

Evaluating procedures for updating catch advice between stock assessments of reef fishes with management strategy evaluation

Nikolai Klibansky, Shannon Calay, Rob Cheshire,
Cassidy Peterson, Kyle Shertzer, Erik Williams, and Matt Vincent

Southeast Fisheries Science Center,
National Marine Fisheries Service, NOAA,
101 Pivers Island Road, Beaufort, North Carolina, 28516, USA

April 15, 2022

Abstract

We conducted a management strategy evaluation (MSE) to investigate how management procedures that adjust catch advice between stock assessments performed compared with existing management procedures. We built operating models (OM) for four reef fish species from the US Southeast Atlantic, based on recent stock assessments including Black Sea Bass, Red Porgy, Snowy Grouper, and Vermilion Snapper. These OM contained parameters and data specific to each stock, associated fisheries, and the sampling programs that monitor them. The analysis assumed efficient implementation of management, such that observed catch was equal to total allowable catch (TAC). Our analysis focused on a base scenario intended to most closely characterize the reality of each stock. We also developed multiple alternative scenarios to investigate the sensitivity of the analysis to deviations from the base configuration. A set of management procedures (MP) was applied independently in closed loop simulation for each species and scenario, with many replicate runs. The MP varied in terms of how often stock assessments were conducted (every 1, 5, or 10 years), and how catch advice (i.e. TAC) was adjusted between stock assessments. Between assessments, TACs were either fixed, adjusted based on projections, or adjusted based on a reference index of abundance. Results varied among species and scenarios, but generally showed that healthy stock and fishery status ($SSB > SSB_{MSY}$ and $F < F_{MSY}$) and comparable levels of total catch could be maintained with stock assessments conducted every 1, 5 or 10 years, whether TACs were fixed, projected, or adjusted based on indices of abundance. But these management procedures vary in terms of average annual variability in yield (AAVY) which was highest when TACs were adjusted based on indices of abundance and lowest when TACs were fixed between assessments.

INTRODUCTION

Motivation

Stock assessments are highly resource intensive processes. Recent NMFS guidance has prioritized increasing assessment throughput and reducing assessment frequency where appropriate for managed species (ICES 2012; Methot 2015; Lynch et al. 2018). Reduced stock assessment frequency would free up assessment scientists to conduct research to advance fisheries management and prioritize unassessed stocks.

One such mechanism of allowing reduced stock assessment frequency while still providing continuous adaptive management is through the use of interim assessment approaches (e.g., Huynh et al. 2020). Interim assessment approaches use an indicator of stock abundance (e.g., index of abundance) to adjust total allowable catch (TAC) between full stock assessments. Interim assessments would still require updated data and analytical input from assessment scientists, but the computational load and time commitment would be substantially diminished. Interim assessments have been found to be conceptually sound (Huynh et al. 2020), and here we seek to identify whether these methods are suitable for stocks in the South Atlantic.

MSE Background

The current stock assessment paradigm and the interim assessment approach are each examples of a management procedure. A management procedure, or management strategy, is a formal rule that defines how fisheries are managed and includes (1) a clearly defined data-generating protocol, (2) an estimating model that describes the current state or condition of the stock (e.g., stock assessment), (3) a predefined catch (harvest) control rule

that describes how catch advice should be adjusted given the state of the stock, and (4) an implementation rule which describes how the catch advice will be applied (Sainsbury et al. 2000; Punt 2010). Management procedures are designed to be adaptive to changes in stock biomass or trajectory (e.g., as estimated biomass declines, allowable catch declines; Edwards 2016). Consider how the current assessment paradigm can be presented within the management procedure framework, for example, as it includes: (1) a data-collection enterprise to collect abundance, biological, and catch information from the stock; (2) the stock assessment, or estimating, model which uses scientifically collected data to provide an estimate of stock status and biological reference points, (3) catch control rule, which varies with each FMC, but examples include the P^* approach (Shertzer et al. 2008) or the 40-10 rule (PFMC 2020), and (4) the implementation of the catch control rule-defined management action, often in the form of TACs.

Because in situ scientific testing of alternate management procedures is logistically unfeasible in practice, we test management procedures using management strategy evaluation (Butterworth and Punt 1999; De Oliveira et al. 2008). MSE is a framework for testing candidate management procedures across a range of uncertainties using closed-loop simulation (Fig. 1; Holland 2010; Punt et al. 2016).

Management procedure performance is measured based on pre-specified performance metrics, which are designed to reflect the management objectives of the system. Typically management objectives include increasing total allowable catch and reducing risk to the resource. When management objectives are not already defined, as in a Fishery Management Plan, stakeholder involvement is often necessary to clarify the goals of the fishery. Understandably, not all performance metrics can be simultaneously maximized (e.g., increasing TAC increases risk to the resource). MSE therefore serves to lay bare the societal

trade offs associated with fisheries management (Punt 2017).

An operating model is created that represents the 'true' system dynamics, and the full management procedure is applied to the stock(s) of interest simulated in the operating model. However, there are many assumptions that must be made when generating a simulation of a stock and fishery. To ensure that none of these assumptions or decision-points impact the results, particularly those that are most uncertain, multiple operating models can be generated representing different scenarios (Butterworth and Punt 1999; De Oliveira et al. 2008; Punt et al. 2016). For example, if natural mortality is uncertain, multiple scenarios can be built with different realistic assumptions of natural mortality (high, medium, and low).

We then measure the effects of the management procedure on the 'true' populations simulated by the operating models. If the management procedure cannot perform well across all operating models (e.g., natural mortality scenarios), then it likely is not a good candidate for implementation, because we are unlikely to know the real natural mortality rate in practice. The presence of multiple operating models serves as a test of the management procedure against many possible realities, ensuring that the chosen management procedure is robust to the uncertainties that we have in the system (Butterworth and Punt 1999; Rademeyer et al. 2007; Punt et al. 2016). These uncertainties can extend to future projections, accordingly monitoring how management procedures would perform in the face of nonstationarity.

Interim assessment approach

While setting interim-period catches based on assessment model forecasts can yield improved performance over fixed catches, Huynh et al. (2020) noted that forecasts are made under assumptions of future stock dynamics and are not responding to true observed changes in

stock size. The interim approach is a modification of an empirical, indicator-based management procedure, in which catch advice is adjusted based on an indicator of stock abundance (e.g., index of abundance; Geromont and Butterworth (2015); Carruthers et al. (2016)). Within the interim approach, TACs set by the stock assessment model are adjusted in interim years based on the behavior of the stock indicator. Unlike traditional stock assessment-based TACs, interim assessment TAC adjustments may be implemented immediately with no time-lag (Huynh et al. 2020).

Huynh et al. (2020) used MSE to test the interim assessment approach across 3 species of differing life histories (short-, moderate-, long-lived), each with 6 unique operating model scenarios (Base, Hyperstable, Hyperdeplete, Depleted, Lightly fished, Episodic M), using the openMSE software suite [openMSE, Hordyk et al. (2021); DLMtool, Carruthers et al. (2021); MSEtool, Hordyk et al. (2022); SAMtool, Huynh et al. (2022)]. Authors found that the use of an interim assessment approach markedly improved management performance over using a fixed TAC between assessments and more closely approximated the management performance of using assessment projections to specify interim period catches. Furthermore, management resulting from the interim assessment approach was not adversely impacted by increased interim period length and hyperstable/deplete scenarios in which the assumption that the relative abundance index is proportional to stock biomass is violated. Ultimately, the interim approach reduced the need for frequent stock assessments, potentially freeing up resources to focus on other stocks or research questions.

Study Objectives

In this study, we sought to replicate the Huynh et al. (2020) approach by applying it specifically to South Atlantic fishes based on information from the most recent stock

assessments. Representative species analyzed include Black Sea Bass, Red Porgy, Vermillion Snapper, and Snowy Grouper. We used species-specific biological parameters and were able to more accurately describe the data-generating, fishery, and stock assessment dynamics for these species within the openMSE framework, including delays in data collection and management implementation (Shertzer and Prager 2007) and specification of different selectivity functions for catch and abundance indices. Simulations were further specified with empirically estimated levels of observation and process uncertainty. The purpose of this study was to explore whether interim assessment approaches would be suitable for South Atlantic fishes.

MATERIALS AND METHODS

Operating models

We used the most recent stock assessments from US Southeast US Atlantic for Red Porgy (SEDAR 2020), Black Sea Bass (SEDAR 2018*b*), Vermilion Snapper (SEDAR 2018*a*), and Snowy Grouper (SEDAR 2021). Elements of the stock assessments are stored as data sets in the R package **bamExtras** (<https://github.com/nikolaifish/bamExtras>). Management strategy evaluation was conducted with the R package **openMSE** (Hordyk et al. 2021), which includes three sub-packages **DLMtool**, **MSEtool**, and **SAMtool** (Carruthers et al. 2021; Hordyk et al. 2022; Huynh et al. 2022). This package uses a set of data objects with very specific list-like structures (S4; <http://adv-r.had.co.nz/S4.html>), such as operating model objects (OM). Such objects are made up of a set of named 'slots', which contain specific types of information with particular dimensions, to characterize main components of an MSE. Input and output from the Beaufort Assessment Model (BAM) is configured as a list which has a

specific structure, and is stored in an `rdat` file. The most recent `rdat` files are also stored in `bamExtras` as objects (e.g. `rdat_BlackSeaBass`). In order to the transfer information from BAM `rdat` objects to `openMSE` objects, we wrote a set of functions and constructed a new R package, `bamMSE` (<https://github.com/nikolaifish/bamMSE>).

Operating models in `openMSE` (`OM` class `S4` objects) are constructed of four sub-objects: `Stock`, `Fleet`, `Obs`, and `Imp` class objects. Values stored in `Stock` objects describe a fish stock, `Fleet` objects characterize a fishing fleet that fishes that stock, `Obs` objects contain parameters describing how the simulated stock and fleet are observed (e.g. bias and error of catch or relative abundance data), and the `Imp` objects allow the user to set how well managers adhere to management recommendations (i.e. implementation error). A fifth `openMSE` object class `Data`, is used to store various types of data, including data sets typically used in fitting stock assessment models (e.g. indices of abundance, catch, and age-composition time series). Functions in `bamMSE` convert `rdat` objects to `Stock`, `Fleet`, `Obs`, and `Data` objects (`rdat2Stock`, `rdat2Fleet`, `rdat2Obs`, and `rdat2Data`).

We used the function `MSEtool::Assess2OM` to convert BAM results to `OM` objects, with the help of a wrapper function `bamAssess2OM` (not yet added to `bamMSE`). While the values of the steepness parameter h , and unfished recruitment (R_0) for the Beverton-Holt stock-recruit relationship are directly passed from BAM to `Assess2OM`, most other inputs had to be modified from BAM. Since `openMSE` requires the first age in age-structured populations to be age-0 and most BAM assessments start with age-1, age-based data from BAM was linearly extrapolated from age-1 to age-0 before being passed to `Assess2OM`. Such data include several three dimensional numeric arrays with dimensions simulation, age, and year (within recent assessment period): fish weight, fish length, proportion mature, numbers of fish, total fishing mortality rate (F ; including dead discards) , and natural mortality rate (M).

Maturity-at-age is used to compute spawner biomass using methods followed in BAM assessments. For gonochoristic species (separate male and female sexes; Vermilion Snapper) maturity-at-age was based on females. For protogynous hermaphrodite species (females transition to male as they age; Black Sea Bass, Red Porgy, and Snowy Grouper) maturity-at-age was based on the proportion of fish of either sex mature-at-age.

We ran `rdat2Data` on the `rdat` objects, and passed indices of abundance, their associated CVs, and selectivity-at-age to an `openMSE` `cpars` list object. In `openMSE` a `cpars` list object is used to pass custom parameters, typically time-varying, to an operating model. The `rdat2Data` function assumes that indices based on surveys or recreational fisheries are in numbers and indices based on commercial fisheries are in pounds. Only indices available in the terminal year of the assessments were passed to `cpars` and the `OM` objects (Fig. 2). Estimates of growth parameters (K , L_∞ , and t_0) from BAM were also passed to the `OM` objects. Selectivity-at-age of removals is estimated annually during the empirical assessment period by `Assess2OM`, and the average of the last three years of the assessment period was passed to `cpars` and the `OM` objects to be used in the projection years.

In most cases, empirical values provided to the `OM` objects are used to determine distributions of parameters, data, or residuals. These distributions are then used by `openMSE` when simulating values in the projection years. For example, the number of age samples (numbers of otoliths) available to management procedures during the projection period are determined by drawing from a distribution based on how many age samples were actually available in the most recent SEDAR assessment of that species. However, the time series of residuals for indices of abundance, were generated ahead of time and added to the `OM` objects. Residuals were sampled from a lognormal distribution based on index CVs from the most recent SEDAR assessment. In the base and most other scenarios, life history parameters

were constant over time in the simulated populations. Maximum F was limited in the simulations by rounding the maximum observed F up to the next integer value. This limit prevents catch limits from being set unreasonably high.

Management procedures

The entire MSE was run repeatedly and independently, applying one of nine different management procedures (MP). In all of these MP, age-structured stock assessments were conducted at a frequency of every 1, 5, or 10 yr. For all MPs, in an assessment year, reference points were estimated (e.g. MSY , F_{MSY}), and a simple harvest control rule was applied to determine the TAC for the following year. The TACs were computed from the assessment output using the `SAMtool:HCR_MSY` function which projects abundance one year beyond the terminal year of the assessment, and applies the catch equation used in the assessment and the target level of fishing mortality (F_{target}). In the base scenario, F_{target} was set equal to F_{MSY} from the simulated assessment. We also limited the maximum F to $F_{max} = 3$ to avoid simulating unreasonably high catch levels. Between assessments, catch was set in different ways.

In the fixed TAC MPs [SCA (1), SCA (5), and SCA (10)] stock assessments were conducted every 1, 5, or 10 yr respectively, and the harvest control rule was applied to determine the TAC. Between assessments, the TAC was fixed.

In the projection MPs, [Projection (5) and Projection (10)], stock assessments were conducted every 5 or 10 yr, respectively. In these MPs a projection analysis is simulated in order to set the TACs for the interval between assessments. The population is projected forward for the interval between assessments F_{target} is applied to compute catch. These catches are then used as TACs in the simulated management during the interval. These

TACs will tend to exhibit a temporal trend since projected abundance is not usually constant.

In the average index interim MPs [Average Index (5) and Average Index (10)] stock assessments were conducted every 5, or 10 yr respectively. In an assessment year, the TAC is computed as in other MPs, and is treated as a reference catch (C_{ref}) in the interim procedure. The value of a selected index of abundance is also recorded and treated as a reference value (I_{ref}). Between assessments, a TAC in year $y + 1$ is set by according to Eq. 1:

$$\text{TAC}_{y+1} = C_{\text{ref}} \frac{\tilde{I}_{\text{recent}}}{I_{\text{ref}}} \quad (1)$$

This calculation scales C_{ref} by the ratio of an average of the three most recent years of the abundance index ($\tilde{I}_{\text{recent}}$) and the reference year (I_{ref}). Using the average of recent years of the index instead of a single recent value effectively smooths the index. This serves to avoid a noisy index leading to highly varying TACs.

In the buffered index interim MPs, Buffered Index (5) and Buffered Index (10), stock assessments were conducted every 5, or 10 yr respectively. Between assessments, a TAC in year $y + 1$ is set by according to Eq. 2:

$$\text{TAC}_{y+1} = C_{\text{ref}} \frac{I_y + \hat{\sigma}}{I_{\text{ref}} + \hat{\sigma}} \quad (2)$$

where I_y is the value of the index in the most recent year and $\hat{\sigma}$ is the standard deviation of the index. Including $\hat{\sigma}$ in this calculation assures that the noisier the index, the smaller the adjustment to the TAC, limiting variation in TACs.

Statistical catch-at-age model configuration

Management procedures that ran statistical catch-at-age models utilized the `SAMtool::SCA` function. The model run in this function was configured as similar to BAM as the function would allow. The configuration was specific to each species' operating model where applicable.

The range of ages modeled in the assessments is from age-0 to the maximum age modeled in the BAM assessment for each species. Life history is age-structured such that, length, weight, maturity, and natural mortality vary with age. Years of data included in the time series span from the earliest historical year to two years prior to the current year (i.e. assessment year). Management changes go into effect the year following the assessment. Effectively, the last two years of available data are excluded from the stock assessment, to simulate a typical data lag observed in SEDAR assessments. For interim procedures, we assumed that management would be put in place the next year, and therefore did not impose any additional data lag. Applying data lags is not a standard option in `openMSE` so the lag option was written into wrapper functions for `SAMtool` functions `SCA`, `make_MP`, `make_interim_MP`, and `make_projection_MP`, for this project.

The `SCA` models included one combined set of removals (landings and discards). Although most of the empirical BAM assessments include multiple fleets of removals with landings and discards separate, this is not an option when using `runMSE`¹. The model also fits to one or more indices of abundance, depending on the species, and one set of simulated catch-at-age data corresponding to the landings.

Selectivity-at-age associated with removals was estimated within `SCA` with a logistic

¹The flexibility to add multiple landings fleets has recently been added to `openMSE` using the `multiMSE` function, but would require a nontrivial amount of coding to incorporate into the current analysis.

selectivity function. Selectivity of the indices was provided to **SCA** as fixed information.

The **SCA** model used a Beverton-Holt stock-recruit function, the Baranov catch equation, and a multinomial distribution for fitting catch-at-age information. Several parameters are fixed by default in **SCA** including Beverton-Holt steepness, equilibrium F , CV associated with catch, and the standard deviation of recruitment residuals (i.e. "rec sigma"). Recruitment deviations were estimated for one generation prior to the first year that age-composition data were available. Toward the end of each assessment, recruitment deviations were not estimated for several years and recruitment was predicted from the mean value of the stock-recruit relationship. The number of years for which recruitment deviations were not estimated was based on an option within the **SCA** function (`late_dev="comp50"`) which specifies it as the: "number of ages (smaller than the mode) for which the catch-at-age matrix has less than half the abundance than that at the mode" (**SAMtool::SCA** help file).

Scenarios

The configuration of the MSE described above was considered the base scenario. Various modifications to the configuration were made to create other scenarios described below.

Many other scenarios were explored but did not change the conclusions of our analysis, so they are not included in the current report.

1. Index high CV (ucvhi): Uncertainty (coefficient of variation; CV) in primary index of abundance of abundance is twice the CV of the base scenario
2. Index low CV (ucvlo): Uncertainty (CV) in primary index of abundance of abundance is half the CV of the base scenario
3. Index is biased (ubias): Primary index of abundance increasingly underestimates population size (i.e. decreasing catchability)
4. High TAC (tachi): $TAC = 1.25MSY$ ($F_{target} = 1.25F_{MSY}$ in harvest control rule)
5. Low TAC (taclo): $TAC = 0.75MSY$ ($F_{target} = 0.75F_{MSY}$ in harvest control rule)

For each management procedure, we ran the replicate simulations 250 times for the base scenario and 50 times for each alternative scenario. In preliminary analysis of the base scenario, we found that results were approximately the same with 50 runs as 250 runs, although output plots were smoother when more runs were completed. However, output objects from 250 are rather large (200-400 Mb) and become cumbersome to work with.

