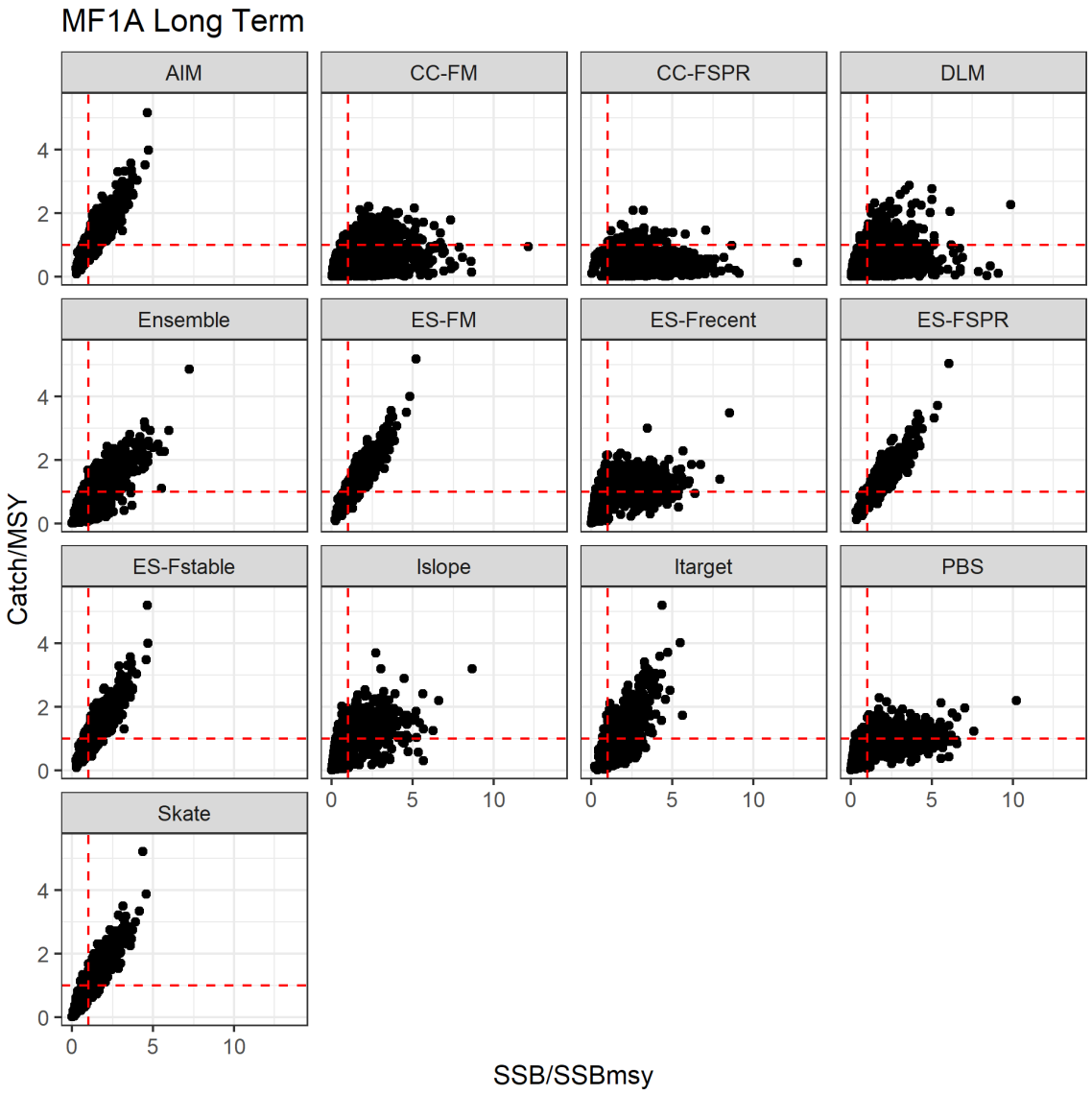


Figure A6.47. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

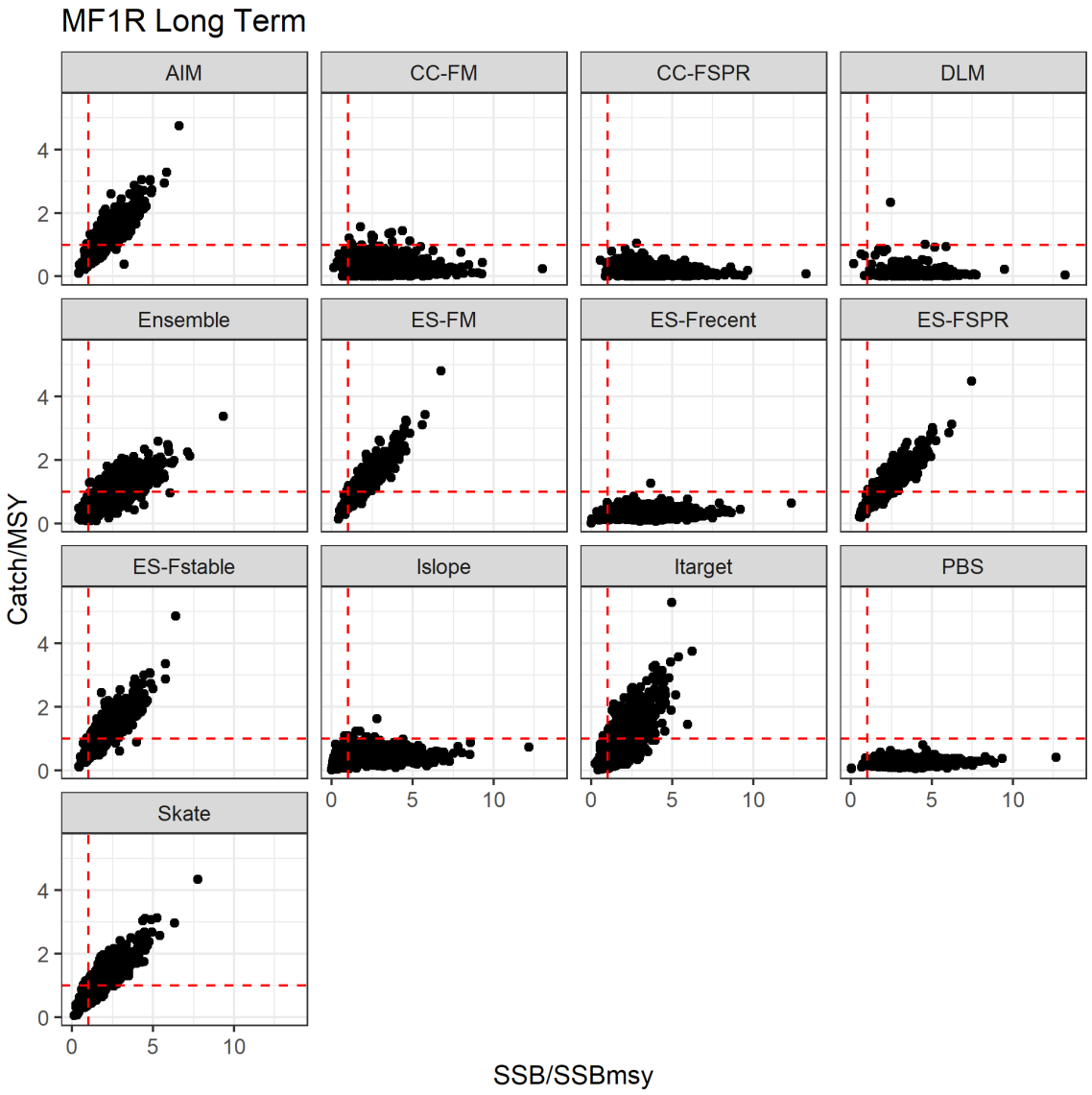


Figure A6.48. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

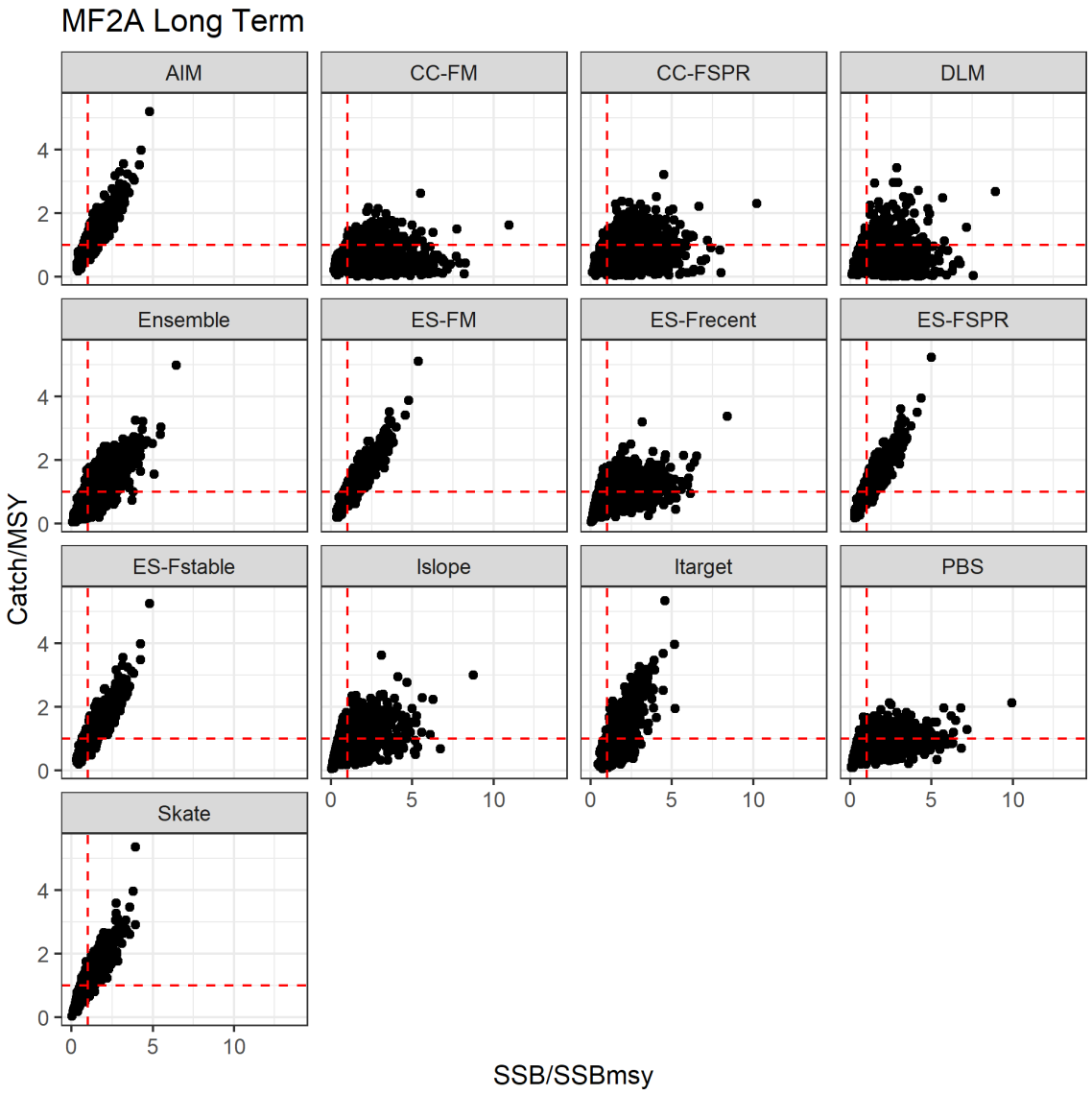


Figure A6.49. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

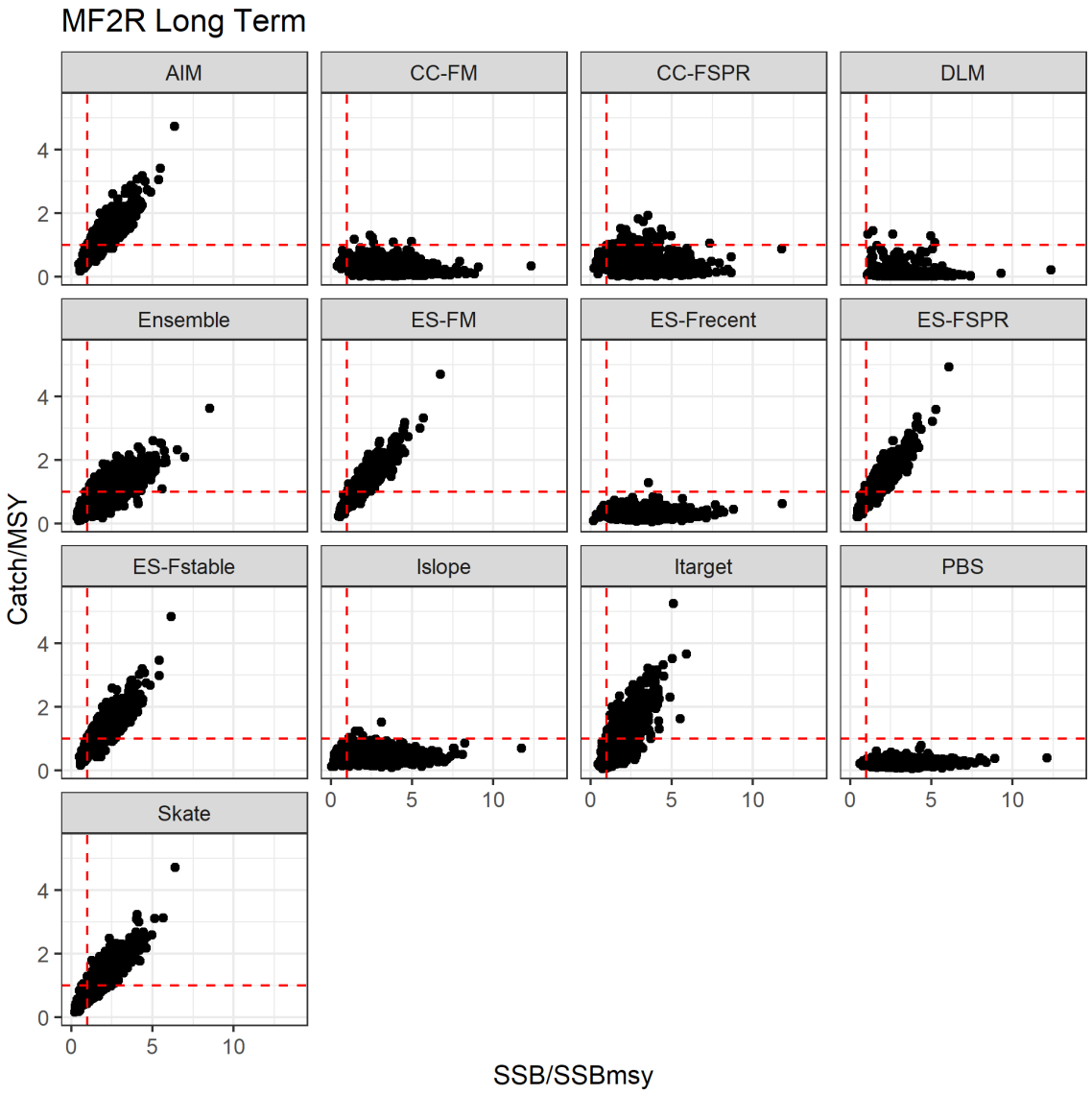


Figure A6.50. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

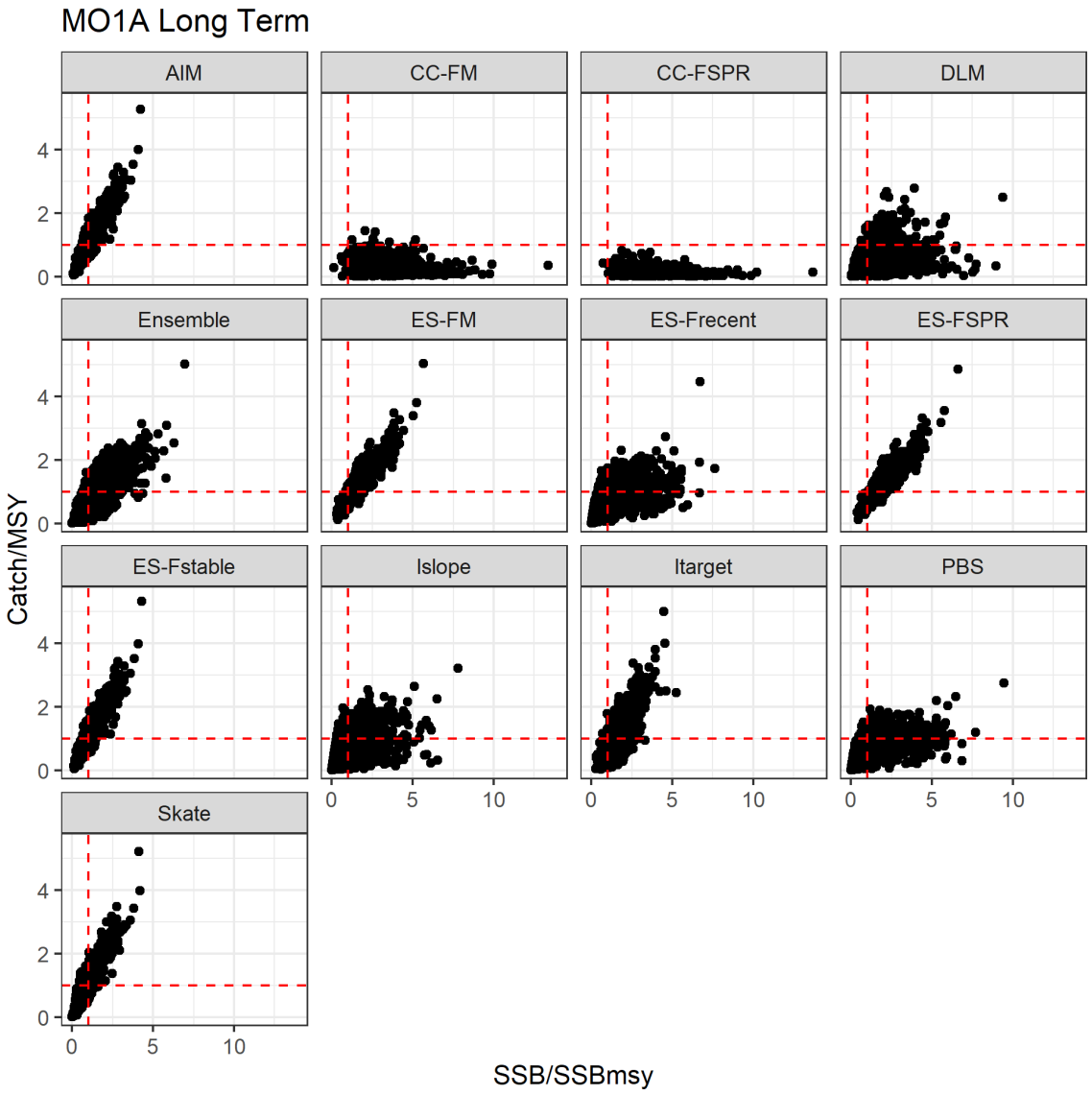


Figure A6.51. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

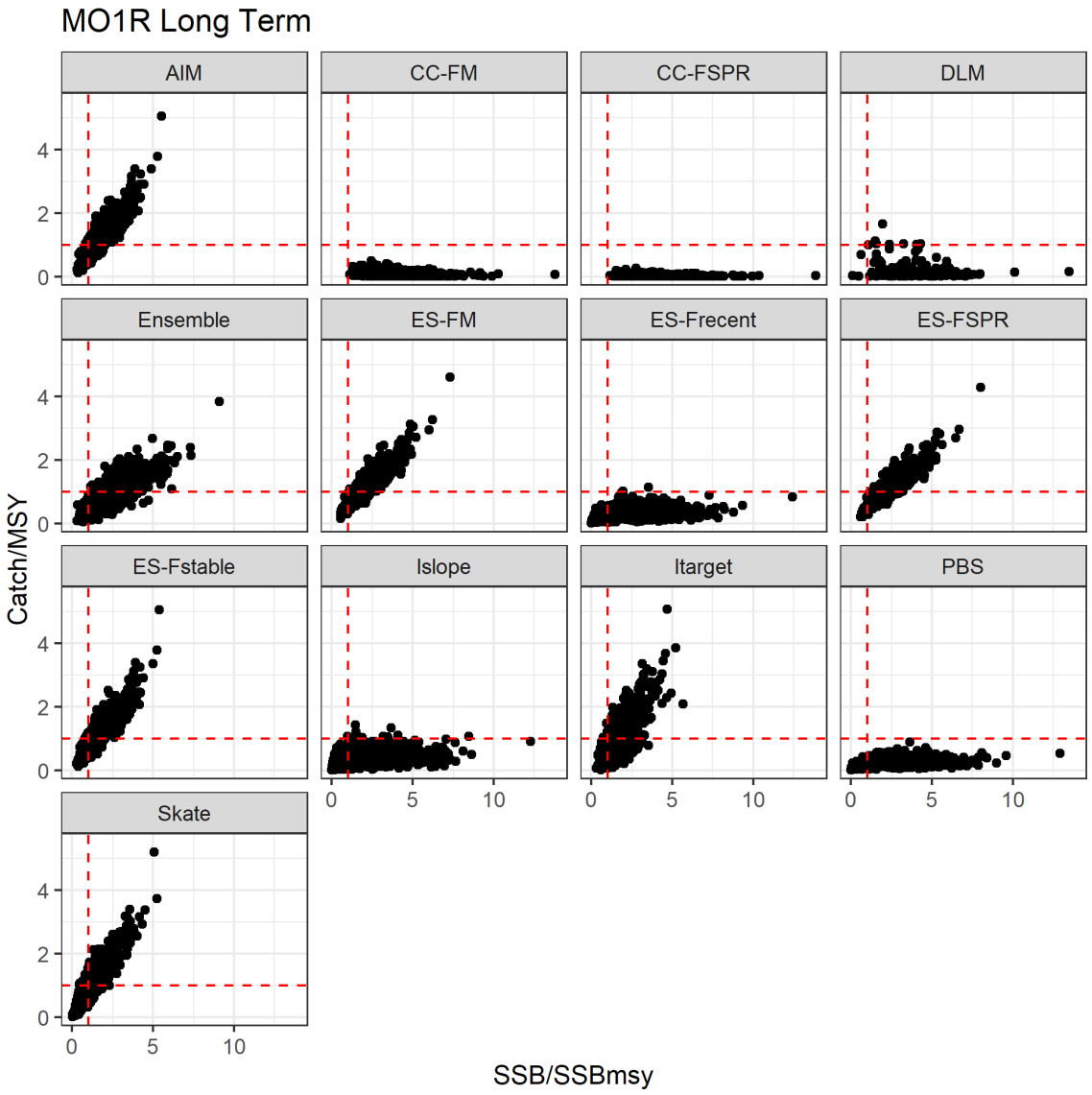


Figure A6.52. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

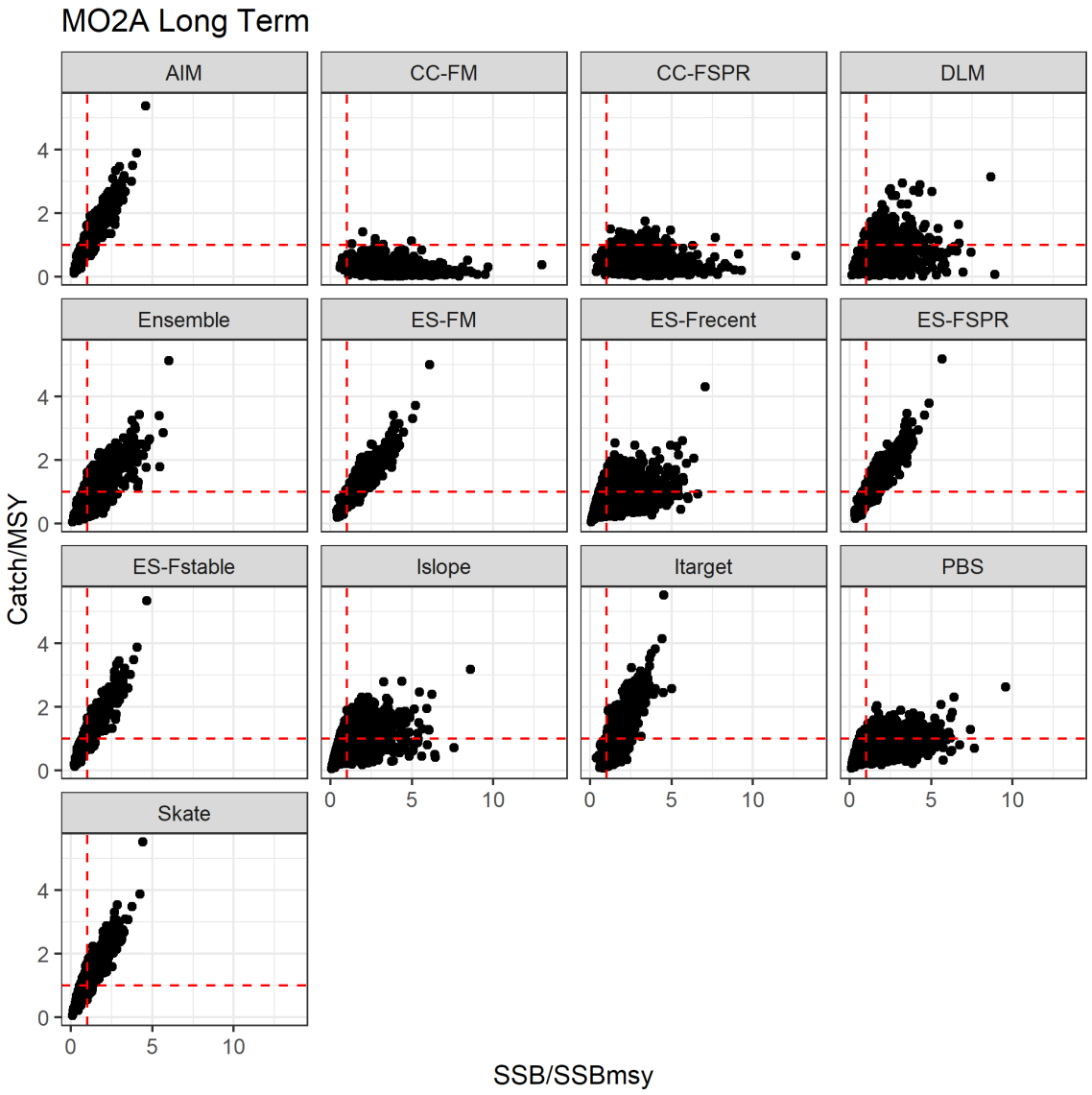


Figure A6.53. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

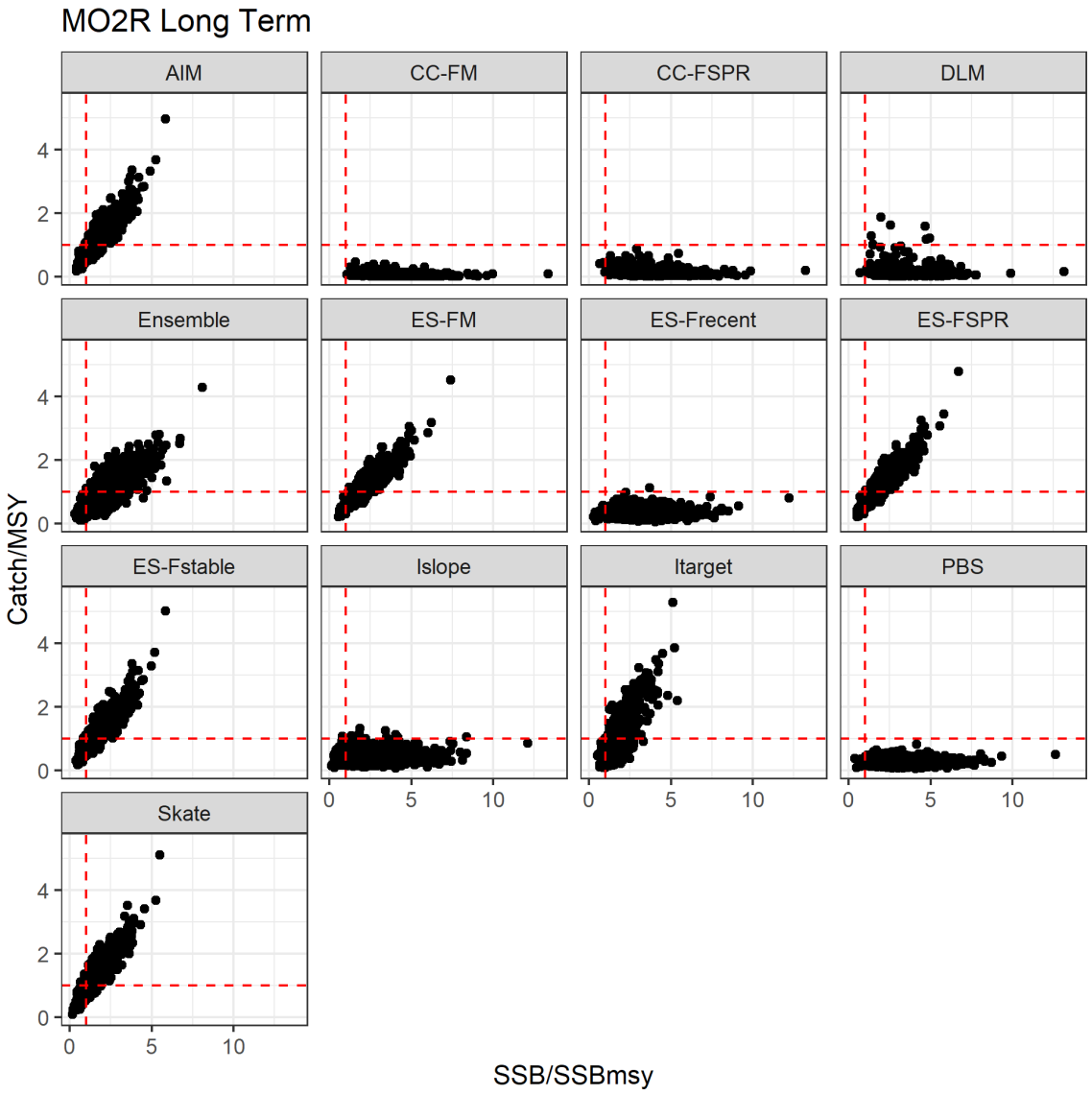


Figure A6.54. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

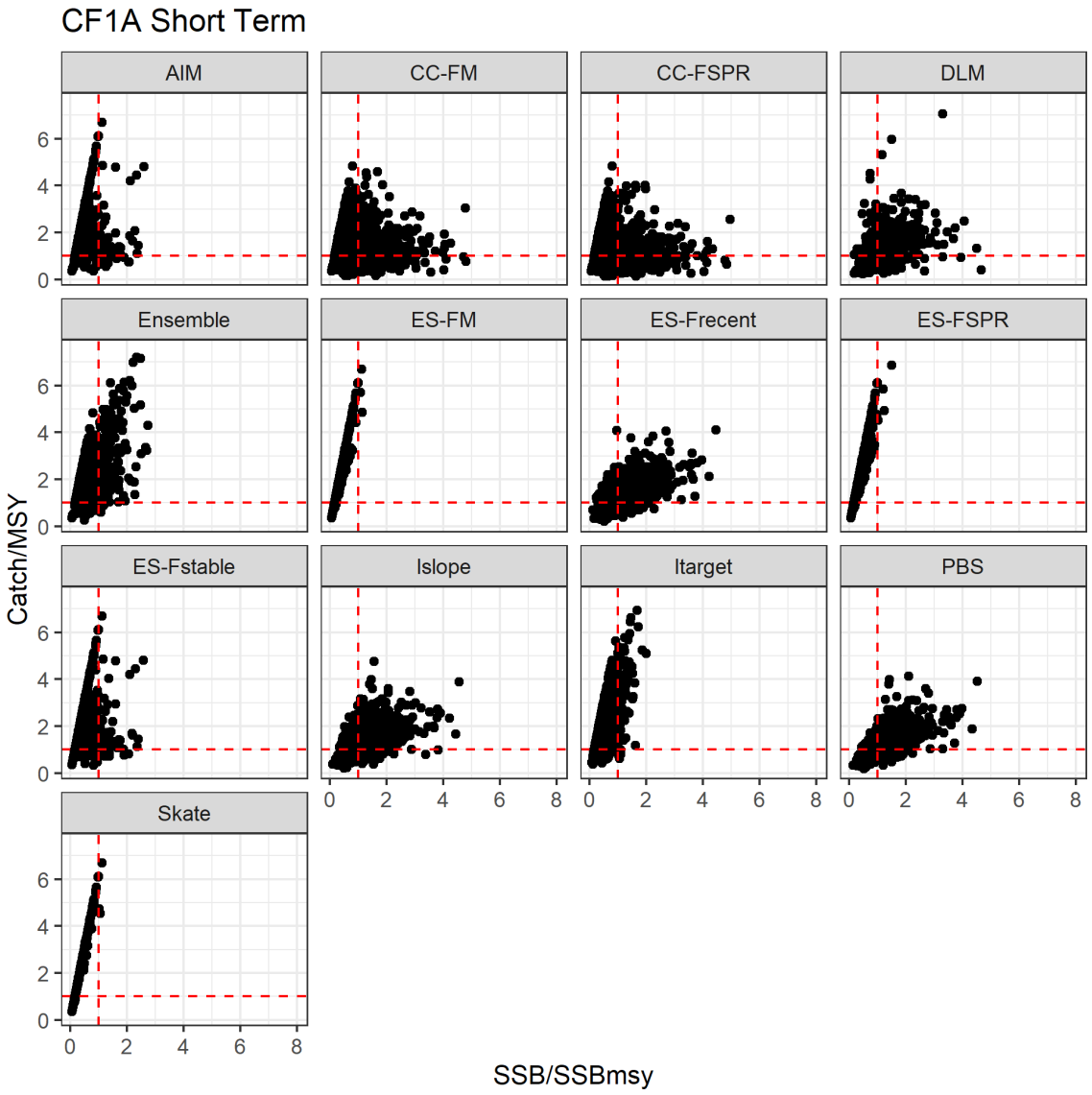


Figure A6.55. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

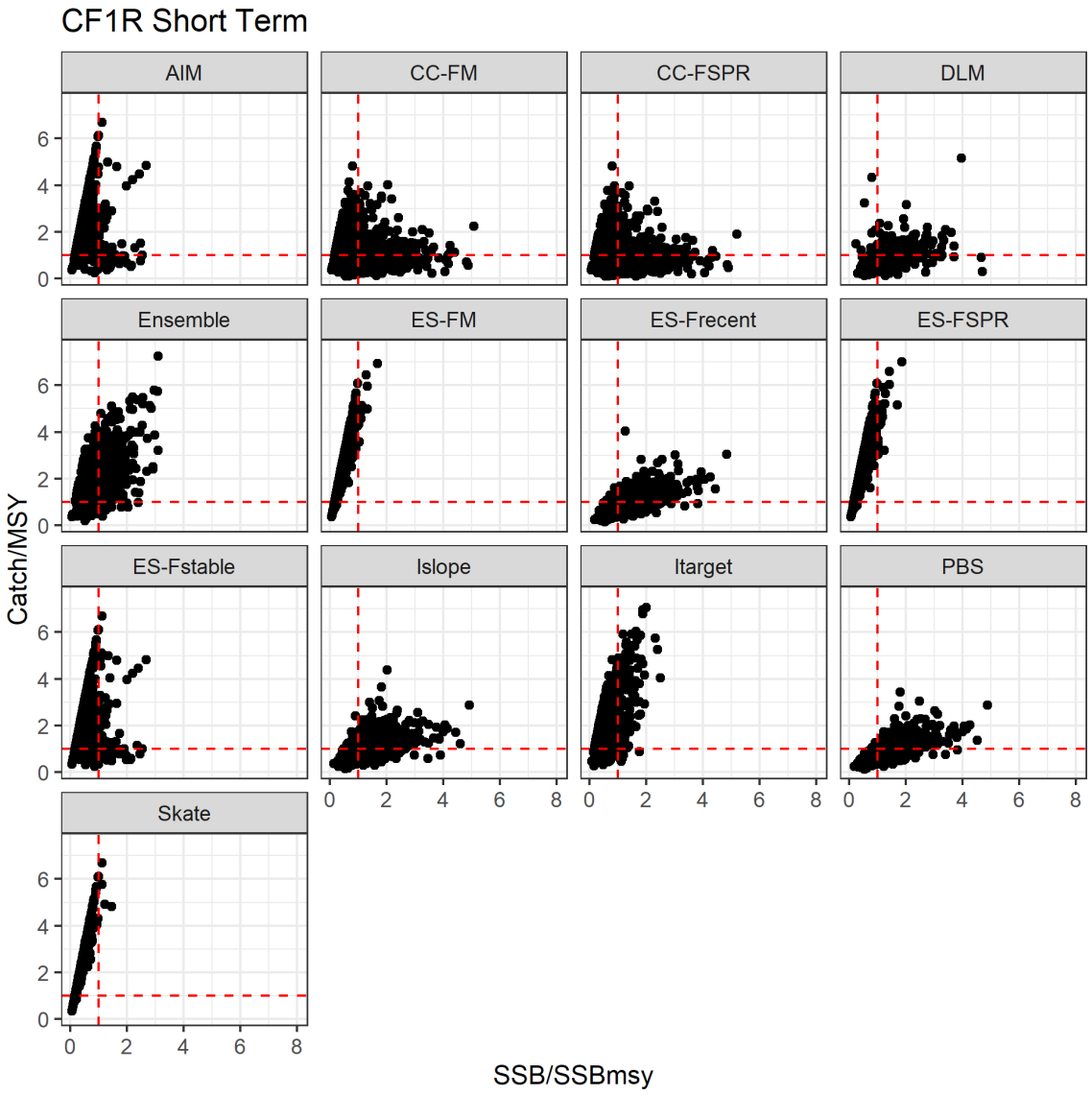


Figure A6.56. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

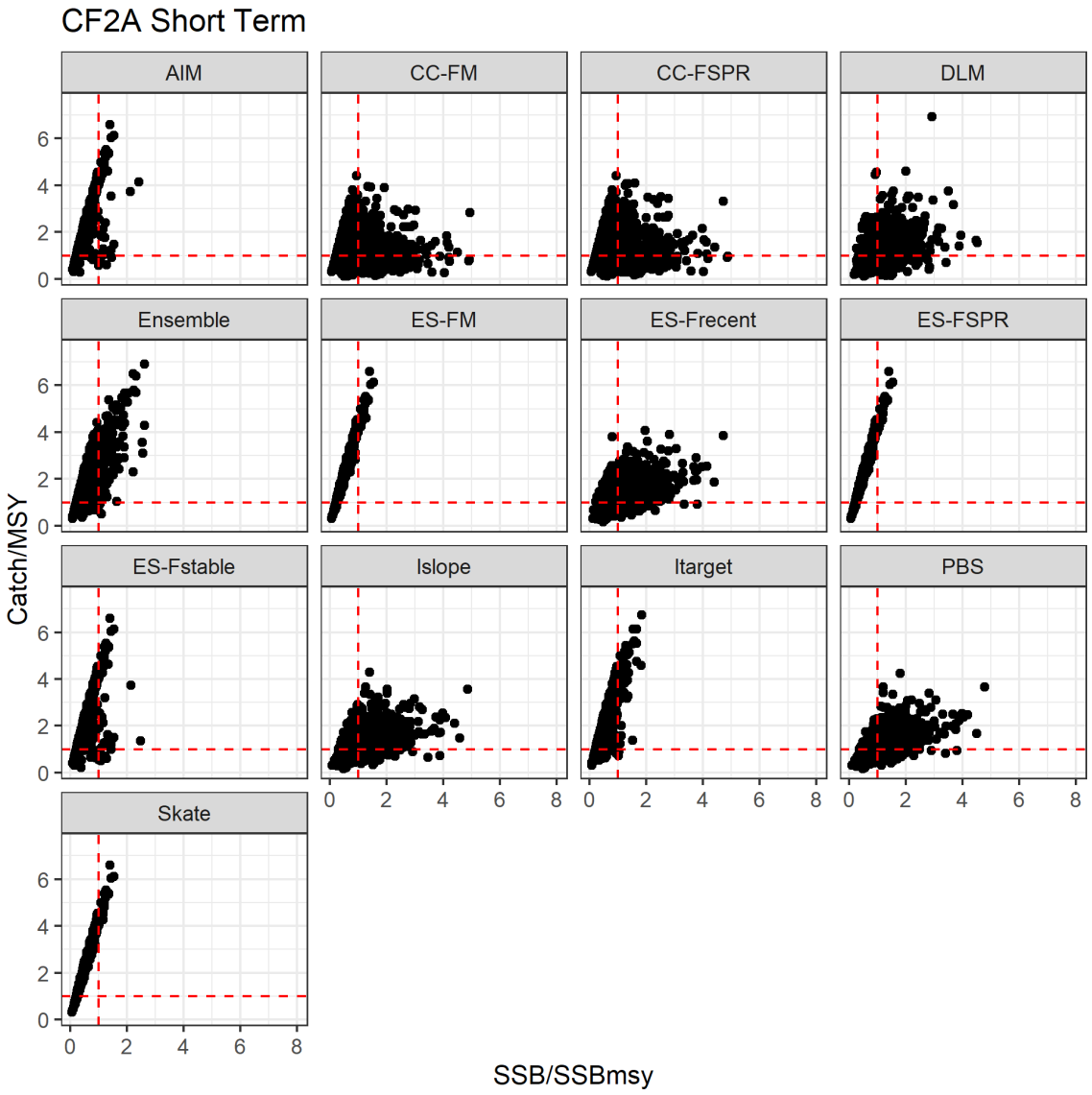


Figure A6.57. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

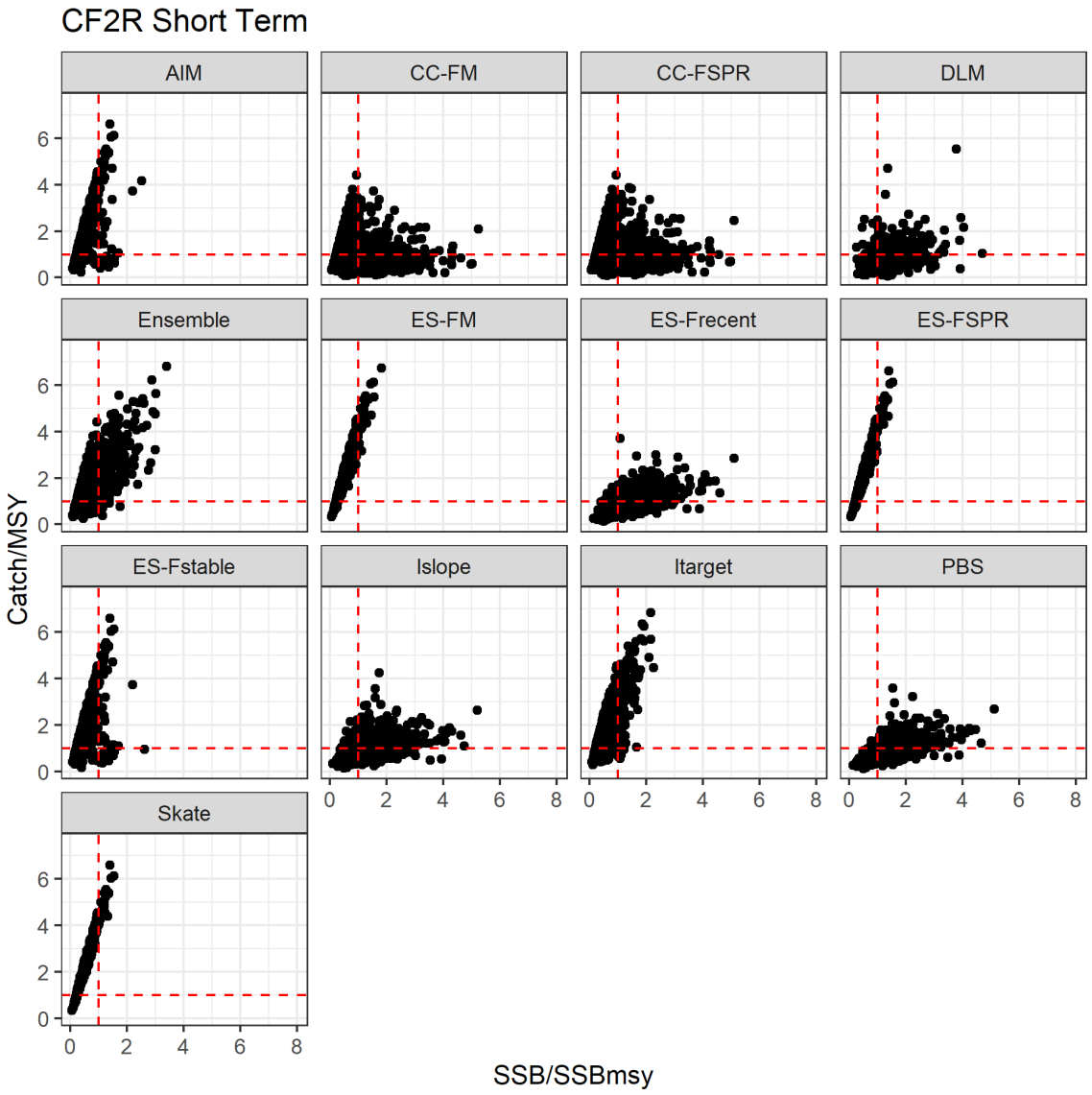


Figure A6.58. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

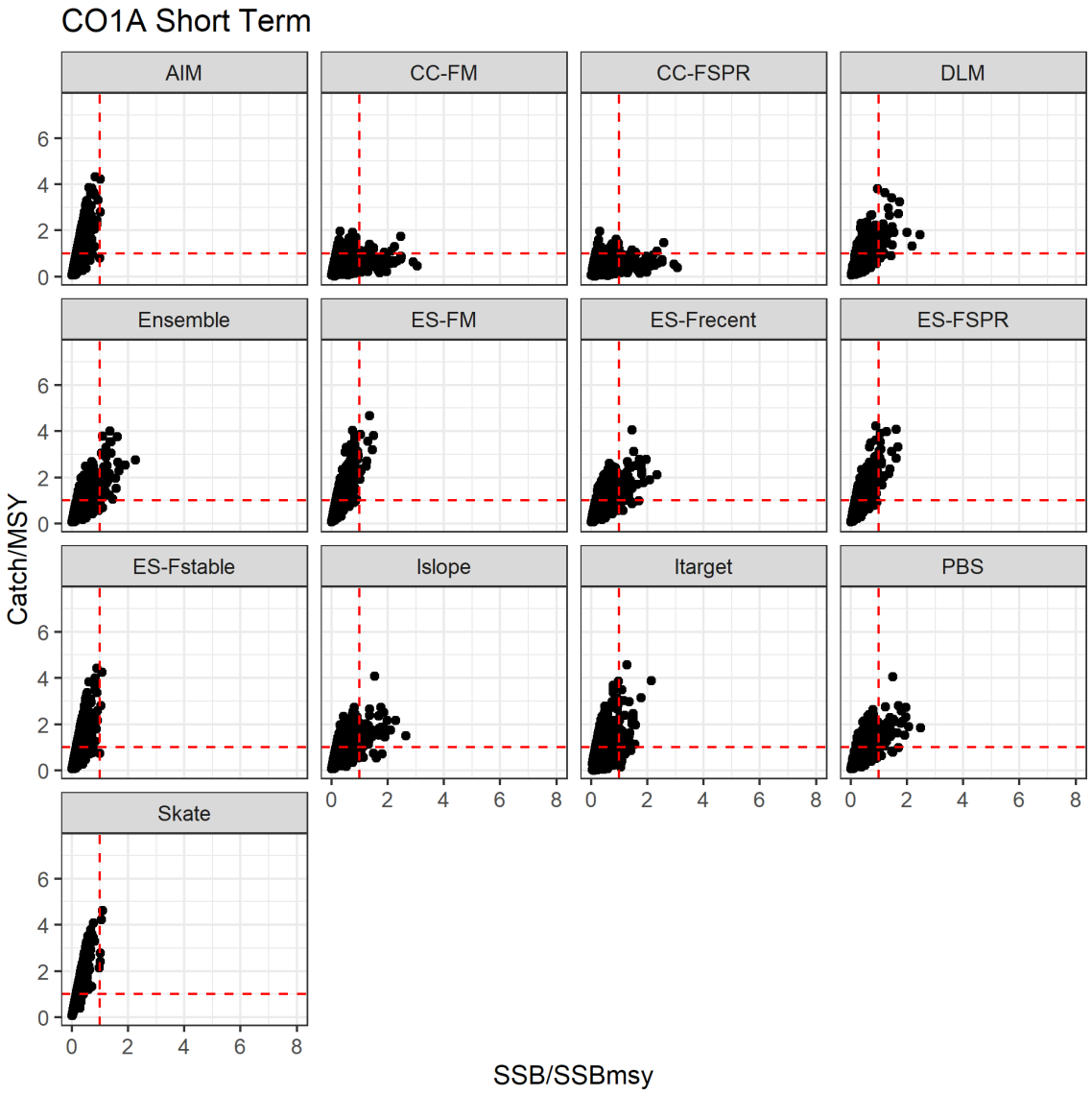


Figure A6.59. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

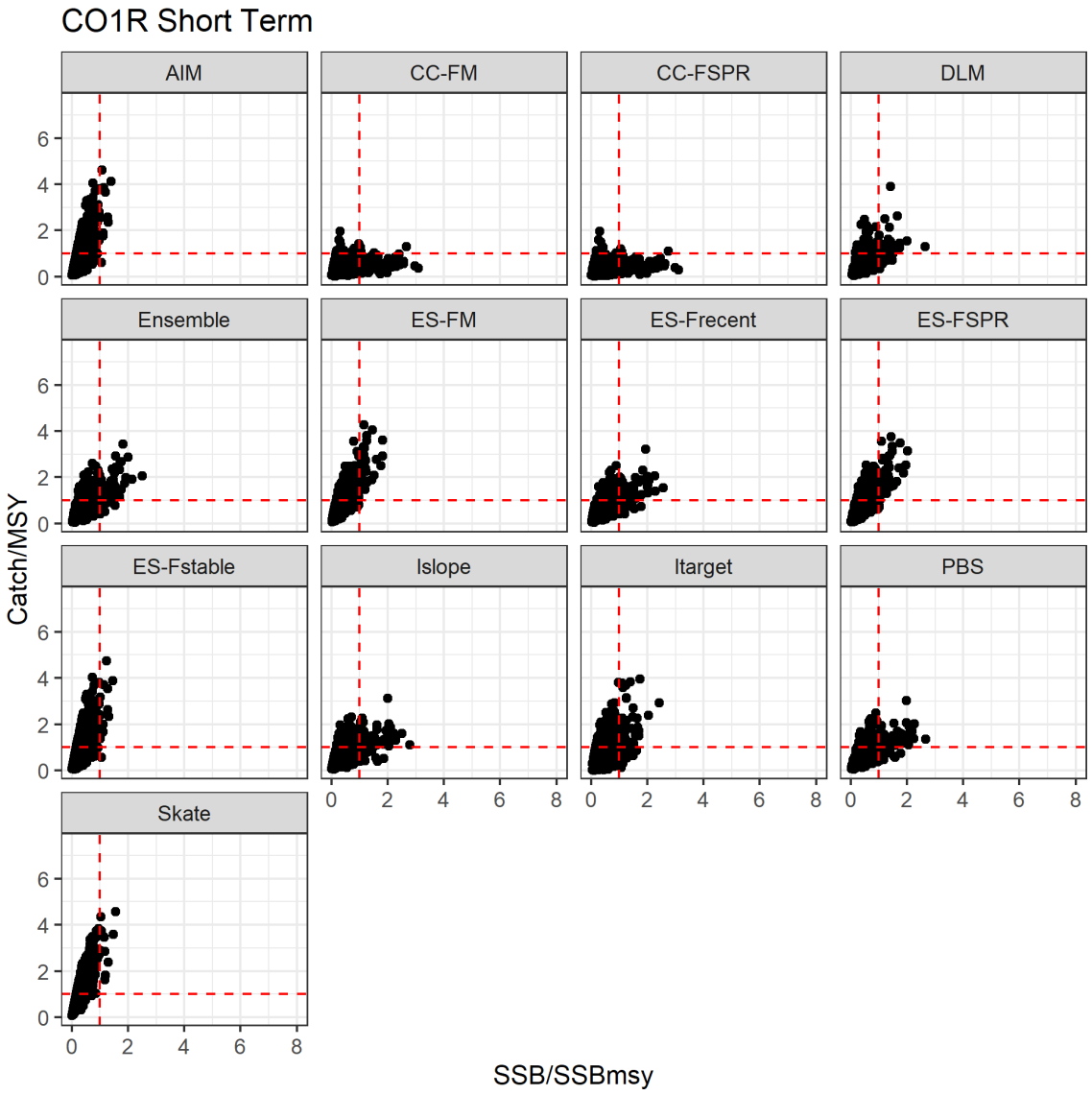


Figure A6.60. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

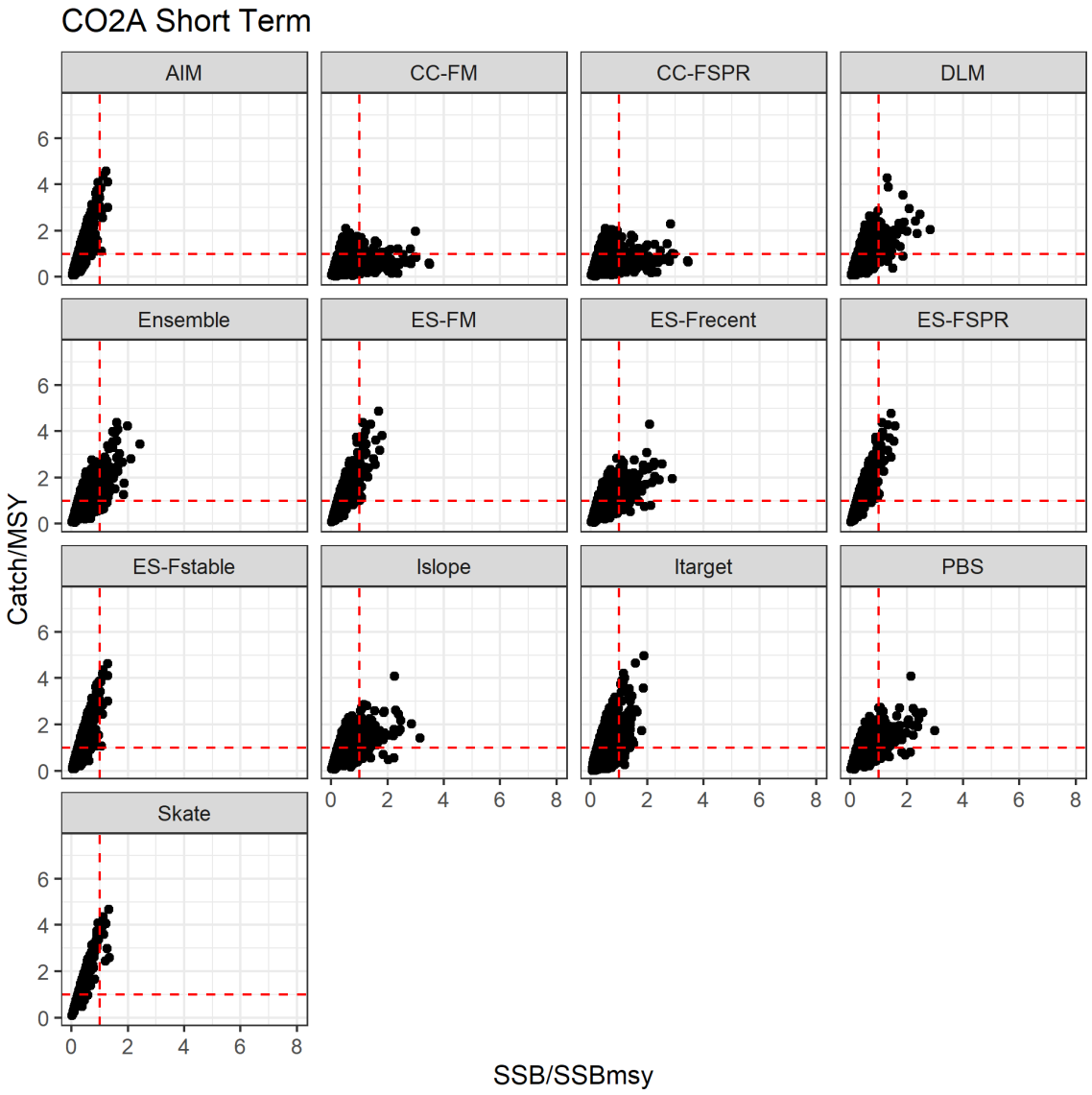


Figure A6.61. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

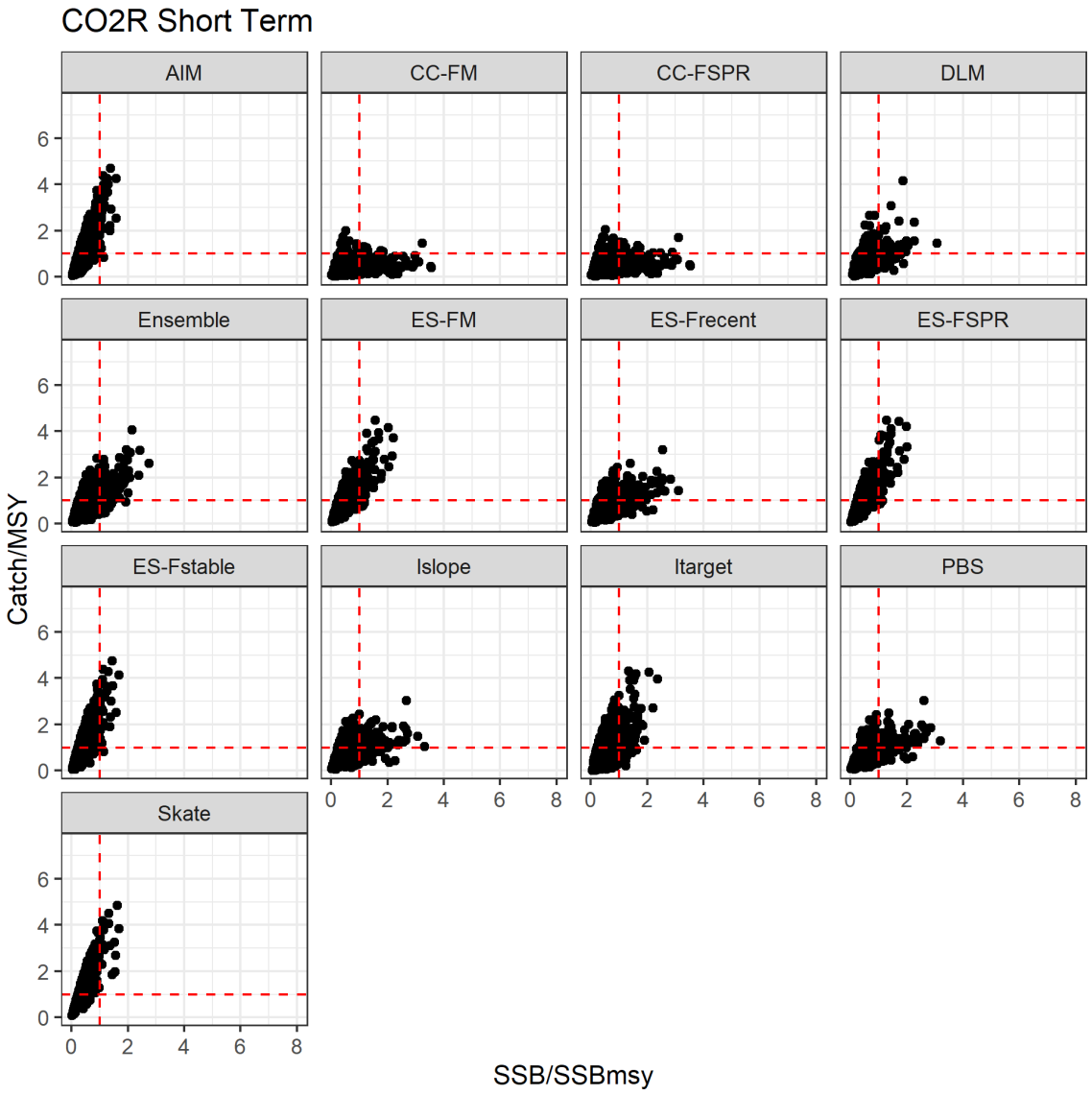


Figure A6.62. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

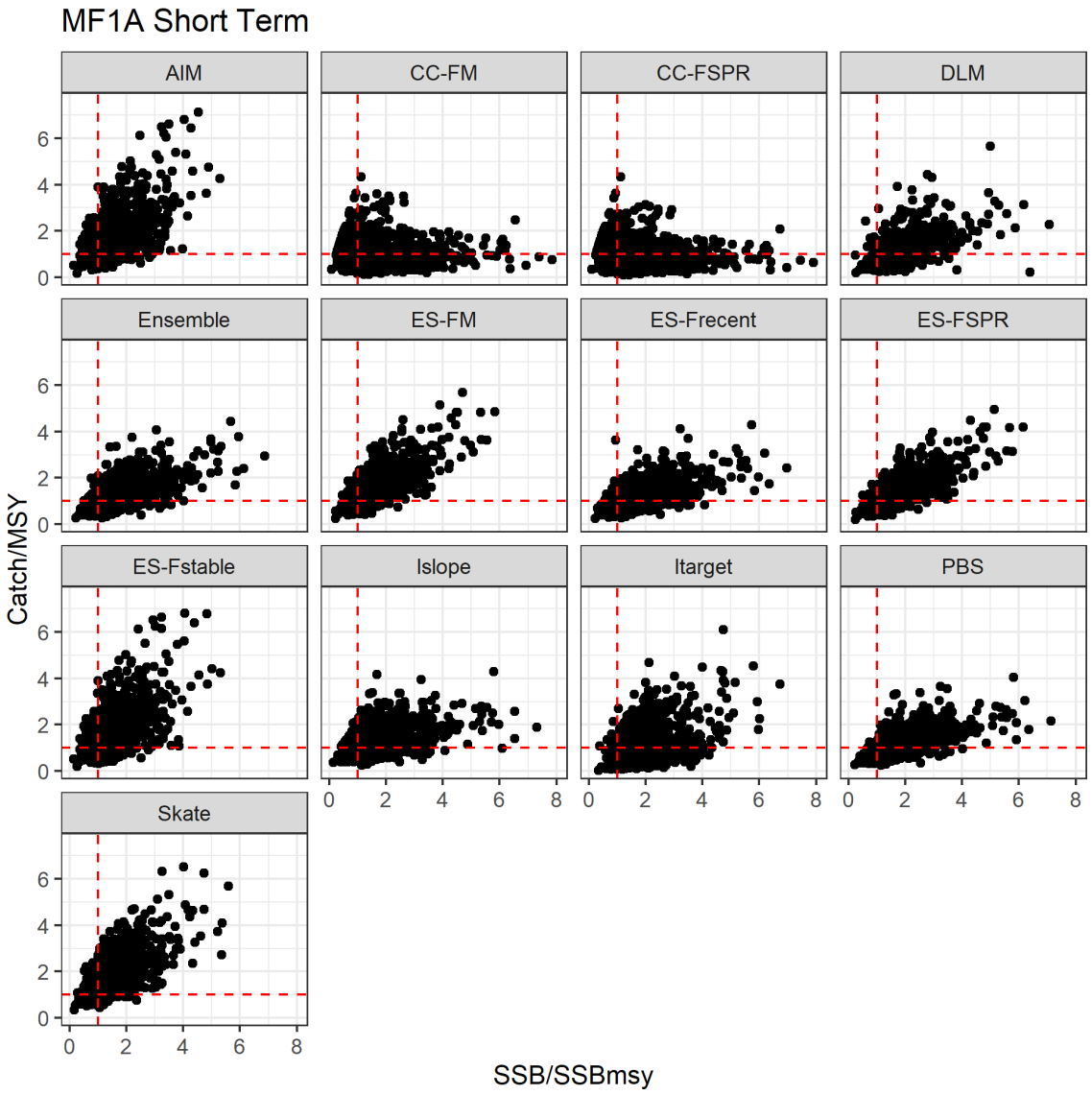


Figure A6.63. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

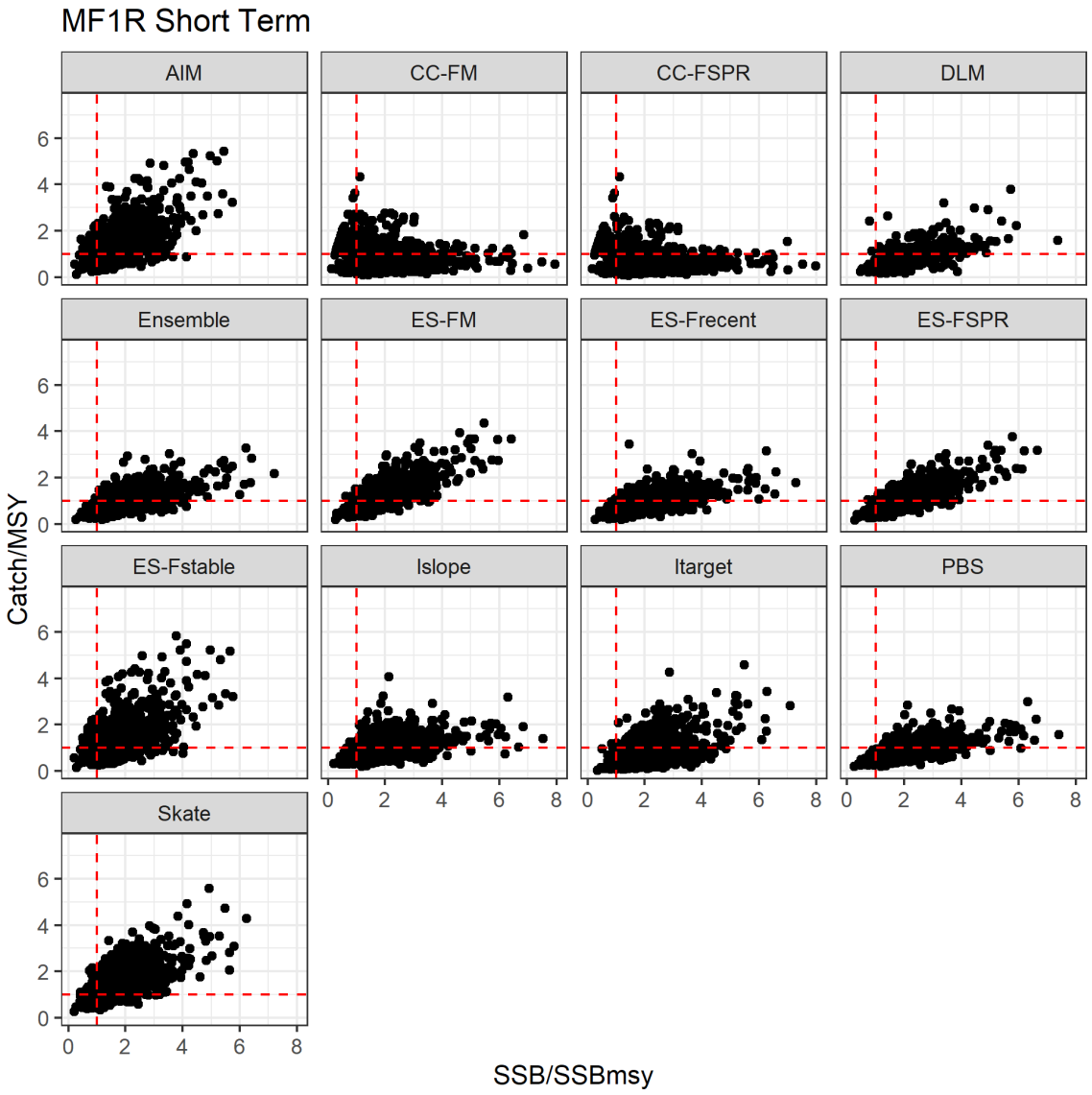


Figure A6.64. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

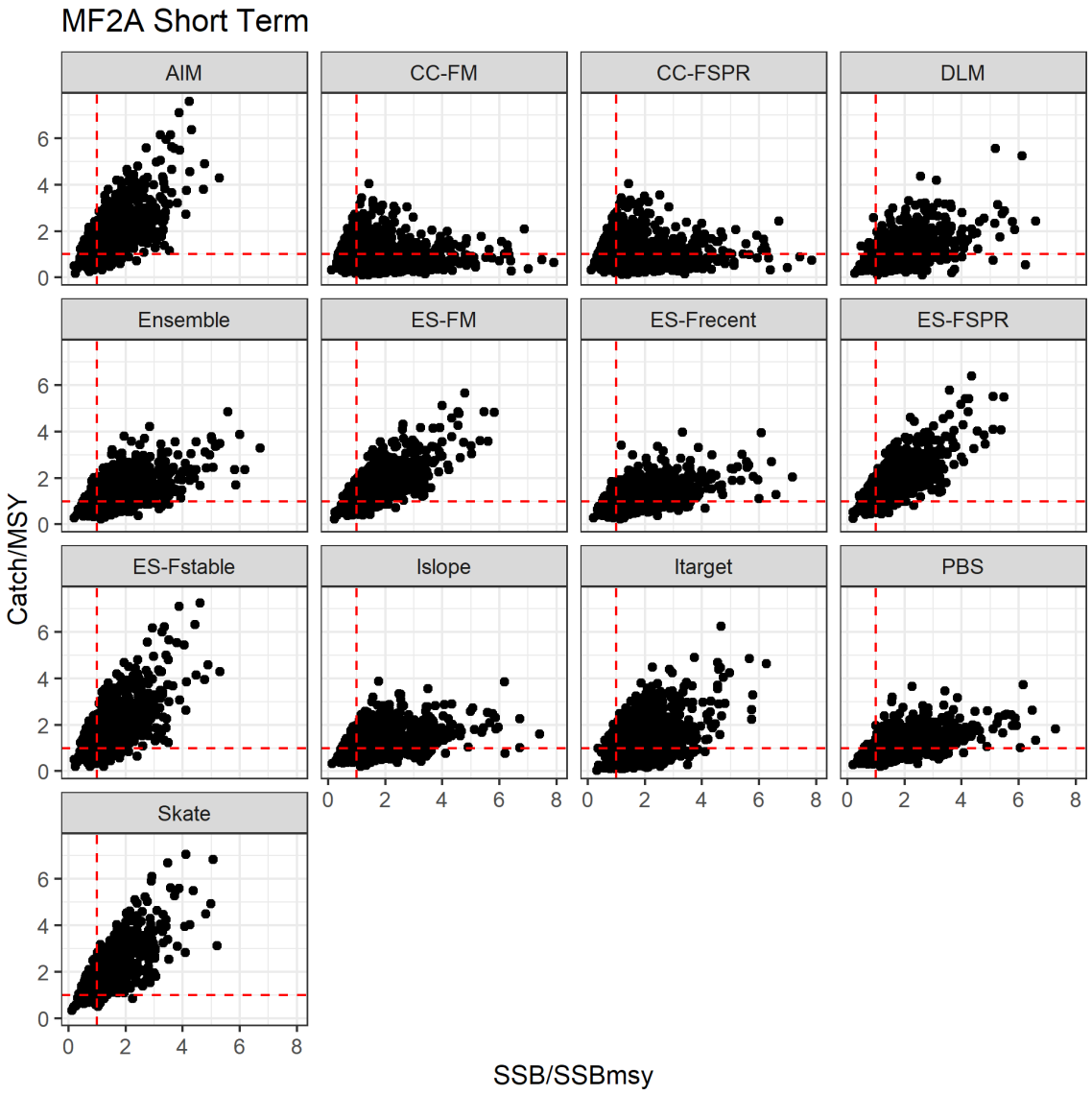


Figure A6.65. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

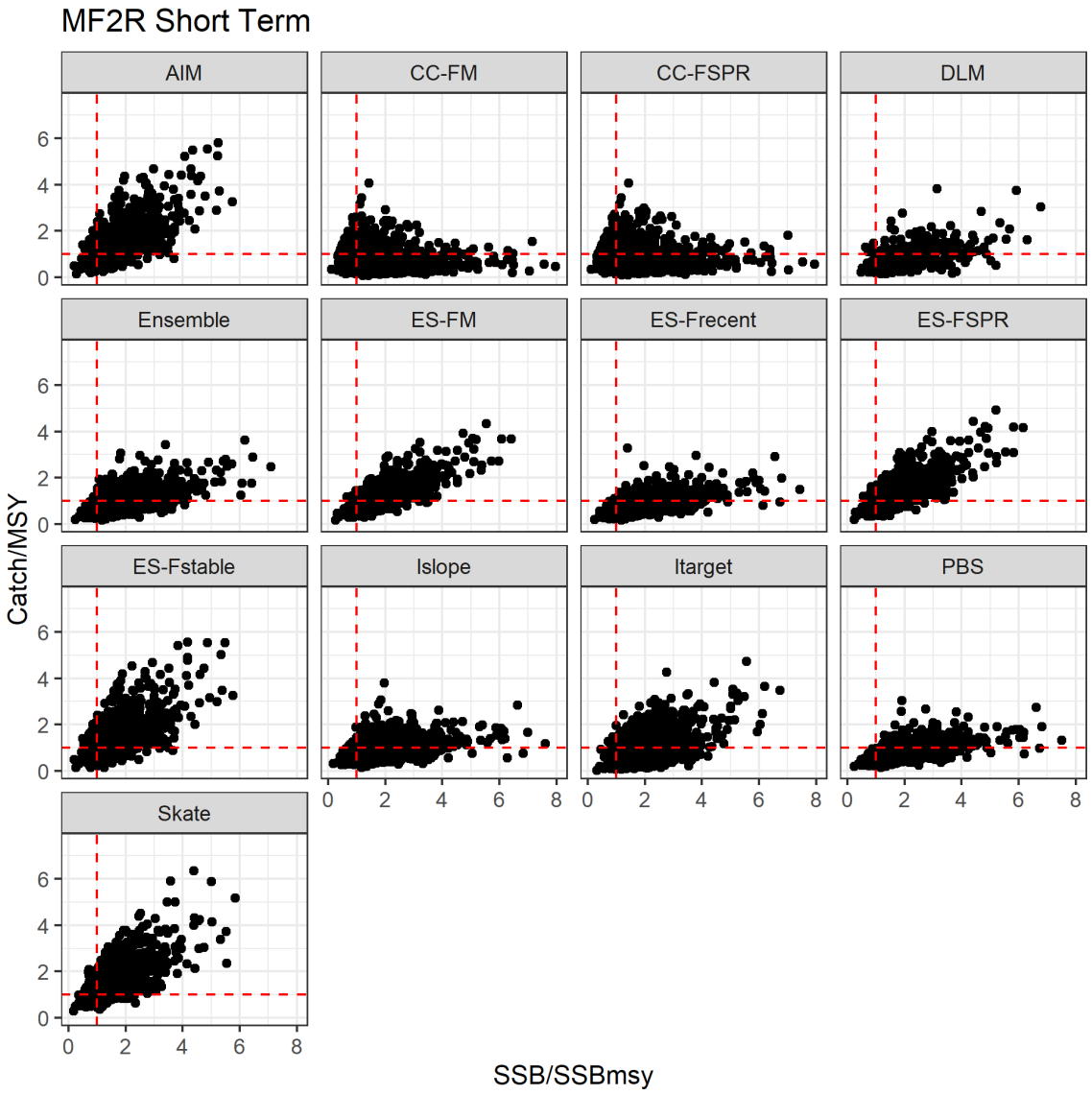


Figure A6.66. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

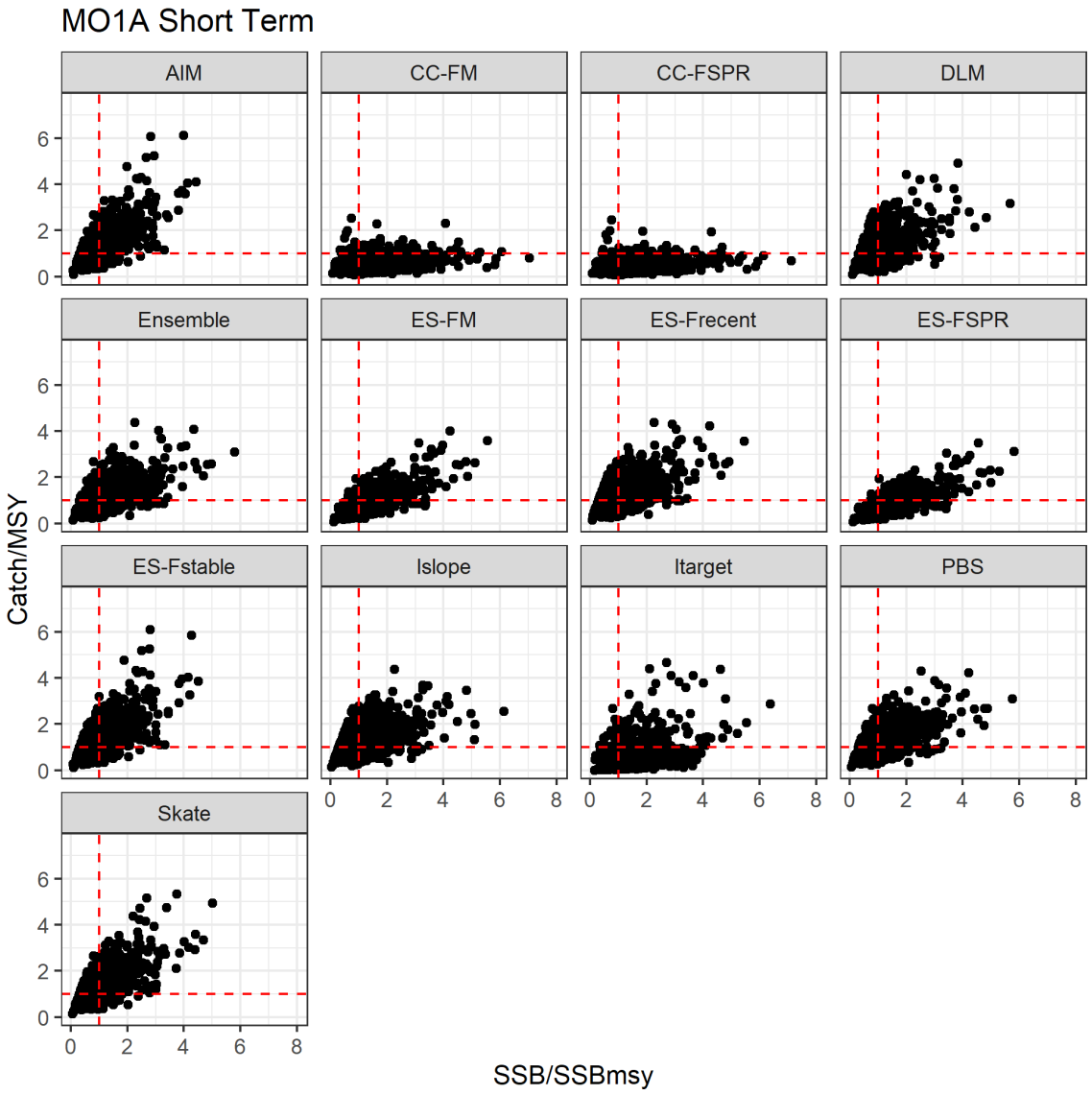


Figure A6.67. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

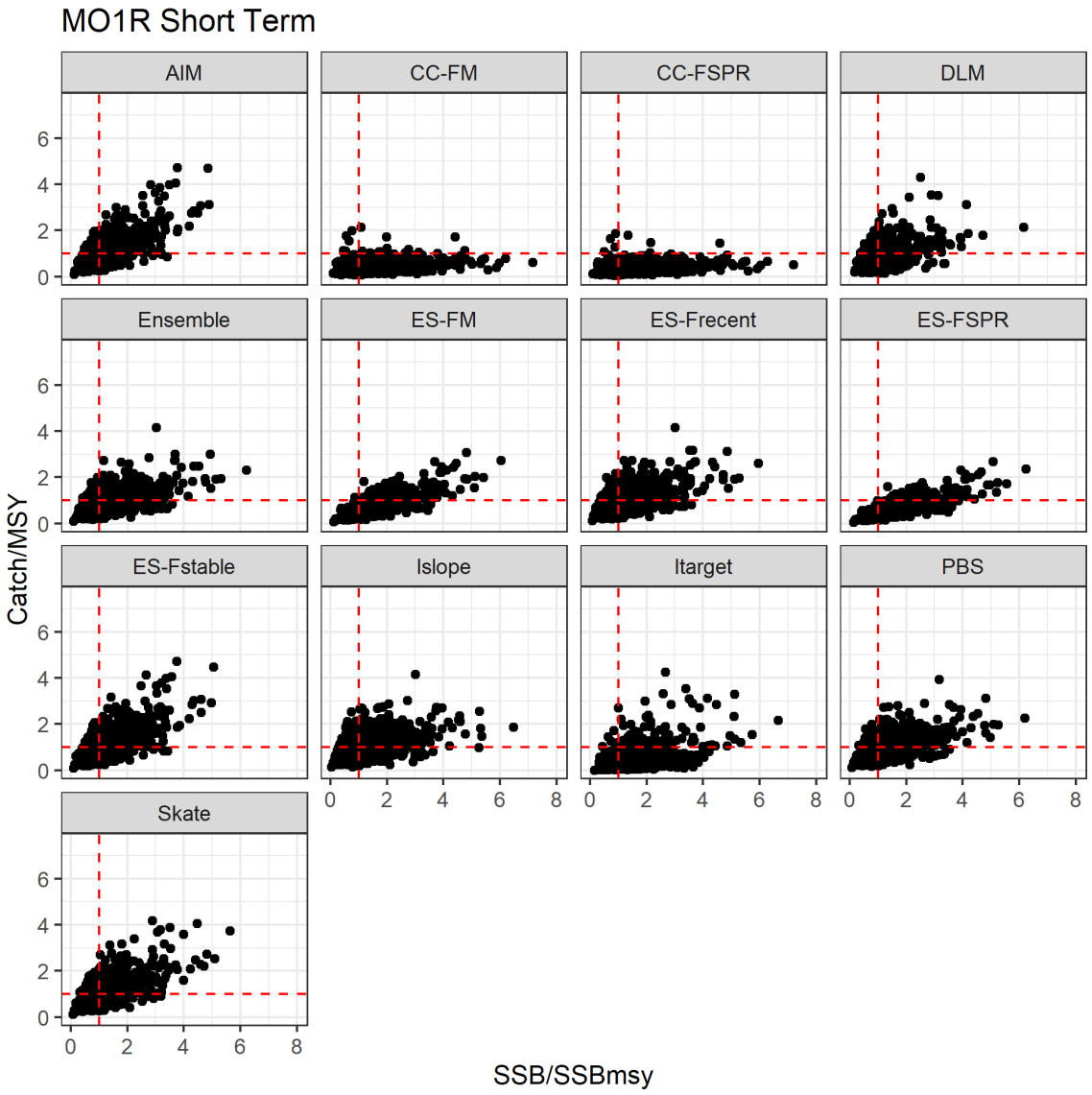


Figure A6.68. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

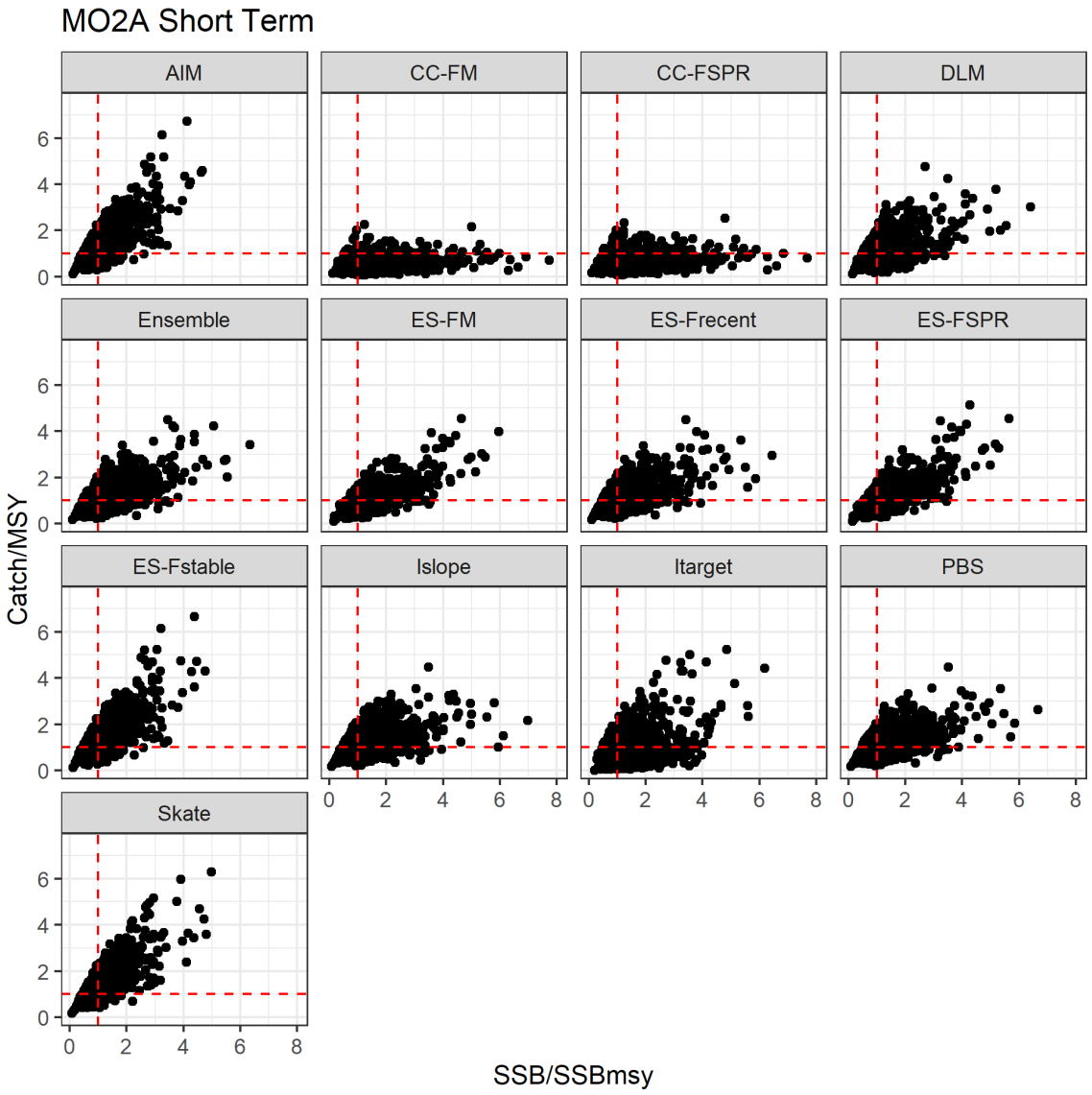


Figure A6.69. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

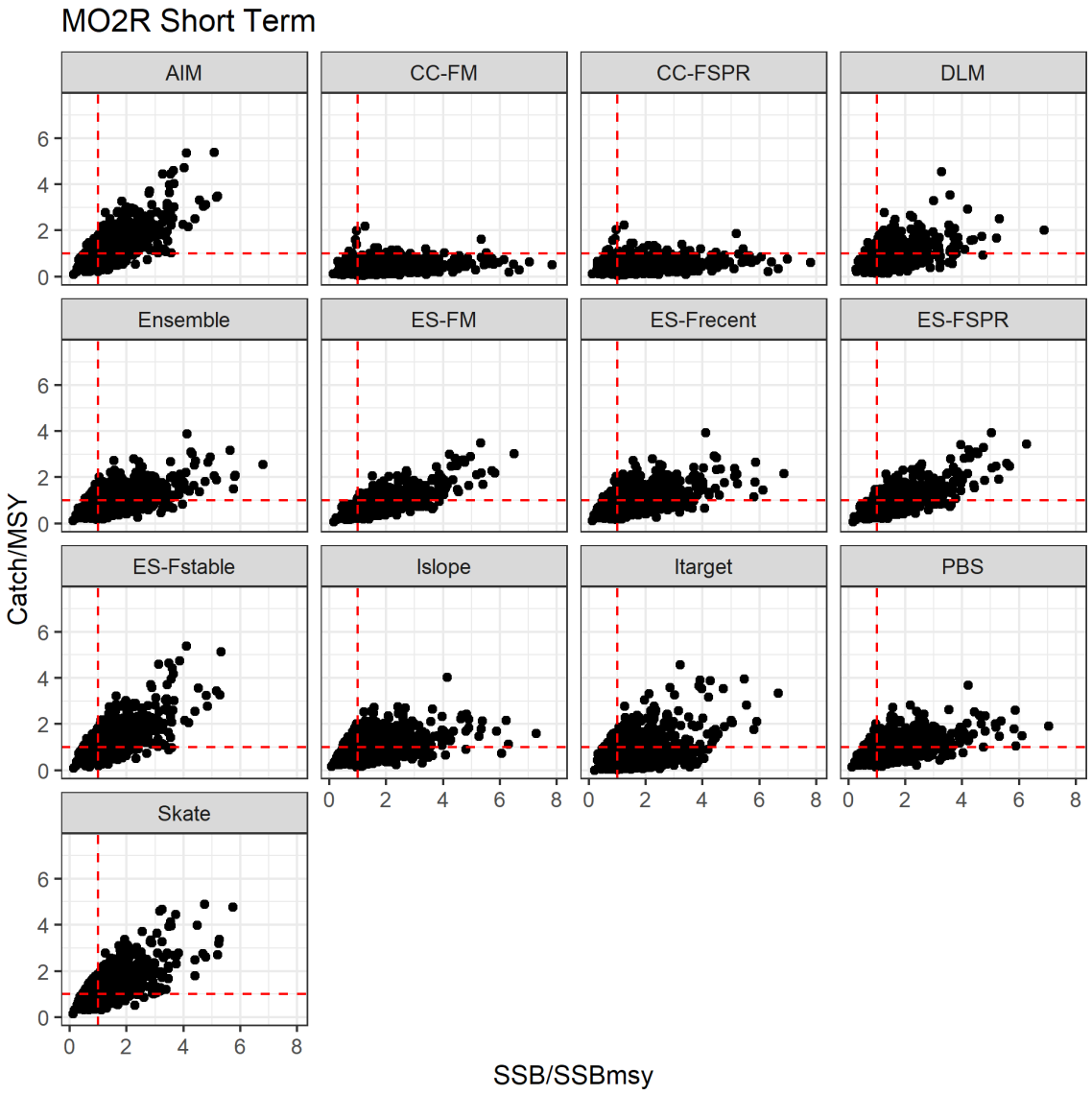


Figure A6.70. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

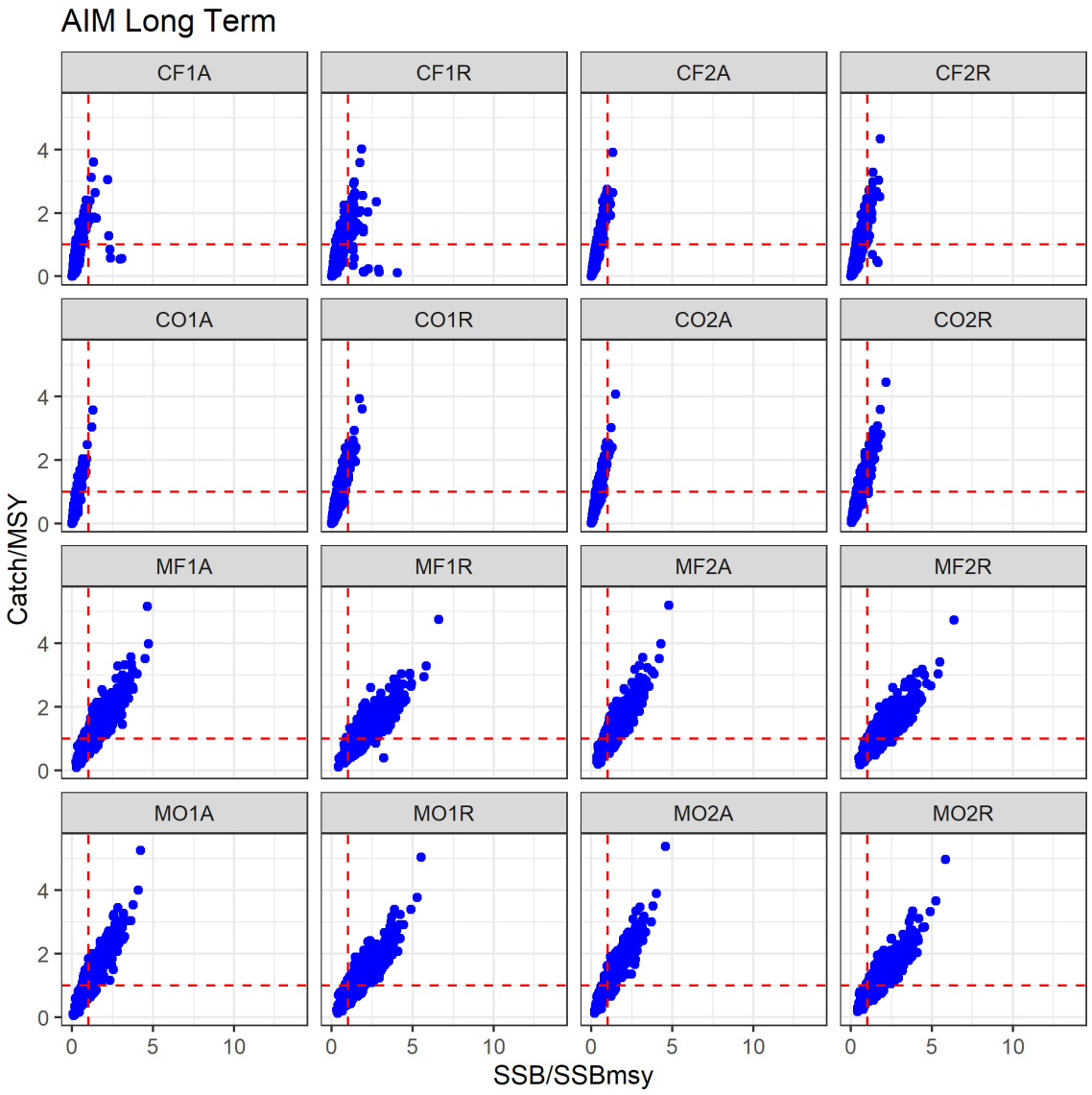


Figure A6.71. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

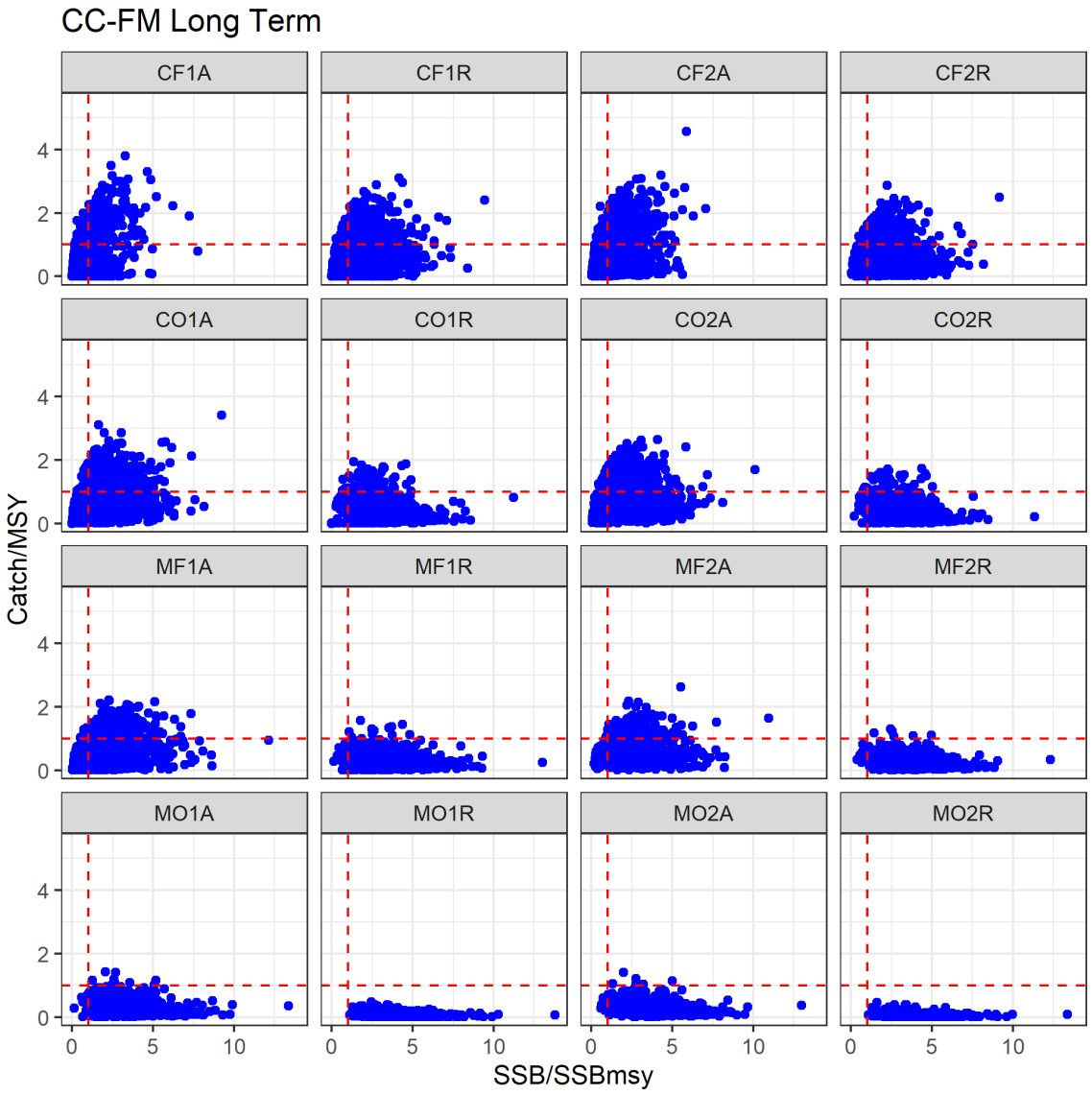


Figure A6.72. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

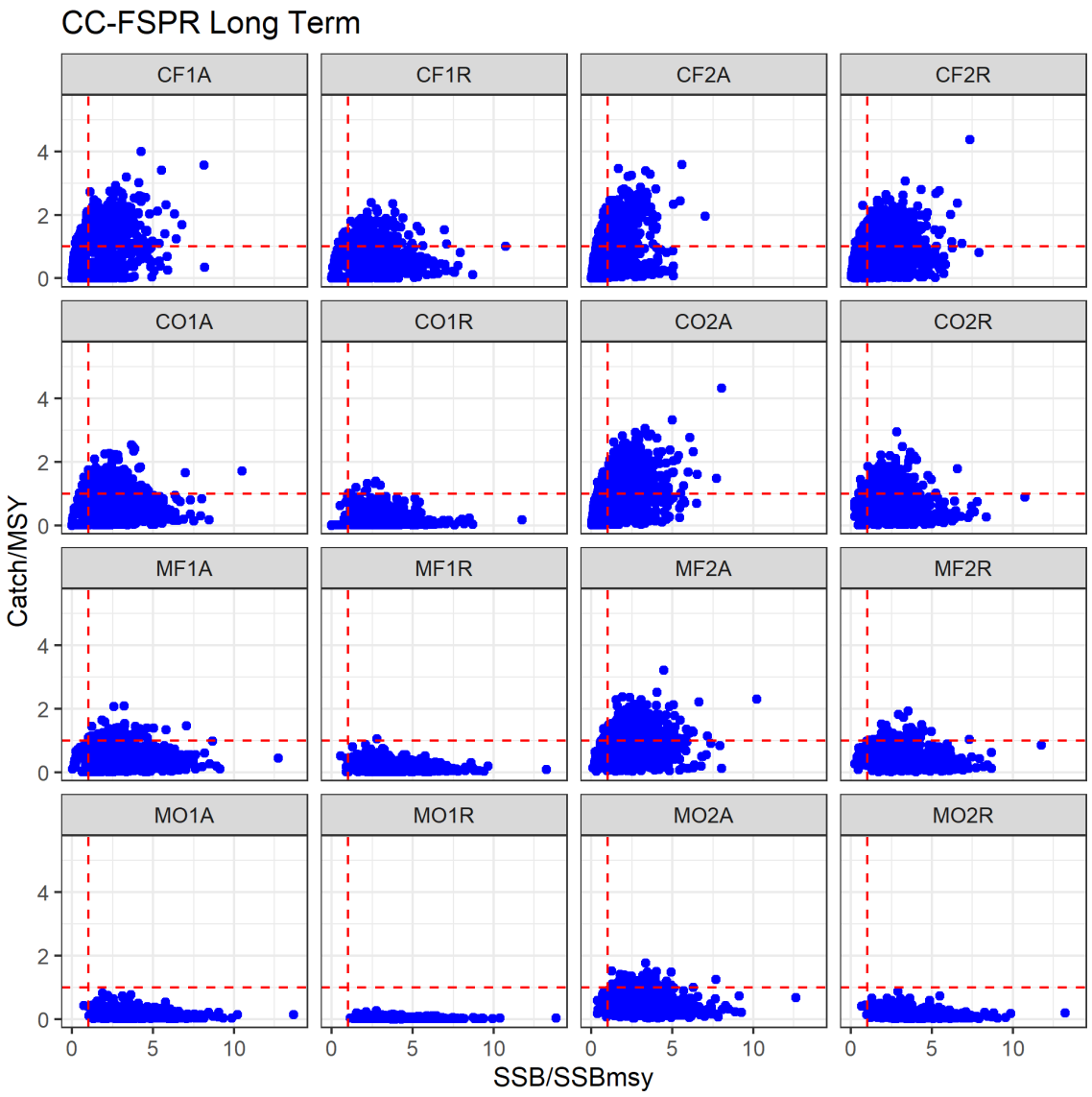


Figure A6.73. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

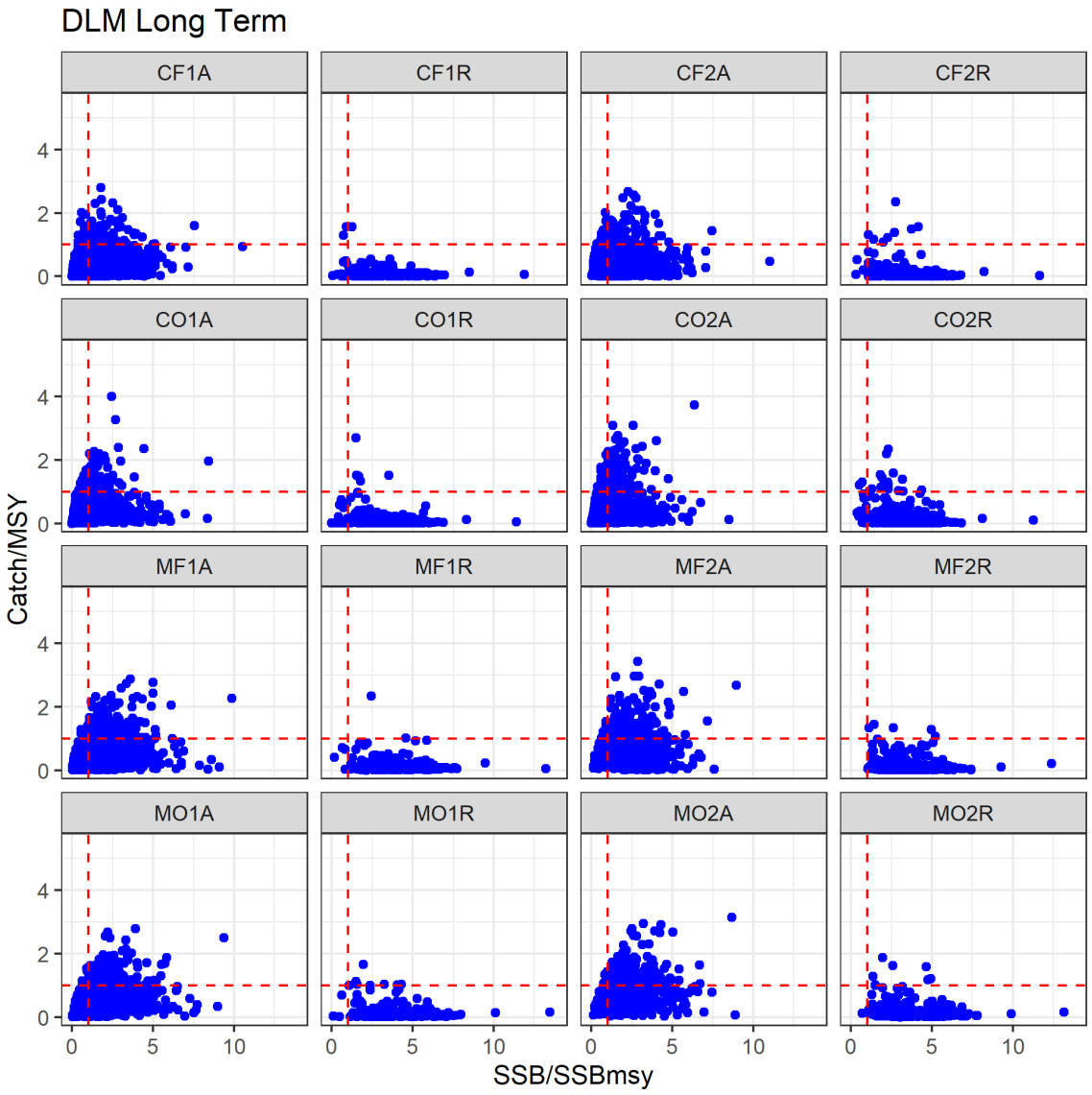


Figure A6.74. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

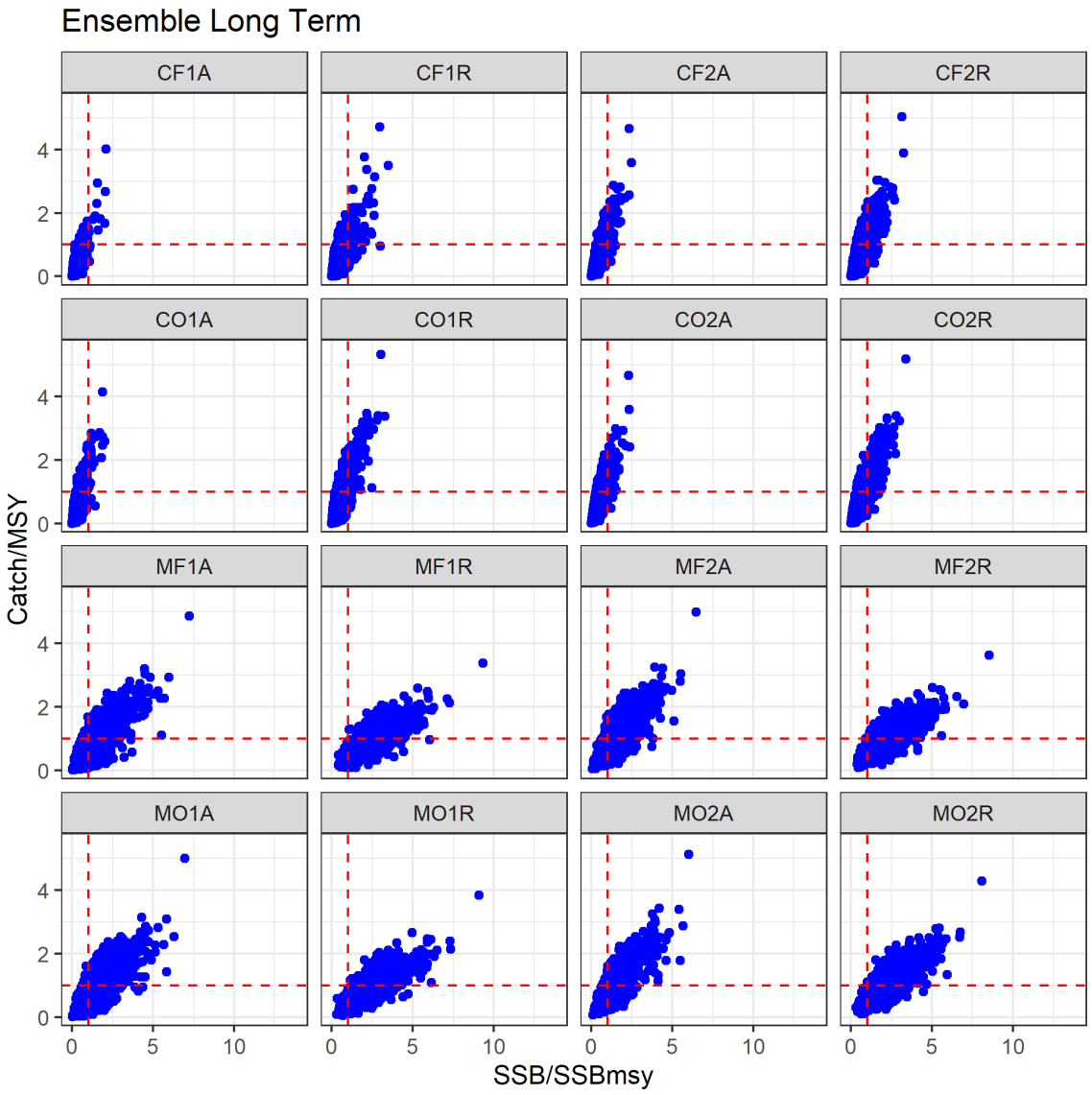


Figure A6.75. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

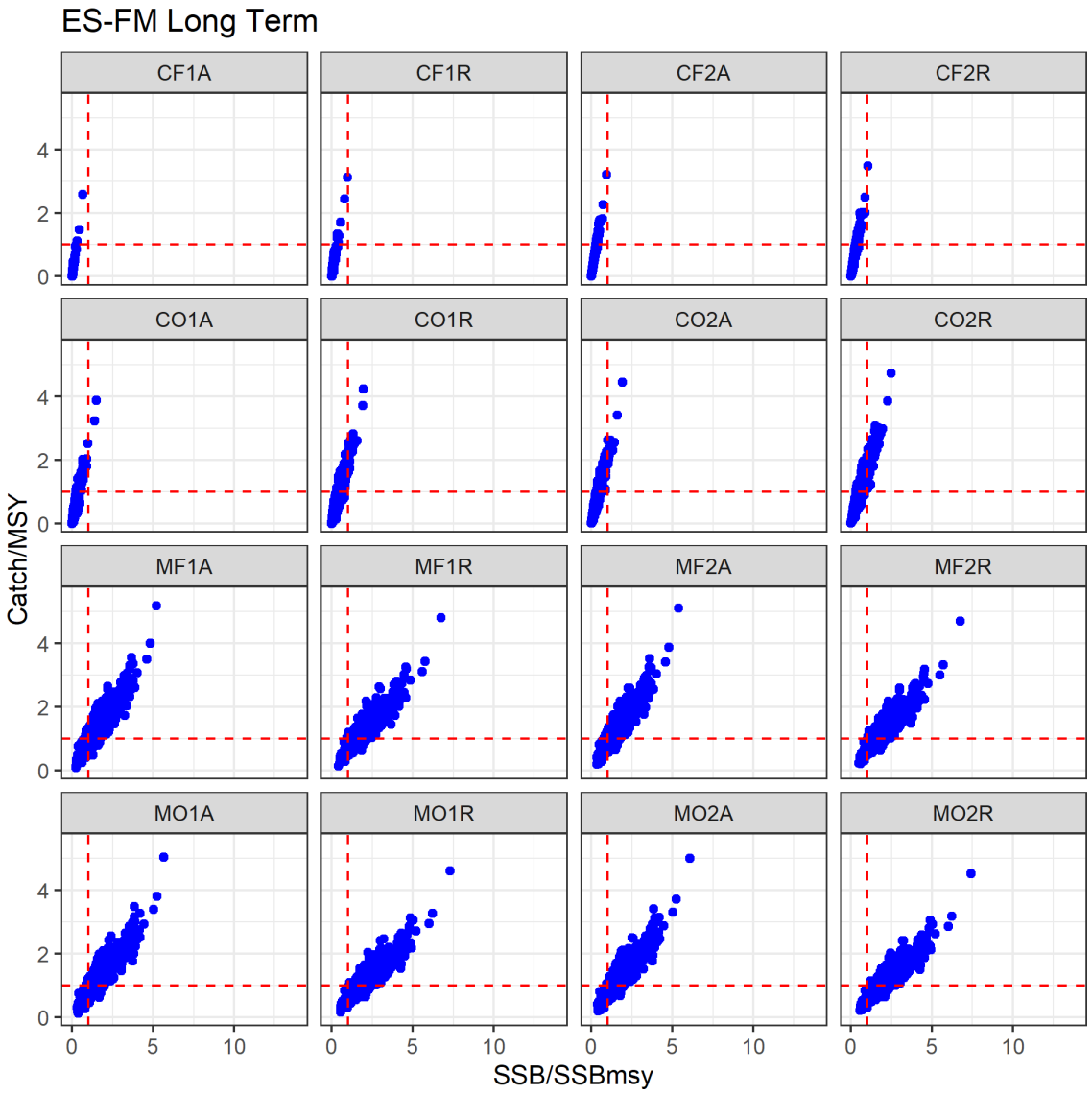


Figure A6.76. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

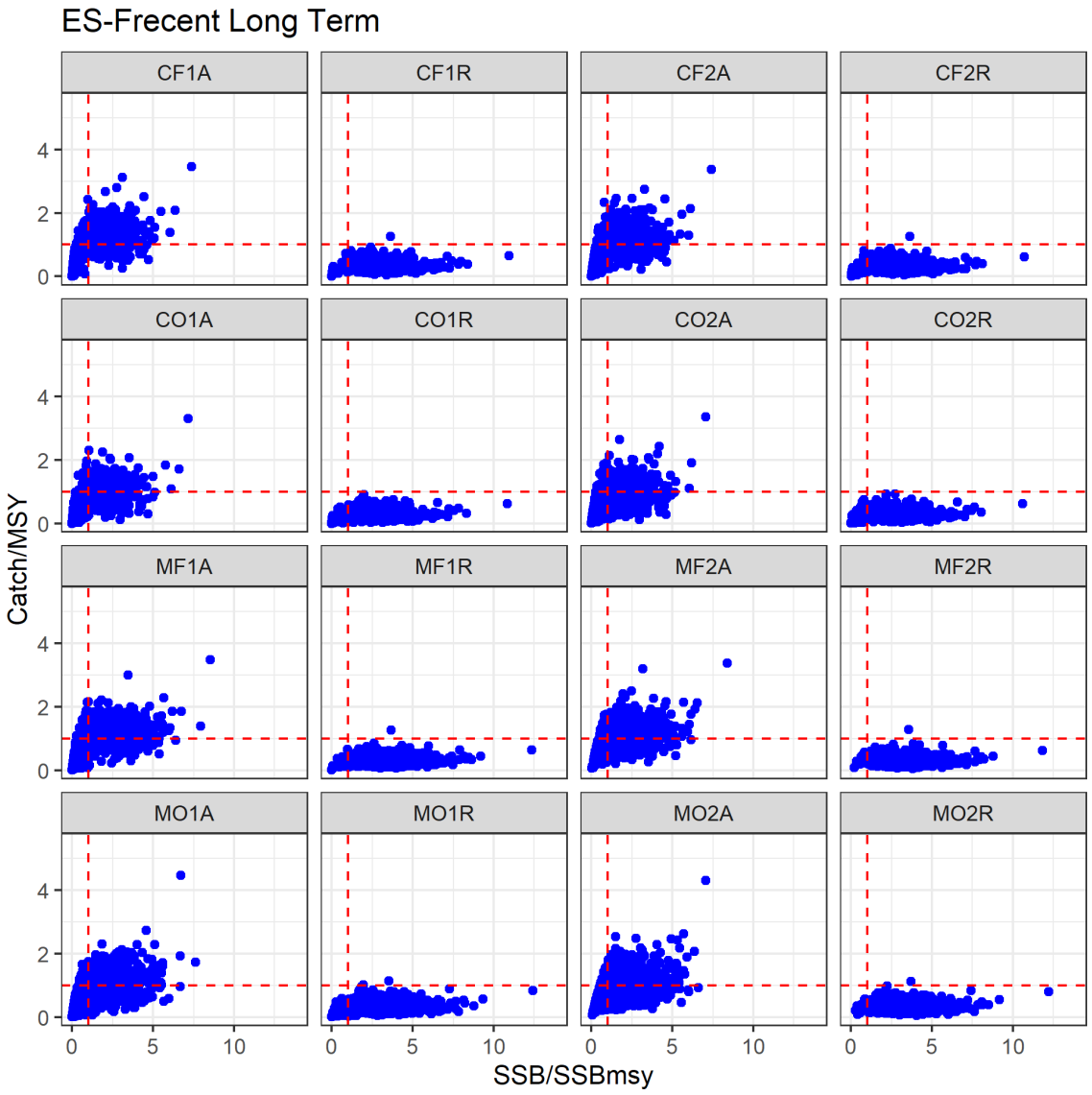


Figure A6.77. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

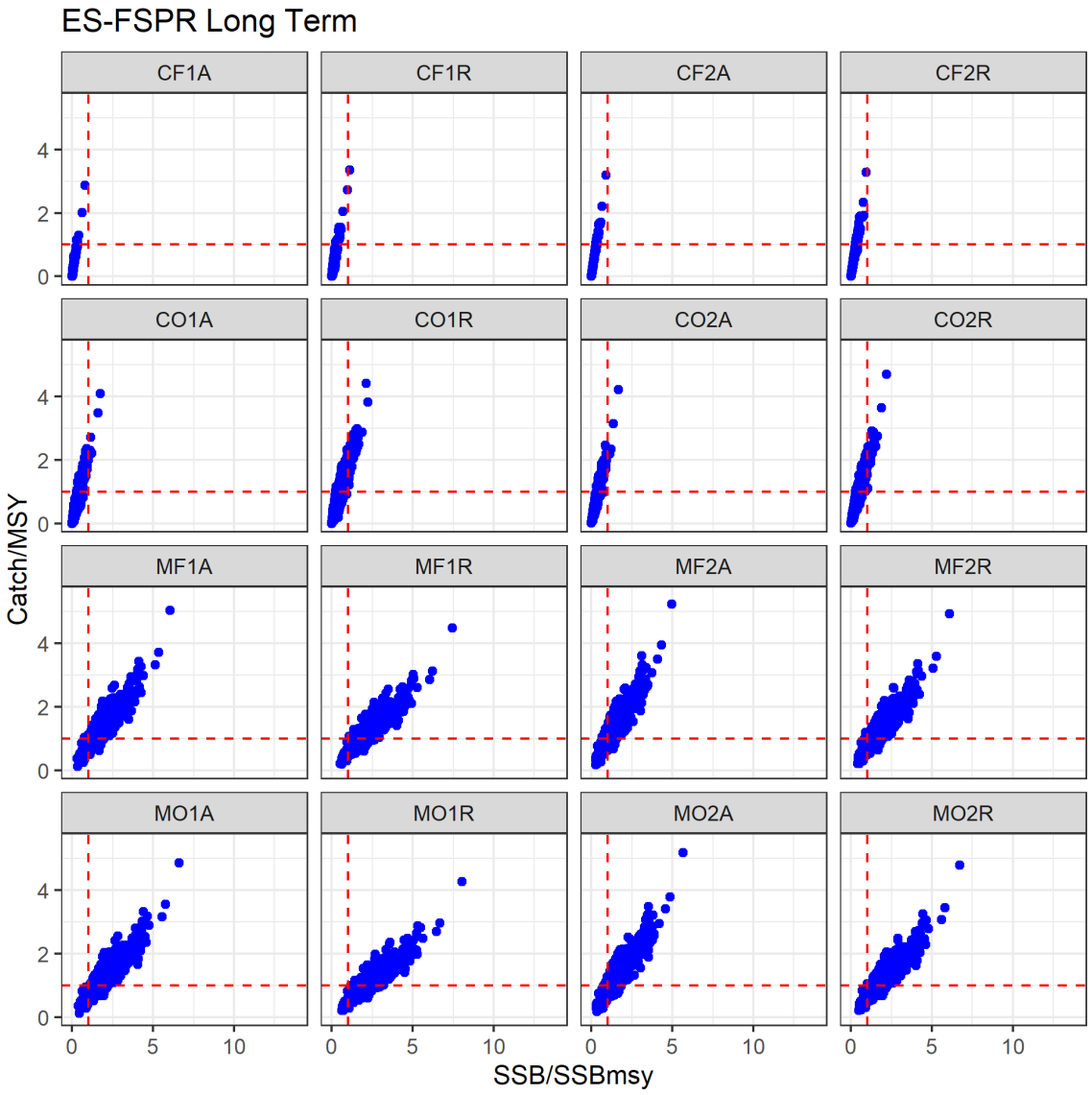


Figure A6.78. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

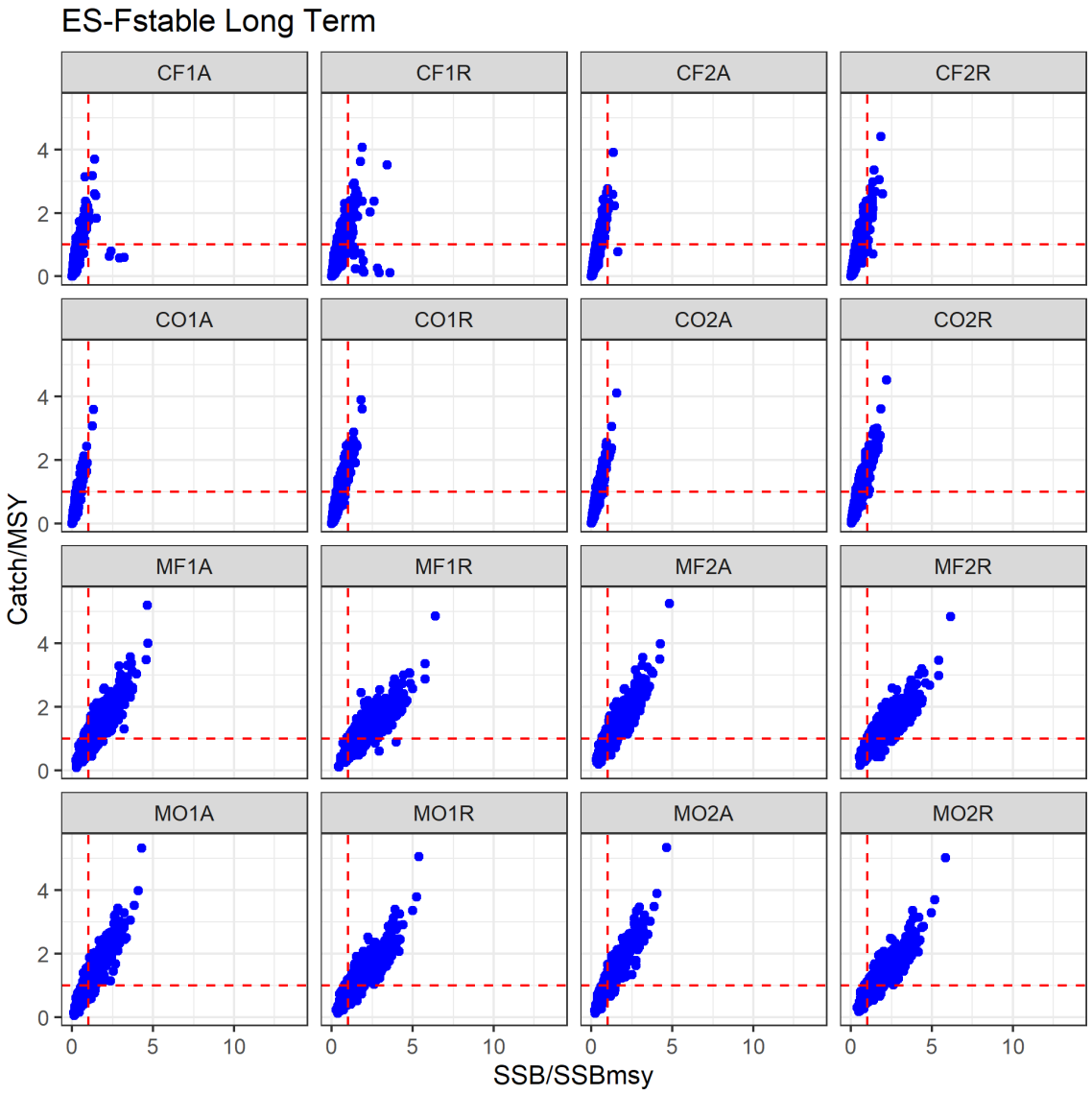


Figure A6.79. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

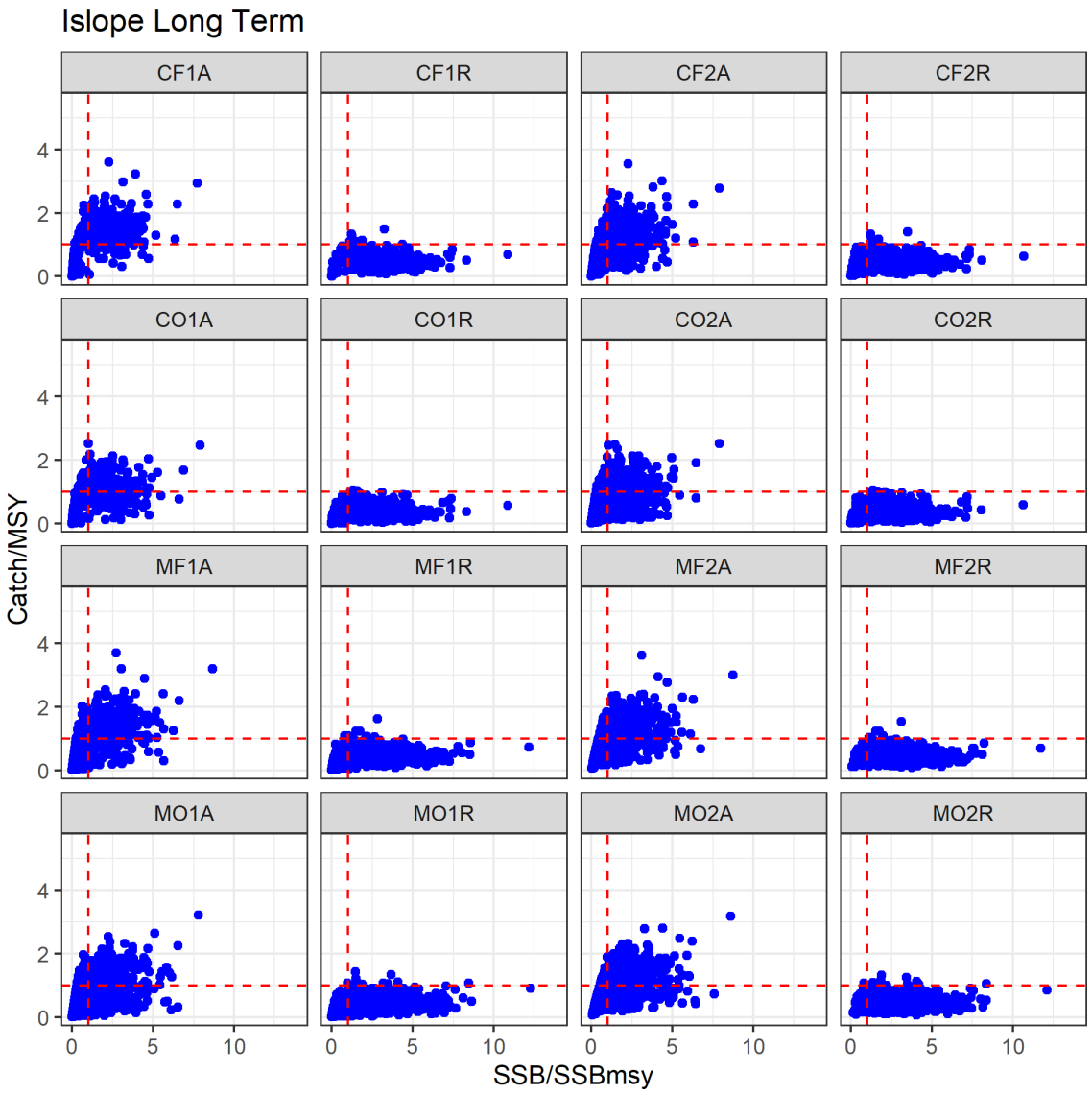


Figure A6.80. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

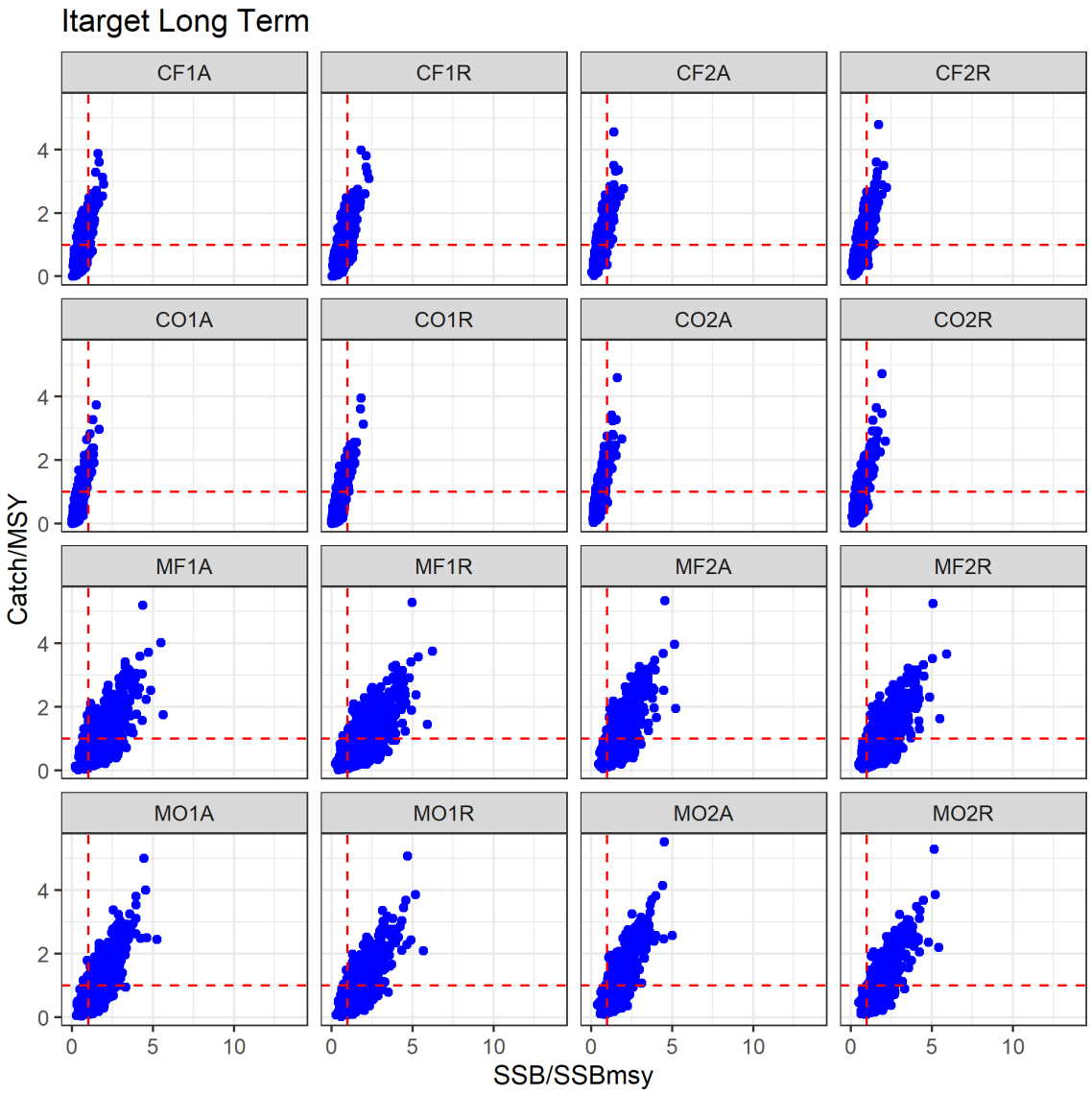


Figure A6.81. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

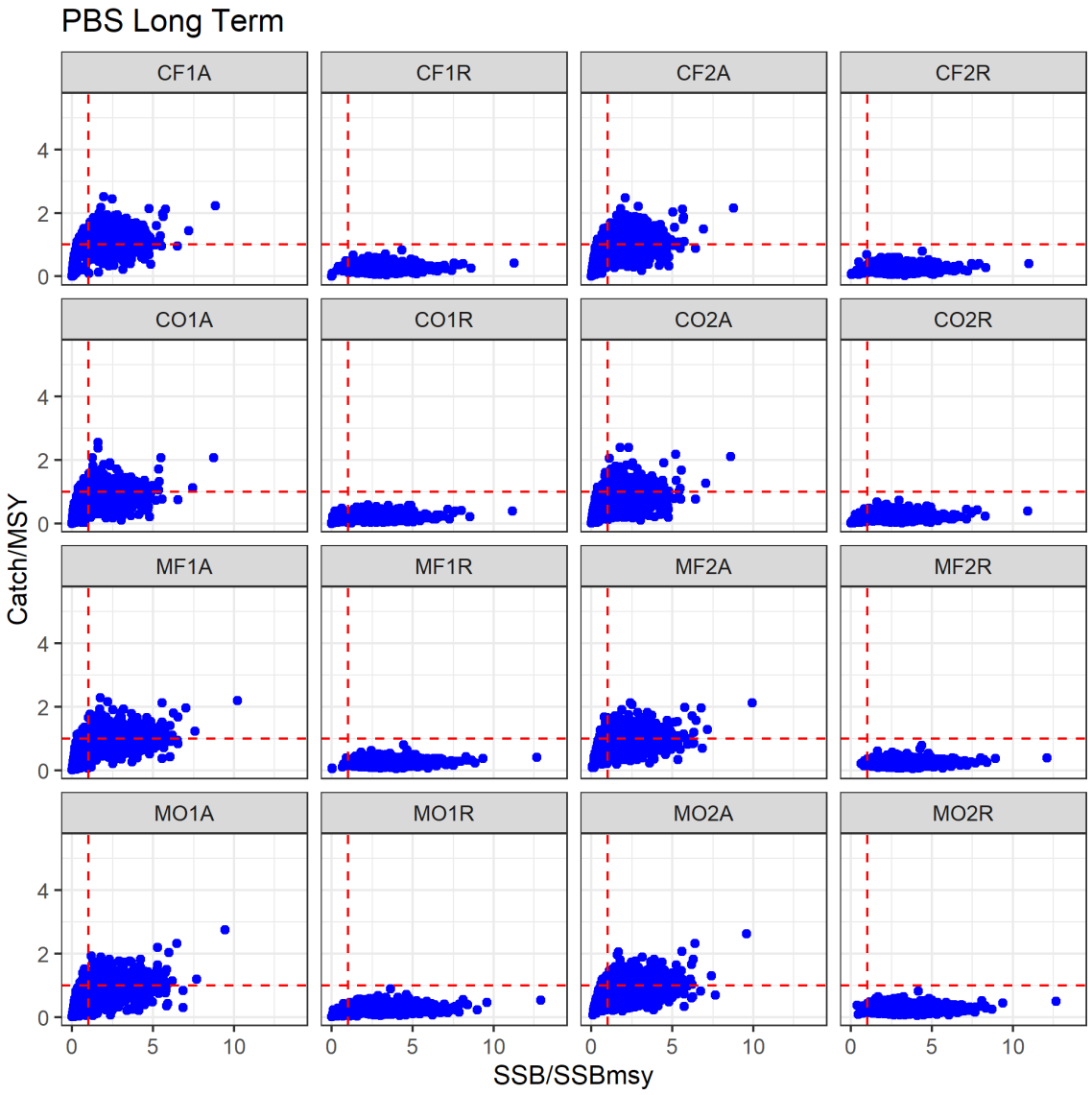


Figure A6.82. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

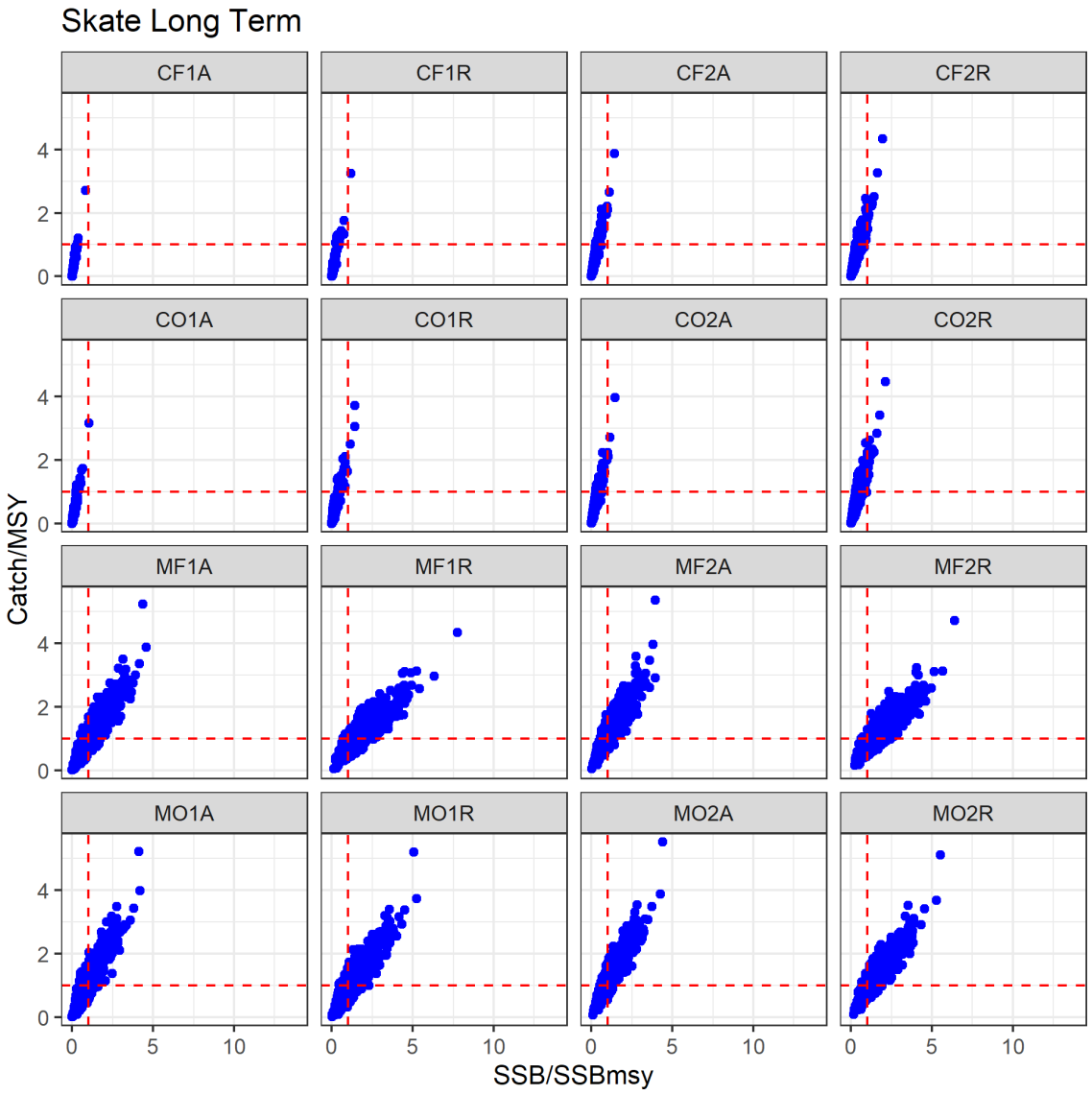


Figure A6.83. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

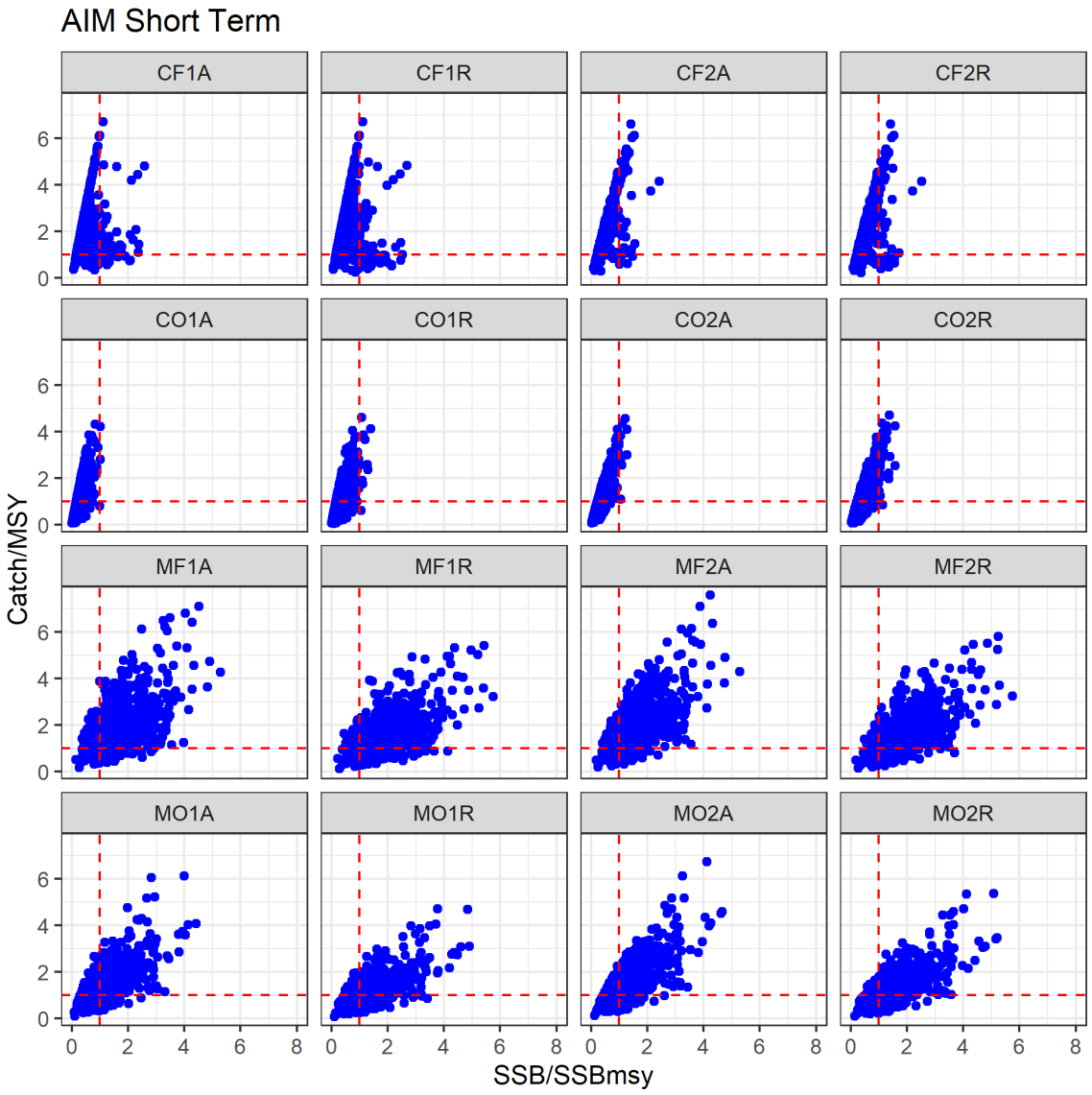


Figure A6.84. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

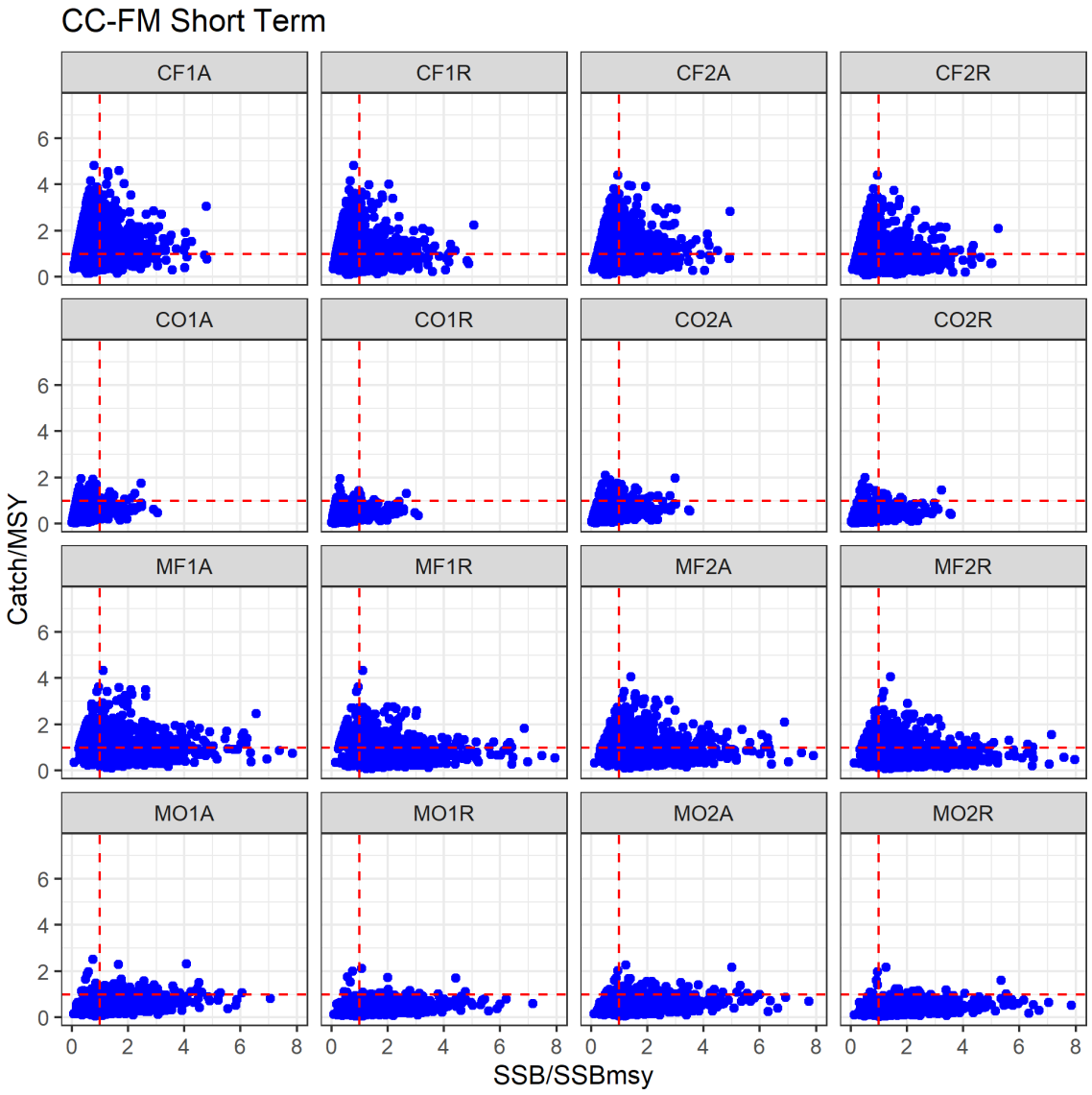


Figure A6.85. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

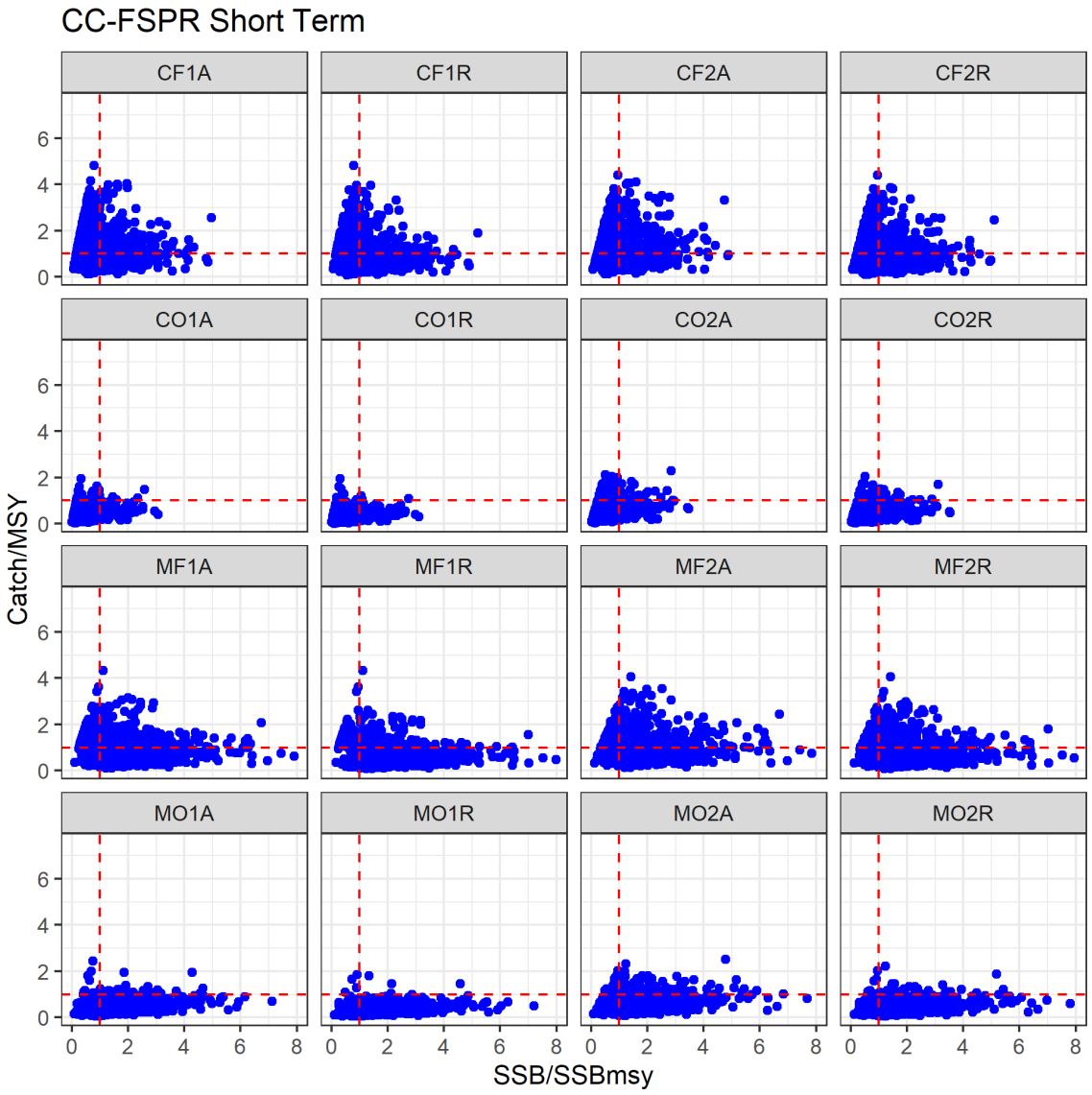


Figure A6.86. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

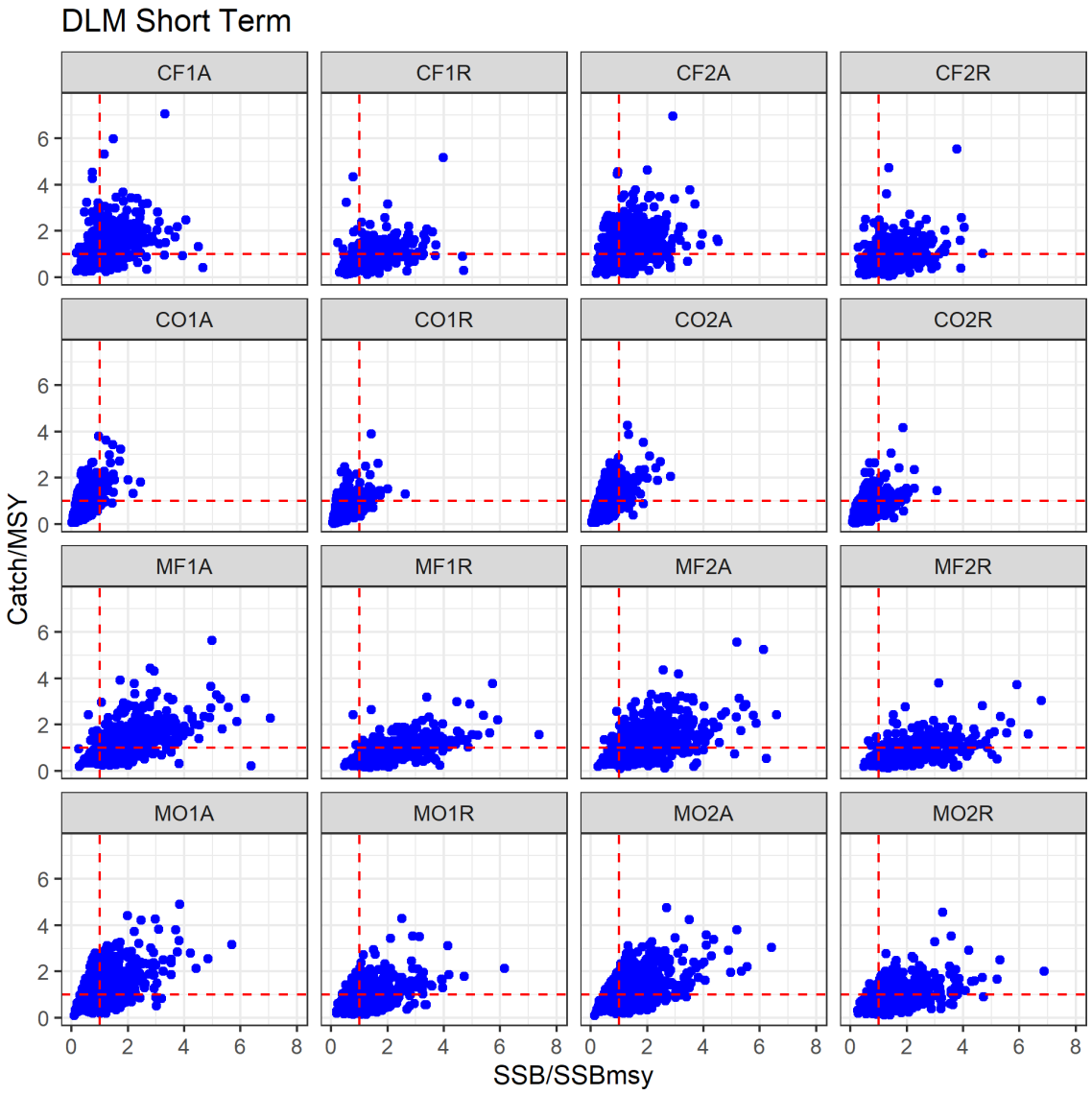


Figure A6.87. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

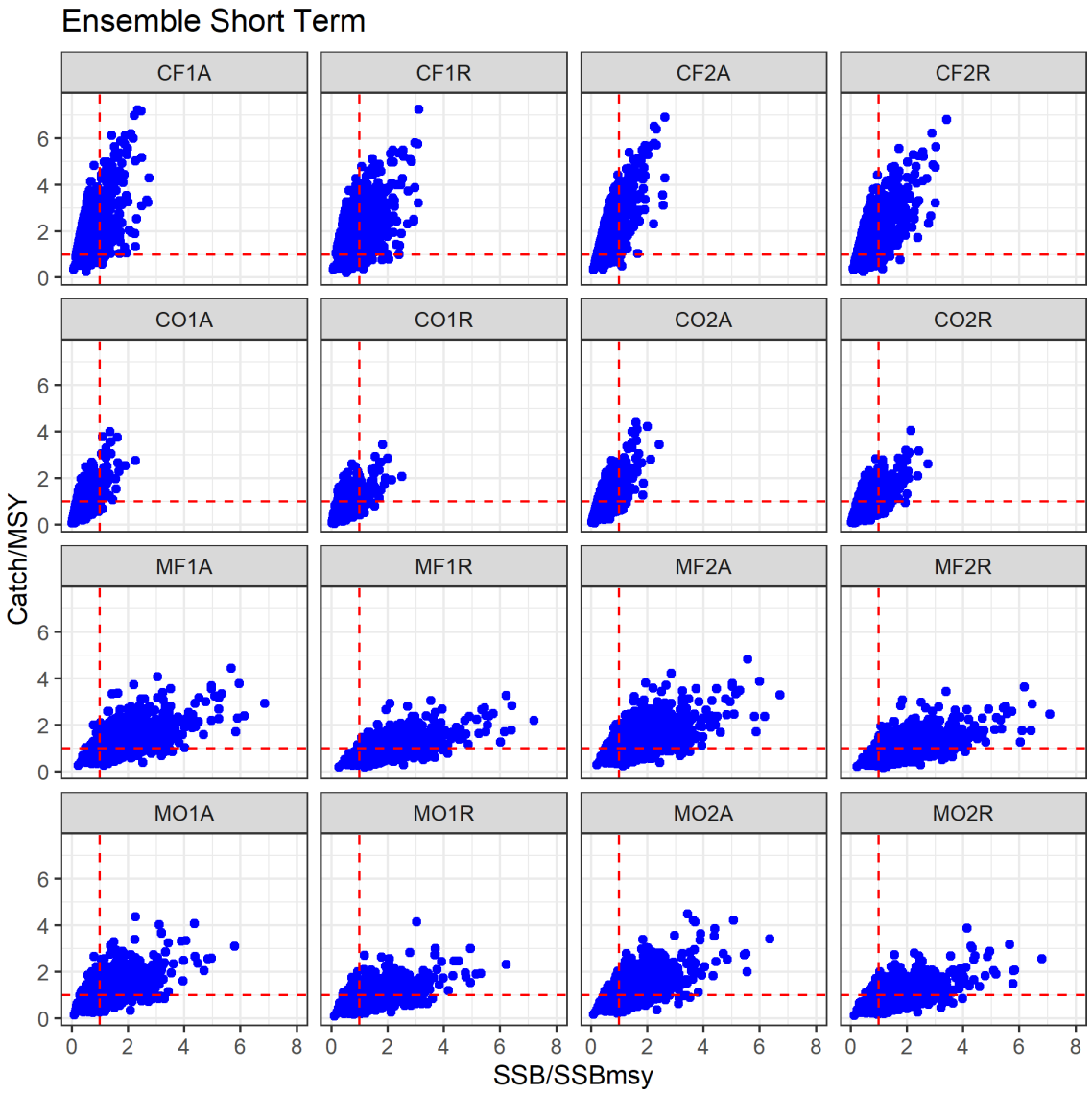


Figure A6.88. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

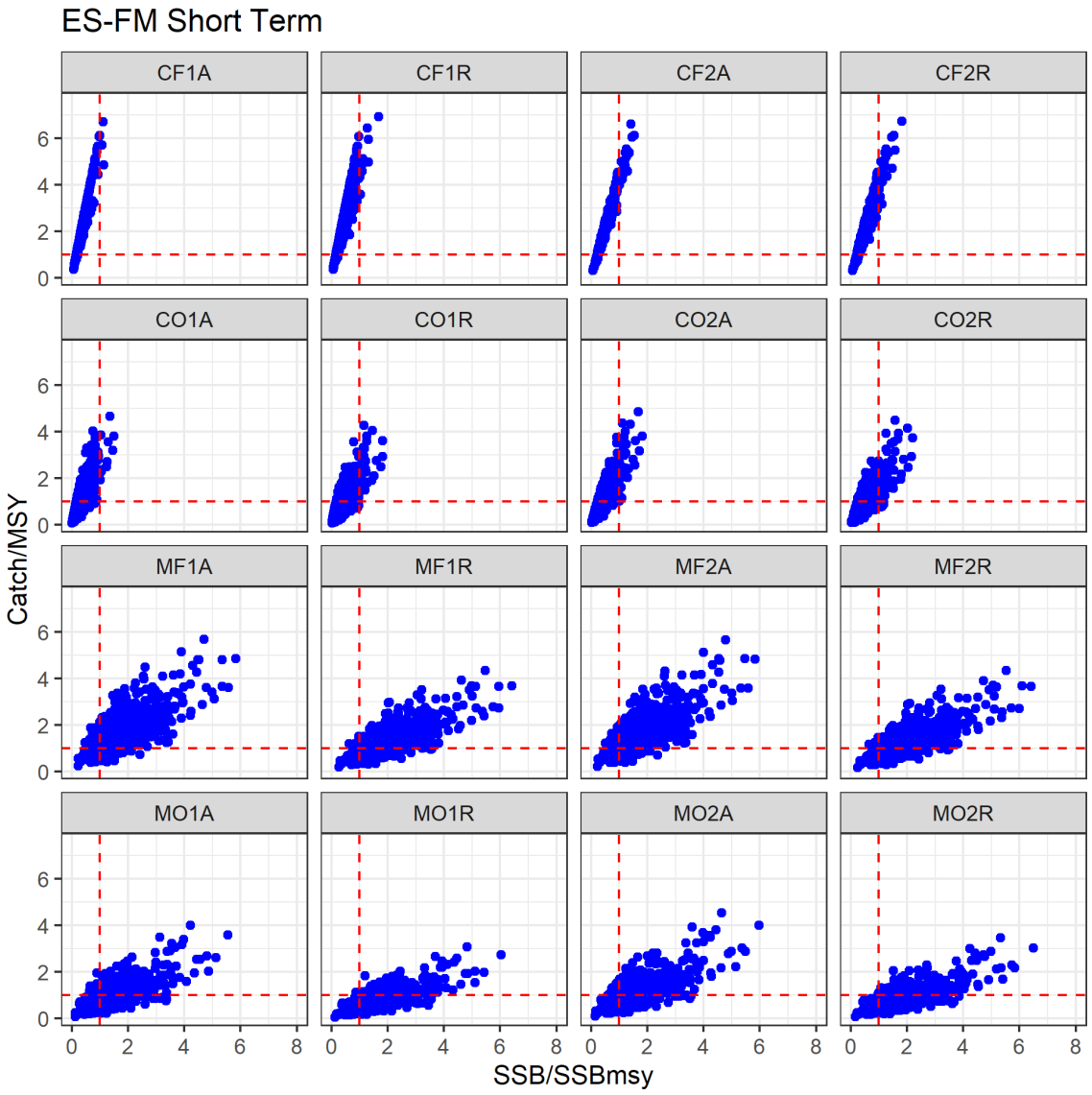


Figure A6.89. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

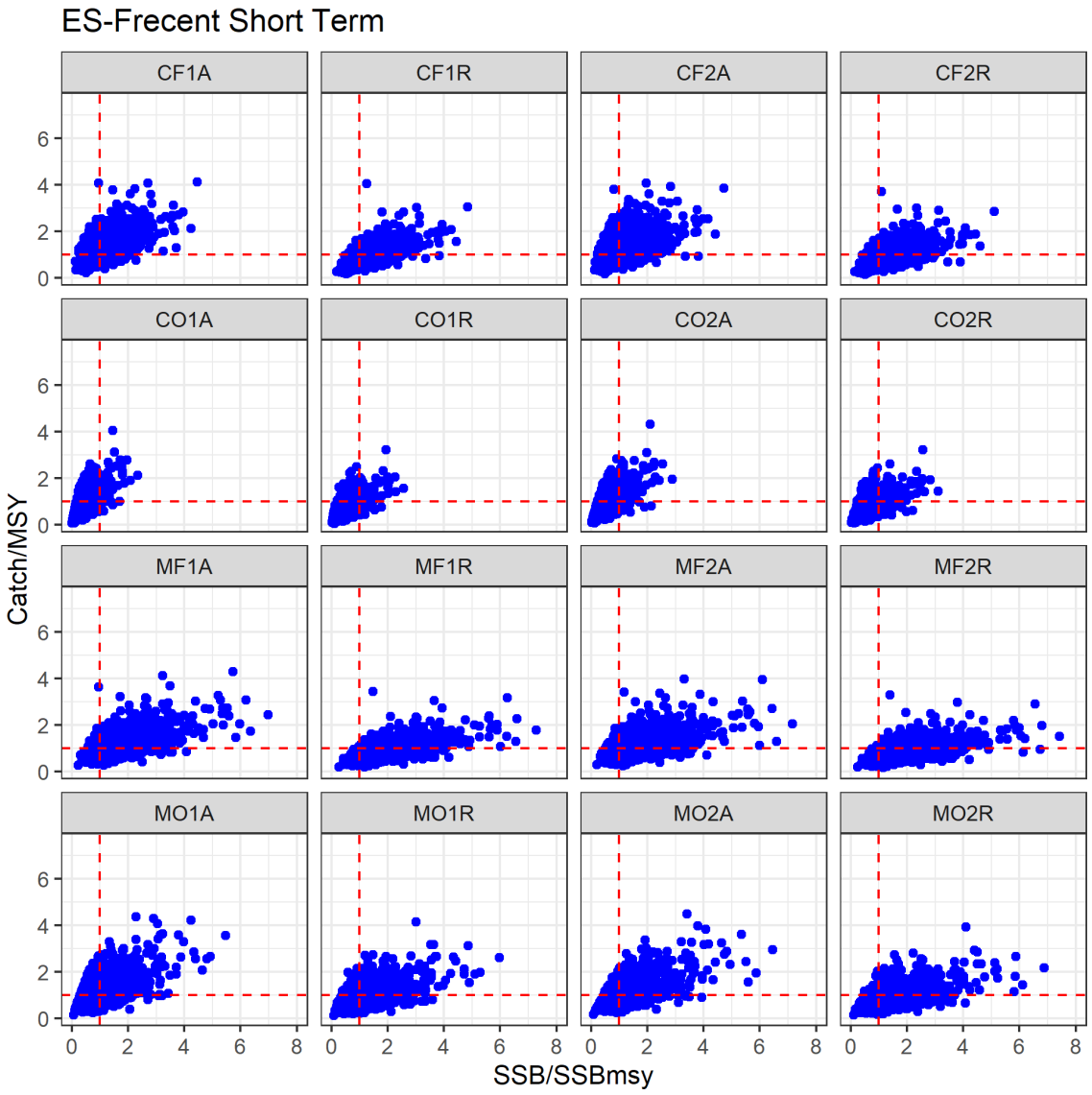


Figure A6.90. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

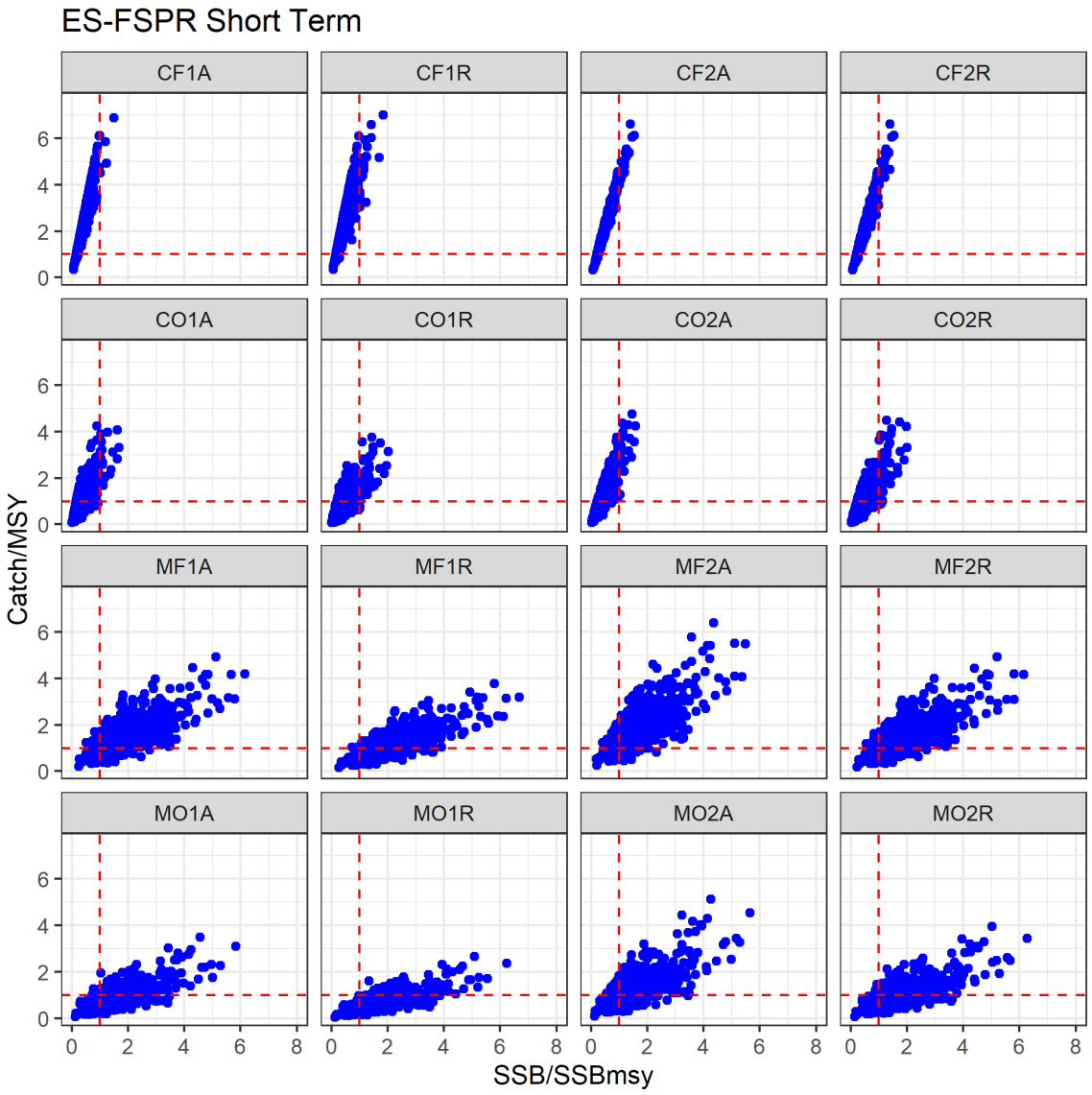


Figure A6.91. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

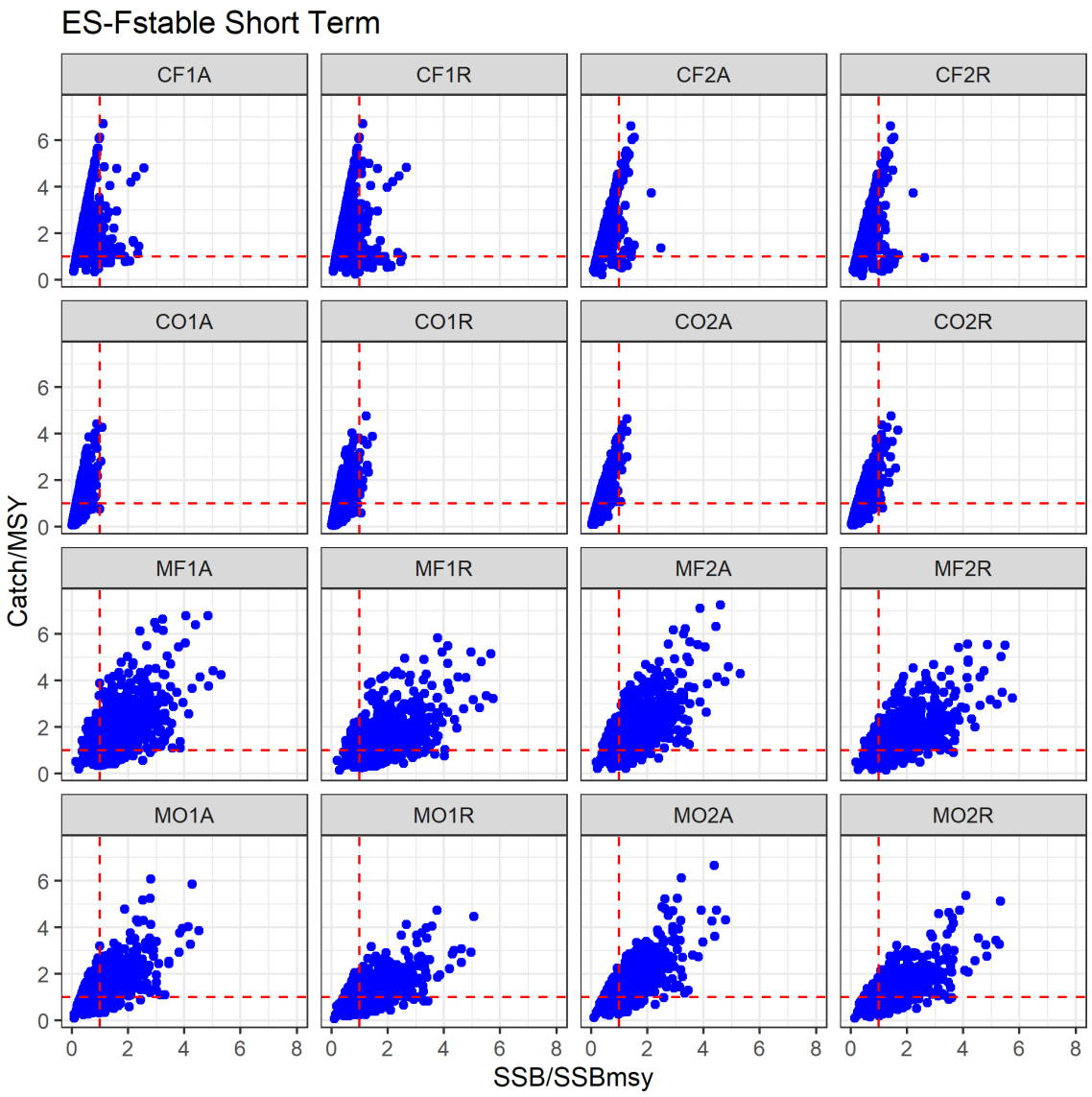


Figure A6.92. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

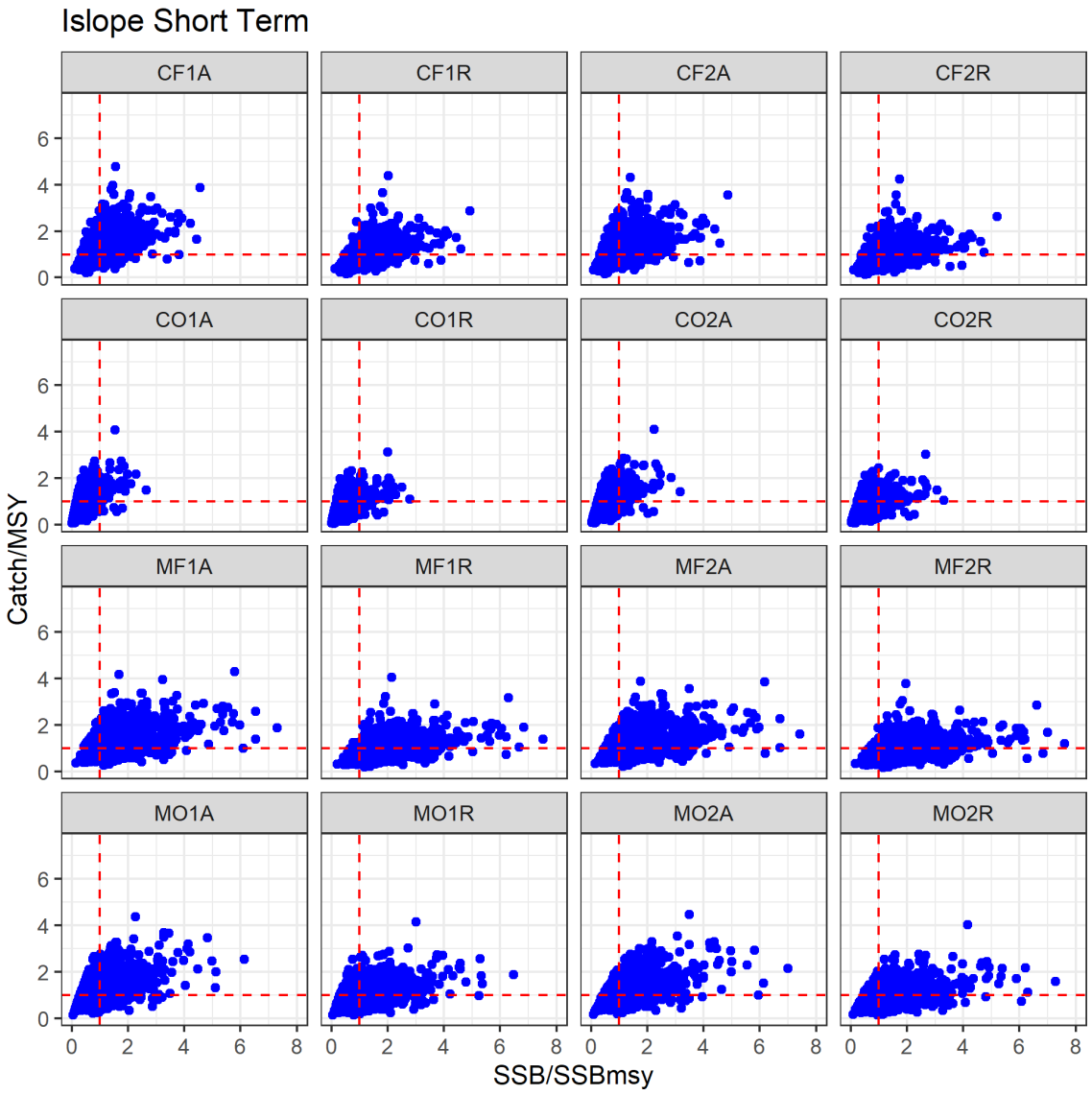


Figure A6.93. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

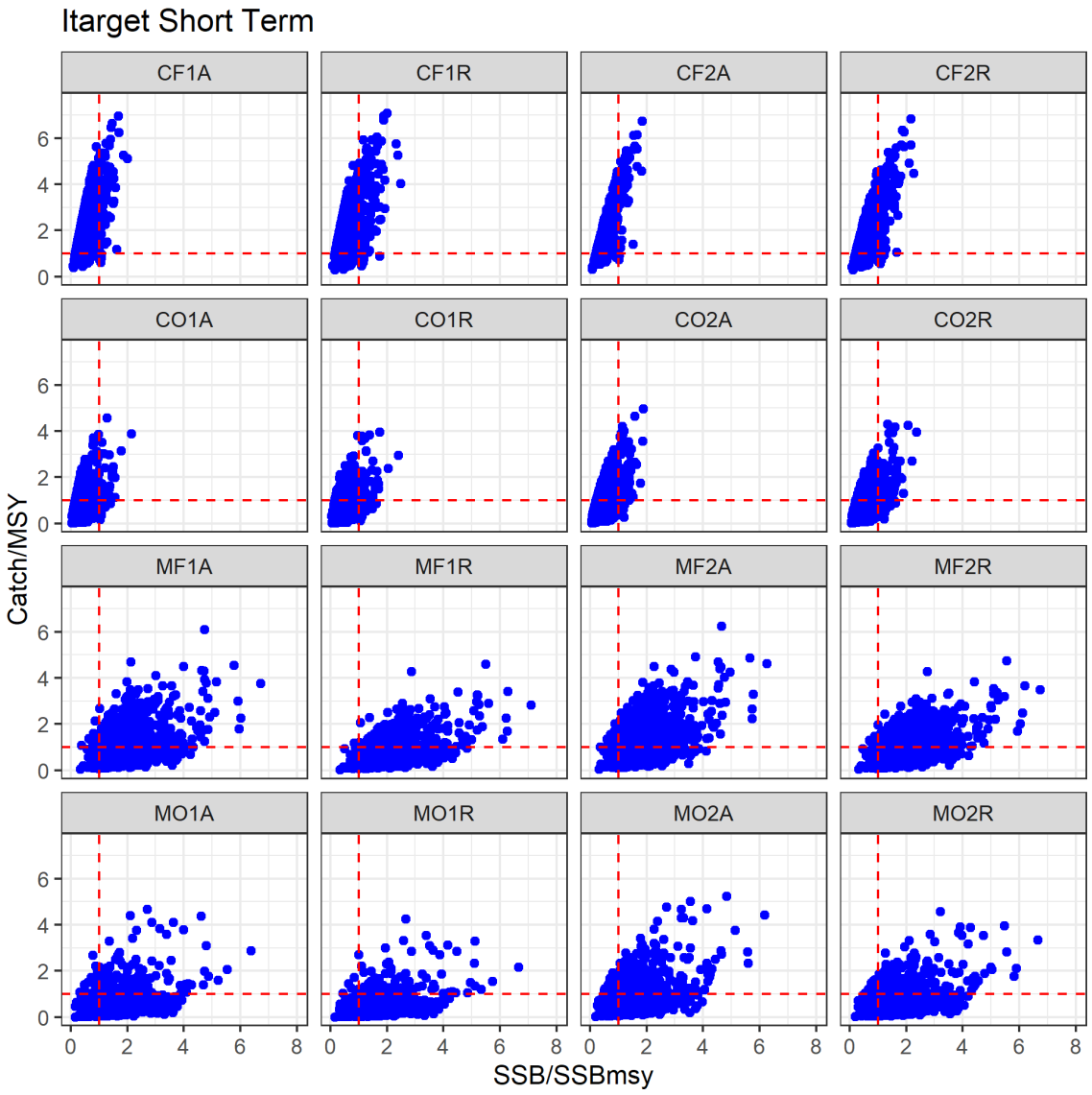


Figure A6.94. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

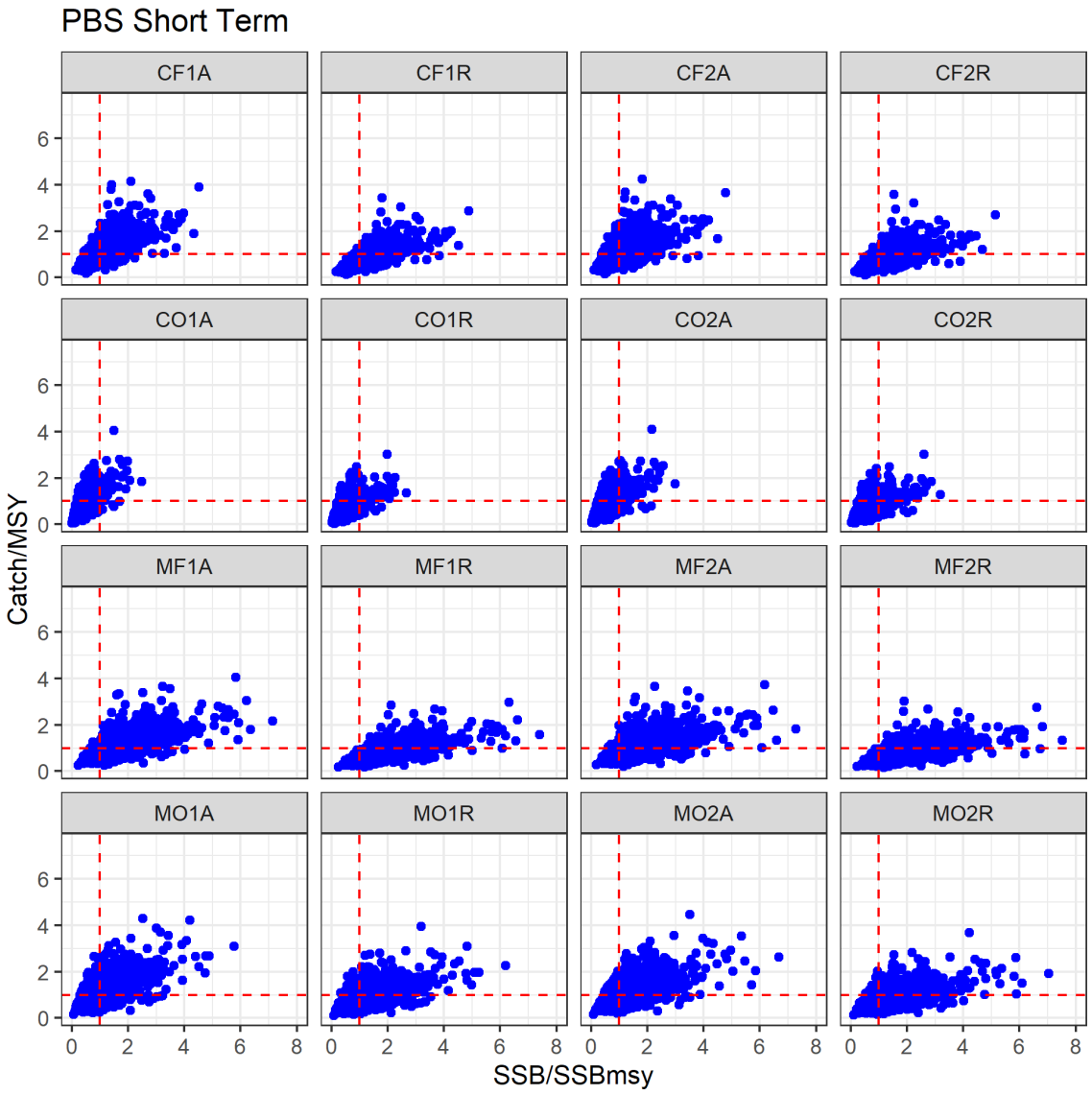


Figure A6.95. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

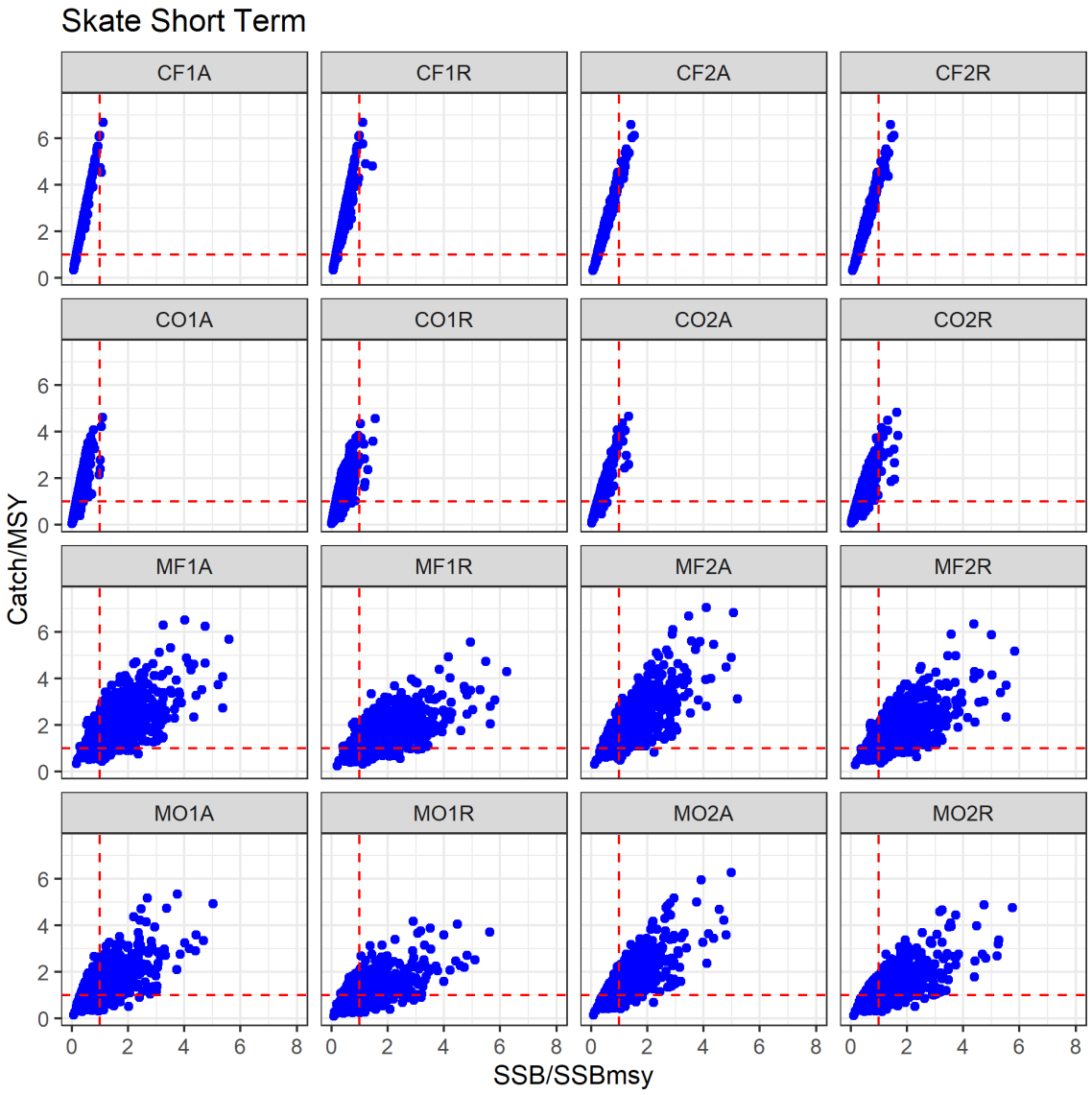


Figure A6.96. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

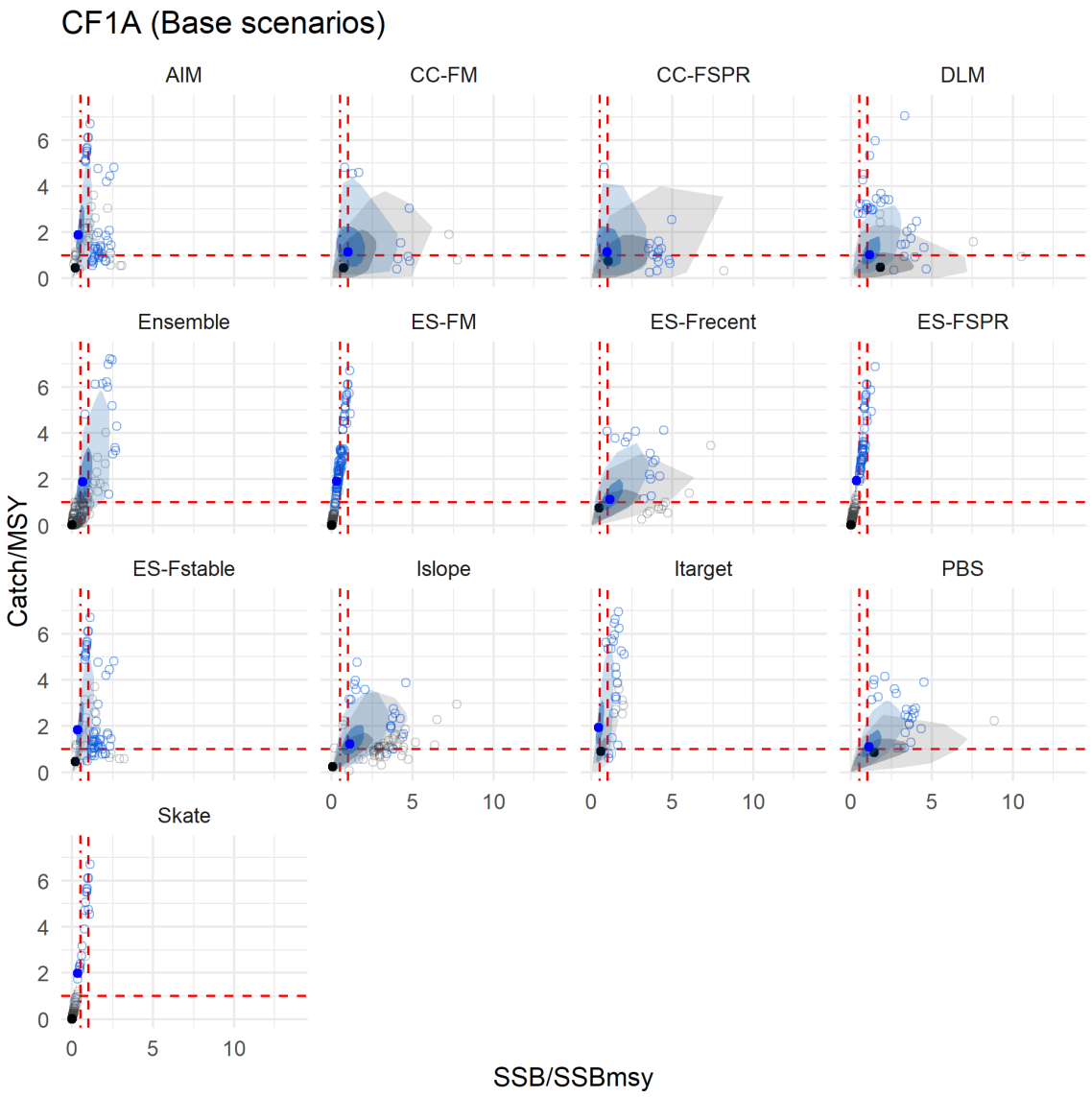


Figure A6.97. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

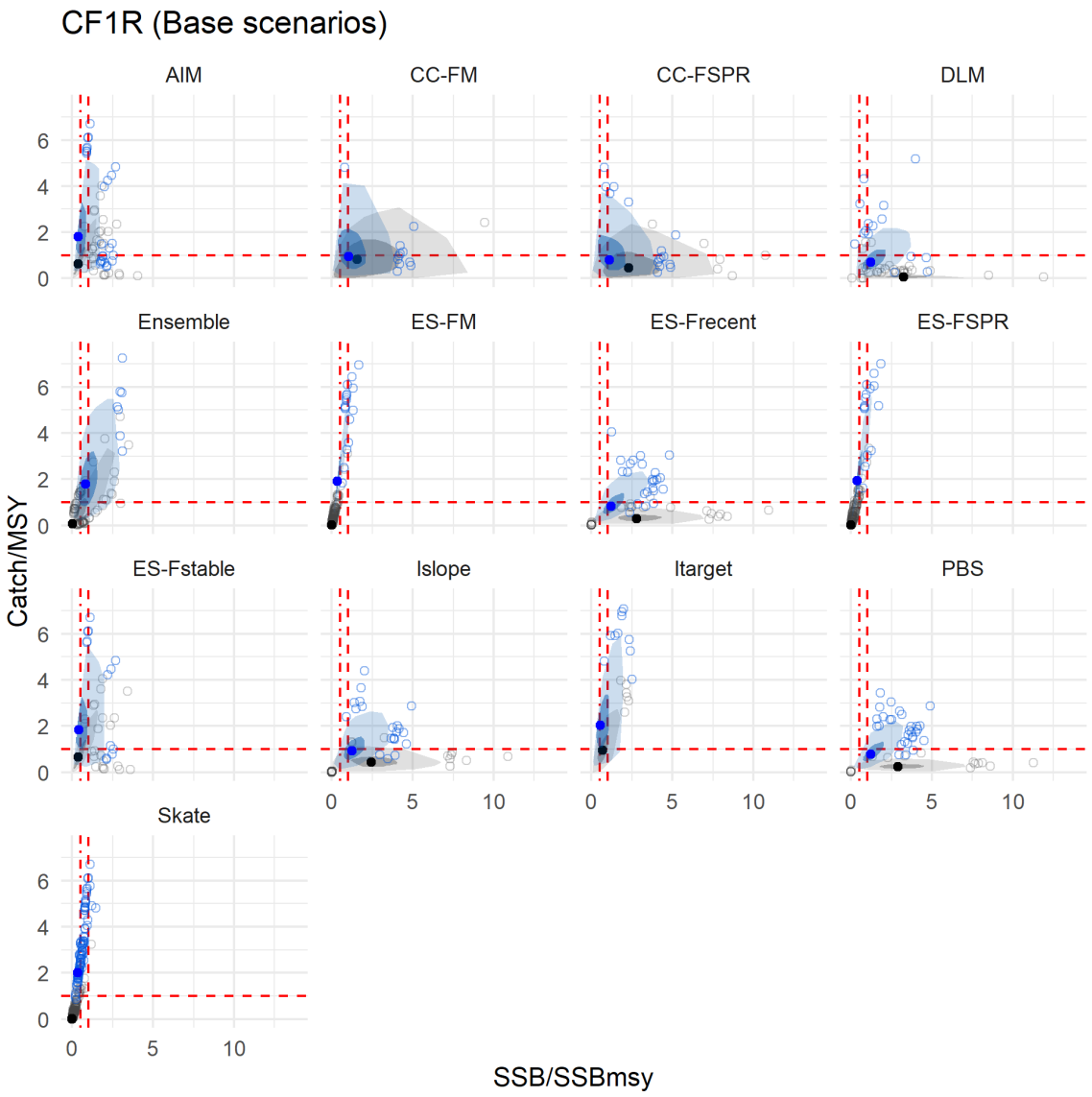


Figure A6.98. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

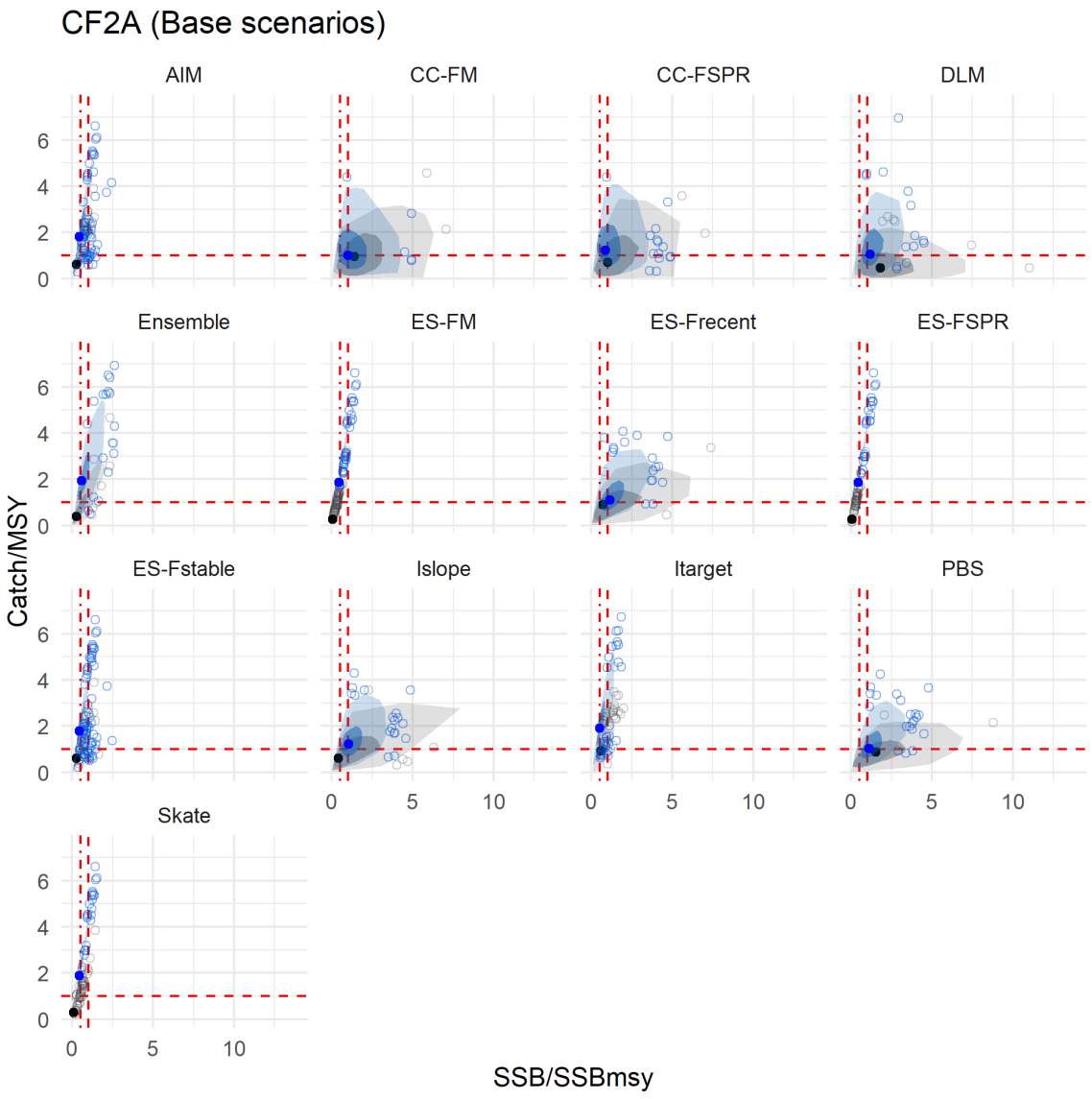


Figure A6.99. Bagplots (a bivariate generalization of the boxplot) for longterm (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

