# Performance of catch-based and length-based methods in data-limited fisheries

Maite Pons 1, Jason M. Cope 1,2, and Laurence T. Kell 3

1 School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, Washington 98195-5020, USA

2 National Oceanic and Atmospheric Administration Fisheries, Northwest Fisheries Science Center, Fishery Resource Analysis and Monitoring Division, 2725 Montlake Boulevard East, Seattle, Washington 98112-2097, USA.

3 Centre for Environmental Policy, Imperial College London, London SW7 1NE, UK.

***Abstract***

The quality of data for many small scale fisheries and by caught species around the world is insufficient to use conventional stock assessments methods. Therefore, recently many methods have been developed to assess fisheries when data are limited. For example catch-based models can be used when only total catch data are available and length-based models when only samples of the length composition are taken from the catch. Here, we evaluated the performance of both catch-based and length-based models, using simulation testing to estimate the exploitation status of species with contrasting life histories under different harvest scenarios. For unassessed fisheries where reconstructing time series of catch is possible, catch-based methods such as Depletion Based Stock Reduction Analysis (DBSRA) and Simple Stock Synthesis (SSS) seemed to be the best approach to assess stocks. For fisheries that are still developing and time series of catch are not available, obtaining length-composition data could give a good approximation of the exploitation status of the stocks. In many of the scenarios tested, length-based models such as Length Based Spawning Potential Ratio (LBSPR) performed as well as catch-based methods.

*Keywords*: data-limited assessment methods, simulation testing, SPR, harvest rates

## Introduction

The provision of fisheries management advice requires the assessment of stock status relative to reference points, the prediction of the response of a stock to management, and checking that predictions are consistent with reality (Kell at al., 2016).,. Major commercial species usually have substantial data to inform complex stock assessments models (e.g. Methot and Wetzel 2013); this includes long time series of total removals, catch-at-age data, relative abundance indices, fishing effort, size and/or age composition, and information on life-history parameters. Most of the datasets required for these such stock assessments are unavailable however, for most small-scale fisheries around the world. Fisheries and stocks lacking these multiple data types are commonly known as “data-poor” or “data-limited” fisheries (Costello *et al.*, 2012; Dowling *et al.*, 2015). Recently, many data-limited approaches have been developed to meet an increase demand for science-based fisheries management for unassessed fisheries, stocks and species where resources are limited (Wetzel and Punt 2011; Costello *et al.* 2012; Dowling *et al.* 2015, 2016; Chrysafi and Kuparinen 2016; Rosenberg *et al.* 2017).

Assessing stocks using only catch data started many years ago with the development of Stock Reduction Analysis, (SRA;Kimura and Tagart, 1982; Kimura *et al.*, 1984). Since then, the method has been extended to estimate productivity and reconstruct historical abundance trends by making assumptions about final biomass relative to unfished or initial biomass (i.e., stock depletion; Thorson and Cope 2015). SRA has subsequently been further extended to incorporate stochastic variability in population dynamics (Stochastic-SRA; Walters *et al.* 2006), a flexible shape for the production function (Depletion-Based SRA; Dick and MacCall 2011), prior information regarding resilience and population abundance at the start of the catch time series (Catch-MSY; Martell and Froese 2013; Froese *et al.* 2017), and age-structured population dynamics (Simple Stock Synthesis; Cope 2013). Despite these differences, this family of catch-only models shares a common dependence upon prior assumptions about final stock depletion. Simulation testing indicates that these methods perform well only when assumptions regarding final relative abundance are met and although they might be appropriate to predict sustainable catch or biomass, but not to reconstruct abundance time series (Carruthers *et al.*, 2012; Wetzel and Punt, 2015).

For many small-scale fisheries, obtaining reliable information on historical total catch is difficult, while collecting length measurements from samples of the catch is easier. Mean-length mortality estimators (Beverton and Holt, 1957) assume that fishing mortality directly influences the mean length of the catch under equilibrium conditions. Length-based spawning potential ratio (LBSPR, Hordyk *et al.* 2015a) and length-based Integrated Mixed Effects (LIME, Rudd and Thorson 2017) models, have recently been developed allowing the estimation of instantaneous fishing mortality (*F*) and spawning potential ratio (SPR) when basic biological parameters are known. SPR is the proportion of the unfished reproductive potential per recruit under a given level of fishing pressure (Goodyear 1993). Both methods have the same data-requirements, but LIME does not assume equilibrium conditions; the mixed-effects aspect of LIME extends length-based methods by estimating changes in recruitment and fishing mortality over time (Rudd and Thorson, 2018).

There is increasing interest in developing new methodologies to quantitatively assess data-limited fisheries to manage them and prevent overfishing. Usually, these assessment method performances are tested using simulation experimentation (Cope, 2008). Carruthers *et al.* (2016) used a closed-loop simulation approach to compare a range of management procedures for setting catch limits in data-limited fisheries. They found that data-limited methods using observations of stock depletion offer the best overall performance across life history types, data quality and autocorrelation in recruitment strength. However, these management procedures are based on setting catch limits and were designed for use in data-limited fisheries for which annual catch data are available, sometimes together with a relative abundance index (delay-difference stock assessment, Carruthers *et al.* 2014). In many data poor fisheries, measuring total removals is difficult, as is enforcing catch limits. Hordyk *et al.* (2015b) tested some harvest strategies using a simulation approach to assess the utility of LBSPR as a tool for management in data-limited fisheries using an effort-based harvest control rule.

No studies, however, have compared the performance of both, length-based and catch-based methods as estimation models using the same simulated populations, a possible reason is because finding a common metric between catch-based and stock status metrics is difficult. Therefore we use simulation testing where an operating model (OM) is used to represents the main sources of uncertainty and to generate data for use in data limited stock assessment methods. We then

use a common metric for comparison across models, namely exploitation or harvest rates, to evaluate the performance of catch-based and length-based models using simulation testing to estimate exploitation status for three different fish stocks with contrasting life history strategies under three possible exploitation scenarios.

## Methods

Three different exploitation rate (*U*) scenarios were considered ,which correspond to historical fishing mortality histories commonly seen in many fisheries. In the first, *U* increases until it reaches a maximum and start declining afterwards; this is a classic example where management measures were implemented to reduce fishing pressure. The second scenario assumes that *U* increases and remains constant after reaching a maximum; this could be due to implementation of effort limits for example. The third scenario has constantly increasing *U*, which would occur for fisheries that are still developing.

In addition, three population life history types of varying longevity and somatic growth rate were simulated: (i) a short-lived fast-growth species, pacific chub mackerel, *Scomber japonicus*, (ii) a medium-lived medium-growth fish, albacore tuna, *Thunnus alalunga*, and (ii) a longer-lived slow-growth species, canary rockfish, *Sebastes pinniger* (Table 1).