Evaluating performance

To evaluate the relative performance of the MPs, we consider a set of reference points and performance metrics. We computed total catch, relative catch (C/MSY), stock status (SSB/SSB_{MSY}), fishery status (F/F_{MSY}), and the average annual variability in yield (AAVY). The AAVY is the absolved relative difference in catch between two consecutive years, calculated as in Eq. 3:

$$AAVY = \frac{\sum_{y=1}^Y \sqrt{\left(\frac{C_y - C_{y+1}}{C_{y+1}}\right)^2}}{Y} \quad (3)$$

where C_y and C_{y+1} are catches in years y and $y + 1$, and Y is the total number of years in the projection period.

We also computed several performance metrics, focusing on long term values, measured during the last 10 years of the projection period. These included average long term yield, the probability $P()$ of healthy stock status $P(SSB > SSB_{MSY})$, the probability of not overfishing $P(F < F_{MSY})$, and the probability of stable yield $P(AAVY < 20\%)$.

RESULTS

Base scenario

Response variables for the base scenario varied among time periods over the projection years for all metrics. Stock status (SSB/SSB_{MSY}) increased over time among all nine SCA-based MPs for all four species, but was similar between MPs during the same time interval (Fig. 3). Fishery status (F/F_{MSY}) was somewhat more variable within and between MPs, but still largely consistent among SCA-based MPs, for all species (Fig. 4). Total catch was also similar among SCA-based MPs for all species, for similar time periods (Fig. 5). It was relatively less consistent for Black Sea Bass, but variability in total catch within MPs was also more variable for that stock. Average annual variability in yield (AAVY) showed more substantial variation among MPs (Fig. 6). For Vermilion Snapper and Snowy Grouper, the fixed TAC and projection MPs showed substantially lower AAVY than the interim MPs. Red Porgy showed a similar pattern, although AAVY for the Average Index (5) and Average Index (10) was relatively lower. For Black Sea Bass, there was relatively little difference in AAVY among MPs compared to the other species.

Time series of relative catch (catch/MSY), averaged across simulation runs, are generally most stable when assessments are run annually, and show the largest interannual variations when projection MPs are used (Figs. 7 and 8). For the projection MPs, the difference in catch between consecutive years is actually modest for most years, but tends to be very large between a stock assessment year and the next year, leading to periodic spikes. For Black Sea Bass, the fixed TAC MP was often well above or below the other MPs with a 10 yr assessment interval, but catches were otherwise very steady because TACs were fixed

during the intervals. The index-based approaches tended to produce similar catches which were more consistent than the projection and fixed TAC MPs. Relative catches were all more similar to the annual assessment MP when the assessment interval was 5 yrs. Patterns in F/F_{MSY} time series reflected dynamics similar to the catch. However the fluctuations in the projection MP were less pronounced (Figs. 9 and 10) and tended to be lower than other MPs, representing a more conservative strategy. Despite differences in catch and relative F among MPs, trends in $\text{SSB}/\text{SSB}_{\text{MSY}}$ were similar among all MPs for Red Porgy and Snowy Grouper. For Vermilion Snapper only the projection MP deviated from the others, with $\text{SSB}/\text{SSB}_{\text{MSY}}$ tending to be slightly higher (Figs. 11 and 12). Trends in $\text{SSB}/\text{SSB}_{\text{MSY}}$ for Black Sea Bass varied among MPs with a 10 yr assessment interval, with the Fixed TAC MP showing the widest fluctuations. But with a 5 year assessment interval, $\text{SSB}/\text{SSB}_{\text{MSY}}$ trends were fairly similar among MPs for all species including Black Sea Bass.

Trade off plots display bivariate relationship between two performance metrics for multiple management procedures, to identify differences between MPs in performance-metric space. We considered several important trade offs to look for differences in performance among MPs, plotting mean values across simulation runs, for each MP. In particular we looked at trade offs between yield (catch), variability in yield, or fishery status, versus stock status. We limited data used to produce these plots to the final 10 years of the projection, to look at long term values.

Trade off plots of probability that $\text{AAVY} < 20\%$ versus the probability that $\text{SSB} > \text{SSB}_{\text{MSY}}$ (Fig. 13), the probability that $F < F_{\text{MSY}}$ (Fig. 14), and mean relative long term yield (Fig. 15), show little differentiation among MPs along the vertical axis. The MPs mostly differ in the probability of that $\text{AAVY} < 20\%$. As suggested in other plots, SCA (5),

and SCA (10) produce more consistent yields, which is indicated in trade off plots as having higher probabilities that $AAVY < 20\%$. Performance of other MPs in terms of probability of $AAVY < 20\%$ varied by species. Trade off plots of mean relative long term yield versus the probability that $SSB > SSB_{MSY}$ (Fig. 16) and probability that $F < F_{MSY}$ versus probability that $SSB > SSB_{MSY}$ (Fig. 17) do not indicate any consistent differentiation among MPs.

We also produced phase plots of F/F_{MSY} versus SSB/SSB_{MSY} (Fig. 18) including median values among all simulation runs, as well as points for each run. Again we limited data to the final 10 years of the projection periods. Phase plots also do not indicate any consistent differentiation among MPs. However the projection MP resulted in slightly higher stock status for Vermilion Snapper and Red Porgy and the fixed TAC MP resulted in slightly higher F -status in Black Sea Bass.

Alternative scenarios

Although results for alternative scenarios differed from the base in various ways, the overall results are the same (Figs. S1-S15). Management procedures perform similarly except for variation in variability of catch. One exception to this generality is for Red Porgy in the scenario where the harvest control rule set the TAC to $1.25 \times MSY$. Under this scenario, the Red Porgy could not recover from its already depleted state and only the annual assessment MP was able to keep the stock from going completely crashing in more than 50% of simulation runs (Fig. S15).

DISCUSSION

Overall, our analysis suggested that all of the management procedures we considered resulted in similar trends in SSB/SSB_{MSY} for a given stock and scenario. With a 10 yr assessment interval the Black Sea Bass stock, showed the most variability among MPs, and only the annual assessment MP produced stable SSB/SSB_{MSY} until late in the projection period; but with 5 yr assessment intervals, all MPs produced similar results. In this stock the projection MP tended to be a bit conservative, resulting in fairly high SSB/SSB_{MSY} , but even this MP could be tuned by applying a more aggressive harvest control rule if that was desired.

Even this variation in SSB/SSB_{MSY} , evident in the central tendency of the time series, is less apparent when comparing distributions in boxplots by time period. This is generally true for F/F_{MSY} as well. Distributions of total catch are also similar among MPs, comparing the same time periods.

The performance metric that really separates the MPs is variability in yield. The simplest, fixed TAC, MPs correspond to the lowest variability in yield, followed by the projection MPs which set TACs that change between assessments but follow a simple trajectory. For the index-based approaches, TACs vary in more complex ways between assessments based on changes in abundance, leading to greater variability in catch.

An ideal interim management procedure would allow for stock assessments to be conducted less frequently while still achieving management goals. While we found index-based interim MPs performed somewhat better than fixed TAC MPs when comparing time series of SSB/SSB_{MSY} at the assessment frequency, there is no clear improvement when comparing interim procedures conducted with an assessment frequency of 10 yrs and fixed

TAC MPs with a assessment frequency of 5 yrs.

It may be possible to improve upon the index-based MPs, although we did not investigate this. In particular, the computation used to adjust the TACs can be easily modified to decrease variation in TACs and presumably catch. But it's likely that this modification would trade off with the ability of the MP to maintain desirable stock status. Such an MP might be expected to occupy a location in performance space intermediate between the current index-based MPs and the fixed TAC MPs.

In order for a new MP to be favored over an existing one, it would likely need to exhibit a more favorable ratio of performance to effort. We have focused on assessing performance, but can make some basic conjectures about effort that might be useful to guiding decisions about whether or not interim procedures are favorable. If we assign 1 unit of scientific effort to conducting a stock assessment (including analysis and process; e.g. SEDAR), then conducting one assessment every 5 years would be 0.2 units per year (upy) and every 10 years would be 0.1 upy. If an index-based TAC adjustment were to be applied every year, it would be completed nine times in 10 years. Therefore the relative effort of completing such an analysis was equal to $0.1/9$ upy, it would take just as much effort as doing one additional assessment, and therefore the effort would be equivalent to doing assessments every 5 years. It is likely that the analytical effort of updating an index of abundance and applying the adjustment procedure is substantially less than the effort that goes into an entire assessment, and presumably the assessment process would be simpler. But it seems less clear that the effort would be less than $\frac{1}{9}$ th as much.

In addition, we should consider the management effort required to review and adjust catch advice, similarly assigning 1 unit of management effort per assessment and 0.2 upy

when conducting assessments every 5 years. But in this case, the management effort required to review and adjust catch advice based on an updated index is probably similar to the management effort of a full assessment. Therefore the management effort of applying an index-based MP with an assessment interval of 10 years would be about 1 upy, or about $5\times$ the management effort of applying a fixed TAC or projection MP with an assessment interval of 5 years.

This consideration of effort is clearly hypothetical, but attempts to outline some of the costs that an index-based MP might have to overcome to be preferred over the status quo approach. In the present analysis we found limited improvement of the index-based MPs over status quo approaches in terms of maintaining healthy stocks, and poorer performance in terms of more variable catches. So if the net performance of the index-based MPs could even be considered similar to the status quo, the actual effort associated with conducting and implementing an index-based MP will have to be very low for the approach to be preferred.

ACKNOWLEDGEMENTS

We extend special thanks to Quang Huynh for answering many questions about the `openMSE` software and writing new features and functions into the software in response to our inquiries. Thanks also to Adrian Hordyk and Tom Carruthers at Blue Matter for additional help with `openMSE`. Thank you to Kate Siegfried and Lew Coggins for thoughtful suggestions. Thanks to the Fischmeister 5000 uber computer for running computations day and night, without complaint.

LITERATURE CITED

- Butterworth, D., and A. Punt. 1999. Experiences in the evaluation and implementation of management procedures. *ICES Journal of Marine Science* 56:985–998.
- Carruthers, T., Q. Huynh, and A. Hordyk, 2021. DLMtool: Data-Limited Methods Toolkit. URL <https://CRAN.R-project.org/package=DLMtool>.
- Carruthers, T. R., L. T. Kell, D. D. Butterworth, M. N. Maunder, H. F. Geromont, C. Walters, M. K. McAllister, R. Hillary, P. Levontin, and T. Kitakado. 2016. Performance review of simple management procedures. *ICES Journal of Marine Science* 73:464–482.
- De Oliveira, J. A., L. T. Kell, A. E. Punt, B. A. Roel, and D. S. Butterworth. 2008. Managing without best predictions: the management strategy evaluation framework. *Advances in fisheries science* 50:104–134.
- Edwards, C. T. 2016. Feedback control and adaptive management in fisheries. *Management Science in Fisheries: An introduction to simulation-based methods* page 16.
- Geromont, H., and D. Butterworth. 2015. Generic management procedures for data-poor fisheries: forecasting with few data. *ICES Journal of Marine Science* 72:251–261.
- Holland, D. S. 2010. Management strategy evaluation and management procedures: tools for rebuilding and sustaining fisheries .
- Hordyk, A., Q. Huynh, and T. Carruthers, 2021. openMSE: Easily Install and Load the 'openMSE' Packages. URL <https://CRAN.R-project.org/package=openMSE>.
- Hordyk, A., Q. Huynh, and T. Carruthers, 2022. MSEtool: Management Strategy Evaluation Toolkit. URL <https://CRAN.R-project.org/package=MSEtool>.
- Huynh, Q., T. Carruthers, and A. Hordyk, 2022. SAMtool: Stock Assessment Methods Toolkit. URL <https://CRAN.R-project.org/package=SAMtool>.
- Huynh, Q. C., A. R. Hordyk, R. E. Forrest, C. E. Porch, S. C. Anderson, and T. R. Carruthers. 2020. The interim management procedure approach for assessed stocks: Responsive management advice and lower assessment frequency. *Fish and Fisheries* 21:663–679.
- ICES. 2012. Report of the workshop on frequency of assessments (WKFREQ), Copenhagen, Denmark .
- Lynch, P. D., R. D. Methot, and J. S. Link. 2018. Implementing a Next Generation Stock Assessment Enterprise: An Update to the NOAA Fisheries Stock Assessment Improvement Plan. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-183 page 127.
- Methot, R. D. 2015. Prioritizing fish stock assessments. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO152 page 31.

- PFMC. 2020. Pacific coast groundfish fishery management plan for the California, Oregon, and Washington groundfish fishery. Pacific Fishery Management Council, Portland, OR.
- Punt, A. E. 2010. Harvest control rules and fisheries management. *Handbook of marine fisheries conservation and management* pages 582–594.
- Punt, A. E. 2017. Strategic management decision-making in a complex world: quantifying, understanding, and using trade-offs. *ICES Journal of Marine Science* 74:499–510.
- Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. A. De Oliveira, and M. Haddon. 2016. Management strategy evaluation: best practices. *Fish and Fisheries* 17:303–334. URL <http://dx.doi.org/10.1111/faf.12104>.
- Rademeyer, R. A., É. E. Plagányi, and D. S. Butterworth. 2007. Tips and tricks in designing management procedures. *ICES Journal of Marine Science* 64:618–625.
- Sainsbury, K. J., A. E. Punt, and A. D. M. Smith. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. *ICES Journal of Marine Science: Journal du Conseil* 57:731–741.
- SEDAR. 2018*a*. SEDAR 55 - South Atlantic Vermilion Snapper Assessment Report. SEDAR, North Charleston SC page 170. URL <http://sedarweb.org/sedar-55>.
- SEDAR. 2018*b*. SEDAR 56 - South Atlantic Black Seabass Assessment Report. SEDAR, North Charleston SC page 164. URL <http://sedarweb.org/sedar-56>.
- SEDAR. 2020. SEDAR 60 South Atlantic Red Porgy Stock Assessment Report. SEDAR, North Charleston SC page 181. URL <http://sedarweb.org/sedar-60>.
- SEDAR. 2021. SEDAR 36 Update - South Atlantic Snowy Grouper Stock Assessment Report. SEDAR, North Charleston SC page 118. URL <http://sedarweb.org/2020-update-sedar-36-update-assessment-report-south-atlantic-snowy-grouper-revised-january-2021-0>.
- Shertzer, K. W., and M. H. Prager. 2007. Delay in fishery management: diminished yield, longer rebuilding, and increased probability of stock collapse. *ICES Journal of Marine Science* 64:149–159.
- Shertzer, K. W., M. H. Prager, and E. H. Williams. 2008. A probability-based approach to setting annual catch levels. *Fishery Bulletin* 106:225–232.

FIGURES

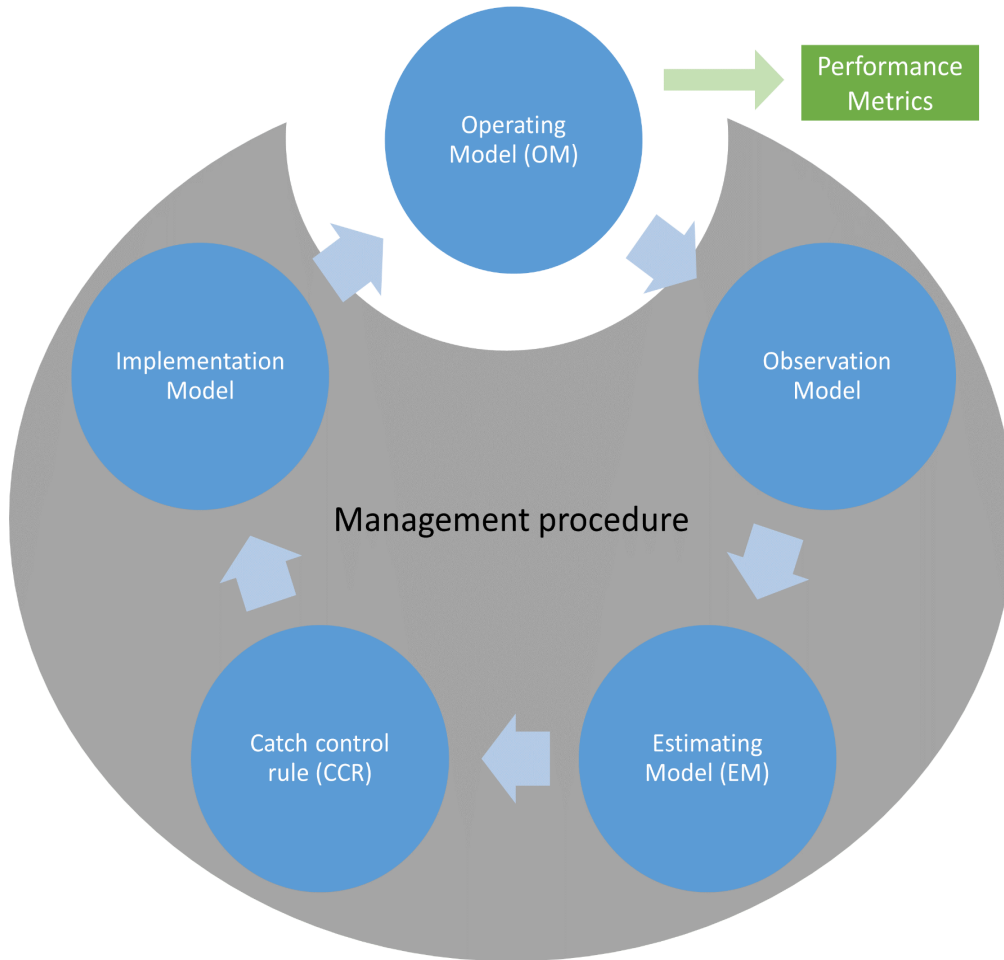


Fig. 1 Diagram of the MSE process. The operating model is the simulated "true" dynamics of the system; the observation, or data-generating, model simulates the process of scientific data collection from the stock and fishery; the estimating model is used to estimate the status of the stock; the catch control rule is a pre-defined decision rule that adaptively adjusts management measures based on the status of the stock; the implementation model implements the management recommendations from the catch control rule back into the operating model with implementation and management uncertainty. Note that all model components highlighted in gray comprise the management procedure and performance metrics (or management procedure performance) is measured from the operating model.

