Supply of Research Services to establish MSY proxies for data-limited stocks (2017-18) to the Marine Institute, Rinville, Oranmore, Co. Galway. (Ref: ITT17-015)

Laurence Kell, Cóilín Minto

October 31, 2019

Contents

SUMMARY

Trends and fluctuations in populations are determined by complex interactions between extrinsic forcing and intrinsic dynamics. For example, stochastic recruitment can induce lowfrequency variability, i.e. cohort resonance, which can induce apparent trends in abundance and may be common in agestructured populations; such lowfrequency fluctuations can potentially mimic or cloak critical variation in abundance linked to environmental change, overexploitation or other types of anthropogenic forcing (Bjrnstad, 2004). Although important, these effects can be difficult to disentangle. The simulations so far show that life histories are important and should be used to help condition operating models to ensure robust feedback-control rules. MSE is important to help develop these robust feedback control rules and to help identify appropriate observational systems. Although the performance of the HCR depended on the life-history characteristic, it was not in the way initially expected, i.e. the outcomes could not be grouped solely by whether the Operating Models (OMs) represented fast growing vs. late maturing species or demersal vs. pelagic stocks. What was important was the nature of the dynamics, i.e. how variable was the stock between years; for example, a stock could exhibit high interannual variability if natural mortality and recruitment variability was high, regardless of the values of k, Linf, L50. The nature of the indices is also important; for example, even if a stock had low interannual variability, an index could be highly variable if it was based on juveniles or there were large changes in spatial distribution between years. It is therefore necessary to look at the robustness of HCRs to the nature of the time-series of the stock (as represented by the OM) and to the characteristics of the data collected from it (as represented by the Observation Error Model). This will require tuning by constructing a reference set of OMs and then tuning the HCR to secure the desired trade-offs. The work so far can be considered as focusing first on developing HCR that perform satisfactorily for a reference set, the next step is to develop case-specific HCRs.

8 *Aspects to consider for the 3.2.1 rule by the next meeting would be:*

- 8.1 *Investigating the impact of relative weighting of the r, f and b components of the rule on the performance of the rule;*
- 8.2 *Investigating more extensively the time-lag properties of the r component, including alternative formulations;*
- 8.3 *Setting of appropriate reference levels in the f and b component of the rules, and the extent to which this could be done with tuning that depends on life-history traits and/or the nature of the time-series;*
- 8.4 *Investigation of the use of trends in an index without a reference level.*

8 *Longer term aspect to consider for data-limited rules:*

- 9.1 *Focusing on the nature of time-series and developing diagnostics that could help determine the rules that would work well under alternative characterisations of the nature of the time-series, and aspects such as quality of data used by the rules (and hence ability to detect signals), ability to set appropriate reference points, etc.;*
- 9.2 *Linking life-history traits, the form of density-dependence and fishery characteristics (e.g. including fishery selectivity) to the nature of resulting time-series;*
- 9.3 *Develop guidance for use of catch rules by linking (a) and (b);*
- 9.4 *Avoiding the shot-gun approach to simulation testing e.g. by making more extensive use of sensitivity (elasticity) analysis to highlight factors that are most important in determining the time-series behaviour of stocks;*
- 9.5 *Investigating the implications of how the operating models are set up (fishing history, depletion levels, selectivity assumptions, mortality) on the behaviour of the stock and on the performance of the catch rule.*

1. The Call

The overall aim of the project is to develop and test a range of assessment models and methods to establish MSY reference points (or proxy MSY reference points) across the spectrum of data-limited stocks. There is a requirement for the following research services over a 24 months period between May 2017 and April 2019:

Task 1: ***Stock prioritisation*** *A number of example stocks have been identified (**Table 3.**). The final list of stocks will be prioritised using criteria like: economic value of the stock; importance of the species to the ecosystem (key-stone species); sensitivity to the impacts of fishing; available data.*

Task 2: ***Data collation*** *To run in parallel with other tasks. The project relies on existing data sets, however these data need to be collated in a usable form. Most datasets are available from the Marine Institute, or are publicly available, but others may only exist in other European labs/agencies.*

Task 3: ***Method and simulation framework development and implementation*** *A number of data-limited methods exist. In order to compare the performance of these methods it would be useful to implement them all in the same framework, e.g. R. New methods may also be developed in the same framework.*

Task 4: ***Method performance appraisal*** *Develop a set of diagnostics that can be applied across range of models. Also assess the stability of the model, sensitivity to assumptions and bias in the advised catch.*

Task 5: ***Reference point comparisons*** *Once reference points have been identified, their performance should be evaluated through simple management strategy evaluations.*

Task 6: ***Liaison with Marine Institute*** *The service provider is expected to meet on a regular basis with Marine Institute staff involved in the project: Monthly update meetings at the Marine Institute premises in Oranmore Galway*

Task 7: ***Linkage with other projects*** *The service provider is required to link research output to the following projects:*

- *The International Council of the Exploration of the Sea (ICES) is in the process of developing methods to identify MSY proxy reference points for data-limited stocks (WKLIFE and WKPROXY series of workshops). The service provider is required to contribute to this process by proposing and testing new assessment models and methods of establishing reference points and will be expected to attend up to 4 one-week meetings at ICES headquarters in Copenhagen. However there are key differences with the ICES approach:*
  This research contract will include stocks not currently assessed by ICES; this research contract will focus on the available data for each stock first and on the methods second; the ICES approach focuses on the methods first and then applies a limited number of methods to a large number of stocks.
- *Marine Institute research and development on data poor stocks which includes the biology, stock dynamics and Management Strategy Evaluation (MSE) for Pollock. It is expected that the service provider will collaborate closely with the team developing assessment methods for the pollock stock.*
- *Galway Mayo Institute of Technology GMIT had been awarded a Cullen fellowship for a PhD project on management strategy evaluation for monkfish. It is expected that the Cullen PhD and service provider will closely collaborate on tasks like data collation, assessment model implementation, simulation model development and management strategy evaluation.*

2. Project Plan

The principles of the European Union Common Fisheries Policy (CFP), which has driven the management of Europe's common fisheries resources since 1983, are to manage the activities of fishing fleets aims to ensure sustainable exploitation of the ocean's living resources, the provision of important food resources to humankind, and the profitability of an industry that is an important economic and social activity in many areas of Europe and elsewhere. The overall aim of the project is to support the CFP by developing and testing a range of assessment models and methods to establish MSY reference points (or proxy MSY reference points) across the spectrum of data-limited stocks.

Quantitative scientific advice is at the heart of fisheries management regulations, providing estimates of the likely current and future status of fish stocks through statistical population models, termed stock assessments, but also probabilistic comparisons of the expected effects of alternative management procedures. Management Strategy Evaluation (MSE) uses stochastic simulation to incorporate both the inherent variability of natural systems, and our limited ability to model their dynamics, into analyses of the expected effects of a given management intervention on the sustainability of both fish stocks and fleets.

Following the adoption of the precautionary approach [PA, Garcia, 1996] by many fisheries organisations, biological reference points have become central to management. Reference points are used as targets to maximise surplus production and limits to minimise the risk of depleting a resource to a level where productivity may be compromised. They must integrate biological processes such as growth, recruitment, mortality and connectivity into indices for productivity and spawning reproductive potential [Kell et al., 2015b] to provide limits and targets for exploitation. They are increasingly required for by-caught, threatened, endangered, and protected species where data and knowledge are limited, not just for the main commercial stocks, where analytical assessments are available [Sainsbury and Sumaila, 2003].