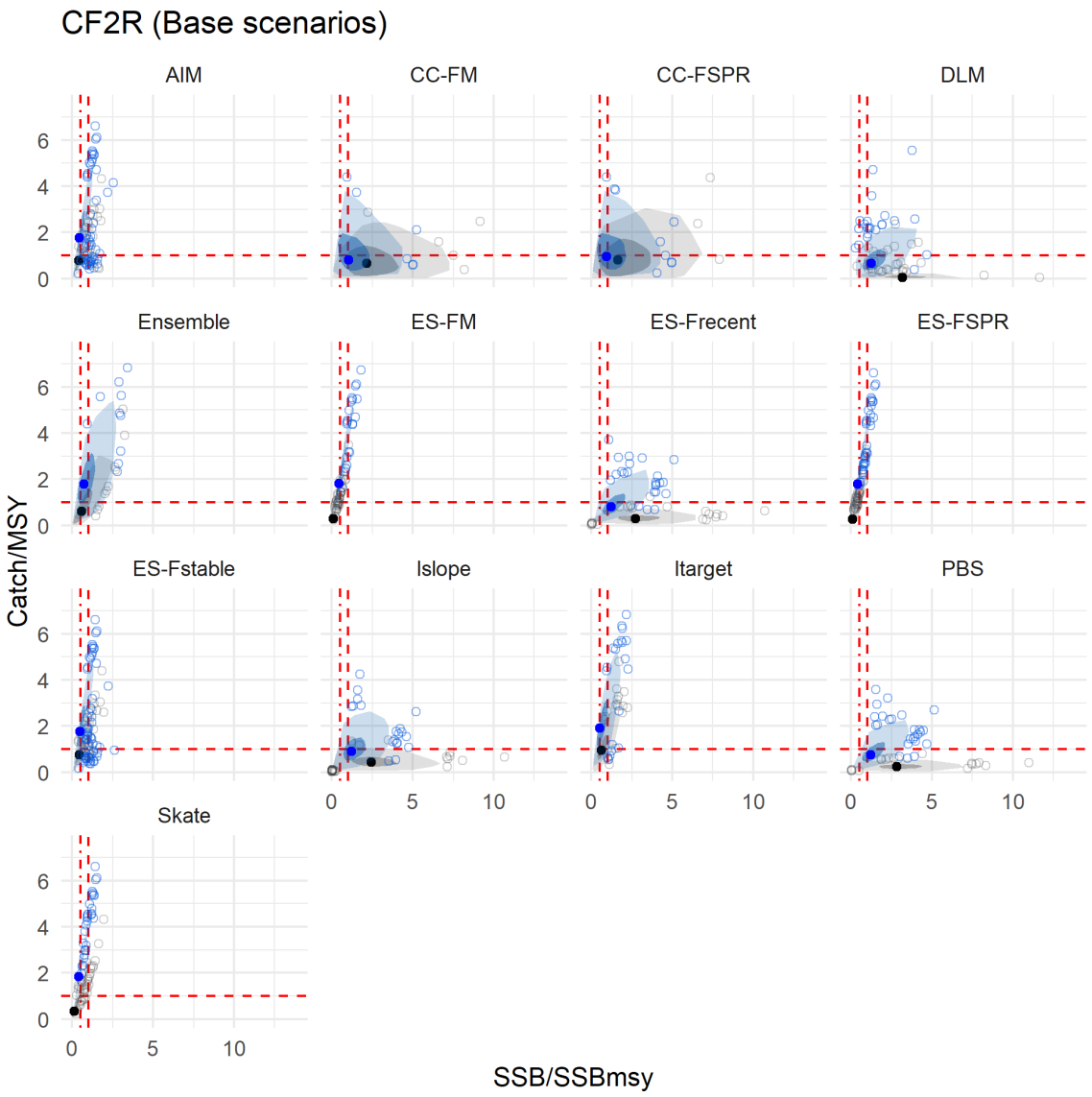


Figure A6.100. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

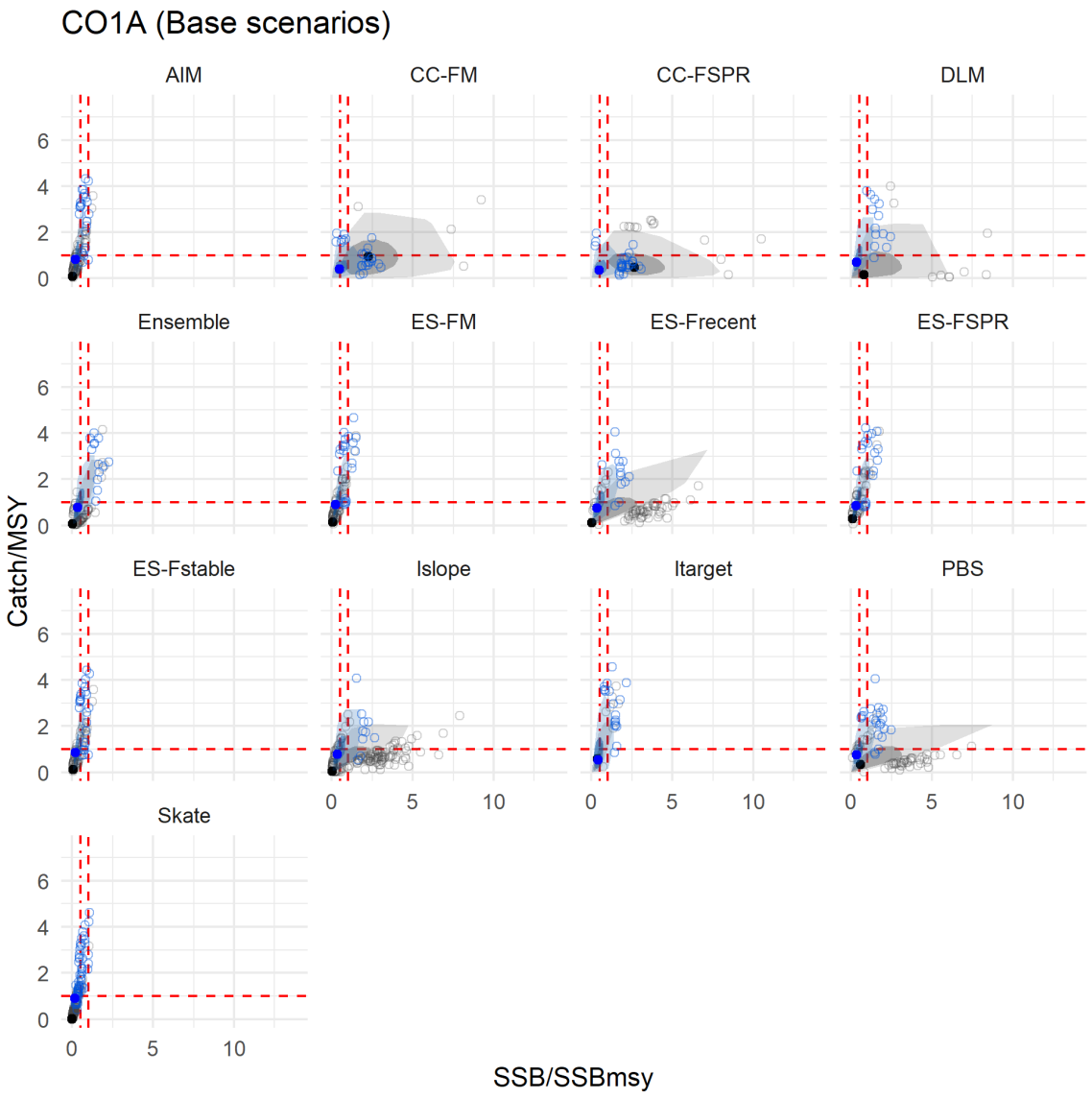


Figure A6.101. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

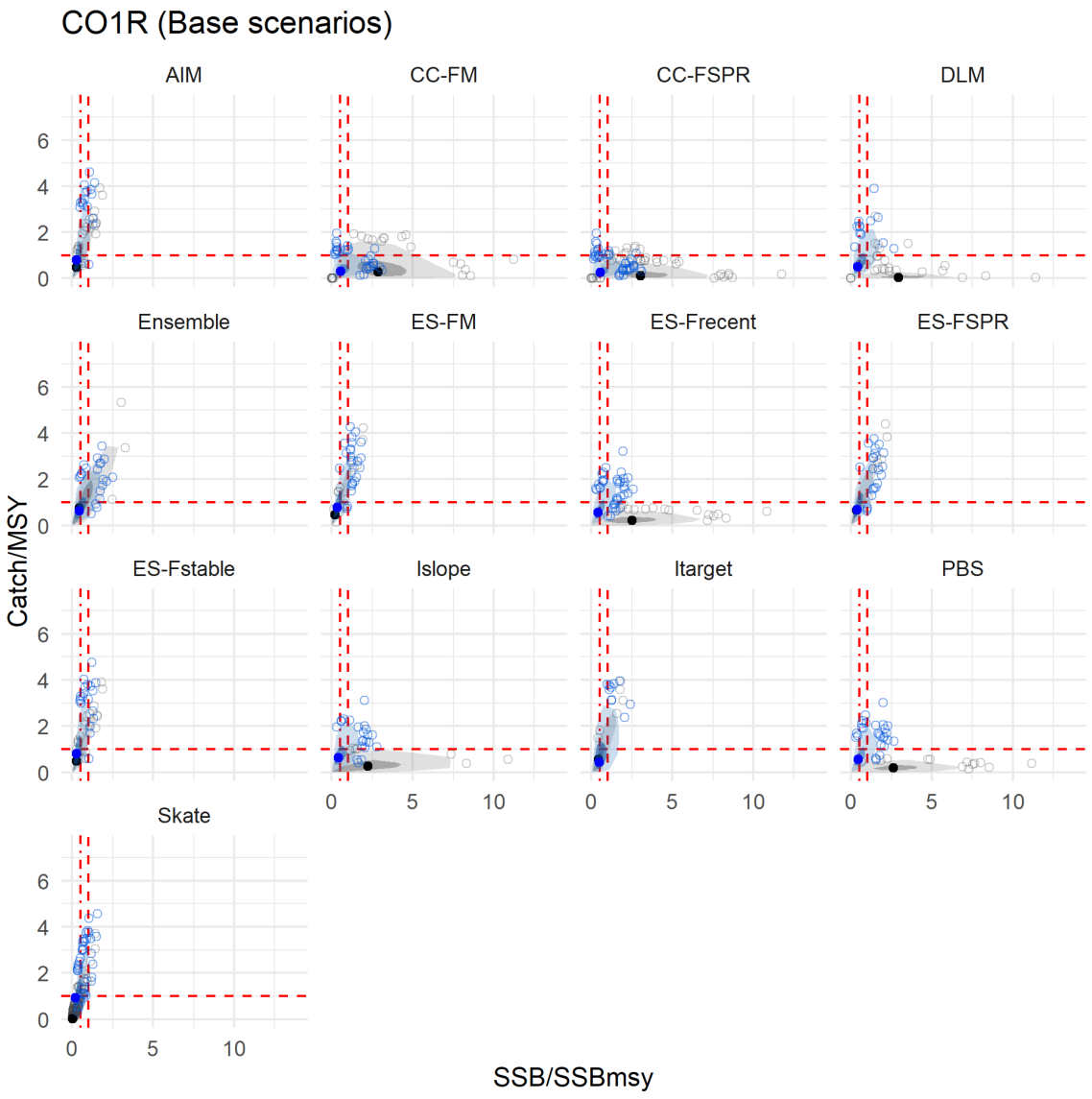


Figure A6.102. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

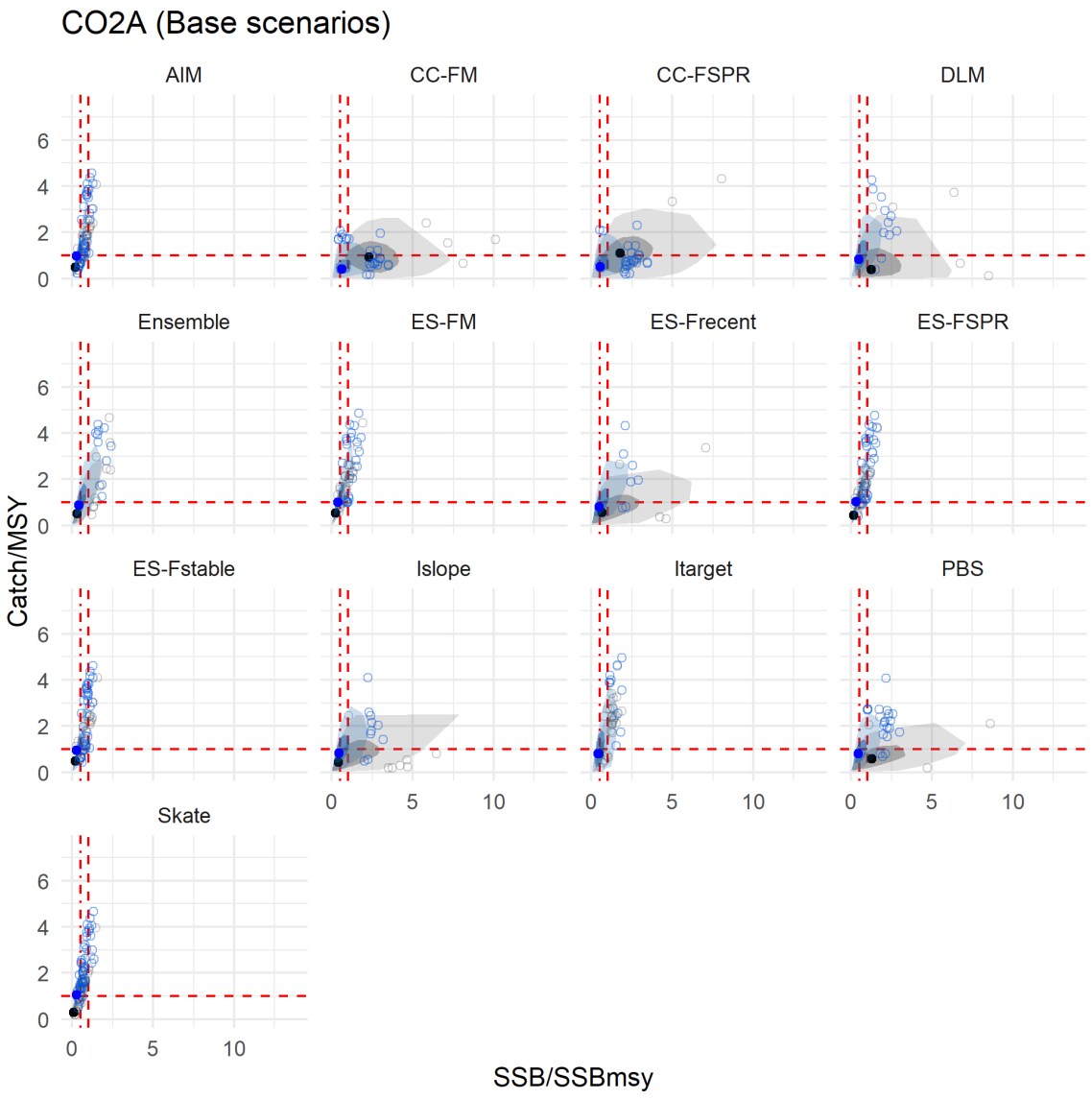


Figure A6.103. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

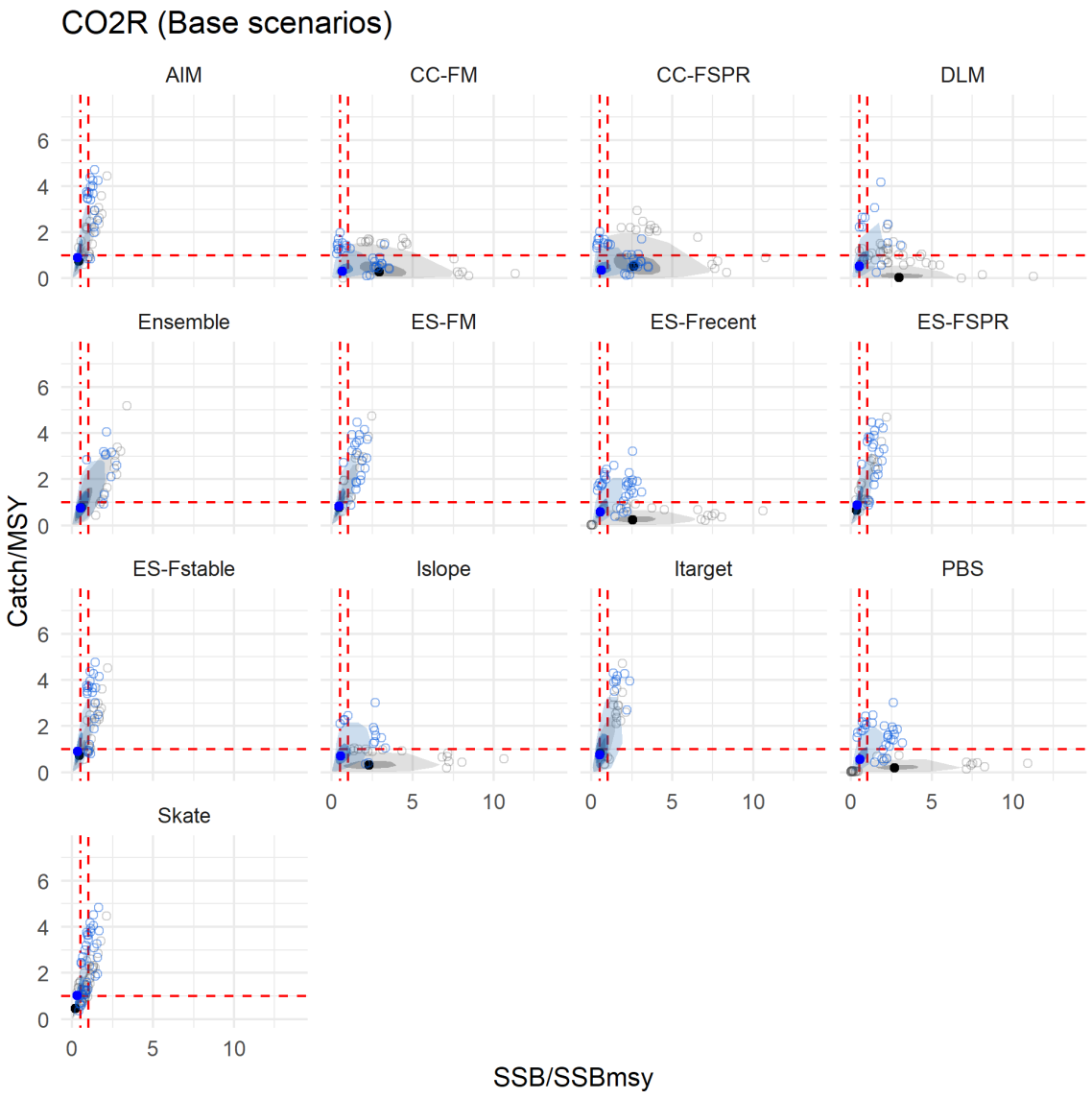


Figure A6.104. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

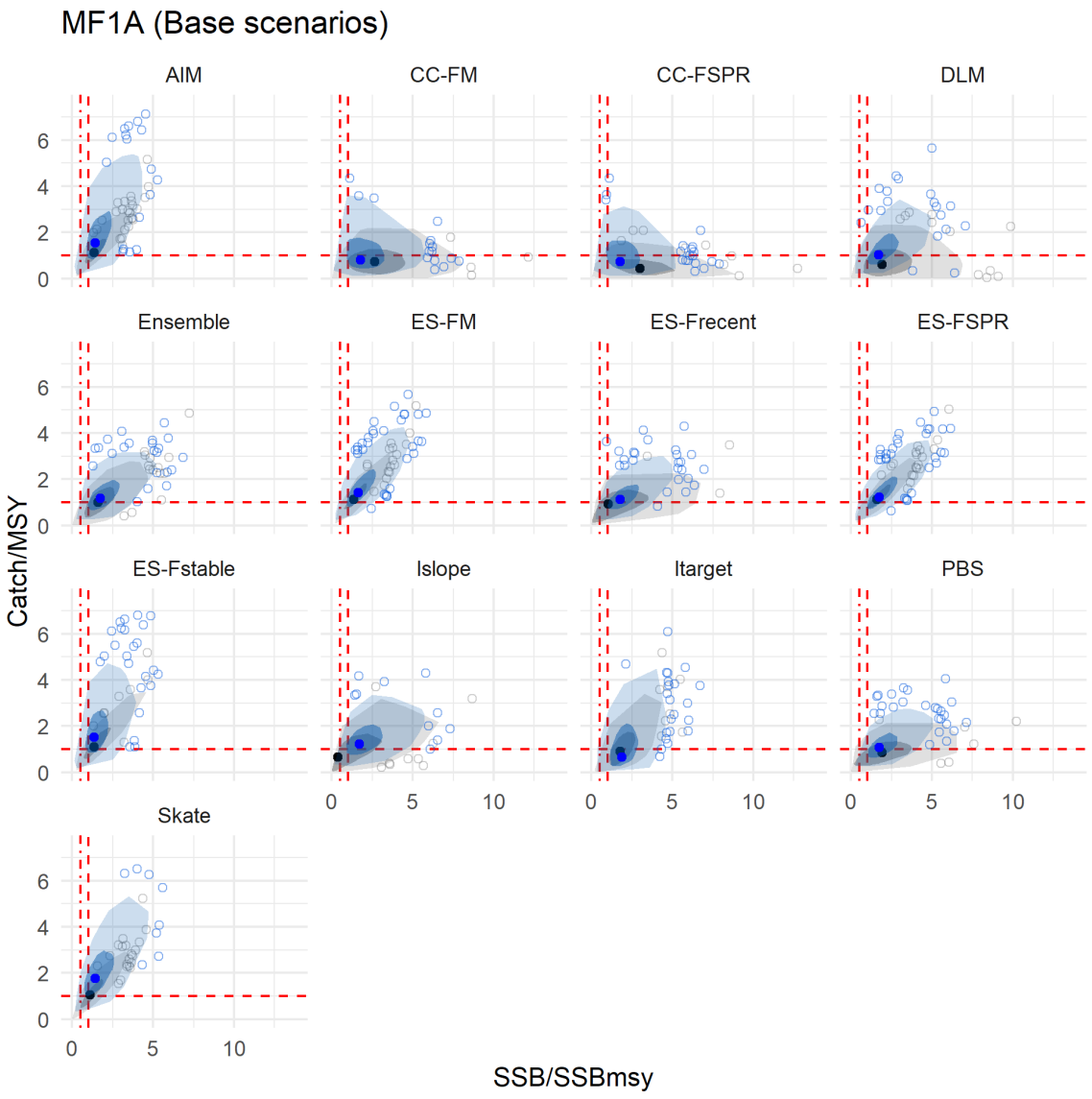


Figure A6.105. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

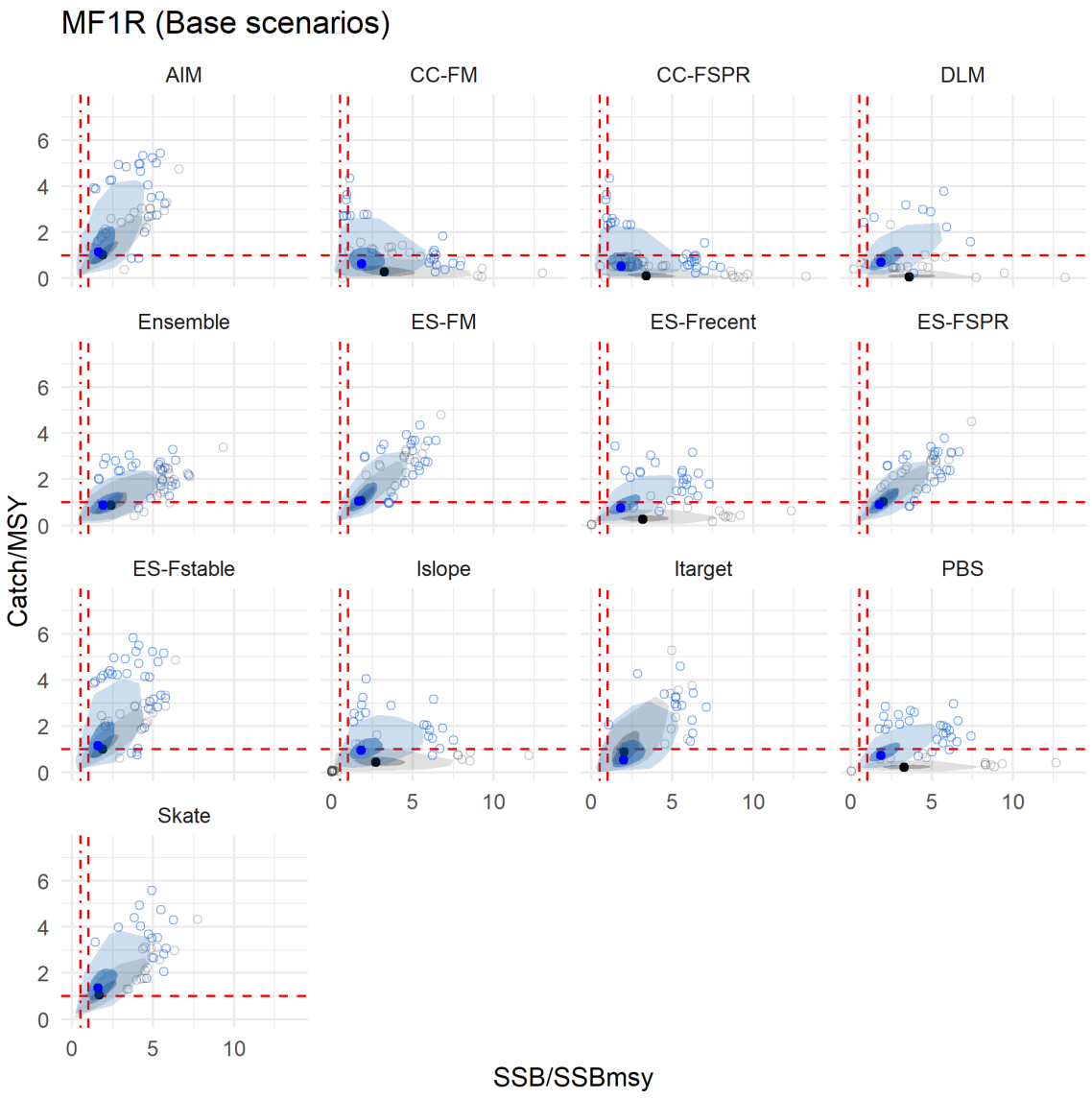


Figure A6.106. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

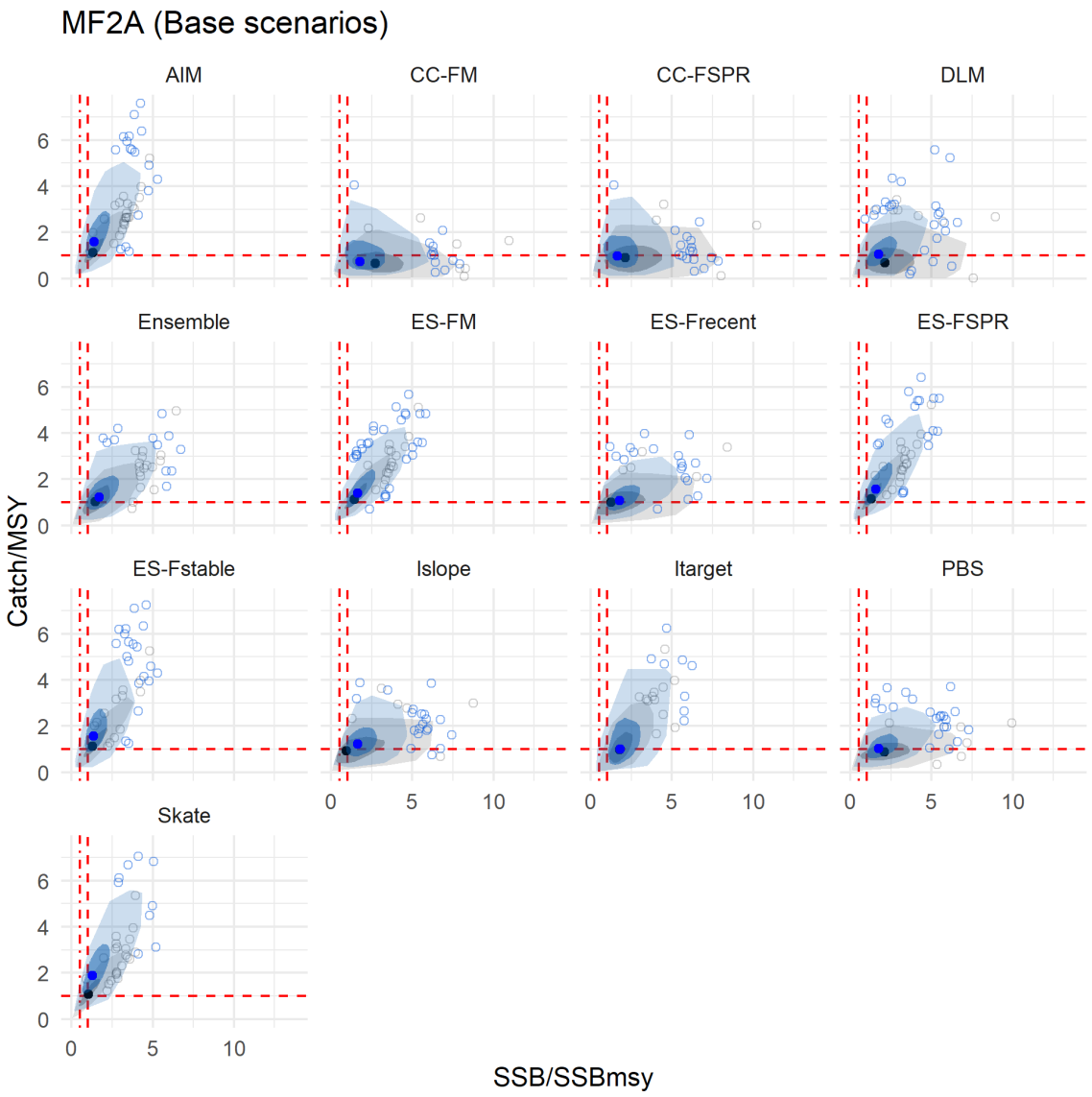


Figure A6.107. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

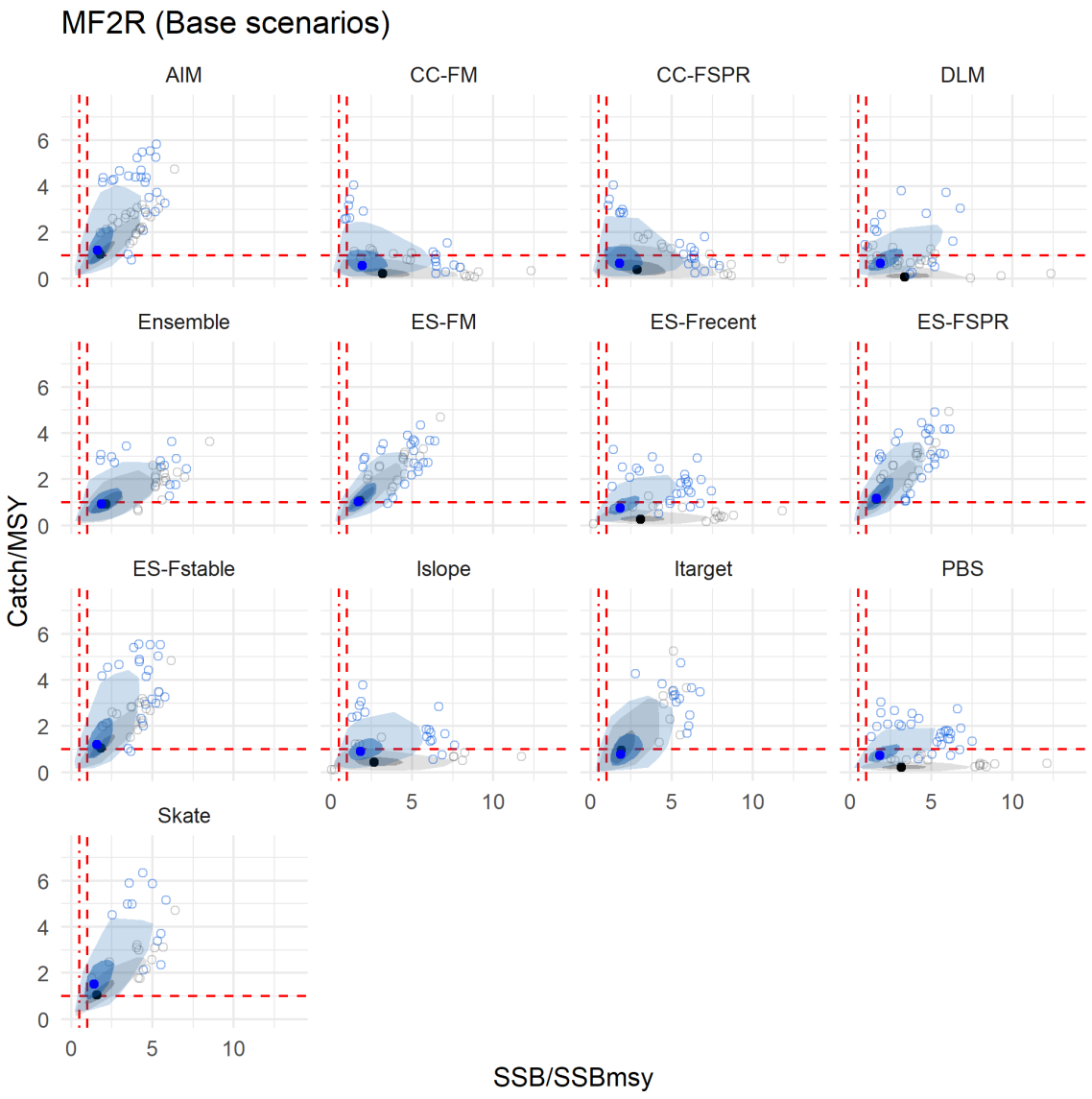


Figure A6.108. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

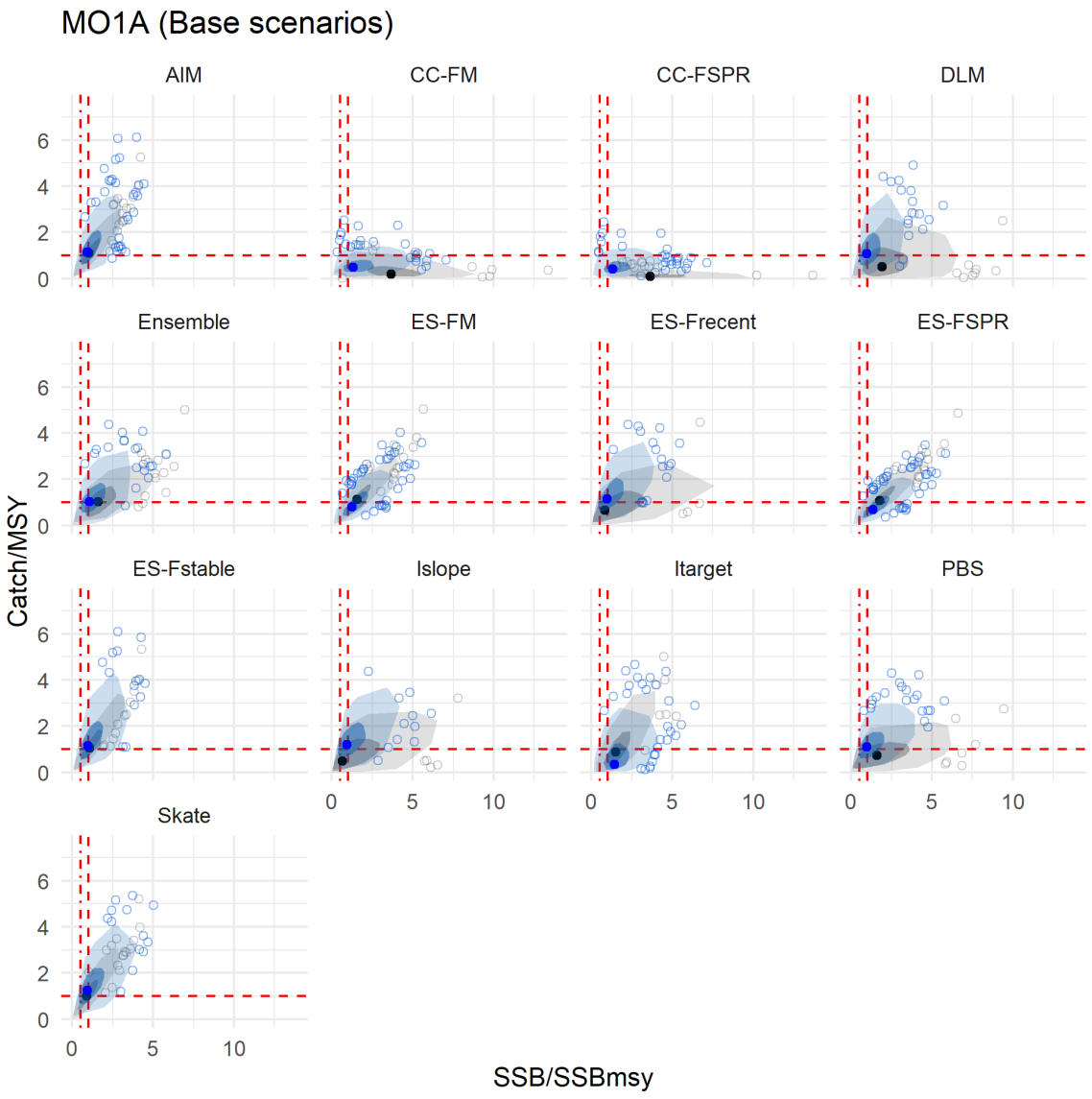


Figure A6.109. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

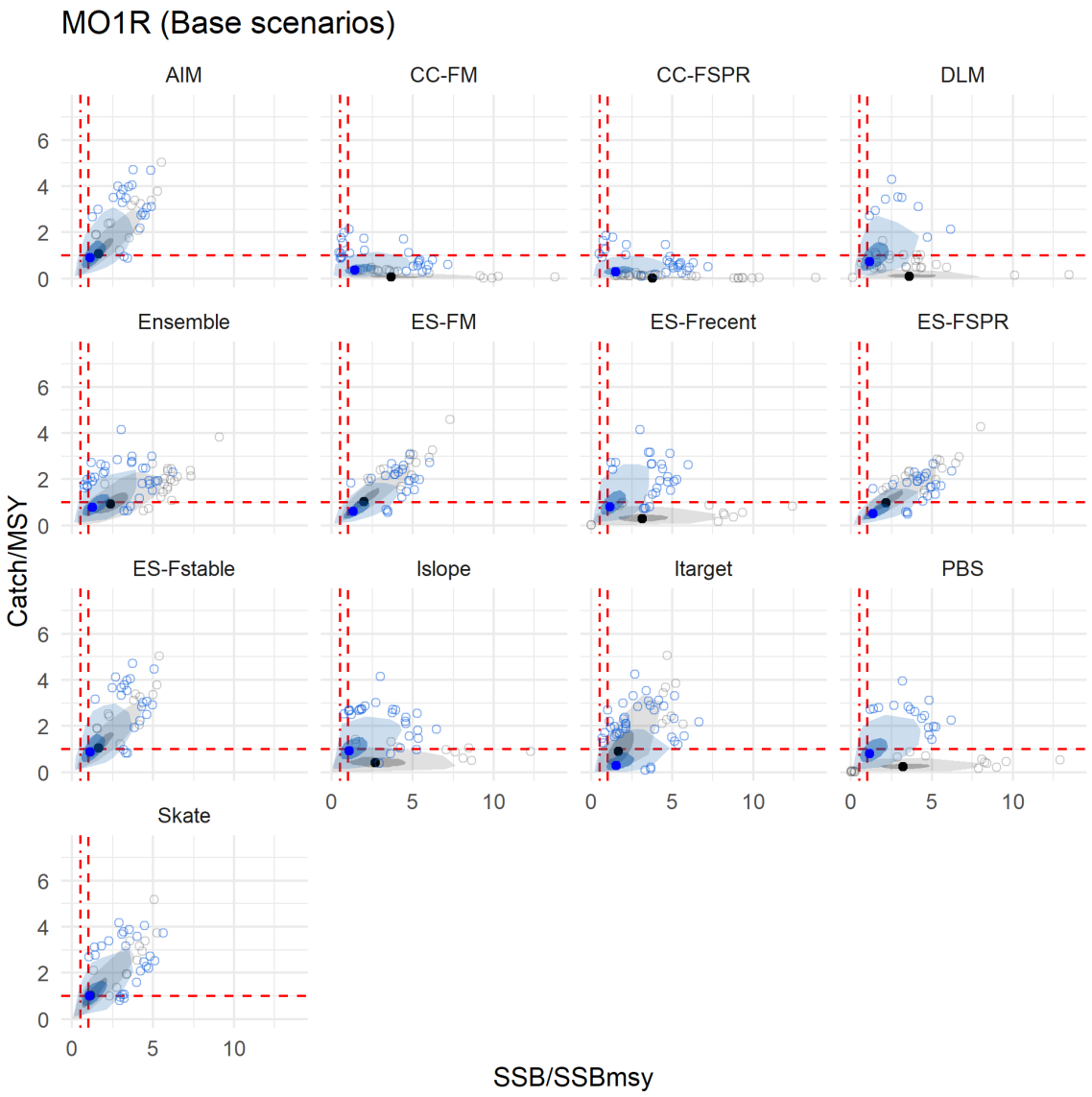


Figure A6.110. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

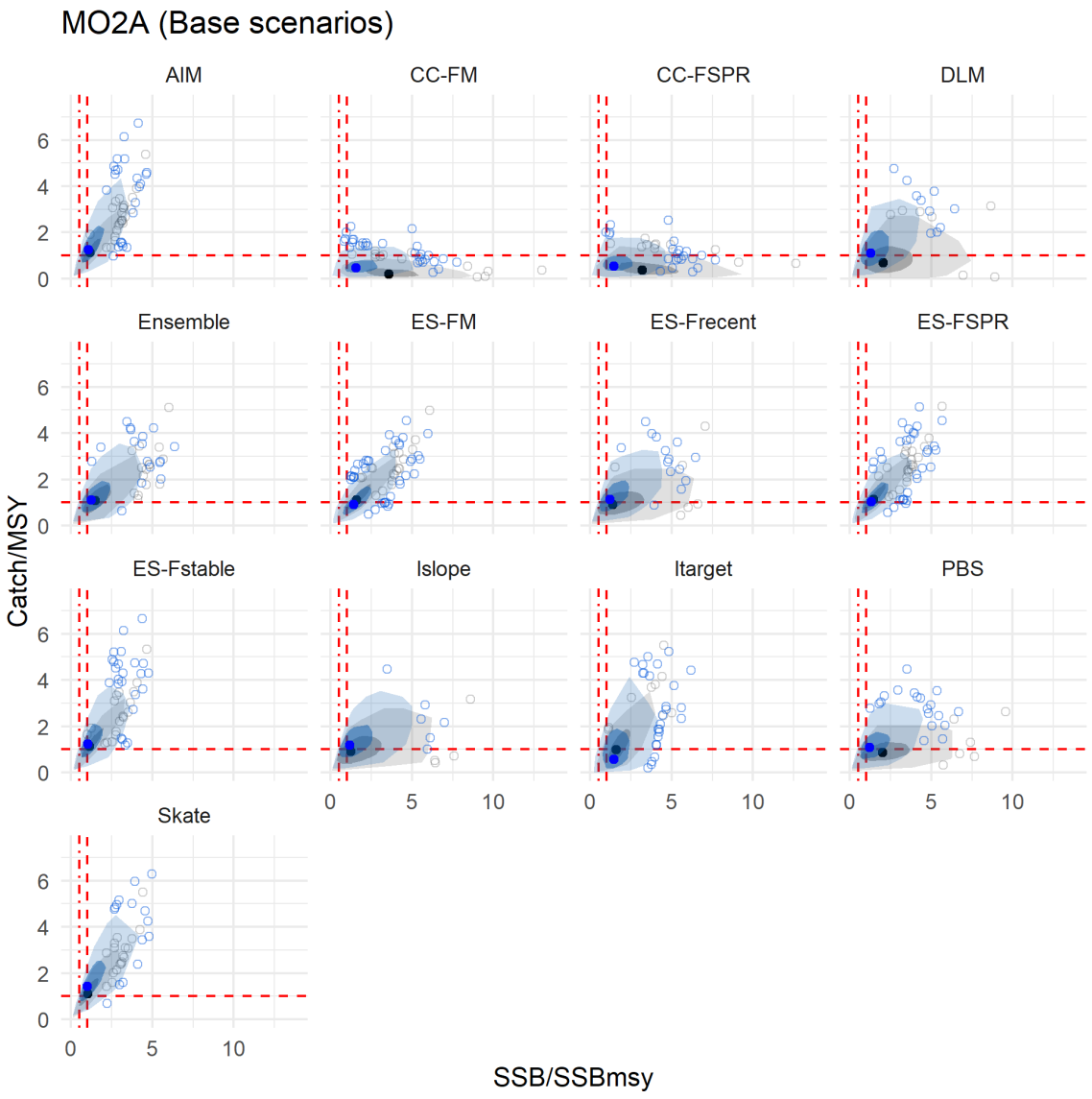


Figure A6.111. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

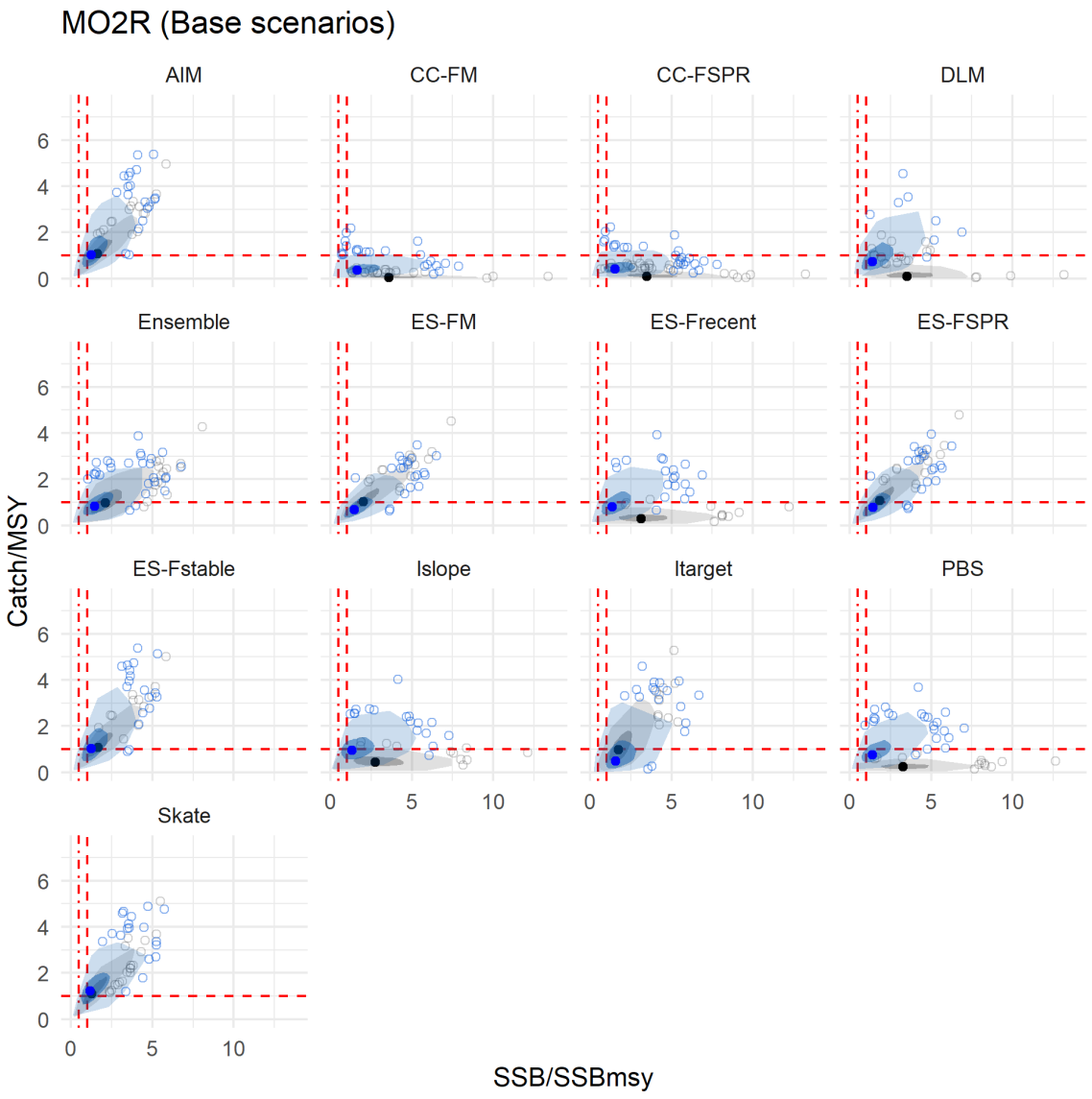


Figure A6.112. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

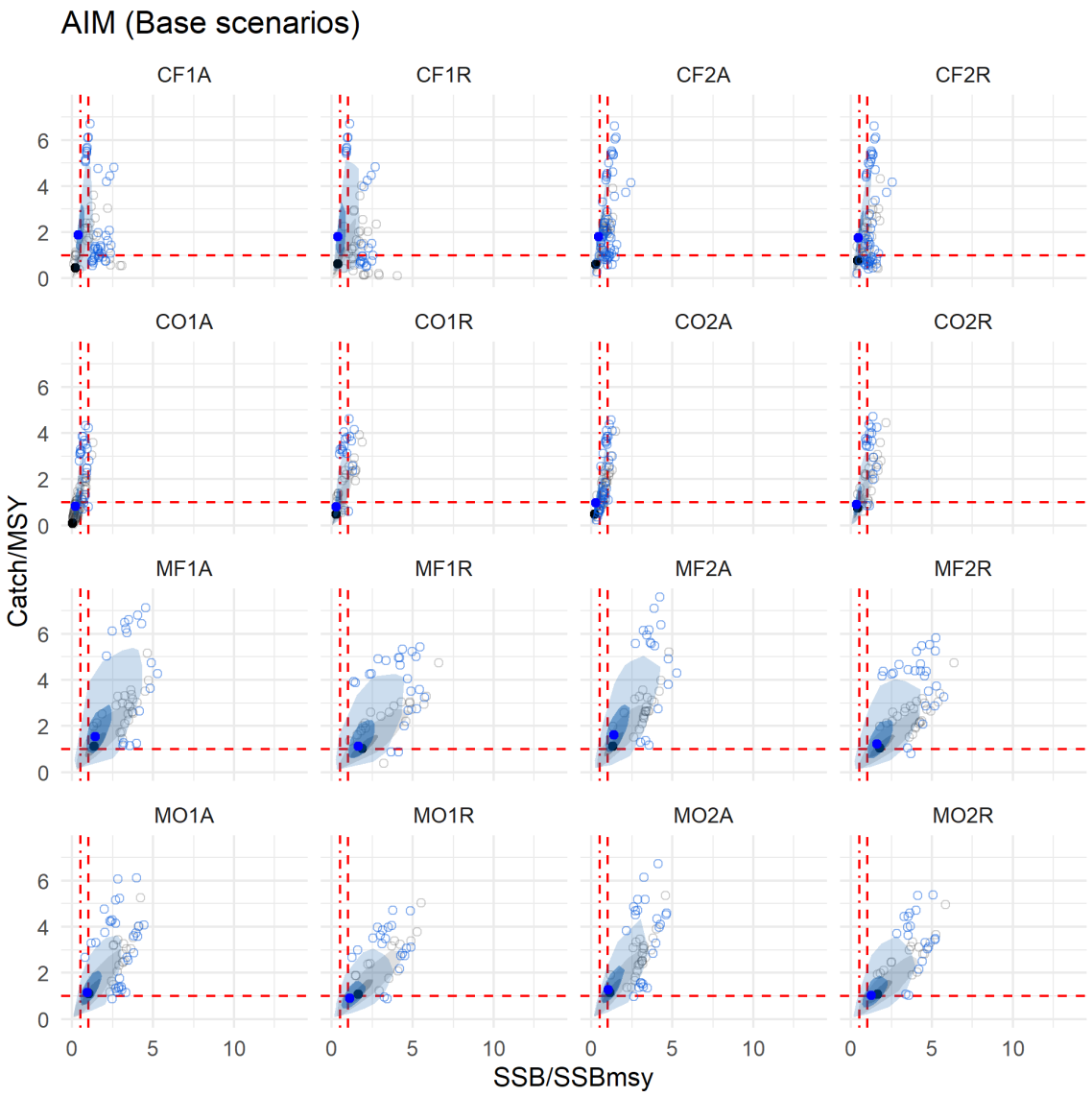


Figure A6.113. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

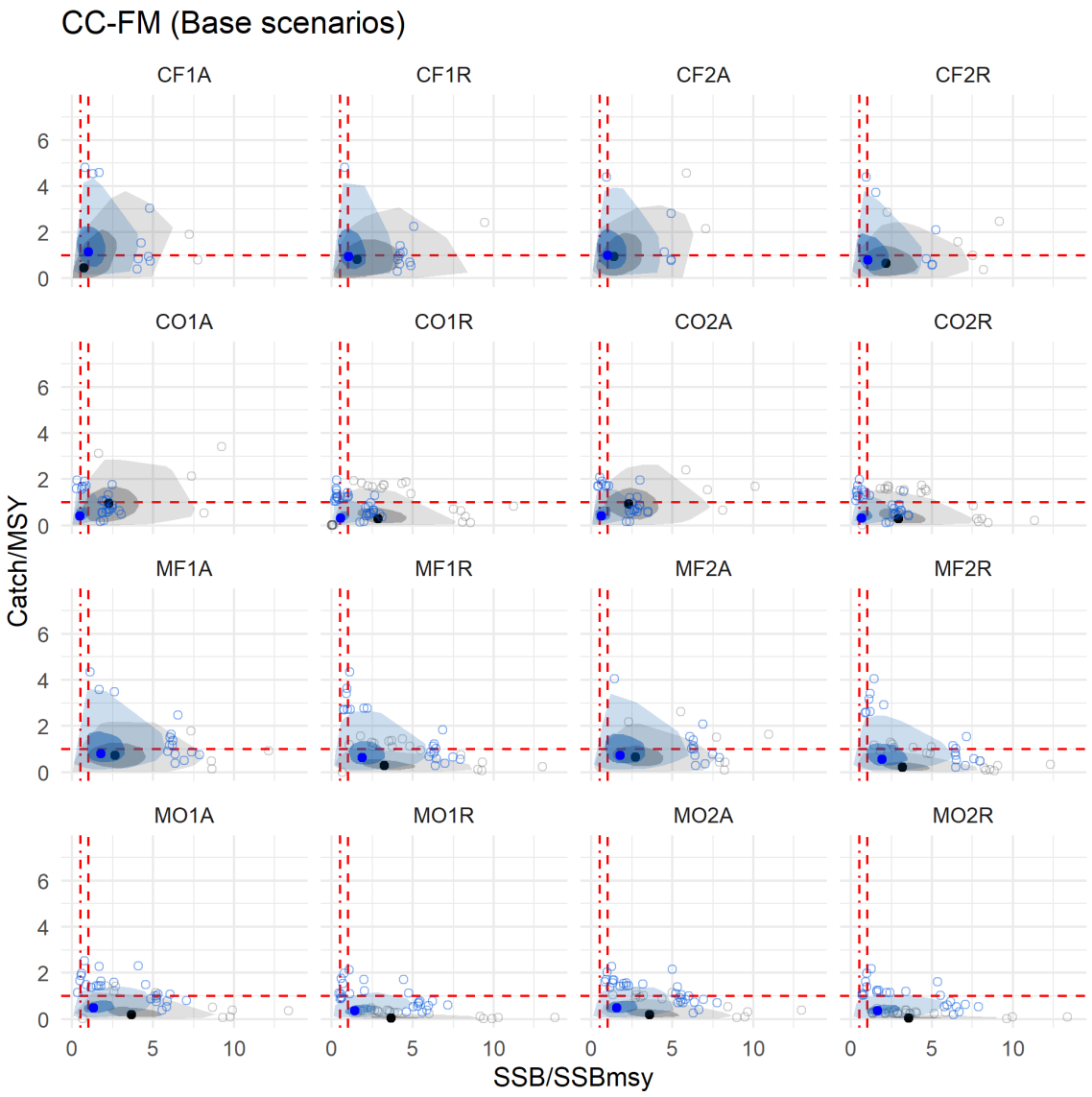


Figure A6.114. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

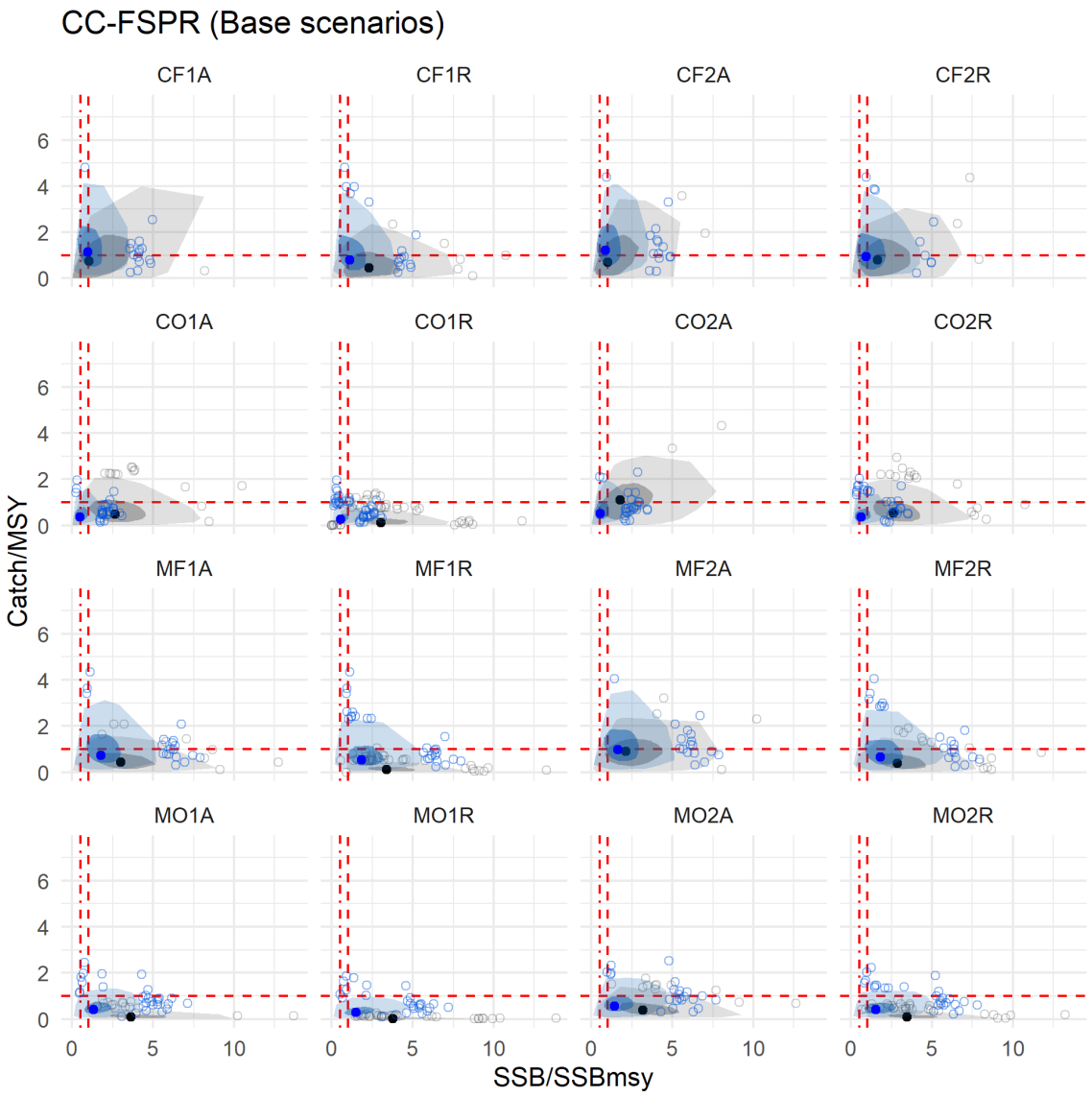


Figure A6.115. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

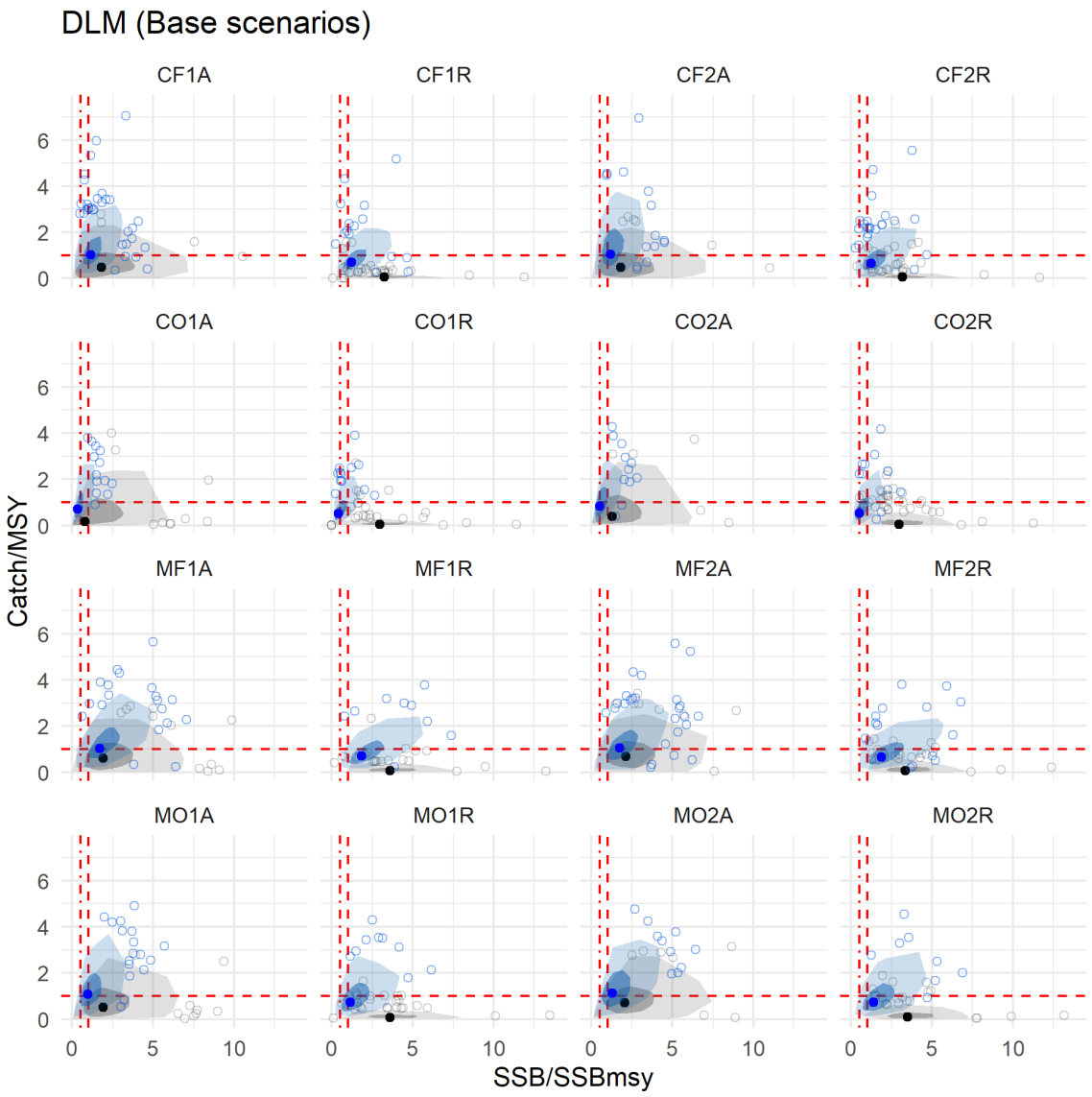


Figure A6.116. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

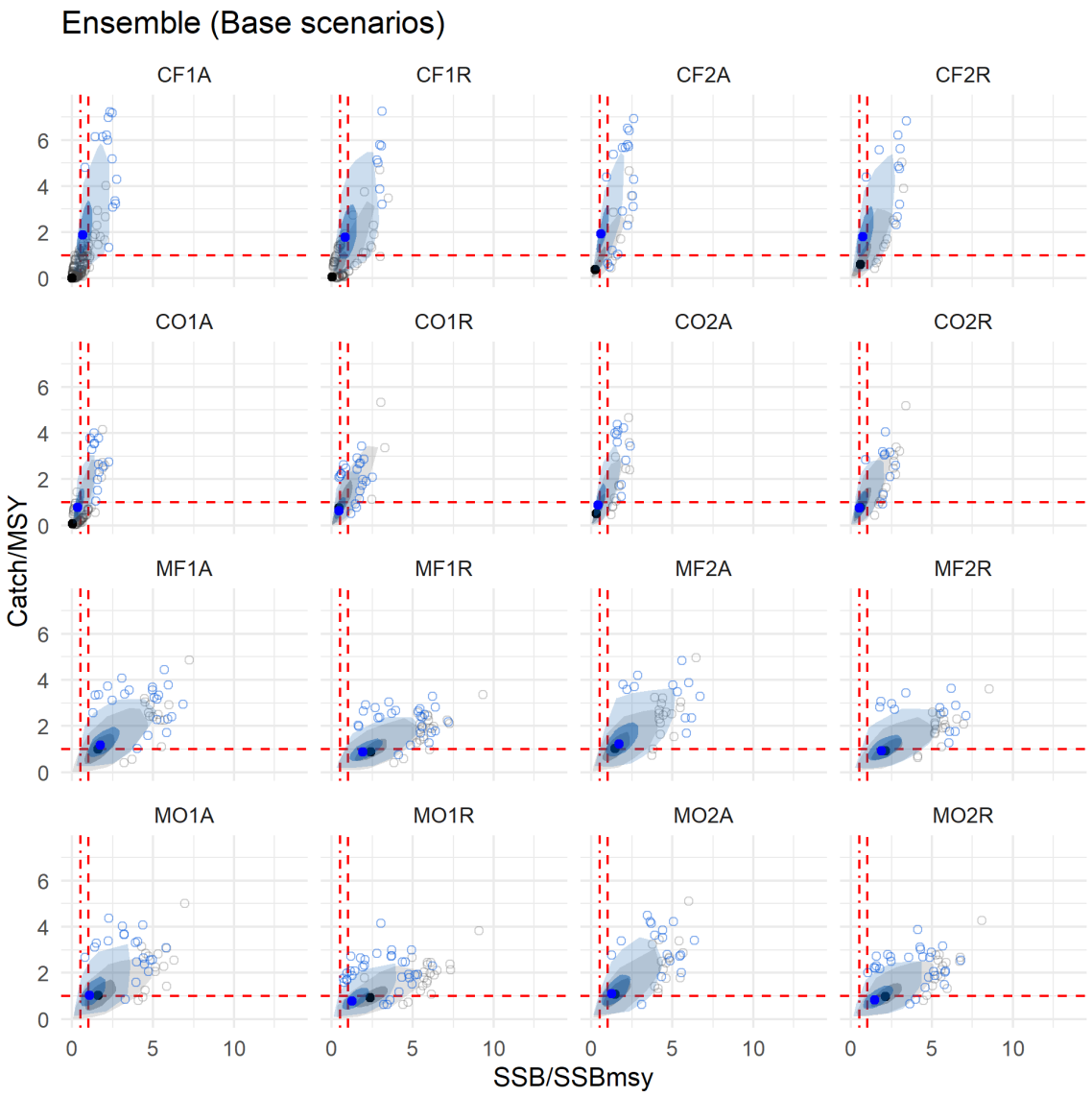


Figure A6.117. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

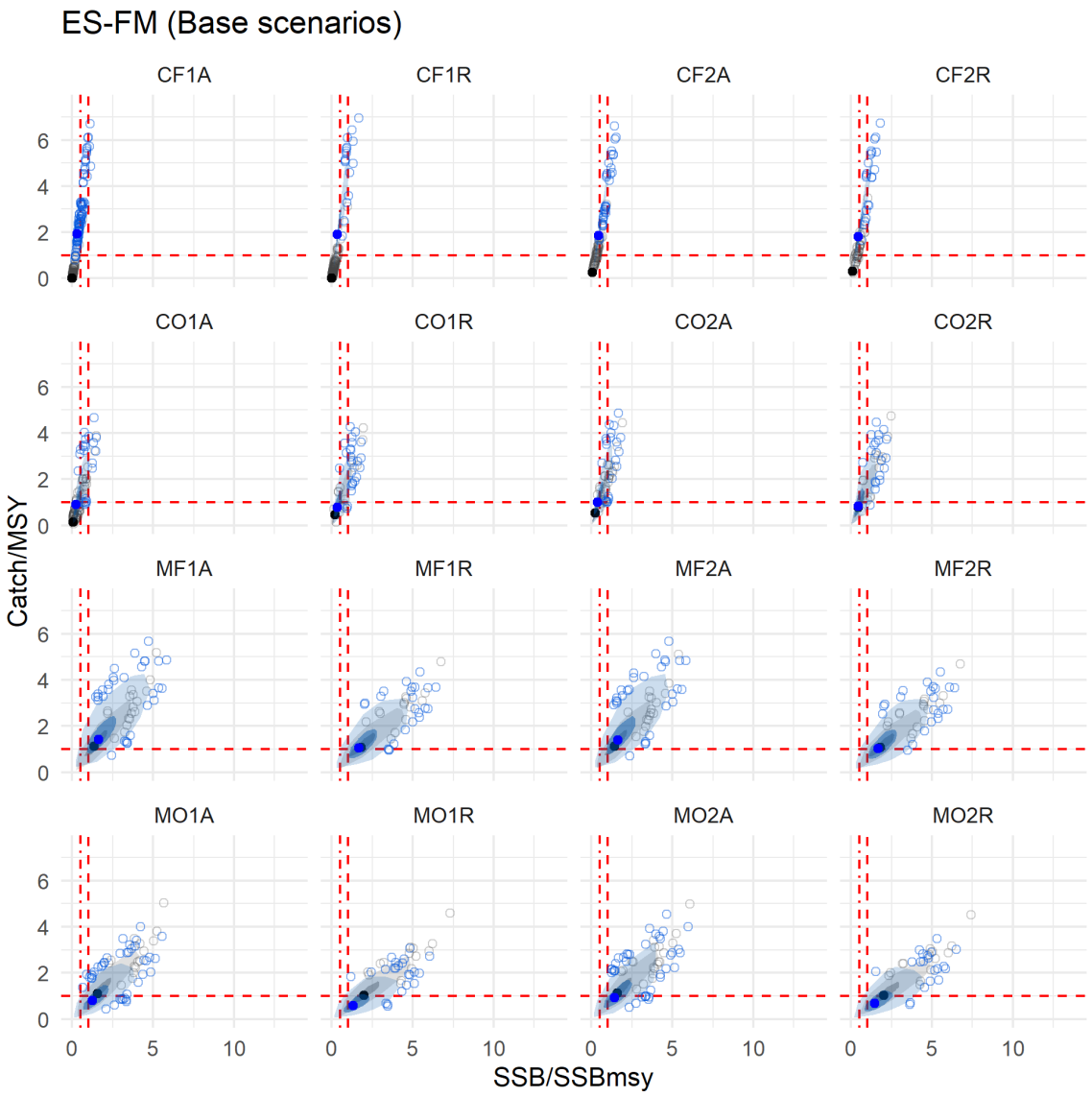


Figure A6.118. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

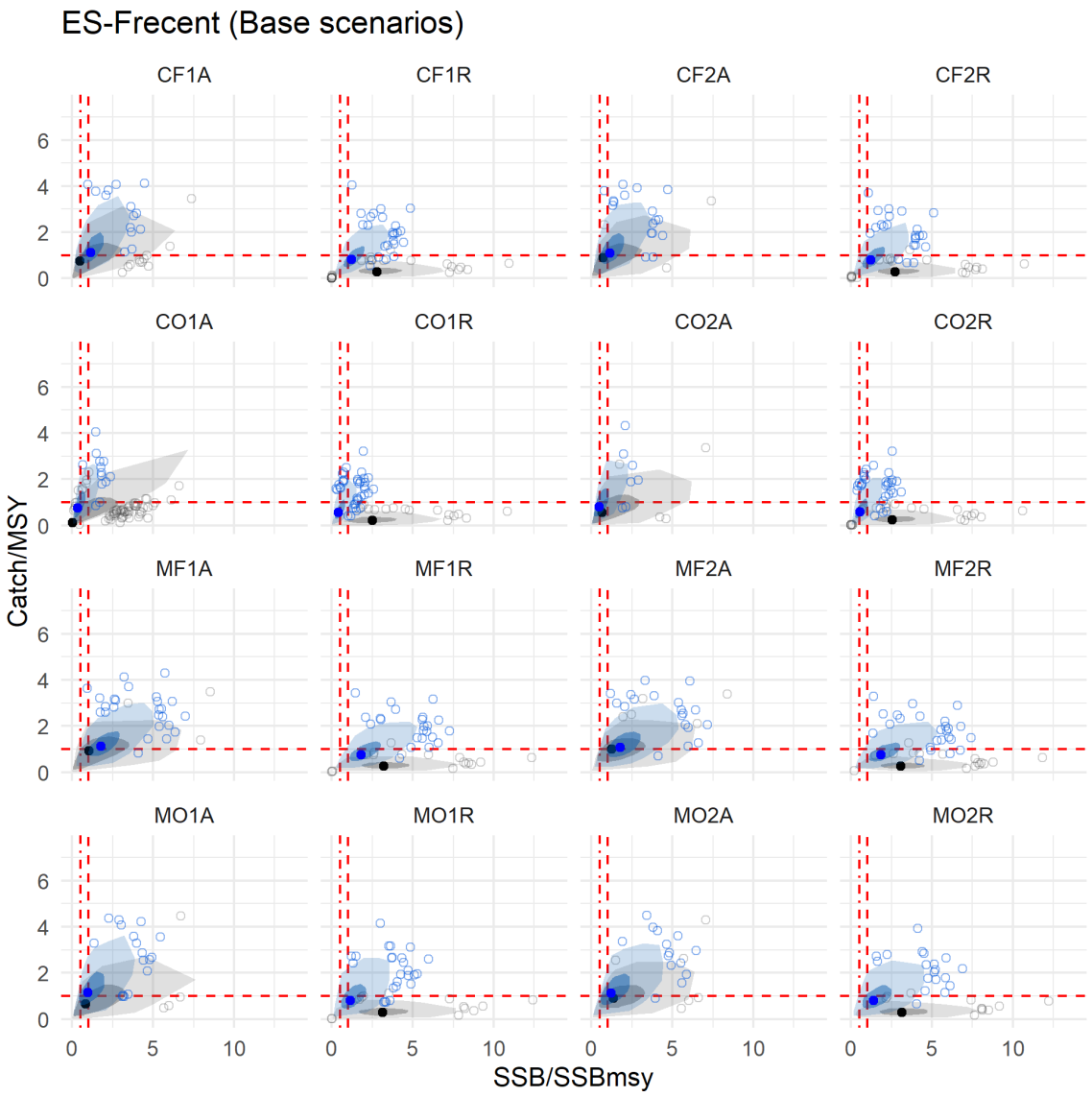


Figure A6.119. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

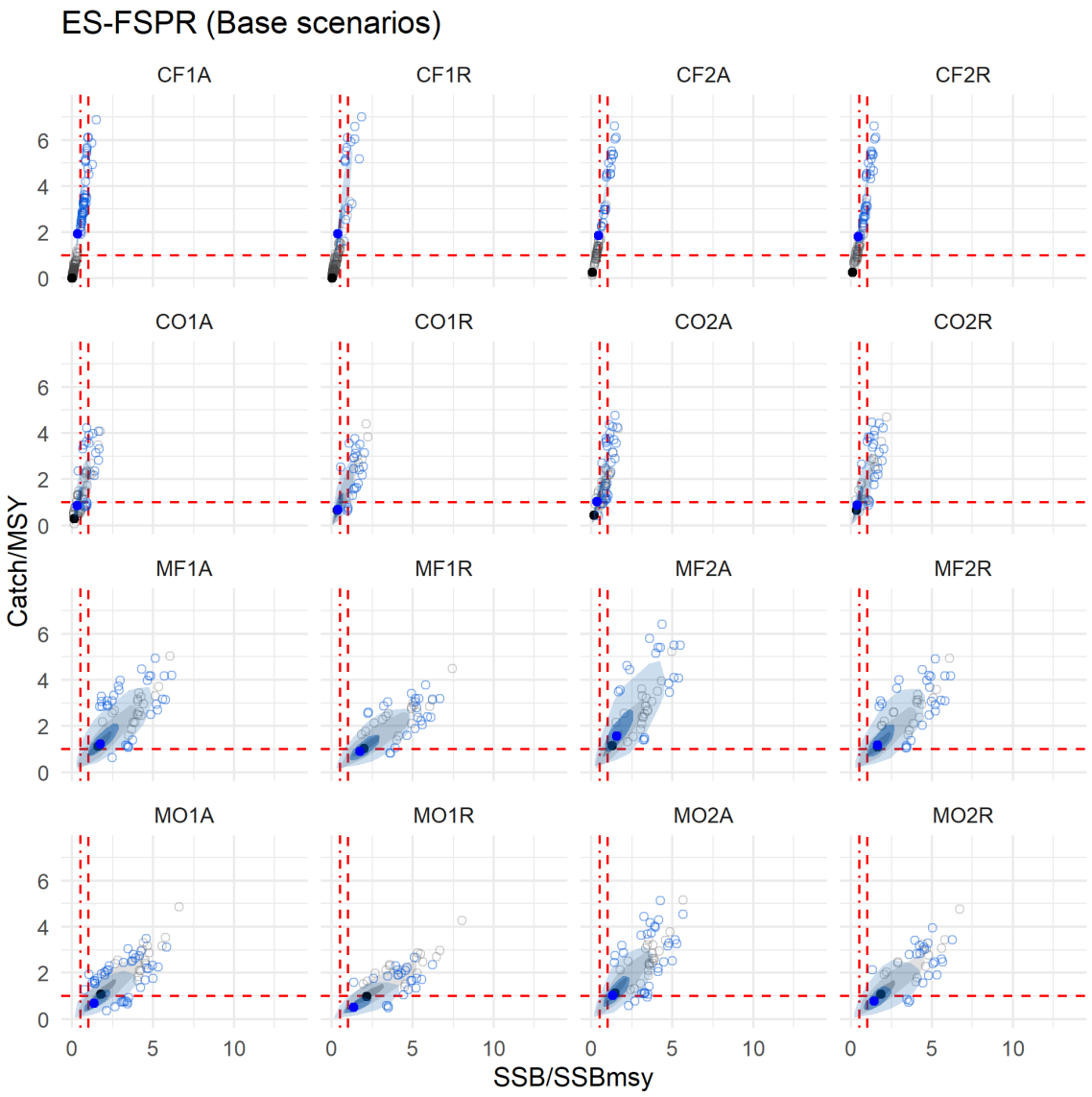


Figure A6.120. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

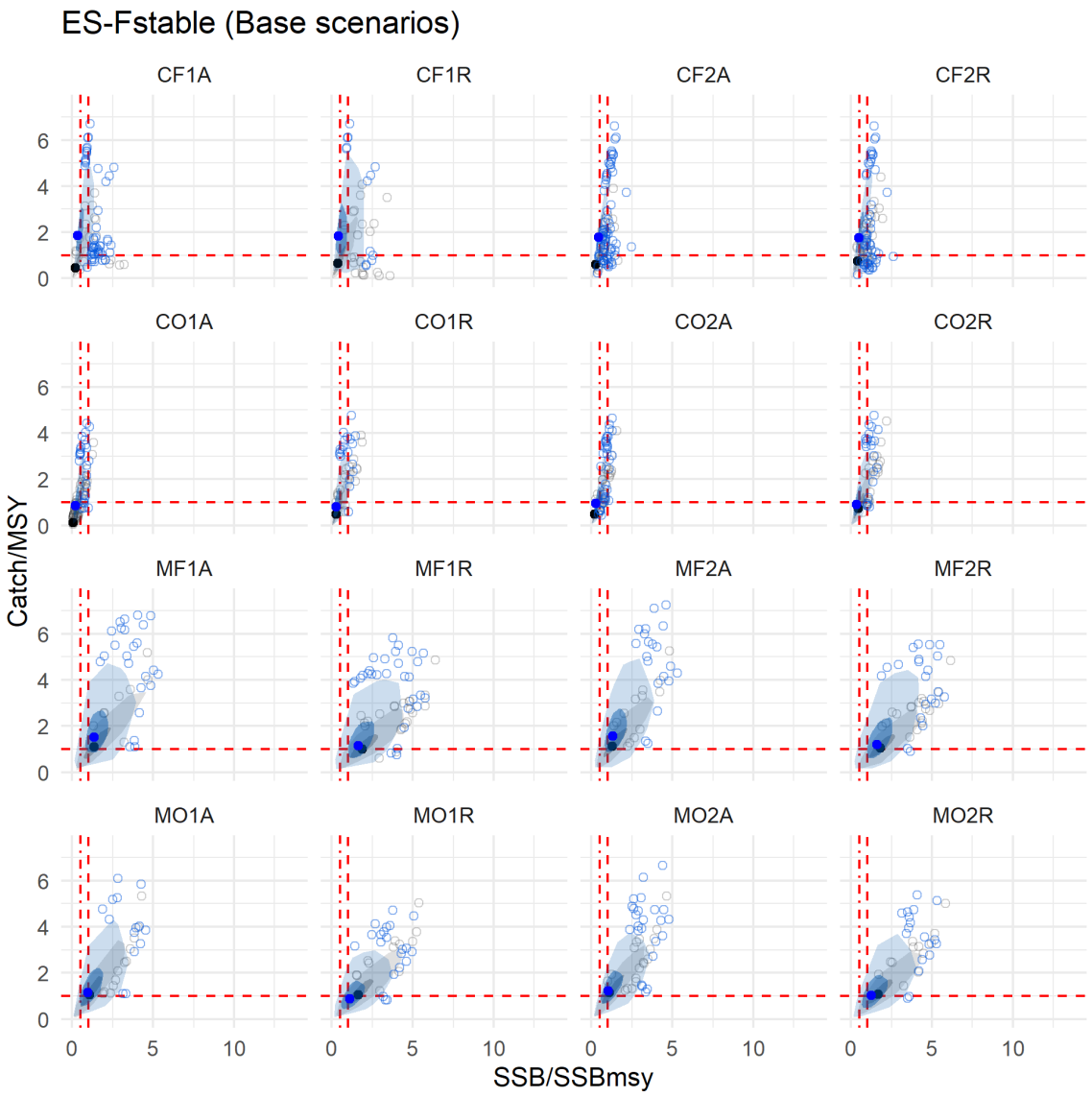


Figure A6.121. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

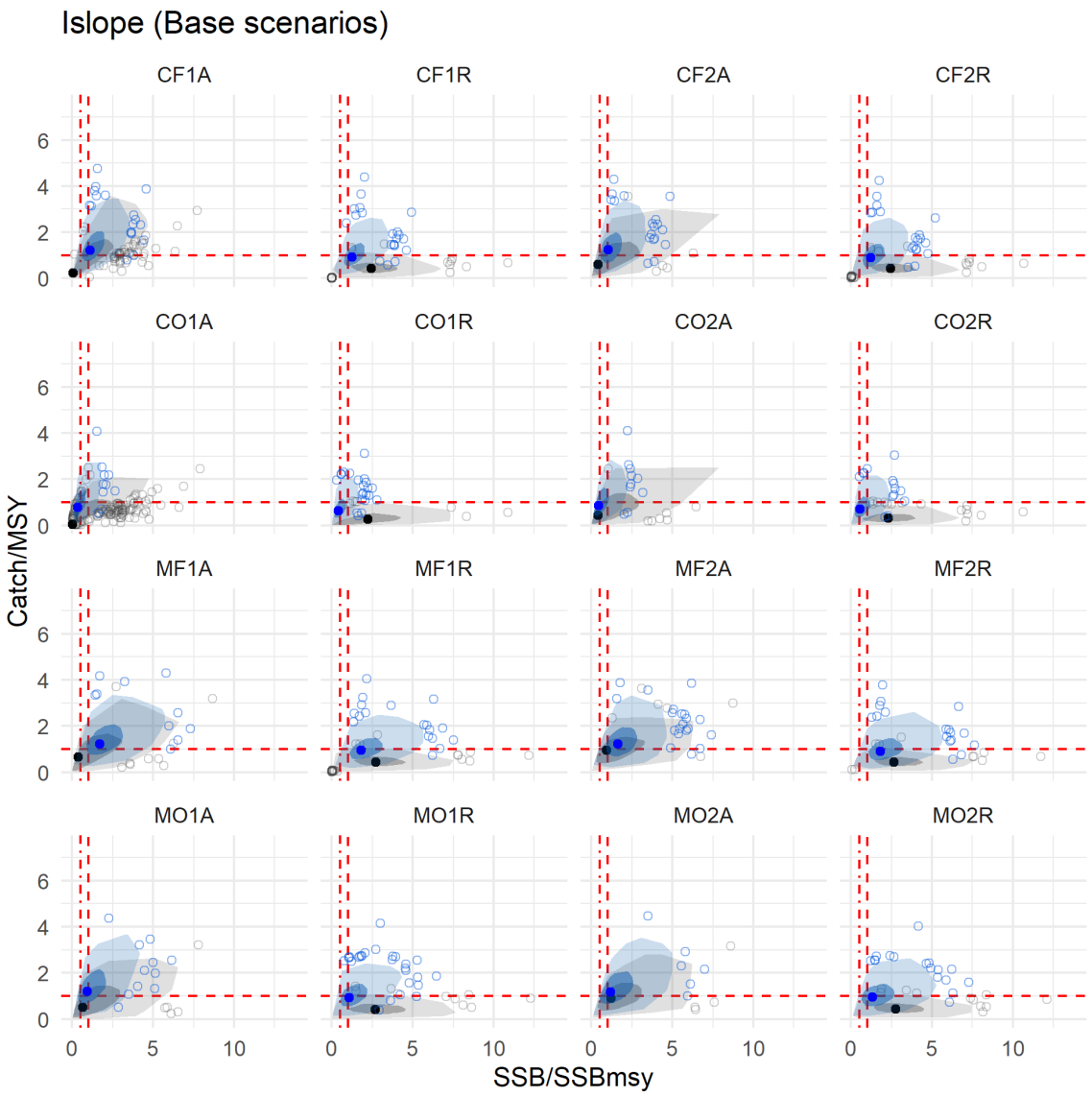


Figure A6.122. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

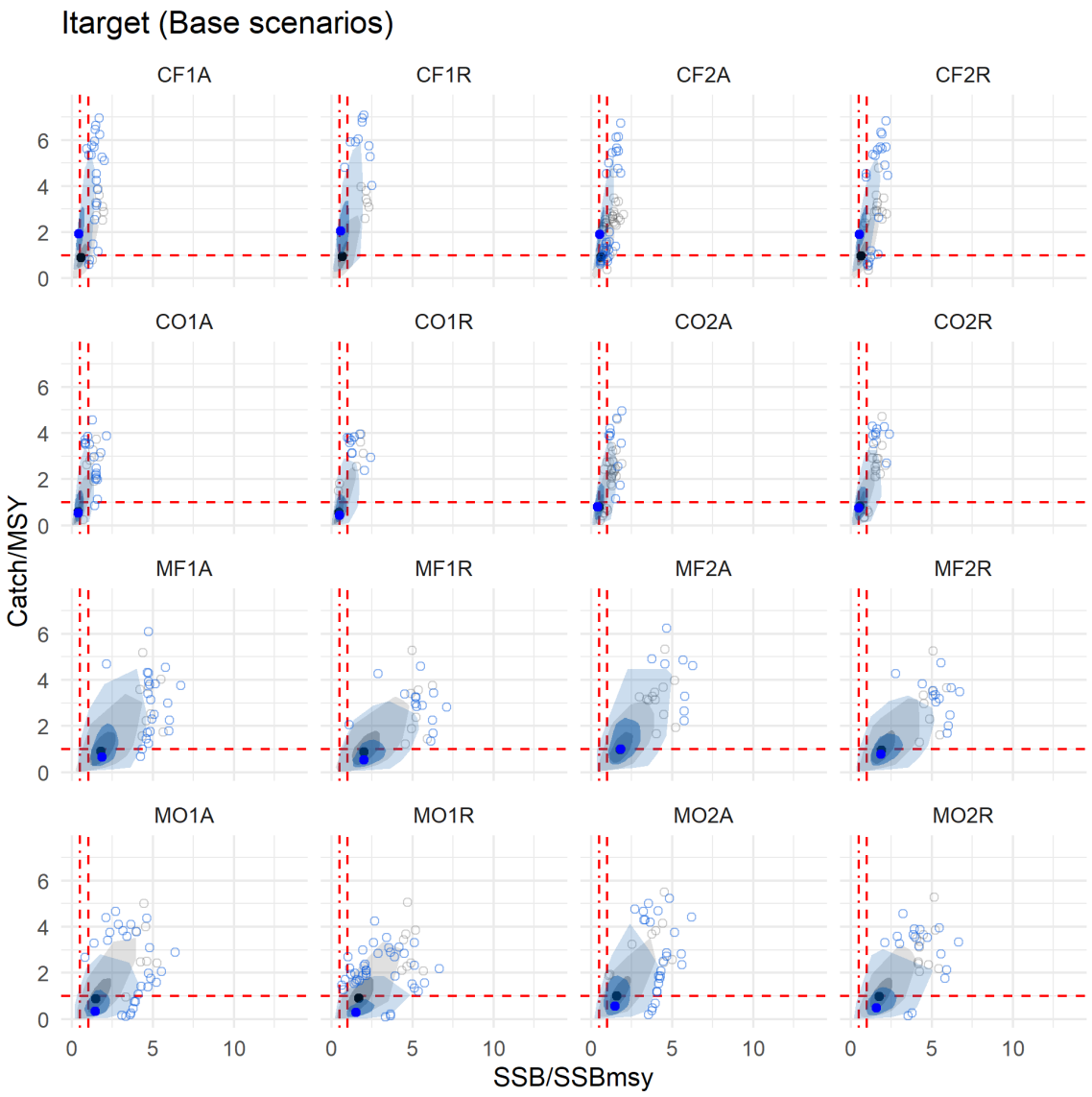


Figure A6.123. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

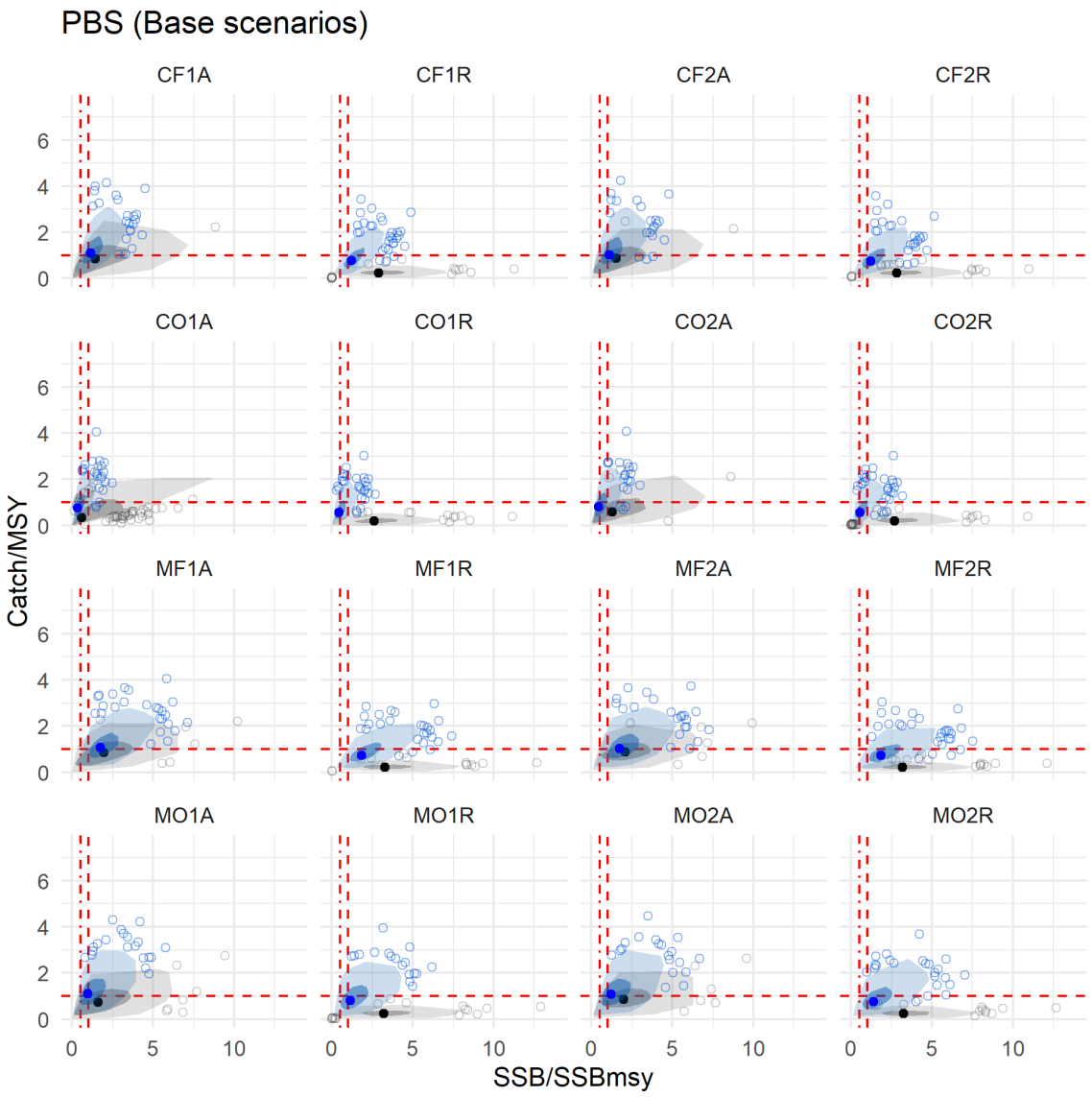


Figure A6.124. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

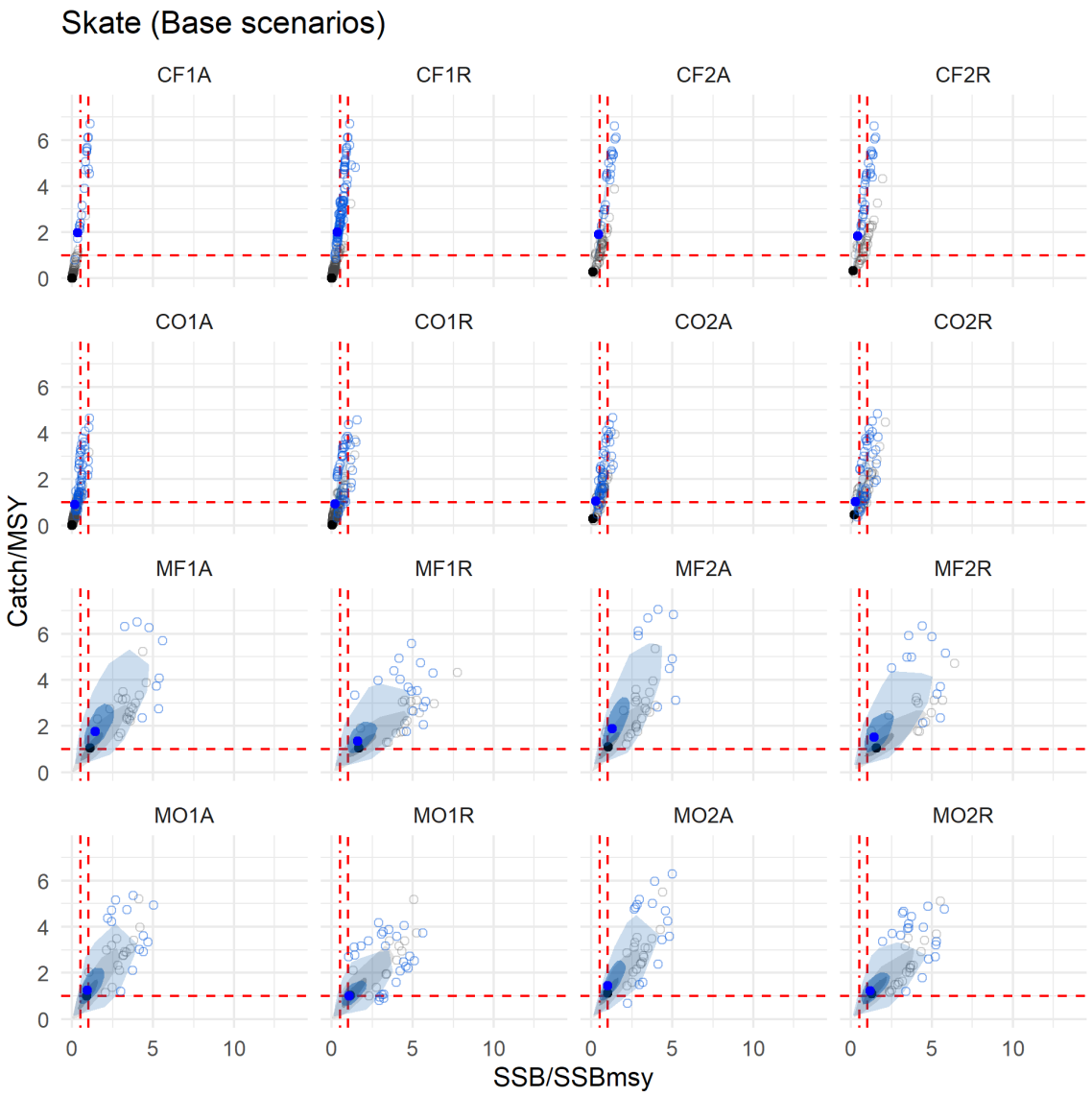


Figure A6.125. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

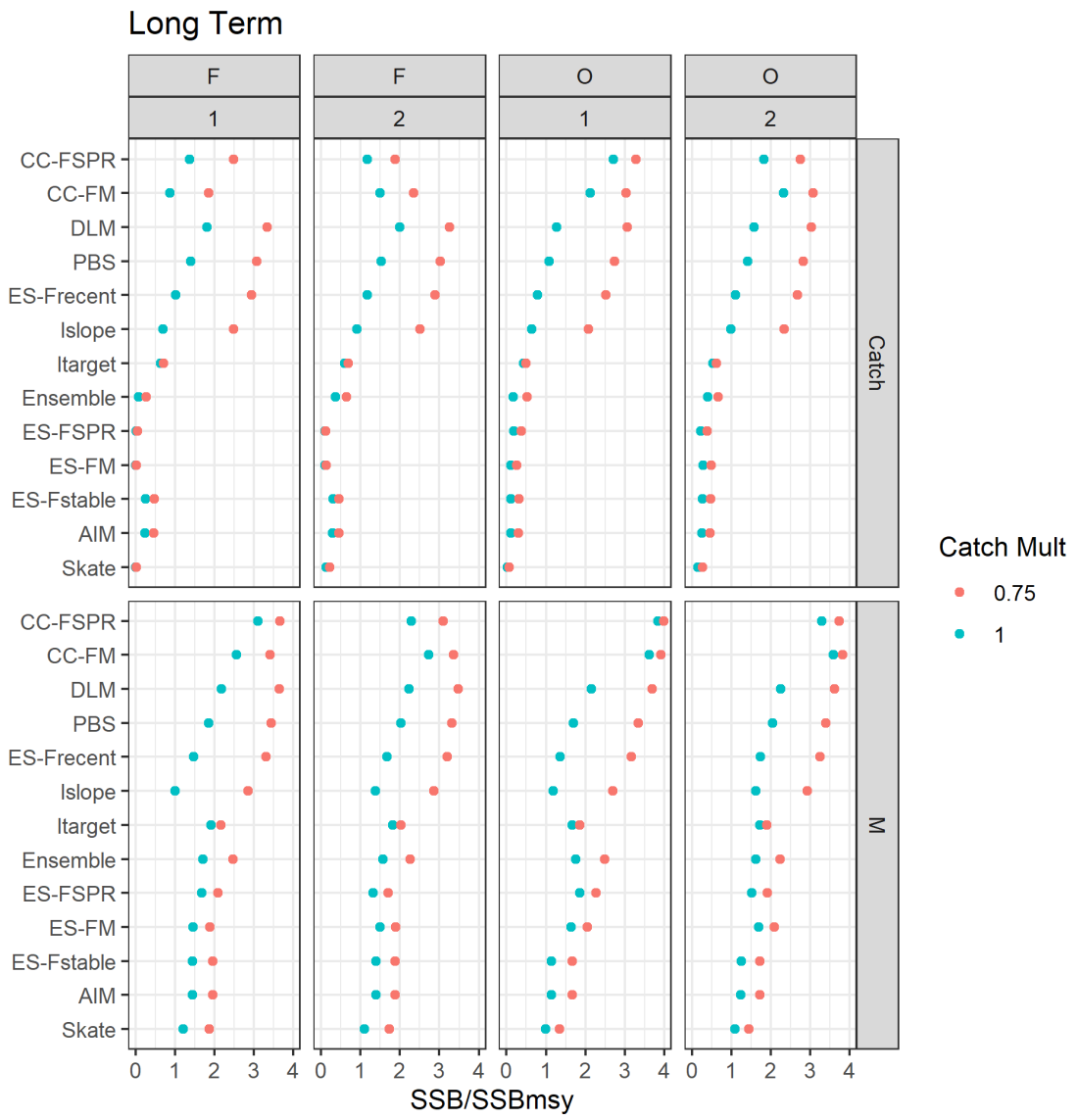


Figure A6.126. Mean values of spawning stock biomass (SSB) relative to its maximum sustainable yield reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

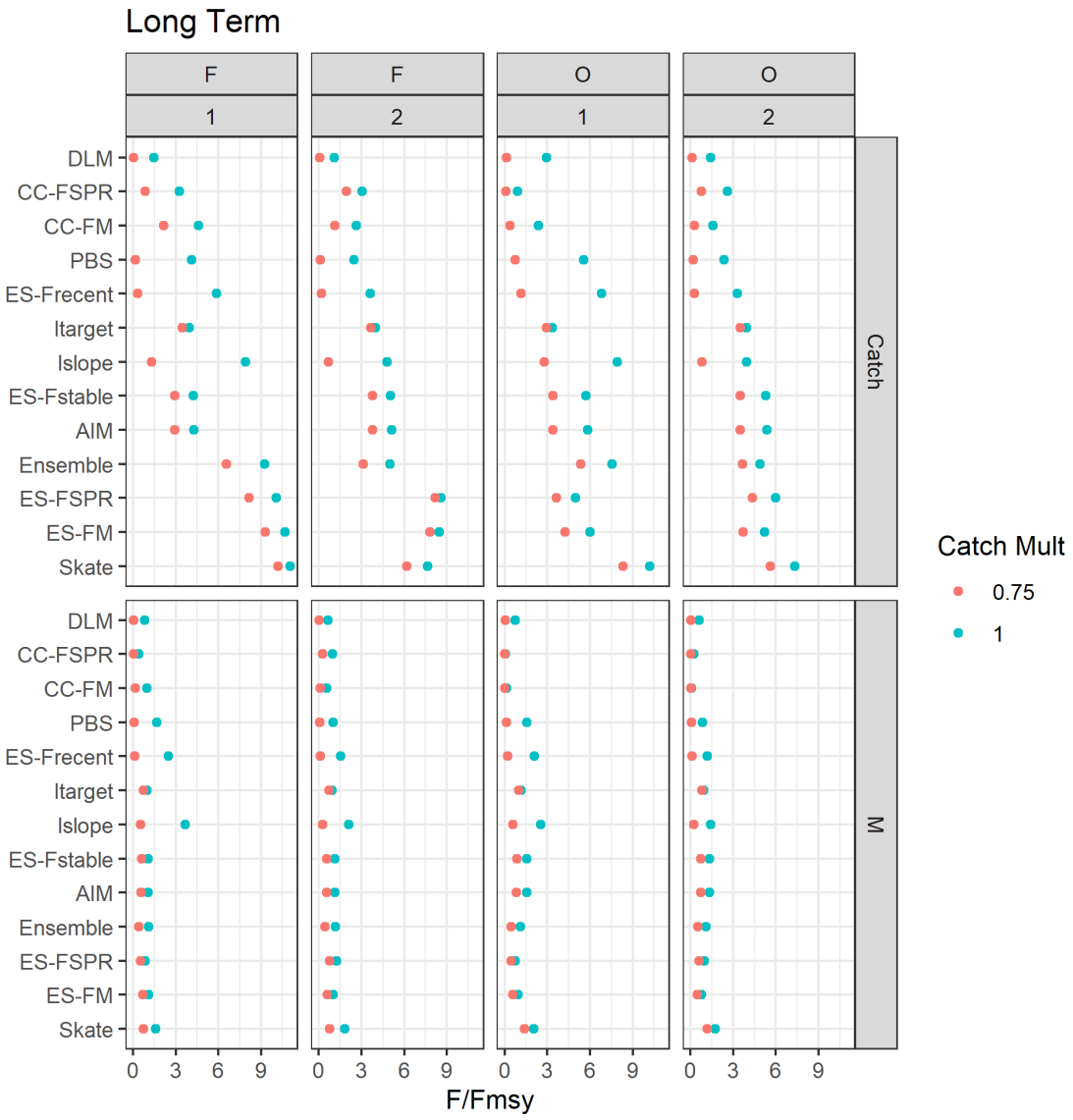


Figure A6.127. Mean values of fishing mortality rate (F) relative to its maximum sustainable yield reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of F history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

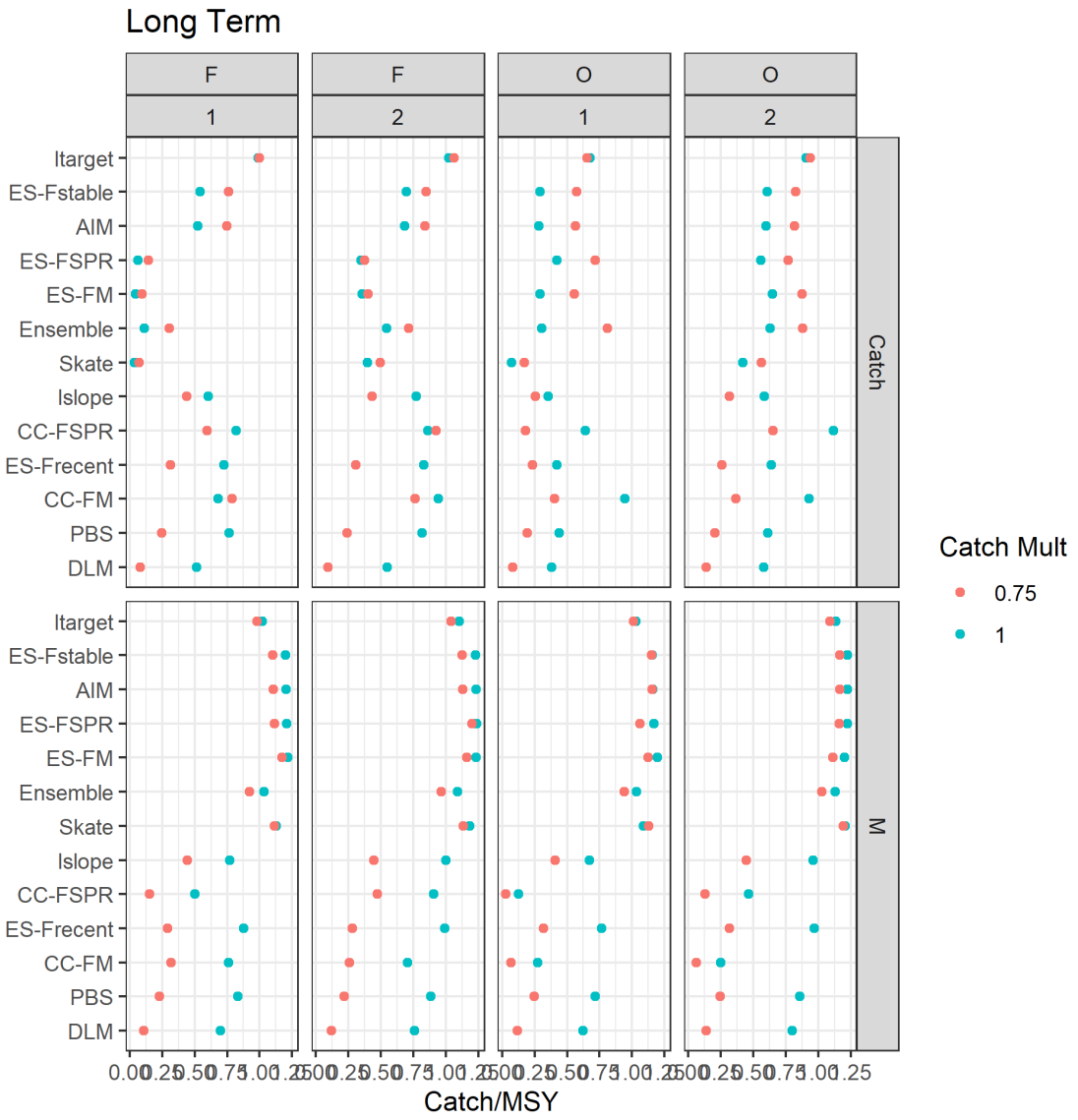


Figure A6.128. Mean values of catch relative to its maximum sustainable yield reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

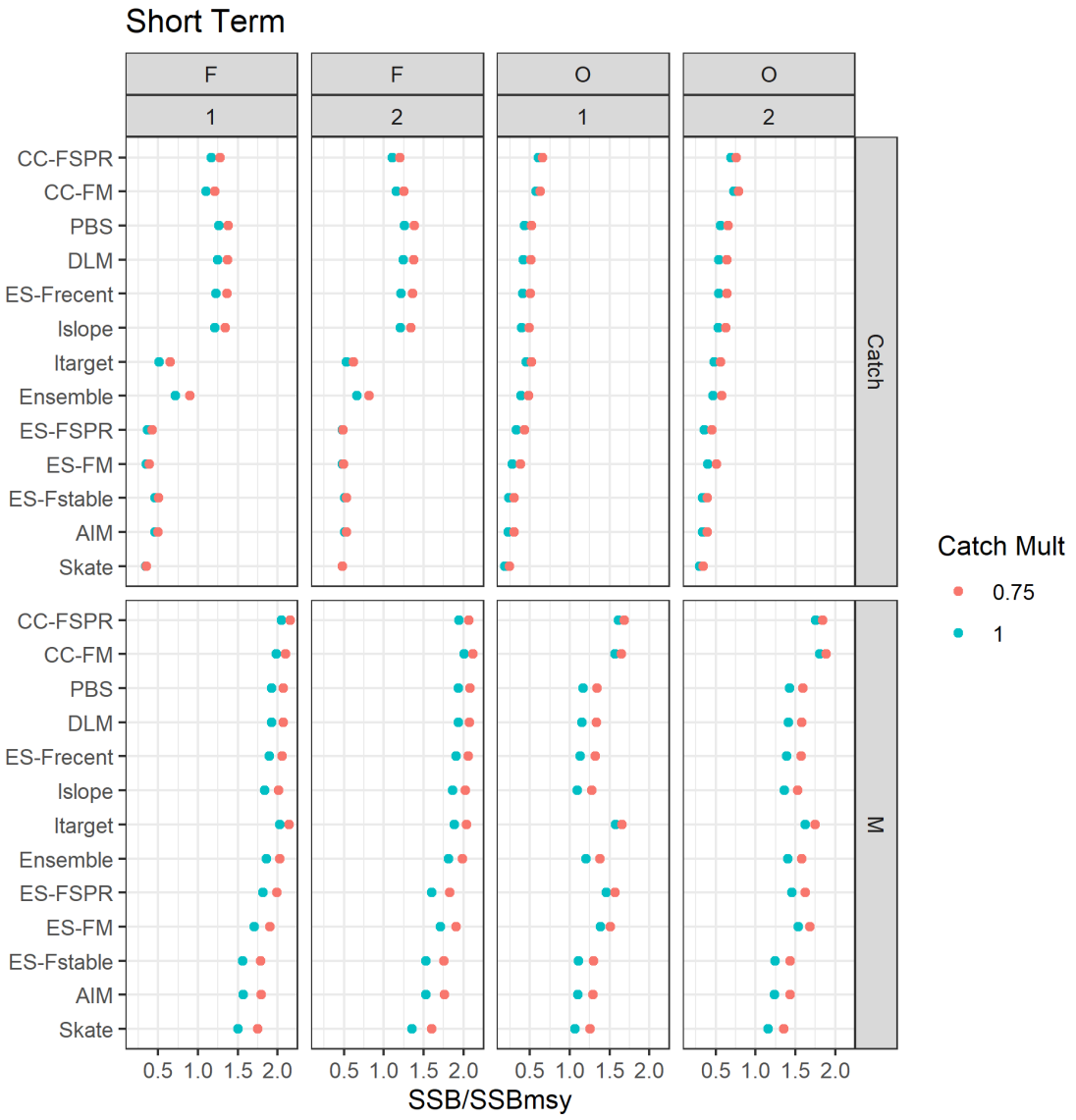


Figure A6.129. Mean values of spawning stock biomass (SSB) relative to its maximum sustainable yield reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

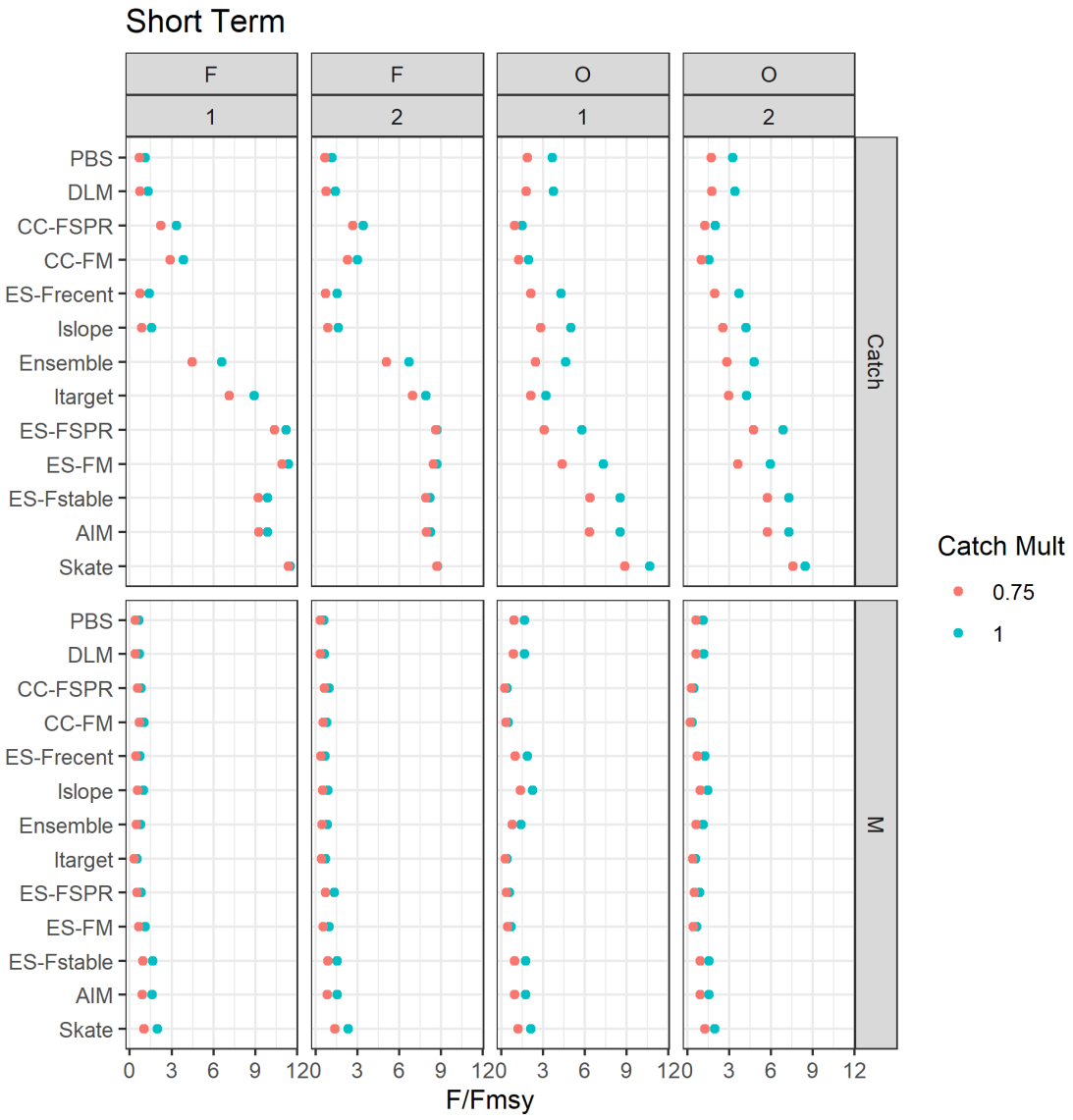


Figure A6.130. Mean values of fishing mortality rate (F) relative to its maximum sustainable yield reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of F history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

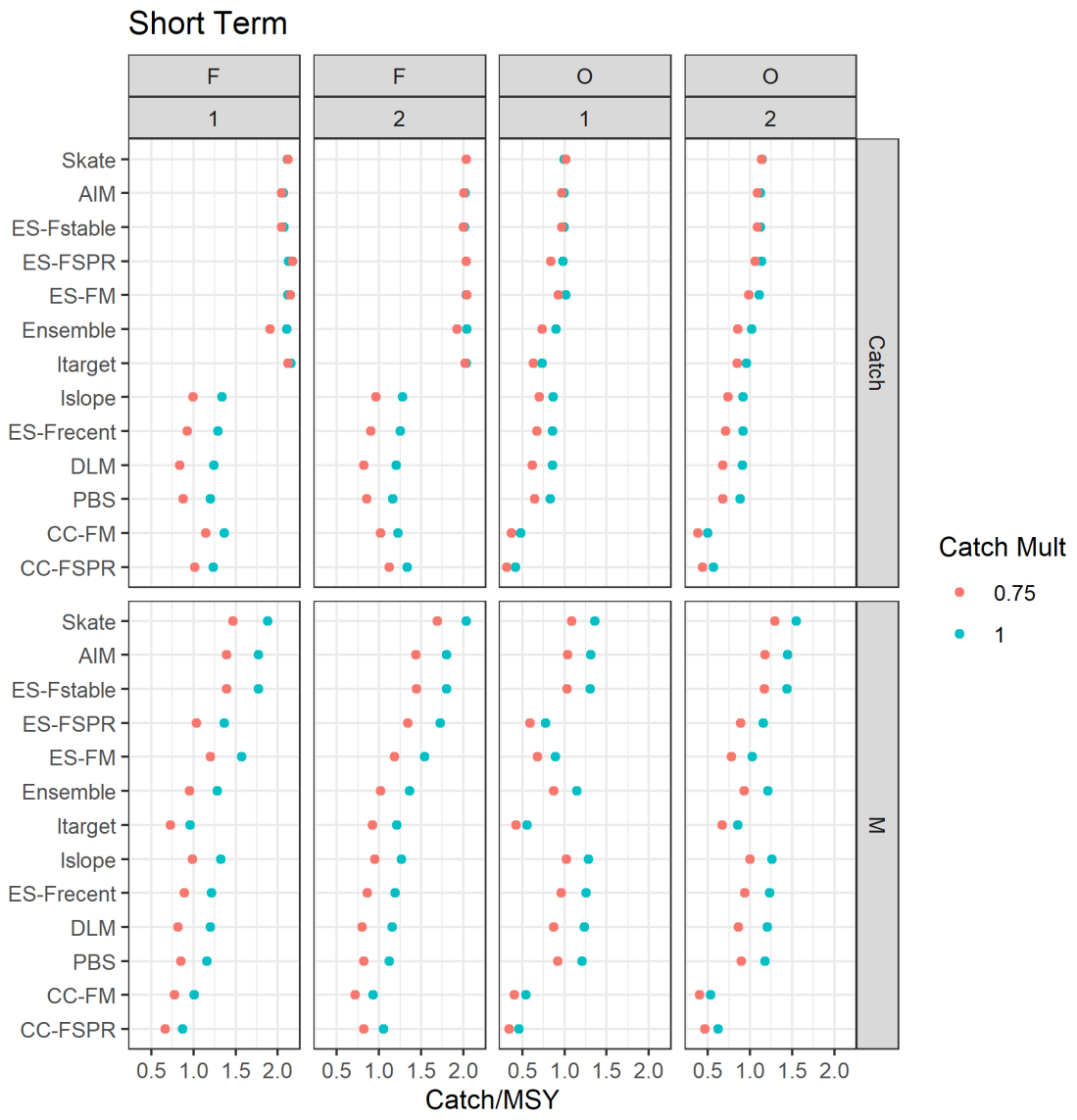


Figure A6.131. Mean values of catch relative to its maximum sustainable yield reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

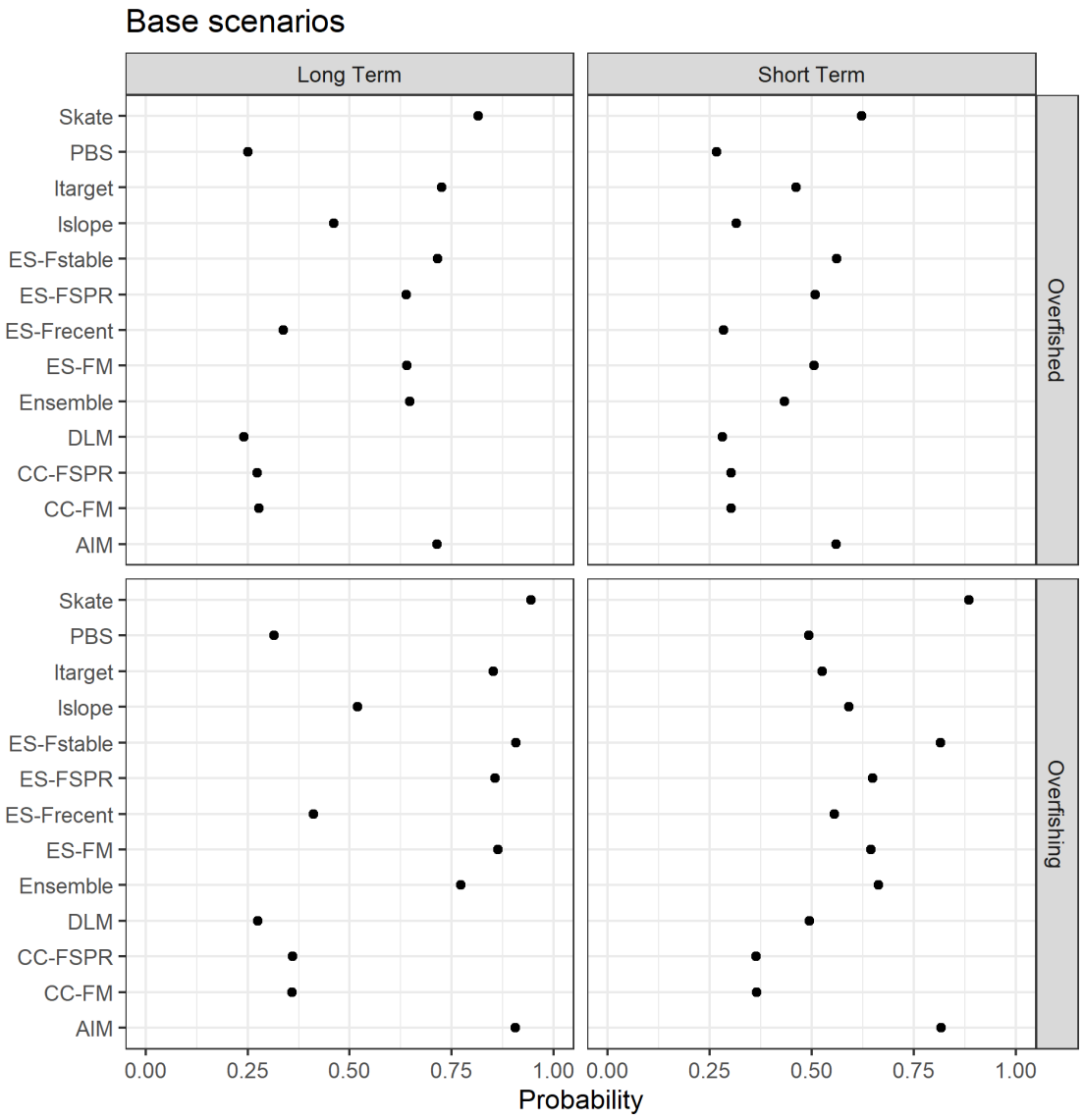


Figure A6.132. Probability stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all base analyses.

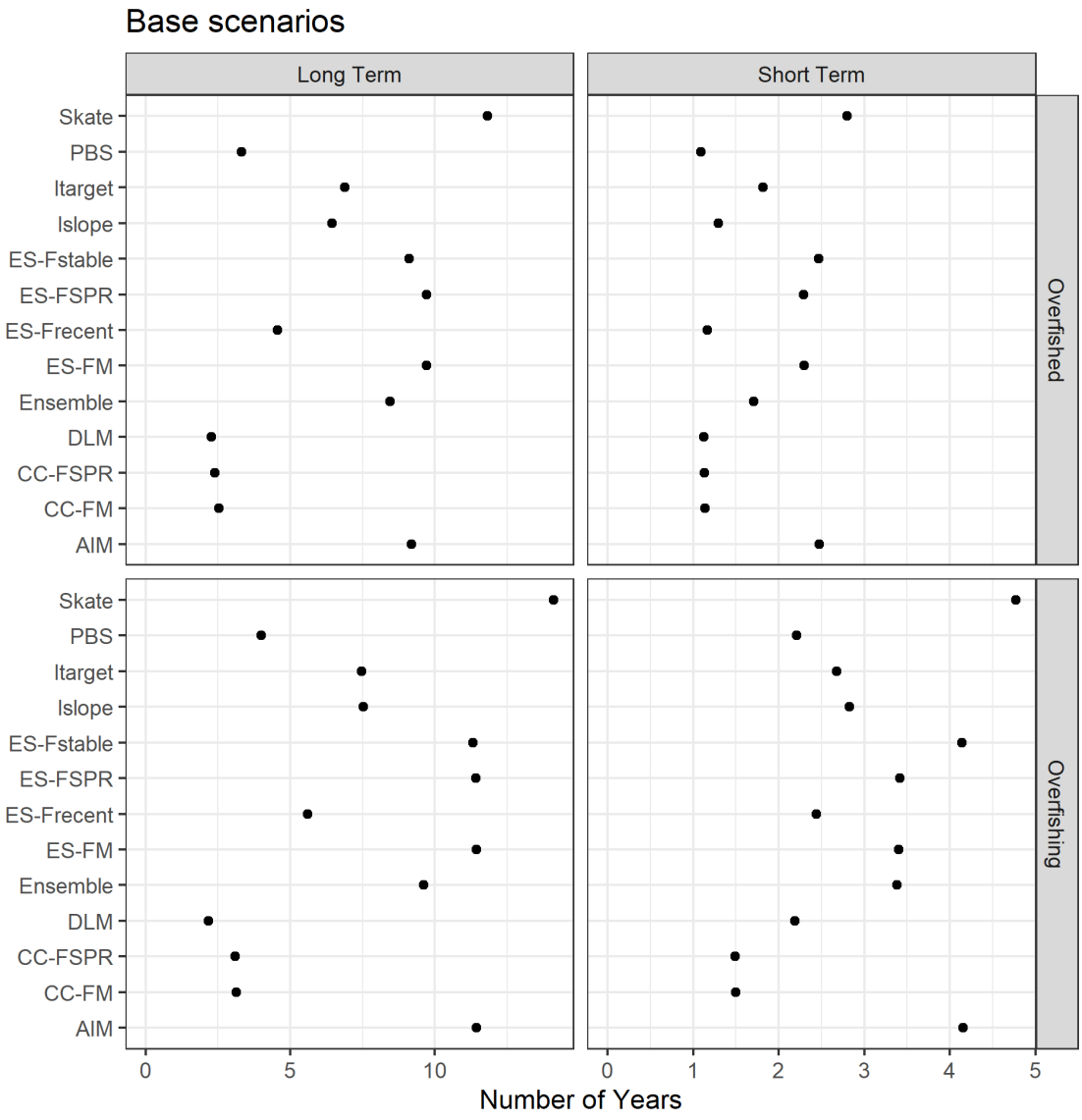


Figure A6.133. Number of years stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all base analyses.