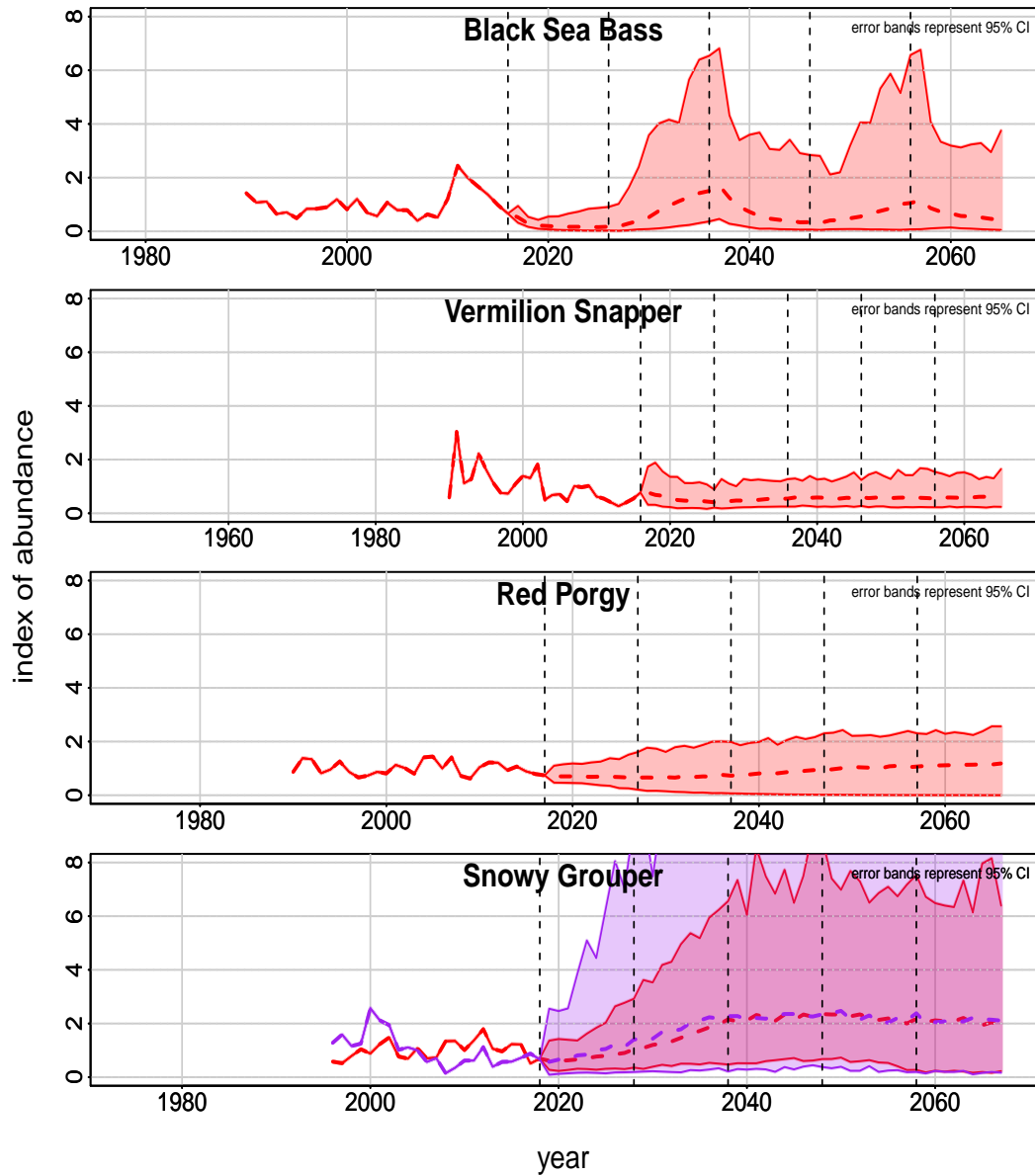


Fig. 2 Indices derived from BAM stock assessments available during the projection period, for the SCA (10) MP in the Base scenario, where stock assessments are conducted every 10 years during the projection period (dashed vertical lines). Shaded areas represent 95% CI for indices among simulation runs.

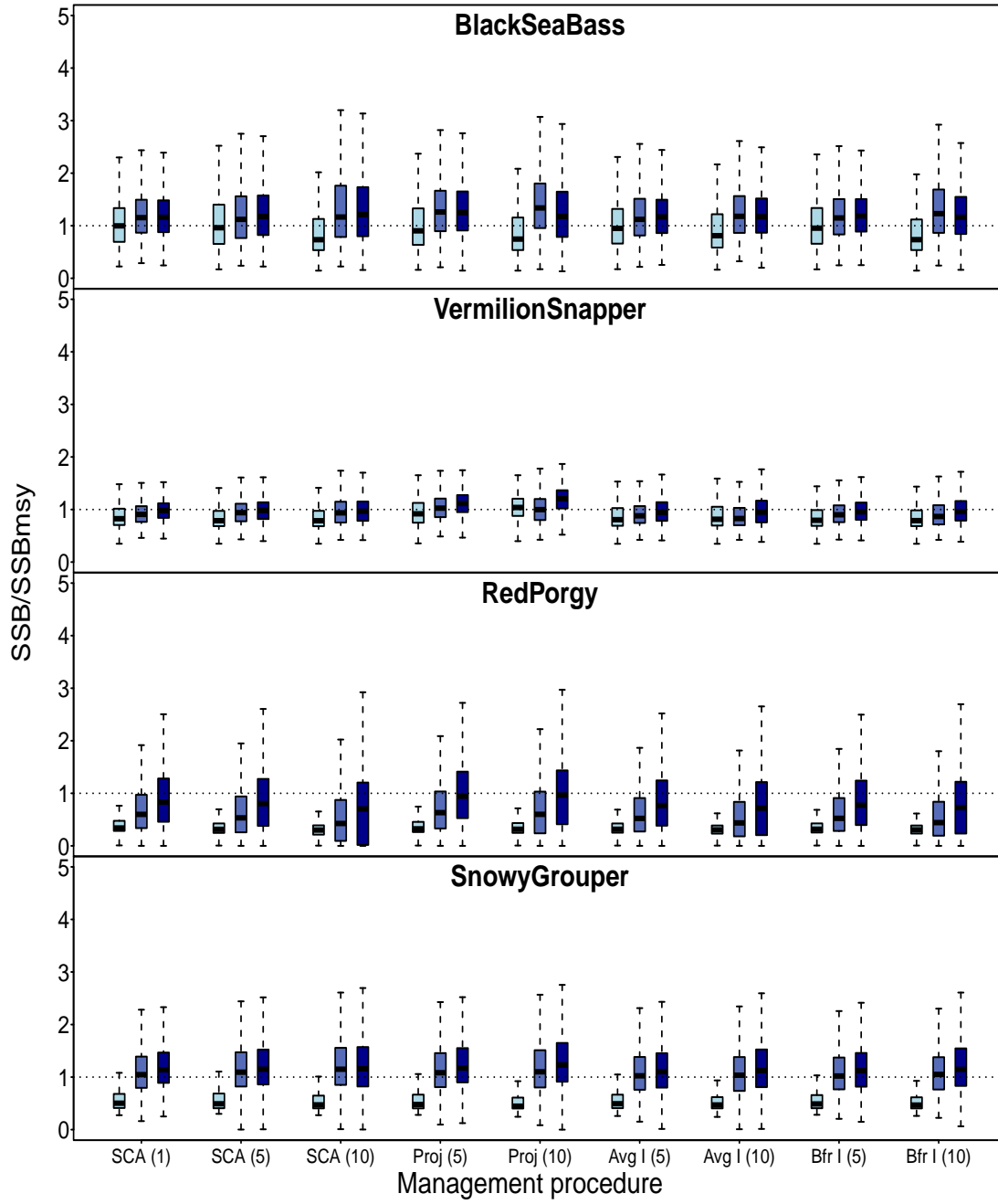


Fig. 3 Box plots of SSB/SSB_{MSY} for the base scenario. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times IQR$.

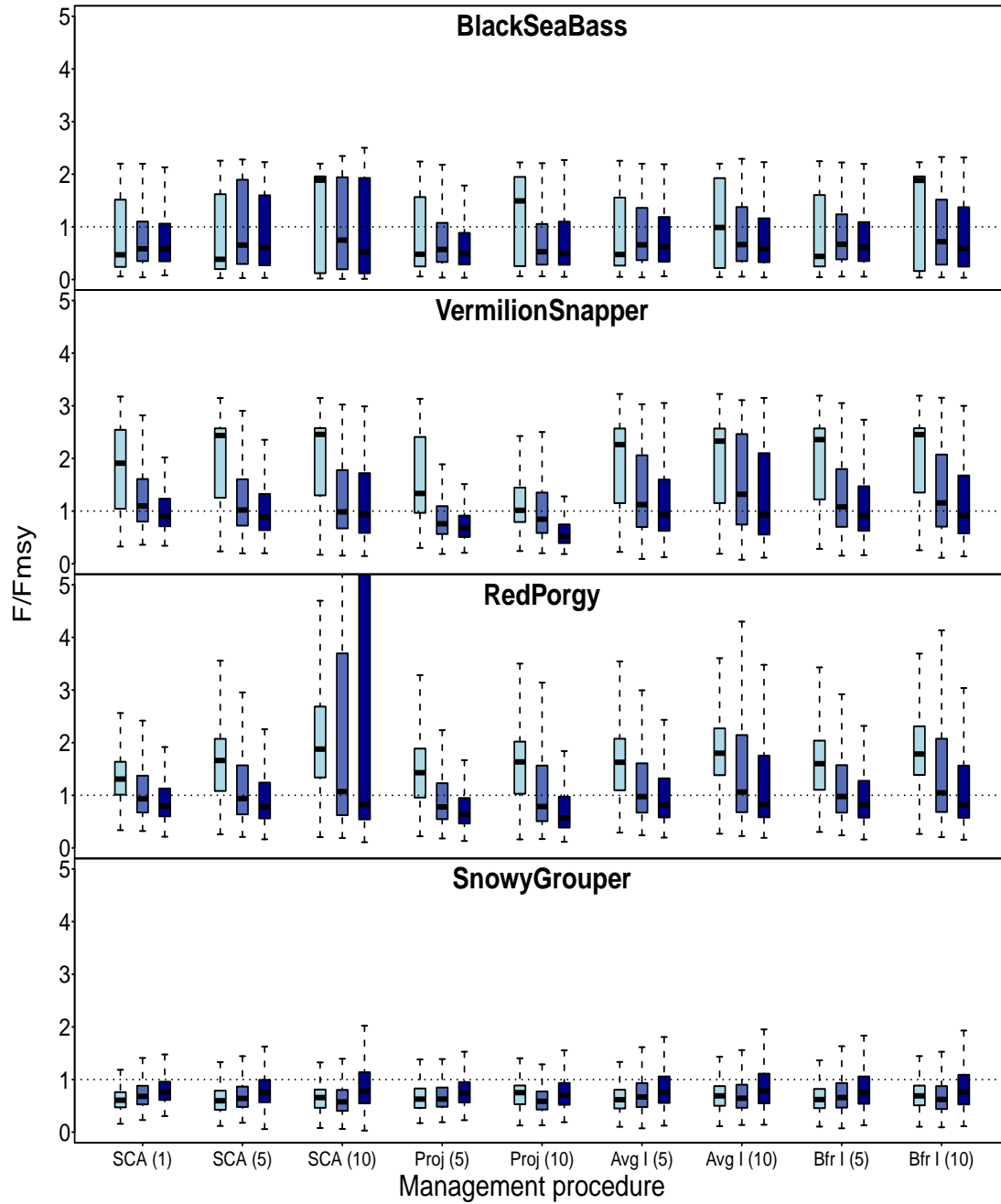


Fig. 4 Box plots of F/F_{MSY} for the base scenario. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times IQR$.

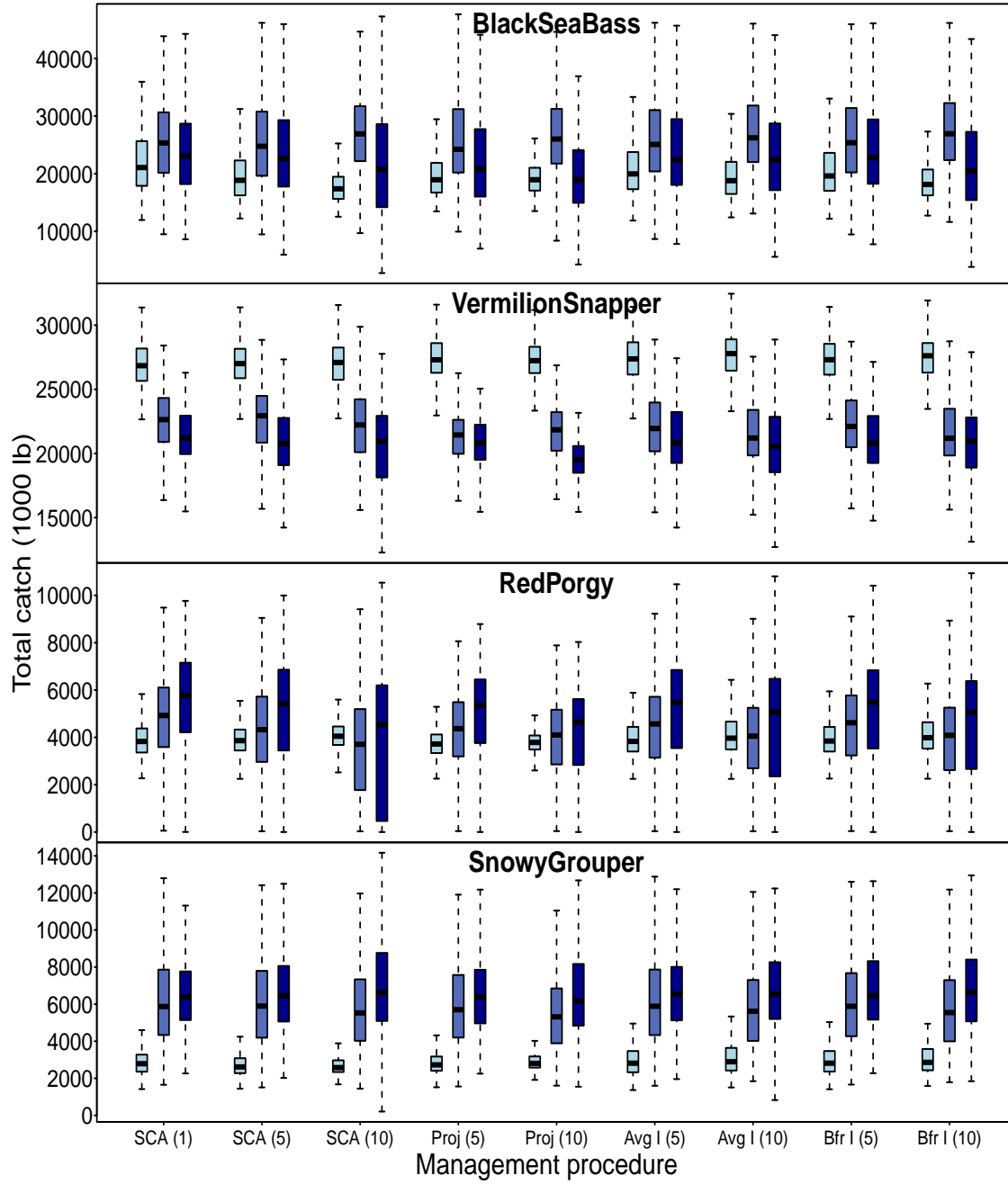


Fig. 5 Box plots of total catch (annual catches summed across each time period) for the base scenario. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times \text{IQR}$.

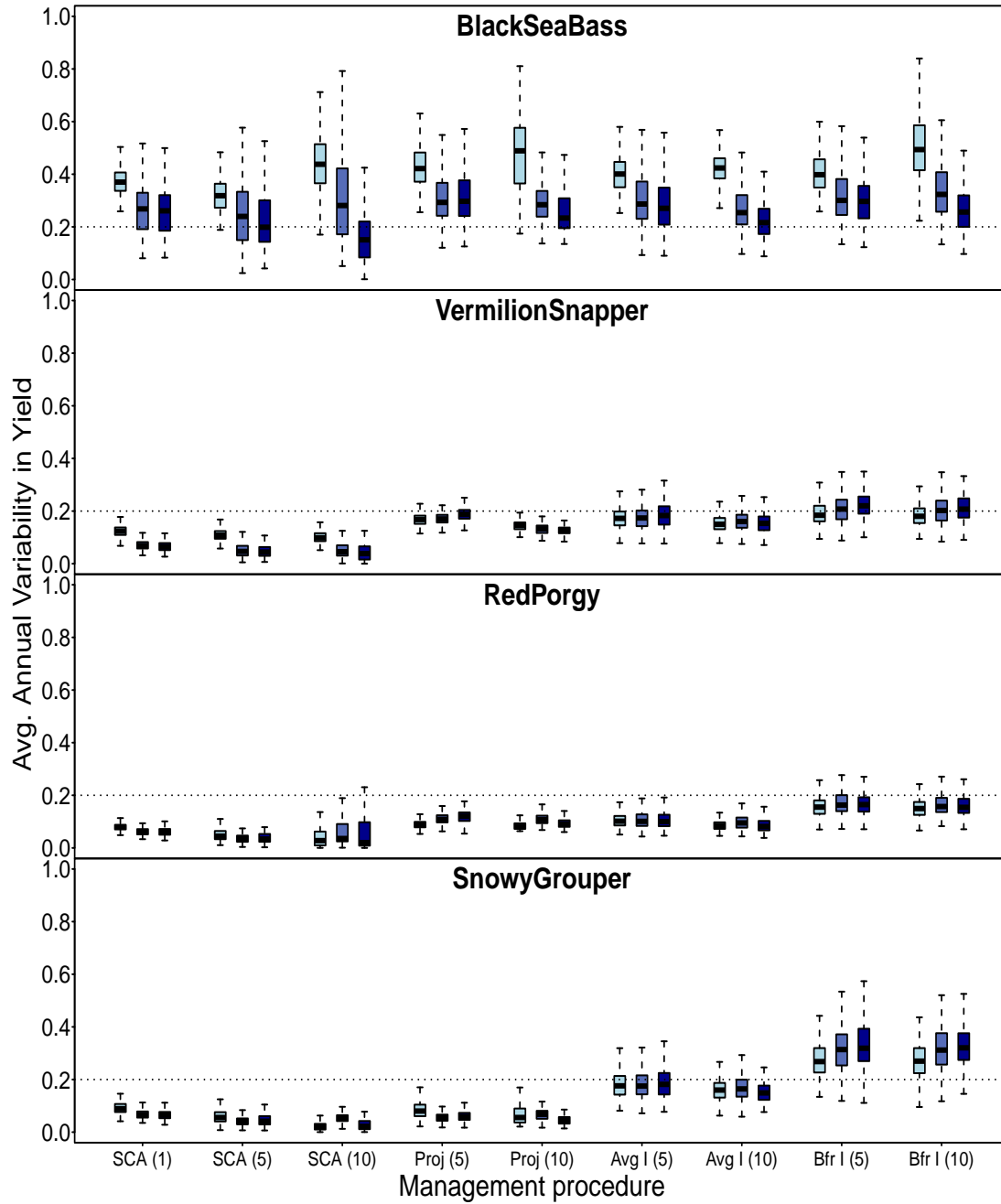


Fig. 6 Box plots of average annual variation in yield (AAVY) for the base scenario. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times \text{IQR}$.