### Operatingmodel specifications

The OMwas implemented using Stock Synthesis (SS) Version 3.30.10 (Methot and Wetzel 2013; Methot *et al.* 2018) in order to simulate age structured populations. SS assumes that the absolute level of catch is known well enough to allow the model to calculate the level of fishing intensity needed to obtain that level of catch conditioned on the model estimates of age-specific population abundance and selectivity (Methot and Wetzel, 2013). Fishing intensity in SS is estimated to match the observed catch; the harvest rate is therefore the total annual catch divided by the total abundance of the exploited biomass. Different catch histories thus lead to different exploitation histories, and the different scenarios could affect the performance of a data-limited methods. In order to simulate different exploitation histories in each run, we added a random observation error to the catch of 0.1 around the observed catch value in each year.

To condition the OM published life history parameter values (Table 1) from formal stock assessments were used (ICCAT, 2014; Crone and Hill, 2015; Thorson and Wetzel, 2015). Each population was assumed to be targeted in a single area, by only one fleet with a selectivity pattern that was logistic and constant through time. This controls for any complications arising from having multiple fleets and allow focus to be on method performance, not model specification. Each simulated population began at an unfished biomass and all catch scenarios terminate at the same stock depletion level. No indices of abundance were included, but we defined a final stock depletion implemented through the use of a survey index equal to 1 at the beginning of the time series and 0.4 in the last year, so spawning biomass (SB) in the last year is 0.4 SB0 (Cope 2013). In each OM, all parameters were fixed, except *R0* (which allow the model to set the absolute biomass of the population), and annual lognormal recruitment deviations were assumed (Table 1). A Beverton–Holt spawner–recruit function (Beverton and Holt, 1957) was assumed.

To simulate catch length frequencies, the expected age-length key output from SS was used to assign a length distribution to each age. Summing within a length bin then gave the length distribution of the catch. Length bins were defined at every 2 cm from 30 to 150 cm for Albacore, from 12 to 76 cm for Pacific Chub Mackerel and from 8 to 60 cm for Canary Rockfish. To obtain samples of the length composition of the catch to use in the length-based assessment models, 1000 fish/year were drawn using a multinomial distribution from the catch at length, using the probability of being caught at each length bin in each year.

OMs were generated using a factorial design encompassing 9 factors; namely (*i*) three scenarios of harvest rates (Figure 1 to 3) and (*ii*) life history with three levels (Table 1). For each OM 100 datasets were simulated for harvest rate (*U*), total biomass (*B*) and SPR (Figure 1 to 3).

### Estimation models

Each catch or length-based method evaluated is described in detail below:

#### *Catch-based data-limited methods*

**Catch-MSY** (CMSY; Martell and Froese 2013). It is a SRA approach with a Schaefer biomass dynamic model. As input data, it requires a time series of removals, prior ranges of the population rate of increase (*r*) and carrying capacity (*K*), and possible ranges of relative stock sizes in the final year of the time series. Probable ranges for *r* and *K* are filtered with a Monte Carlo approach to detect ‘viable’ *r-K* pairs. A parameter pair is considered ‘viable’ if the corresponding biomass trajectories calculated from a production model are compatible with the observed catches, so that the population abundance never falls below 0, and is compatible with the prior assumption of relative biomass (i.e., stock depletion; Martell and Froese 2013). The *r-K* pairs are drawn from uniform prior distributions and the Bernoulli distribution is used as the likelihood function for accepting each *r-K* pair. CMSY uses catch and productivity to estimate MSY. However, here we used the modified version of CMSY by Rosenberg *et al.* (2017) to extract biomass trends from all viable *r-K* pairs. Then the biomass trajectory is calculated as the median of all viable biomass trajectories generated during the Monte Carlo process. We used the R package *datalimited* version 0.1.0 (Anderson *et al.* 2016) available at <https://github.com/datalimited/datalimited>.

**State-space catch-only model** (SSCOM). This is a hierarchical model based on a coupled harvest-dynamics model. The model is a Bayesian state-space model that integrates across three stochastic functional forms: variation in effort, population dynamics and fishing efficiency (Thorson *et al.*, 2013). SSCOM can reconstruct biomass time series from catch data whenever fishing mortality follows semi-predictable dynamics over time. The different types of population and effort dynamics can be extracted from the same catch stream using nonlinear models for population-dynamics as a function of biomass and linear models for effort dynamics as a function of log-scaled biomass for example. We used the package *datalimited* version 0.1.0 (Anderson *et al.*, 2016) to run this model. We modified the code to extract biomass trajectories and to use a lognormal distribution for depletion (Table 1). However, the effort dynamic priors were set as in Anderson *et al.* (2017).

**Depletion based stock reduction analysis** (DBSRA). DBSRA (Dick and MacCall, 2011) modifies the SRA approach as it uses Monte Carlo draws from four parameter distributions (*M*, *FMSY/M*, *BMSY/B0* and *depletion*) while using age at maturity (*Amat*) to separate the biomass into immature and mature biomass (fishery selectivity is also assumed to have an identical pattern to the age-at-maturity ogive). It uses a delay-difference production model with a time lag for recruitment and mortality as:

where *Bt* is the biomass at the start of the year *t*, *M* is the instantaneous rate of natural mortality, and is the latent annual production based on a function of adult biomass in year *t-Amat*. Biomass in the first year (*B0*) is assumed equal to *K*. The package *fishmethods* version 1.10-3 was used to perform this analysis (Nelson, 2017).

**Simple Stock Synthesis** (SSS). This method is based on the Stock Synthesis package (Methot and Wetzel, 2013). Like DBSRA, SSS uses Monte Carlo draws of *M*, steepness (*h*), and initial recruitment (*R0*) while fitting to an artificial abundance survey representing stock depletion (Cope, 2013). All fixed values are drawn from prior distributions to represent uncertainty in model-derived outputs. The code for running SSS can be found at <https://github.com/shcaba/SSS>.

#### *Length-based data-limited methods*

**Length based spawning potential ratio** (LBSPR). In LBSPR, SPR in an exploited population is a function of the ratio of fishing mortality to natural mortality (*F/M*), and the two life history ratios *M/K* and *Lm/L∞*; *k* is the von Bertalanffy growth coefficient, *Lm* is the size of maturity and *L∞* is asymptotic size (Hordyk *et al.*, 2015a). The inputs to LBSPR are: *M/k*, *L∞*, the variability of length-at-age (*CVL∞*), which is normally assumed to be around 10%; and length at maturity specified in terms of *L50* and *L95* (the size at which 50% and 95% of a population matures). Given the assumed values for the *M/K* and *L∞* parameters and length composition data are from an exploited stock, the LBSPR model uses maximum likelihood methods to estimate the selectivity ogive, which is assumed to be a logistic curve defined by the selectivity-at-length parameters *S50* and *S95* (the size at which 50% and 95% of a population is retained by the fishing gear), and the relative fishing mortality (*F/M*), and these are used to calculate SPR (Hordyk *et al.*, 2015 a,b). Estimates of SPR are primarily determined by the length of fish relative to *L*50 and *L∞*. Dependent on the selectivity, if a reasonable proportion of fish in a sample attain sizes approaching *L∞*, a high estimate of SPR will be derived. LBSPR is an equilibrium based method with the following assumptions: (i) asymptotic selectivity, (ii) growth is adequately described by the von Bertalanffy equation, (iii) a single growth curve can be used to describe both sexes which have equal catchability, (iv) length at-age is normally distributed, (v) rates of natural mortality are constant across adult age classes, (vi) recruitment is constant over time, and (vii) growth rates remain constant across the cohorts within a stock (Hordyk *et al.*, 2015a). Analyses were conducted using LBSPR package version 0.1.2 in R (Hordyk 2017).