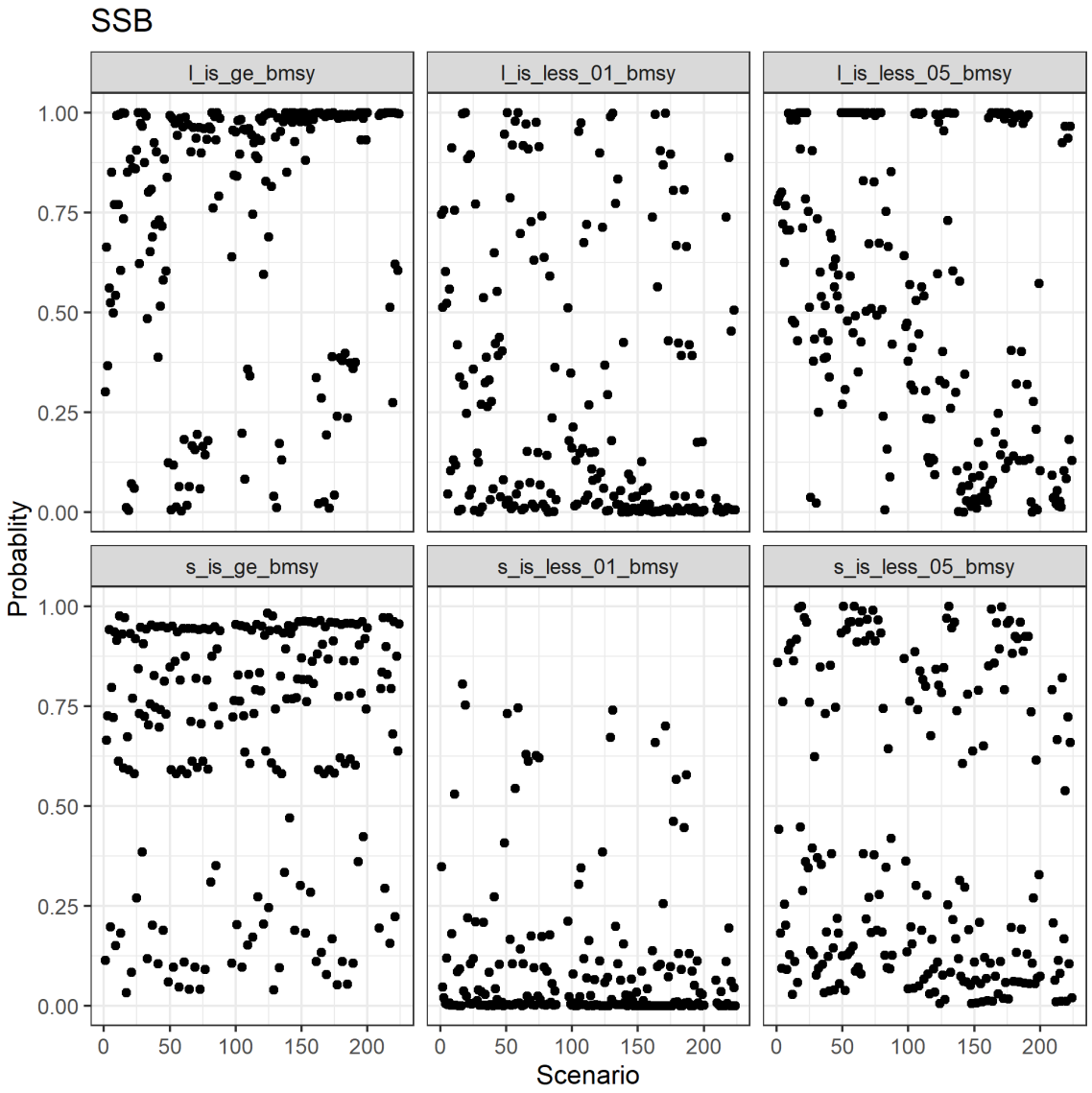


Figure A6.134. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

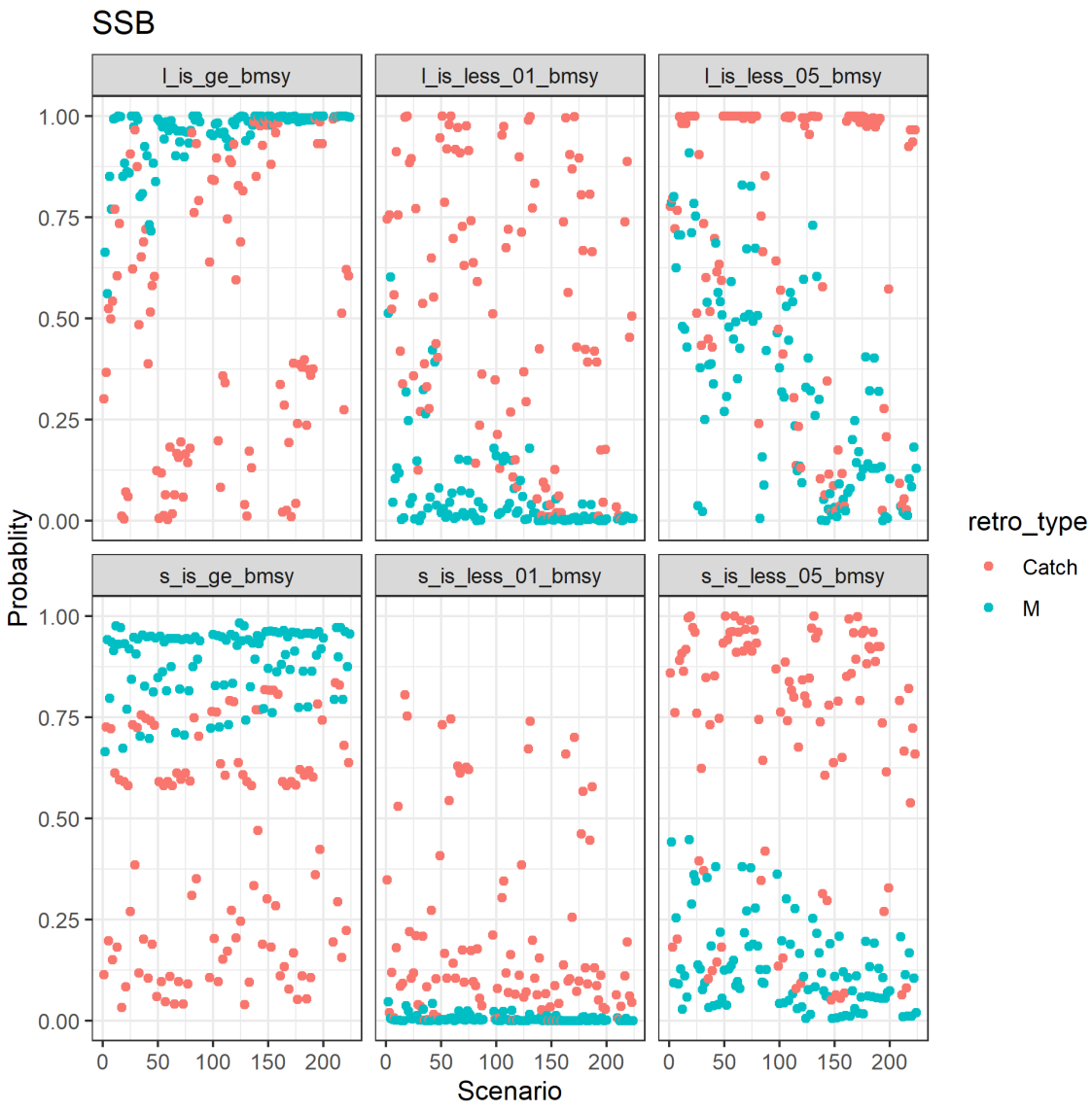


Figure A6.135. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

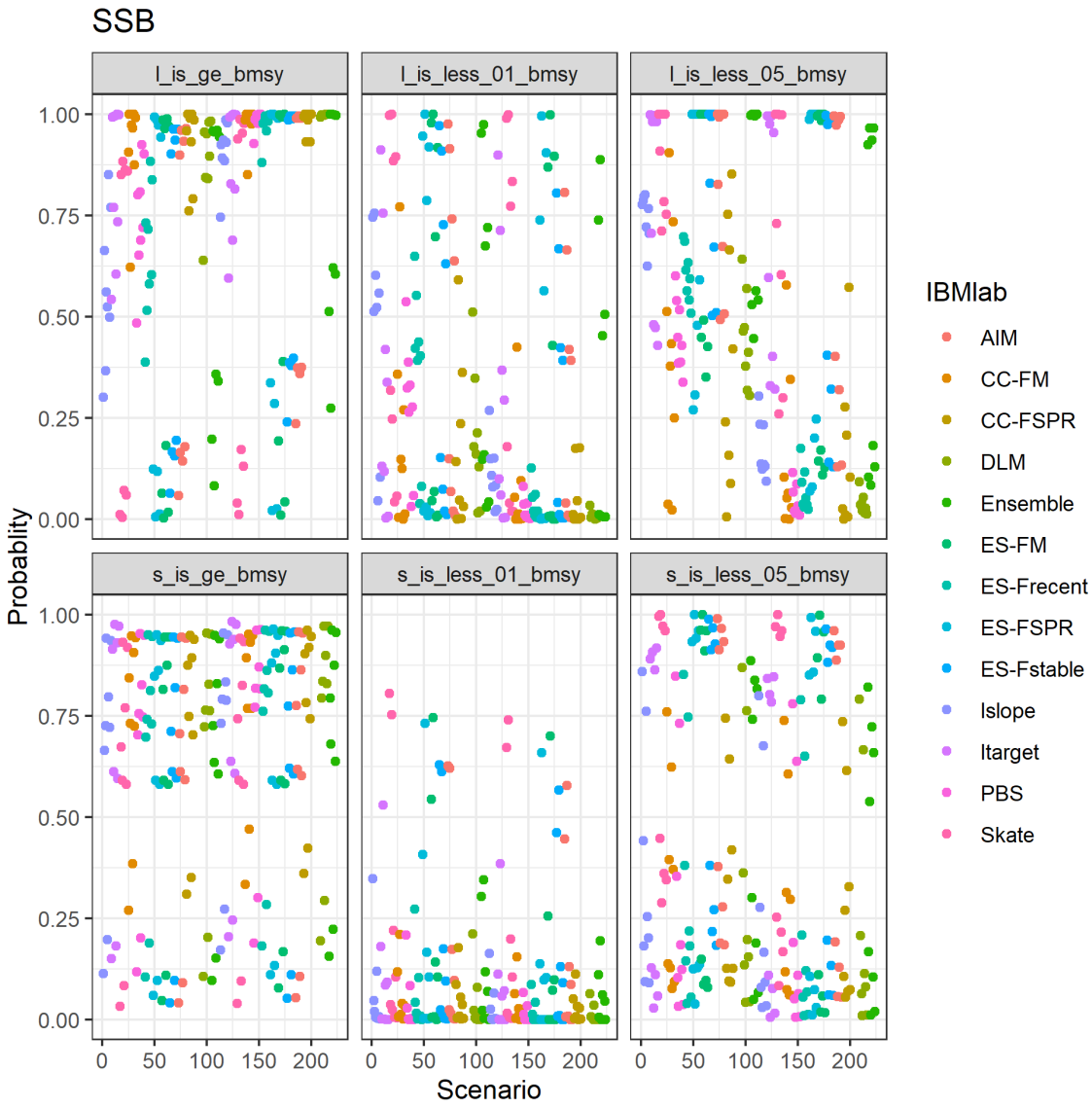


Figure A6.136. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

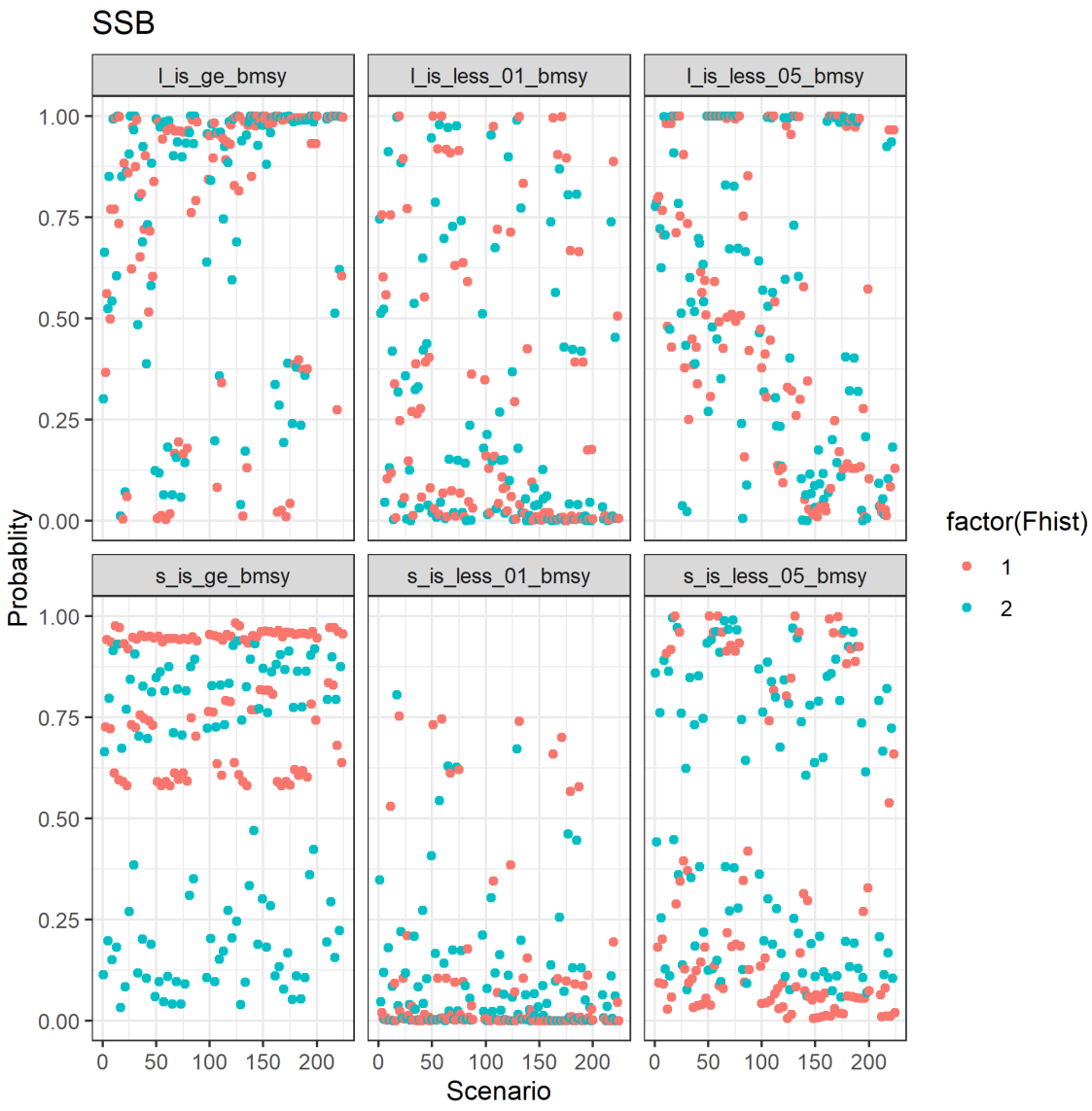


Figure A6.137. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

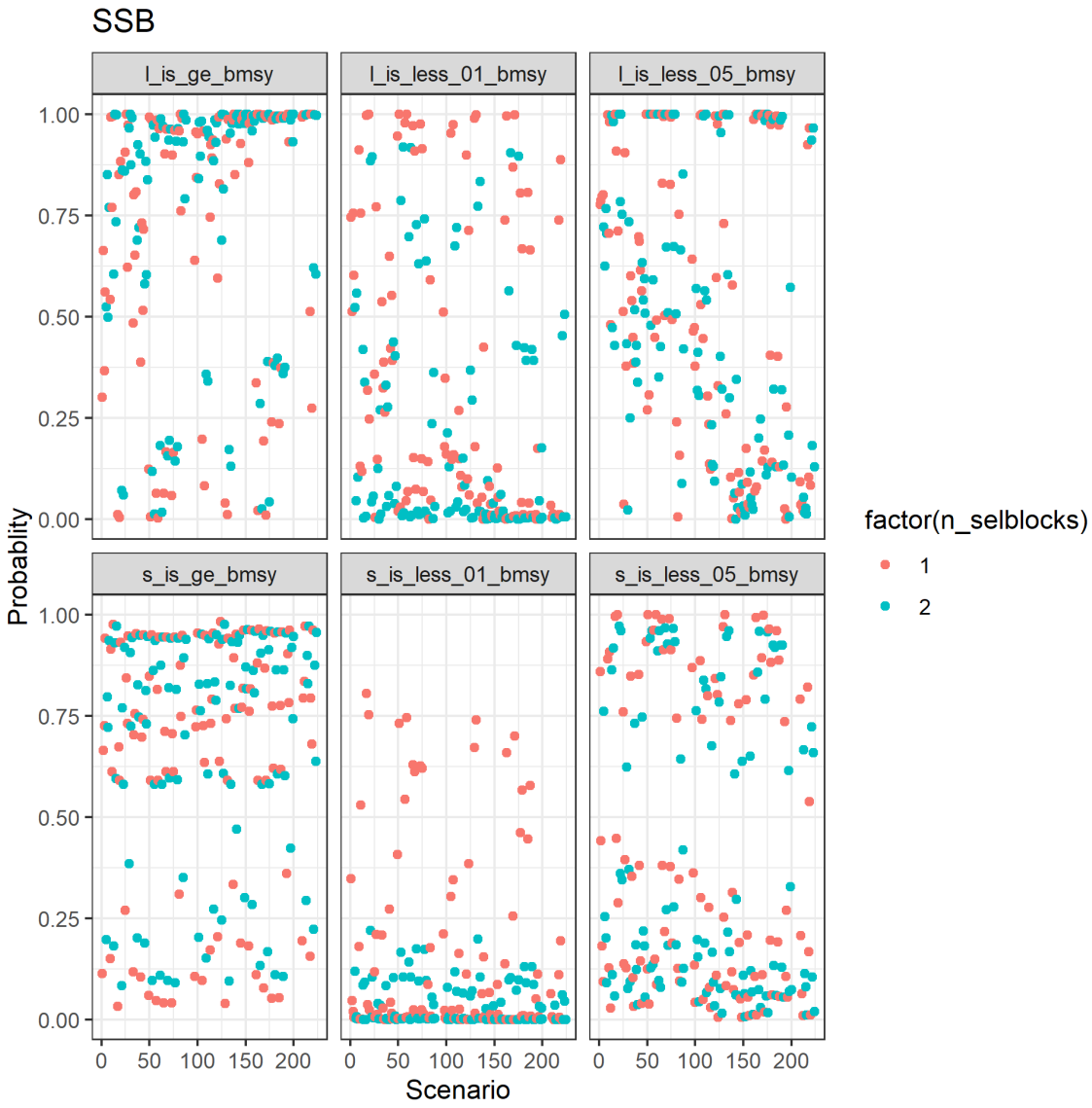


Figure A6.138. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

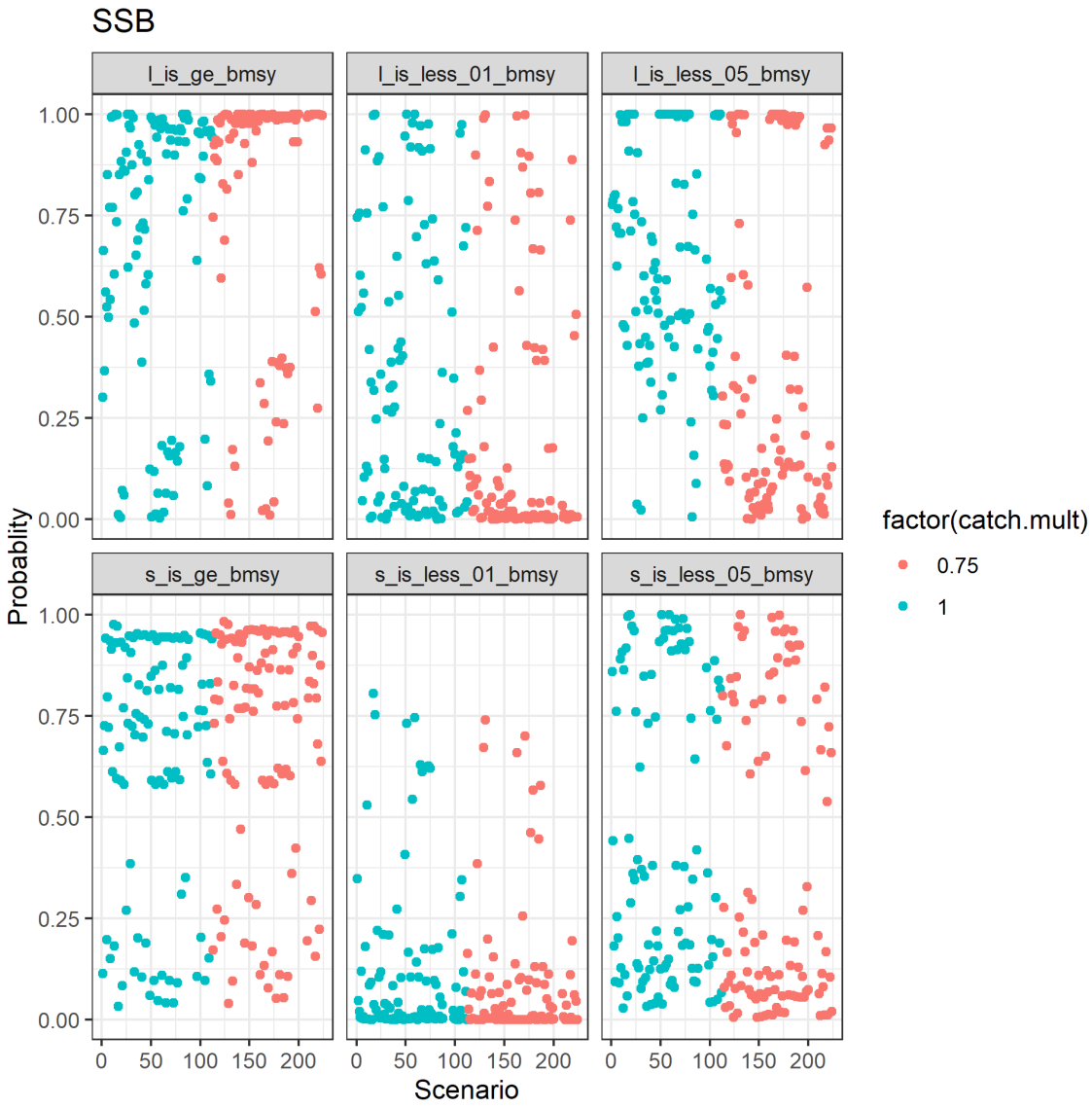


Figure A6.139. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

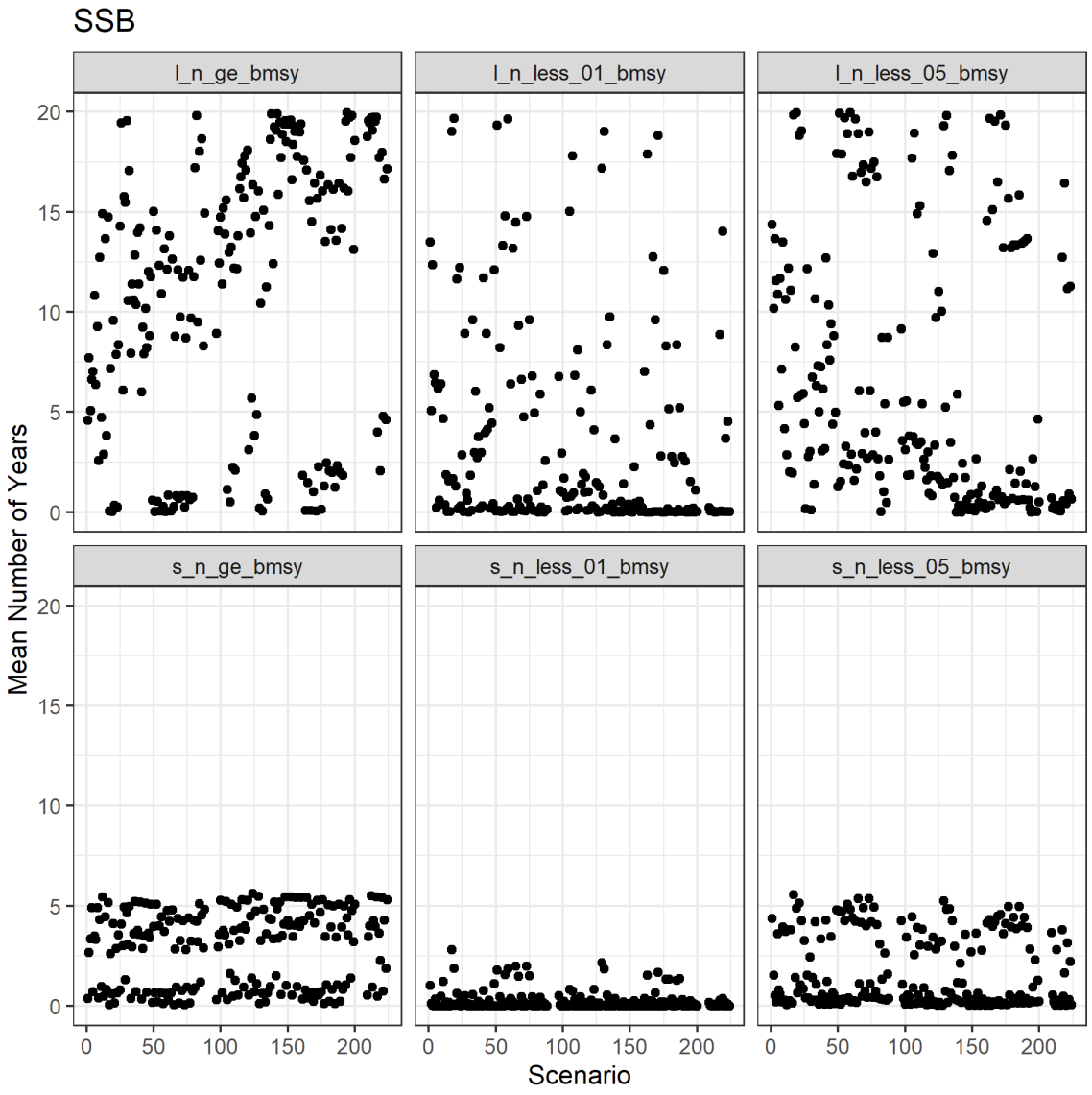


Figure A6.140. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

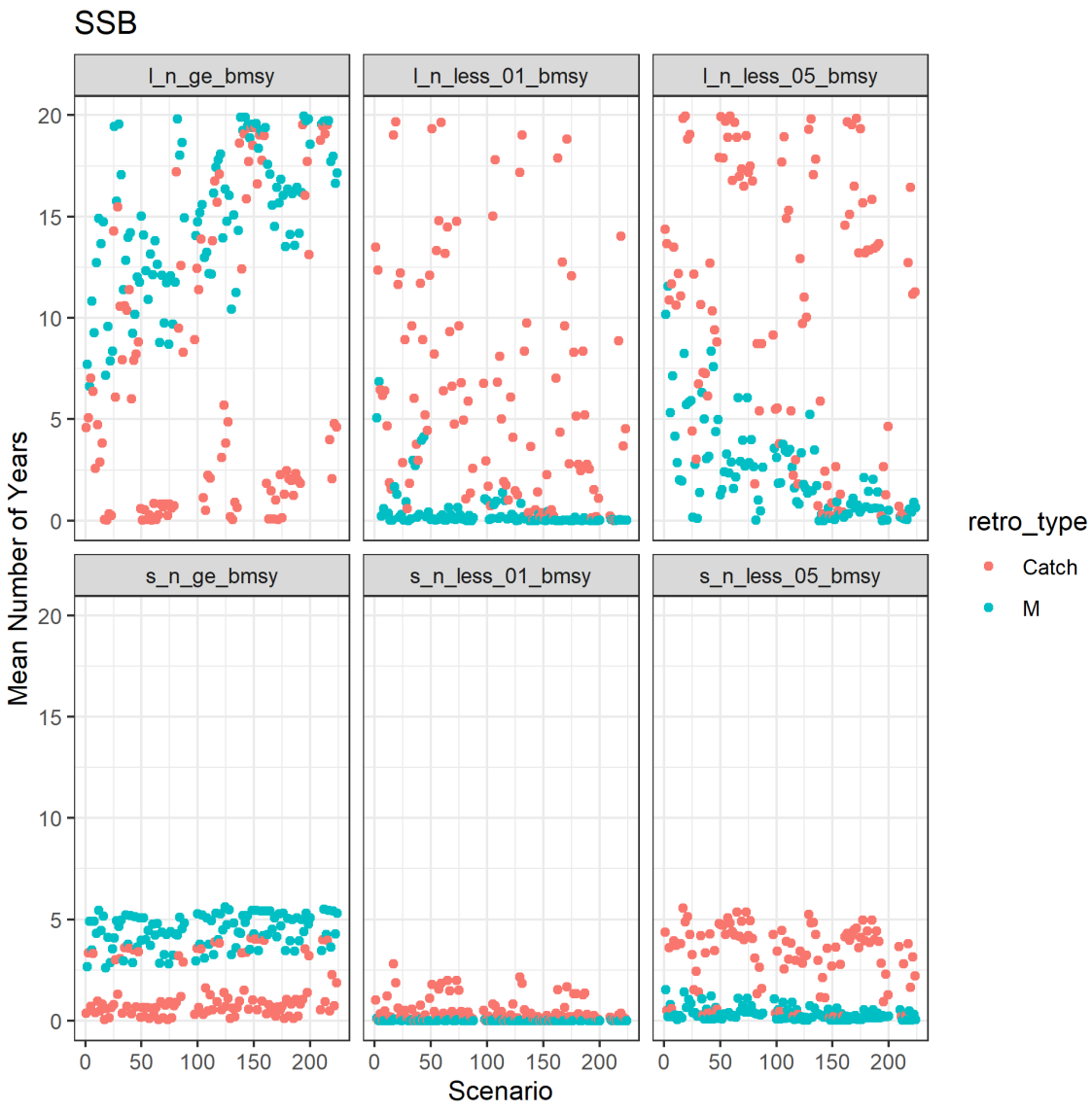


Figure A6.141. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

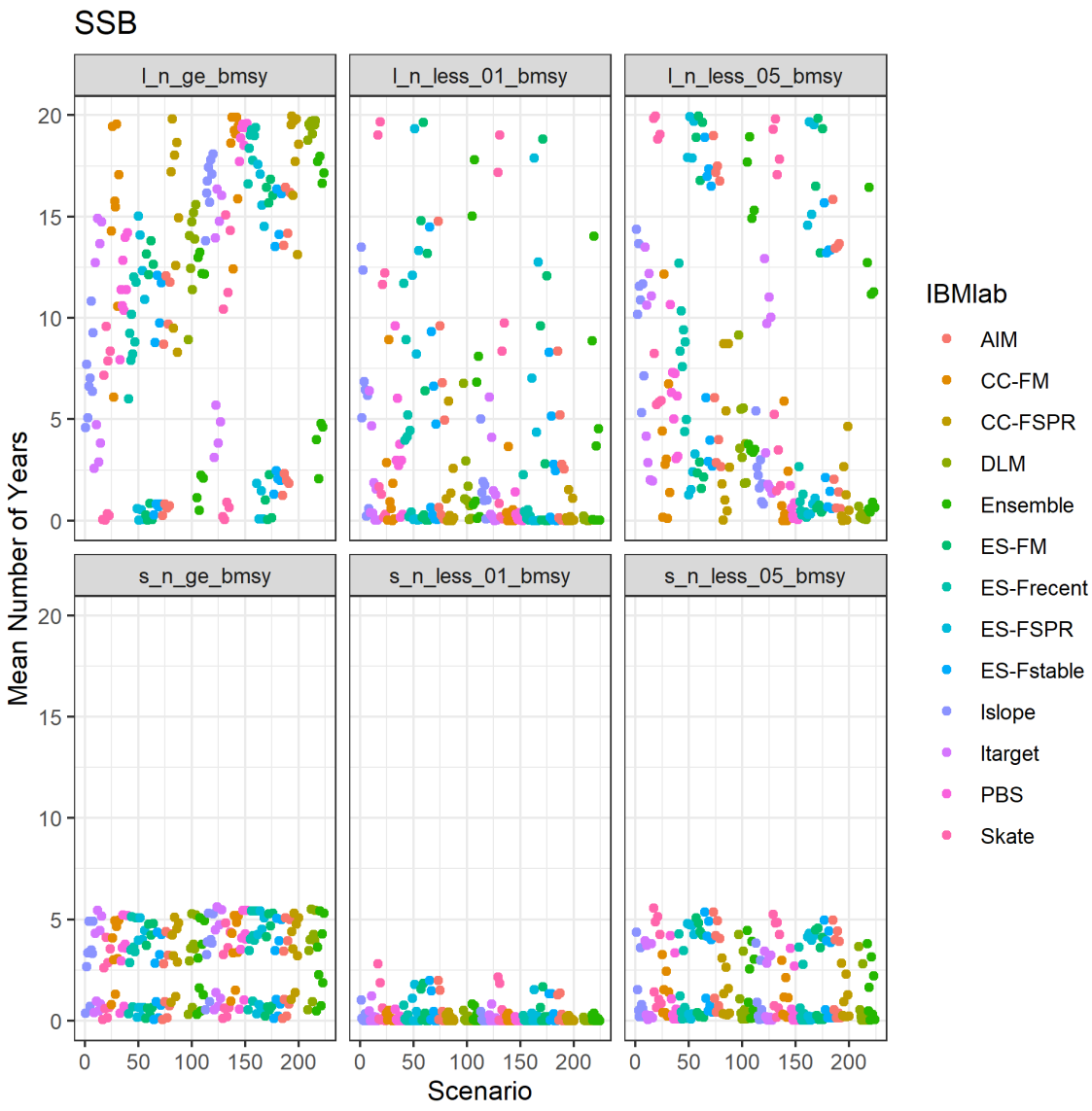


Figure A6.142. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

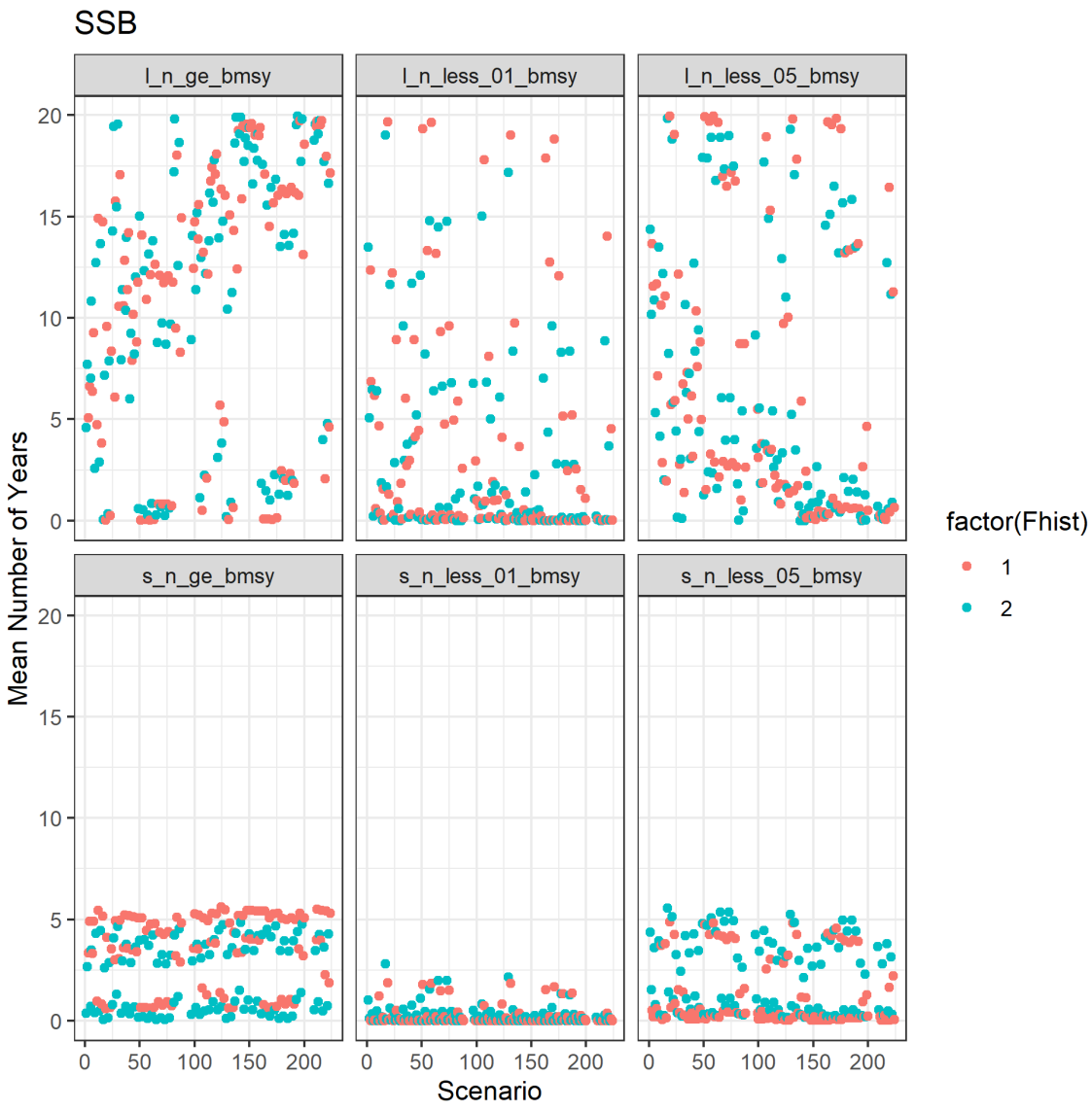


Figure A6.143. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

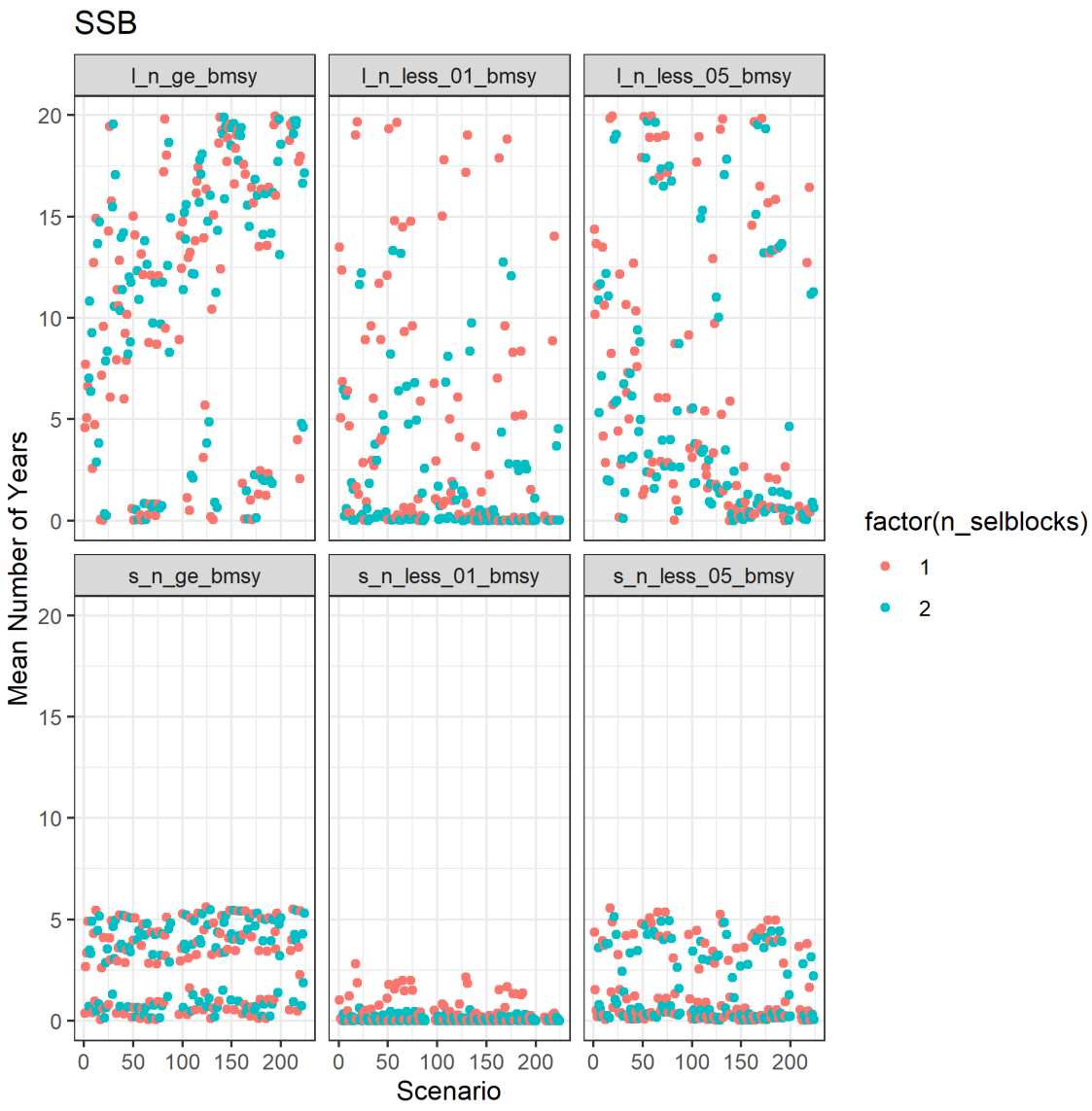


Figure A6.144. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

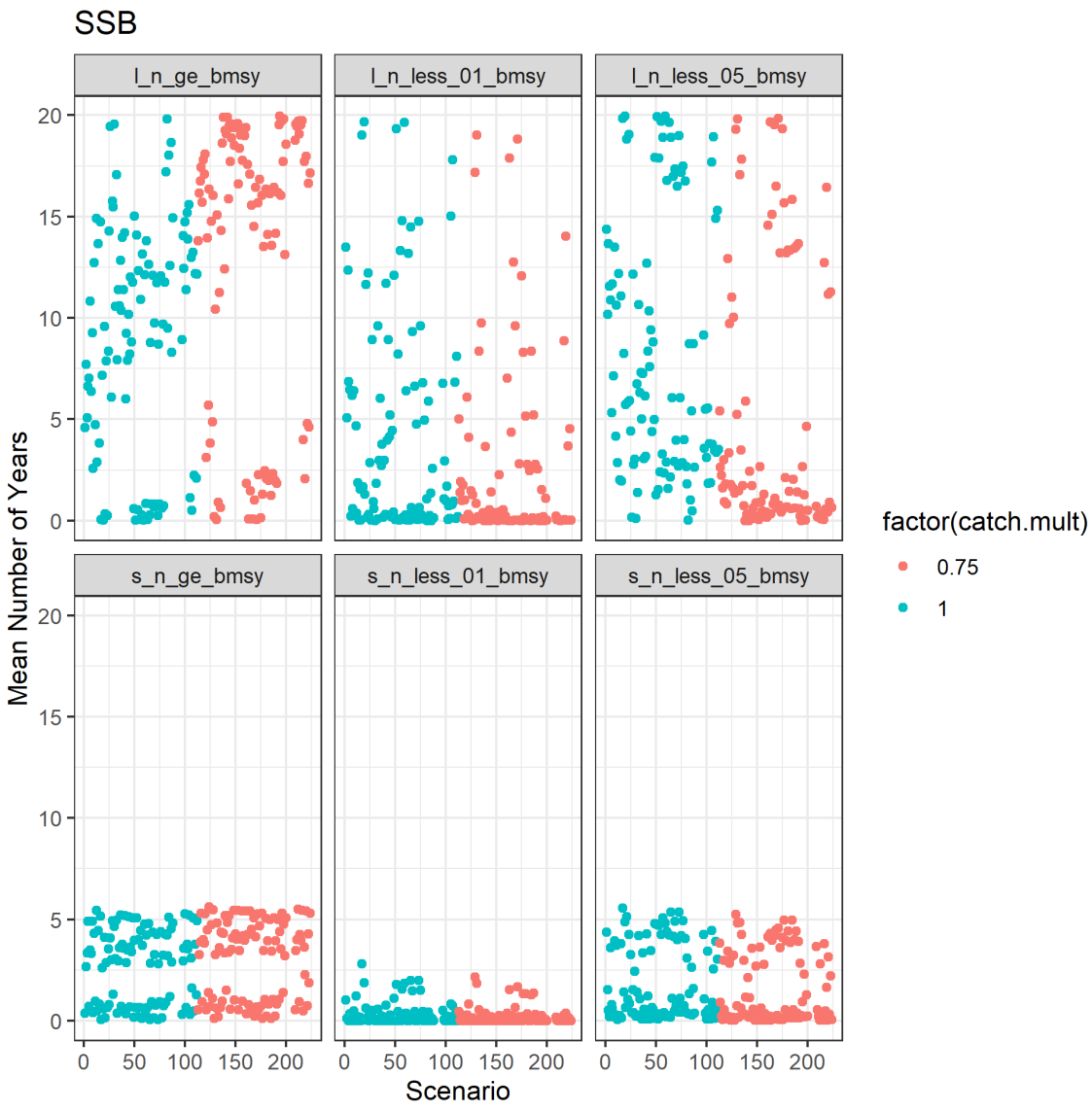


Figure A6.145. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

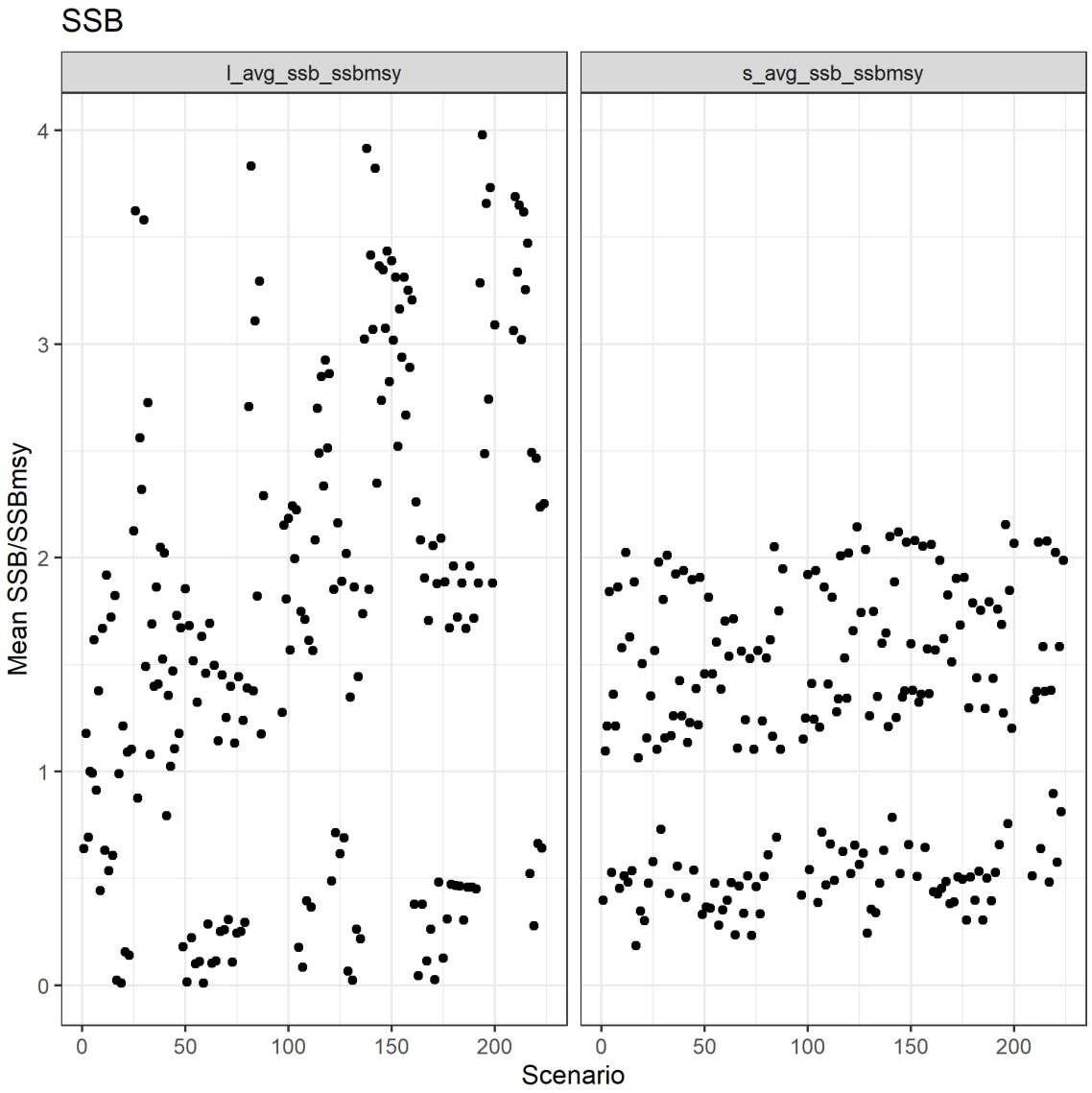


Figure A6.146. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

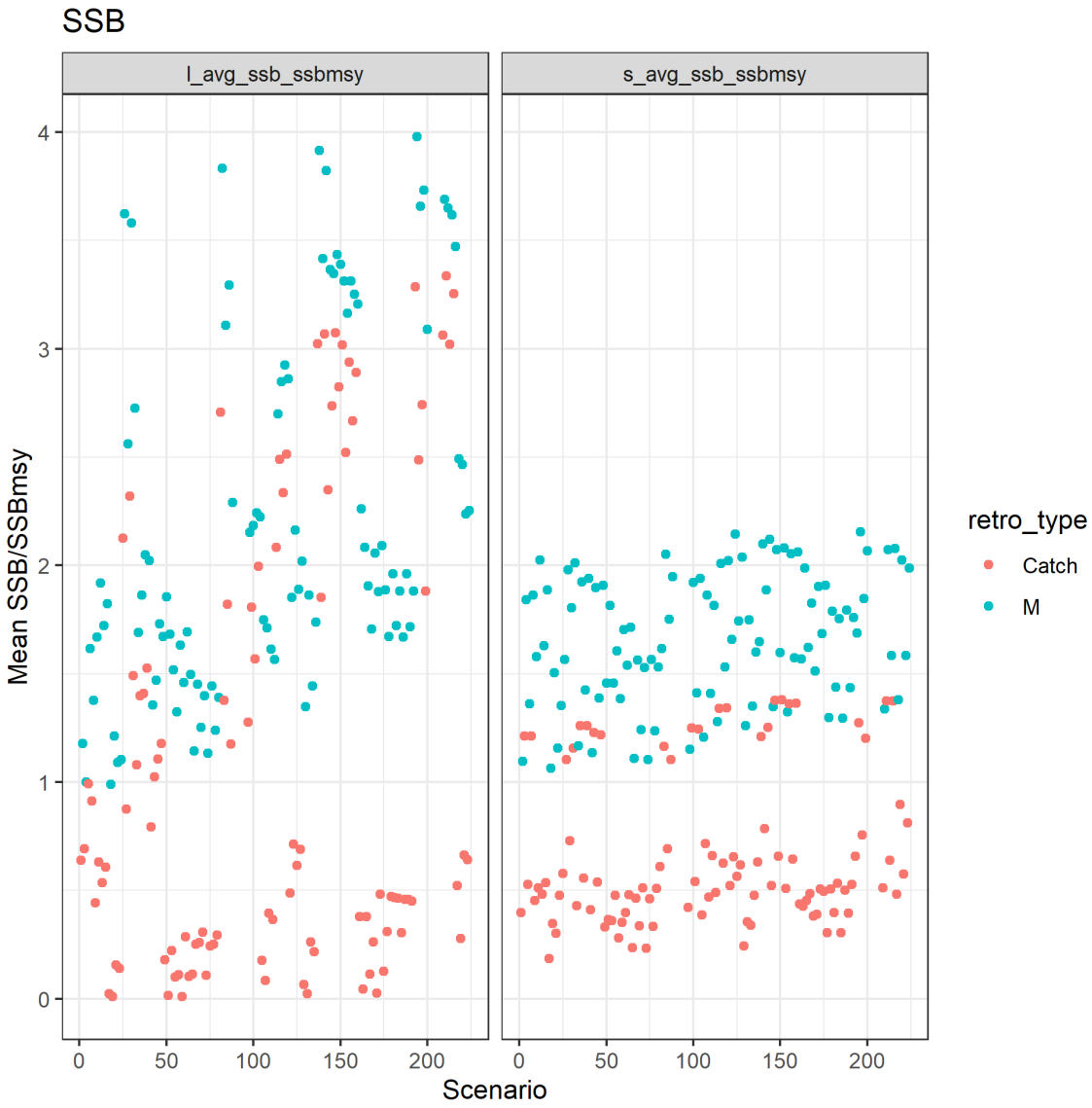


Figure A6.147. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

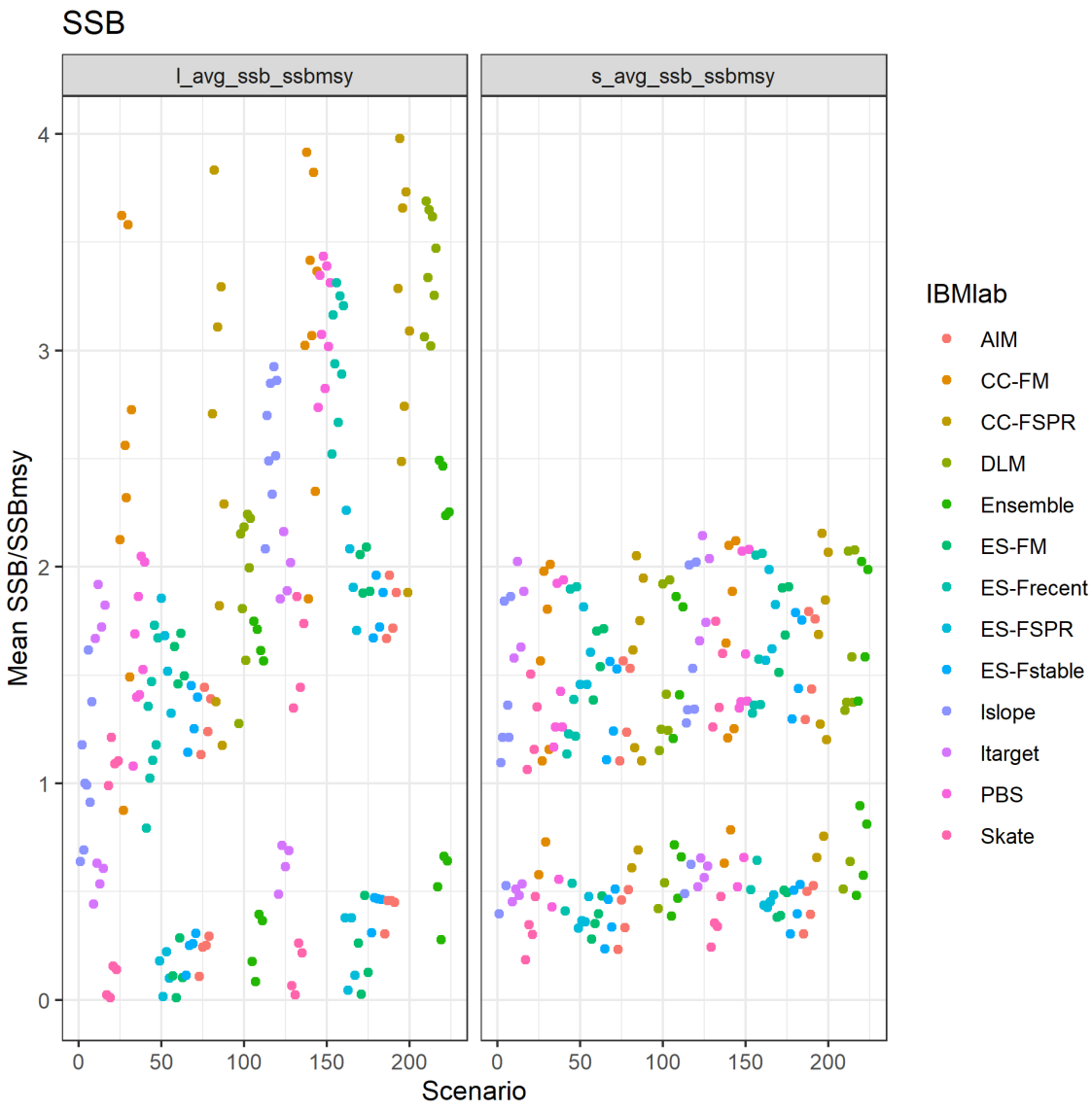


Figure A6.148. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

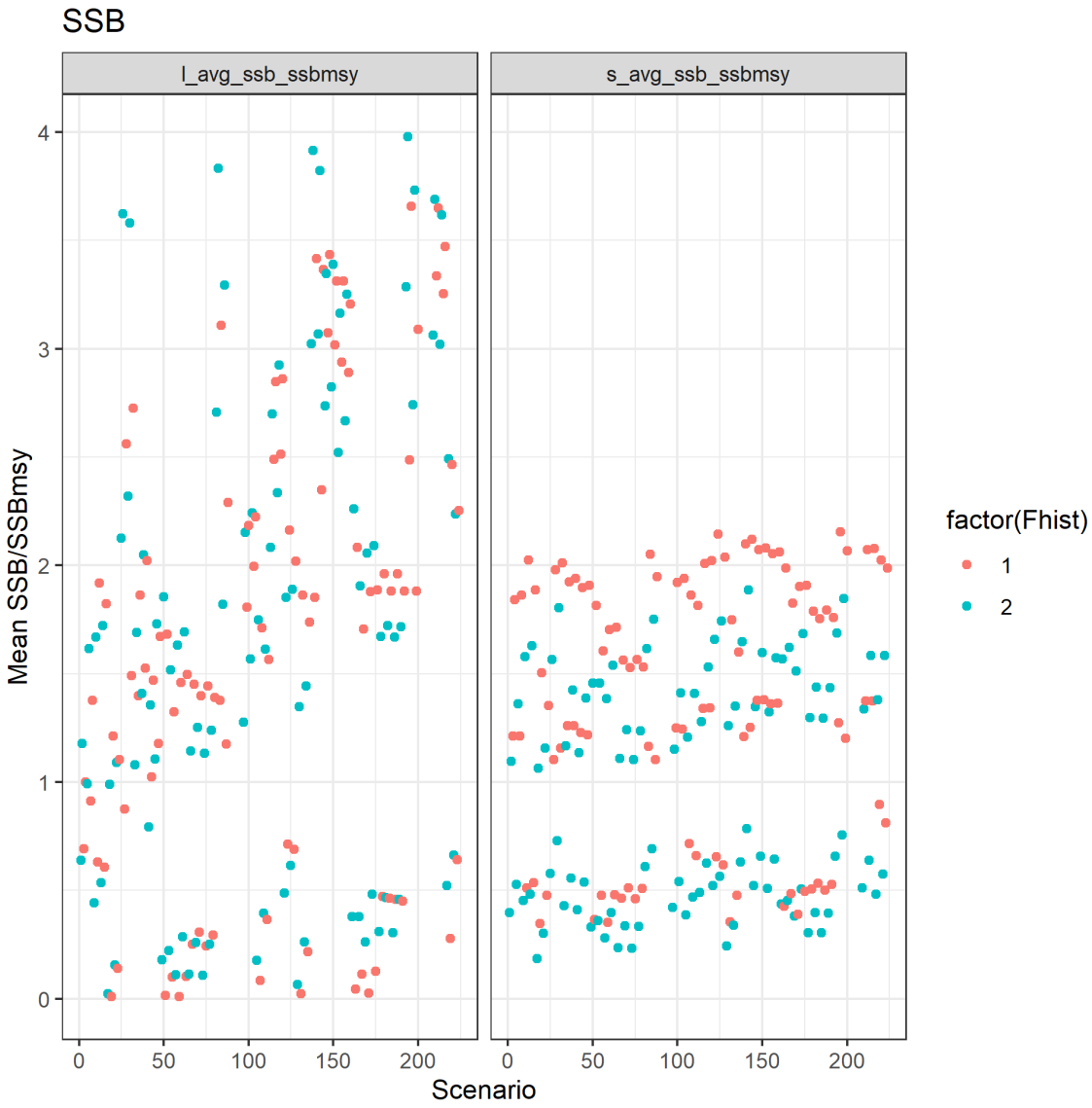


Figure A6.149. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

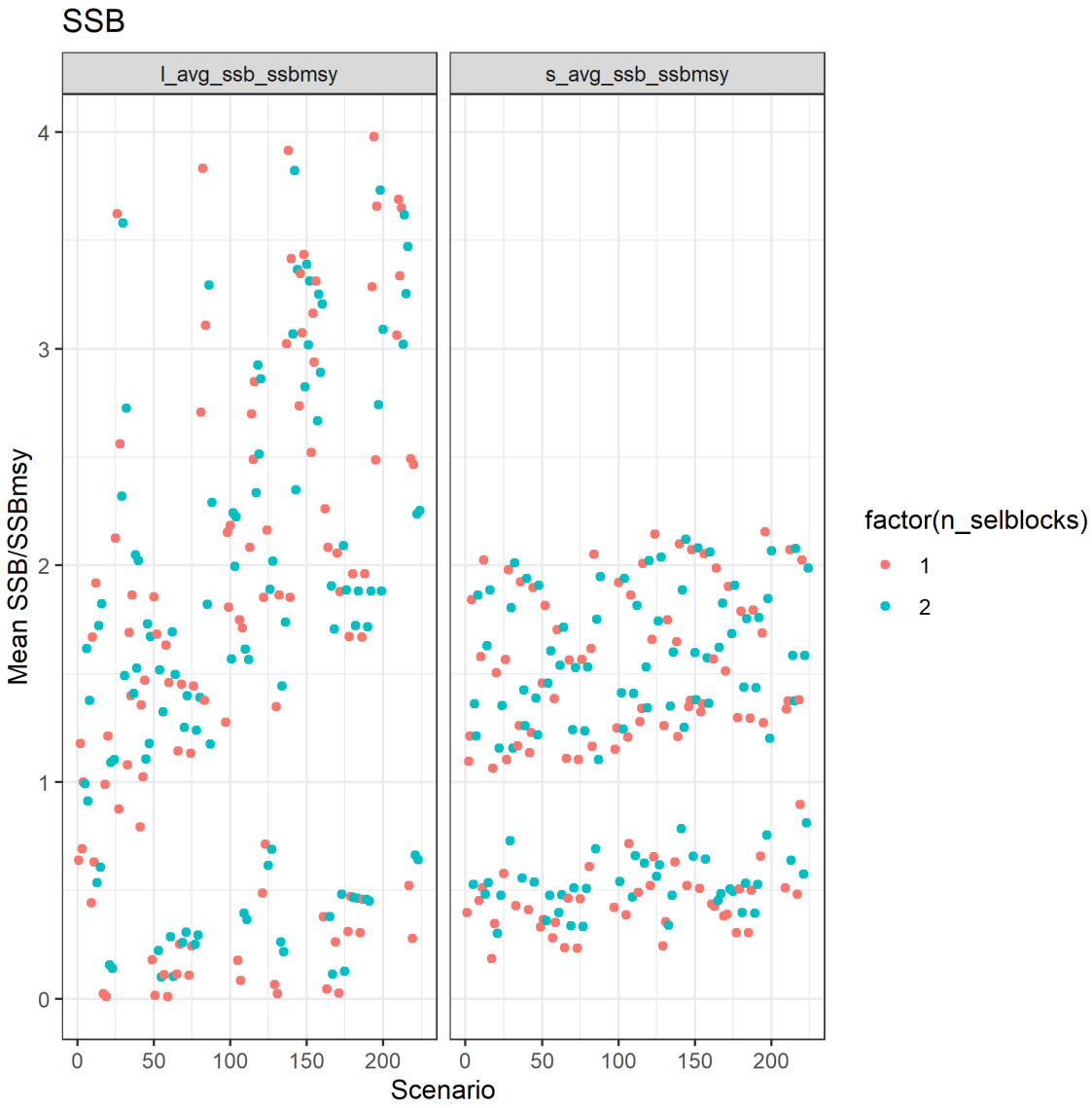


Figure A6.150. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

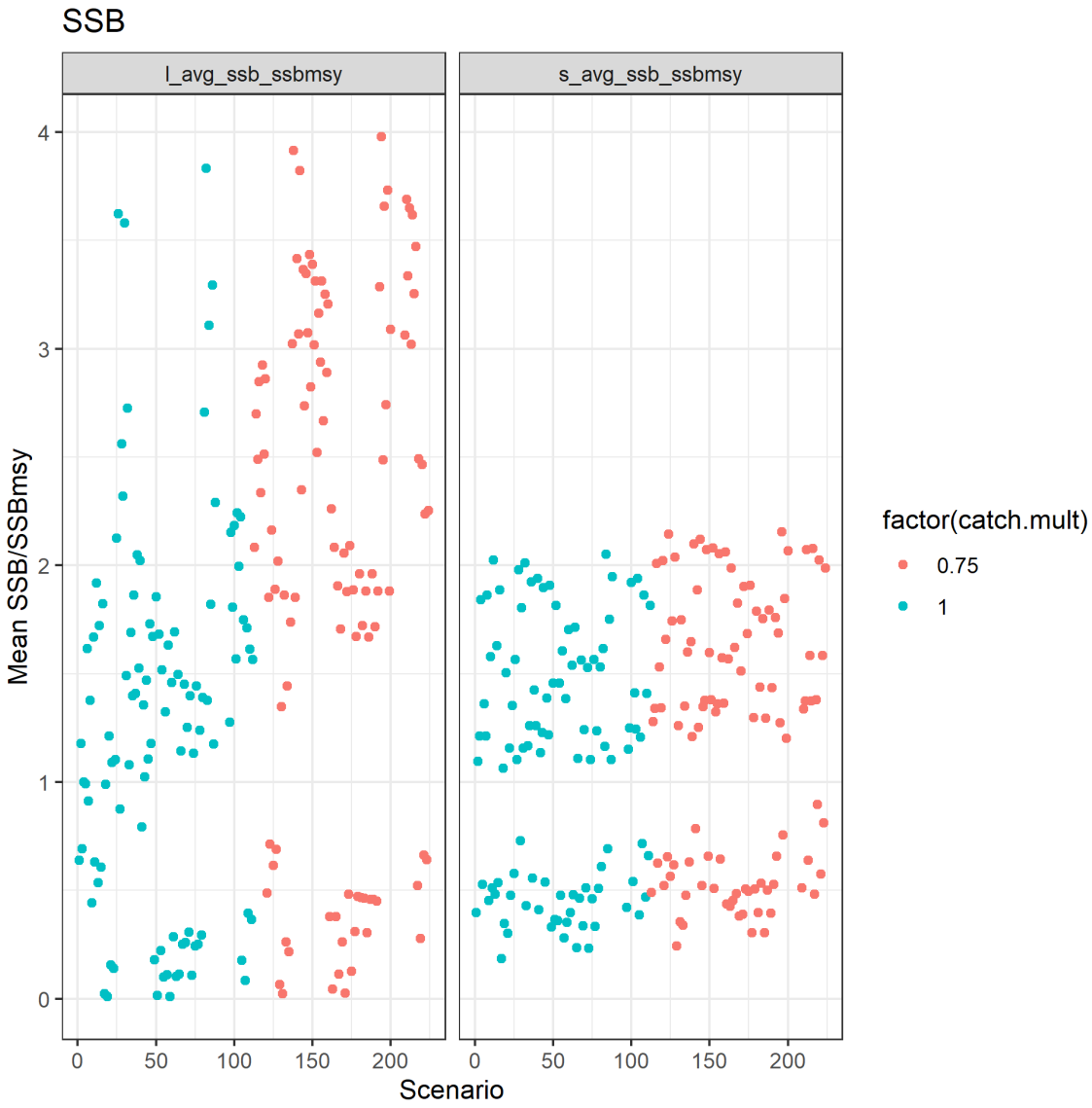


Figure A6.151. Confetti plot of the mean values for some of the spawning stock biomass (SSB) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

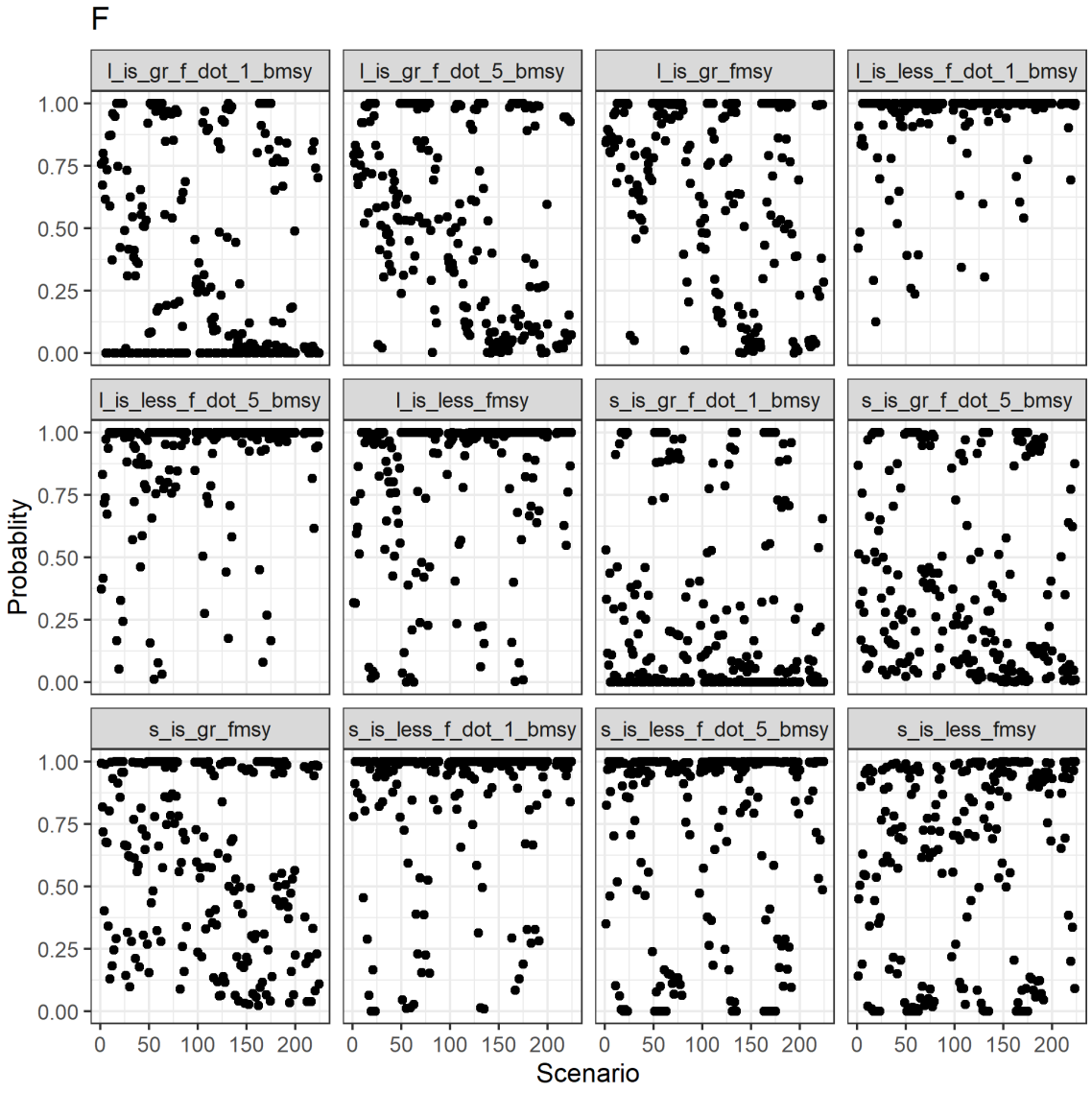


Figure A6.152. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

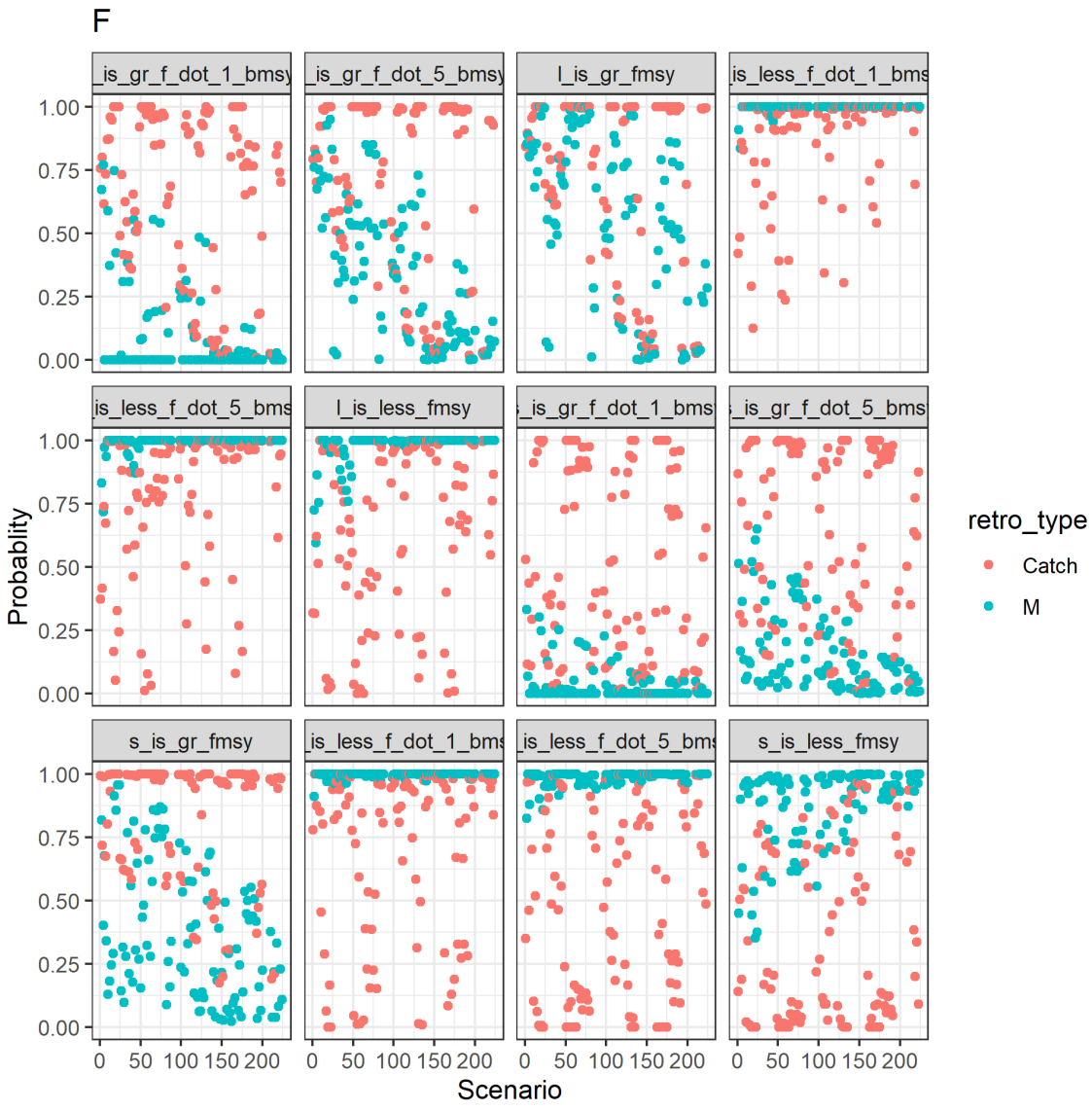


Figure A6.153. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

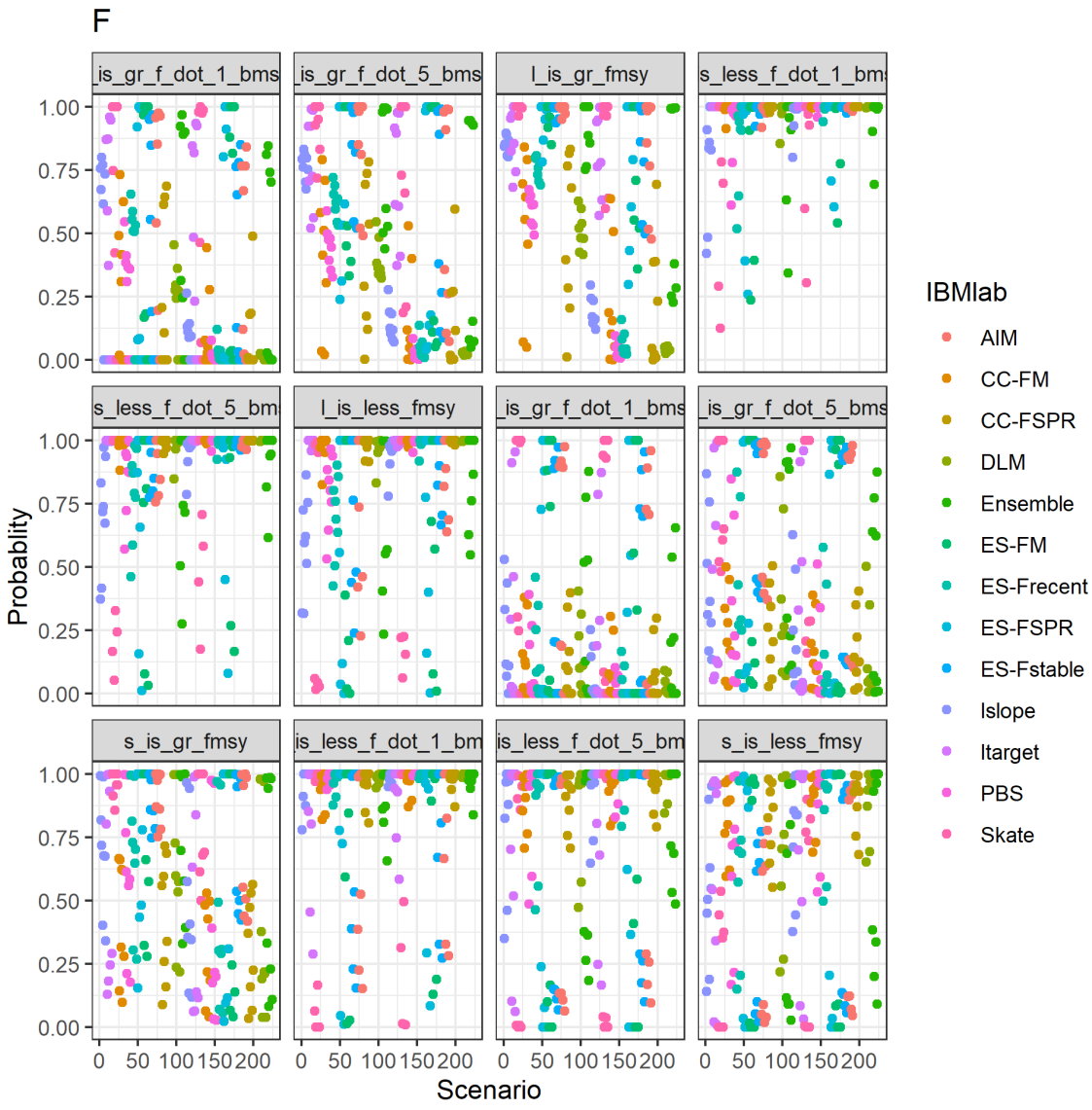


Figure A6.154. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

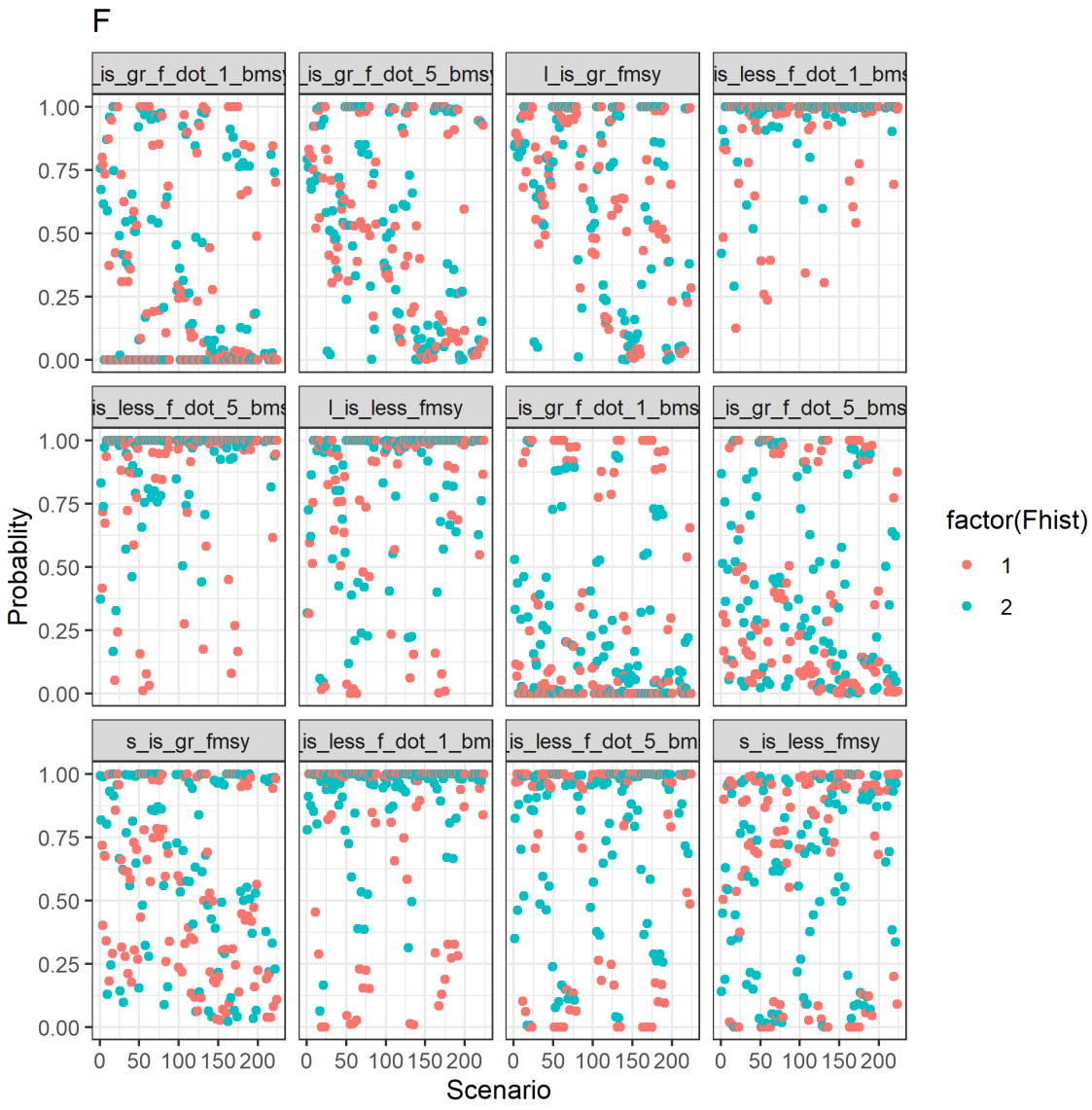


Figure A6.155. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

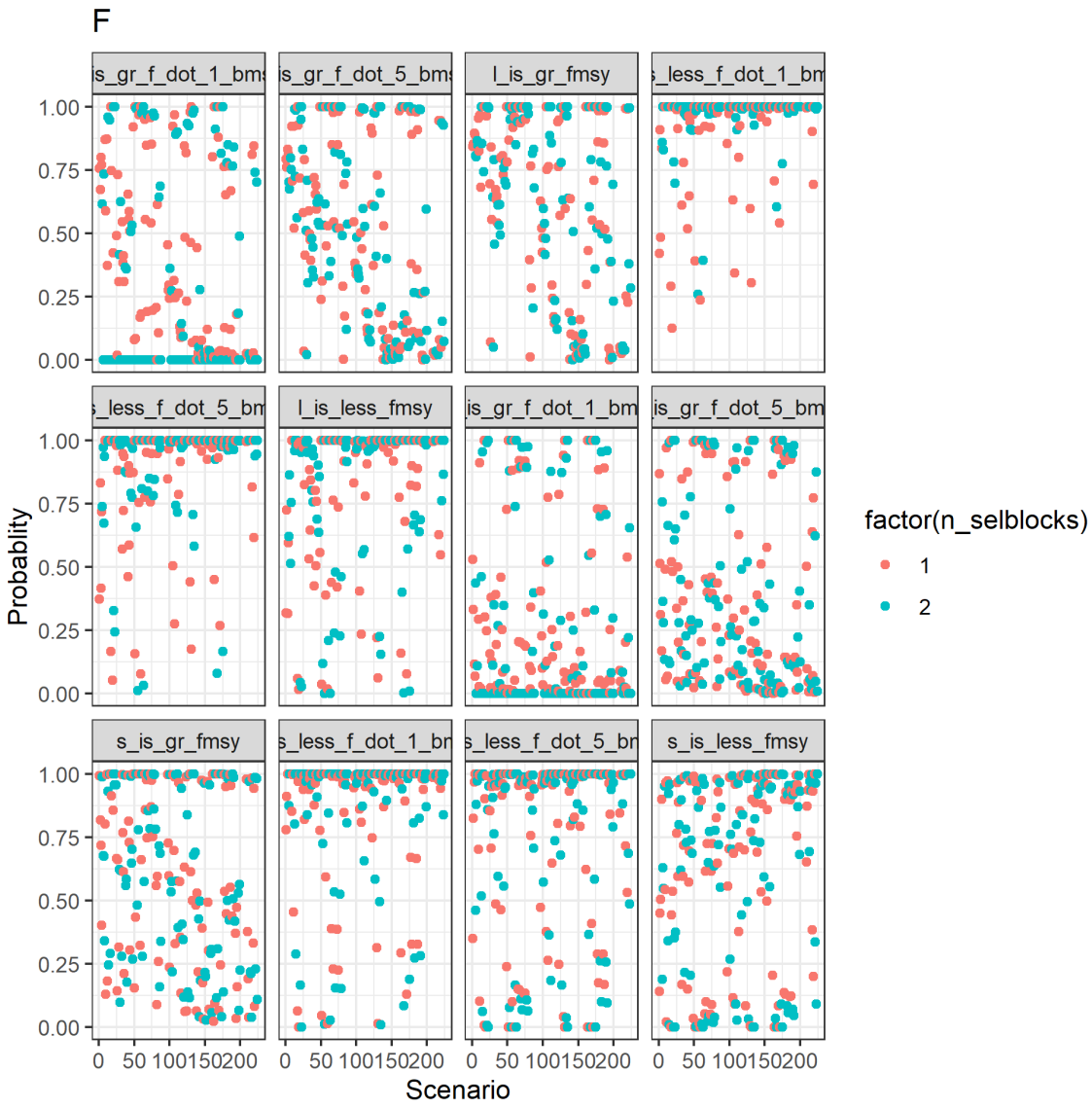


Figure A6.156. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

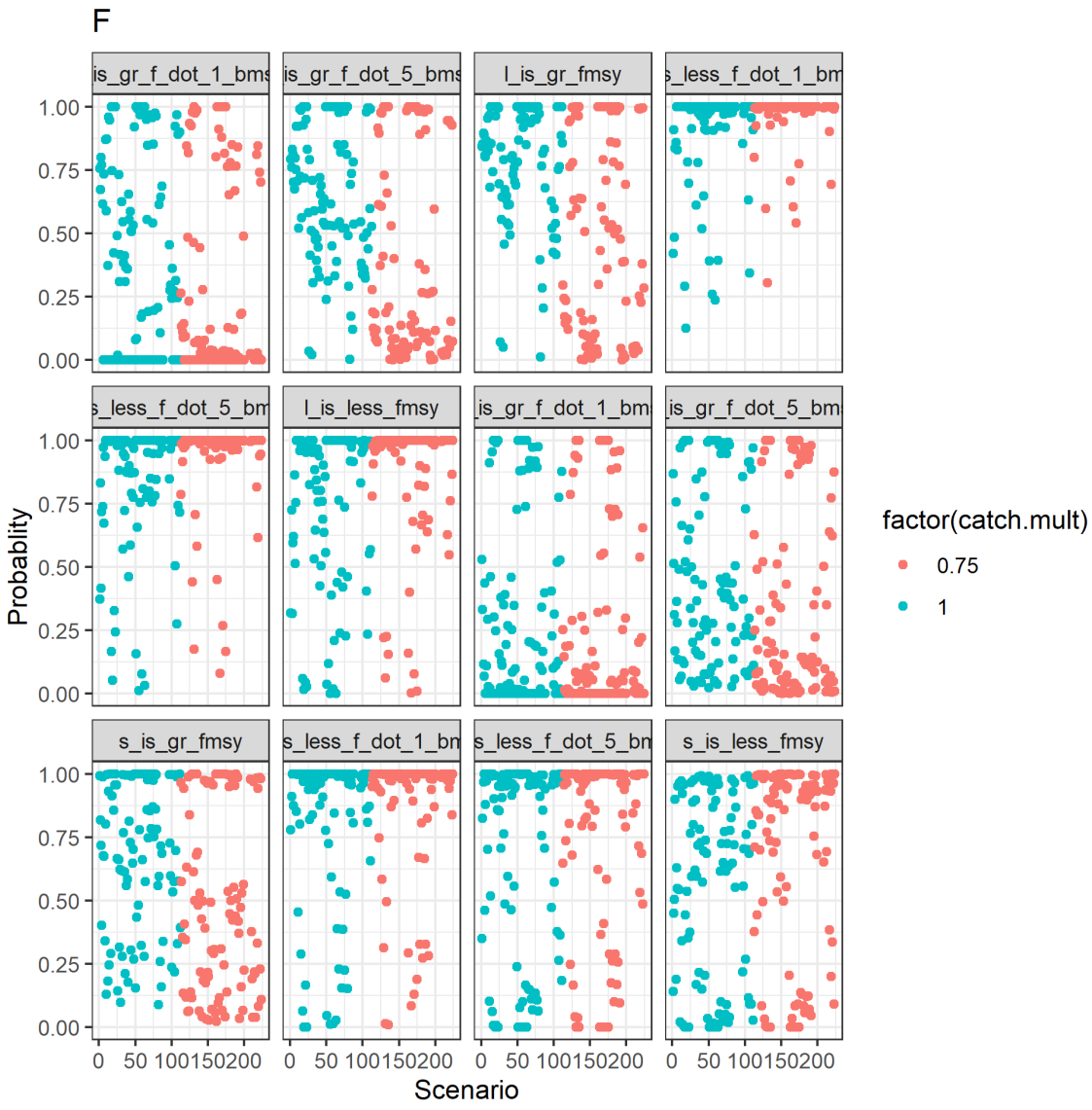


Figure A6.157. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

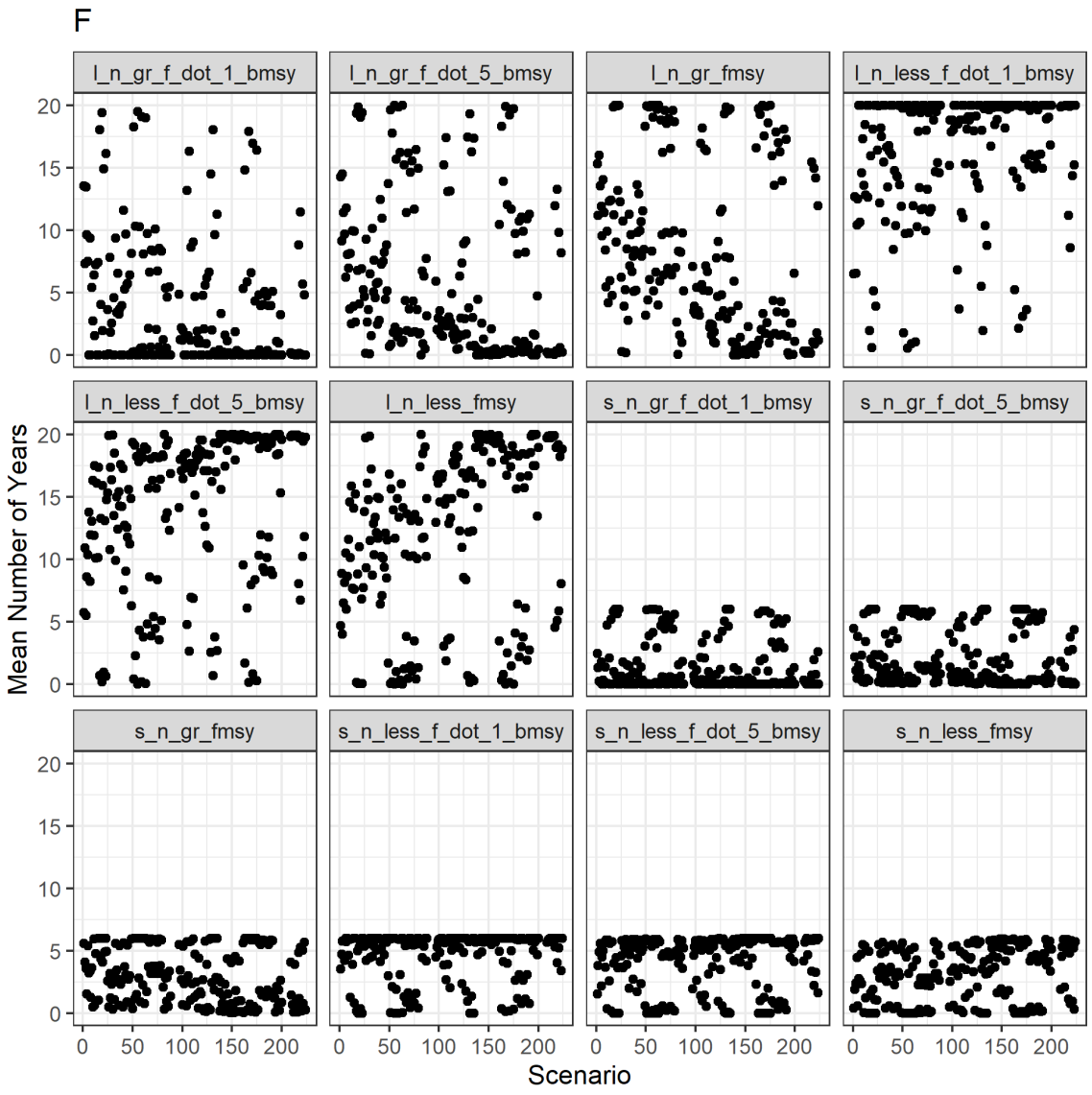


Figure A6.158. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

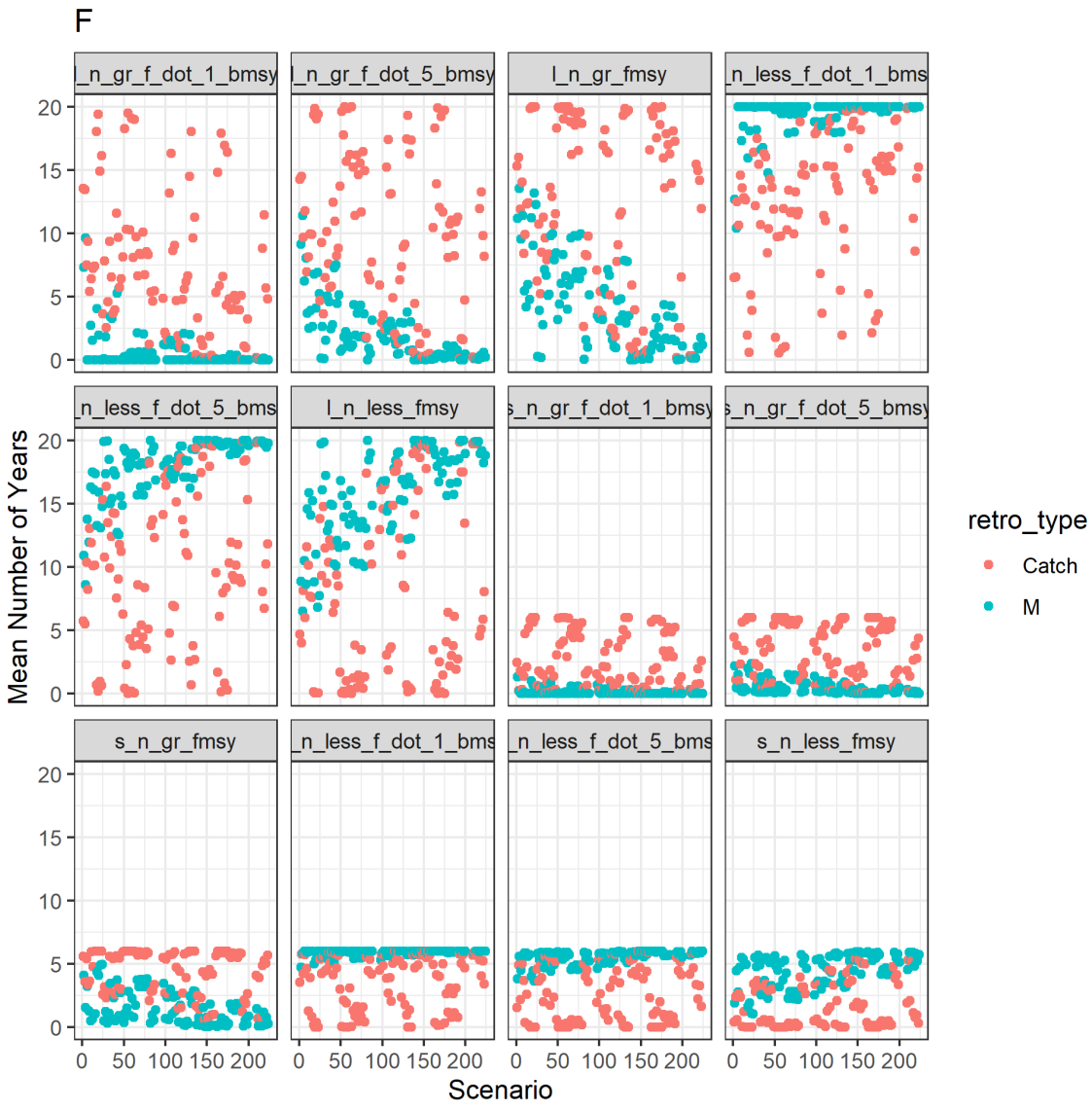


Figure A6.159. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

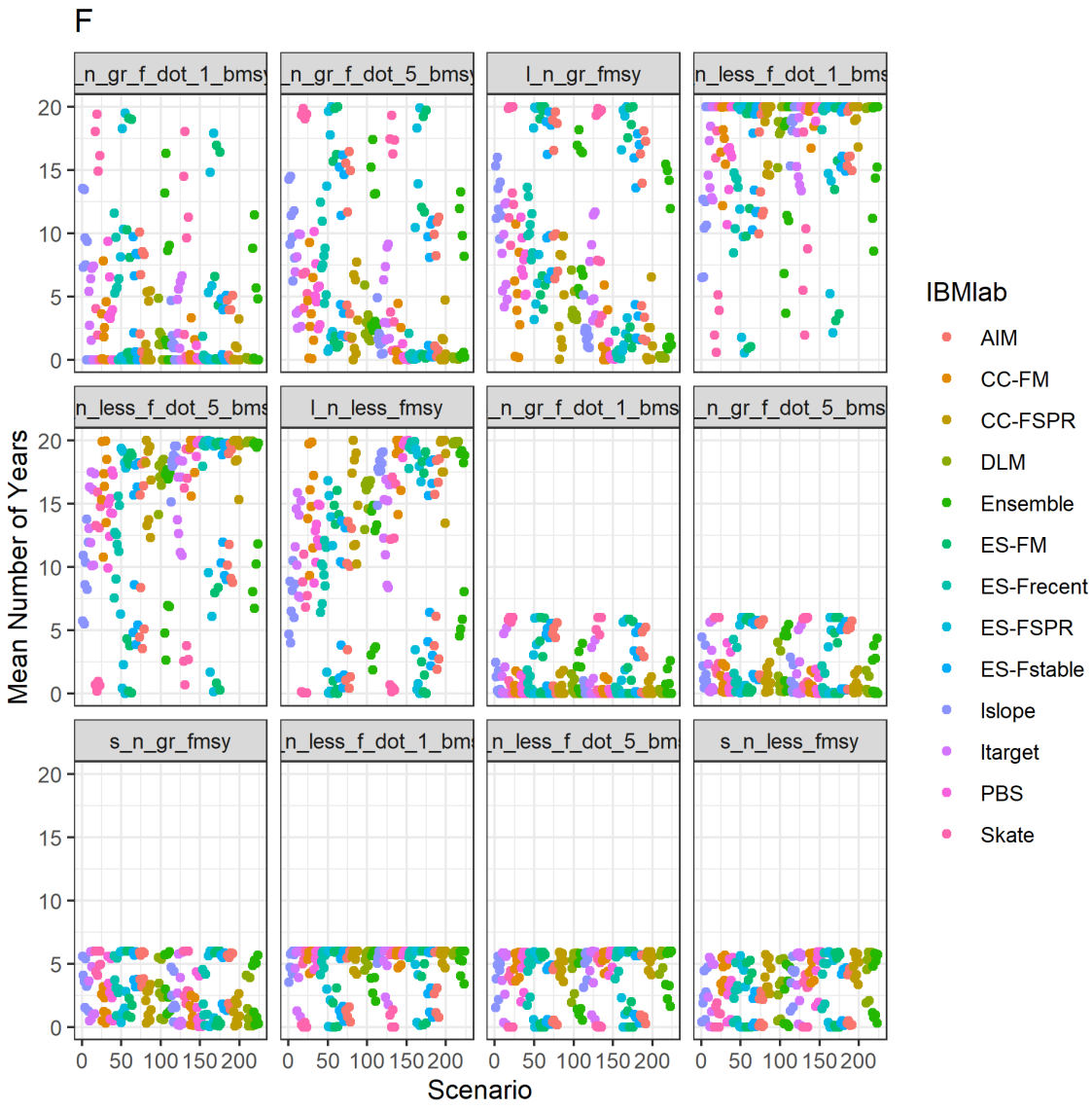


Figure A6.160. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

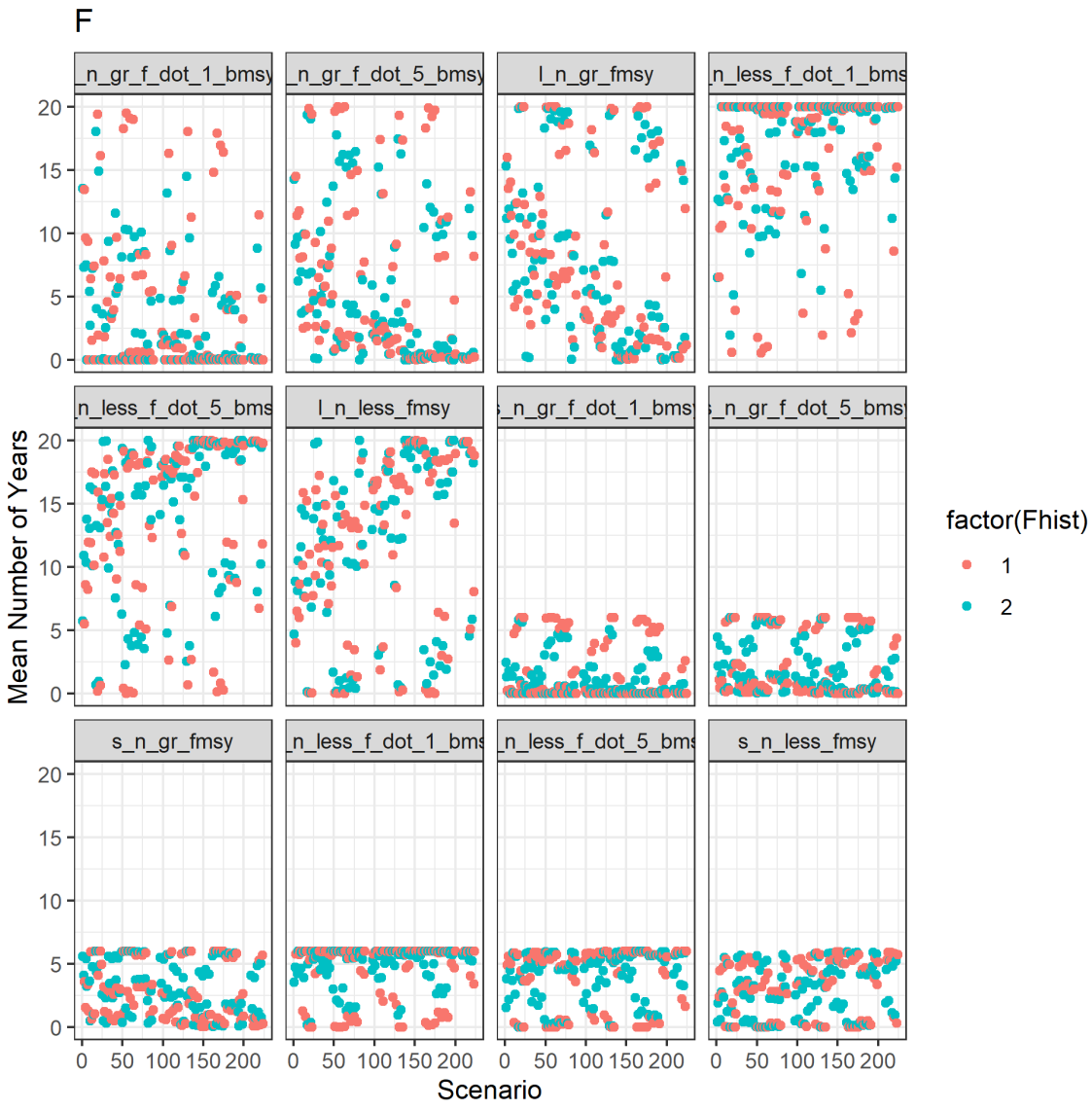


Figure A6.161. Confetti plot of the mean values for some of the fishing rate mortality (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

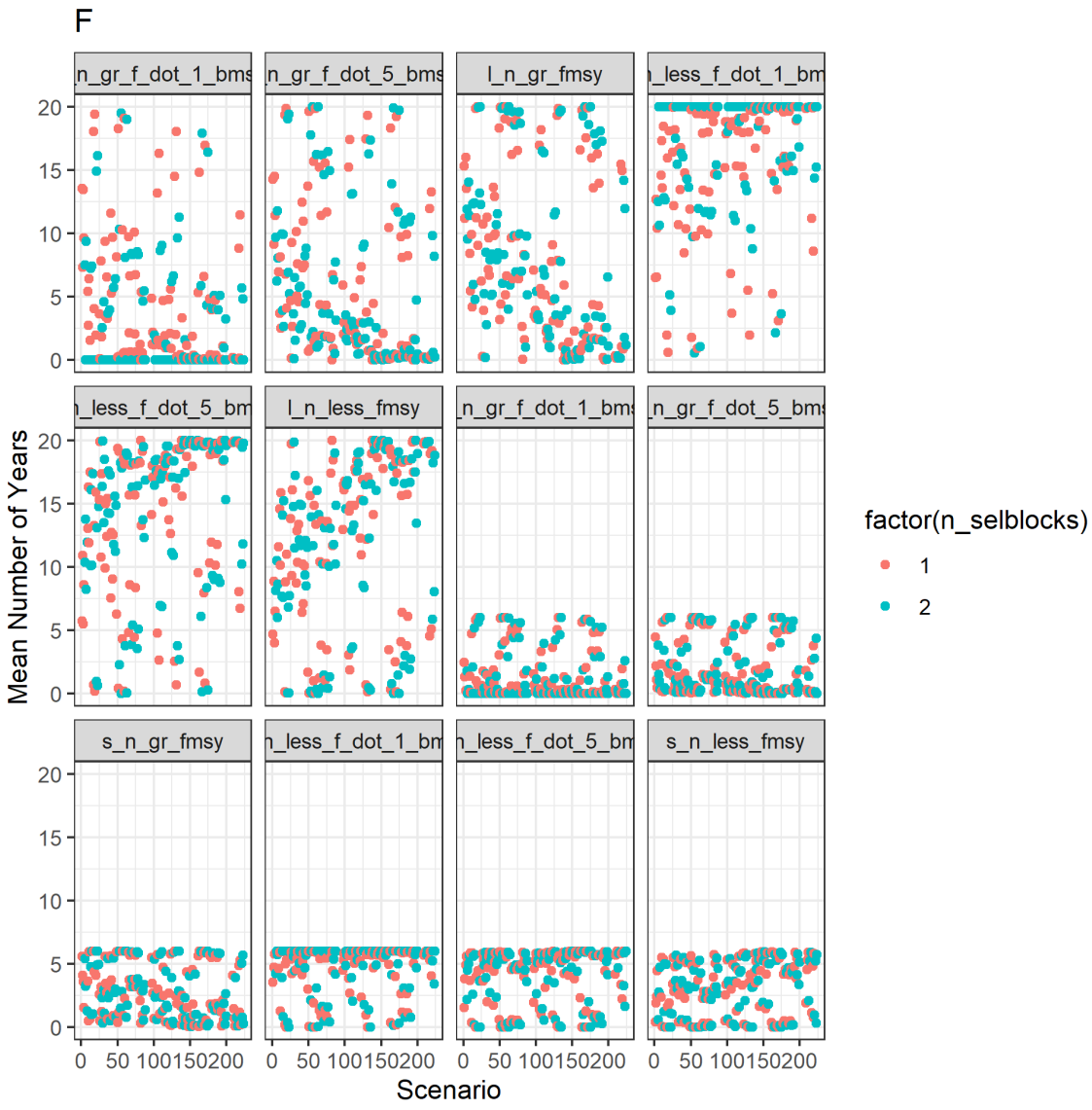


Figure A6.162. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

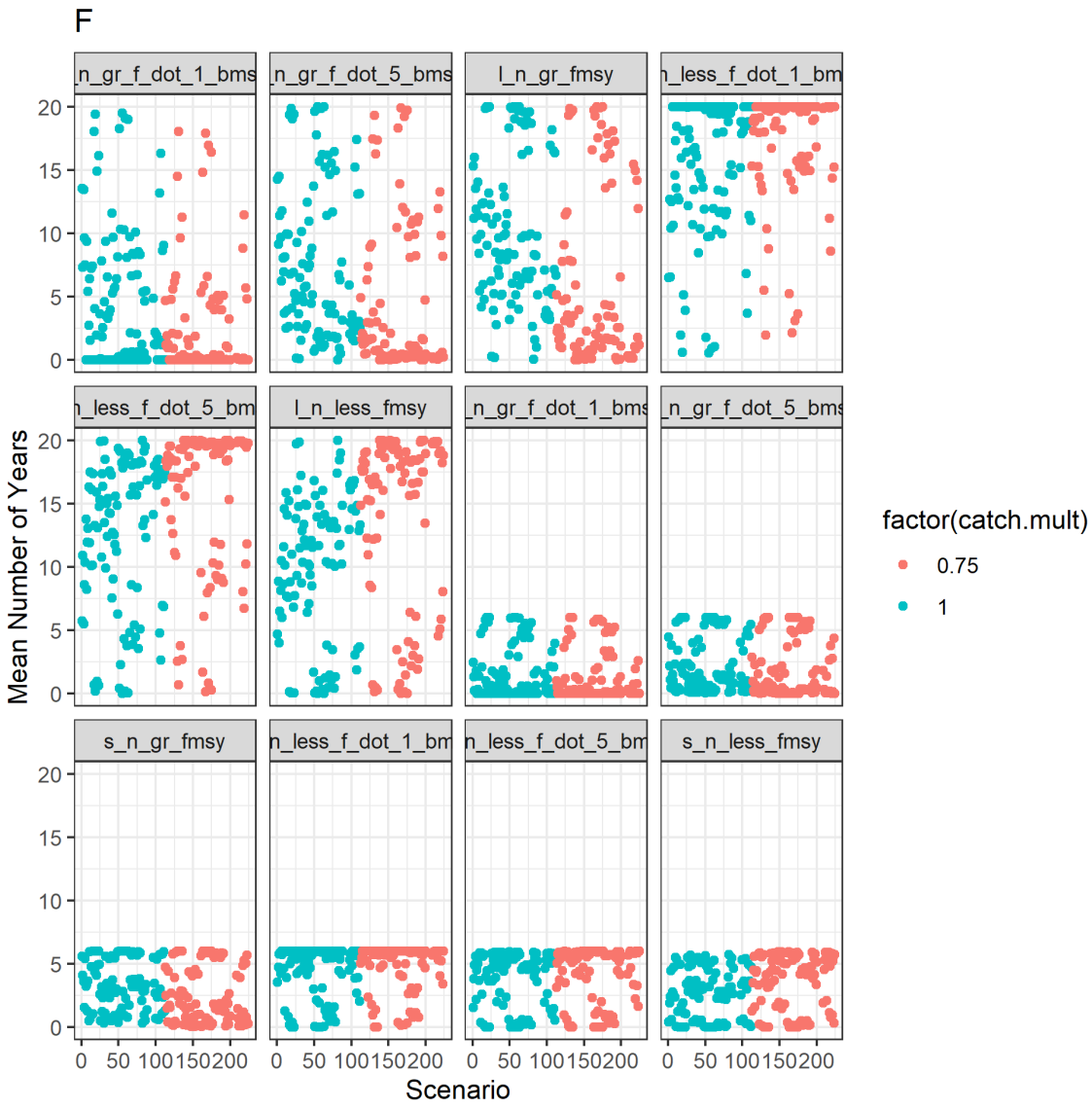


Figure A6.163. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

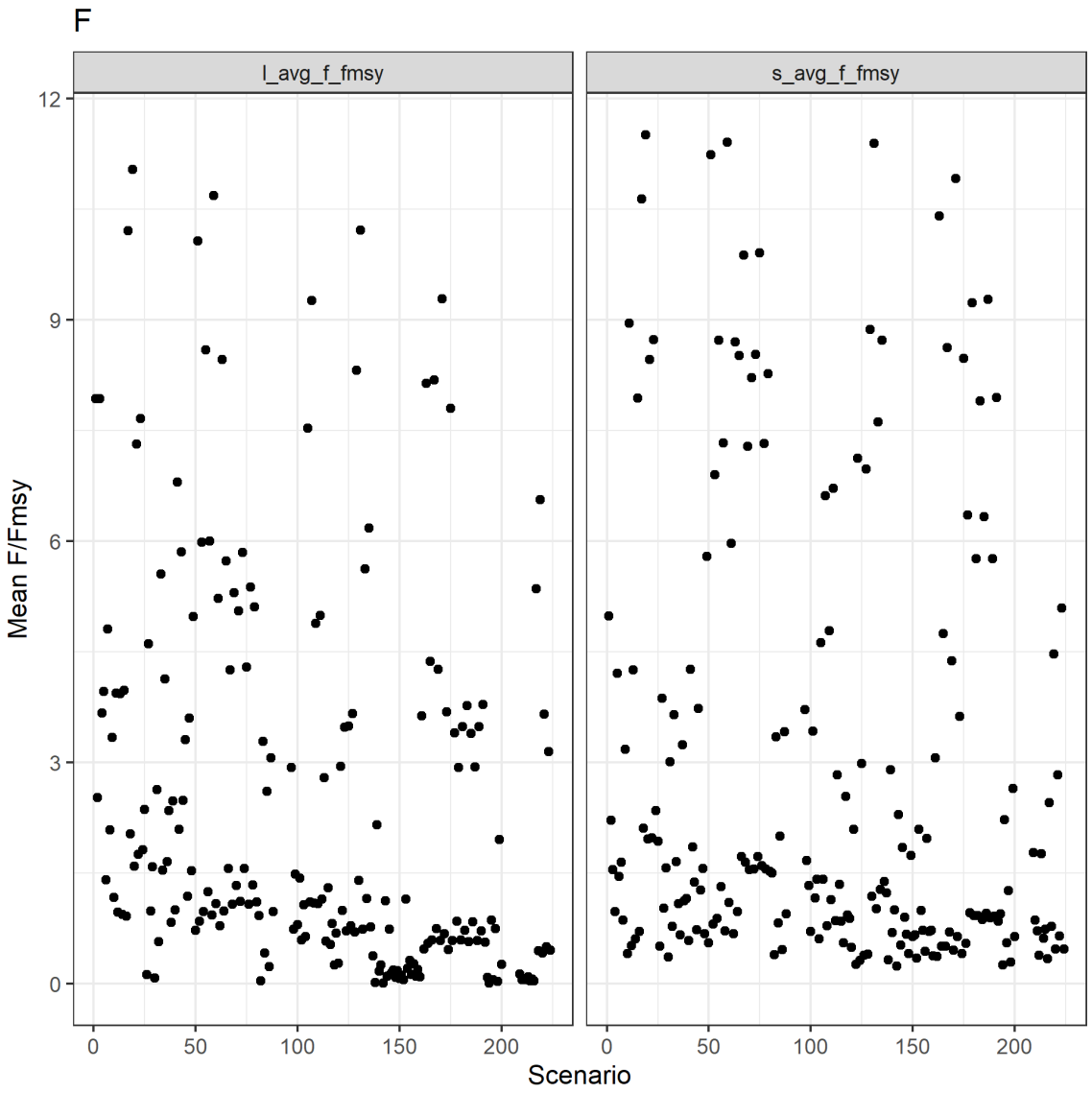


Figure A6.164. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

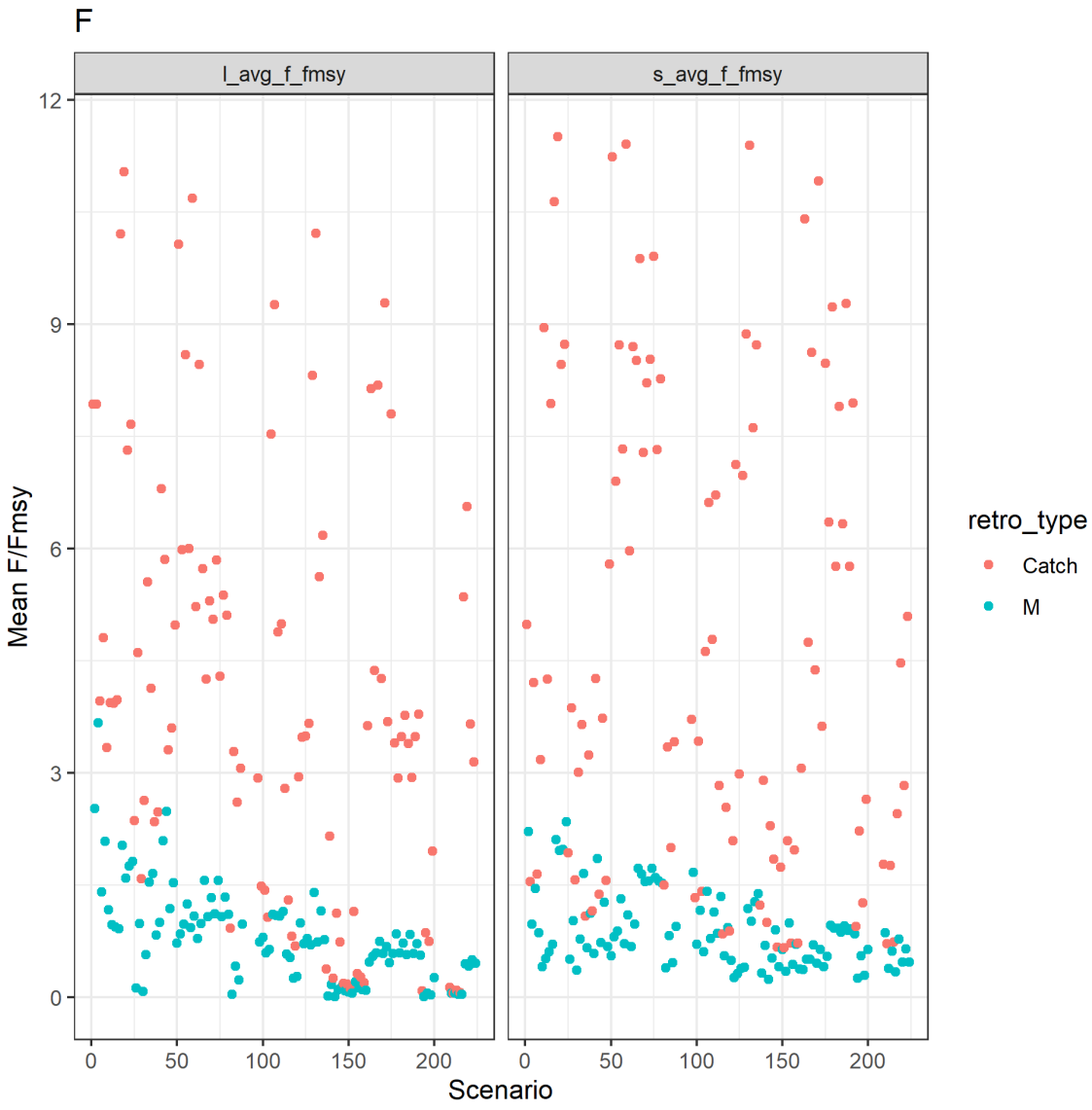


Figure A6.165. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

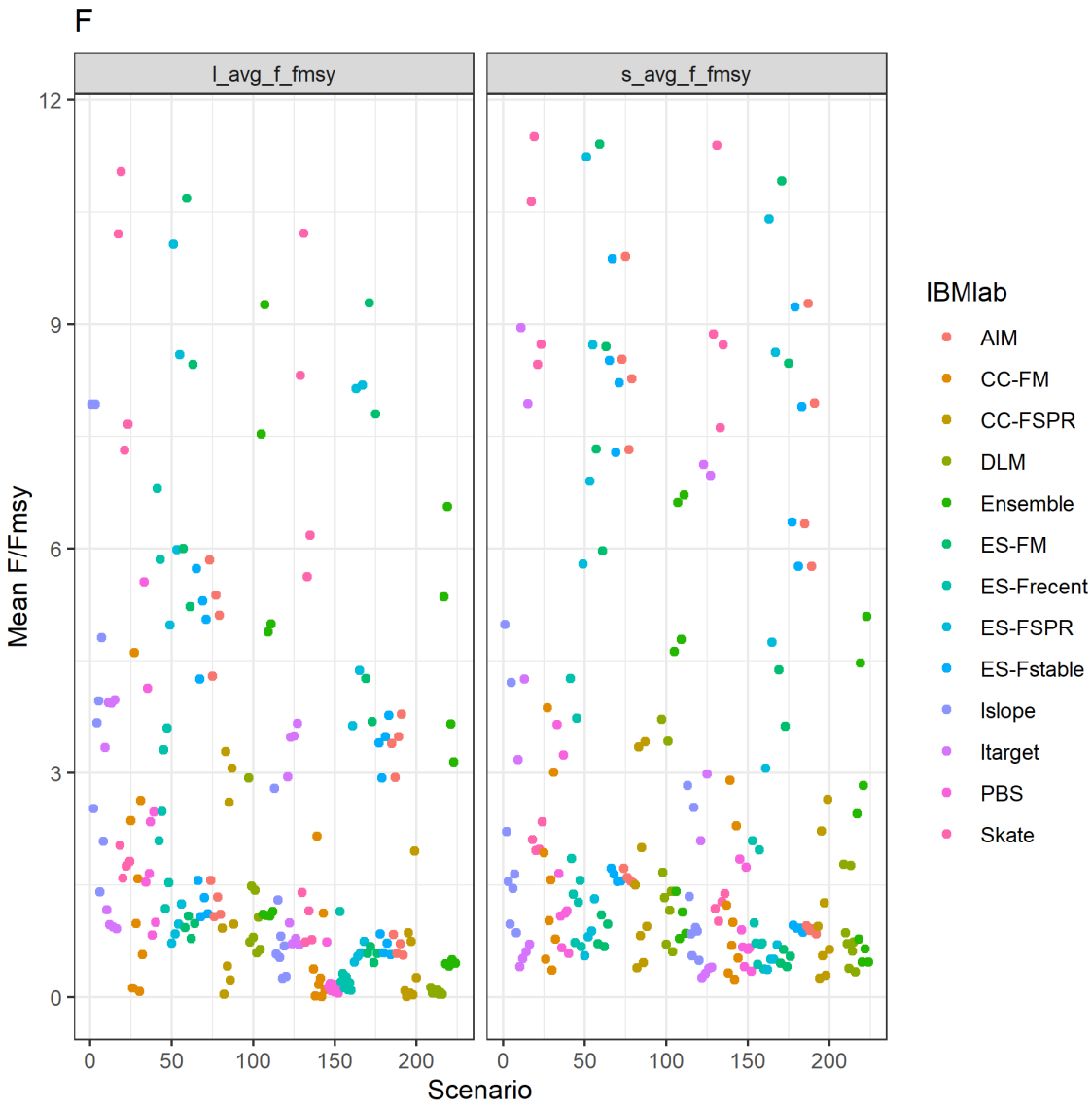


Figure A6.166. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

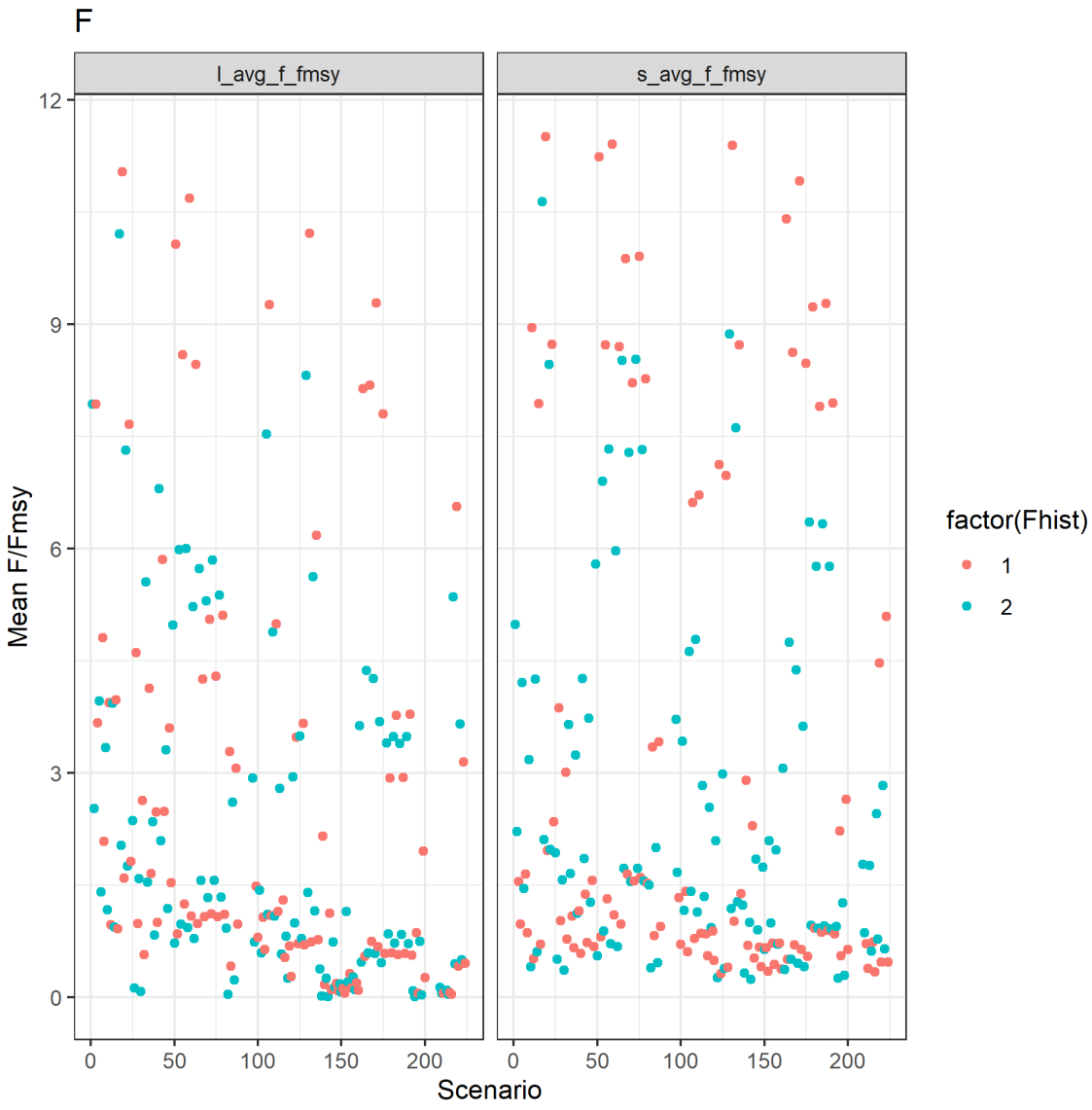


Figure A6.167. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

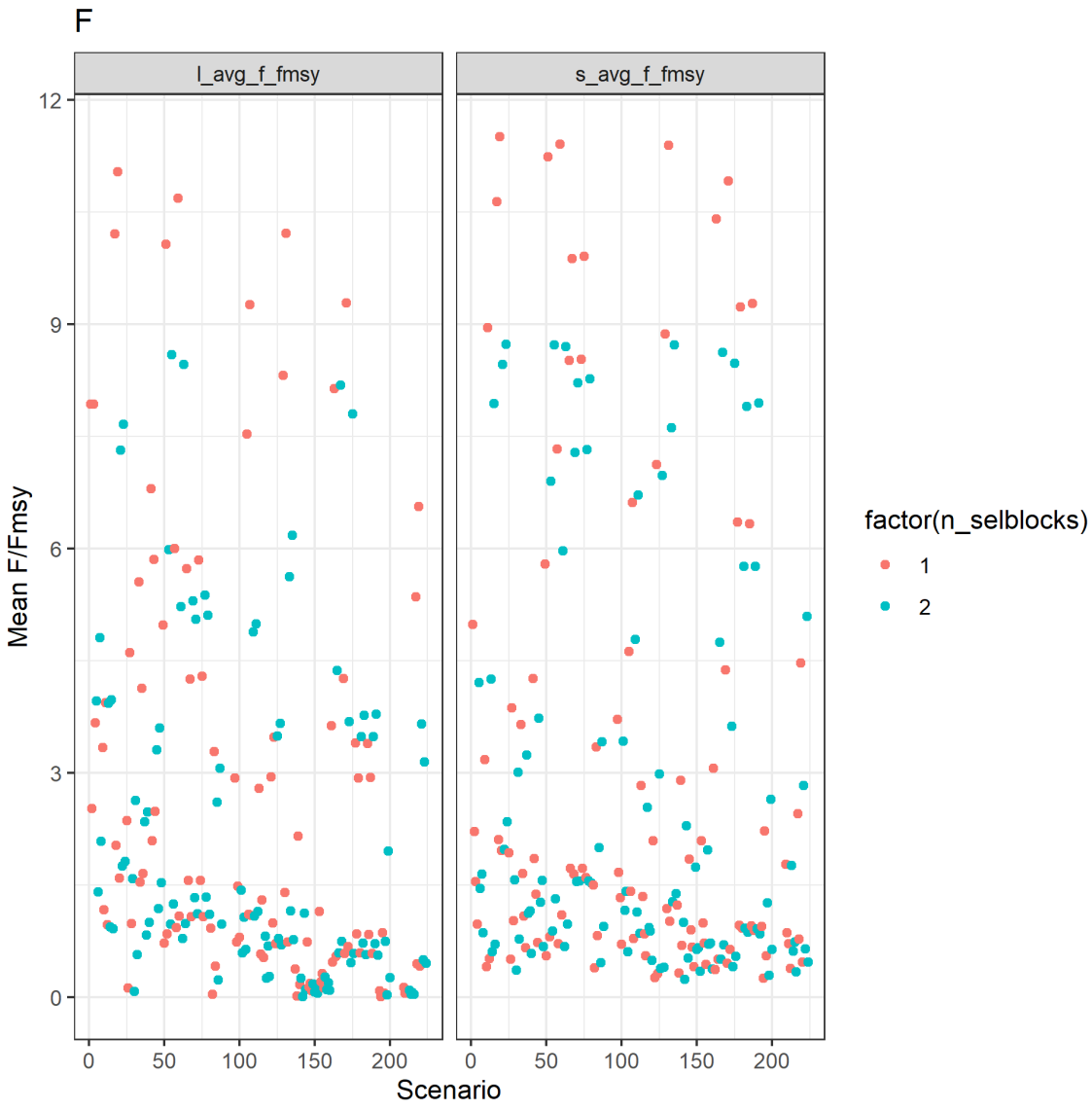


Figure A6.168. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

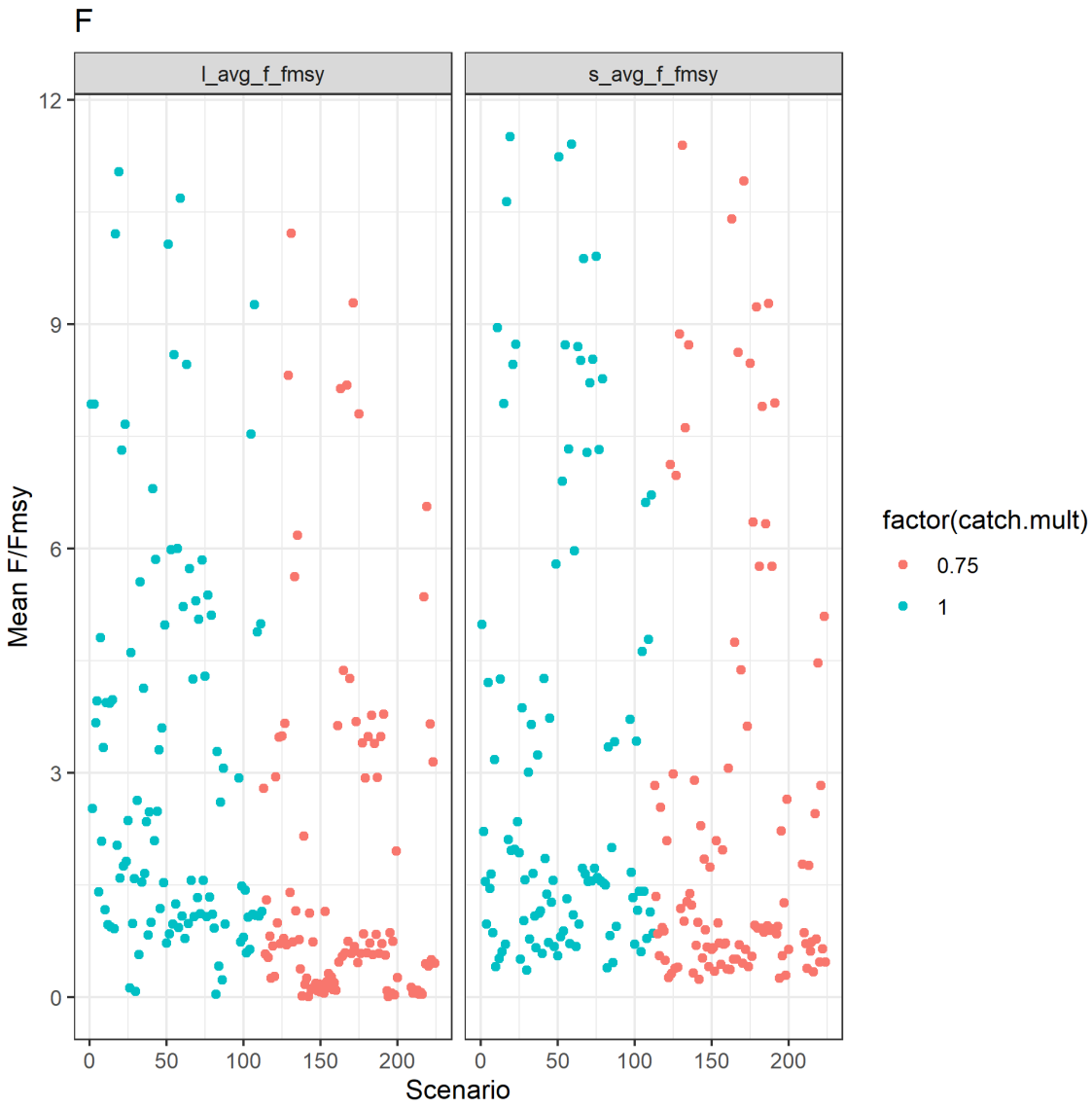


Figure A6.169. Confetti plot of the mean values for some of the fishing mortality rate (F) metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

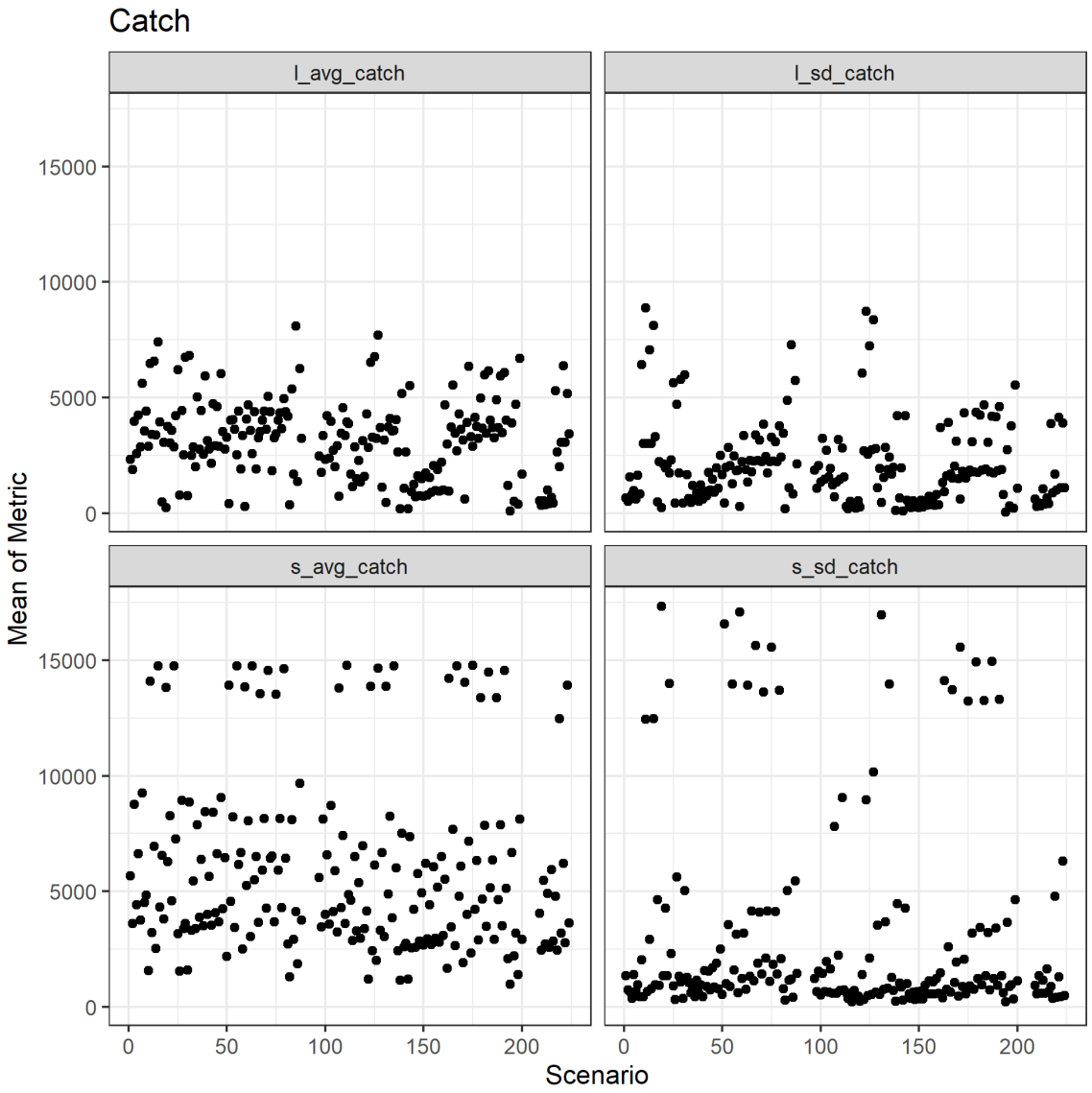


Figure A6.170. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

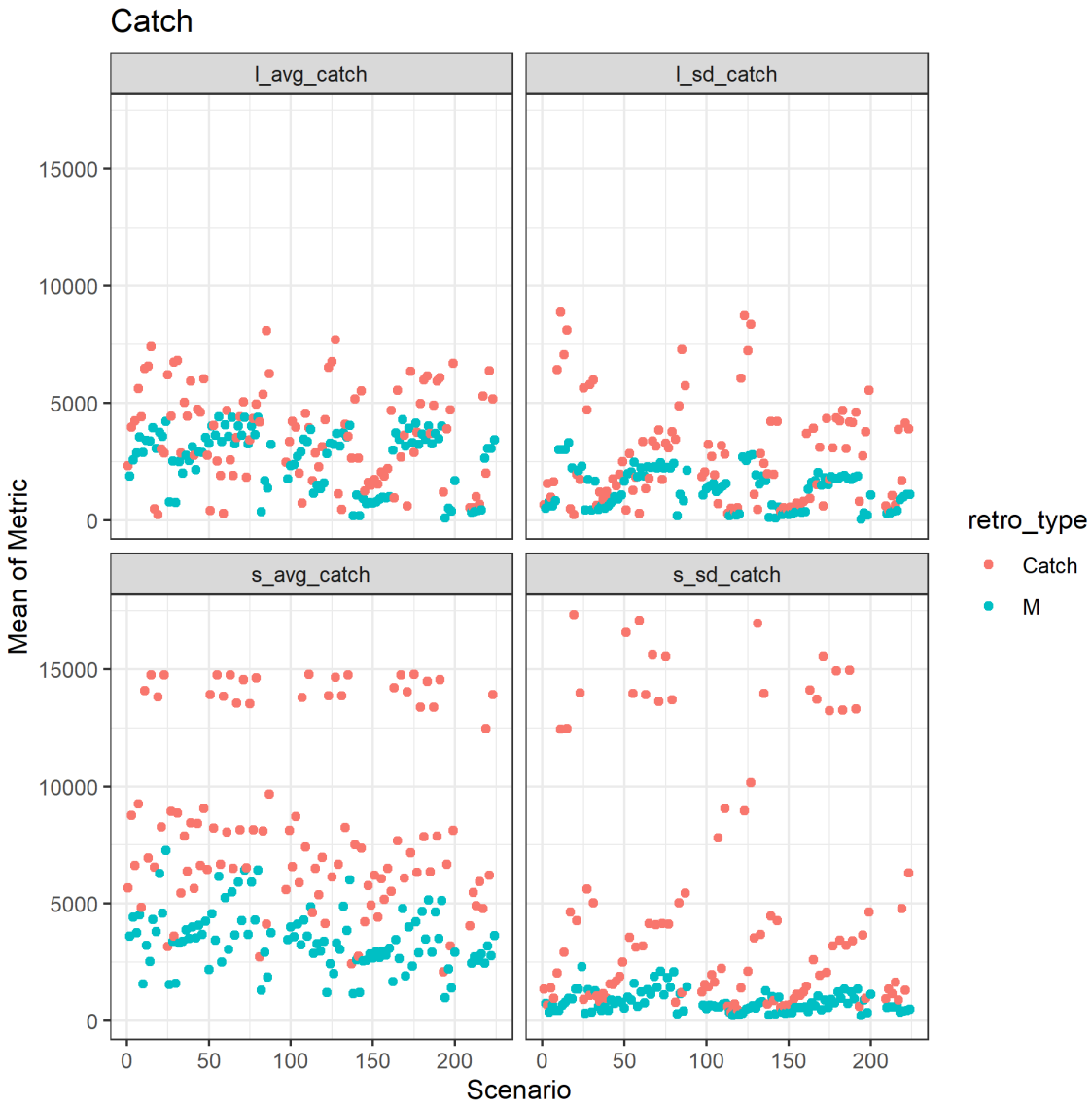


Figure A6.171. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

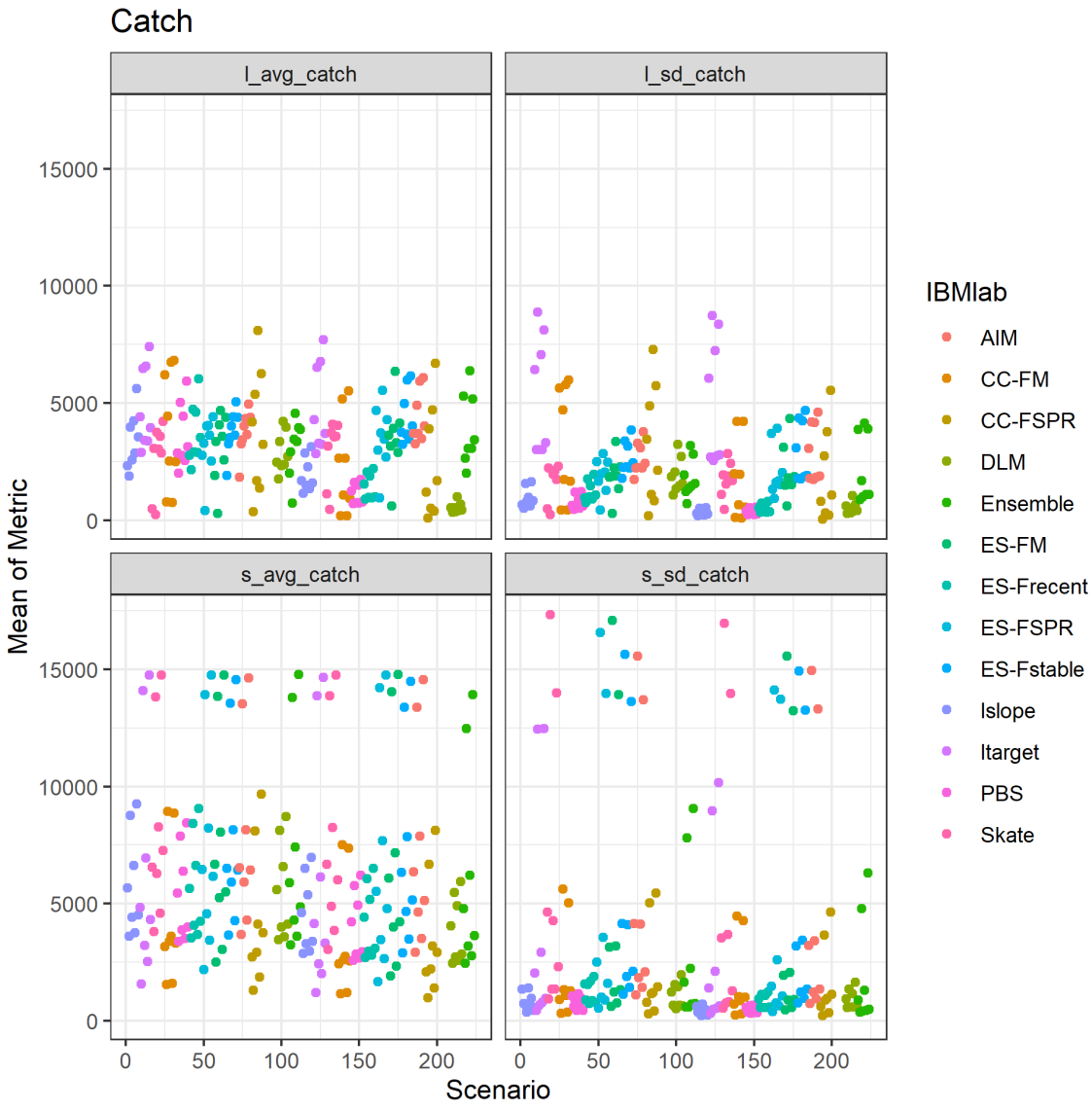


Figure A6.172. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

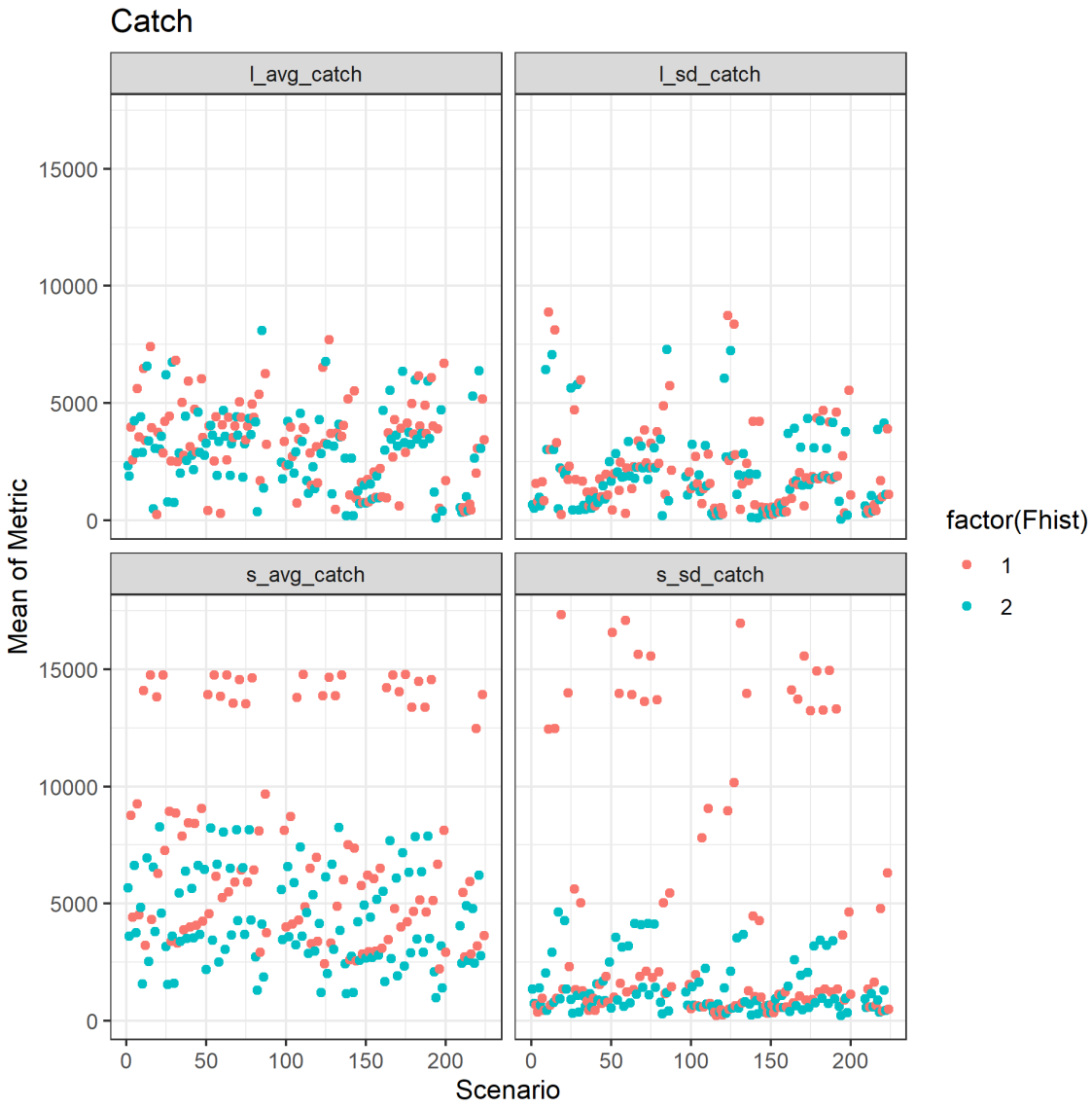


Figure A6.173. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

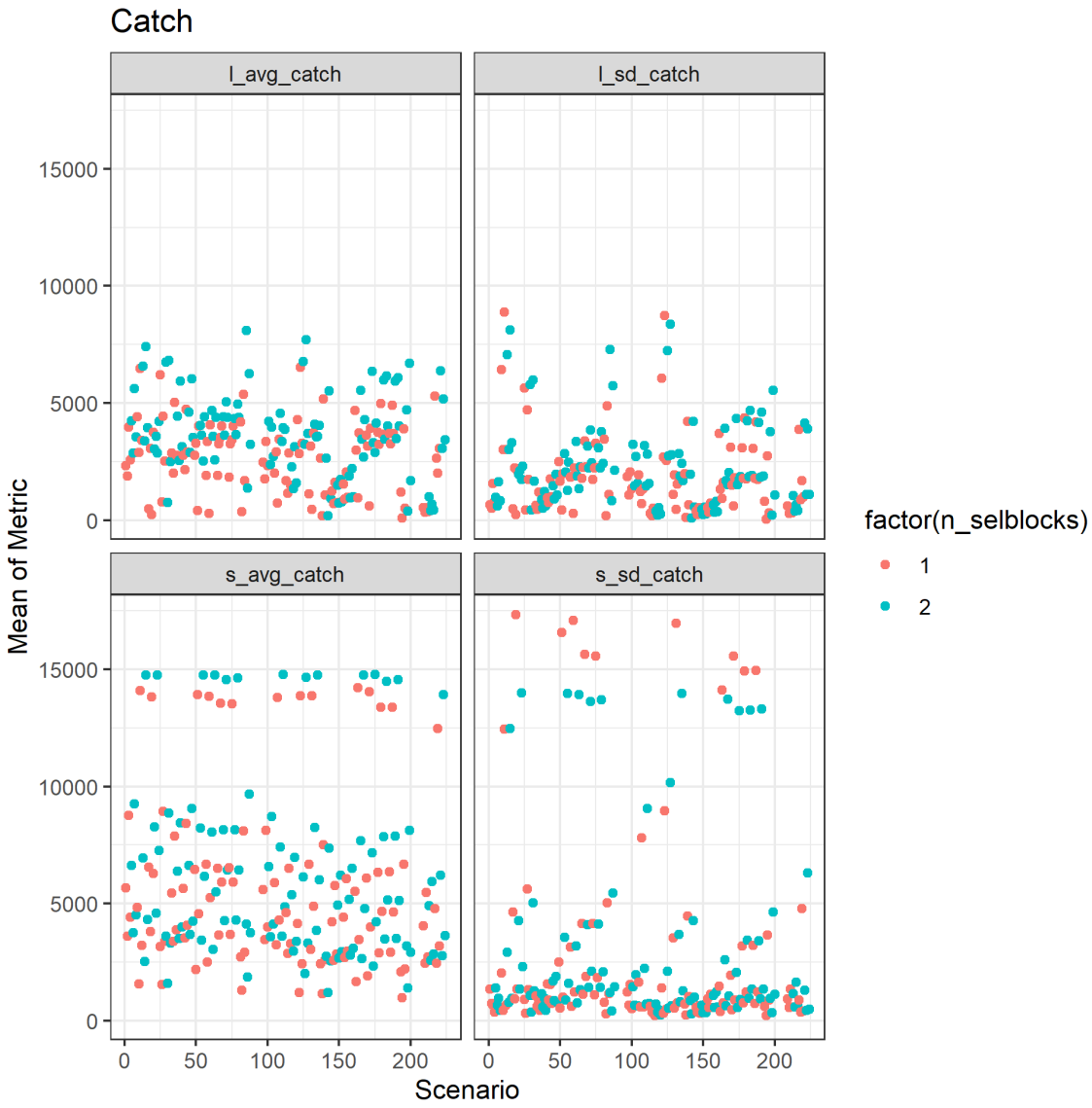


Figure A6.174. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

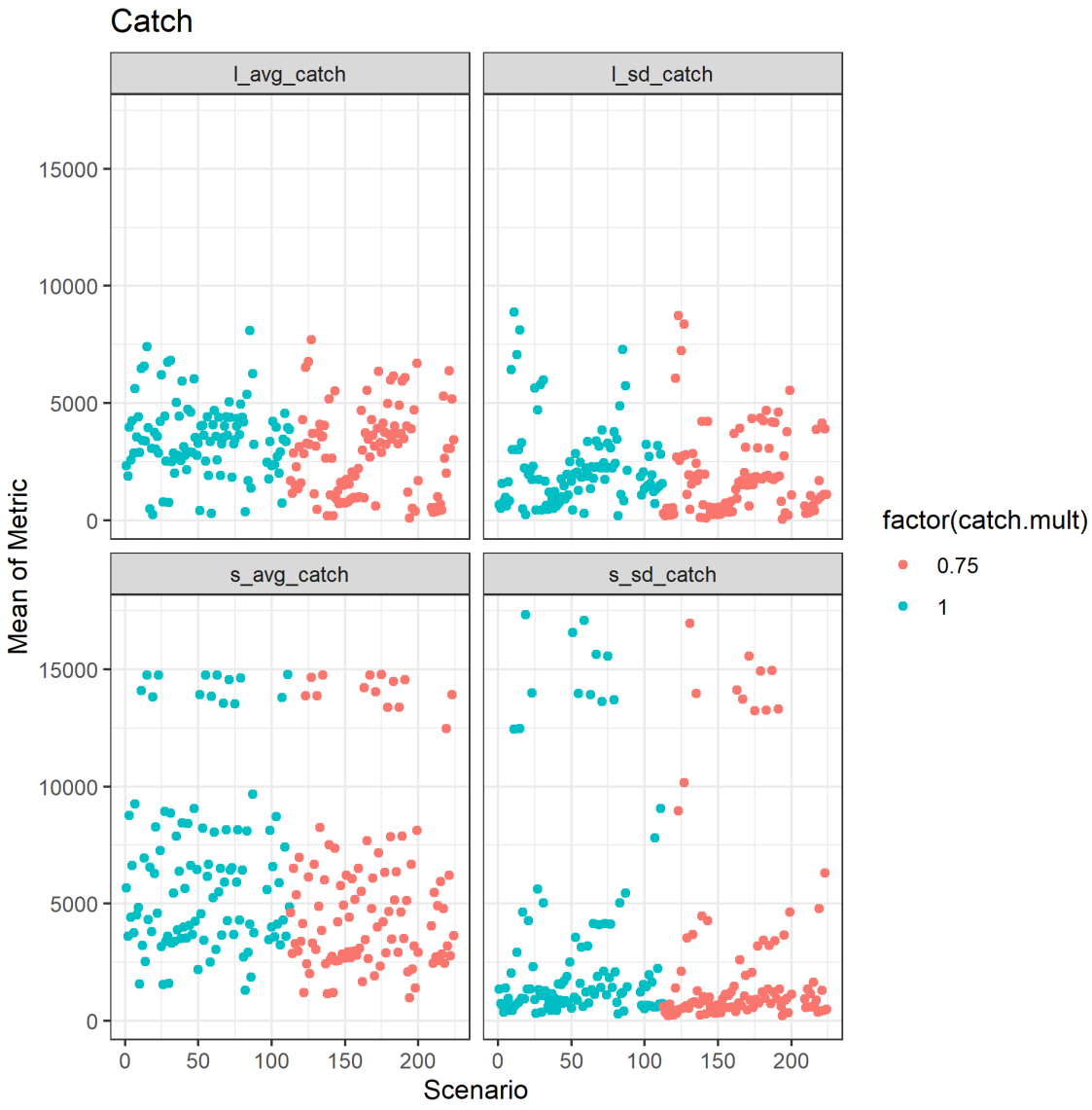


Figure A6.175. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

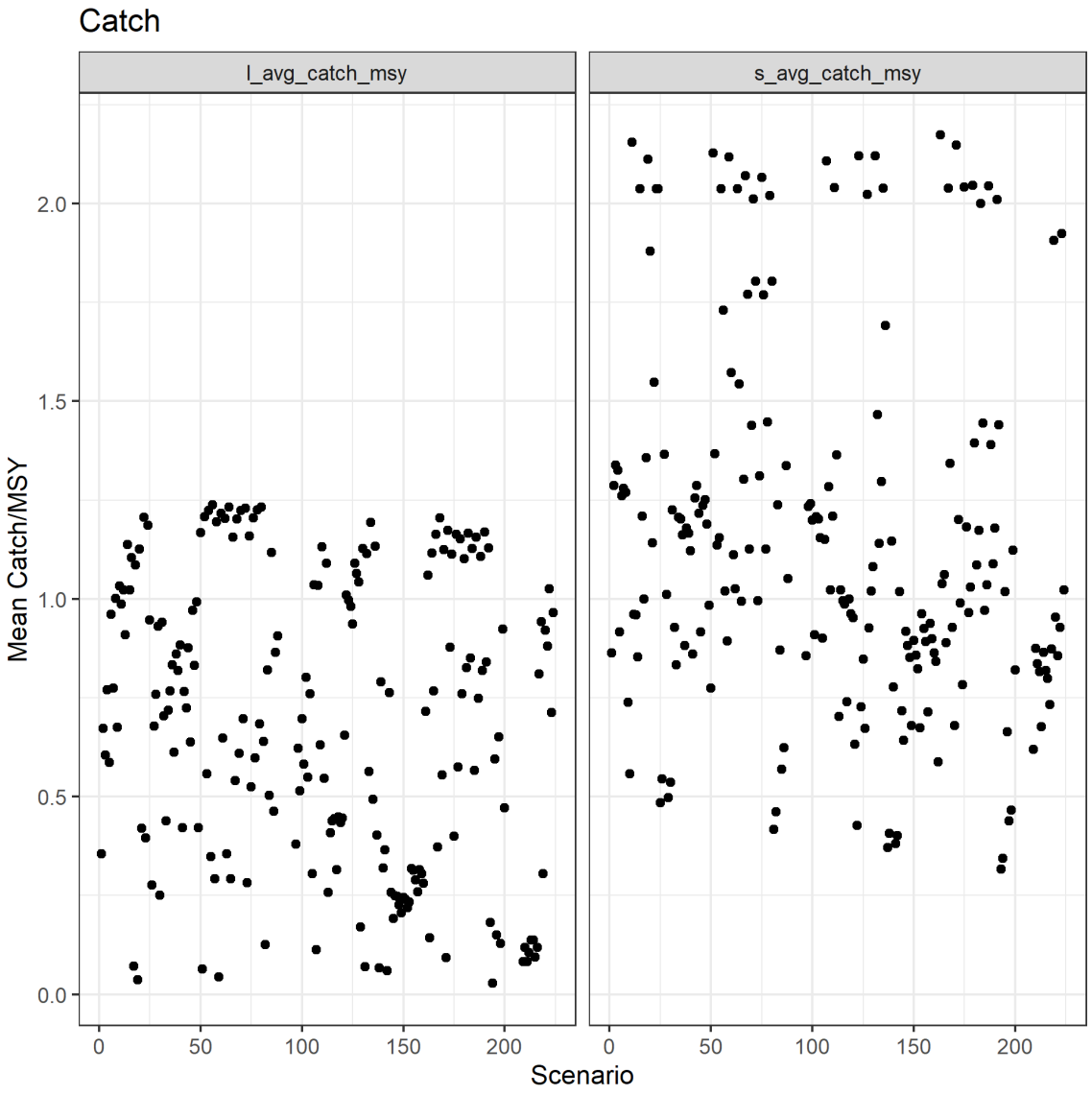


Figure A6.176. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

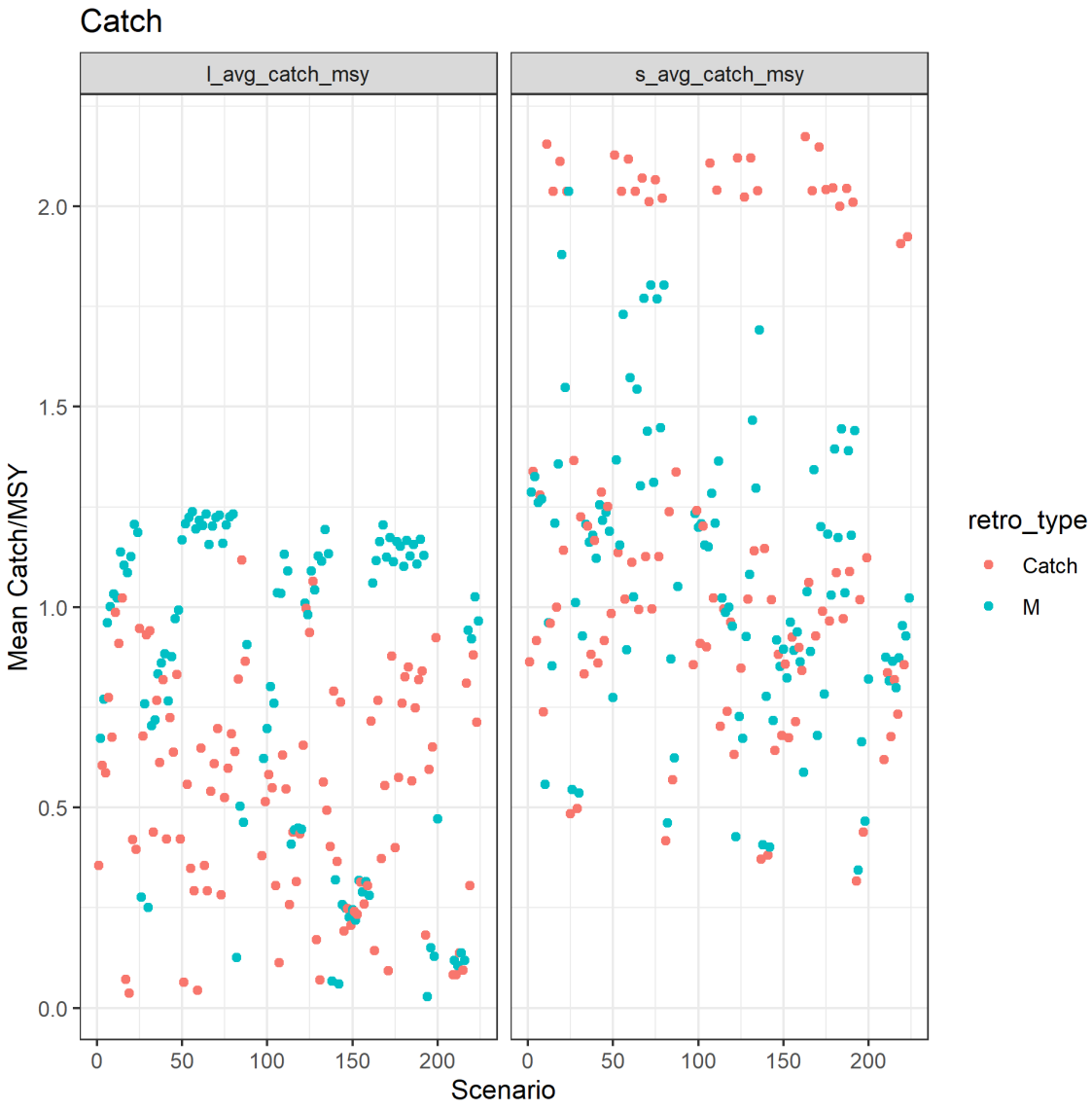


Figure A6.177. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

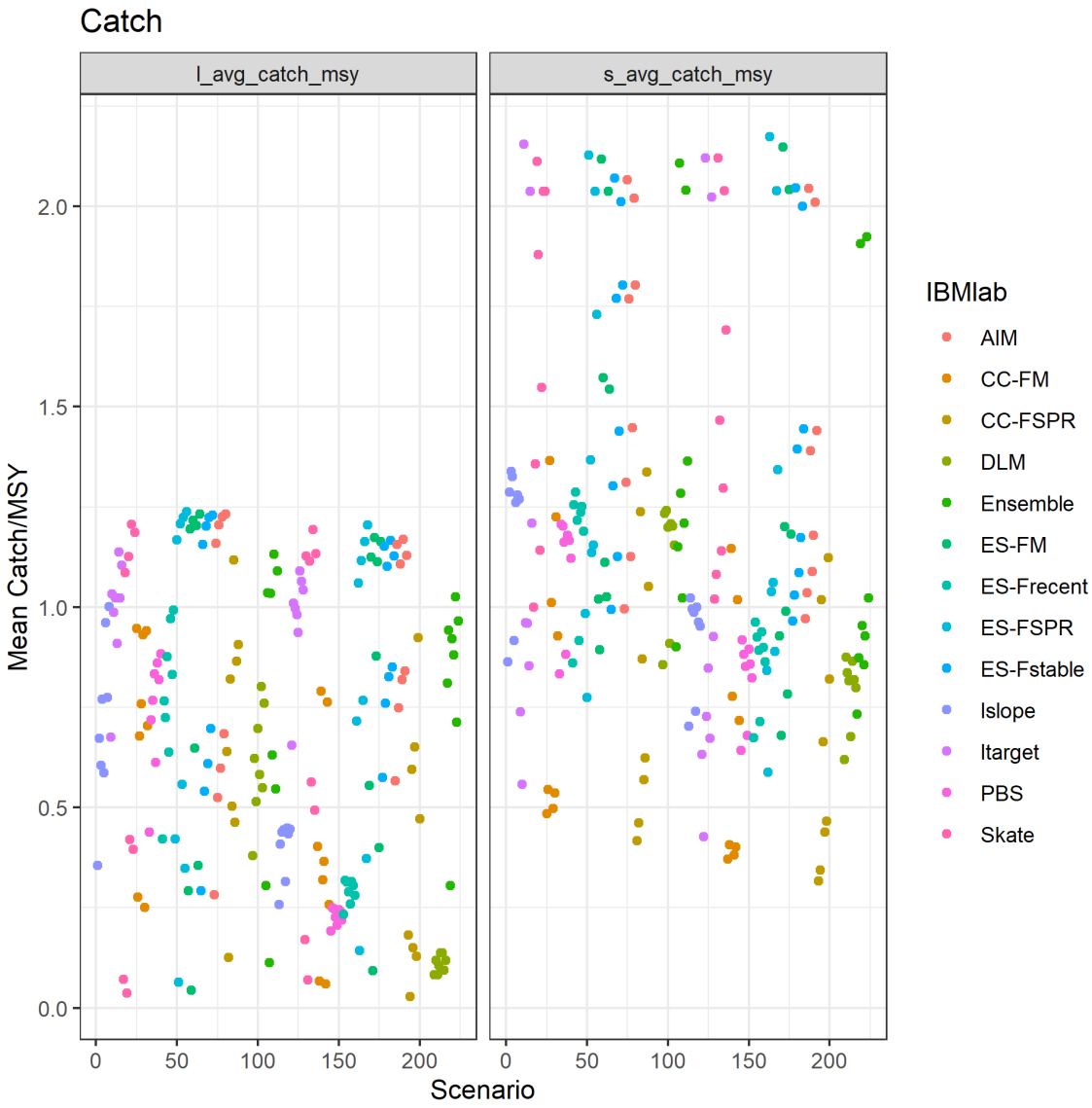


Figure A6.178. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

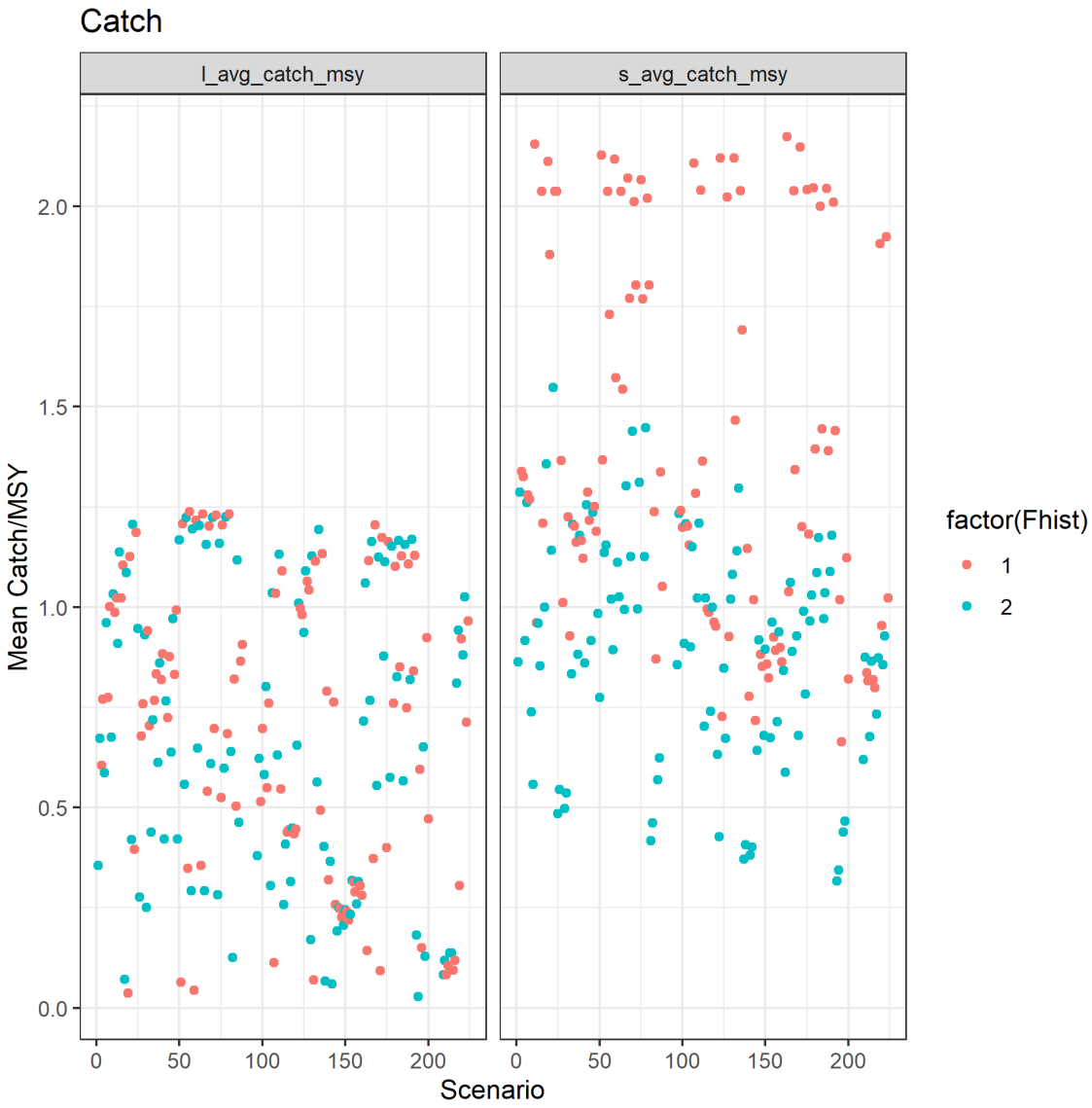


Figure A6.179. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

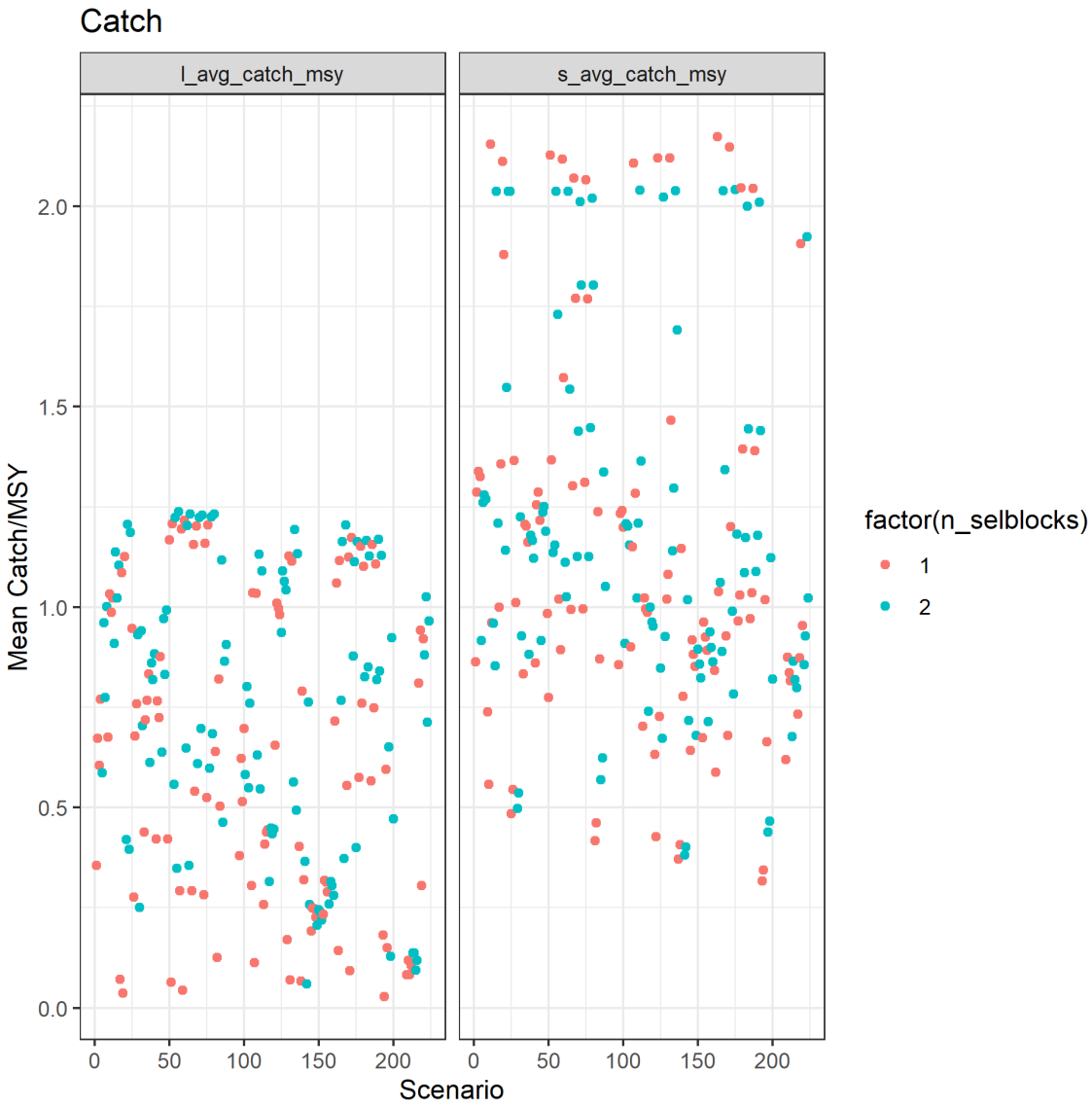


Figure A6.180. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

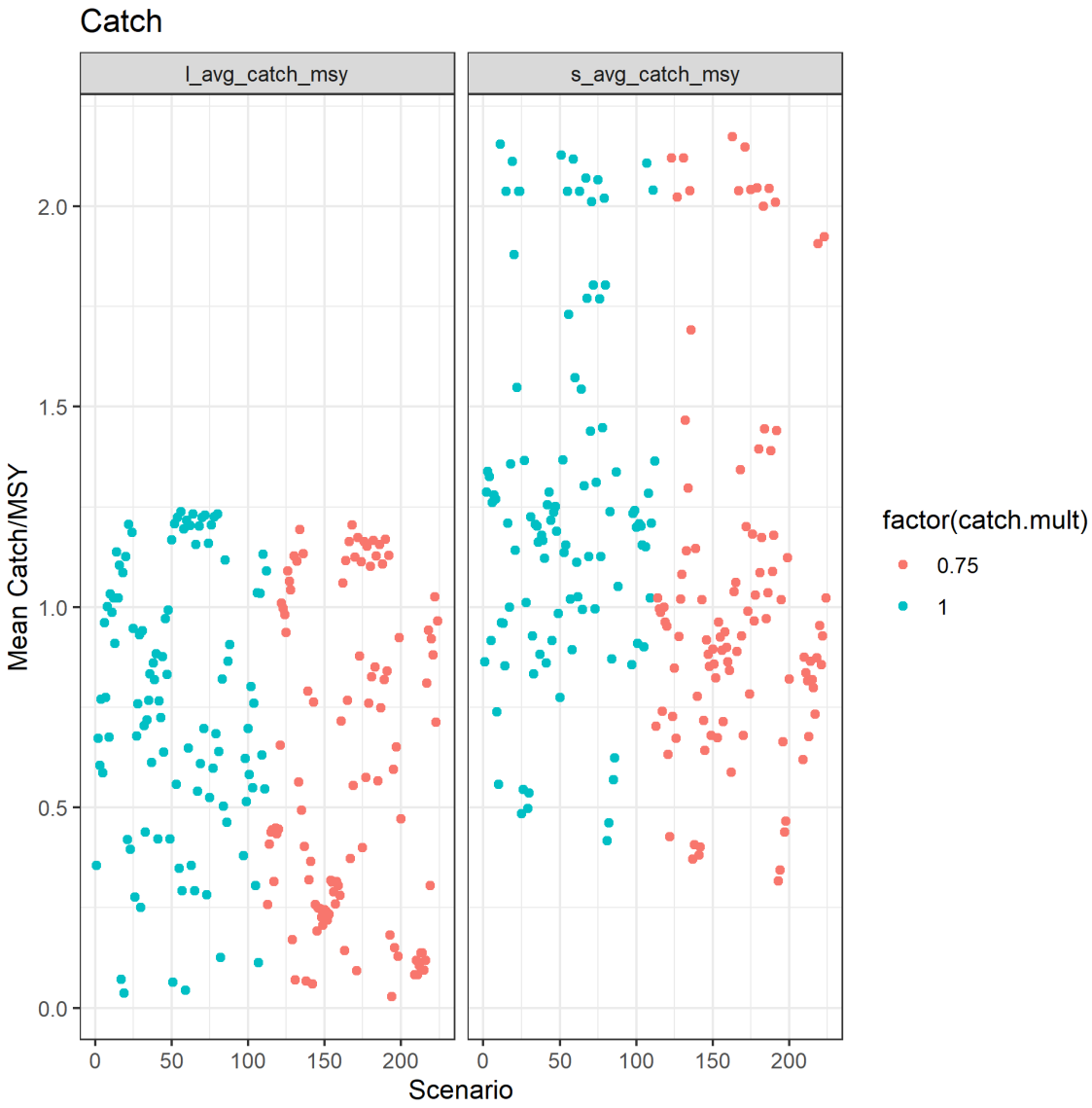


Figure A6.181. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

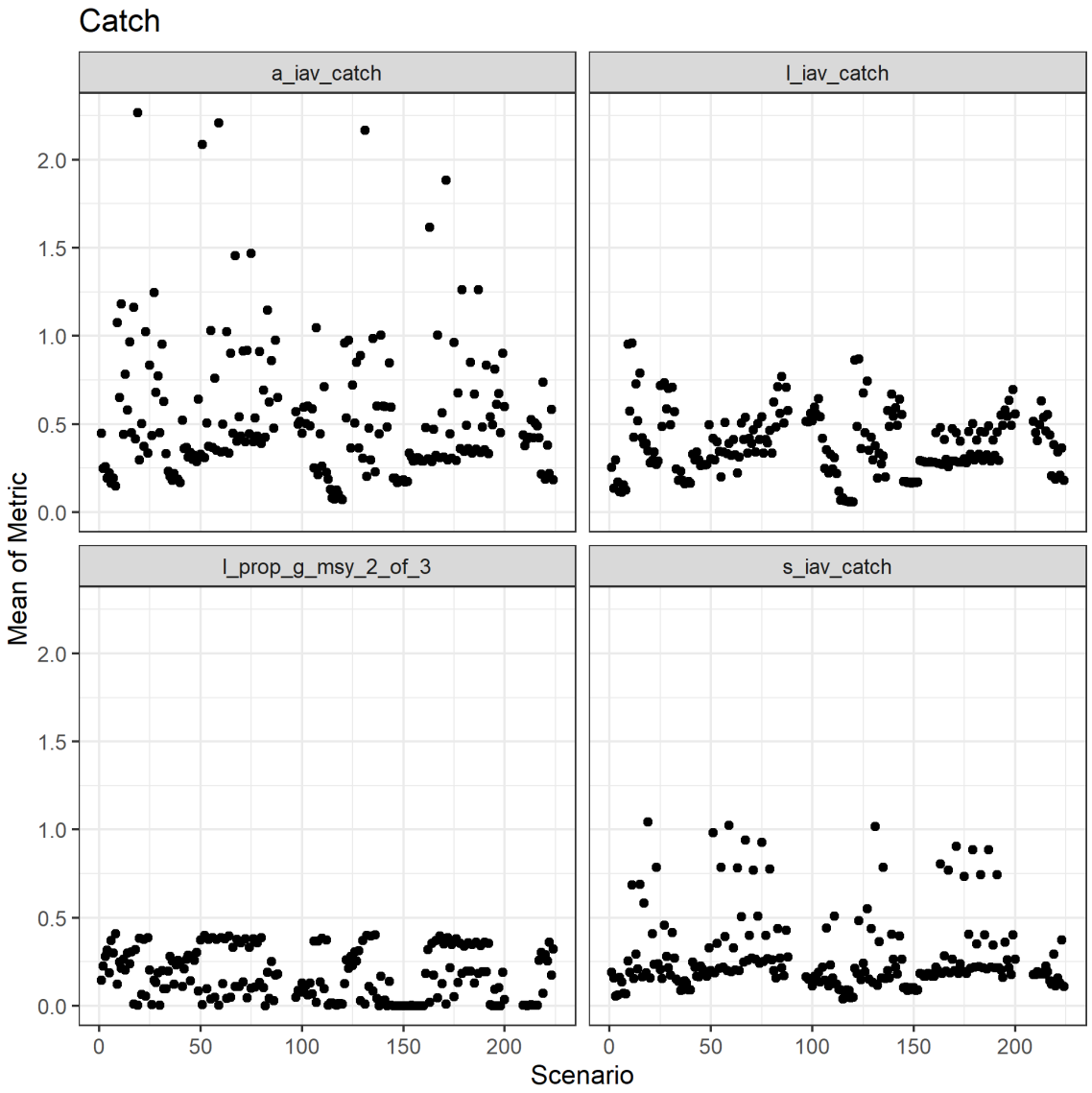


Figure A6.182. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

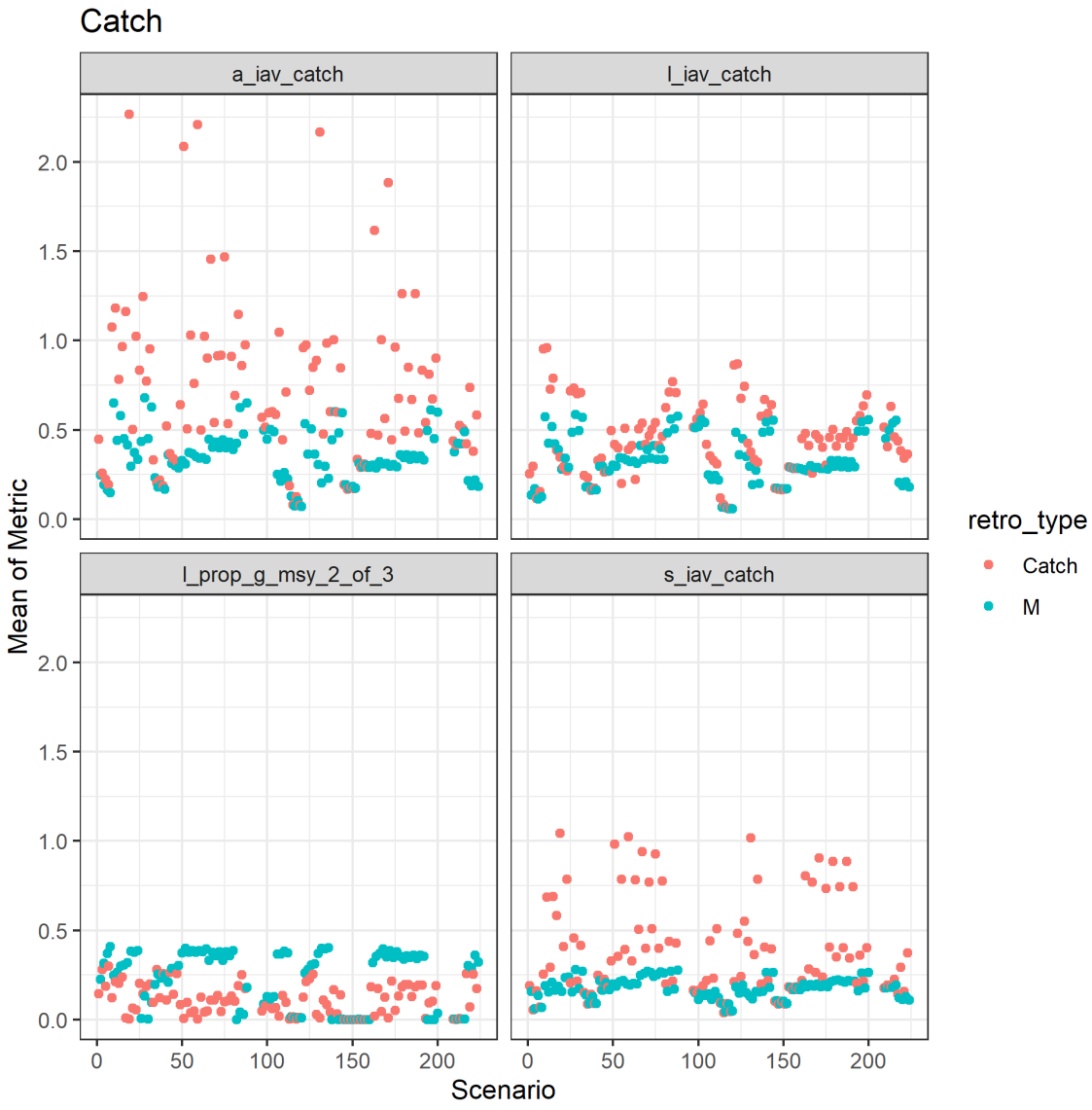


Figure A6.183. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

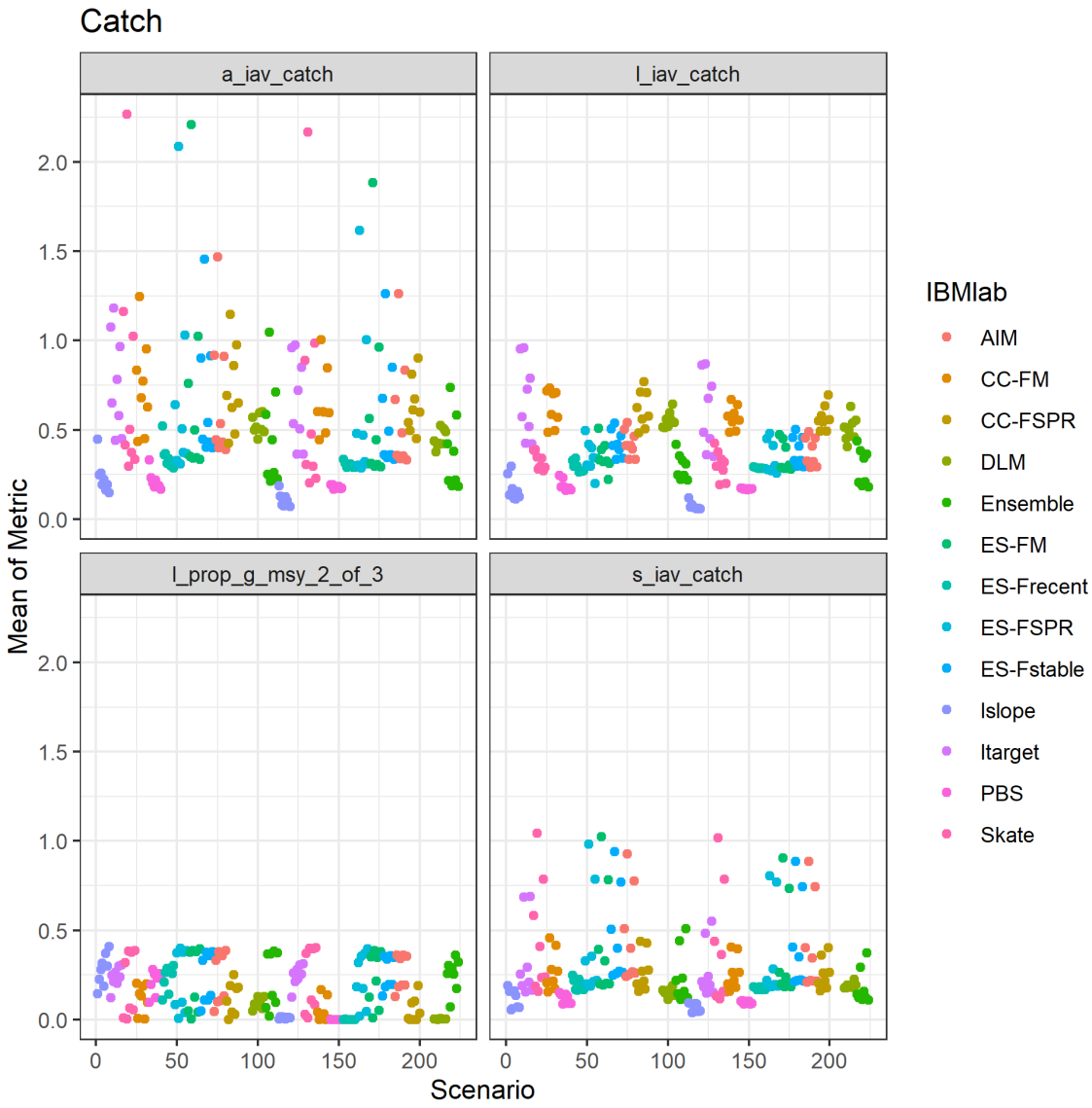


Figure A6.184. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

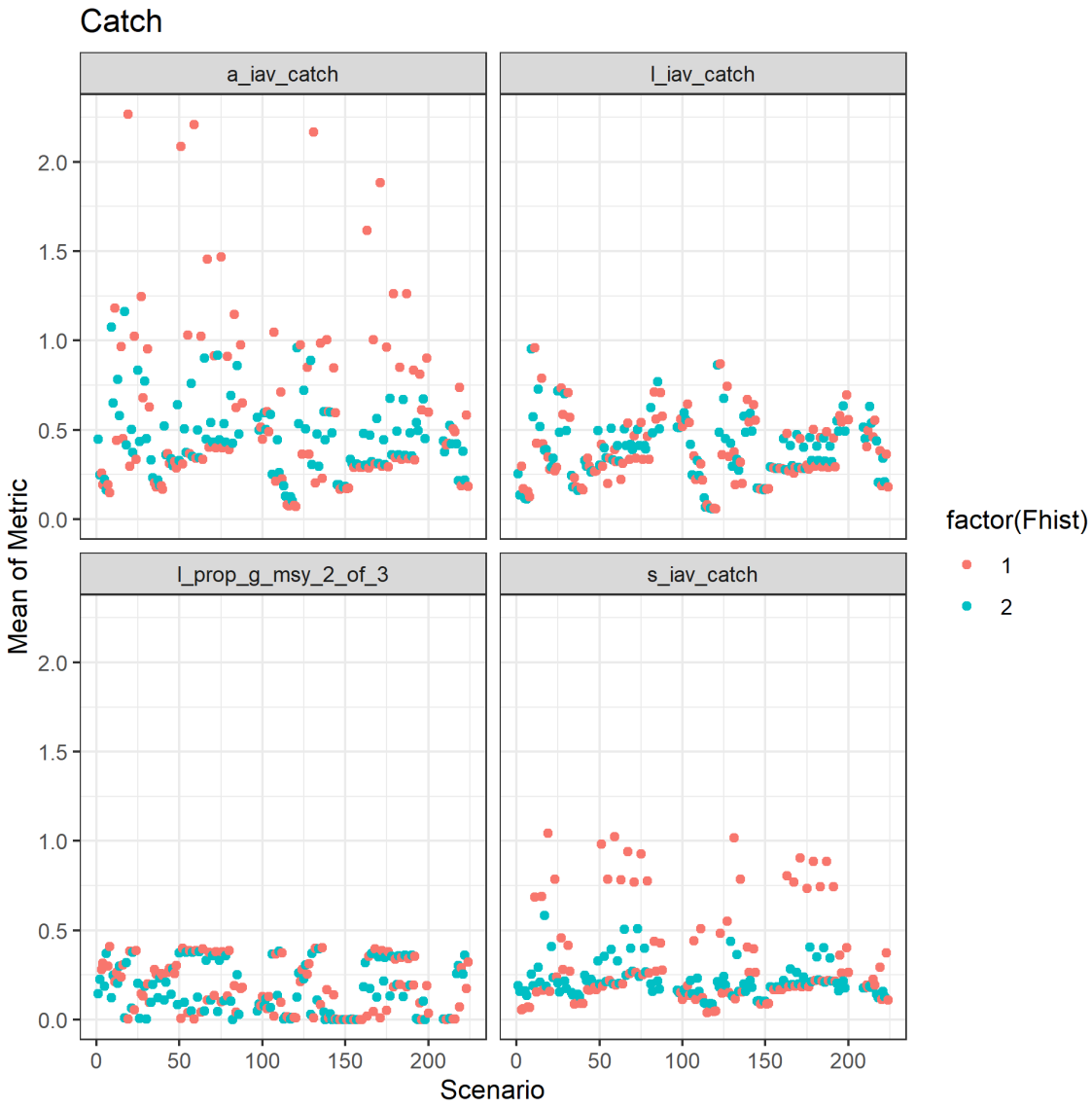


Figure A6.185. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

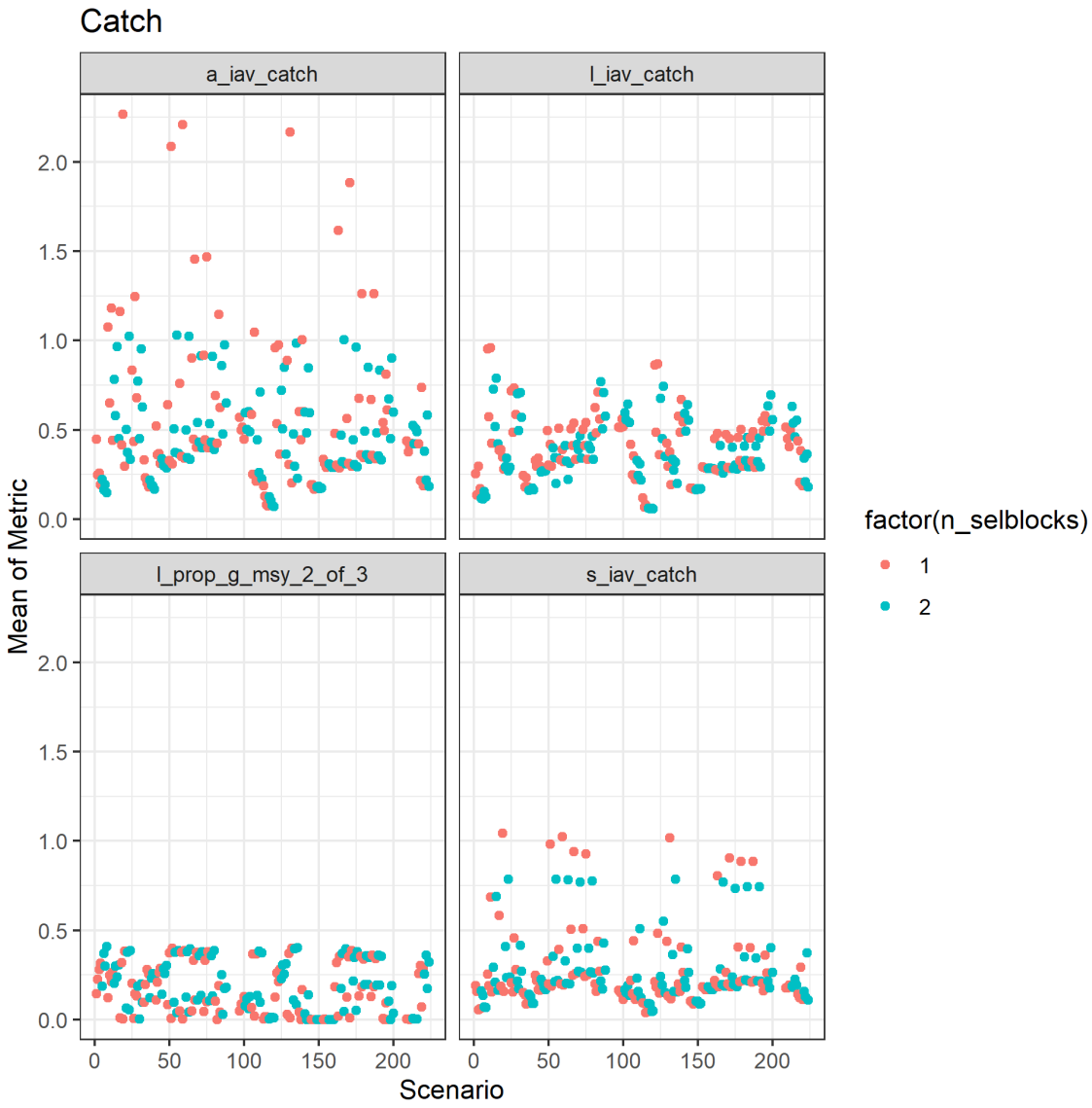


Figure A6.186. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

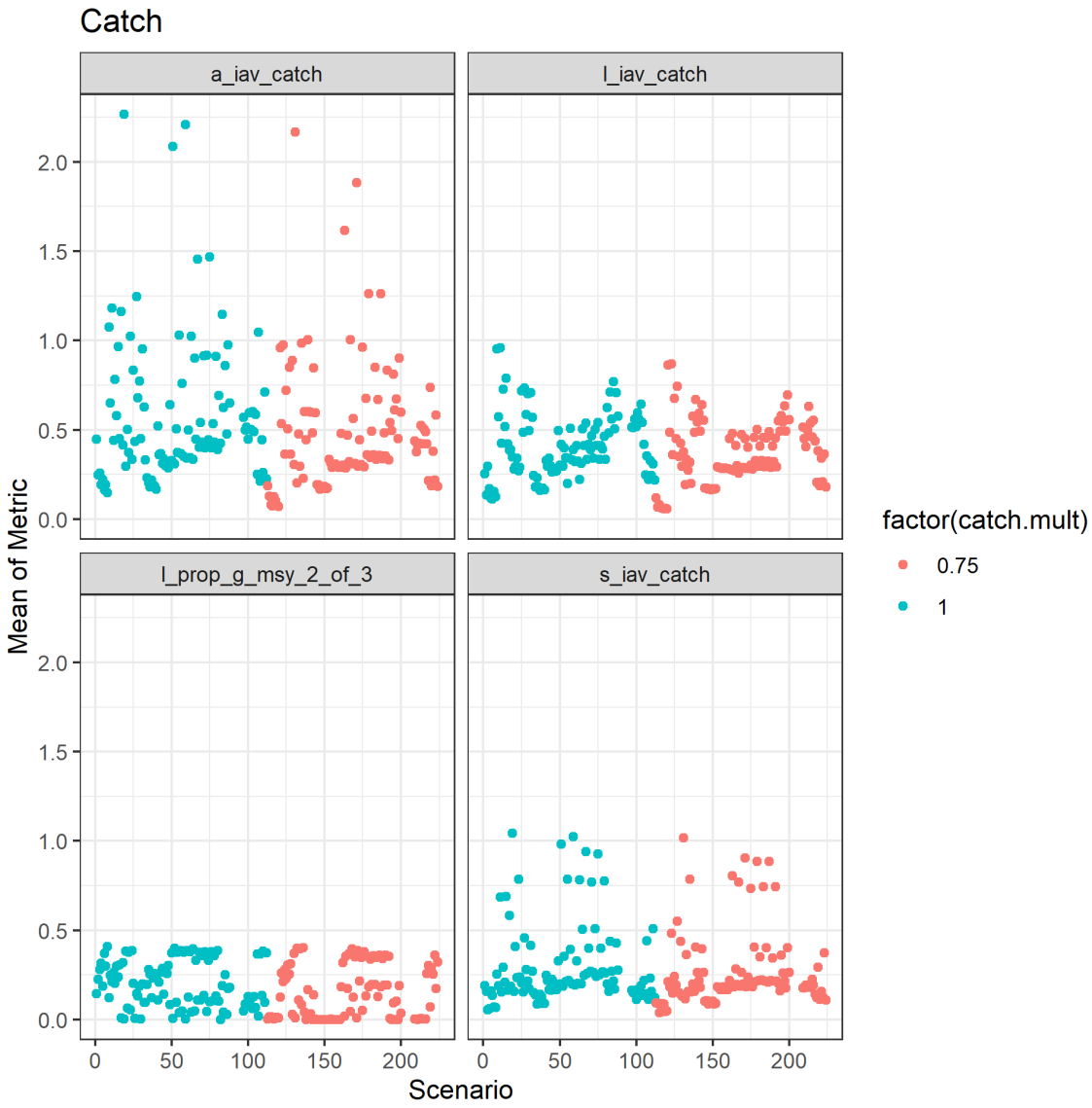


Figure A6.187. Confetti plot of the mean values for some of the catch metrics in the base analyses by the combinations of index-based method (IBM) and scenario (denoted scenario in these plots). If colored, the colors denote different levels of 1 factor.

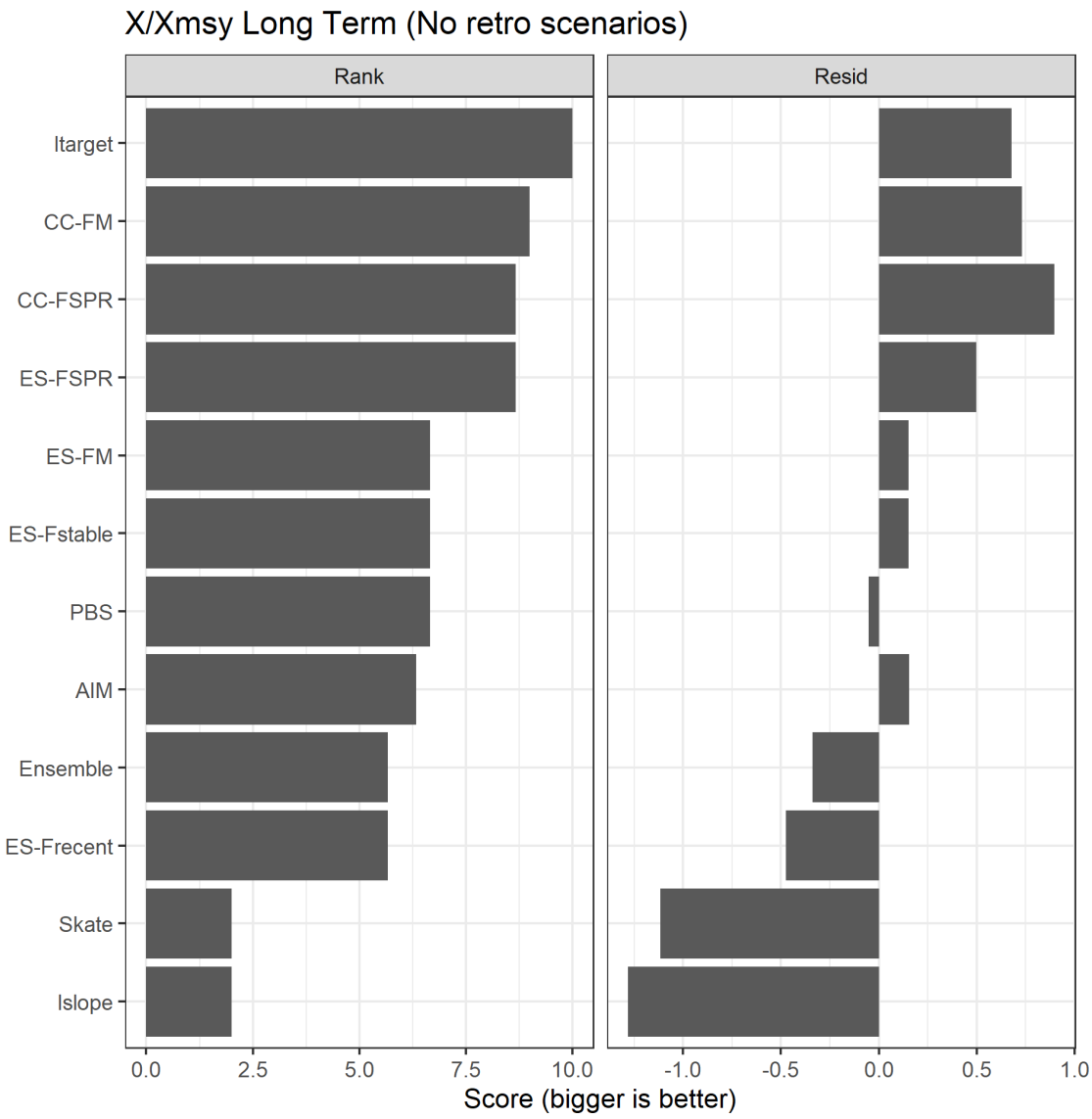


Figure A6.188. Two sets of scores, Rank and Resid, for the no retro analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in the long term.

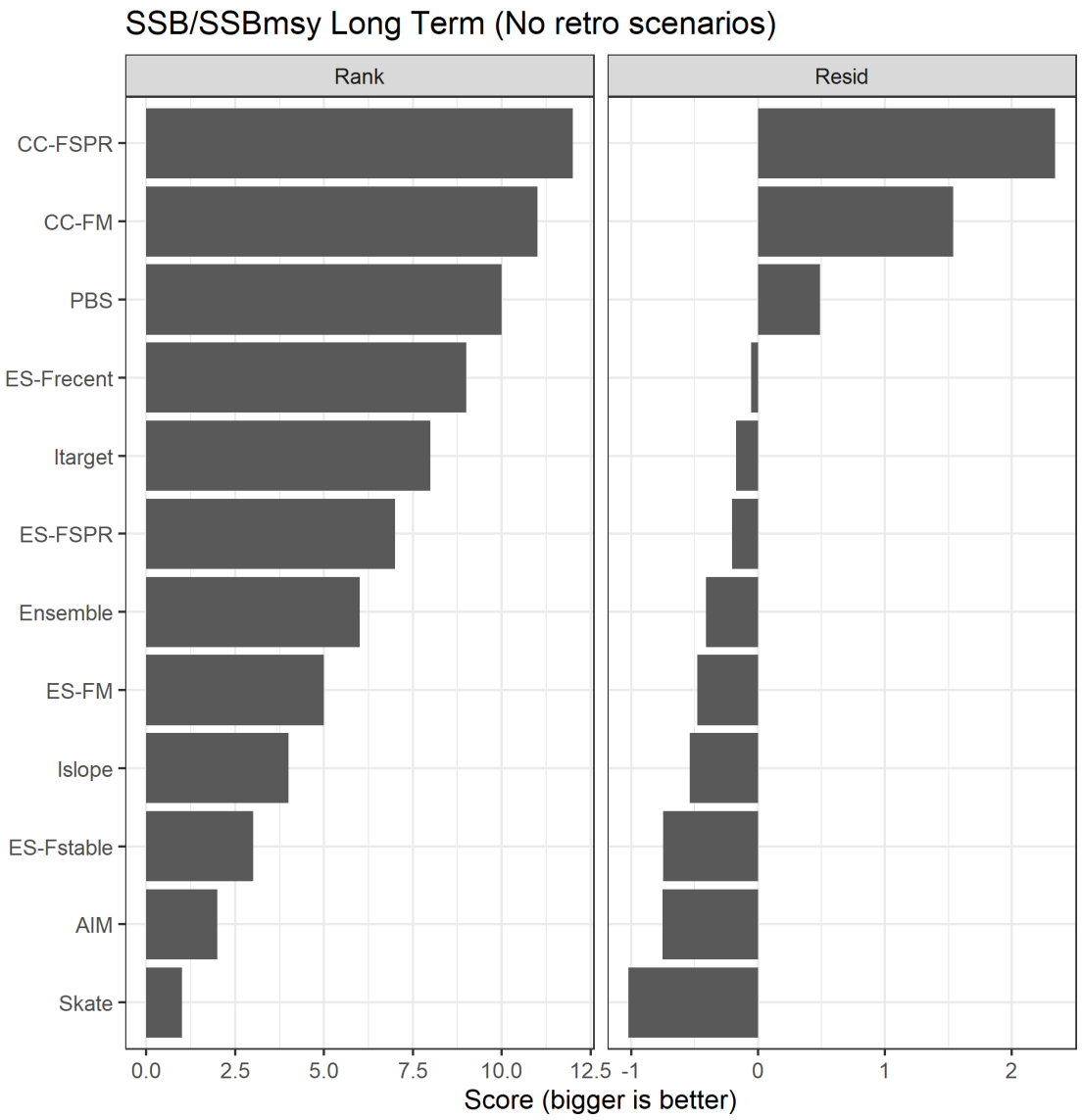


Figure A6.189. Two sets of scores, Rank and Resid, for the no retro analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in the long term.

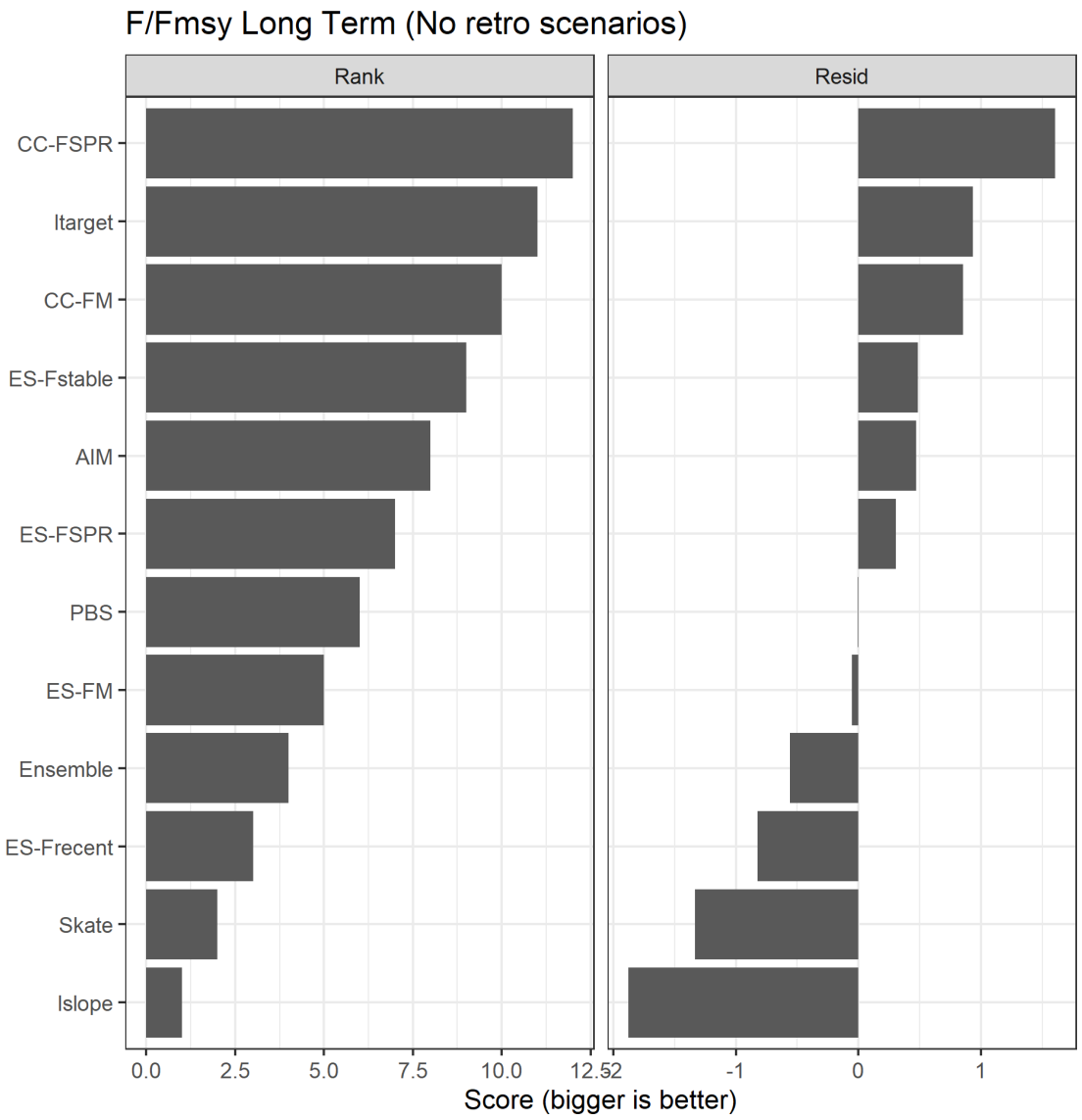


Figure A6.190. Two sets of scores, Rank and Resid, for the no retro analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in the long term.

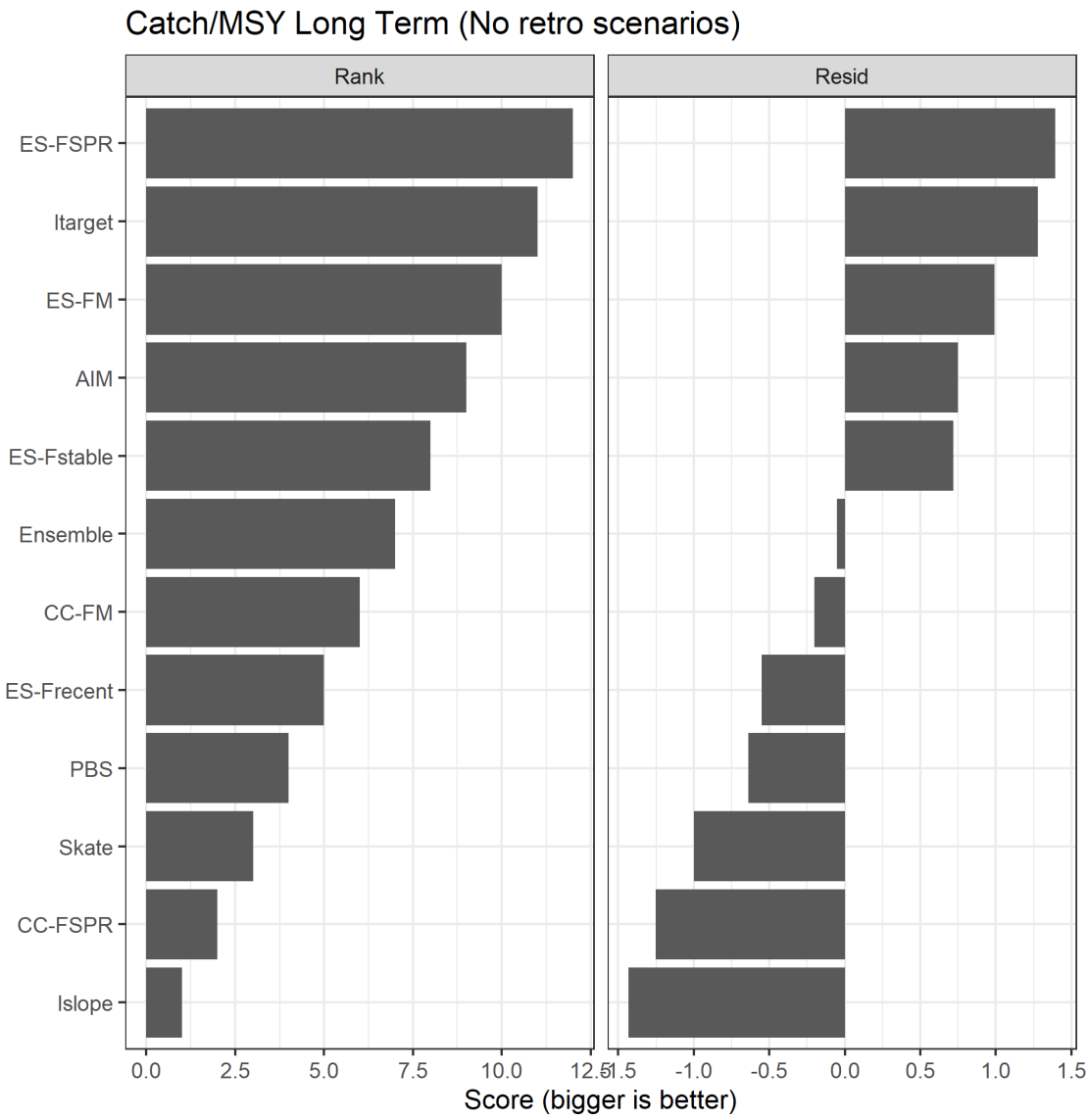


Figure A6.191. Two sets of scores, Rank and Resid, for the no retro analyses for the catch relative to maximum sustainable yield (MSY) in the long term.

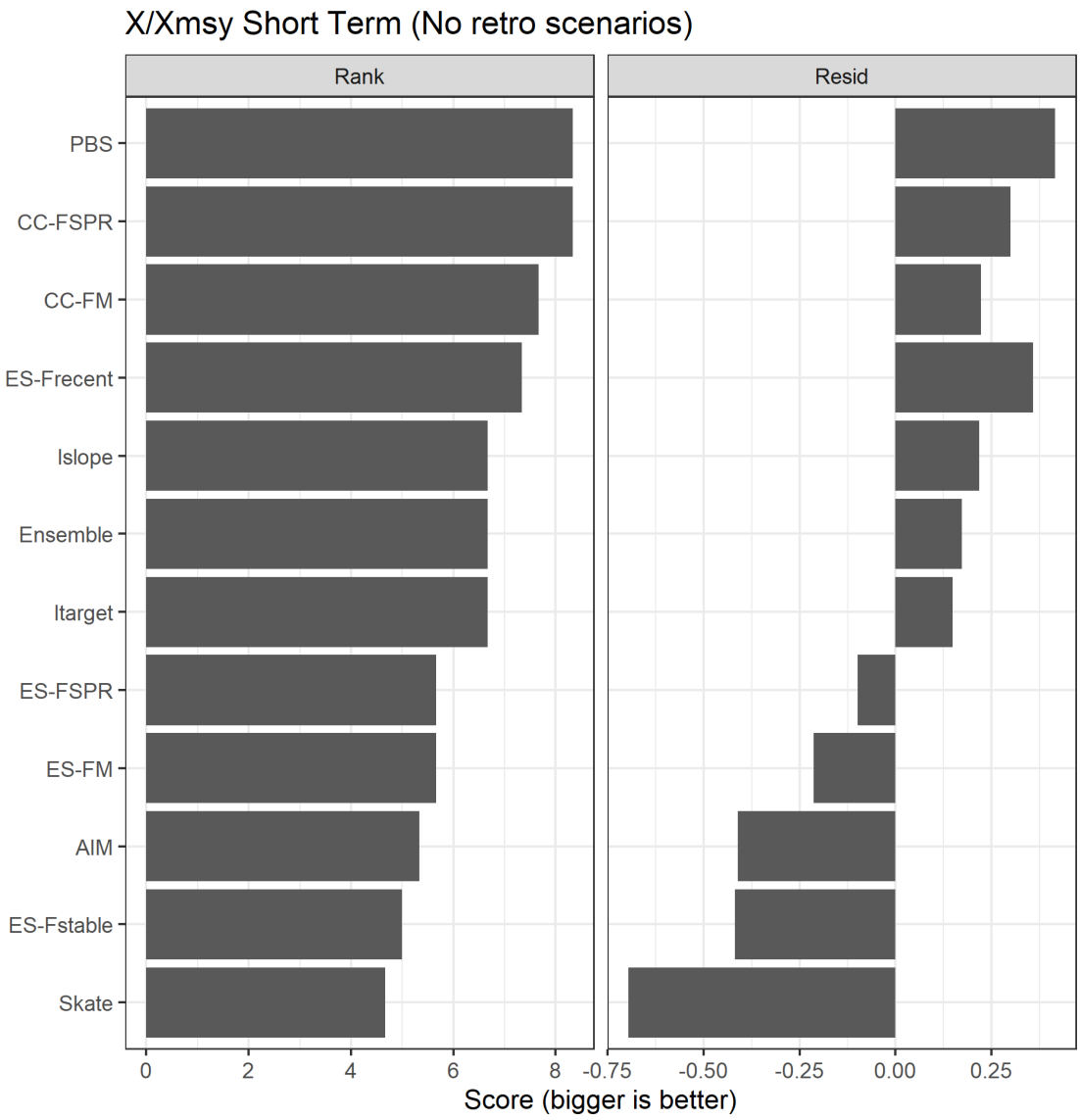


Figure A6.192. Two sets of scores, Rank and Resid, for the no retro analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in the short term.