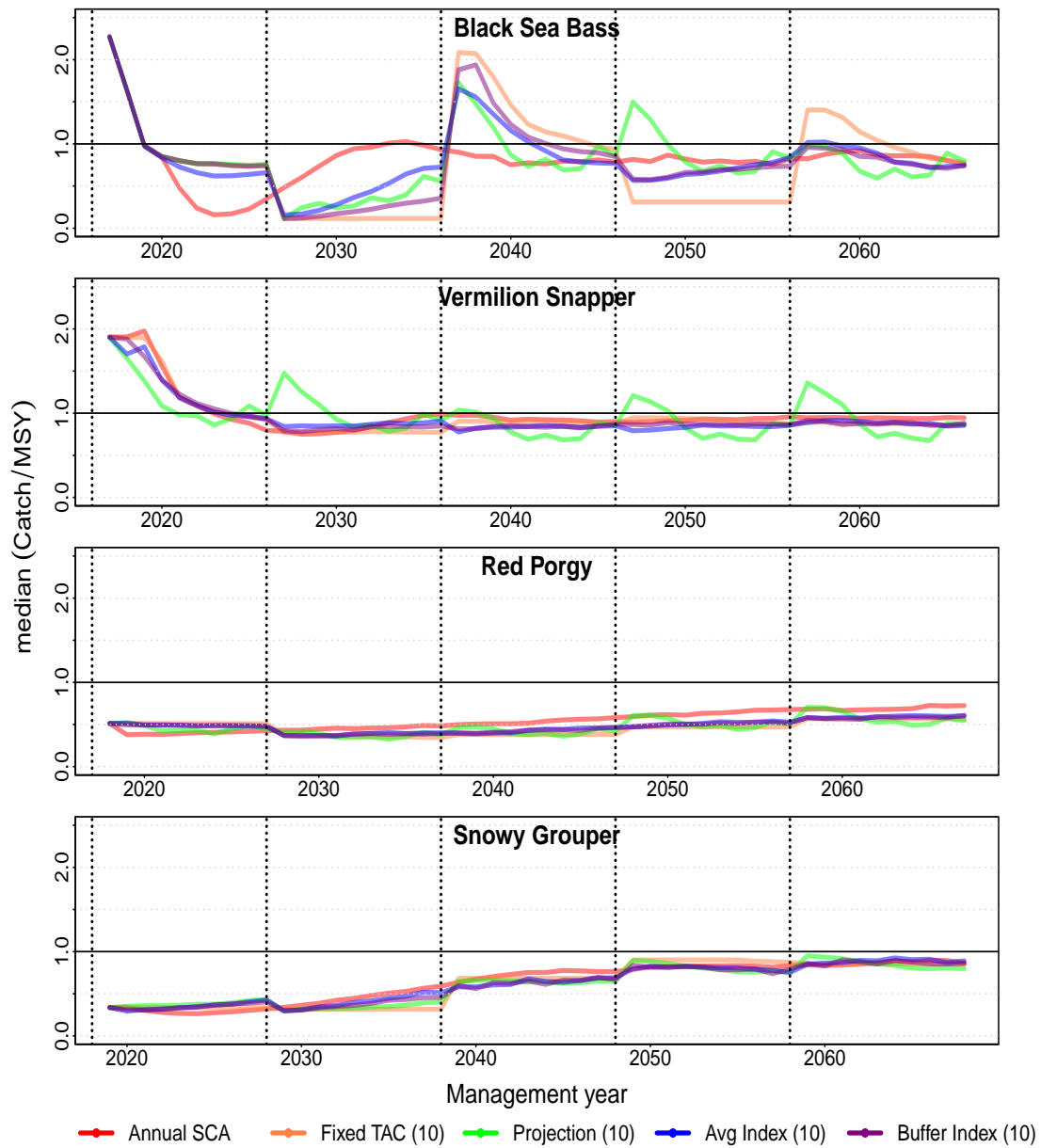


Fig. 7 Time series of median catch/MSY among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 10 yr intervals.

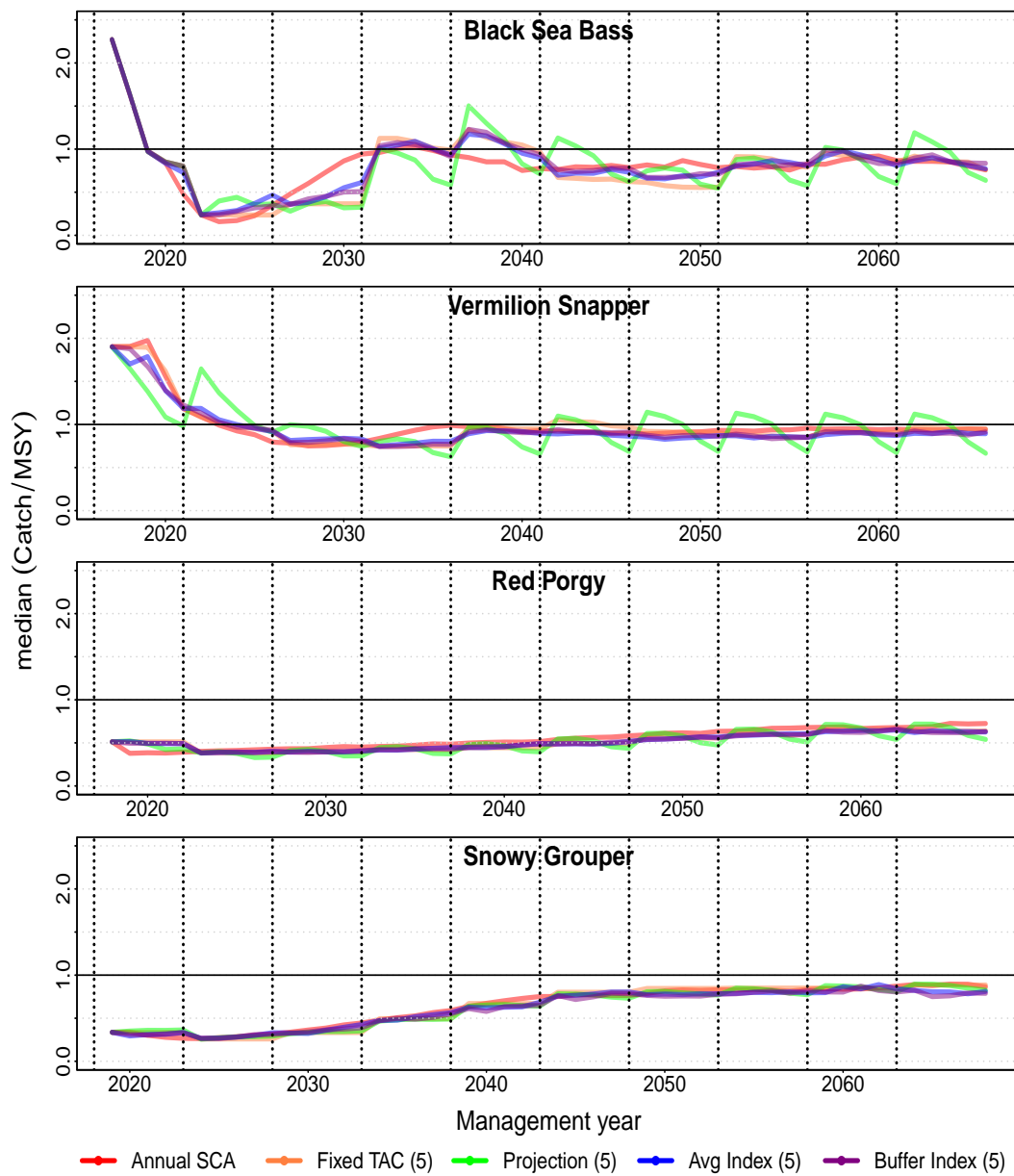


Fig. 8 Time series of median catch/MSY among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 5 yr intervals.

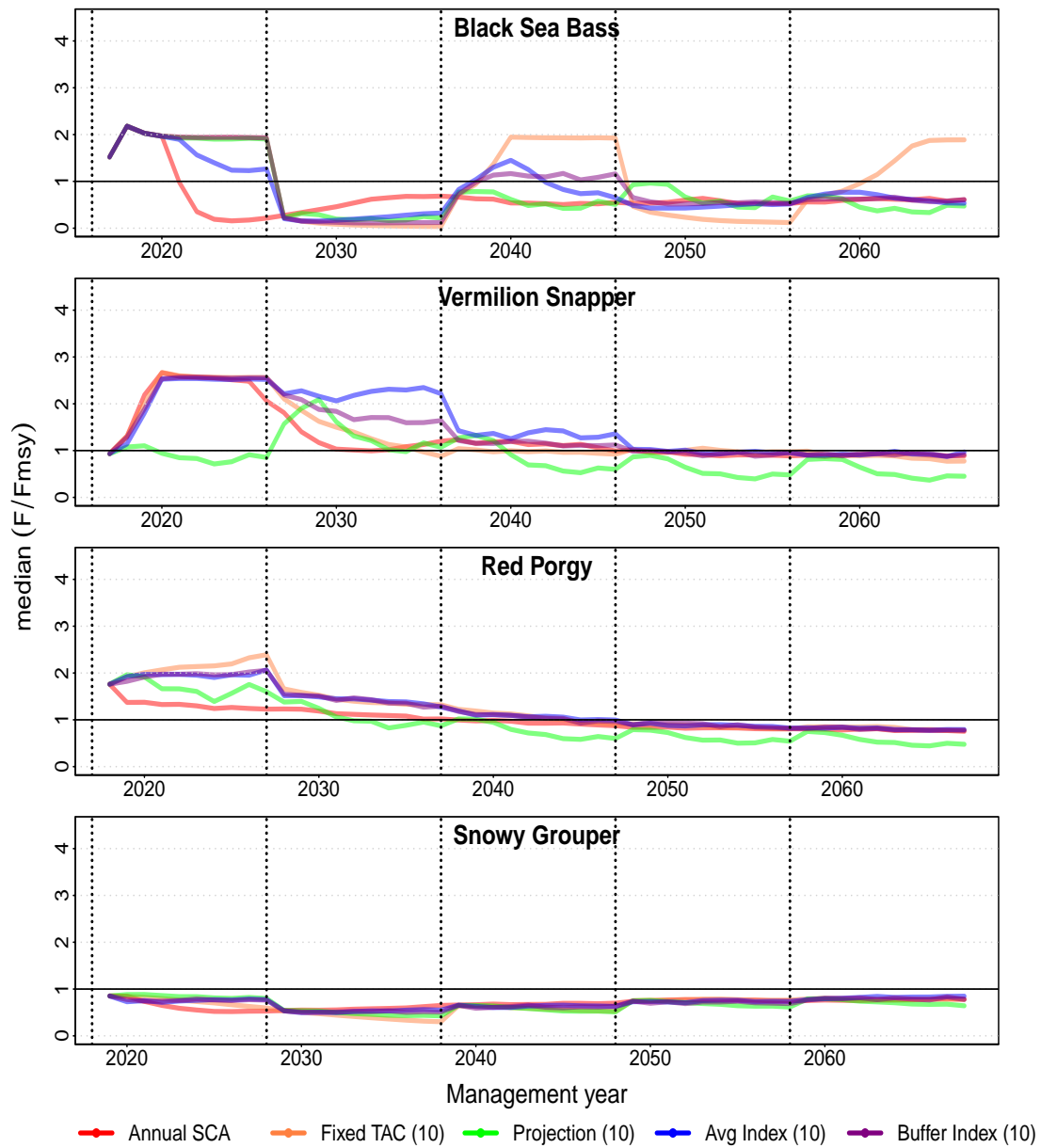


Fig. 9 Time series of median F/F_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 10 yr intervals.

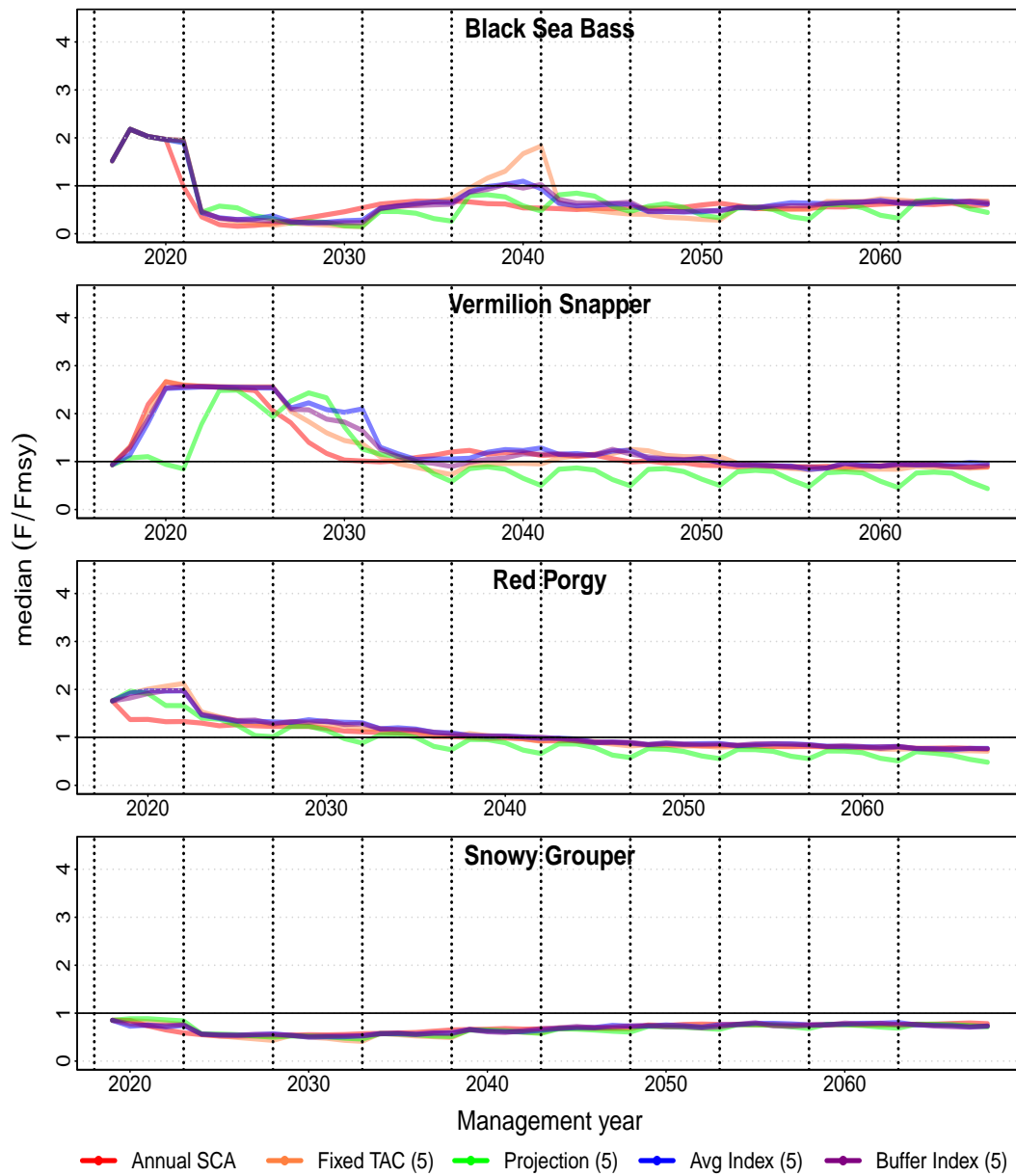


Fig. 10 Time series of median F/F_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 5 yr intervals.

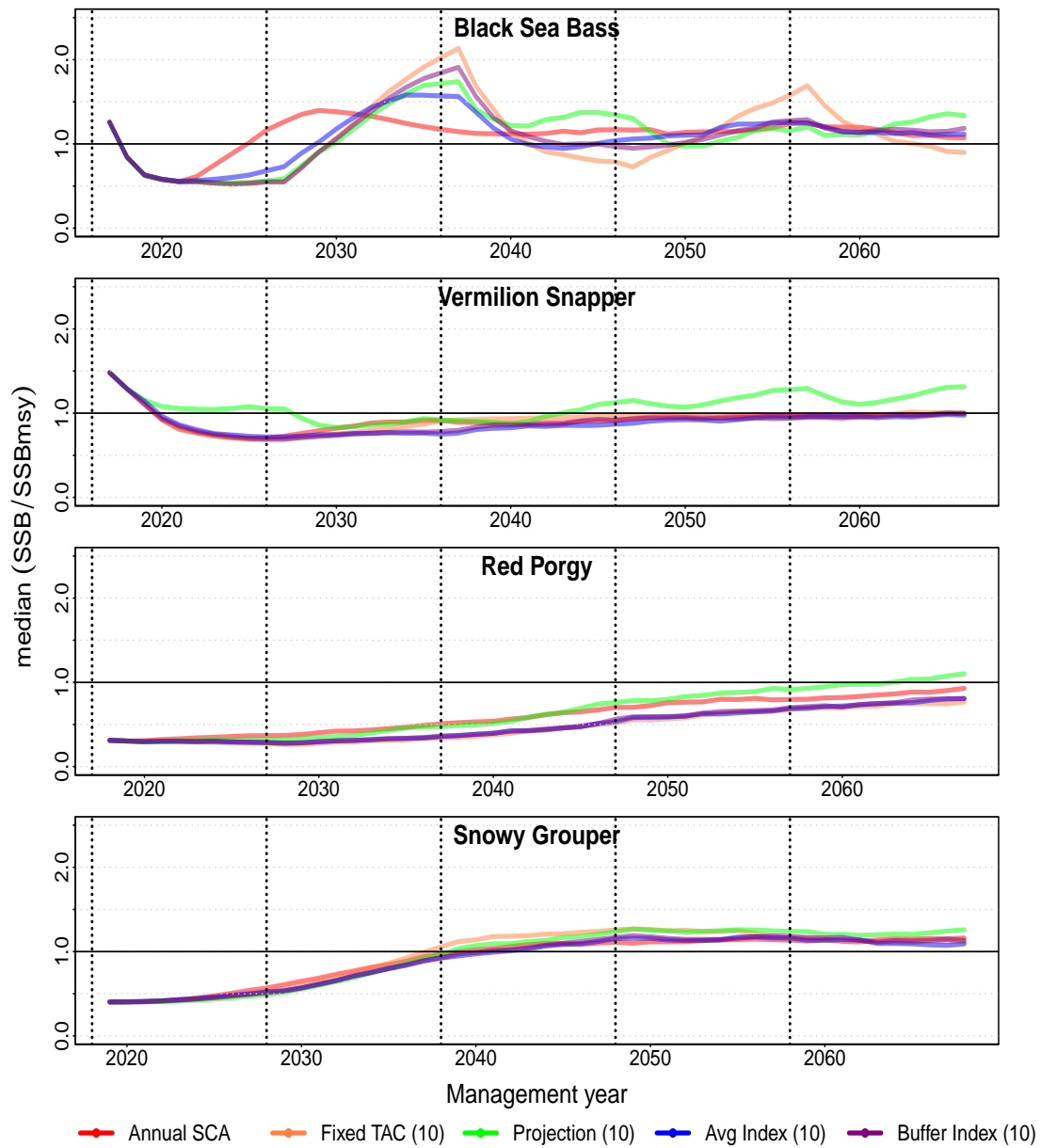


Fig. 11 Time series of median SSB/SSB_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 10 yr intervals.