A main objective of reference points is to prevent overfishing, e.g. growth, recruitment, economic and target overfishing. Growth and recruitment overfishing are generally associated with limit reference points, while economic overfishing may be expressed in terms of either targets or limits. The difference between targets and limits is that indicators may fluctuate around targets but in general limits should not be crossed. Target overfishing occurs when a target is overshot, although variations around a target is not necessarily considered serious unless a consistent bias becomes apparent. In contrast even a single violation of a limit reference point may indicate the need for immediate action. Therefore to achieve MSY requires limit as well as target reference points.

2.1 Workplan

A variety of reference points and methods for deriving them are used for both data rich and poor stocks. In a data rich situtation reference points may be derived directly from a stock assessment model, e.g. in a biomass dynamic model where MSY is a function of the estimated parameters (r and K); ad-hoc approaches when using age methods such as Virtual Ppopulation Analysis, where assumptions about the stock recruitment relationship and future selection and biological parameters have to be made after fitting the assessment model; or in a state space formulation [Nielsen and Berg, 2014] which actually estimates a prediction mechanism and reference points.

In data poor situations a wide variety of statistical methods have been used or proposed to estimate stock status, productivity, fishing rates and reference points, for example using samples of length-composition [Kokkalis et al., 2015, Prince et al., 2015], age-composition [Thorson and Cope, 2015], fishery catch and fishing effort data [Roa-Ureta et al., 2015], abundance indices [Needle, 2015] or simple length-based reference points [Cope and Punt, 2009]. Before being able to make management recommendations, a link between a trigger reference point and stock status has to be identified, e.g. so a harvest control rule (HCR) can be used to link removals to the current state of the resource [Restrepo and Powers, 1999]. Cope and Punt [2009] proposed a way to do this for catch-based length indicators, using a decision tree and a risk assessment. This approach could be used for a range of indicators. However, [Cope and Punt, 2009] also noted that a full examination of such an approach requires a management strategy evaluation.

In an MSE setting reference points are tuned (i.e. chosen) to meet management objectives. Harvest strategies (i.e. HCRs) can either be model based or empirical Dowling et al. [2015], in the former a stock assessment is used to estimate stock status and reference points while in the later management is based on trends in the data directly. The Commission for the Conservation of Southern Bluefin Tuna (CCSBT) provides a model-free example of a MP [Hillary et al., 2015] that is based on year-to-year changes and trends in empirical indicators (i.e. CPUE and fisheries independent indices); reference levels are then tuned to meet management objectives using MSE, where tuning refers to adjusting the parameters of the MP to try and achieve the stated objectives represented by the OM. Model-based MPs, for example those based on a stock assessment model, may include the estimation of MSY-based reference points, but the values of F, F_{MSY}, B and B_{MSY} from the OM do not need to be equivalent to their proxies in the MP (e.g. if a stock assessment models used in the MP is structually different from that used to condition the OM).

WKLIFE was tasked with developing operational methods for setting proxy reference points for stocks where survey based assessments indicate trends (category 3) and for which reliable catch data are available (category 4). Methods so far examined include length-based indicators, spawning potential ratio (SPR), catch and cpue, and catch only based methods. These methods are now being implemented by the ICES study group WKPROXY, and along with others [Thorson et al., 2015, Carruthers et al., 2014], will be first evaluated using simulation, e.g. using crosstesting where a data rich stock assessment is used to generate data and then fitted to a data poor model and estimates of stock status compared. This approach requires a data rich assessment and that the stock assessment represents the stock dynamics. Alternatively life history theory and relationships can be used to simulate stocks and fisheries for a variety of hypotheses about their dynamics, i.e. using the FLife package [see Rosenberg et al., 2014]. To ensure a management advice framework based on a stock assessment and reference points (i.e. a control system) is robust requires showing that it still functions correctly in the presence of uncertainty or stressful environmental conditions [Radatz et al., 1990]. This requires an Operating Model (OM, i.e. a simulation model that represents hypotheses about resource dynamics) conditioned on ecological processes that affect the behaviour of management systems, i.e the focus is on the future, not on fitting historical data as when conditioning an OM on a stock assessment. This is a less data, and more hypothesis-orientated approach [Kell et al., 2006a].

To conduct MSE requires six steps [Punt and Donovan, 2007]; namely

Identification *of management objectives and mapping these to performance measures to quantify* *how well they are achieved*

Selection *of hypotheses about system dynamics for building **O**perating **M**odels (i.e. Simulation Models)*

Building *the simulation models, i.e. conditioning them on data and knowledge, and rejecting and weighting different hypotheses.*

Identifying *alternative management strategies, (i.e.the combination of pre-defined data, stock assessment methods, reference points and HCRs.*

Running *the simulations using the HCRs as feedback control procedures; and*

Agreeing *the Management Strategies that best meet management objectives.*

The MSE will be conducted using FLR[1] *[Kell et al., 2007] a family of packages in R for conducting MSE and stock, following the tasks below. Task 1 will help identify stocks and management objectives; Task 2 will allow the OMs to be conditioned; Task 3 will allow the OM to be implemented and psuedo data simulated to evaluate the proposed stock assessment methods and reference points; Task 4 will screen the candidate stock assessment methods (both those considered by WKLIFE but others to ensure that methods are state-of-the-art); and Task 5 will conduct MSE. Task 6 will ensure that the project delivers tools that make a major contribution to the management of data poor stocks; and Task 7 will help in dissemination, and ensure that methods developed tested across a wide range of case studies.*

[1] http://www.flr-project.org/

2.2 Tasks

2.21 Task 1: Stock prioritisation (Laurie)

A number of example stocks have been identified. The final list of stocks will be prioritised using criteria like: economic value of the stock; importance of the species to the ecosystem (key-stone species); sensitivity to the impacts of fishing; available data.

The final choice of stock will be made after the award of the contract based on their economic, social and ecological importance but also to reflect a contrasting range of life histories, fisheries and datasets. We will however focus on the following stocks:

- *Sprat in the Celtic Sea and West of Scotland Sprat (Sub-area VI & Divisions VIIa-c and f-k)*
- *Grey gurnard VI & VII (excl. VIId)*
- *Ling IIIa, IVa, VI, VII, VIII, IX, XII, and XIV*
- *Rays, primarily in areas VIIa,f,g*
- *John dory in ICES Sub-area VII and Divisions VIIIa,b and d (Northeast Atlantic)*
- *In collaboration with Newport STO:*
  - *Saithe VII, VIII, IX, X*
  - *Pollock VII*
- *Turbot VIIe,f,j,h and sub area VIII and IXa*
- *Brill VII (or suitably defined)*

A meta-database will be created identifying the data sources and relevant publications for the potential stocks of interest but also for related stocks and species. A reason for considering related stocks and species is because it is possible to compare the performance of data poor and data rich methods using cross-testing. In a cross-test data and population estimates from a data rich assessment are used to simulate data poor datasets which can then be used to test data poor assessment methods, allowing candidate methods to be identified. While a Robin Hood approach can be used to take information from data rich stock to inform assessments of information poor species, i.e. taking from the rich to help the poor. Since similarities in taxonomy, life-history or ecology allows the information from data-rich stocks to be utilised as prior distributions or penalty functions for the data poor species. Hierarchical Bayesian methods allow poor-data species to borrow strength from species with good-quality data. Jiao et al. [2011] for example used hierarchical Bayesian state-space surplus production models and showed estimates were considerably more robust than those of the nonhierarchical models.