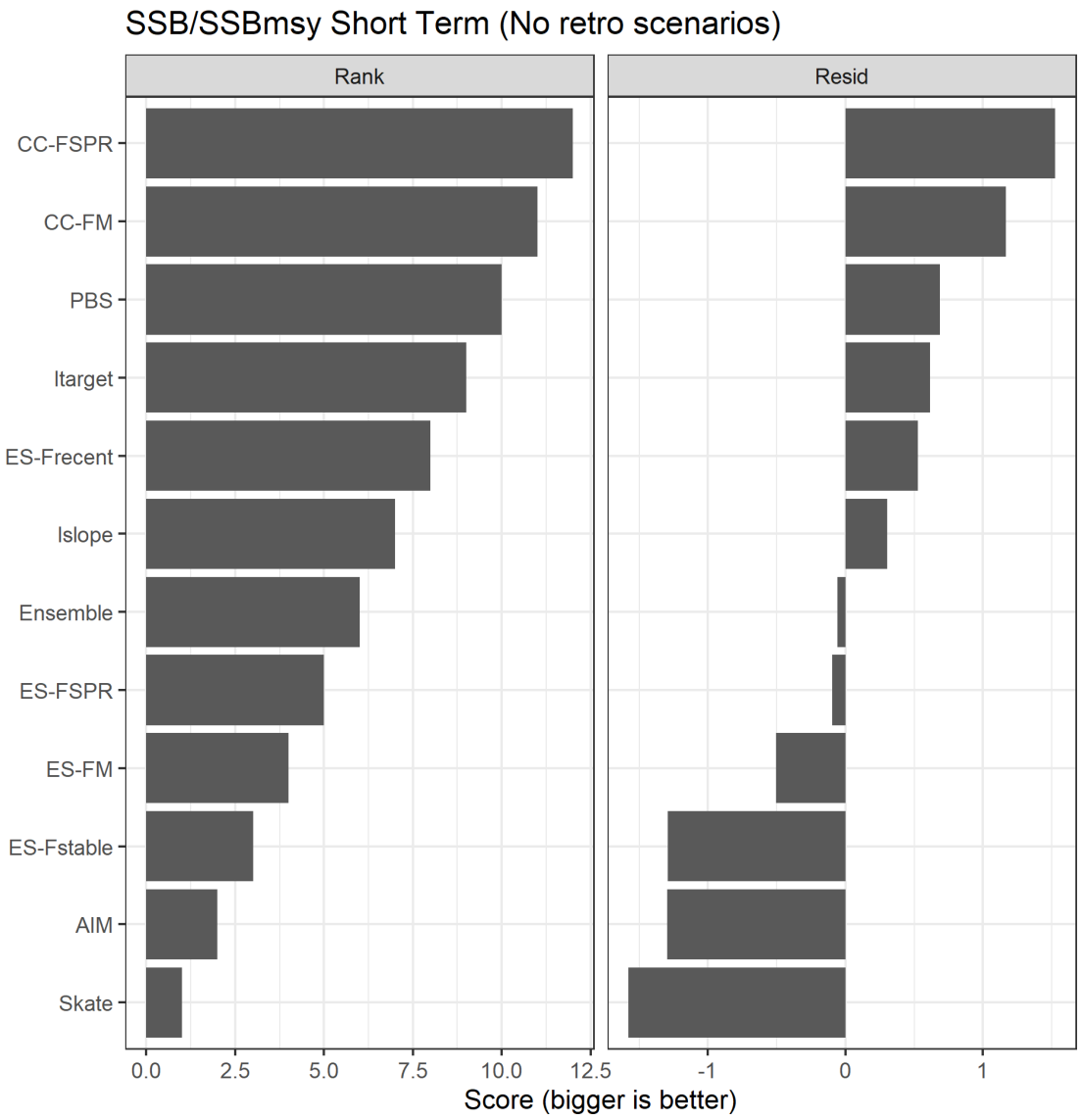


Figure A6.193. Two sets of scores, Rank and Resid, for the no retro analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in the short term.

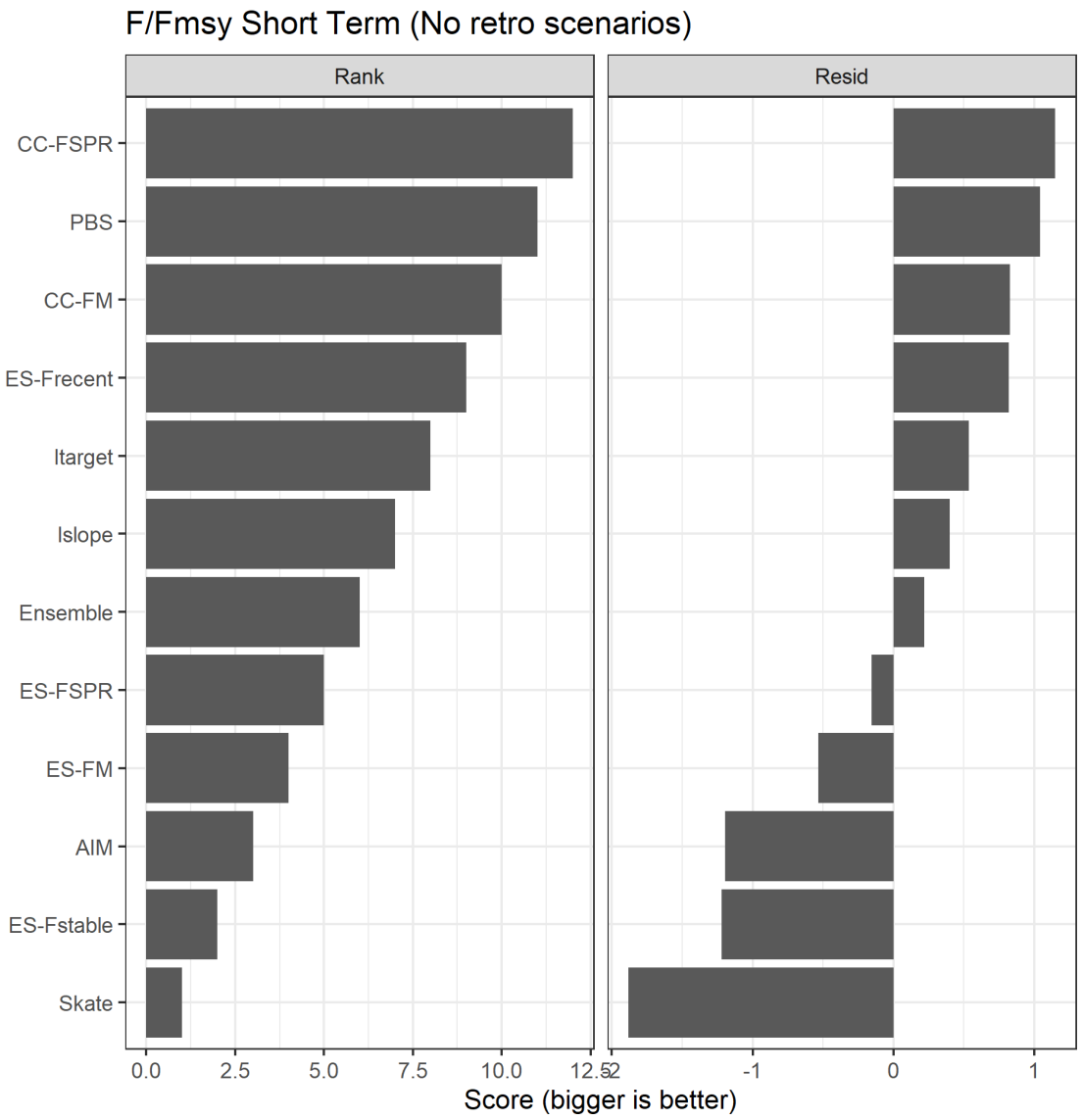


Figure A6.194. Two sets of scores, Rank and Resid, for the no retro analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in the short term.

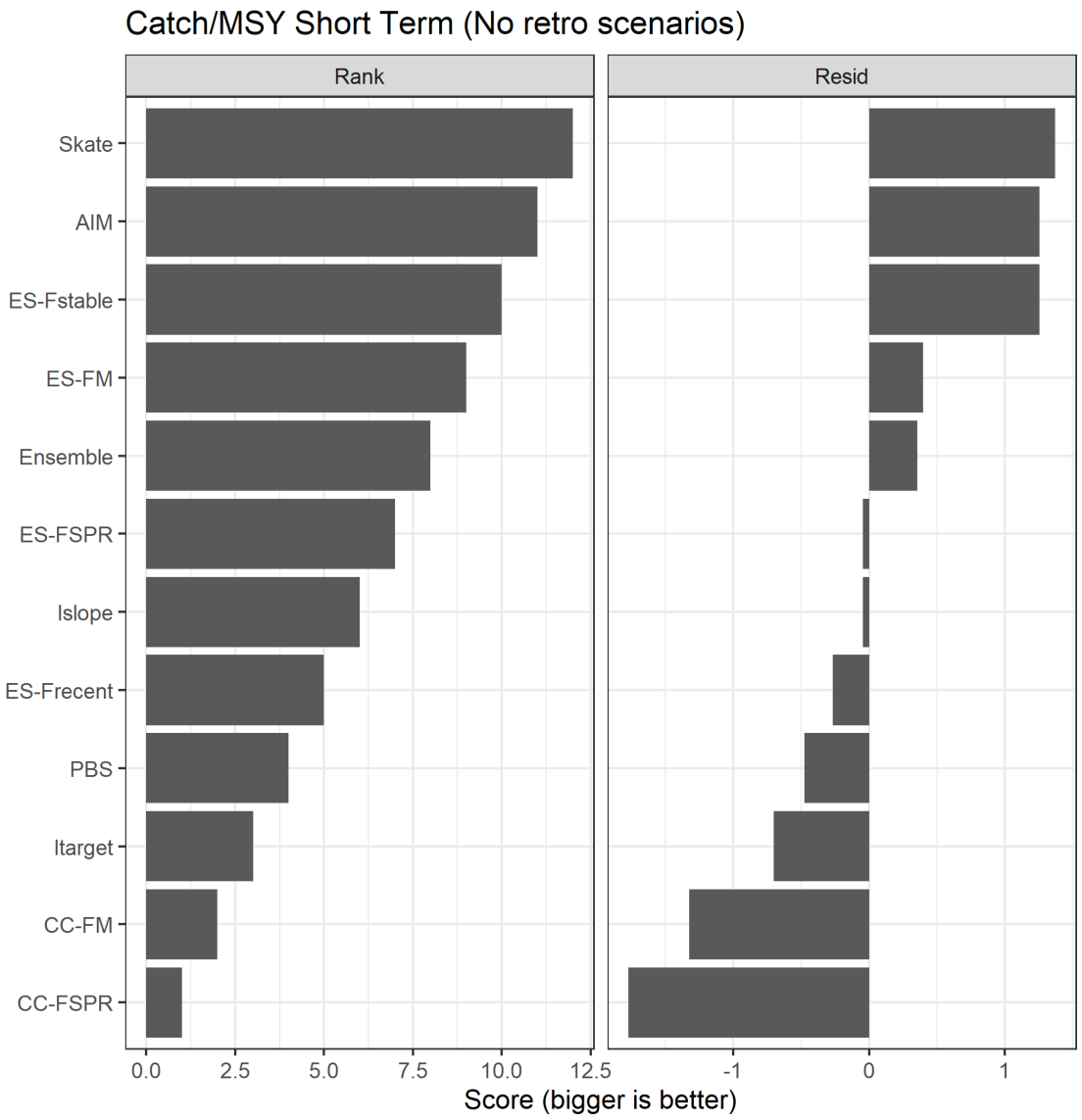


Figure A6.195. Two sets of scores, Rank and Resid, for the no retro analyses for the catch relative to maximum sustainable yield (MSY) in the short term.

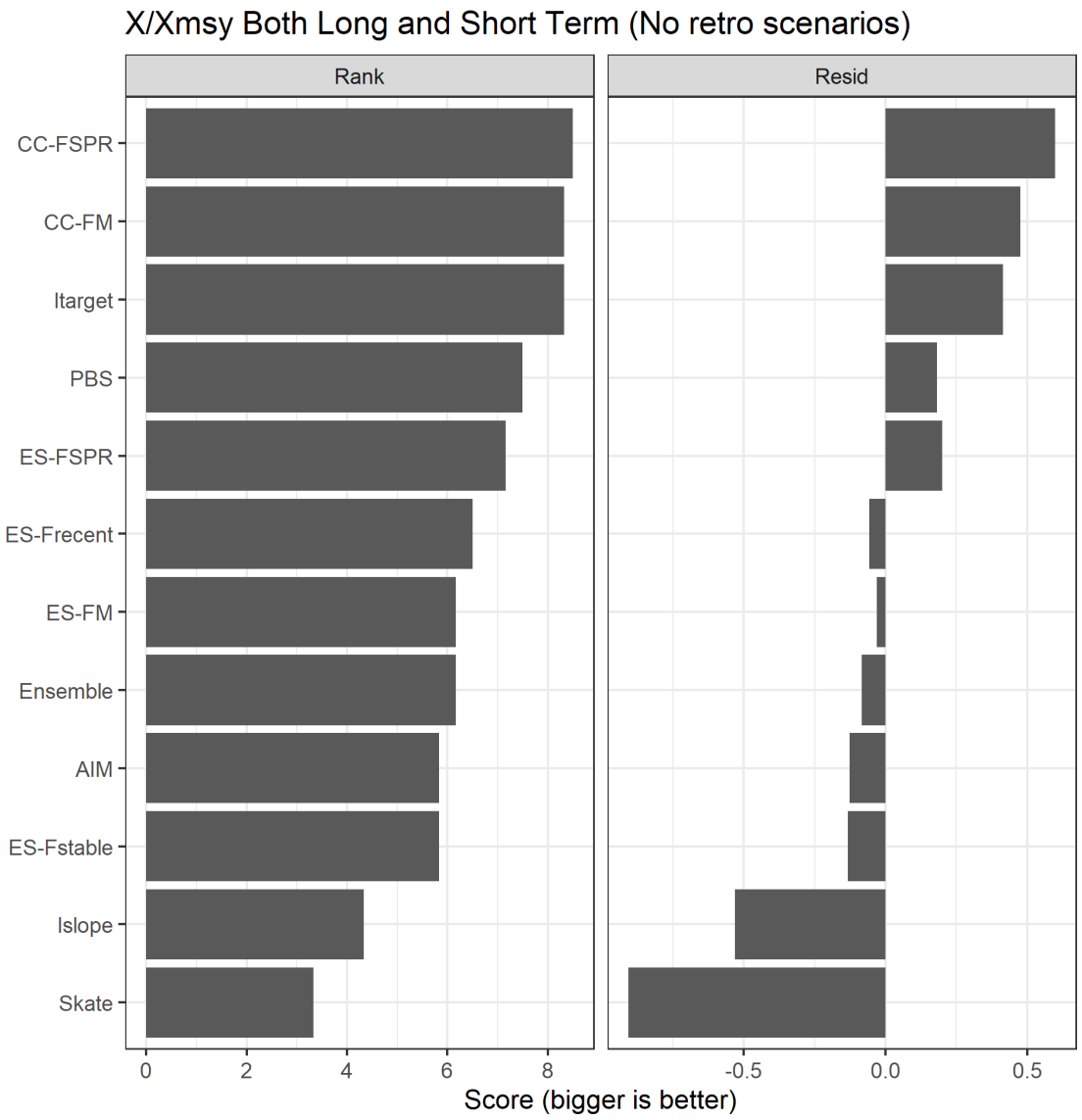


Figure A6.196. Two sets of scores, Rank and Resid, for the no retro analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in both the long and short term.

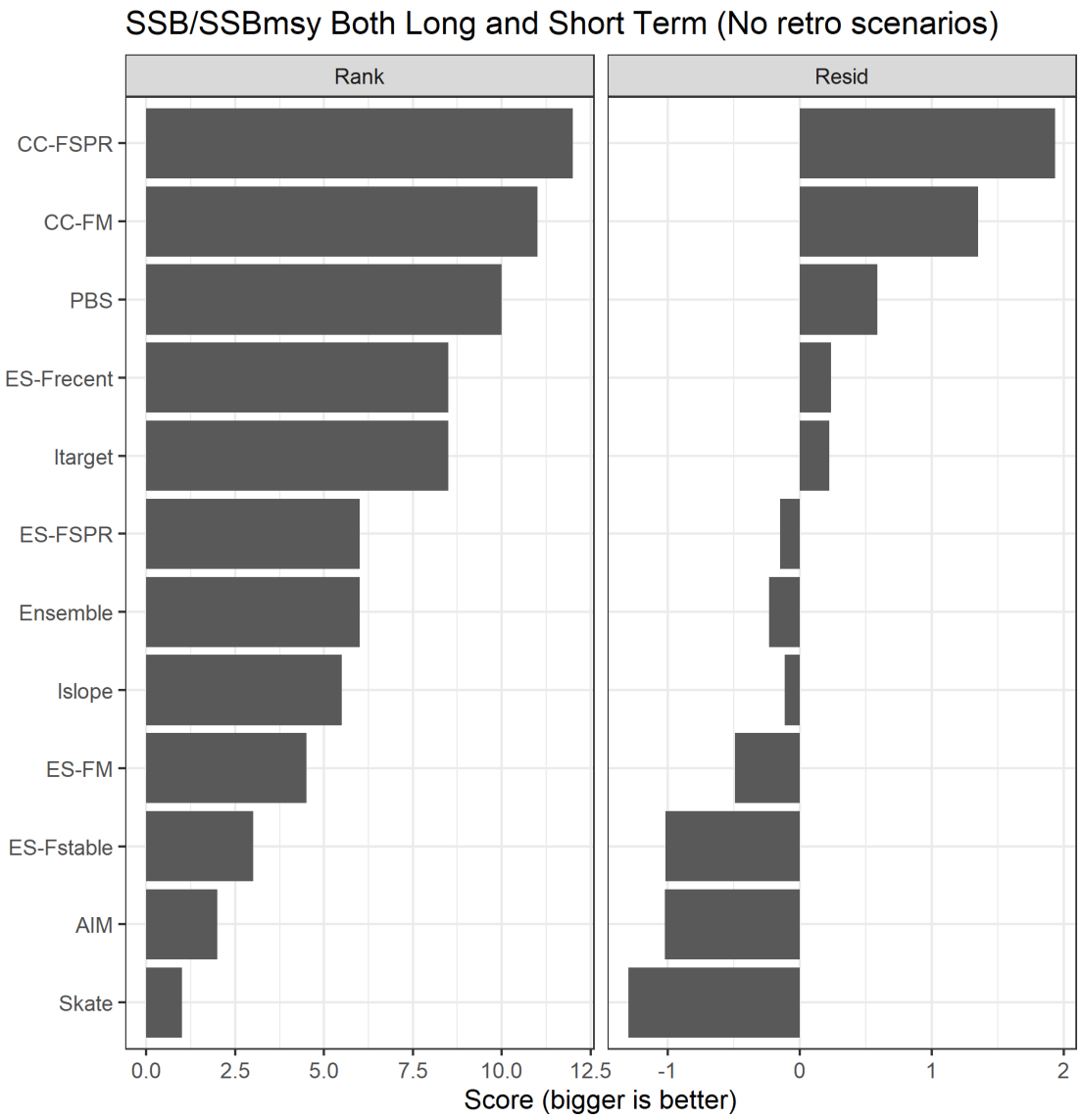


Figure A6.197. Two sets of scores, Rank and Resid, for the no retro analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in both the long and short term.

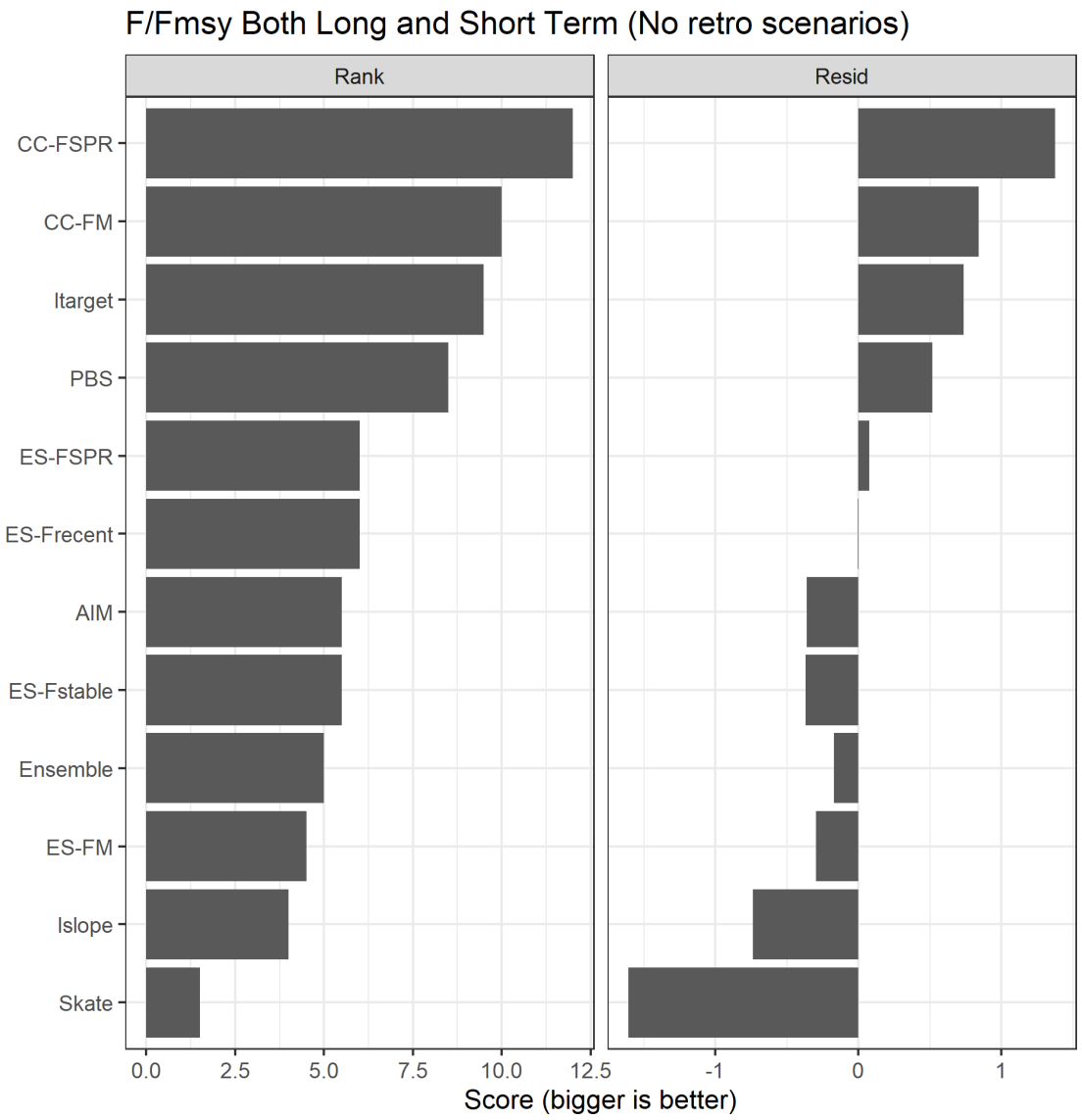


Figure A6.198. Two sets of scores, Rank and Resid, for the no retro analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in both the long and short term.

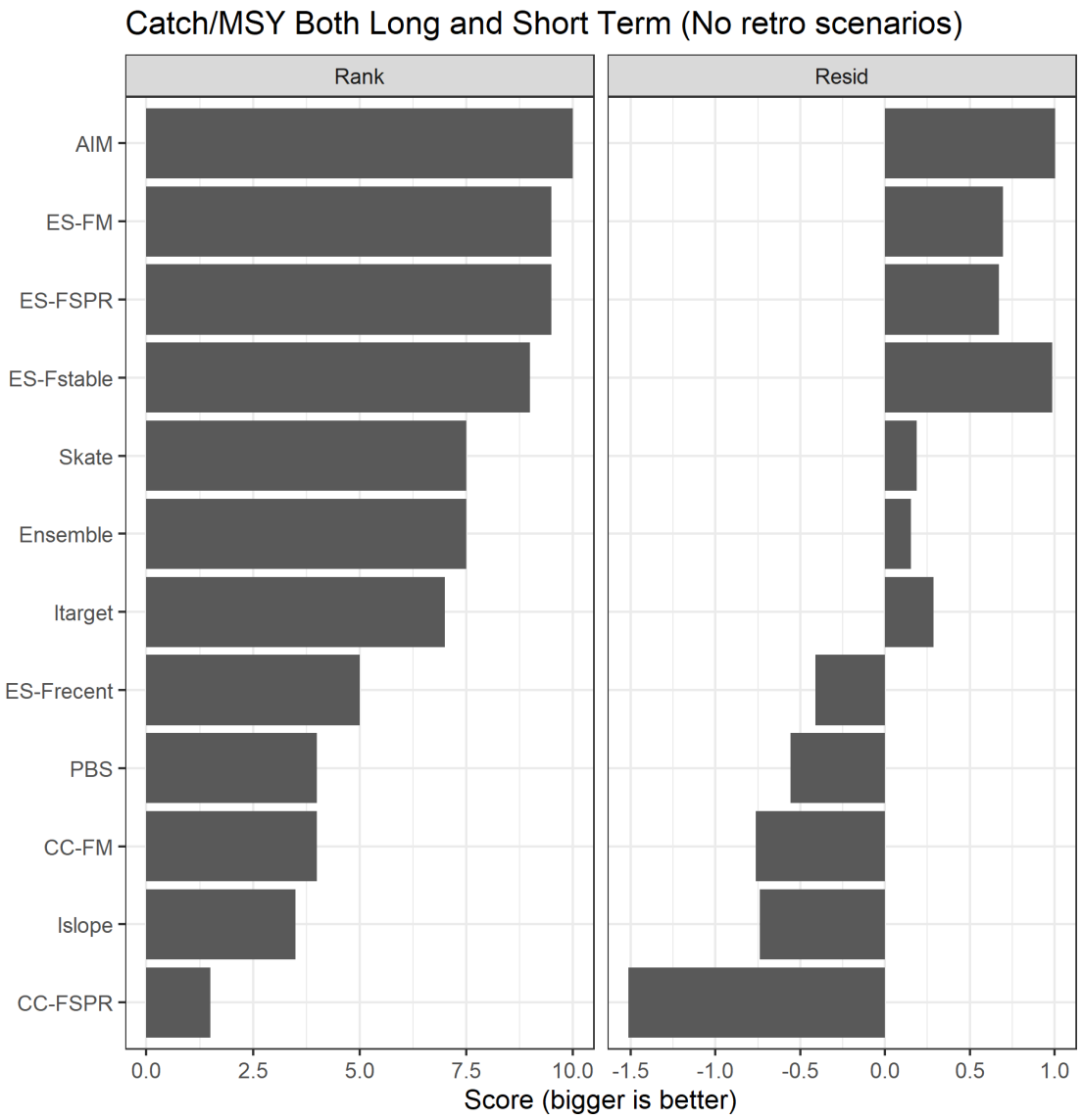


Figure A6.199. Two sets of scores, Rank and Resid, for the no retro analyses for the catch relative to maximum sustainable yield (MSY) in both the long and short term.

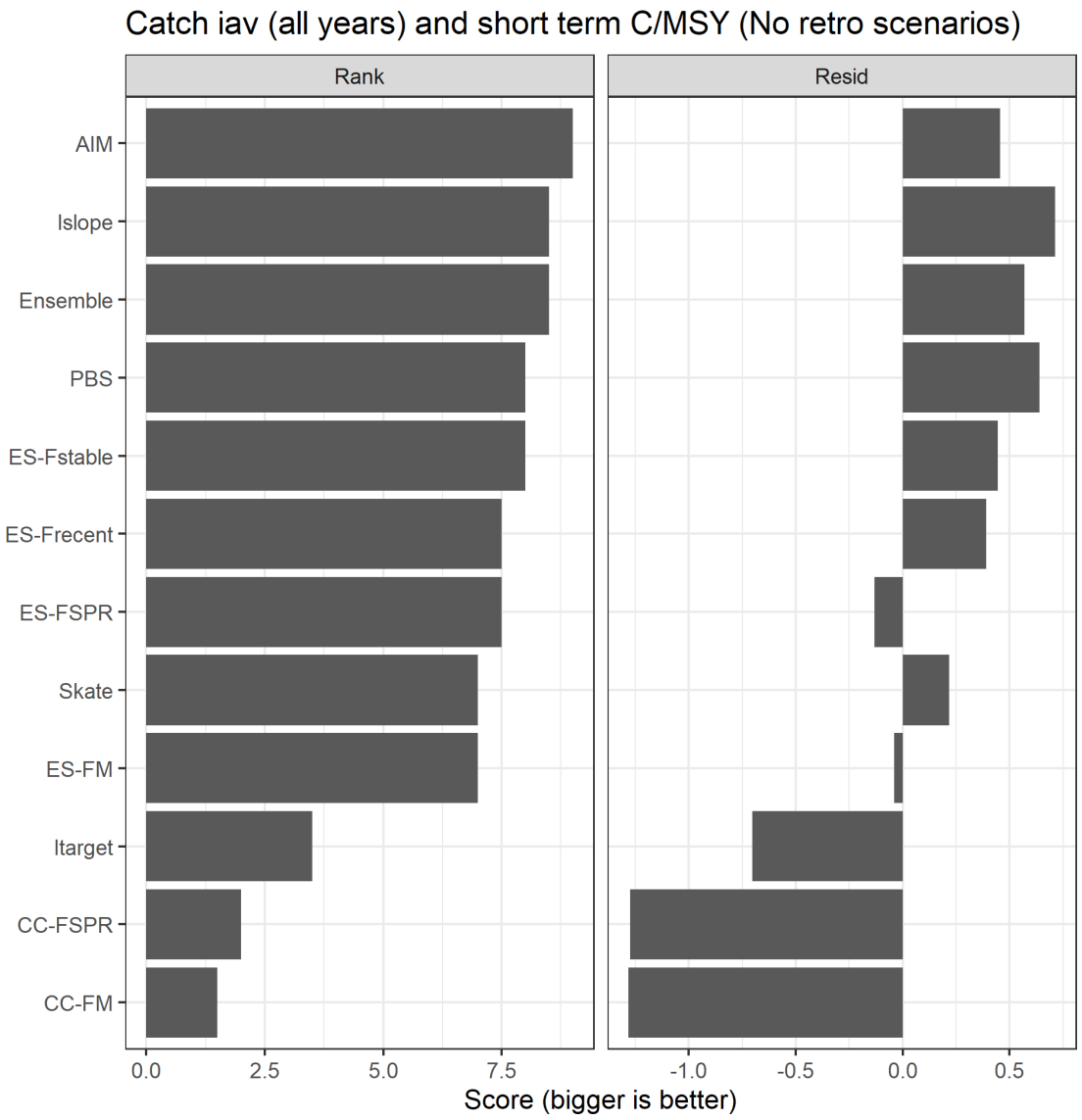


Figure A6.200. Two sets of scores, Rank and Resid, for the no retro analyses for the 2 metrics of interannual variability in catch over the entire feedback period and the short term mean catch/maximum sustainable yield (MSY).

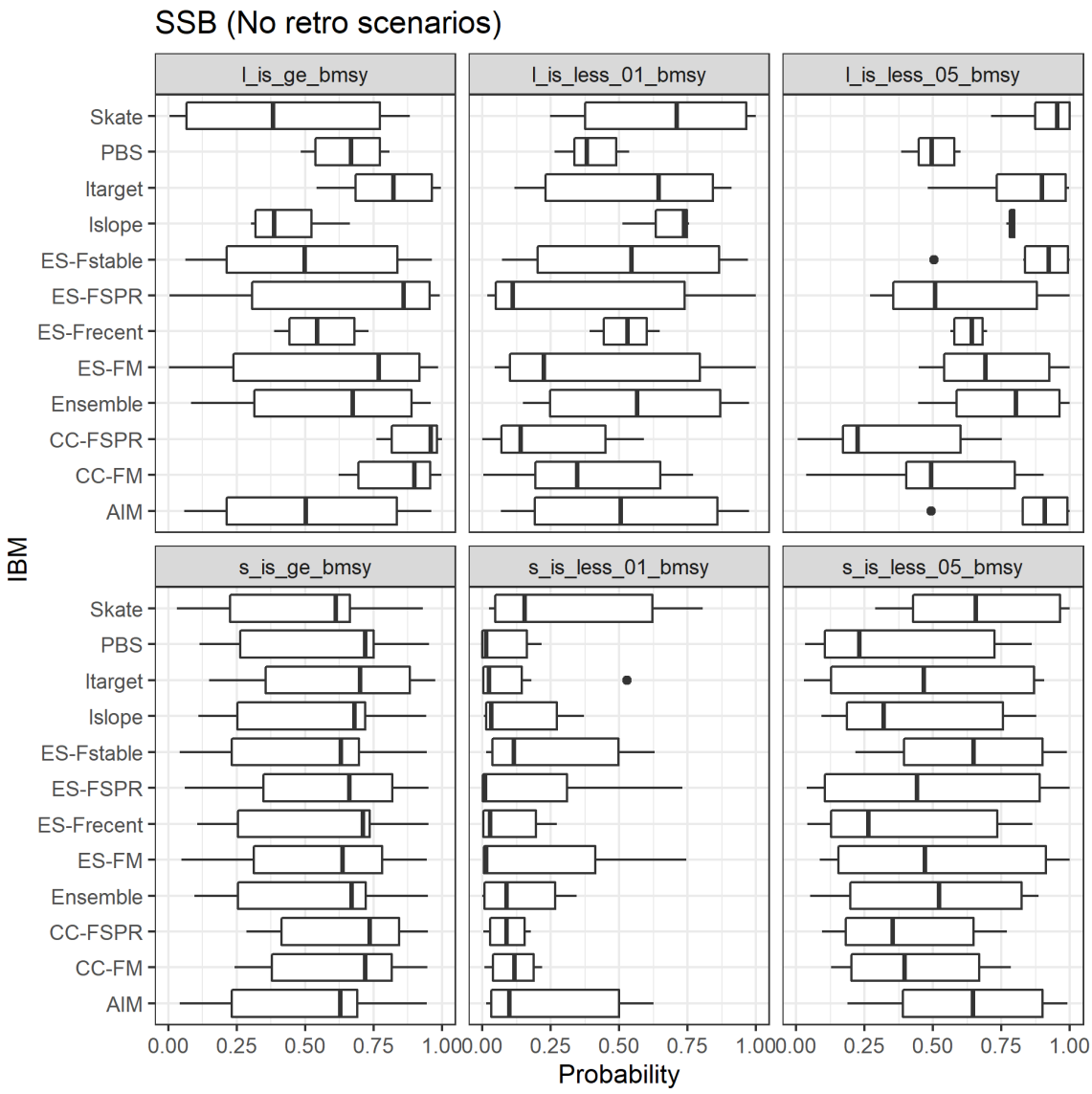


Figure A6.201. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the no retro analyses by index-based method (IBM).

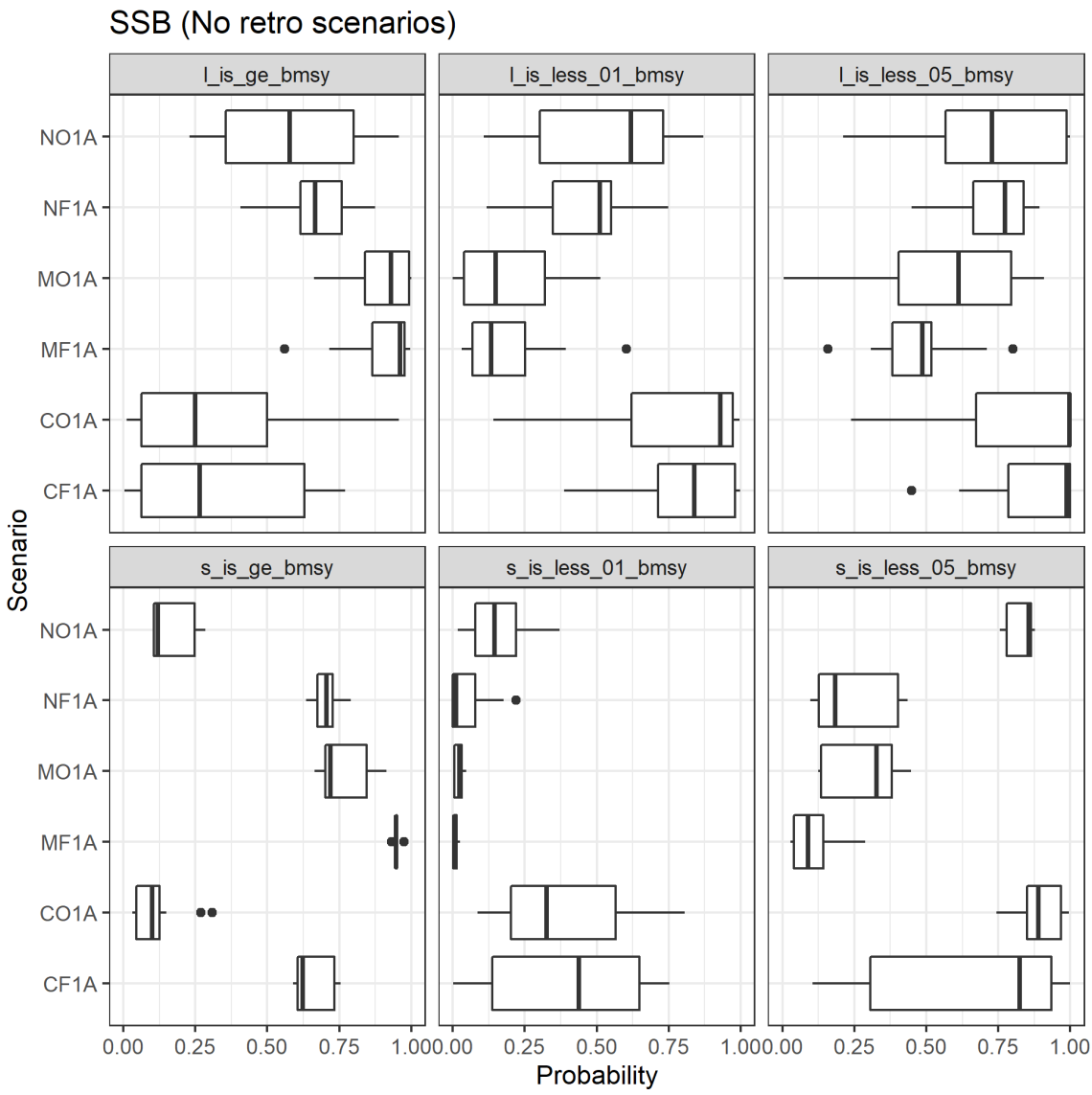


Figure A6.202. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the no retro analyses by scenario.

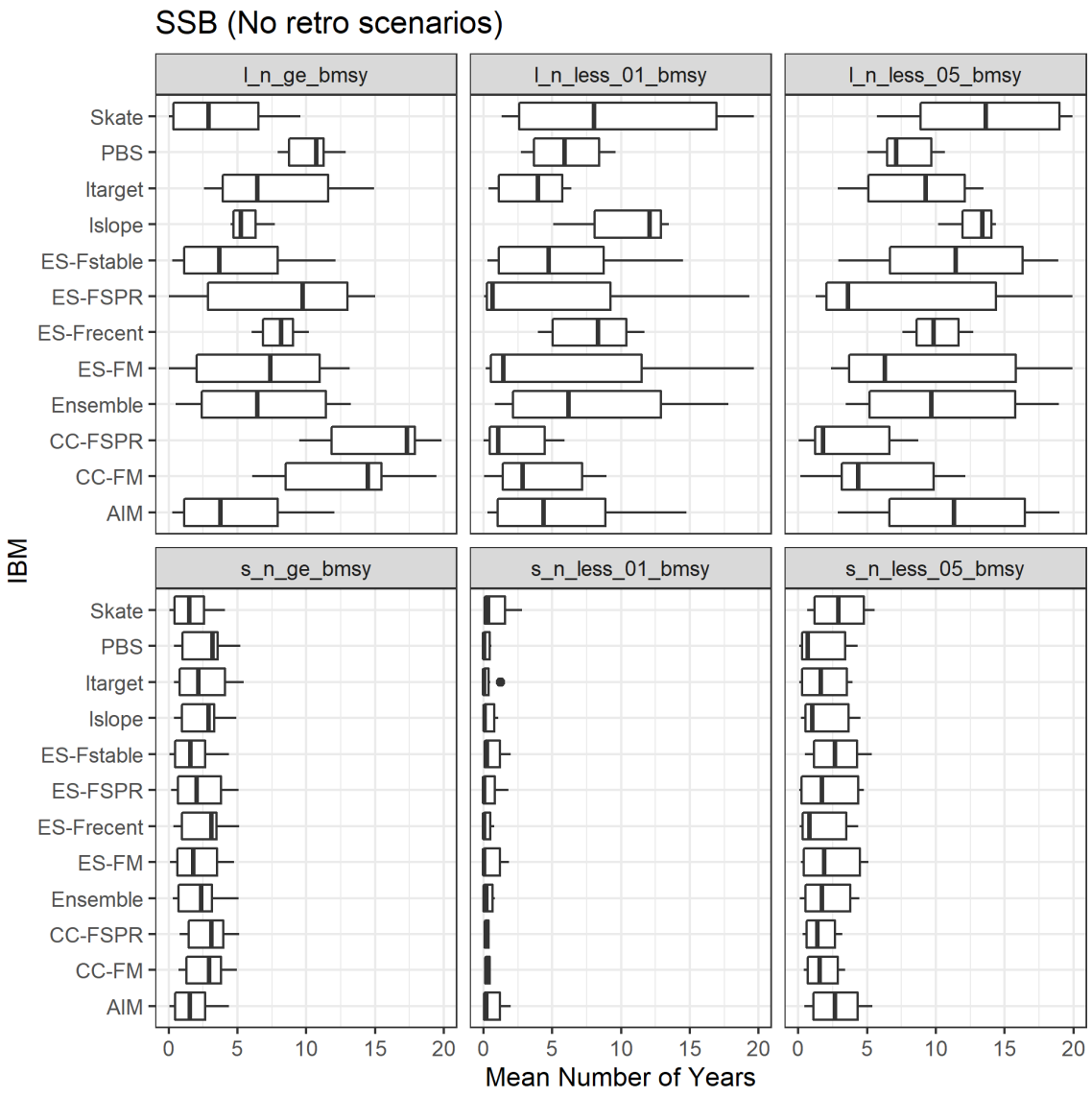


Figure A6.203. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the no retro analyses by index-based method (IBM).

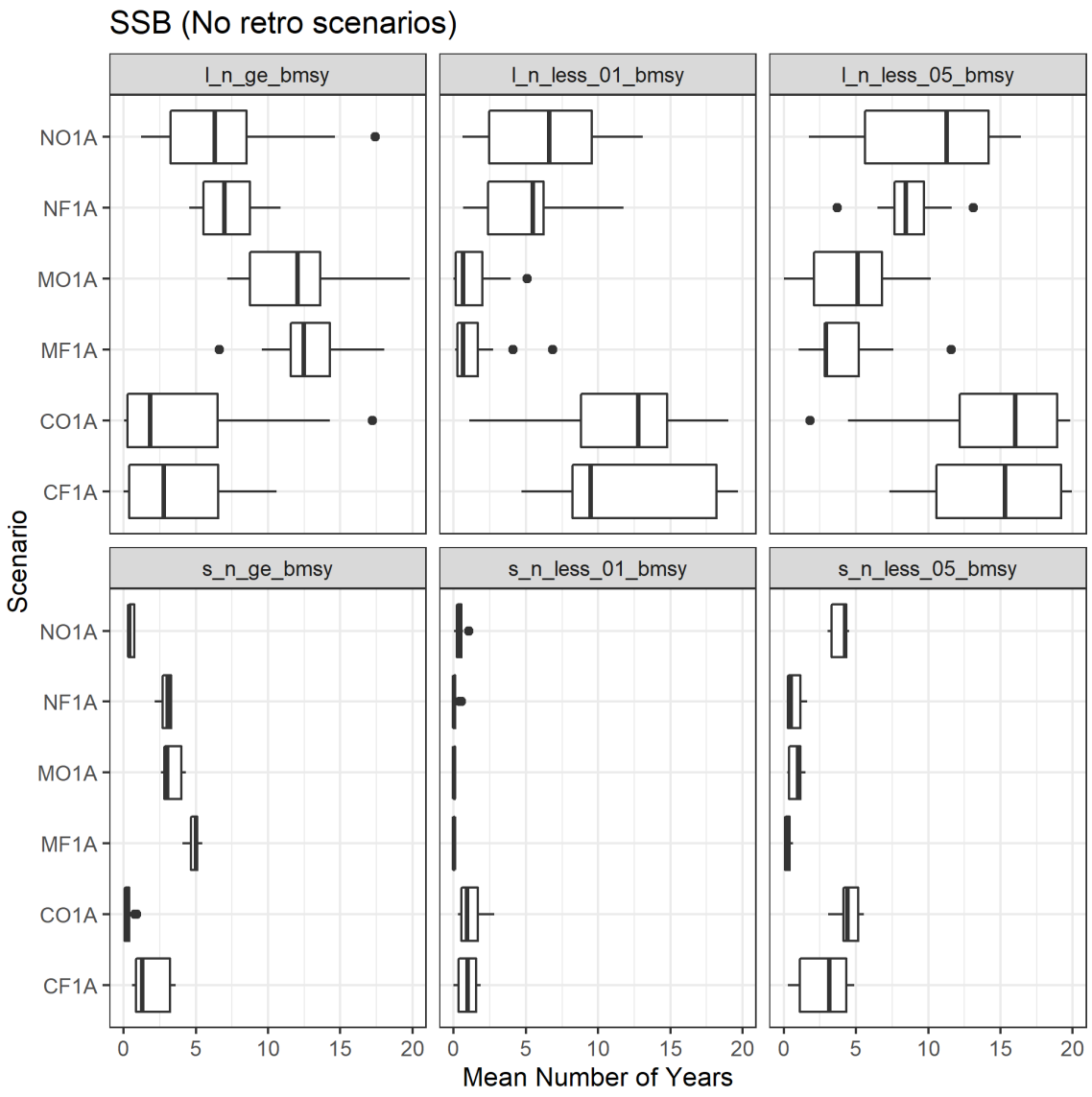


Figure A6.204. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the no retro analyses by scenario.

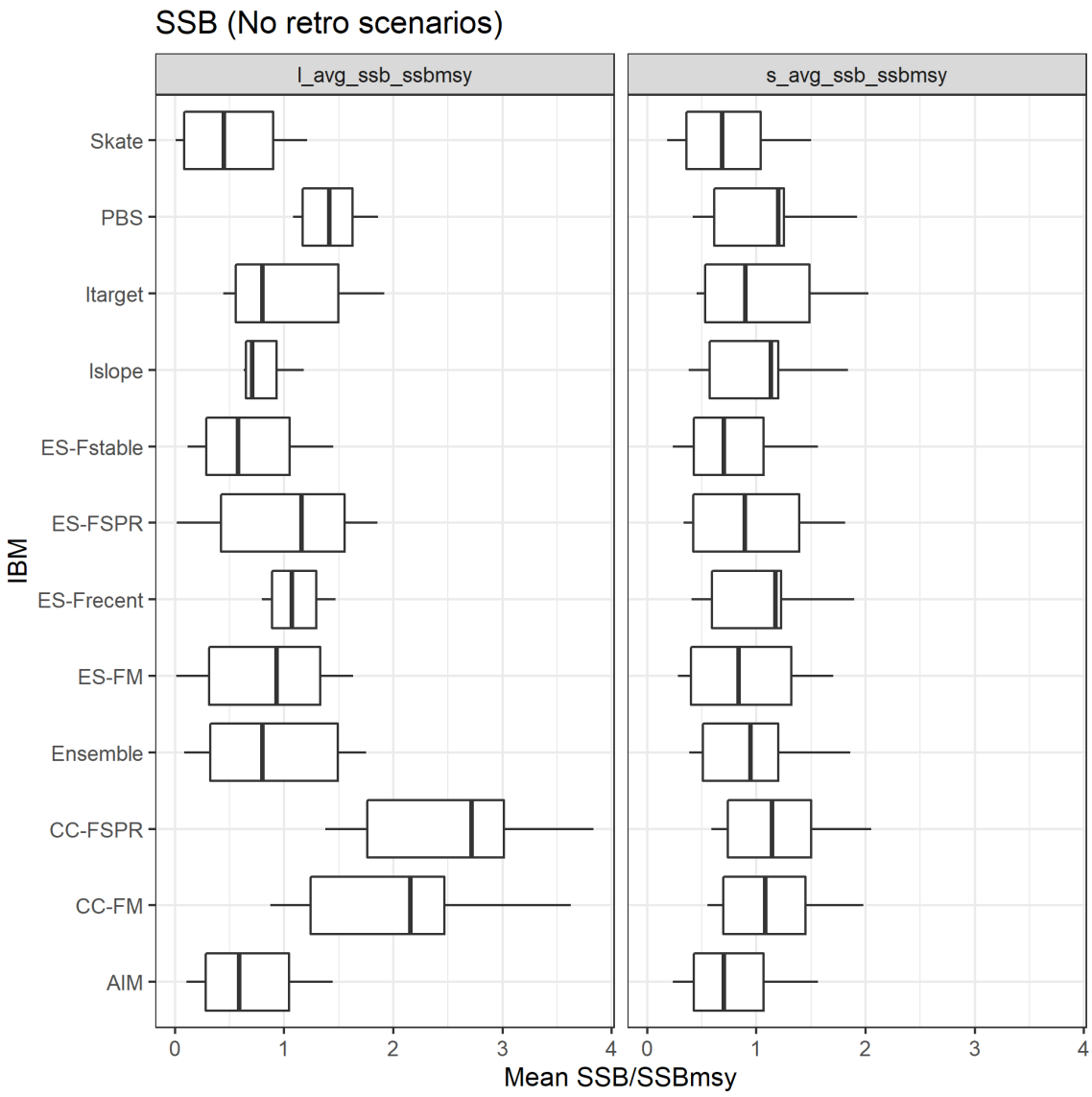


Figure A6.205. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the no retro analyses by index-based method (IBM).

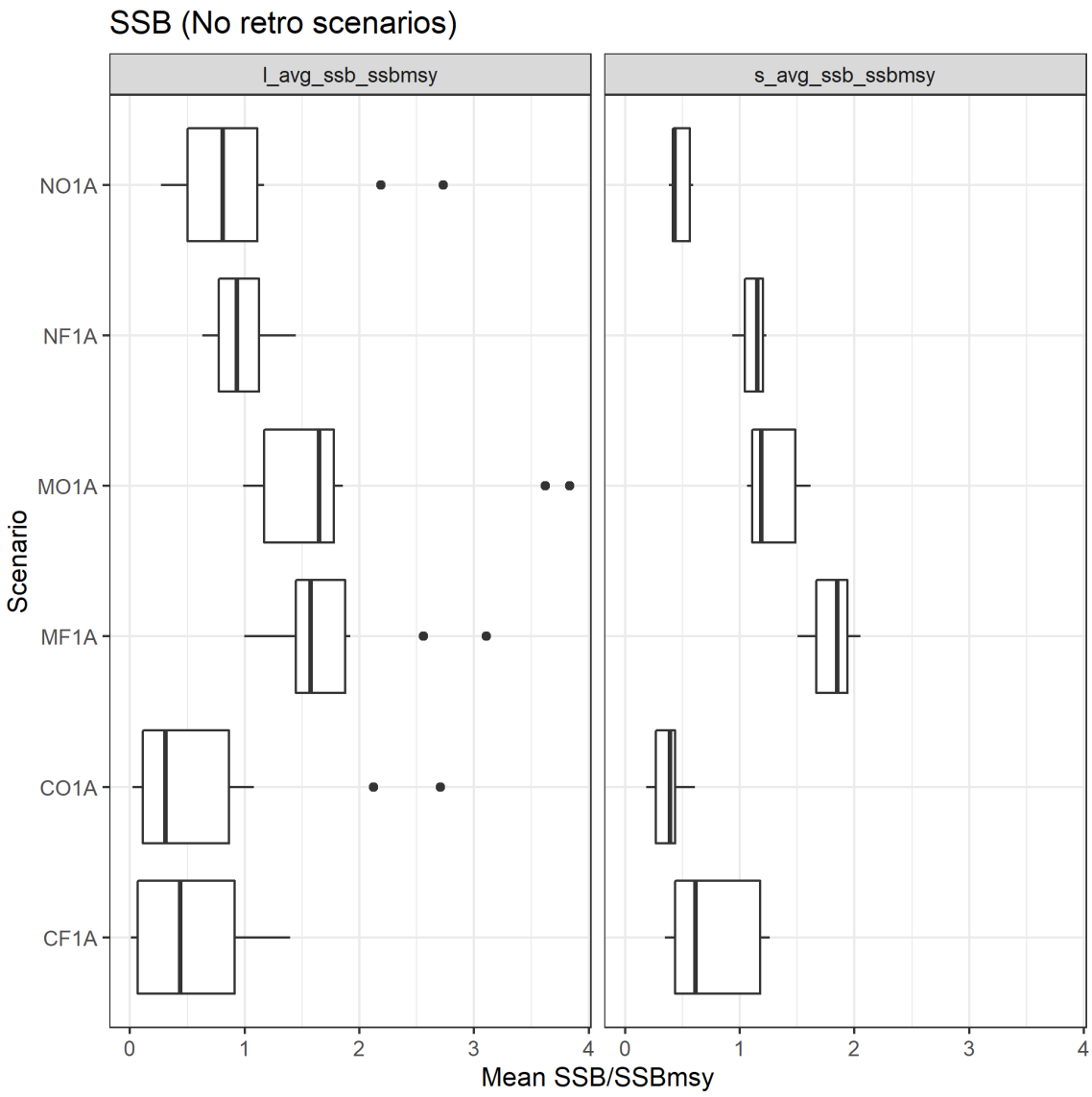


Figure A6.206. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the no retro analyses by scenario.

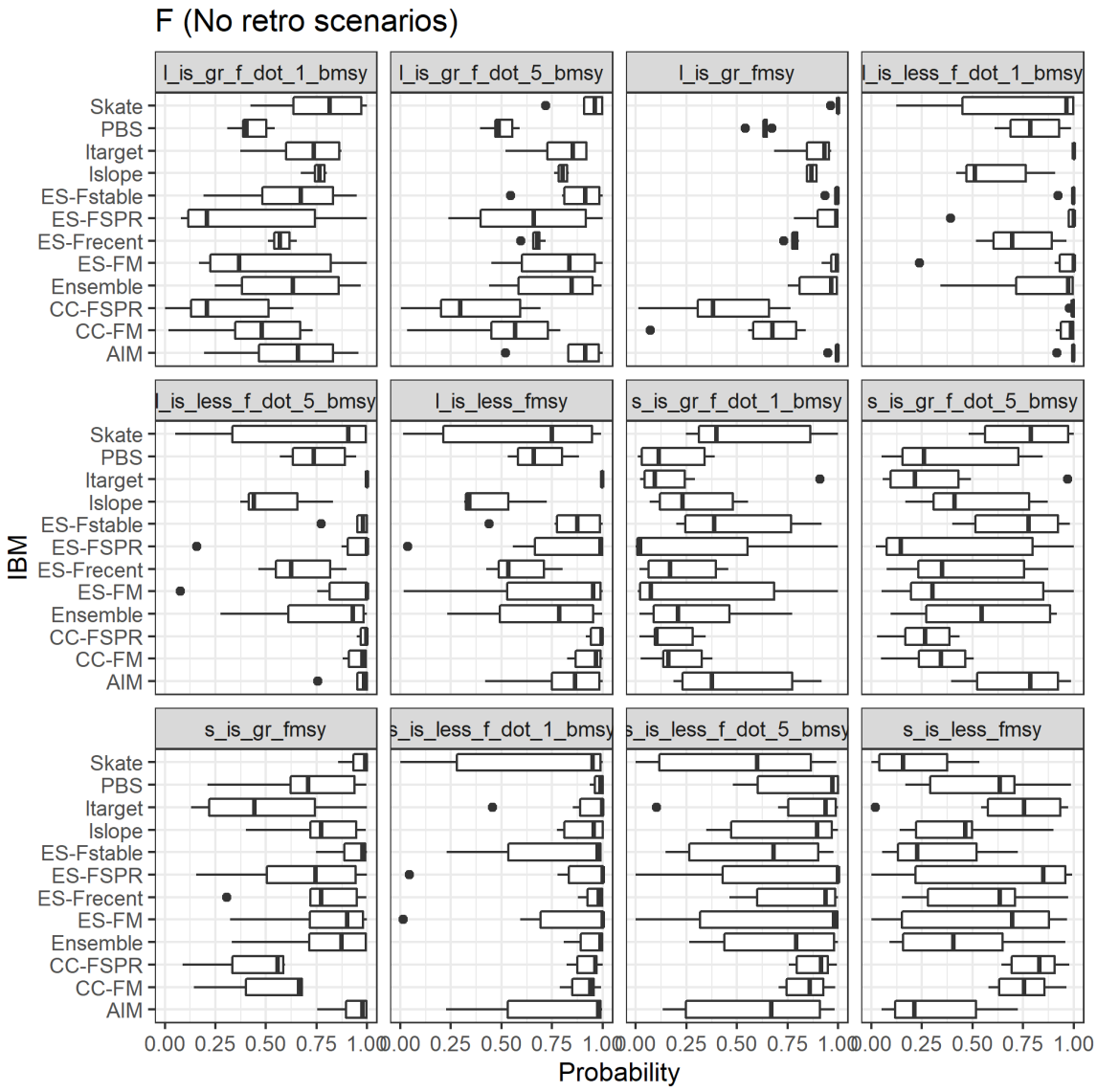


Figure A6.207. Boxplot of the mean values for the fishing mortality rate (F) metrics in the no retro analyses by index-based method (IBM).

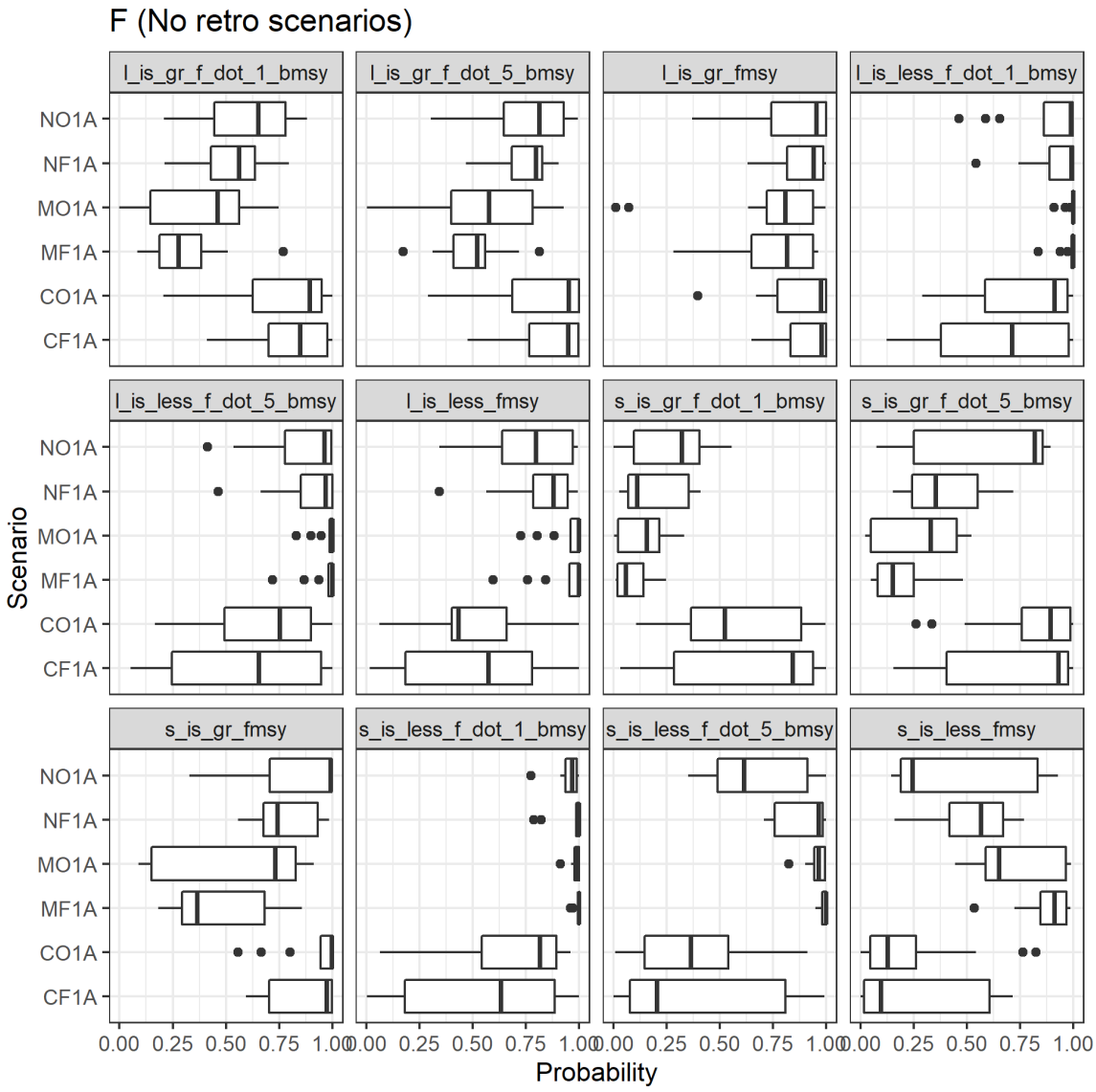


Figure A6.208. Boxplot of the mean values for the fishing mortality rate (F) metrics in the no retro analyses by scenario.

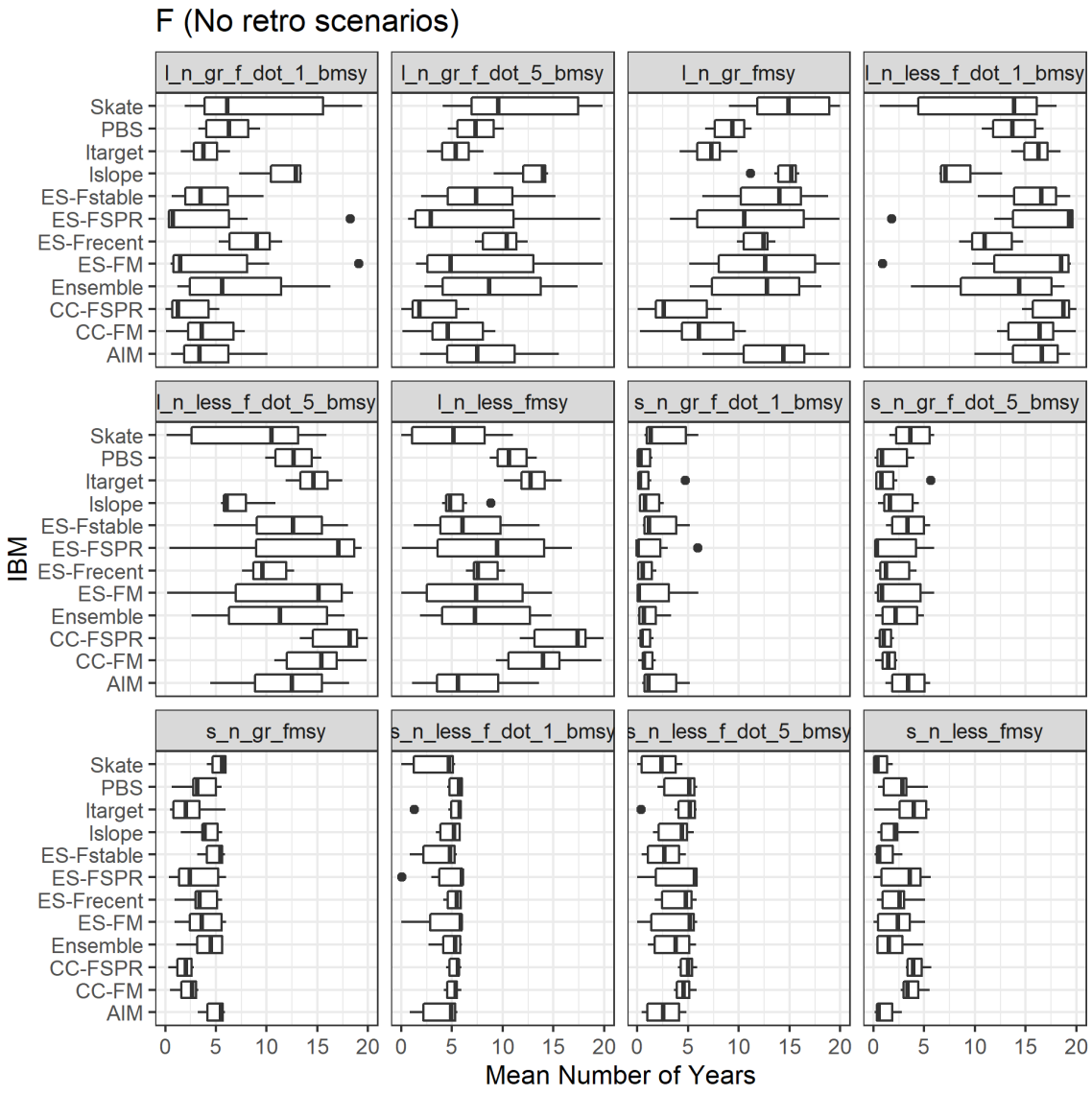


Figure A6.209. Boxplot of the mean values for the fishing mortality rate (F) metrics in the no retro analyses by index-based method (IBM).

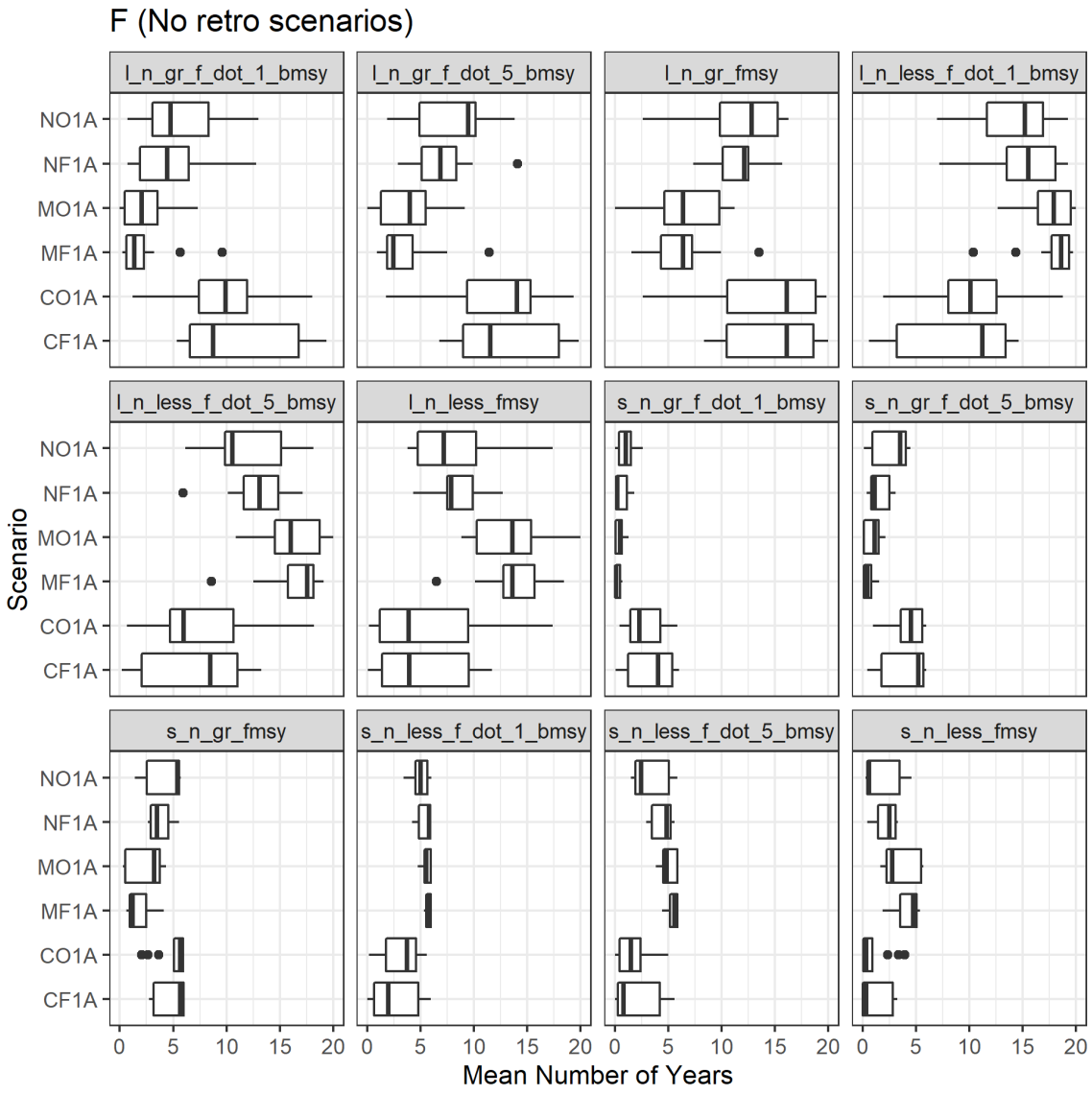


Figure A6.210. Boxplot of the mean values for the fishing mortality rate (F) metrics in the no retro analyses by scenario.

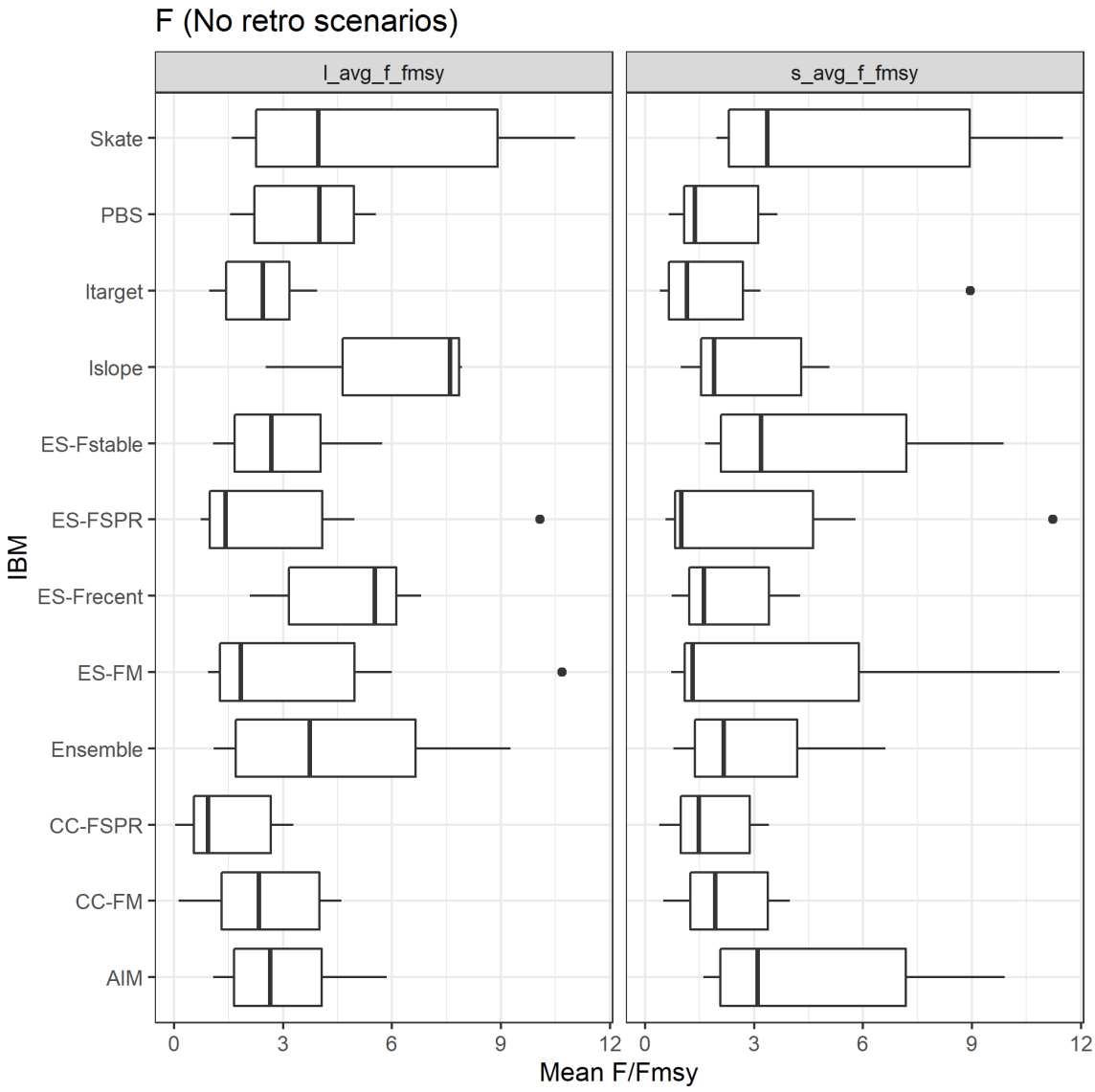


Figure A6.211. Boxplot of the mean values for the fishing mortality rate (F) metrics in the no retro analyses by index-based method (IBM).

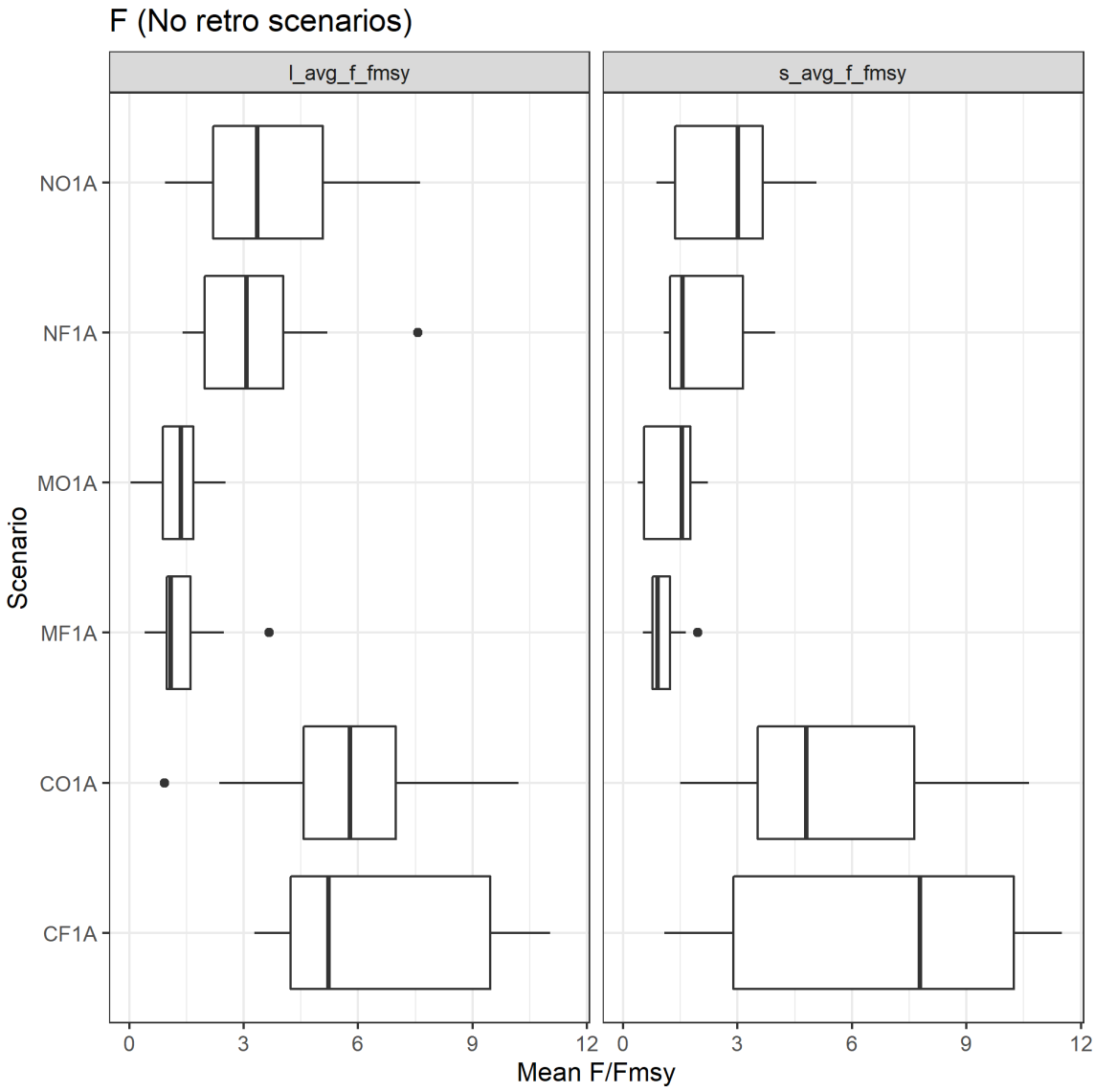


Figure A6.212. Boxplot of the mean values for the fishing mortality rate (F) metrics in the no retro analyses by scenario.

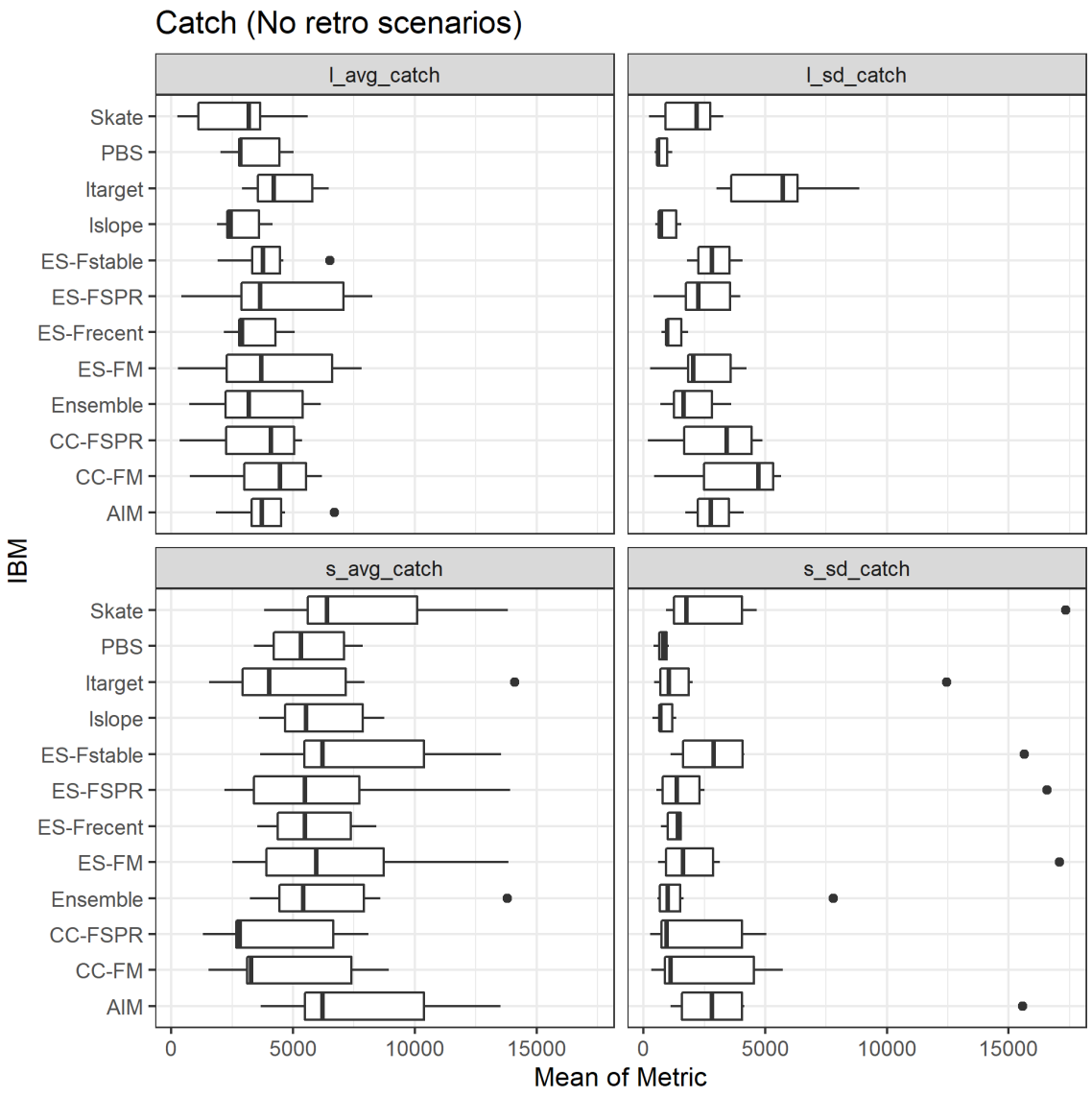


Figure A6.213. Boxplot of the mean values for the catch metrics in the no retro analyses by index-based method (IBM).

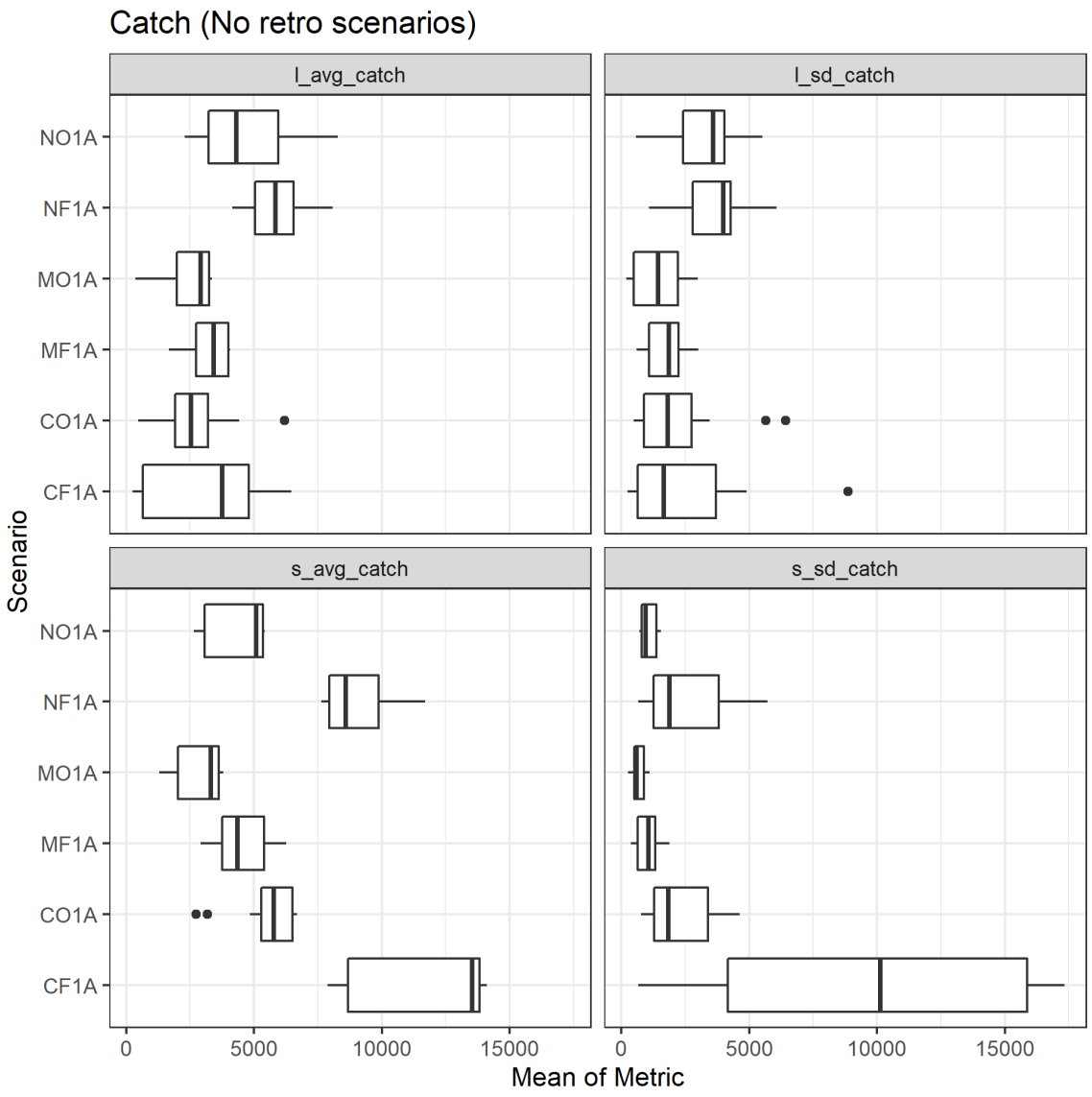


Figure A6.214. Boxplot of the mean values for the catch metrics in the no retro analyses by scenario.

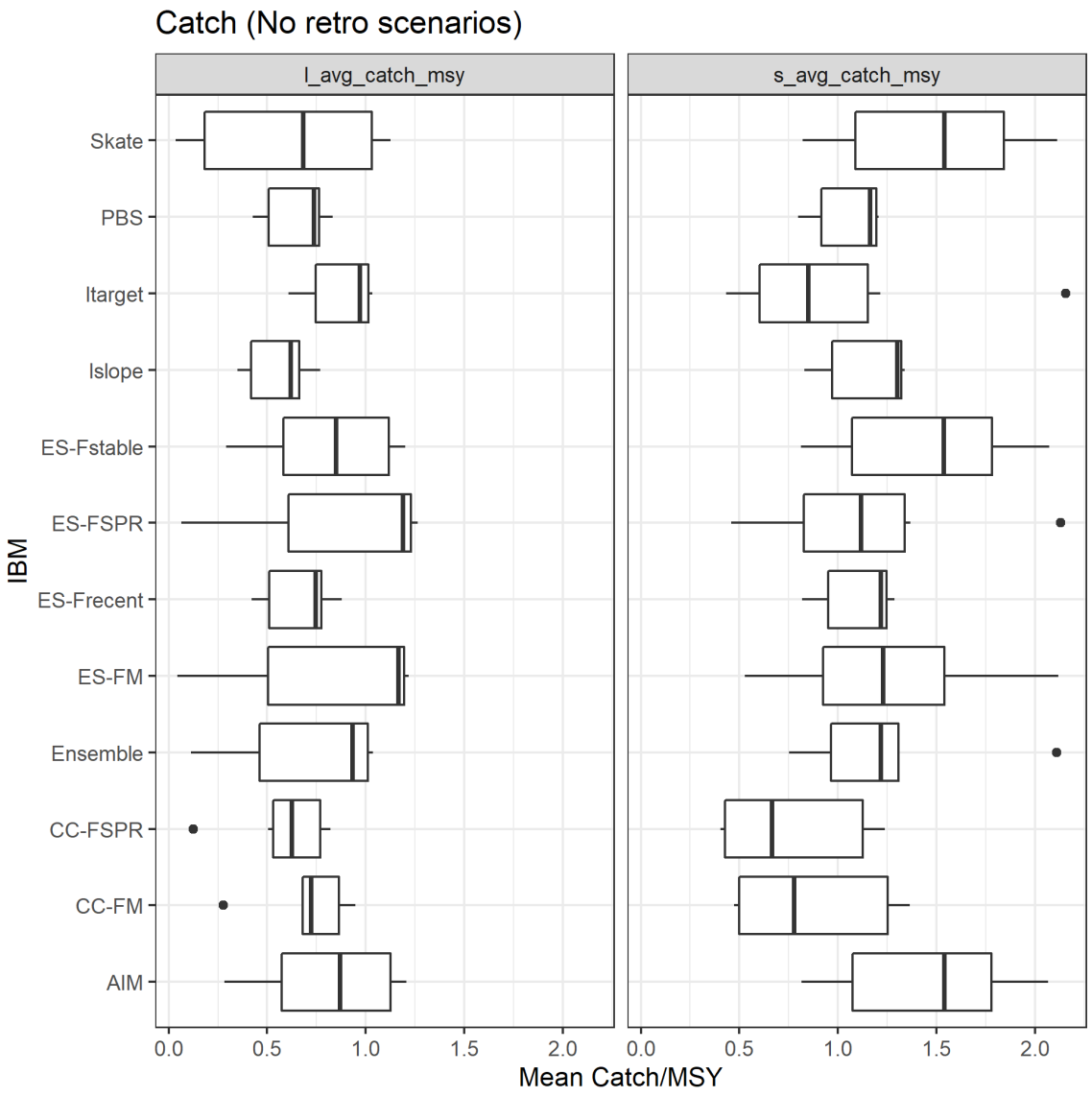


Figure A6.215. Boxplot of the mean values for the catch metrics in the no retro analyses by index-based method (IBM).

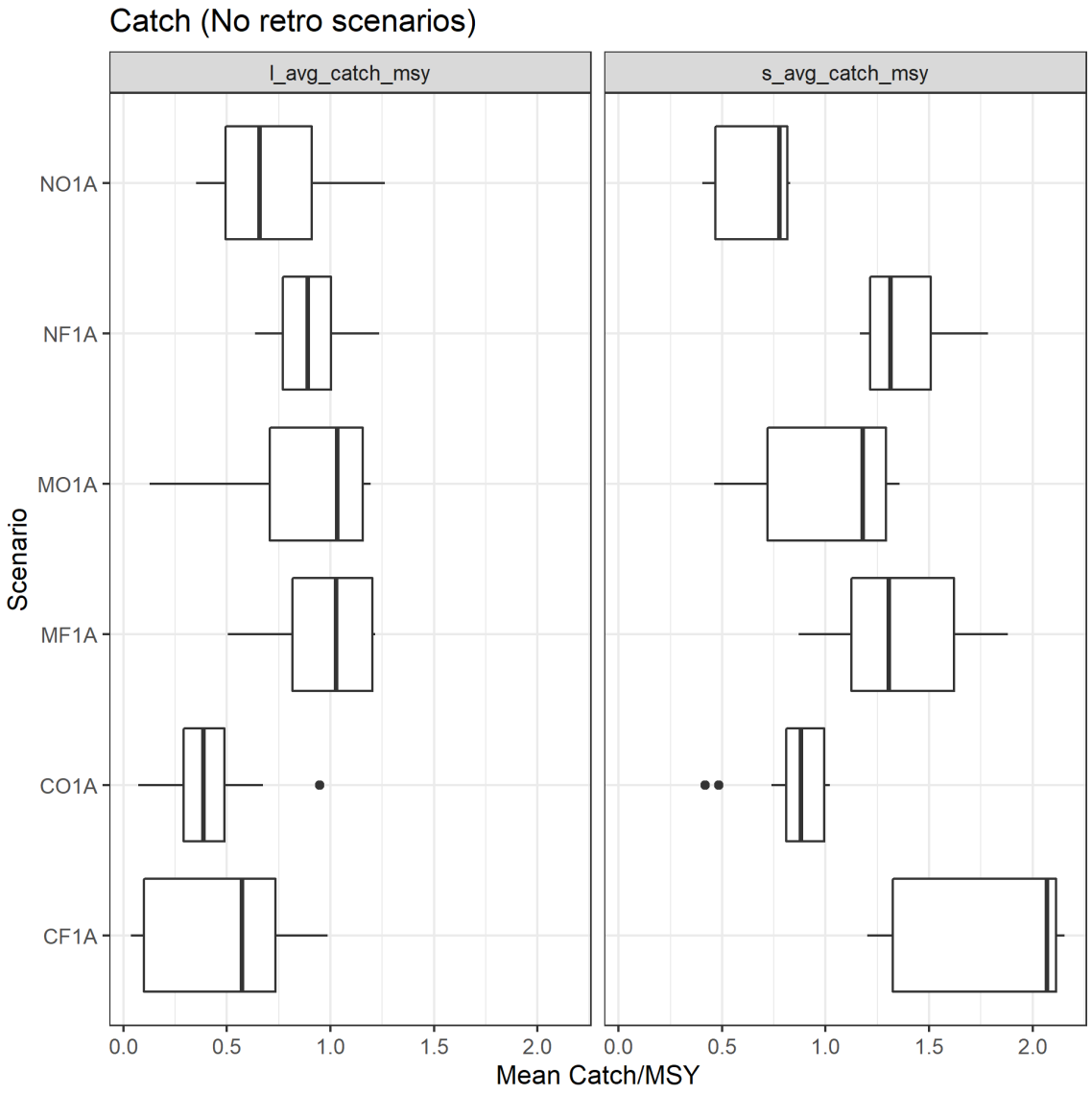


Figure A6.216. Boxplot of the mean values for the catch metrics in the no retro analyses by scenario.

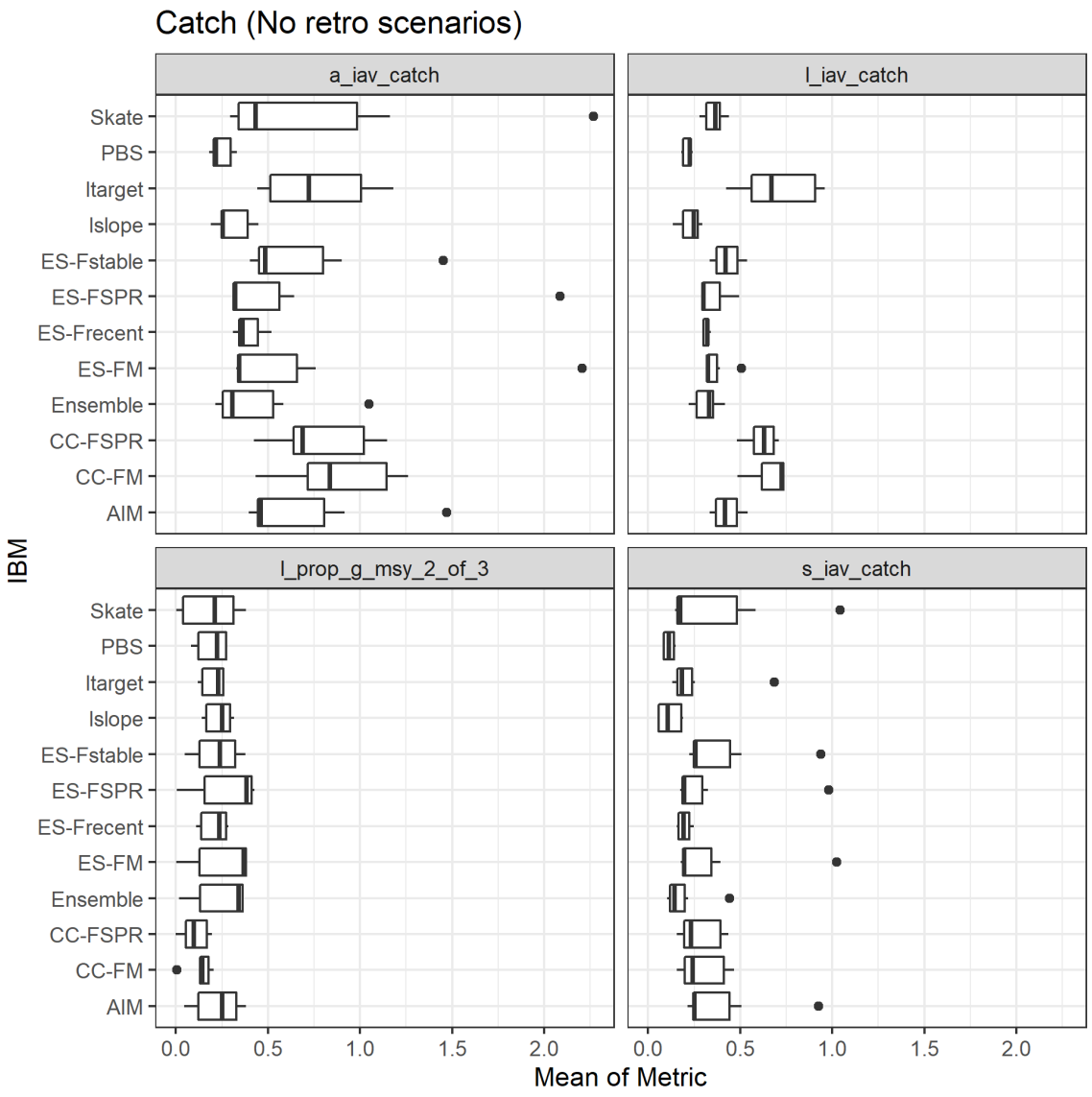


Figure A6.217. Boxplot of the mean values for the catch metrics in the no retro analyses by index-based method (IBM).

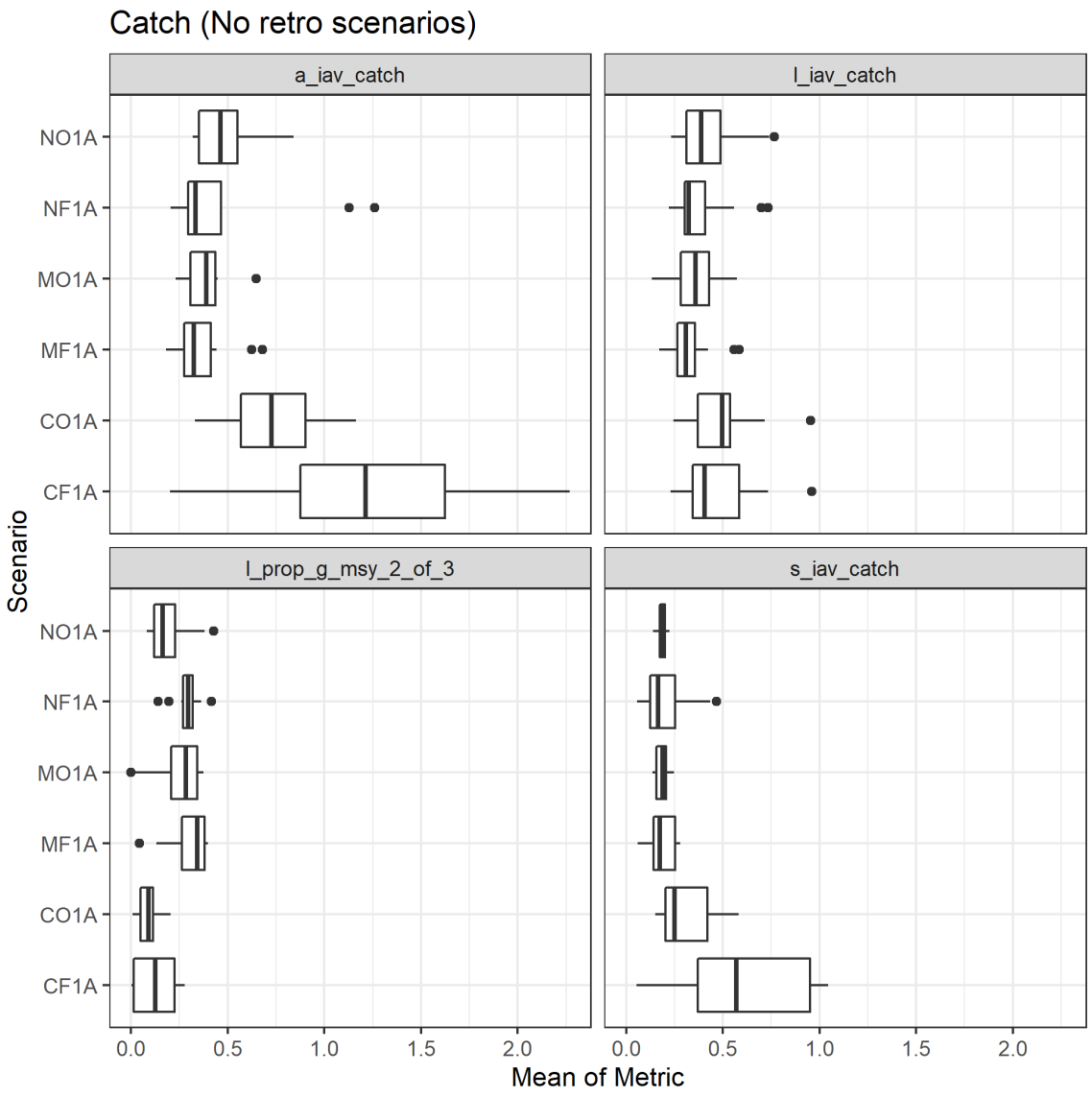


Figure A6.218. Boxplot of the mean values for the catch metrics in the no retro analyses by scenario.

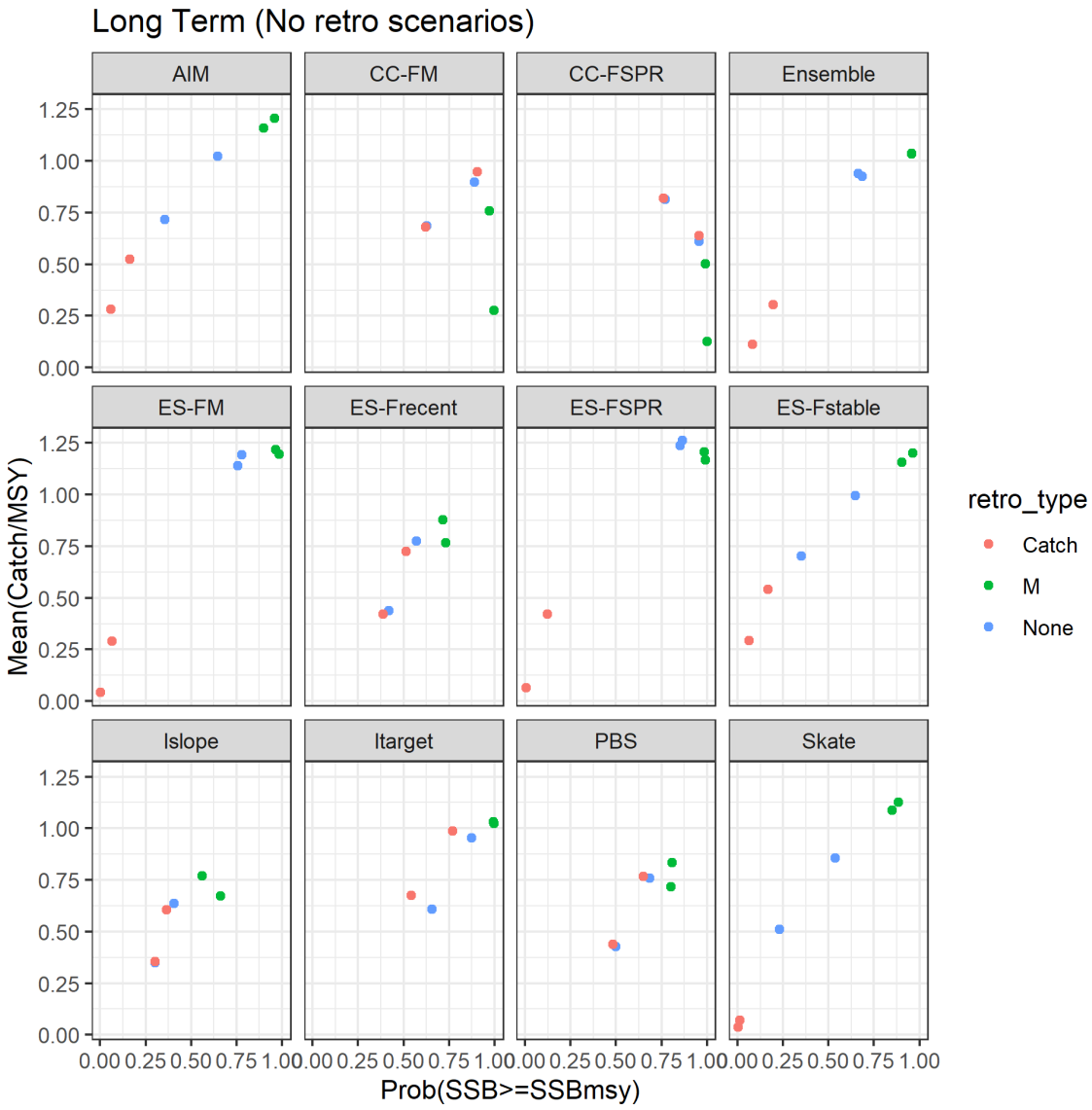


Figure A6.219. Trade off plot by index-based method (IBM) for the no retro analyses between the probability of a rebuilt stock and the mean catch relative to maximum sustainable yield (MSY) in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

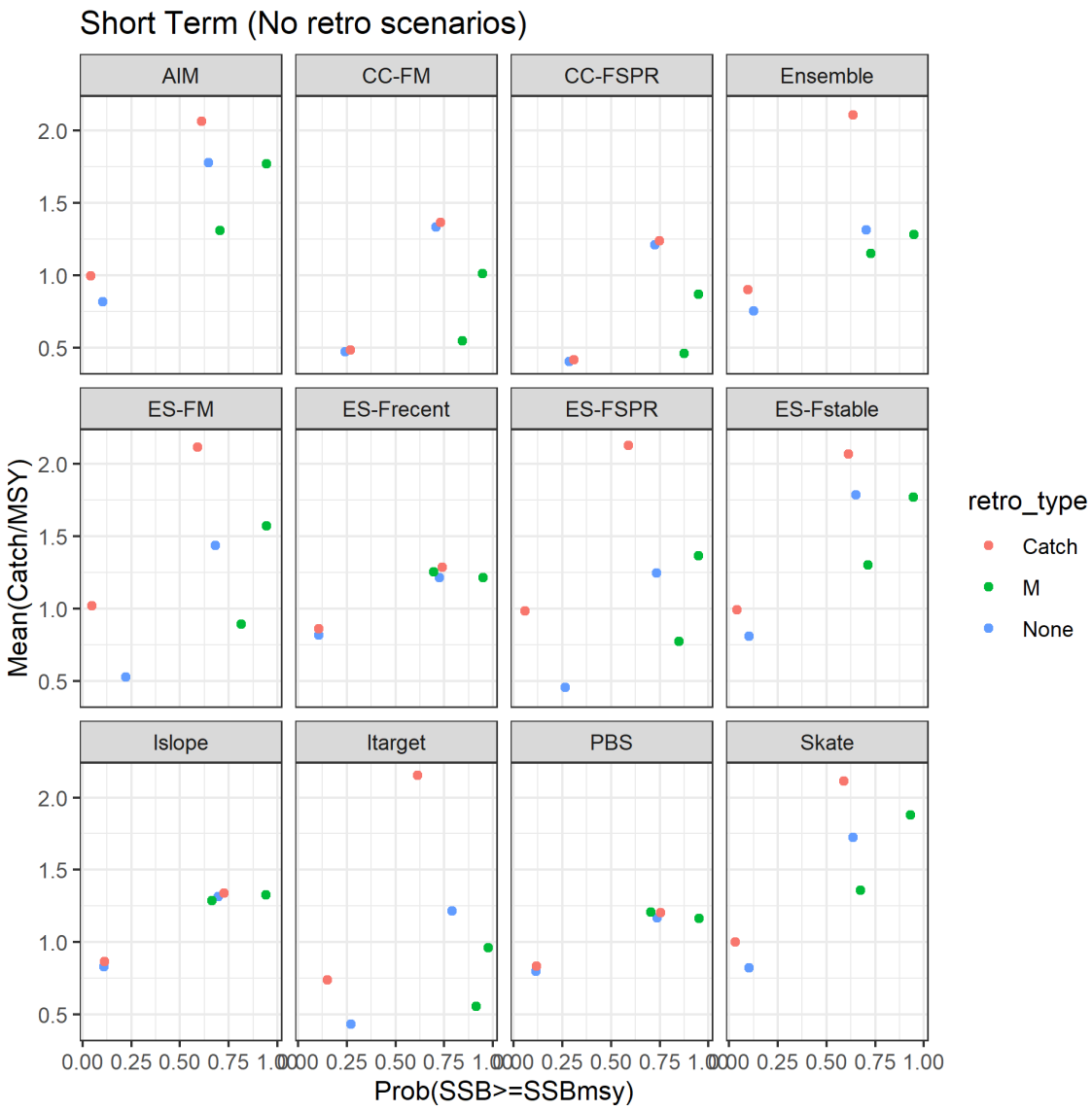


Figure A6.220. Trade off plot by index-based method (IBM) for the no retro analyses between the probability of a rebuilt stock and the mean catch relative to maximum sustainable yield (MSY) in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

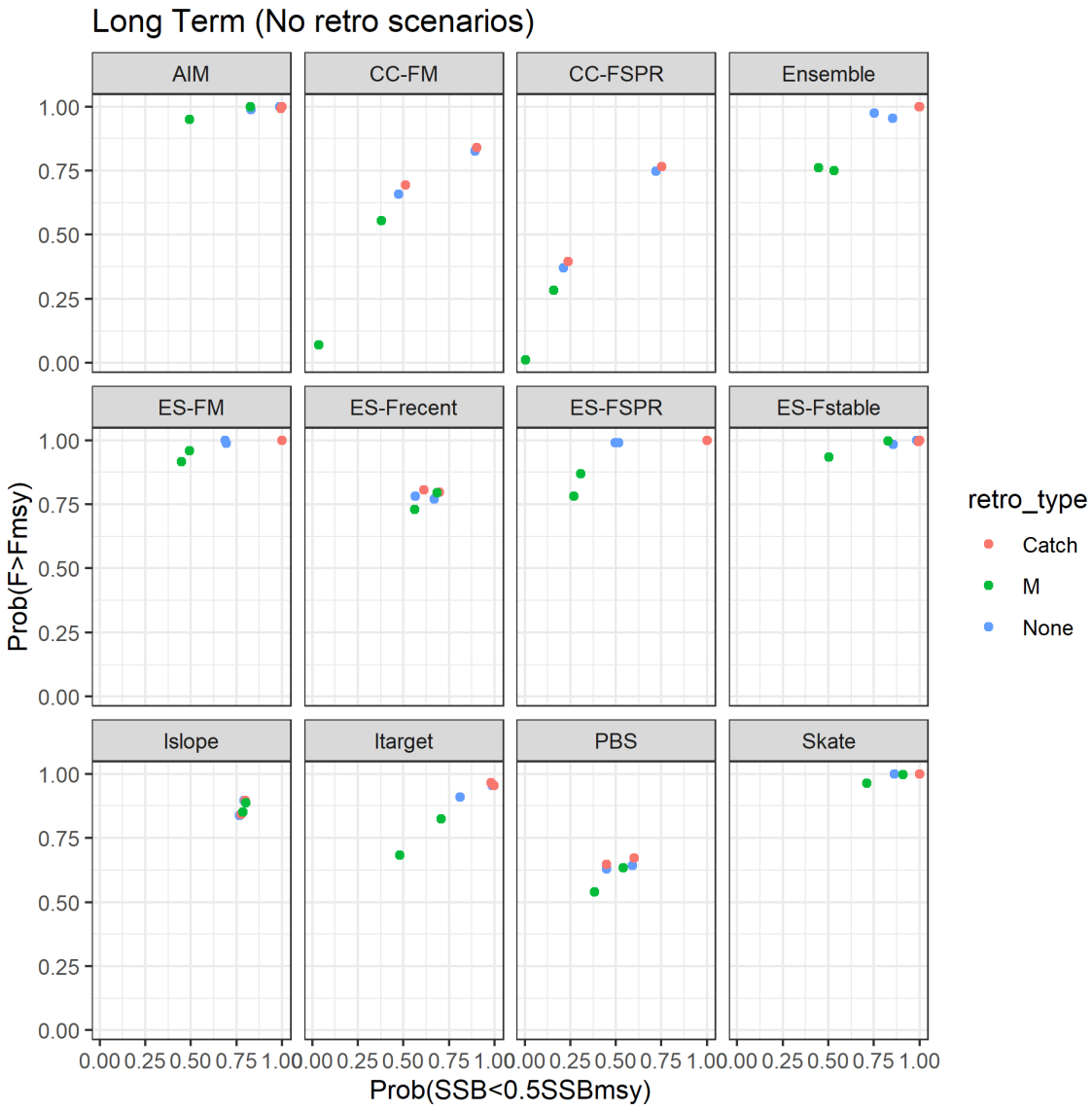


Figure A6.221. Trade off plot by index-based method (IBM) for the no retro analyses between the probability the stock is overfished and the probability that overfishing is occurring in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

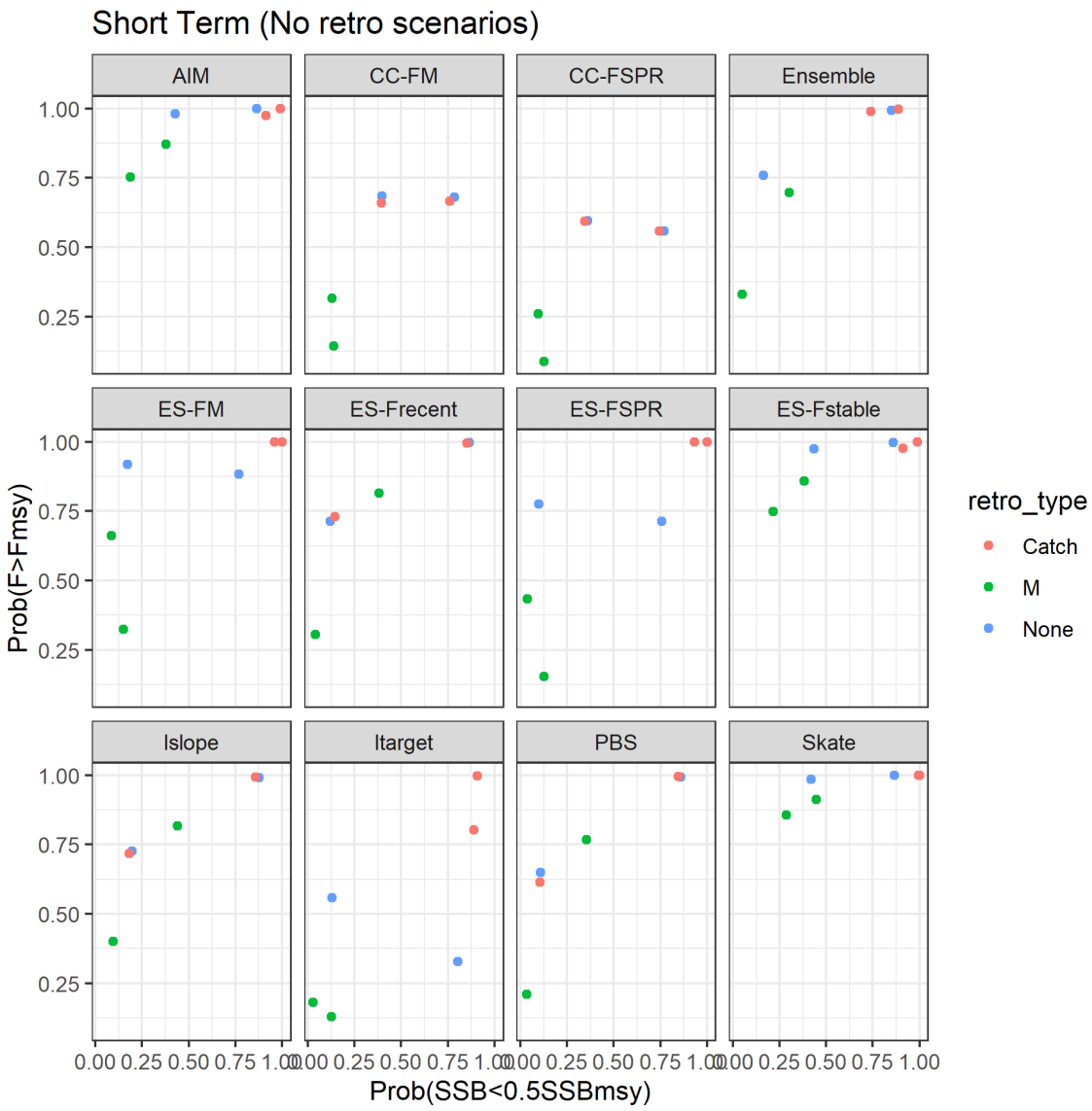


Figure A6.222. Trade off plot by index-based method (IBM) for the no retro analyses between the probability the stock is overfished and the probability that overfishing is occurring in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

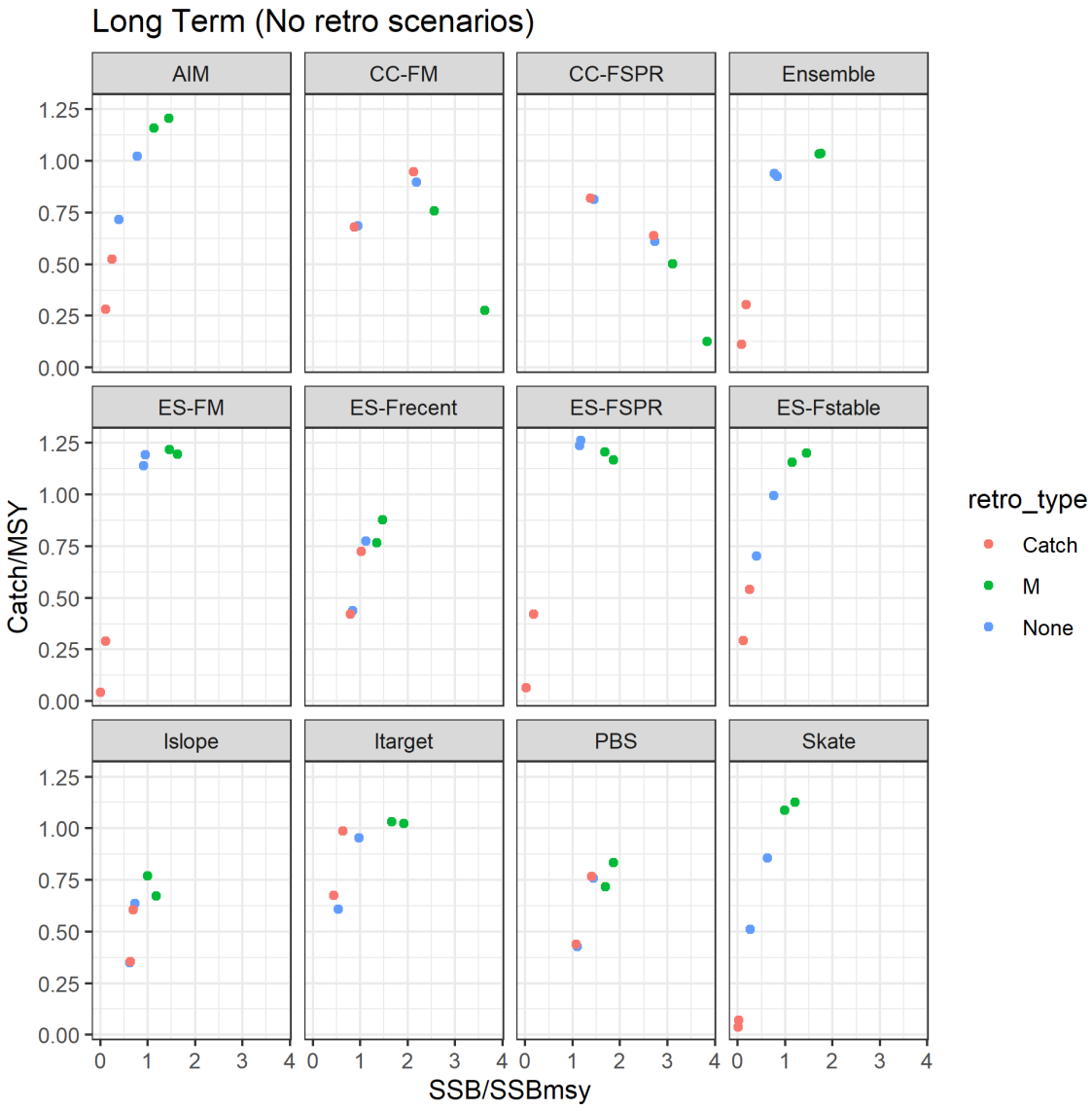


Figure A6.223. Trade off plot by index-based method (IBM) for the no retro analyses between the average spawning stock biomass (SSB) and catch relative to their maximum sustainable yield (MSY) reference points in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

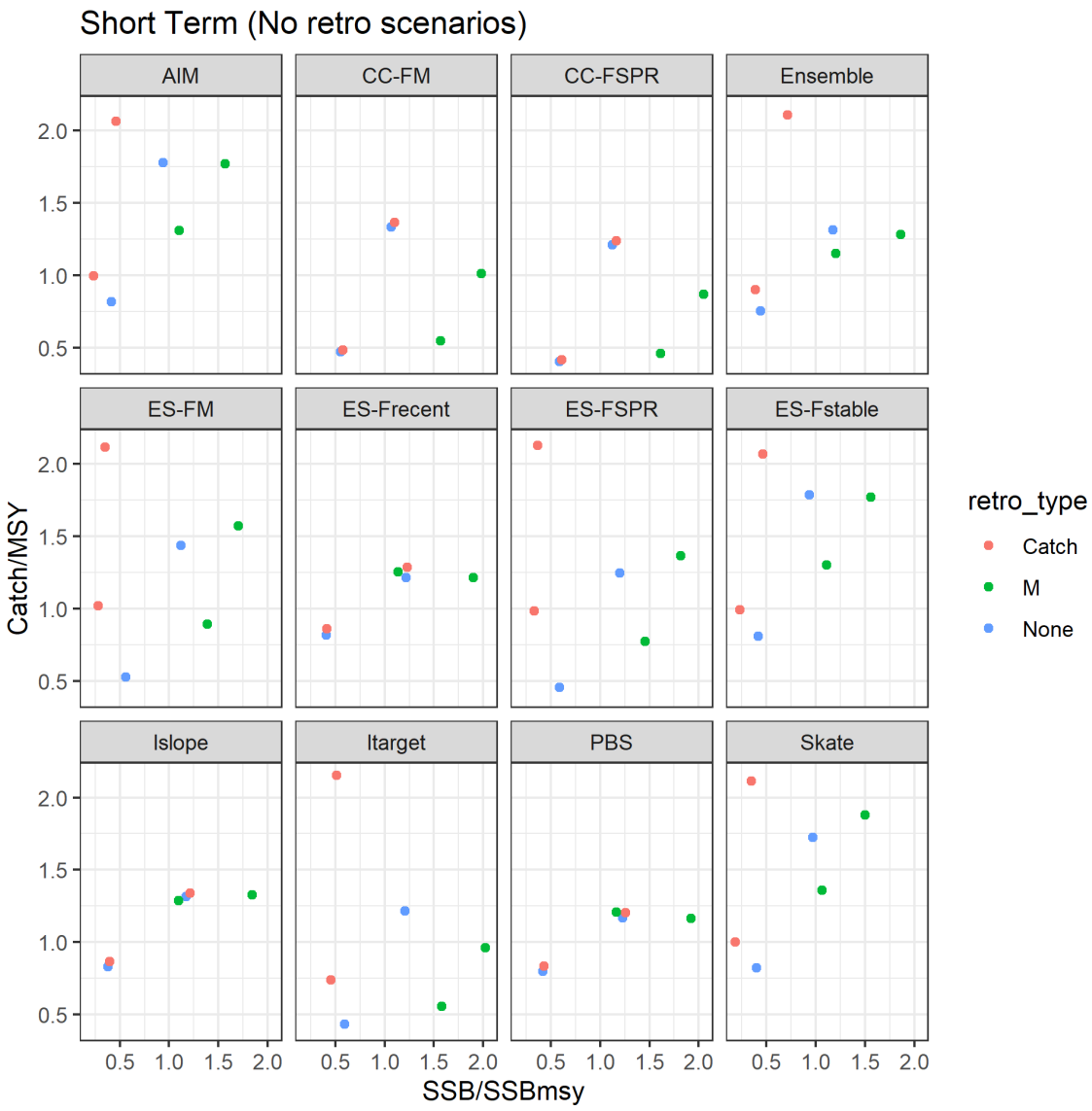


Figure A6.224. Trade off plot by index-based method (IBM) for the no retro analyses between the average spawning stock biomass (SSB) and catch relative to their maximum sustainable yield (MSY) reference points in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.



Figure A6.225. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

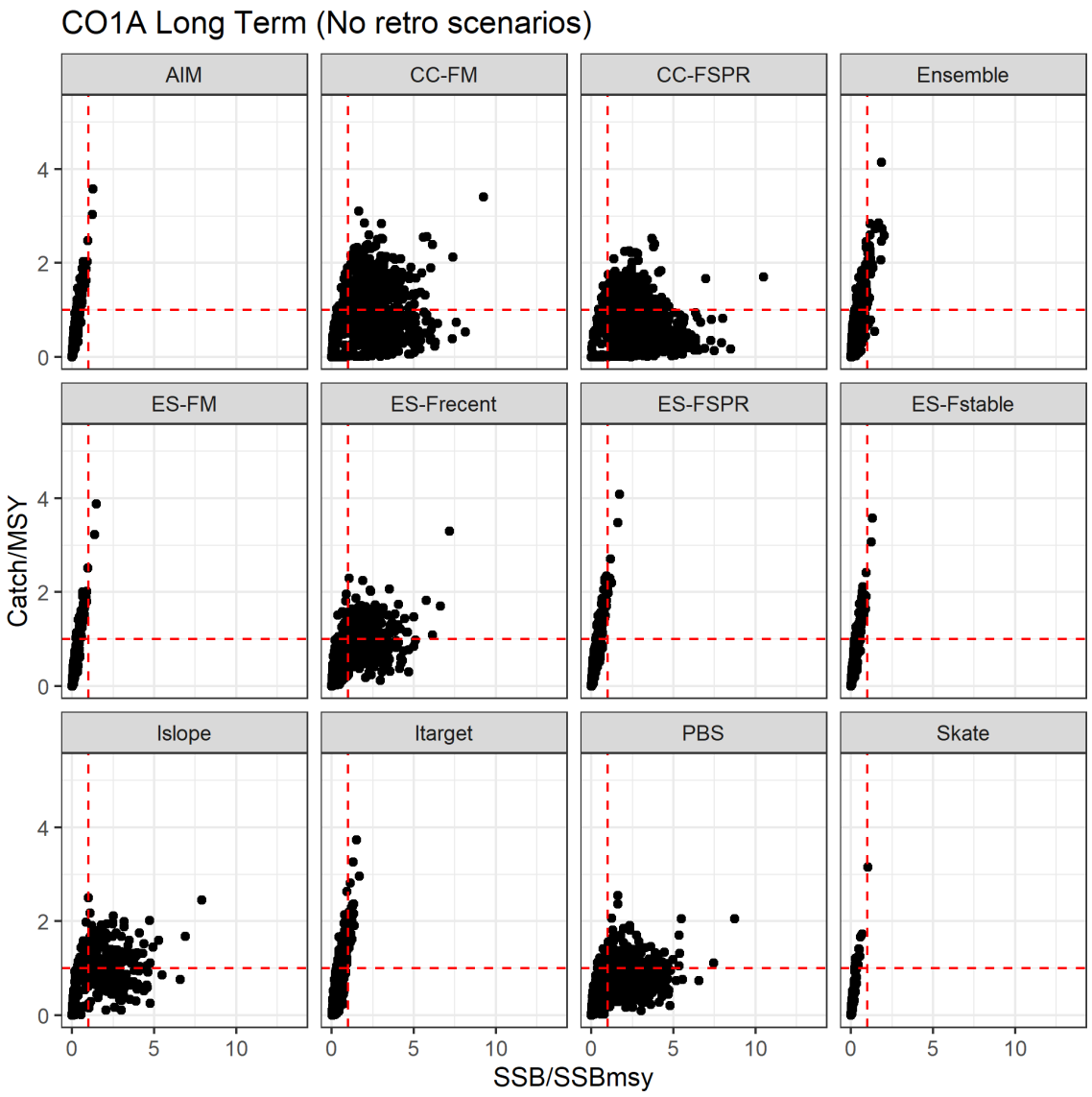


Figure A6.226. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

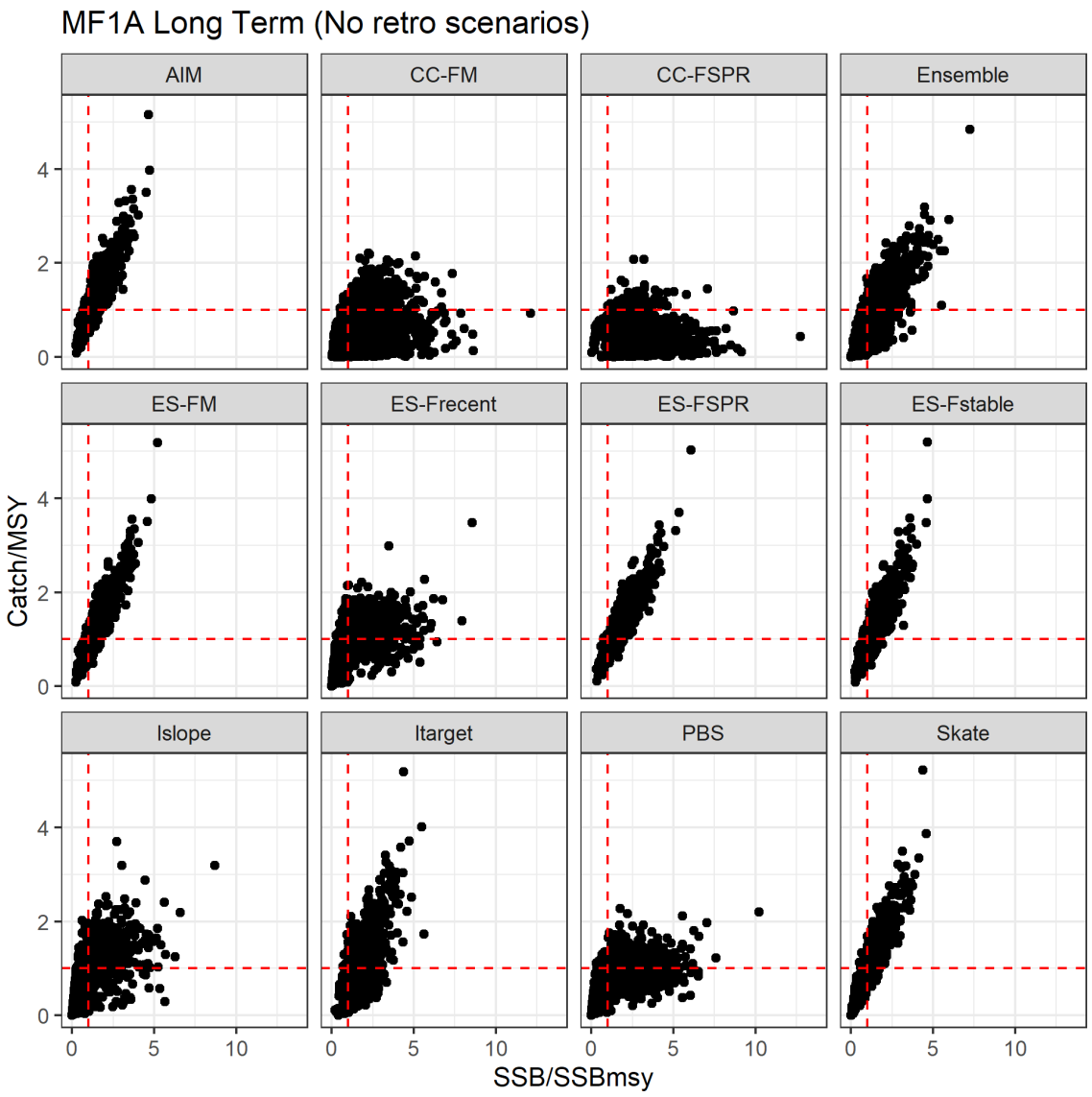


Figure A6.227. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

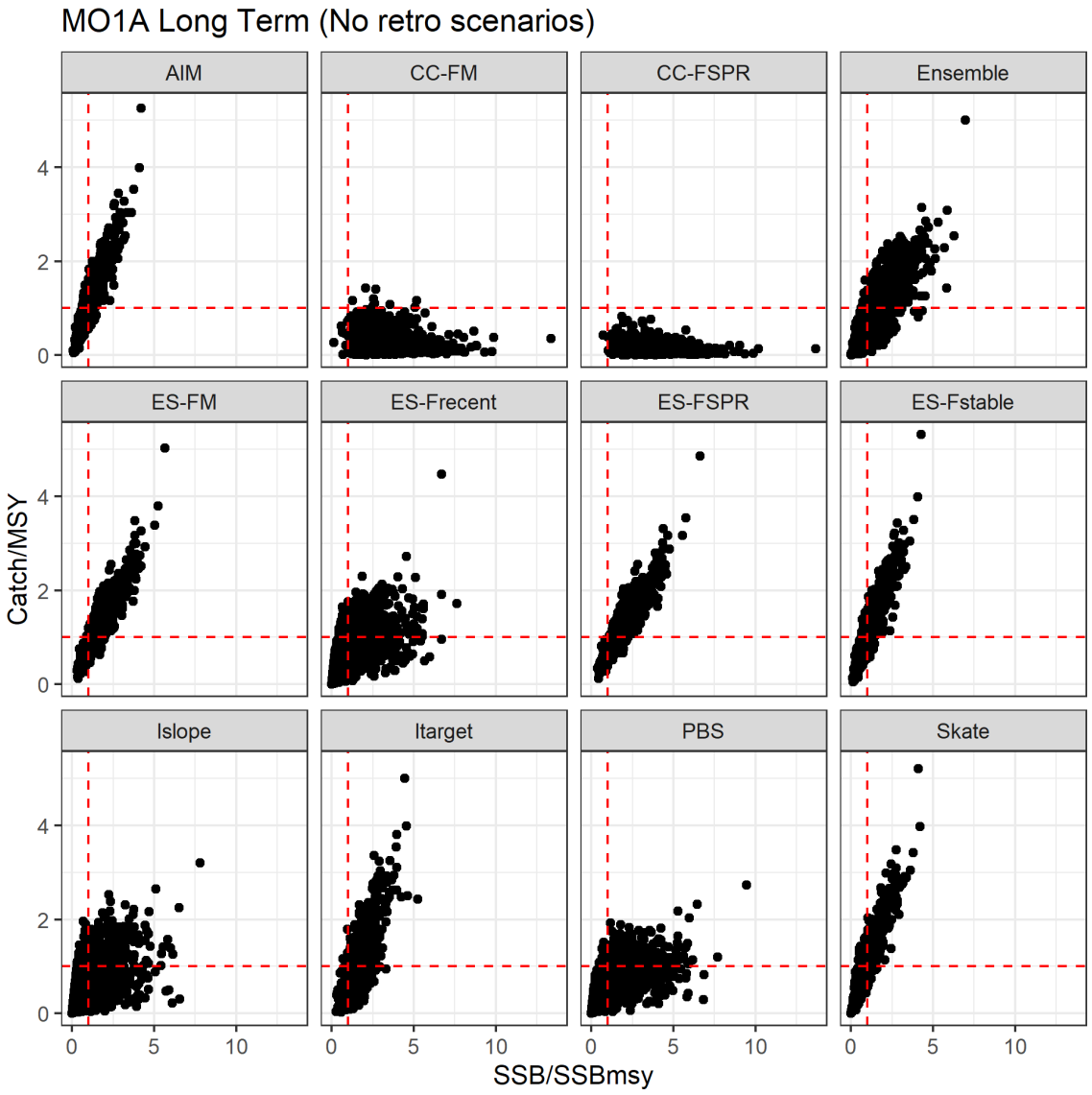


Figure A6.228. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

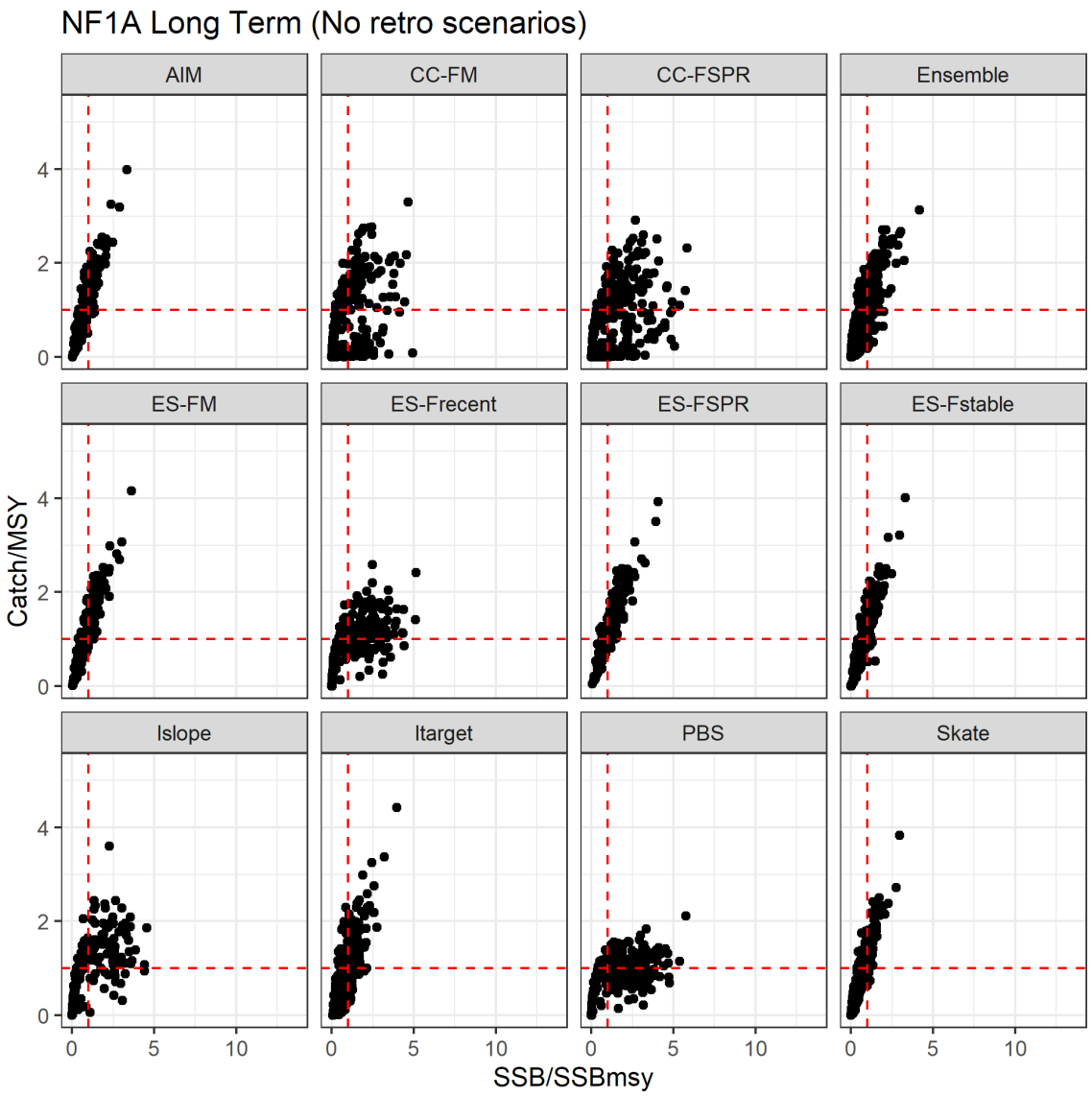


Figure A6.229. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

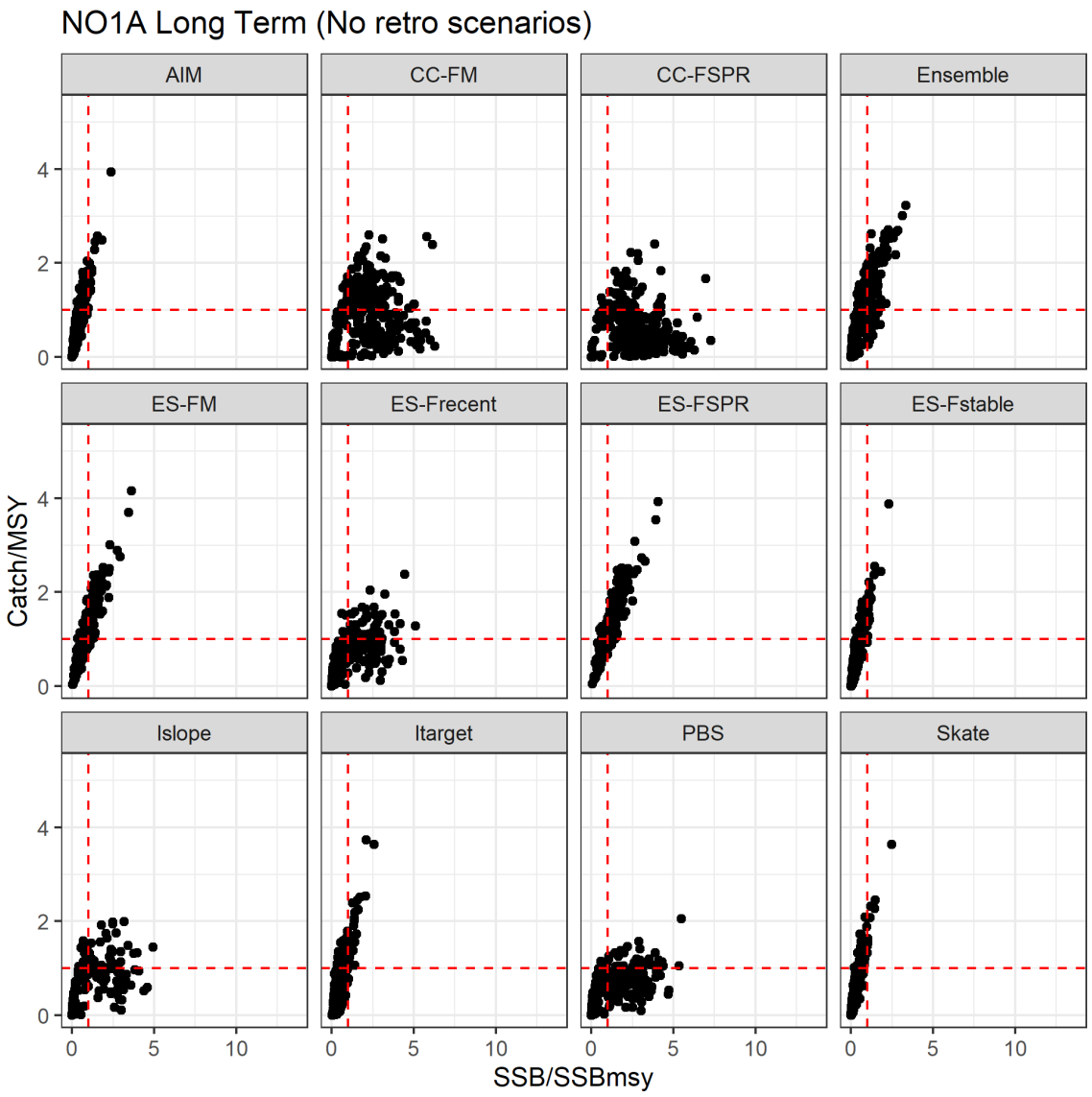


Figure A6.230. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

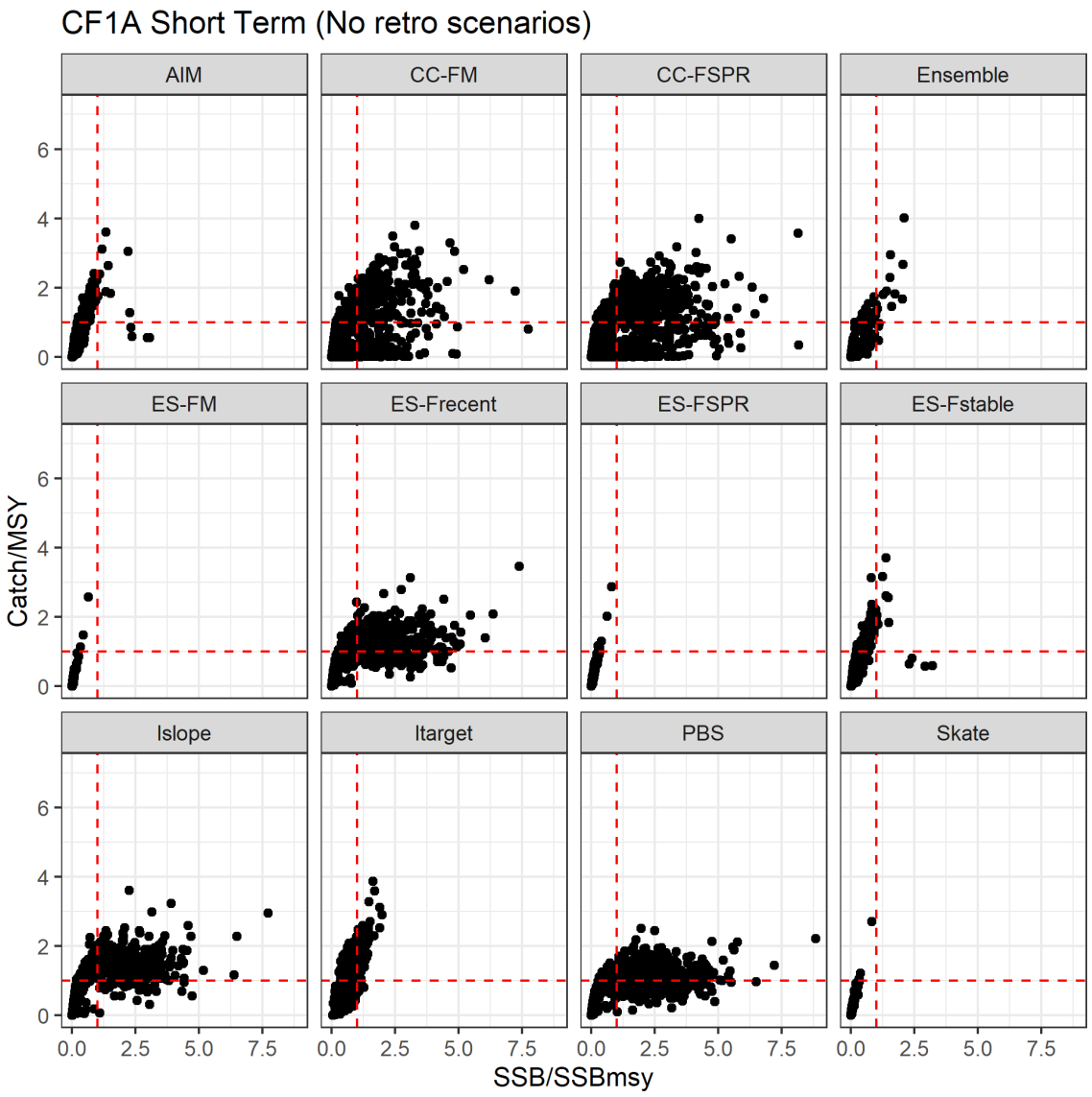


Figure A6.231. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

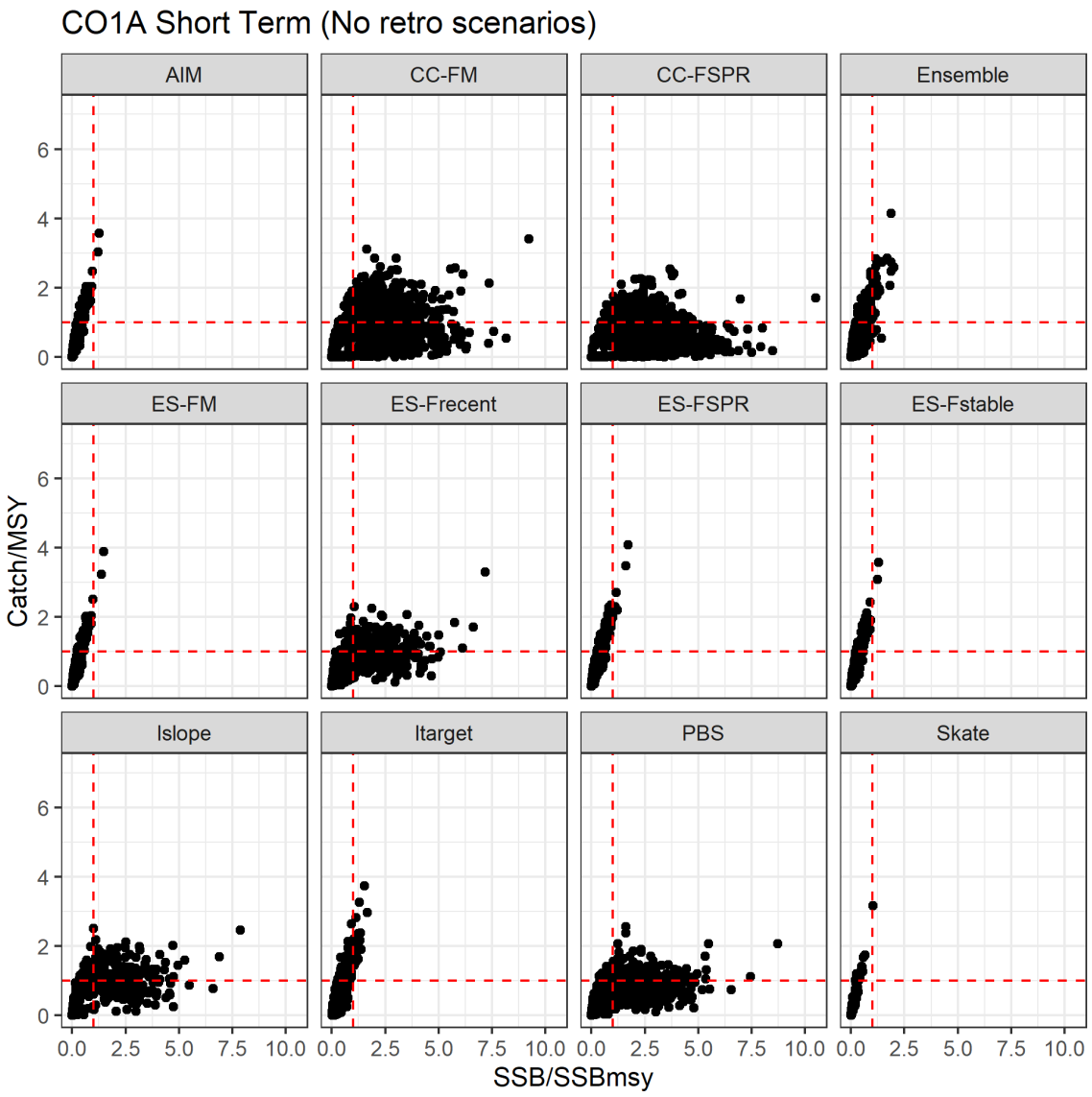


Figure A6.232. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

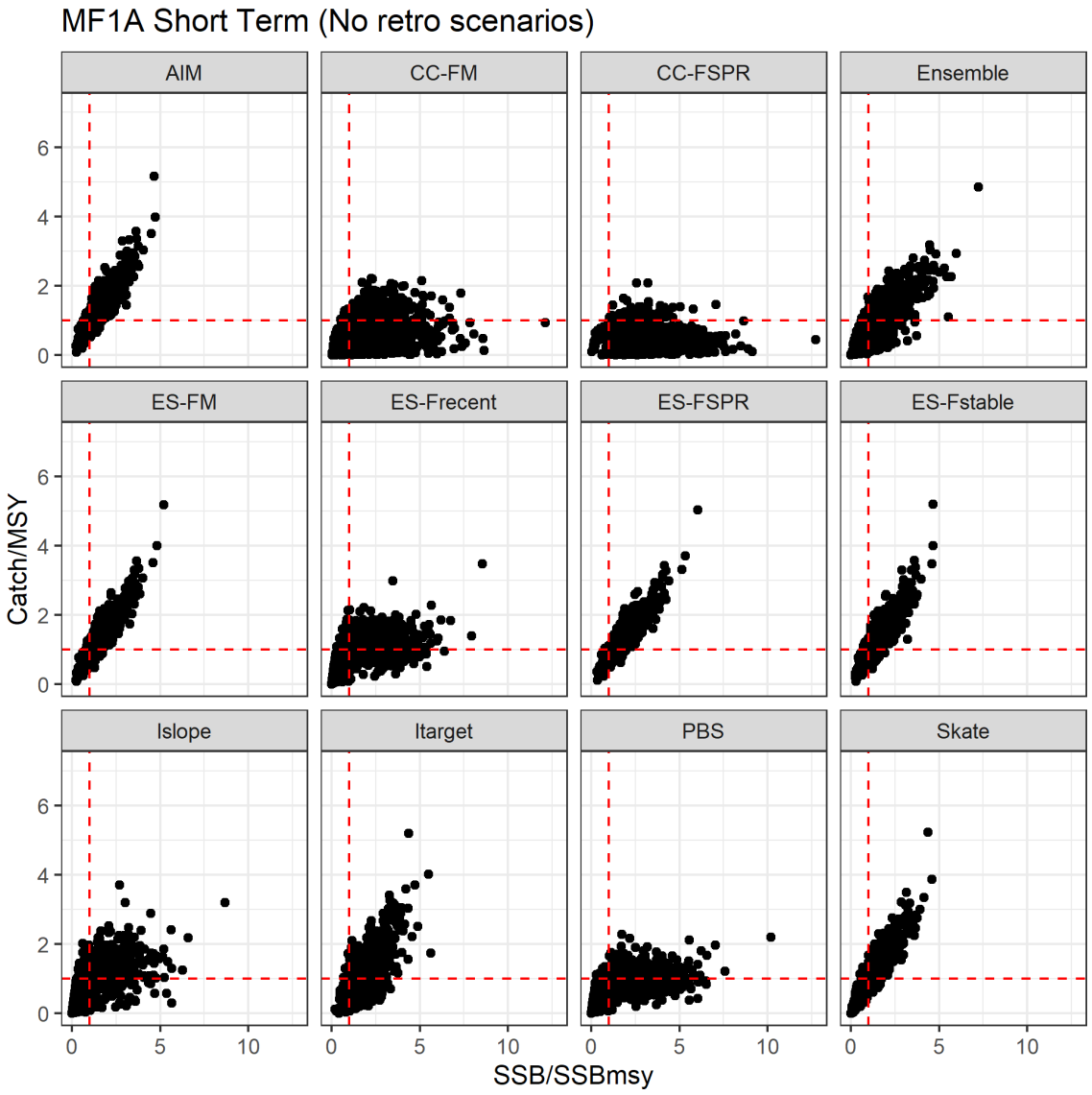


Figure A6.233. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

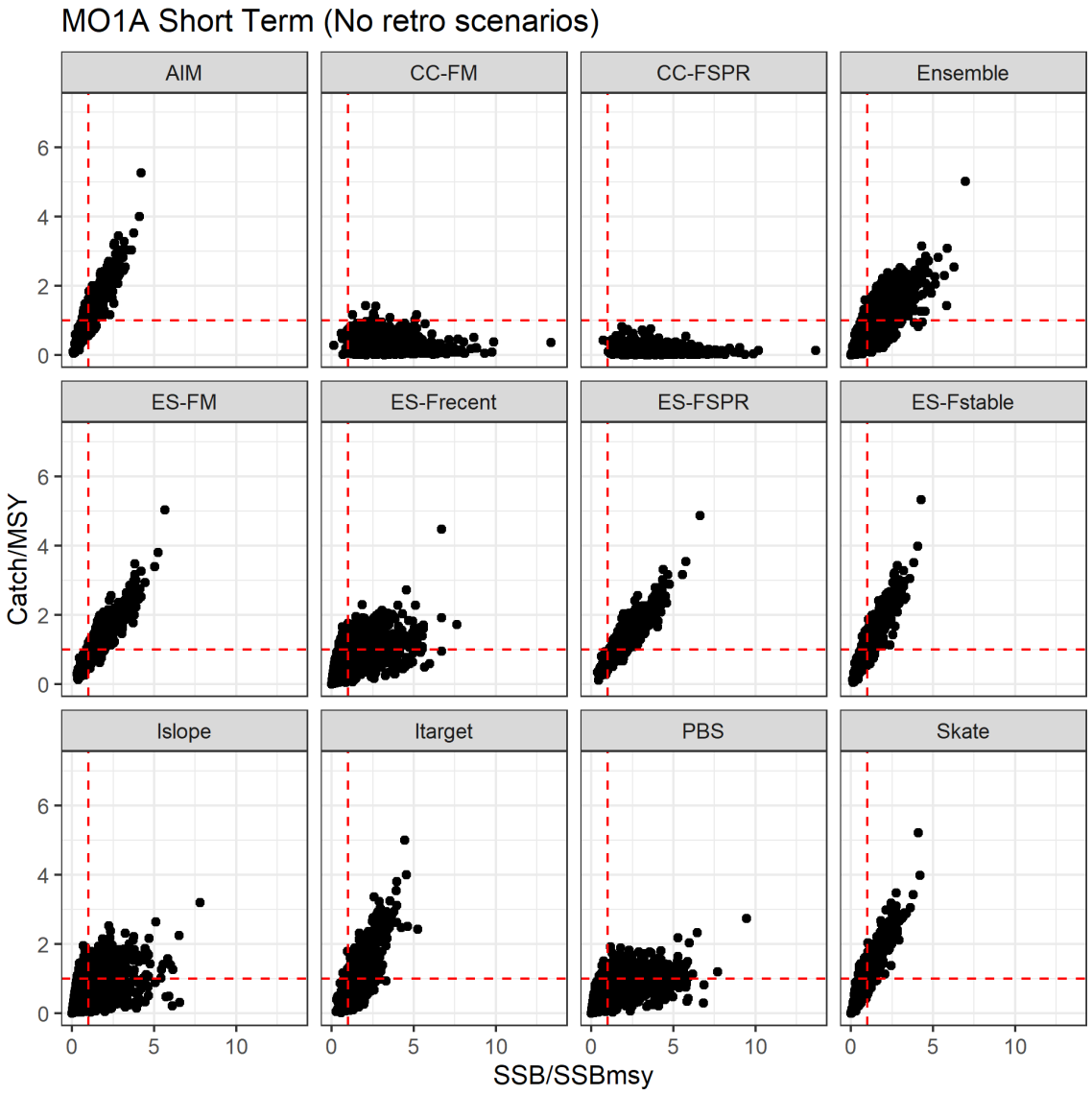


Figure A6.234. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

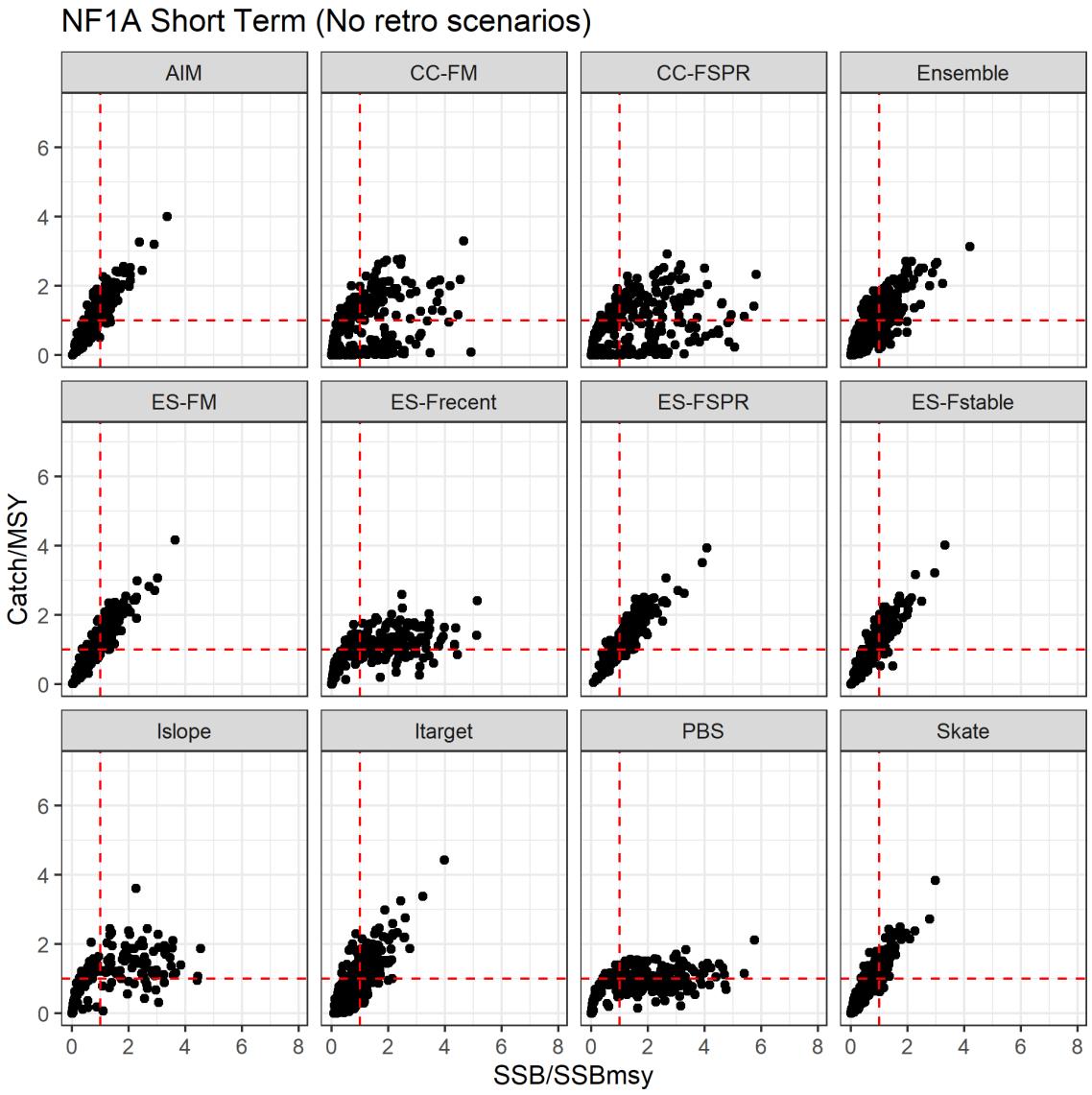


Figure A6.235. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

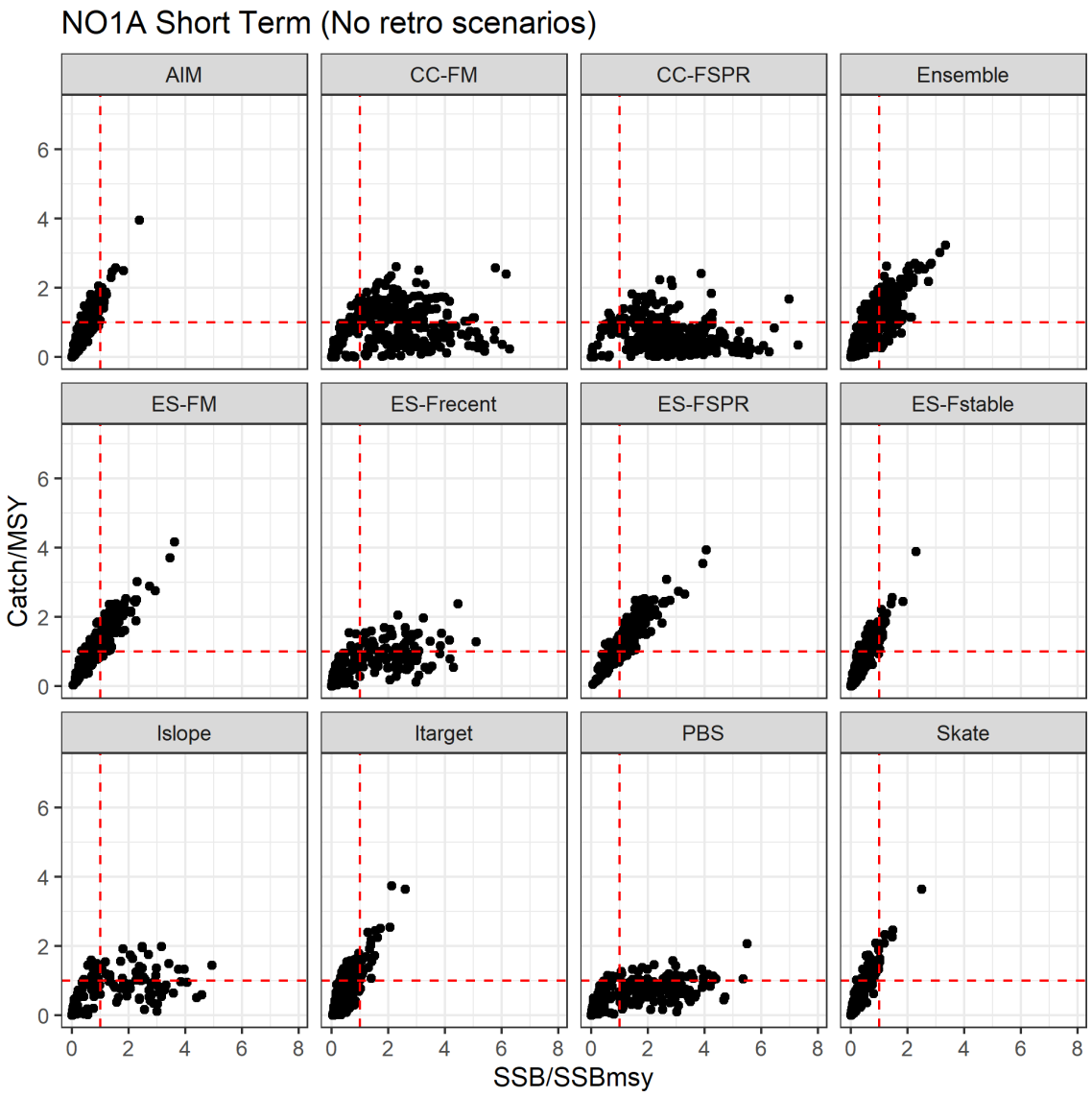


Figure A6.236. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

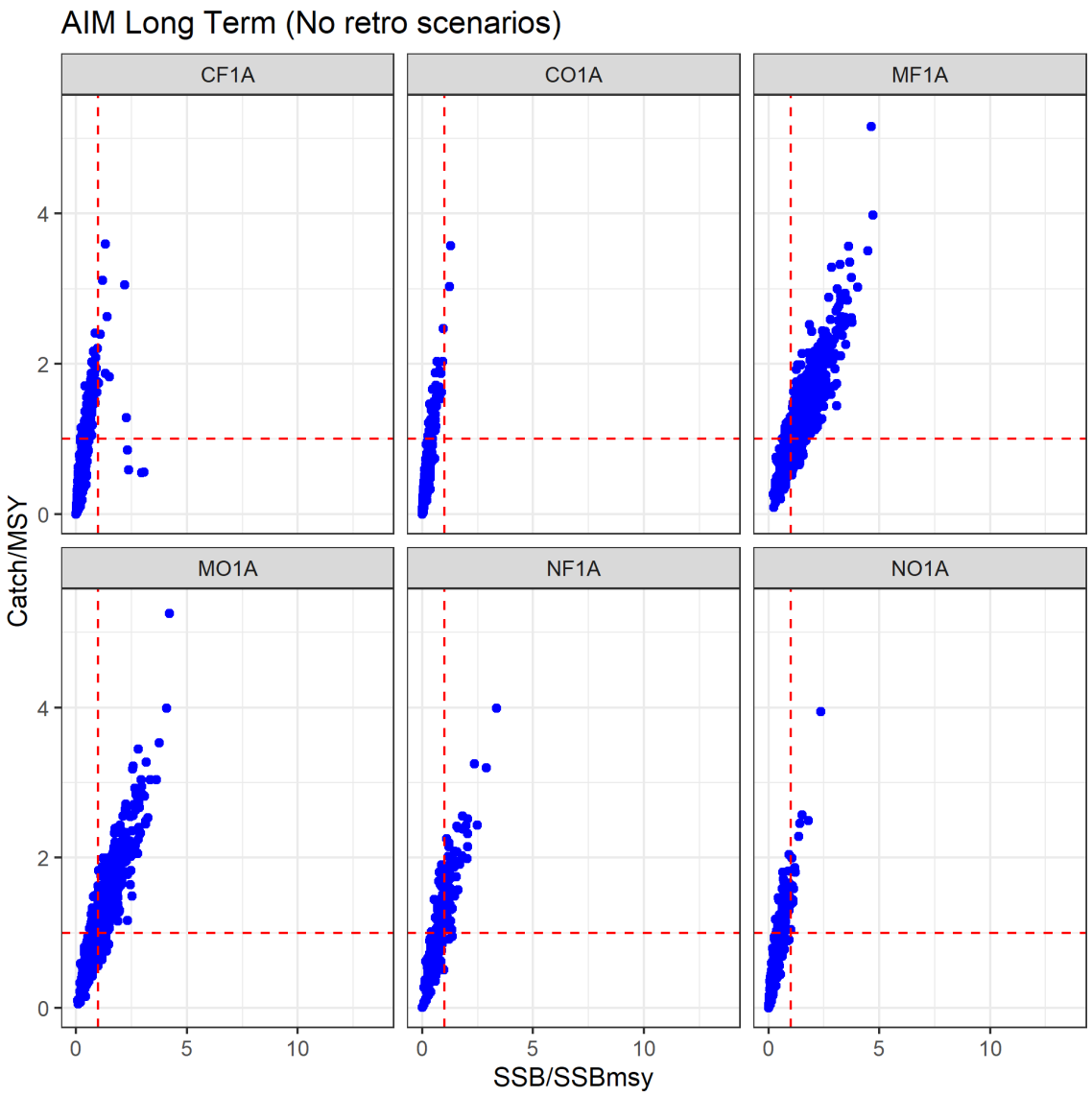


Figure A6.237. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

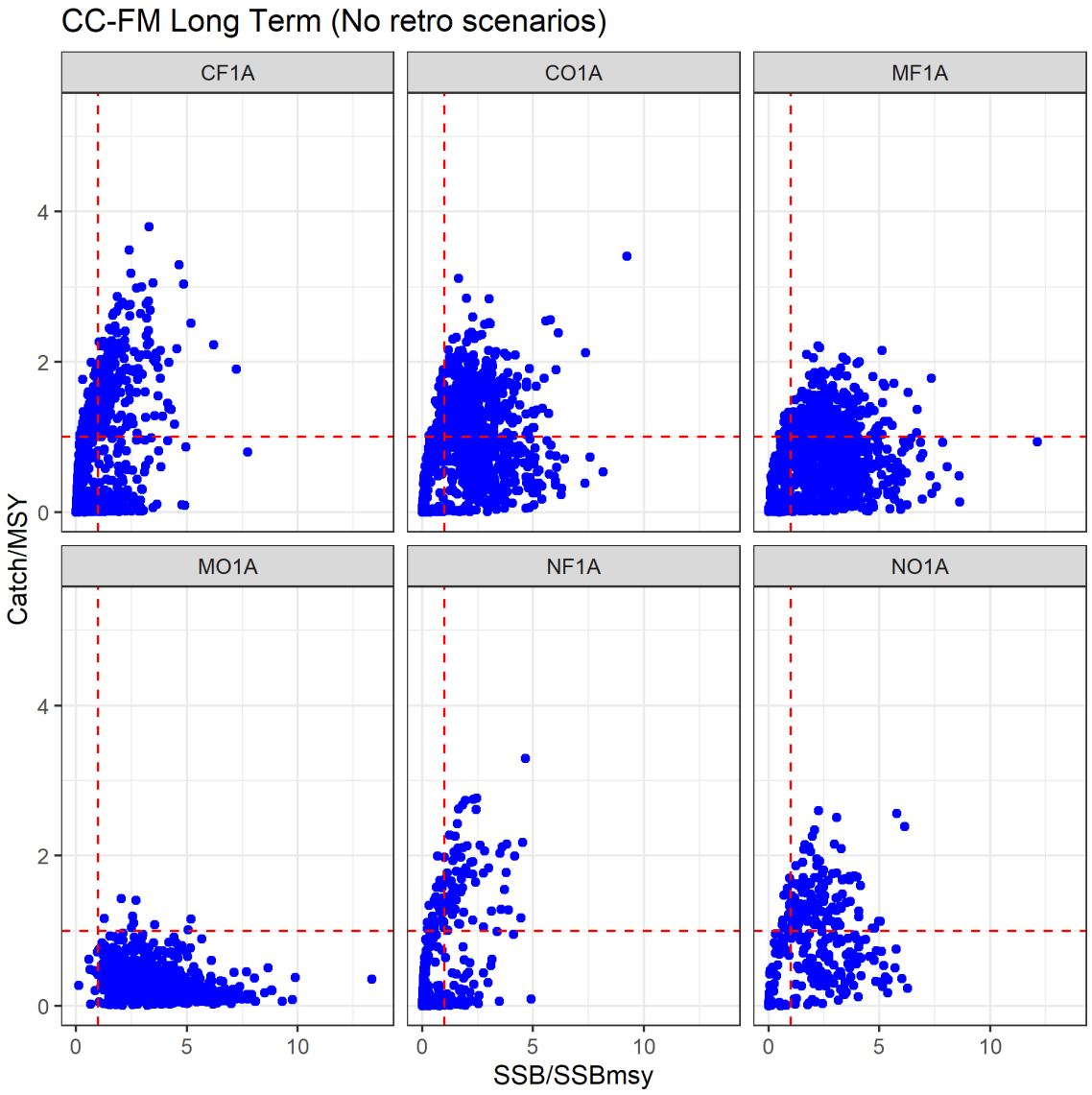


Figure A6.238. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

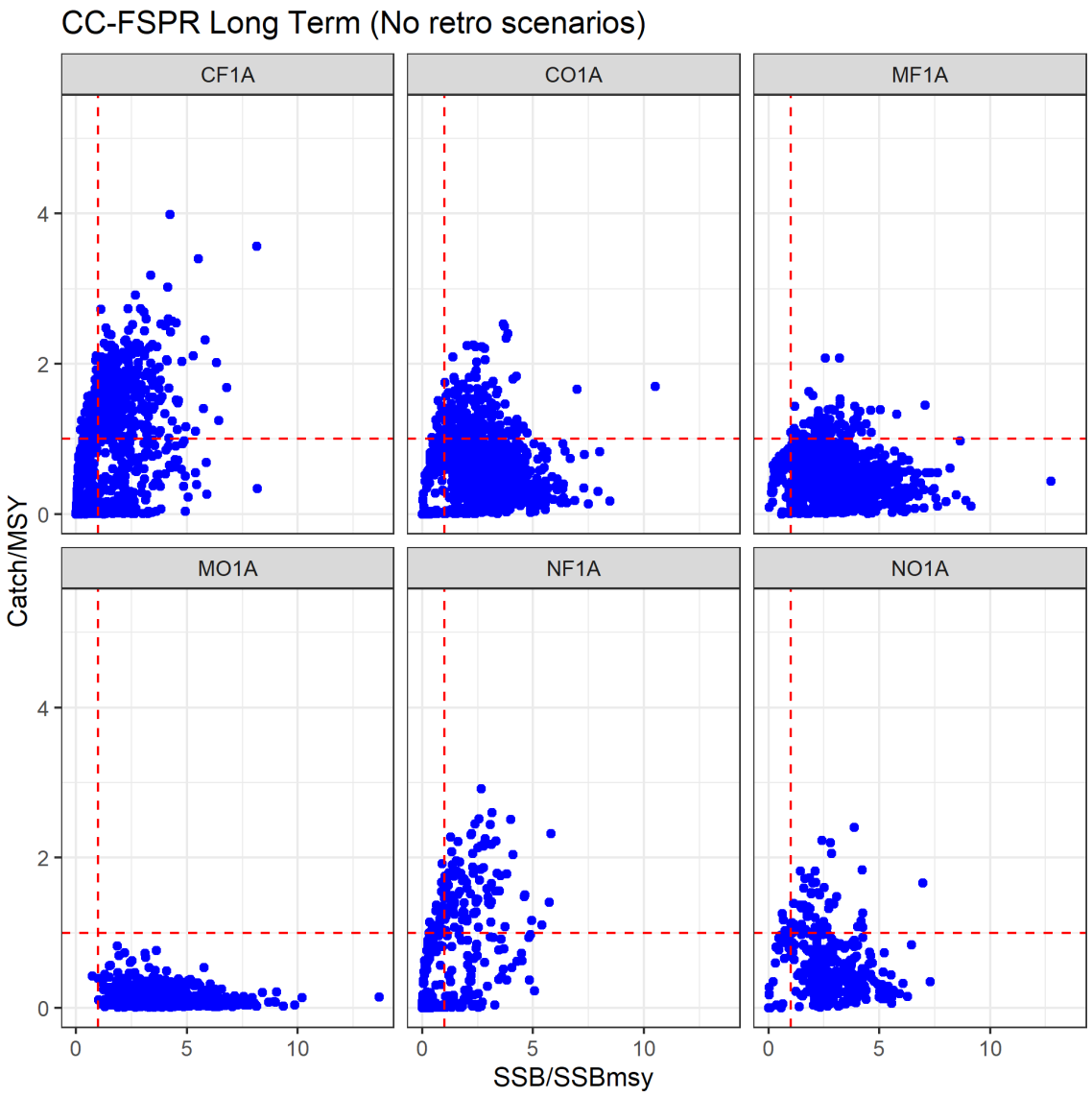


Figure A6.239. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

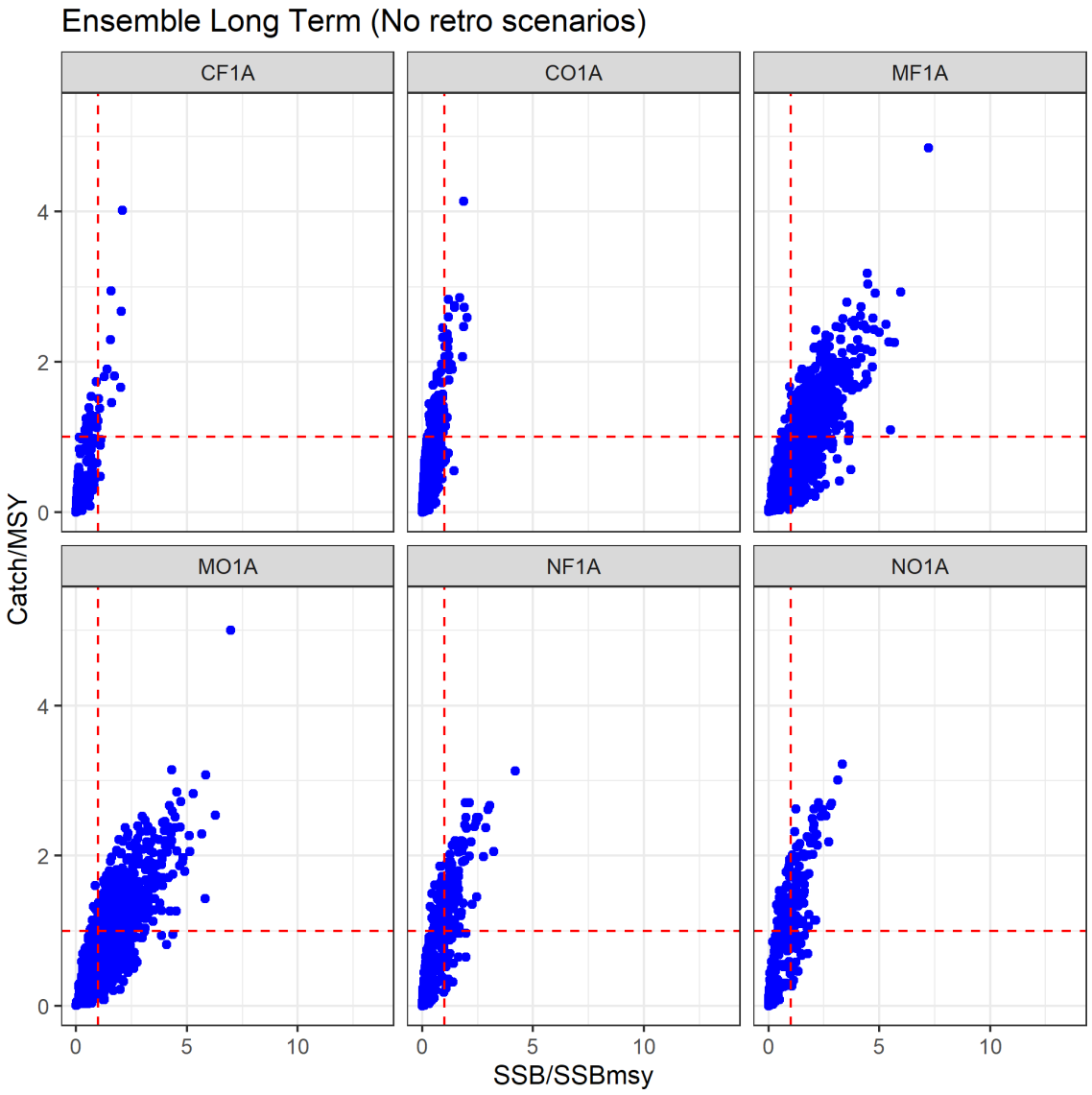


Figure A6.240. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

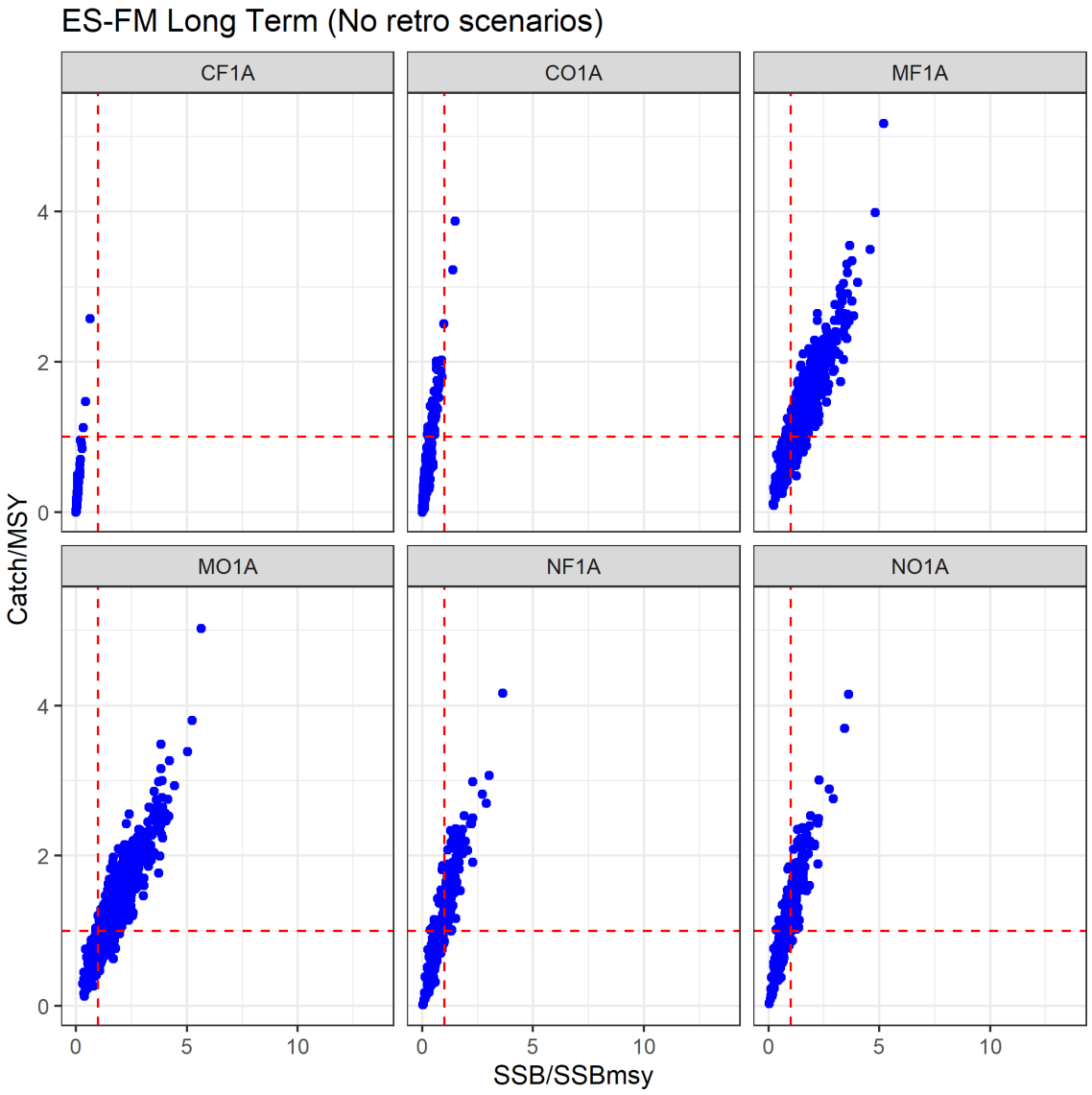


Figure A6.241. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

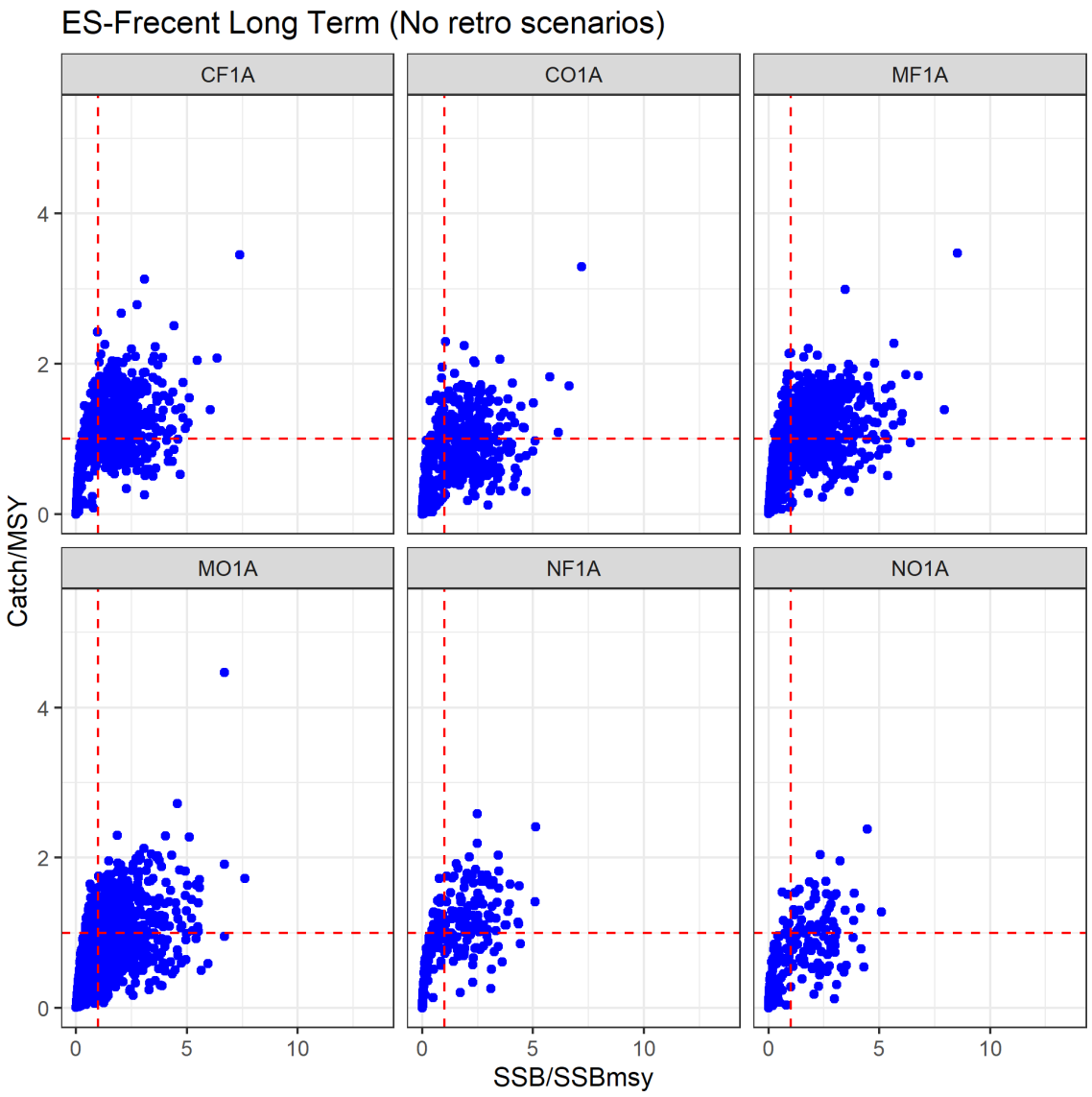


Figure A6.242. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

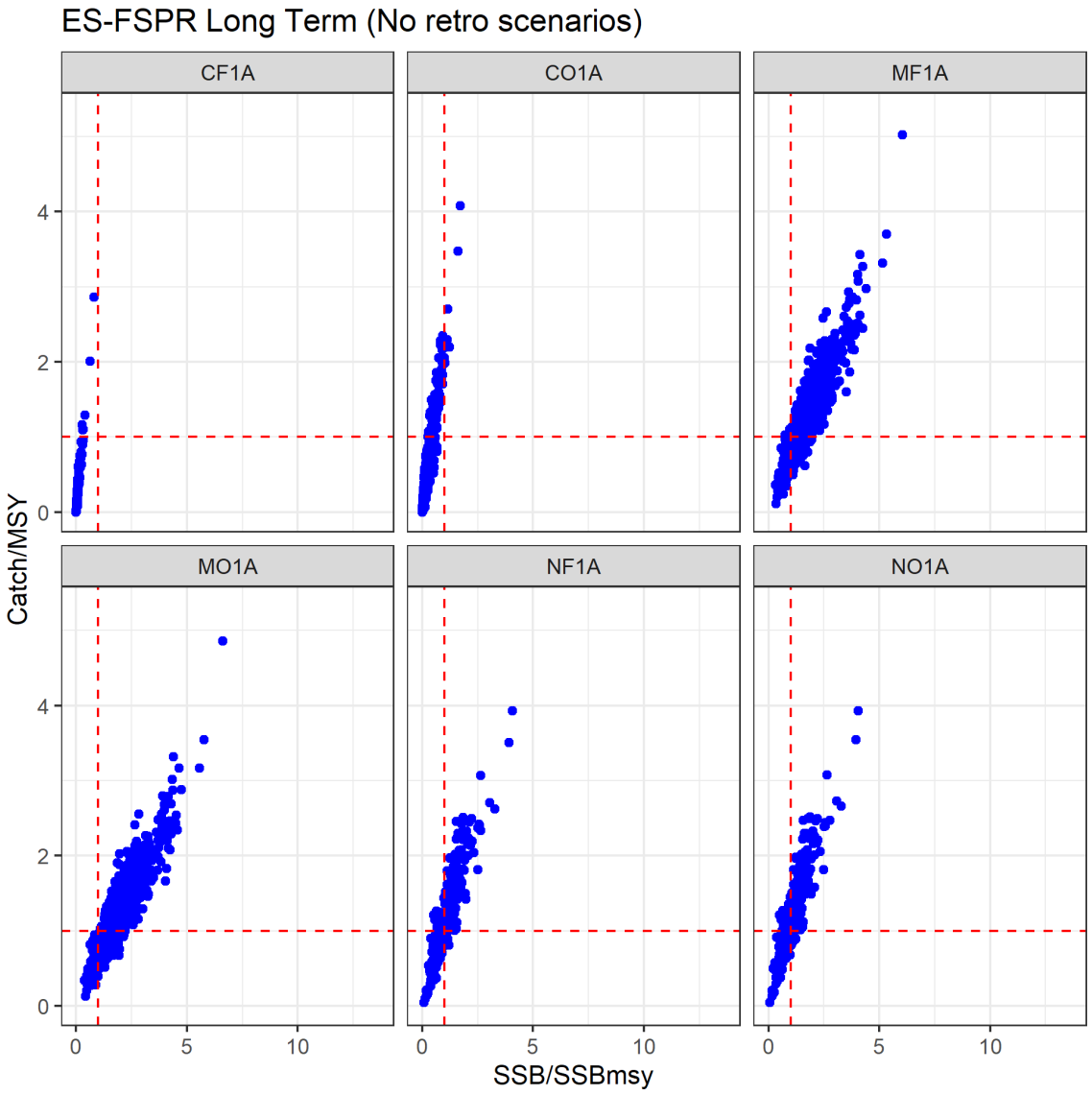


Figure A6.243. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

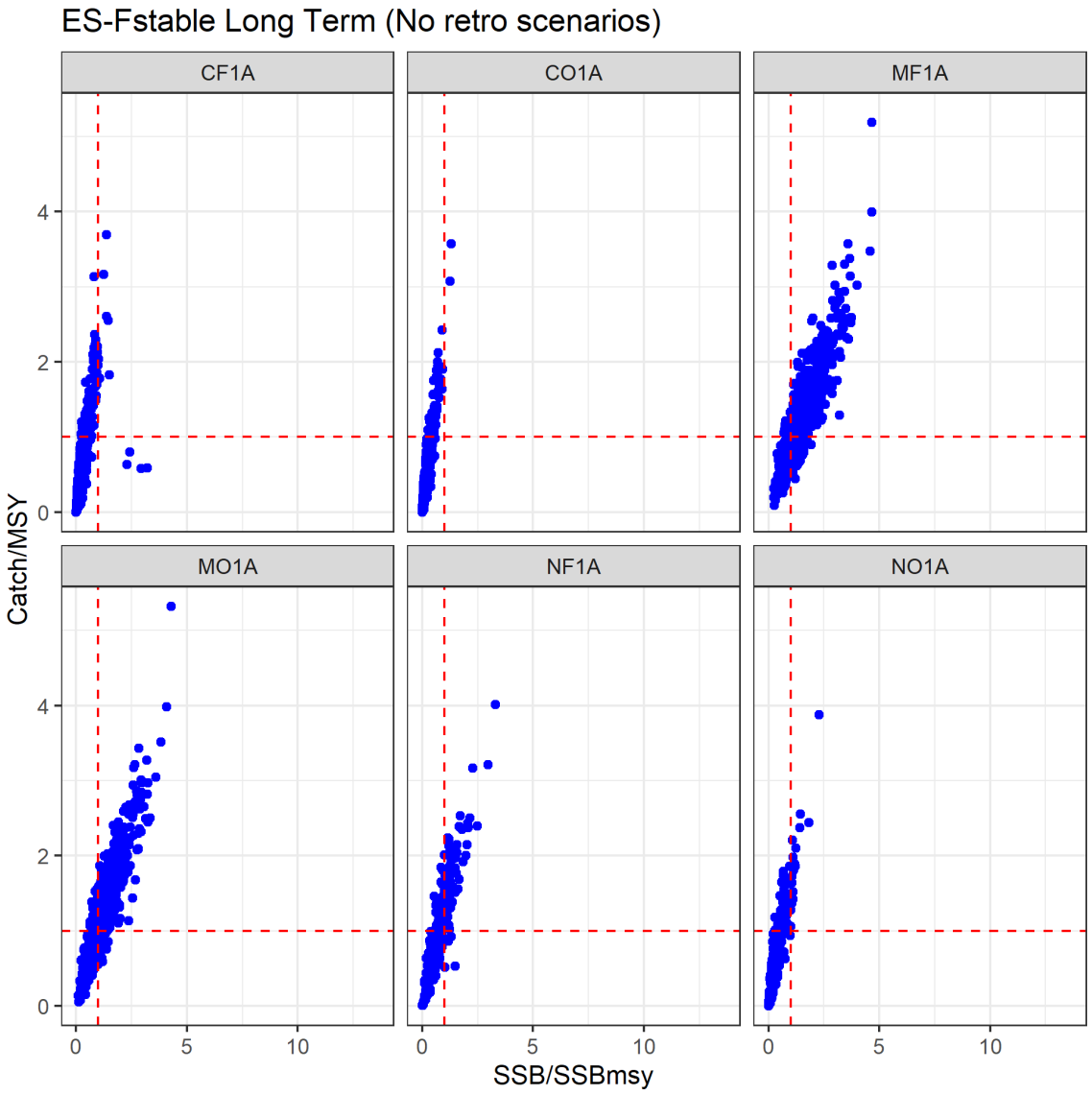


Figure A6.244. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.