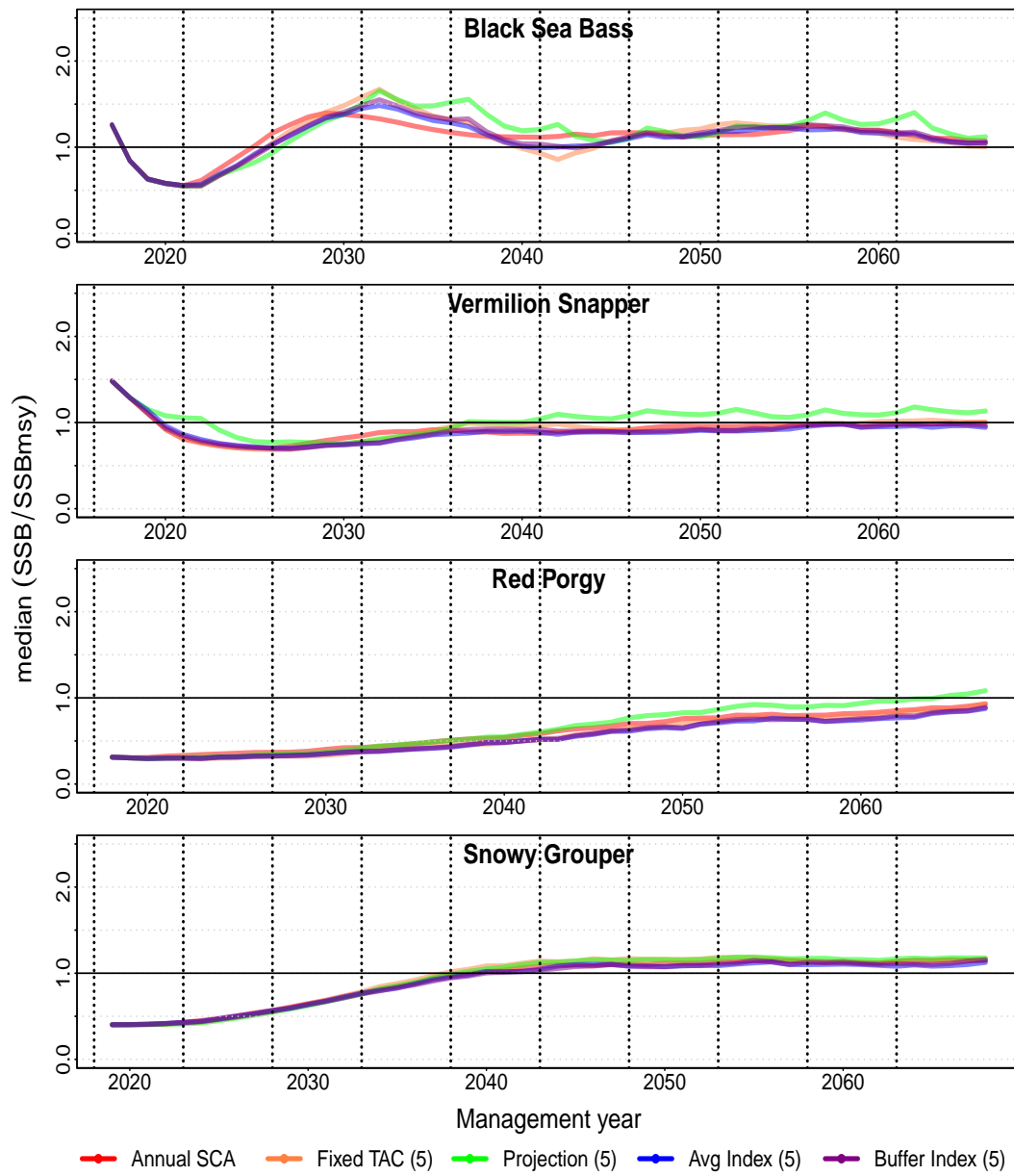


Fig. 12 Time series of median SSB/SSB_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 5 yr intervals.

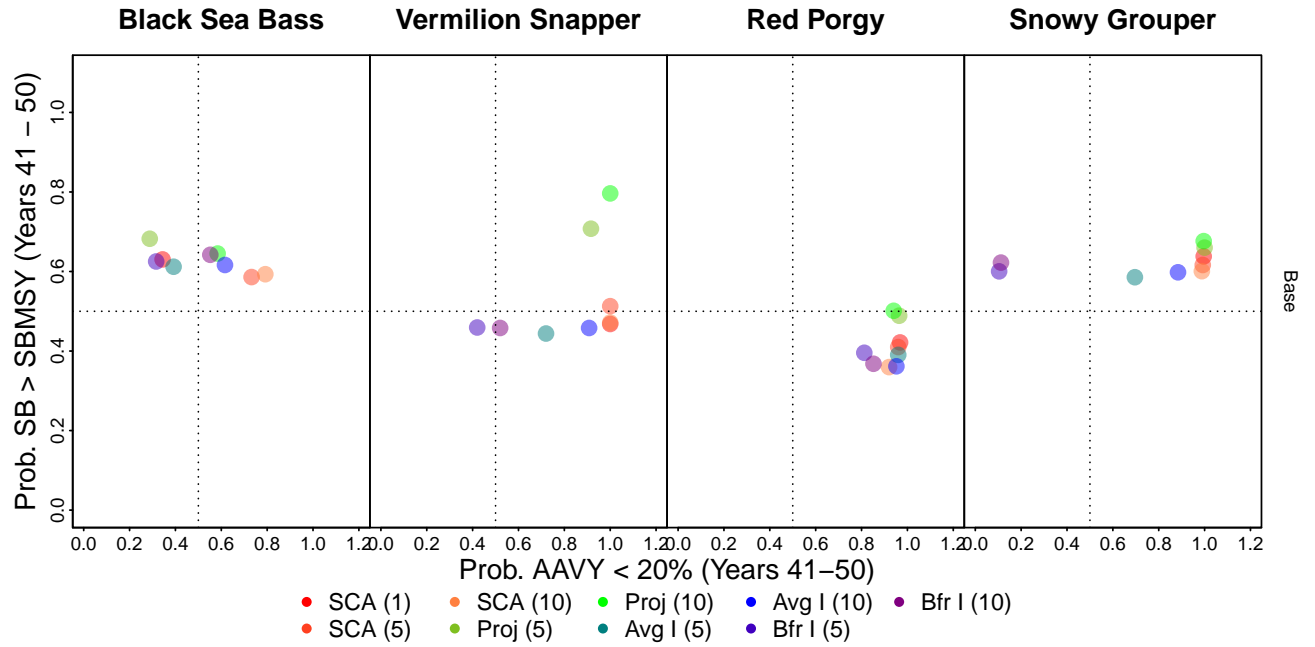


Fig. 13 Trade off plots of the probability that average annual variability in yield (AAVY) <20%, versus the probability that $SSB > SSB_{MSY}$ among simulation runs for the base scenario. Plots are based on data for the last 10 years of the projection period.

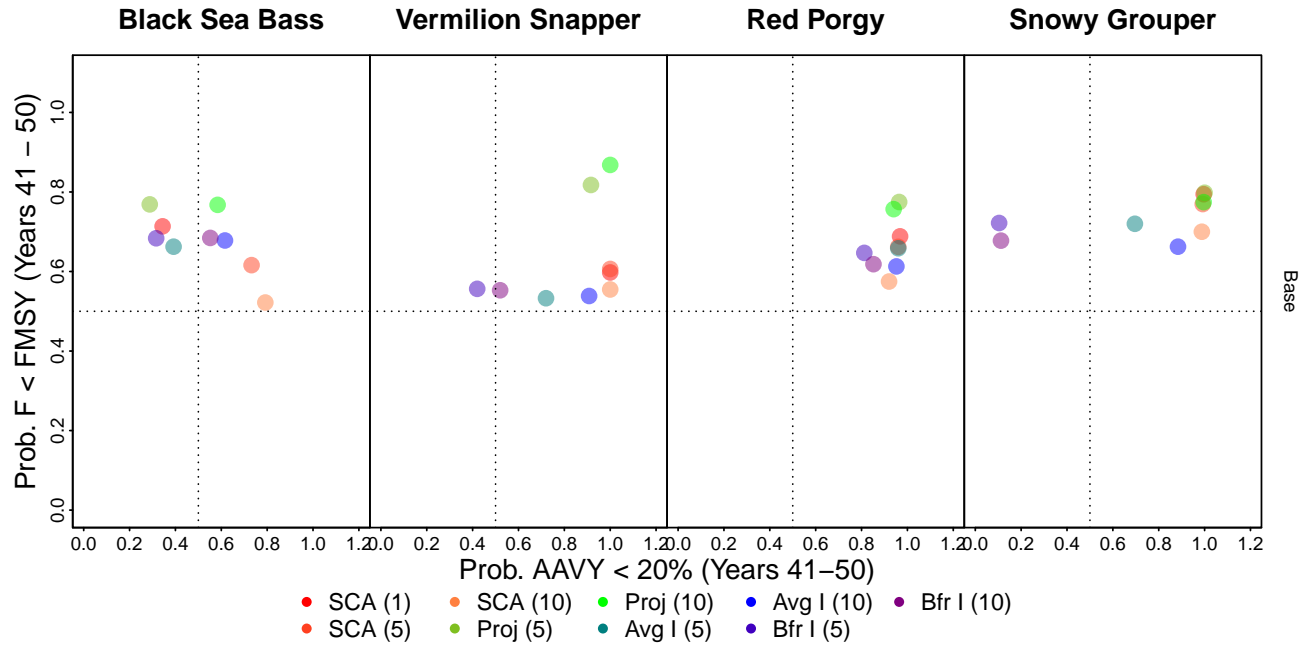


Fig. 14 Trade off plots of the probability that average annual variability in yield (AAVY) <20%, versus the probability that $F < F_{MSY}$ among simulation runs for the base scenario. Plots are based on data for the last 10 years of the projection period.

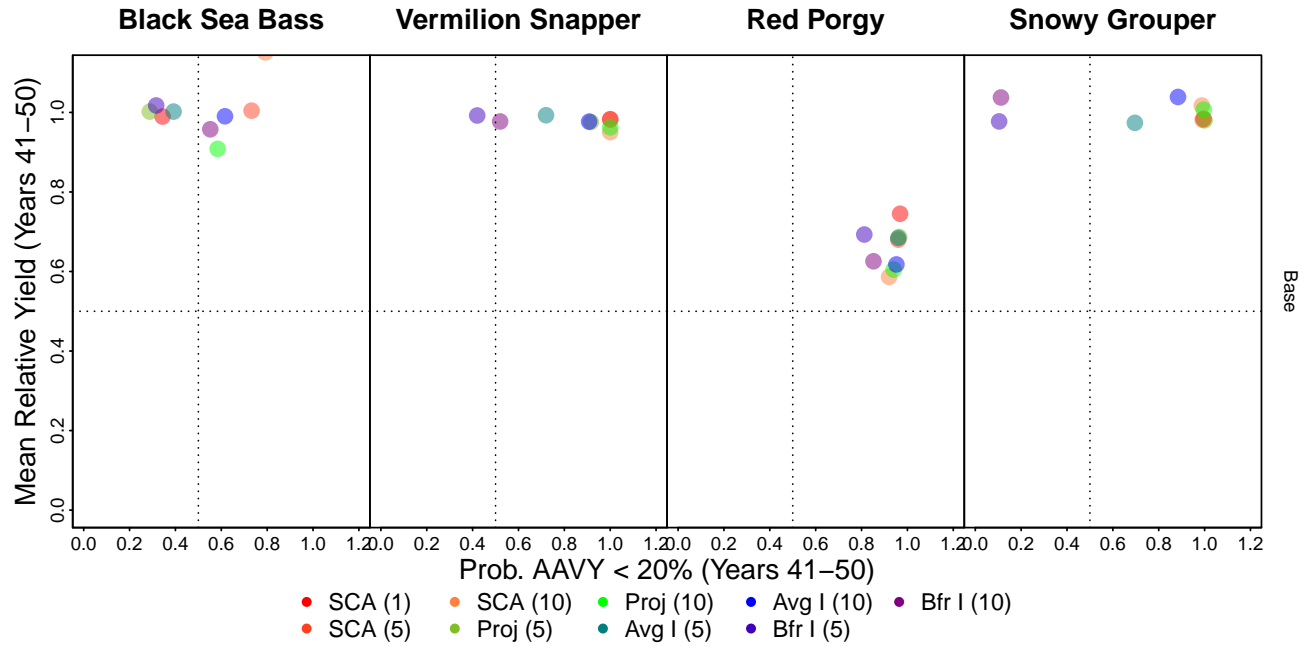


Fig. 15 Trade off plots of the probability that average annual variability in yield (AAVY) <20%, versus longterm yield among simulation runs for the base scenario. Plots are based on data for the last 10 years of the projection period.

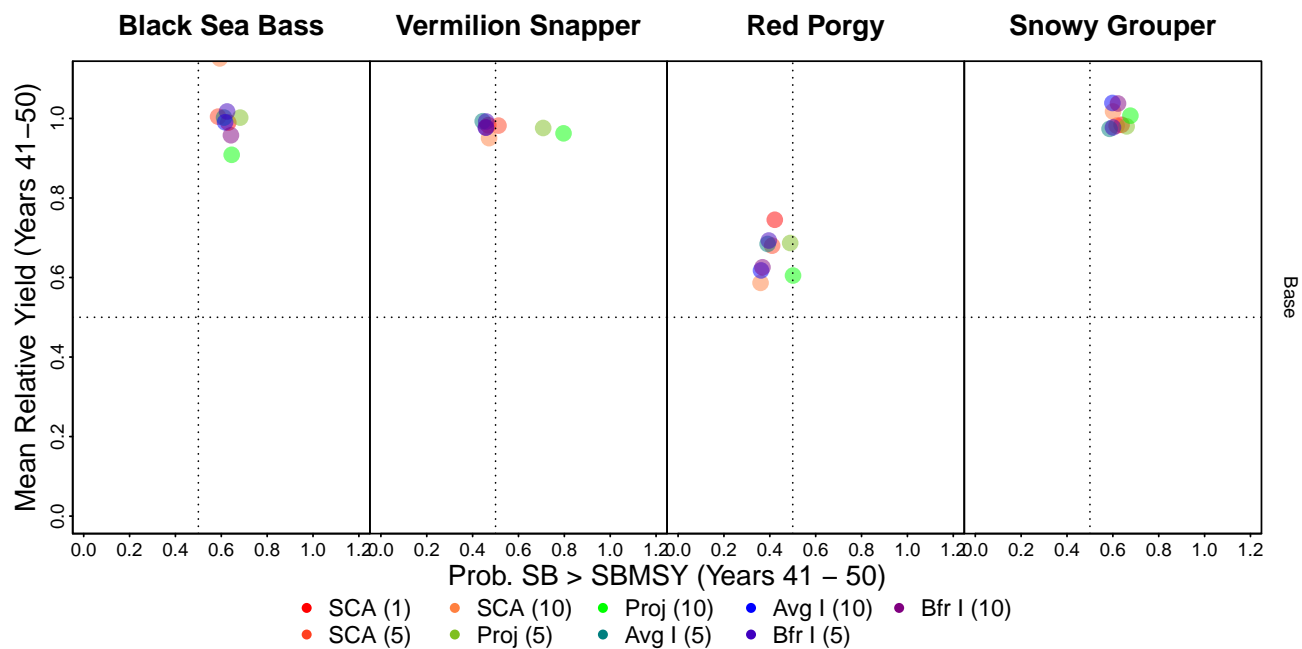


Fig. 16 Trade off plots of the probability that $SSB > SSB_{MSY}$ versus longterm yield among simulation runs for the base scenario. Plots are based on data for the last 10 years of the projection period.

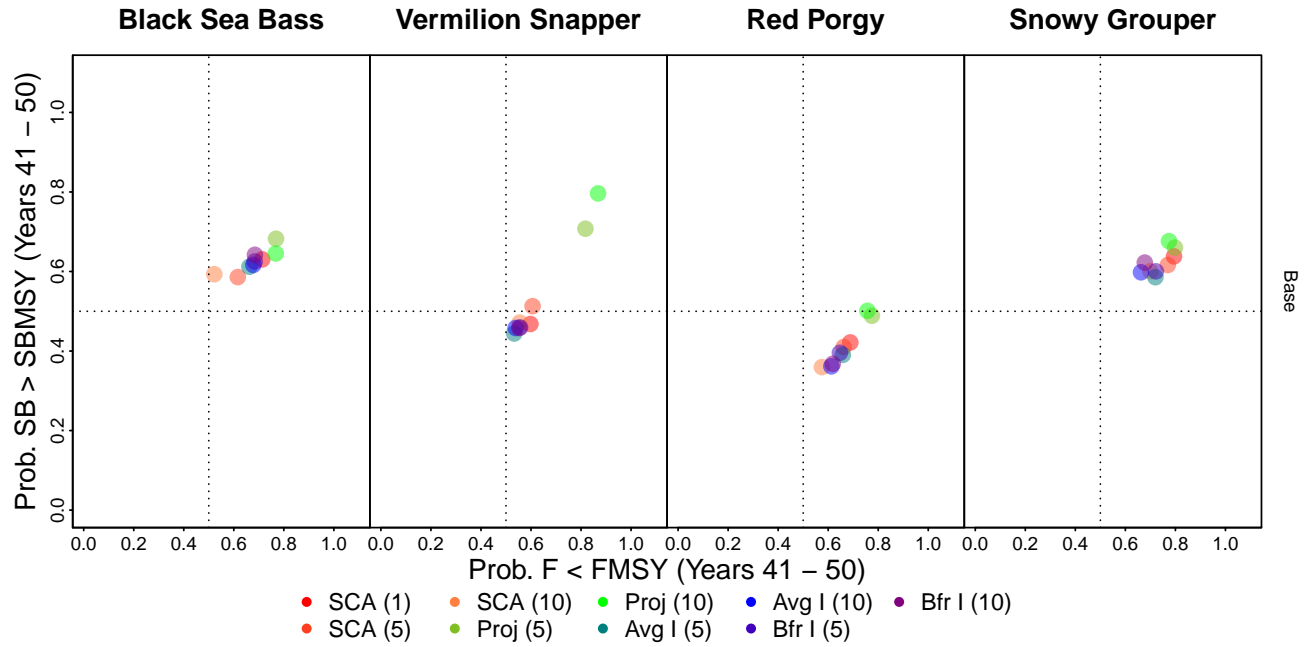


Fig. 17 Trade off plots of the probability that $F < F_{MSY}$ versus the probability that $SSB > SSB_{MSY}$ among simulation runs for the base scenario. Plots are based on data for the last 10 years of the projection period.