Once the meta-database has been prepared (end of month 1) then a list of study stocks will be agreed, following which a database will be designed and an attempt to aquire the data made, following which the final list will be agreed (end of month 2).

2.22 Task 2: Data collation (Laurie)

Data collation (to run in parallel with other tasks) The project relies on existing data sets, however these data need to be collated in a usable form. Most datasets are available from the Marine Institute, or are publicly available, but others may only exist in other European labs/agencies.

Datasets will include, stock assesment datasets (ICES, NAFO, STECF[2], ...), fisheries, life history parameters (Fishbase[3], fishnets[4]), surveys (MI), commercial sampling sets (MI), and economic data (e.g. BIM).

Life history data will also be compiled since many studies have shown the relationships between life history traits for processes such as growth, maturity and natural mortality. Life history has also been

[2] https://stecf.jrc.ec.europa.eu/dd/medbs/ram
[3] http://www.fishbase.org/search.php
[4] https://github.com/fishnets/fishnets

used to develop priors in stock assesments for difficult to estimate parameters. e.g. the population growth rate (r) in data poor assessments. Life history parameters can also be used to develop an Operating Model.

The meta-database will be extended with code to read the data into a common format, e.g. data frames and other objects in R and FLR to model stocks, populations and fisheries. It is important to ensure that data are easily available and that any processing steps are well documented and standardised this is true for both the basic data and the results from the MSE (e.g. datasharing[5]). s.

2.23 Task 3: Framework Development (Laurie)

Method and simulation framework development and implementation A number of data-limited methods exist. In order to compare the performance of these methods it would be useful to implement them all in the same framework, e.g. R. New methods may also be developed in the same framework.

*All methods will be implemented in R and be compatible with The Fishery Library in R (**FLR**) [Kell et al., 2007], a project that has for the last ten years been building an extensible toolset of statistical and simulation methods for quantitative fisheries science, with the overarching objective of enabling fisheries scientists to carry out analyses of management procedures in a simplified and robust manner through the MSE approach.*

***FLR** has become widely used in many of scientific bodies providing fisheries management advice, both in Europe and elsewhere. The evaluation of the effects of elements of the revised CFP, the analysis of the proposed fisheries management plans for the North Sea, or the comparison of management strategies for Atlantic tuna stocks, among others, have used the **FLR** tools to advice managers of the possible courses of action to favour the sustainable use of many marine fish stocks.*

*The **FLR** toolset is currently composed of a variety of packages, covering the various steps in the fisheries advice and simulation workflow. They include a large number of S4 classes, and more recently Reference Classes, to model the data structures that represent each of the elements in the fisheries system. Class inheritance and method overloading are essential tools that have allowed the **FLR** packages to interact, complement and enrich each other, while still limiting the number of functions an user needs to be aware of. Methods also exist that make use of R's parallelization facilities and of compiled code to deal with complex computations.*

*Using **FLR** means that advantage can be taken of existing methods, and that dissemination and support is easier and will be maintained after the life of the project. Under the project development will be on mainly on four packages **FLife**, **mpb**, **hcr** and **oem**. In addition new methods will be added to other **FLR** packages as appropriate.*

***FLife** is a package for modelling life history relationships, e.g. for developing priors, estimating quantities such a Z from length data, deriving reference points and indicators and for creating OMs.*

***mpb** is a package for modelling biomass dynamic based management procedures, it has also been used to assess a variety of stocks, e.g. Atlantic bigeye and North Atlantic albacore and swordfish. It also forms the basis of the North Atlantic albacore MP evaulated using MSE.*

***hcr** will be a new package that will include a variety of emprirical HCRs, e.g. those used by CCSBT, proposed by ICES and from a review of the literature [e.g. Pomarede et al., 2010].*

***oem** will be a new package that implements a variety of observation error models, that will be used to simulate a variety of datasets from OMs that will be used by the data poor methods.*

*In addition Dr Kell is the developer of two packages that are used to provide a common set of diagnostics across stock assessment methods (**diags**) and to summarise stock status relative to reference points and for summarising the results from MSE (**kobe**)*

There are a variety of R packages, that already implement a variety of data poor methods, e.g. https://github.com/quang-huynh/MLZ, https://github.com/AdrianHordyk/LBSPR https://github.com/cran/DLMtool, wherevever possible these packages will be used and we will collaborate with their developers.

[5] https://github.com/AdrianHordyk/datasharing

2.24 Task 4: Method Performance Appraisal (Laurie)

Method performance appraisal Develop set of diagnostics that can be applied across range of models. Also assess the stability of the model, sensitivity to assumptions and bias in the advised catch.

A range of summary statistics will be required to illustrate trade-offs between multiple potentially conflicting objectives. Although there are many potential summary statistics so that decision makers can choose between tangible options on the basis of actual projections rather than abstract concepts and performance statistics, however, should ideally be few, informative and based axes such as stock status, 'safety', 'stability' and 'yield'. It is also necessary to distinguish between techincal summary statistics (i.e. those required to evaluate model fits and performance) and those required to evaluate management objectives.

Results will be presented using an interactive app, based on shiny, that can be used to evaluate the performance of the models.

2.25 Task 5: Reference Point Comparisons (Laurie, Cóilín)

Reference point comparisons (across candidate methods) Once reference points have been identified, their performance should be evaluated through simple management strategy evaluations. *A set of appropriate stock assessment models will be fit to the available data (Task 1 and 2). Methods used will reflect the available data on a stock-by-stock basis (Task 2) and also methods available or developed in Tasks 3 and 4. Performance diagnostics (e.g., residual inspection, retrospective patterns) will be run for each assessment model fit.*
We will provide a review of reference points and indicators, based on a variety of assumptions, e.g. MSY based on biomass dynamic stock assessment models, and indicators such as L_{opt},L_{50} *and* L_{mega} *for length-based methods. Then compare these using simulation, e.g. cross-testing and simulations based on FLife. This will allow the power of the methods to detect whether a stock has achieved its targets and avoid its limits. Based on this screening process a set of candidate reference points and assessment methods will be proposed for MSE.*

MSE will include a Value-of-infomation analysis, where the benefits of collecting better data and new infomation will be evaluated.

2.26 Task 6: Liaison with Marine Institute (Cóilín, Laurie, Laurie)

The service provider is expected to meet on a regular basis with Marine Institute staff involved in the project: Monthly update meetings at the Marine Institute premises in Oranmore Galway

The proposal is ambitious but achievable. However there needs to be good communication between the consortum and the Marine Instiutute to ensure the project keeps focused and delivers. Therefore we will arrange monthly face-to-face meetings, make all code, data and results available on the cloud and in a suitable repository (e.g. github) and provide a web based interface for model results.
Wider project 6-monthly progress reports and meetings at the Marine Institute will ensure the over-arching goals of the project are achieved.

2.27 Task 7: Linkage with other Projects (Cóilín)

- *Monkfish*
  This project will develop in close collaboration with the Cullen Fellowship of Mr Luke Batts, co-supervised by Dr Hans Gerritsen (MI) and Dr Cóilín Minto (GMIT). Active collaboration will occur with Tasks 3–5, as these are similarly proposed in the Cullen Fellowship where they are applied specifically to Lophius budegassa and Lophius piscatorius stocks in ICES areas VII-VIII.
- *Pollock*
  Active collaboration exists between GMIT and the Newport Research Cluster (e.g., Unlocking the Archive *project). Further collaboration and linkages will be built around data-poor assessment of pollock (liasing with the dedicated Scientific and Technical Officer working on pollock at the*

Furnace research facility). Both visits to Newport and group attendance at the monthly meetings will facilitate collaboration and crossover.