**Length-based integrated mixed effects** (LIME). The model uses length data and biological information to estimate *F* and SPR. LIME has the same data-requirements as LBSPR, but does not assume equilibrium conditions; the mixed effects aspect of LIME extends length-based methods by estimating changes in recruitment and fishing mortality over time (Rudd and Thorson, 2018). LIME uses automatic differentiation and Laplace approximations as implemented in Template Model Builder (TMB; Kristensen *et al*. 2016) to calculate the marginal likelihood for the mixed-effects. All other assumptions are the same as LBSPR but LIME estimates one selectivity curve for the entire time series of length data while LBSPR estimates one selectivity curve for each year since each time step estimation in LBSPR is independent (Hordyk *et al.*, 2015a).

### Comparing methods outputs

One of the challenges when comparing catch and length-based methods is they produce different model outputs. Catch-only models estimate total and/or spawning stock biomass and sustainable catches, whereas length-based models estimate exploitation and transient SPR, which is similar to relative stock status. These are fundamentally different measures of the population. Given the challenge of having different metrics, the performance of a method can be compared to the OM and described as the Relative Error (RE), where RE = (*estimated-true*)/*true*. This allows the measure of uncertainty, in both bias and precision, in the methods under each scenario, and is used as a standardized metric of model performance. Bias in this study is how far, on average, the performance measure from each estimation model is from the true value. Imprecision is related to the variability around that estimated average value. We used as a performance the exploitation rate (*U*) catch/biomass for the catch-based models and, for the length-based models, to scale the estimated fishing mortality (*F*) between 0 and 1 we used this transformation: *U* = 1 - *e* (- *F*). In addition, we presented the average RE across a period of time equal to the generation length of each species (see Table 1).

## Results

### Differences among harvest rate scenarios

In general, catch-based models were more biased and less precise when there was no contrast in the time series of catch data (Scenario 2 and 3, Figure 1 to 3). For example, for the fast-growth mackerel, the median RE for CMSY in Scenario 1 was 1.3 (range: 1.1 to 1.5), but it increased to 3.2 in Scenario 3 (range: 2.1 to 4.2). However, LIME showed similar patterns in exploitation rates in both bias and precision for each species across scenarios and LBSPR seems to be more positive biased in Scenario 1 than in the Scenarios 2 and 3 for all species (Figure 4).

### Catch-based models

We expected that the various catch based models considered in this study to perform differently since they have different model structure and assumptions. SSS and DBSRA were, in general, less biased and more precise than other catch-based models, in particular when comparing with both catch-MSY models. CMSY showed the worst performance among all catch-based models tested in this study. CMSY presented the highest RE for mackerel in Scenario 3 (median RE=3.1) and the lowest RE for Albacore in Scenario 1 (median RE=0.49). Even in this case, harvest rates were estimated to be 50% higher than the truth. CMSY was less biased when the catch time series had an increasing followed by a decreasing trend (Scenario1) and highly biased when there was no contrast in the time series of catch (Scenario 2 and 3) for the three species (Figure 4).

SSCOM in general was less biased than CMSY and in some cases less biased than DBSRA or SSS (i.e. for mackerel). However, SSCOM was less precise than any other catch-based model, showing a broader range of RE in most of the cases (Figure 4).

SSS estimated unbiased exploitation rates across different scenarios of harvest trends and life histories, except for the case of the fast-growth mackerel. It was positively biased in those cases; the harvest rates were estimated to be 66% higher than the true values (median RE=0.66) in Scenario 1 and around 38% in Scenario 2 (median RE=0.38) and 30% Scenario 3 (median RE=0.30). For this species, DBSRA and SSCOM was less biased in the three scenarios. SSS was the less biased estimation method for Albacore and Rockfish for the three scenarios. The median RE range between -0.02 in Scenario 3 to 0.15 in Scenario 1 for Albacore, and between -0.09 in Scenario 3 and 0.15 in Scenario 1 for Rockfish. DBSRA seemed to be more sensitive to the different trends in harvest rates for the medium-growth Albacore (Figure 4).

In summary, among catch-based models, CMSY was the least precise and most positively biased, particularly in Scenarios 2 and 3 for the three species. SSCOM was also less precise but less biased than the CMYS and DBSRA in general. The method that performed best in terms of bias and precision was SSS most of the time (Figure 4). The age-structured aspect of SSS has also been shown elsewhere to be better suited for slower life histories (Wetzel and Punt 2015).

### Length-based models

In some cases, length models gave a less biased estimation than catch-based models. LBSPR was generally less biased for slow-growth species like rockfish and highly positively biased for fast-growth species like mackerel. LIME was positively biased in all cases and highly imprecise in general but in particular for the slow-growth species (Figure 4).

LIME did not converge in many cases, however, between 32% of the times in Scenario 1 and 9% in Scenario 3 for Albacore, between 27% in Scenario 1 and 4% in Scenario 2 for Mackerel, and between 61% in Scenario 1 and 67% in Scenario 3 for Rockfish. In all cases, runs that did not converge were not included. For slow-growth species LIME had more difficulties in converging and in this case also showed the greateest imprecision in estimates of harvest rates (Figure 4).

In summary, between the 2 length-based models, LBSPR was more precise than LIME in general. Both showed similar performance for slow-growth and medium-growth species and very different and opposite performance for fast-growth species. For short lived species LIME was less biased. On the other hand, for slow growth species LBSPR was less biased (Figure 4).

In Scenarios 2 and 3, where there is no contrast in the time series of catch, length-based methods performed better than CMSY. But, in general, all data-limited models tested here performed worse for the fast-growth species (Figure 4).

## Discussion

Simulation studies commonly use different model specifications from those of the methods to allow robustness to be evaluated; in some cases, however, often the same population model is used for simulation and estimation, i.e. self-testing. Using the same model for simulation and estimation could result in optimistic results that might not be true under many scenarios (Francis, 2012). For example it is not possible to explore the robustness to model structure and assumptions when the model used for simulation and estimation is the same. Our approach evaluated multiple data-limited assessment methods that assume different population dynamics, uncertainty and fishing effort dynamics. It is to be expected, due to these differences, that the various methods would performed differently. Rosenberg *et al.* (2017) used four catch-based data limited models and found that frequently models disagreed about population status estimations with no model showing good performance across all fish stocks, i.e. high precision and low bias, and the performance of the models depended on the scenario considered. When scenarios are chosen to represent uncertainty about the dynamics this infers that the method may not be robust and when a scenario represents a particularly stock or fishery then it may be difficult to draw any general conclusions. It was found that model performance is highly dependent on the life history of the species of concern, the dynamic of the population and the fishing intensity. This is why scenarios were chosen to represent different harvest rates and species with contrasting life histories.