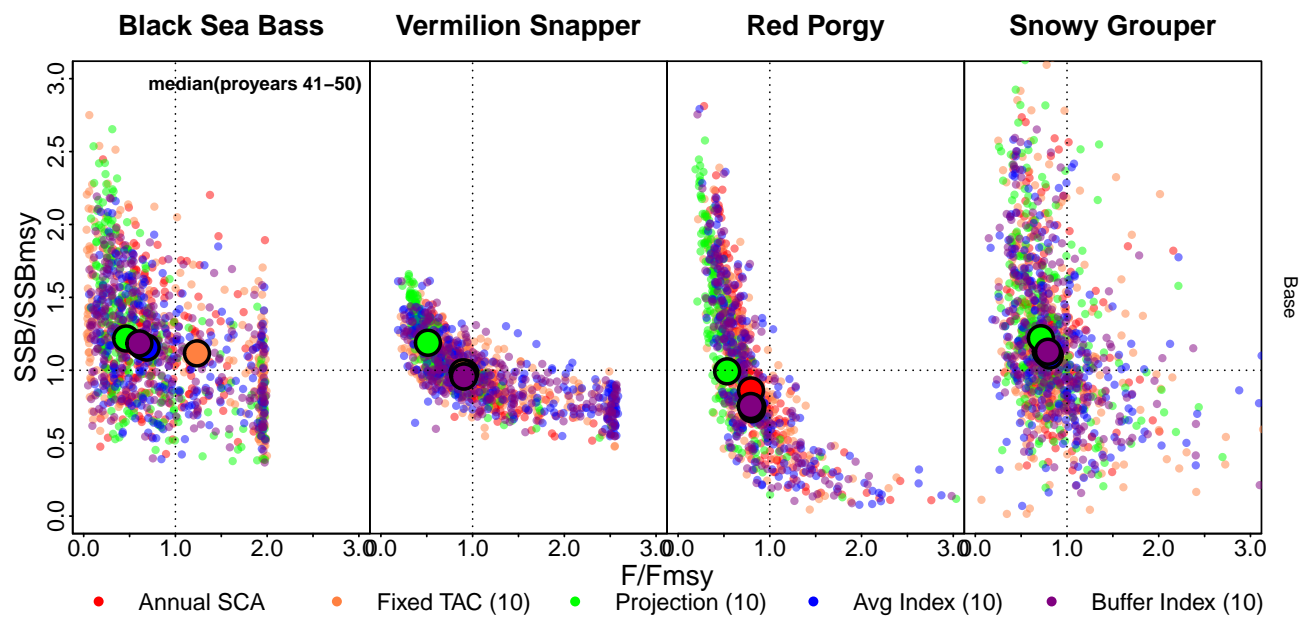


Fig. 18 Phase plots of the probability that $F > F_{MSY}$ versus the probability that $SSB > SSB_{MSY}$ among simulation runs for the base scenario.

APPENDIX

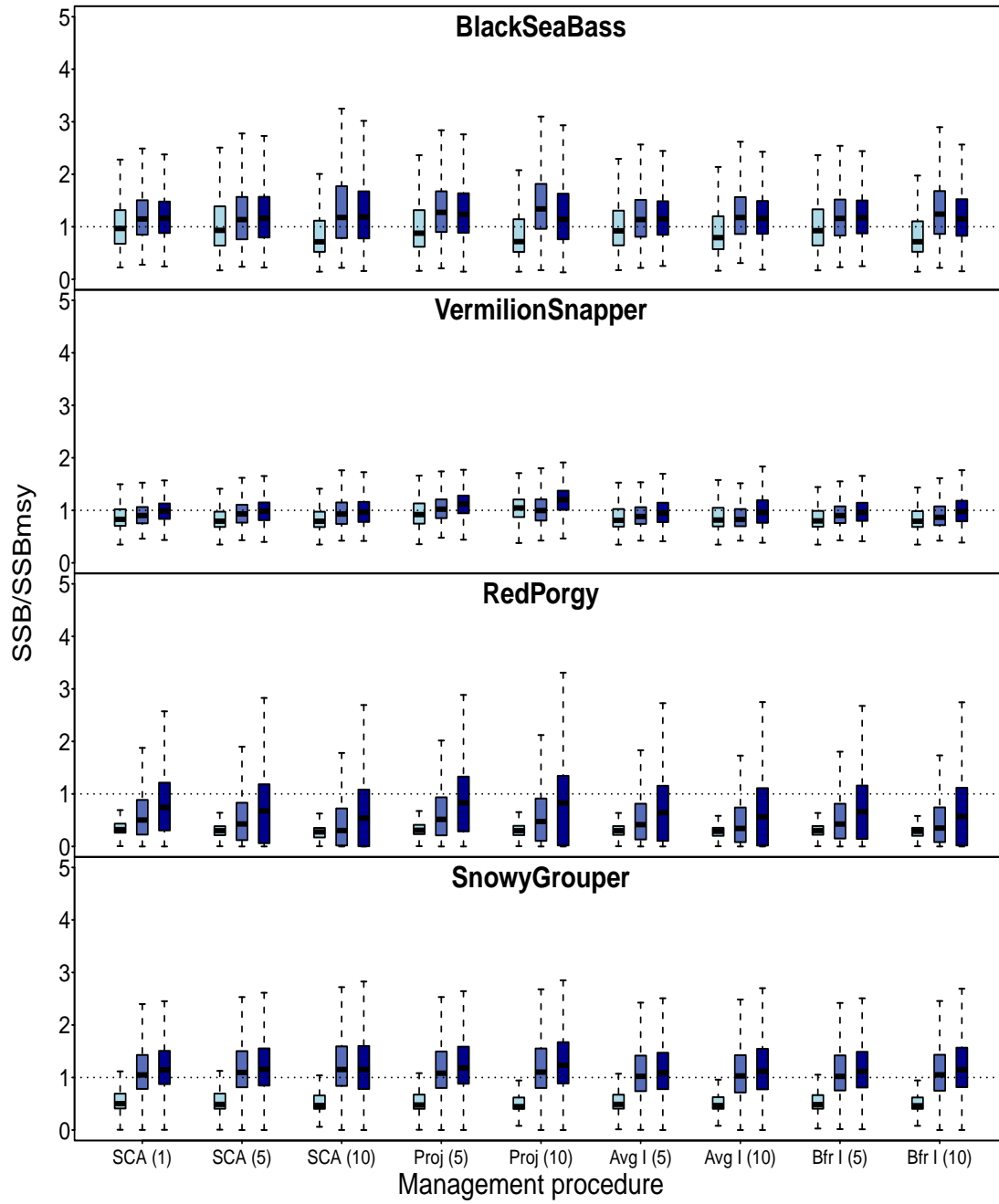


Fig. S1 Box plots of SSB/SSB_{MSY} for the base and alternative scenarios. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times \text{IQR}$.

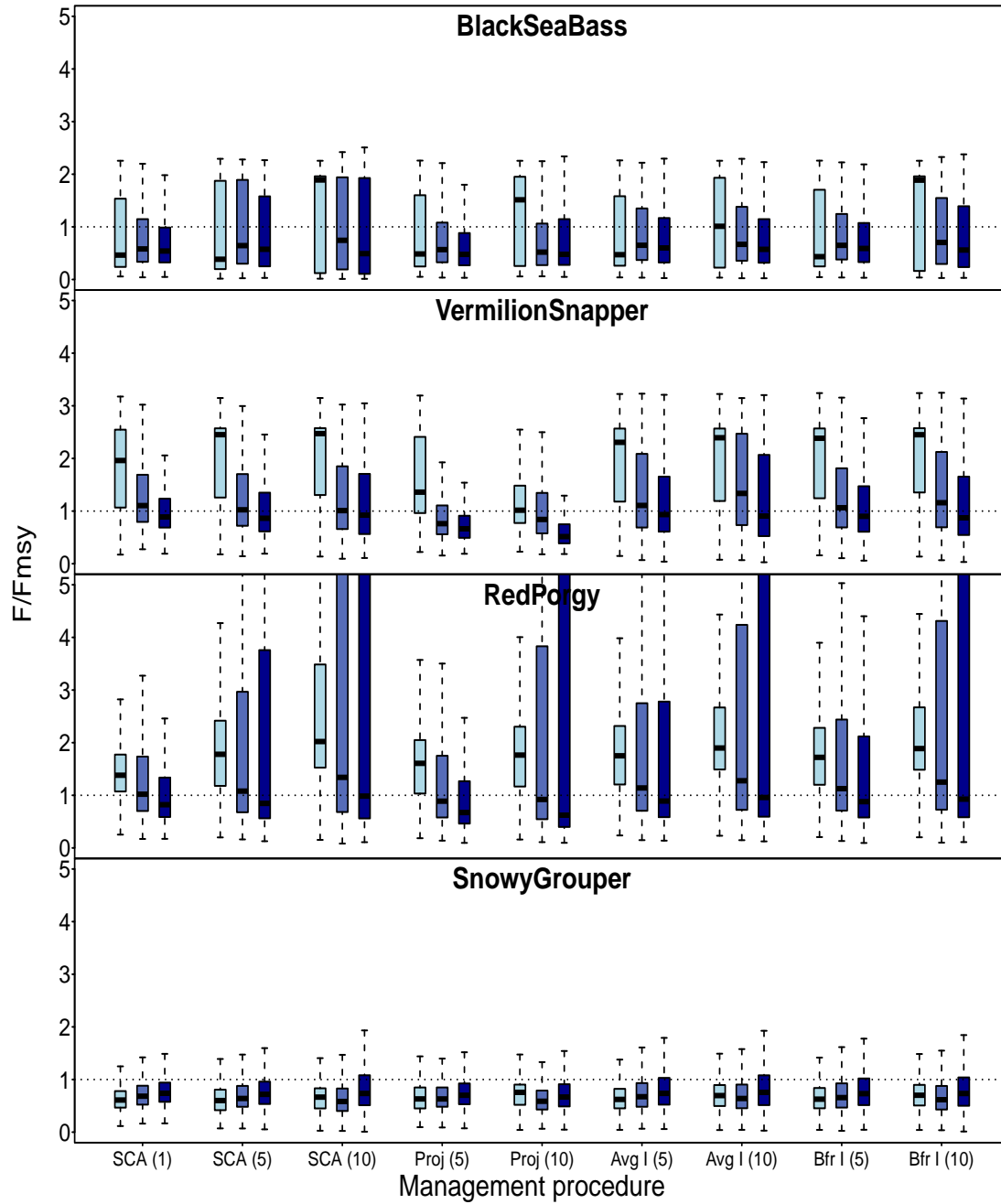


Fig. S2 Box plots of F/F_{MSY} for the base and alternative scenarios. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times IQR$.

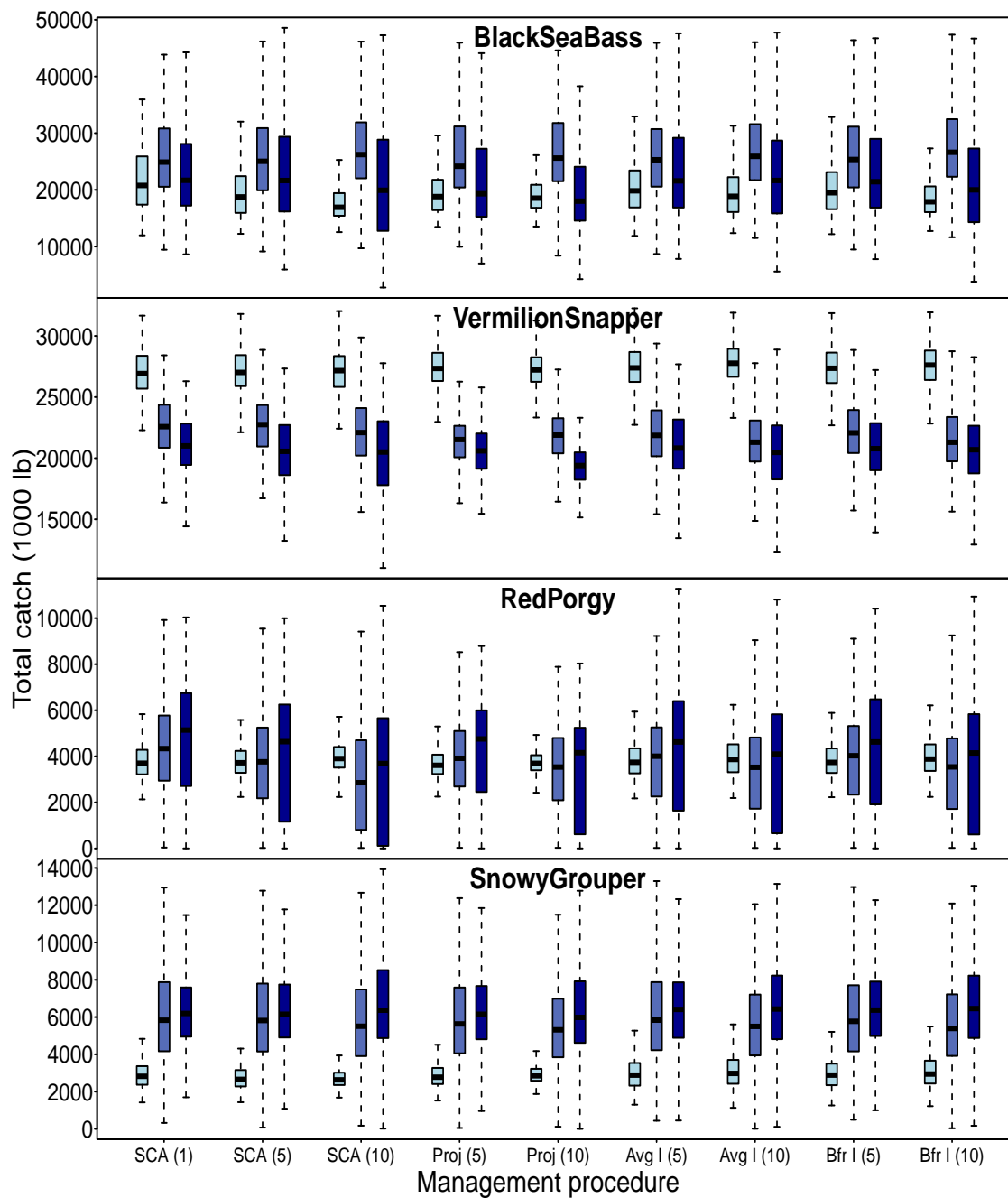


Fig. S3 Box plots of total catch (annual catches summed across each time period) for the base and alternative scenarios. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times \text{IQR}$.

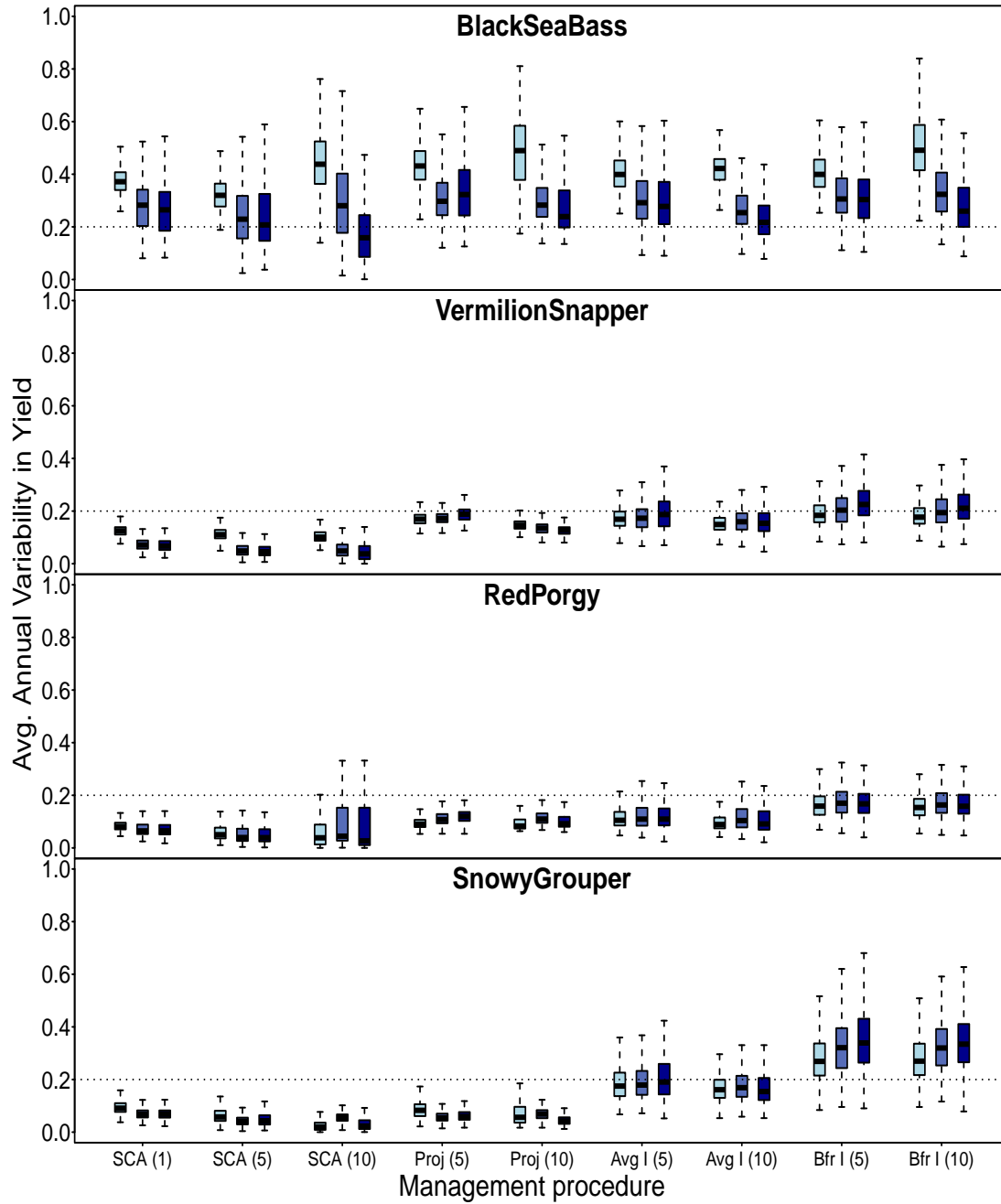


Fig. S4 Box plots of average annual variation in yield (AAVY) for the base and alternative scenarios. Grouped boxes represent sequential time periods during the projection period. Boxes represent interquartile range (IQR). Whiskers are drawn to $1.5 \times \text{IQR}$.