Figure A6.245. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

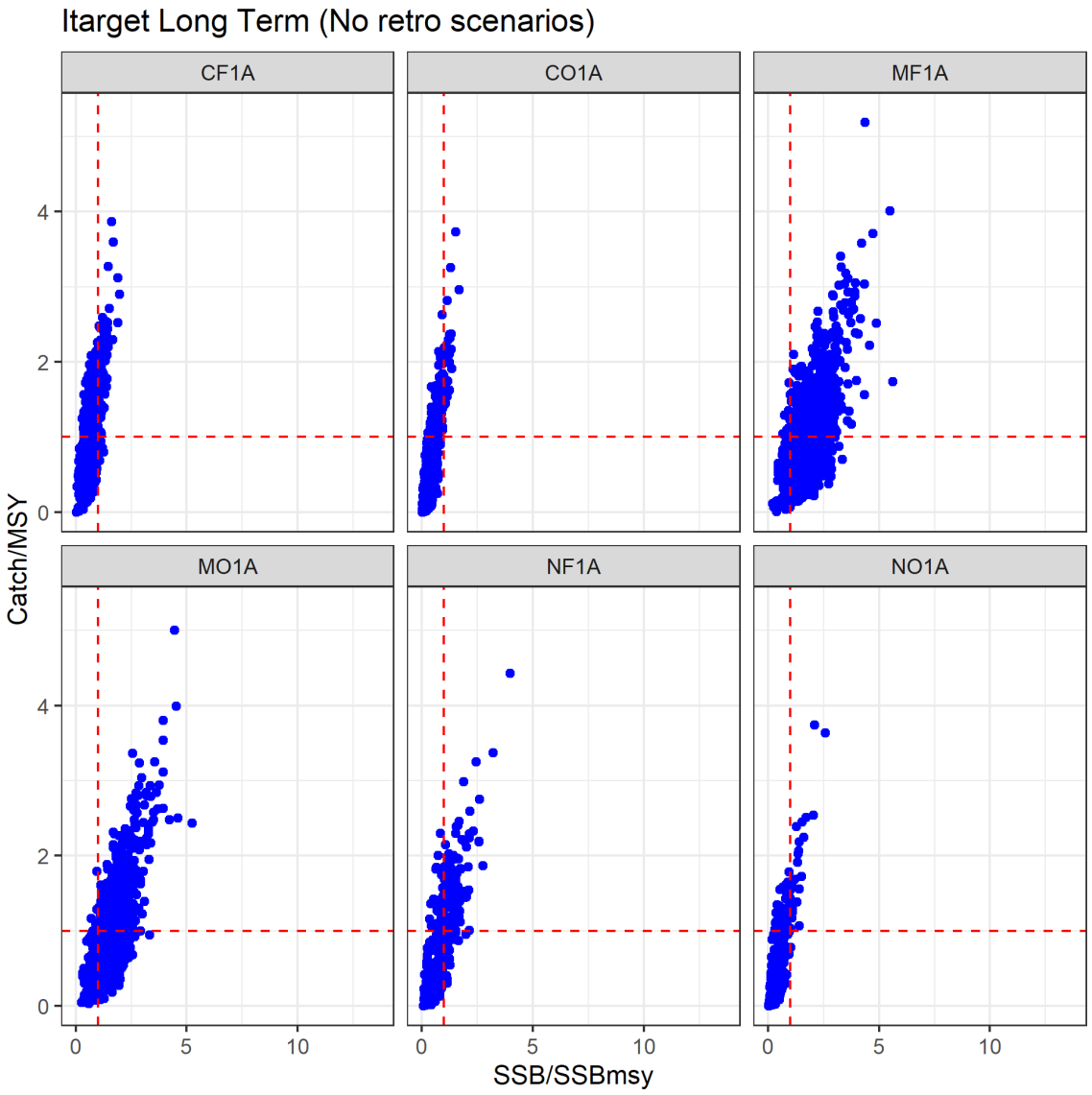


Figure A6.246. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

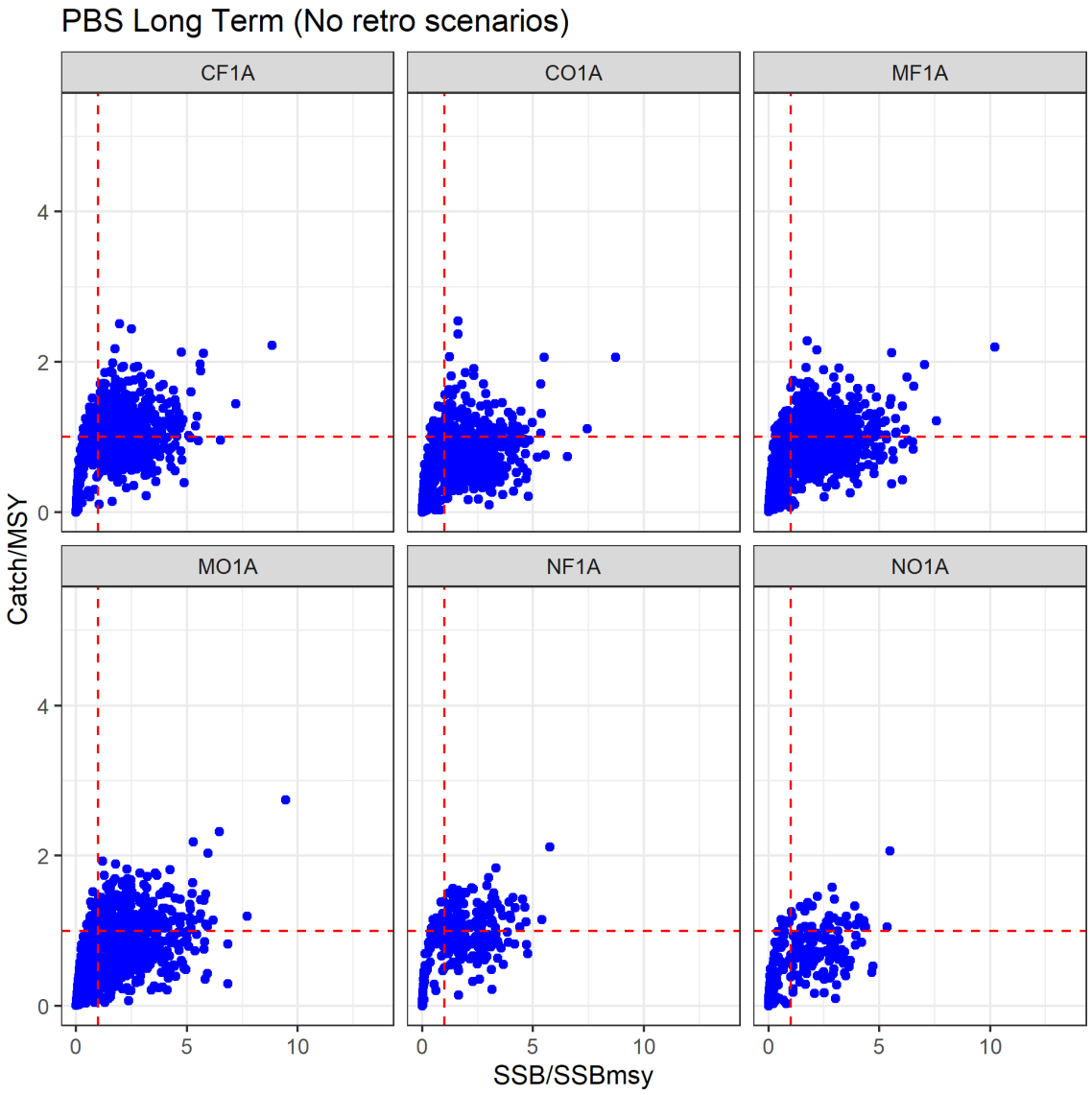


Figure A6.247. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

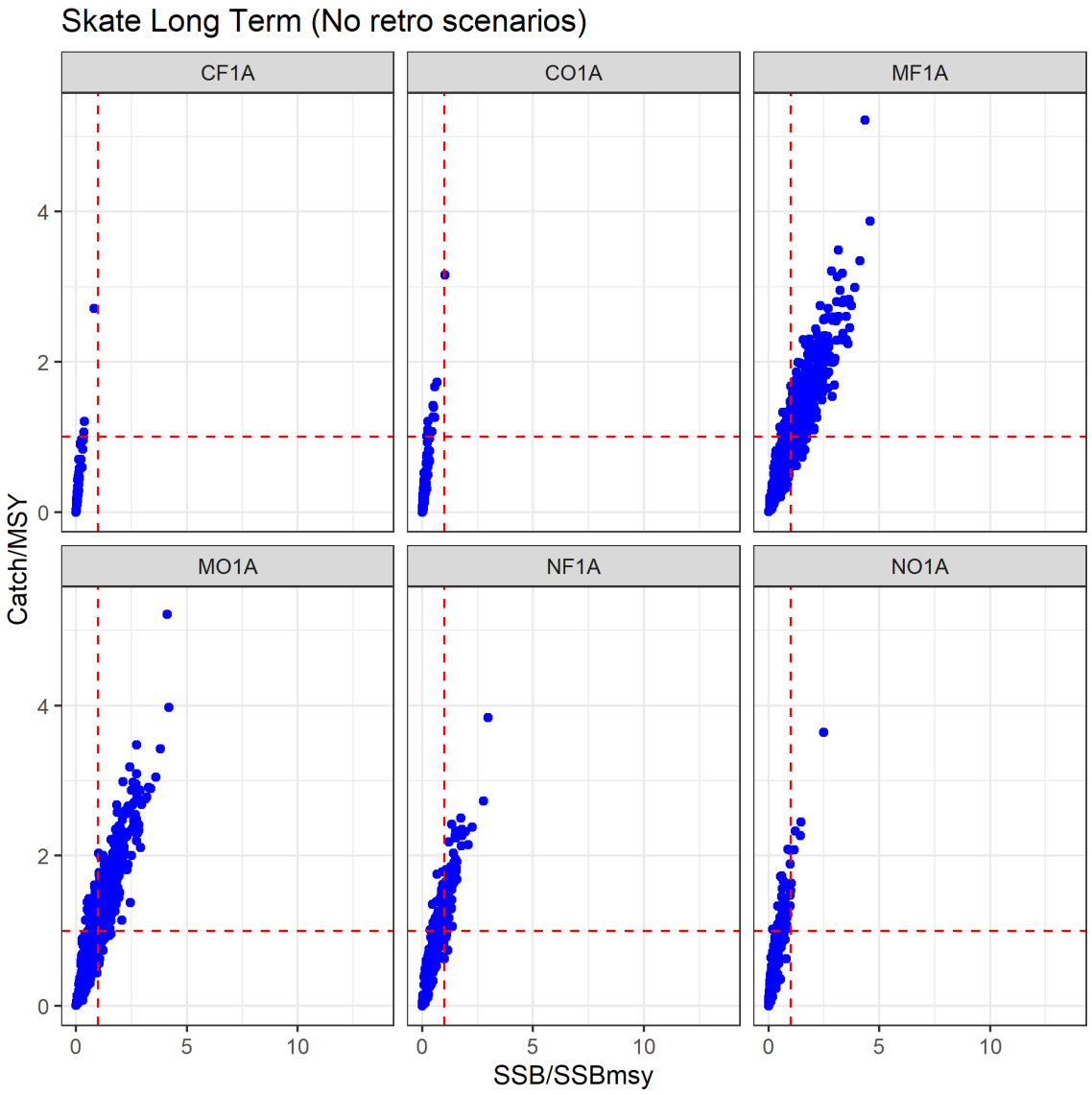


Figure A6.248. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

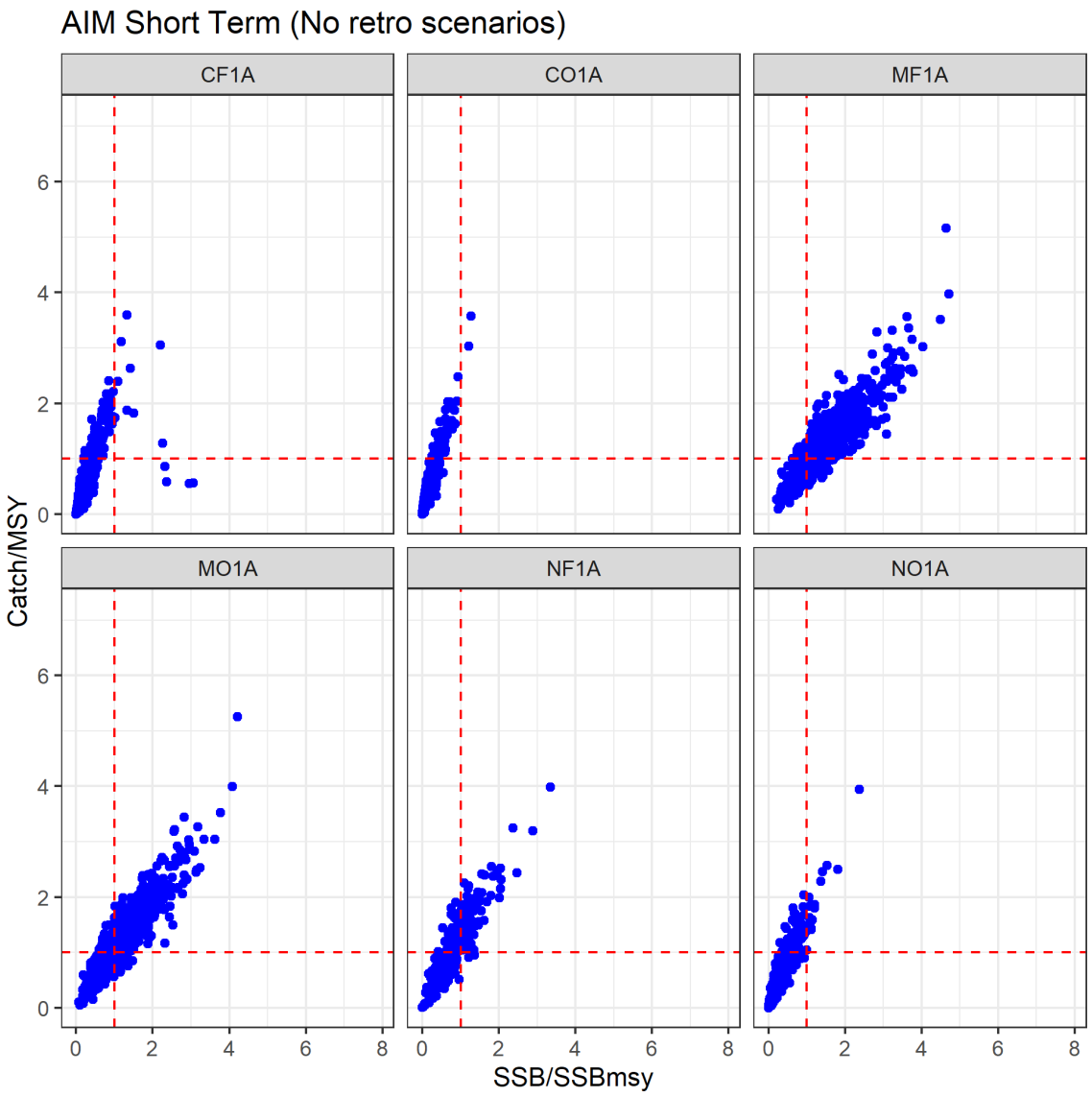


Figure A6.249. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

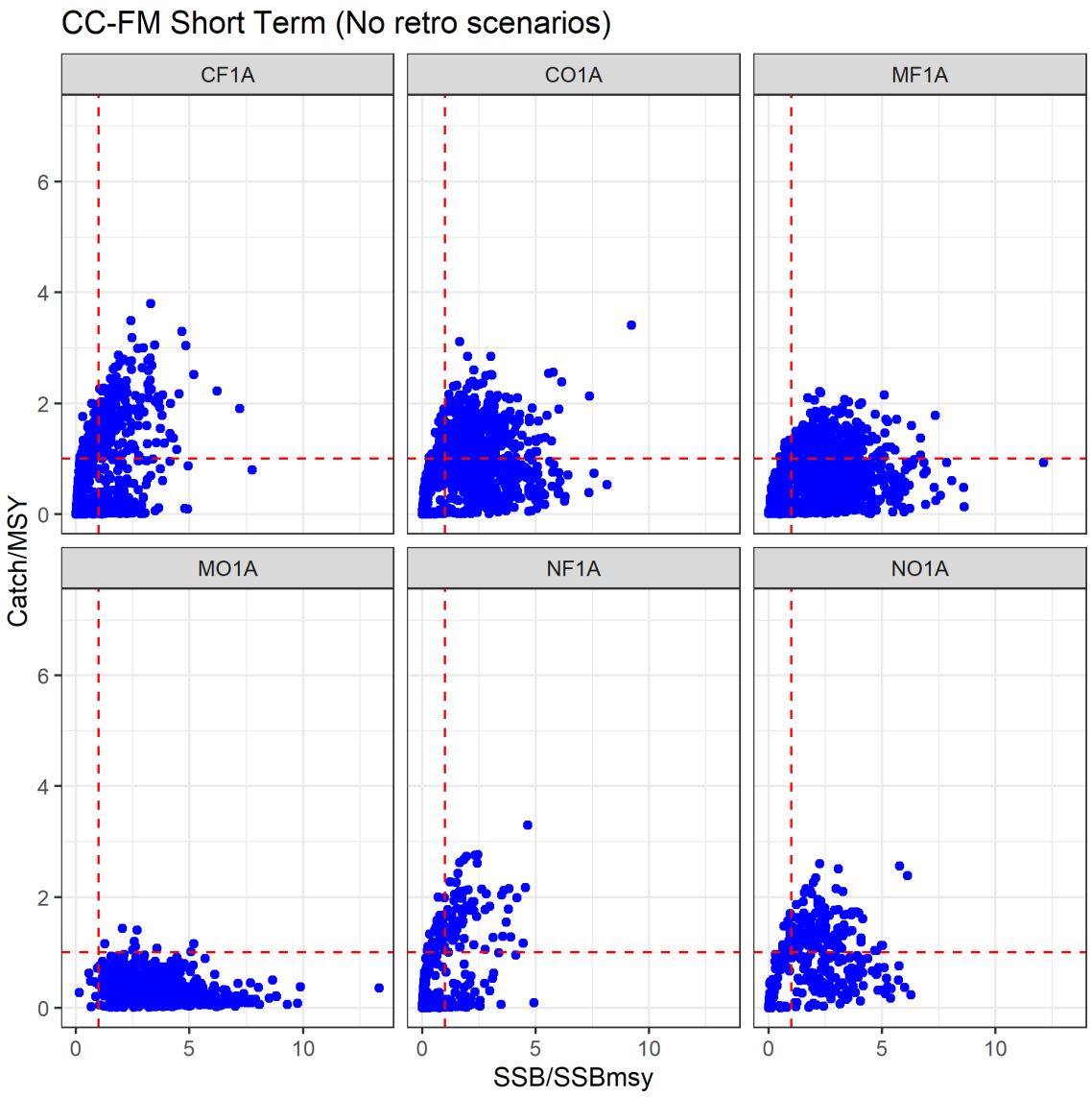


Figure A6.250. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

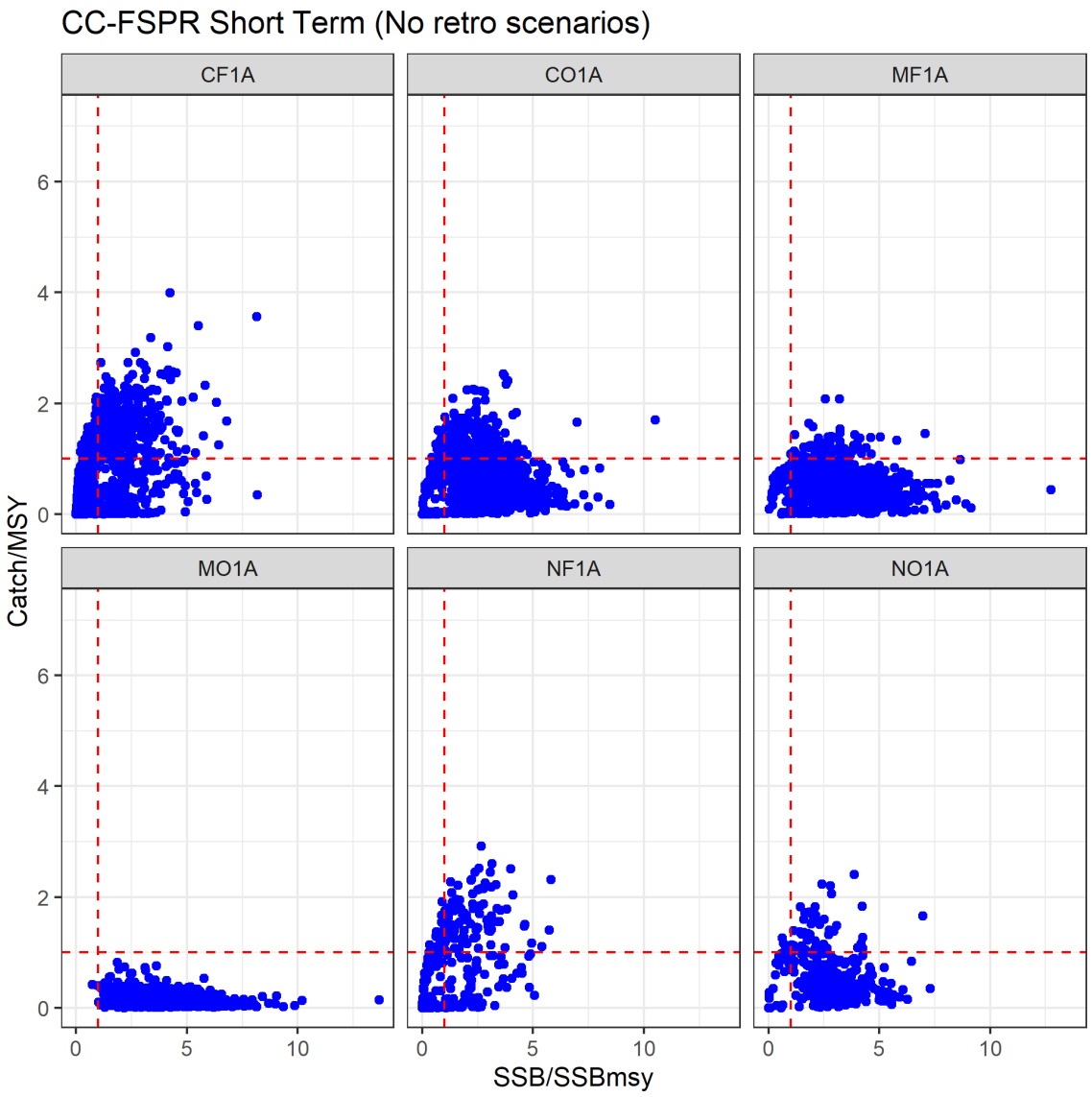


Figure A6.251. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

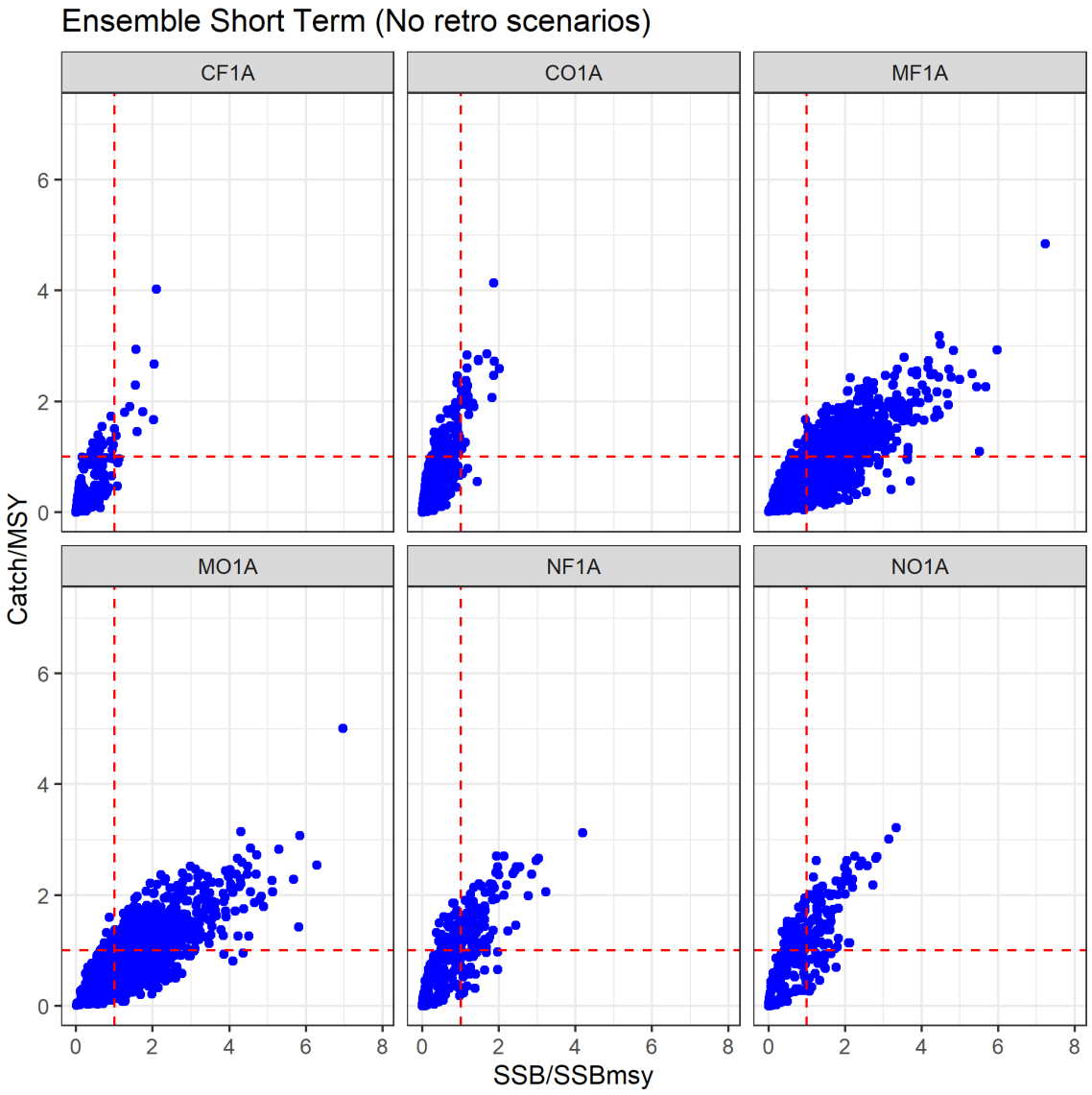


Figure A6.252. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

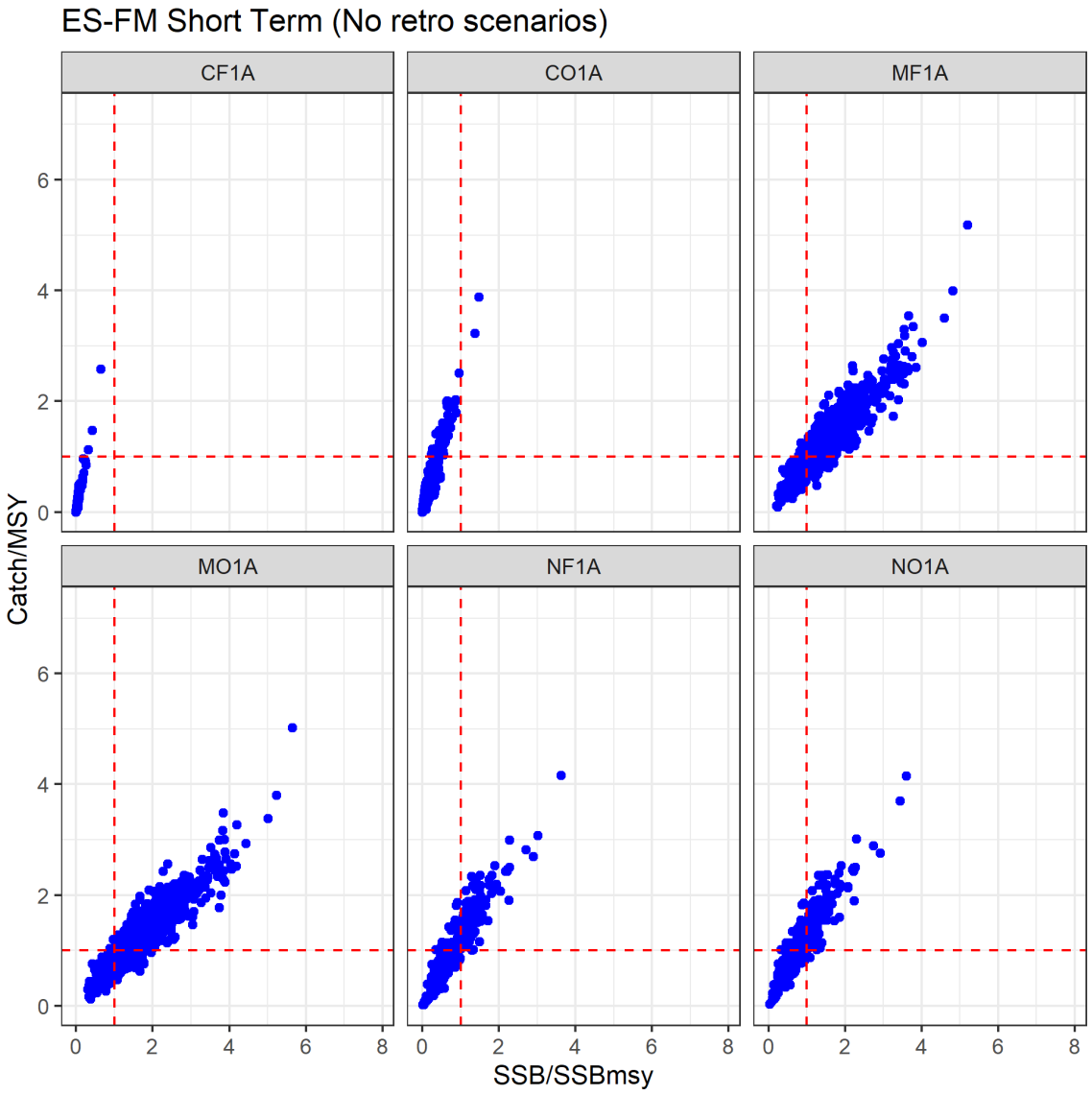


Figure A6.253. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

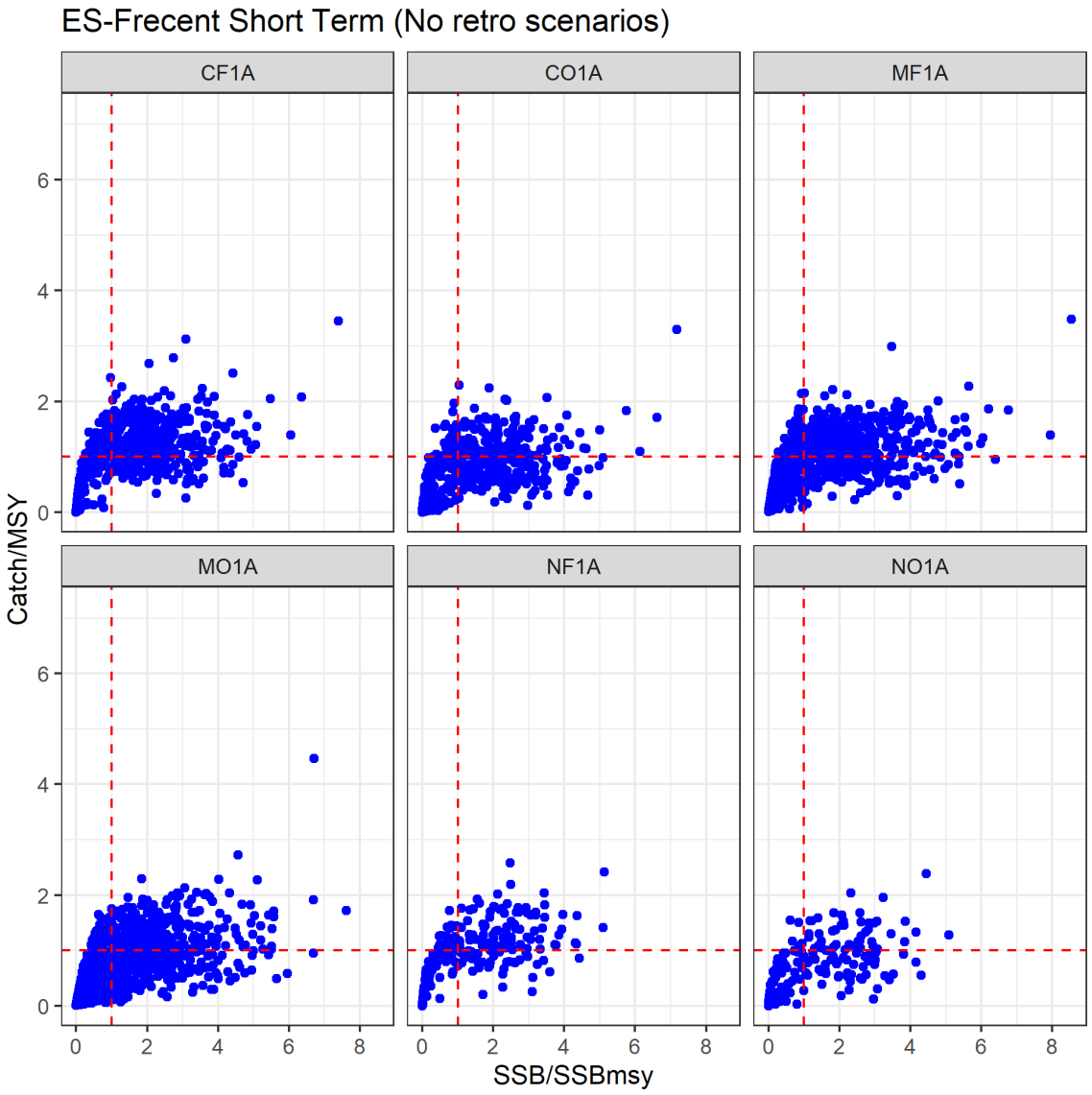


Figure A6.254. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

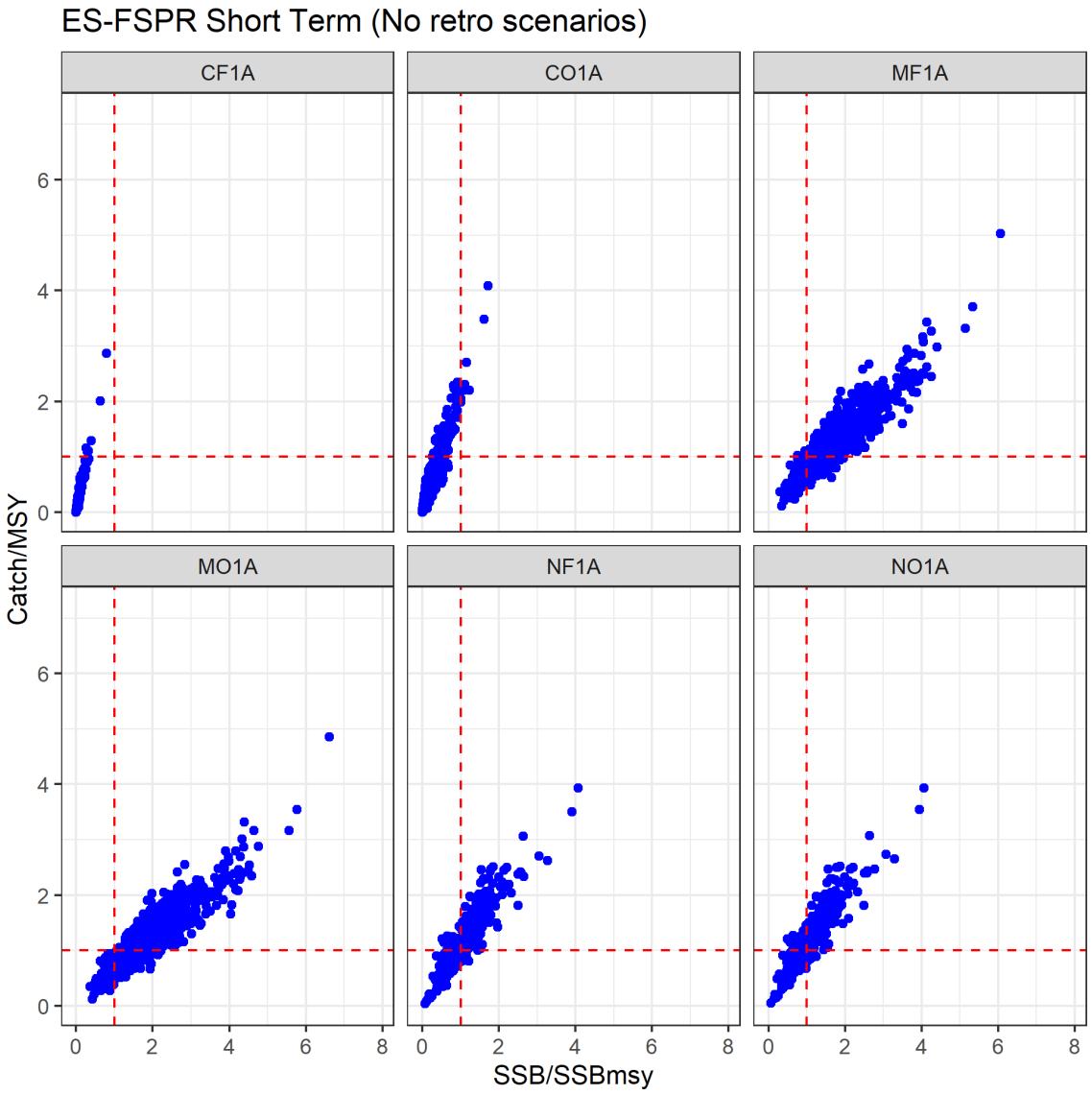


Figure A6.255. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

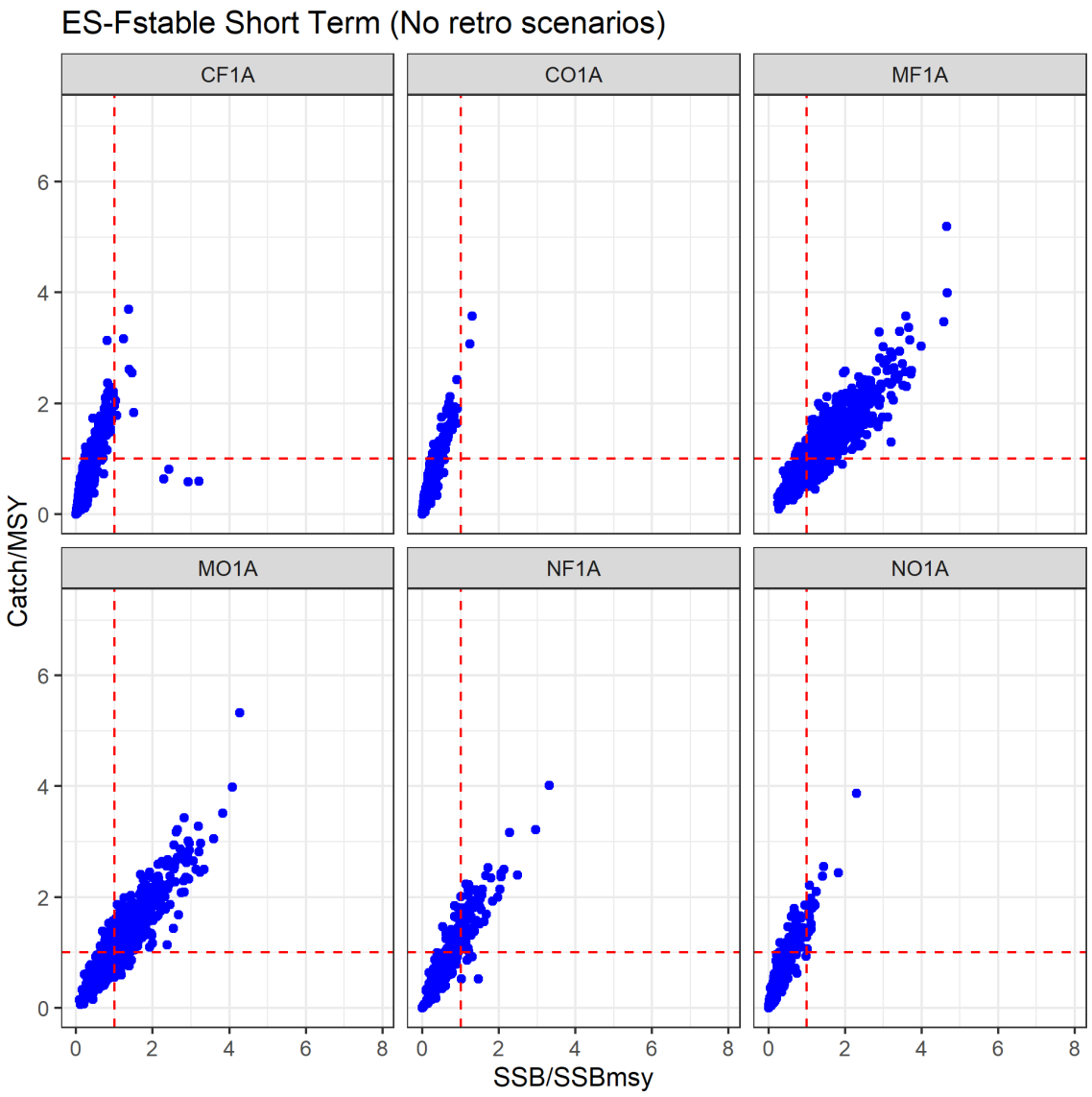


Figure A6.256. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

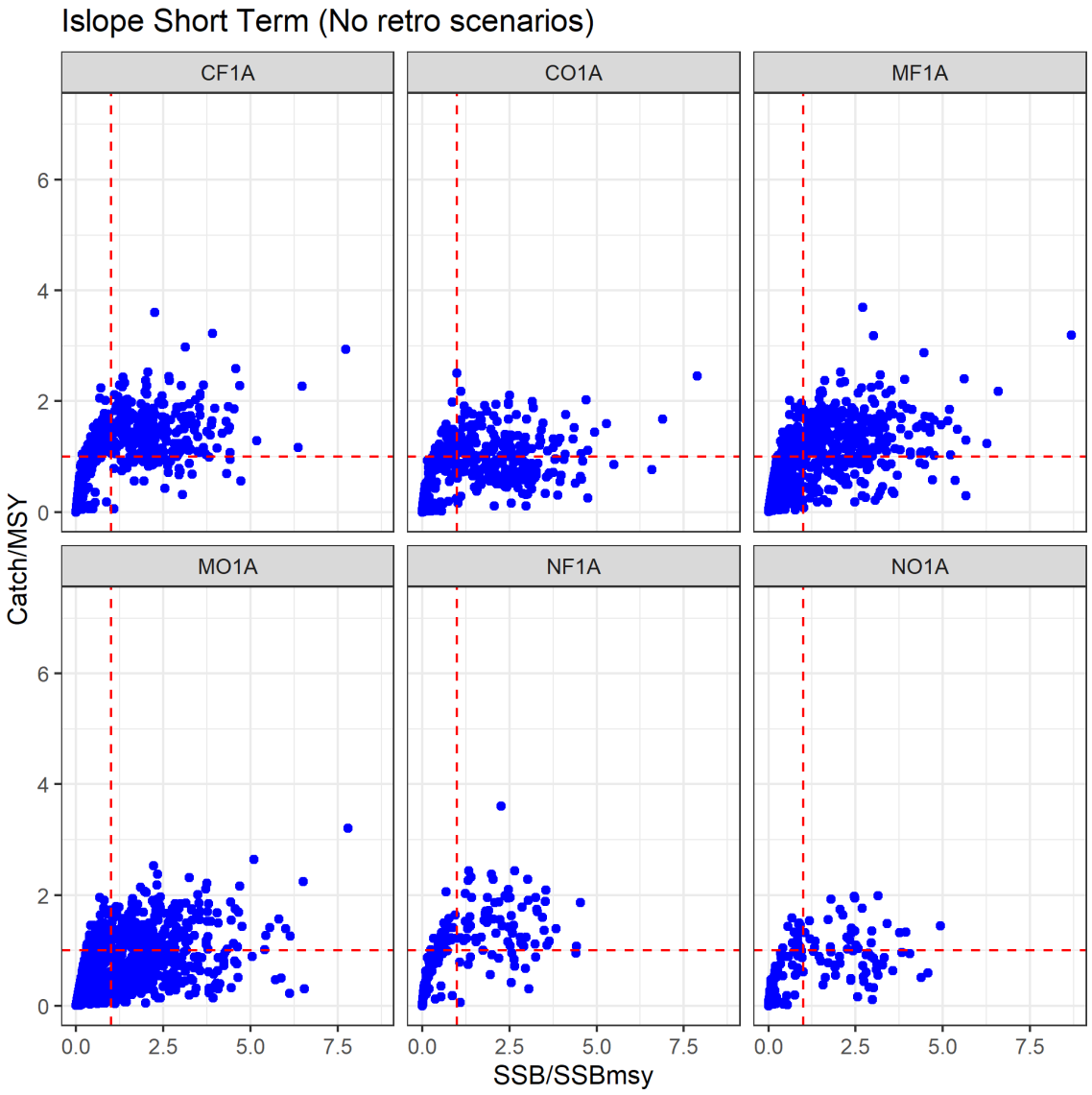


Figure A6.257. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

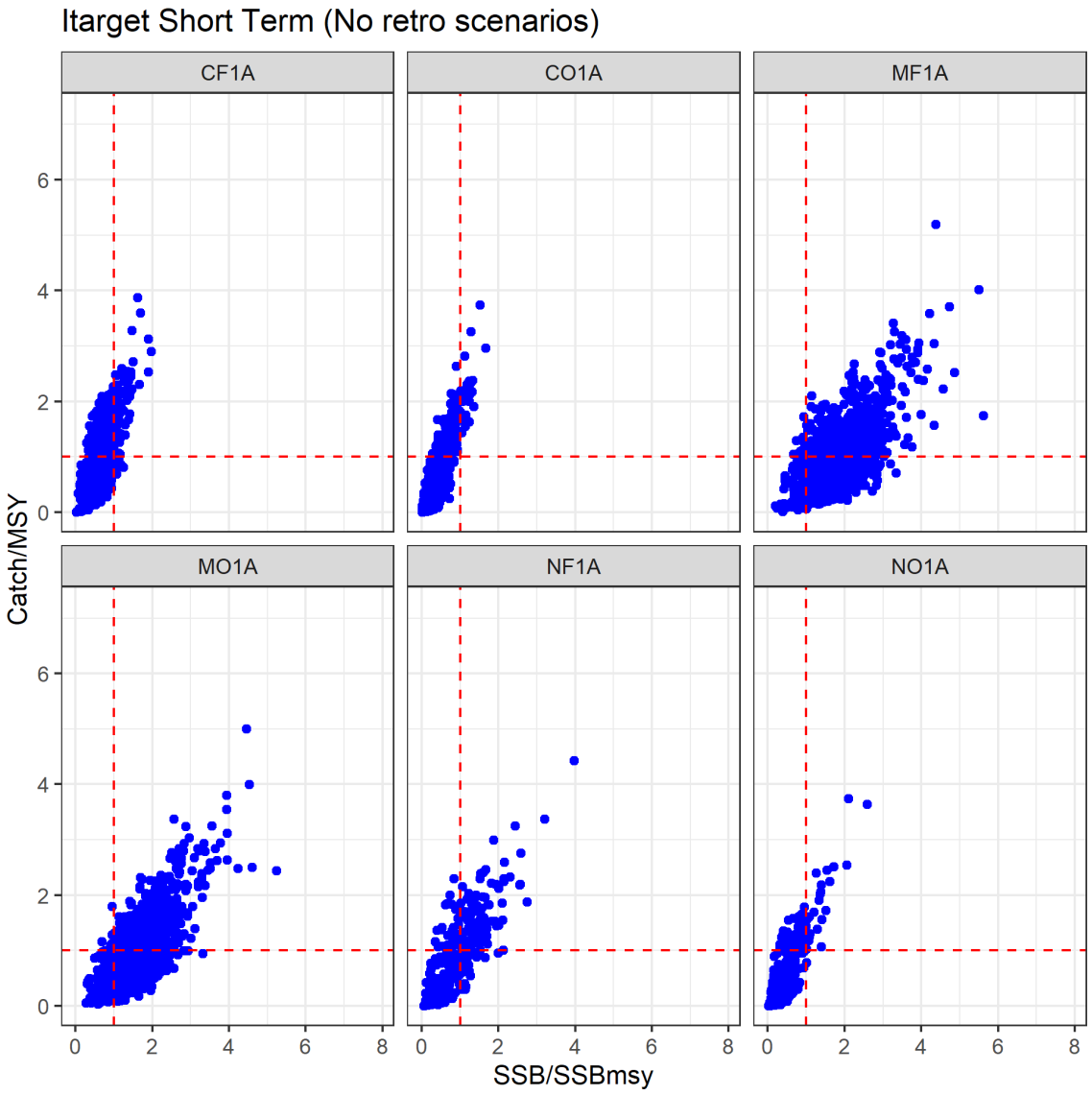


Figure A6.258. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

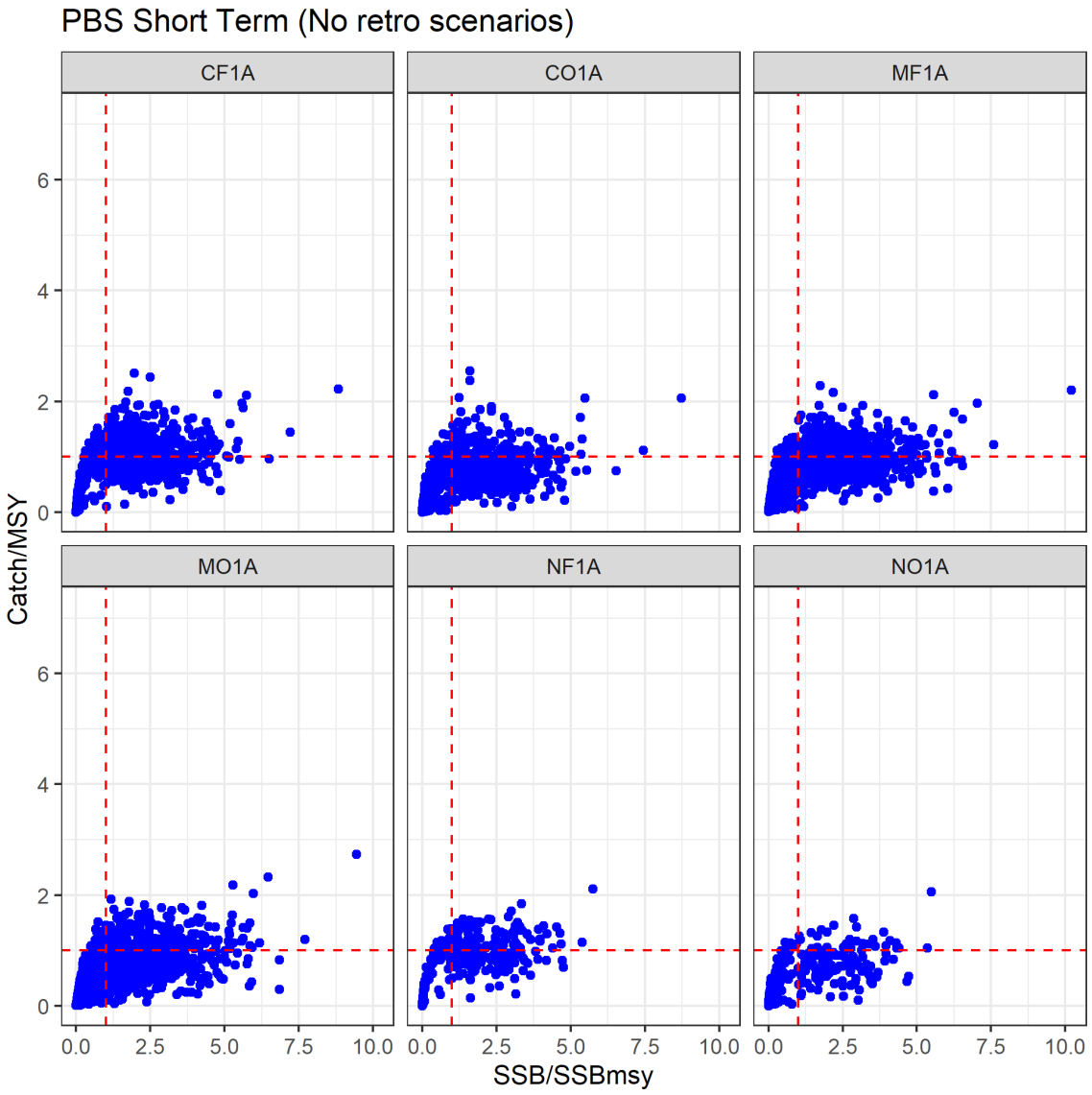


Figure A6.259. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

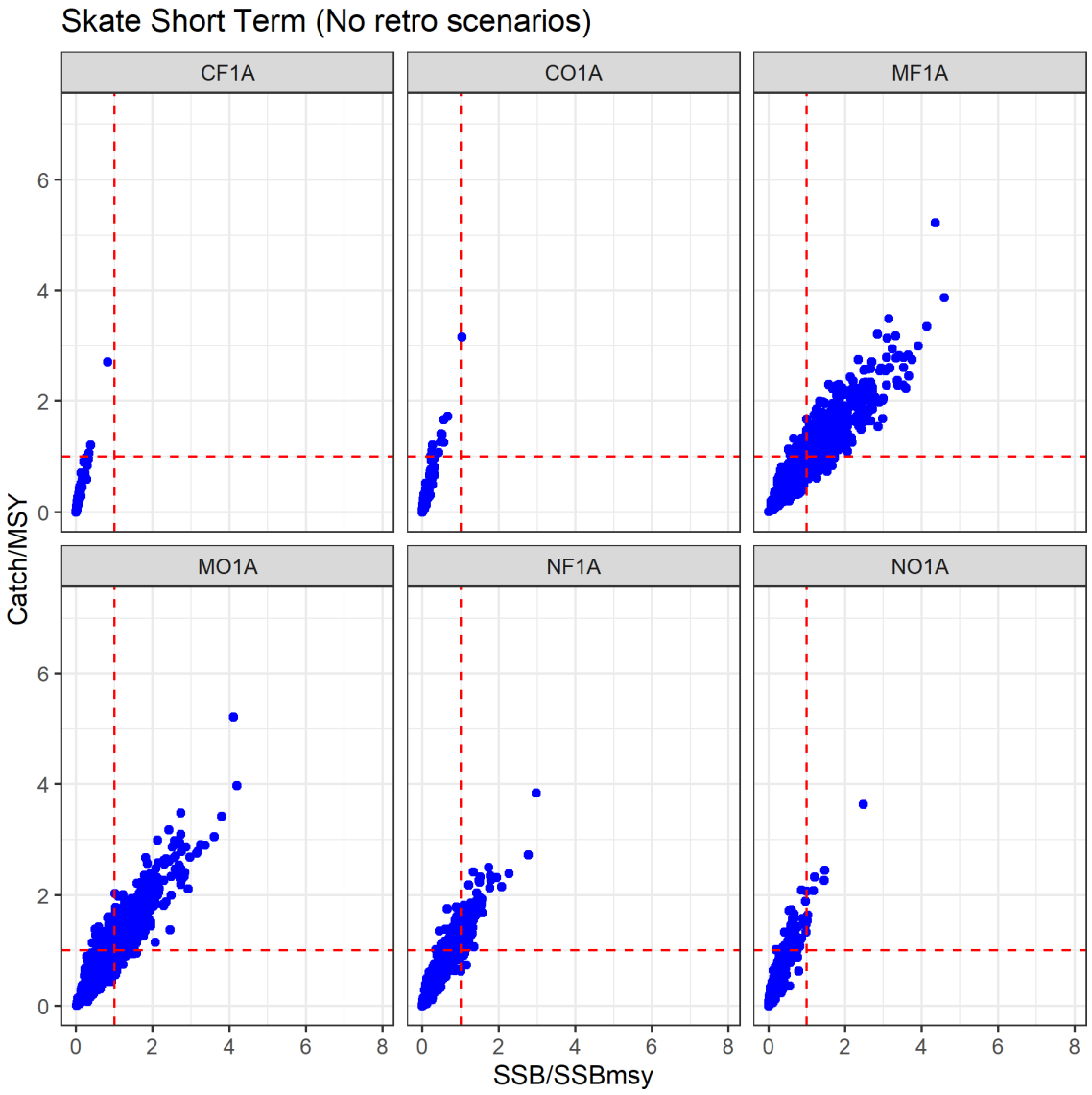


Figure A6.260. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

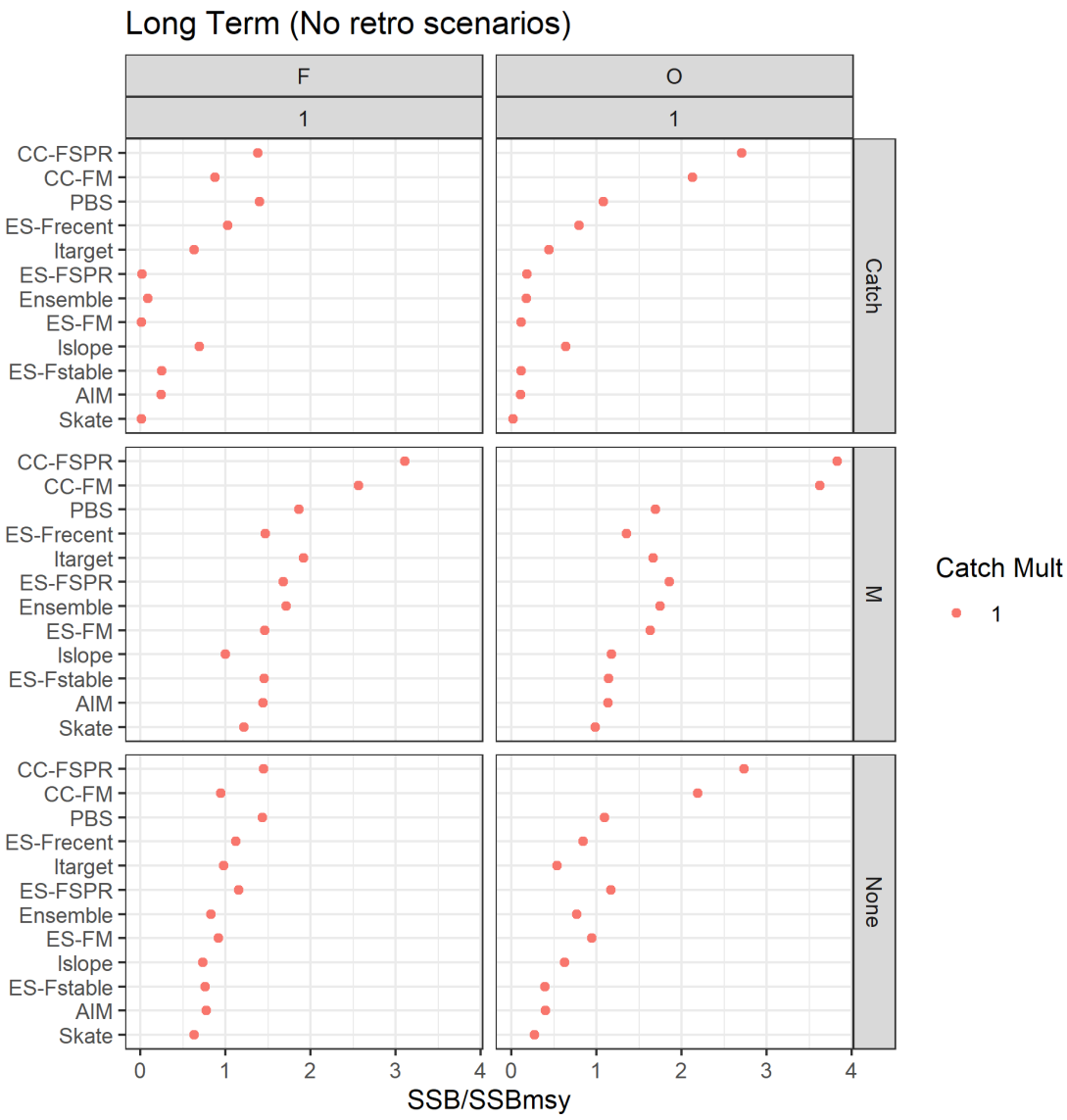


Figure A6.261. Mean values of spawning stock biomass (SSB) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

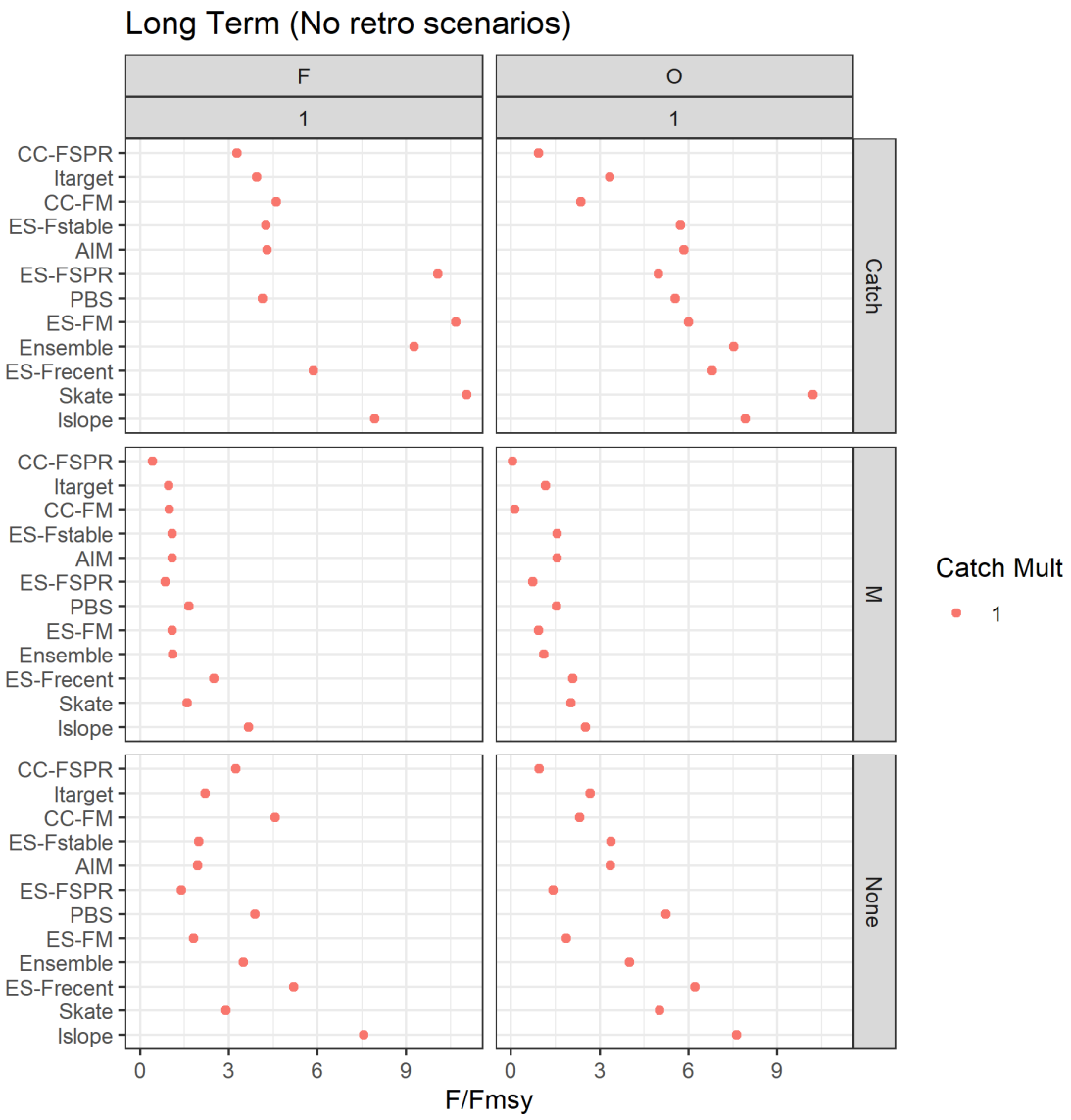


Figure A6.262. Mean values of fishing mortality rate (F) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of F history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

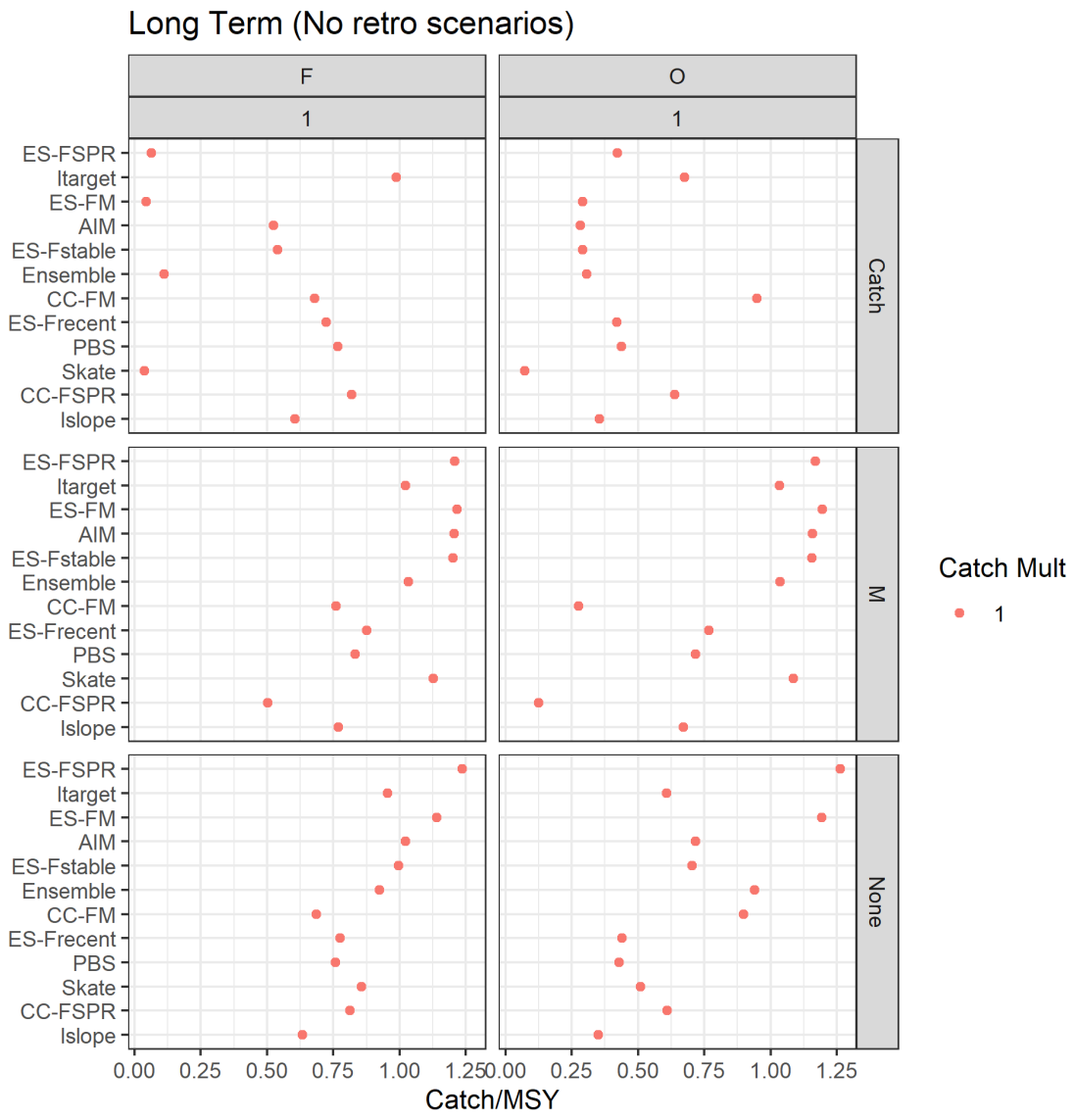


Figure A6.263. Mean values of catch relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

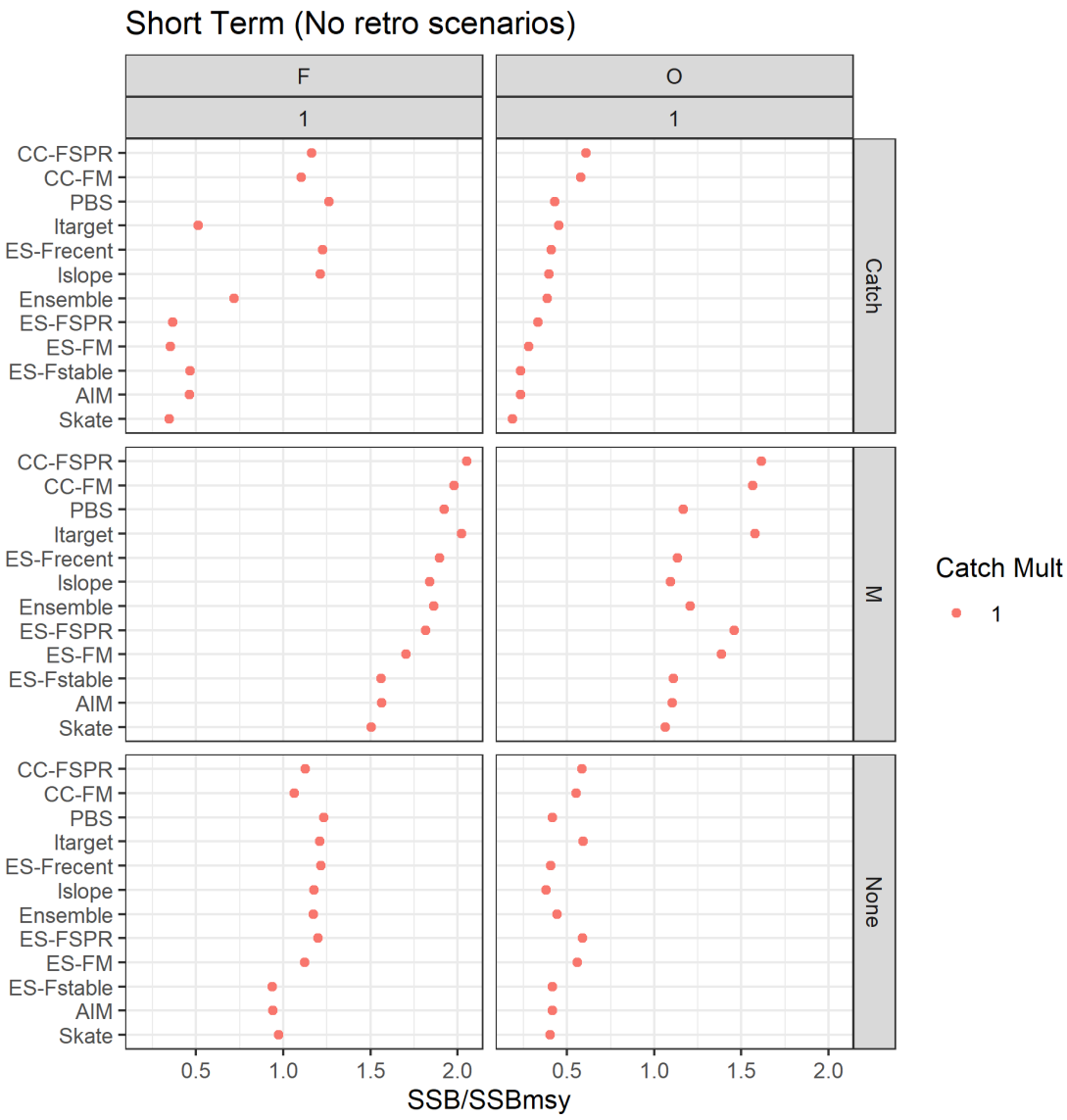


Figure A6.264. Mean values of spawning stock biomass (SSB) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

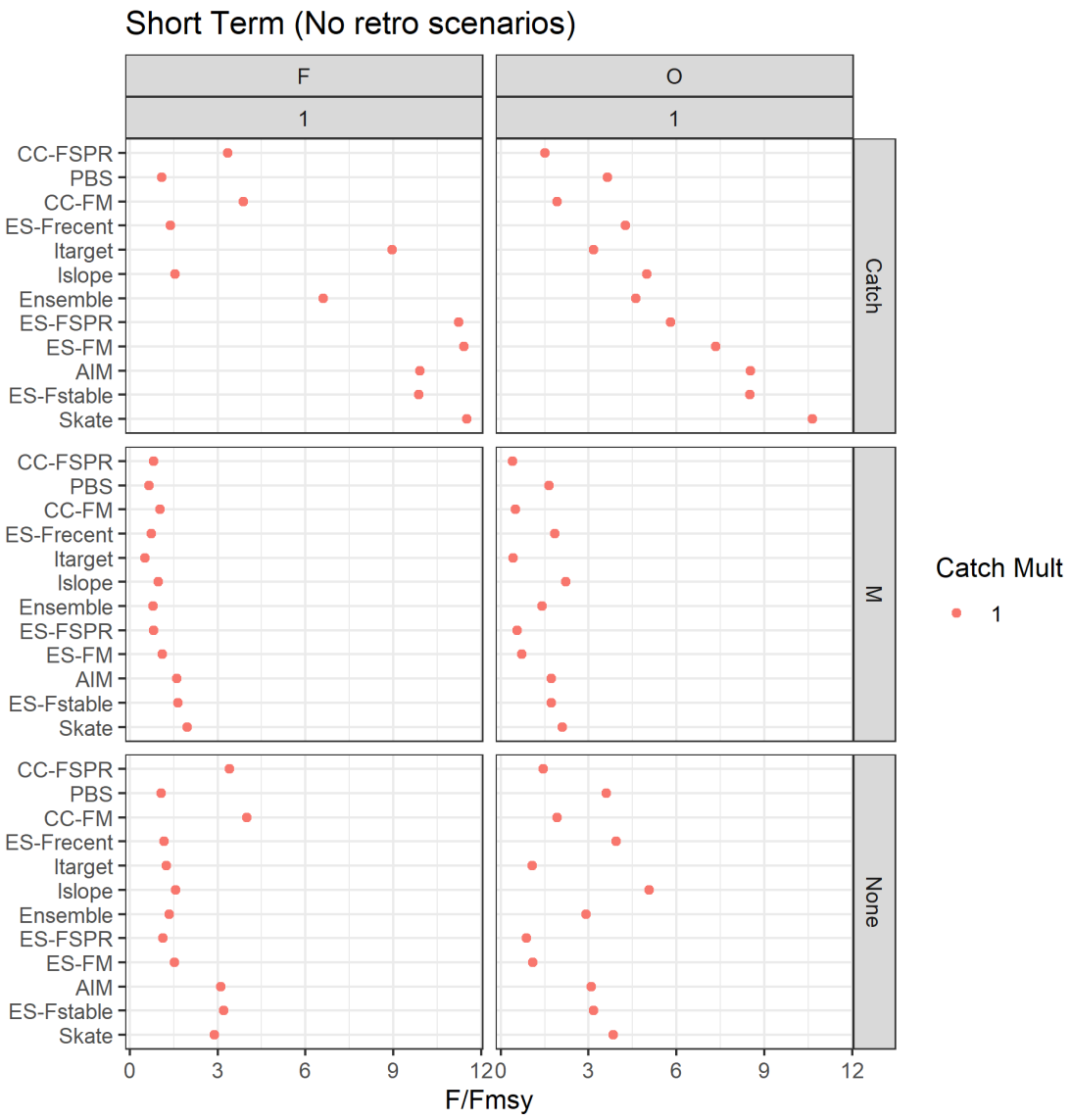


Figure A6.265. Mean values of fishing mortality rate (F) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of F history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

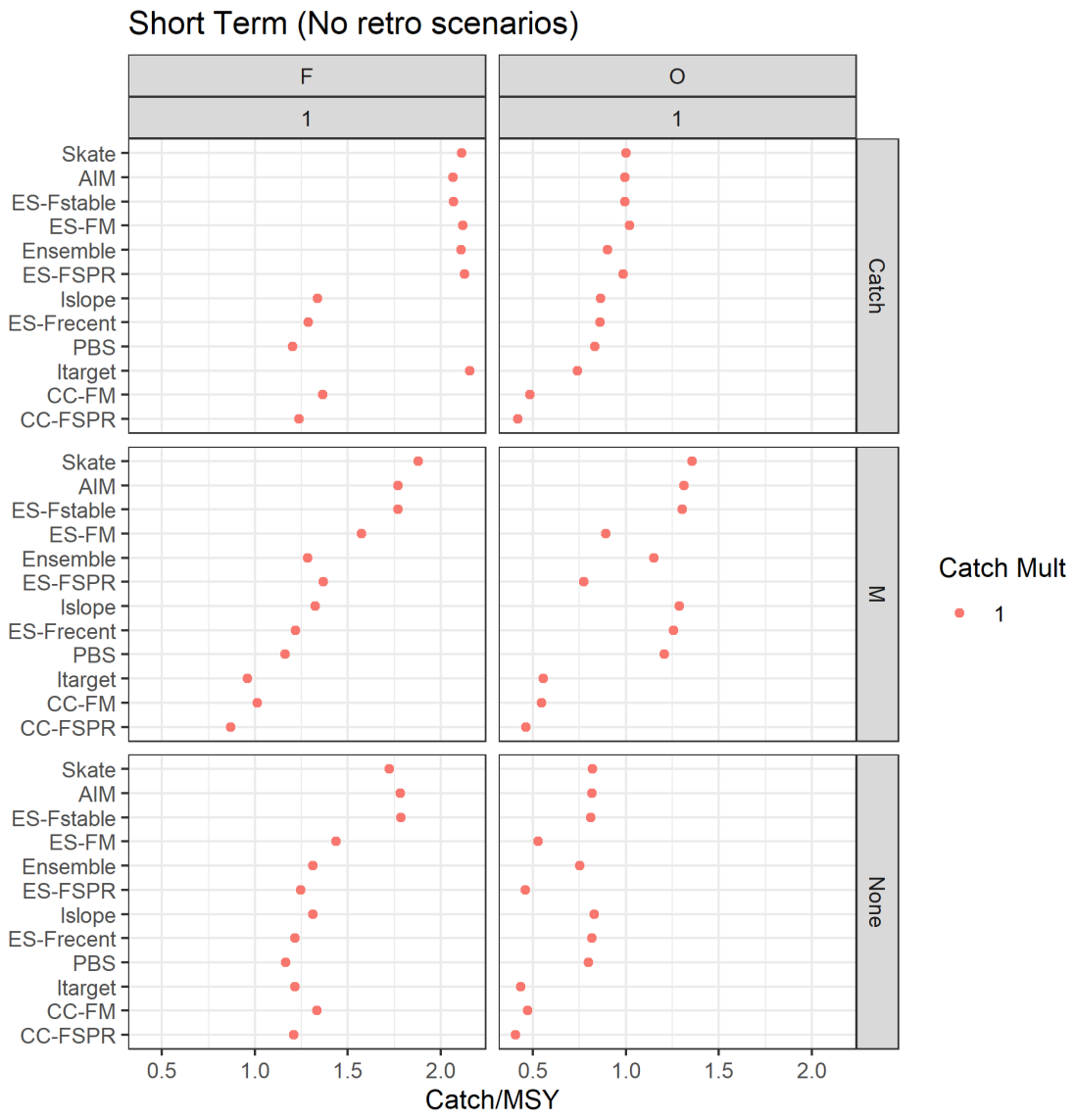


Figure A6.266. Mean values of catch relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

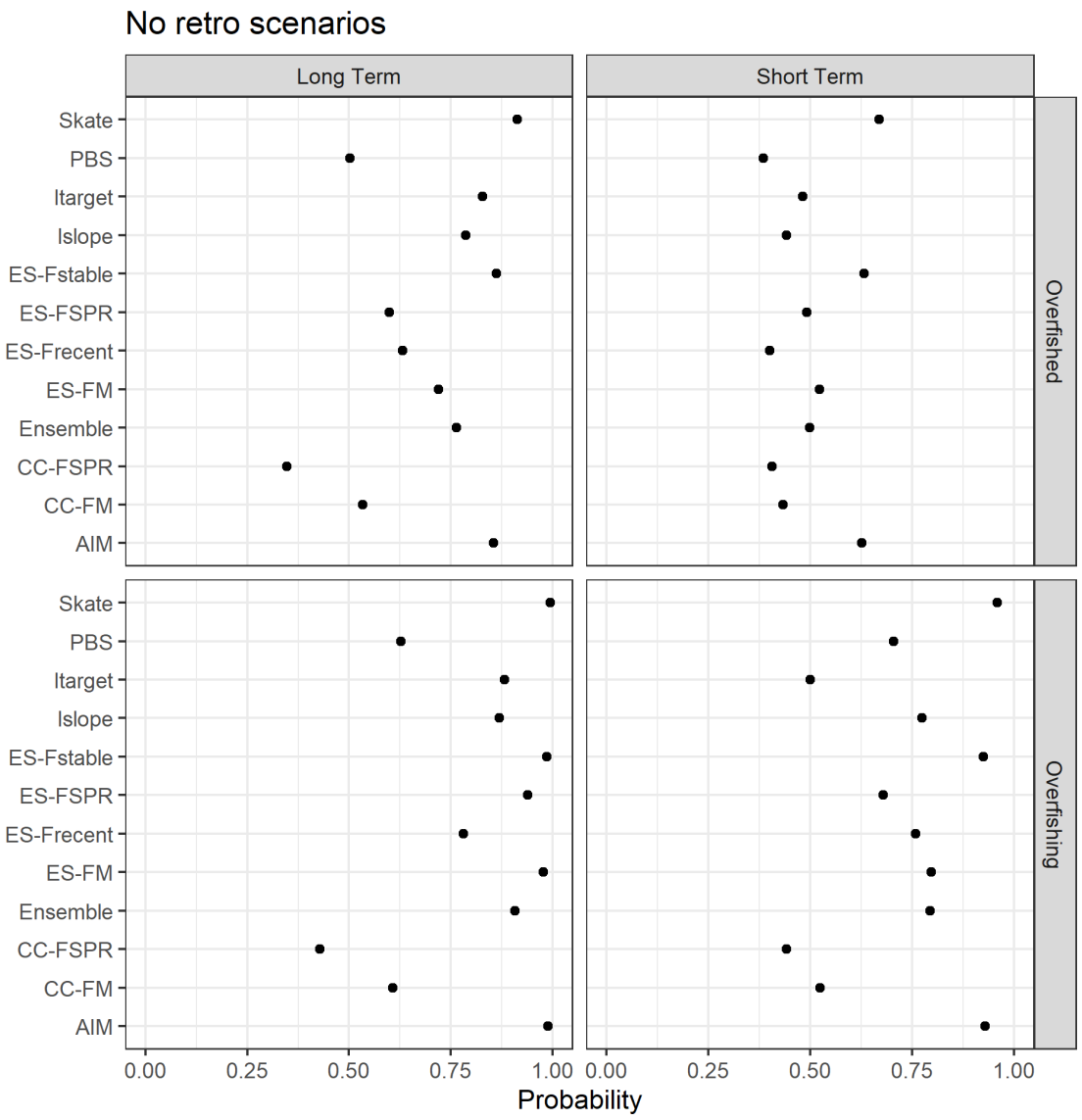


Figure A6.267. Probability stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all no retro analyses.

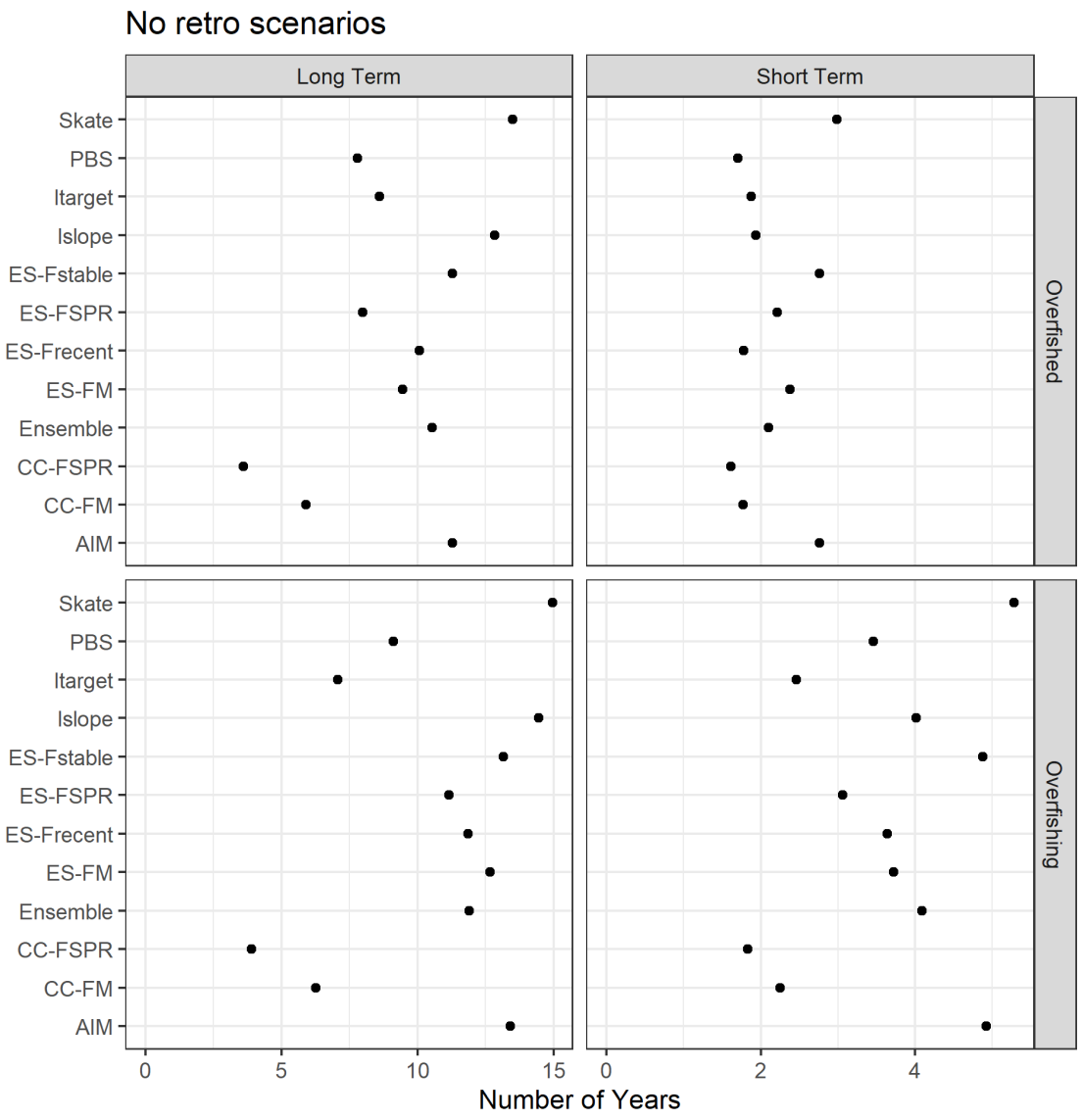


Figure A6.268. Number of years stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all no retro analyses.

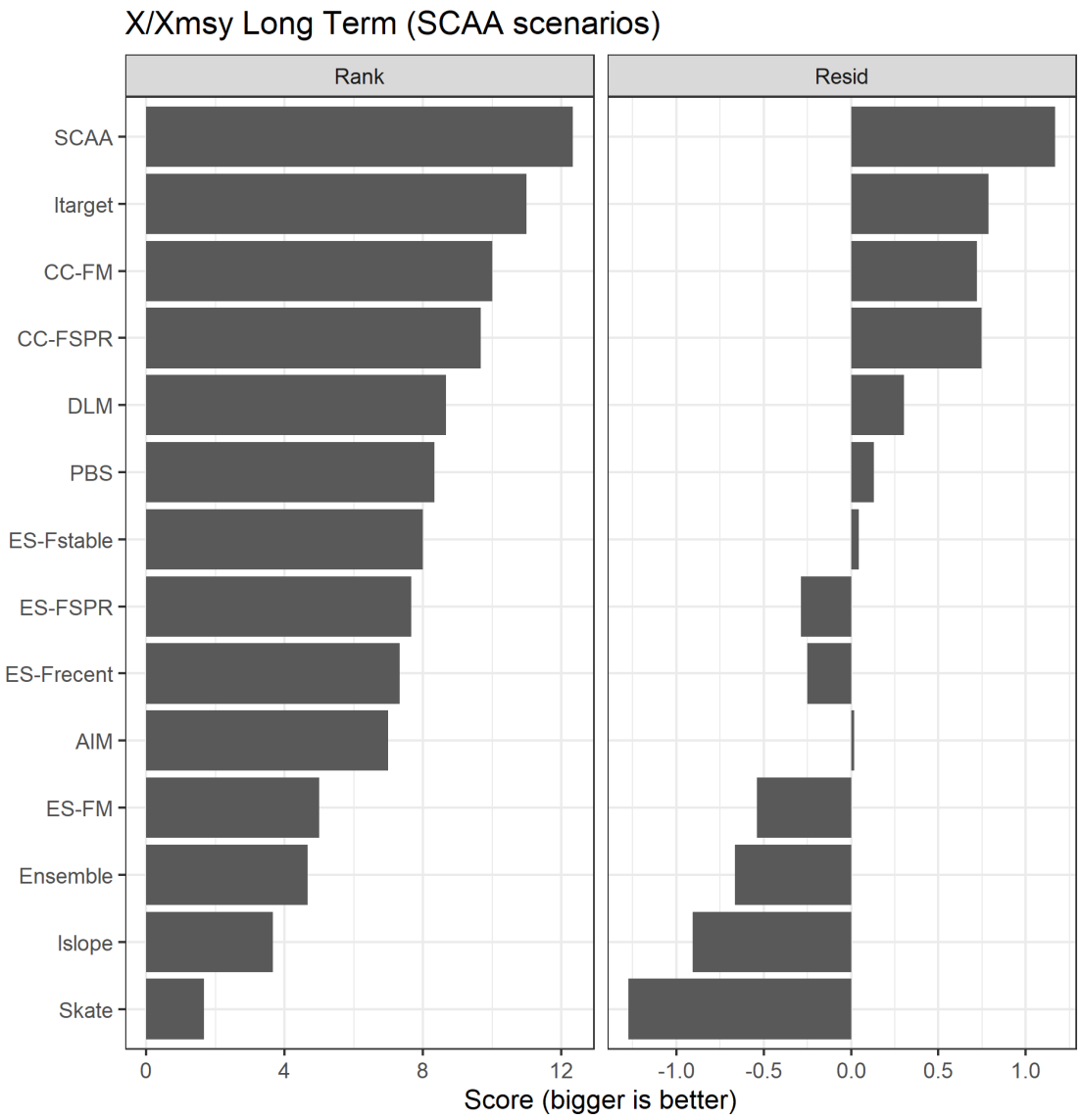


Figure A6.269. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in the long term.

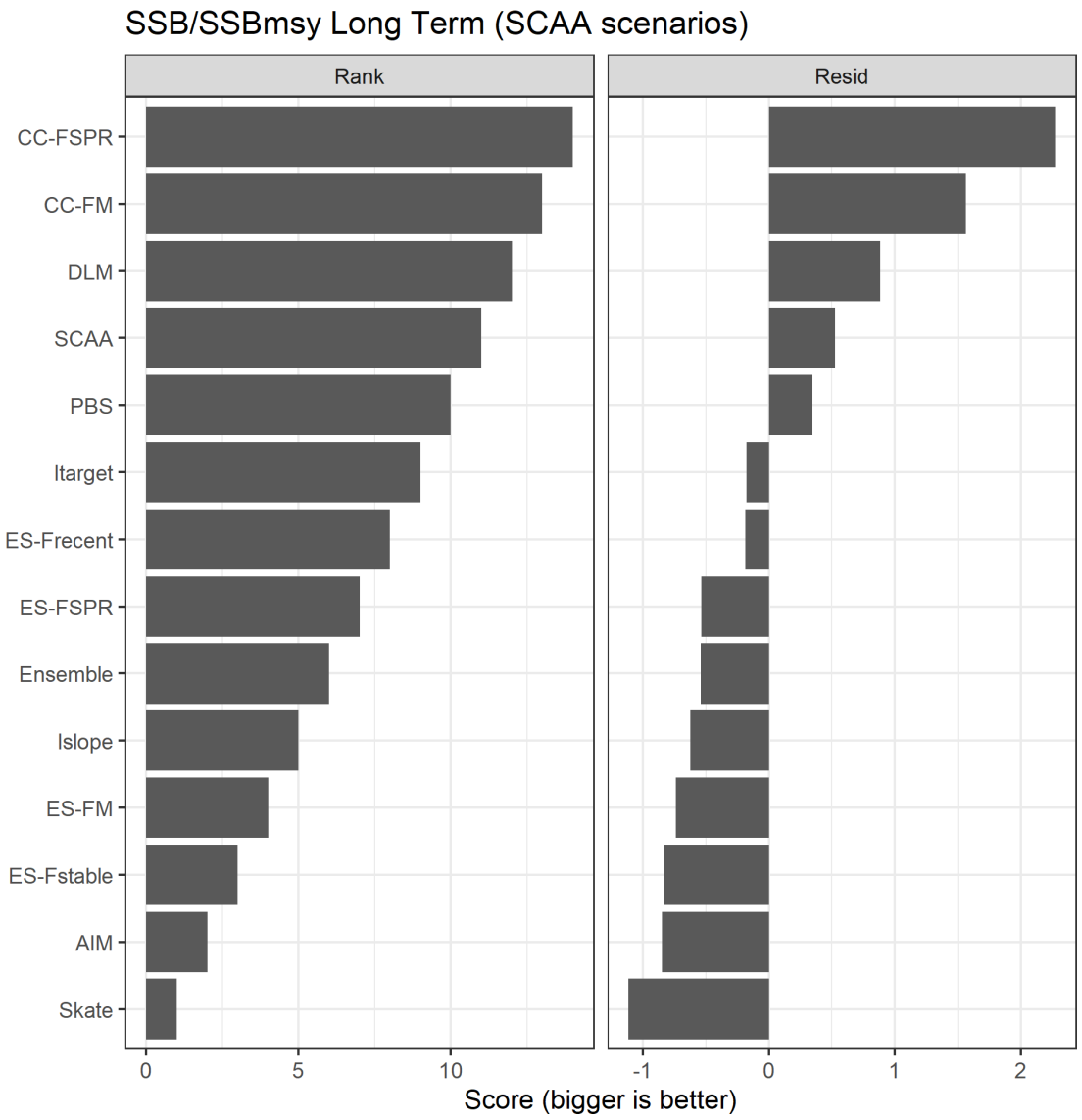


Figure A6.270. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in the long term.

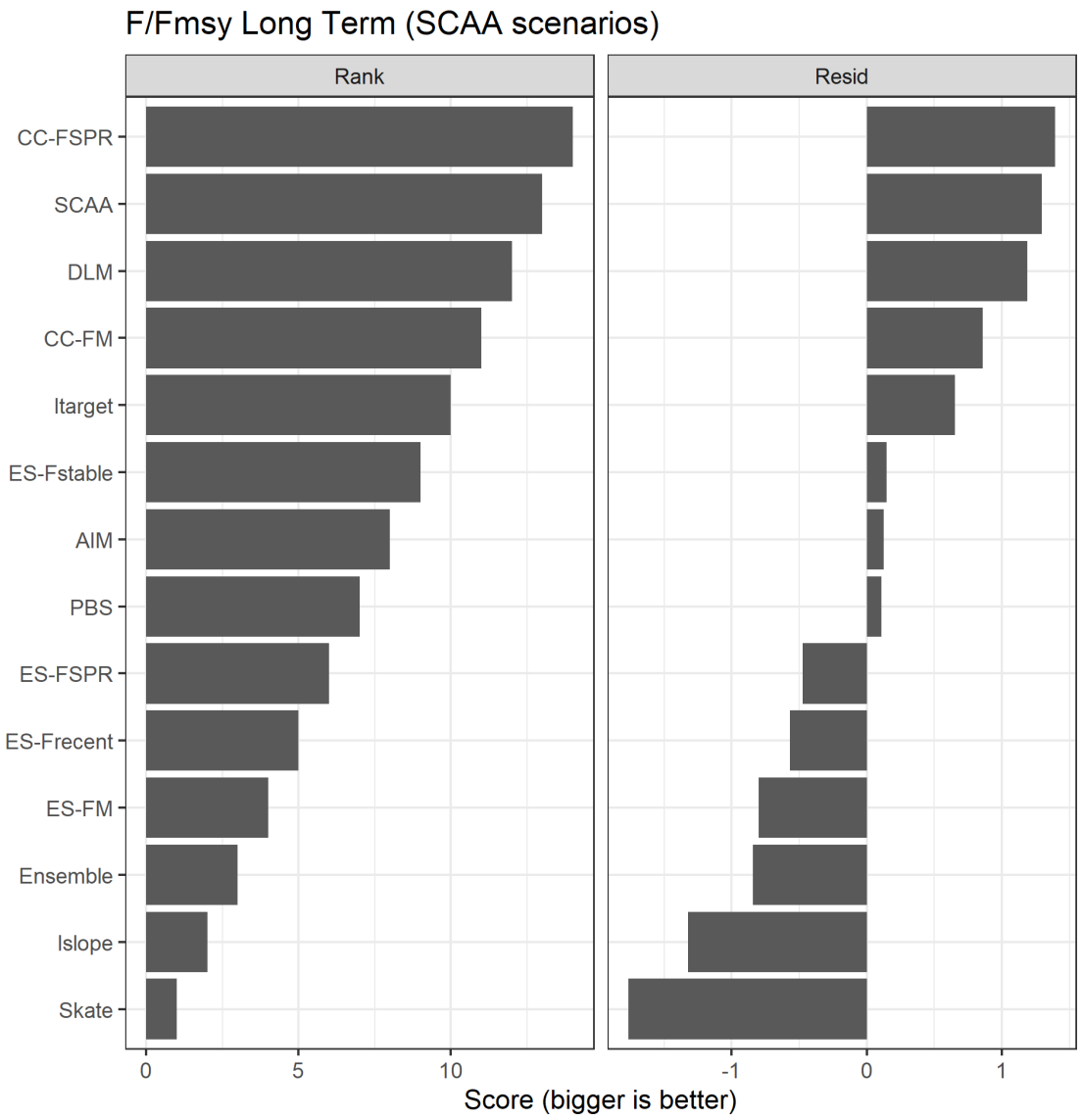


Figure A6.271. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in the long term.

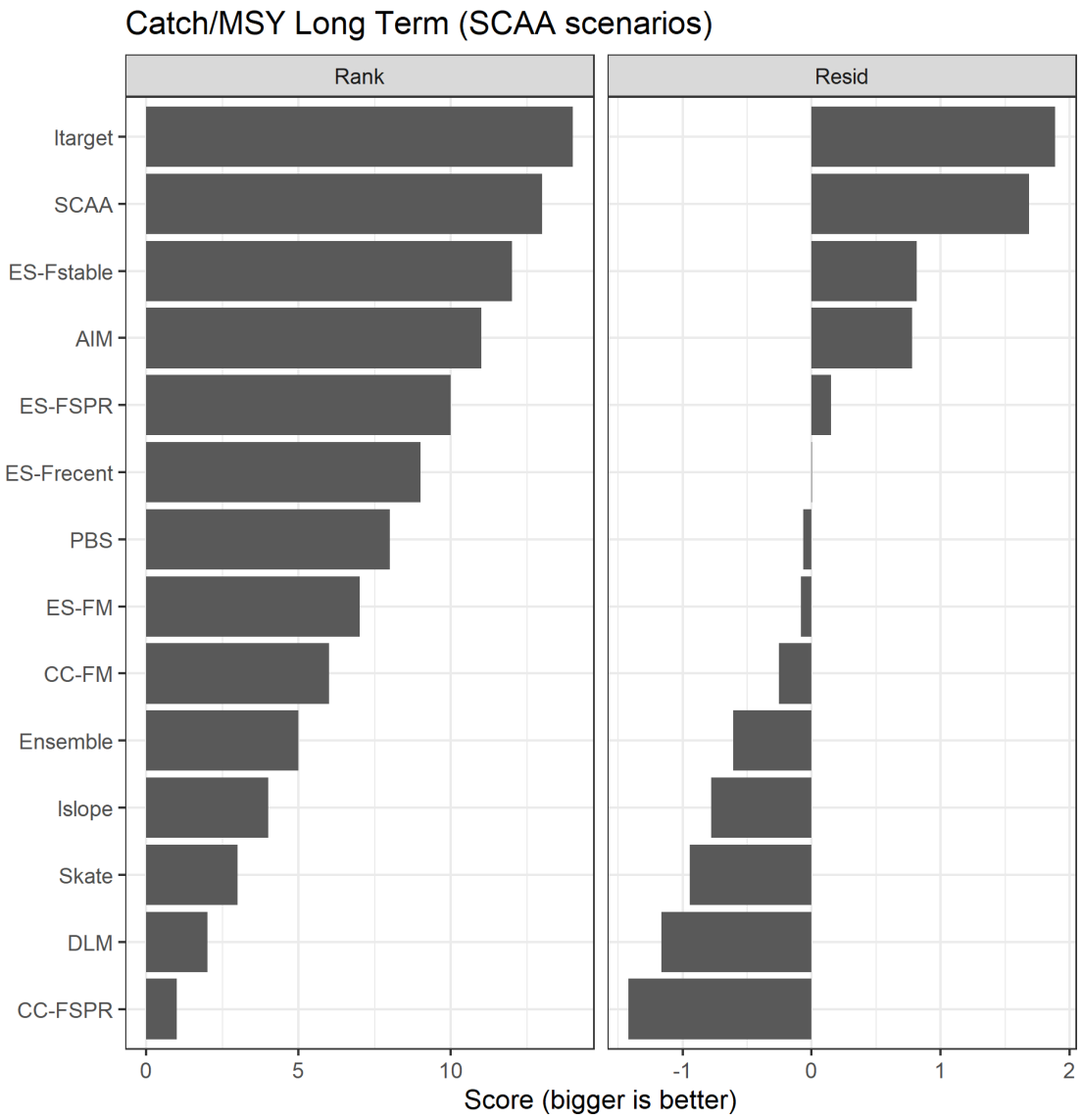


Figure A6.272. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the catch relative to maximum sustainable yield (MSY) in the long term.

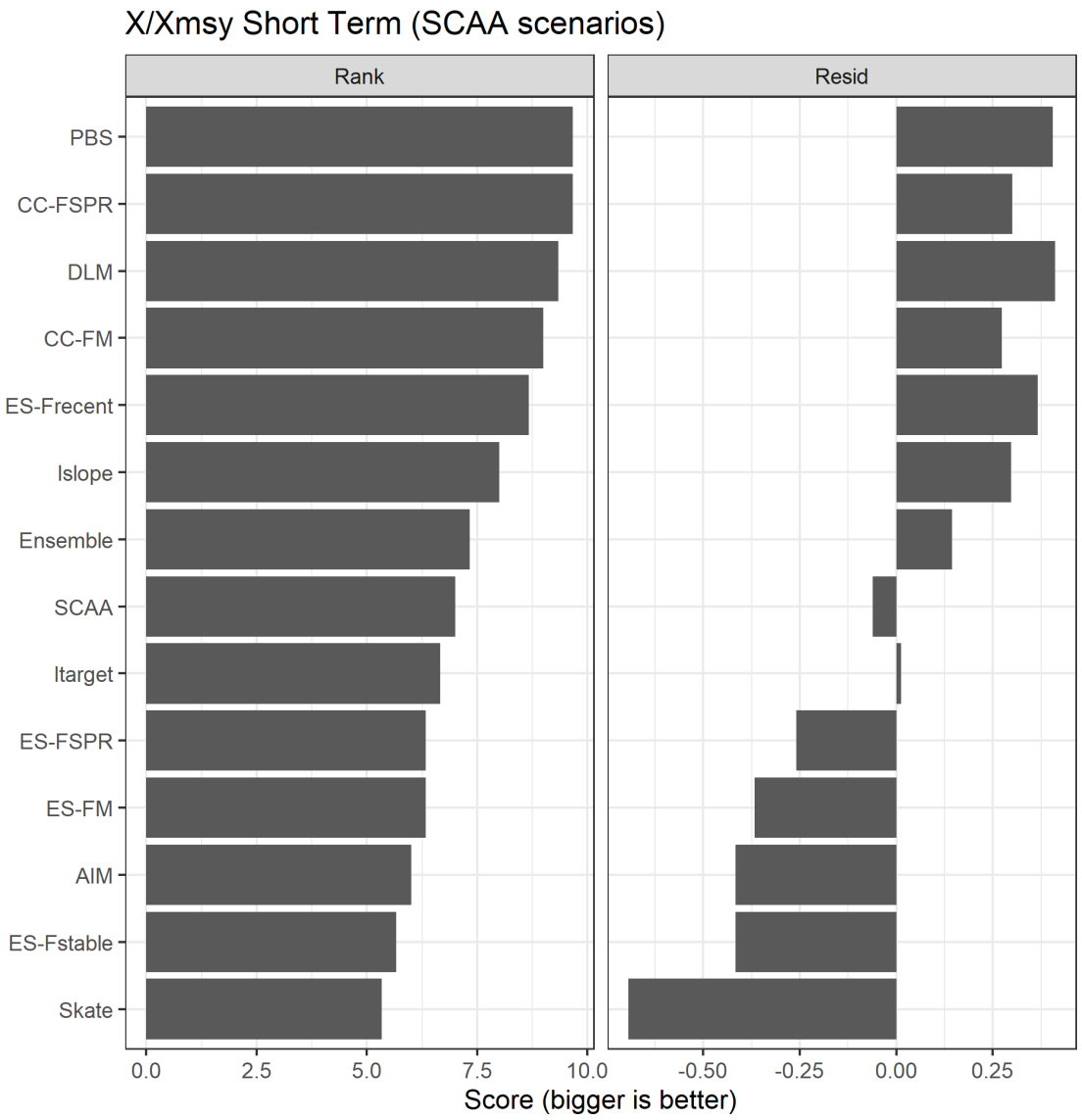


Figure A6.273. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in the short term.

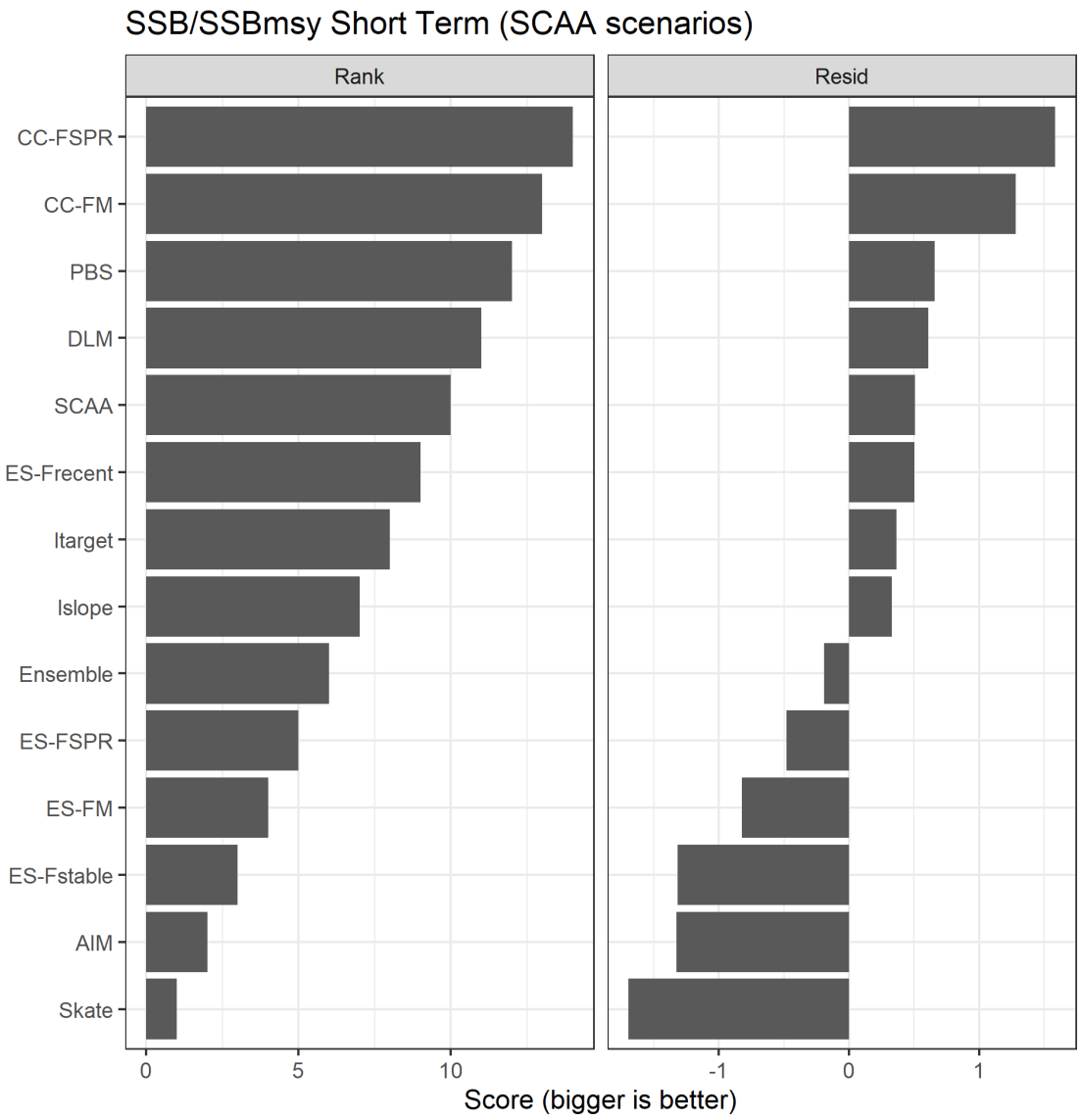


Figure A6.274. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in the short term.

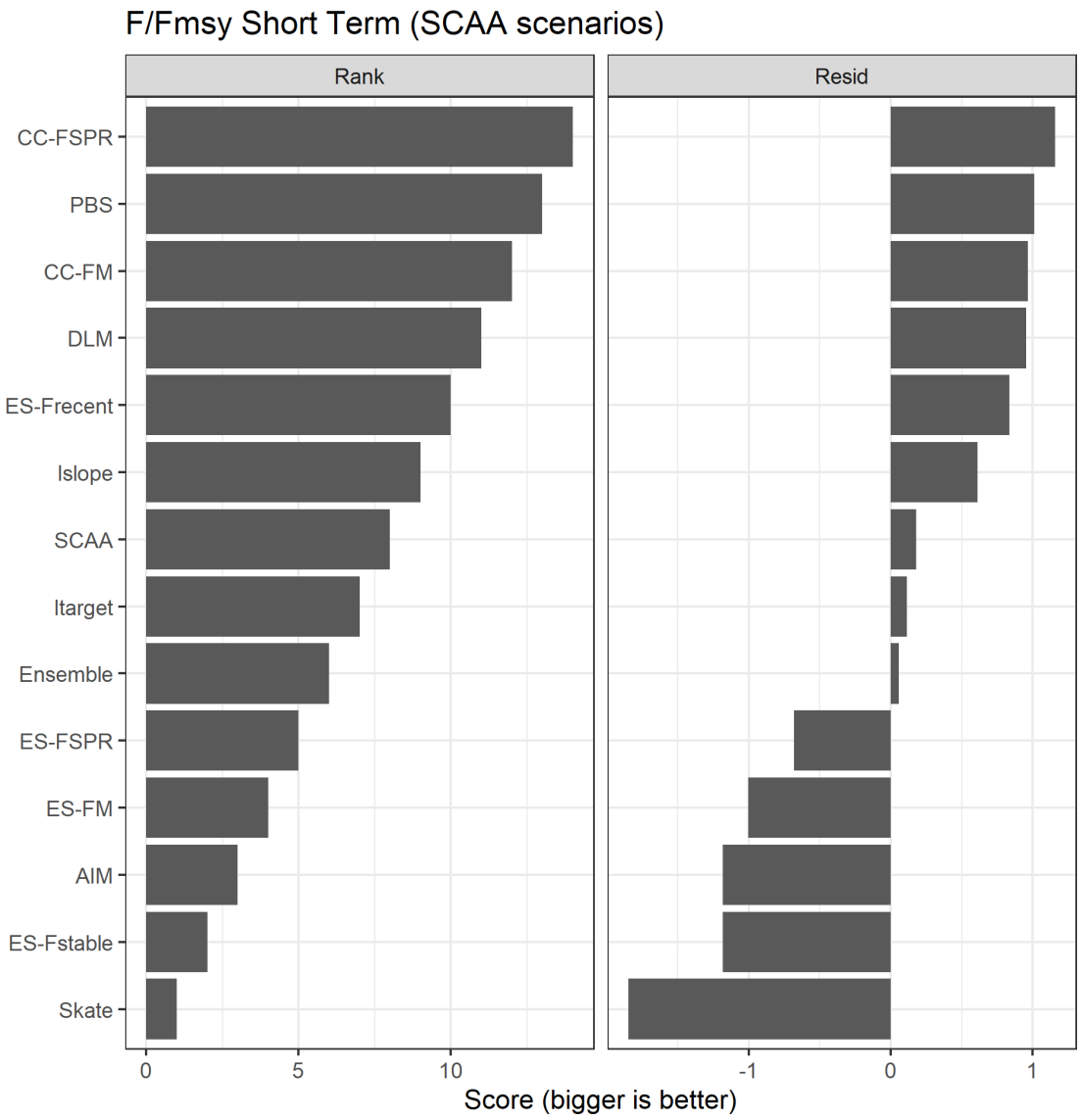


Figure A6.275. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in the short term.

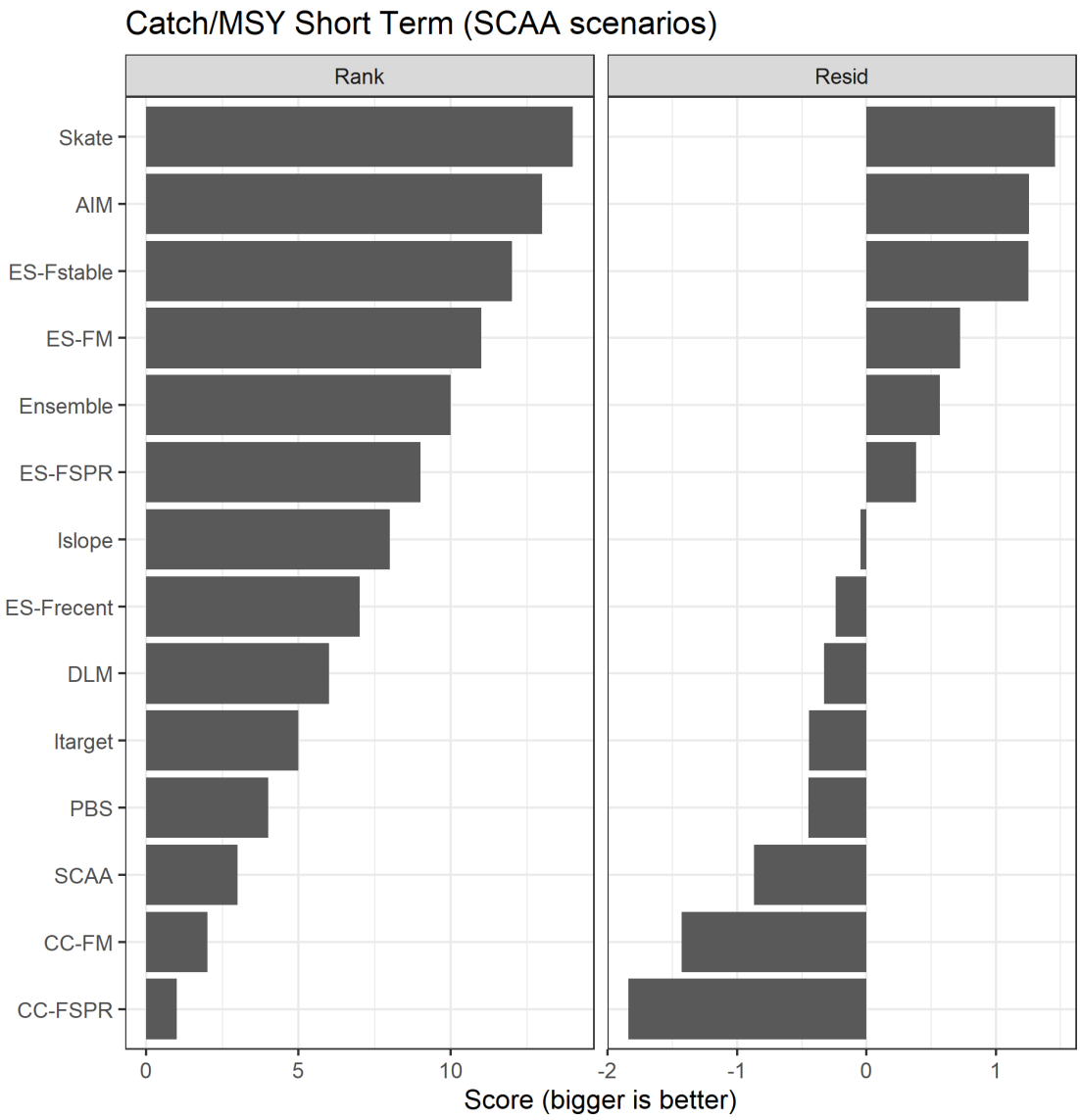


Figure A6.276. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the catch relative to maximum sustainable yield (MSY) in the short term.

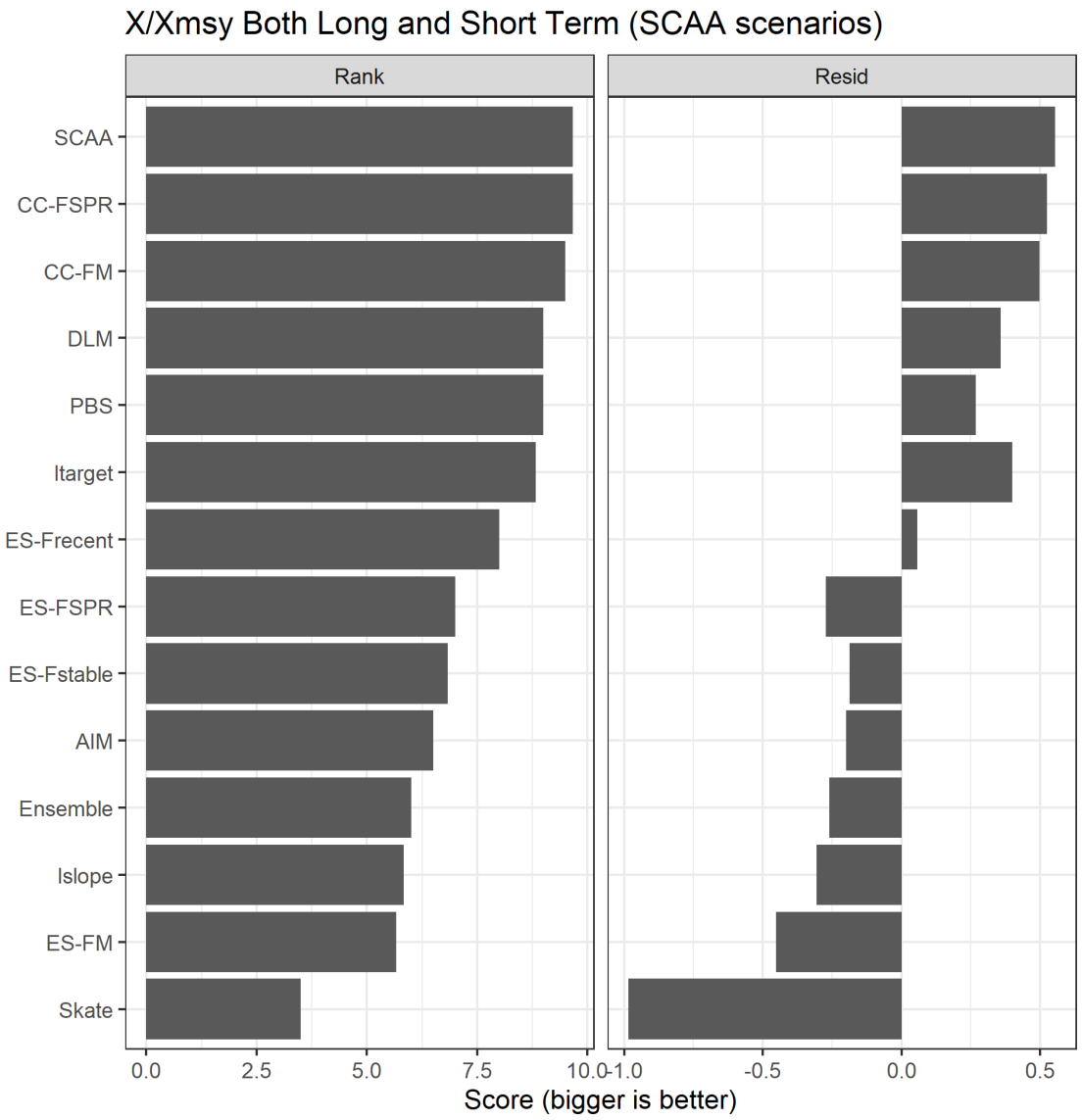


Figure A6.277. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the 3 metrics of spawning stock biomass (SSB), fishing mortality rate (F), and catch relative to their maximum sustainable yield (MSY) reference points (denoted X/X_{MSY}) in both the long and short term.

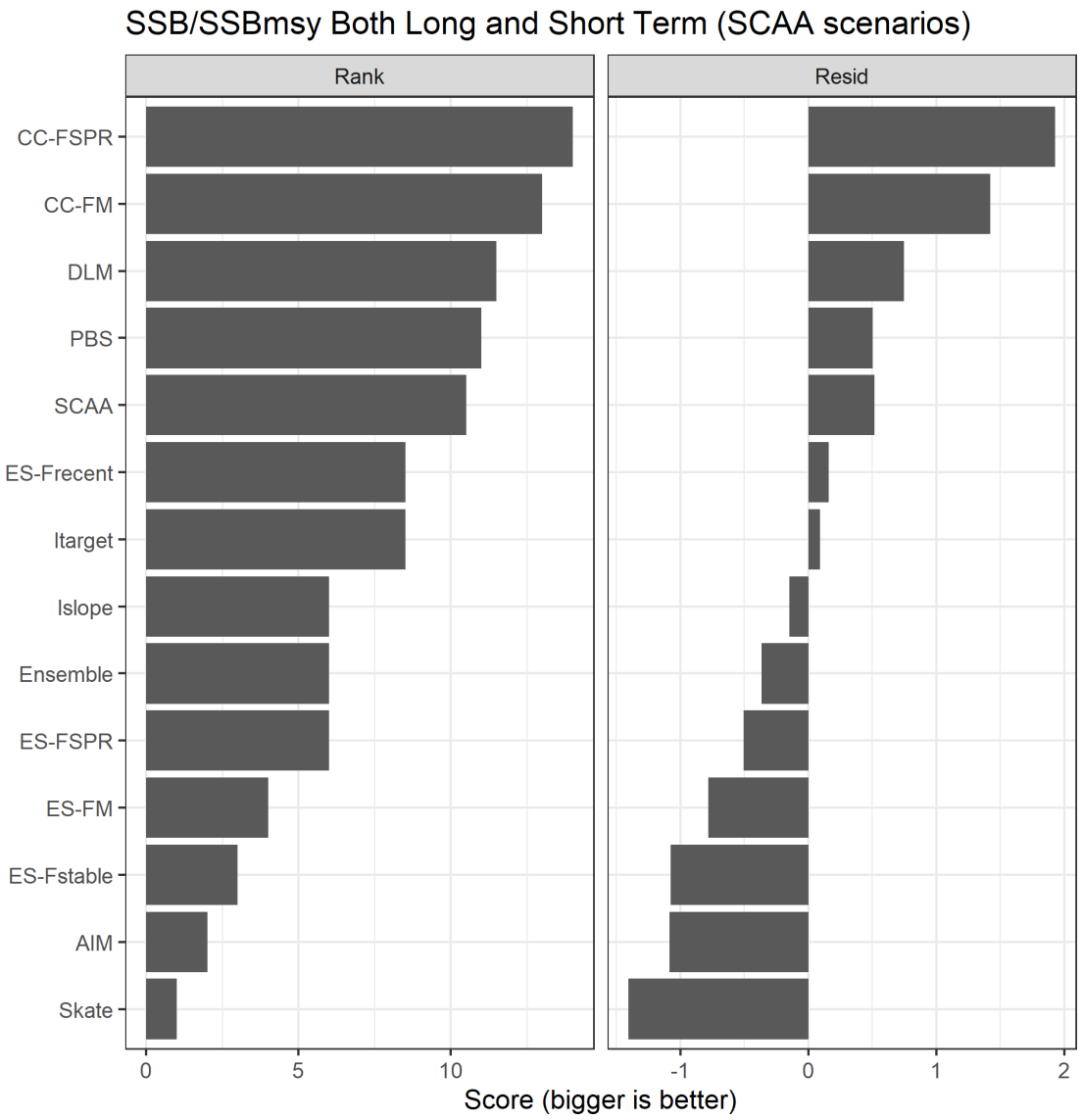


Figure A6.278. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the spawning stock biomass (SSB) relative to maximum sustainable yield (SSB_{MSY}) in both the long and short term.

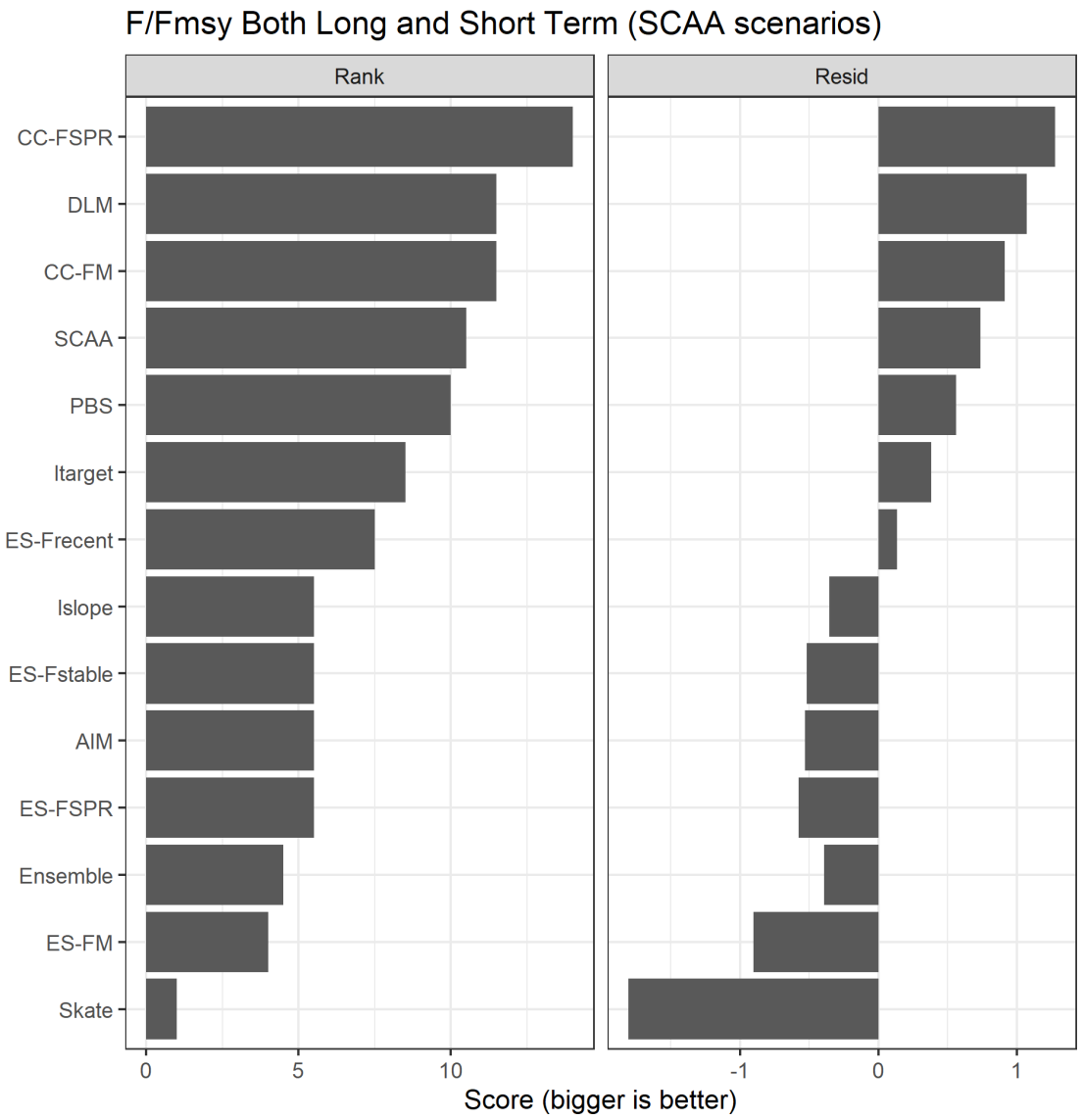


Figure A6.279. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the fishing mortality rate (F) relative to maximum sustainable yield (F_{MSY}) in both the long and short term.

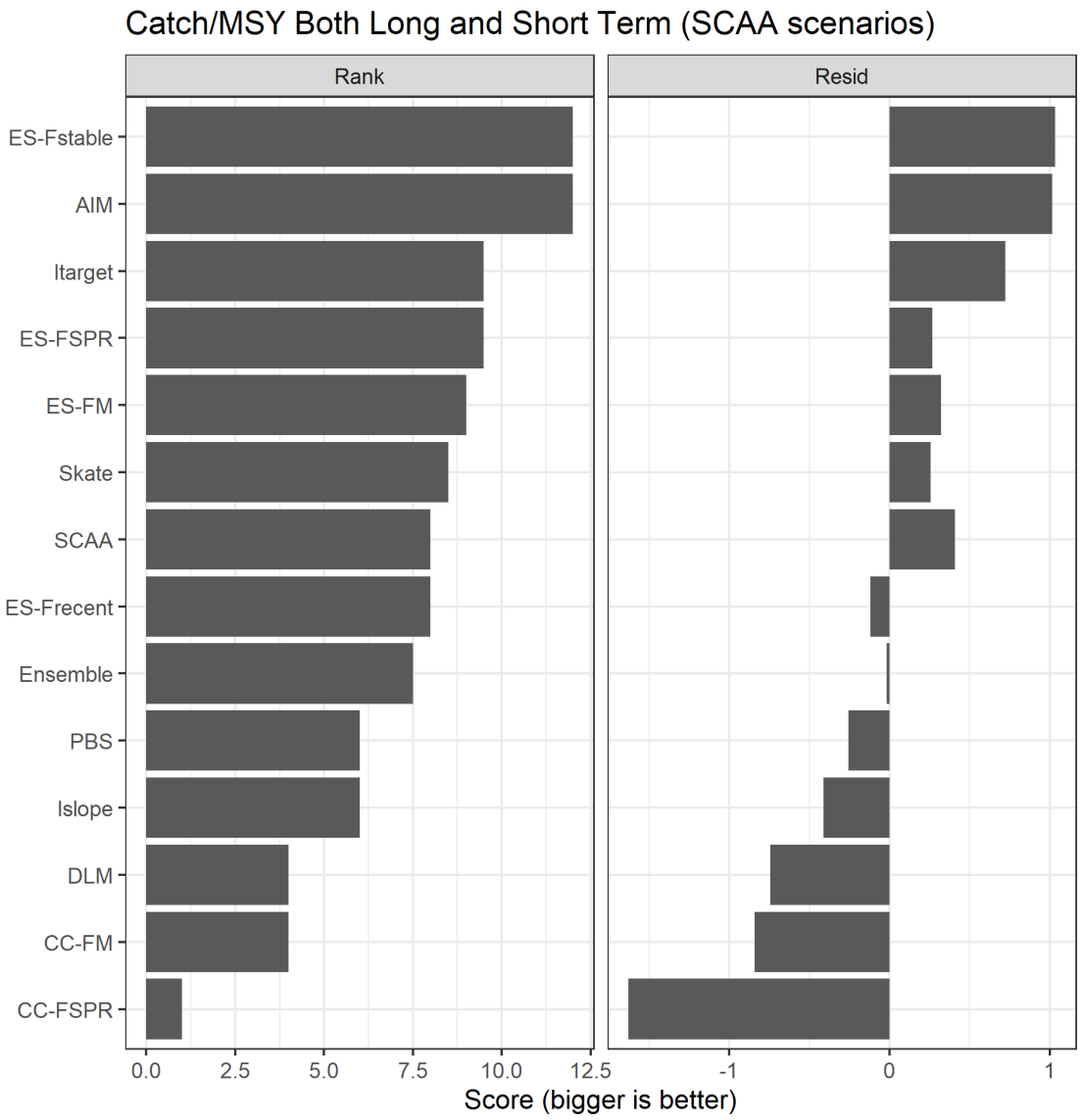


Figure A6.280. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the catch relative to maximum sustainable yield (MSY) in both the long and short term.

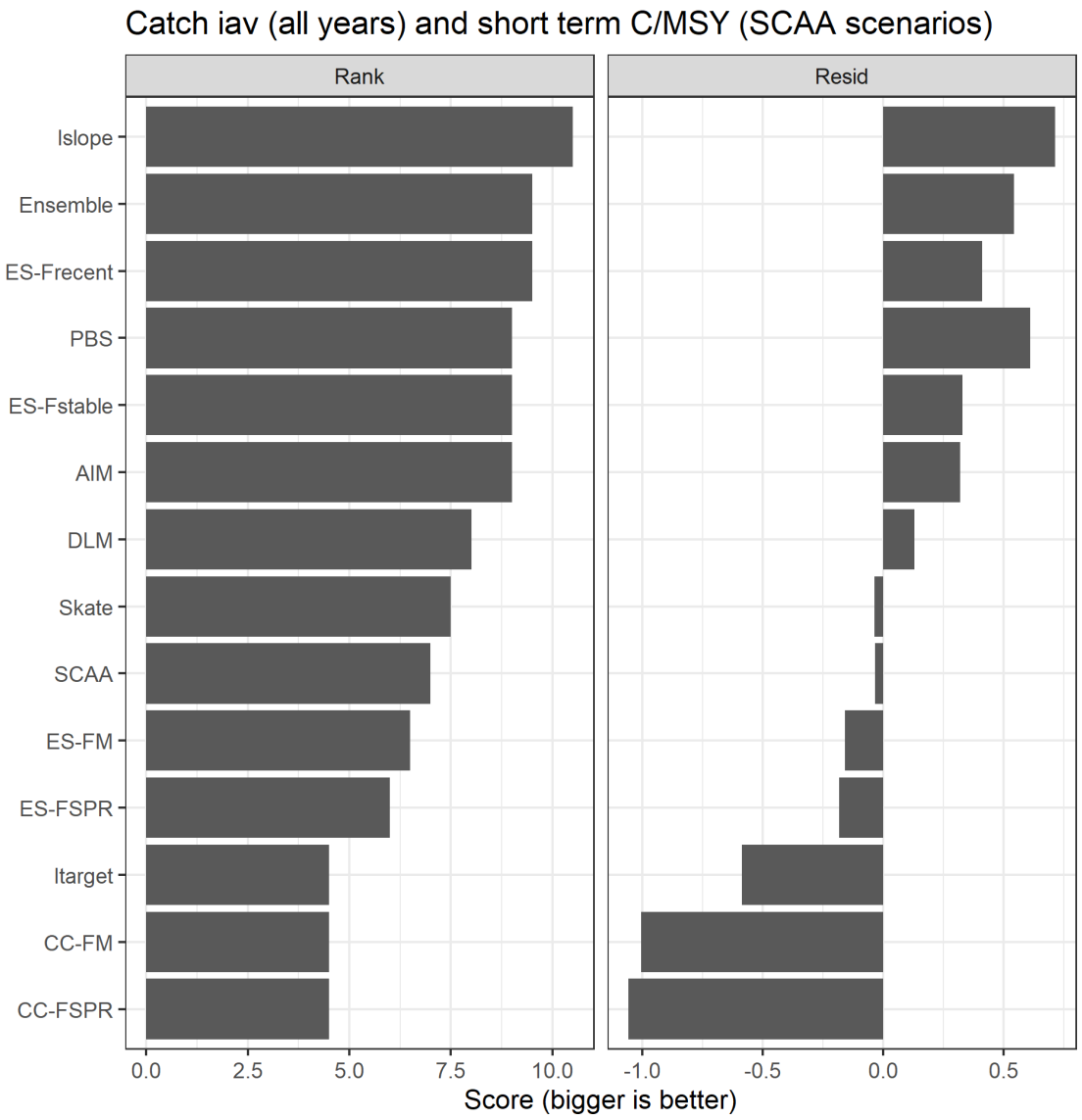


Figure A6.281. Two sets of scores, Rank and Resid, for the statistical catch-at-age (SCAA) analyses for the 2 metrics of interannual variability in catch over the entire feedback period and the short term mean catch/maximum sustainable yield (MSY).

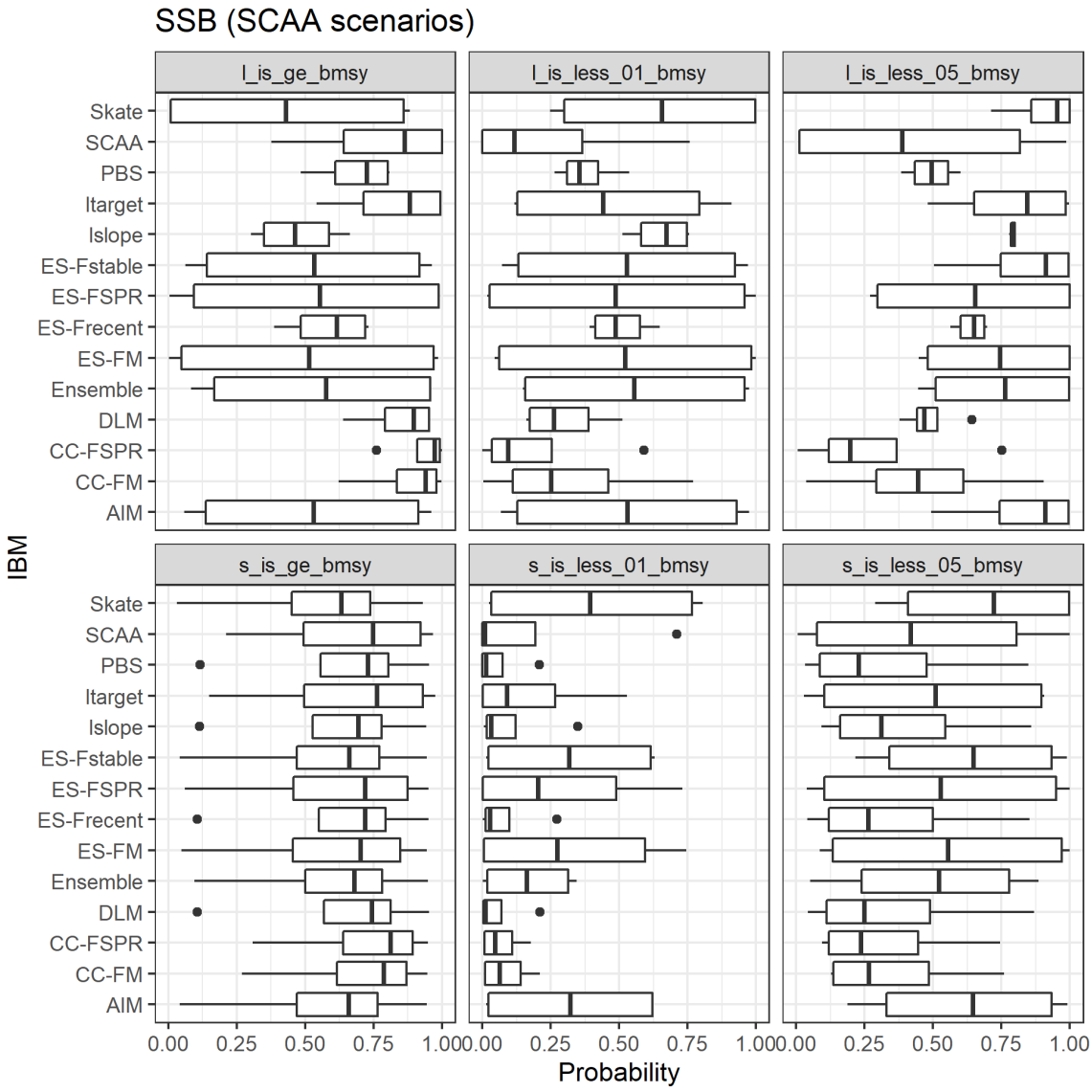


Figure A6.282. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

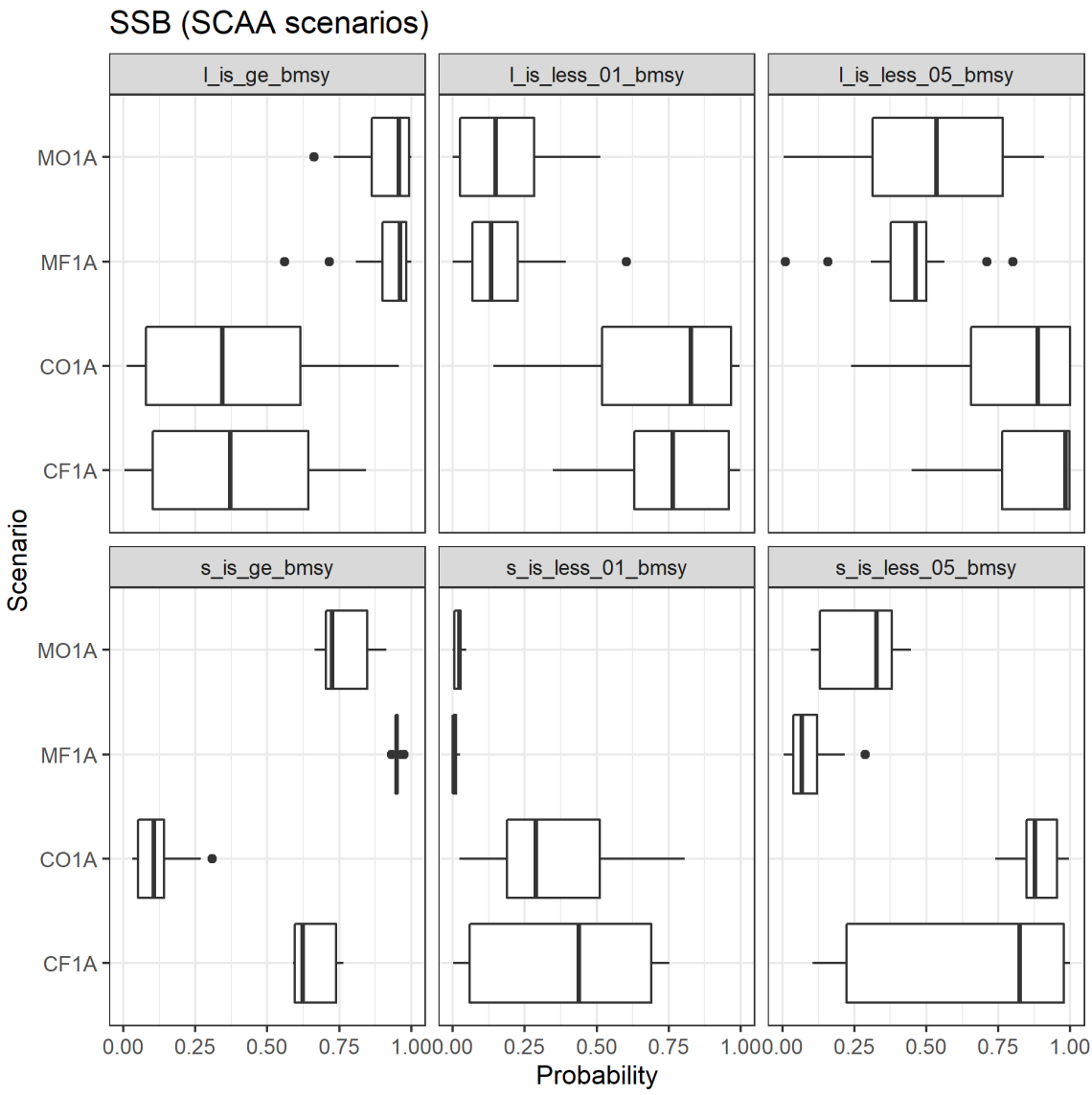


Figure A6.283. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the statistical catch-at-age (SCAA) analyses by scenario.

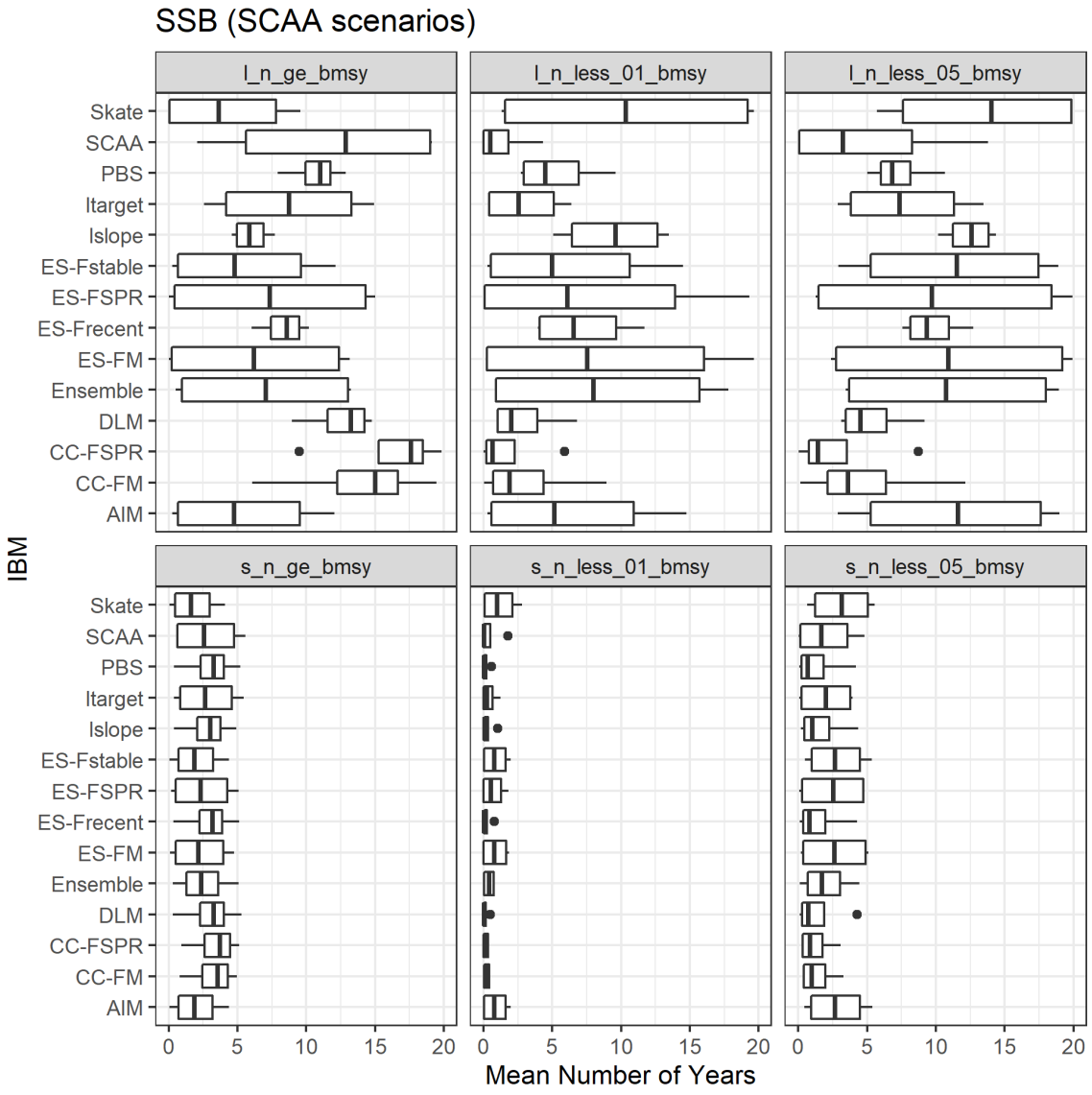


Figure A6.284. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

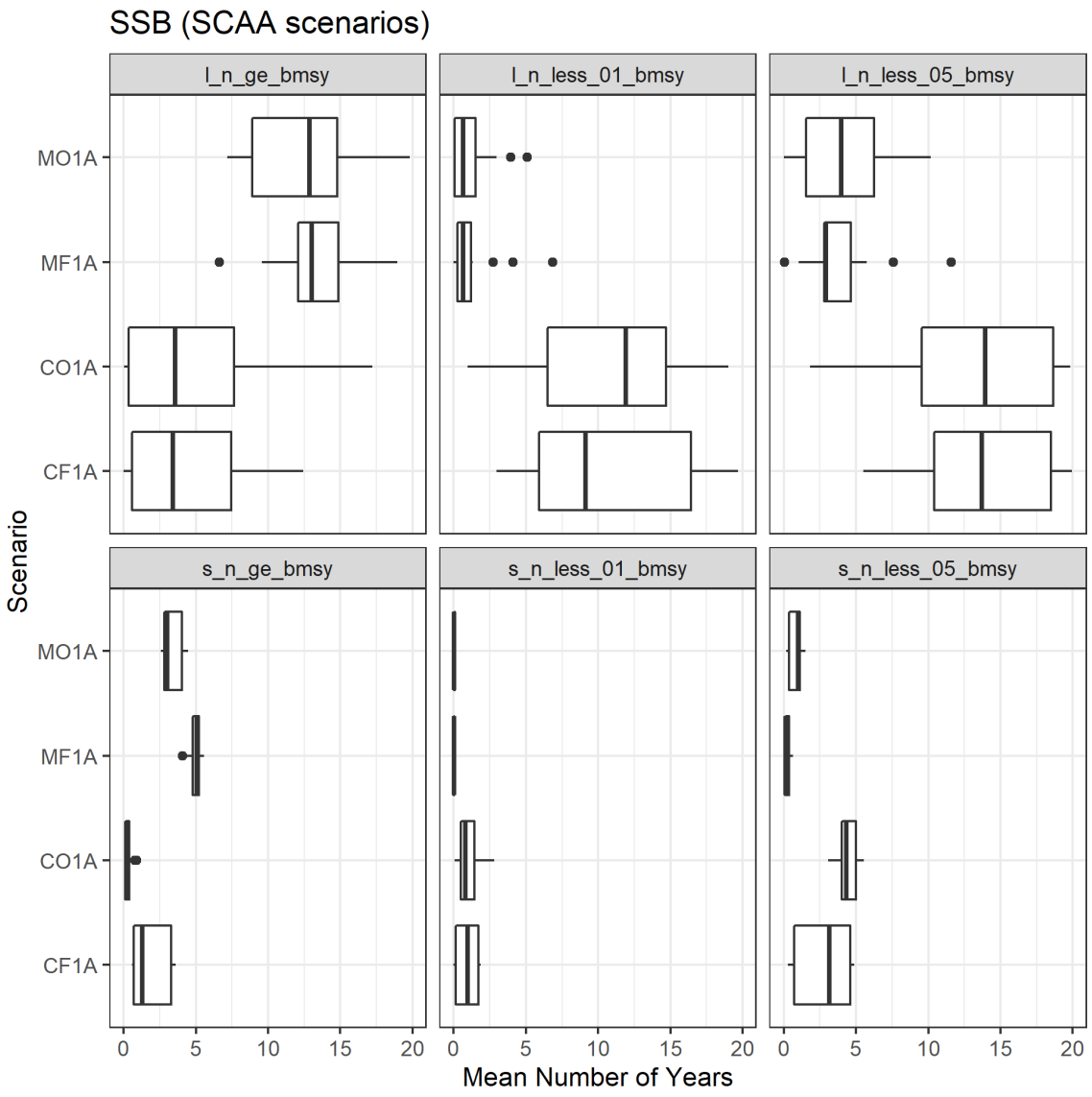


Figure A6.285. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the statistical catch-at-age (SCAA) analyses by scenario.

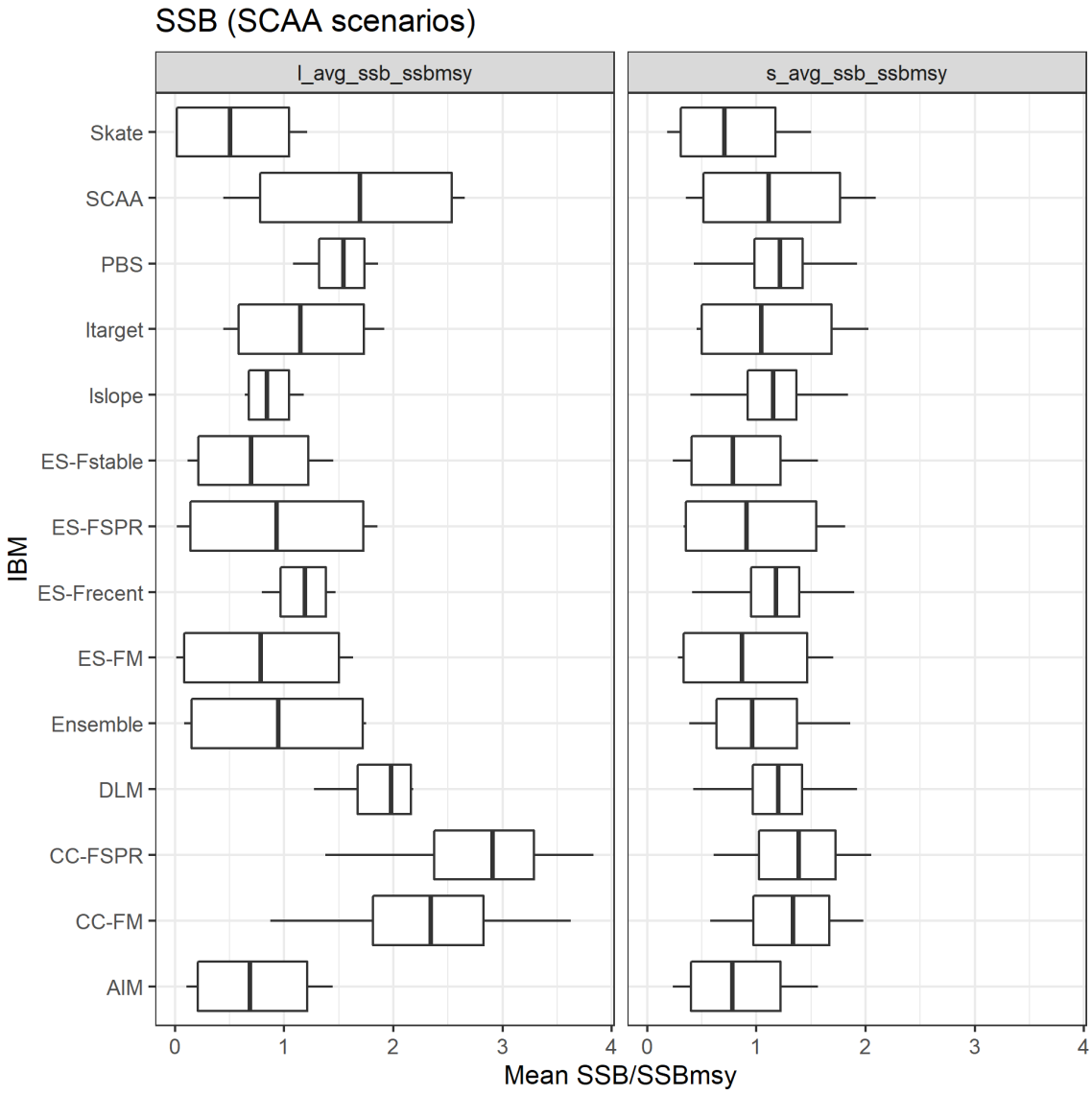


Figure A6.286. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

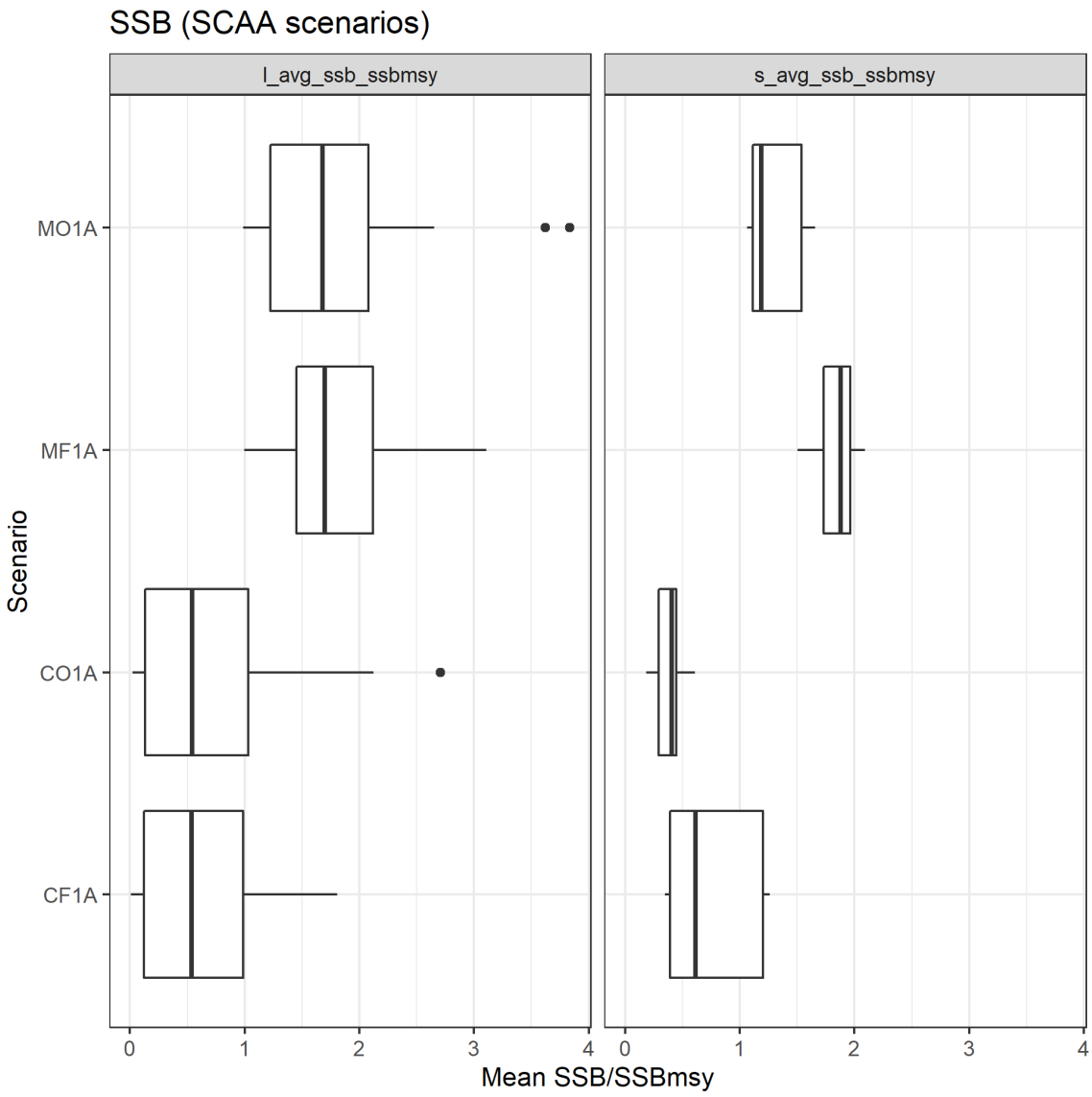


Figure A6.287. Boxplot of the mean values for the spawning stock biomass (SSB) metrics in the statistical catch-at-age (SCAA) analyses by scenario.

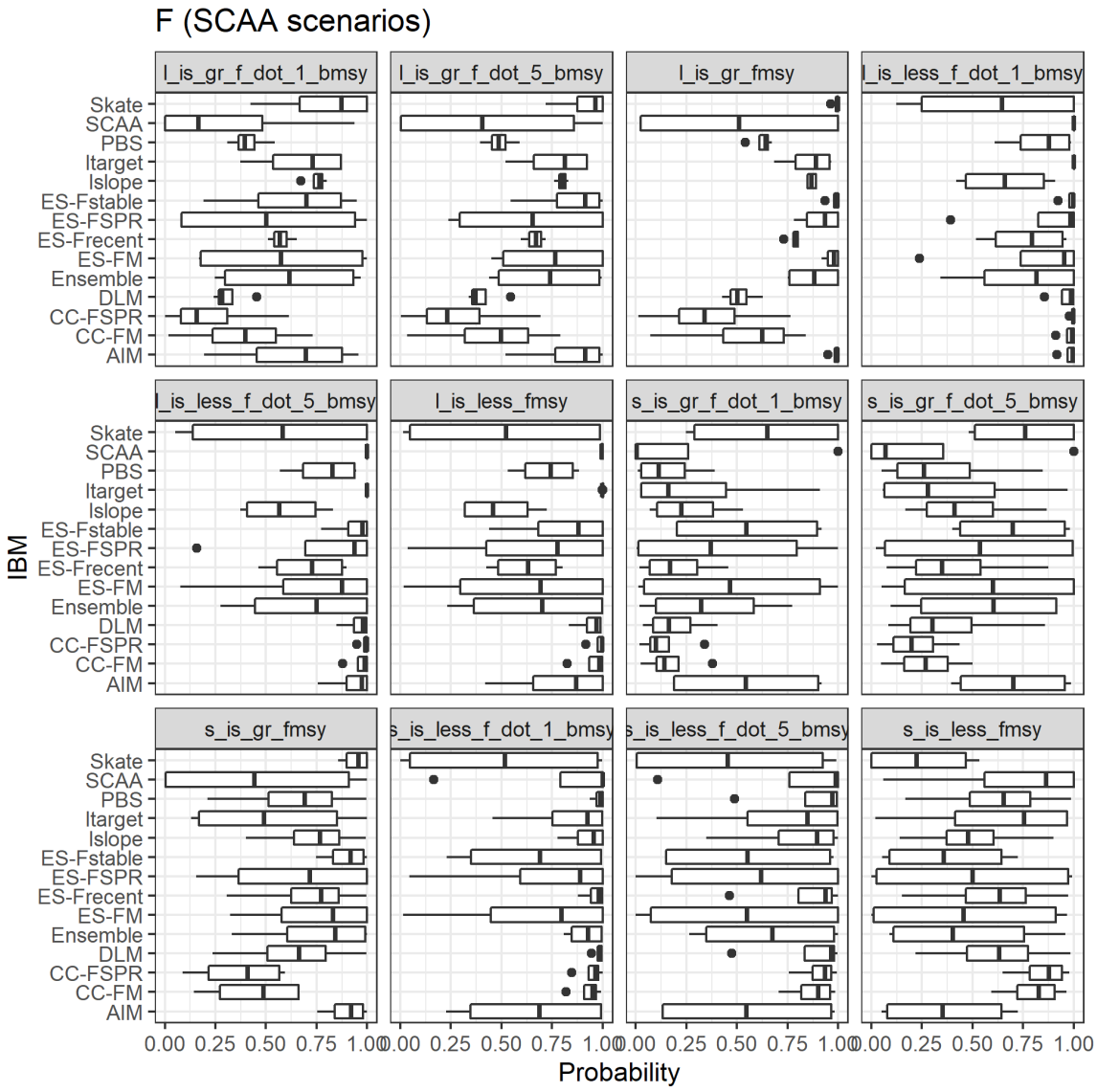


Figure A6.288. Boxplot of the mean values for the fishing mortality rate (F) metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

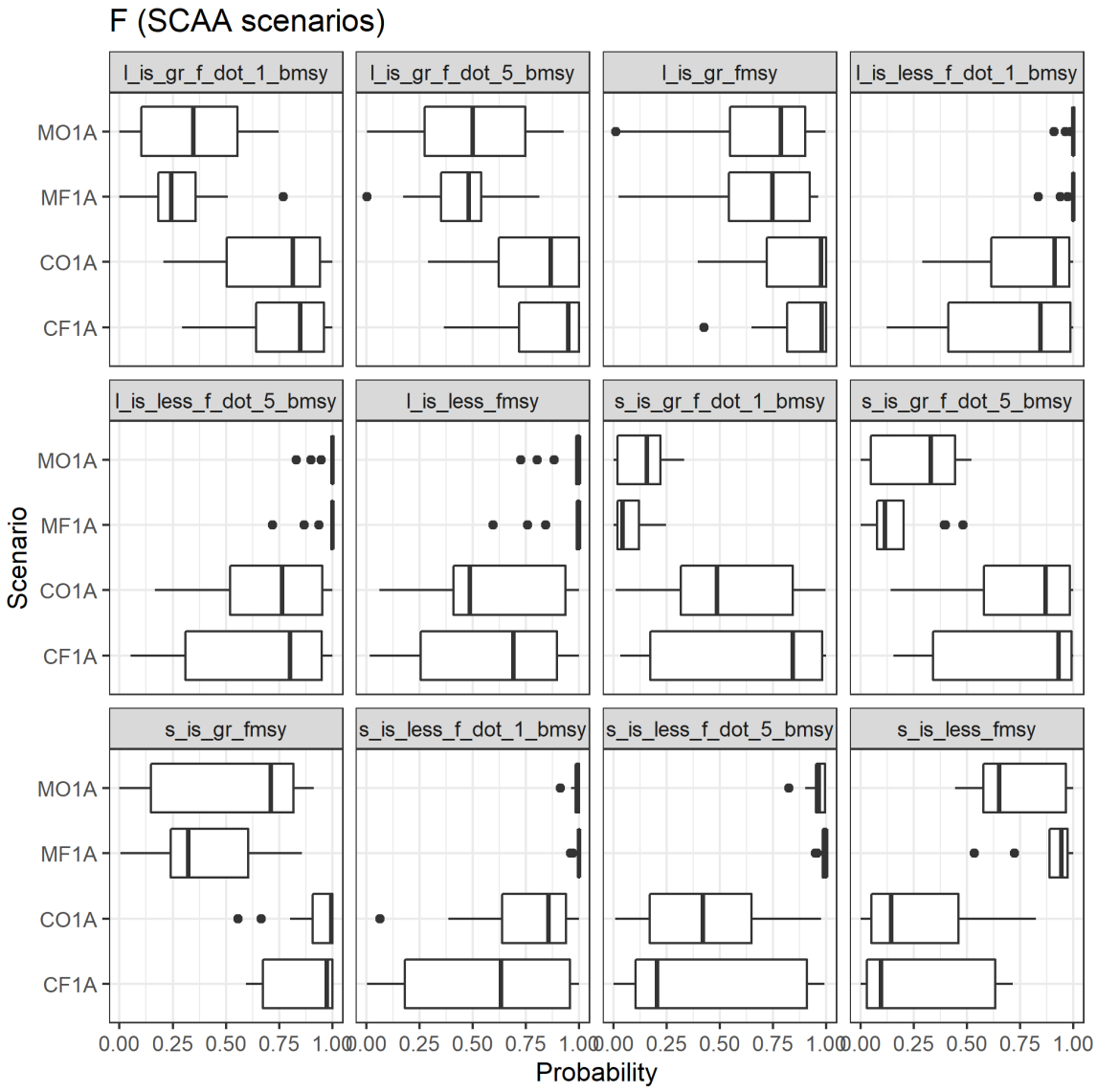


Figure A6.289. Boxplot of the mean values for the fishing mortality rate (F) metrics in the statistical catch-at-age (SCAA) analyses by scenario.

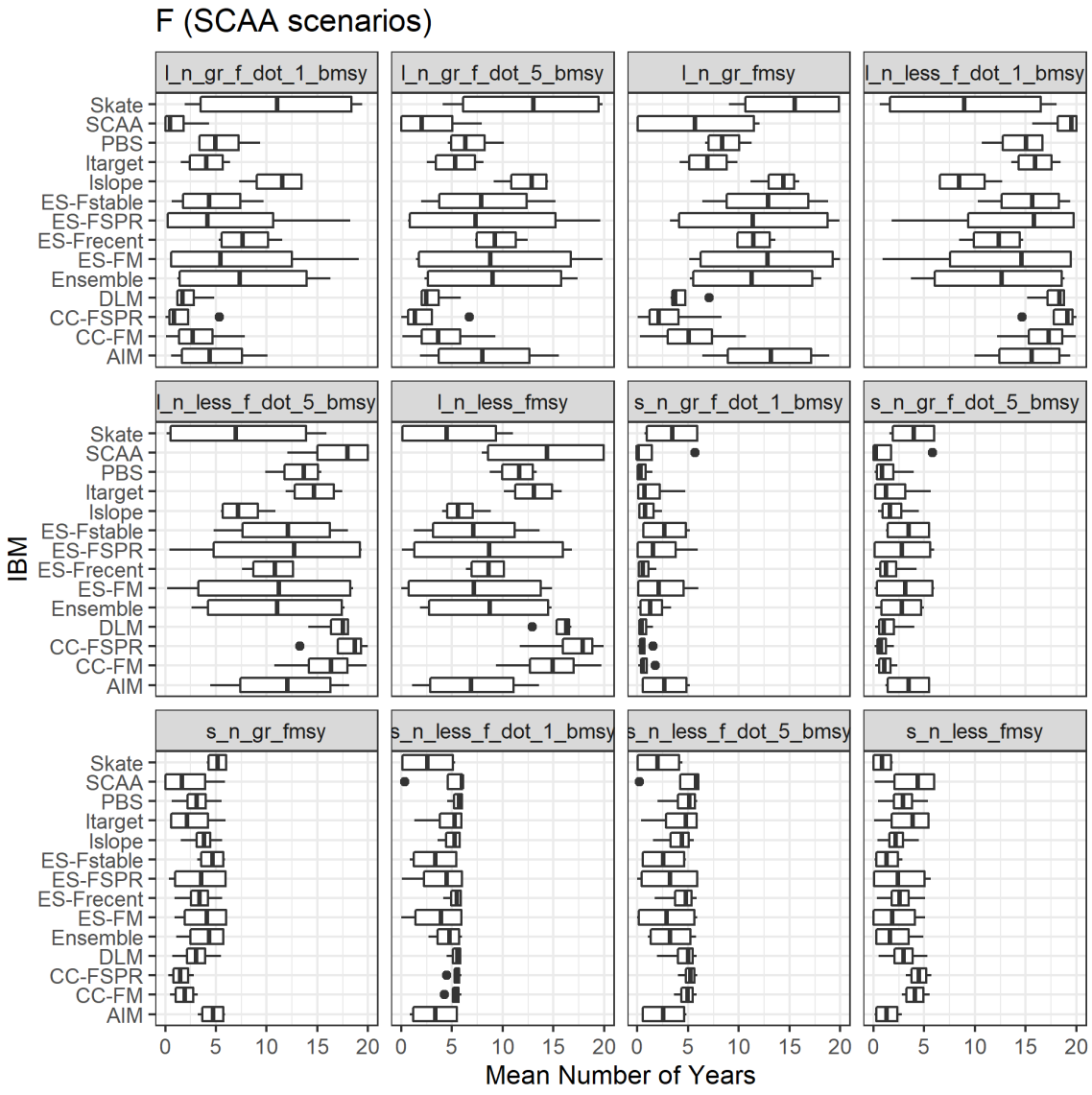


Figure A6.290. Boxplot of the mean values for the fishing mortality rate (F) metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

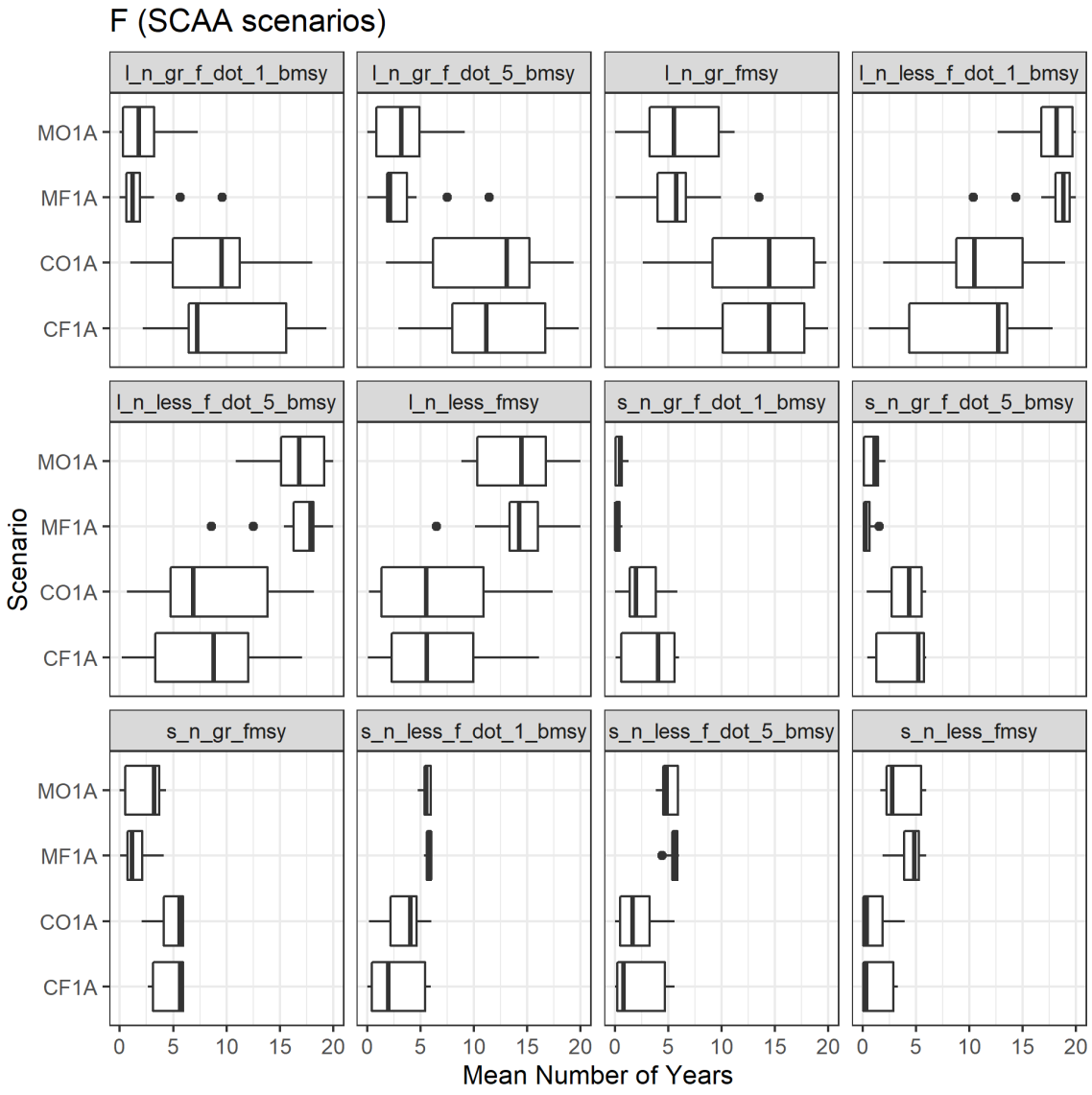


Figure A6.291. Boxplot of the mean values for the fishing mortality rate (F) metrics in the statistical catch-at-age (SCAA) analyses by scenario.