- *DRuMFISH project*
  The project will also link with the DGMARE project: "Study on approaches to management for data-poor stocks in mixed fisheries (DRuMFISH)" to which GMIT is a partner in the consortium. Methodological development from DRuMFISH (e.g., hierarchical methods) will be directly relevant to the present proposal.
- *CPV codes 71354500-9 Marine survey services 73112000-0 Marine research services 90712300-4 Marine conservation strategy planning 98360000-4 Marine services 77700000-7 Services incidental to fishing 73000000-2 Research and development services and related consultancy services 73110000-6 Research services 73200000-4 Research and development consultancy services 73210000-7 Research consultancy services.*
- *There are also other projects worldwide that can be link to e.g. the global group on stock assessment methods and the tRFMO MSE WG.*

2.3 Project Management and Milestones

To achieve the goals of the project requires good communication between the consortum and the Marine Instiutute. Therefore we will arrange monthly face-to-face meetings and 6 monthly meetings throughout the lifetime of the project. In addition all code will be available via a repository such as github, databases will be accesible via the cloud and we provide a web based interface (such as a shiny app) for model results.

The 6 monthly meetings, i.e. kick-off, 6^{th} *month,* 12^{th} *month and the final meeting will provide the milestones. Outputs will include working papers for the two ICES groups WGCSE, and WKPROXY. These will document the methods and also present case study applications which will be agreed at the 6 monthly meetings and developed on the cloud. To help with this a wiki will be used.*

2.31 Task Summary

A full break down of tasks is given below. Each task is assigned to a lead who will take responsibility for delivery, however, the task overlap and members of the consortium will be involved across the range of tasks as required, see figure **??**.

Task 1 *Stocks (Laurie)*

- *Identify managment objectives and current reference points.*
- *Create a meta-database, that will identify data sources and allow case studies to be choosen.*
- *The meta-database should include related species and stocks, as well as those listed below, to allow OMs to be conditioned, cross-testing to be conducted and priors to be developed.*

Task 2 *Data (Laurie)*

- *Design a DB to hold the data refered to in the meta-database*
- *Develop R/SQL scripts to access and read the data*
- *Develop tools to summarise, check the data and conduct meta-analyses.*
- *Summarise the error structure of the data series to allow a variety of OEM to be implemented*
- *Develop tools to help in conditioning OMs*

Task 3 *Framework (Laurie)*

- *Review appropriate assessment methods, used by ICES and Regional Fisheries Management Organisations*
- *Identify code and R packages*
- *Implement methods as required in R using S4 classes and methods and packages such as TMB, Stan, Bugs.*
- *Develop OMs based on life history and/or data rich stocks*
- *Develop OEMs for the range of data (e.g. CPUE, catch, length, size composition) used in the data poor assessments.*

Task 4 *Methods (Laurie)*

- *Compare methods to determine data and infomation needs, e.g. using simulation without feedback control to conduct power analyses where the probabilities of achiveing targets and avoiding limits is evaluated.*

Task 5 *Reference Points (Laurie)*

- *Fit and compare appropriate stock assessments that reflect the available data (Task 2) and methods developed (Tasks 3 and 4)*
- *Run MSE using the OM, OEM and MPs based on Task 4.*
- *Conduct a Value-of-Infomation analysis, i.e. what are the benefits of reducing CVs or collecting alternative data series*

Task 6 *Liason (Cóilín)*

- *Monthly meetings to monitor progress and get feedback .*
- *Six monthly meetings to review progress and agree next stage of work.*
- *Attendance at ICES WGs to provide advice and obtain feedback.*

Task 7 *Other Projects (Cóilín)*

Active collaboration will be developed and maintained with the following ongoing projects

- *Pollock assessment and MSE at Newport*
- *Cullen fellowship on Monkfish MSE*
- *DRuMFISH project on data-poor stock assessment in mixed fisheries*

2.4 Outputs

The expected outputs are a a collection of existing and new assessment models for data-limited stocks, all implemented in the same framework (e.g. R) with a set of diagnostic tools that can be applied to all models.

*Methods will be implemented as R packages and/or methods; this will include full documentation, online help vignettes and tests. As well as conducting cross-testing to validate models a full set of diagnostic tools will be provided, inlcuding checks for convergence using methods such as likelihood profiling; identification of violation of assumptions by checking residuals to fits using the **diags** package; and use methods such as the jack knife or bootstrap to identify problems with the data and model specifications; and conduct hindcasting to evaluate prediction ability [Kell et al., 2016a].*

A set of proposed reference points for a range of stocks with associated management strategy evaluations to contribute to sustainable management of these stocks.

We will review the reference points currently used in the management of stocks under the CFP and estimated by ICES, in addition we will review reference points used elsewhere by other management bodies and RFMOs. Then compare datapoor and traditional target and limit reference points using crosstesting, in consultation with the Marine Institute and ICES WGs we will come up with a set of candidate reference points, classified by data reqirements, and evaluate these using MSE.

Working documents describing the methods and findings to relevant ICES groups (e.g WGCSE; WKPROXY).

Dr Kell will attend the Methods related WGs to develop the novel approaches, while Dr Alexander Tidd will attend stock assessment WGs to present applications. All papers will be coauthored with Galway-Mayo Institute of Technology and Marine Institute staff.

Publication(s) in peer-reviewed journals on new methods/tools/evaluations.

There are a variety of manuscripts that the project will produce, i.e. on new methods, applications, and a potential review of the relative value-of-information to the value-of-Control.

2.5 Deliverables

Data *Case study data*

- **Meta-database** *identifying data sources for candidate stocks, this will be updated through the life of the project and be made available via the web as a living document.*
- **Stock database** *to hold the data required to condition the OMs and to parameterise the OEM, with tools for analysis and summary.*
- **Summary database** *For performance statistics from the MSE.*

R Packages *These will form part of the **FLR** family of packages*

- **FLCore** *will be updated to include data poor methods as appropriate; in particular the following packages*
  - **kobe** *Package with methods for summary statistics*
  - **diags** *Package for stock assessment diagnostics*
  - **FLife** *Package for simulation basd on life histories, e.g. for conditioning OMs*
  - **mpb** *Package for management procedures based on biomass based methods*
  - **hcr** *Package for emprirical MPs*
  - **oem** *Observation error models.*

Dissemination ICES WGs *Attendance at WGCSE and WKPROXY*

- **Web based tools** *Shiny app to summaries results*

Manuscripts *A folio of papers will be agreed at the 1st 6 montly meeting, this will include at least 1 paper each on*

- **Paper 1** *Methods*

Paper 2 *Applications and*

Paper 3 *Review, e.g. of the relative value-of-infomation and control.*

3. Tables

tablePreliminary list of stocks

Species	*TAC*	*Commercial Catch*	*Data*	*Comments*
Sprat	*No*	*Targeted species for small fleet*	*Poor; mainly landings weights*	*Key-stone prey fish*
Gurnards	*No*	*Nearly 100% discarded*	*Reasonable discard and survey data. No age data*	*Key-stone prey widely distributed and abundant*
Saithe Pollock Ling	*Yes*	*Mixed fishery*	*Some port sampling, observer and survey data. Very limited age data*	*Key-stone predator*
Rays Skates	*Yes*	*Targeted and mixed fishery*	*Some port sampling, observer and survey data. No age data*	*Sensitive species slow reproduction*
John Dory	*No*	*Mixed fishery but can be targeted to an extent*	*Some port sampling, observer and survey data. No age data*	*Sensitive species valuable non-TAC species (not protected by fisheries management)*
Turbot Brill	*No*	*Mixed fishery but can be targeted to an extent*	*Some port sampling, observer and survey data. Very limited age data*	*Sensitive species valuable non-TAC species (not protected by fisheries management)*