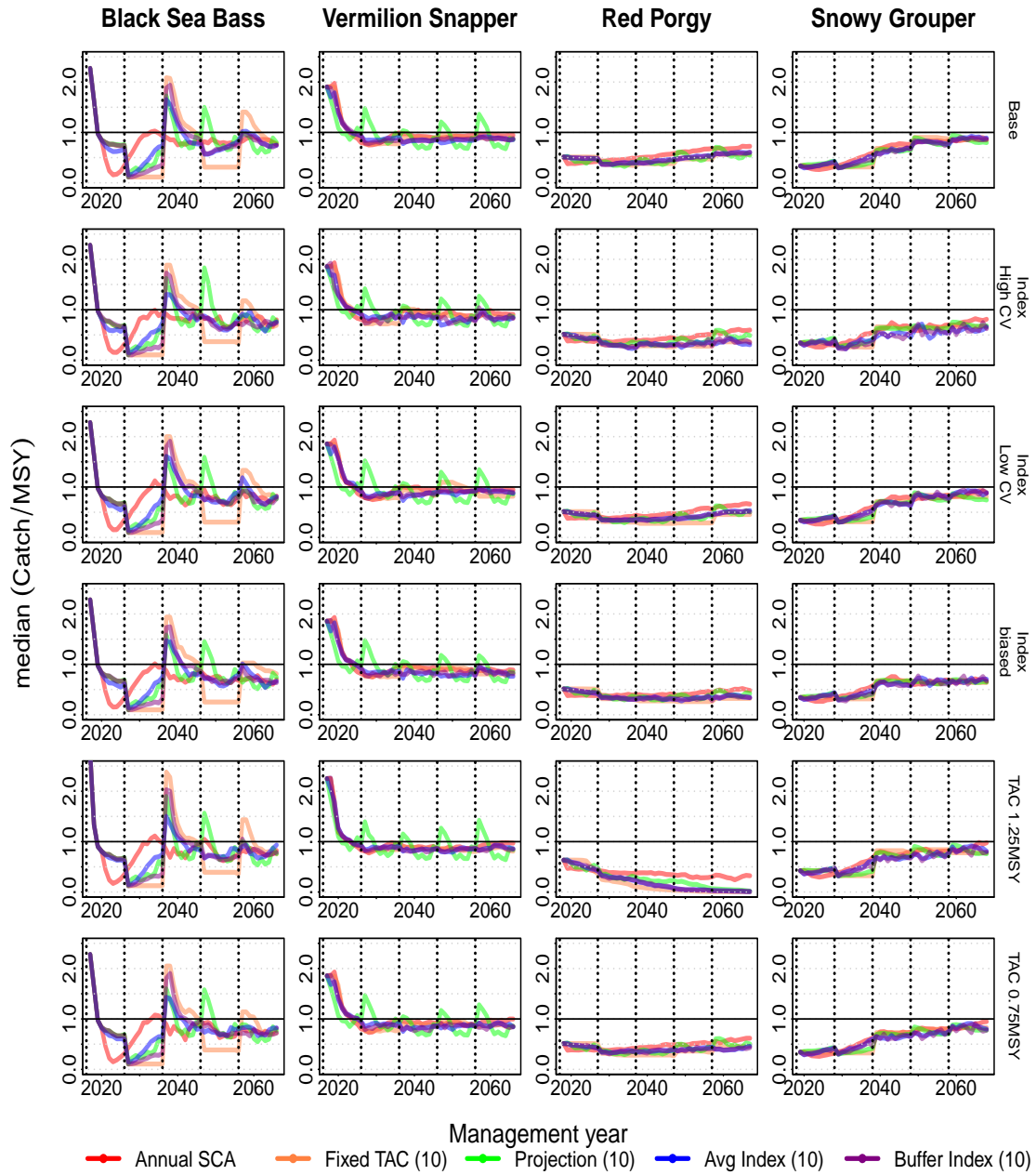


Fig. S5 Time series of median catch/MSY among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 10 yr intervals.

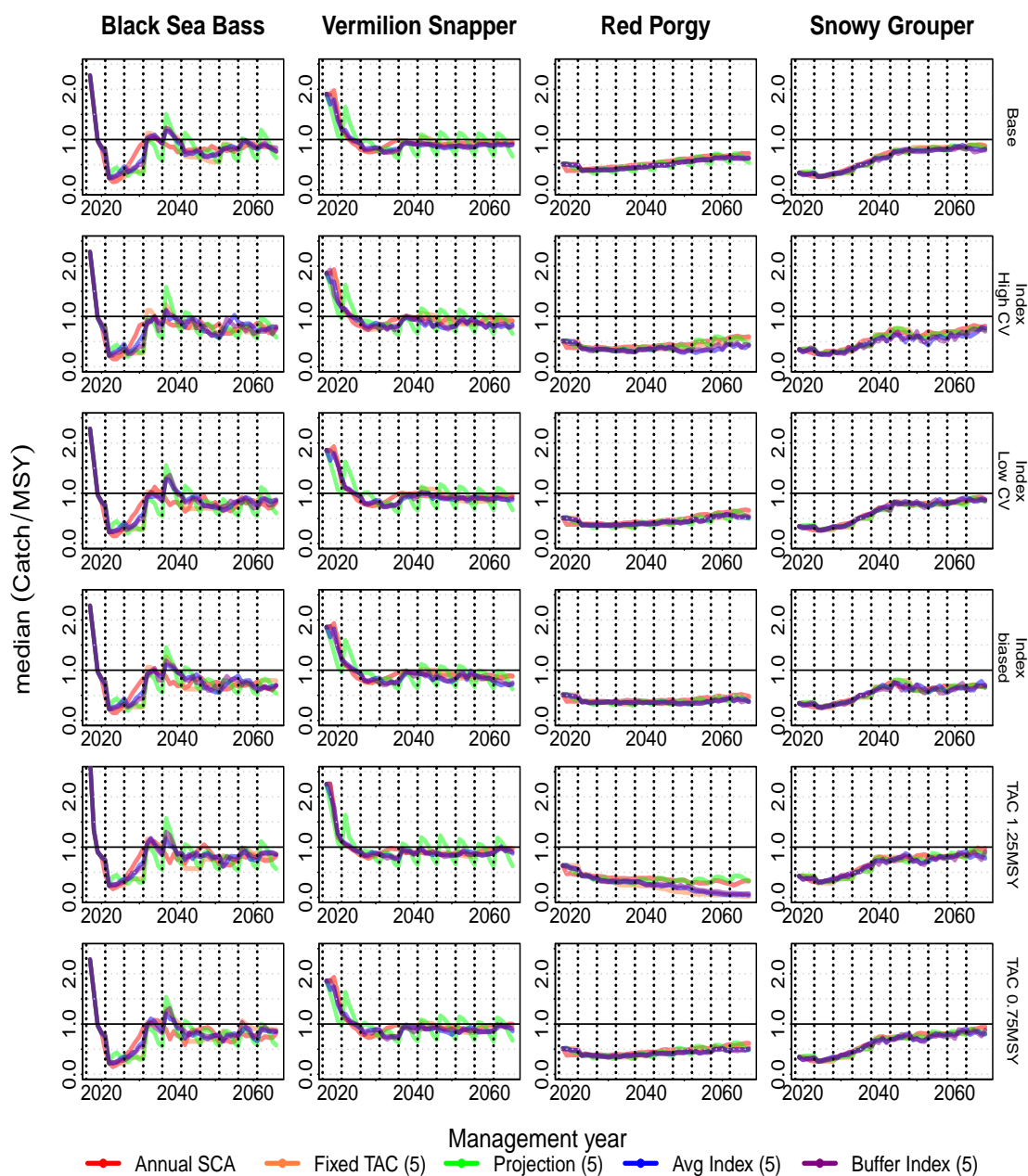


Fig. S6 Time series of median catch/MSY among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 5 yr intervals.

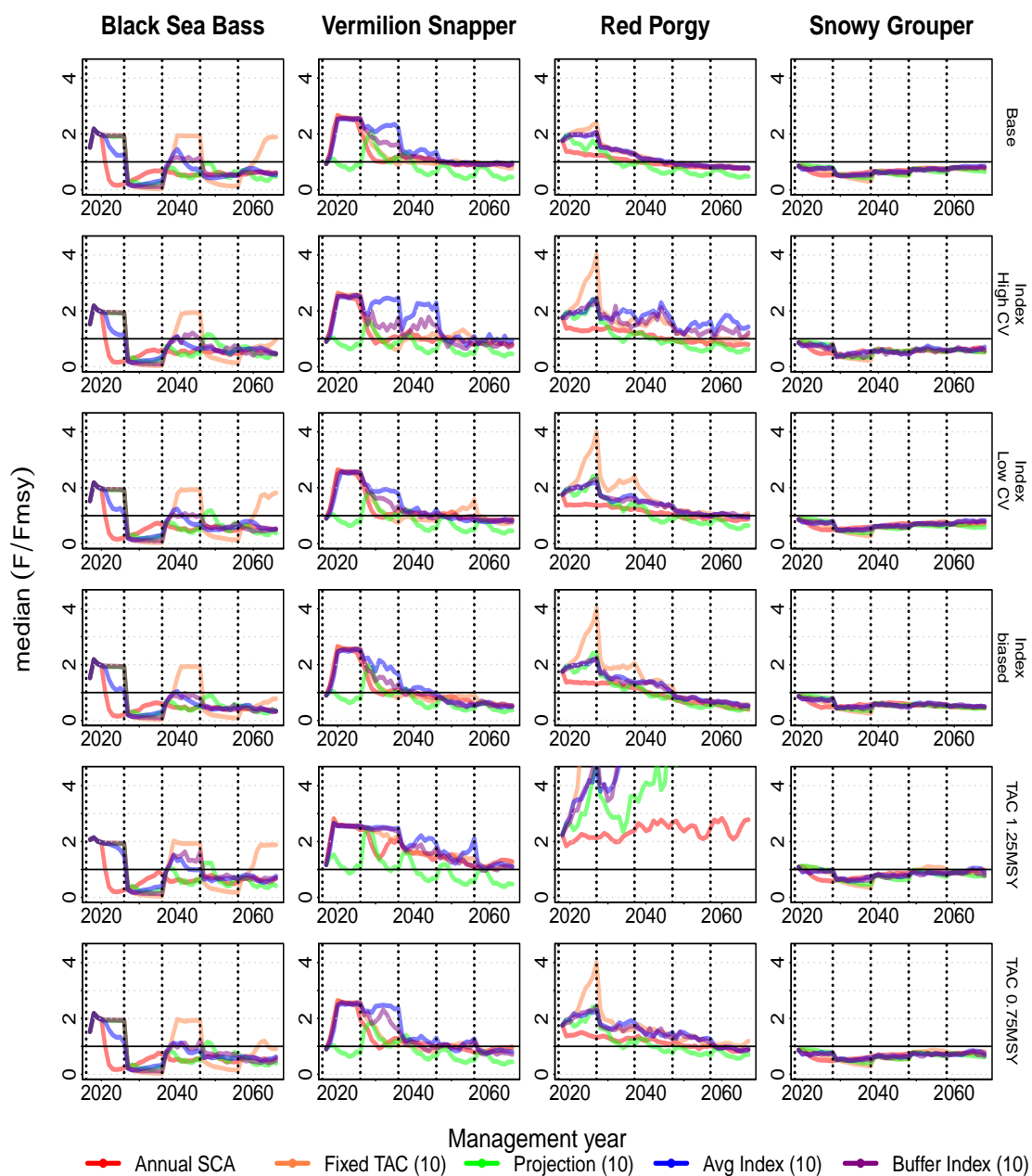


Fig. S7 Time series of median F/F_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 10 yr intervals.

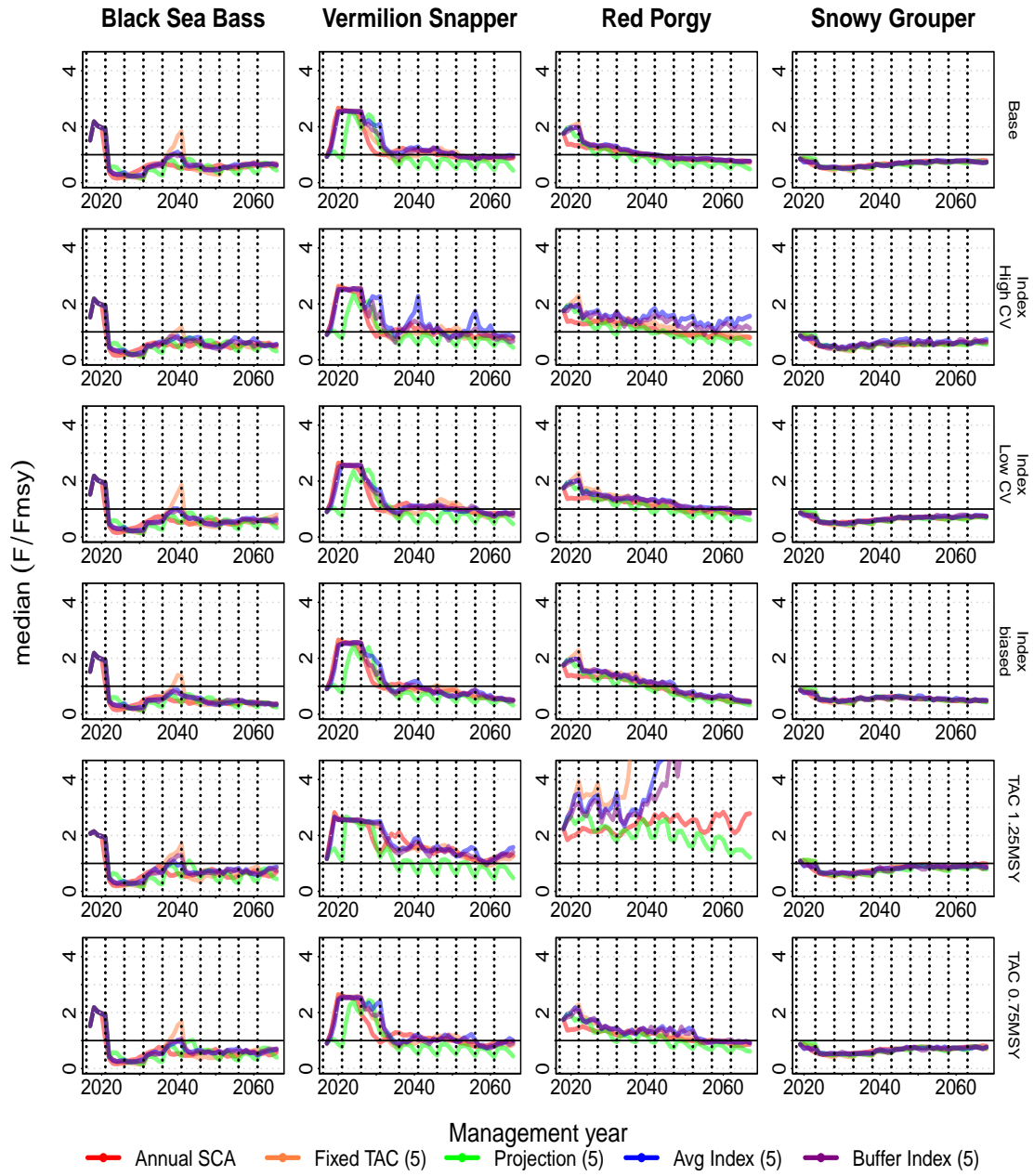


Fig. S8 Time series of median F/F_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 5 yr intervals.

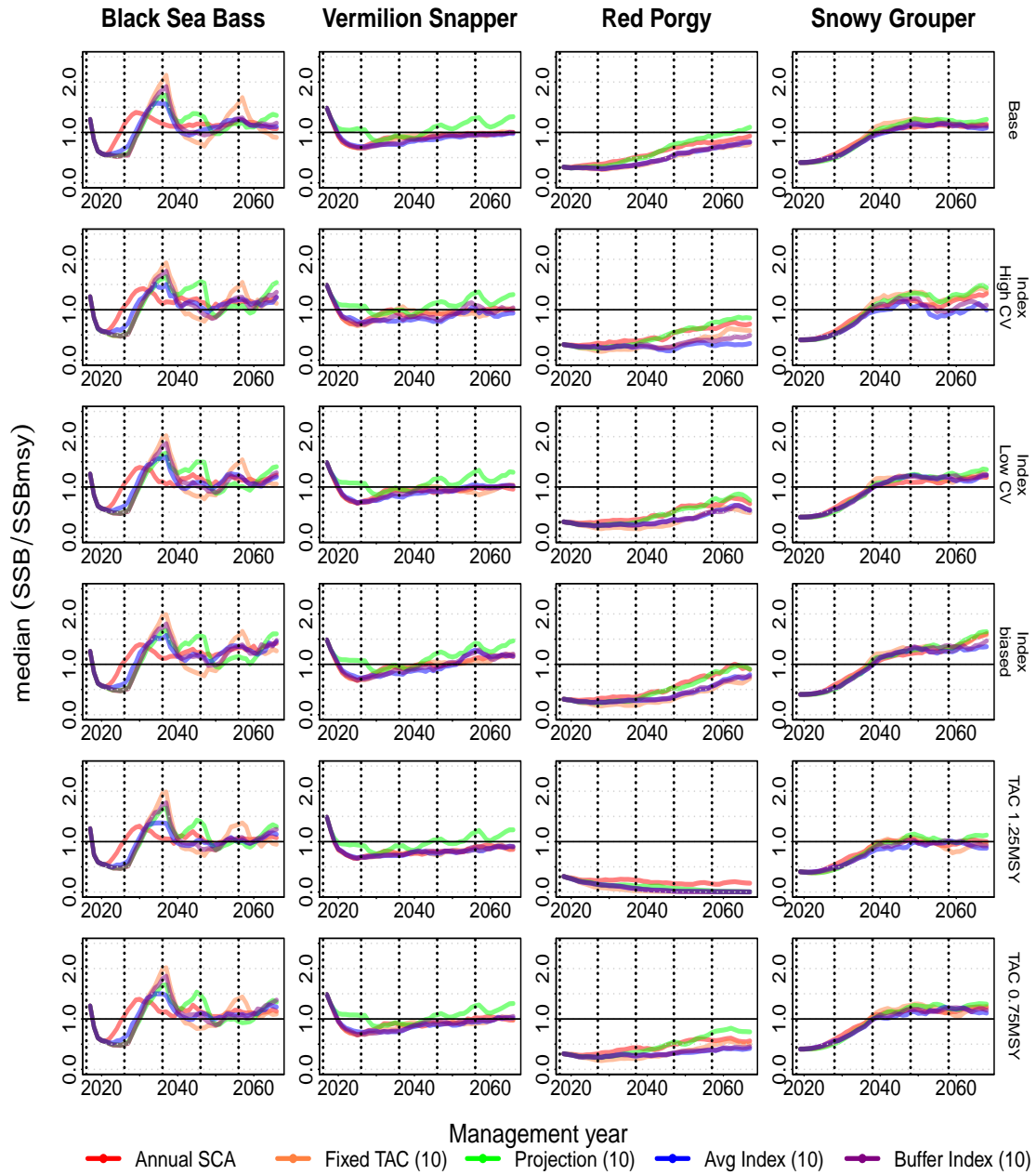


Fig. S9 Time series of median SSB/SSB_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 10 yr intervals.