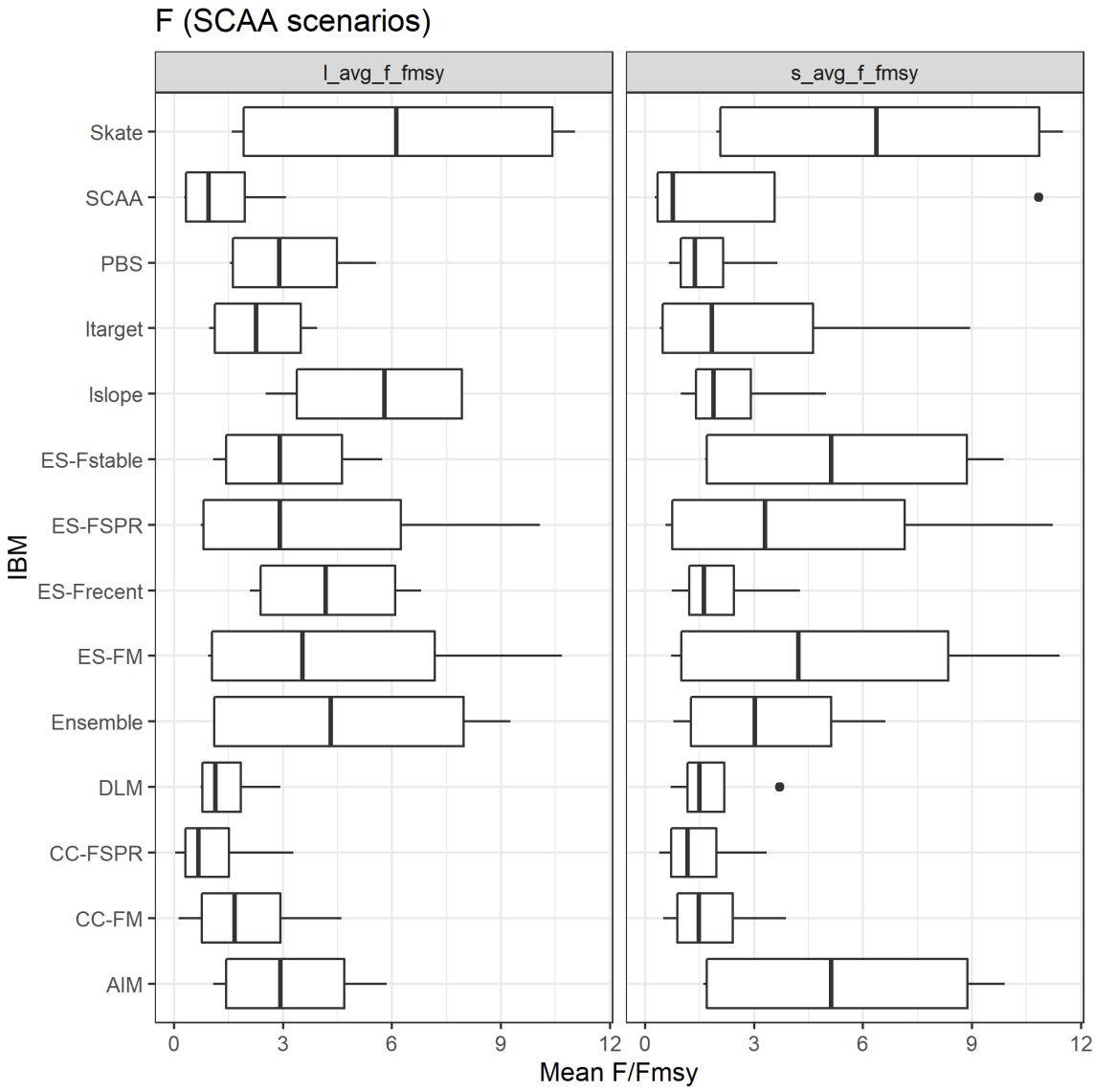


Figure A6.292. Boxplot of the mean values for the fishing mortality rate (F) metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

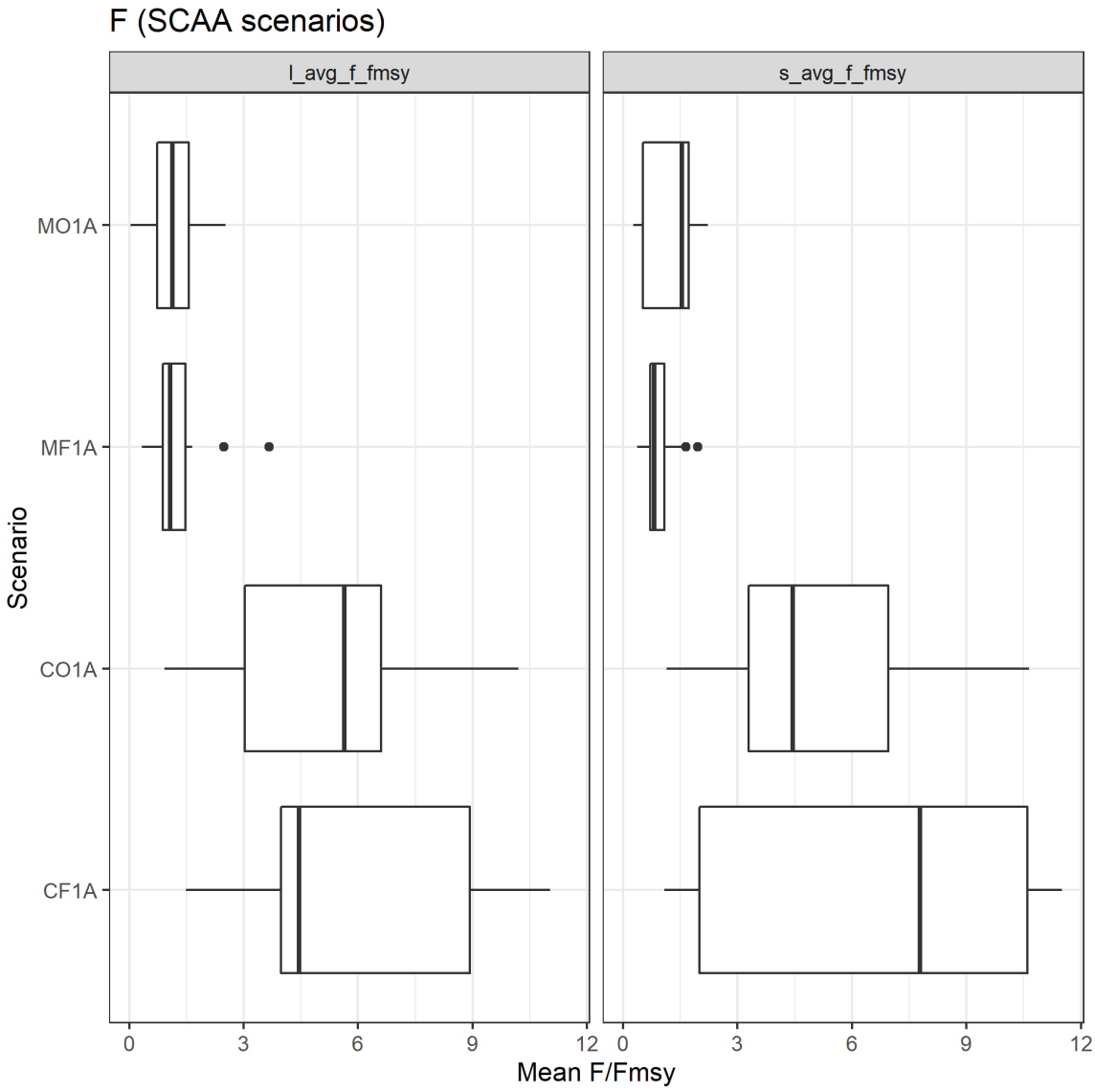


Figure A6.293. Boxplot of the mean values for the fishing mortality rate (F) metrics in the statistical catch-at-age (SCAA) analyses by scenario.

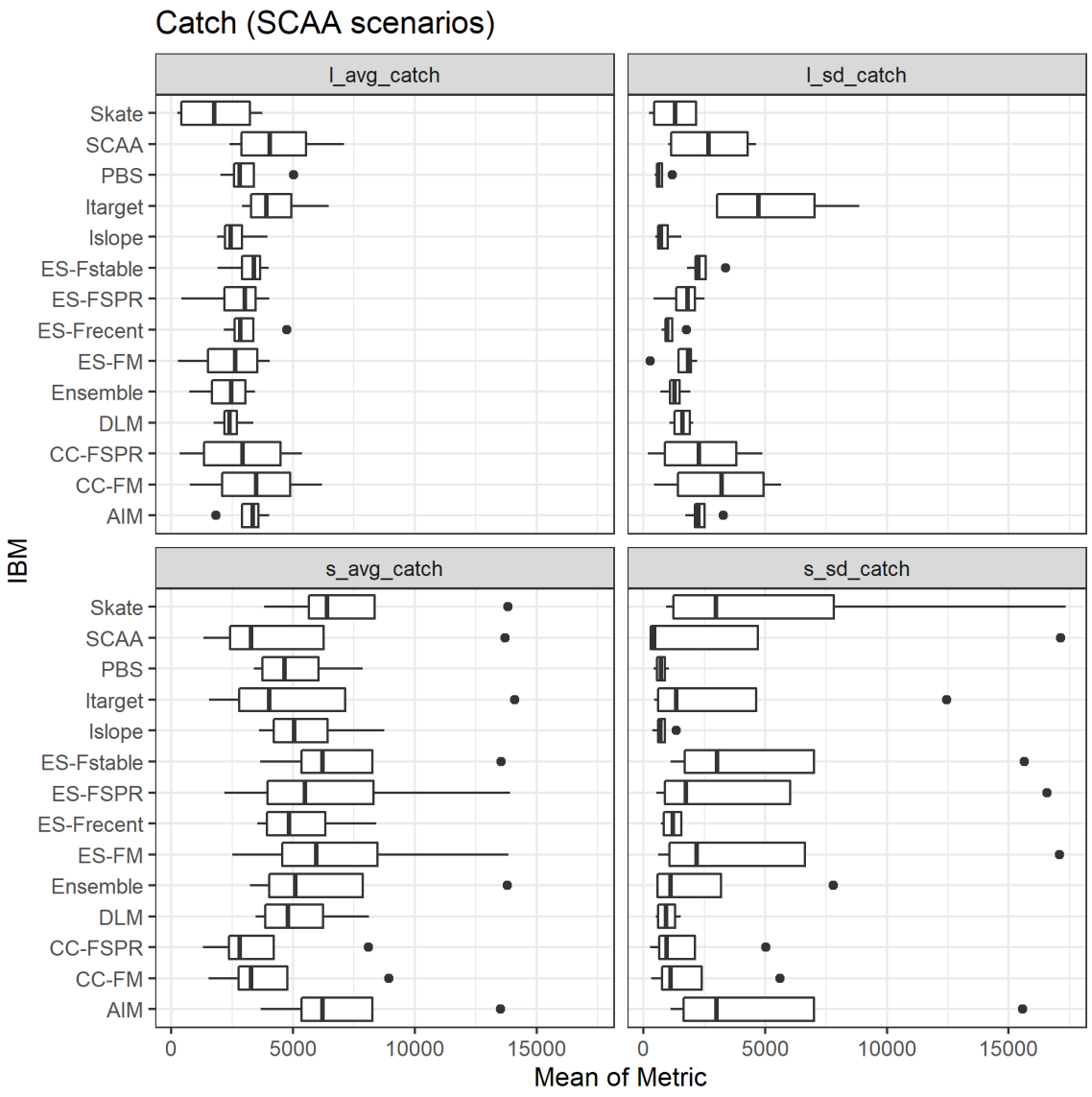


Figure A6.294. Boxplot of the mean values for the catch metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

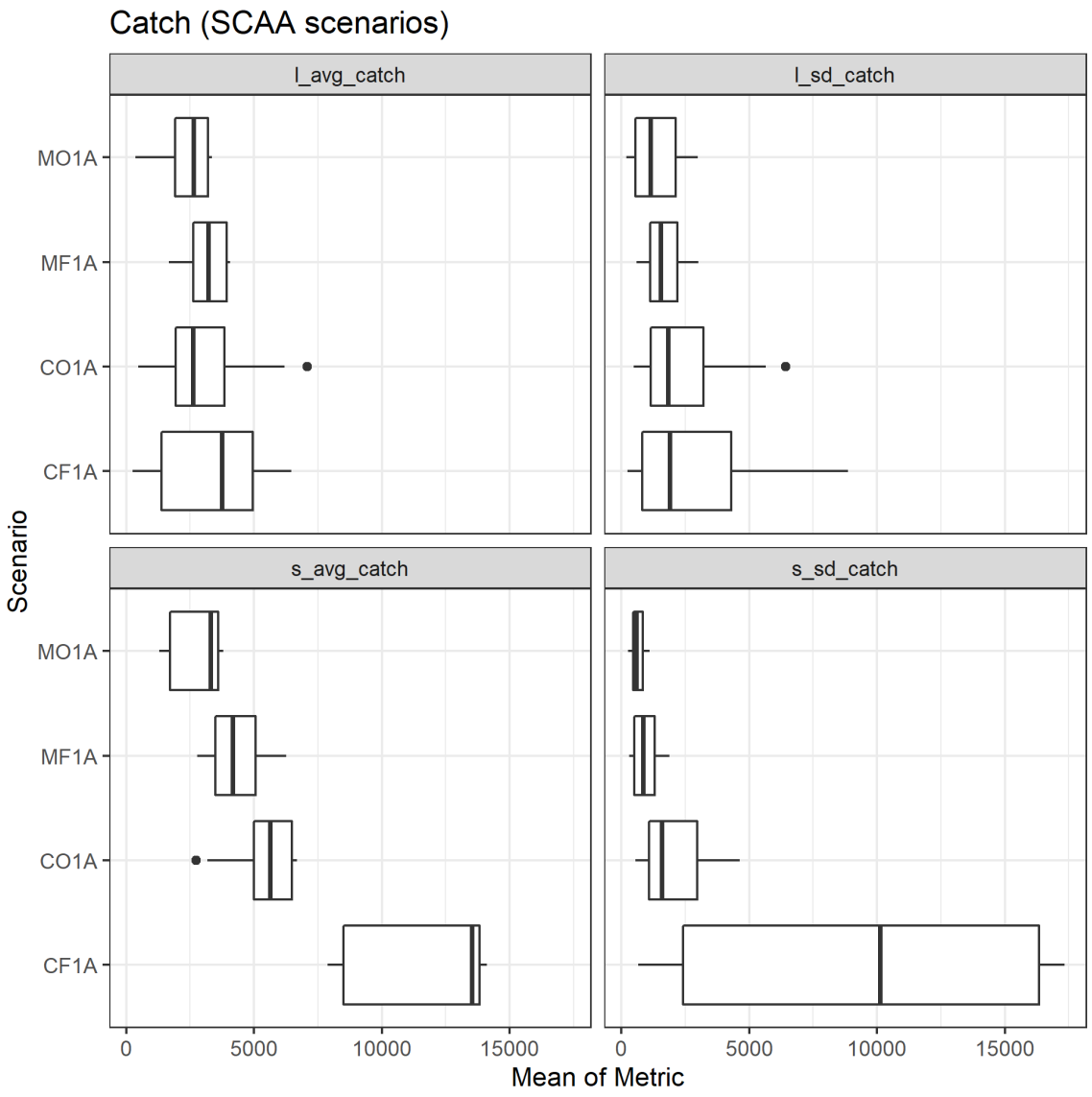


Figure A6.295. Boxplot of the mean values for the catch metrics in the statistical catch-at-age (SCAA) analyses by scenario.

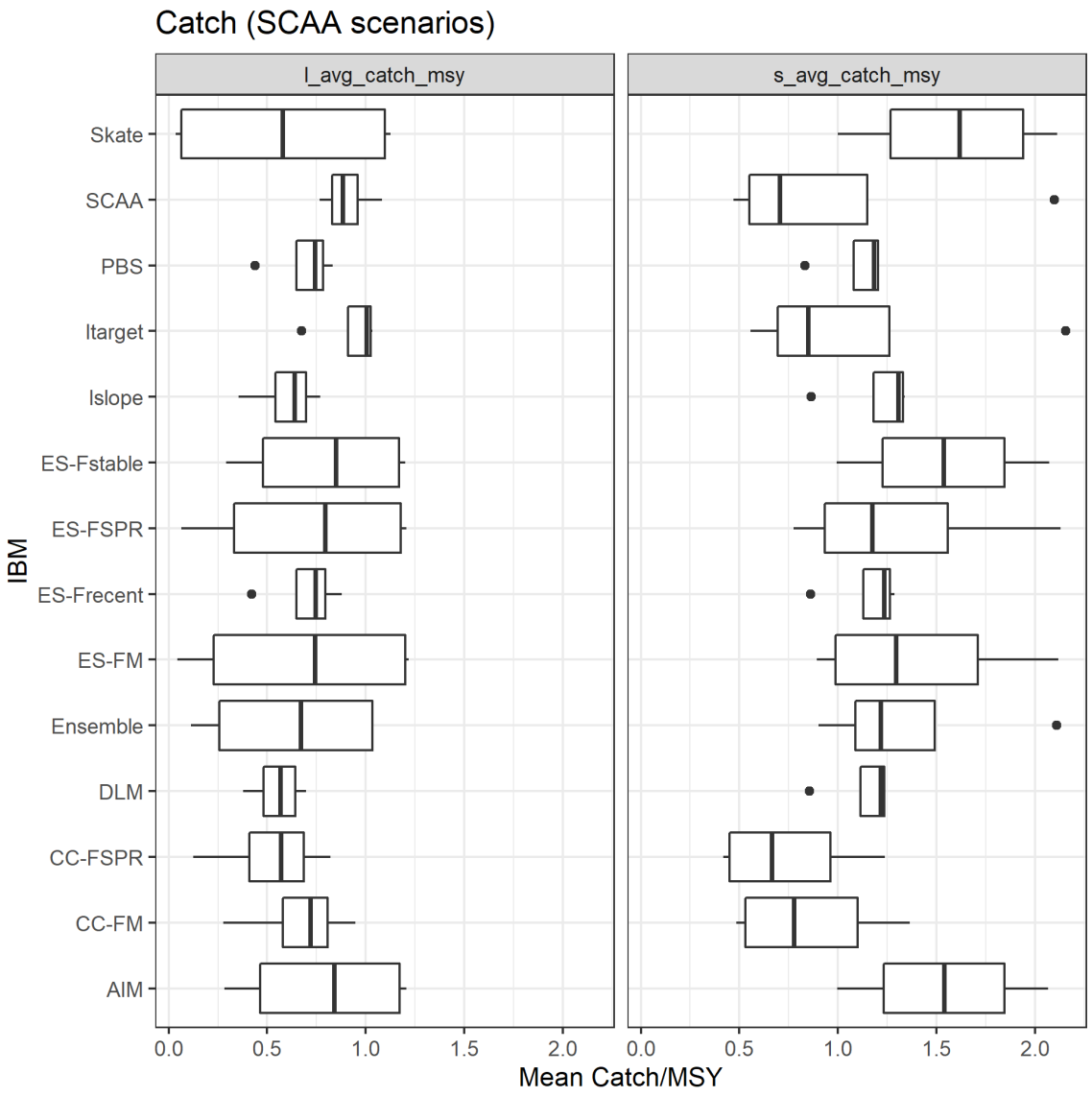


Figure A6.296. Boxplot of the mean values for the catch metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

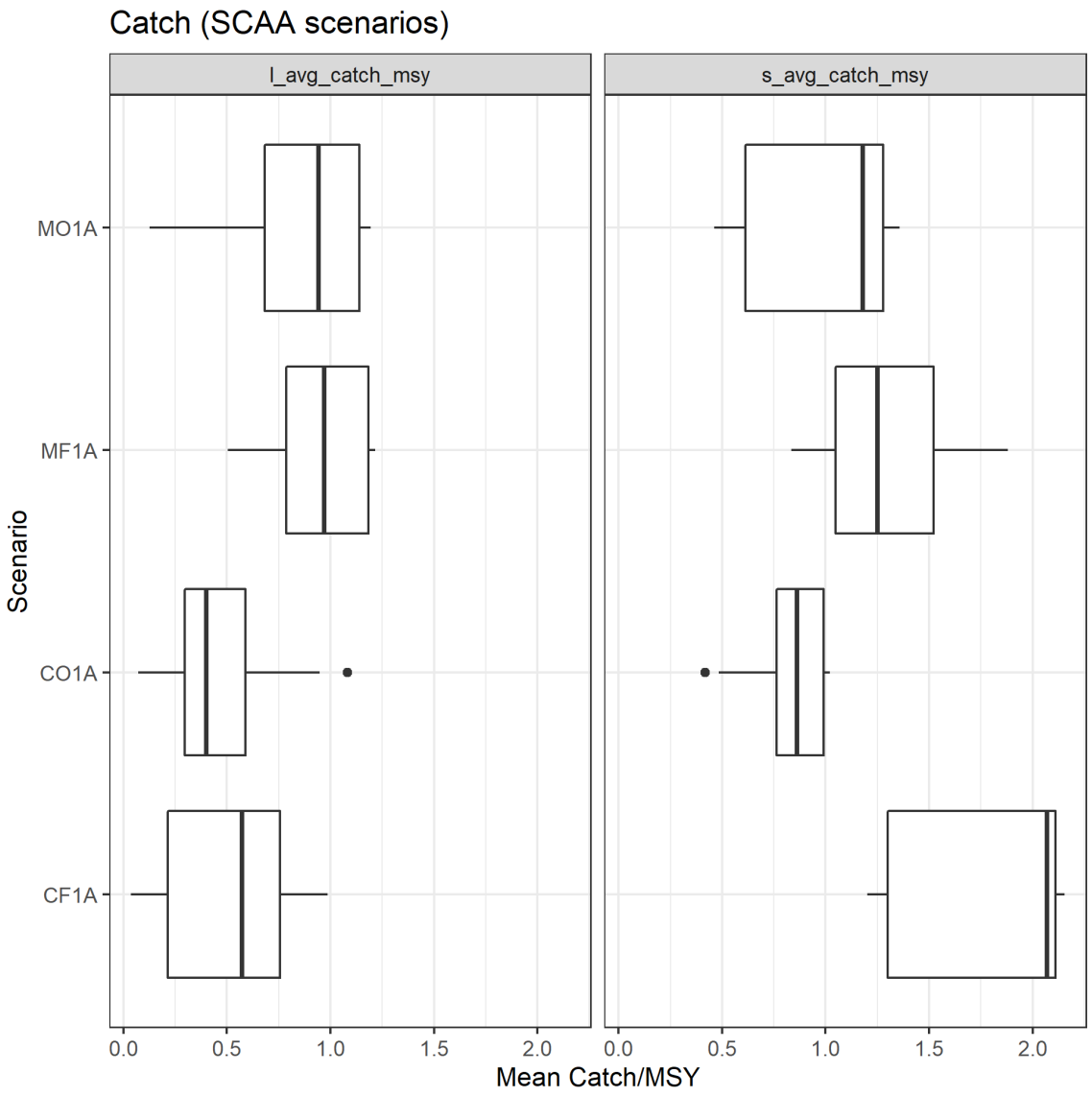


Figure A6.297. Boxplot of the mean values for the catch metrics in the statistical catch-at-age (SCAA) analyses by scenario.

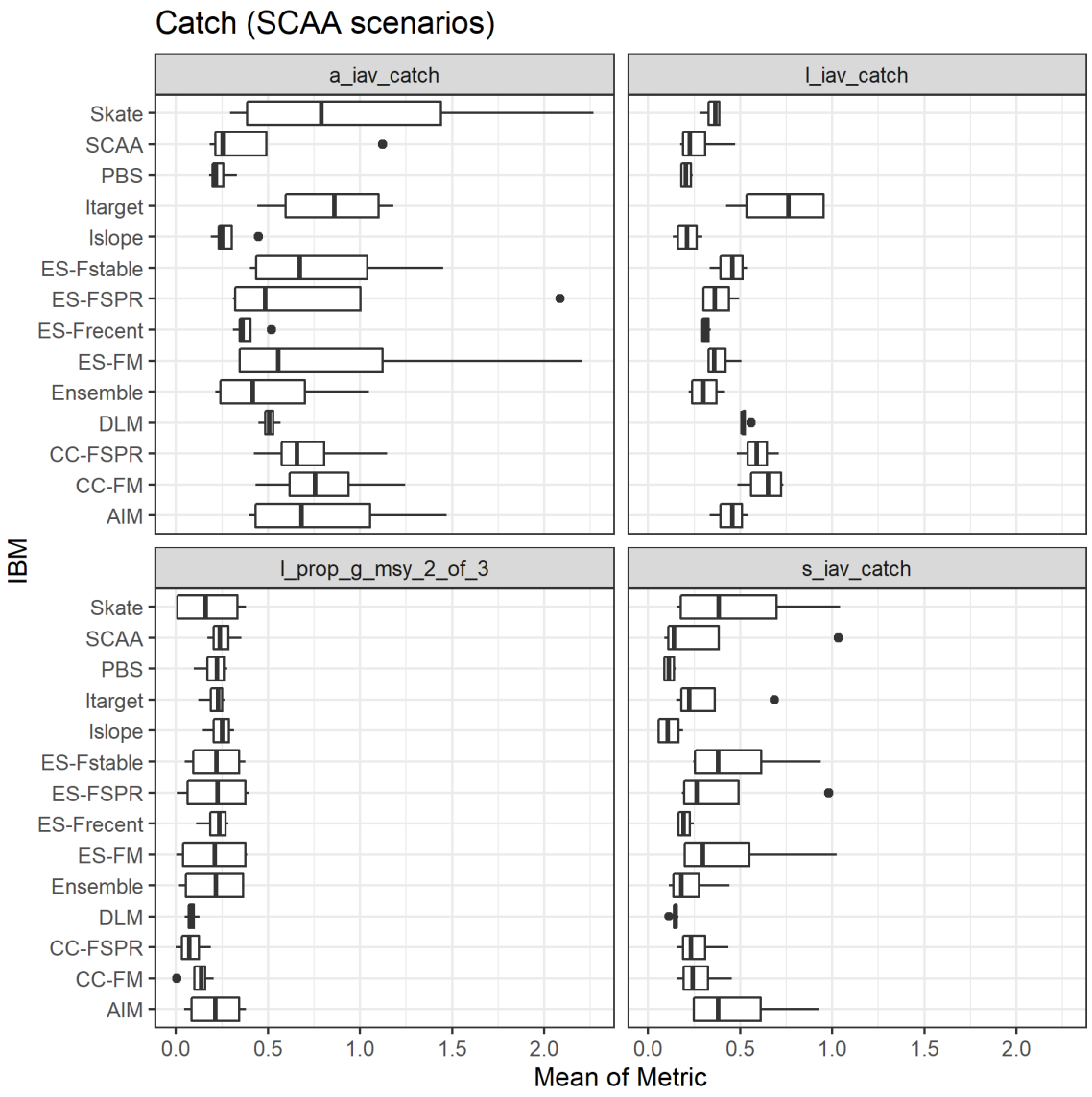


Figure A6.298. Boxplot of the mean values for the catch metrics in the statistical catch-at-age (SCAA) analyses by index-based method (IBM).

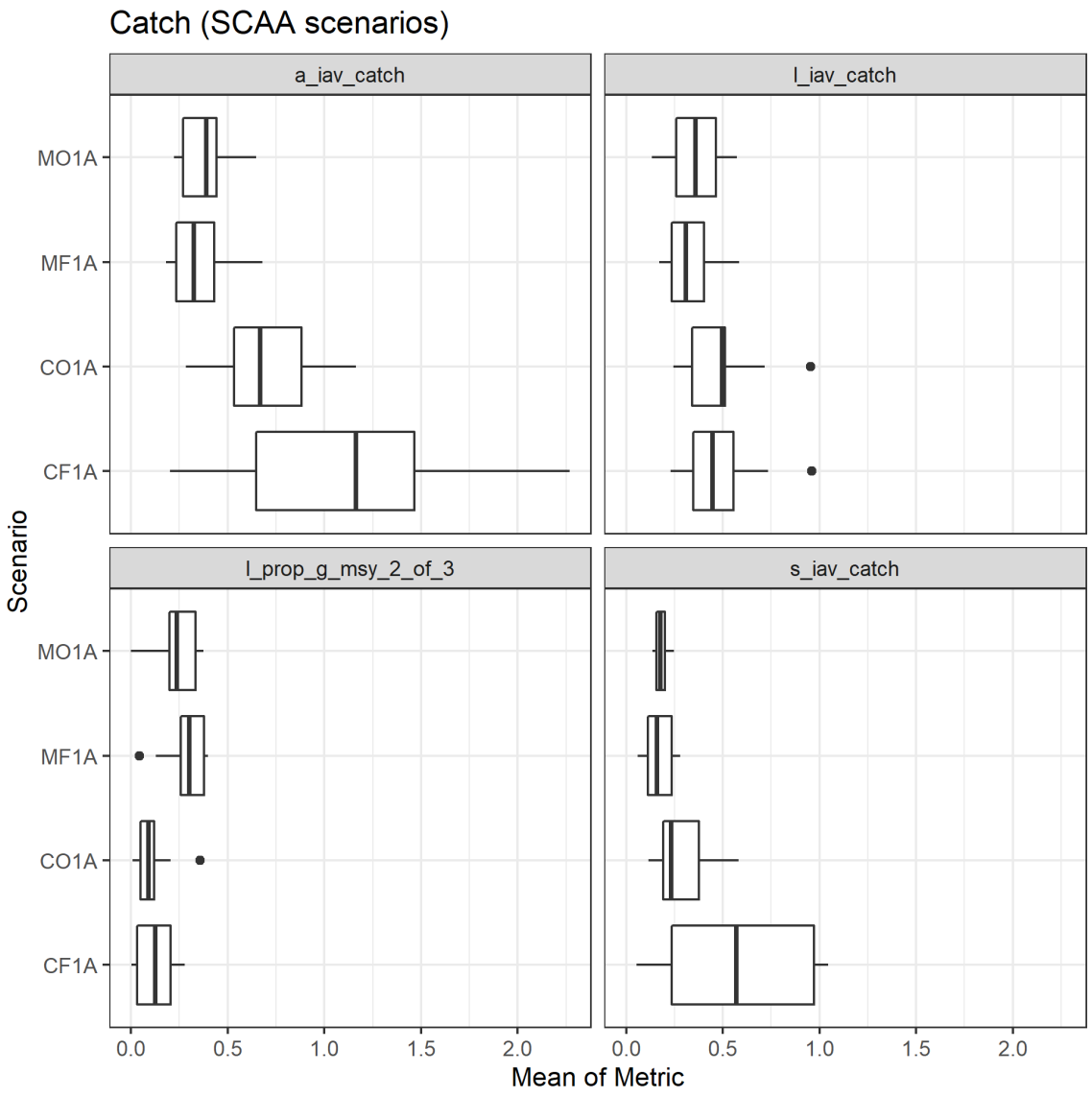


Figure A6.299. Boxplot of the mean values for the catch metrics in the statistical catch-at-age (SCAA) analyses by scenario.

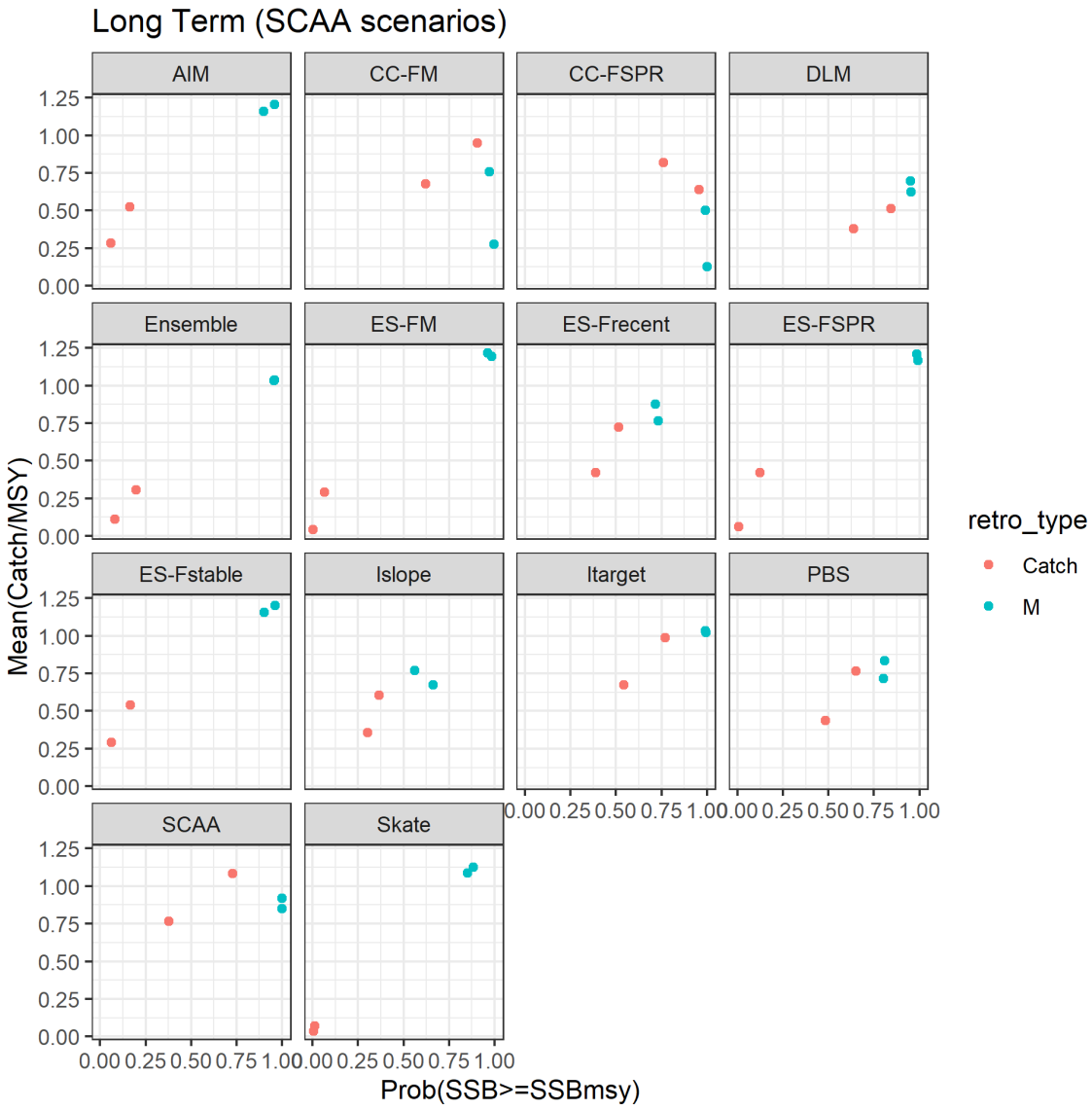


Figure A6.300. Trade off plot by index-based method (IBM) for the statistical catch-at-age (SCAA) analyses between the probability of a rebuilt stock and the mean catch relative to maximum sustainable yield (MSY) in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

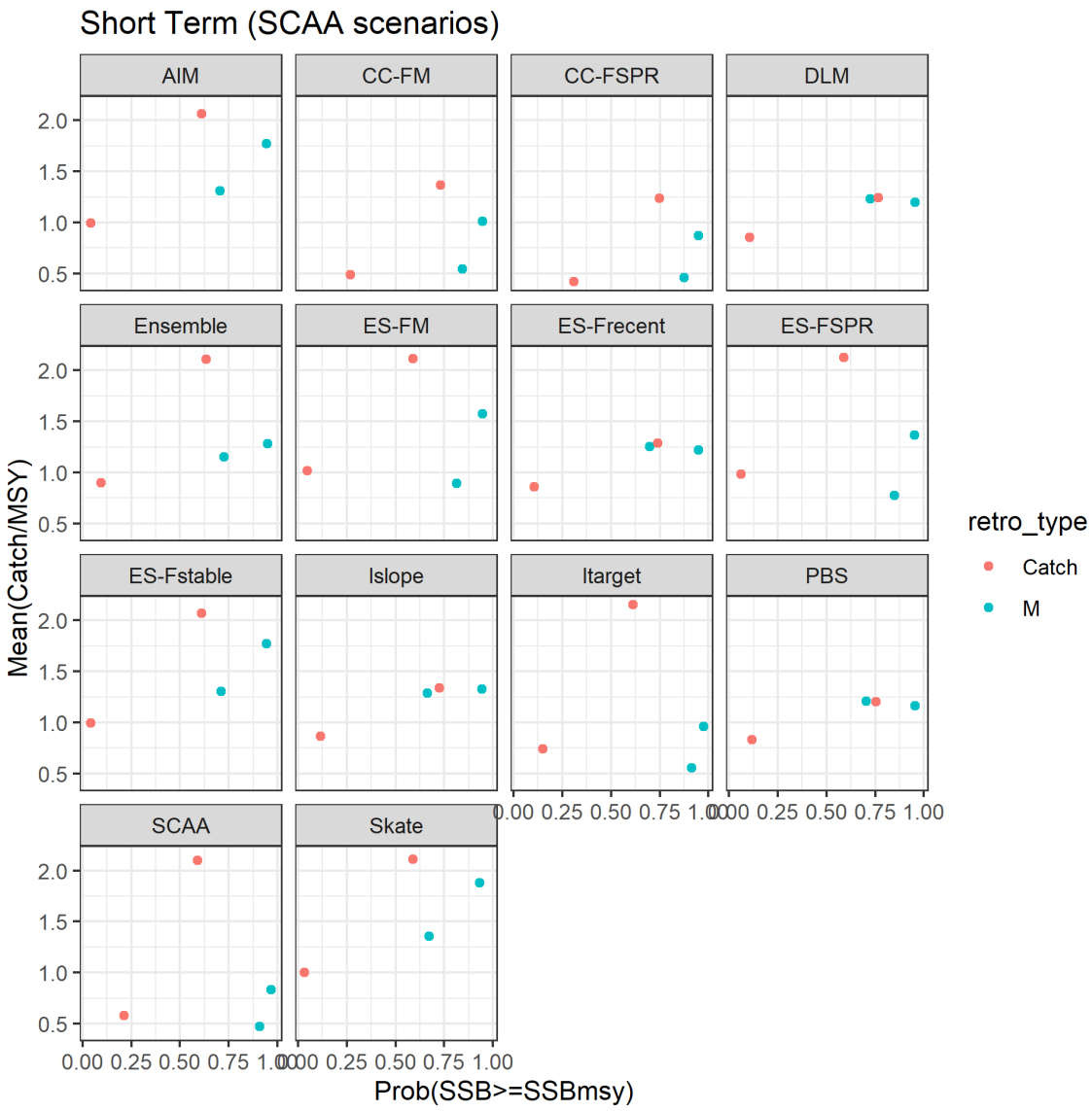


Figure A6.301. Trade off plot by index-based method (IBM) for the statistical catch-at-age (SCAA) analyses between the probability of a rebuilt stock and the mean catch relative to maximum sustainable yield (MSY) in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

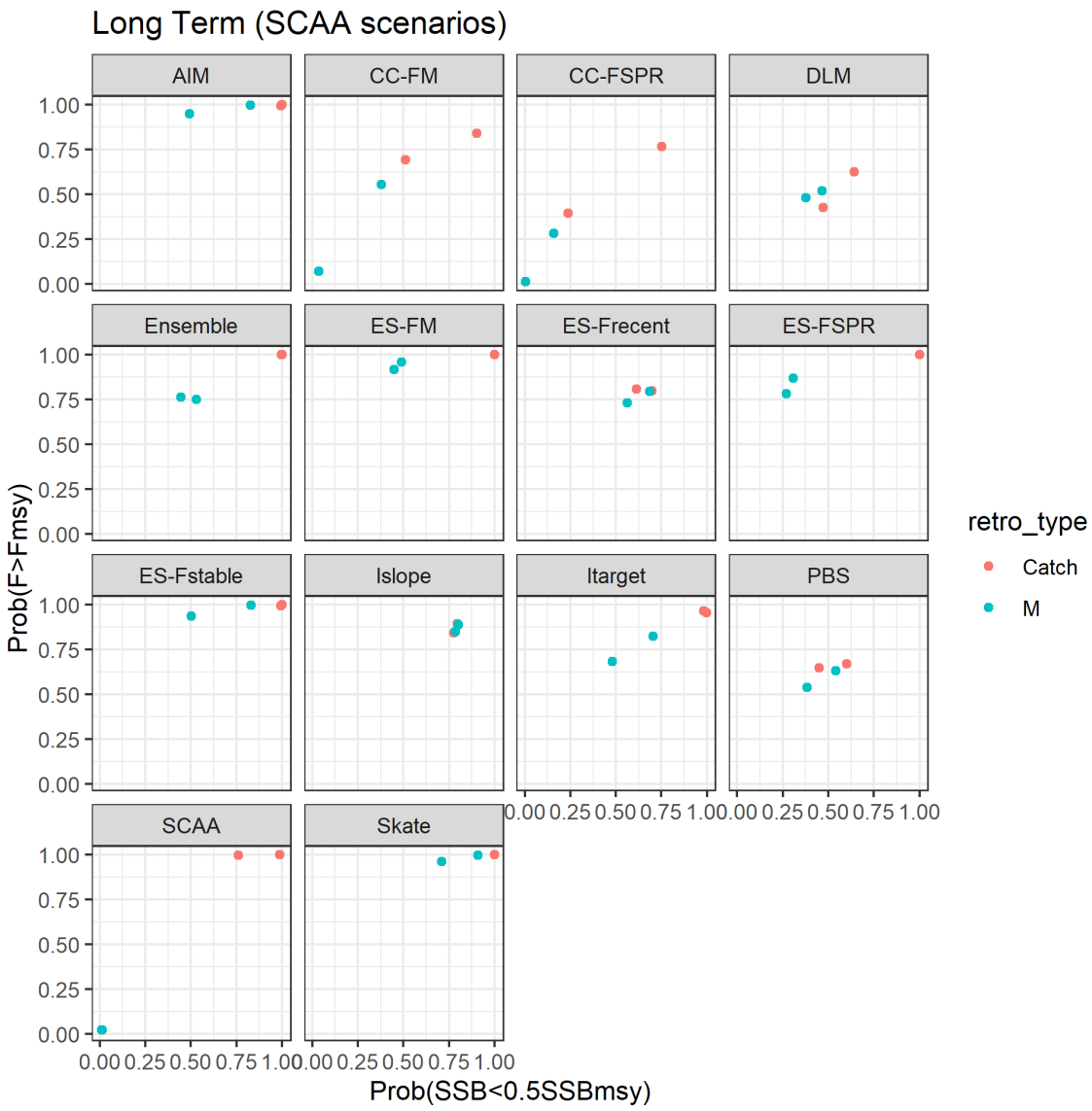


Figure A6.302. Trade off plot by index-based method (IBM) for the statistical catch-at-age (SCAA) analyses between the probability the stock is overfished and the probability that overfishing is occurring in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

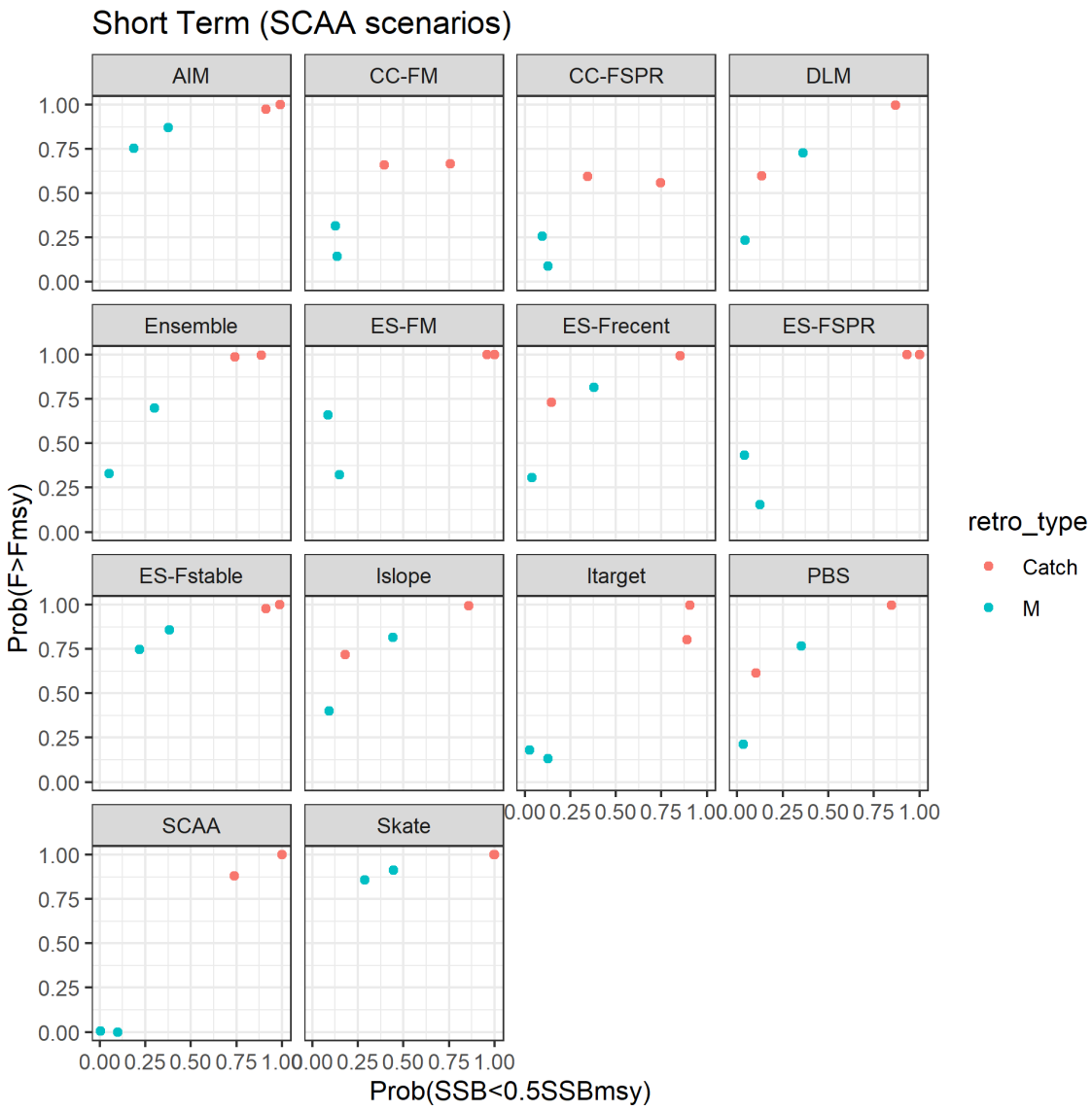


Figure A6.303. Trade off plot by index-based method (IBM) for the statistical catch-at-age (SCAA) analyses between the probability the stock is overfished and the probability that overfishing is occurring in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

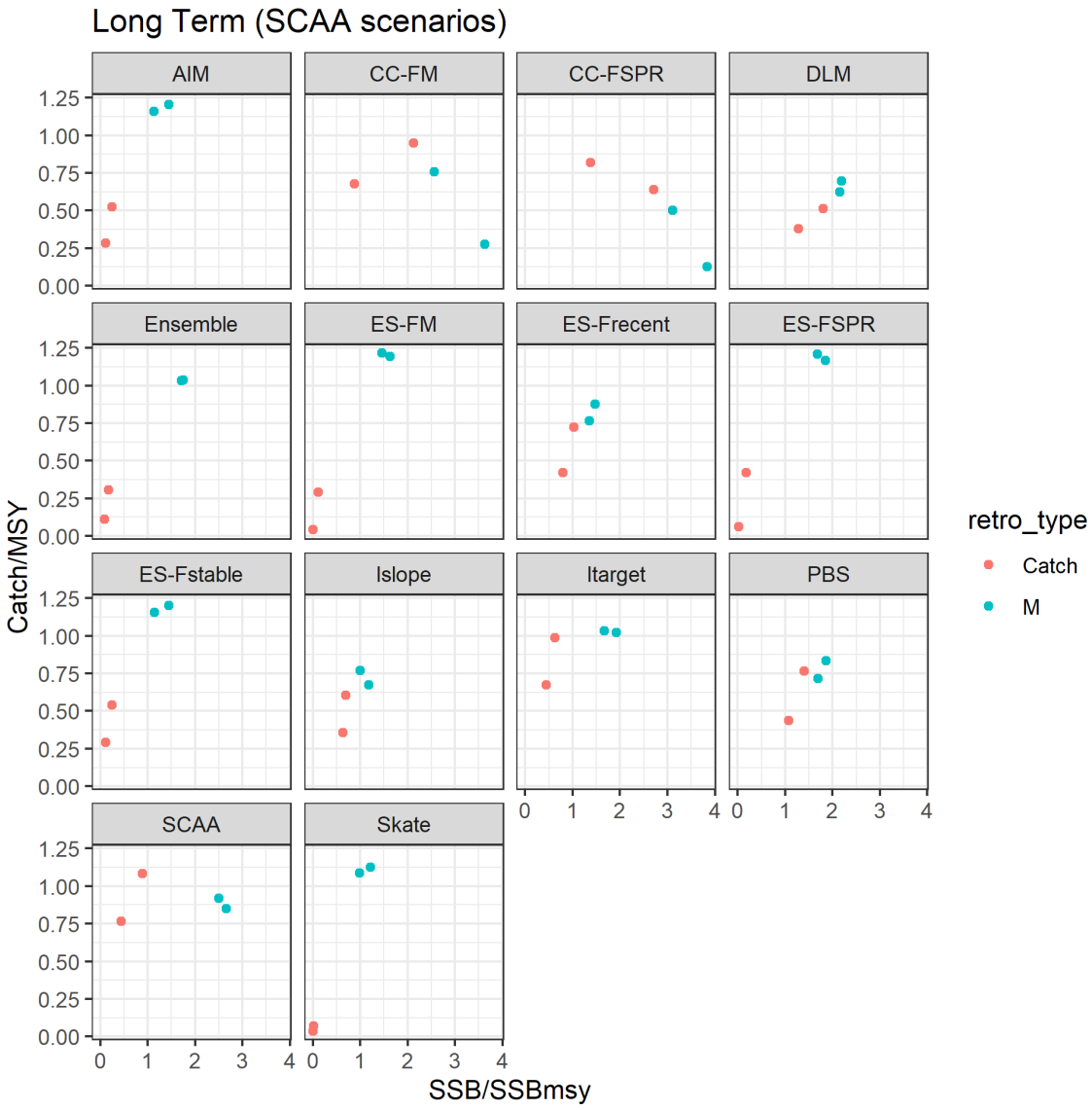


Figure A6.304. Trade off plot by index-based method (IBM) for the statistical catch-at-age (SCAA) analyses between the average spawning stock biomass (SSB) and catch relative to their maximum sustainable yield (MSY) reference points in the long term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

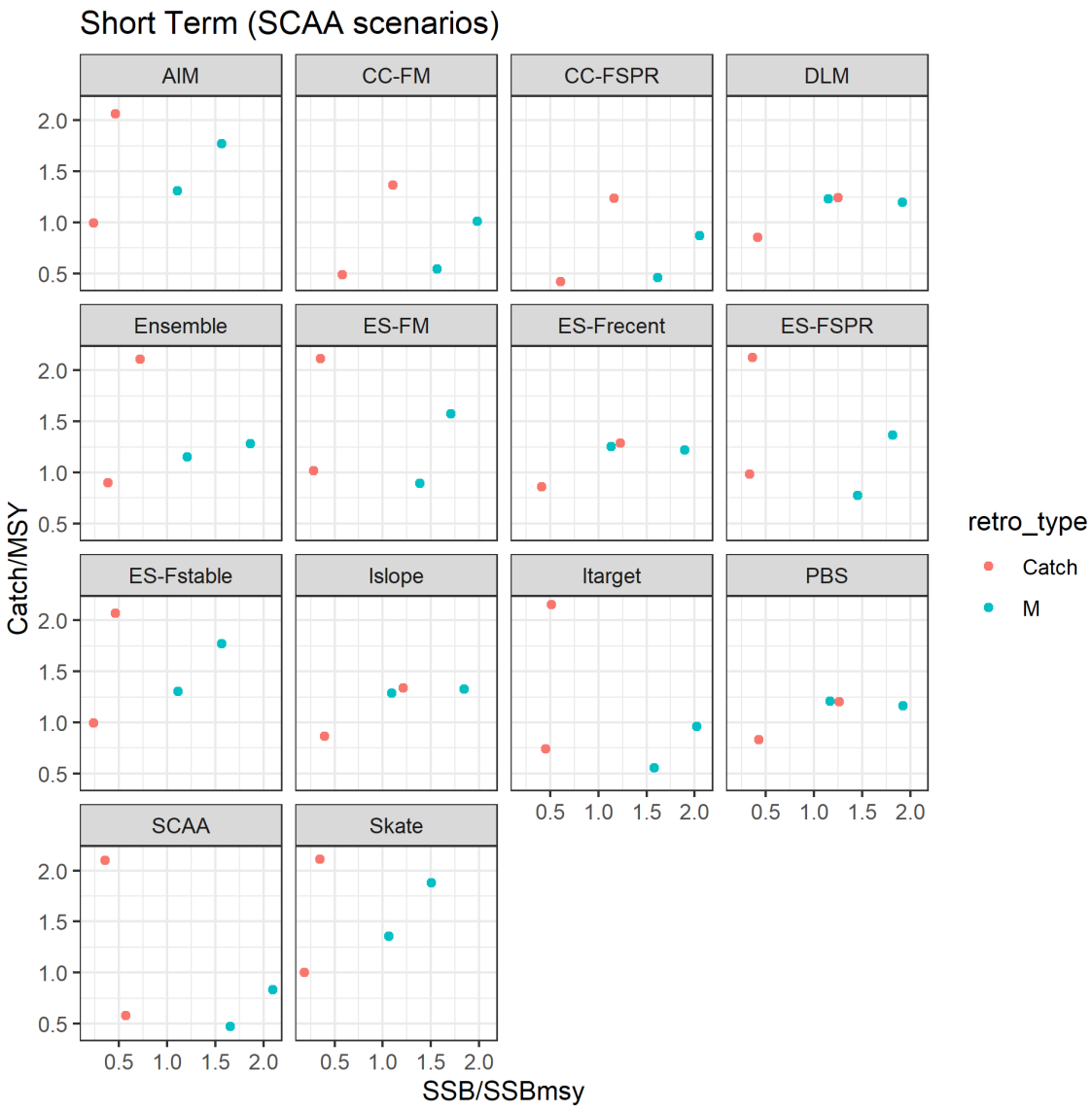


Figure A6.305. Trade off plot by index-based method (IBM) for the statistical catch-at-age (SCAA) analyses between the average spawning stock biomass (SSB) and catch relative to their maximum sustainable yield (MSY) reference points in the short term. Each point represents 1 scenario with the color indicating the source of the retrospective pattern for that scenario.

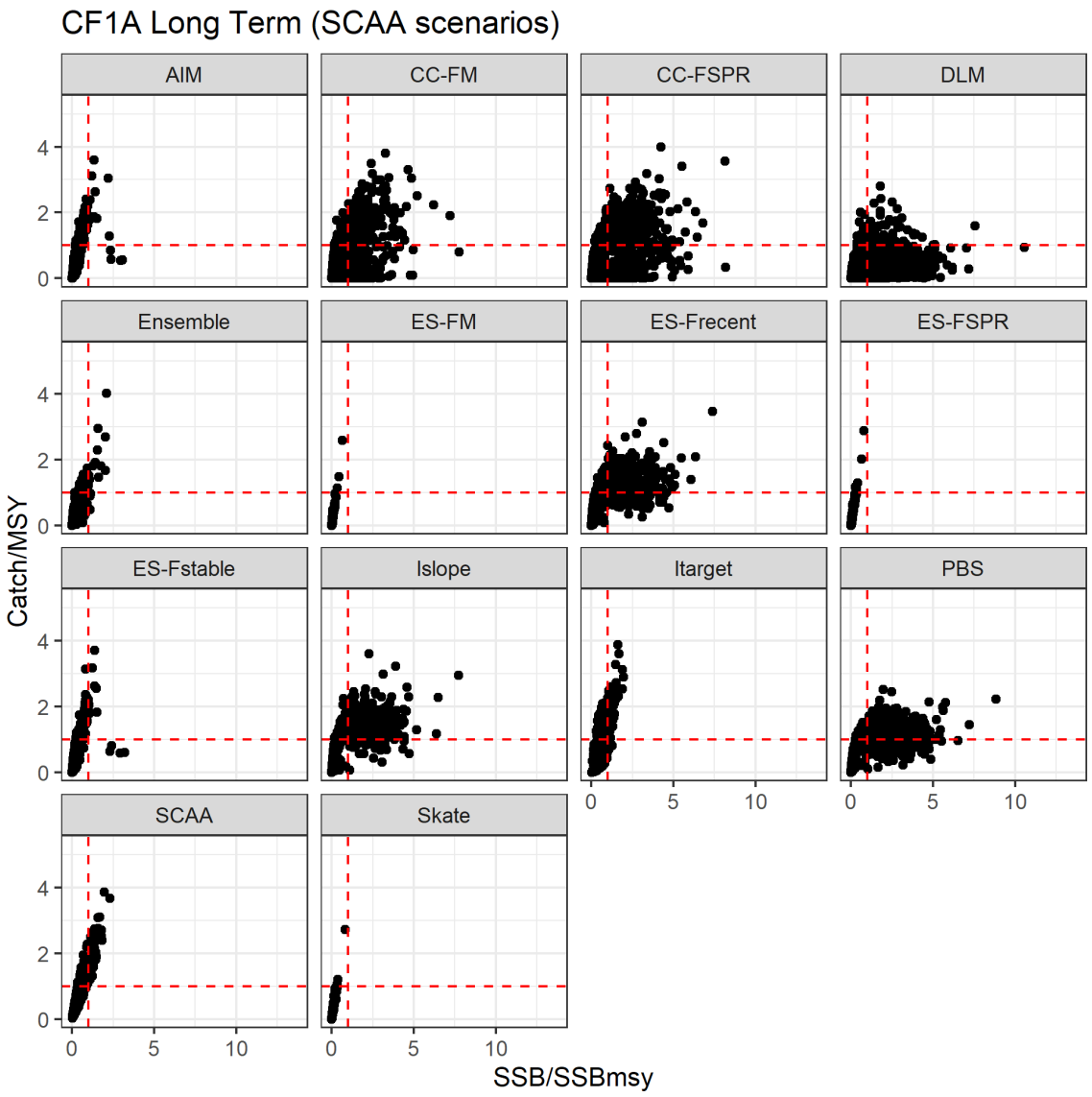


Figure A6.306. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

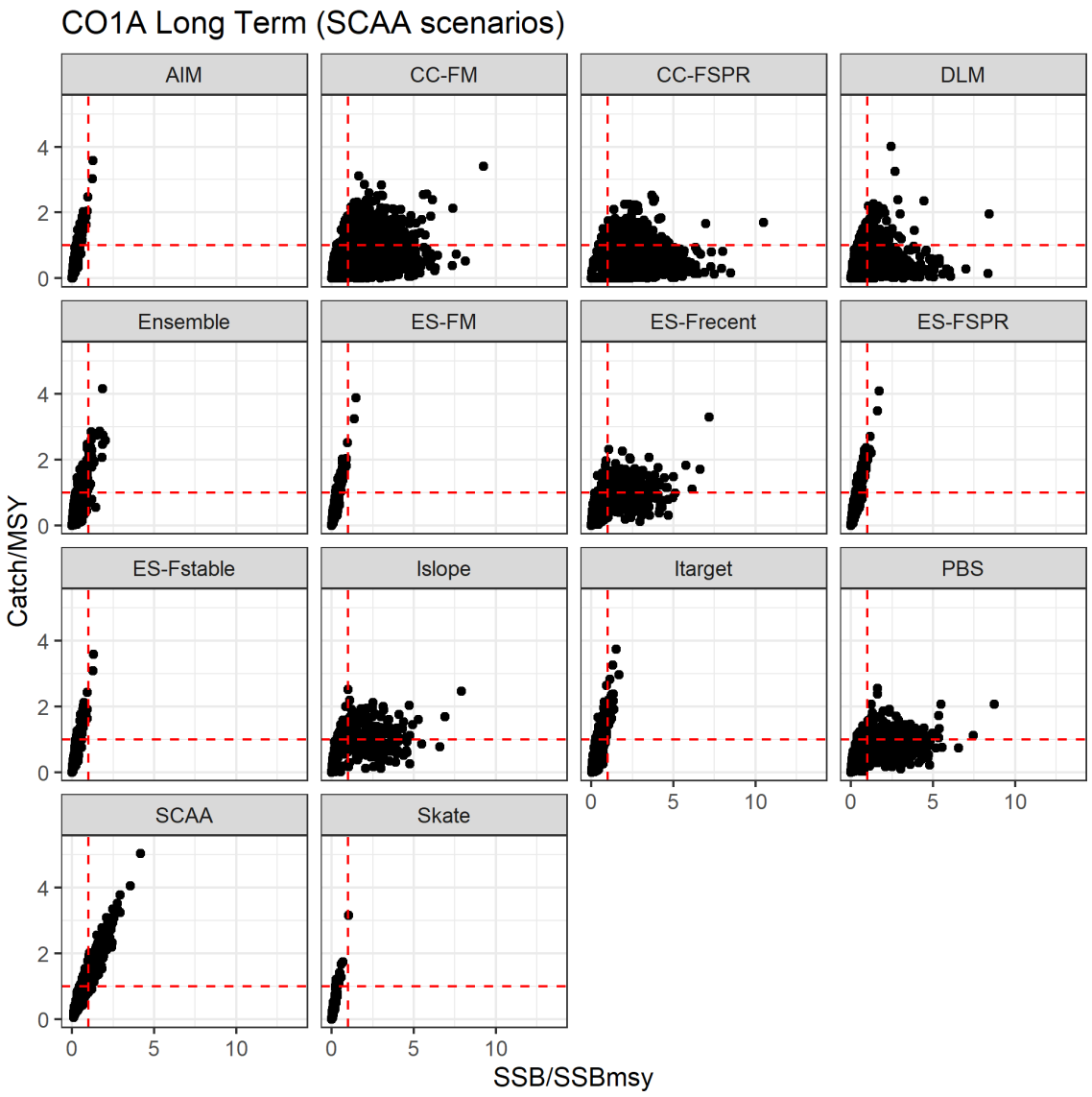


Figure A6.307. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

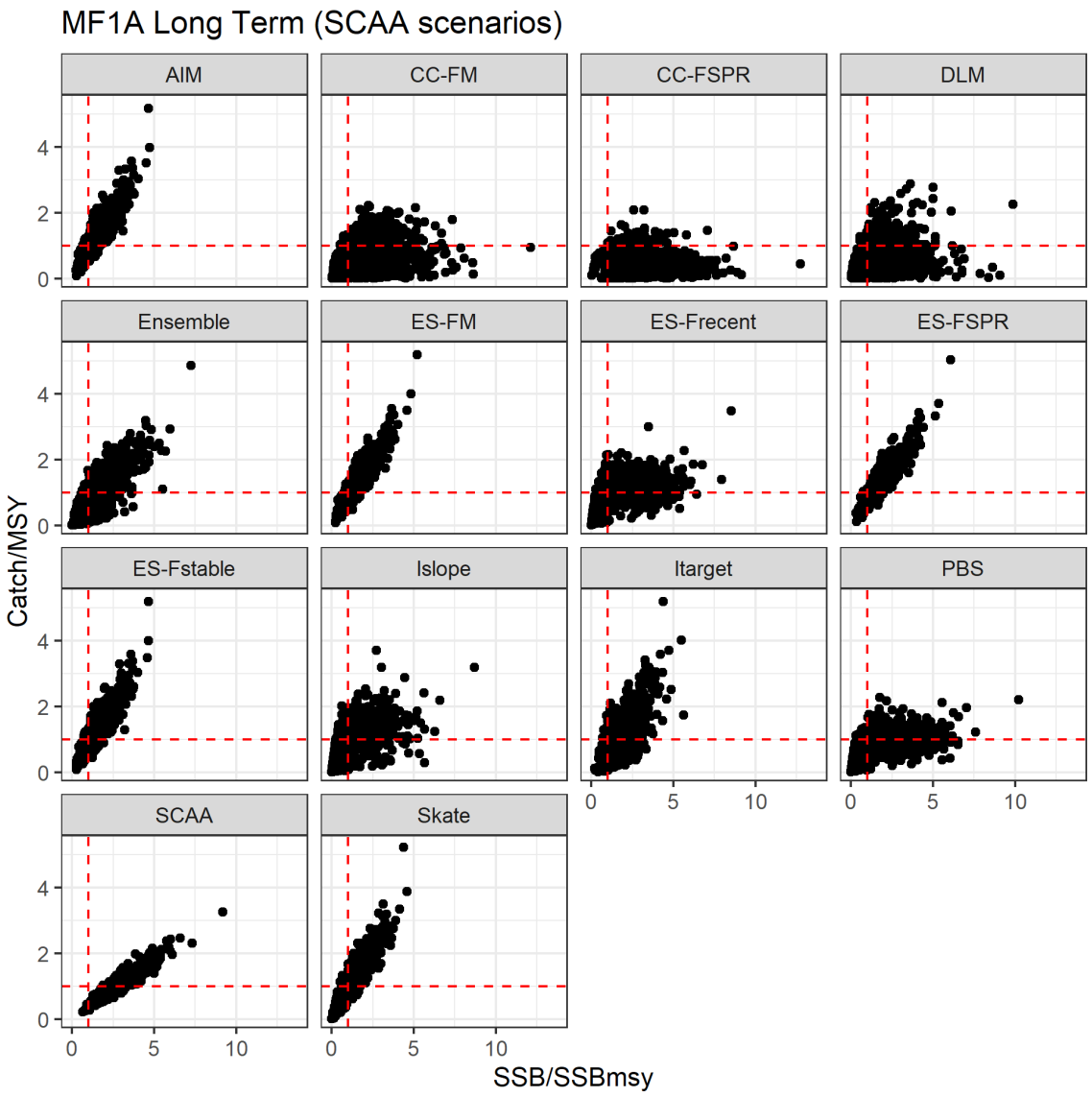


Figure A6.308. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

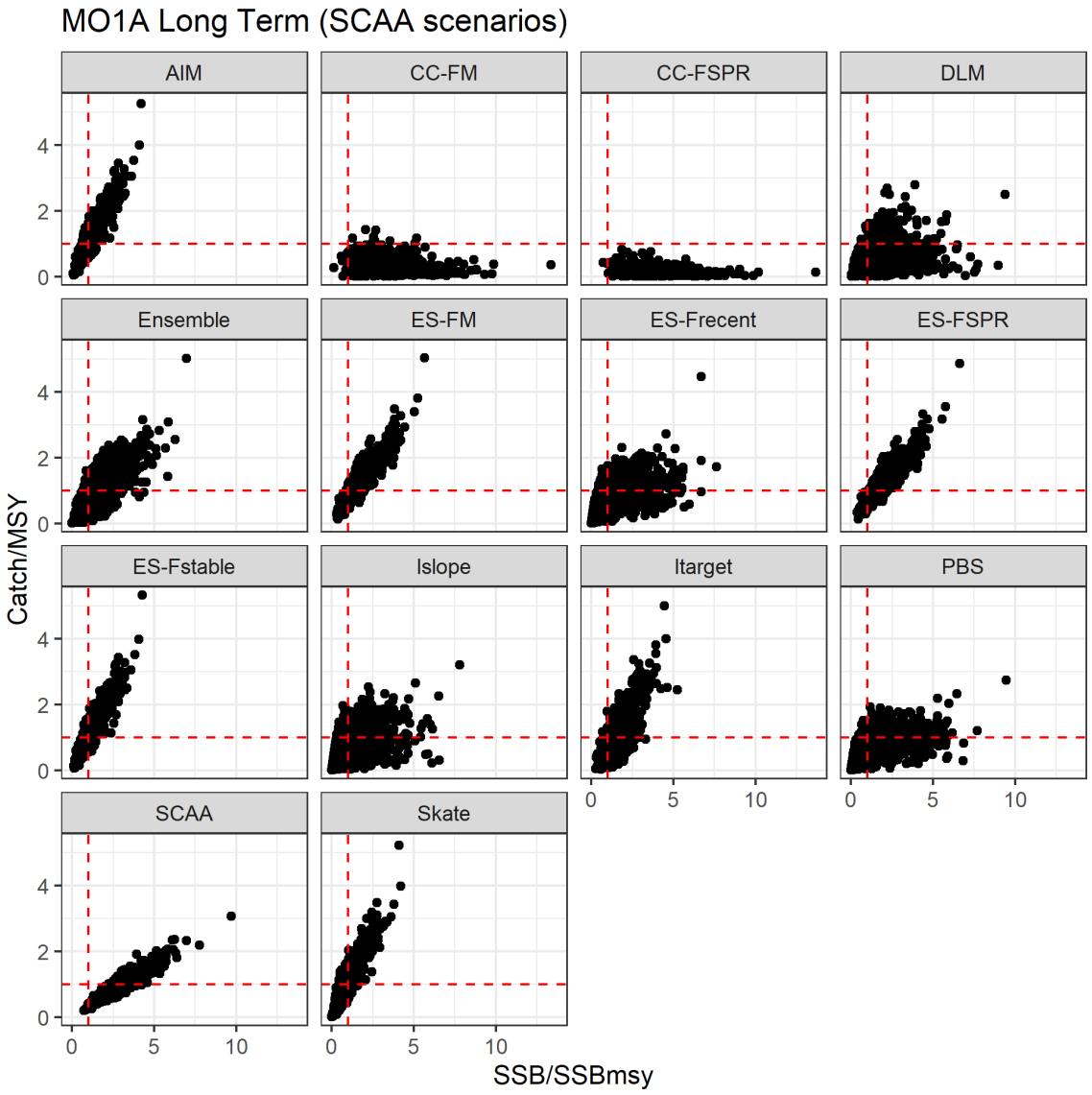


Figure A6.309. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

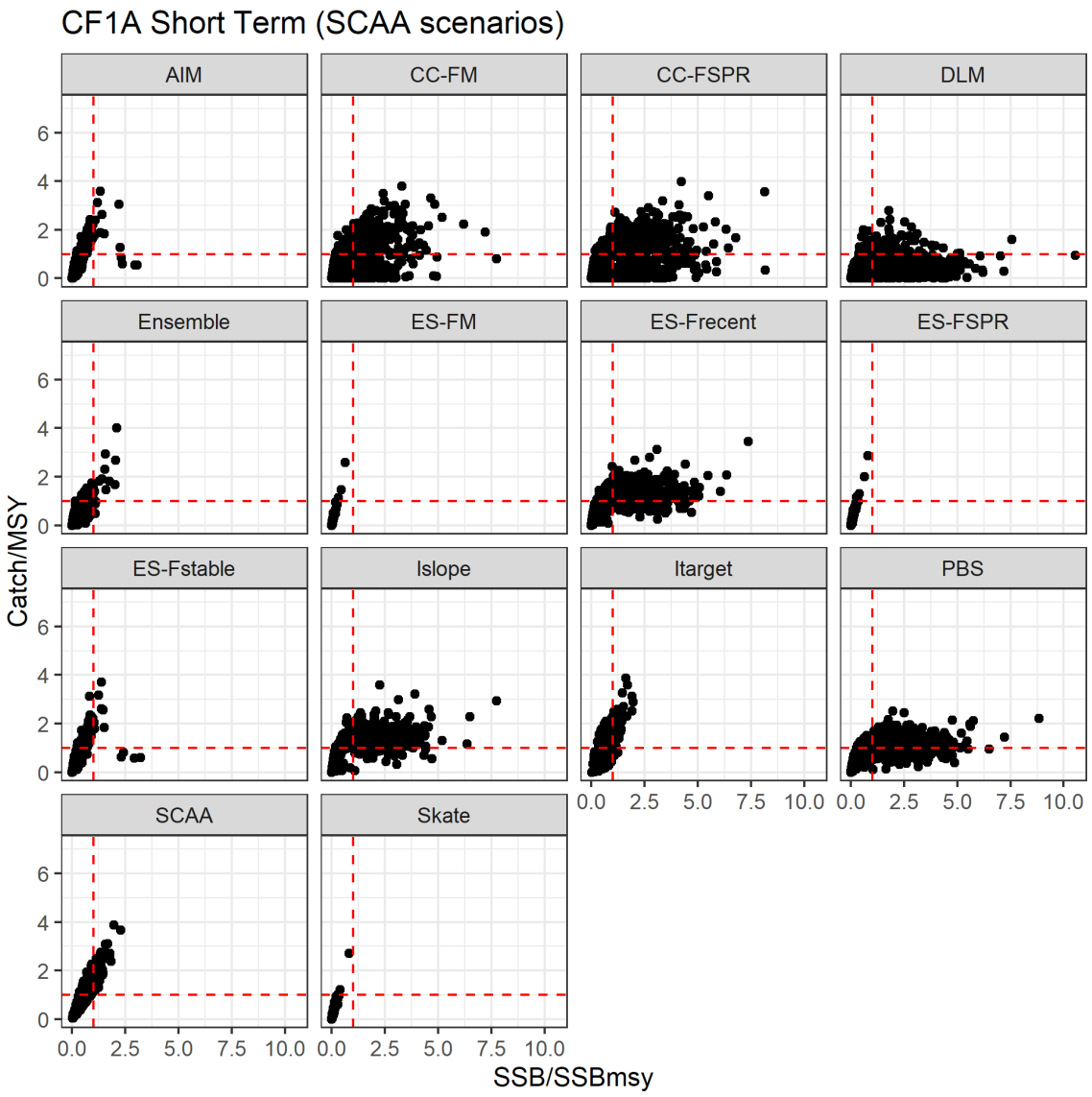


Figure A6.310. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

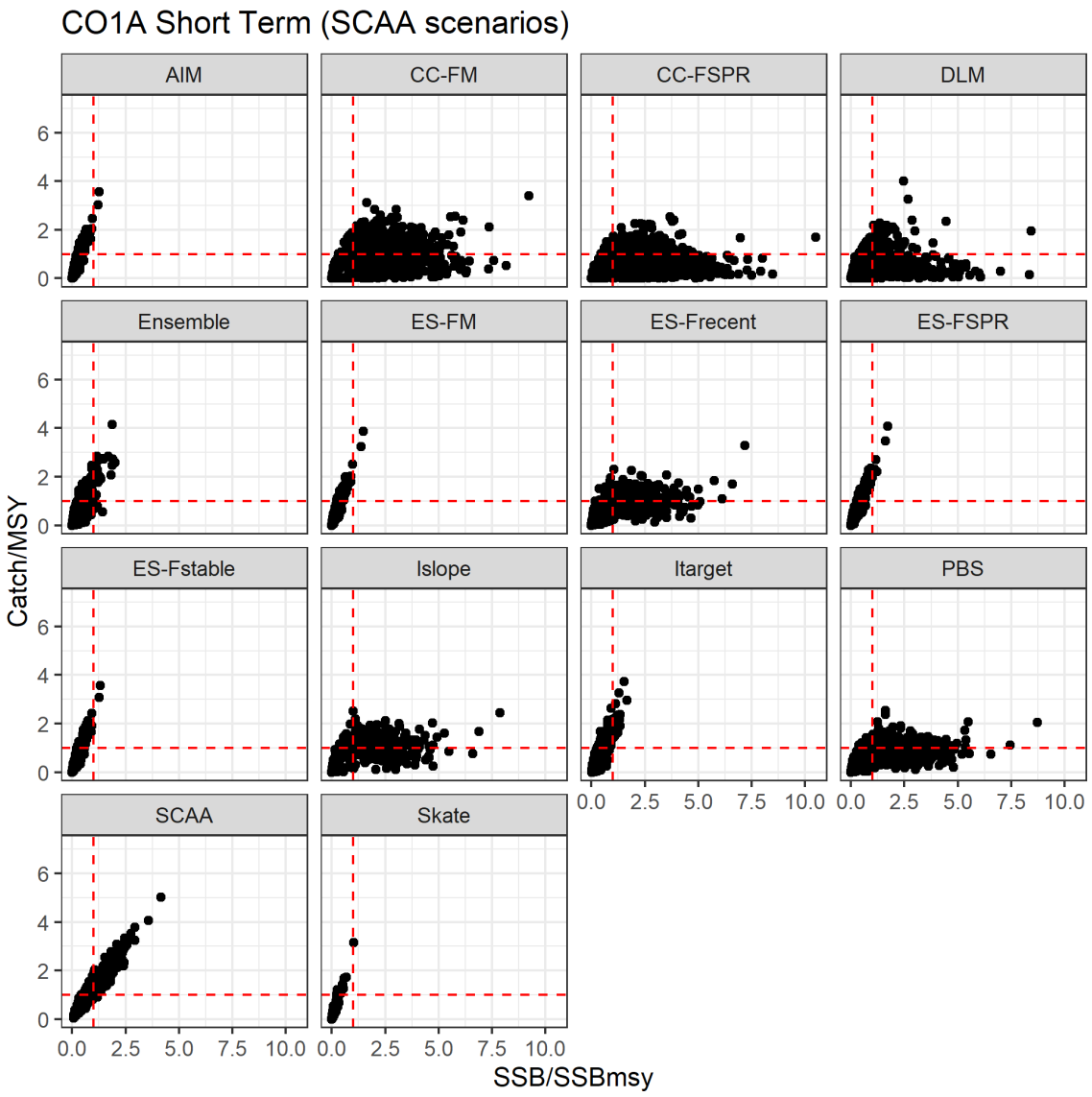


Figure A6.311. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

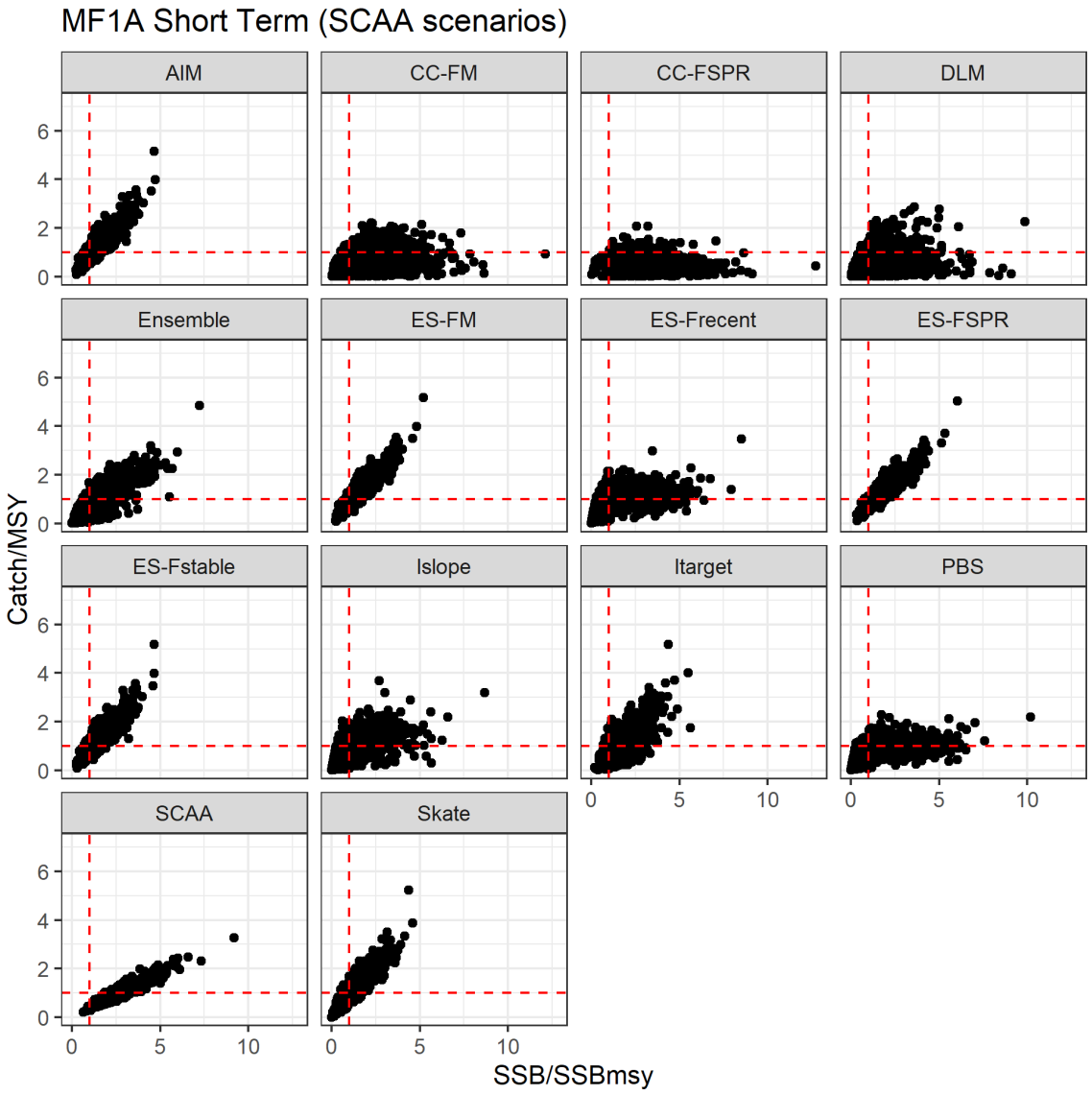


Figure A6.312. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

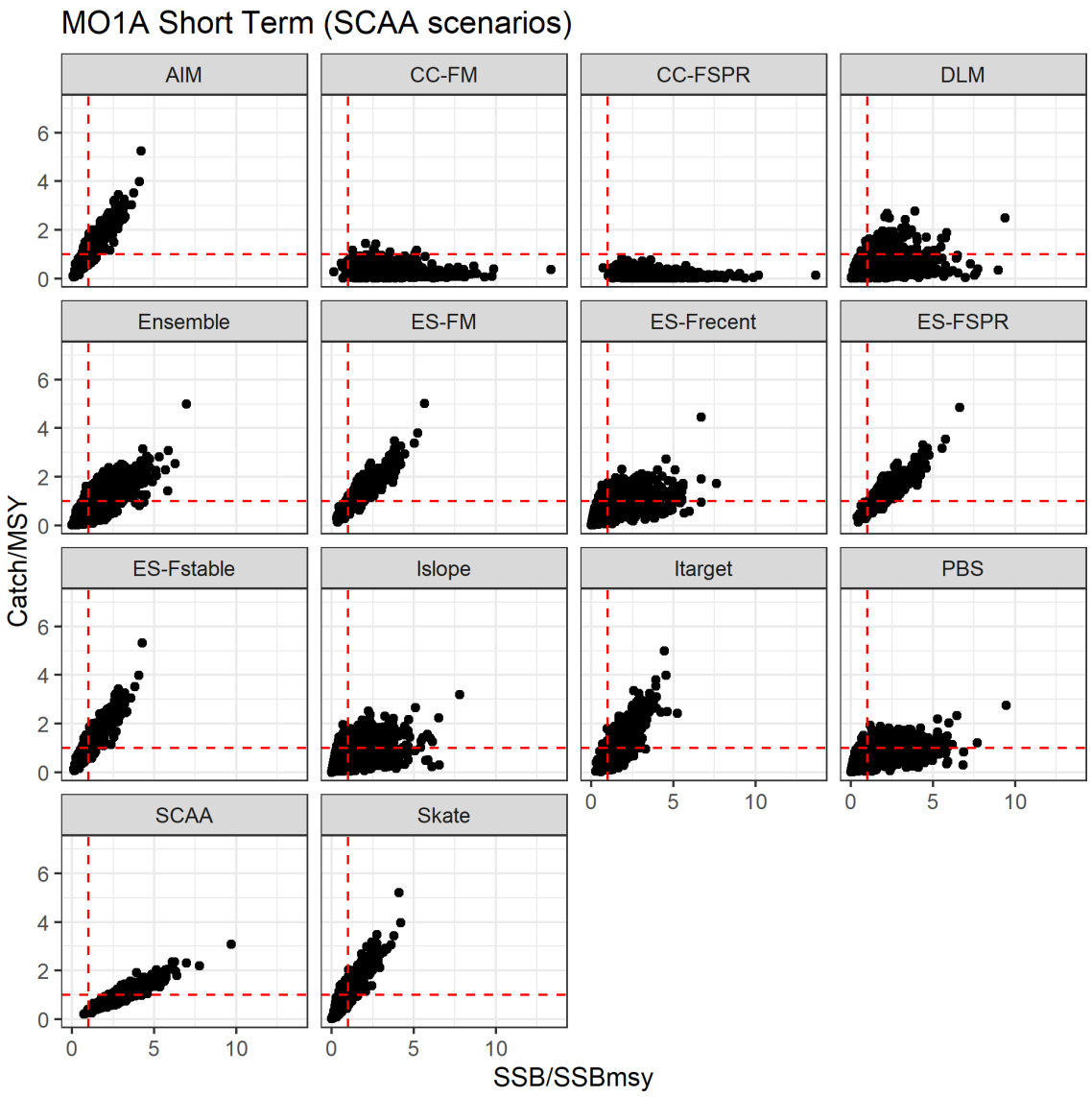


Figure A6.313. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the scenario defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

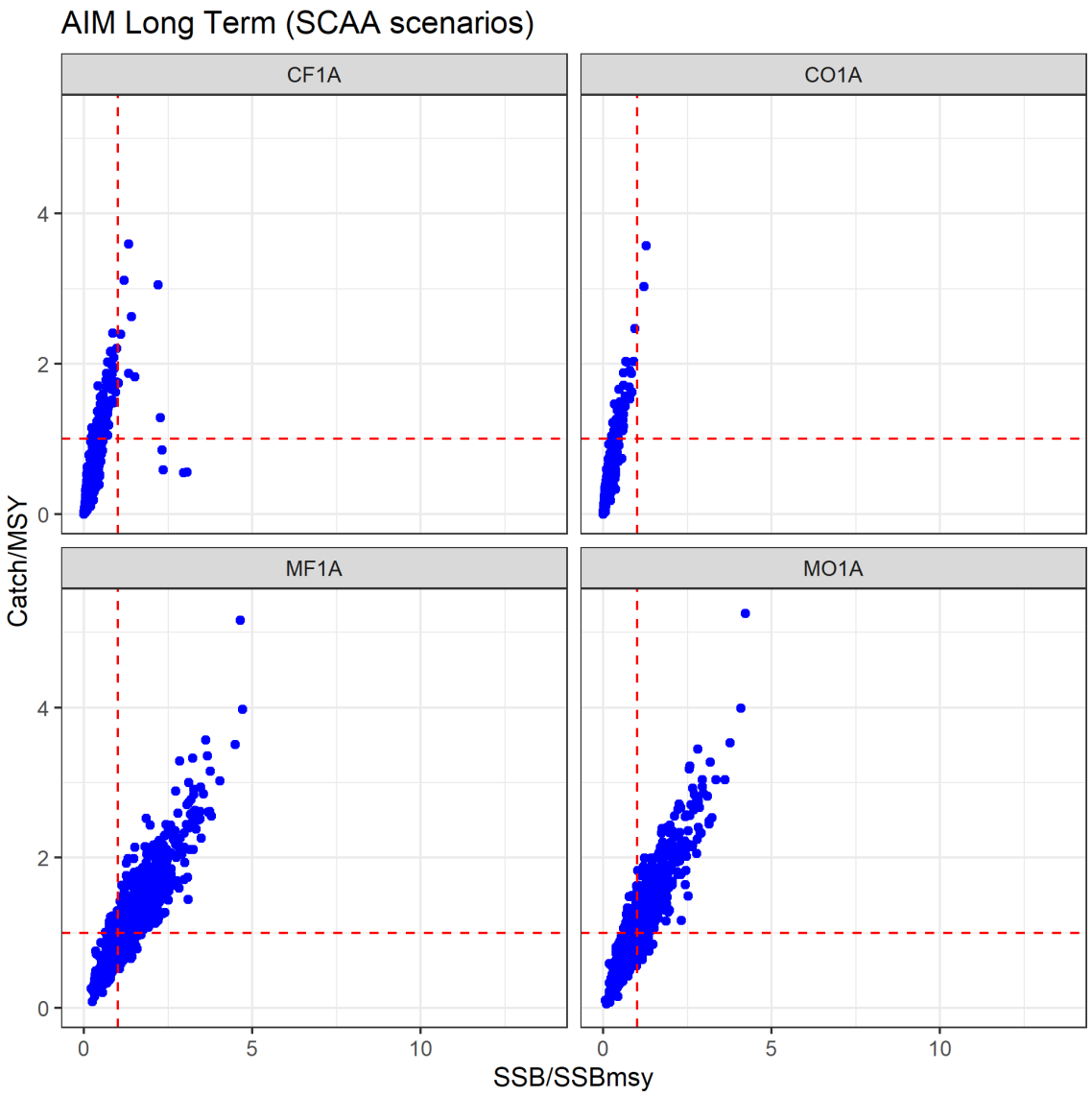


Figure A6.314. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.



Figure A6.315. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.



Figure A6.316. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

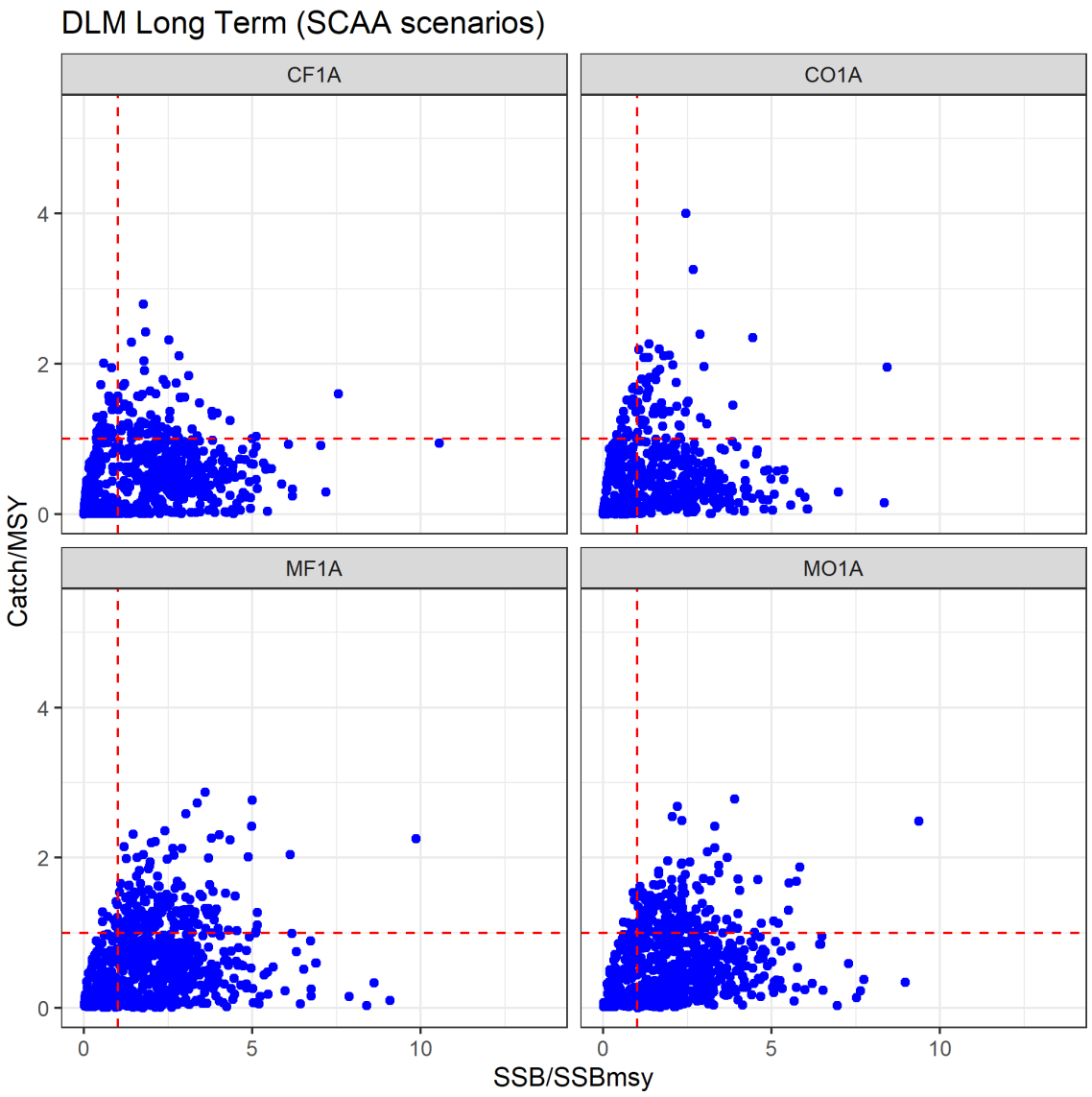


Figure A6.317. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

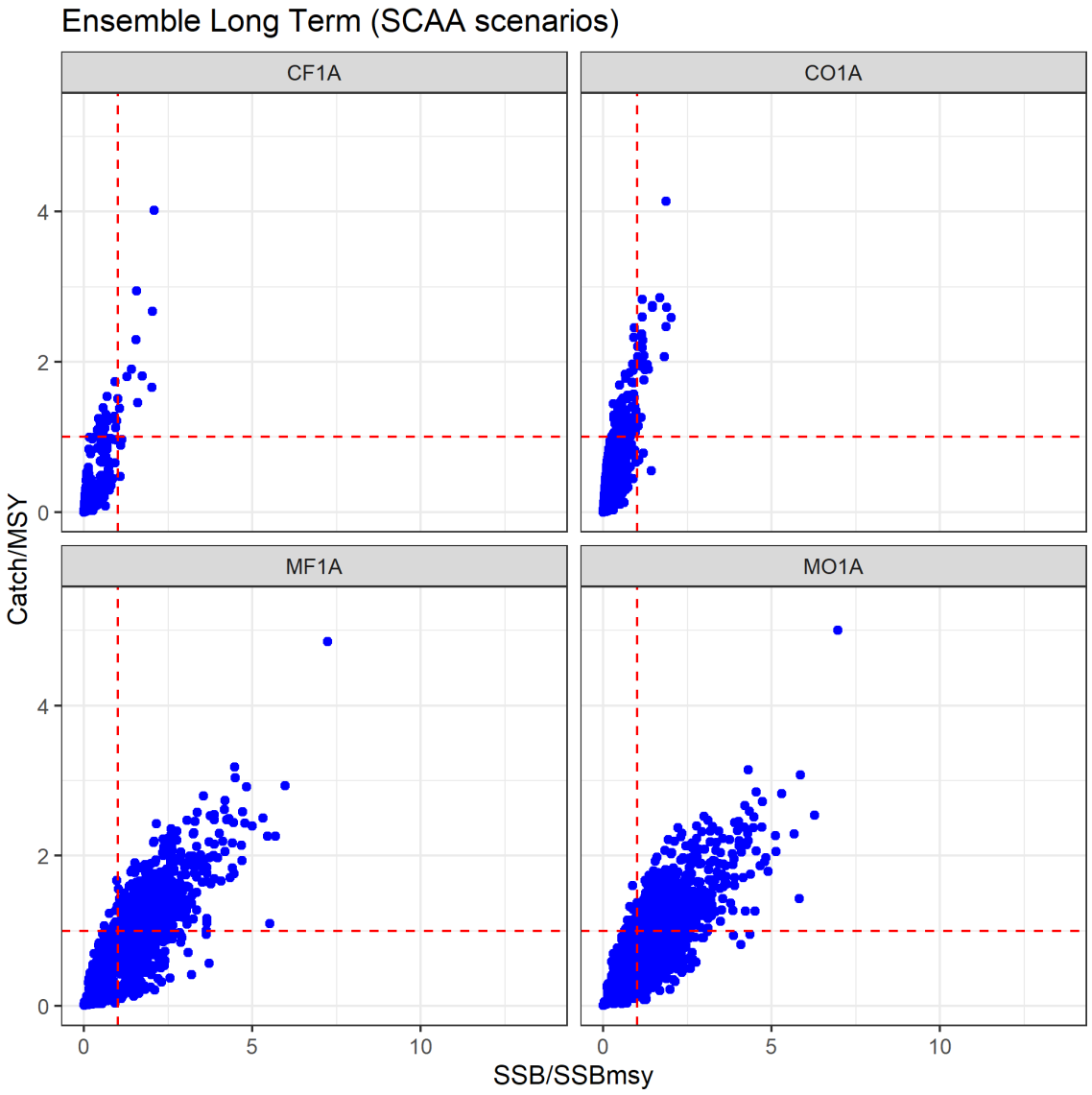


Figure A6.318. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

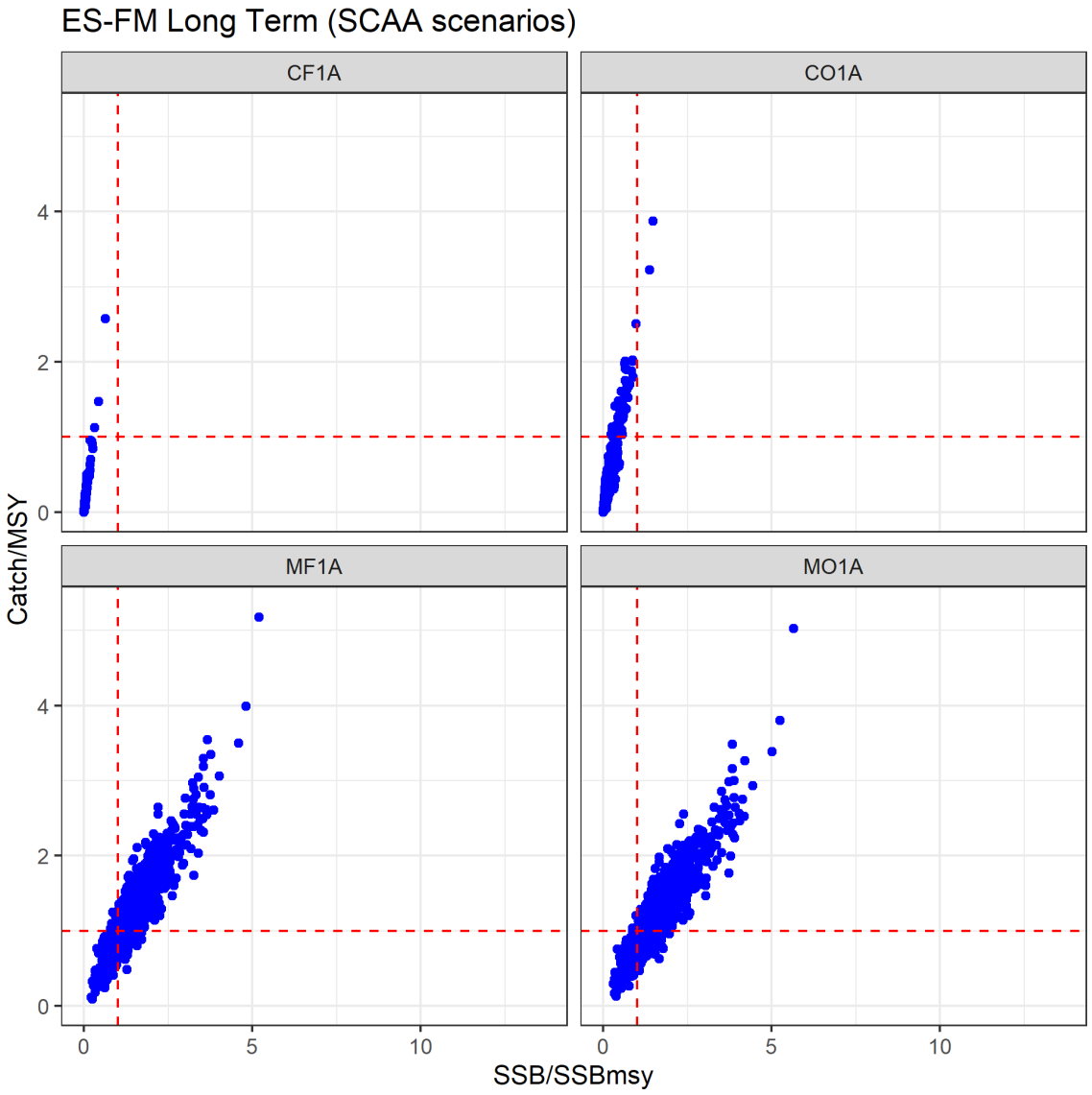


Figure A6.319. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

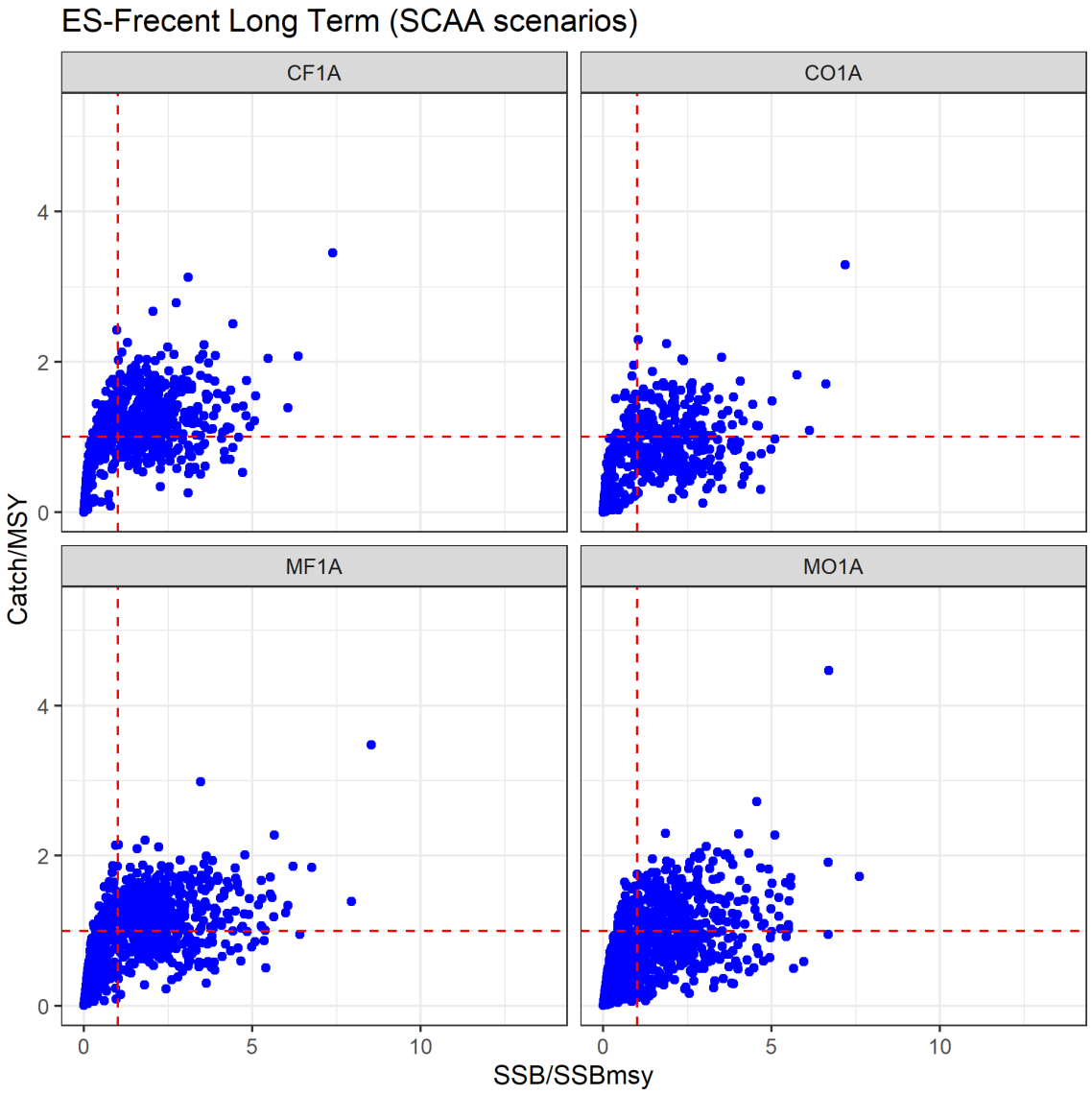


Figure A6.320. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

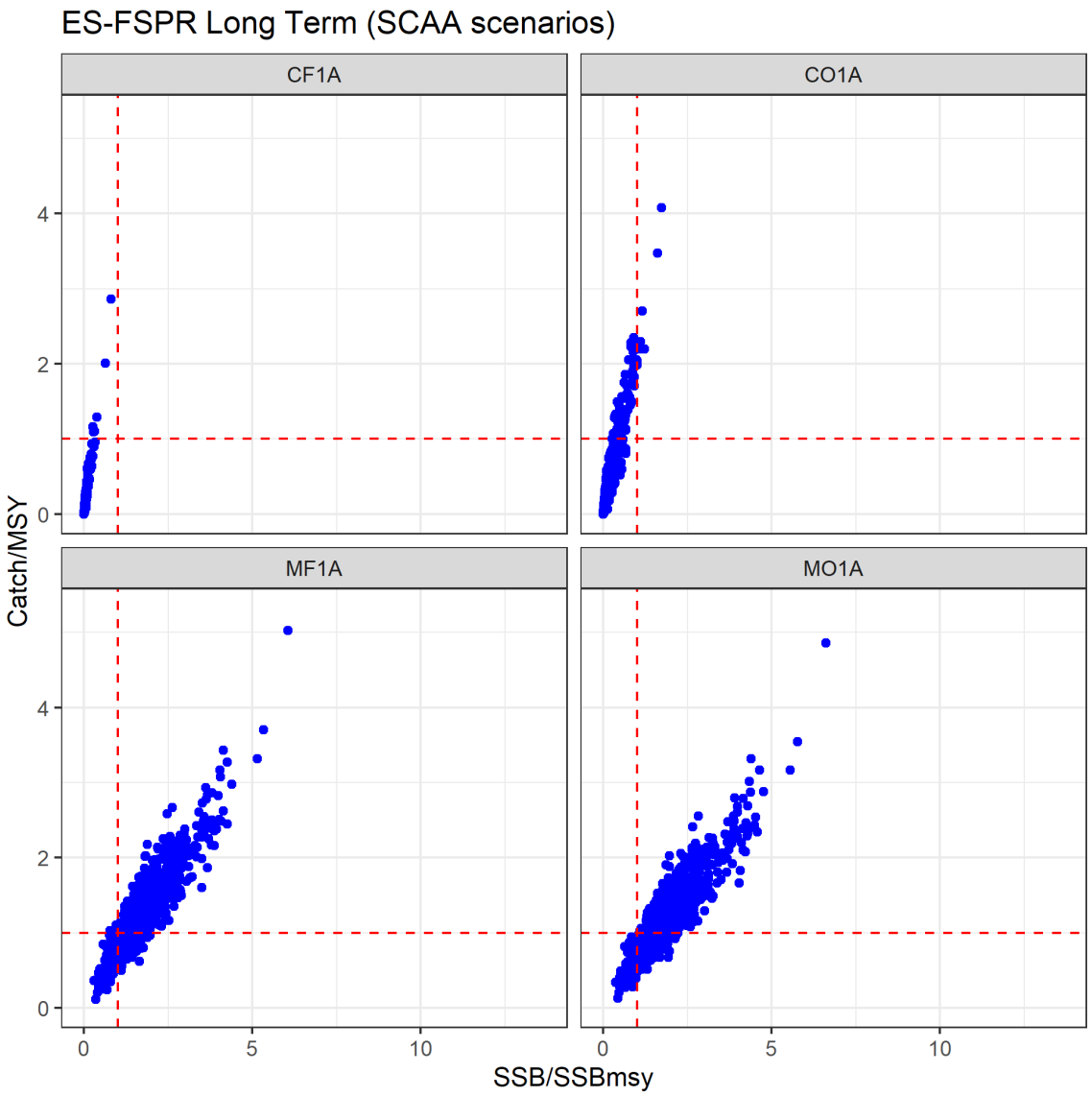


Figure A6.321. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

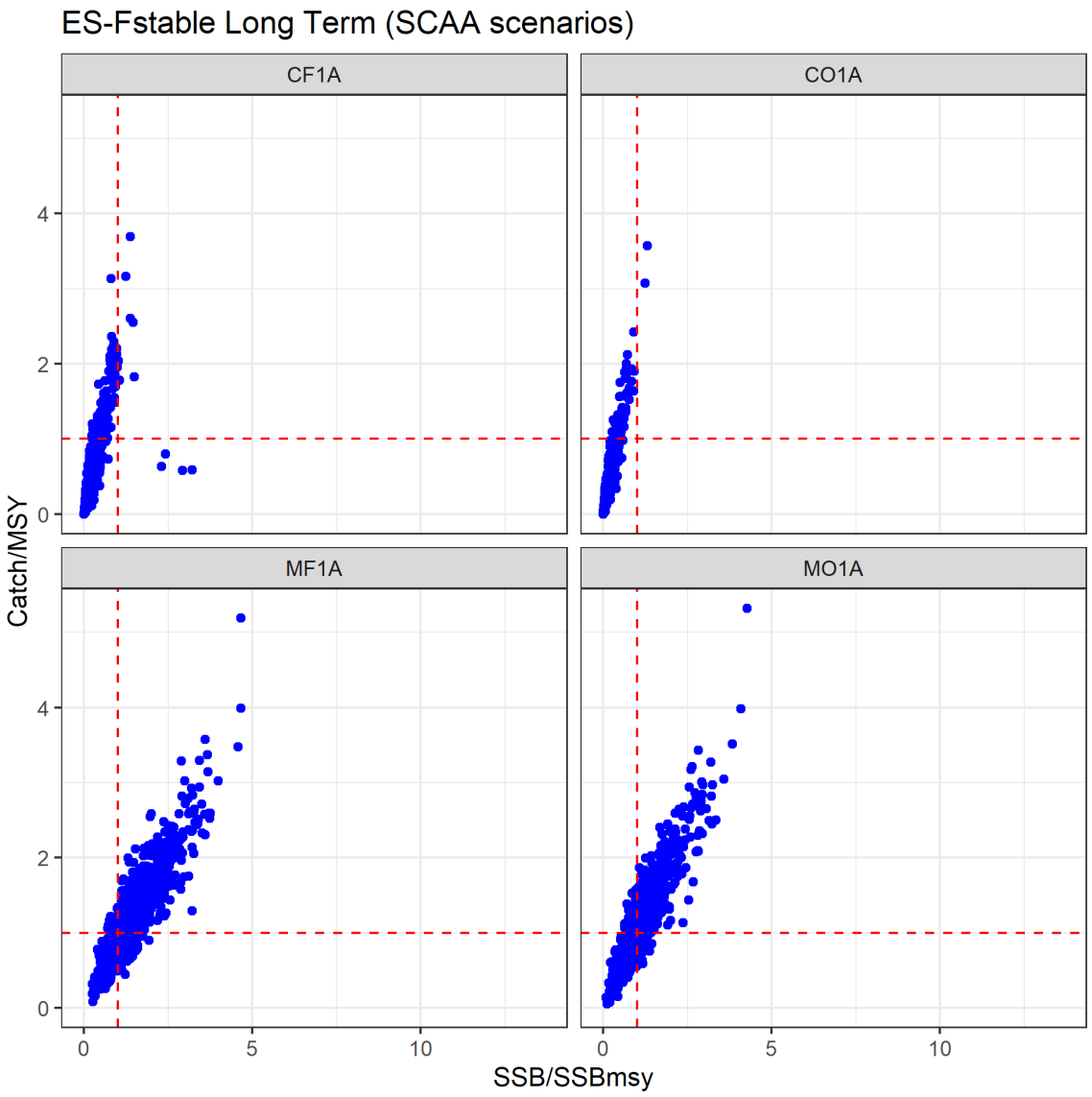


Figure A6.322. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

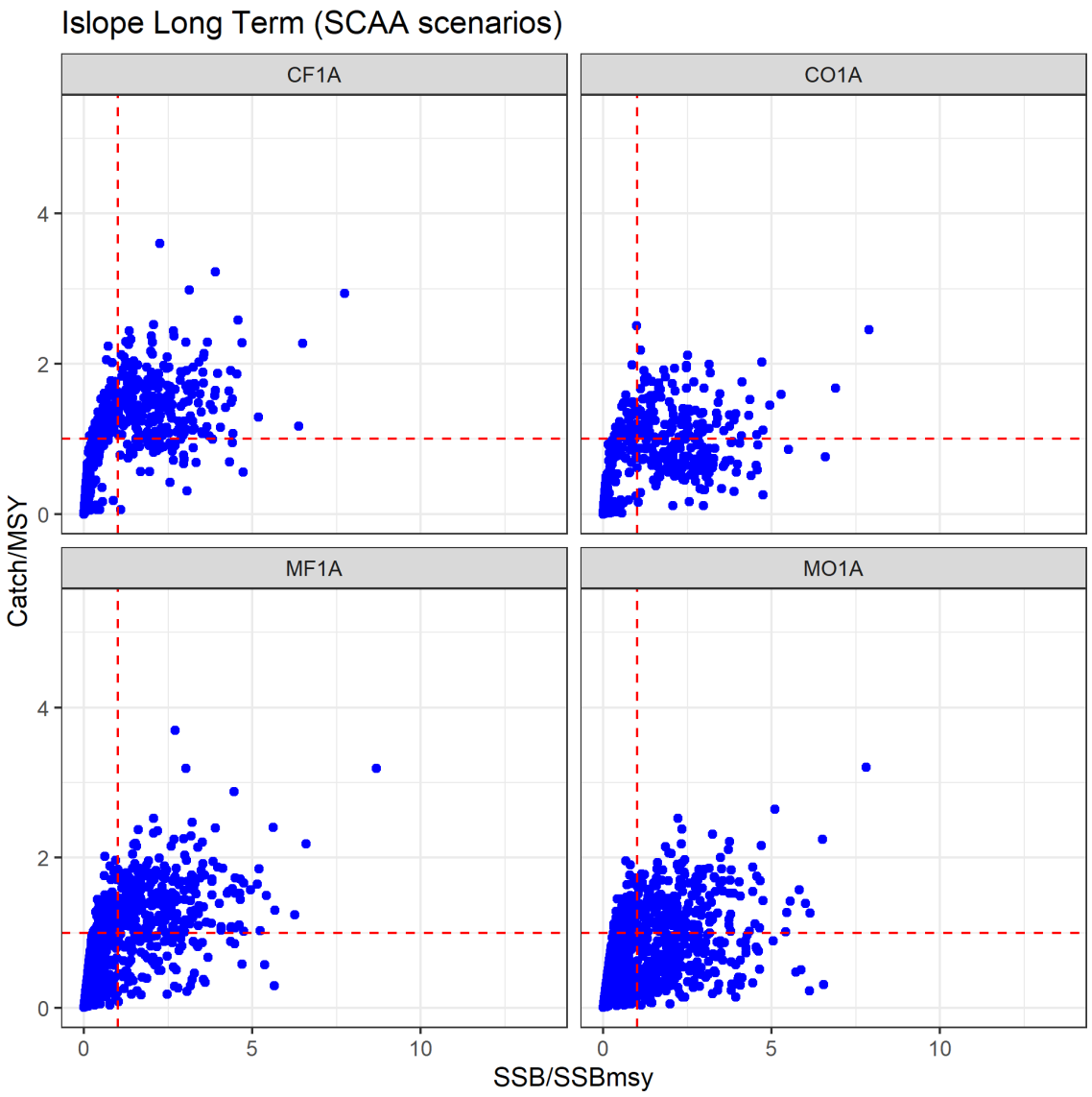


Figure A6.323. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

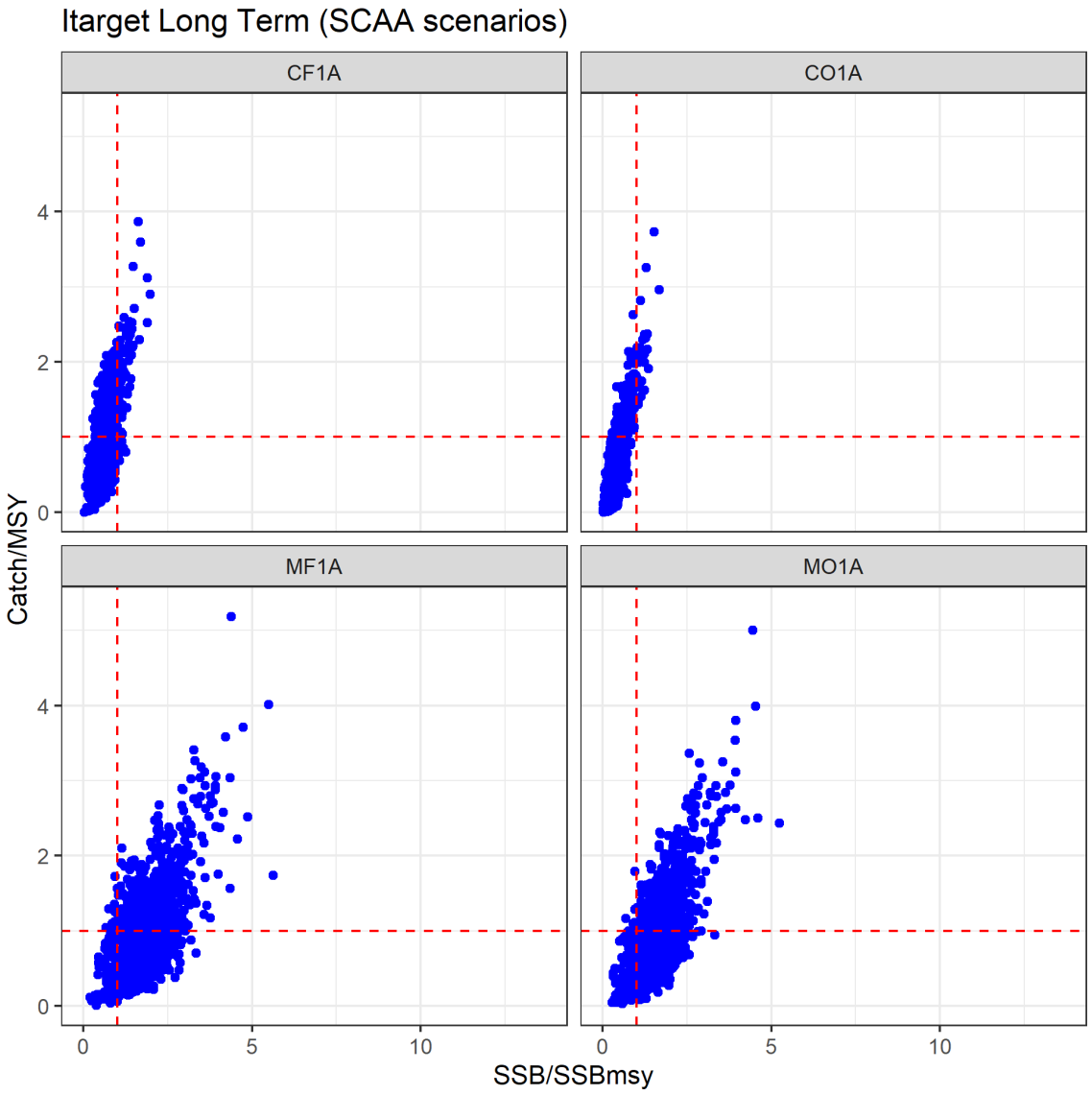


Figure A6.324. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

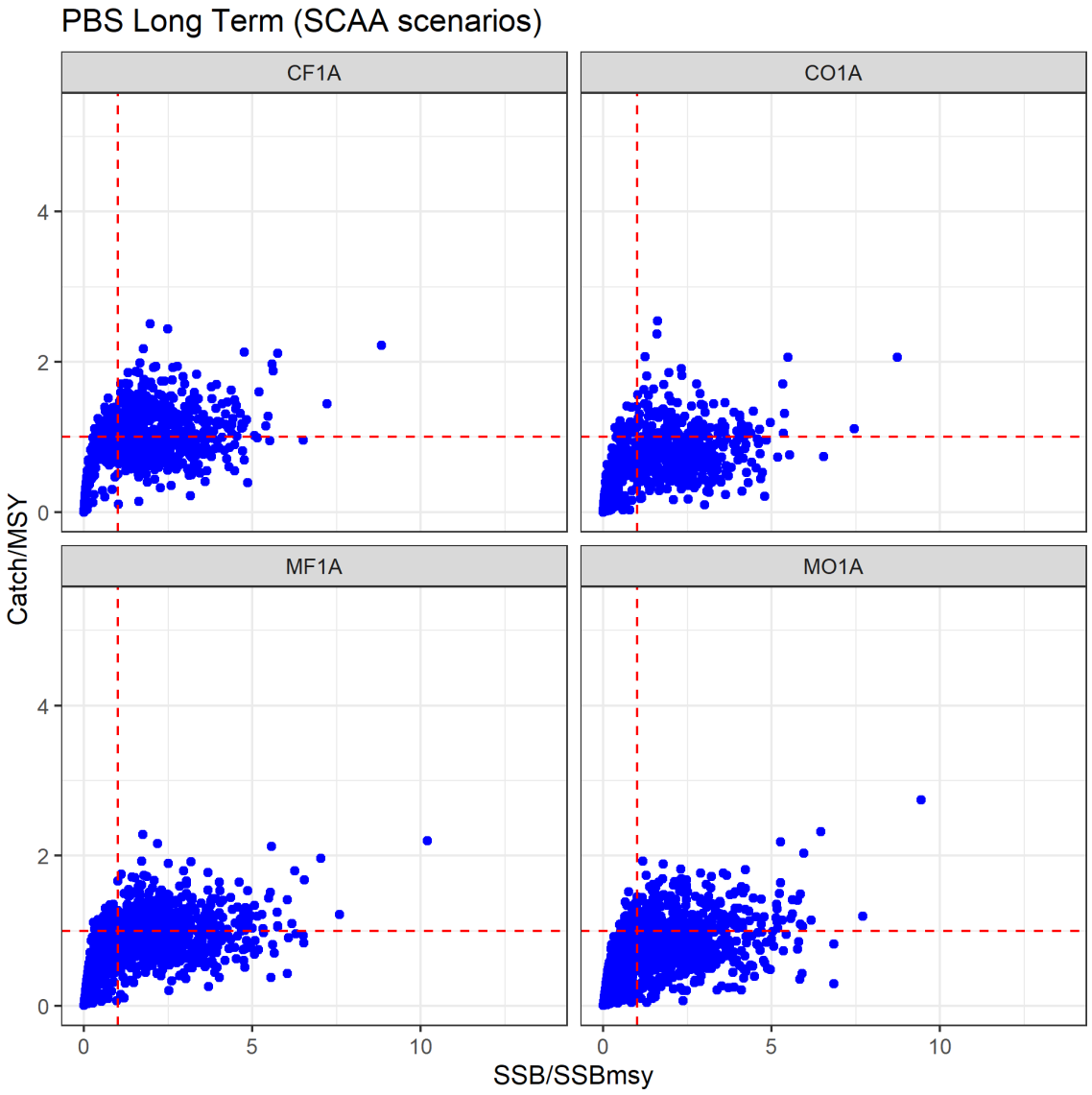


Figure A6.325. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

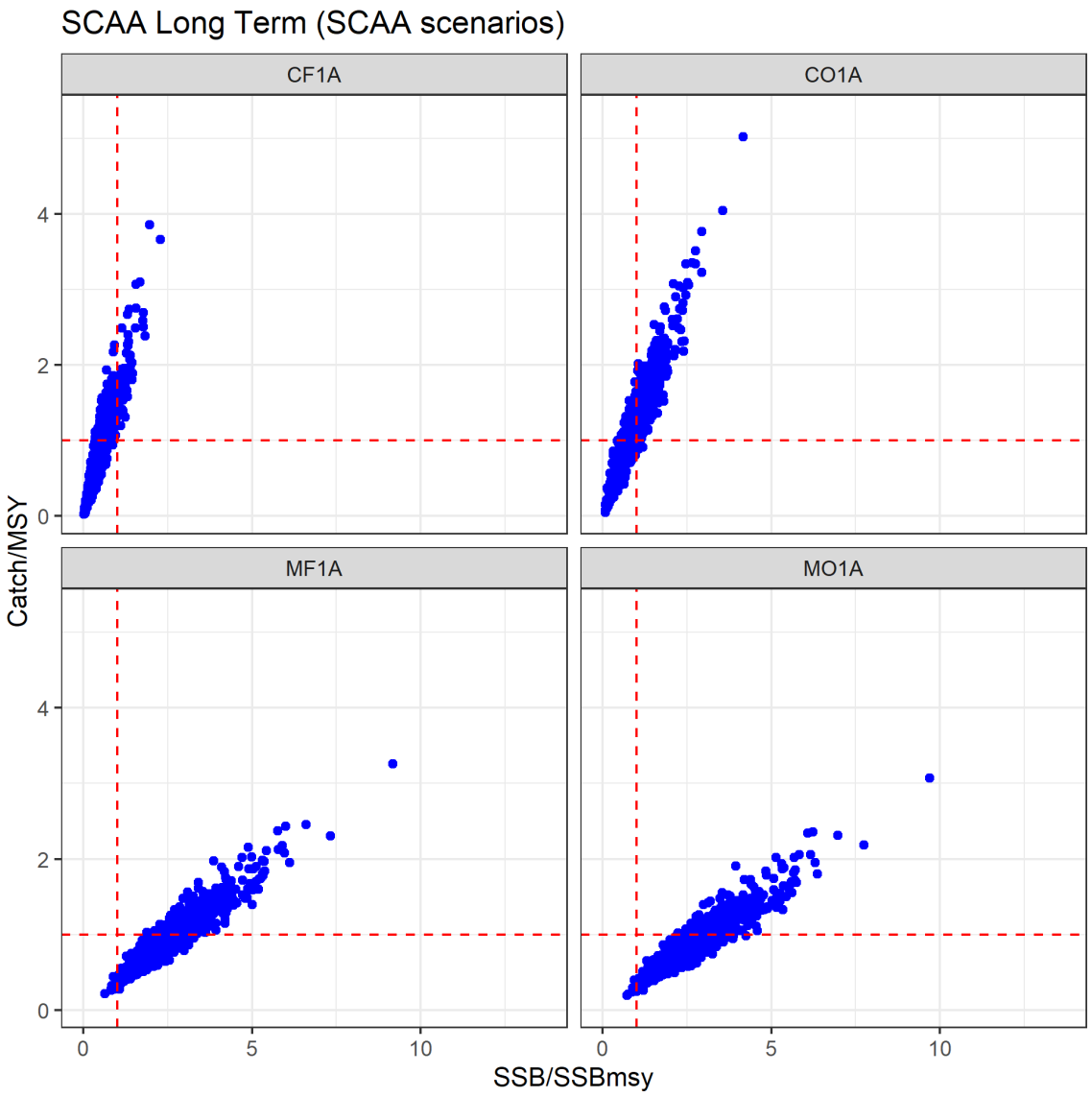


Figure A6.326. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

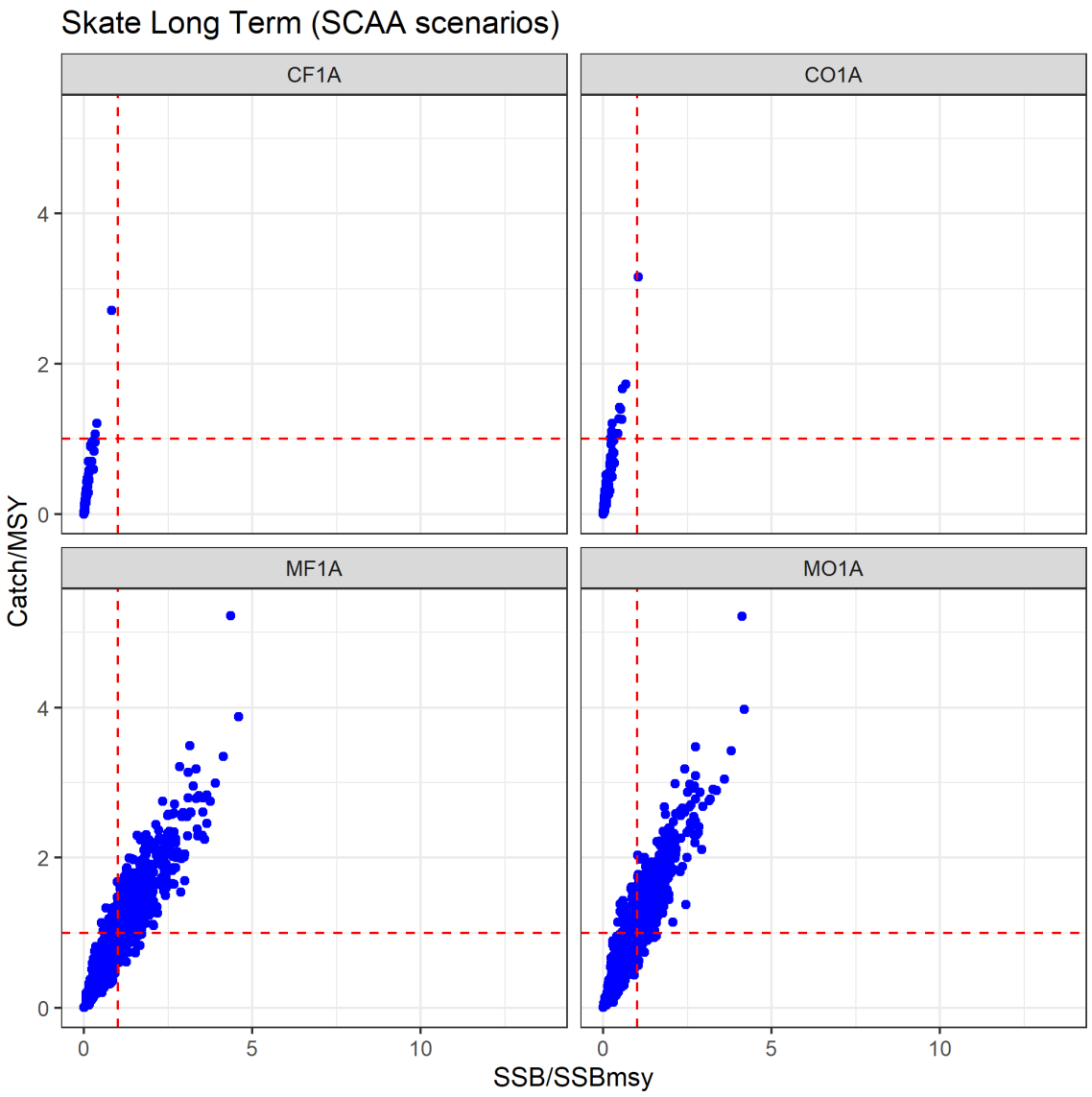


Figure A6.327. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the long term with each dot representing 1 of the 1,000 simulations.

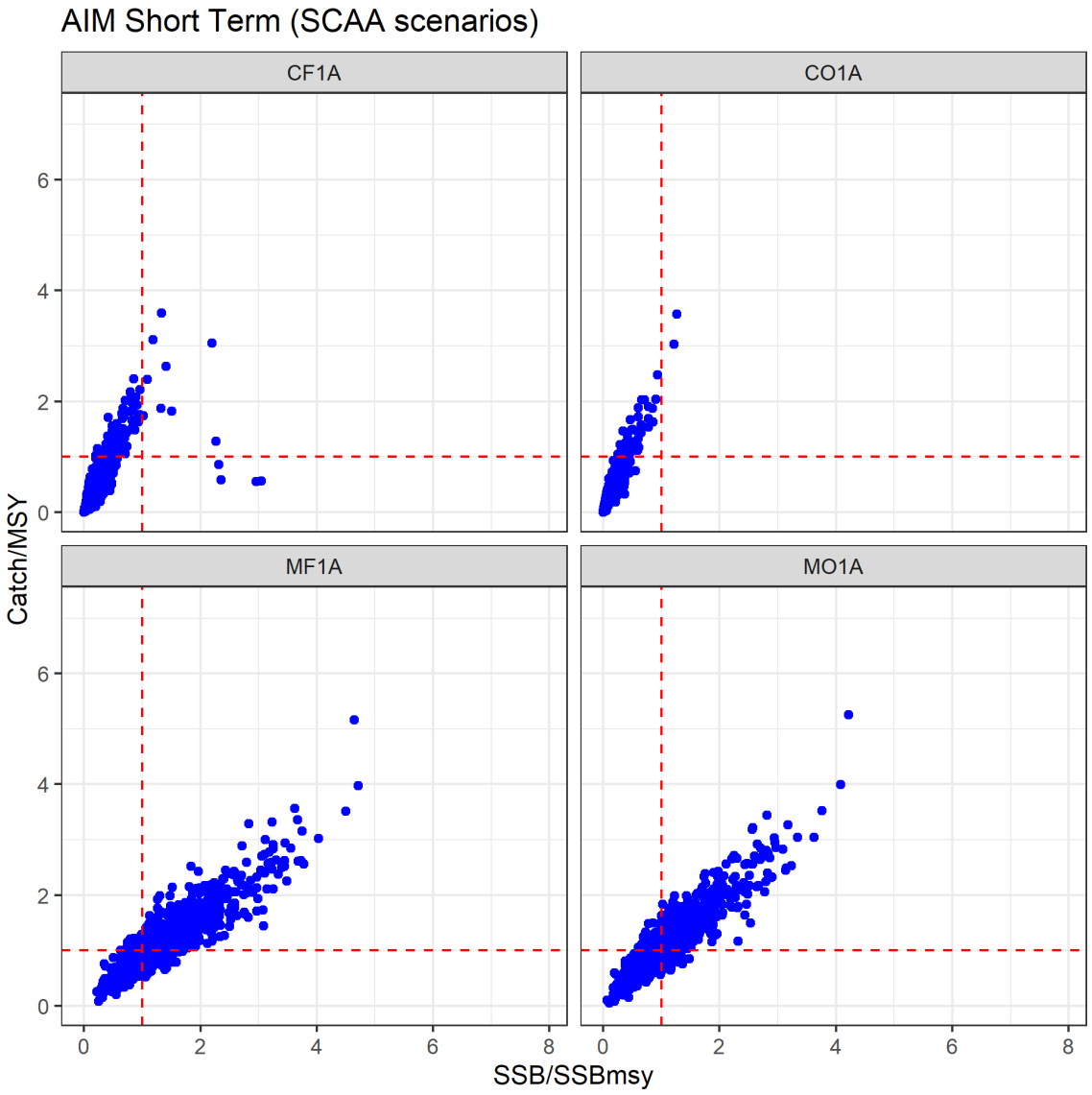


Figure A6.328. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

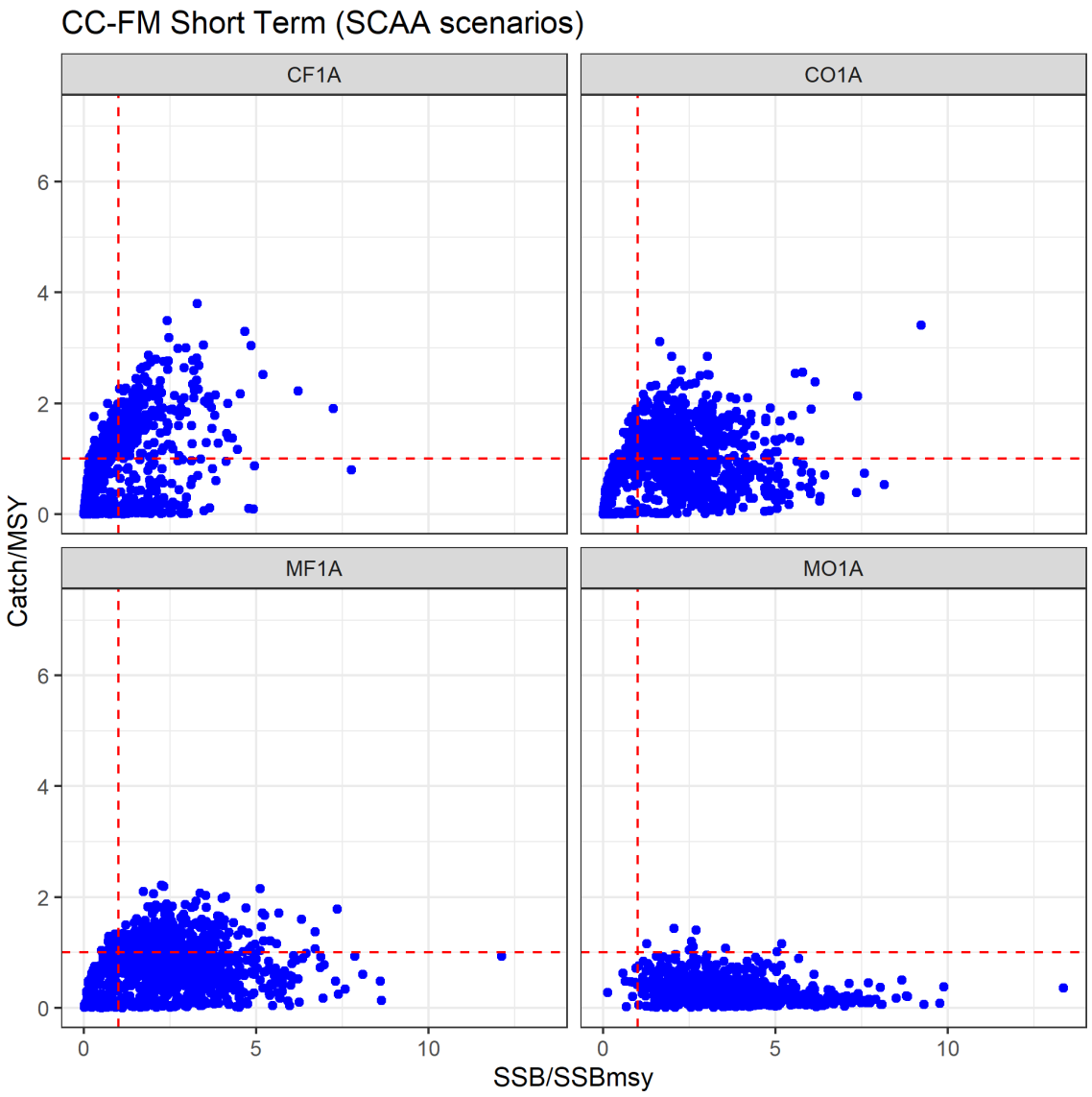


Figure A6.329. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

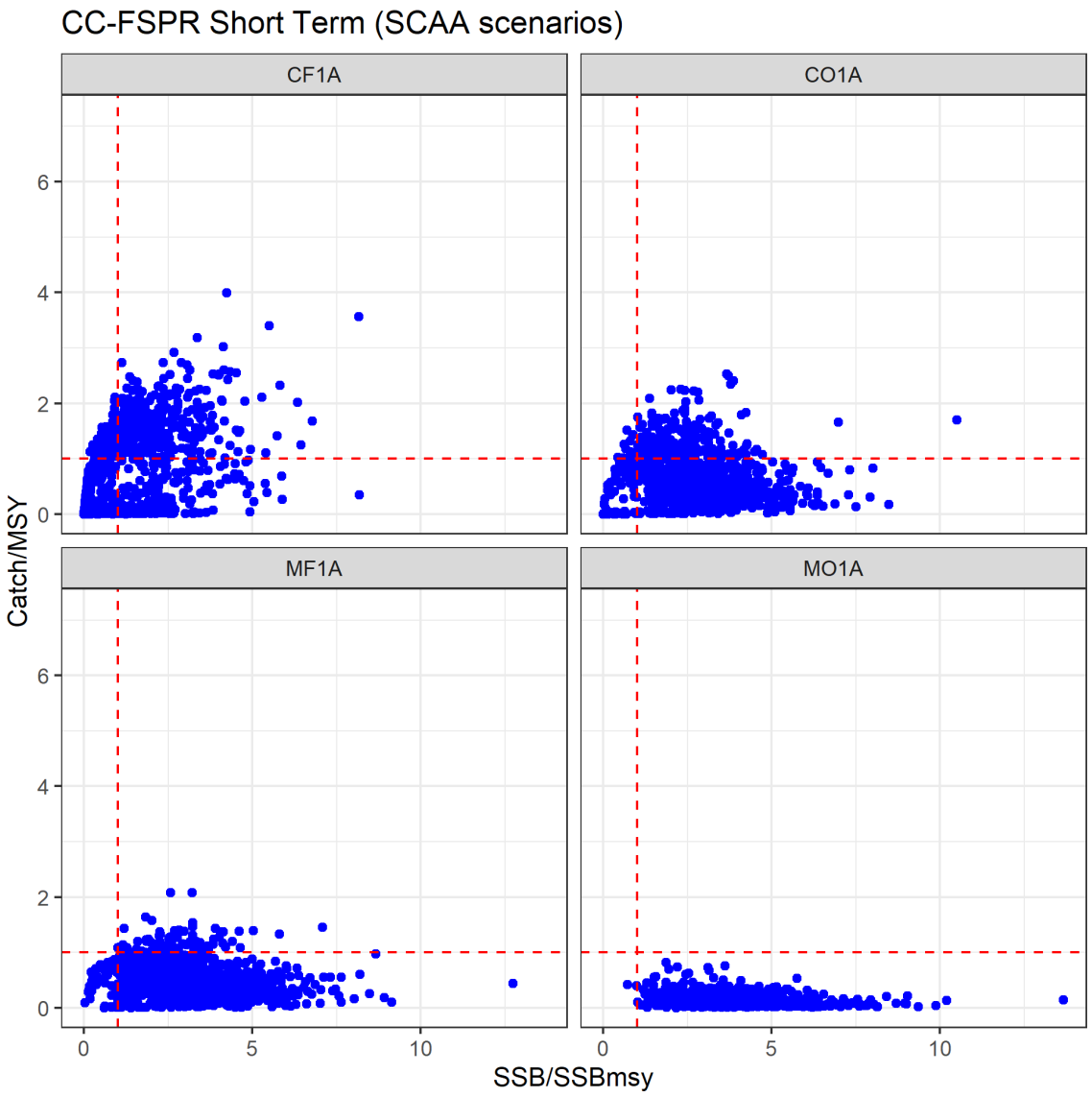


Figure A6.330. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

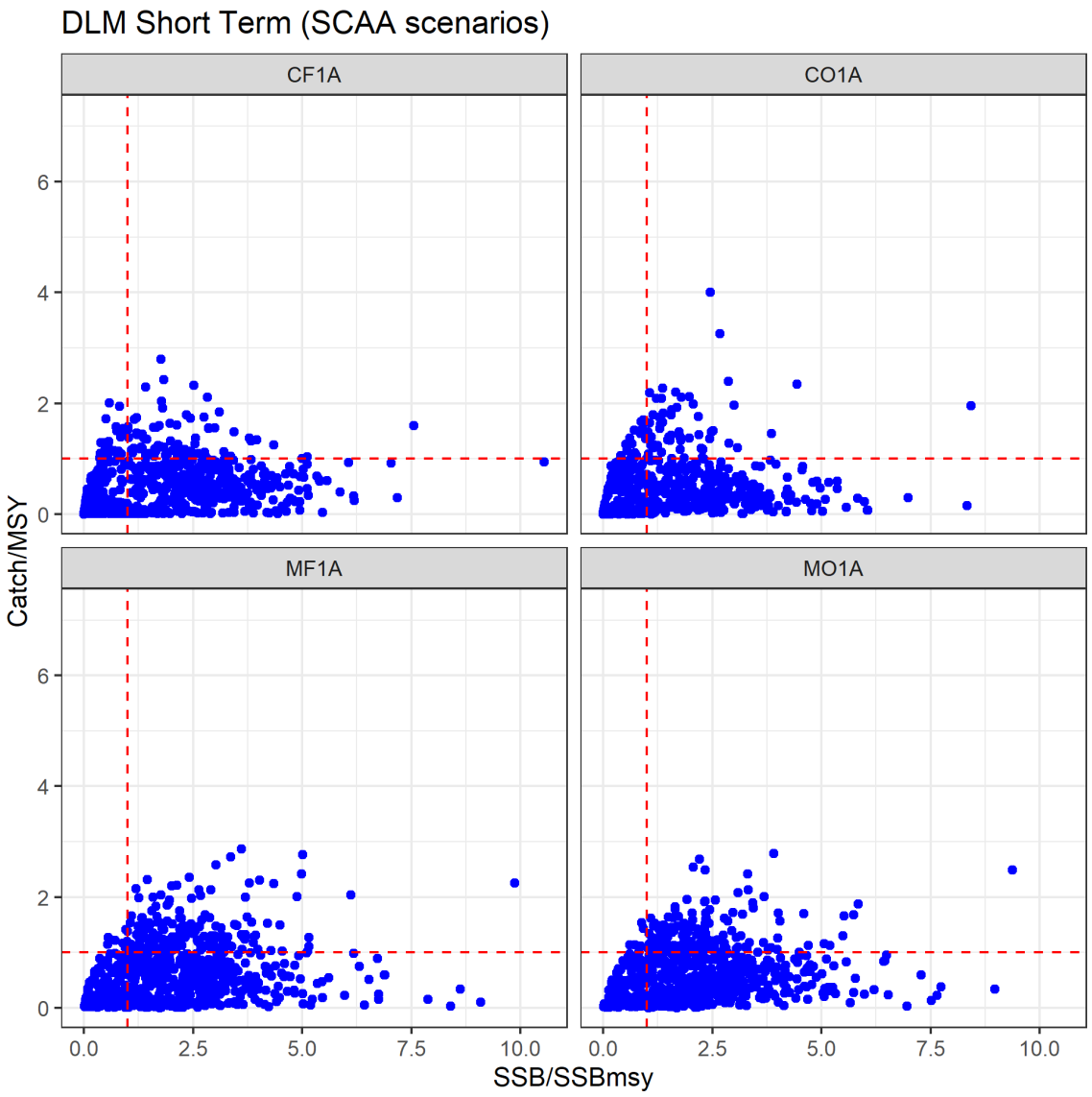


Figure A6.331. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

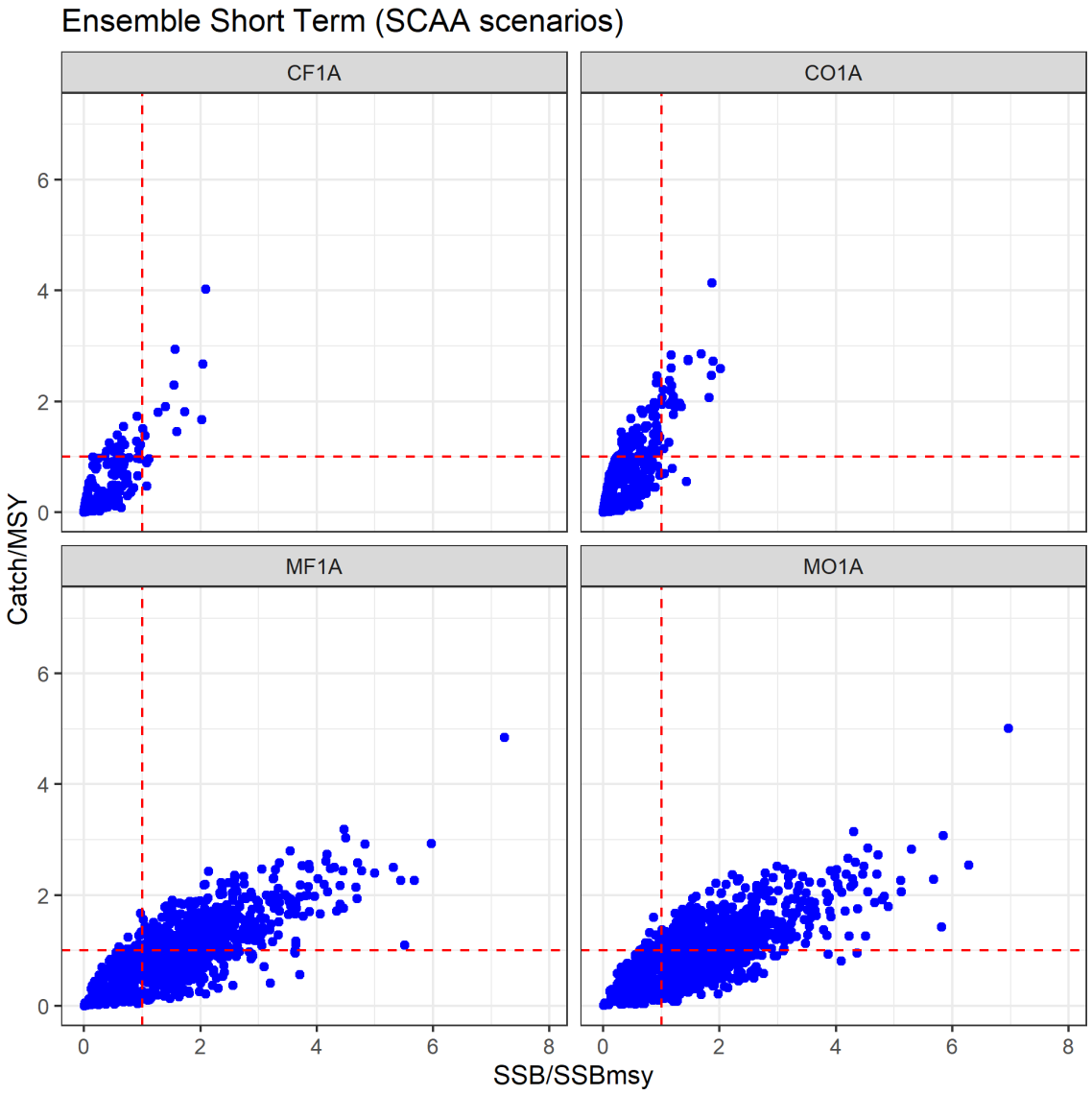


Figure A6.332. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

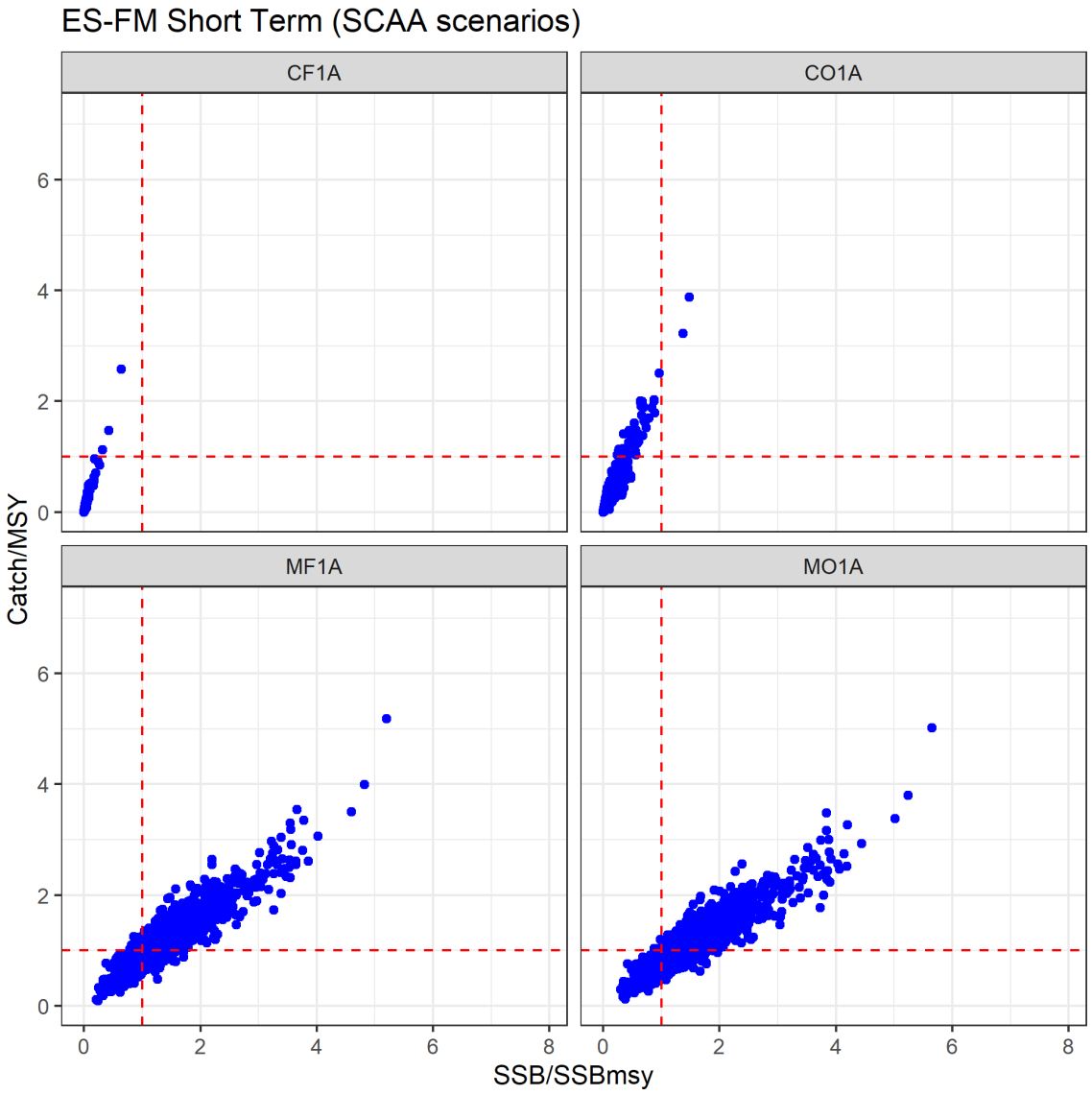


Figure A6.333. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.



Figure A6.334. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

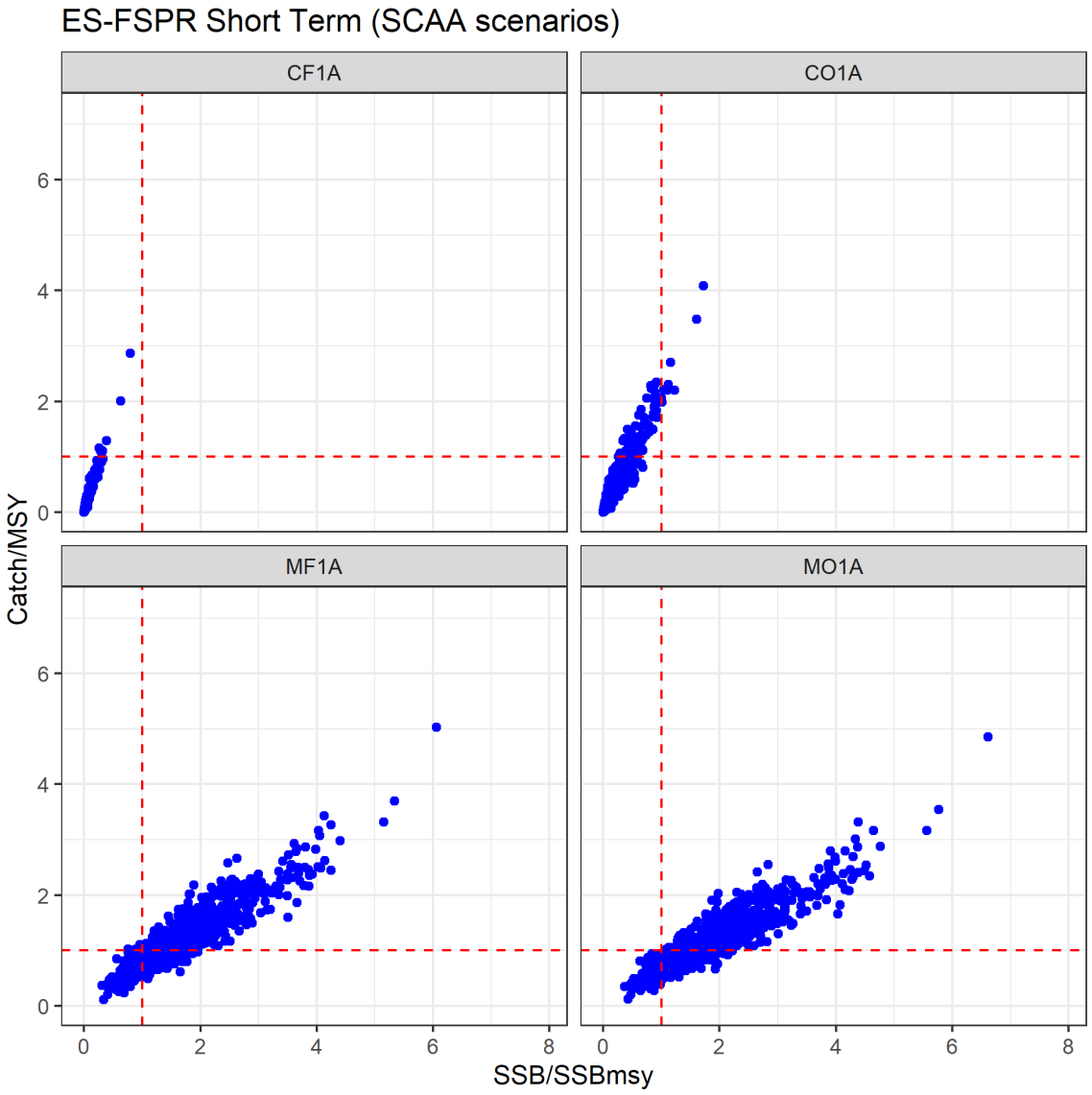


Figure A6.335. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

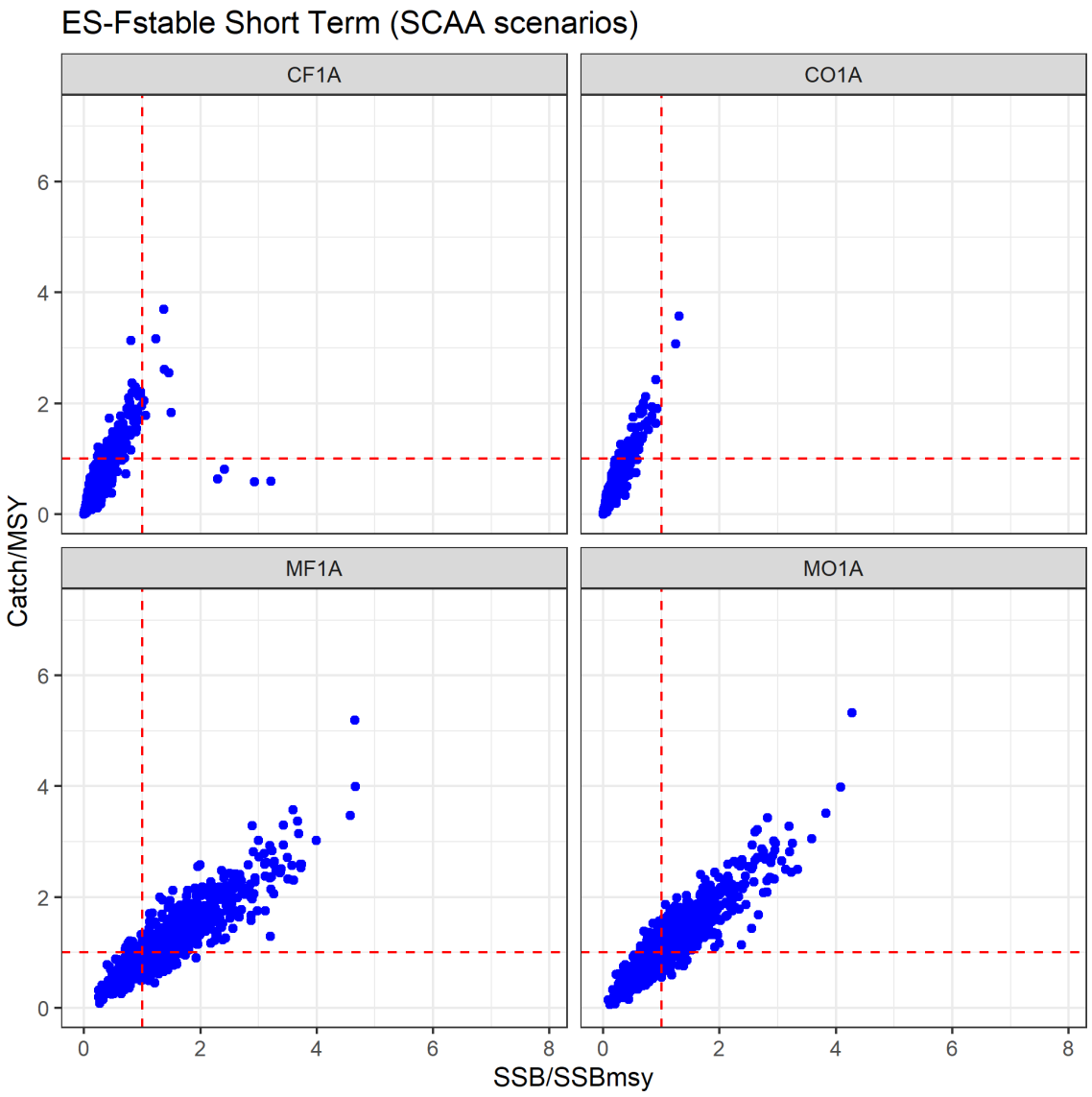


Figure A6.336. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

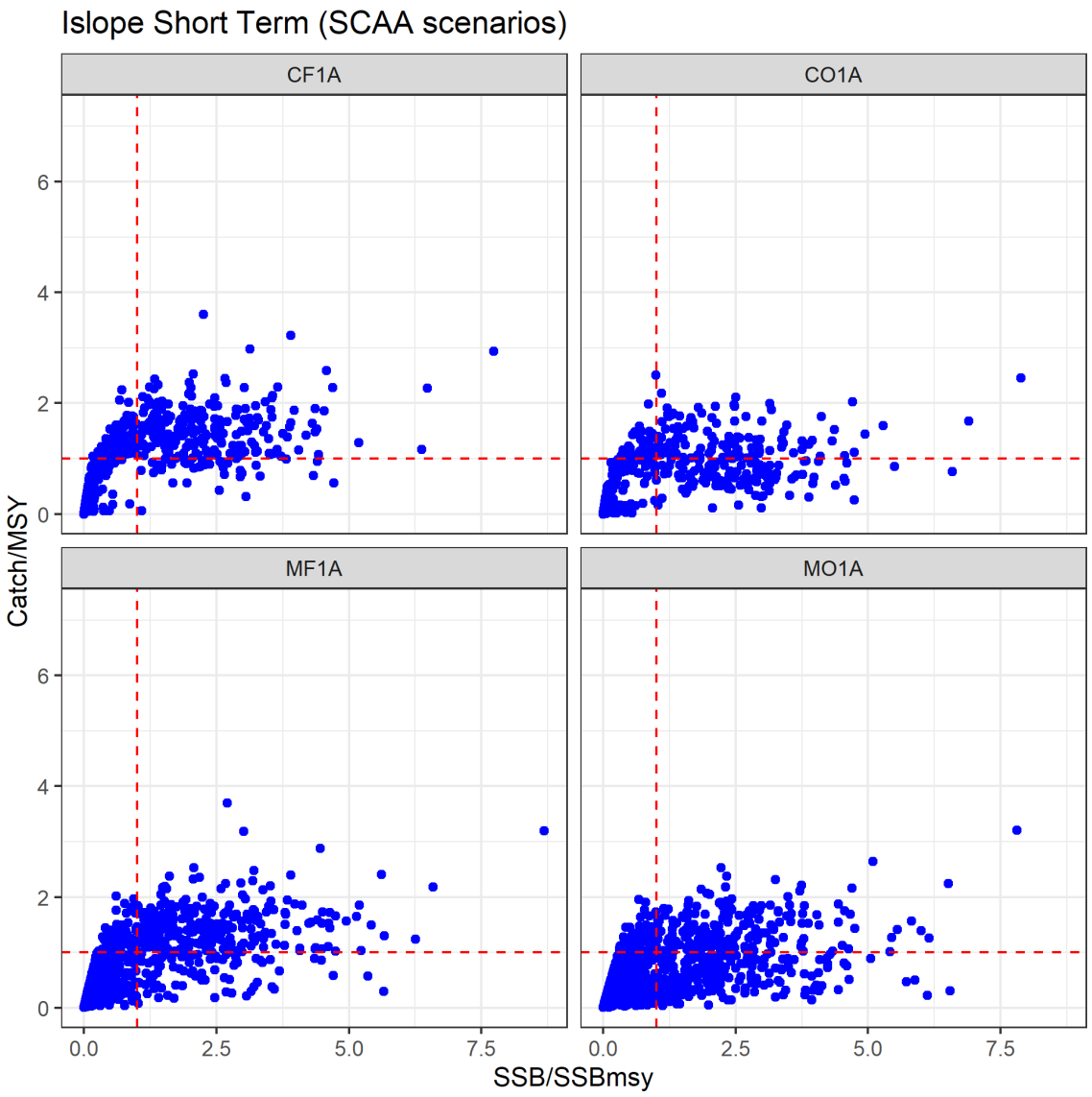


Figure A6.337. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

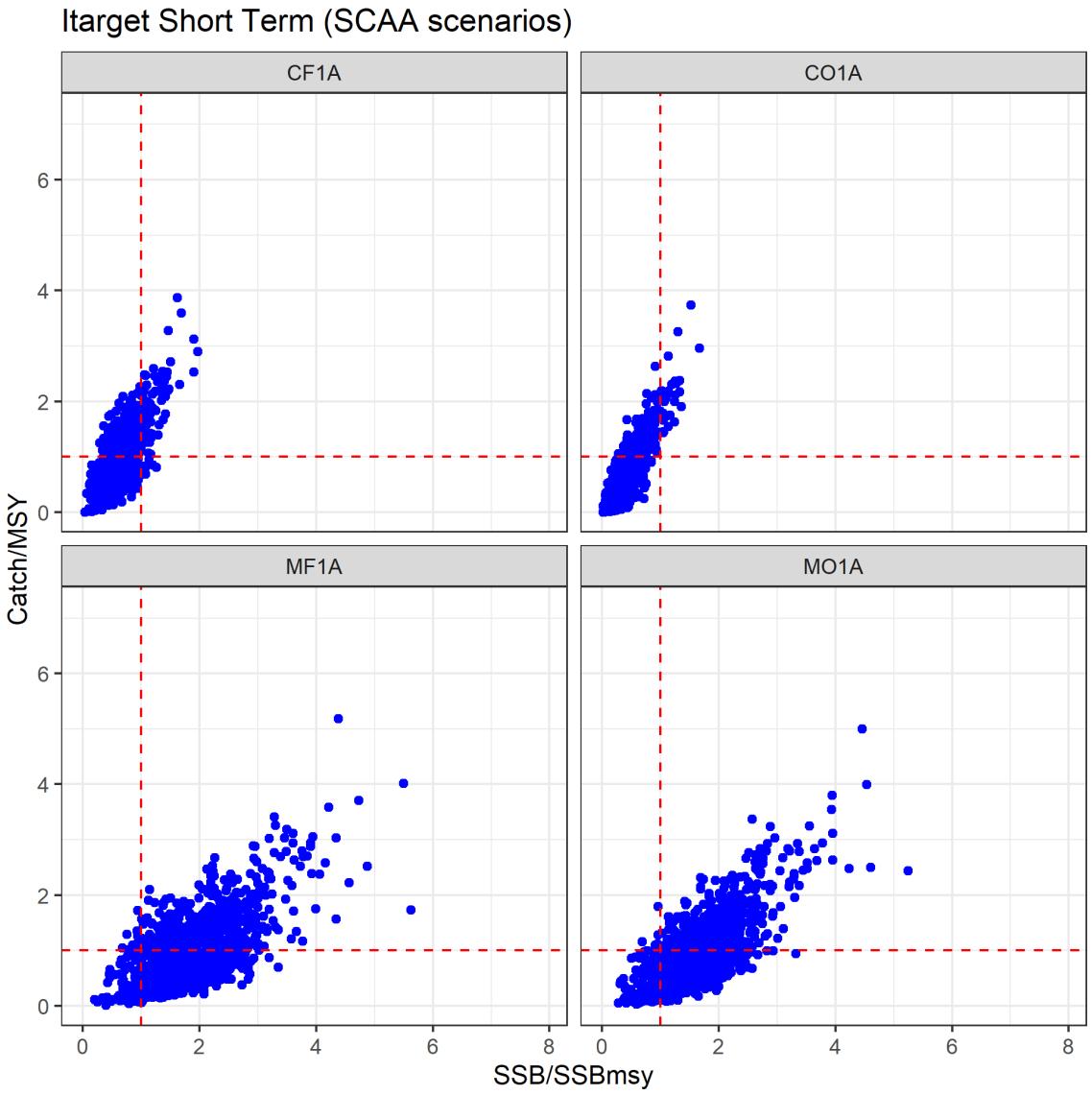


Figure A6.338. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

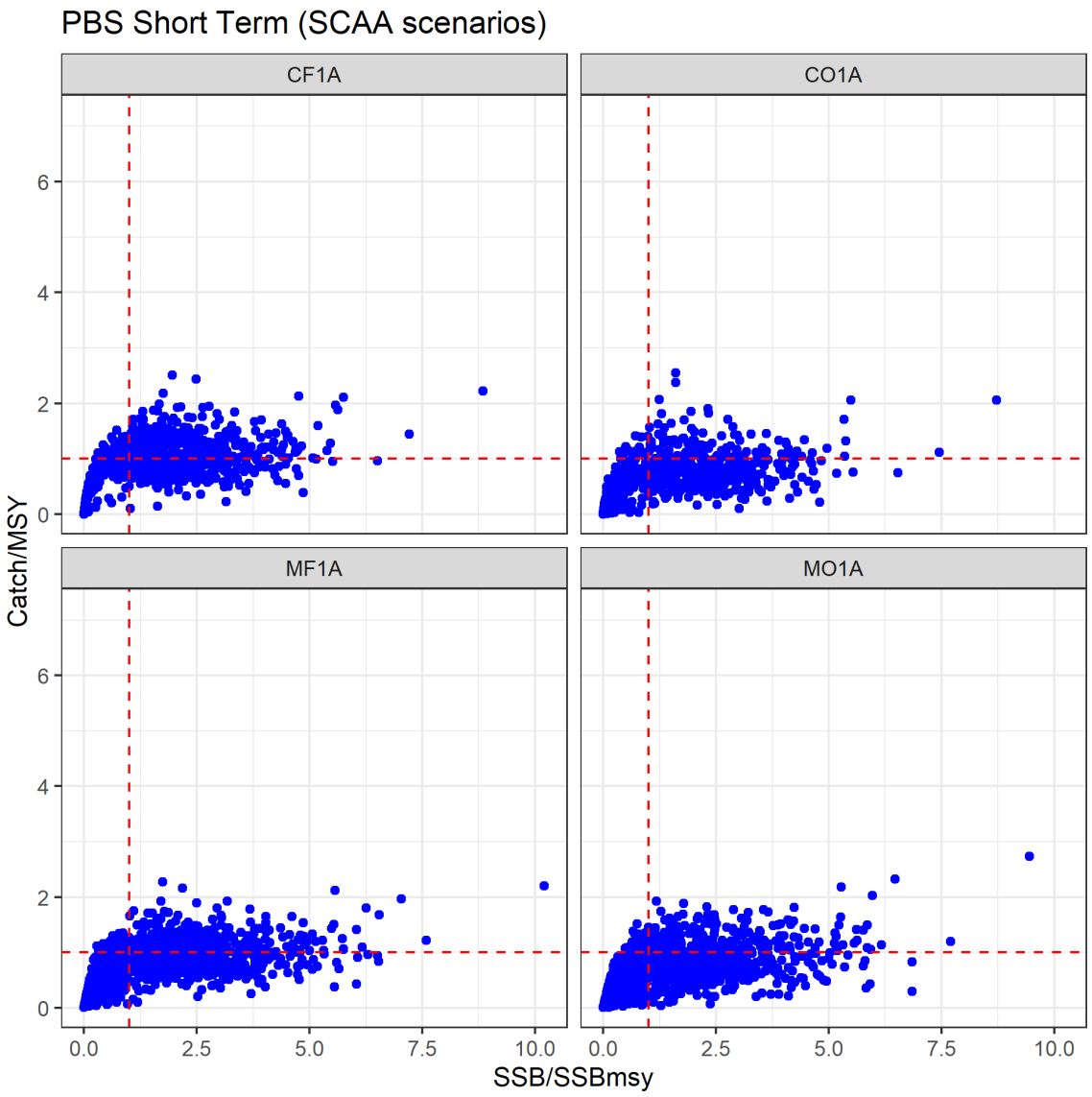


Figure A6.339. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

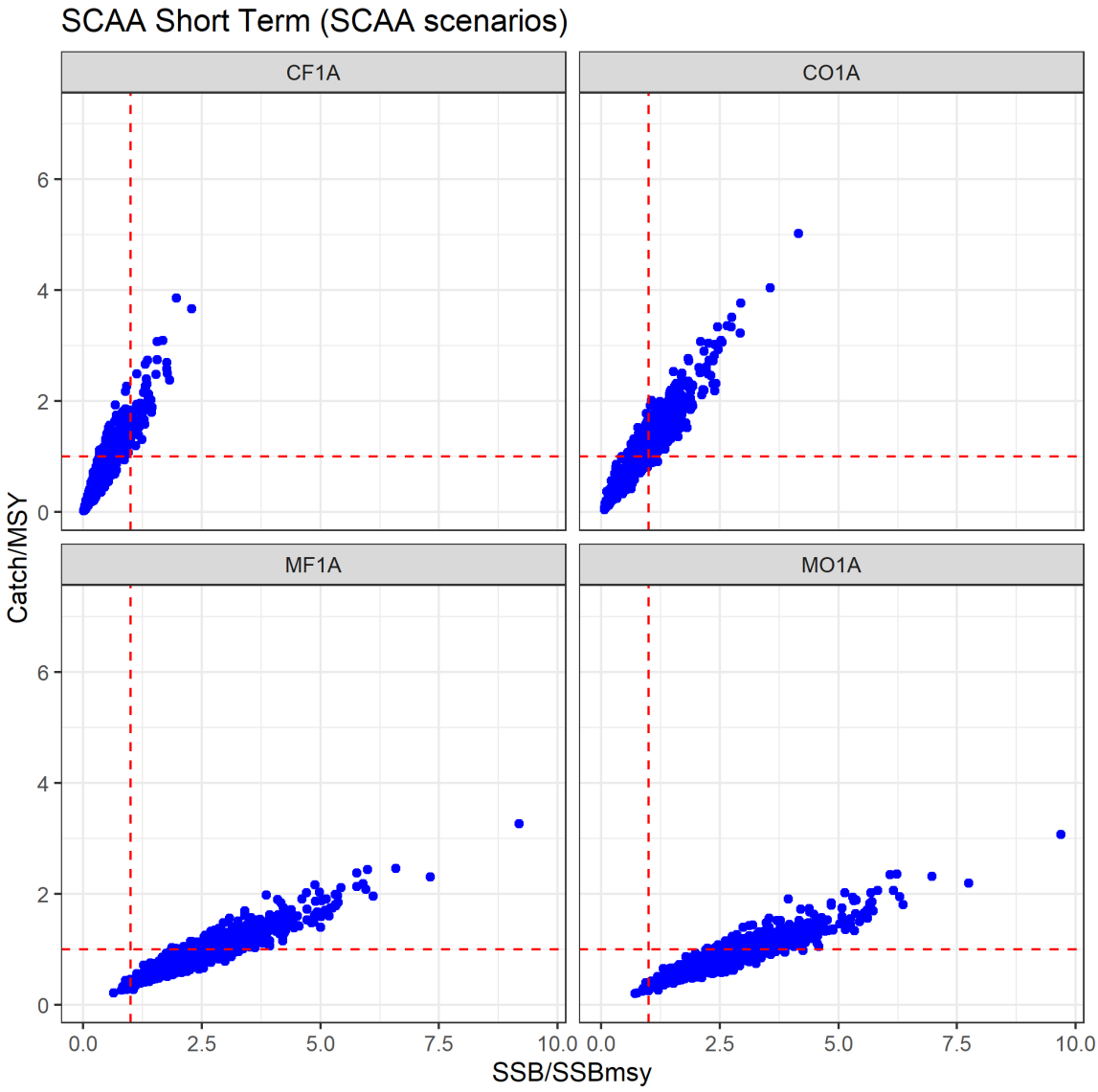


Figure A6.340. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

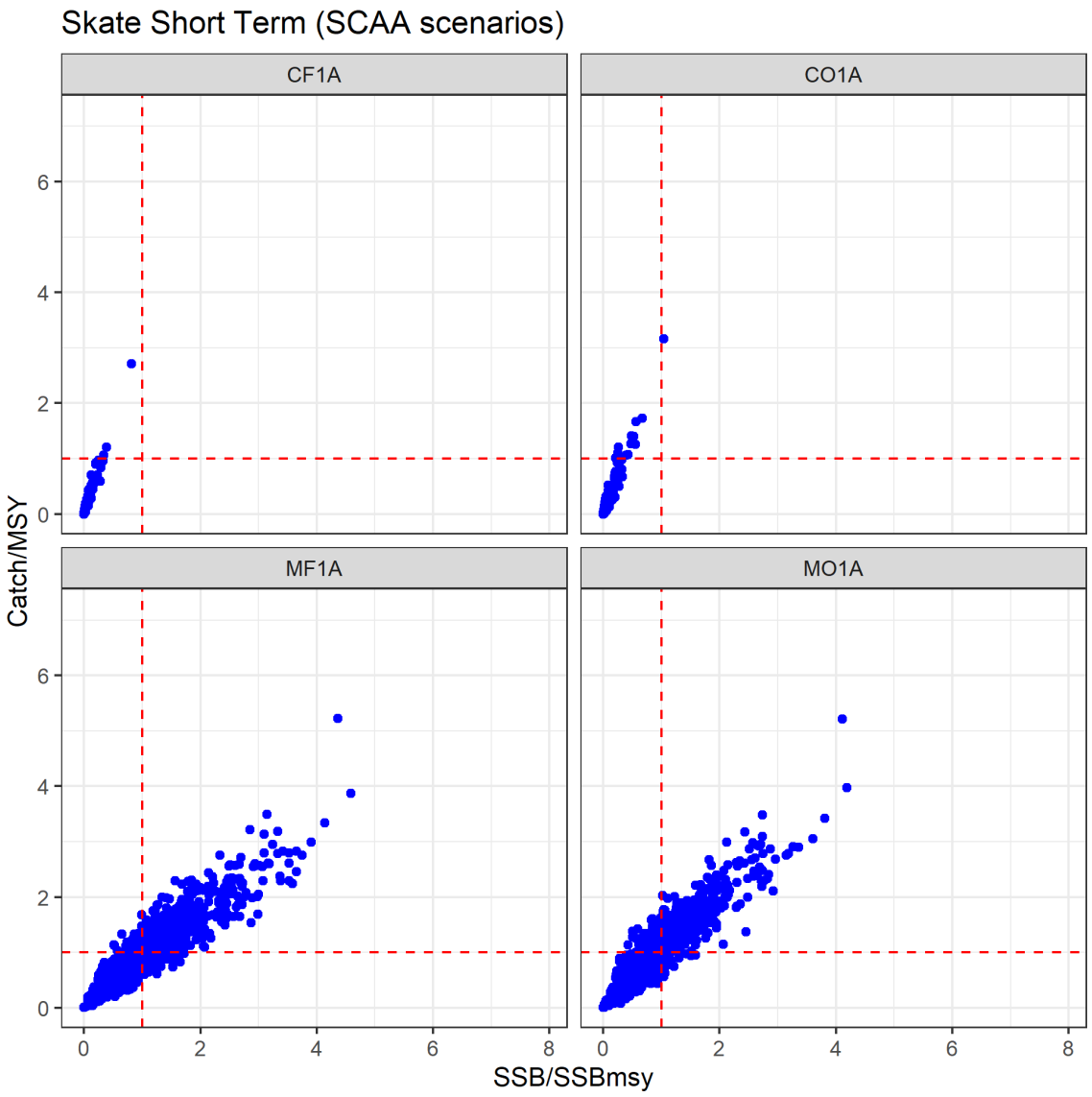


Figure A6.341. Spawning stock biomass (SSB) relative to the SSB at maximum sustainable yield (SSB_{MSY}) and catch relative to MSY for the index-based method (IBM) defined in the top left in the short term with each dot representing 1 of the 1,000 simulations.