References

H. Arrizabalaga, P. De Bruyn, G. Diaz, H. Murua, P. Chavance, A. de Molina, D. Gaertner, J. Ariz, J. Ruiz, and L. Kell. Productivity and susceptibility analysis for species caught in Atlantic tuna fisheries. Aquatic Living Resources*, 24(01):1–12, 2011.*

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

T. R. Carruthers, L. T. Kell, D. D. Butterworth, M. N. Maunder, H. F. Geromont, C. Walters, M. K. McAllister, R. Hillary, P. Levontin, T. Kitakado, et al. Performance review of simple management procedures. ICES J. Mar. Sci.*, 73(2):464–482, 2016.*

T. Catchpole, C. Frid, and T. Gray. Discards in north sea fisheries: causes, consequences and solutions. Marine Policy*, 29(5):421–430, 2005.*

J. M. Cope and A. E. Punt. Length-based reference points for data-limited situations: applications and restrictions. Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science*, 1(1):169–186, 2009.*

J. Deroba, D. Butterworth, R. Methot, J. De Oliveira, C. Fernandez, A. Nielsen, S. Cadrin, M. Dickey-Collas, C. Legault, J. Ianelli, L. Kell, et al. Simulation testing the robustness of stock assessment models to error: some results from the ices strategic initiative on stock assessment methods. ICES J. Mar. Sci.*, 72(1):19–30, 2015.*

N. Dowling, C. Dichmont, M. Haddon, D. Smith, A. Smith, and K. Sainsbury. Empirical harvest strategies for data-poor fisheries: A review of the literature. Fisheries Research*, 171:141–153, 2015.*

C. M. Fortuna, L. Kell, D. Holcer, S. Canese, E. Filidei Jr, P. Mackelworth, and G. Donovan. Summer distribution and abundance of the giant devil ray (mobula mobular) in the adriatic sea: Baseline data for an iterative management framework. Scientia Marina*, 78(2):227–237, 2014.*

F. L. Frédou, T. Frédou, D. Gaertner, L. Kell, M. Potier, P. Bach, P. Travassos, F. Hazin, and F. Ménard. Life history traits and fishery patterns of teleosts caught by the tuna longline fishery in the south atlantic and indian oceans. Fisheries Research*, 179:308–321, 2016.*

J.-M. Fromentin, S. Bonhommeau, H. Arrizabalaga, and L. L. Kell. The spectre of uncertainty in management of exploited fish stocks: the illustrative case of Atlantic bluefin tuna. Marine Policy*, 47:8–14, 2014.*

S. Garcia. The precautionary approach to fisheries and its implications for fishery research, technology and management: an updated review. FAO Fisheries Technical Paper*, pages 1–76, 1996.*

H. Glenn, D. Tingley, S. S. Maroño, D. Holm, L. Kell, G. Padda, I. R. Edvardsson, J. Asmundsson, A. Conides, K. Kapiris, et al. Trust in the fisheries scientific community. Marine Policy*, 36(1): 54–72, 2012.*

R. M. Hillary, A. L. Preece, C. R. Davies, H. Kurota, O. Sakai, T. Itoh, A. M. Parma, D. S. Butterworth, J. Ianelli, and T. A. Branch. A scientific alternative to moratoria for rebuilding depleted international tuna stocks. Fish and Fisheries*, 2015.*

Y. Jiao, E. Cortés, K. Andrews, and F. Guo. Poor-data and data-poor species stock assessment using a bayesian hierarchical approach. Ecological Applications*, 21(7):2691–2708, 2011.*

L. Kell, M. Pastoors, R. Scott, M. Smith, F. Van Beek, C. O'Brien, and G. Pilling. Evaluation of multiple management objectives for northeast Atlantic flatfish stocks: sustainability vs. stability of yield. ICES J. Mar. Sci.*, 62(6):1104–1117, 2005a.*

L. Kell, G. Pilling, G. Kirkwood, M. Pastoors, B. Mesnil, K. Korsbrekke, P. Abaunza, R. Aps, A. Biseau, P. Kunzlik, et al. An evaluation of the implicit management procedure used for some ICES roundfish stocks. ICES J. Mar. Sci.*, 62(4):750–759, 2005b.*

L. Kell, J. A. De Oliveira, A. E. Punt, M. K. McAllister, and S. Kuikka. Operational management procedures: an introduction to the use of evaluation frameworks. Developments in Aquaculture and Fisheries Science*, 36:379–407, 2006a.*

L. Kell, G. Pilling, G. Kirkwood, M. Pastoors, B. Mesnil, K. Korsbrekke, P. Abaunza, R. Aps, A. Biseau, P. Kunzlik, et al. An evaluation of multi-annual management strategies for ices roundfish stocks. ICES J. Mar. Sci.*, 63(1):12–24, 2006b.*

L. Kell, I. Mosqueira, P. Grosjean, J. Fromentin, D. Garcia, R. Hillary, E. Jardim, S. Mardle, M. Pastoors, J. Poos, et al. FLR: an open-source framework for the evaluation and development of management strategies. ICES J. Mar. Sci.*, 64(4):640, 2007.*

L. T. Kell, P. Levontin, C. R. Davies, S. Harley, D. S. Kolody, M. N. Maunder, I. Mosqueira, G. M. Pilling, and R. Sharma. The quantification and presentation of risk. Management Science in Fisheries: An Introduction to Simulation-Based Methods*, page 348, 2015a.*

L. T. Kell, R. D. Nash, M. Dickey-Collas, I. Mosqueira, and C. Szuwalski. Is spawning stock biomass a robust proxy for reproductive potential? Fish and Fisheries*, 2015b.*

L. T. Kell, A. Kimoto, and T. Kitakado. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries Research*, 183:119–127, 2016a.*

L. T. Kell, P. Levontin, C. R. Davies, S. Harley, D. S. Kolody, M. N. Maunder, I. Mosqueira, G. M. Pilling, and R. Sharma. The quantification and presentation of risk. Management Science in Fisheries: An Introduction to Simulation-based Methods*, page 348, 2016b.*

A. Kokkalis, U. H. Thygesen, A. Nielsen, and K. H. Andersen. Limits to the reliability of size-based fishing status estimation for data-poor stocks. Fisheries Research*, 171:4–11, 2015.*

A. Leach, P. Levontin, J. Holt, L. Kell, and J. Mumford. Identification and prioritization of uncertainties for management of eastern Atlantic bluefin tuna (Thunnus thynnus). Marine Policy*, 48: 84???92, 2014.*

C. L. Needle. Using self-testing to validate the surbar survey-based assessment model. Fisheries Research*, 171:78–86, 2015.*

A. Nielsen and C. W. Berg. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research*, 158:96–101, 2014.*

G. M. Pilling, L. T. Kell, T. Hutton, P. J. Bromley, A. N. Tidd, and L. J. Bolle. Can economic and biological management objectives be achieved by the use of msy-based reference points? a north sea plaice (pleuronectes platessa) and sole (solea solea) case study. ICES Journal of Marine Science: Journal du Conseil*, 65(6):1069–1080, 2008.*