In general, catch-based models were less biased and more precise when there was contrast (e.g., an increase in the catch and then a decrease) in the time series of catch data. Walters *et al.* (2006) suggested that for SRA, stocks that have experienced extensive historical depletion gains precision from a high rate of rejected parameter draws. In Scenario 3, where catch is still increasing, it is very difficult to have a good estimate of the carrying capacity *K*. So, SSS and DBSRA, which use priors in *FMSY/M* and *BMSY/B0* performed better than the models that only relied in *r* and *K,* even with priors for depletion were centered in the true values. Length-based models, on the other hand, were not dependent on the harvest rate trends. This is not surprising for LBSPR since in equilibrium conditions, the estimates are snapshots of the population and independent in each time step. For LIME, which is not an equilibrium model, Rudd and Thorson (2017) also did not find strong differences for alternative fishing mortality scenarios.

The choice of a “best model” also depend on the life history characteristics and the biological information that is available. SSS seems to be the least biased catch-based model. However, unlike other catch-based models, age and growth estimates are needed in SSS to define age structure and remove catch according to age-/size-based selectivity patterns (Cope, 2013). SSS has the same structure as SS and this might be the reason why is the model that performed the best in most cases since the simulation and estimation models have the same parametrization (Francis, 2012).

DBSRA and SSS performed very similar in some cases (Albacore Scenario 1, Mackerel Scenario 2 and 3). In structure, both models are very similar, however there are a few notable differences between the population dynamics models used in DBSRA and SSS that could explain the different results found here (Wetzel and Punt, 2015). The underlying population dynamics model in SSS is fully age-structured whereas DBSRA uses a delay-difference model based on a biomass production function.

We found that SSCOM and DBSRA performed similarly in terms of bias. In the SSCOM model, a prior in depletion is not needed, but it can be included as we did in this study. Thorson *et al.* (2013) explored the effect of specifying a prior on final depletion and compared the results with DBSRA. They suggested that using a strong prior on final depletion in SSCOM would result in similar performance to DBSRA. Both, DBSRA and SSCOM approximate biomass dynamics using a production function expressed as exploitable biomass (which is equivalent to spawning biomass given selectivity and maturity curves are assumed identical), and both assume that biomass starts at average unfished biomass. However, DBSRA uses deterministic biomass dynamics and uses an asymmetric production function (Dick and MacCall, 2011), while SSCOM has stochastic biomass dynamics and uses a Schaefer production function (Thorson *et al.*, 2013), so it unsurprising they did not performed exactly the same. SSCOM was less biased in some cases, but always less precise than DBSRA. Specifying other priors in SSCOM in future studies, for example for effort-dynamics, could increase its precision.

CMSY performed very poorly in all scenarios overestimating harvest rates, even when given a prior for depletion close to the true value. A key point of the CMSY is the ability to define a reasonable prior range for the parameters of the Schaefer model in particular *K*. In our case, we have arbitrarily chosen 50 times the maximum catch as the upper bound for *K*. However, in the Scenario 3, in a developing fishery, or a fishery that has a continuous increase in catch, it will be more difficult to define the upper bound of *K* because the maximum potential has yet to be reached (Martell and Froese, 2013), thus limiting the performance of these methods under this scenario. However, they also performed poorly in Scenario 1 and 2, in particular for long-lived and short-lived species. Other *K* values could be explored in future studies to see if this improves the outcomes, but it remains a very difficult parameter to specify. For example, Martell and Froese (2013) used maximum catch multiplied by 100. Rosenberg et al. (2017) and Free et al. (2017) found that CMSY was the one that performed second best and better than SSCOM in their scenarios. One of the differences with our study is that they considered a uniform prior for depletion in SSCOM and we considered a Lognormal prior centered around the true value, but it is apparent that method performance is sensitive to a variety of scenarios.

For long-lived species, recruitment variability does not affect the length composition of the catch as much as for short lived species. This is why LBSPR performed pretty well for long-lived species and it was highly biased for short-lived species. LIME however, performed better for short-lived species than LBSPR being able to capture changes in the length composition due to recruitment variability. In general, all catch-based and length-based methods seems to perform worse for the faster life history types.

The present study does not look at parameter misspecification, but correct specification (unbiased) in the life histories parameters and known catch histories. With that level of information, the length-based models like LBSPR showed sometimes better performance than the catch-based models, as the latter were more sensitive to the catch history scenarios, and the length based methods were able to integrate the catch scenarios into the length compositions.

### Conclusions

For unassessed fisheries where data are limited, but reconstructing time series of catch is possible, catch-based methods SSS or DBSRA provided the most reliable outputs to for management. However, to apply SSS and DBSRA, not only catch data is needed, these methods also require extensive prior information, such as growth, maturity, *FMSY/M* (Zhou *et al.*, 2012)and *BMSY/B0* (Thorson *et al.*, 2012) parameters. When this prior information is not available, SSCOM could be use with a good prior for depletion. CMSY could also be considered, but with caution, because it has been proved here that they can be highly biased and influenced by catch trends and uncertainty in *K*.

For fisheries that are still developing, where the time series of catch are unavailable, getting length-composition data could give a good approximation of the status of the stock, in particular for medium to long-lived species. It has been shown here that, in some cases, that length-based models can give the same or less biased estimates of exploitation status than catch-based models.

### Future directions

There has been an emerging field of catch methods and assemble of catch-based methods to estimate global stock status (Costello *et al.*, 2012; Anderson *et al.*, 2017; Rosenberg *et al.*, 2017). The super-ensemble method published by Anderson *et al.* (2017) allows for weighting individual models based on their accuracy. In their study, some models had different assumptions about uncertainty and the dynamics of fishing effort, but all assumed the same population dynamics. A new super-ensemble method that includes models that also assume different population dynamics could be developed in the future based on our results. It is important, however, that the behavior of the models in the ensemble are well understood, i.e. their bias, precision and convergence properties. Combining estimates from different methods in a consistent reproducible manner may provide more stability in the advice for managers. The best performing methods can then be tested using management strategy evaluation (MSE) to specify Management Procedures (MPs) that can help ensure robust and sustainable fisheries management. Where a MP is the combination of pre-defined data, together with an algorithm to which such data are input to provide a value for a management control measure. The importance of considering assessment methods as part of a MP is that a method that provides biased estimates with high precision may be better for setting management regulations than a unbiased but imprecise estimator. Also if a method only provides estimates of exploitation level or MSY then management controls may be different, i.e. based on a total allowable catch (TAC) or effort. MSE also allows a broad range of uncertainty, since traditional stock assessment and advice based upon it, mainly considers measurement and process error when uncertainty about the actual dynamics has a larger impact on achieving management objectives (Punt, 2008) \citep{punt2008refocusing}.