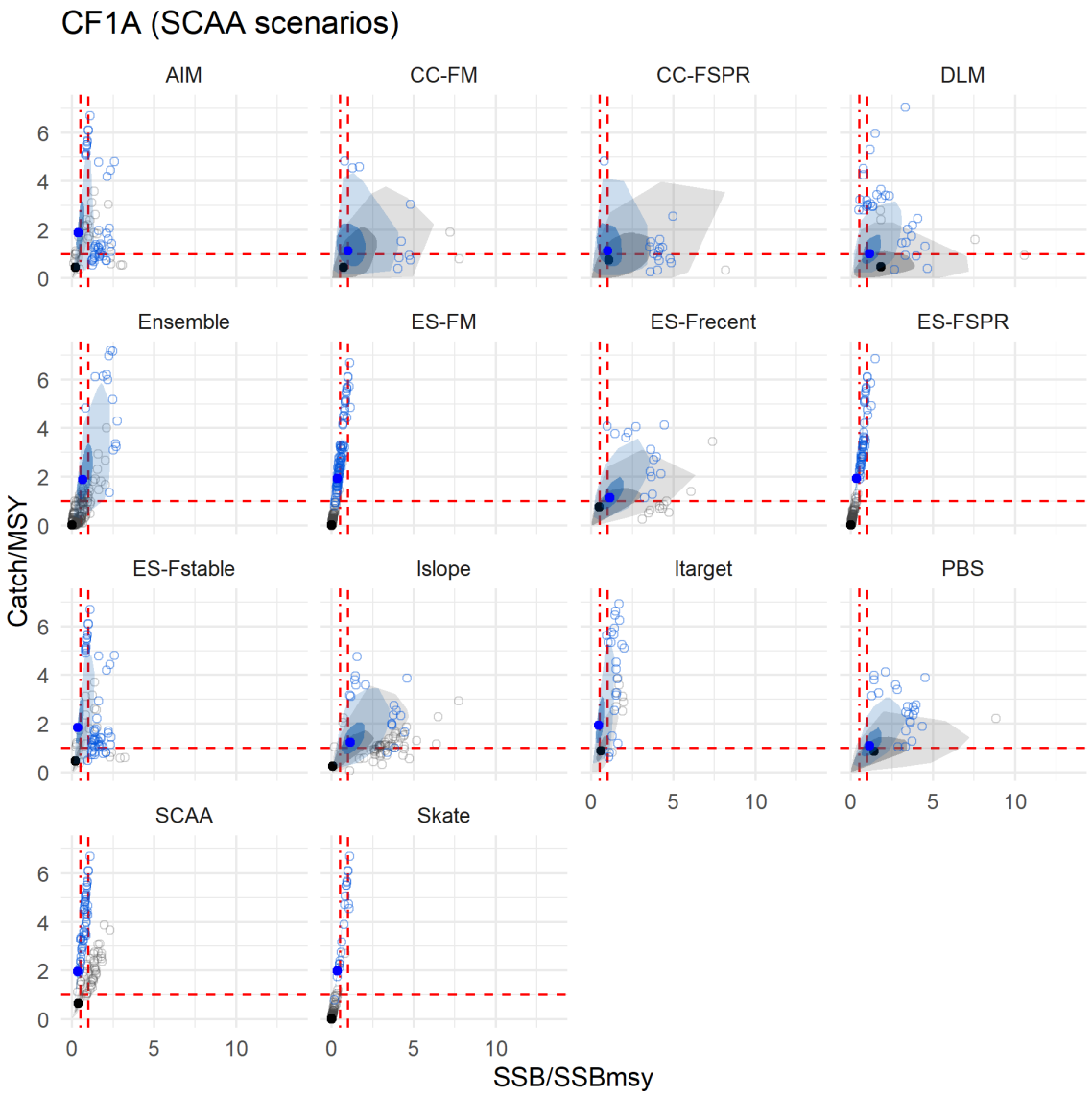


Figure A6.342. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

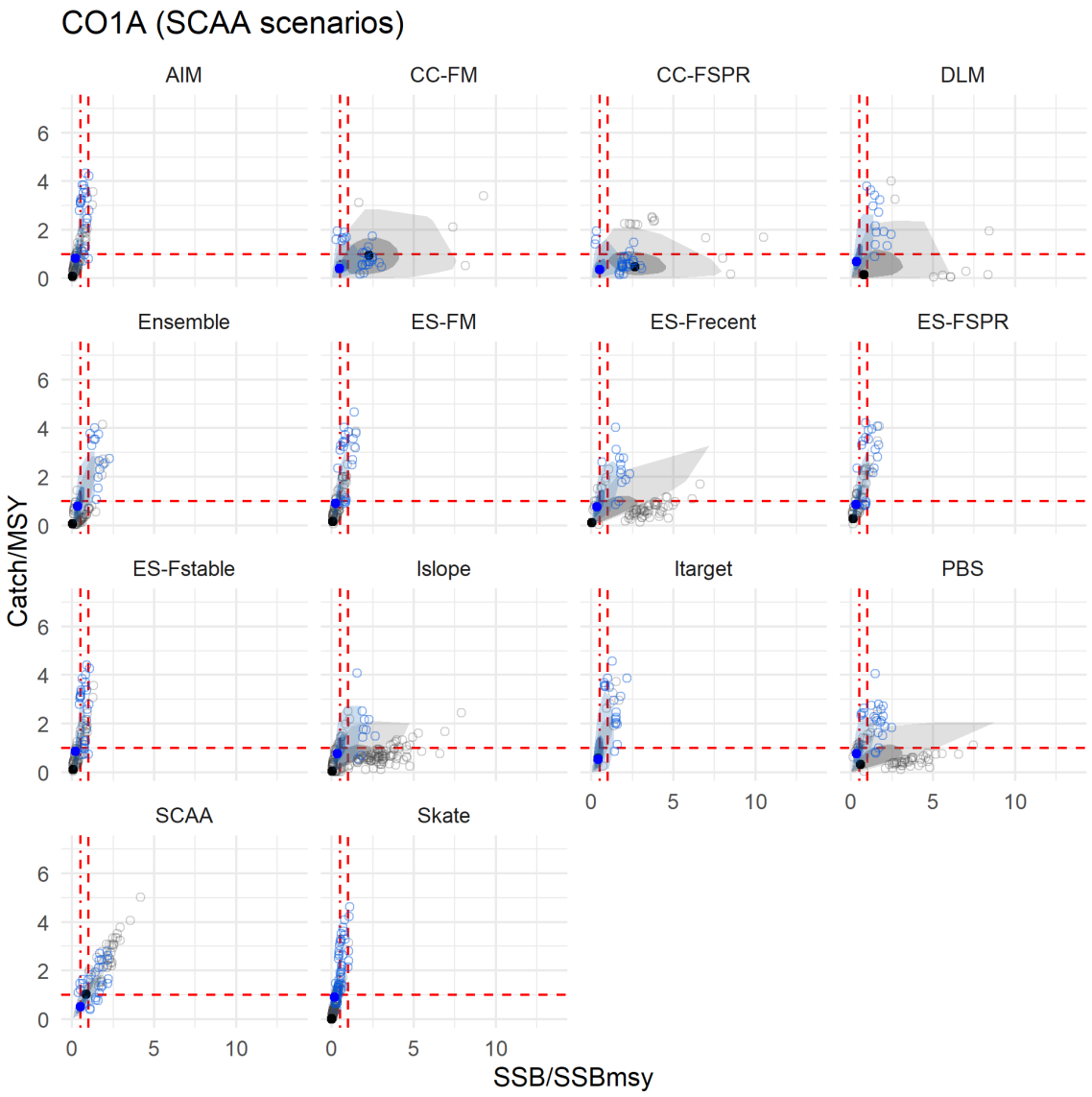


Figure A6.343. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

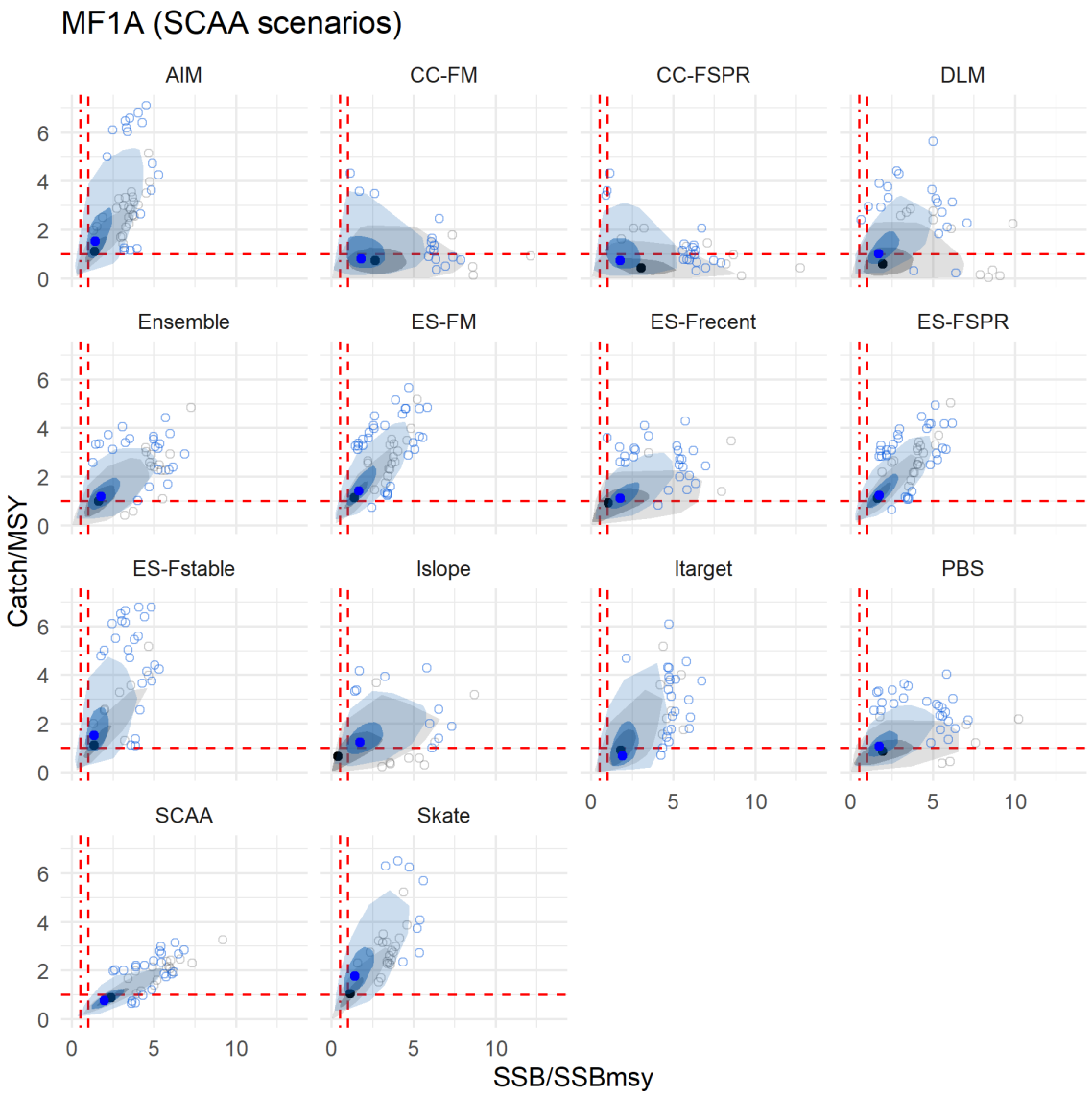


Figure A6.344. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

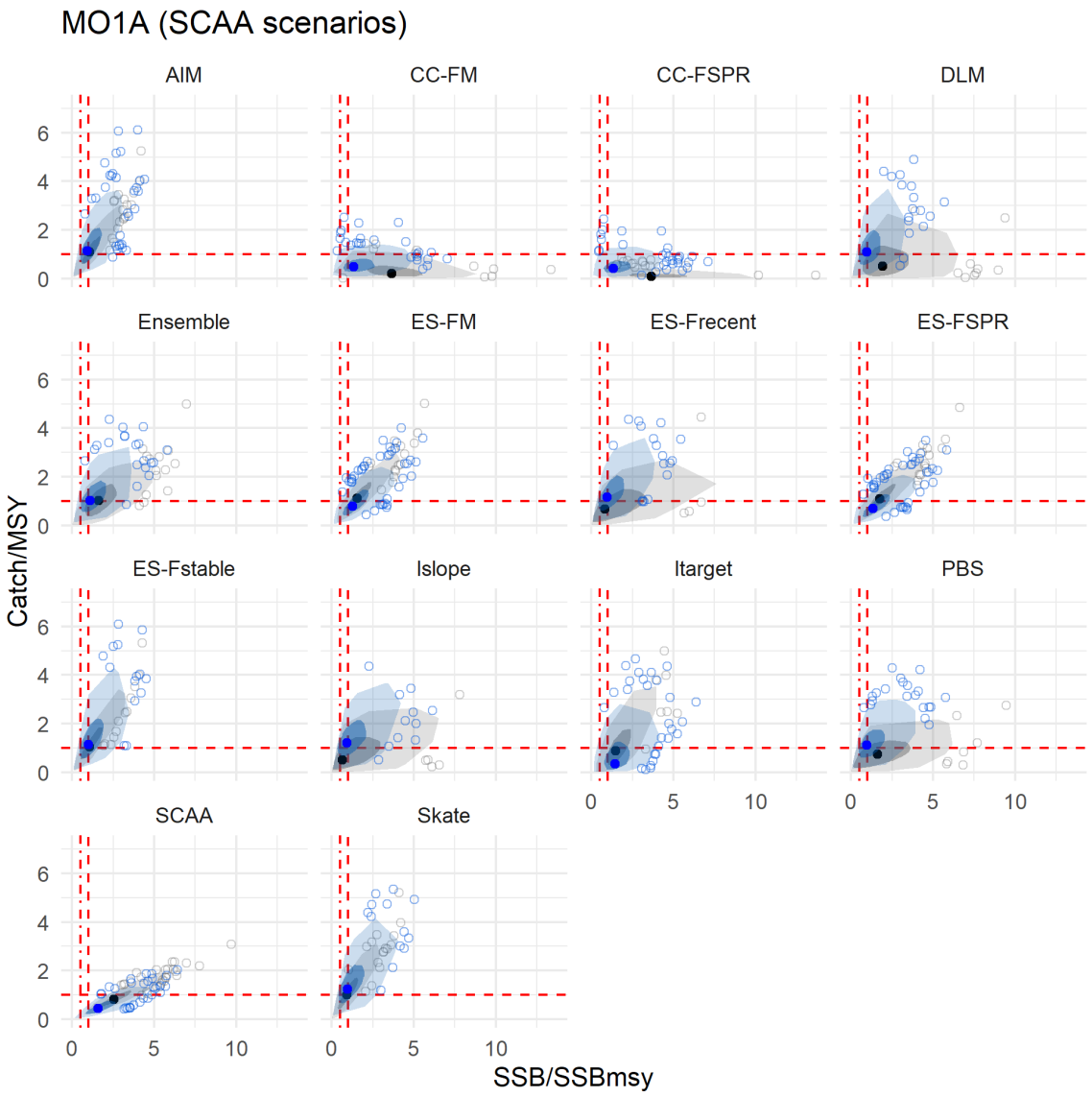


Figure A6.345. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each index-based method (IBM) in the scenario defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

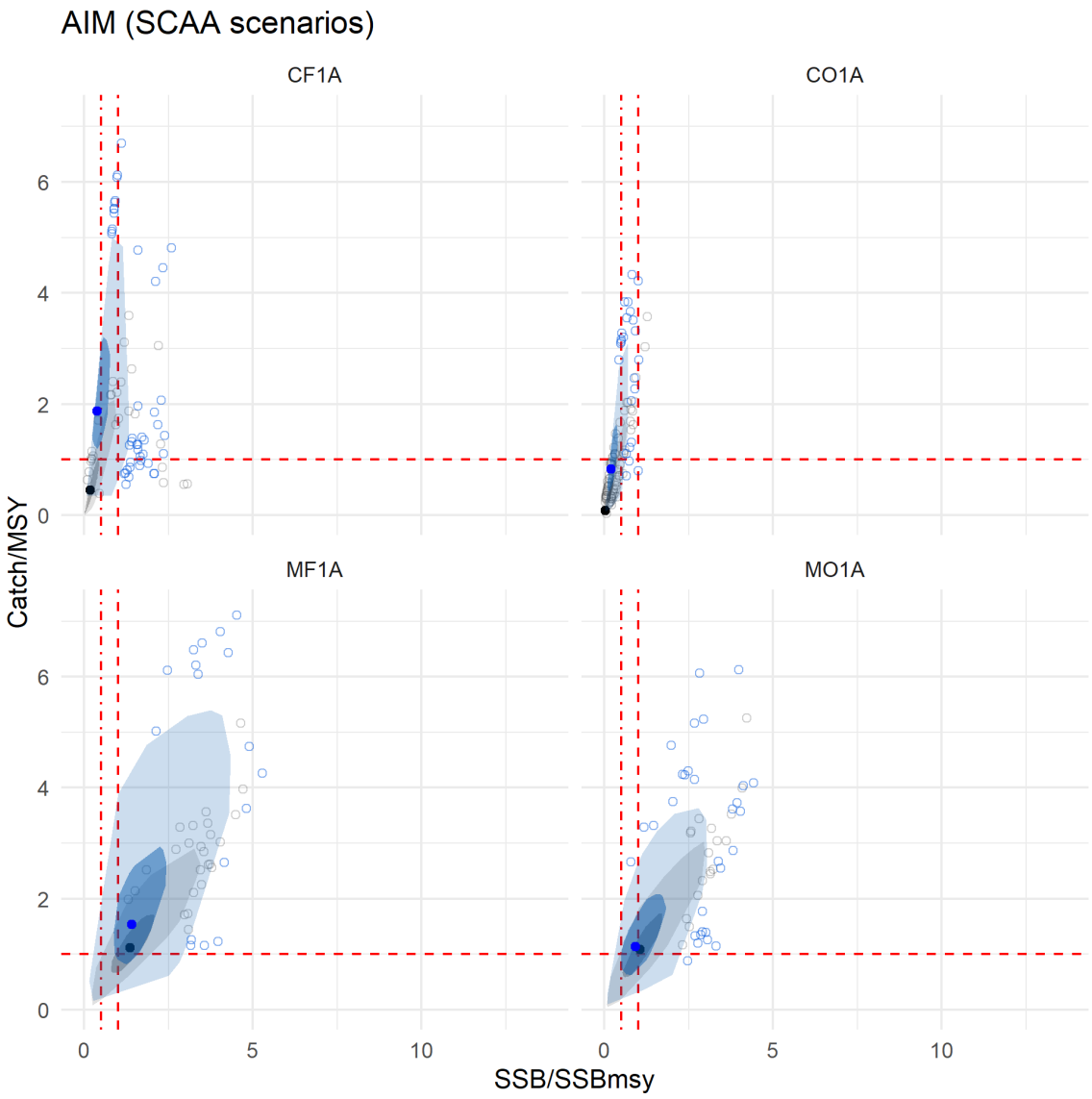


Figure A6.346. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

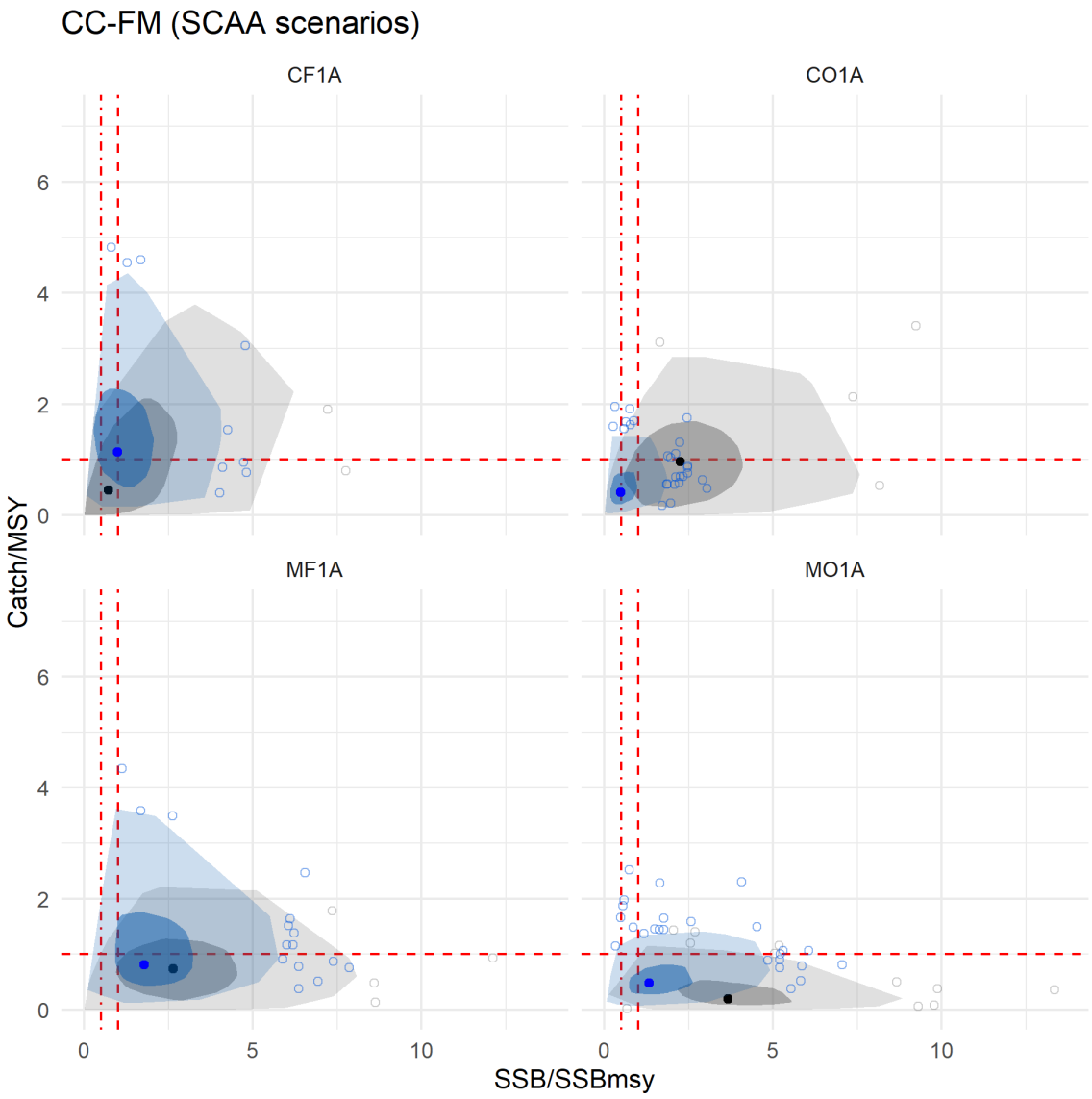


Figure A6.347. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

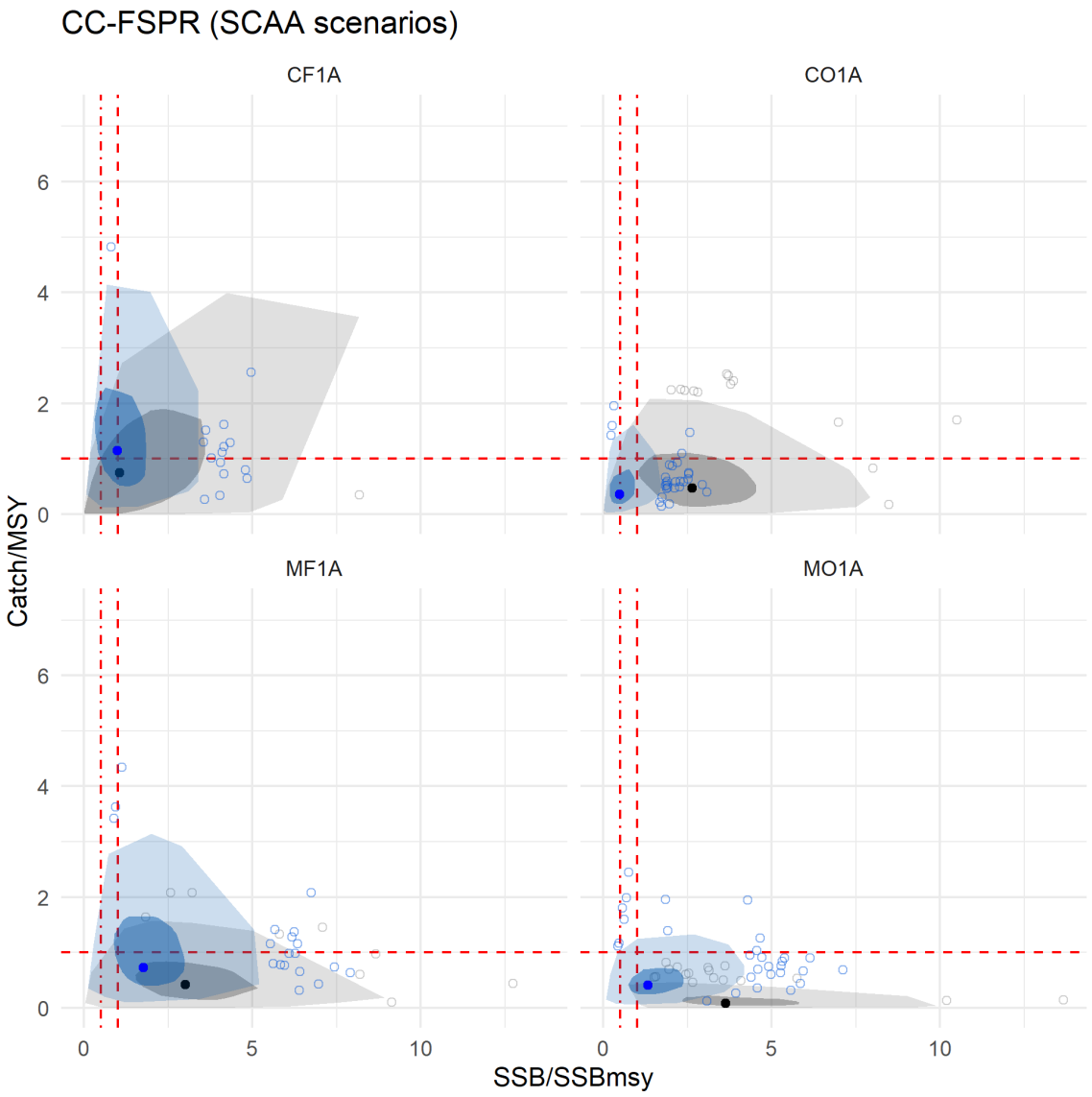


Figure A6.348. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

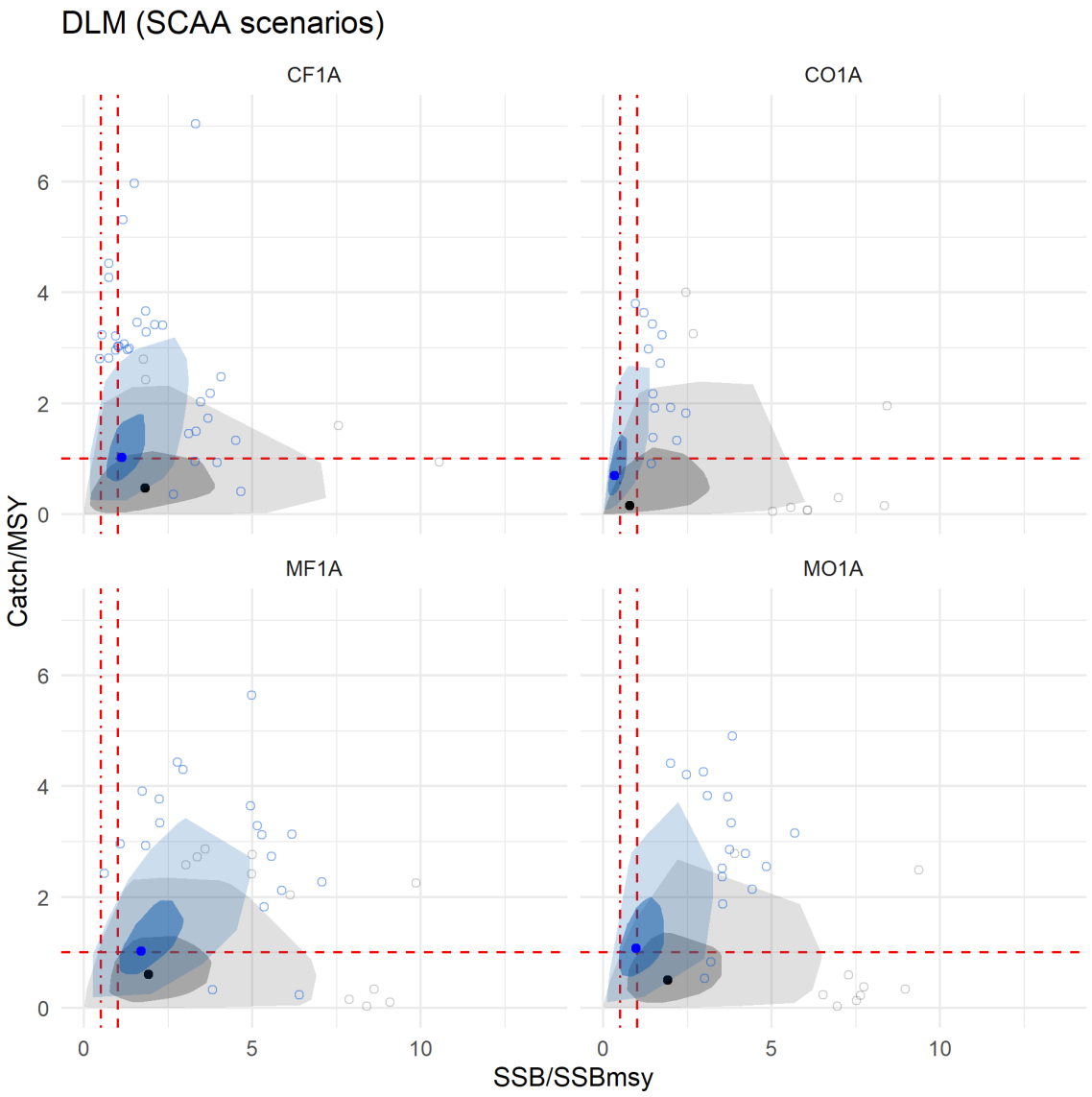


Figure A6.349. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

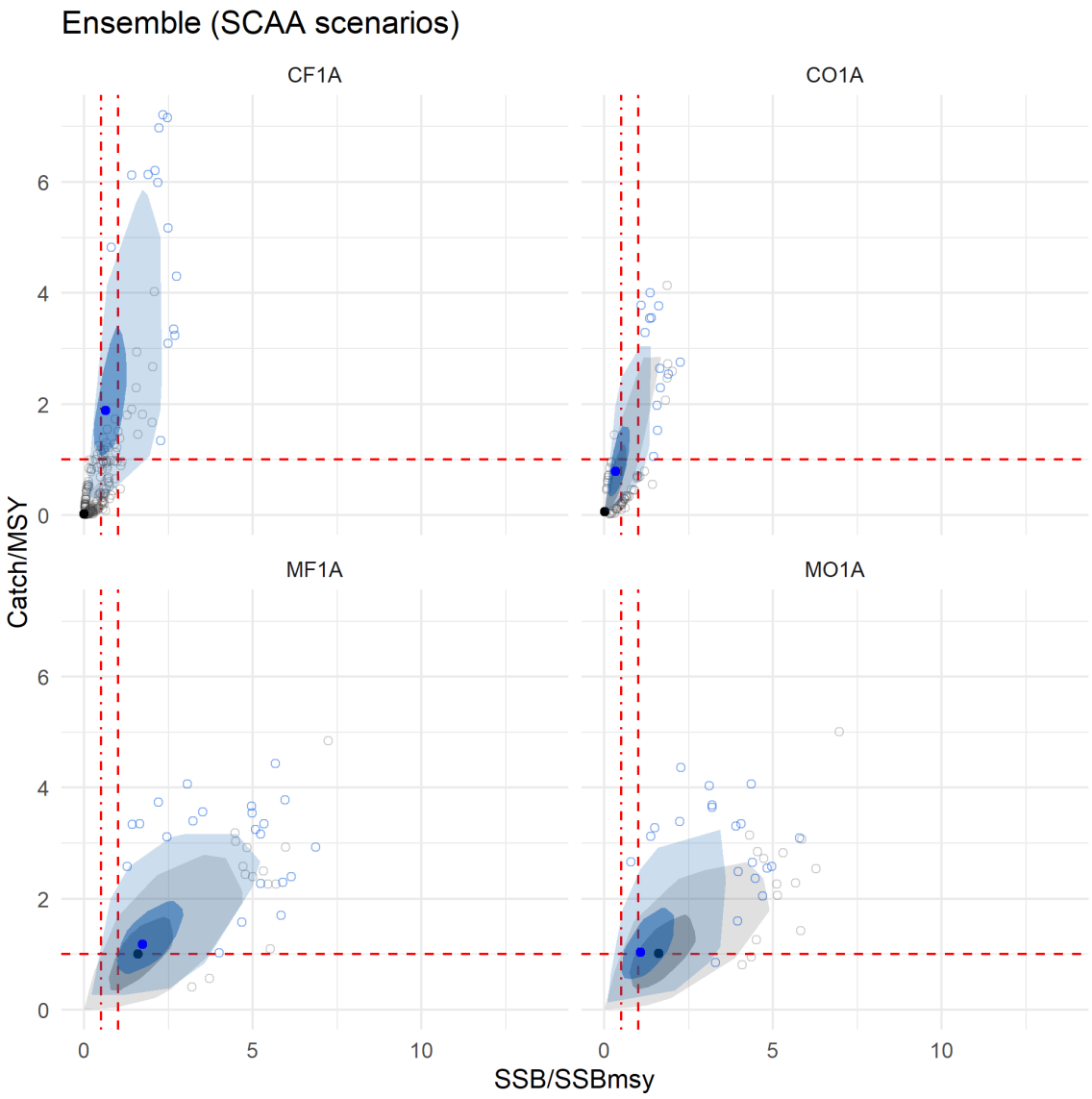


Figure A6.350. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

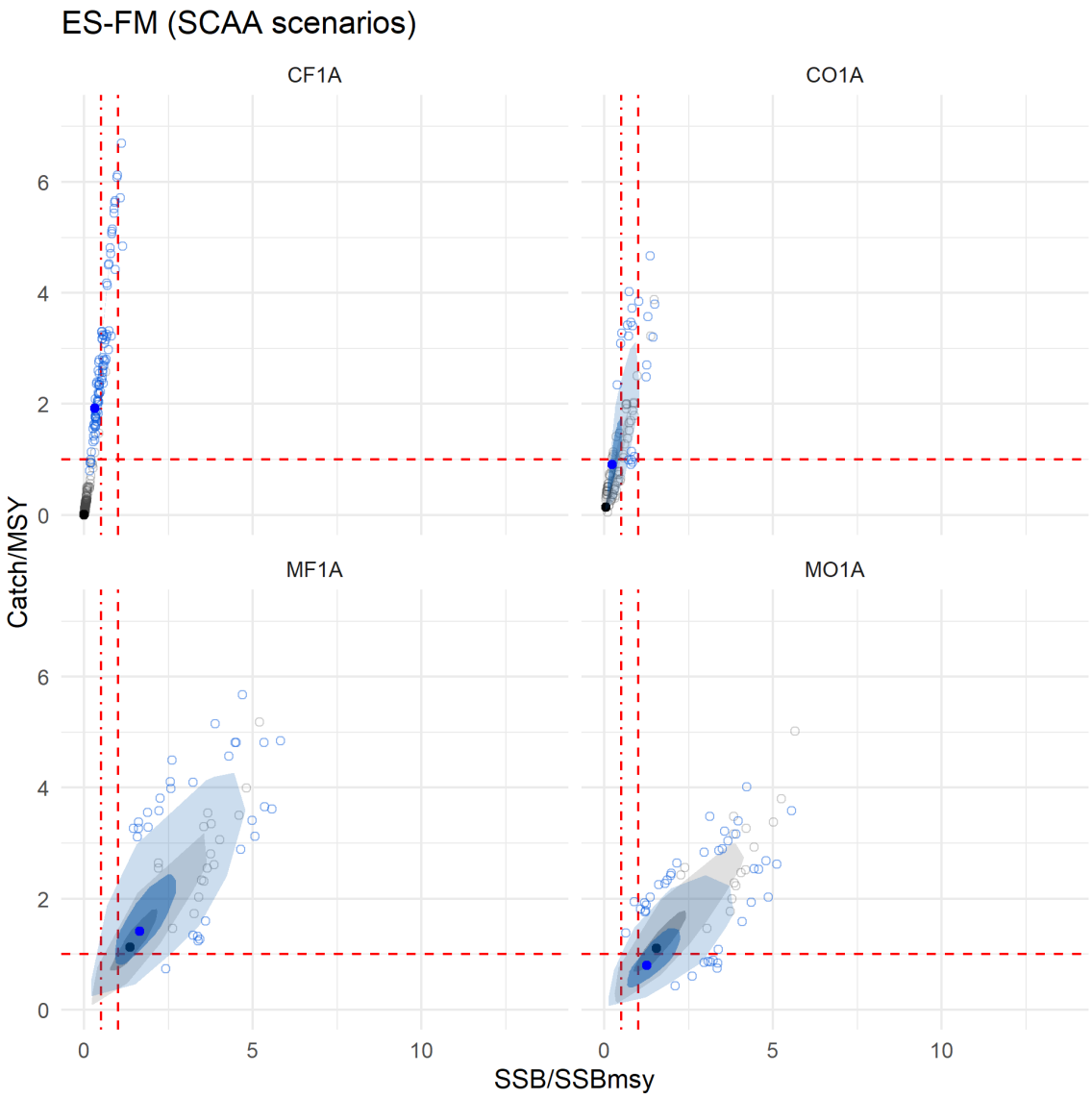


Figure A6.351. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

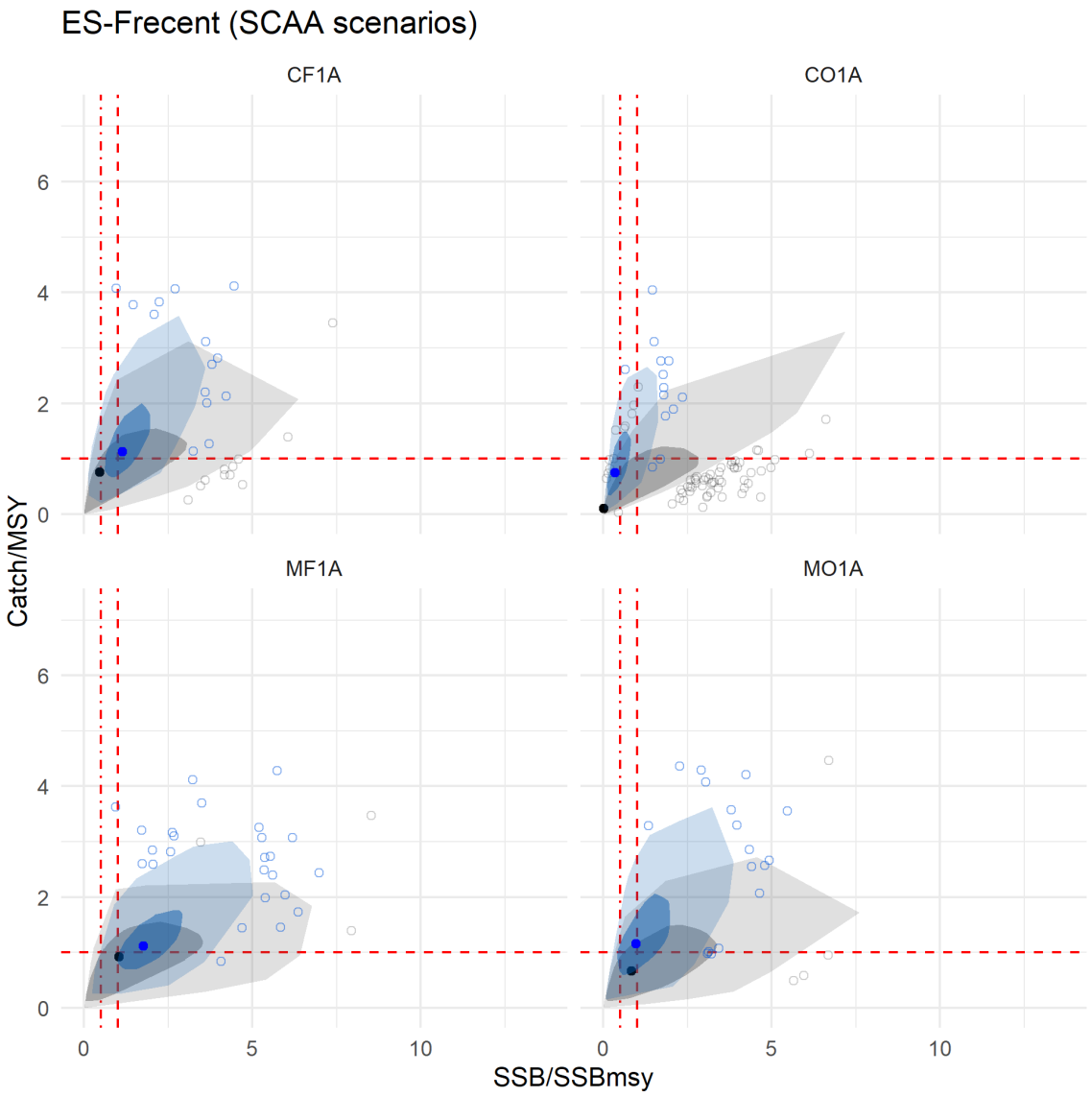


Figure A6.352. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

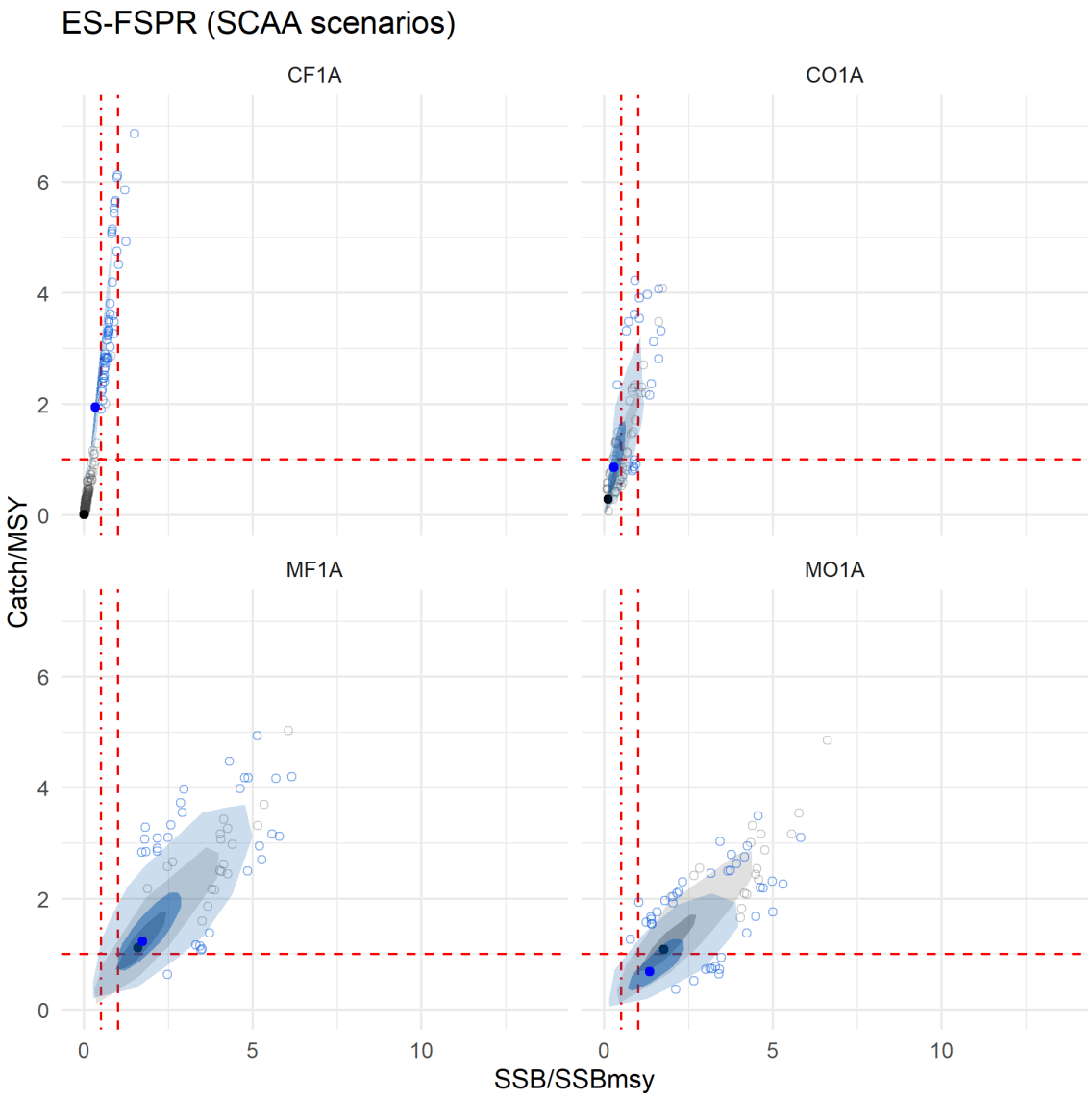


Figure A6.353. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

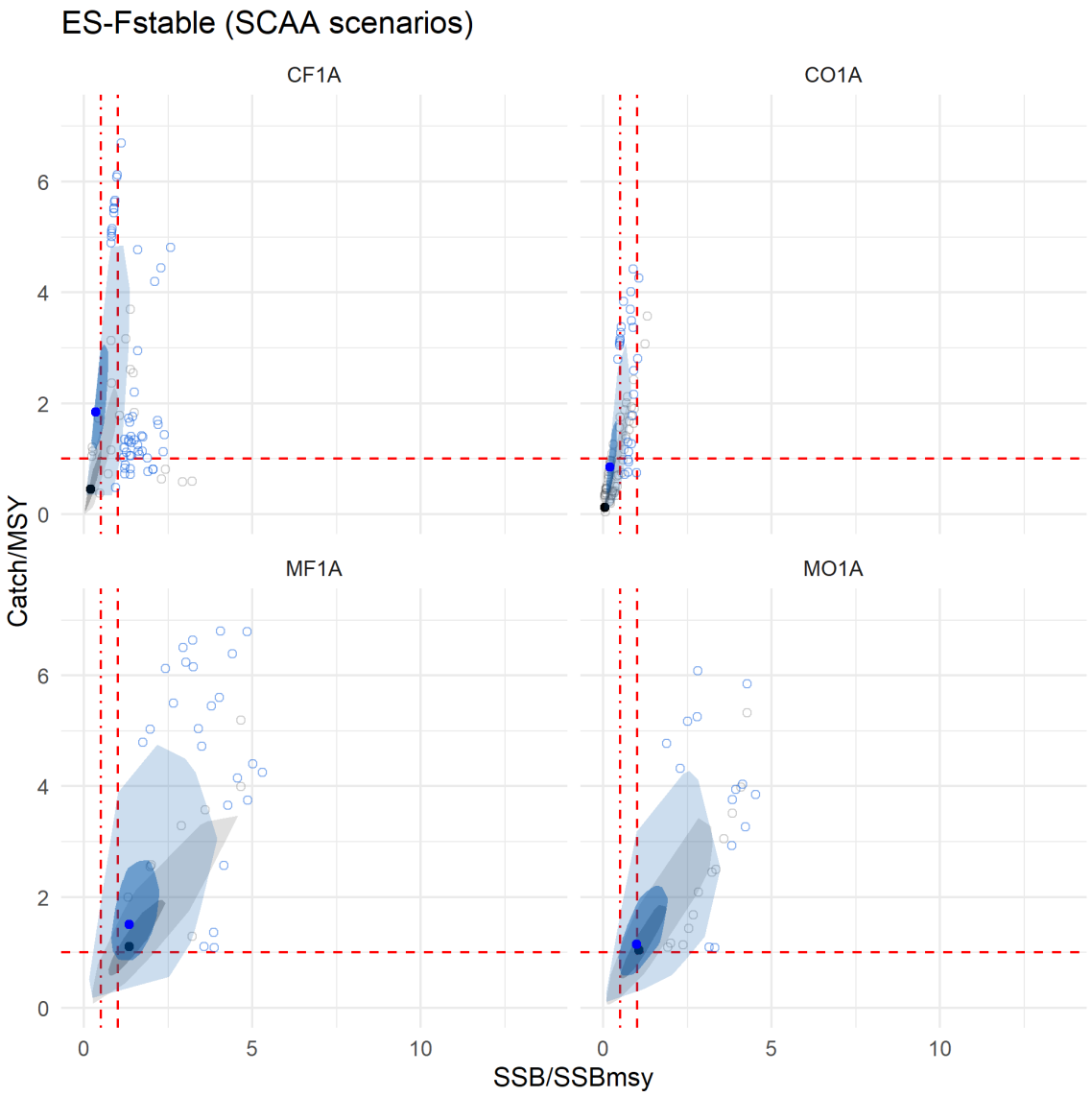


Figure A6.354. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

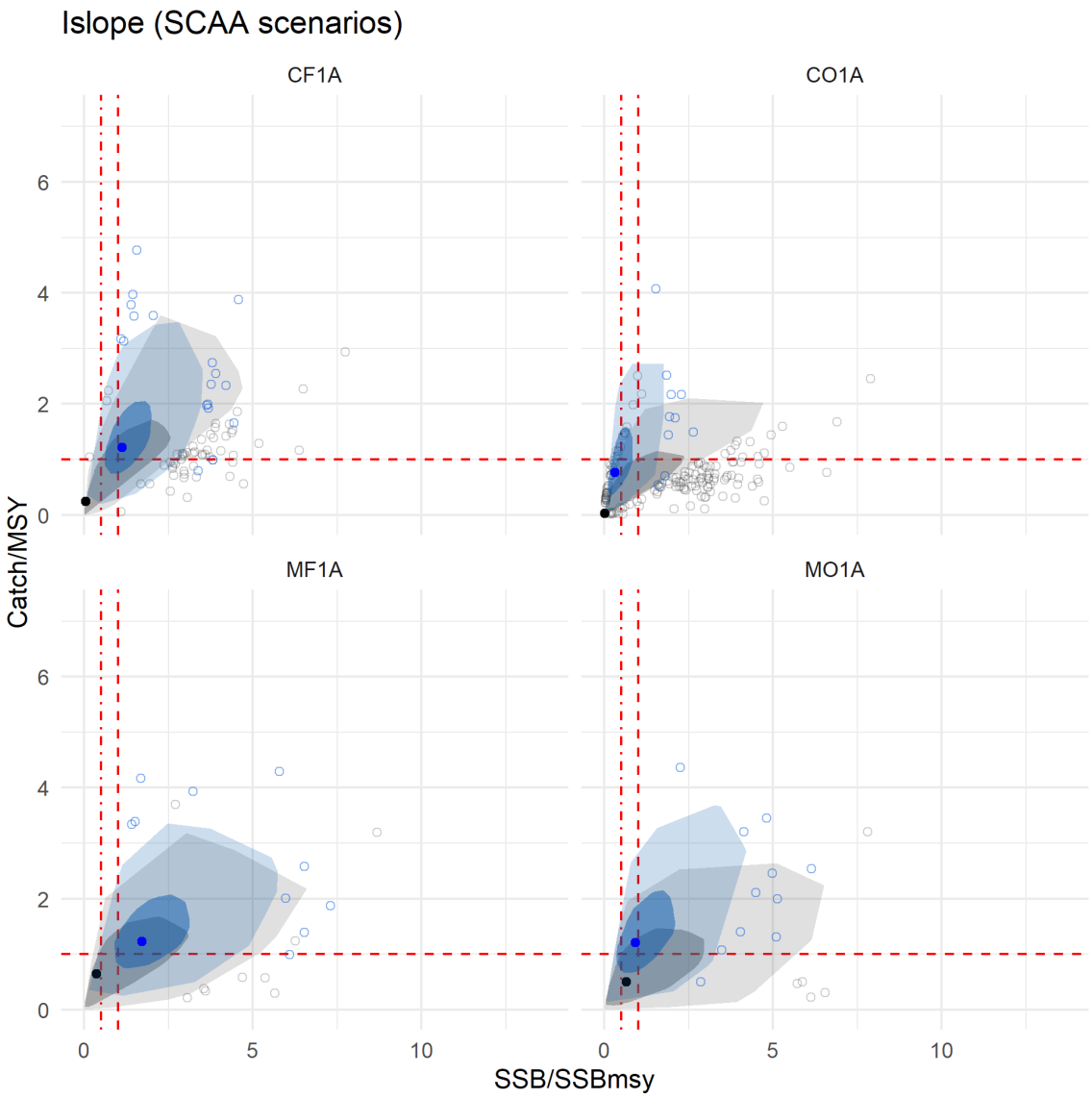


Figure A6.355. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

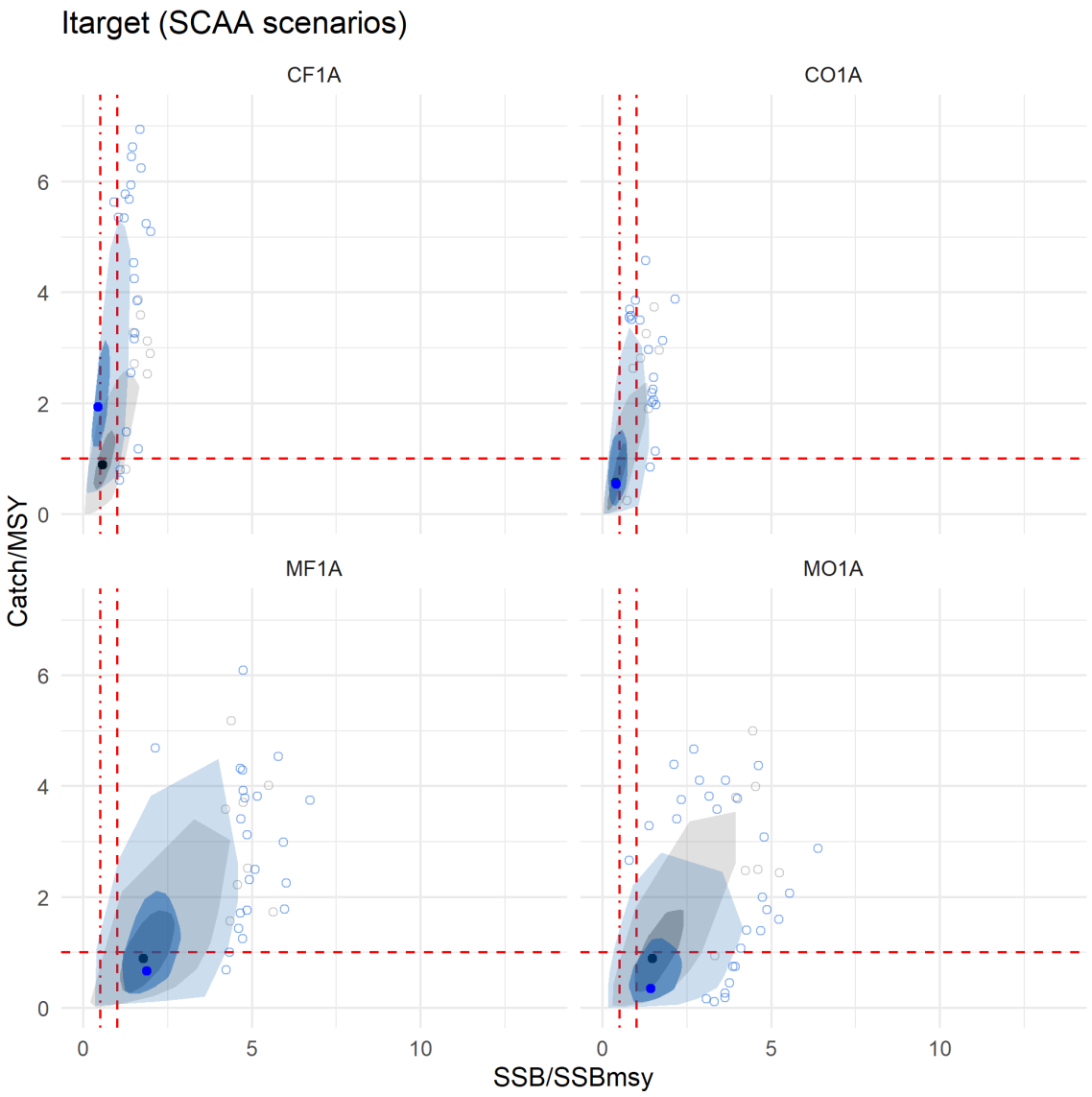


Figure A6.356. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

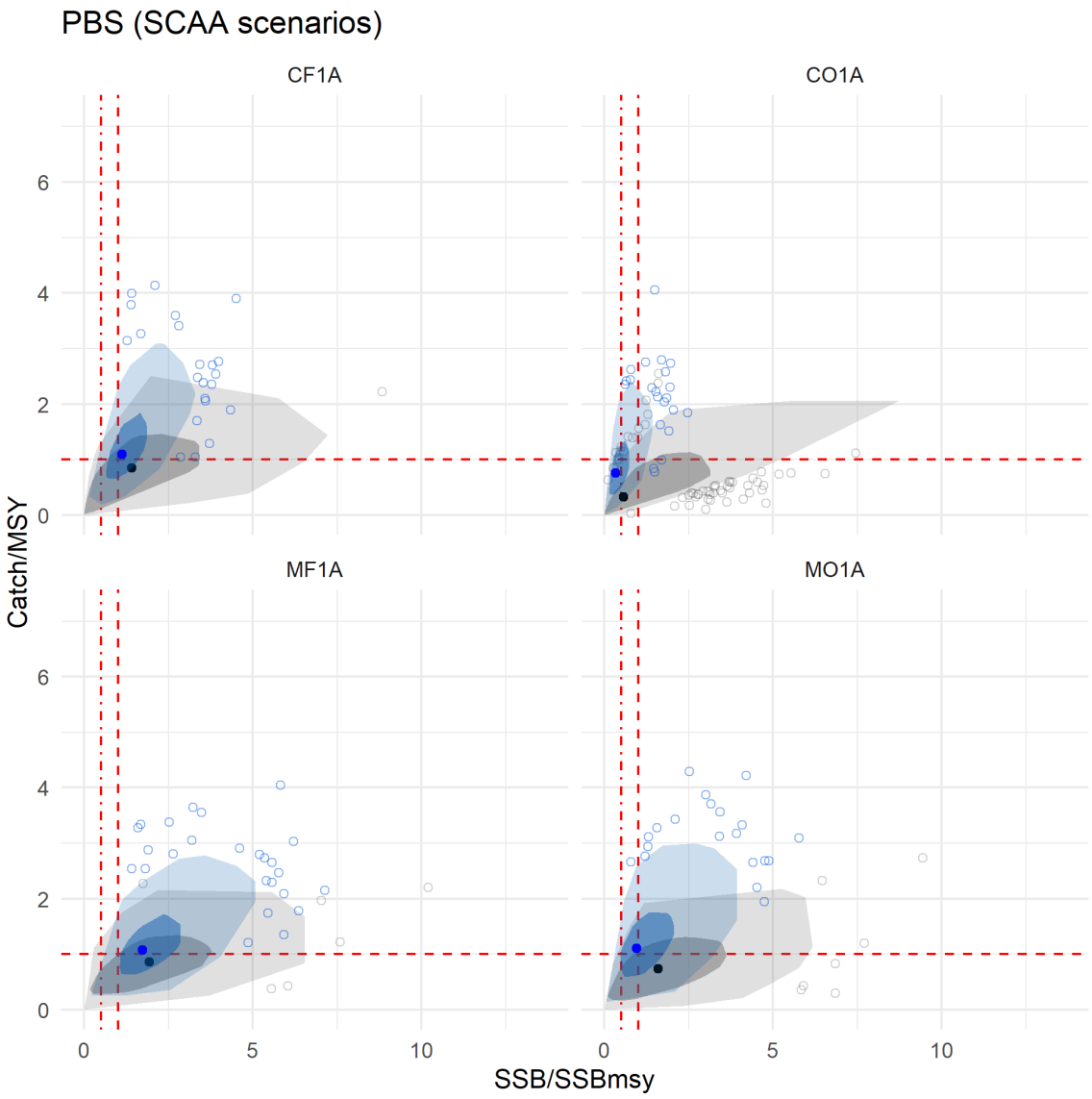


Figure A6.357. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

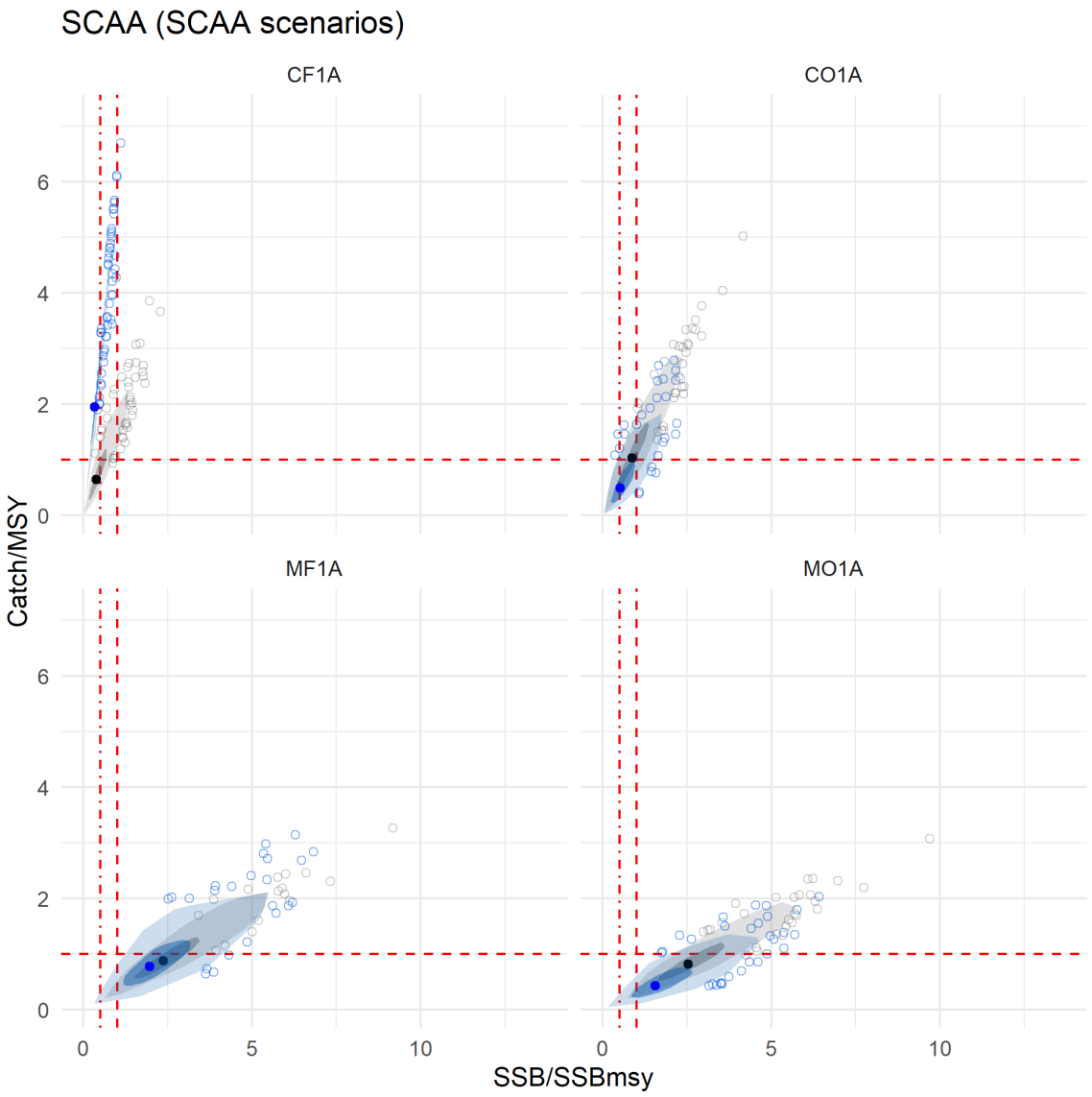


Figure A6.358. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

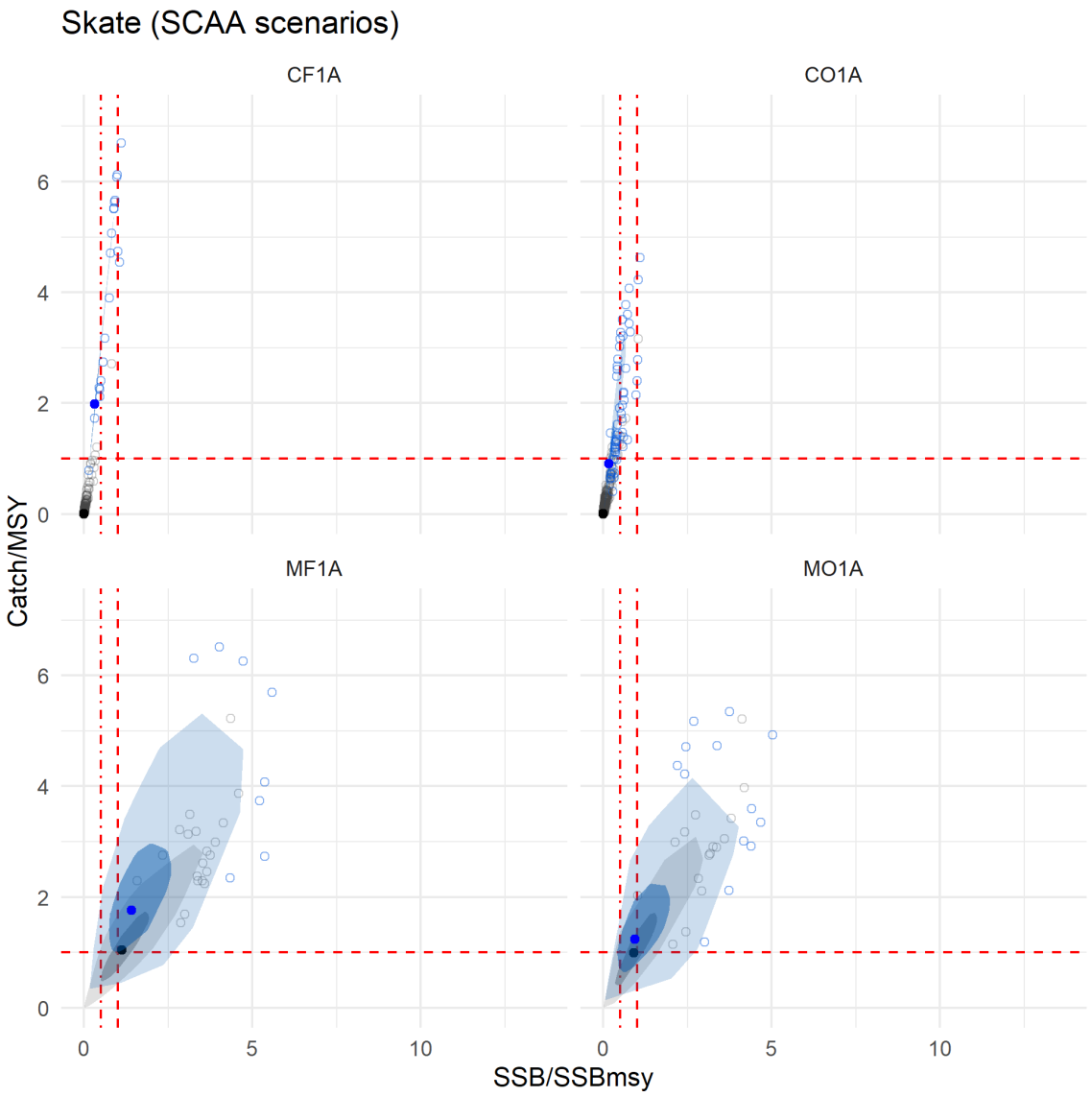


Figure A6.359. Bagplots (a bivariate generalization of the boxplot) for long-term (black) and short-term (blue) spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) and catch/MSY for each scenario using the index-based method (IBM) defined in the top left. The solid dot is the median, the dark shading is the 2-D equivalent of the inner quartile range, the light shading encompasses an area 3 times the bag, and the unfilled dots are outliers.

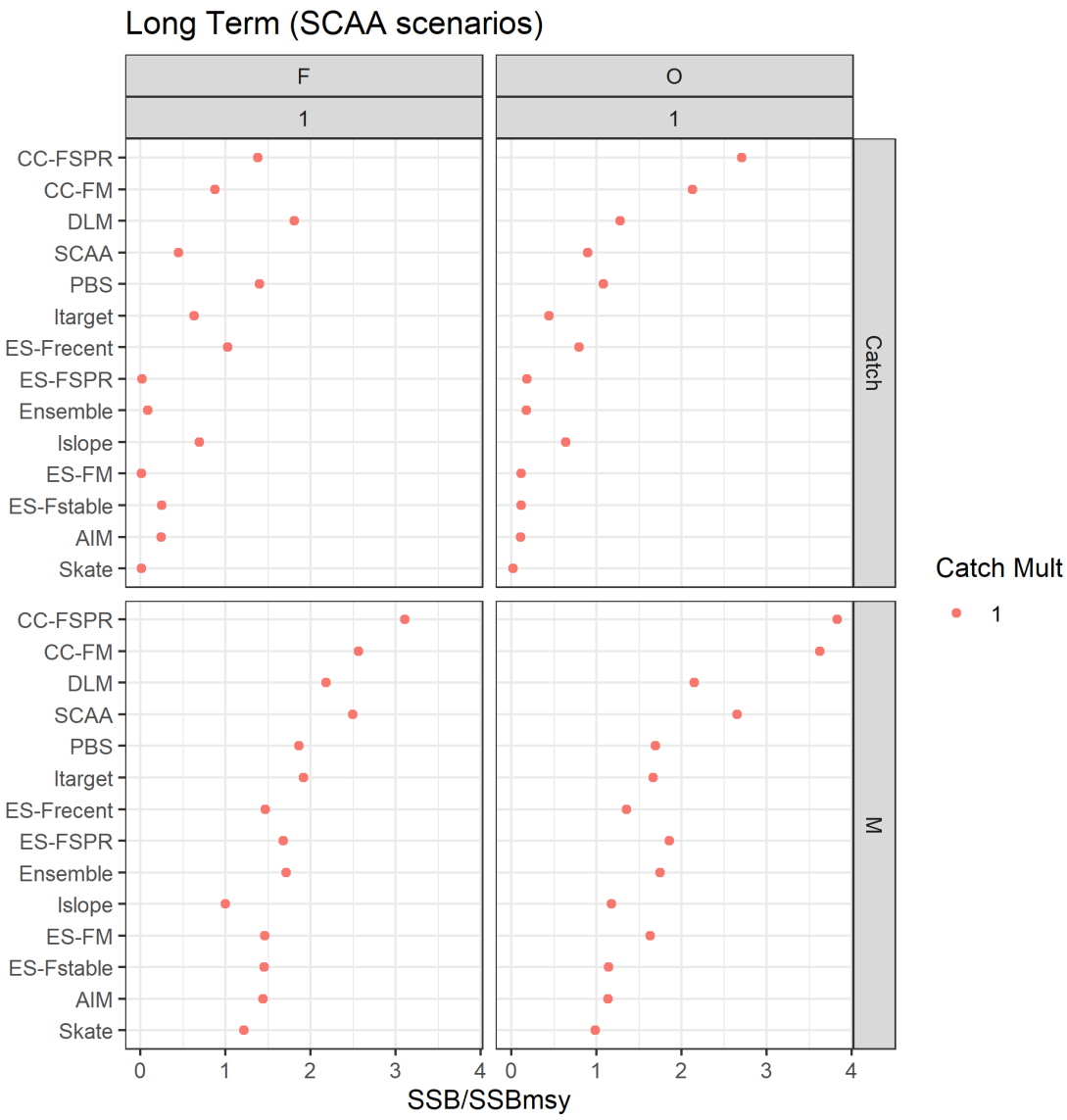


Figure A6.360. Mean values of spawning stock biomass (SSB) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

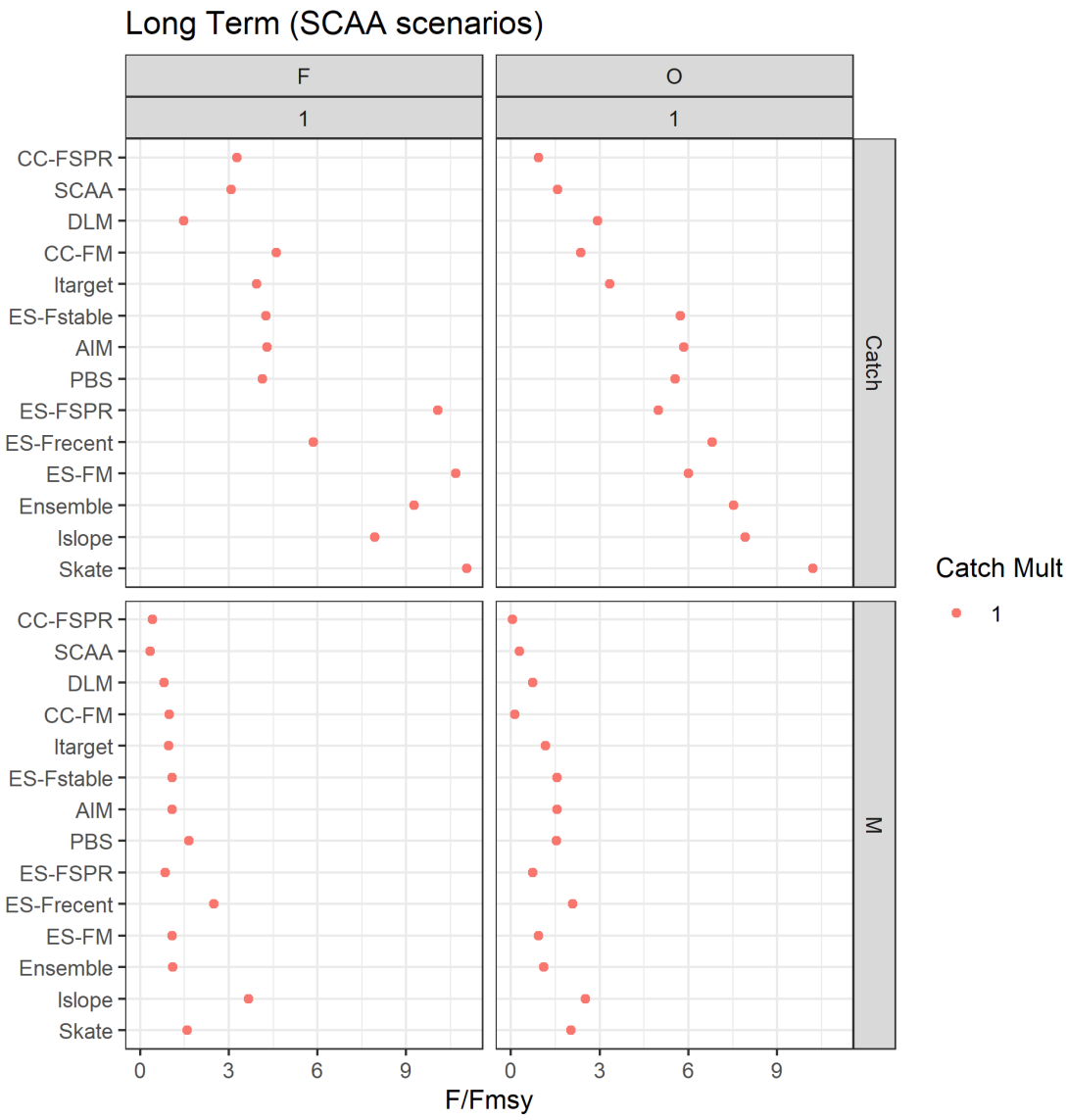


Figure A6.361. Mean values of fishing mortality rate (F) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of F history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

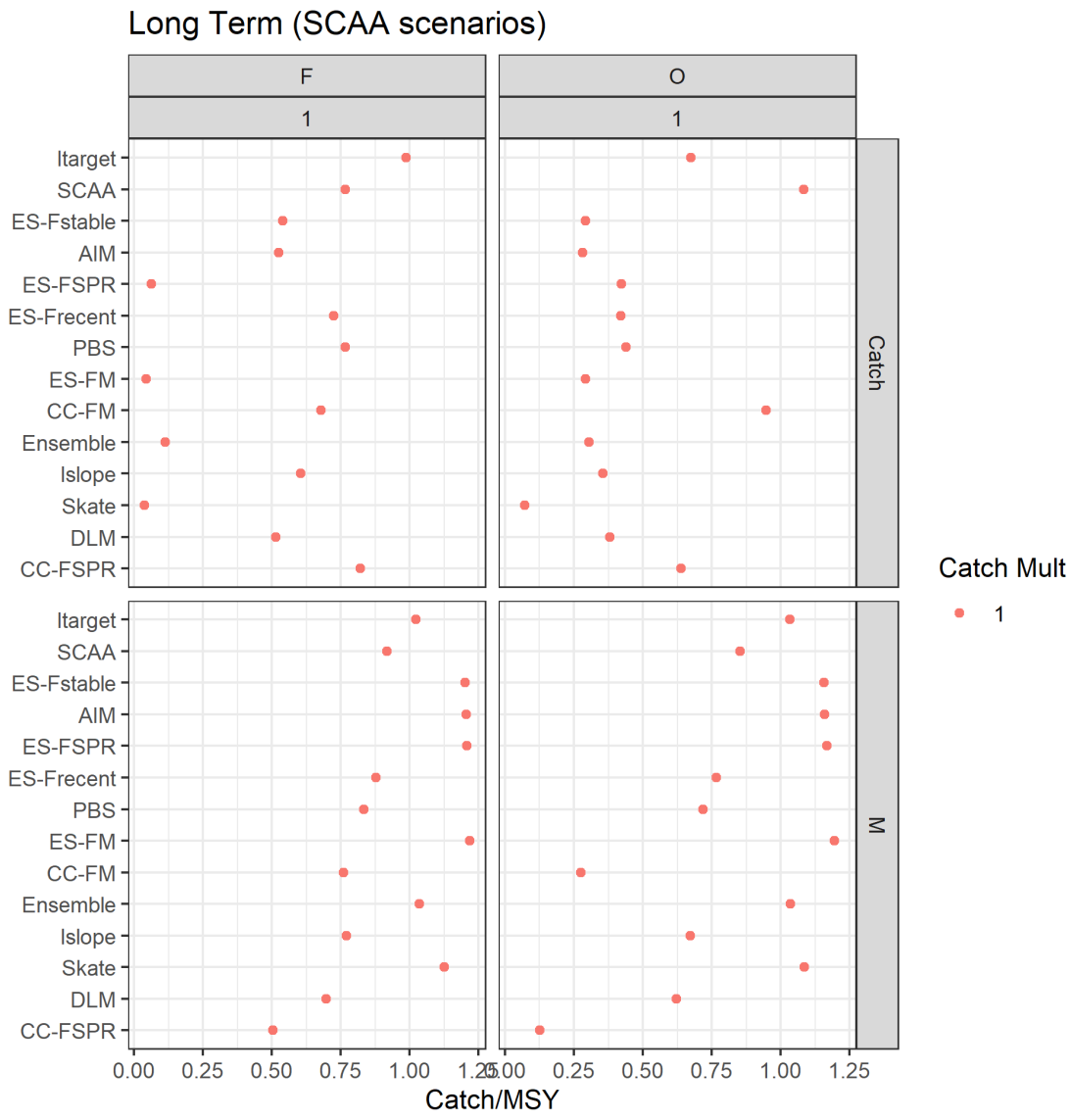


Figure A6.362. Mean values of catch relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the long term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

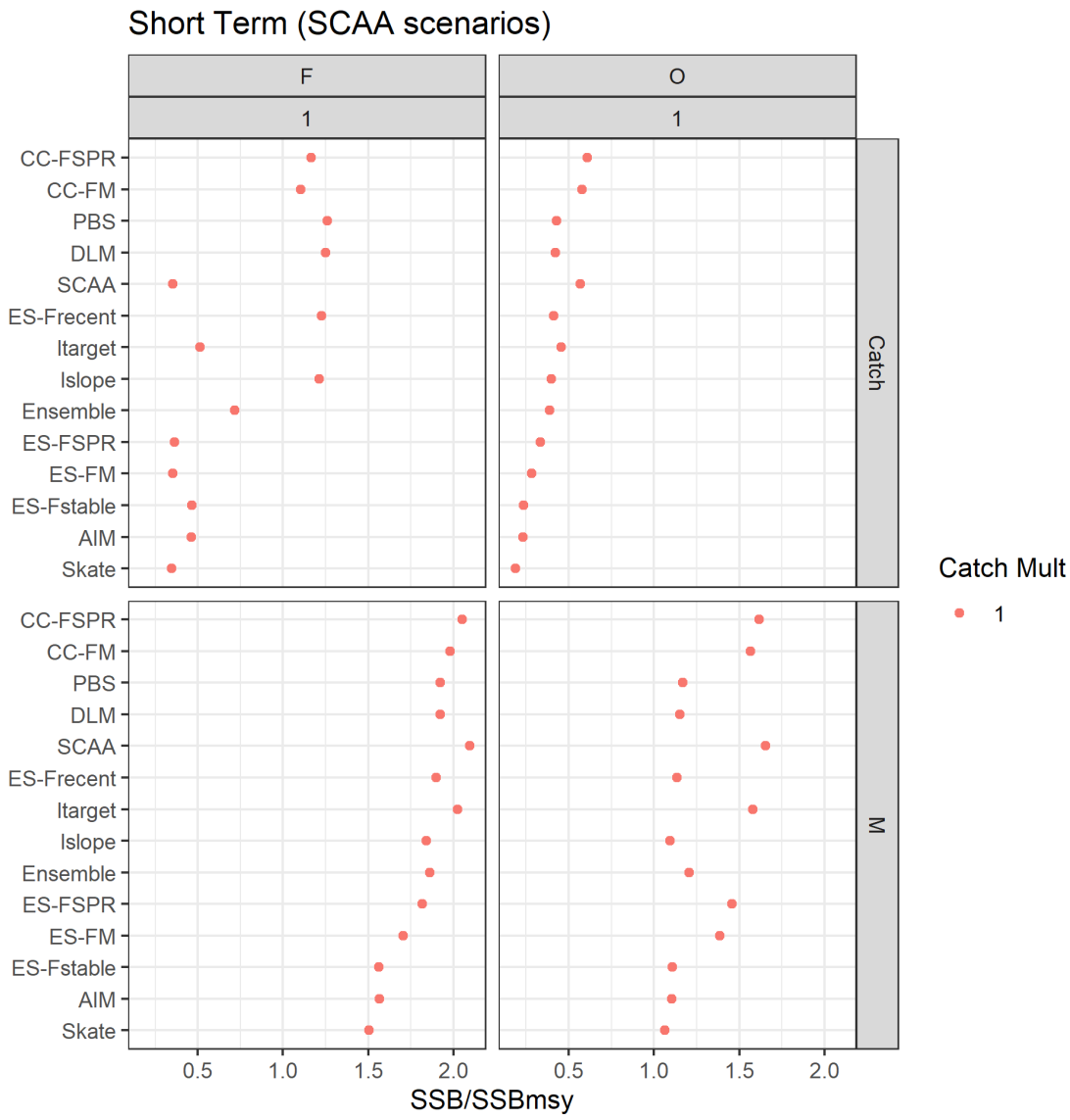


Figure A6.363. Mean values of spawning stock biomass (SSB) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

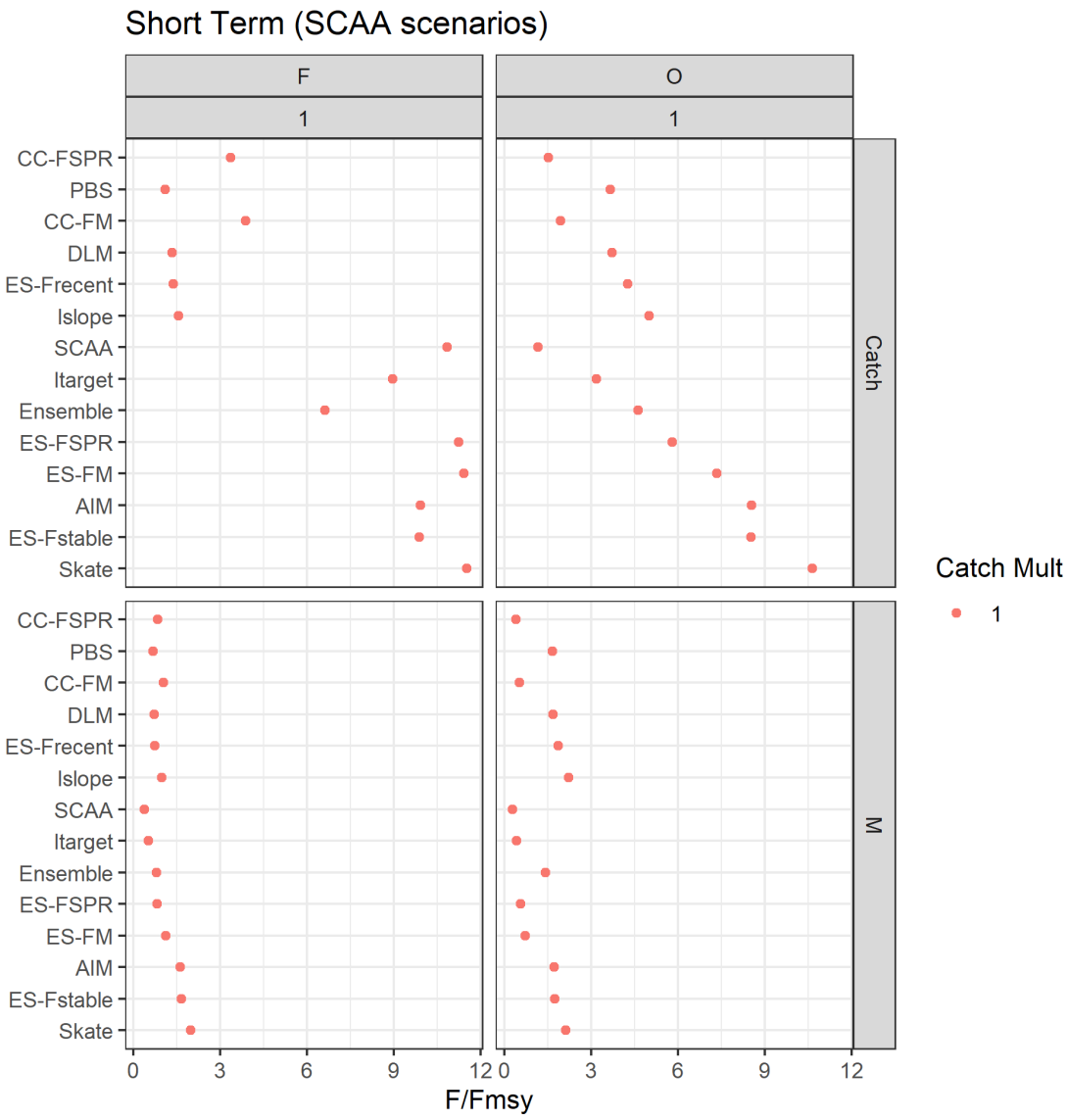


Figure A6.364. Mean values of fishing mortality rate (F) relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of F history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

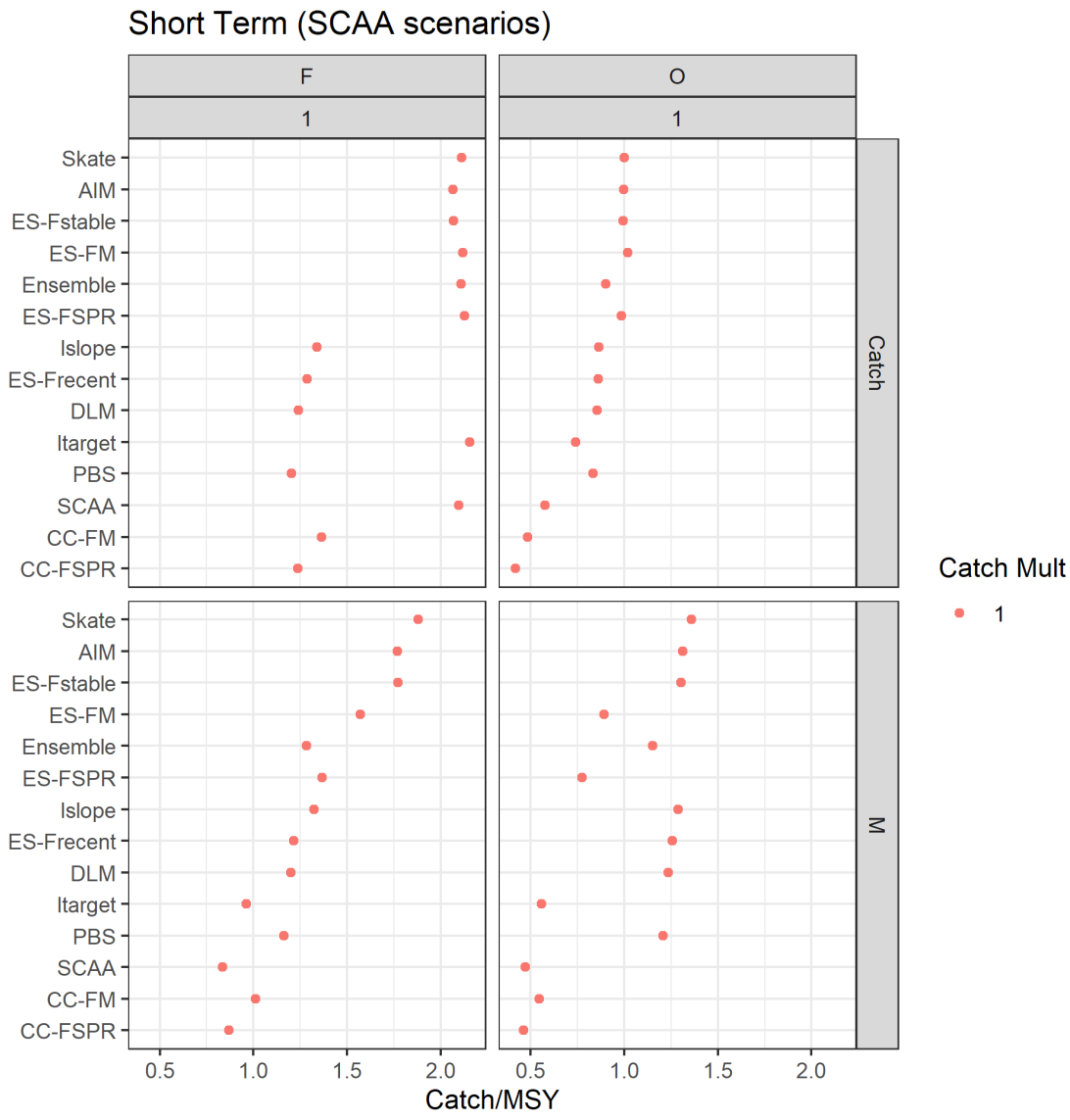


Figure A6.365. Mean values of catch relative to its maximum sustainable yield (MSY) reference point by index-based method (IBM) in the short term with color denoting the catch advice multiplier for the 8 combinations of fishing mortality rate (F) history (F = F_{MSY} in second half of base period, O = overfishing throughout), number of fishery selectivity blocks (1 or 2), and retrospective source (catch or natural mortality [M]). The IBMs are sorted by the mean across the 16 scenarios and arranged so that higher values are at the top.

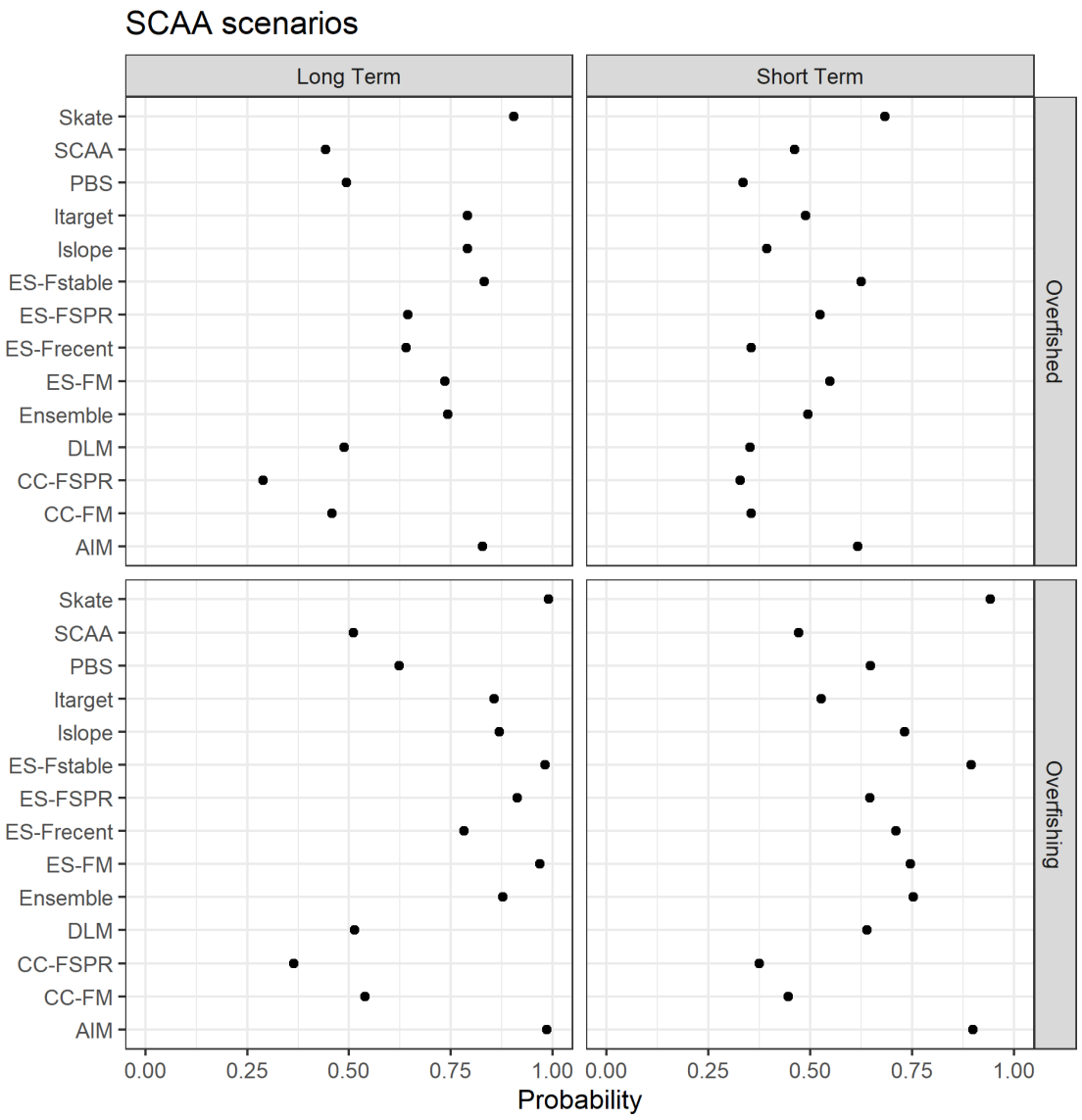


Figure A6.366. Probability stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all statistical catch-at-age (SCAA) analyses.

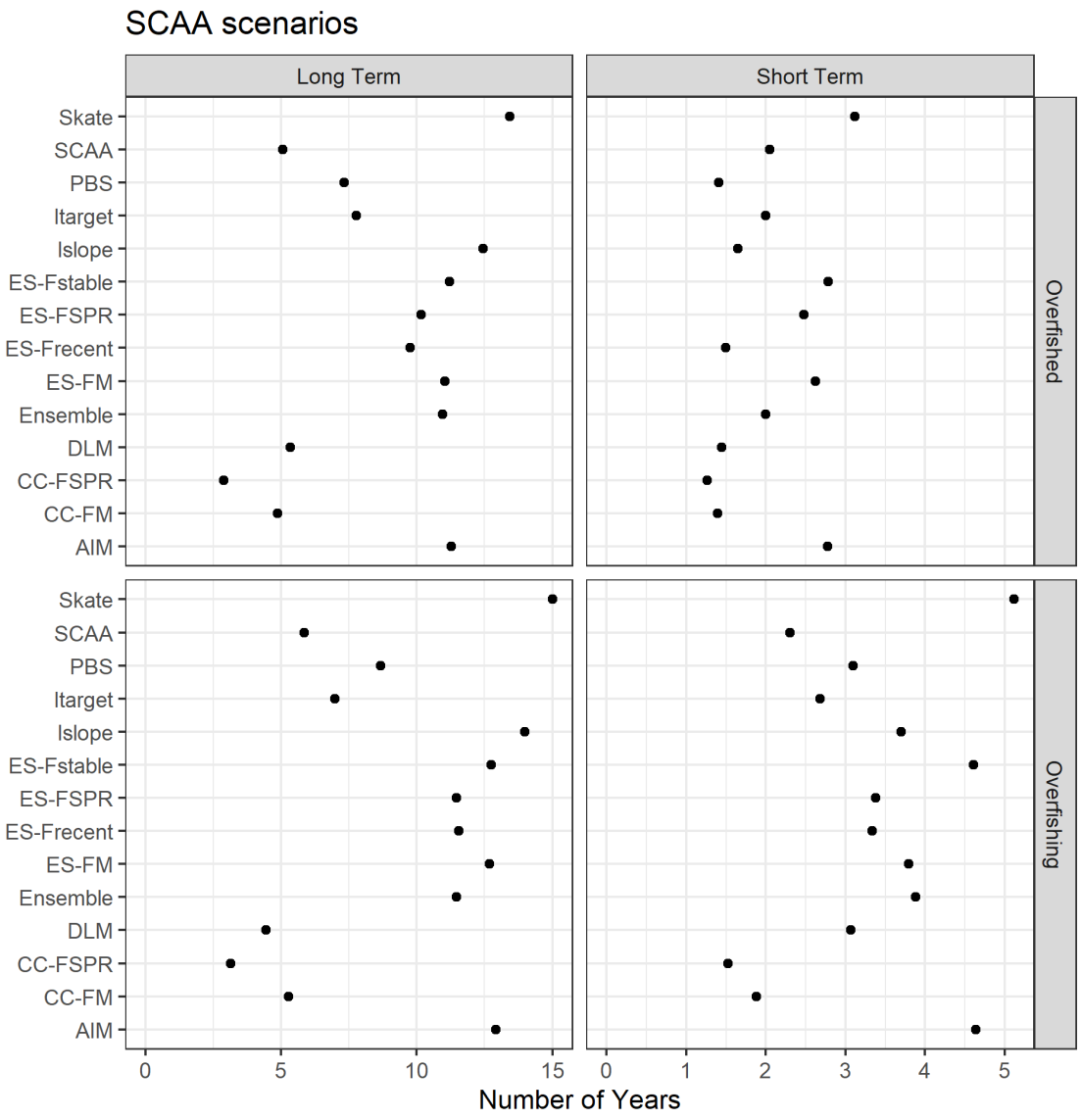


Figure A6.367. Number of years stock is overfished or undergoing overfishing in the short or long term for each index-based method (IBM) averaged across all statistical catch-at-age (SCAA) analyses.

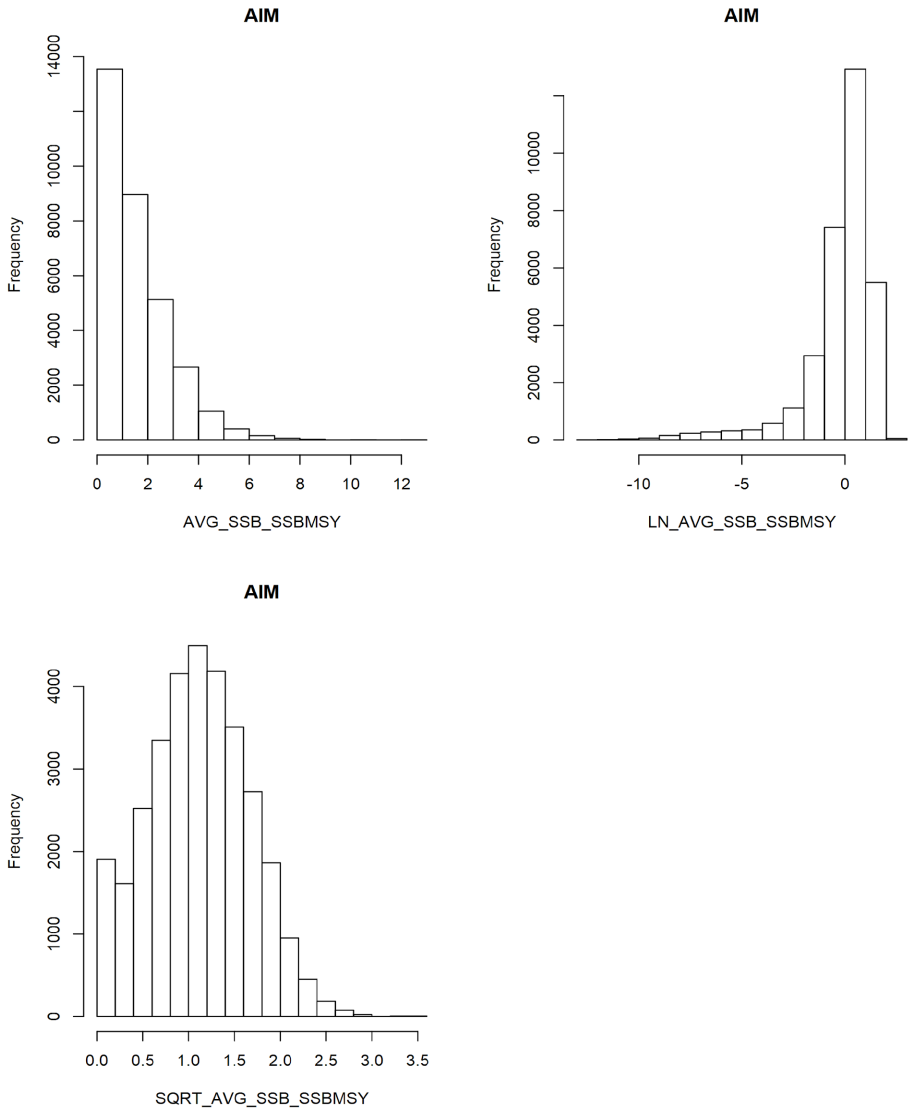


Figure A6.368. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

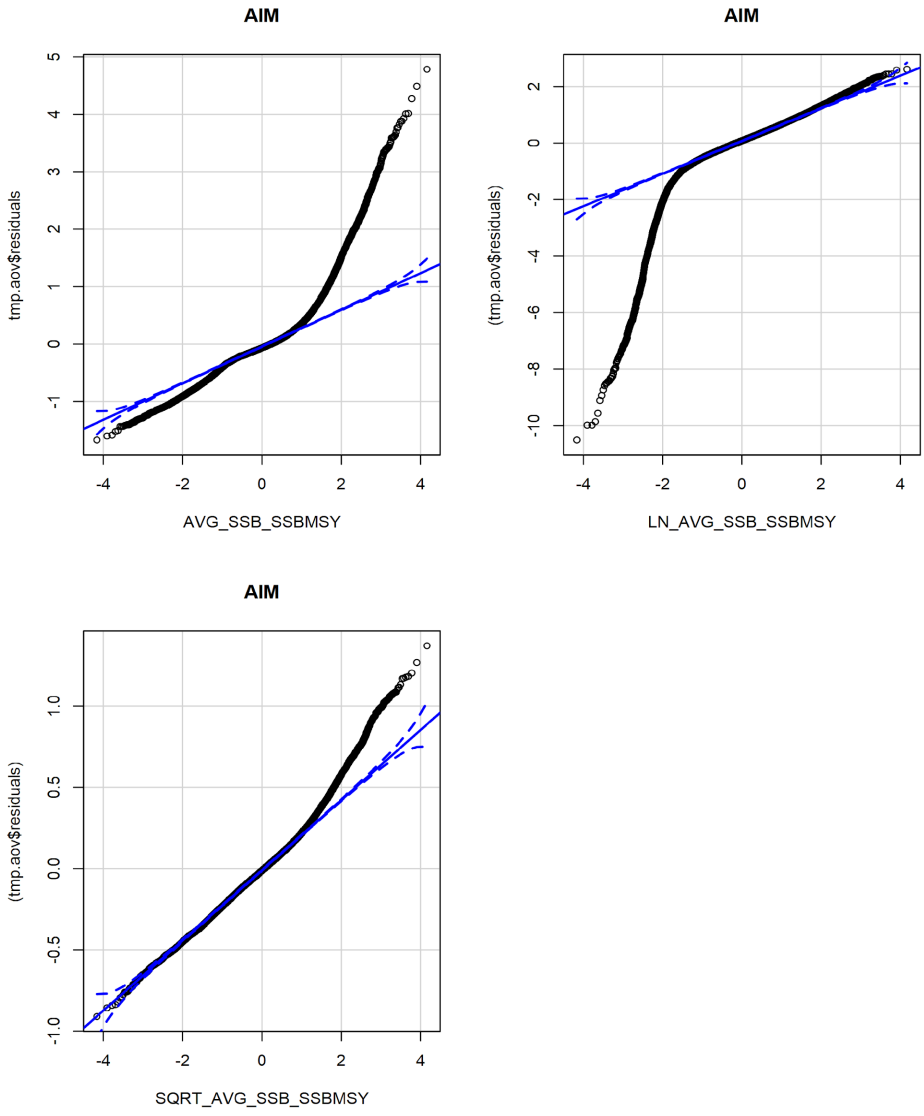


Figure A6.369. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

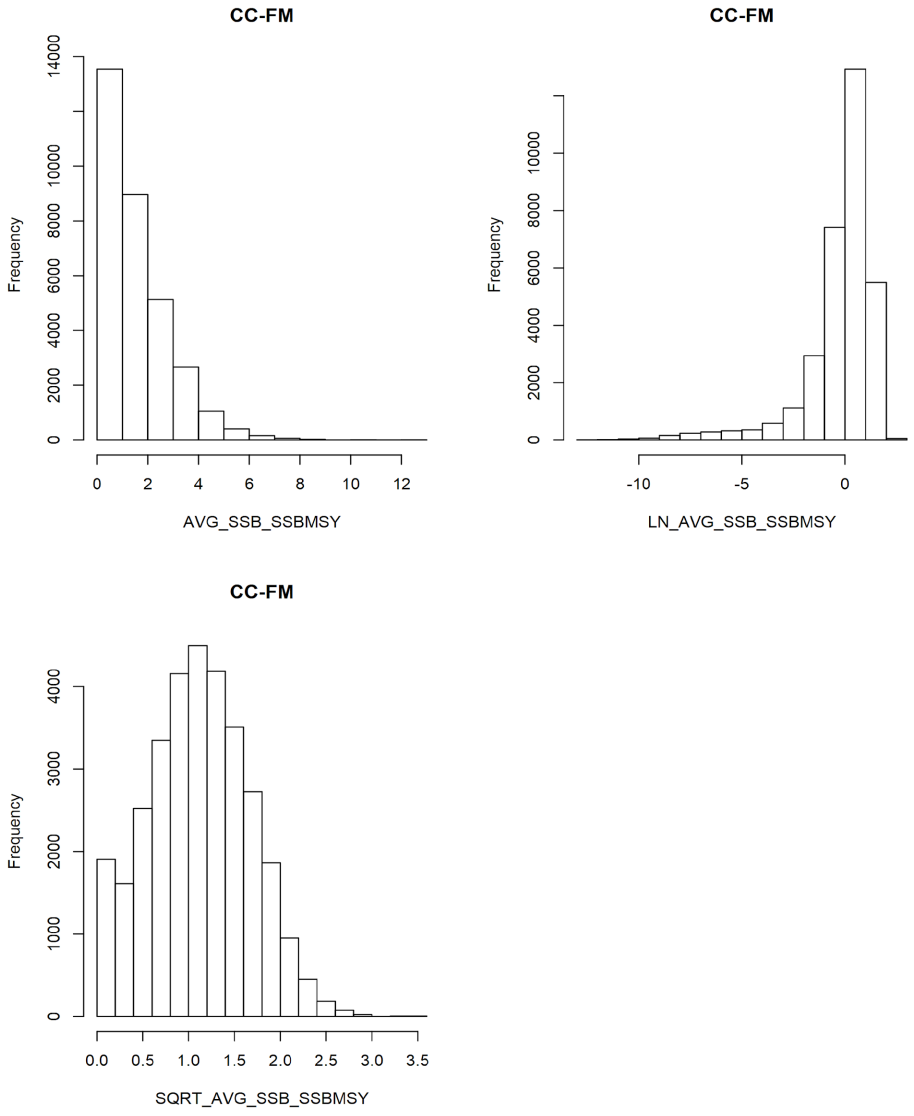


Figure A6.370. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

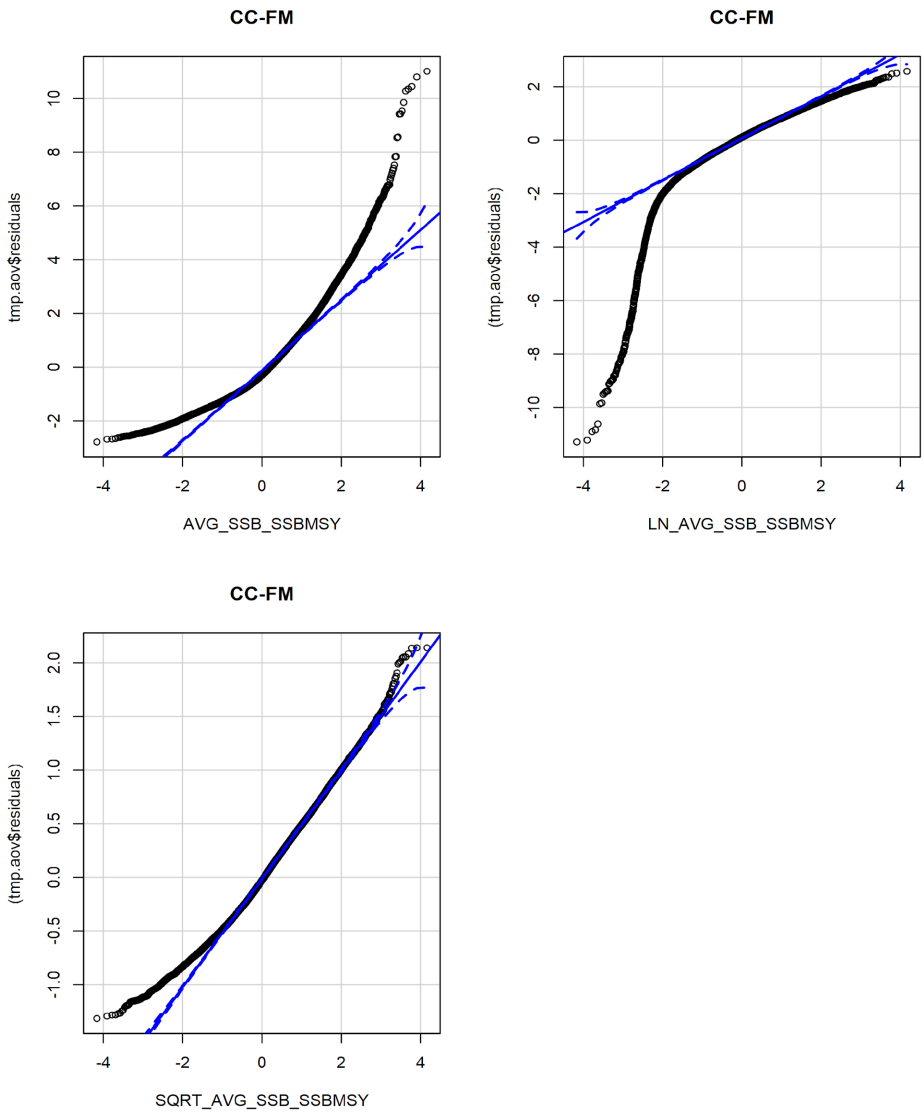


Figure A6.371. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

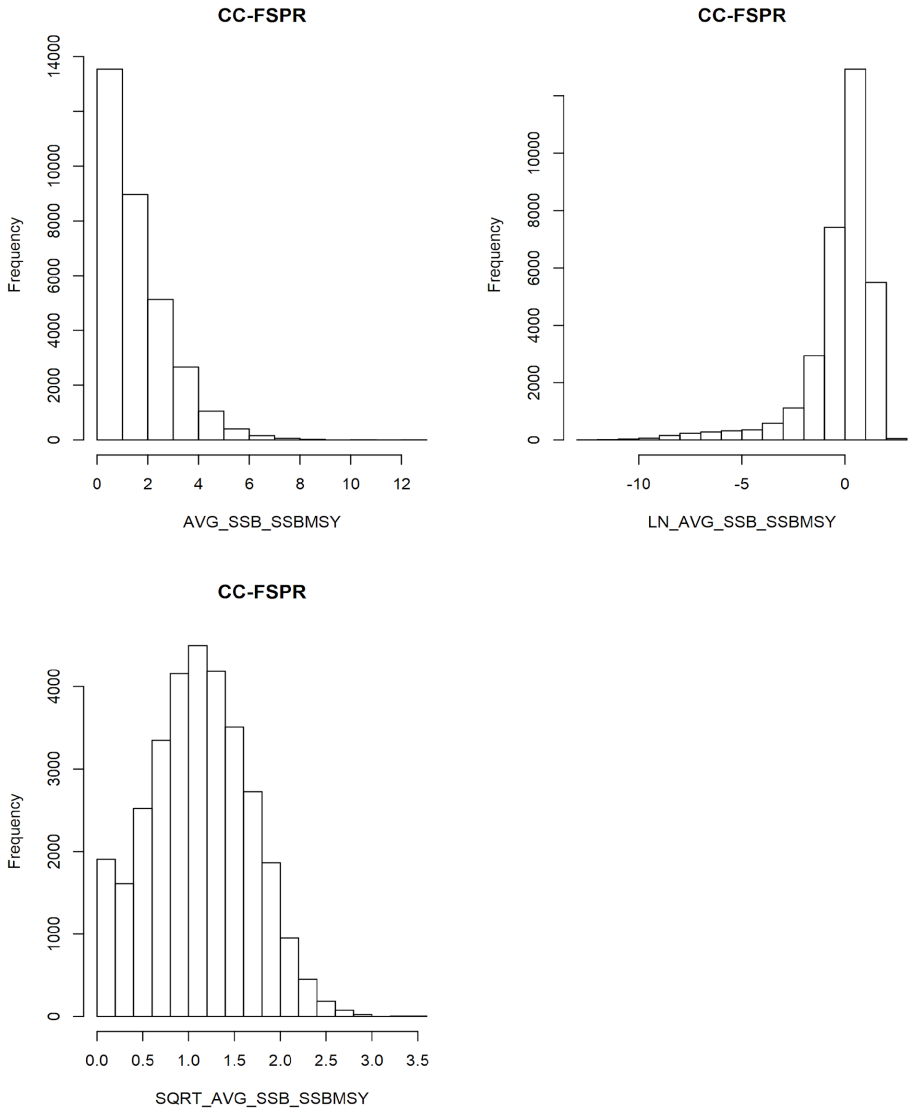


Figure A6.372. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

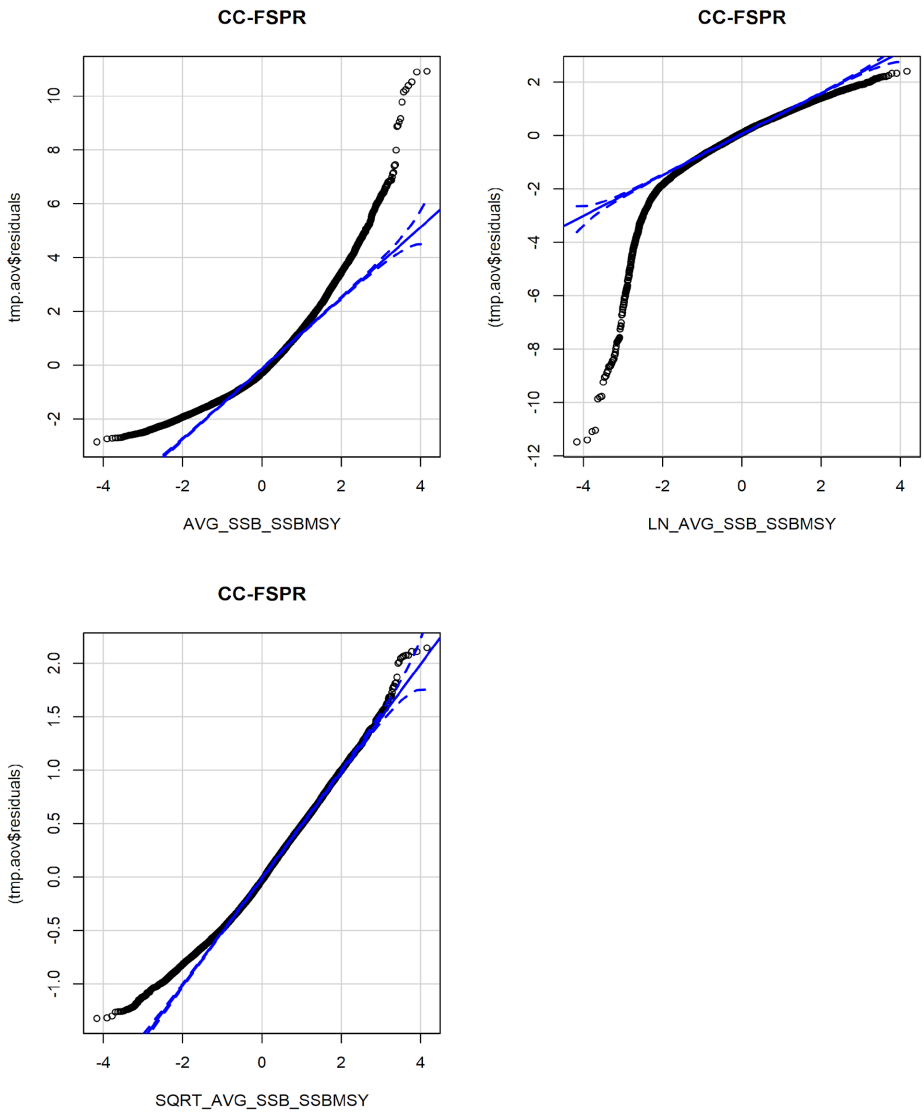


Figure A6.373. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

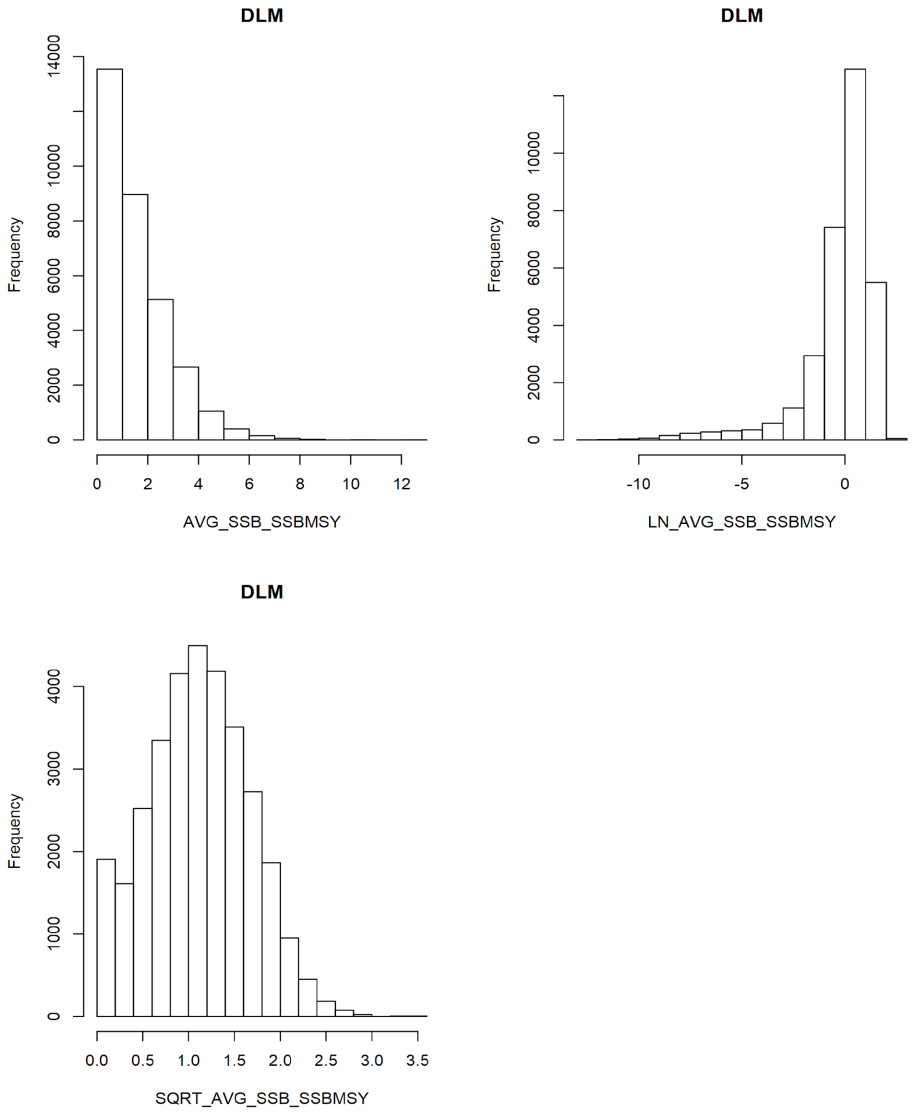


Figure A6.374. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

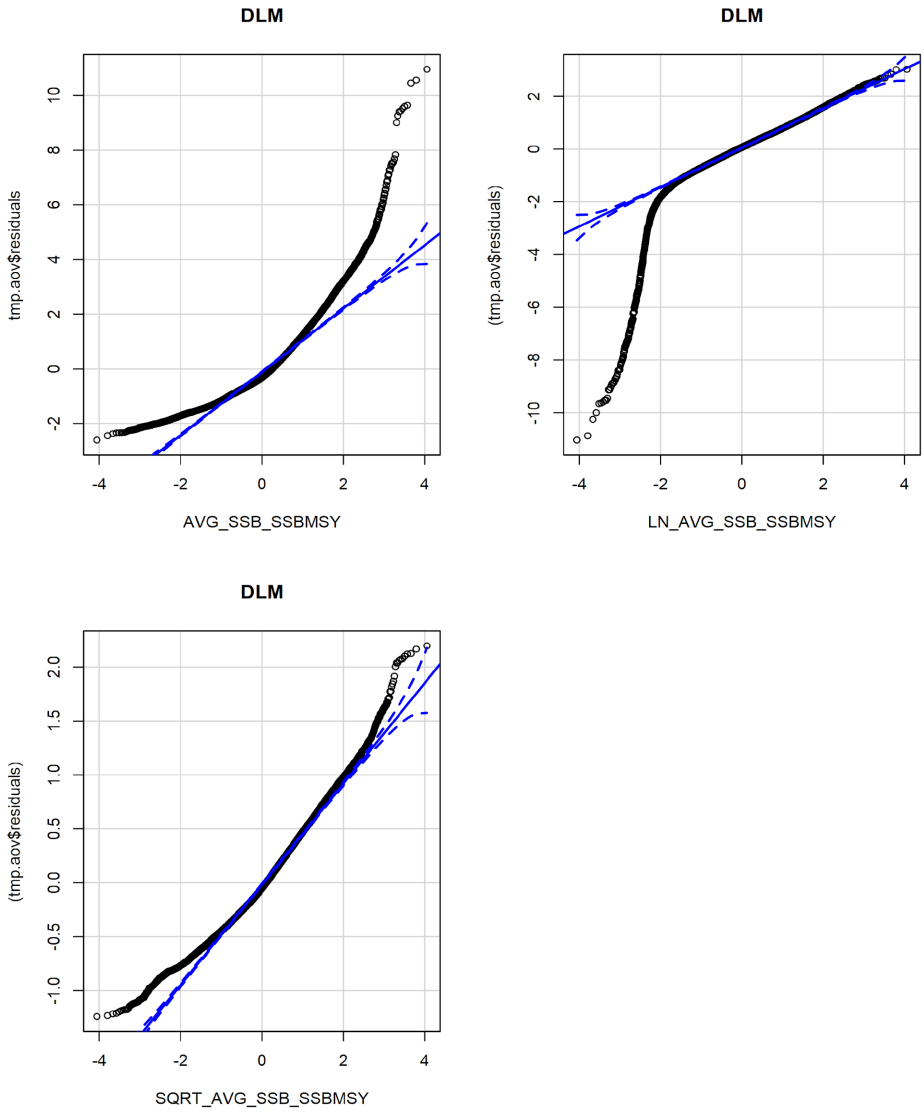


Figure A6.375. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

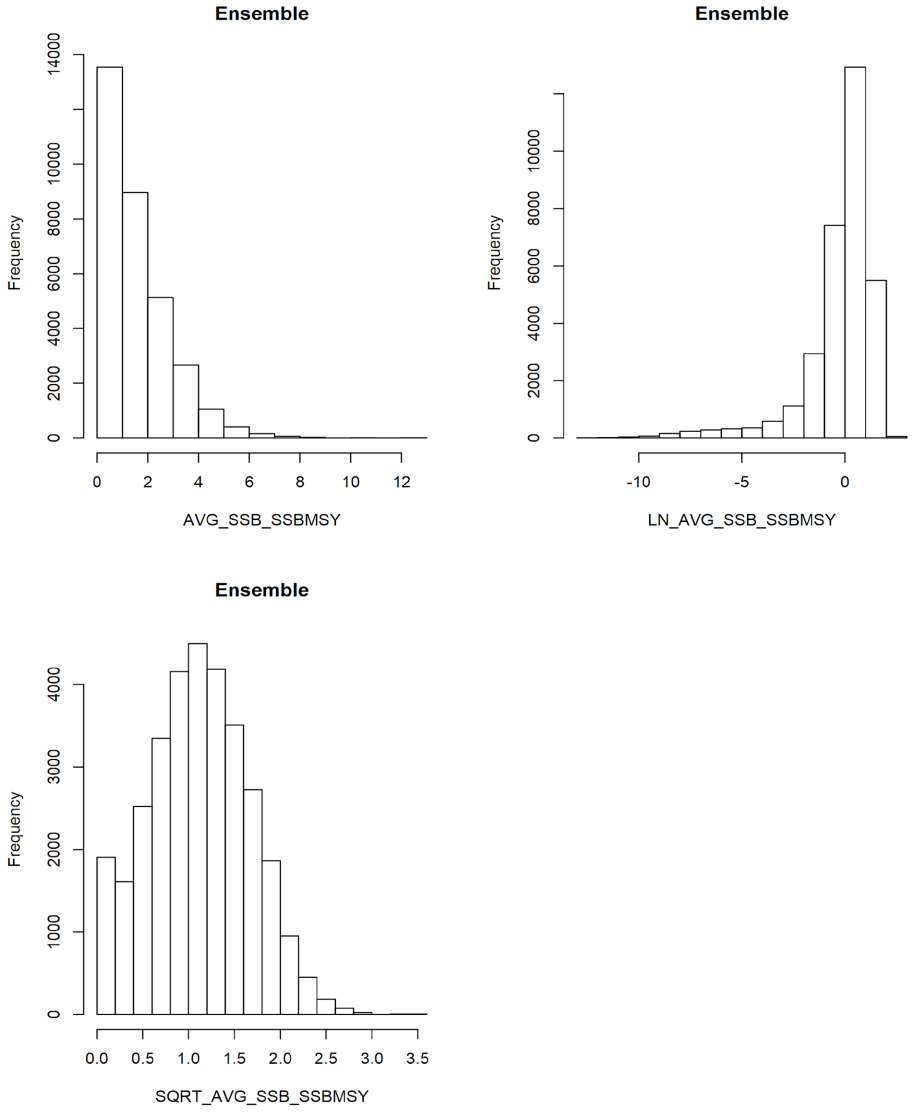


Figure A6.376. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

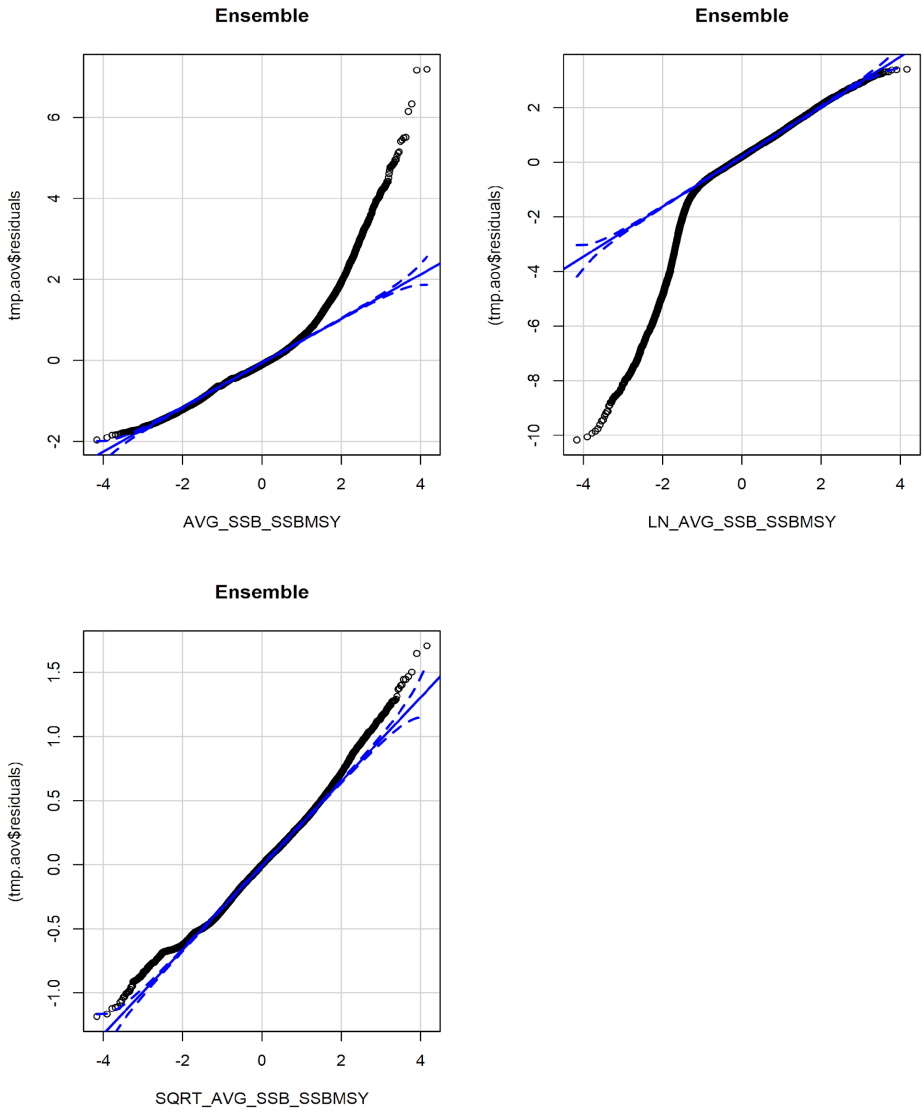


Figure A6.377. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

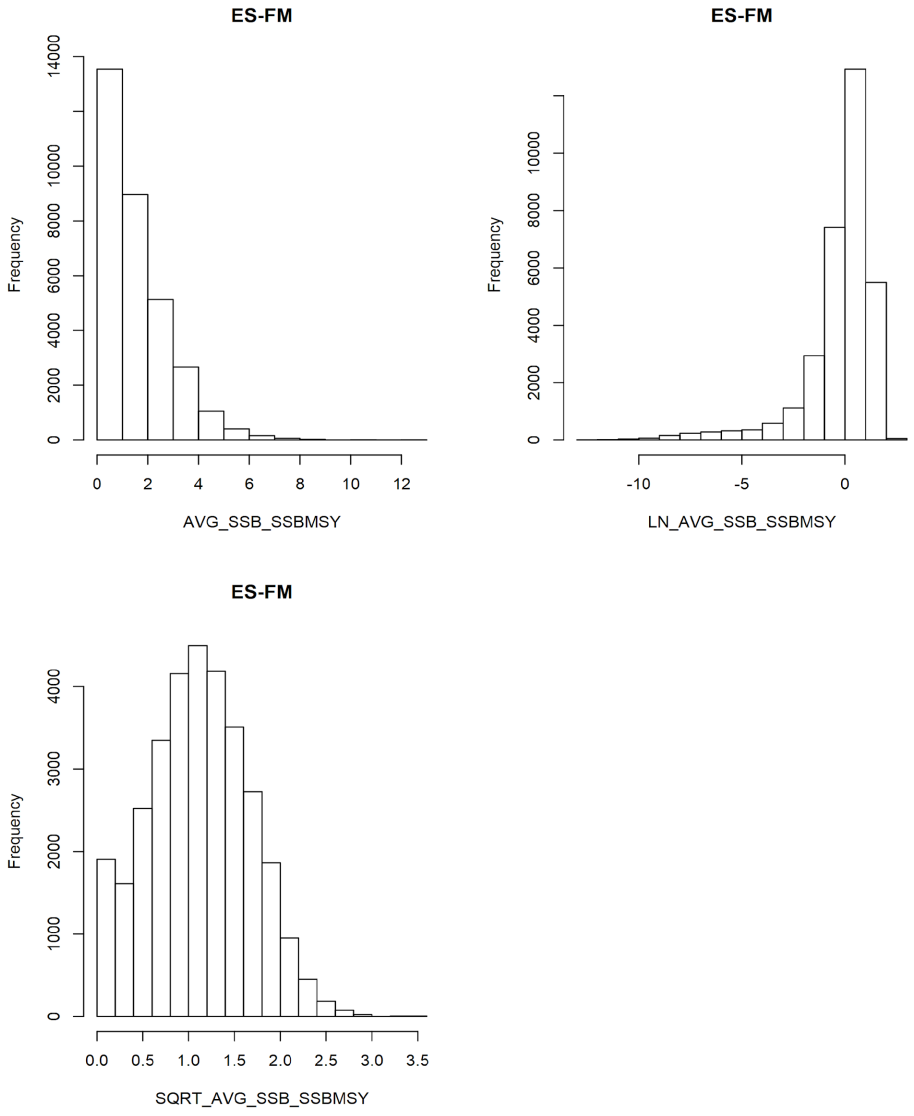


Figure A6.378. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

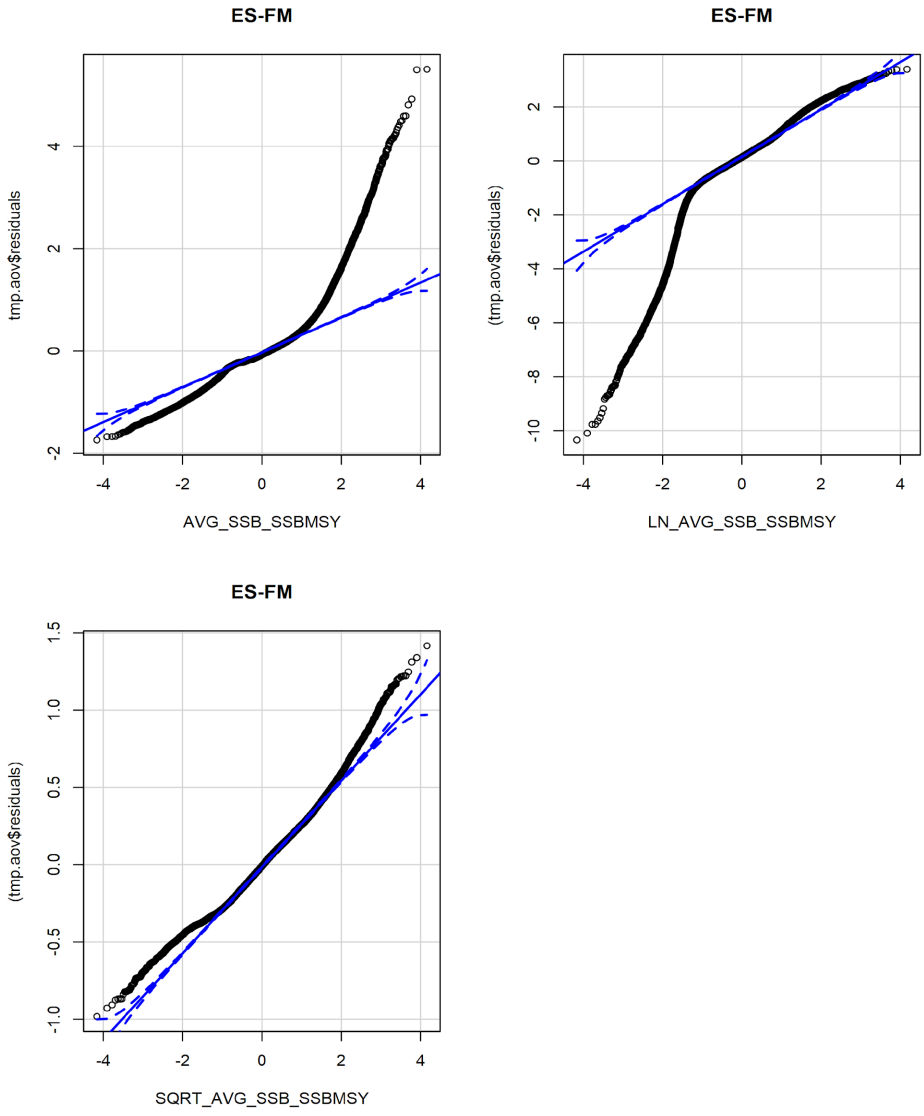


Figure A6.379. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

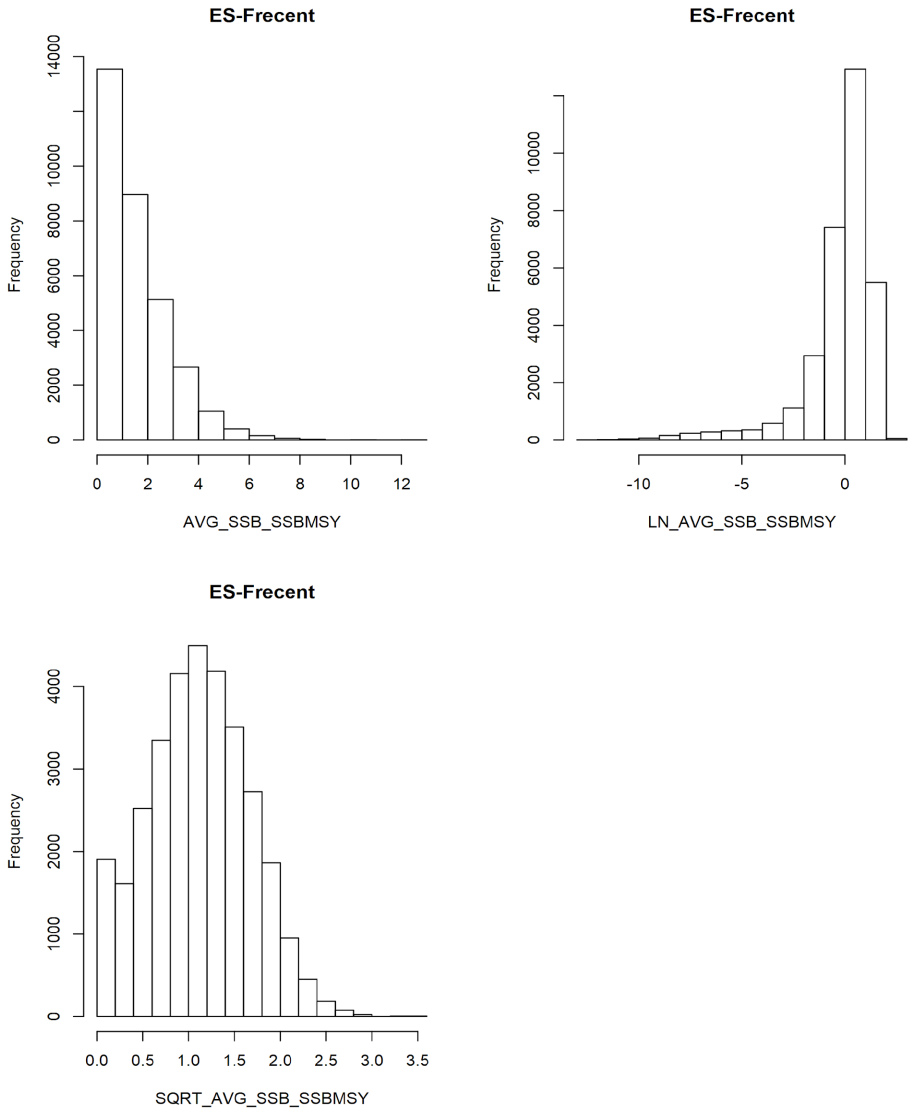


Figure A6.380. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

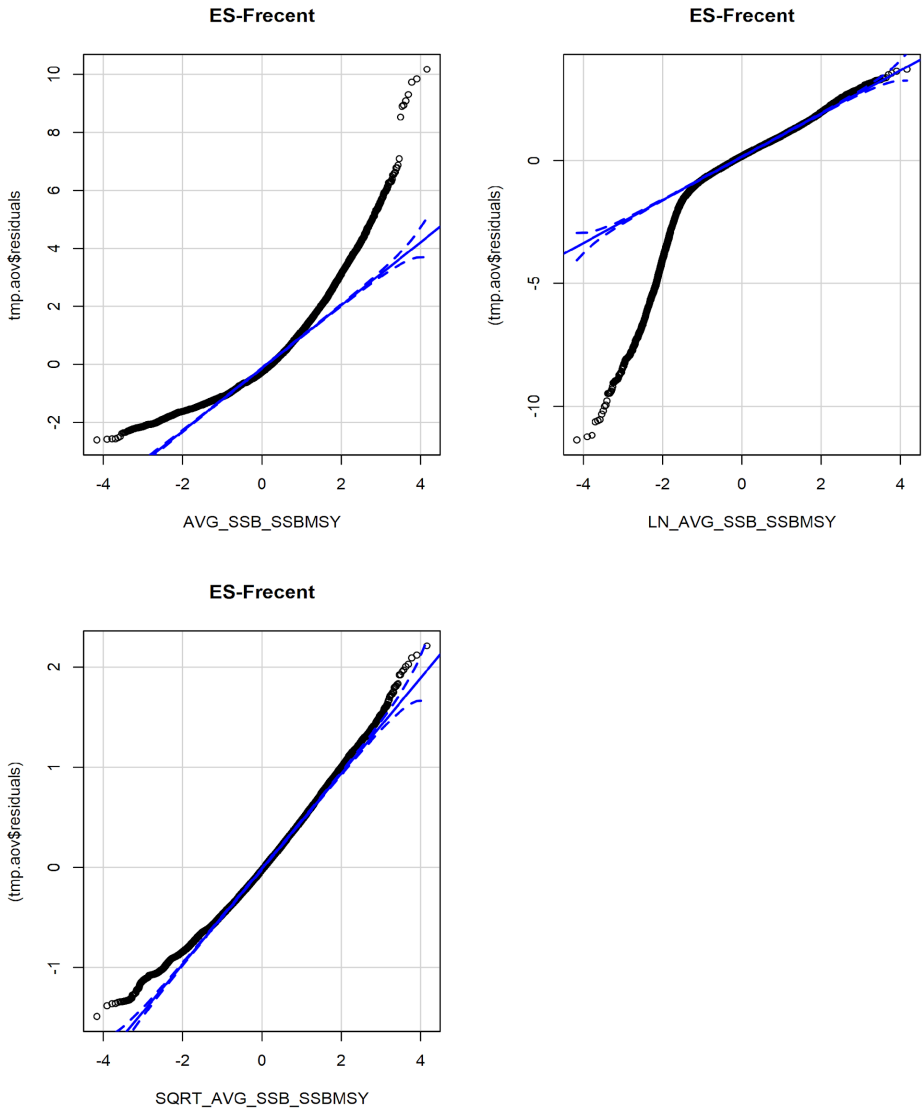


Figure A6.381. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

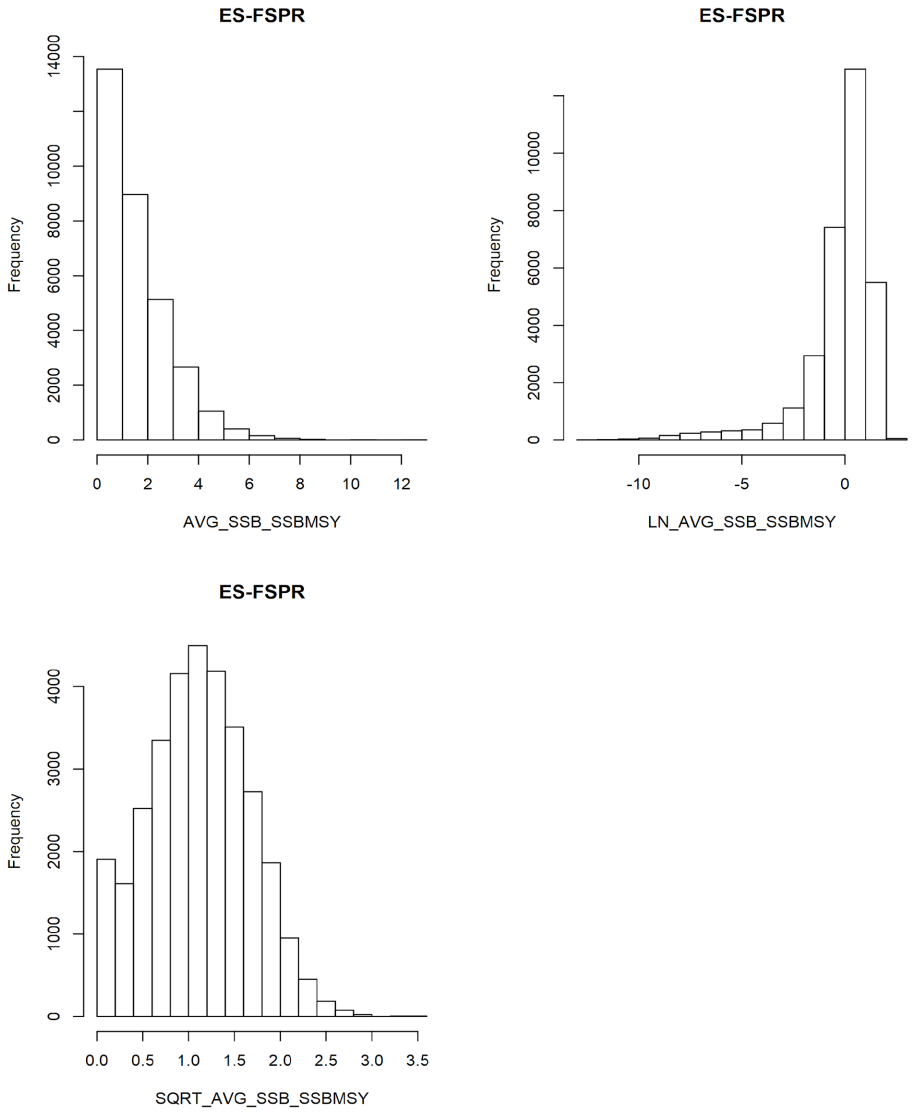


Figure A6.382. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by IBM for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