M. Pomarede, R. Hillary, L. Ibaibarriaga, J. Bogaards, and P. Apostolaki. Evaluating the performance of survey-based operational management procedures. Aquatic Living Resources*, 23(1): 77–94, 2010.*

M. Pons, T. A. Branch, M. C. Melnychuk, O. P. Jensen, J. Brodziak, J. M. Fromentin, S. J. Harley, A. C. Haynie, L. T. Kell, M. N. Maunder, et al. Effects of biological, economic and management factors on tuna and billfish stock status. Fish and Fisheries*, 18(1):1–21, 2017.*

J. Prince, S. Victor, V. Kloulchad, and A. Hordyk. Length based spr assessment of eleven indo-pacific coral reef fish populations in palau. Fisheries Research*, 171:42–58, 2015.*

A. Punt and G. Donovan. Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES J. Mar. Sci.*, 64(4):603–612, 2007.*

J. Radatz, A. Geraci, and F. Katki. Ieee standard glossary of software engineering terminology. IEEE Std*, 610121990:121990, 1990.*

V. Restrepo and J. Powers. Precautionary control rules in us fisheries management: specification and performance. ICES Journal of Marine Science: Journal du Conseil*, 56(6):846–852, 1999.*

R. H. Roa-Ureta, C. Molinet, N. Barahona, and P. Araya. Hierarchical statistical framework to combine generalized depletion models and biomass dynamic models in the stock assessment of the chilean sea urchin (loxechinus albus) fishery. Fisheries Research*, 171:59–67, 2015.*

A. A. Rosenberg, M. Fogarty, A. Cooper, M. Dickey-Collas, E. Fulton, N. Gutiérrez, K. Hyde, K. Kleisner, T. Kristiansen, C. Longo, et al. Developing new approaches to global stock status assessment and fishery production potential of the seas*. Food and Agriculture Organization of the United Nations, 2014.*

K. Sainsbury and U. R. Sumaila. 20 incorporating ecosystem objectives into management of sustainable marine fisheries, Including'Best Practice'Reference points and use of marine protected areas. Responsible fisheries in the marine ecosystem*, page 343, 2003.*

J. T. Thorson and J. M. Cope. Catch curve stock-reduction analysis: An alternative solution to the catch equations. Fisheries Research*, 171:33–41, 2015.*

J. T. Thorson, L. T. Kell, J. A. De Oliveira, D. B. Sampson, and A. E. Punt. Introduction. Fisheries Research*, 171:1–3, 2015.*

A. Tidd, S. Brouwer, and G. Pilling. Shooting fish in a barrel? assessing fisher-driven changes in catchability within tropical tuna purse seine fleets. Fish and Fisheries*, pages n/a–n/a, 2017. ISSN 1467-2979. doi: 10.1111/faf.12207. URL http://dx.doi.org/10.1111/faf.12207.*

A. N. Tidd, T. Hutton, L. Kell, and G. Padda. Exit and entry of fishing vessels: an evaluation of factors affecting investment decisions in the north sea english beam trawl fleet. ICES J. Mar. Sci.*, 68(5):961–971, 2011.*

A. N. Tidd, T. Hutton, L. T. Kell, and J. L. Blanchard. Dynamic prediction of effort reallocation in mixed fisheries. Fish. Res.*, 125:243–253, 2012.*

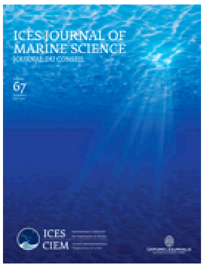
A. N. Tidd, Y. Vermard, P. Marchal, J. Pinnegar, J. L. Blanchard, and E. Milner-Gulland. Fishing for space: fine-scale multi-sector maritime activities influence fisher location choice. PloS one*, 10 (1):e0116335, 2015.*

A. N. Tidd, C. Reid, G. M. Pilling, and S. J. Harley. Estimating productivity, technical and efficiency changes in the western pacific purse-seine fleets. ICES Journal of Marine Science: Journal du Conseil*, 73(4):1226–1234, 2016.*

4. Appendices

5. Appendix 1: Peer Review Papers

5.01 Linking the performance of a data-limited empirical catch rule to life-history traits



Linking the performance of a data-limited empirical catch rule to life-history traits

Journal:	*ICES Journal of Marine Science*
Manuscript ID	ICESJMS-2019-350.R1
Manuscript Types:	Original Article
Date Submitted by the Author:	07-Oct-2019
Complete List of Authors:	Fischer, Simon; Centre for Environment Fisheries and Aquaculture Science; Imperial College London, Centre for Environmental Policy De Oliveira, José; Centre for Environment Fisheries and Aquaculture Science Kell, Laurence; Imperial College London, Centre for Environmental Policy
Keyword:	Management Strategy Evaluation, data-limited, life-history, empirical catch rules, MSY, precautionary, FLR, FLife

Linking the performance of a data-limited empirical catch rule to life-history traits

Simon H. Fischer[1,2], José A. A. De Oliveira[1], Laurence T. Kell[2]

[1] Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK

[2] Centre for Environmental Policy, Imperial College London, Weeks Building, 16-18 Princes Gardens, London SW7 1N, UK

Keywords: Management Strategy Evaluation, data-limited, life-history, empirical catch rules, MSY, precautionary, FLR, FLife.

Journal: ICES Journal of Marine Science

Contact author: Simon H. Fischer, Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK, Email: simon.fischer@cefas.co.uk

ABSTRACT

Worldwide, the majority of fish stocks are data-limited and lack fully quantitative stock assessments. Within ICES, such data-limited stocks are currently managed by setting total allowable catches without the use of target reference points. To ensure that such advice is precautionary, we used Management Strategy Evaluation to evaluate an empirical rule that bases catch advice on recent catches, information from a biomass survey index, catch length frequencies and MSY reference point proxies. Twenty-nine fish stocks were simulated covering a wide range of life-histories. The performance of the rule varied substantially between stocks, and the risk of breaching limit reference points was inversely correlated to the von Bertalanffy growth parameter *k*. Stocks with $k>0.32$ had a high probability of stock collapse. A time-series cluster analysis revealed four types of dynamics, i.e. groups with similar stock trajectories (collapse, B_{MSY}, $2B_{MSY}$, $3B_{MSY}$). It was shown that a single generic catch rule cannot be applied across all life-histories, and management should instead be linked to life-history traits, and in particular, the nature of the time series. The lessons learnt can help future work to shape scientific research into data-limited fisheries management and to ensure fisheries are MSY-compliant and precautionary.

INTRODUCTION

When managing fisheries, decisions must be made with incomplete knowledge, which is why international agreements request the adoption and implementation of the Precautionary Approach (Garcia, 1995). In addition, fish retailers and consumers are increasingly looking for assurances that the food they buy is sustainably produced. Therefore, many regional fisheries management Organisations (RFMOs) have implemented management frameworks based on target and limit reference points to prevent overfishing and ensure targets are achieved. Despite this, most fisheries and commercially exploited stocks still lack reliable estimates of stock status and effective management due to poor data, limited knowledge and insufficient resources (Jardim *et al.*, 2015; Fitzgerald *et al.*, 2018).