(Carruthers *et al.* 2014, 2016) considering, in addition to other uncertainty, model uncertainty. For example although

***Acknowledgments***

The Walton Family Foundation generously funded MP. LK was funded by the Irish exchequer and the European Maritime and Fisheries Fund (EMFF) under the MyDas project. In addition, we would like to acknowledge Merrill Rudd, Chantel Wetzel, Ian Taylor, Andrew Rosenberg, Sean Anderson, Chris Free and James Thorson for their help with SS, r4ss, and other packages and models used in this study.

References

Anderson, S. C., Afflerbach, Jamie Cooper, A. B., Dickey-Collas, M., Jensen, O. P., Kleisner, K. M., Longo, C., Osio, G. C., *et al.* 2016. datalimited: Stock Assessment Methods for Data-limited Fisheries. R package version 0.1.0. https://github.com/datalimited/datalimited.

Anderson, S. C., Cooper, A. B., Jensen, O. P., Minto, C., Thorson, J. T., Walsh, J. C., Afflerbach, J., *et al.* 2017. Improving estimates of population status and trend with superensemble models. Fish and Fisheries, 18: 732–741.

Beverton, R. J. H., and Holt, S. J. 1957. On the Dynamics of Exploited Fish Populations. Fishery Investigations (Great Britain, Ministry of Agriculture, Fisheries, and Food), London. 536 pp.

Carruthers, T. R., Walters, C. J., and McAllister, M. K. 2012. Evaluating methods that classify fisheries stock status using only fisheries catch data. Fisheries Research, 119: 66–79.

Carruthers, T. R., Punt, A. E., Walters, C. J., MacCall, A., McAllister, M. K., Dick, E. J., and Cope, J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fisheries Research, 153: 48–68. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2013.12.014.

Carruthers, T. R., Kell, L. T., Butterworth, D. D. S., Maunder, M. N., Geromont, H. F., Walters, C., McAllister, M. K., *et al.* 2016. Performance review of simple management procedures. ICES Journal of Marine Science, 73: 464–482.

Chrysafi, A., and Kuparinen, A. 2016. Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. Environmental Reviews, 1: 1–44.

Cope, J. M. 2008. Issues and Advances in Data-Limited Stock Assessment: Experimentation through Simulation. Dissertation - Doctor of Philosophy degree: 1–235.

Cope, J. M. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fisheries Research, 142: 3–14. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2012.03.006.

Costello, C., Ovando, D., Hilborn, R., Gaines, S. D., Deschenes, O., and Lester, S. E. 2012. Status and solutions for the world’s unassessed fisheries. Science (New York, N.Y.), 338: 517–20. http://www.ncbi.nlm.nih.gov/pubmed/23019613 (Accessed 10 July 2014).

Crone, P. R., and Hill, K. T. 2015. Pacific mackerel (*Scomber japonicus*) stock assessment for USA management in the 2015-16 fishing year. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, Oregon 97220, USA: 131.

Dick, E. J., and MacCall, A. D. 2011. Depletion-Based Stock Reduction Analysis: A catch-based method for determining sustainable yields for data-poor fish stocks. Fisheries Research, 110: 331–341. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2011.05.007.

Dowling, N. A., Dichmont, C. M., Haddon, M., Smith, D. C., Smith, A. D. M., and Sainsbury, K. 2015. Empirical harvest strategies for data-poor fisheries: A review of the literature. Fisheries Research, 171: 141–153. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2014.11.005.

Francis, C. 2012. Reply to ‘The reliability of estimates of natural mortality from stock assessment models’. Fisheries Research, 119–120: 133–134. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2011.12.005.

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

Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015a. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries, 72: 217–231.

Hordyk, A. R., Loneragan, N. R., and Prince, J. D. 2015b. An evaluation of an iterative harvest strategy for data-poor fisheries using the length-based spawning potential ratio assessment methodology. Fisheries Research, 171: 20–32. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2014.12.018.

ICCAT. 2014. Report of the 2013 ICCAT North and South Atlantic Albacore stock assessment meeting. Collective Volume of Scientific Papers ICCAT, 70: 830–995.

Kell, L.T., Kimoto, A. and Kitakado, T., 2016. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries research, 183, pp.119-127.

Kimura, D. K., and Tagart, J. V. 1982. Stock Reduction Analysis, Another Solution to the Catch Equations. Canadian Journal of Fisheries and Aquatic Sciences, 39: 1467–1472. http://www.nrcresearchpress.com/doi/abs/10.1139/f82-198.

Kimura, D. K., Balsiger, J. W., and Ito, D. H. 1984. Generalized Stock Reduction Analysis. Canadian Journal of Fisheries and Aquatic Sciences, 41: 1325–1333. http://www.nrcresearchpress.com/doi/abs/10.1139/f84-162.

Martell, S., and Froese, R. 2013. A simple method for estimating MSY from catch and resilience. Fish and Fisheries, 14: 504–514. http://doi.wiley.com/10.1111/j.1467-2979.2012.00485.x (Accessed 2 March 2016).

Methot, R. D., and Wetzel, C. R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research, 142: 86–99. Elsevier B.V. http://linkinghub.elsevier.com/retrieve/pii/S0165783612003293.

Methot, R. D., Wetzel, C. R., and Taylor, I. 2018. Stock Synthesis User Manual. V3.30.10: 1–186.

Nelson, G. A. 2017. fishmethods: Fishery Science Methods and Models in R. R package version 1.10-3.: <https://CRAN.R-project.org/package=fishmethods>.

Punt, A.E., 2008. Refocusing stock assessment in support of policy evaluation. Fisheries for Global Welfare and Environment, pp.139-152.

Rosenberg, A. A., Kleisner, K. M., Afflerbach, J., Anderson, S. C., Dickey-Collas, M., Cooper, A. B., Fogarty, M. J., *et al.* 2017. Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries Around the World. Conservation Letters, 00: 1–9. http://doi.wiley.com/10.1111/conl.12363.

Rudd, M. B., and Thorson, J. T. 2018. Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1019–1035.

Thorson, J. T., Cope, J. M., Branch, T. A., Jensen, O. P., and Walters, C. J. 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Canadian Journal of Fisheries and Aquatic Sciences, 69: 1556–1568. http://www.nrcresearchpress.com/doi/abs/10.1139/f2012-077.

Thorson, J. T., Minto, C., Minte-Vera, C. V., Kleisner, K. M., Longo, C., and Jacobson, L. 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Canadian Journal of Fisheries and Aquatic Sciences, 70: 1829–1844. http://apps.webofknowledge.com.proxy.bib.uottawa.ca/full\_record.do?product=WOS&search\_mode=CitingArticles&qid=4&SID=2A4pkoBXbeeUndmjzAG&page=1&doc=3.