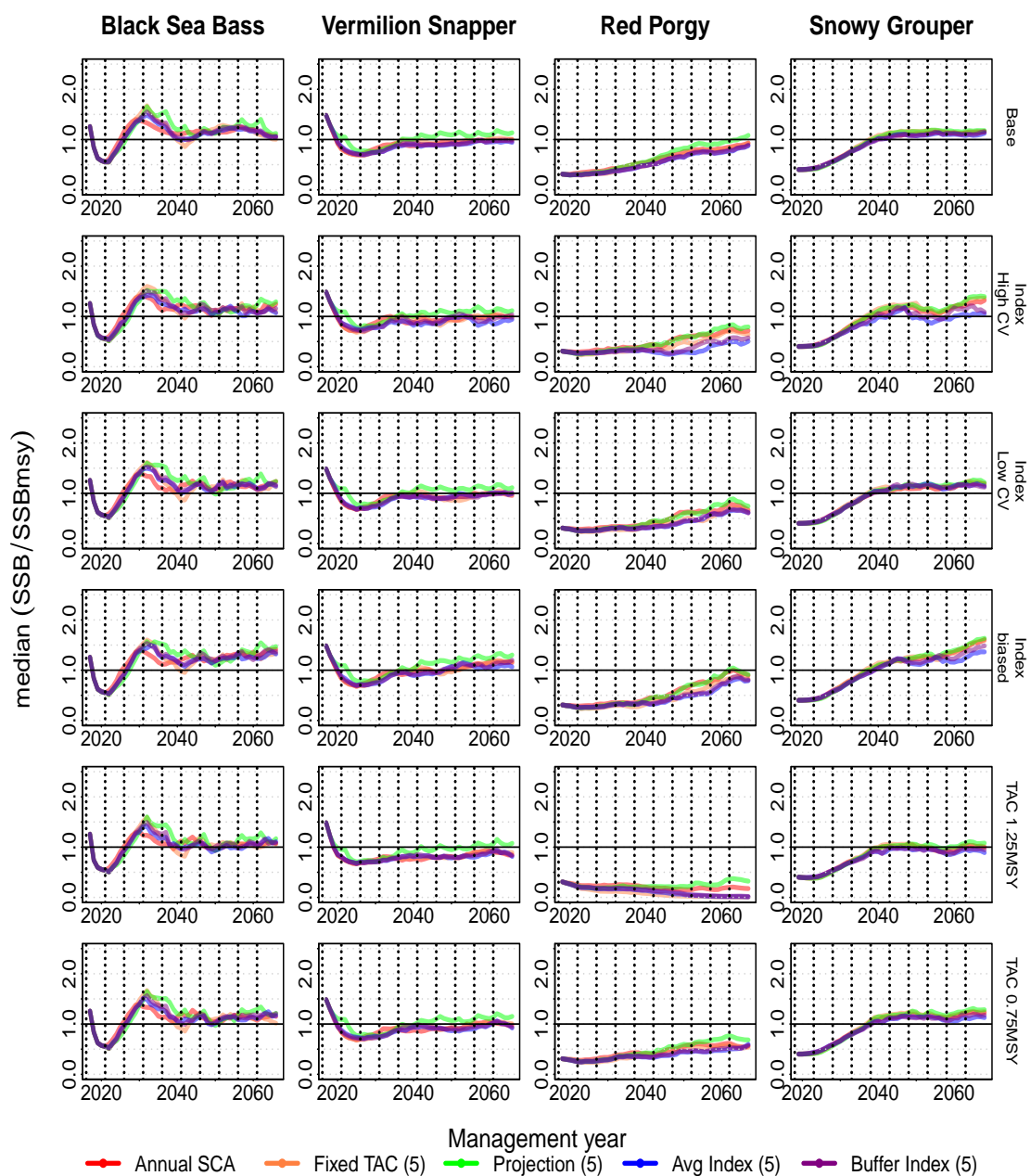


Fig. S10 Time series of median SSB/SSB_{MSY} among simulation runs for the base scenario. Vertical dotted lines indicate assessment years at 5 yr intervals.

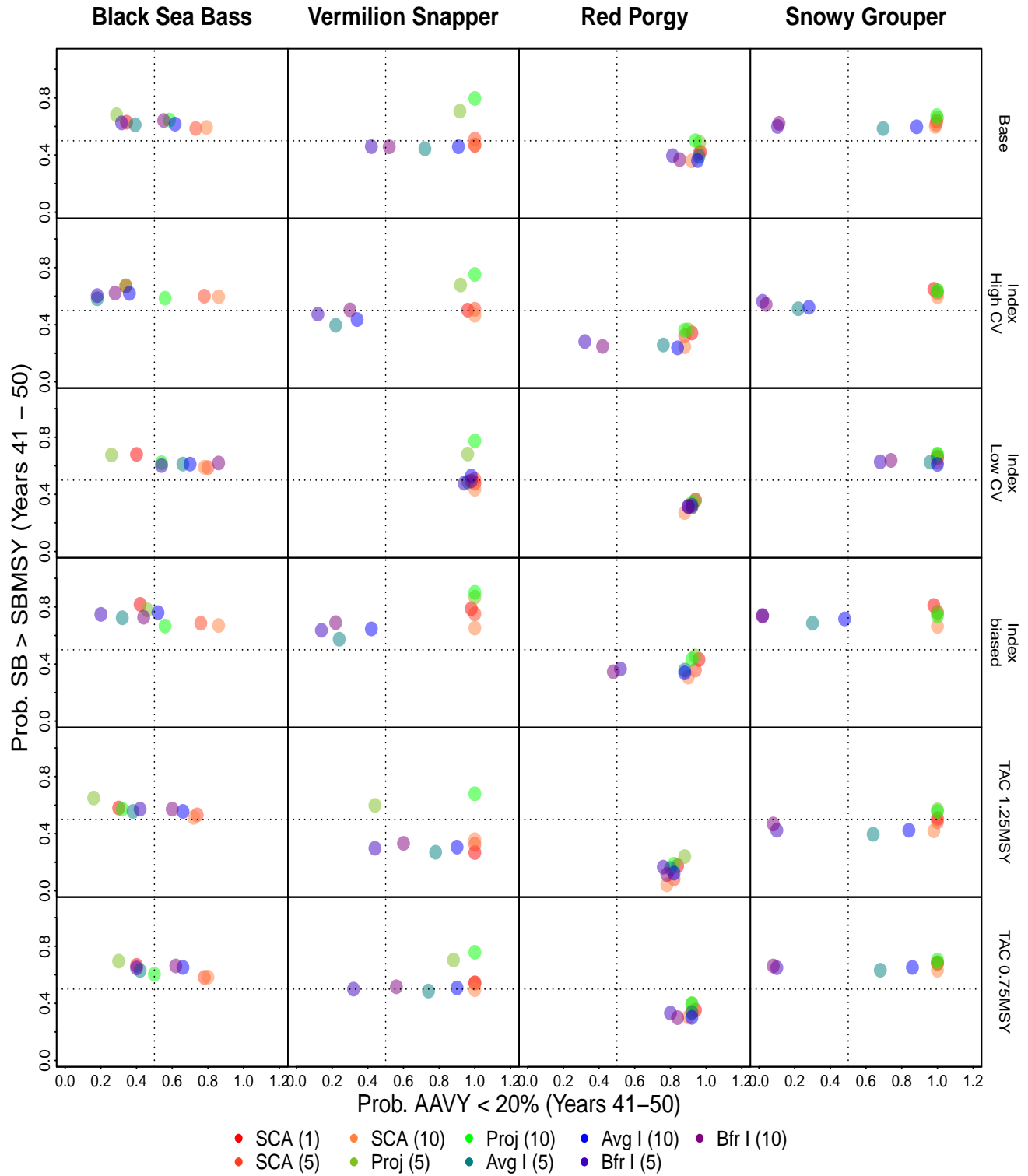


Fig. S11 Trade off plots of the probability that average annual variability in yield (AAVY) <20%, versus the probability that $SSB > SSB_{MSY}$ among simulation runs for the base and alternative scenarios. Plots are based on data for the last 10 years of the projection period.

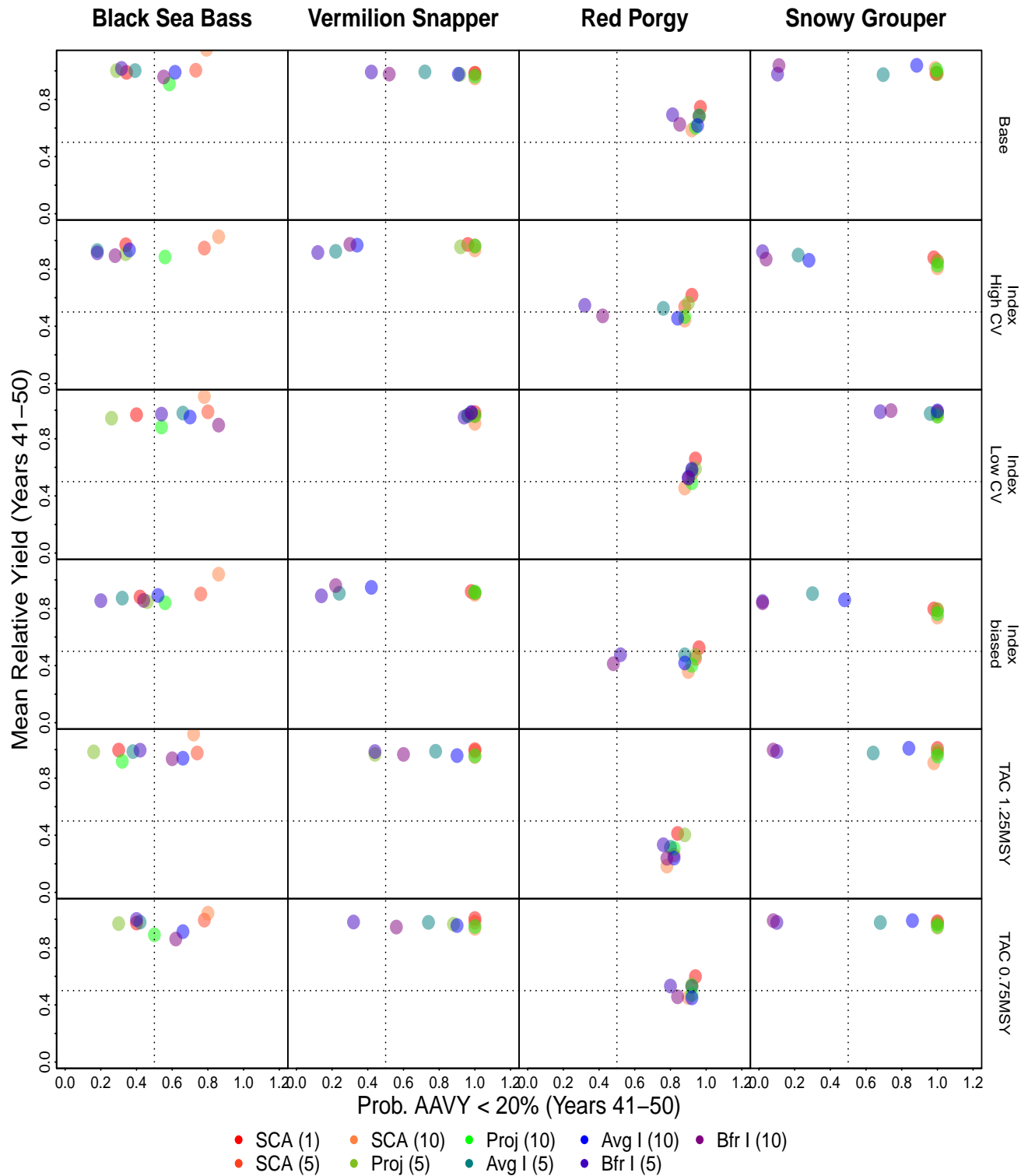


Fig. S13 Trade off plots of the probability that average annual variability in yield (AAVY) <20%, versus longterm yield among simulation runs for the base and alternative scenarios. Plots are based on data for the last 10 years of the projection period.

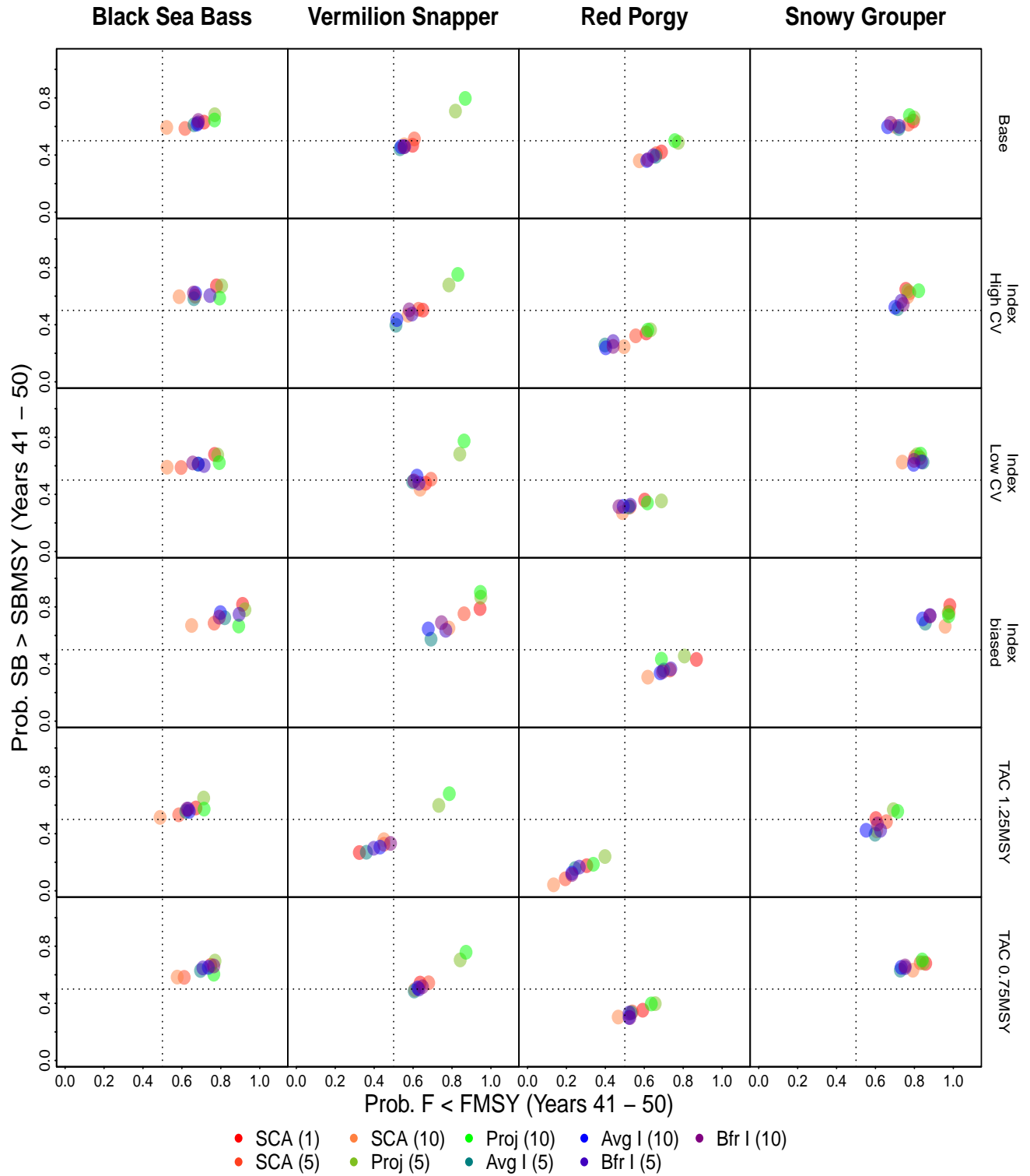


Fig. S14 Trade off plots of the probability that $F > F_{MSY}$ versus the probability that $SSB > SSB_{MSY}$ among simulation runs for the base and alternative scenarios. Plots are based on data for the last 10 years of the projection period.

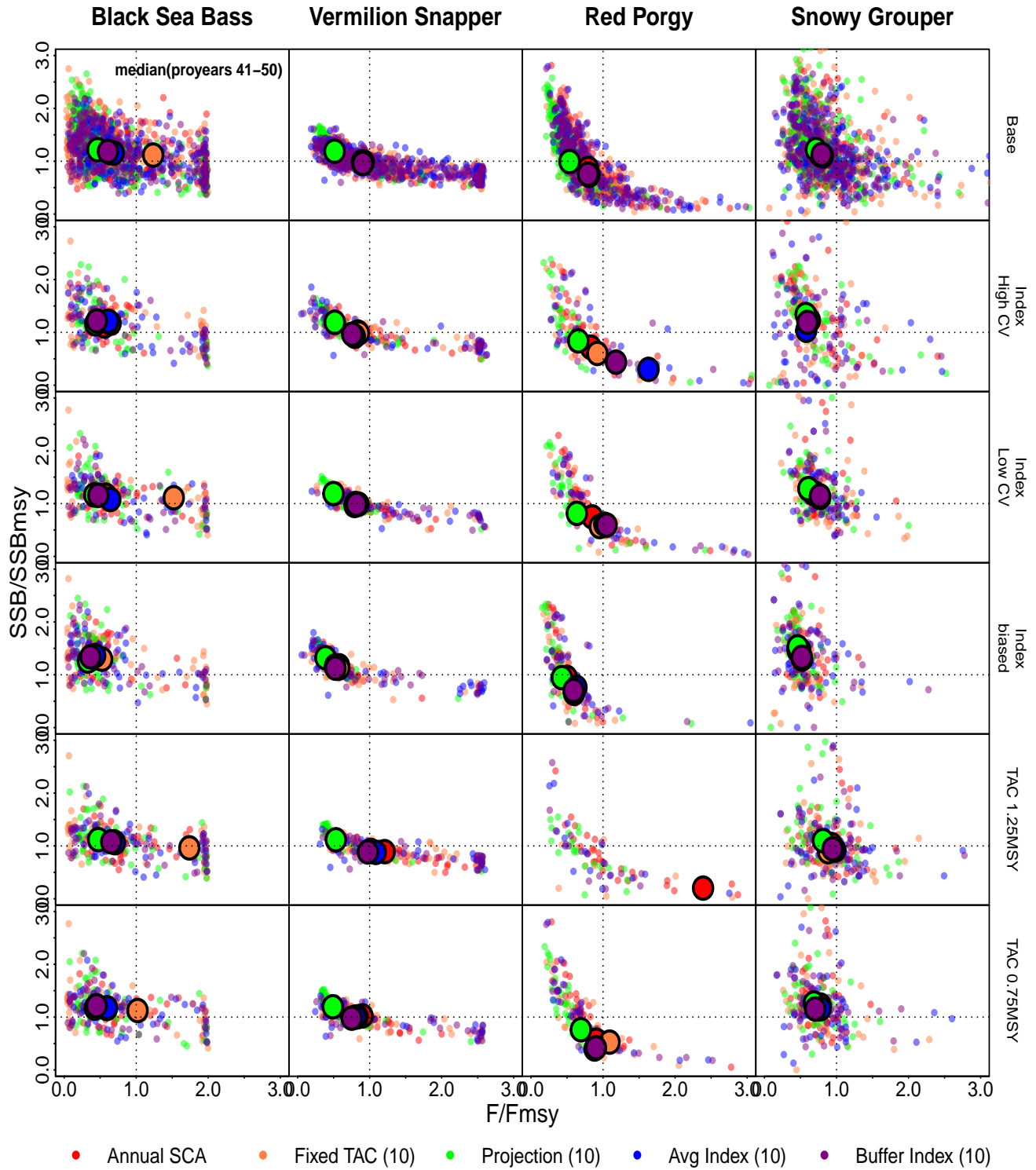


Fig. S15 Phase plots of the probability that $F > F_{MSY}$ versus the probability that $SSB > SSB_{MSY}$ among simulation runs for the base and alternative scenarios.