Since 2012, ICES has applied a framework to provide catch advice for the European data-limited stocks (ICES, 2012a, 2013a). The increasingly sophisticated methods developed for stock assessment are not always suited to data-poor fisheries (Bentley, 2015). Therefore, recently, many data-limited approaches have emerged and re-emerged to meet the increasing demand for science-based fisheries management for data limited stocks (Wetzel and Punt, 2011; Costello *et al.*, 2012; Dowling *et al.*, 2015, 2016; Chrysafi and Kuparinen, 2016; Rosenberg *et al.*, 2018). However, in a review of data-limited methods, Dowling *et al.* (2019) noted the dangers in the indiscriminate use of generic methods and recommended obtaining better data, using care in acknowledging and interpreting uncertainties, developing harvest strategies that are robust to the higher levels of uncertainty and tailoring them to the specific species' and fisheries' data and context.

One way to do this is to evaluate candidate data-limited management frameworks using Management Strategy Evaluation (MSE, Smith, 1994; Punt *et al.*, 2016). MSE uses an Operating Model (OM) to represent a fish stock and the fisheries operating on it. The OM is used to simulate resource dynamics in simulation trials in order to evaluate the performance of a Management Procedure (MP),~~.~~ ~~W~~where the MP is the combination of pre-defined data, together with an algorithm to which such data are input to set a management measure, such as a total allowable catches (TAC). This in turn is converted into a catch that is removed from the operating model in a feedback loop (Punt *et al*., 2016).

The application of MSEs has been mainly focused on data-rich situations, where enough data are available to condition the OM using stock assessment models. An MP may be either model based, where a stock assessment is used to estimate stock status and set management measures (e.g. Kell *et al.*, 2005) or model free, where a trend in an empirical indicator is used to set the catch (Hillary *et al.*, 2016). ~~The~~ MSEs for data-limited purposes are somewhat rarer, although there are notable studies. For example, Carruthers *et al.* (2012) evaluated methods based on catch data alone, and found that catch-based methods were, on average, more negatively biased than stock assessment methods that explicitly model population dynamics and use additional fishing effort data. In a subsequent study, Carruthers *et al.* (2014) found that methods that~~e~~ rely only on historical catches performed worse than maintaining current fishing levels, and that only methods that dynamically accounted for changes in abundance and/or depletion performed well at low stock sizes. Geromont and Butterworth (2015a) tested a range of simple catch rules based on historical catches, length data or survey index data and found that such simple rules perform well and could be used in practice.

Within ICES, simple catch rules have been developed for data limited stocks (ICES, 2012a). For example, the "2 over 3" rule aims to keep stocks at their current level by multiplying recent catches by the trend in a biomass index:

$$C_{y+1} = C_{y-1} \frac{\sum_{i=y-2}^{y-1} I_i/2}{\sum_{i=y-5}^{y-3} I_i/3} \quad (1)$$

where C_{y+1} is the newly advised catch for year $y+1$, C_{y-1} is the last advised catch and I is a biomass index. This rule, in combination with an uncertainty cap (limiting change to no more than 20%) and precautionary buffer (which reduces the catch by 20% if the stock is judged to be outside safe biological levels), is currently applied to give catch advice within ICES for many data limited stocks (ICES, 2018a).

The ICES "2 over 3" rule lacks a management target, can induce oscillatory behaviour resulting in increased biological risk over time, and includes a time lag in the translation of changes in the biological stock into advice (ICES, 2013b, 2017a, 2017b). An alternative catch rule, making use of more data sources, has therefore been proposed (ICES, 2012b):

$$C_{y+1} = C_{y-1}\, r\, f\, b \quad (2)$$

where the advised catch C_{y+1} is based on the previous catch C_{y-1}, multiplied by three components r, f and b, each representing a stock characteristic. Component r corresponds to the trend in a biomass index (I), component f is a proxy for the ratio F_{MSY} divided by the current exploitation based on length data from the catch, and component b is a biomass safeguard which protects the stock once the biomass index drops below a threshold. Initially, this catch rule was merely a concept without specification about what data should be used and how the components could be derived from them (ICES, 2012b). Recently, the rule has been revisited by ICES (2017c) and suggestions made for simulation testing and application to actual stocks. Several options for the three components have been proposed, and initial simulation testing narrowed it down to only one option per component (ICES, 2017b). This catch rule is the focus

100 of the present study, where we aim to (I) establish procedures to simulate data limited fish stocks based on life-history parameters, (II) simulation-test the aforementioned catch rule, (III) associate the performance of the catch rule to life-history parameters, and (IV) provide guidance on the application of the catch rule and thereby advancing the management of data limited stocks.

Jardim *et al.* (2015) tested a simplified version of the rule where components r and f were tested one-at-a-time, and component b excluded, and concluded that the rule based on r (equation 1) performed the poorest, and while the rule based on f was able to reverse decreasing trends in biomass, it resulted in catch levels below MSY and could not prevent some stocks declining when subject to over-exploitation.

As the purpose of this paper is to test catch rules for data limited stocks, assumptions and approximations must be made. We used a similar approach to Jardim *et al.* (2015) where stocks are simulated based on a set of life-history parameters, and where fishing scenarios are developed. The simulations were conducted in the Fisheries Library in R (FLR, Kell *et al.*, 2007) software suite, within an MSE framework originally developed by Jardim *et al.* (2017) for data-rich stock but adapted and extended to accommodate data-limited stocks. Furthermore, the FLR package FLife, is used to simulate stocks based on life-history parameters.

The study stocks are given in Table 1; there are 29 data-limited stocks from European waters (North Sea region, Celtic Sea region, Bay of Biscay and widely distributed stocks) and encompass a wide range of life histories, including roundfish, flatfish, elasmobranchs, shellfish and demersal as well as pelagic species. Jardim *et al.* (2015) used averaged life-history parameters for species to simulate stocks; in contrast, in the present study we chose parameters from particular stock units, so that simulated stocks resemble real stocks in terms of biology (growth, productivity, etc). As this is a data limited simulation approach, however, artificial fishing histories had to be developed.

METHODS

Simulation of stocks

Operating models were conditioned for twenty-nine stocks, simulated based on a limited set of life-history parameters: allometric parameters for length-weight conversion, a and b, von Bertalanffy growth model parameters L_∞, k and t_0 (Von Bertalanffy, 1950), and age at 50% maturity a_{50}. Based on these data, using the FLR (Kell *et al.*, 2007) package FLife, and closely following the approach of Jardim *et al.* (2015), age structured operating models were created. Growth was modelled with the von Bertalanffy growth equation, recruitment by a Beverton-Holt stock recruit function with steepness $h = 0.75$, virgin recruitment set to 1000 (units) for all stocks, the maximum age a_{max} and plus-group set as the age (rounded up) where the stock reached 95% of L_∞, maturity modelled with a sigmoid function centred on a_{50}, and fisheries selectivity modelled with a sigmoid function where the first age at full selectivity equalled a_{50}. Natural mortality M was length-dependent, following Gislason *et al.* (2010). Survey selectivity was modelled with a sigmoid function and the inflection point set to $0.1a_{max}$ and the biomass index was derived by summing the survey catch biomass over all ages. Catch length frequencies were calculated based on the age distribution in the catch. Weights at age were converted first into discrete length classes using the allometric length-weight relationship, then numbers at length were created assuming a normal distribution around the discrete length classes. The final length distribution was derived by aggregating numbers at age in 1cm steps. Full specifications, including equations are given in the online supplementary material.