Thorson, J. T., and Cope, J. M. 2015. Catch curve stock-reduction analysis: An alternative solution to the catch equations. Fisheries Research, 171: 33–41. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2014.03.024.

Thorson, J. T., and Wetzel, C. 2015. The status of canary rockfish (Sebastes pinniger) in the California Current in 2015. Pacific Fishery Management Council: 22.

Walters, C. J., Martell, S. J. D., and Korman, J. 2006. A stochastic approach to stock reduction analysis. Canadian Journal of Fisheries and Aquatic Sciences, 63: 212–223. http://www.nrcresearchpress.com/doi/abs/10.1139/f05-213.

Wetzel, C. R., and Punt, A. E. 2011. Model performance for the determination of appropriate harvest levels in the case of data-poor stocks. Fisheries Research, 110: 342–355. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2011.04.024.

Wetzel, C. R., and Punt, A. E. 2015. Evaluating the performance of data-moderate and catch-only assessment methods for U.S. west coast groundfish. Fisheries Research, 171: 170–187. Elsevier B.V. http://dx.doi.org/10.1016/j.fishres.2015.06.005.

Zhou, S., Yin, S., Thorson, J. T., Smith, A. D. M., Fuller, M., and Walters, C. J. 2012. Linking fishing mortality reference points to life history traits: an empirical study. Canadian Journal of Fisheries and Aquatic Sciences, 69: 1292–1301. http://www.nrcresearchpress.com/doi/abs/10.1139/f2012-060.

Table 1. Life history information and priors for the three species used in the study. Notation: *Lognormal* (µ, σ2); Uniform *U* (a, b). Priors for *K* were Uniform between the maximum catch in the time series and 50 times the maximum catch. \* Generation length (GL) calculated by Stock Synthesis. \*\* For CMSY the depletion priors were Uniform centered in the true value with a minimum of *true - 0.1* and a maximum of *true + 0.1*.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Operating model inputs** | **Symbol** | **Pacific Chub Mackerel** | **Albacore tuna** | **Canary Rockfish** |
| Maximum age | *Agemax* | 12 | 15 | 64 |
| Generation length (years) \* | *GL* | 2 | 5 | 16 |
| Length were 50% of the fish are mature (FL cm) | *L50* | 29 | 90 | 55 |
| Length were 95% of the fish are mature (FL cm) | *L95* | 34 | 100 | 57 |
| Length-weight scaling parameter | *a* | 2.73x10-6 | 1.34x10-5 | 1.80x10-5 |
| Length-weight allometric parameter | *b* | 3.444 | 3.107 | 3.094 |
| Von Bertalanffy Brody growth coefficient (1/years) | *k* | 0.40 | 0.21 | 0.14 |
| Von Bertalanffy asymptotic length (cm) | *L∞* | 38.2 | 122.2 | 60.0 |
| Theoretical age at length=0 | *t0* | -0.6 | -1.3 | -1.9 |
| Coefficient of variation length at age for all ages | *CVL* | 0.1 | 0.1 | 0.1 |
| Natural mortality (1/years) | *M* | 0.60 | 0.30 | 0.05 |
| Relationship between *M* and *k* | *M/k* | 1.50 | 1.40 | 0.35 |
| Steepness | *h* | 0.5 | 0.9 | 0.8 |
| Selectivity at 50% (cm) | *S50* | 18 | 78 | 42 |
| Selectivity at 95% (cm) | *S95* | 25 | 90 | 47 |
| Depletion | *XB0* | 0.4 | 0.4 | 0.4 |
| Survey or depletion standard error | *σS* | 0.01 | 0.01 | 0.01 |
| Observation error in catch | *σC* | 0.1 | 0.1 | 0.1 |
| Recruitment variations | *σR* | 0.3 | 0.4 | 0.5 |
| **Estimation models prior distributions** |  |  |  |  |
| Depletion | *XB0* | *Lognormal* (true, 0.1) \*\* | *Lognormal* (true, 0.1) \*\* | *Lognormal* (true, 0.1) \*\* |
| Carrying capacity | *K* | *U* (max(catch), max(catch)x50) | *U* (max(catch), max(catch)x50) | *U* (max(catch), max(catch)x50) |
| Population rate of increase | *r* | *U* (0.8, 1.2) | *U* (0.2, 0.6) | *U* (0.05, 0.4) |
| Vulnerability | *FMSY/M* | *U* (0, 2) | *U* (0, 2) | *U* (0, 2) |
| Compensation | *BMSY/B0* | *U* (0, 1) | *U* (0, 1) | *U* (0, 1) |

## Figures

Figure 1. Time series of catch, harvest rate (*U*), biomass (*B*) and spawning potential ratio (SPR) for each simulated Pacific Chub Mackerel population for the three harvest rate Scenarios tested. The color lines represent the median value for all runs.

Figure 2. Time series of catch, harvest rate (*U*), biomass (*B*) and spawning potential ratio (SPR) for each simulated North Atlantic Albacore population for the three harvest rate Scenarios tested. The color lines represent the median value for all runs.

Figure 3. Time series of catch, harvest rate (*U*), biomass (*B*) and spawning potential ratio (SPR) for each simulated Canary rockfish population for the three harvest rate Scenarios tested. The color lines represent the median value for all runs.

Figure 4. Exploitation rate relative error for all the catch-based (light colors) and length-based (dark colors) models considered under the three harvest scenarios for the three species. First row: Scenario 1 – ramp shape harvest rate. Second row: Scenario 2 – constant harvest rate. Third row: Scenario 3 – increasing harvest rate.