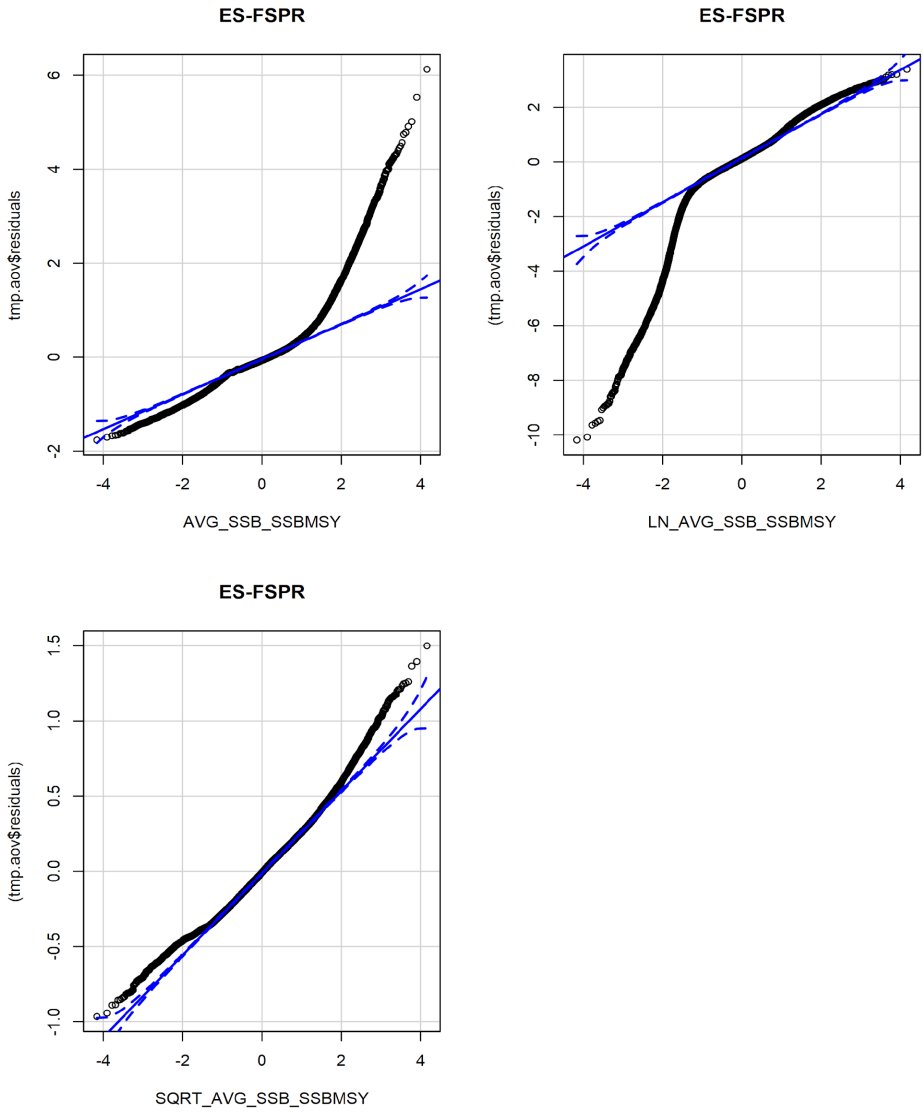


Figure A6.383. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

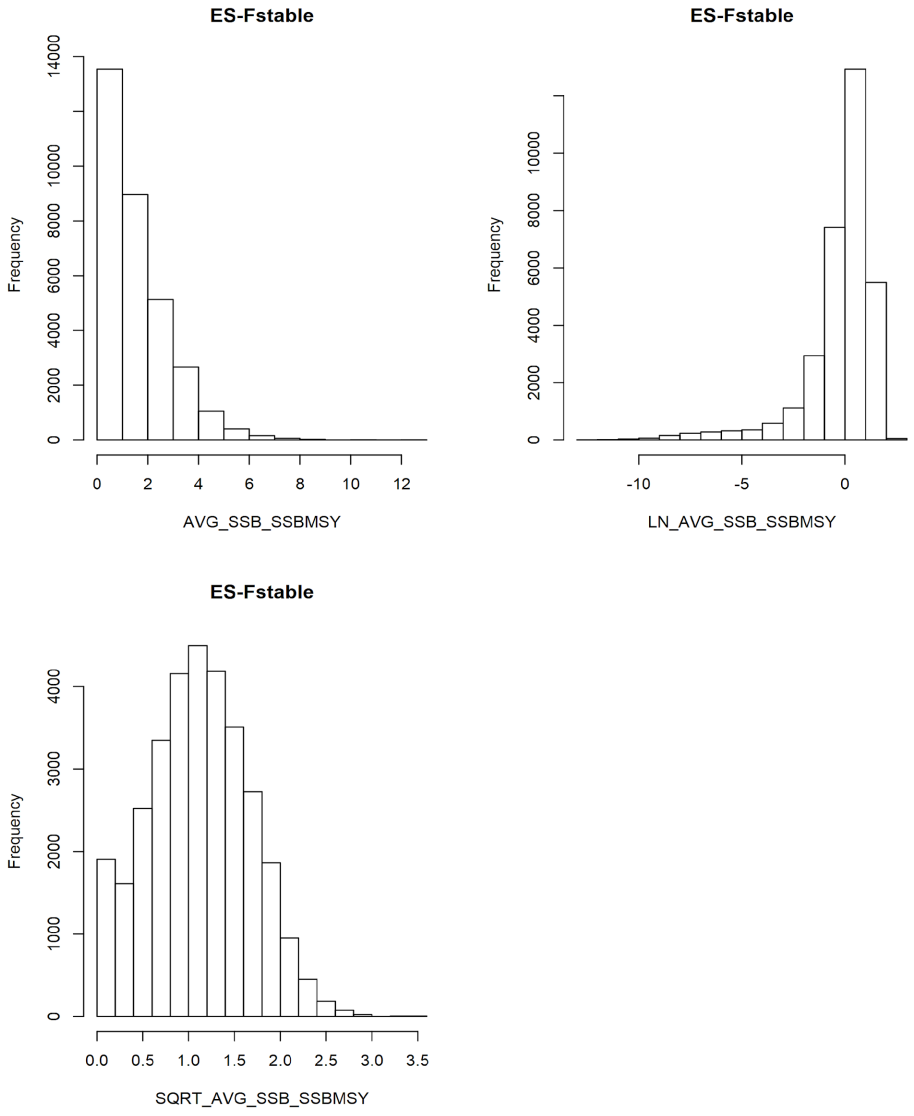


Figure A6.384. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

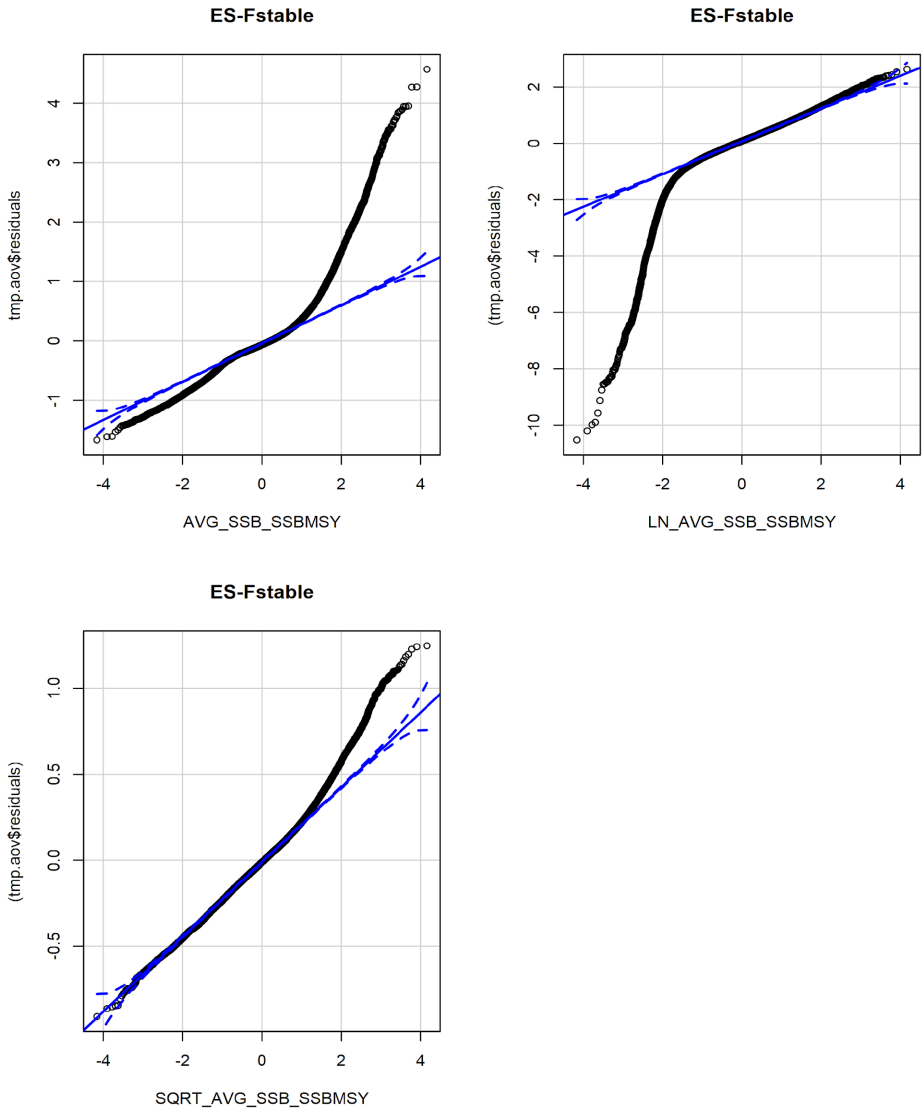


Figure A6.385. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

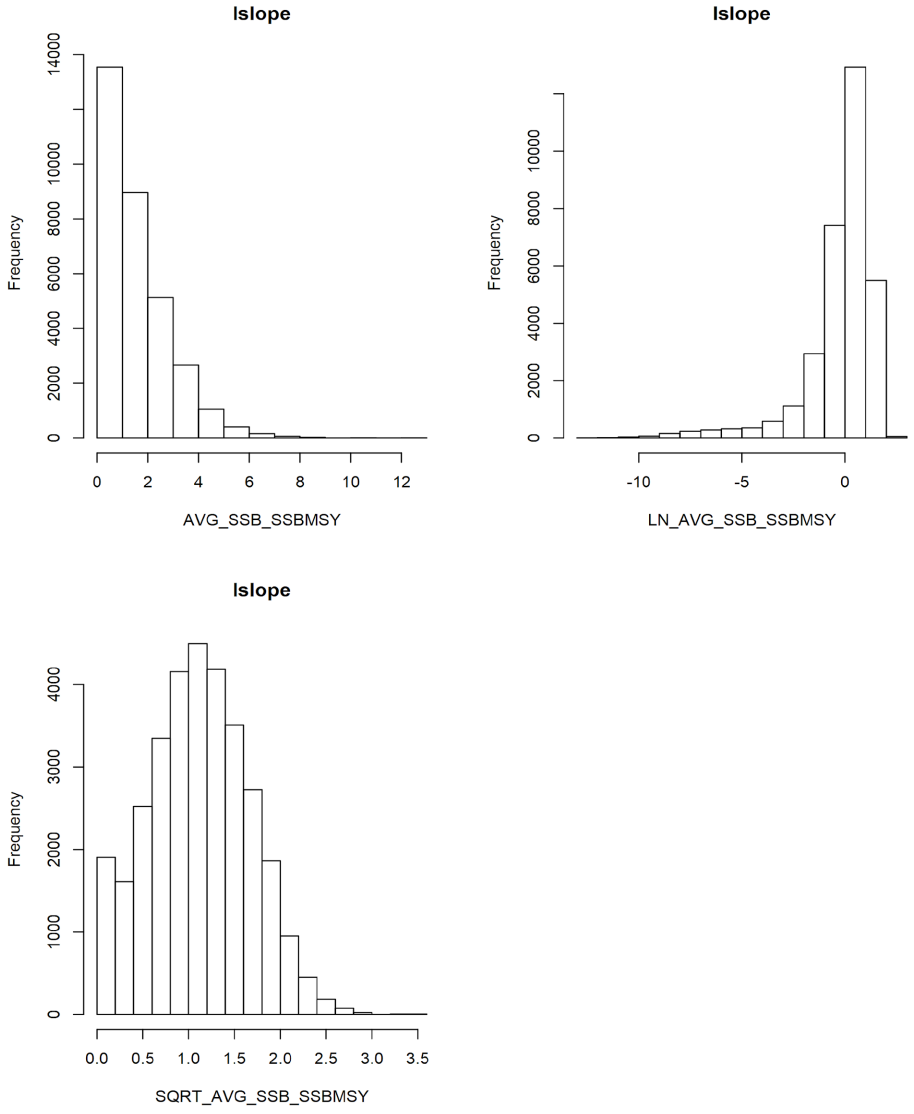


Figure A6.386. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

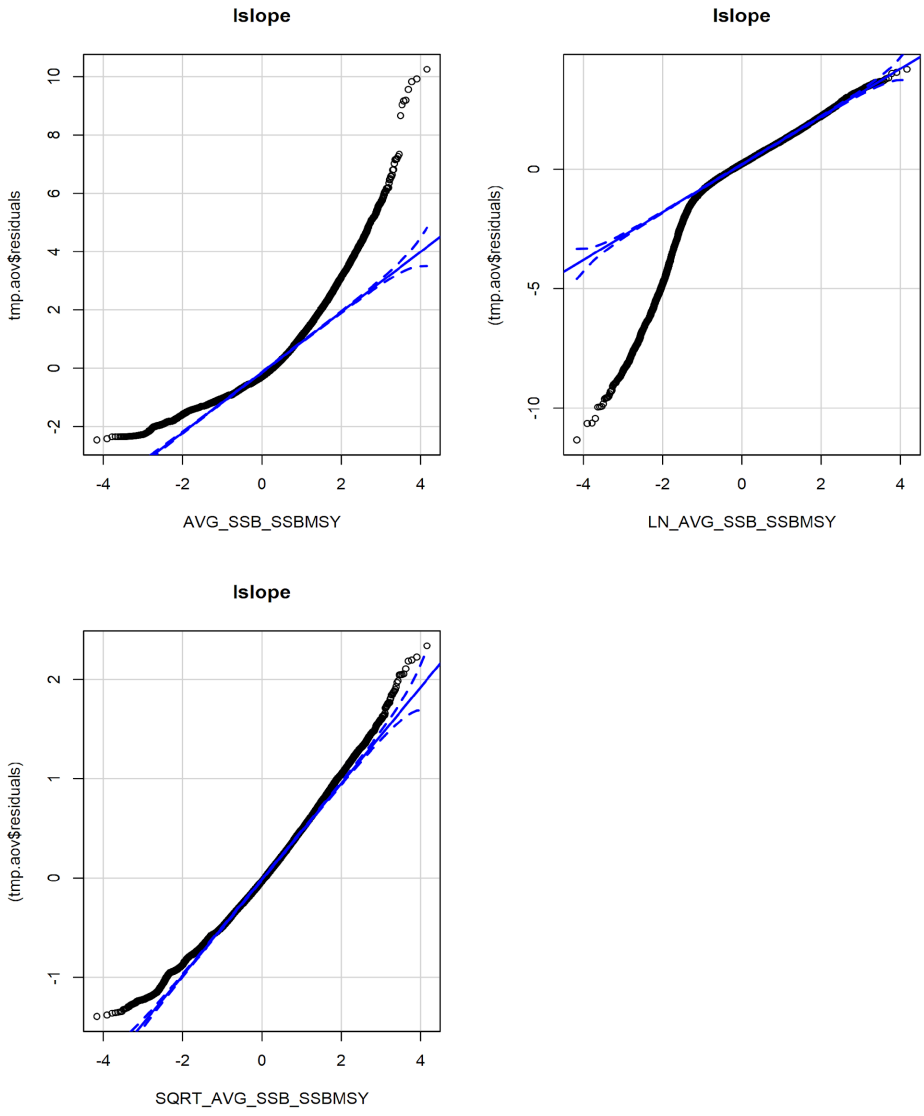


Figure A6.387. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

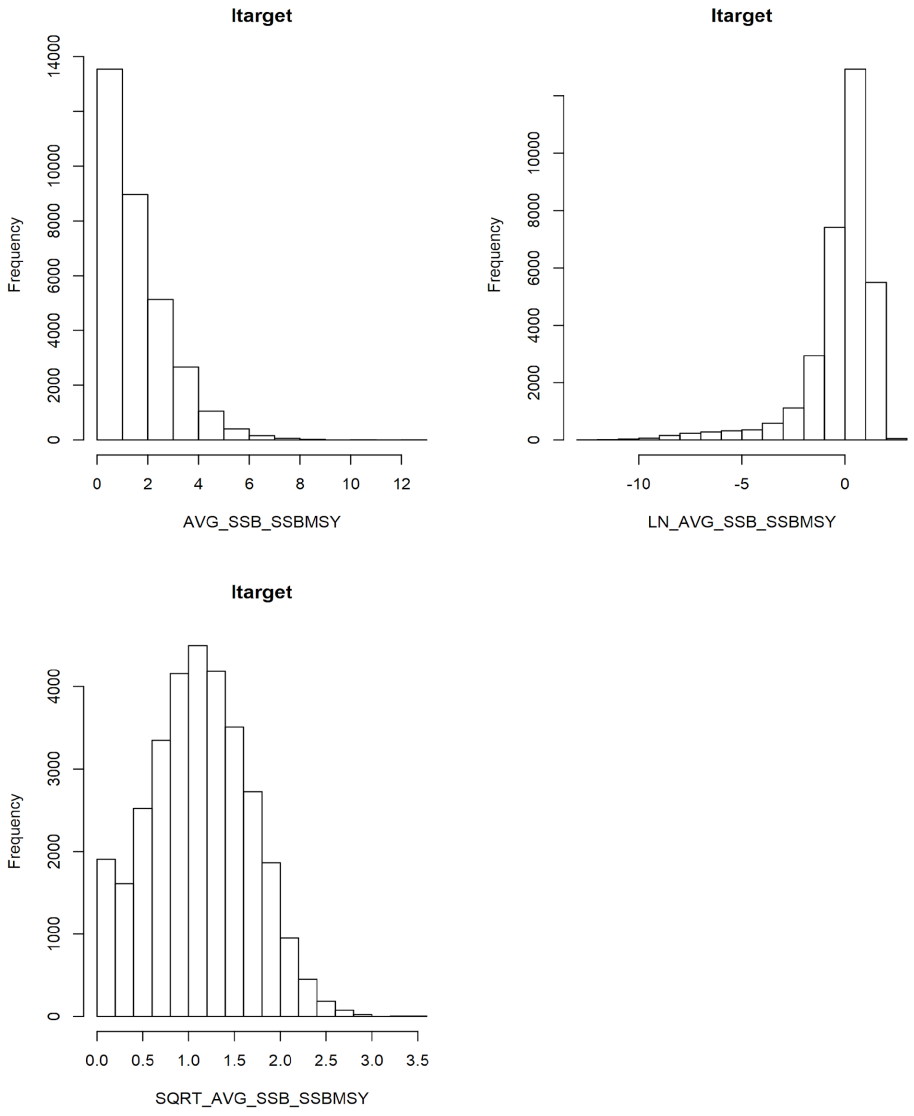


Figure A6.388. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

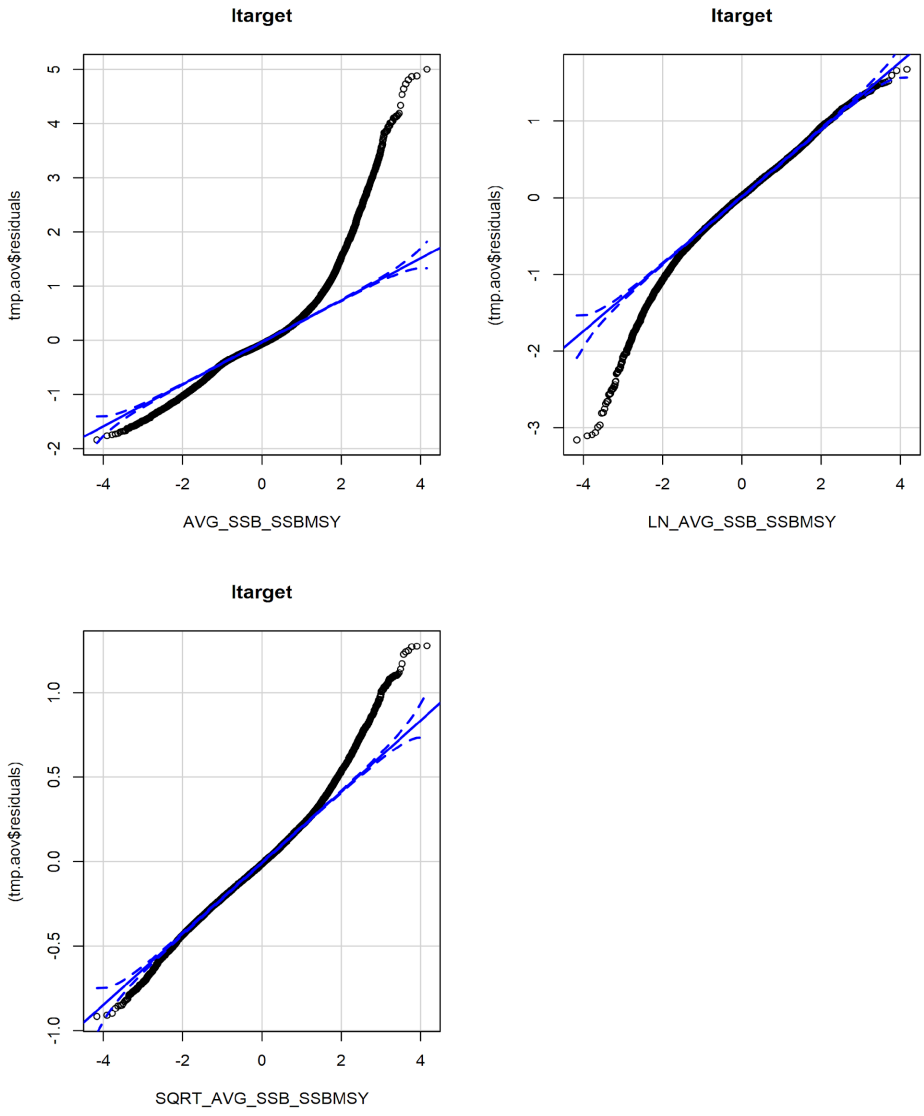


Figure A6.389. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

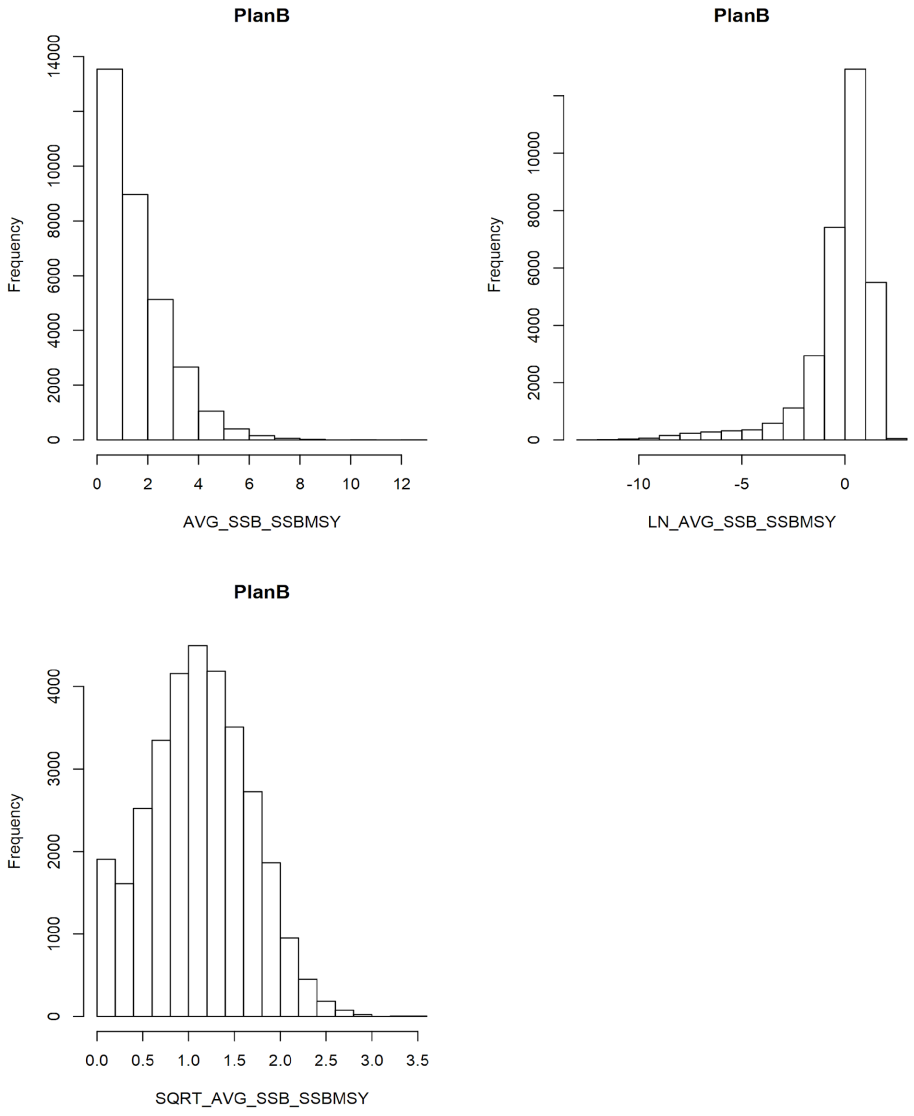


Figure A6.390. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

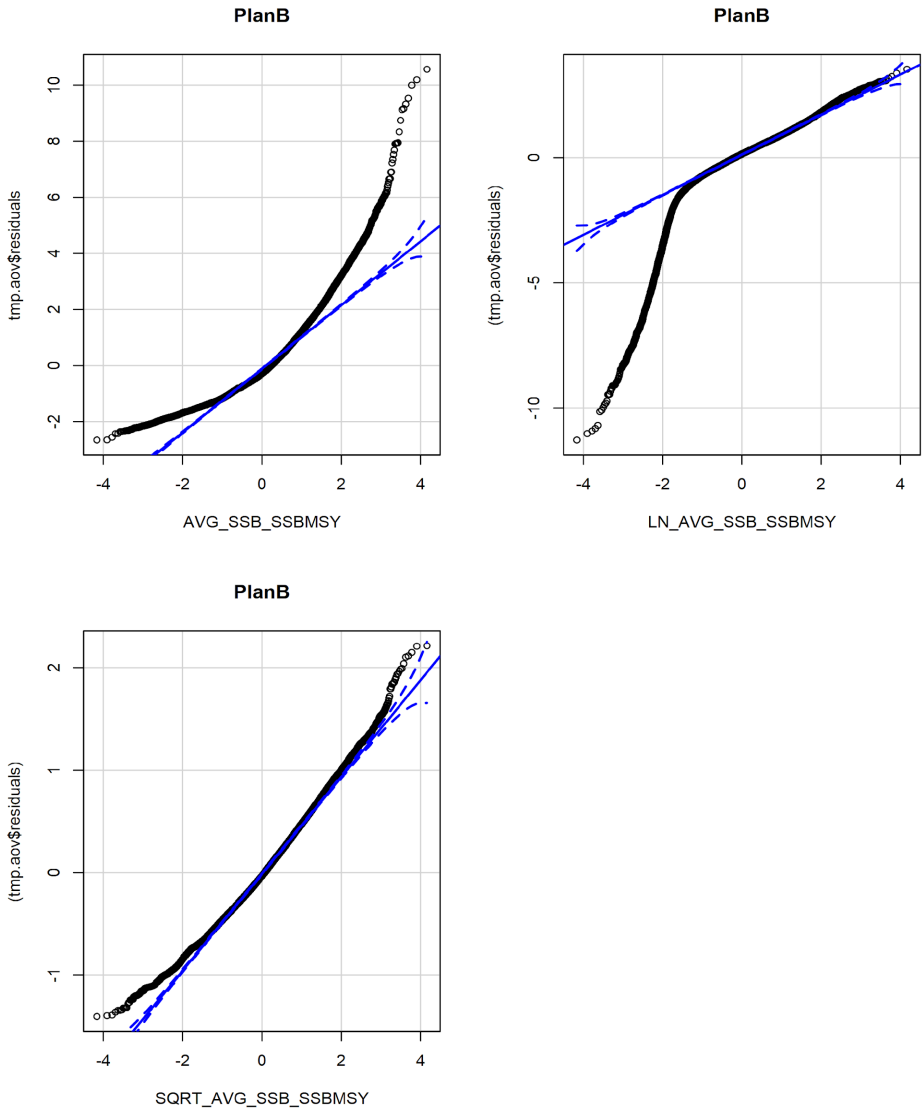


Figure A6.391. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

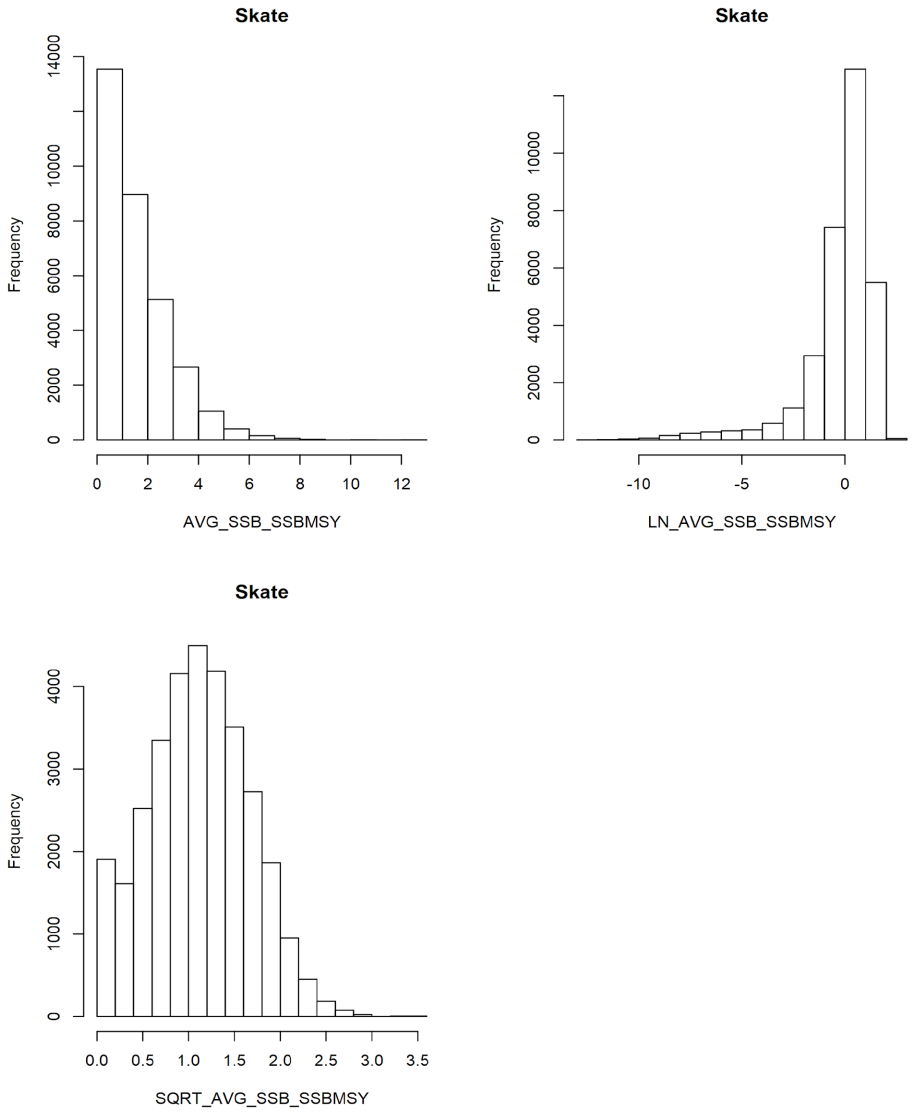


Figure A6.392. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

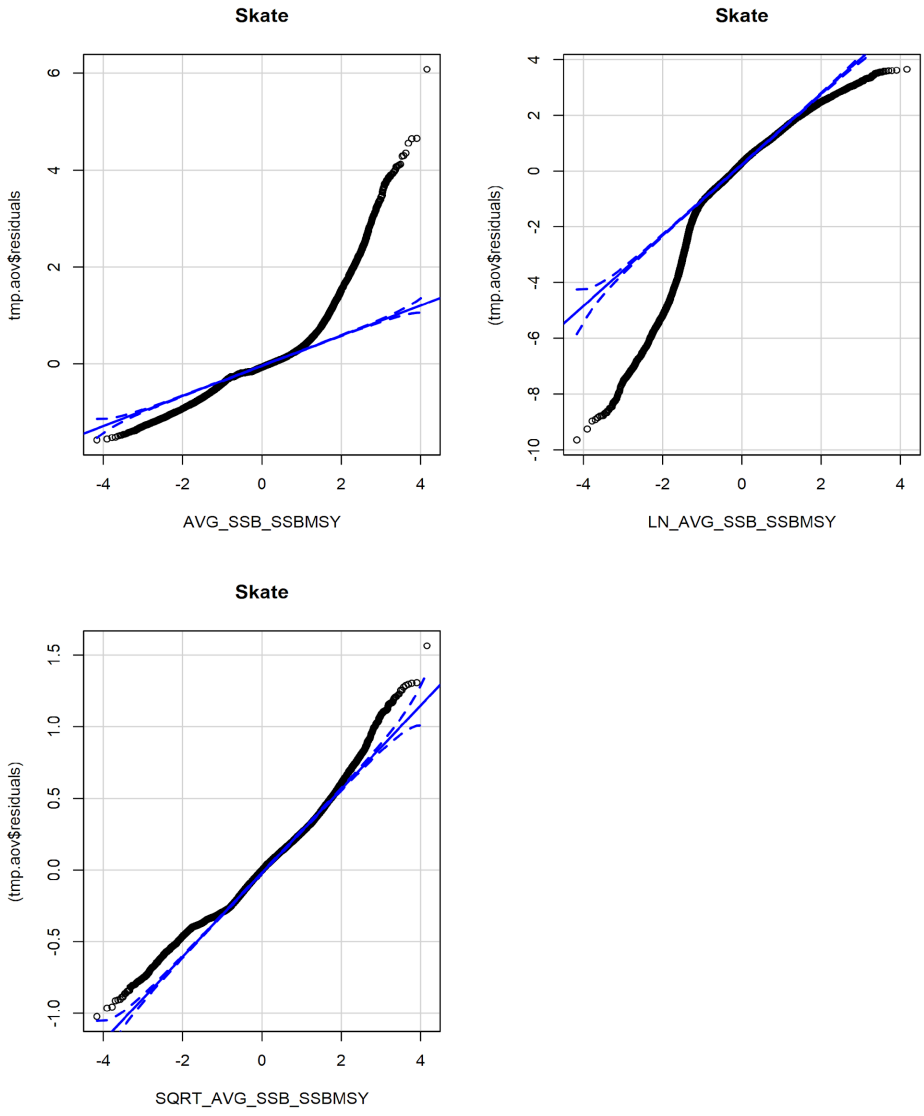


Figure A6.393. Spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

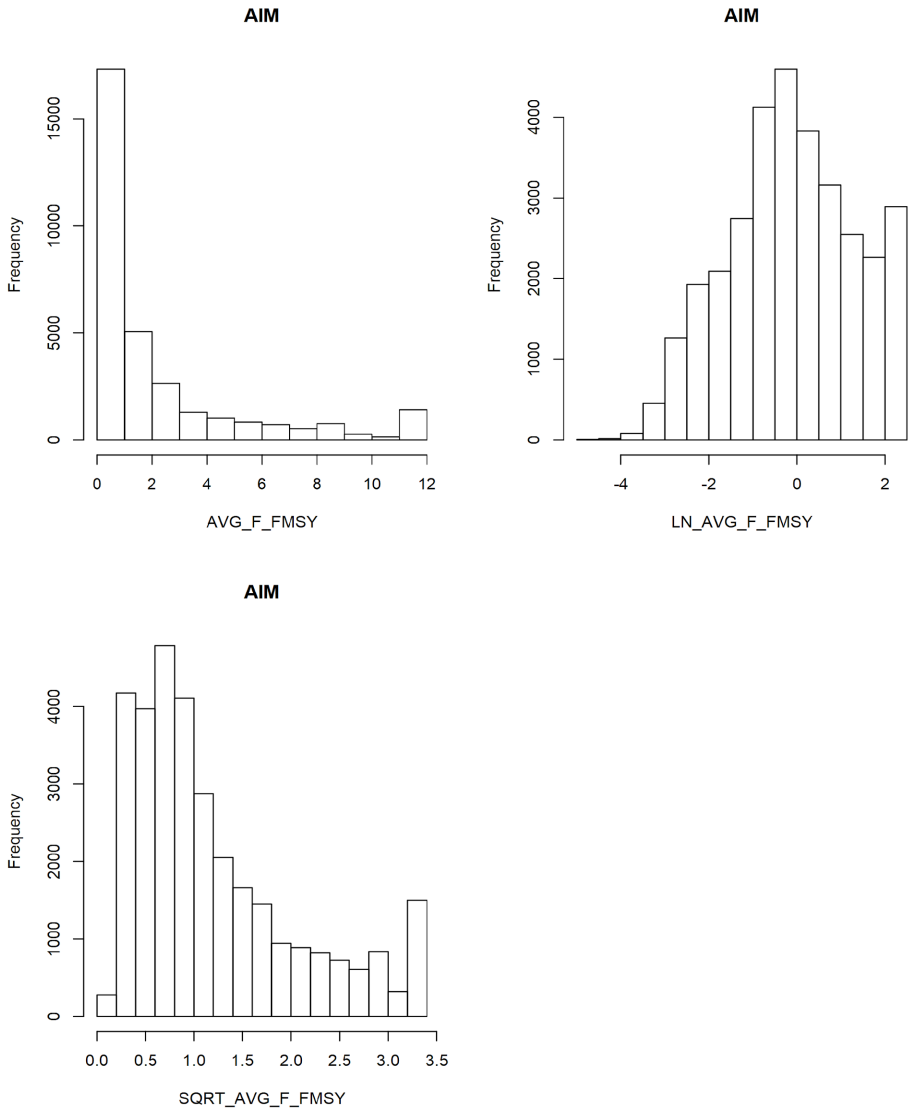


Figure A6.394. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

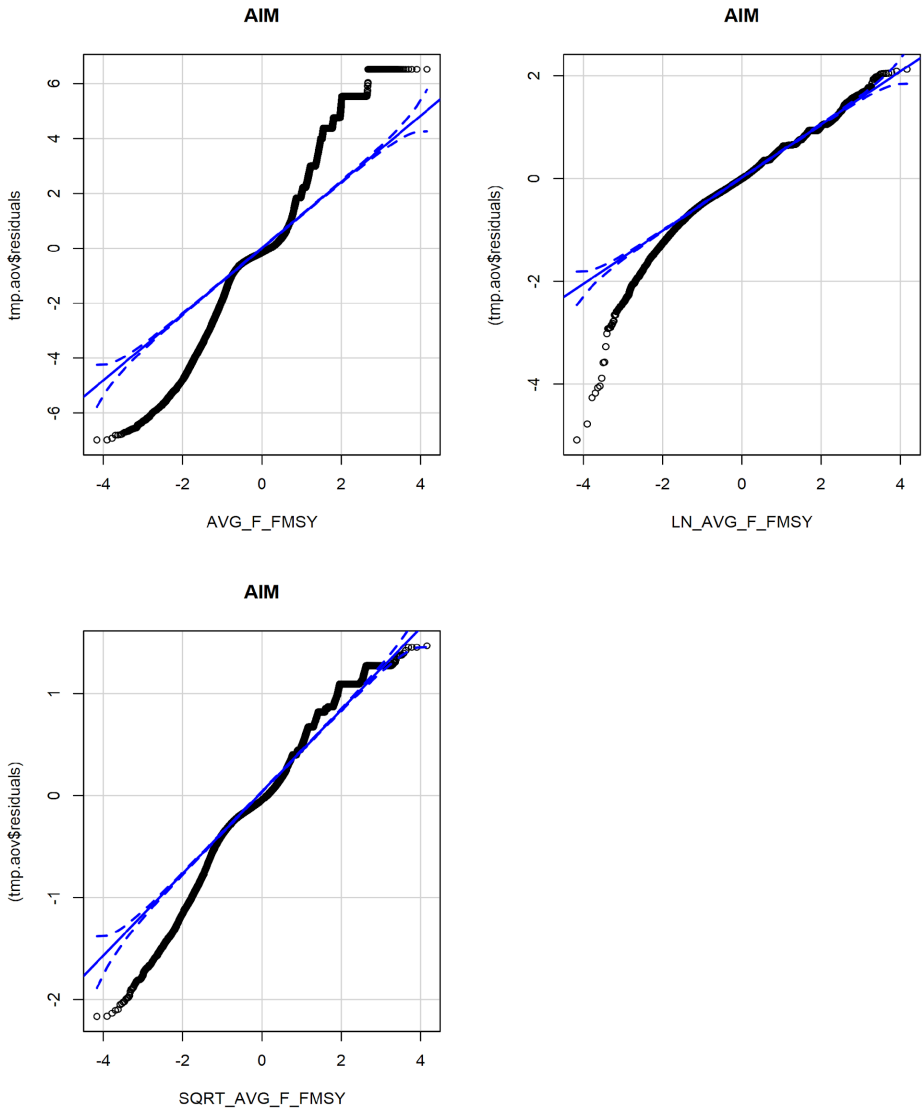


Figure A6.395. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

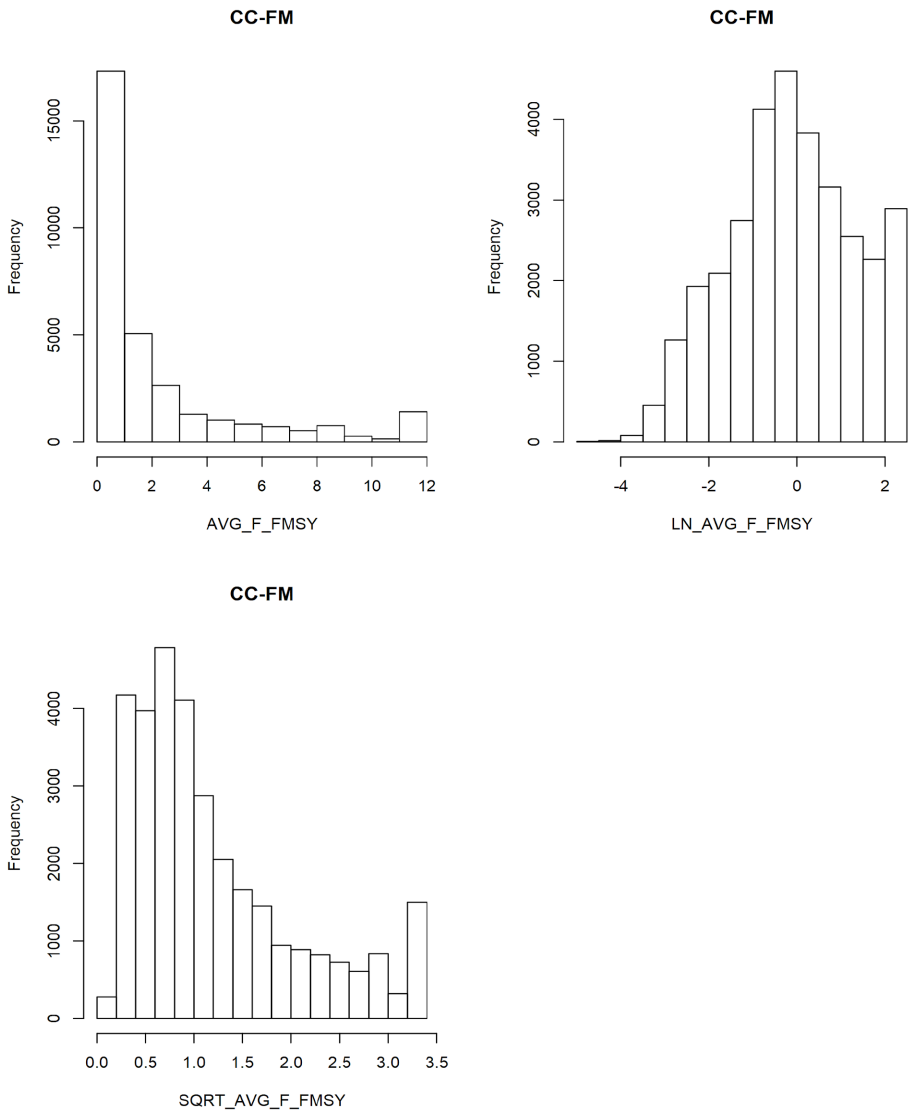


Figure A6.396. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

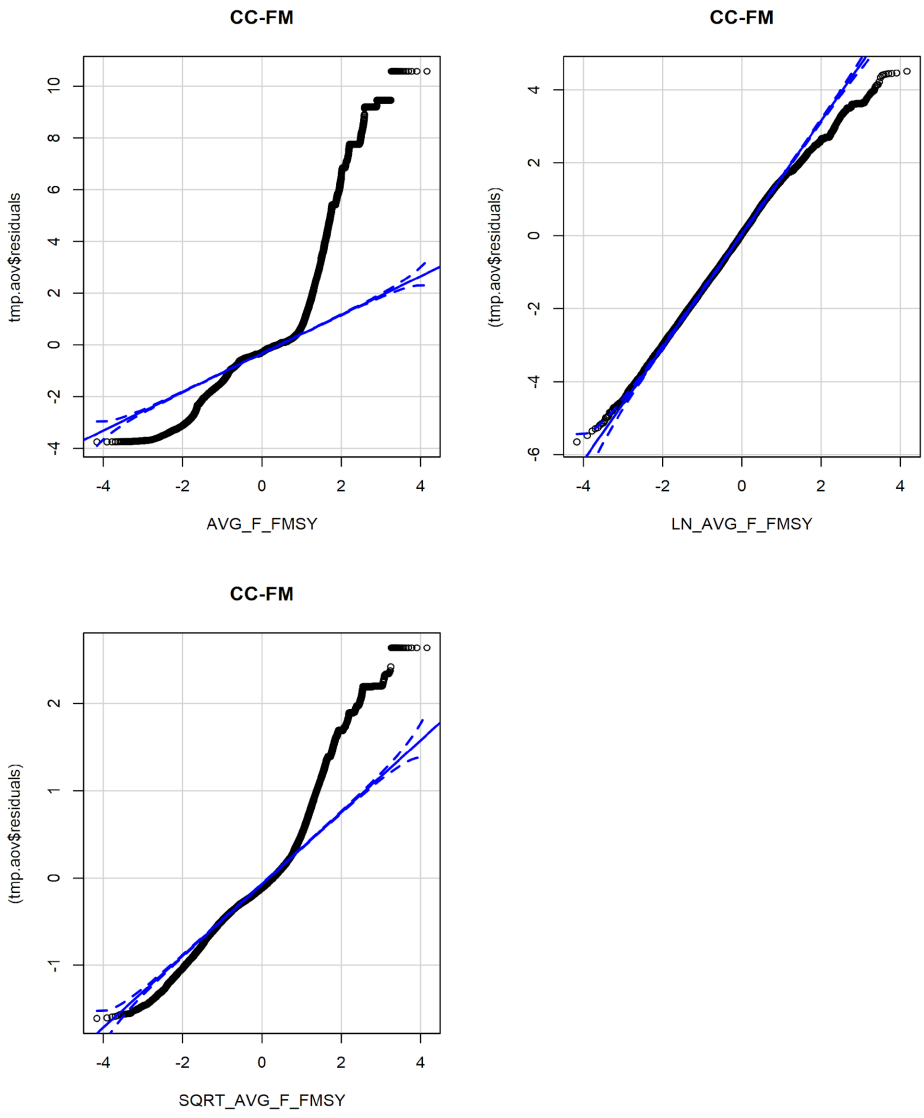


Figure A6.397. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

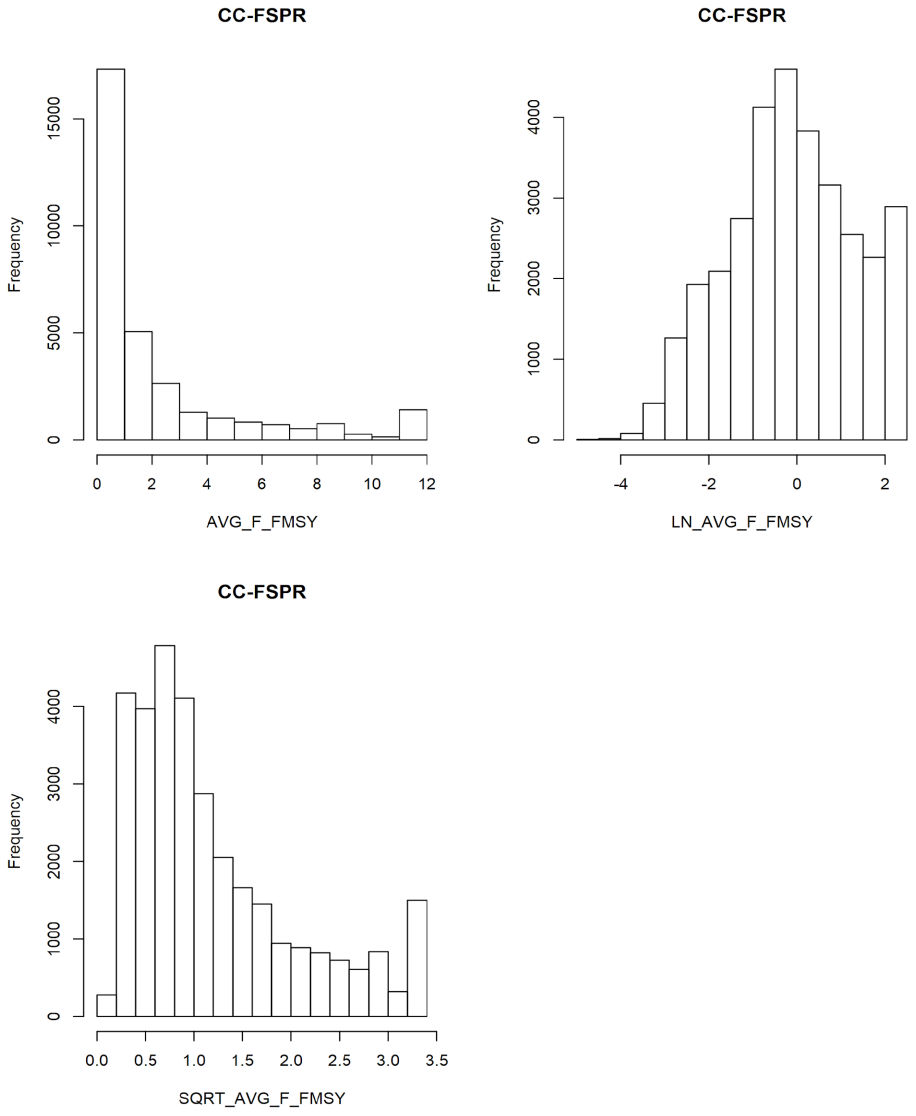


Figure A6.398. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

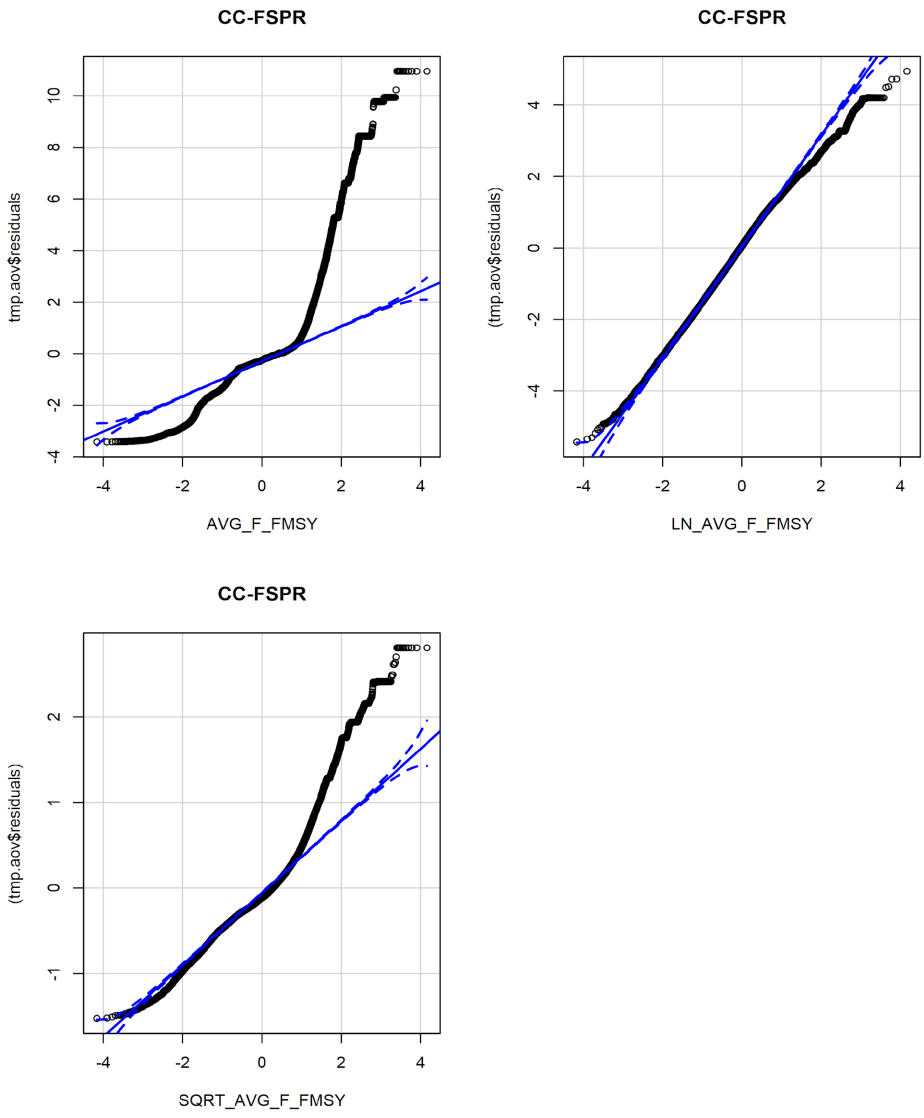


Figure A6.399. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

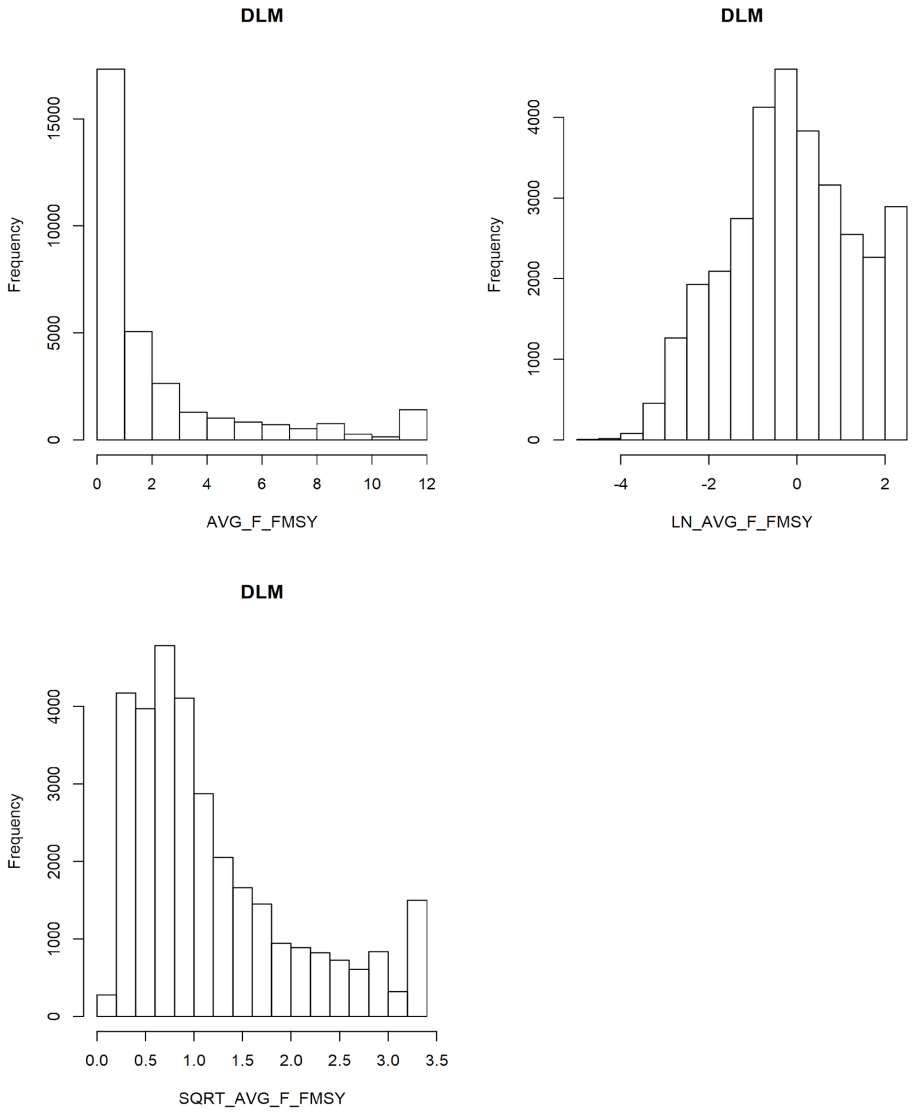


Figure A6.400. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

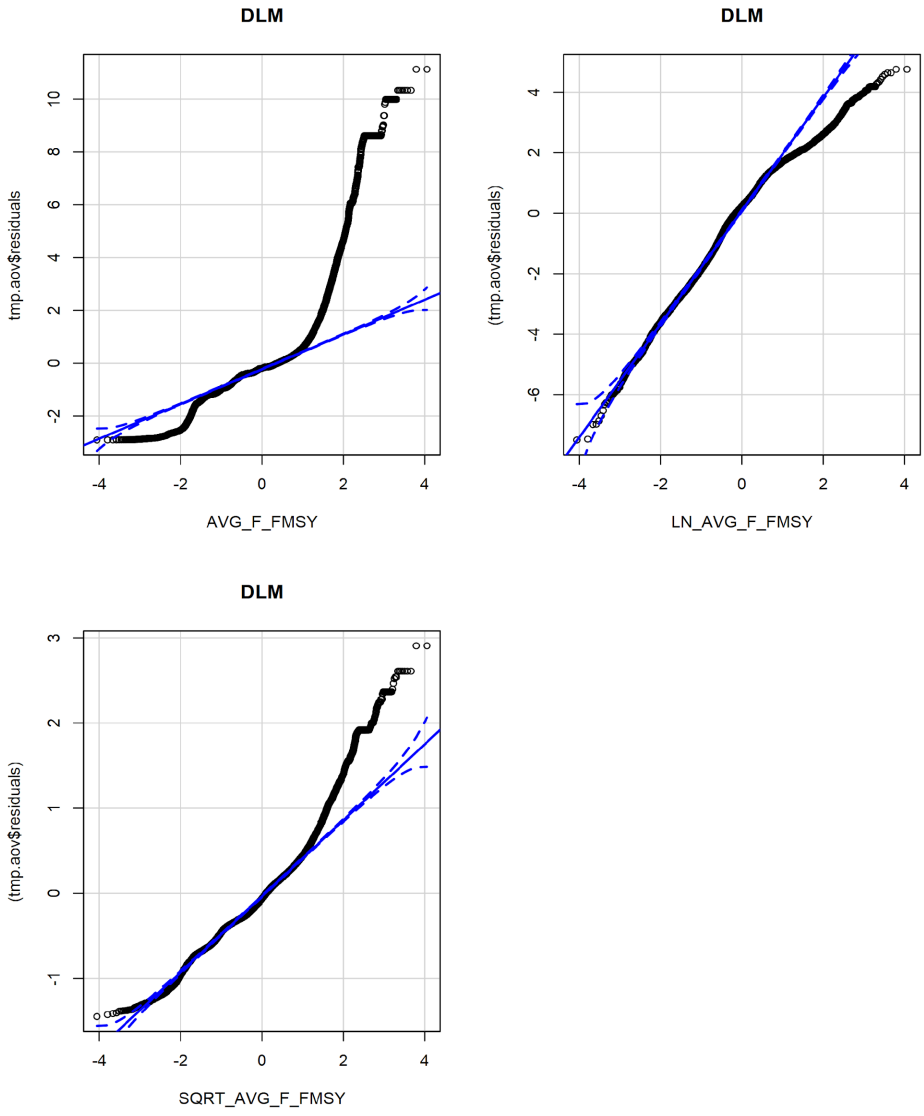


Figure A6.401. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

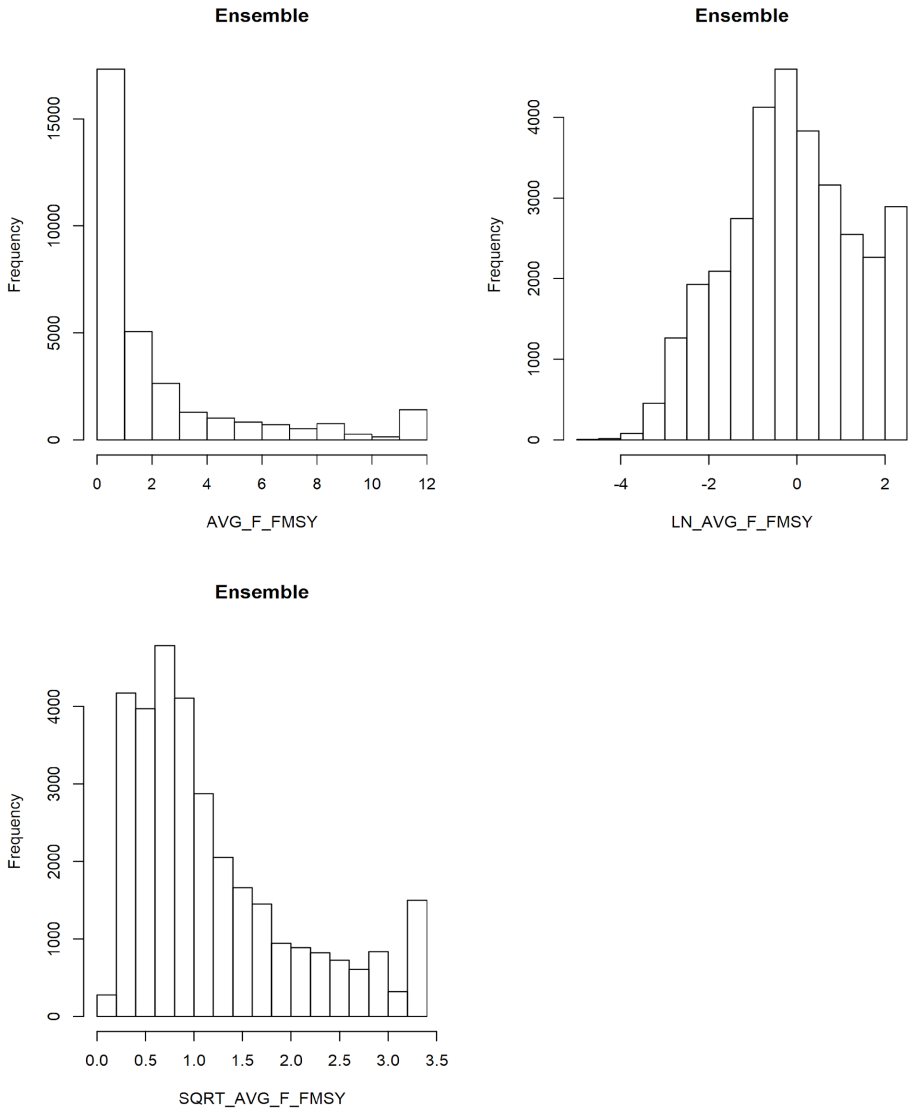


Figure A6.402. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

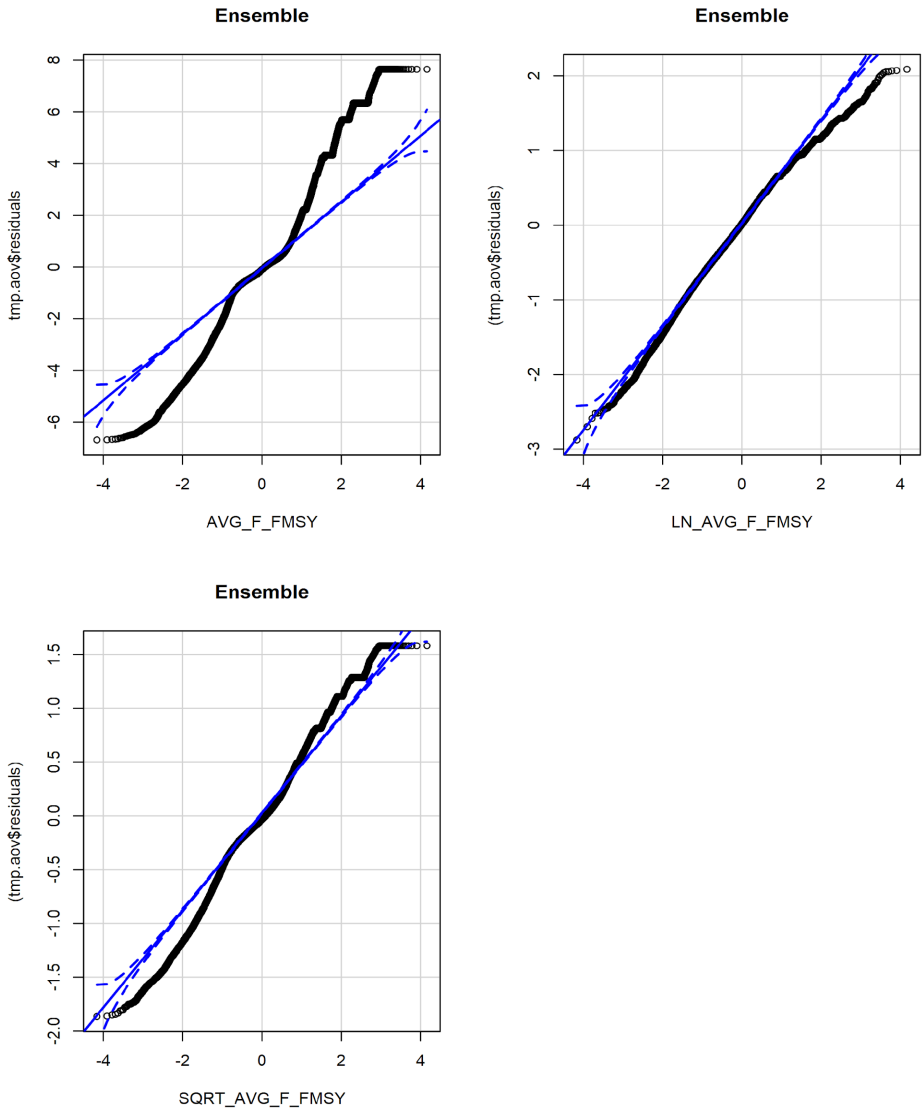


Figure A6.403. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

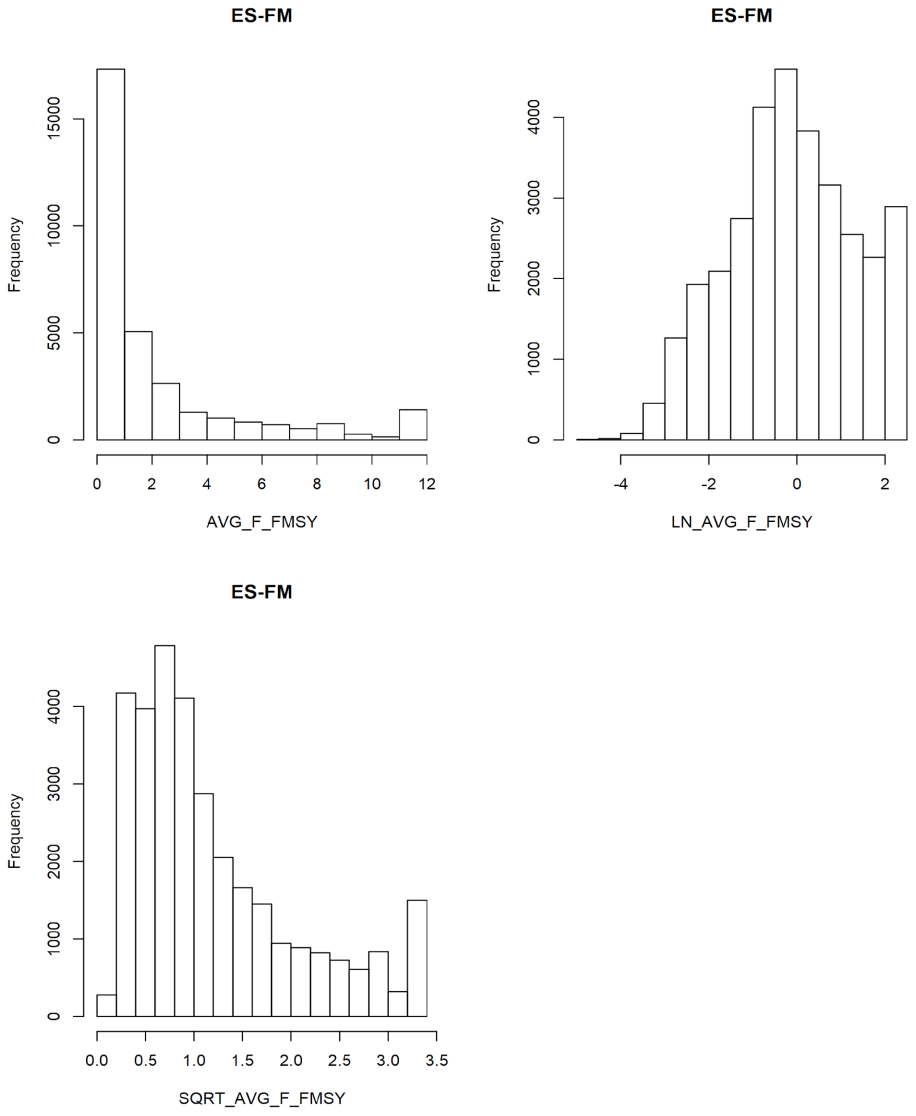


Figure A6.404. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

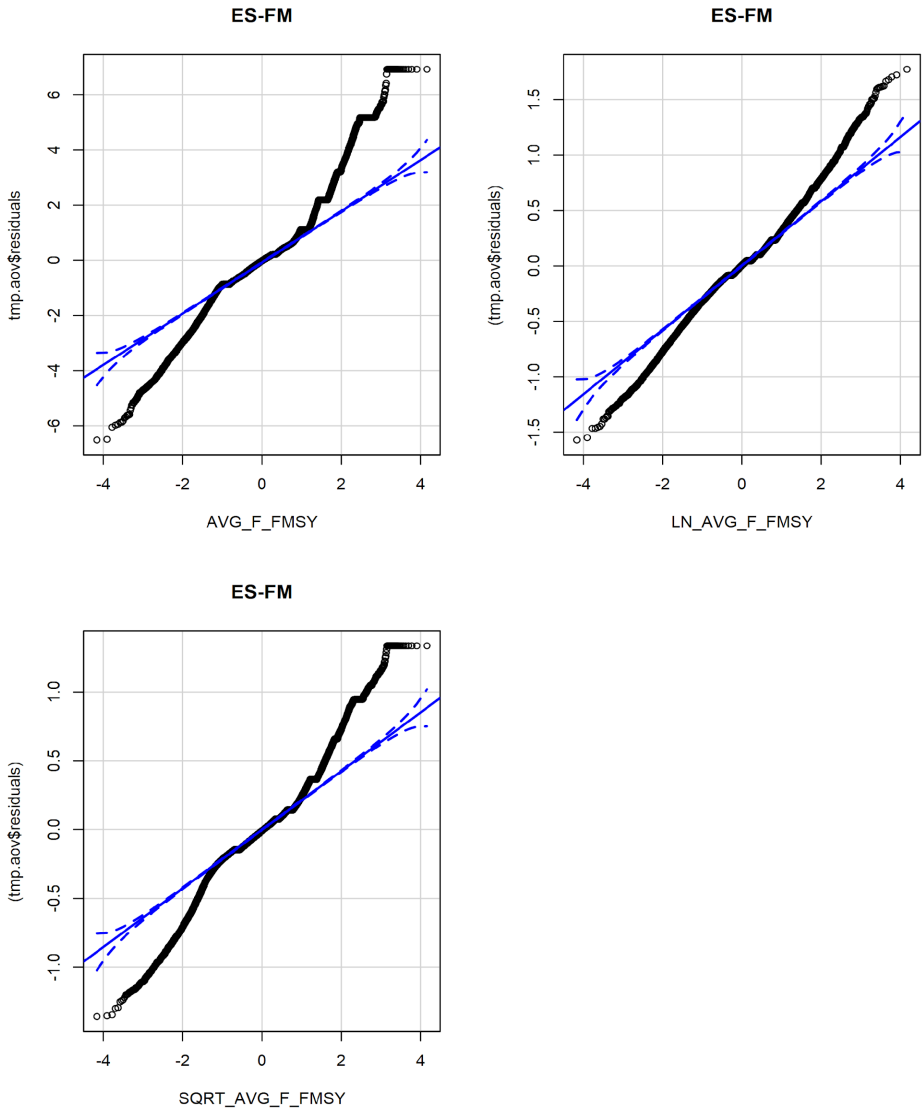


Figure A6.405. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

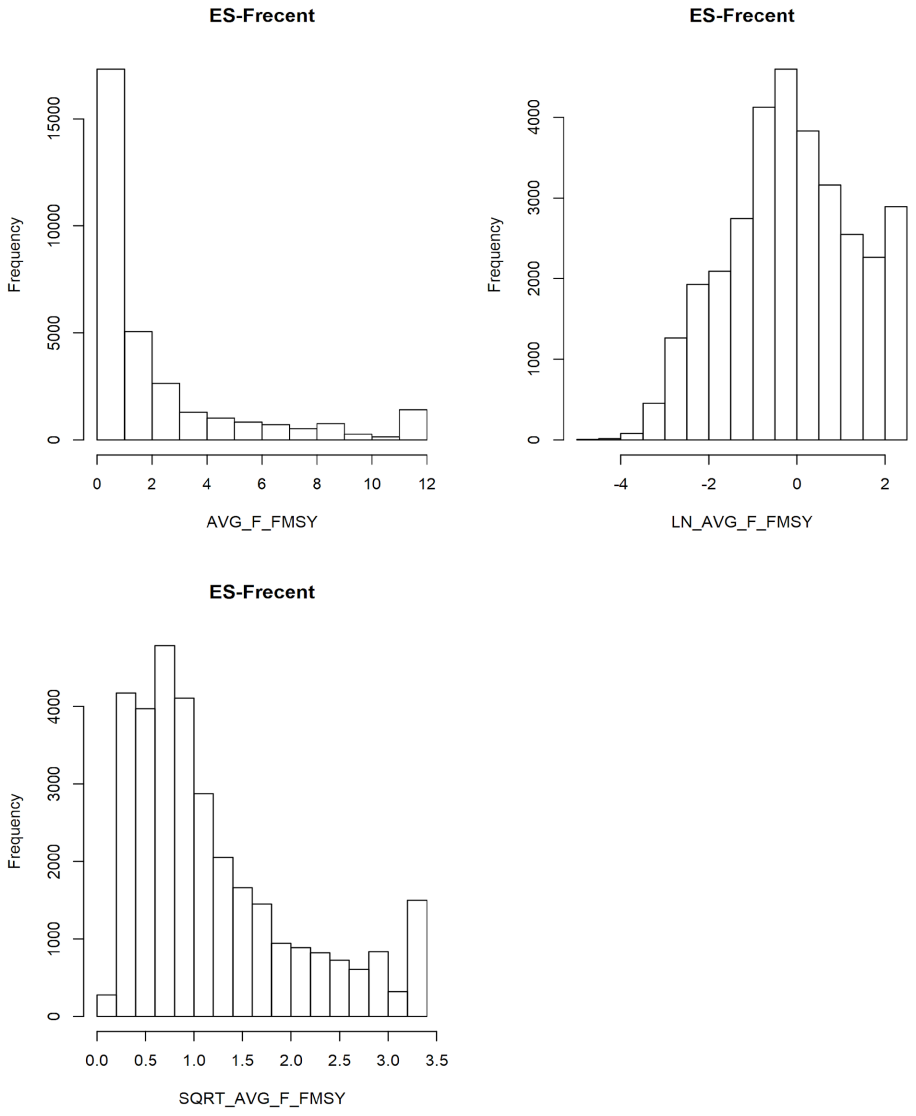


Figure A6.406. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

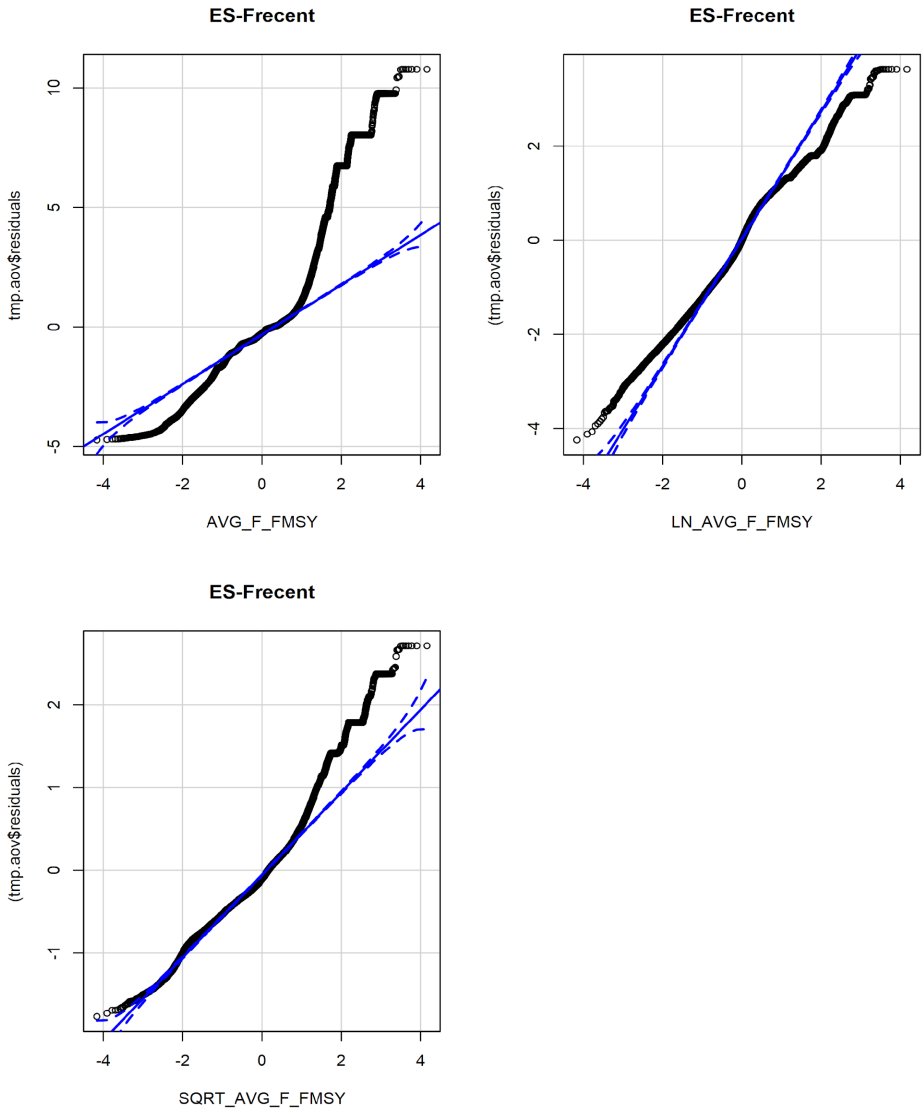


Figure A6.407. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

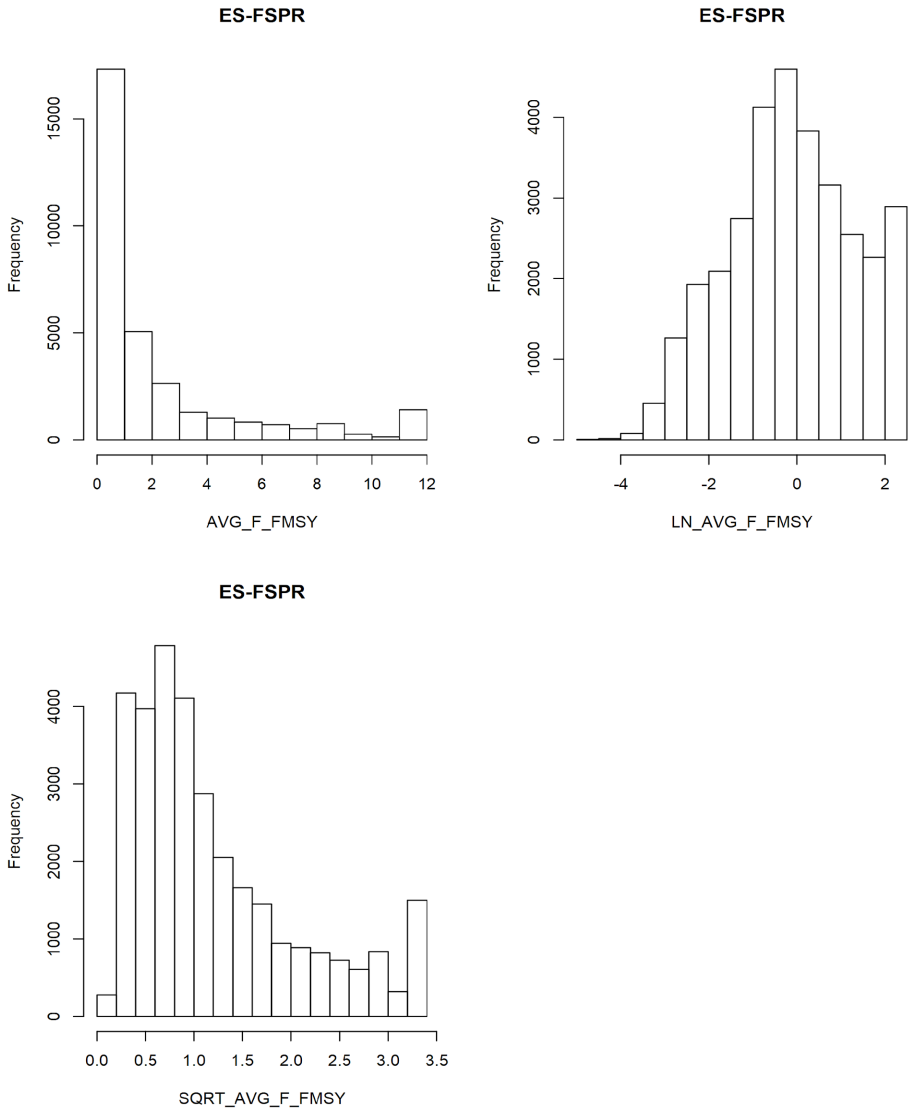


Figure A6.408. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

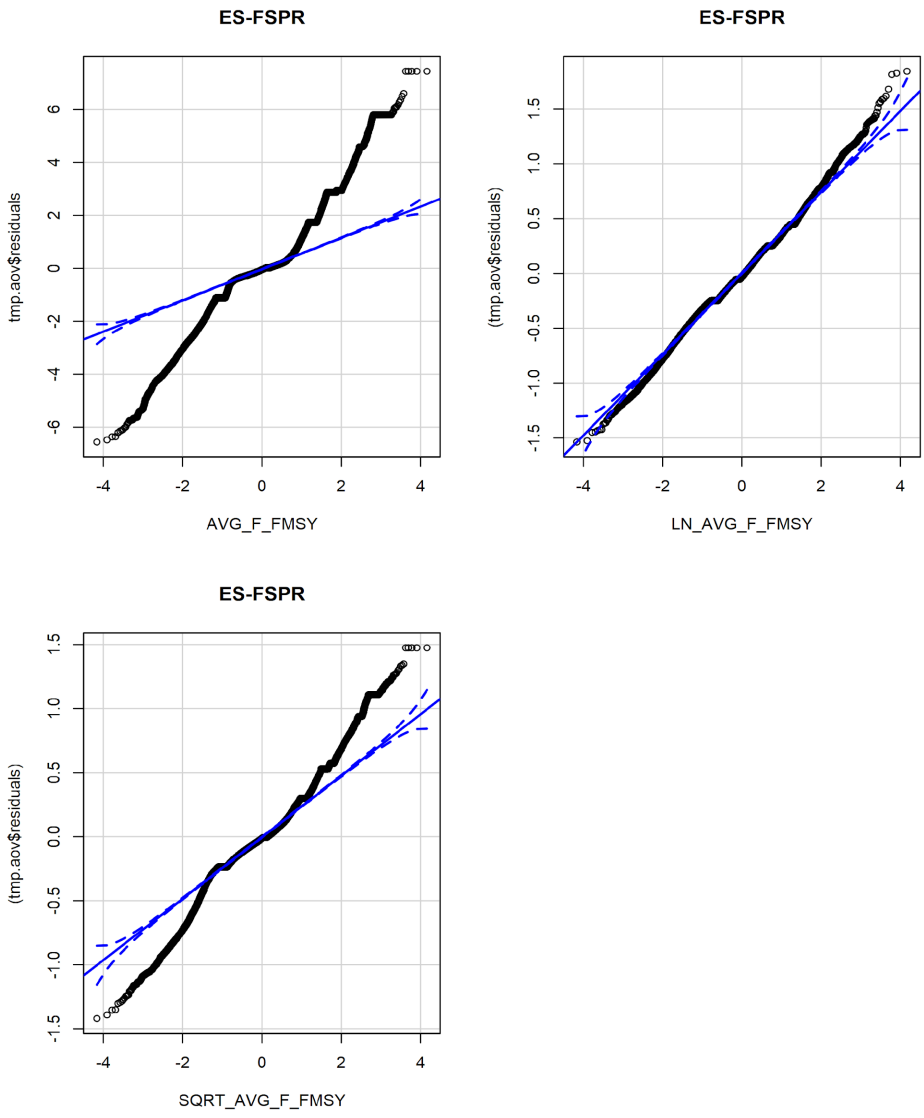


Figure A6.409. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

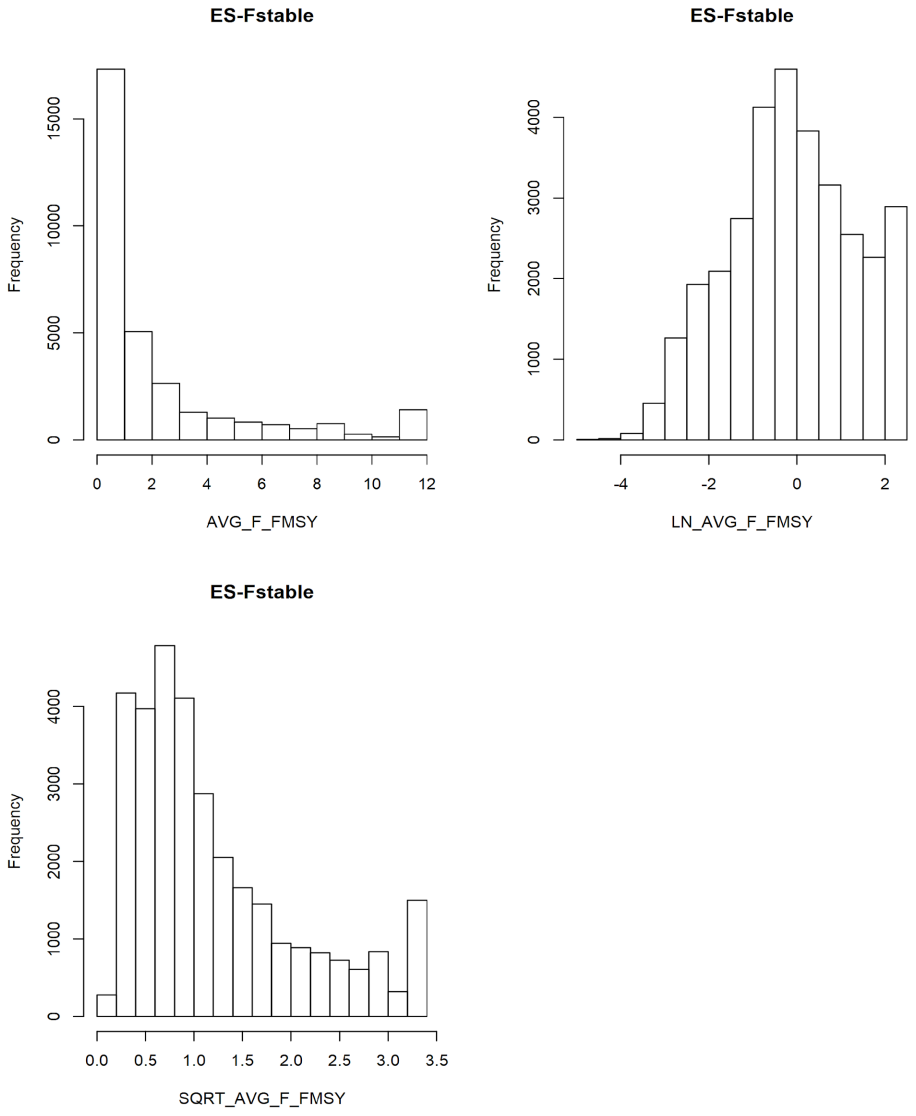


Figure A6.410. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

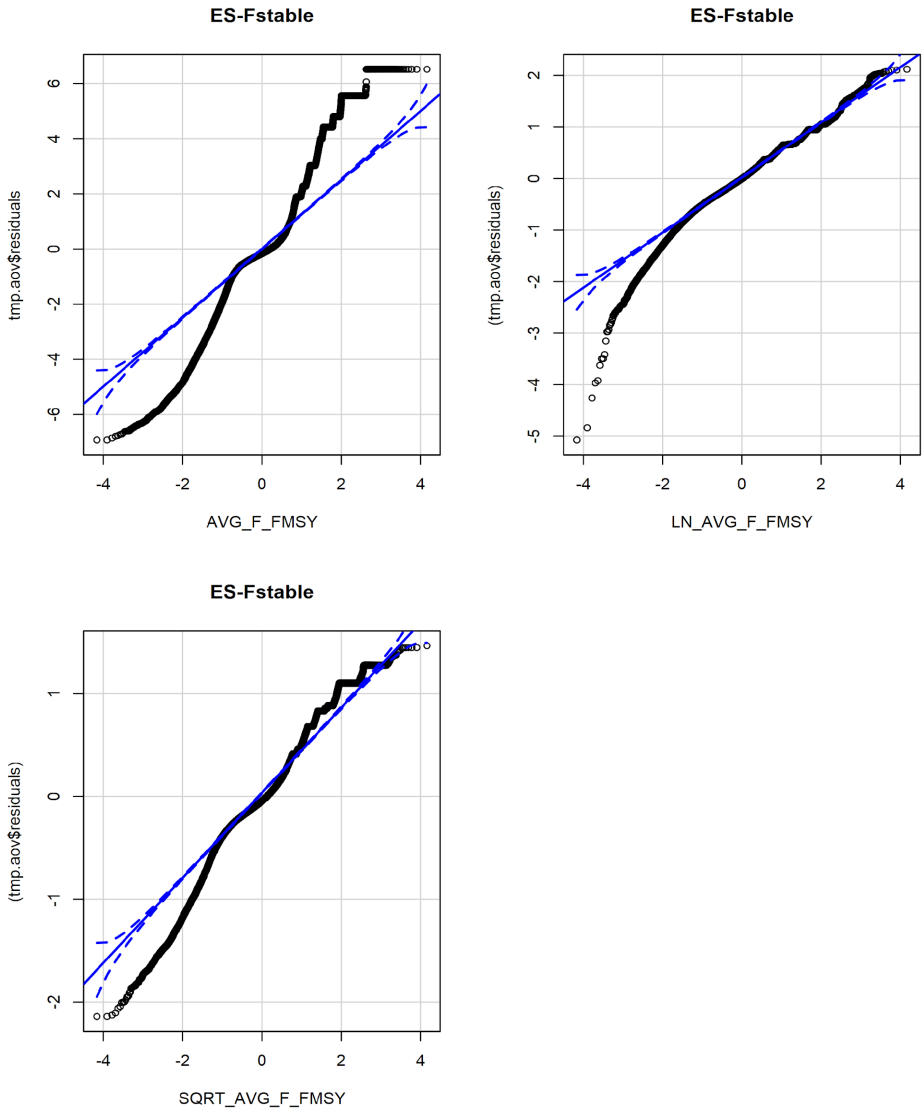


Figure A6.411. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

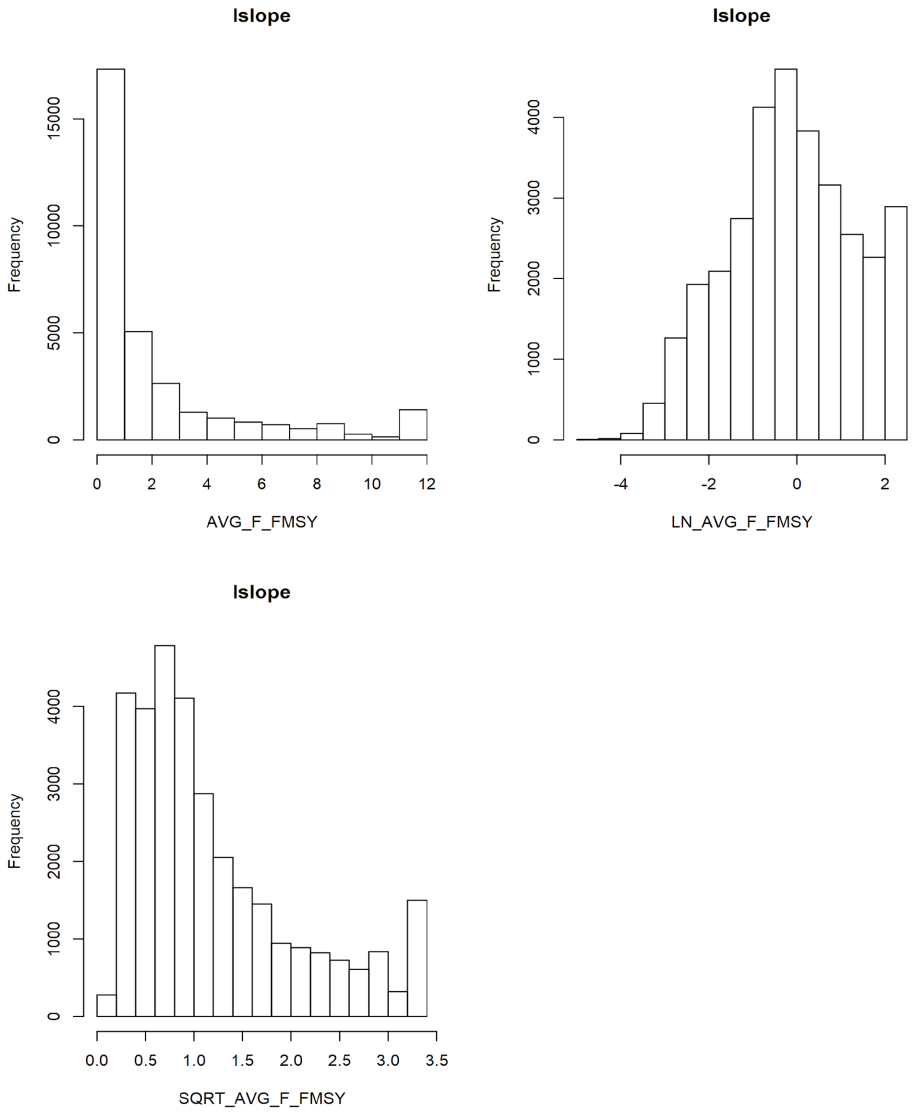


Figure A6.412. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

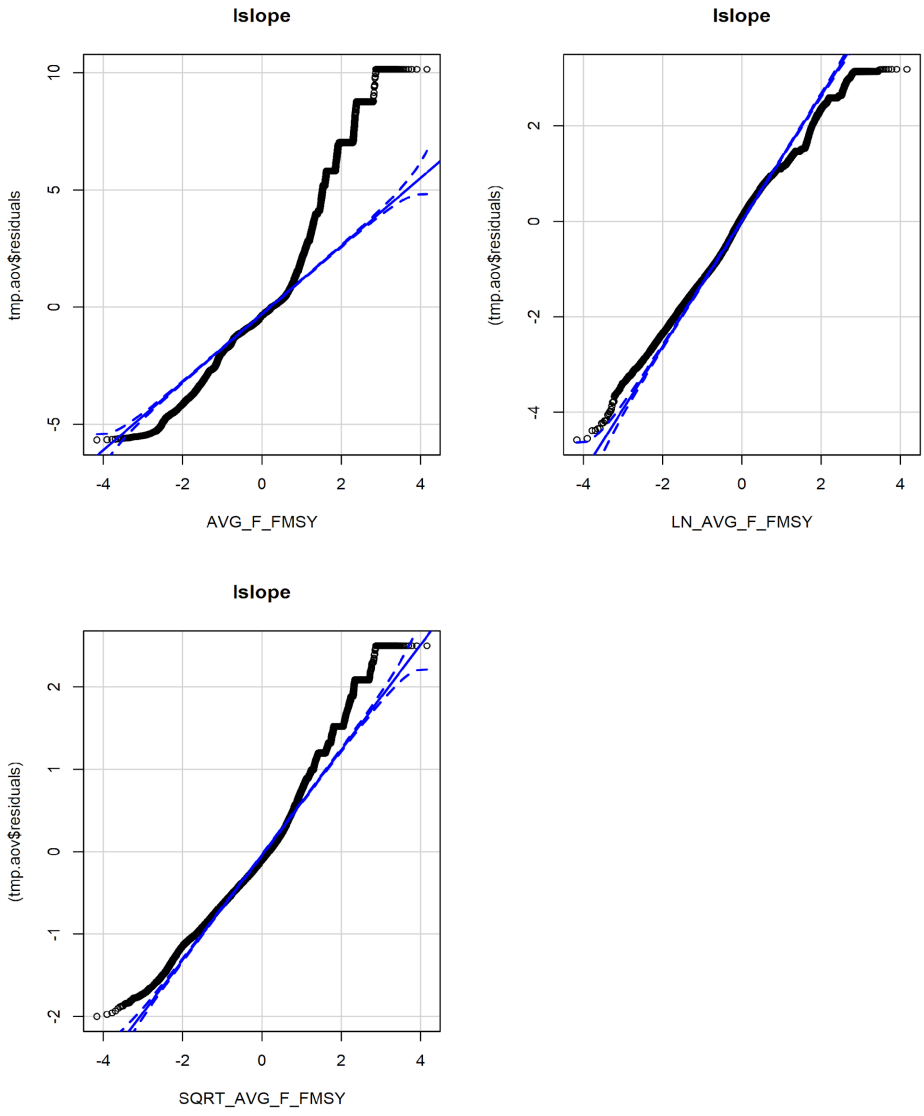


Figure A6.413. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

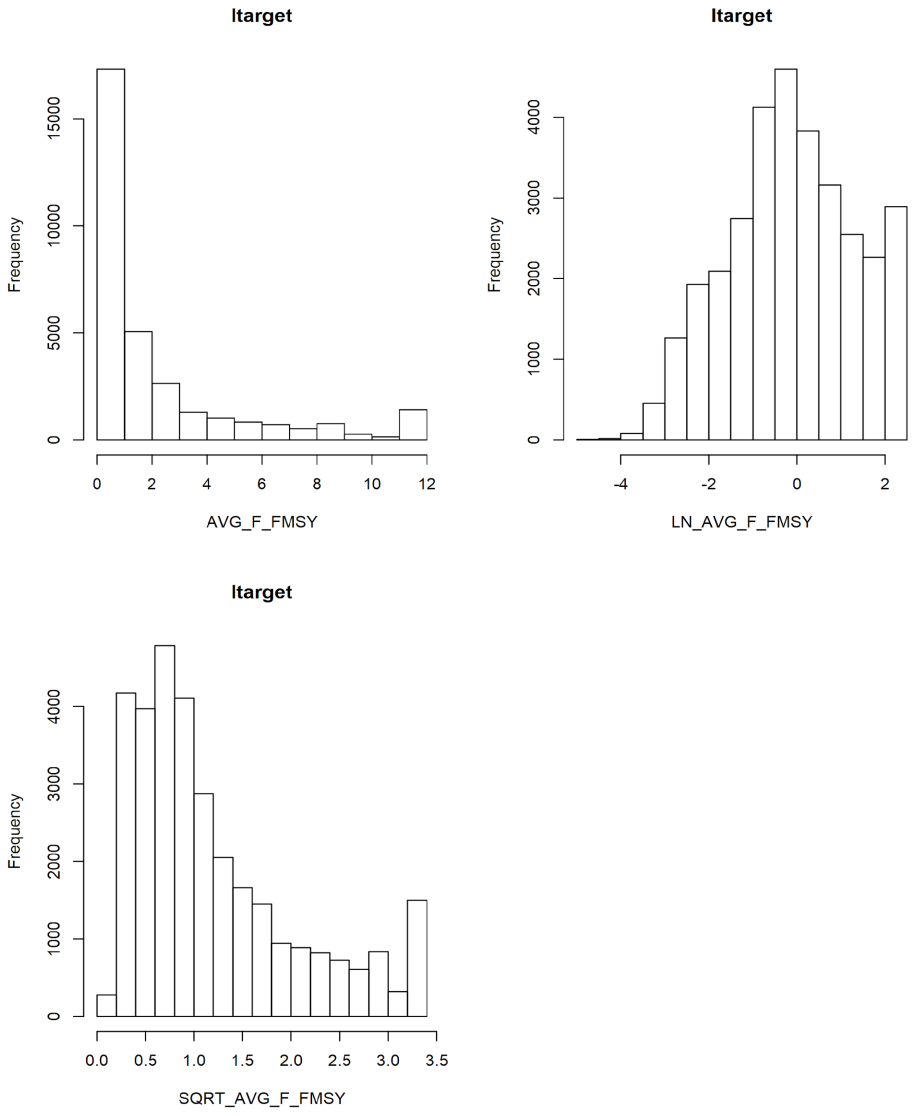


Figure A6.414. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

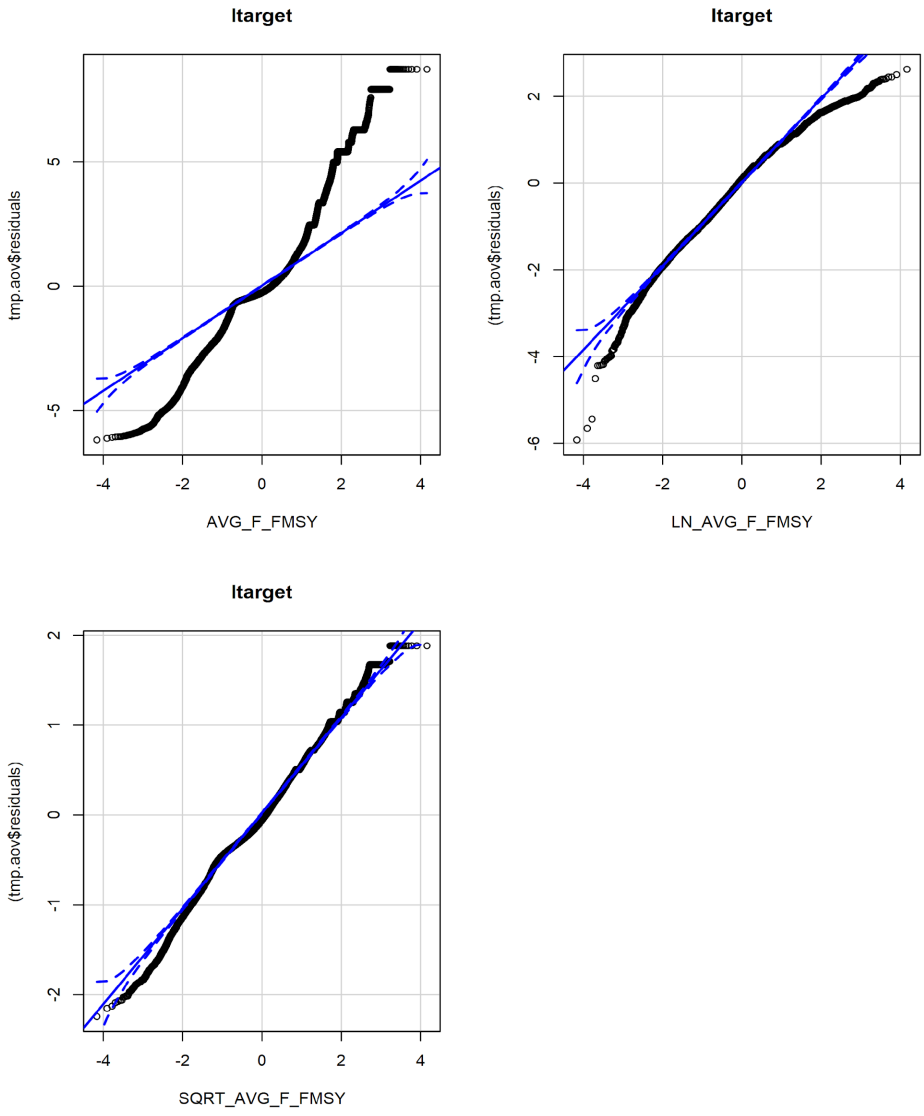


Figure A6.415. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

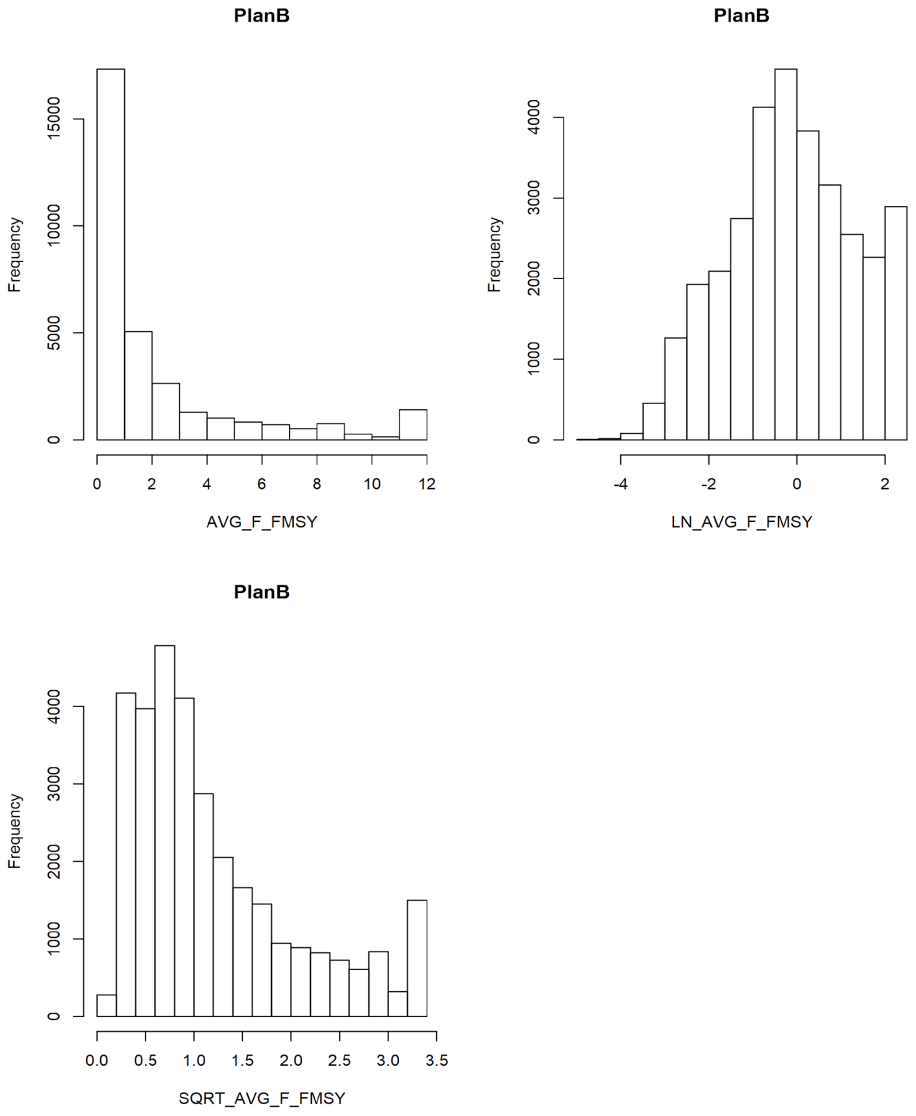


Figure A6.416. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

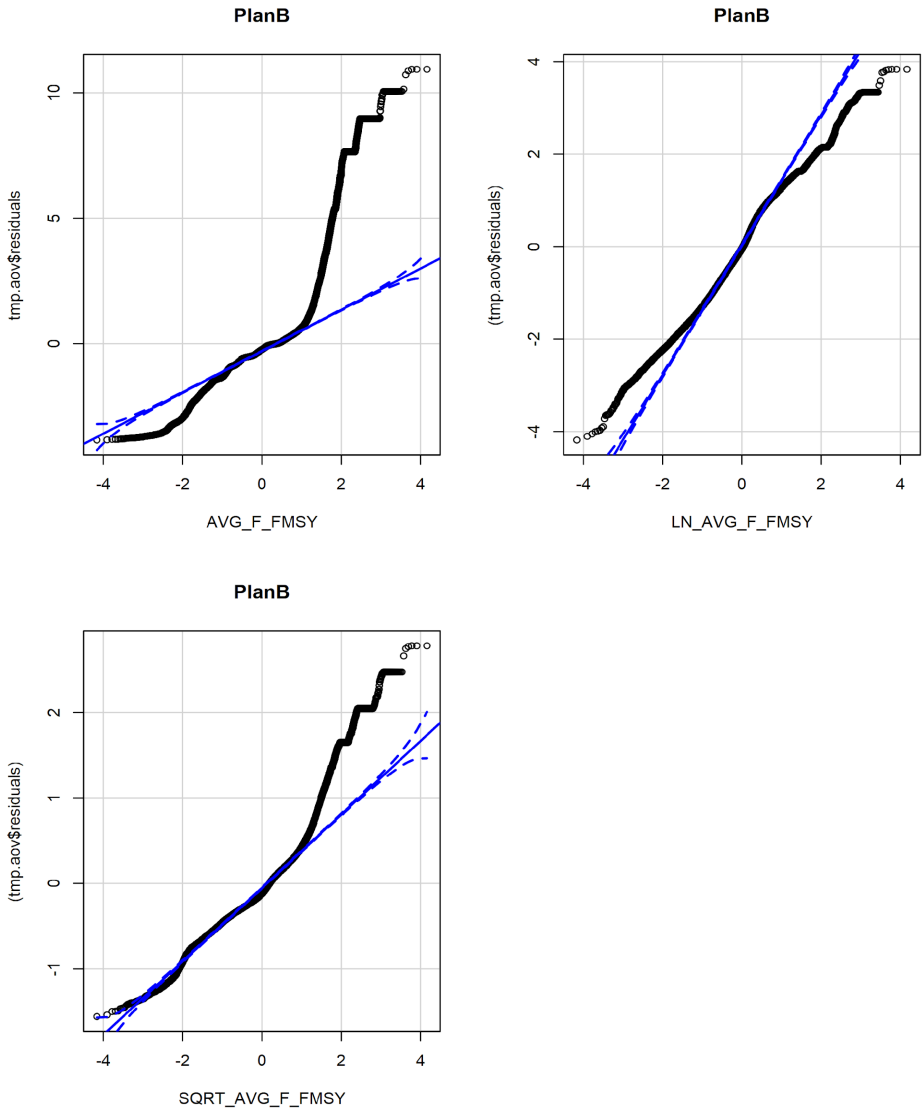


Figure A6.417. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

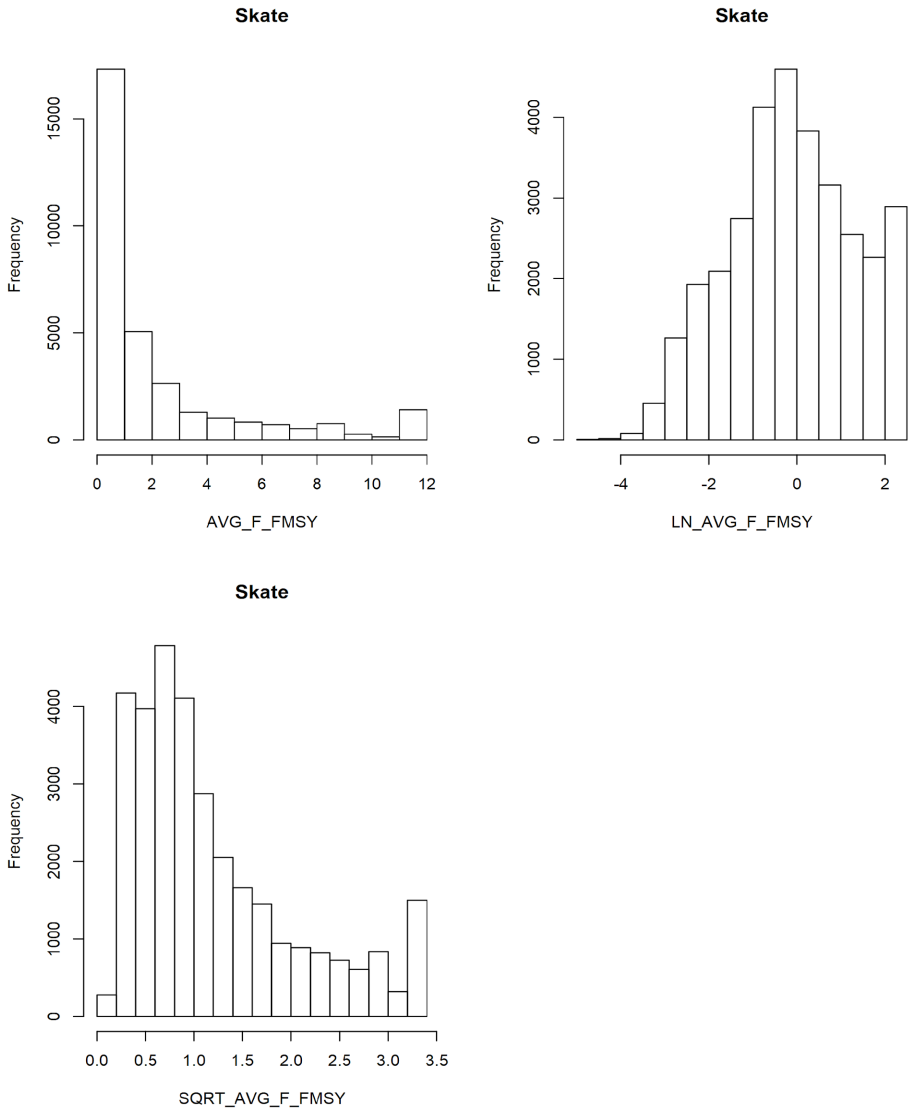


Figure A6.418. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

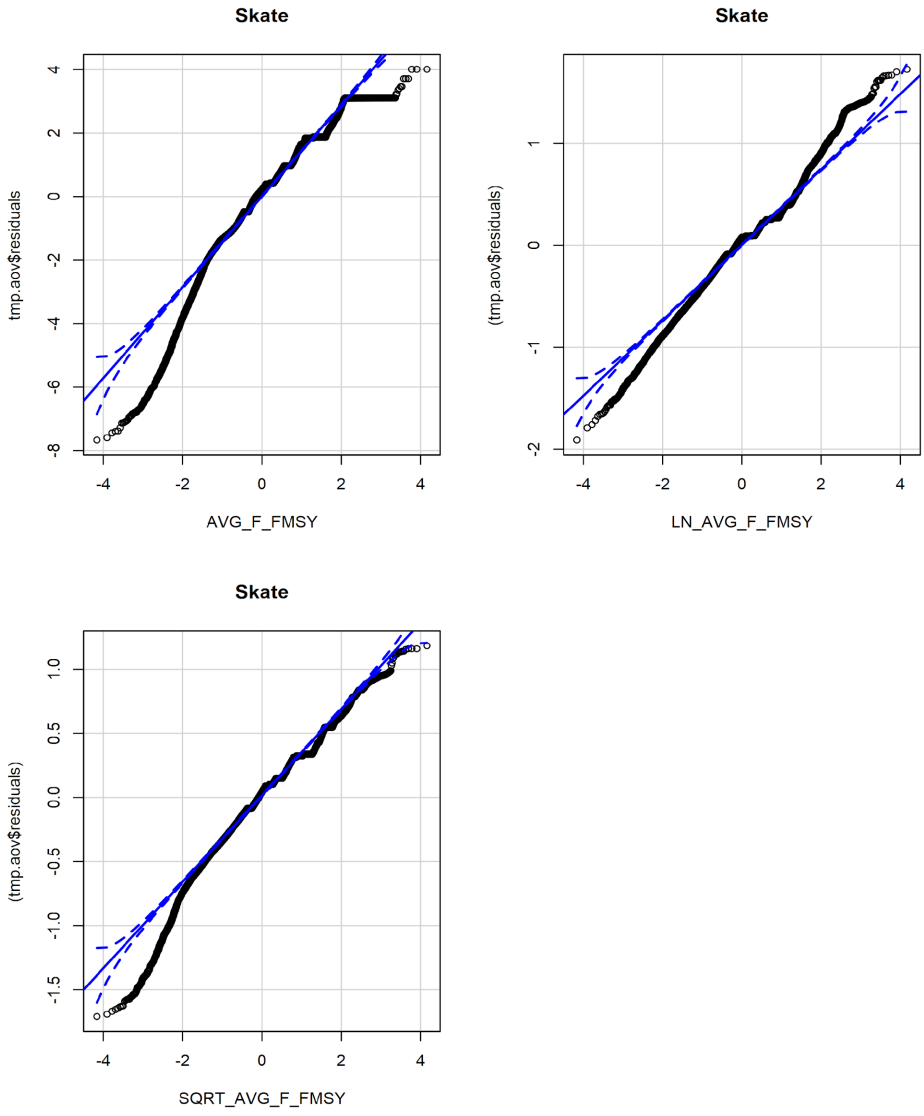


Figure A6.419. Fishing mortality rate (F)/maximum sustainable yield (F_{MSY}) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

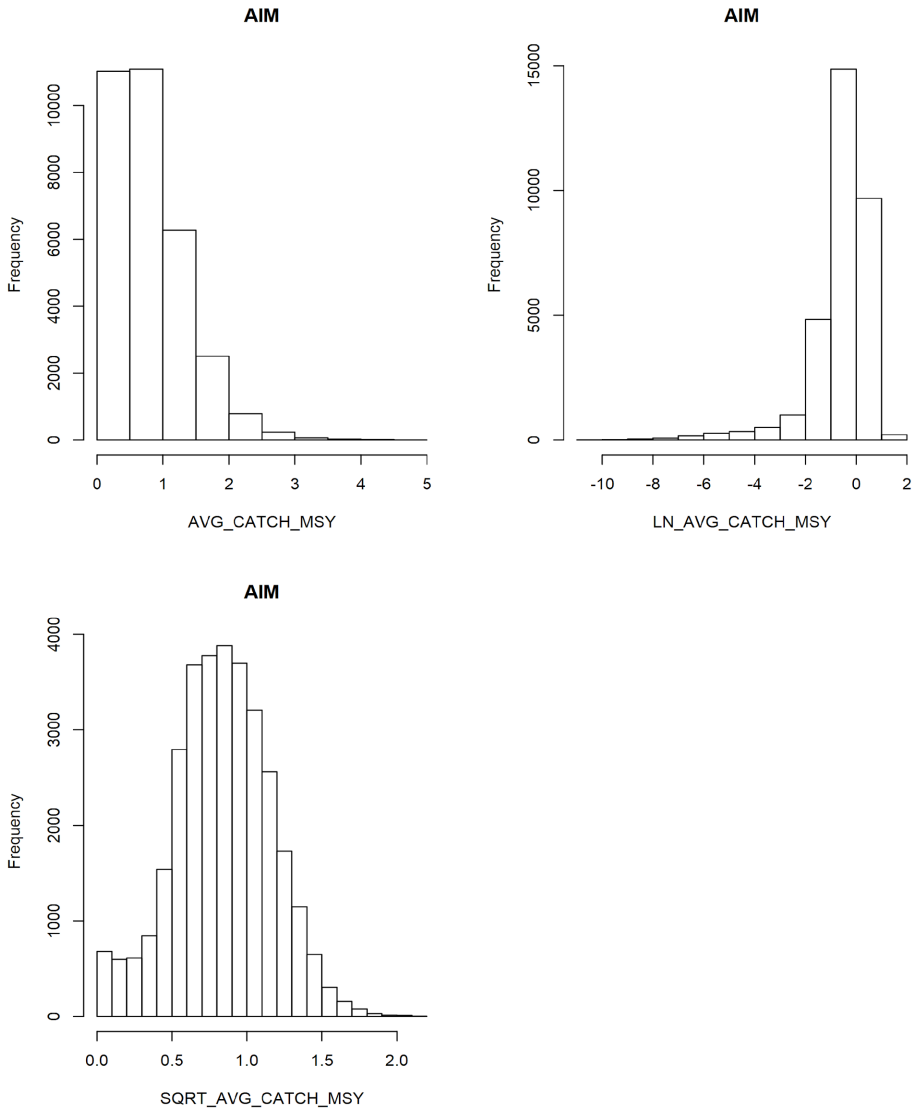


Figure A6.420. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

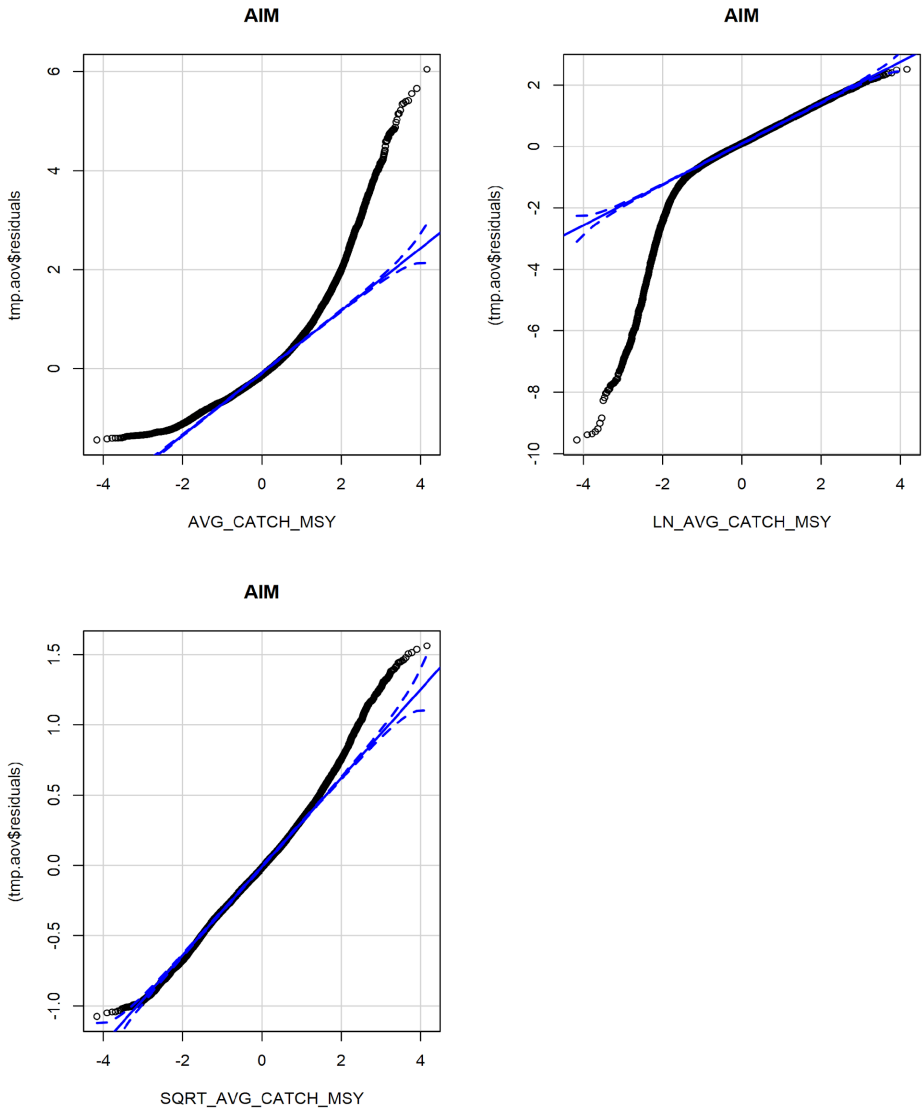


Figure A6.421. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

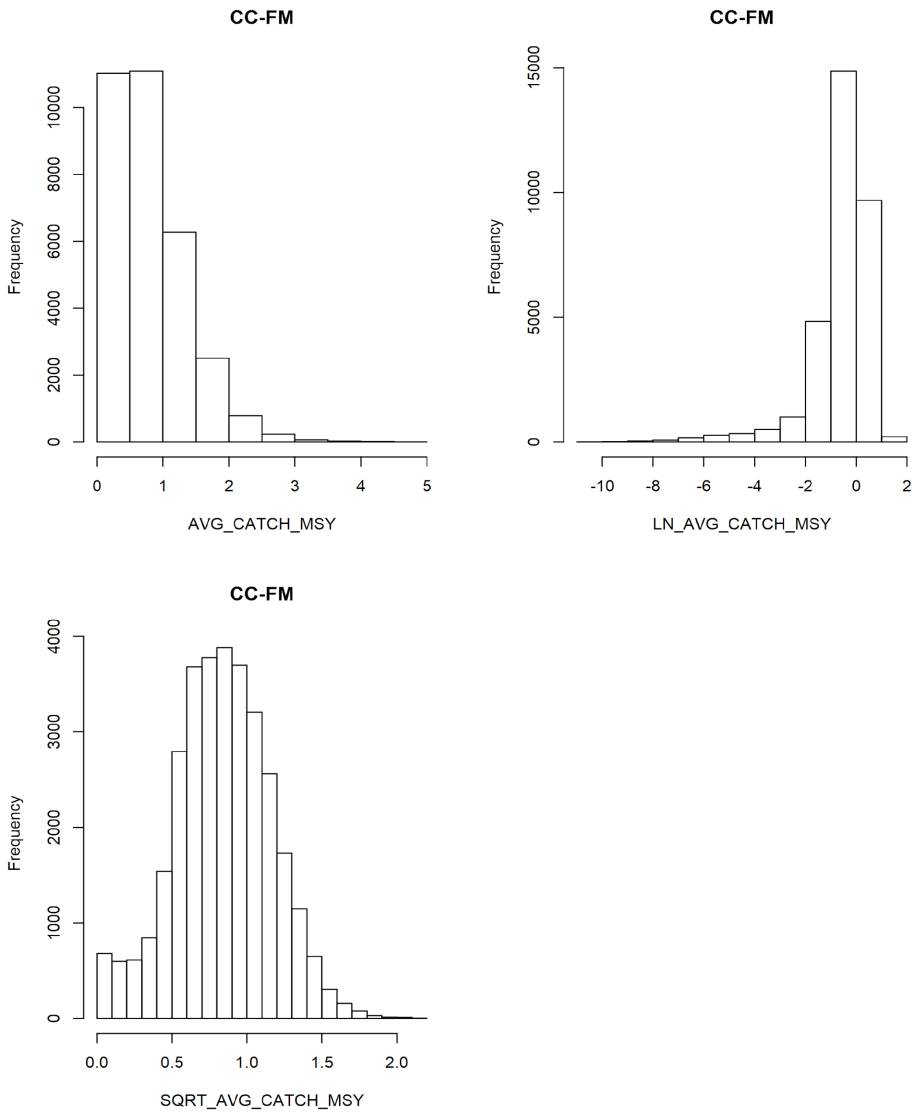


Figure A6.422. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

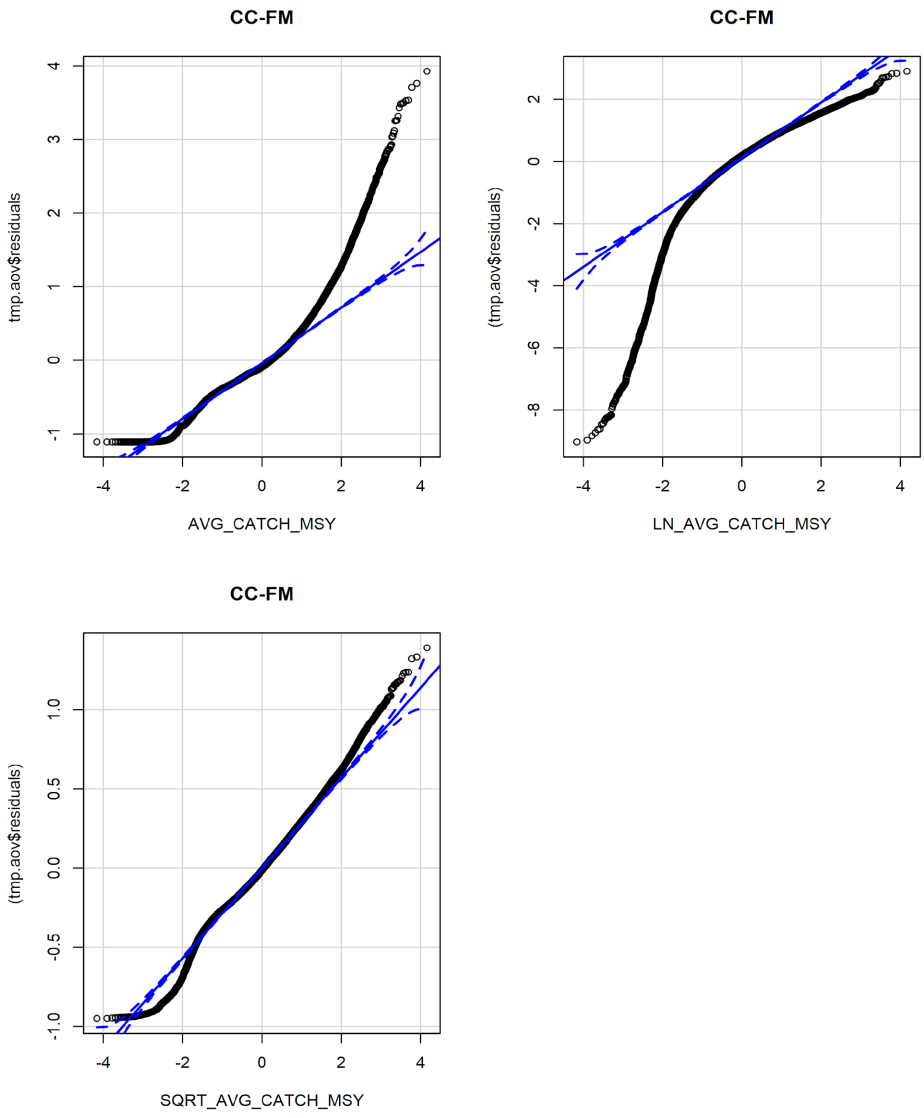


Figure A6.423. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

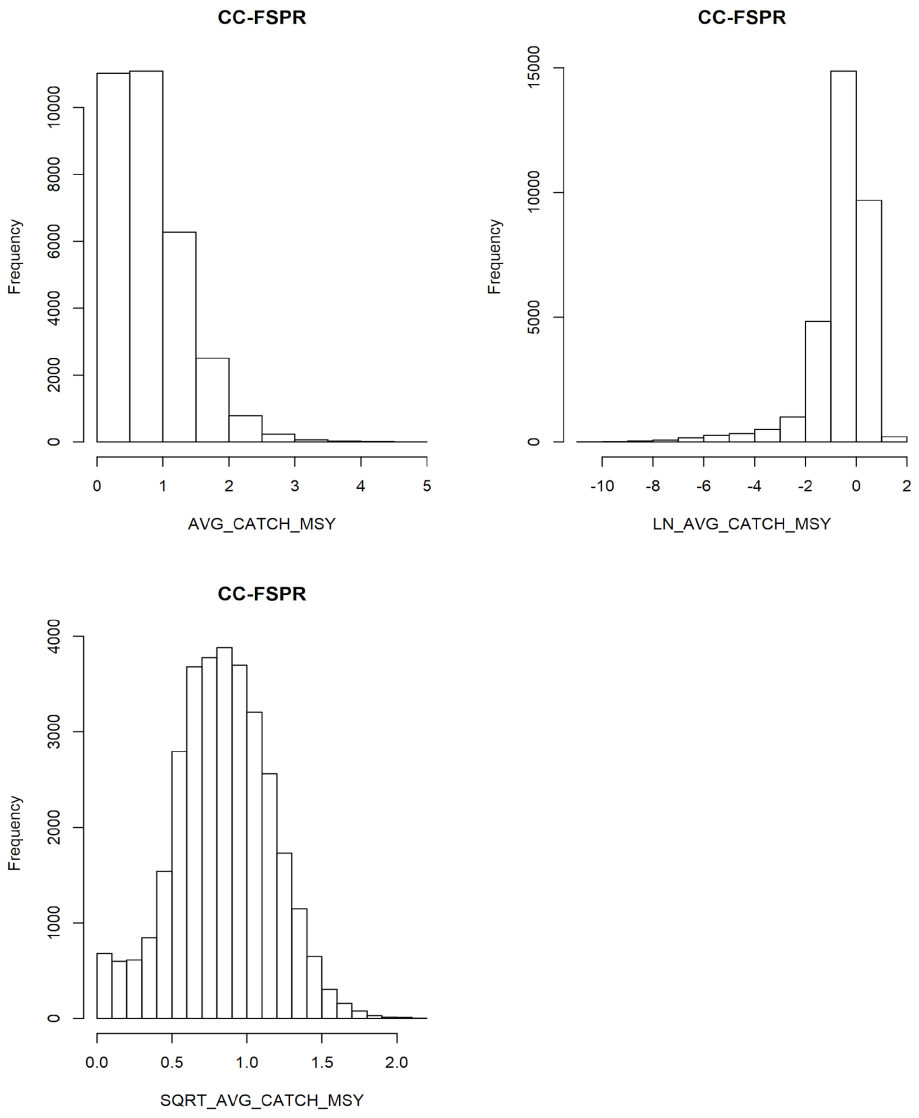


Figure A6.424. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

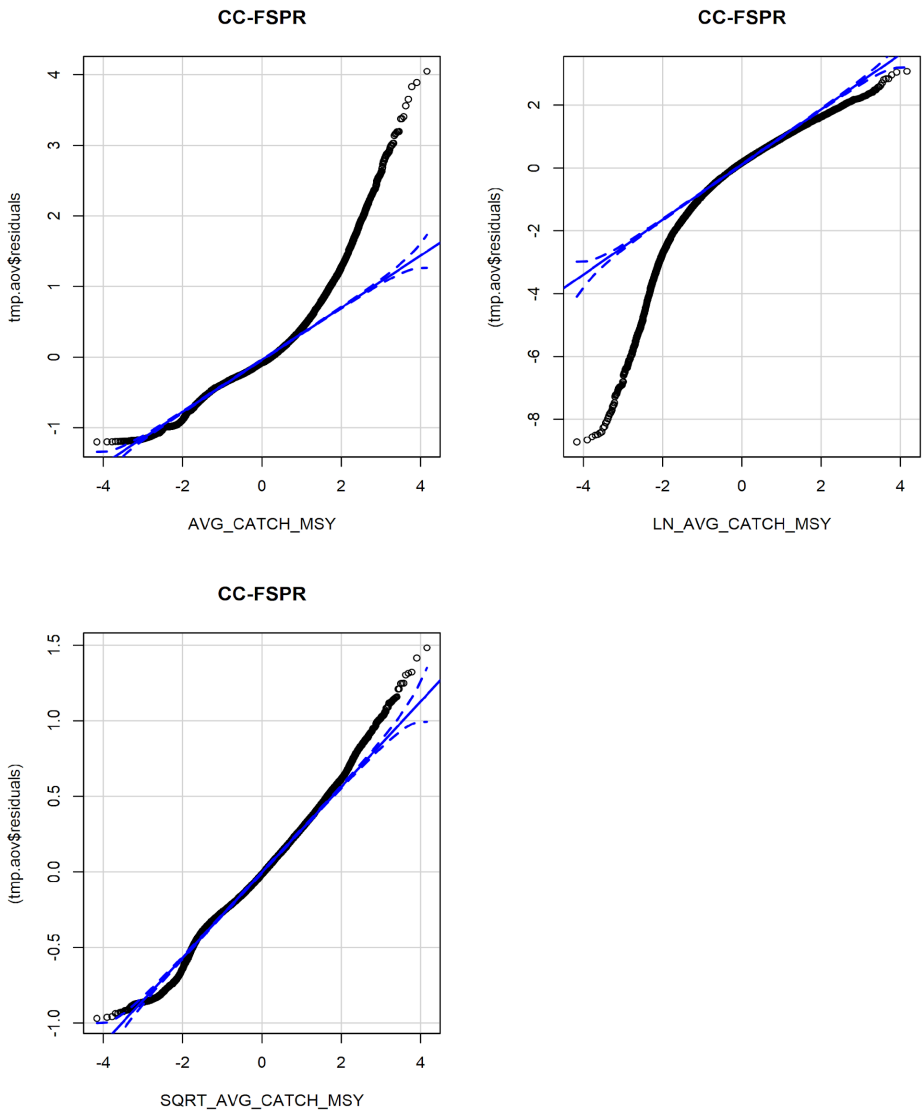


Figure A6.425. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

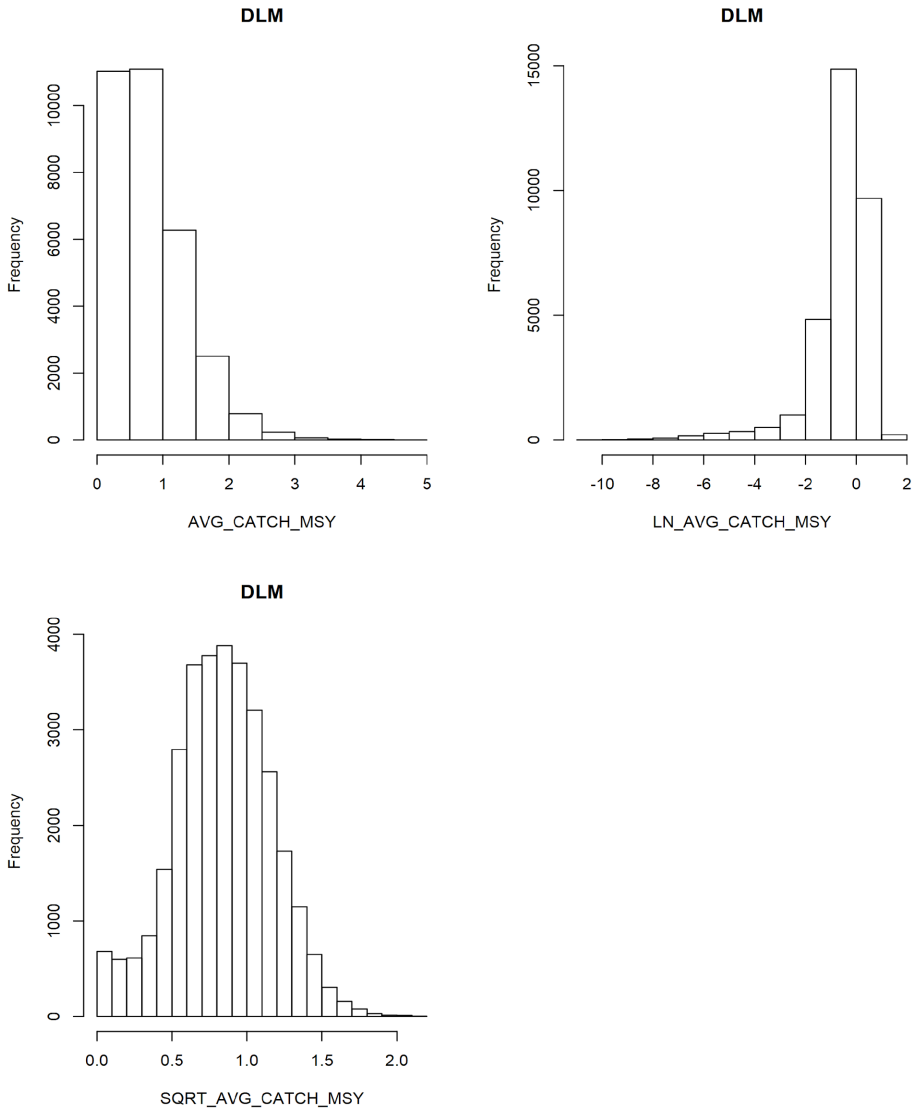


Figure A6.426. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

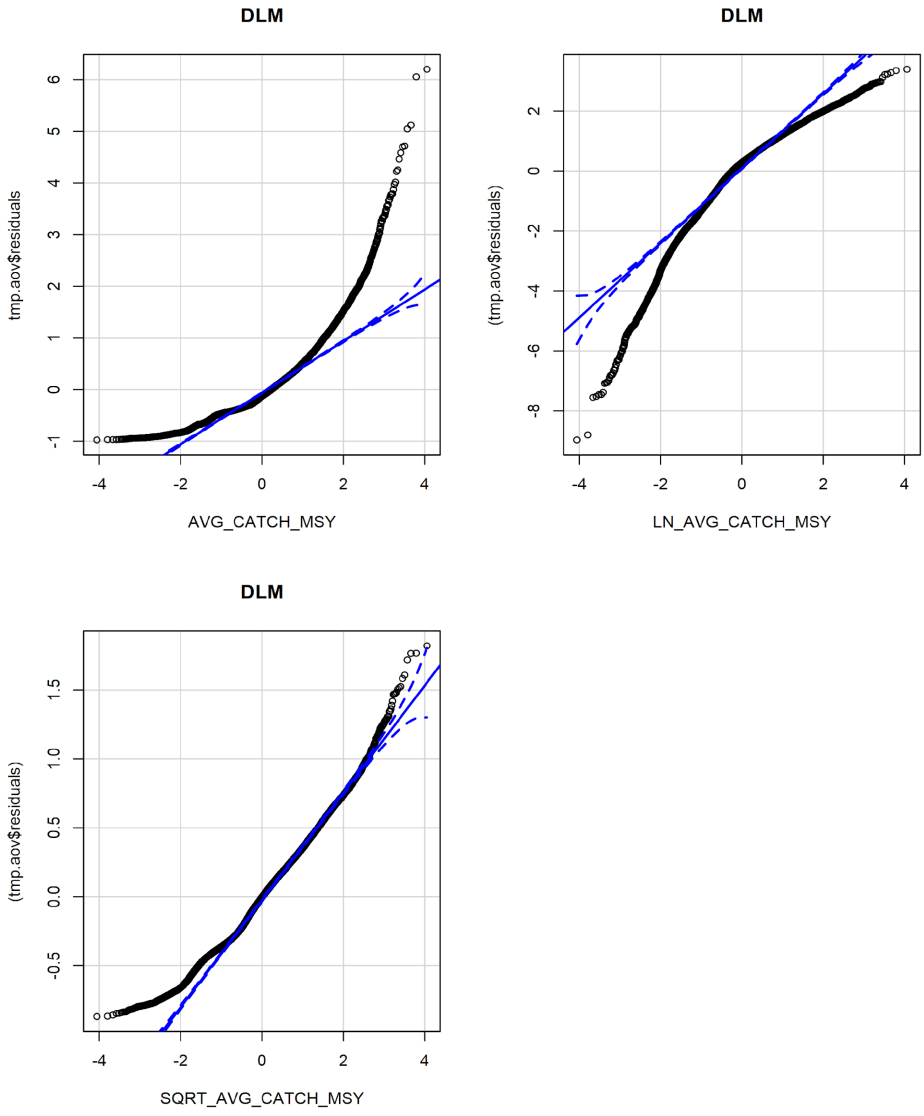


Figure A6.427. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

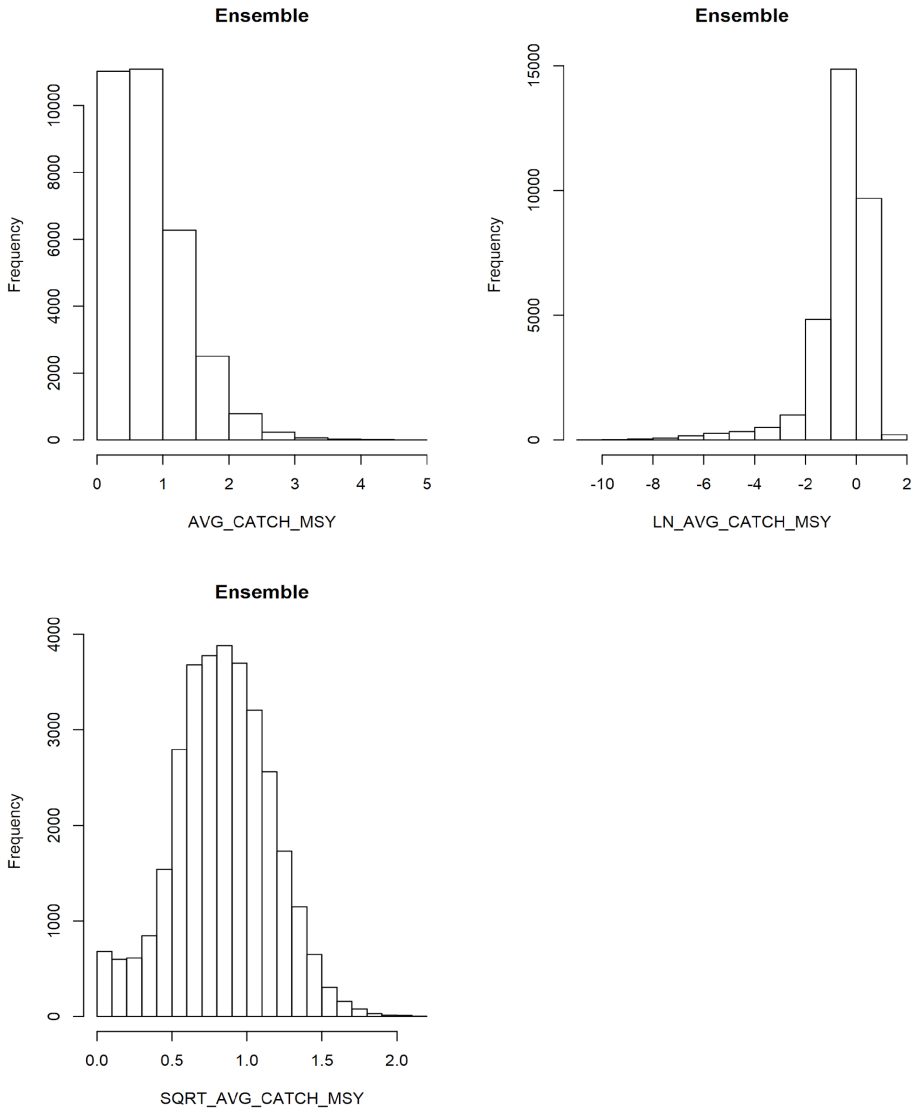


Figure A6.428. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

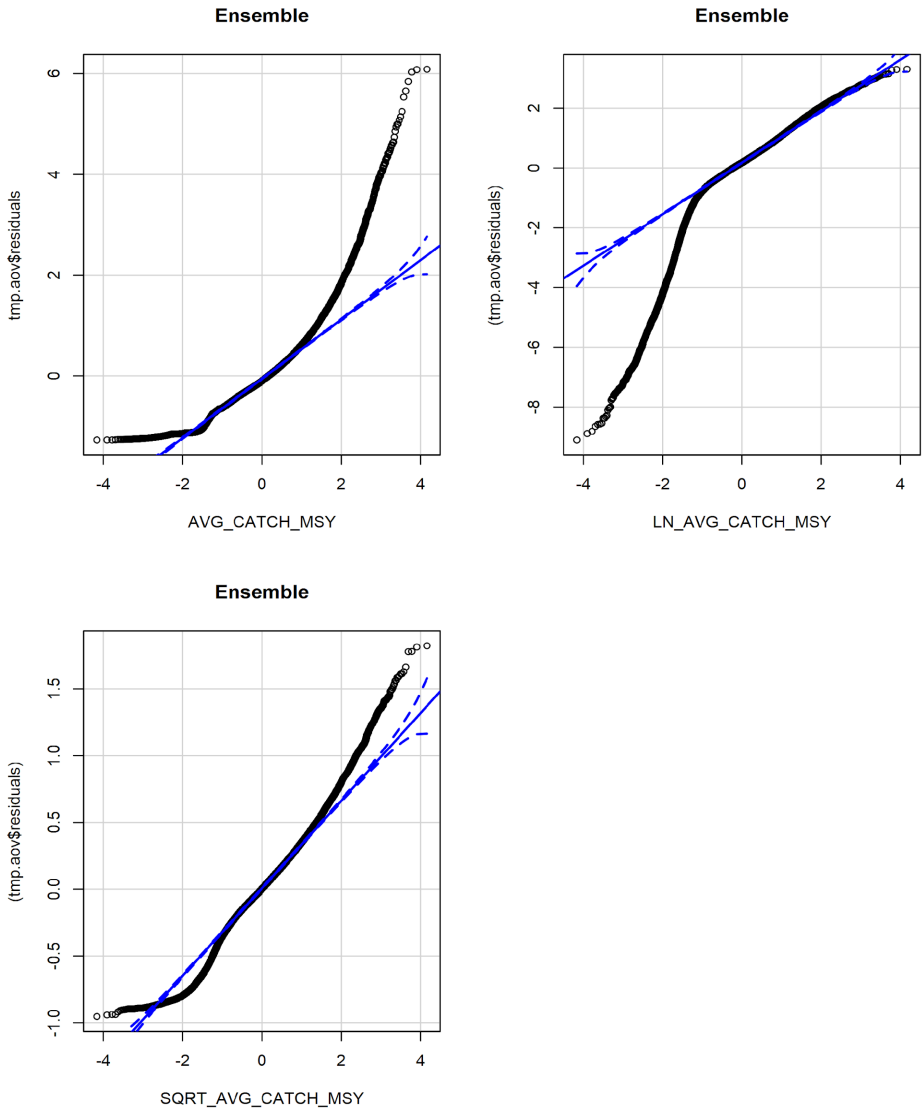


Figure A6.429. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

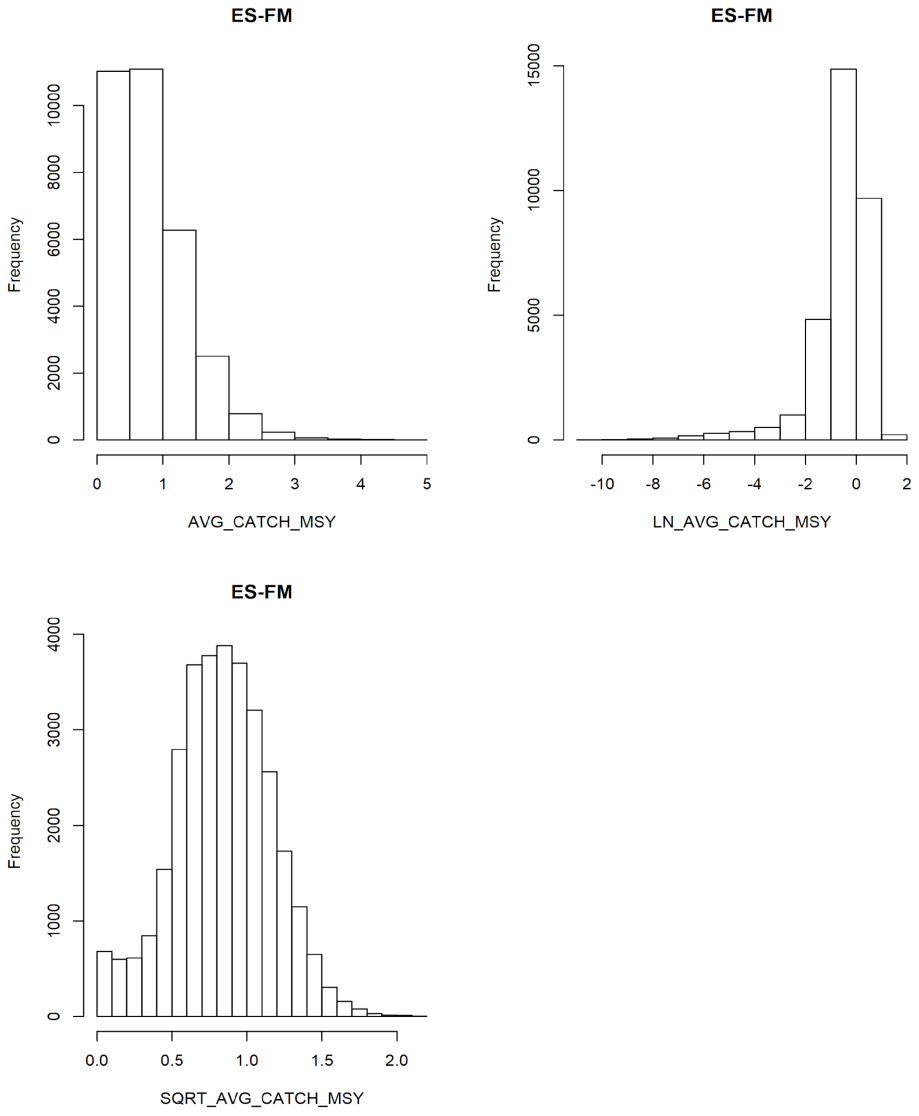


Figure A6.430. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

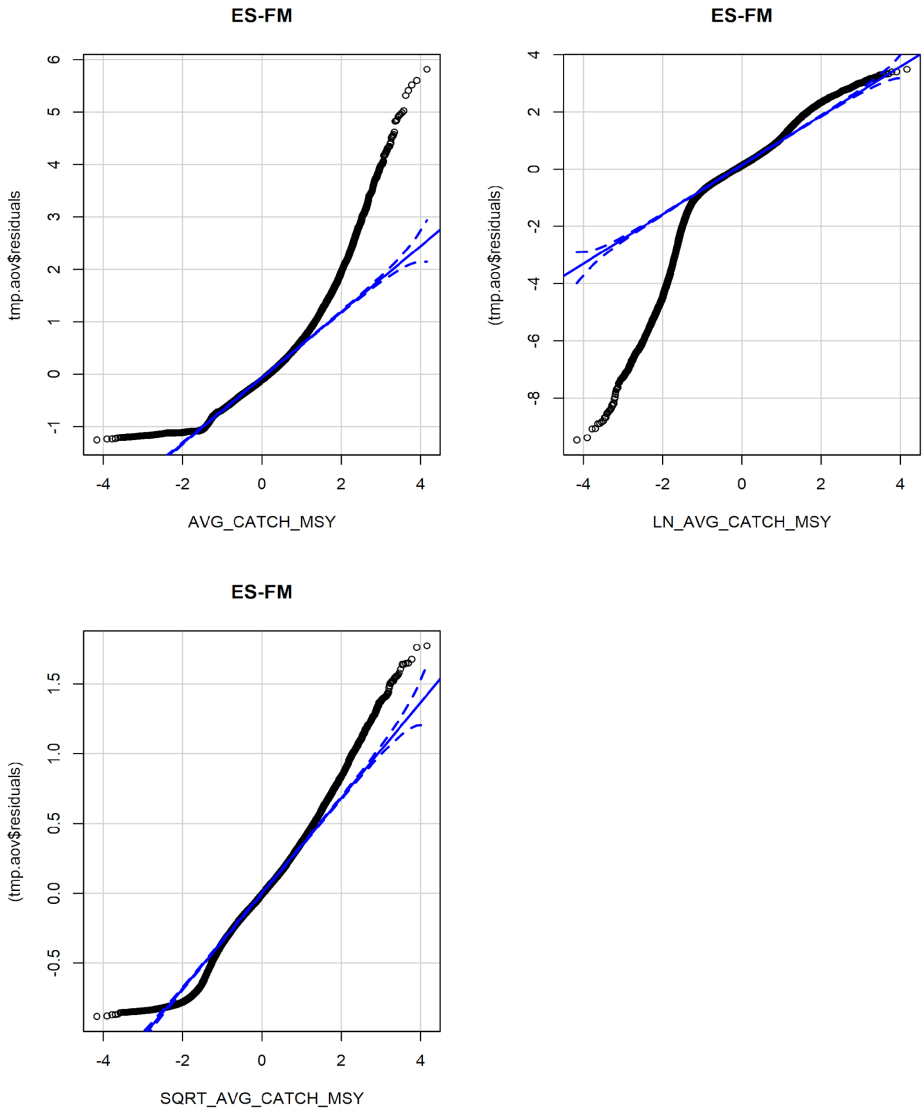


Figure A6.431. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

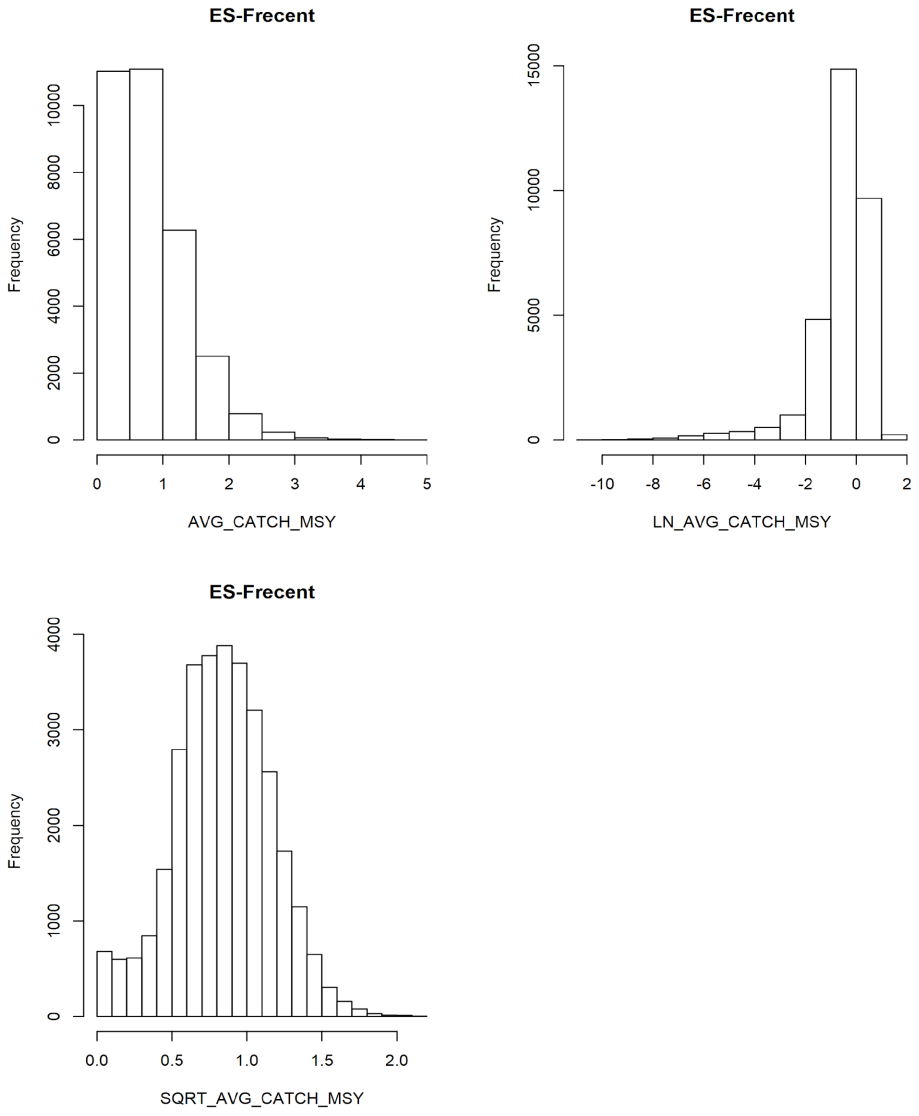


Figure A6.432. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

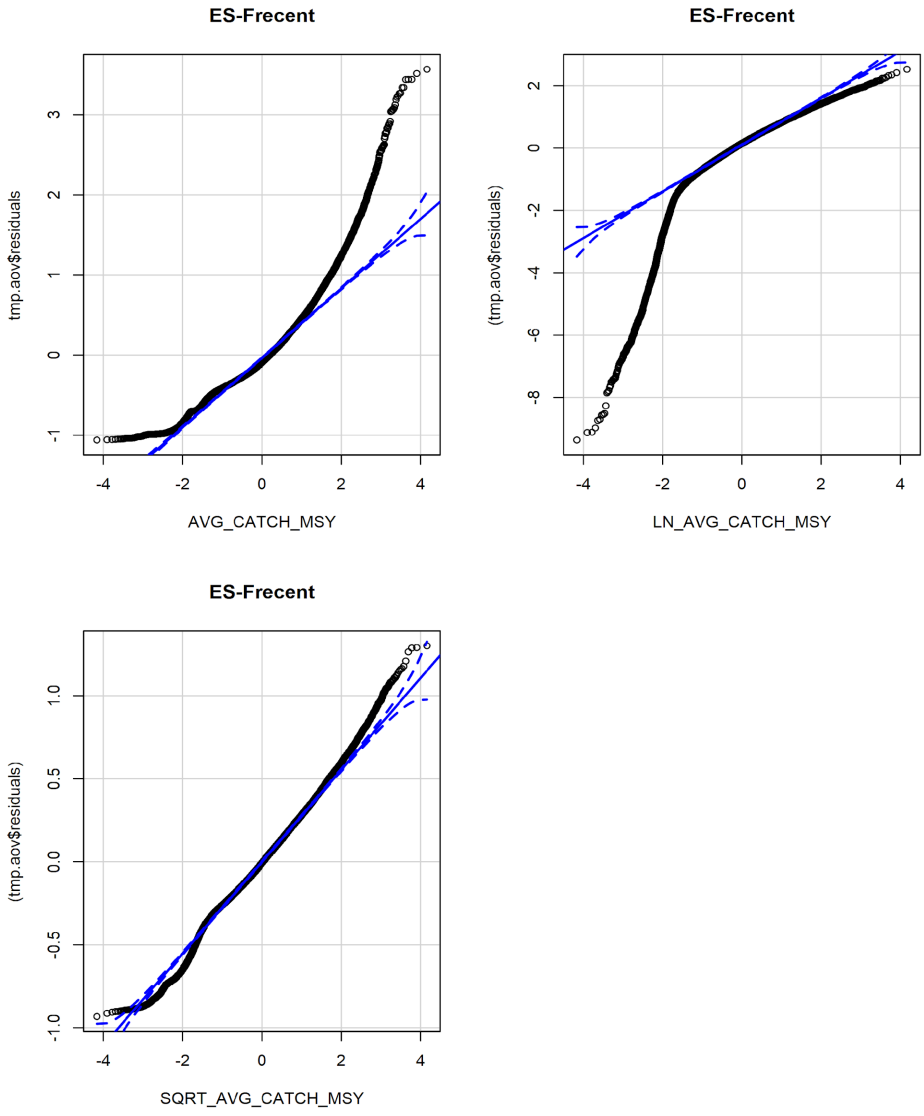


Figure A6.433. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

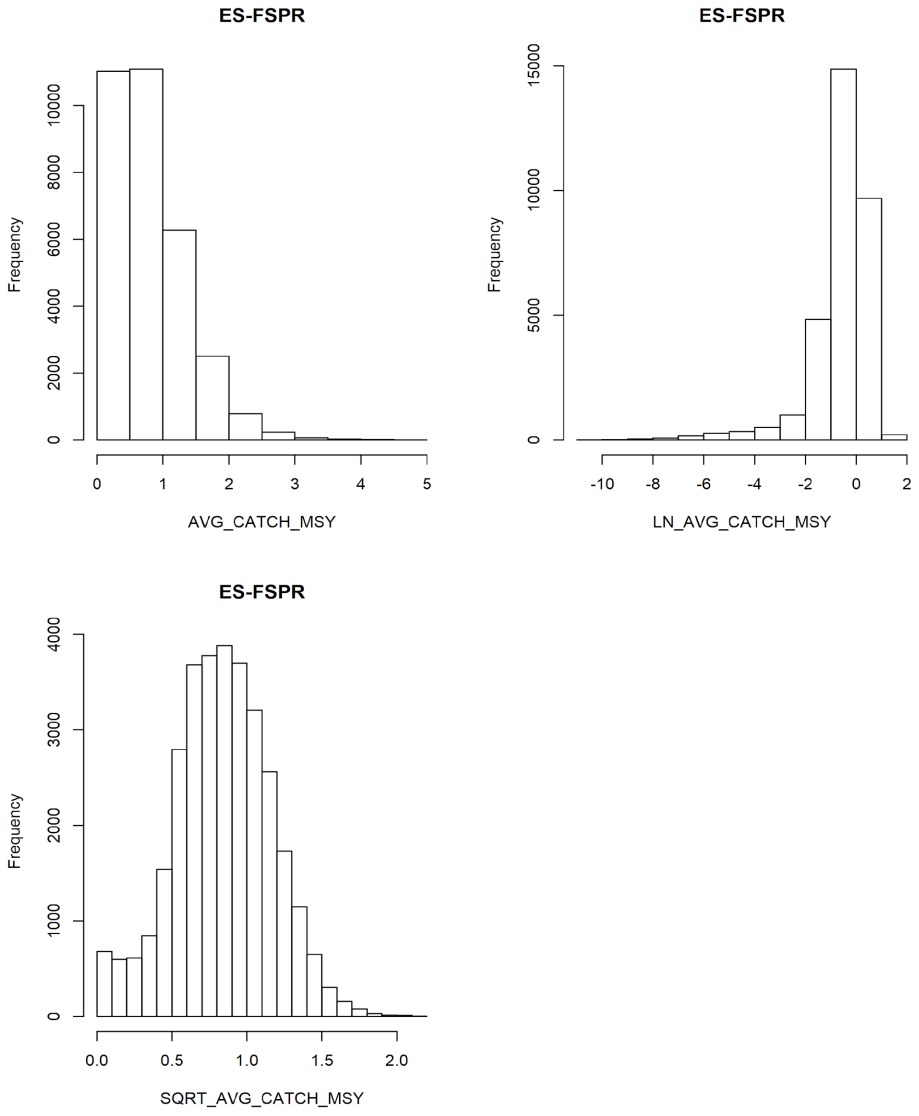


Figure A6.434. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

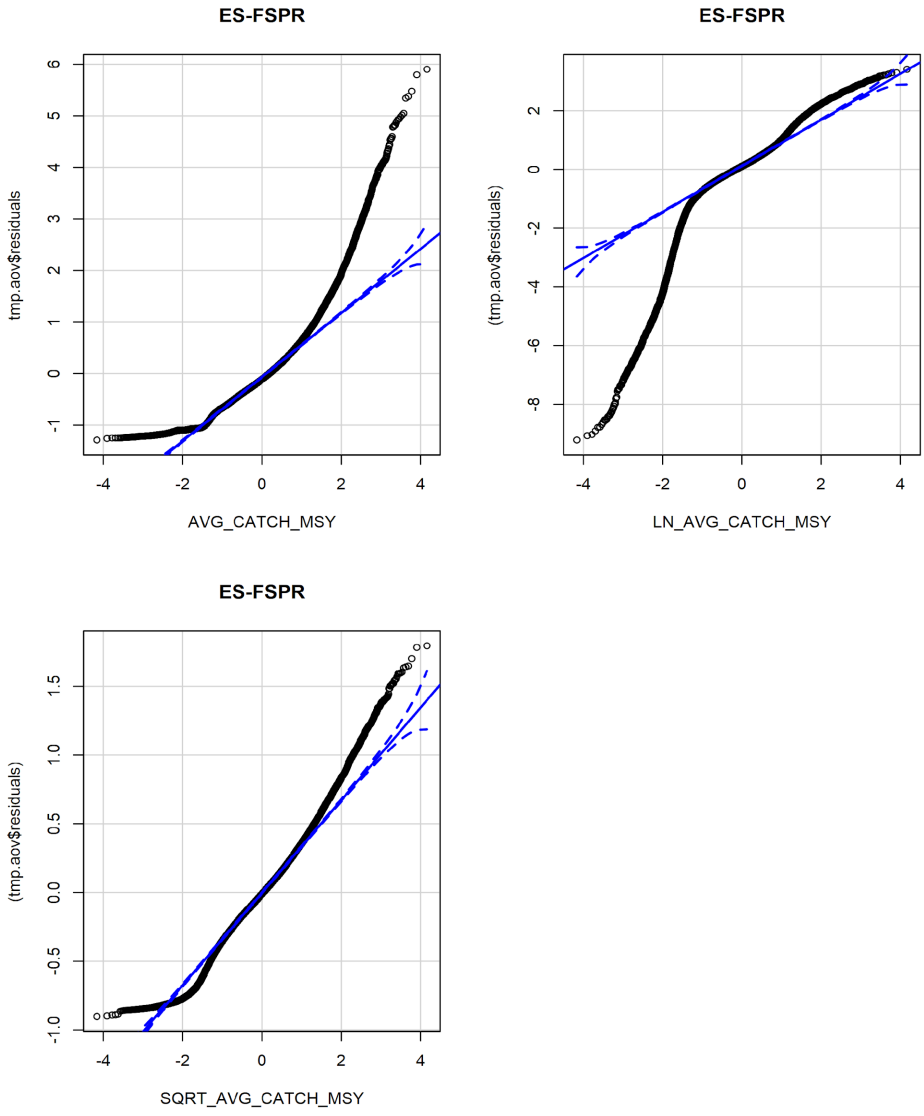


Figure A6.435. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

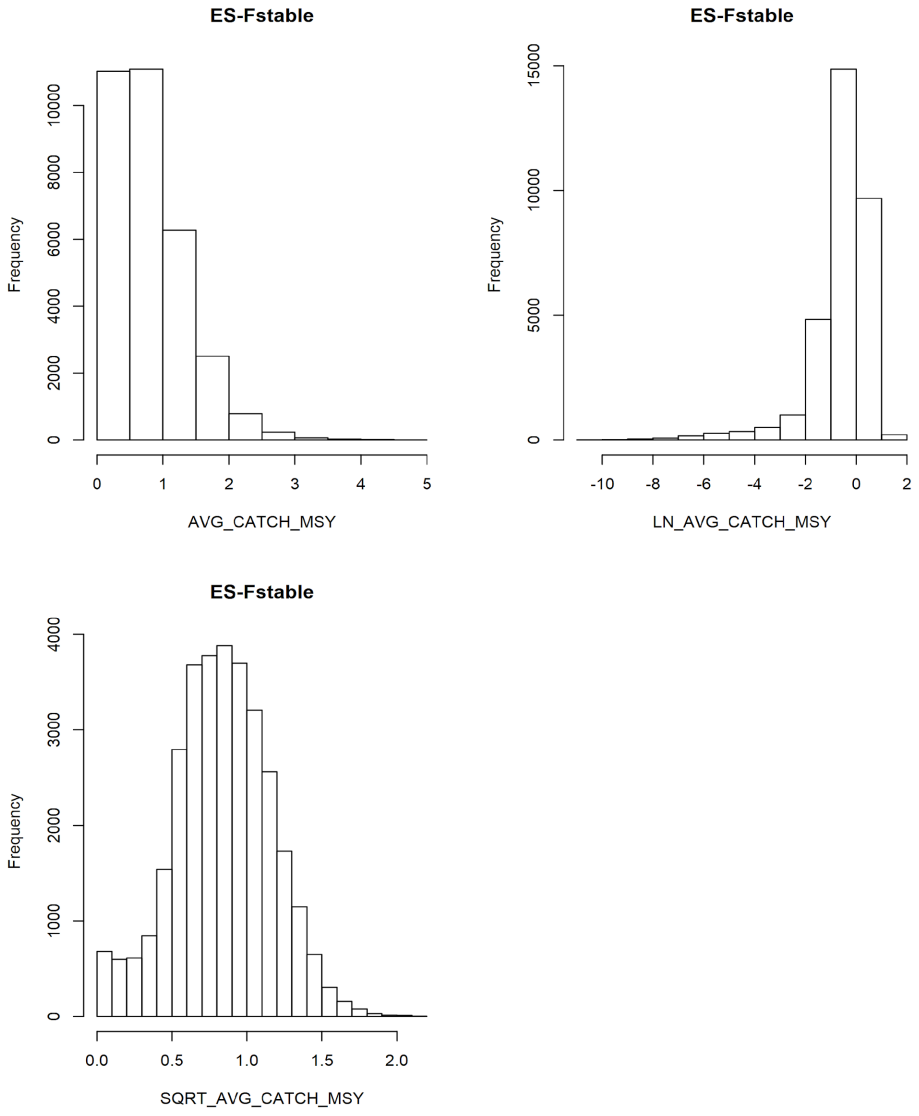


Figure A6.436. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

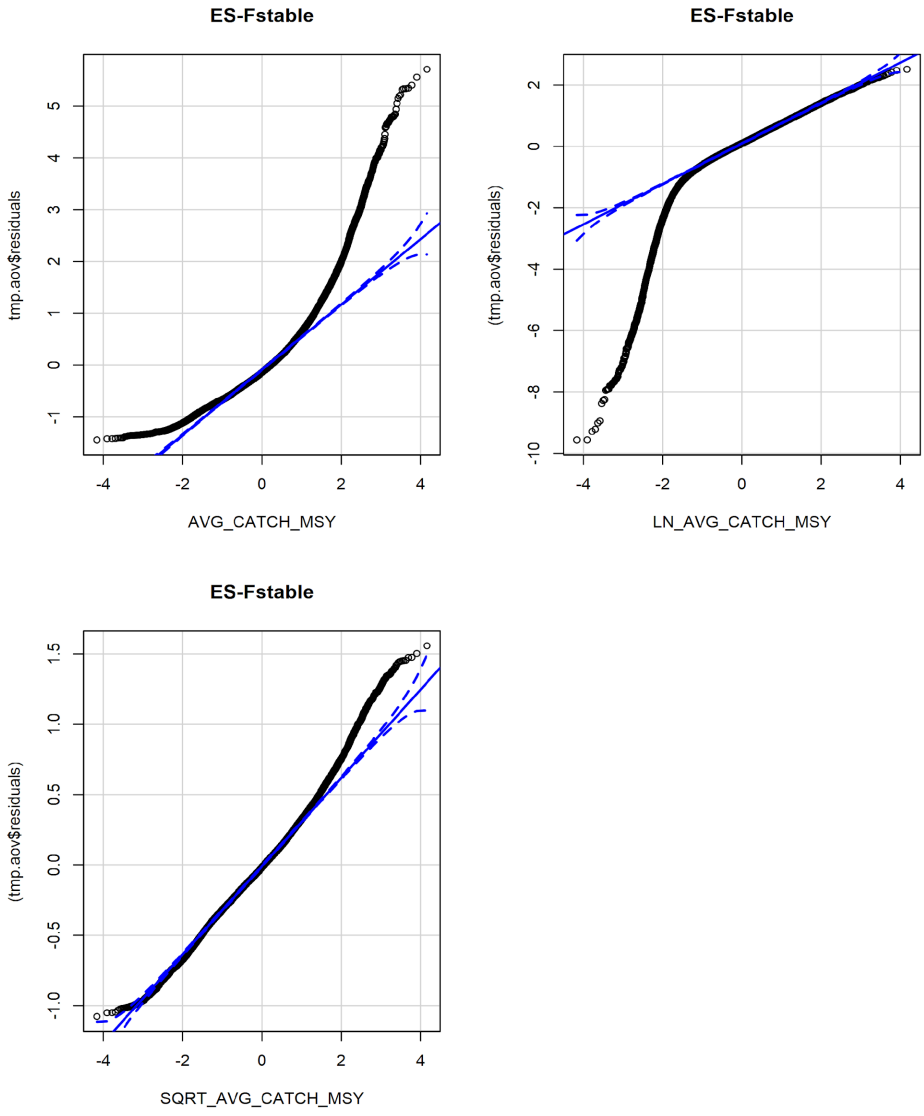


Figure A6.437. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

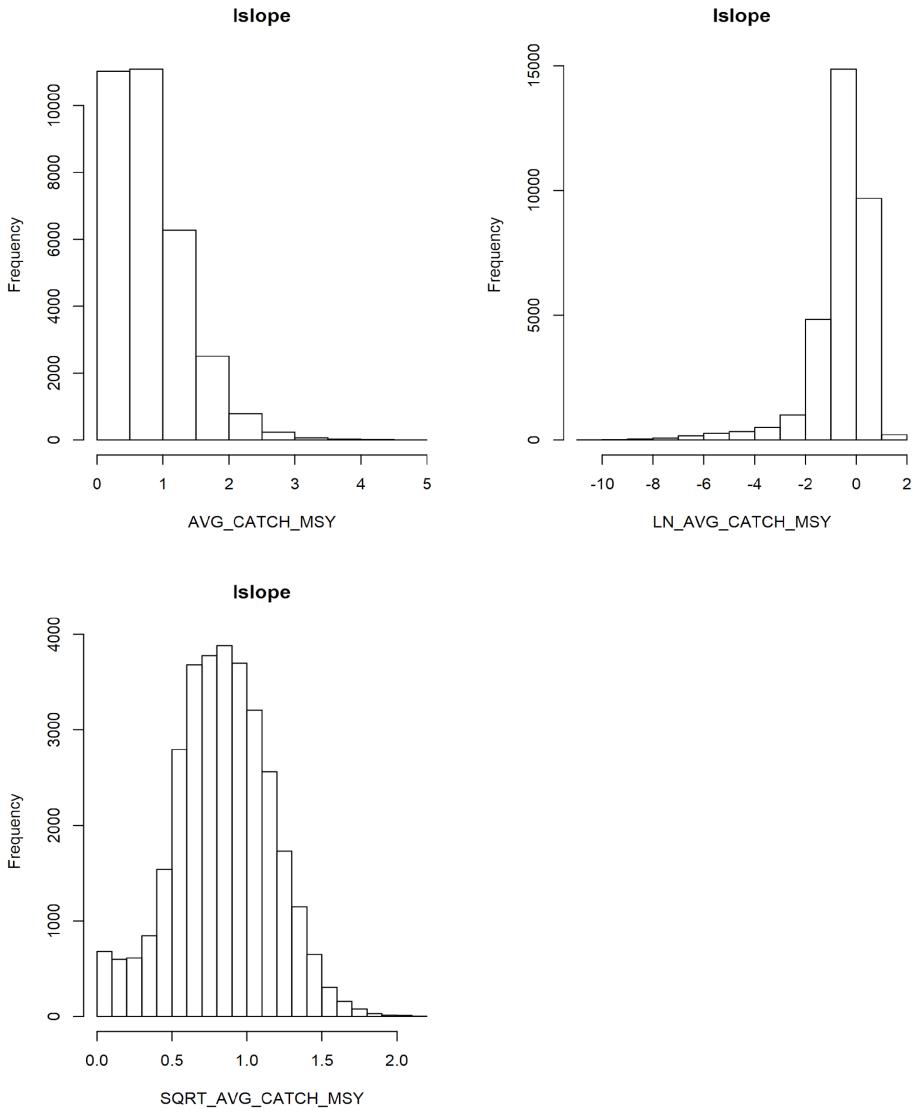


Figure A6.438. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

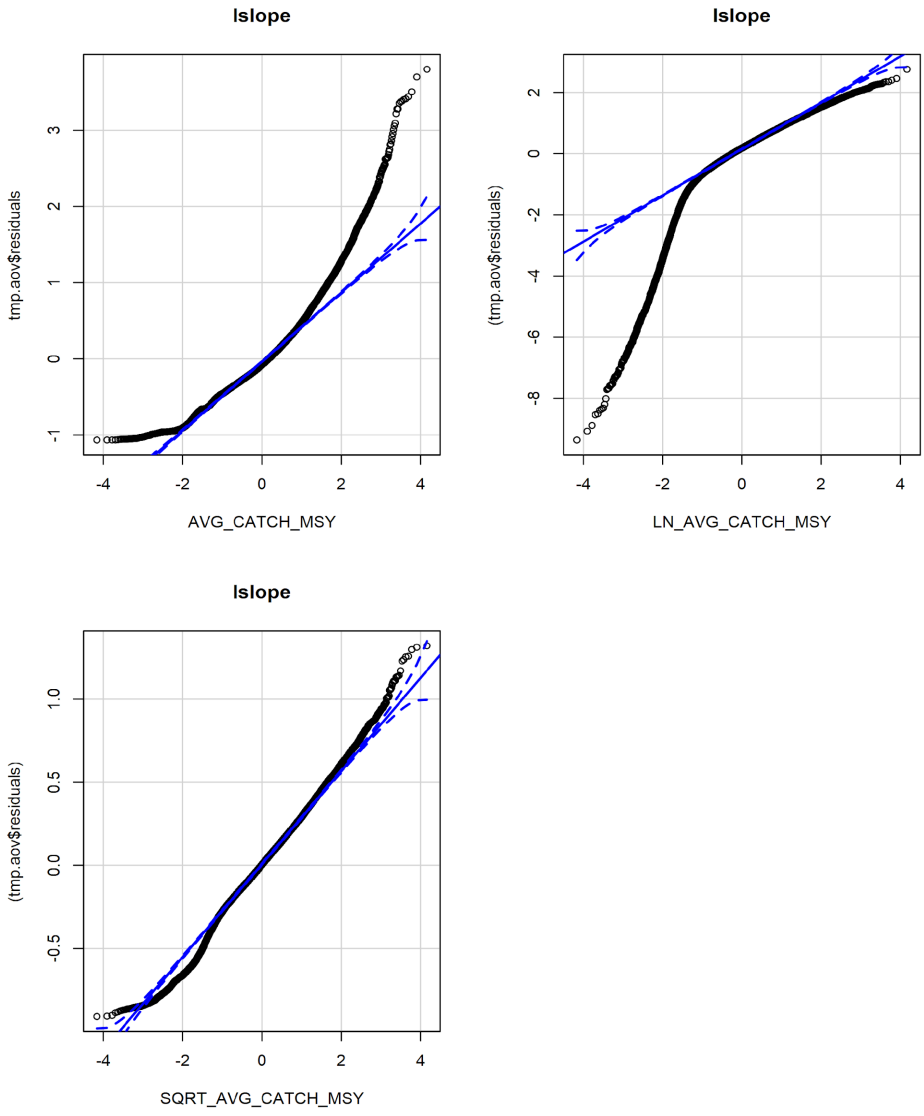


Figure A6.439. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

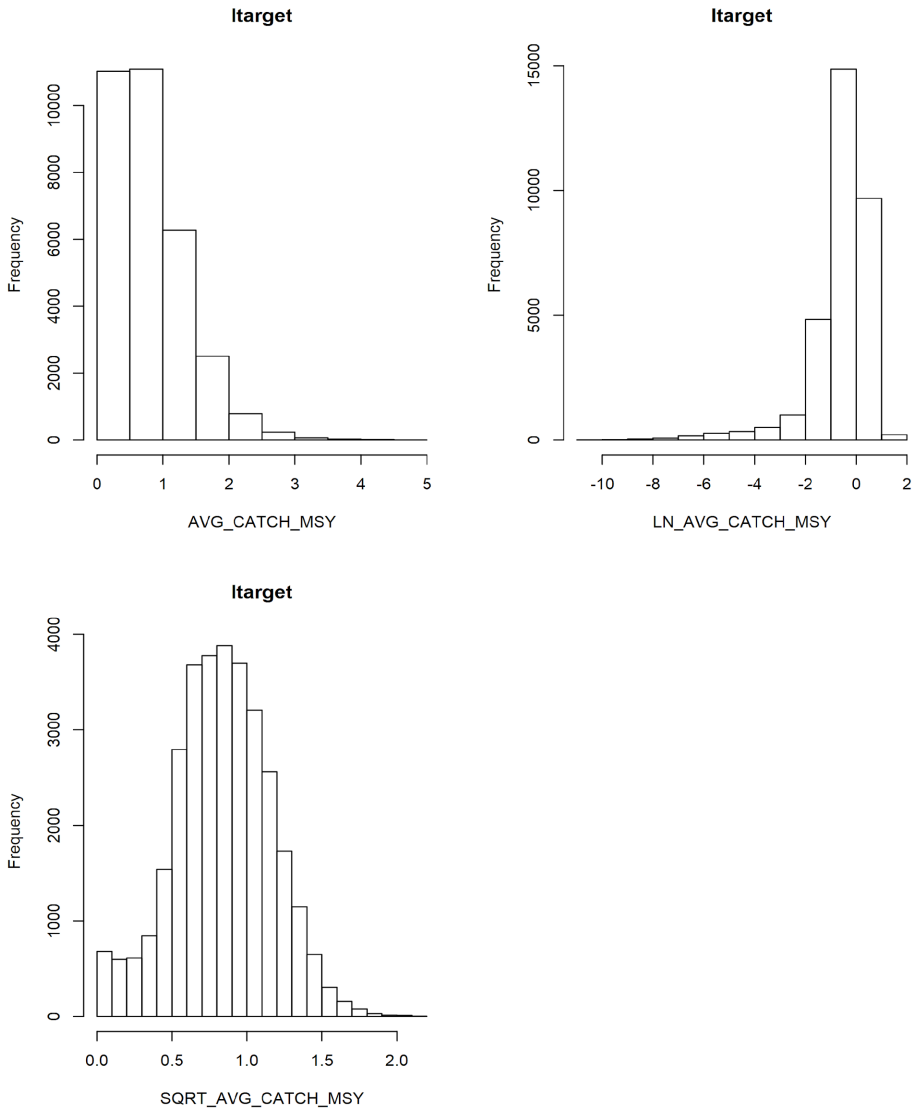


Figure A6.440. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

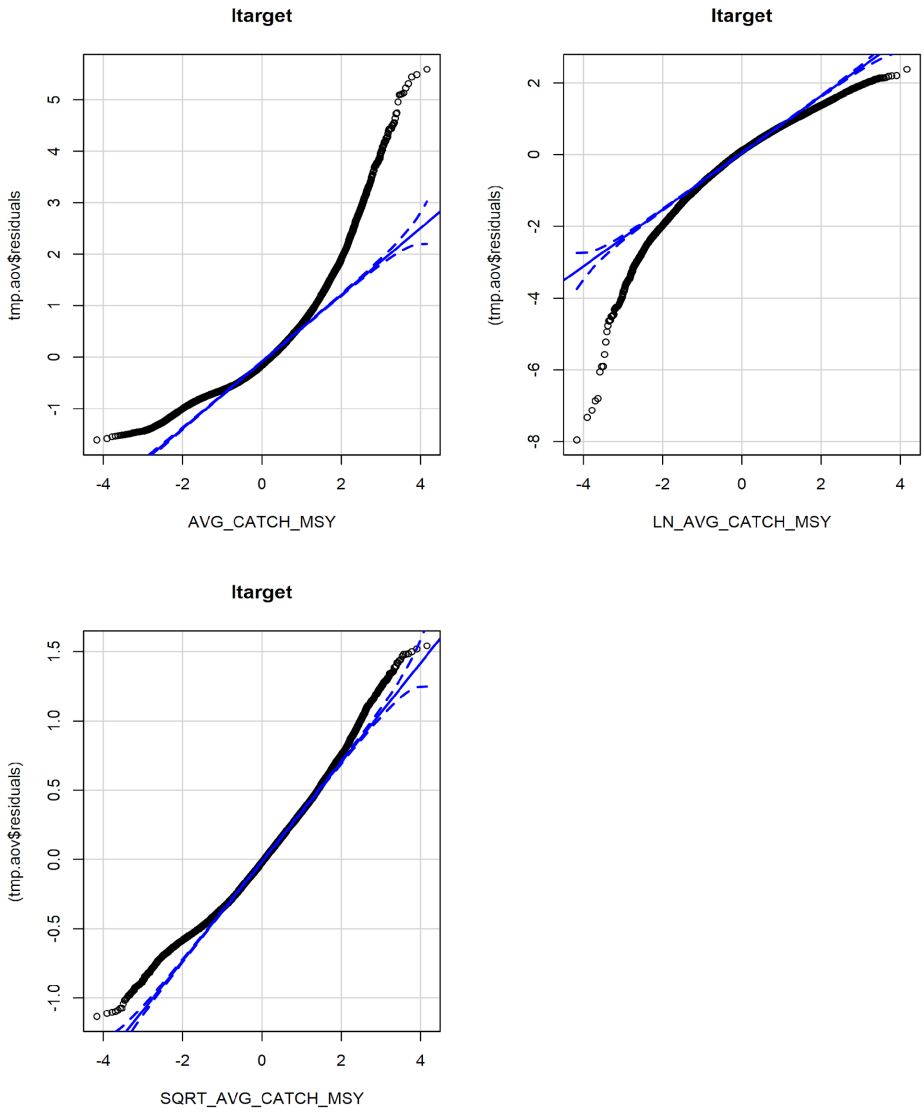


Figure A6.441. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

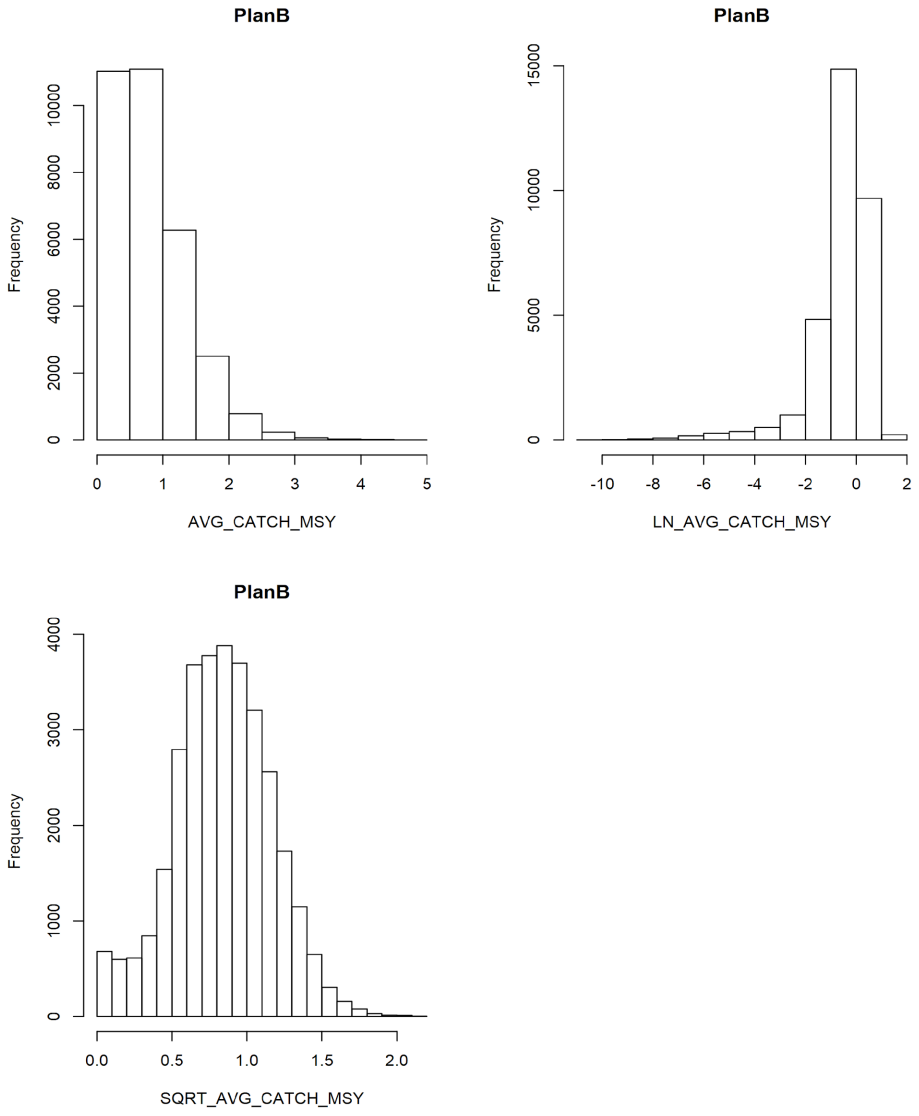


Figure A6.442. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

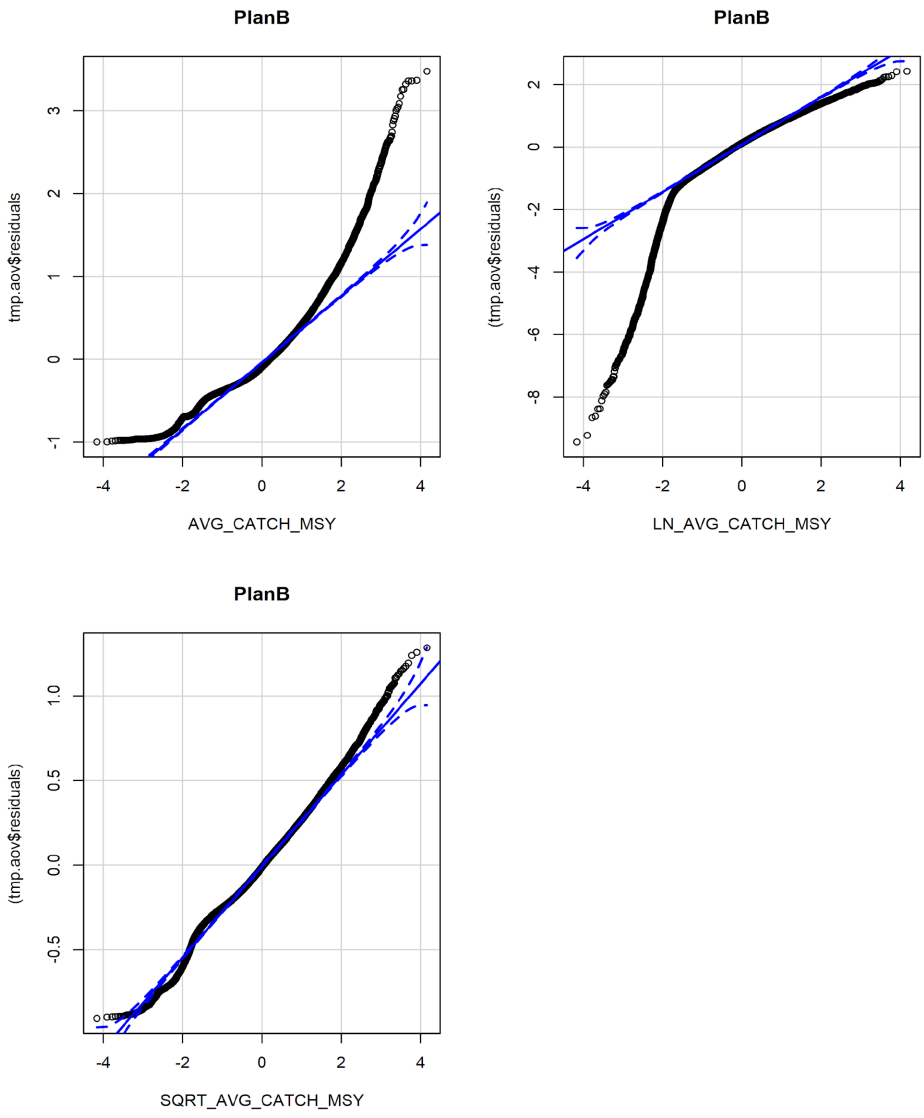


Figure A6.443. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

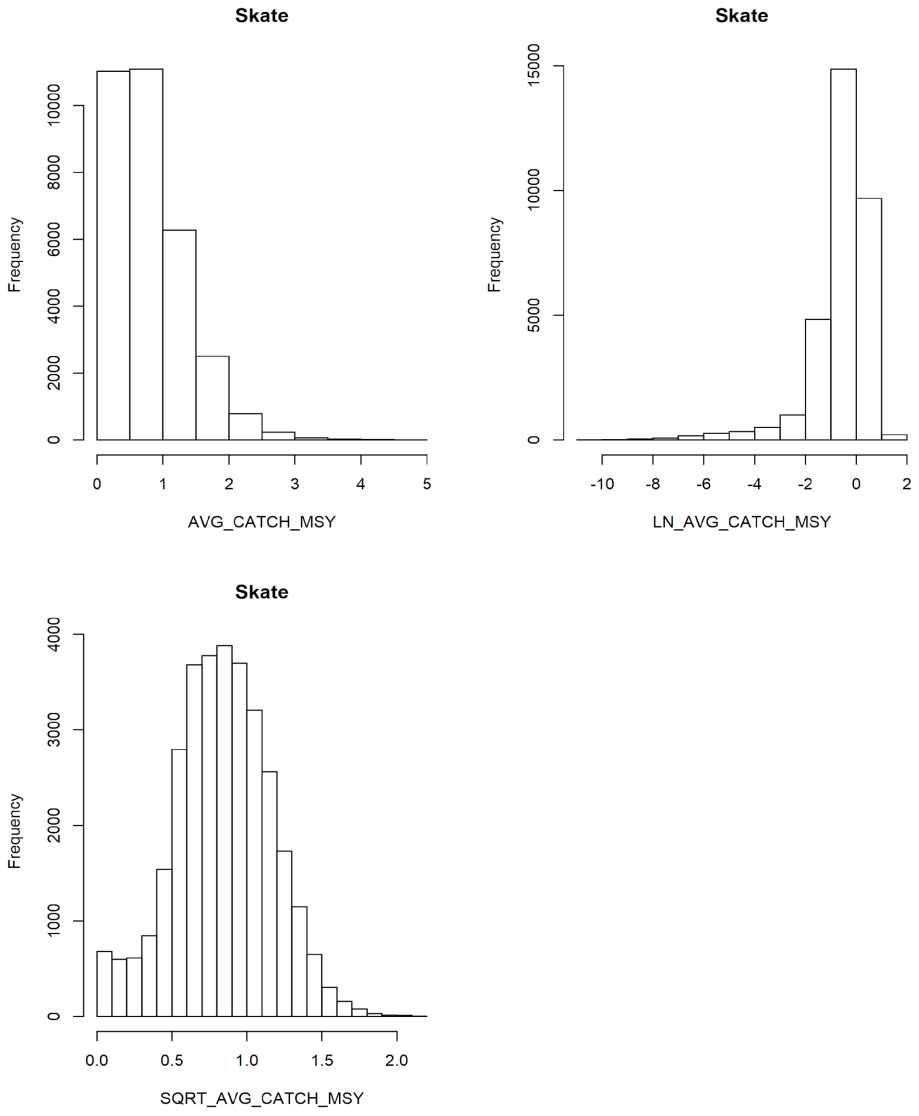


Figure A6.444. Catch/maximum sustainable yield (MSY) histograms for all simulation iterations by index-based method (IBM) for untransformed data (top left), natural logarithm transformed data (top right), and square root transformed data (bottom left).

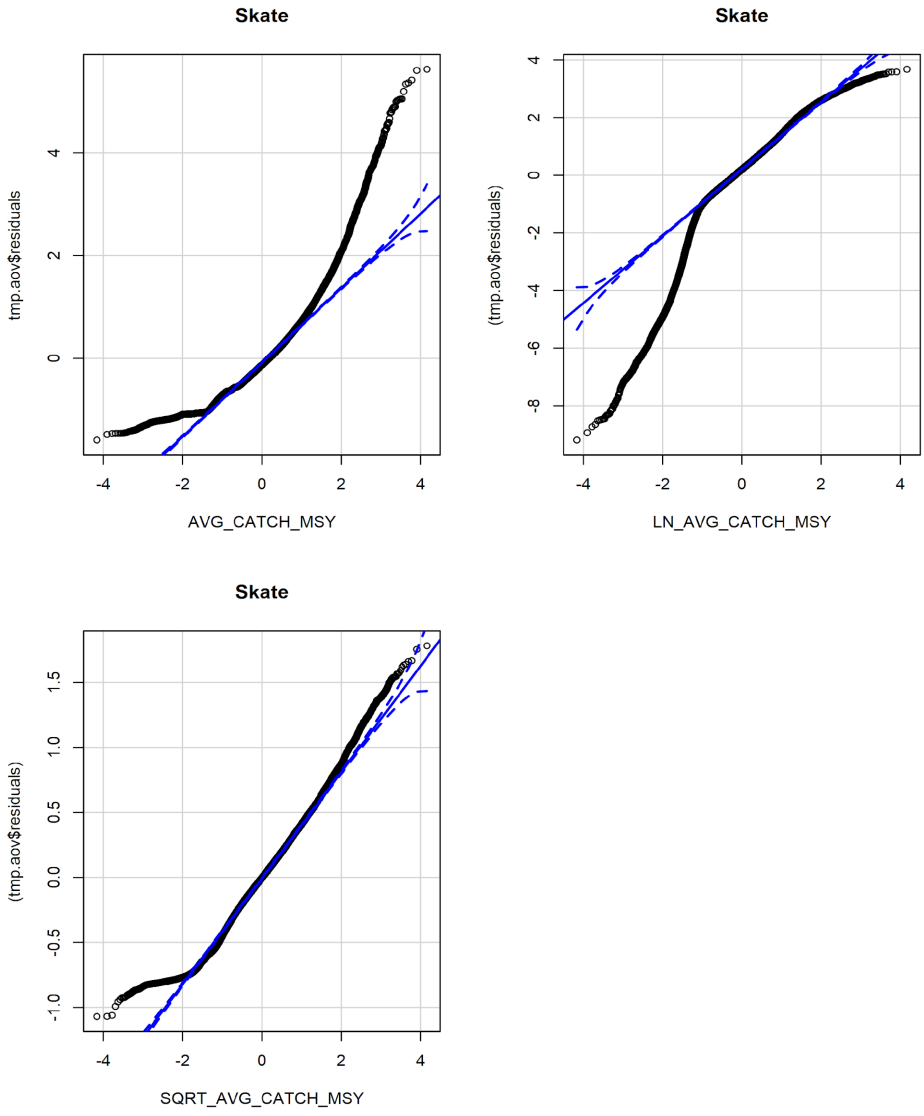


Figure A6.445. Catch/maximum sustainable yield (MSY) qq plots of the normalized residuals from the linear model fit for the analyses of variance (ANOVAs).

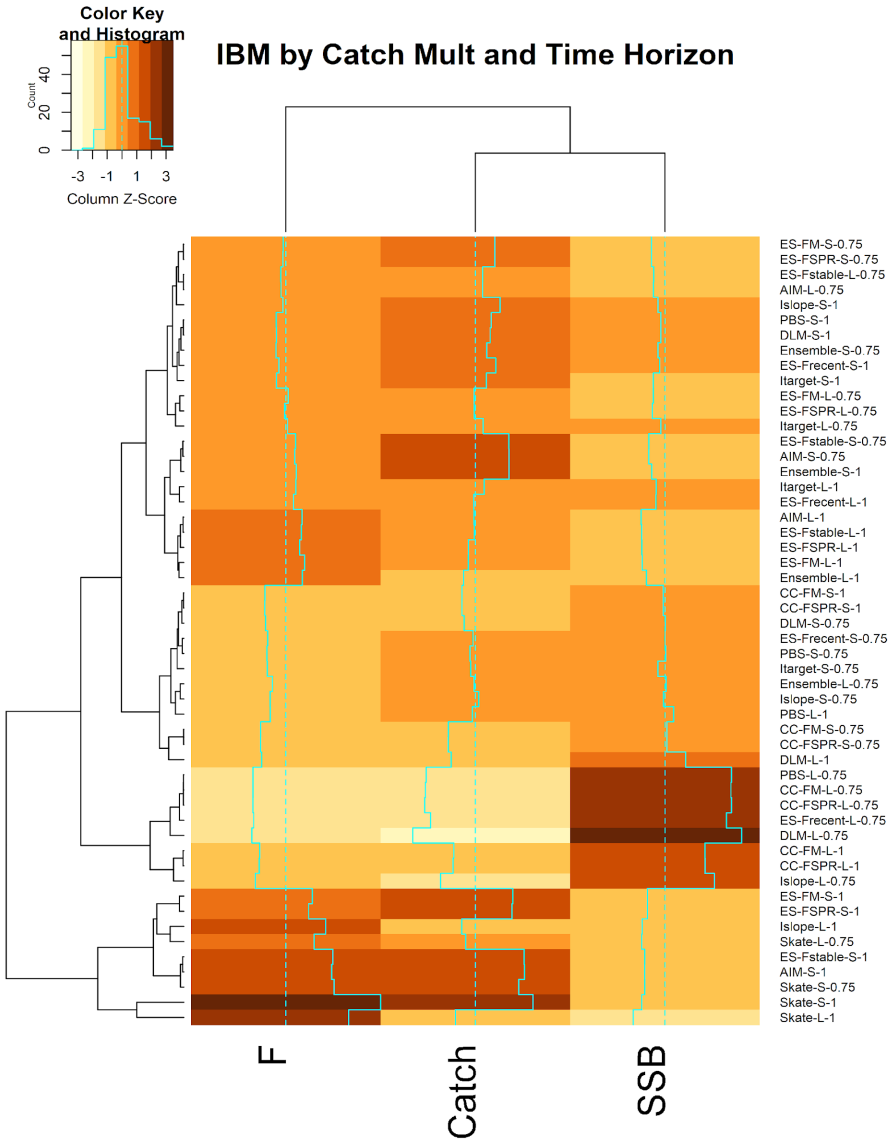


Figure A6.446. Heatmap of median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), time horizon (short- or long-term summary, "S" or "L"), and catch advice multiplier (1 or 0.75). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate position of that cell relative to the mean. Values for each cell in the heatmap can be found in the associated table (rows in the same order as the heatmap).

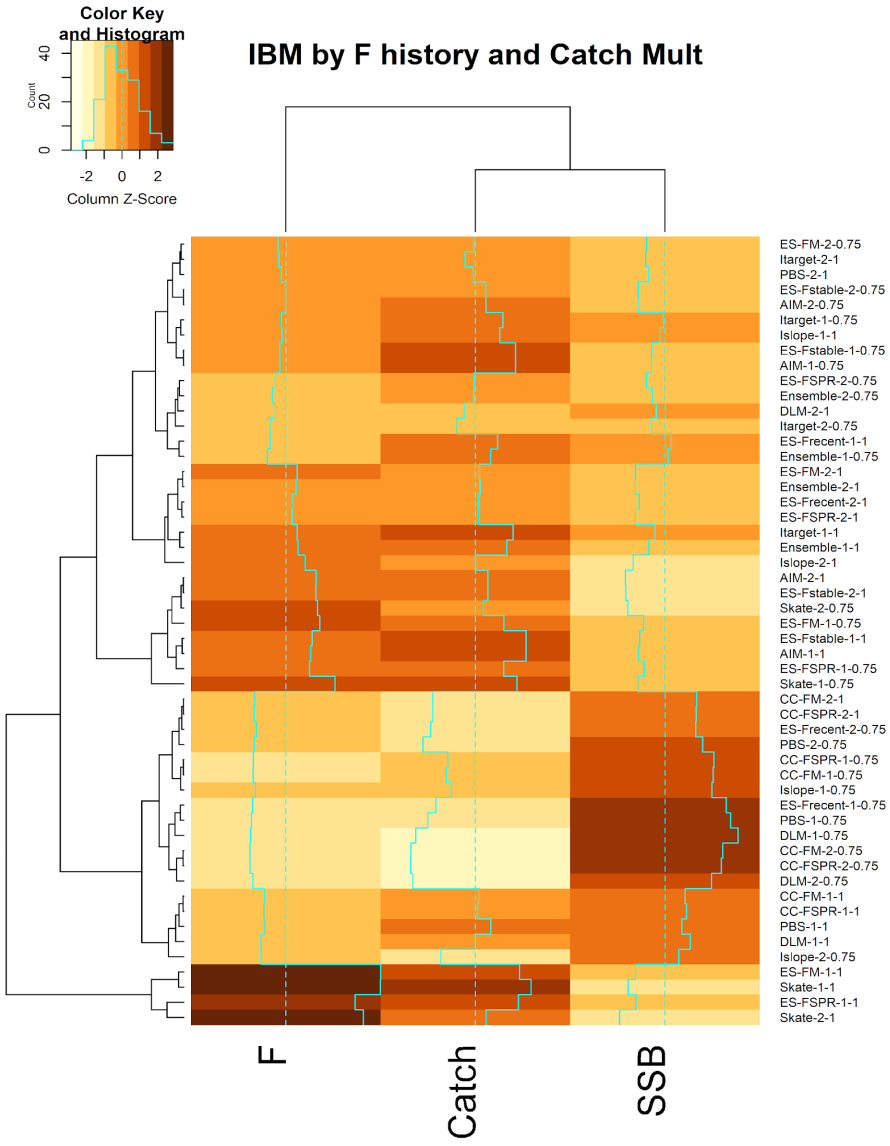


Figure A6.447. Heatmap of median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), catch advice multiplier (1 or 0.75), and fishing history (1=overfishing then F=F_{MSY}, 2=always overfishing). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate position of that cell relative to the mean. Values for each cell in the heatmap can be found in the associated table (rows in the same order as the heatmap).

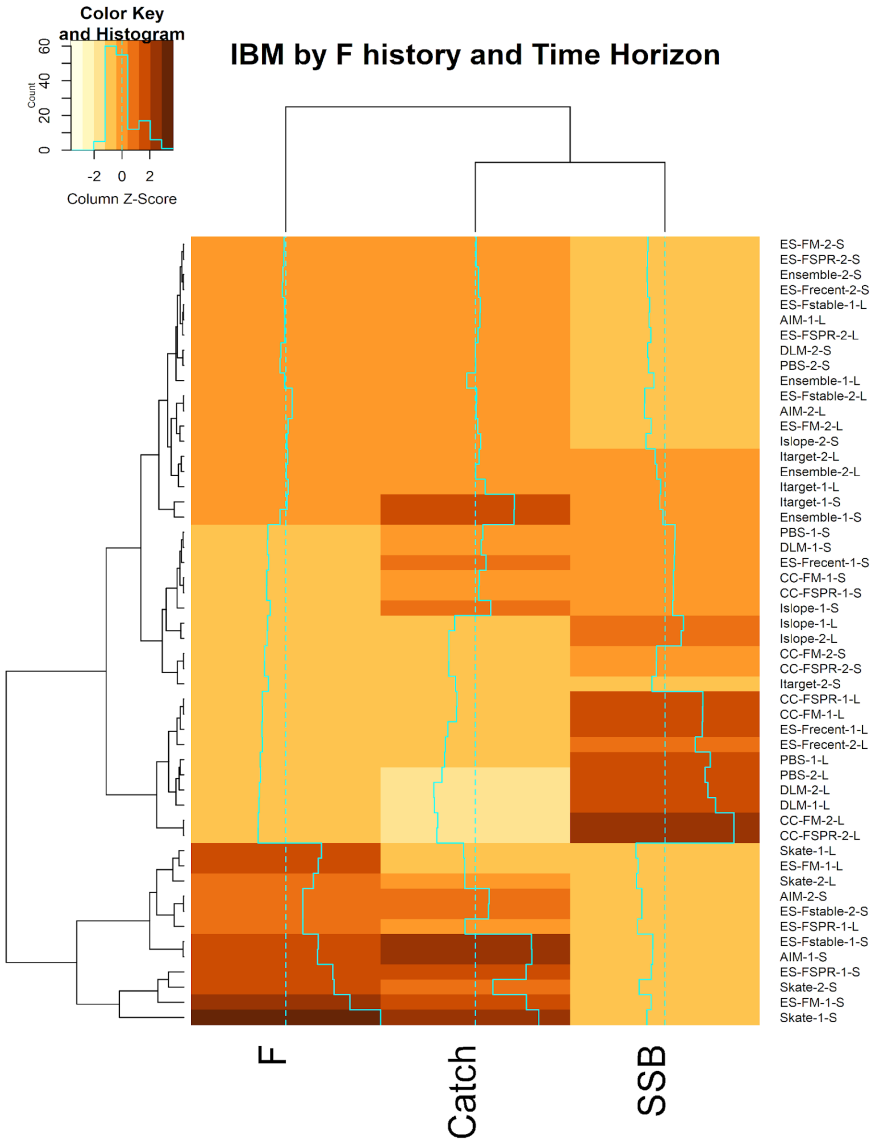


Figure A6.448. Heatmap of median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), fishing history (1=overfishing then F=F_{MSY}, 2=always overfishing), and time horizon (short- or long-term summary, "S" or "L"). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate position of that cell relative to the mean. Values for each cell in the heatmap can be found in the associated table (rows in the same order as the heatmap).

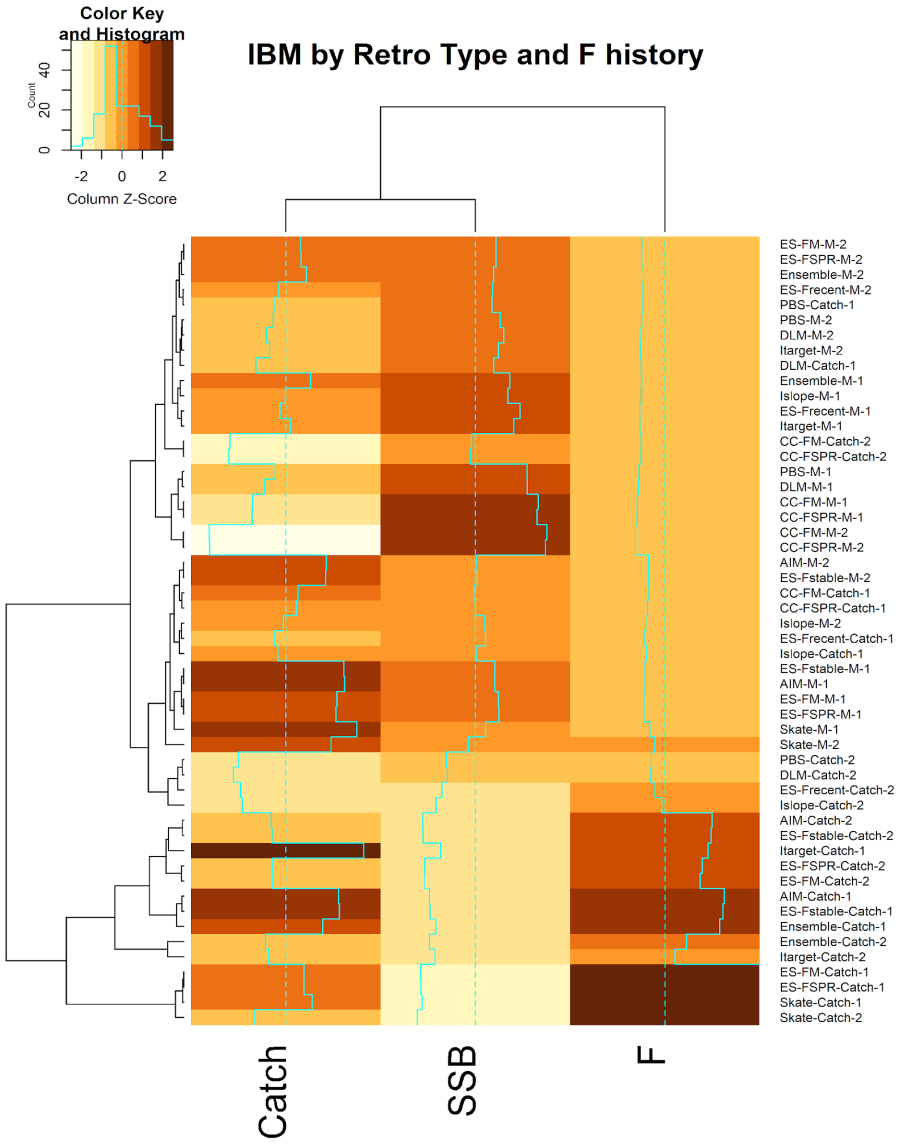


Figure A6.449. Heatmap of median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), source of retrospective pattern (C=catch, M=natural mortality), and fishing history (1=overfishing then F=F_{MSY}, 2=always overfishing). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate position of that cell relative to the mean. Values for each cell in the heatmap can be found in the associated table (rows in the same order as the heatmap).

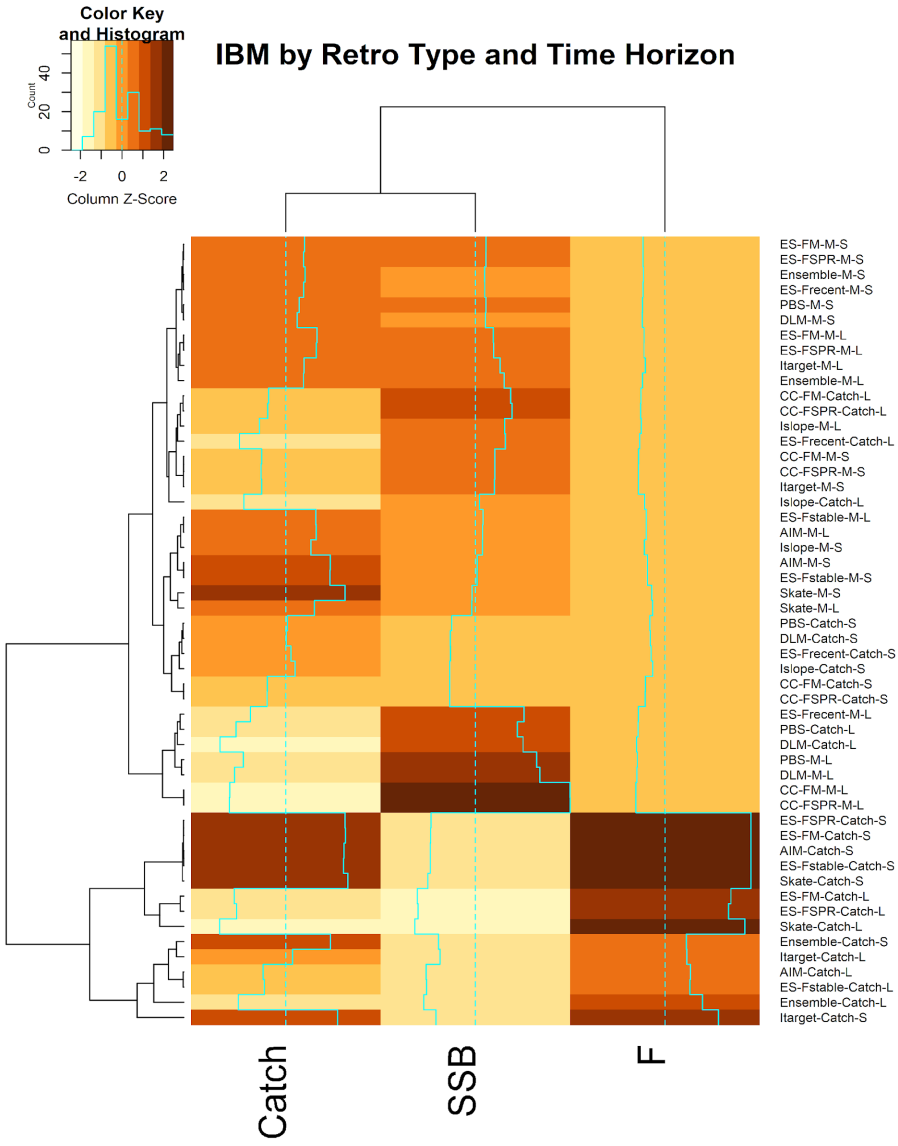


Figure A6.450. Heatmap of median values for spawning stock biomass (SSB)/maximum sustainable yield (SSB_{MSY}), fishing mortality rate (F)/F_{MSY}, and catch/MSY by index-based method (IBM), source of retrospective pattern (C=catch, M=natural mortality), and time horizon (short- or long-term summary, "S" or "L"). The cells are colored according to where they fall in the normalized distribution of all values summarized in the heatmap, and location in that distribution is provided by the key at the top left. A cyan histogram in the key indicates distribution of the data. The cyan lines within the heatmap indicate position of that cell relative to the mean. Values for each cell in the heatmap can be found in the associated table (rows in the same order as the heatmap).

Procedures for Issuing Manuscripts in the Northeast Fisheries Science Center Reference Document (CRD) and the Technical Memorandum (TM) Series

The mission of NOAA's National Marine Fisheries Service (NMFS) is "stewardship of the nation's ocean resources and their habitat." As the research arm of the NMFS's Greater Atlantic Region, the Northeast Fisheries Science Center (NEFSC) supports the NMFS's mission by "conducting ecosystem-based research and assessments of living marine resources, with a focus on the Northeast Shelf, to promote the recovery and long-term sustainability of these resources and to generate social and economic opportunities and benefits from their use." Results of NEFSC research are largely reported in primary scientific media (e.g., anonymously peer-reviewed scientific journals). However, to assist itself in providing data, information, and advice to its constituents, the NEFSC occasionally releases its results in its own series.

NOAA Technical Memorandum NMFS-NE – This series is issued irregularly. The series typically includes: data reports of long-term field or lab studies of important species or habitats; synthesis reports for important species or habitats; annual reports of overall assessment or monitoring programs; manuals describing program-wide surveying or experimental techniques; literature surveys of important species or habitat topics; proceedings and collected papers of scientific meetings; and indexed and/or annotated bibliographies. All issues receive internal scientific review, and most issues receive technical and copy editing.

Northeast Fisheries Science Center Reference Document – This series is issued irregularly. The series typically includes: data reports on field and lab studies; progress reports on experiments, monitoring, and assessments; background papers for, collected abstracts of, and/or summary reports of scientific meetings; and simple bibliographies. Issues receive internal scientific review, and most issues receive copy editing.

CLEARANCE

All manuscripts submitted for issuance as CRDs must have cleared the NEFSC's manuscript/abstract/webpage review process. If your manuscript includes material from another work which has been copyrighted, you will need to work with the NEFSC's Editorial Office to arrange for permission to use that material by securing release signatures on the "NEFSC Use-of-Copyrighted-Work Permission Form."

For more information, NEFSC authors should see the NEFSC's online publication policy manual, "Manuscript/Abstract/Webpage Preparation, Review, & Dissemination: NEFSC Author's Guide to Policy, Process, and Procedure."

STYLE

The CRD series is obligated to conform with the style contained in the current edition of the United States Government Printing Office Style Manual; however, that style manual is silent on many

aspects of scientific manuscripts. The CRD series relies more on the CSE Style Manual. Manuscripts should be prepared to conform with both of these style manuals.

The CRD series uses the Integrated Taxonomic Information System, the American Fisheries Society's guides, and the Society for Marine Mammalogy's guide for verifying scientific species names.

For in-text citations, use the name-date system. A special effort should be made to ensure all necessary bibliographic information is included in the list of references cited. Personal communications must include the date, full name, and full mailing address of the contact.

PREPARATION

Once your document has cleared the review process, the Editorial Office will contact you with publication needs—for example, revised text (if necessary) and separate digital figures and tables if they are embedded in the document. Materials may be submitted to the Editorial Office as email attachments or intranet downloads. Text files should be in Microsoft Word, tables may be in Word or Excel, and graphics files may be in a variety of formats (JPG, GIF, Excel, PowerPoint, etc.).

PRODUCTION AND DISTRIBUTION

The Editorial Office will perform a copy edit of the document and may request further revisions. The Editorial Office will develop the inside and outside front covers, the inside and outside back covers, and the title and bibliographic control pages of the document.

Once the CRD is ready, the Editorial Office will contact you to review it and submit corrections or changes before the document is posted online. A number of organizations and individuals in the Northeast Region will be notified by e-mail of the availability of the document online.