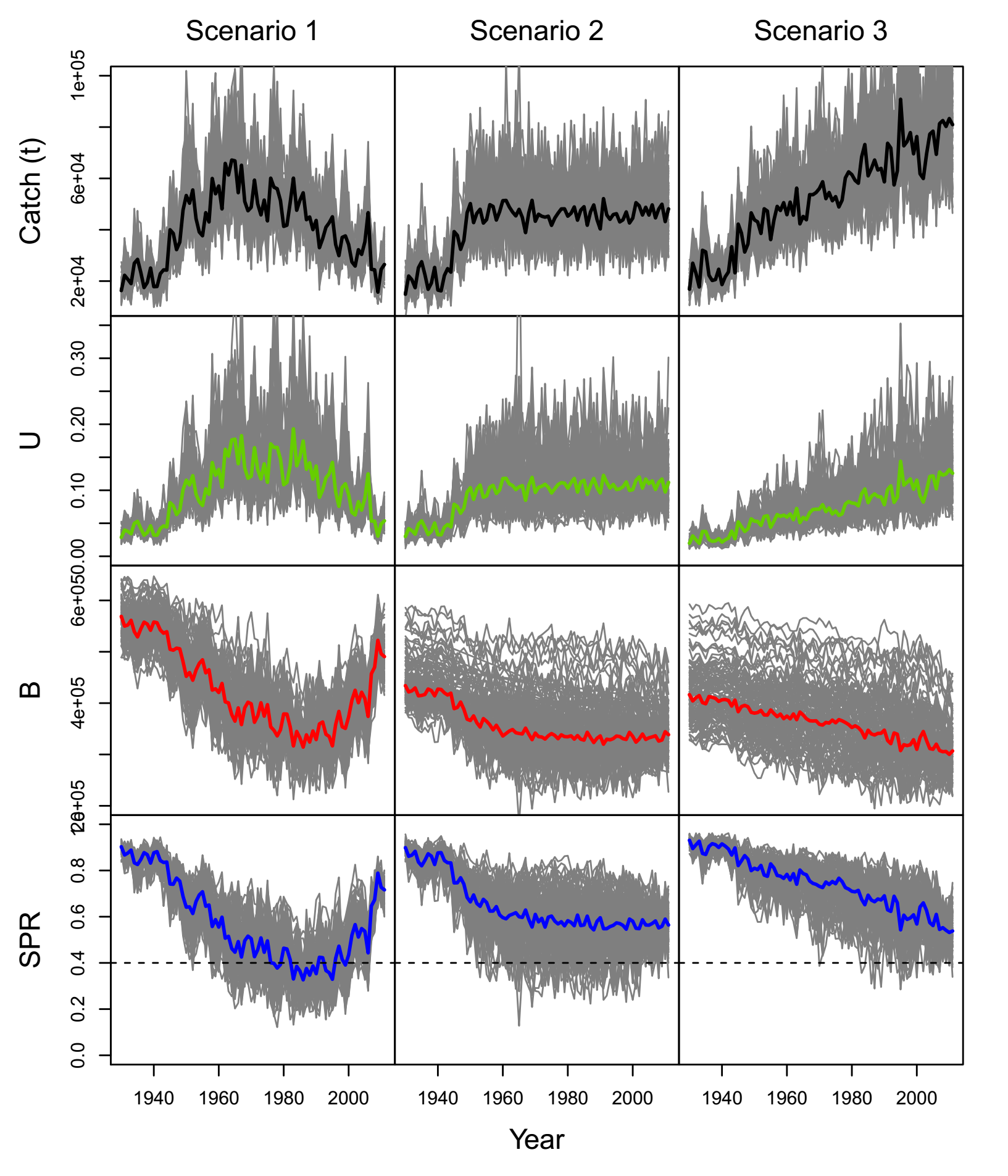


Figure 1.

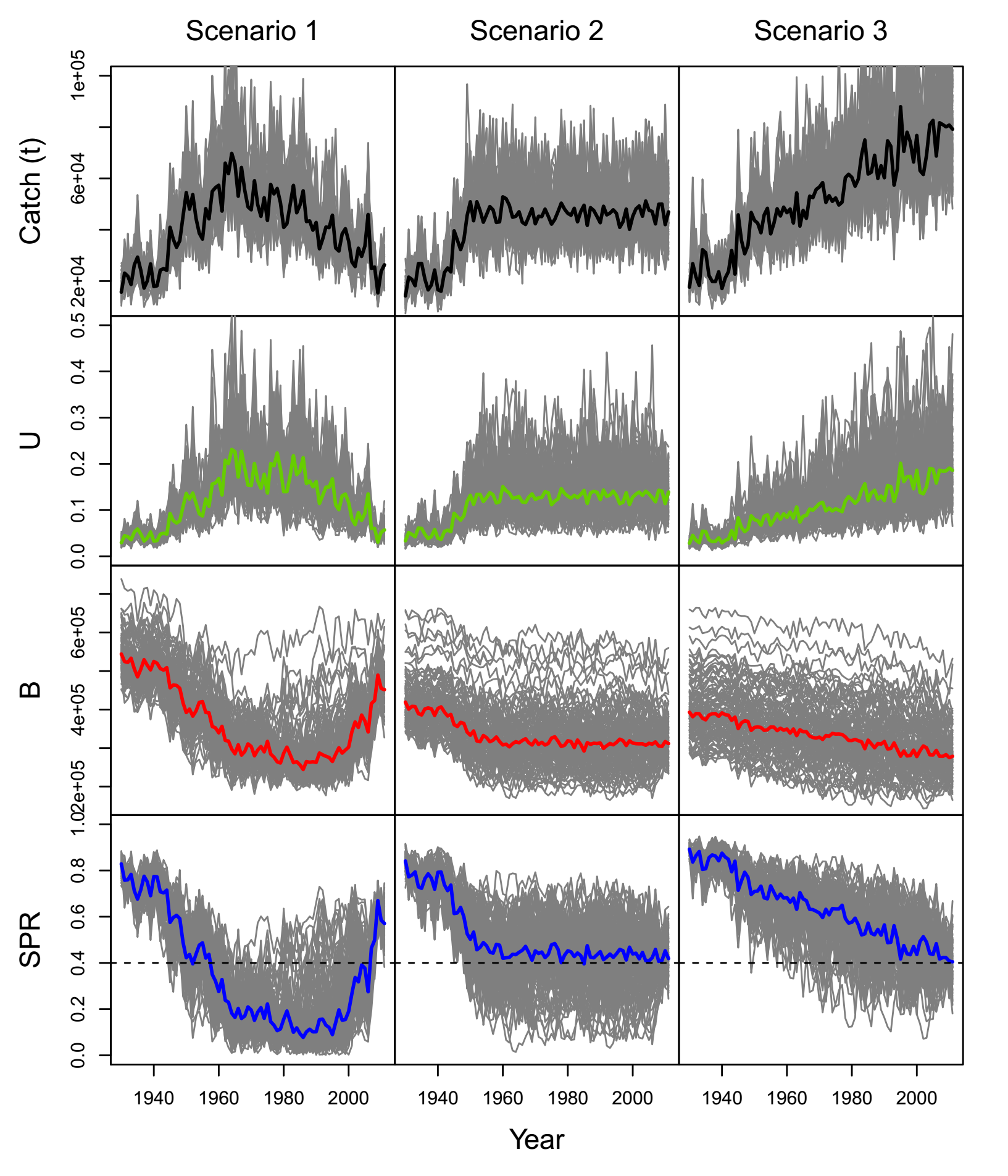


Figure 2.

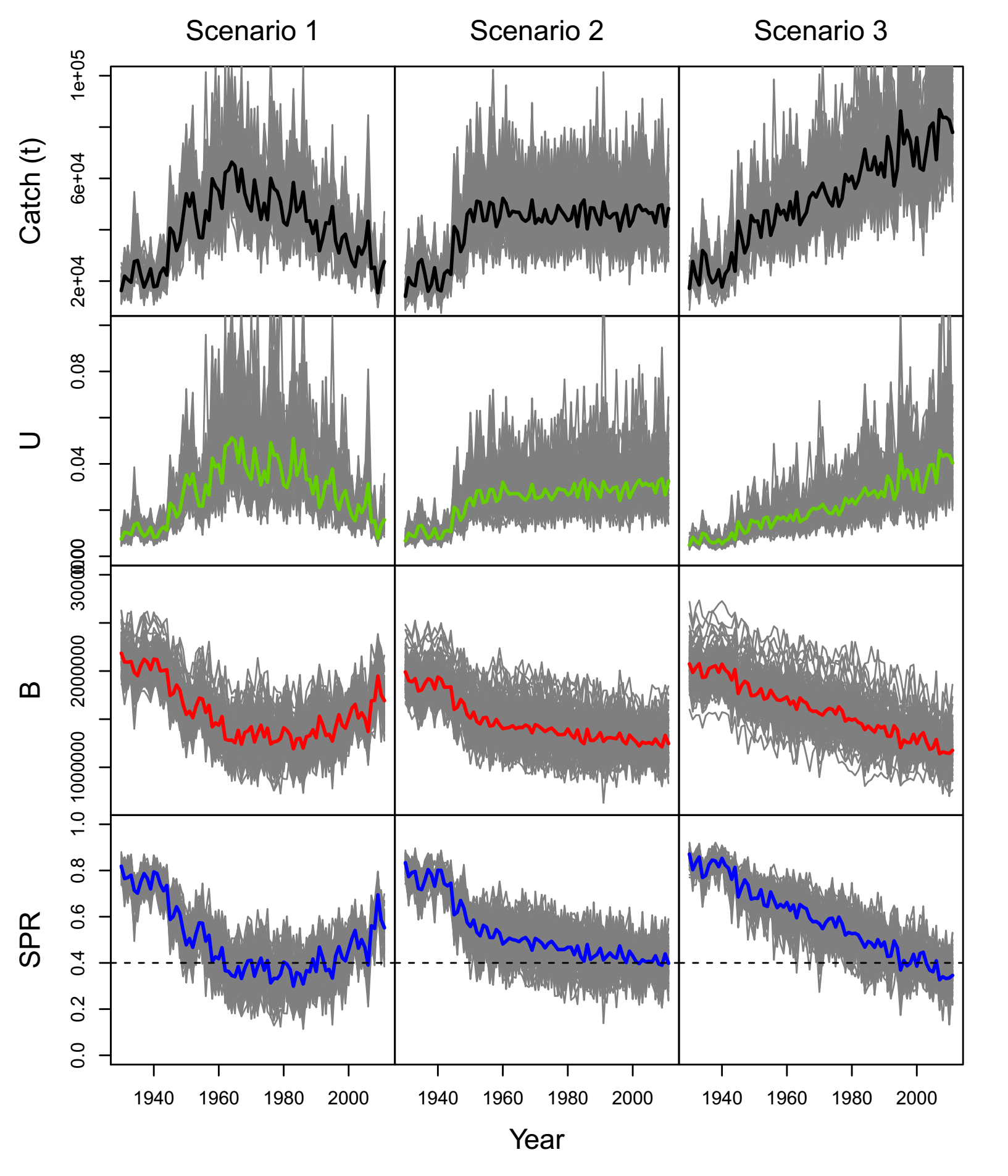
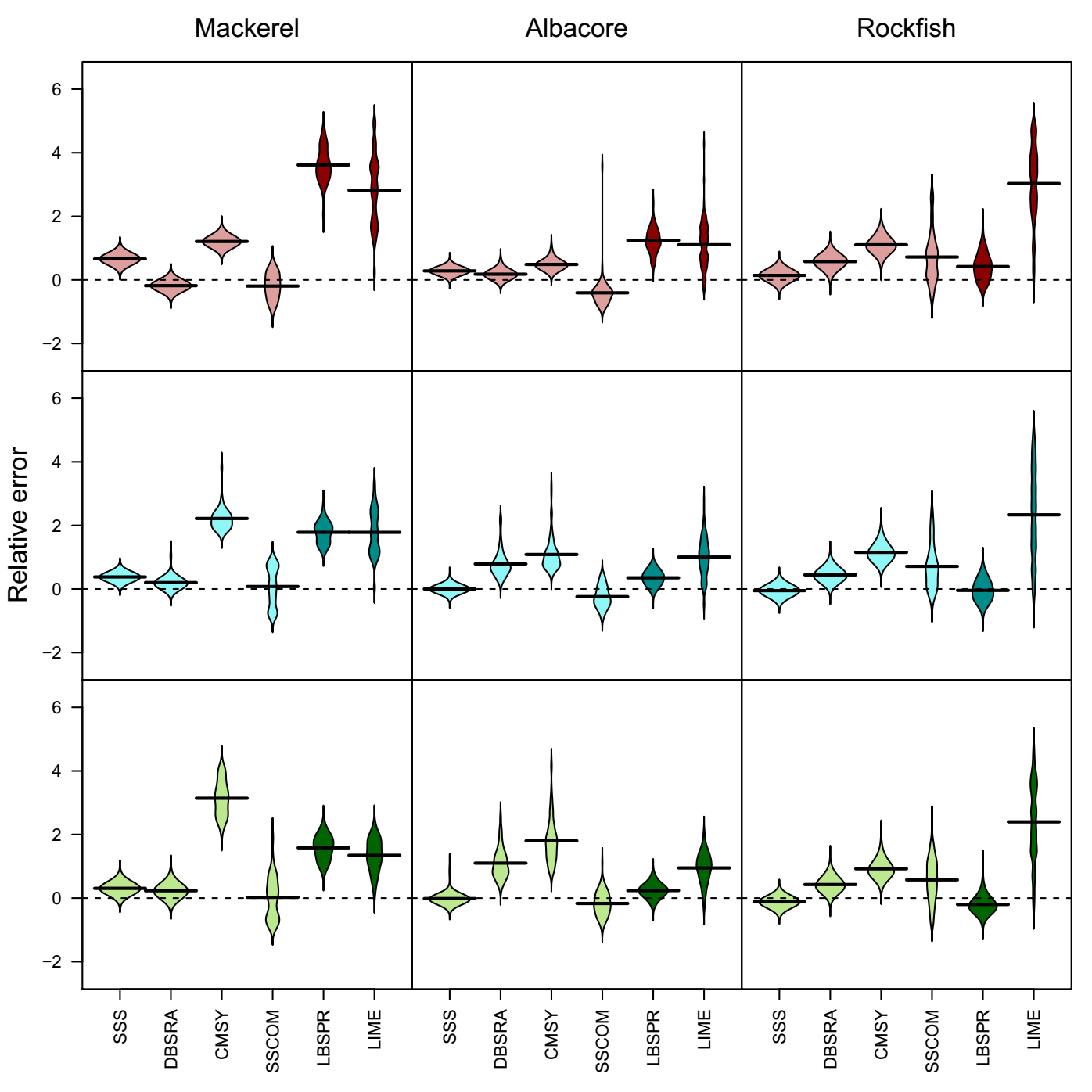


Figure 3.

Figure 4.