Two fishing histories were created for all simulated stocks. Initially, the stocks where fished at $0.5F_{MSY}$ for 75 years and subsequently for another 25 years in a roller-coaster or a one-way fishing scenario (Figure 1). In the one-way scenario, the fishing mortality was increased from $0.5F_{MSY}$ to $0.8F_{crash}$ within 25 years. In the roller-coaster scenario, the fishing mortality was increased from $0.5F_{MSY}$ to $0.75F_{crash}$, kept at $0.75F_{crash}$ for 5 years and then reduced to F_{MSY} by the end of the 25 years. After both fishing scenarios, the stocks were severely depleted; however, in the one-way scenario the stocks were at their lowest levels and declining, whereas in the roller-coaster scenario the stocks had started to recover. This exploitation state was then used as starting point for the MSE simulation.

Catch rule

The main catch rule tested sets catch advice based on the recent catch multiplied by three factors corresponding to perceptions of stock characteristics based on catch and survey data (equation 2). Component r corresponds to the trend in a biomass index, and is based on the “2 over 3” rule (equation 1):

$$r = \frac{\sum_{i=y-2}^{y-1} I_i/2}{\sum_{i=y-5}^{y-3} I_i/3} \qquad (3)$$

where I is the biomass index. Component f is a proxy for the ratio F_{MSY} divided by the current exploitation based on length data from the catch:

$$f = \frac{\bar{L}_{y-1}}{L_{F=M}} \qquad (4)$$

where $\bar{L}_{y-1}$ is the mean length in the catch above the length of first capture (L_c), weighted by catch numbers at length, with L_c defined as the first length class having at least 50% of the mode in the catch length frequency. The reference length $L_{F=M}$ is a proxy for the length at MSY proposed by Beverton and Holt (1957), under the assumption that $F = M$. Using the simplification that $M/k = 1.5$ the reference length can be calculated as:

$$L_{F=M} = 0.75L_c + 0.25L_\infty \qquad (5)$$

Finally, component b of the catch rule is a biomass safeguard protecting the stock when the biomass index drops below a threshold:

$$b = min\left\{1, \frac{I_{y-1}}{I_{trigger}}\right\} \qquad (6)$$

$I_{trigger}$ was based on the lowest historical biomass index value I_{loss} and defined as $I_{trigger} = 1.4I_{loss}$.

Projection

The OM was projected forward for a period of 100 years. Errors were implemented with a log-normal distribution and included for the biomass index ($sd = 0.2$, implemented individually on each age before compiling them into the biomass index), recruitment ($sd = 0.3$ and autocorrelation with $\rho = 0.2$), reference points ($sd = 0.1$), life-history parameters used in the calculation of catch length frequencies ($sd = 0.1$), catch numbers at length ($sd = 0.2$) and implementation of the advice into catch ($sd = 0.1$). The error distributions were set prior to running the simulation and random number deviates were identical for all stocks. Based on these uncertainties, 500 replicates were created for each stock.

Performance of the catch rule

The performance of the catch rule was assessed based on six performance statistics:

- catch/MSY: the catch (averaged over the simulated period), expressed as a proportion of the catch when fishing at F_{MSY},
- collapse risk: risk of stock collapse, i.e. the proportion of the projected stock where the stocks is below 0.1% of virgin spawning stock biomass (SSB),
- B_{lim} risk: risk of the stock falling below B_{lim} (proportion of the projected stock where the stock is below B_{lim}, defined as the stock level where recruitment is at 70% of recruitment achieved at virgin SSB, i.e. 16.3% of virgin SSB for all stocks, because they had the same value of steepness (h) for the Beverton-Holt stock recruitment relationship),
- ICV: inter-annual variability in catch, calculated as follows:

$$\text{ICV} = \text{median over 500 replicates of} \frac{1}{n}\sum_{y \in Y} \left| (C_y - C_{y-1}) / C_{y-1} \right|$$

  where Y is the set of n years in the projection period for which a TAC is set (which could be a biennial TAC)

- SSB/B_{MSY} and F/F_{MSY} : stock status (SSB and F relative to MSY reference points B_{MSY} and F_{MSY}, respectively, averaged over simulation period)

194 Initial analysis revealed that for some stocks and scenarios the stocks collapsed, and catches were reduced to zero as a result. Depending on stock productivity, some stocks subsequently recovered towards virgin biomass due to the zero catch. This behaviour was deemed inappropriate for further exploration of the performance as it implied a reduced risk. Consequently, running the simulations, once a replicate of a scenario had collapsed, the stock level and catch in subsequent simulation years were both set to zero.

200 *Penalized regression*

201 Many of the life-history parameters (both primary parameters used to create stocks, and parameters derived from the simulated stocks) are highly correlated. For example, natural mortality M, von Bertalanffy growth model parameter k, F_{MSY}, MSY, and population growth rate g and conditional growth rate g_c had positive Pearson correlation coefficients $\rho \geq 0.92$ between each other, and k and L_∞ correlated negatively with $\rho = -0.70$. Therefore, in order to determine which of the stock characteristics influenced the performance of the catch rule, ordinary linear models were of limited value for the analysis. Consequently, a penalized regression model was applied (glmnet, Friedman *et al.*, 2010) because this provides procedures for fitting the entire elastic-net regularization path from lasso to ridge regression (Hoerl and Kennard, 1988; Tibshirani, 1996; Zou and Hastie, 2005). A multi-Gaussian model was applied that selected the predictor variable(s) that could explain all six performance statistics (catch/MSY, collapse risk, B_{lim} risk, ICV, SSB/B_{MSY} and F/F_{MSY}). Firstly, only the primary input parameters were used as predictor variables: a, b (length-weight relationship), L_{inf}, k, t_0 (von Bertalanffy growth model parameters), a_{50} (age at 50% maturity). Secondly, the analysis was repeated with additional derived parameters: α, β (Beverton-Holt stock recruitment model parameters), $spr0$ (spawning potential ratio), L_{opt} (mean length when the stock is at MSY level), g, g_c (population growth rate and conditional growth rate at MSY), M (natural mortality), M/k, F_{MSY}/M (F_{MSY} relative to M) and B_{MSY}/B_0 (B_{MSY} relative to virgin biomass, i.e. location of peak in production curve).

Clustering

In order to identify groups of similar-performing stocks for the range of life histories tested, a time-series clustering approach was adopted. The Dynamic Time Warping technique (DTW, Berndt and Clifford, 1994; Aghabozorgi *et al.*, 2015) was selected as a distance measure and clustering performed on the relative stock status SSB/B_{MSY}. Several clustering algorithms (partitional, fuzzy, hierarchical) were trialled. Partitional and fuzzy clustering imply stochasticity, because the results depend on the random location of where the algorithm starts. This proved unreliable for the cluster analysis presented here, because the results were unstable, and even iterating the analysis did not lead to stable clusters. Hierarchical clustering on the other hand does not rely on stochasticity for the formation of the clusters. Additionally, once a hierarchical cluster analysis is conducted, the output can be visualised in a dendrogram and any arbitrary number of clusters can be pursued without having to rely on potentially biased cluster validity indices to select the optimum number of clusters.

Modifications to the catch rule

Various modifications of the catch rule were explored. One option tested was the addition of a multiplier x to the catch rule:

$$C_{y+1} = C_{y-1}\, r\, f\, b\, x. \tag{7}$$

Multipliers of $x =$ 0.5, 0.6, 0.7, 0.8, 0.85, 0.9, and 0.95 were evaluated.

By default, the catch rule does not include any constraints on the catch advice. In order to examine the impact of constraints on the performance of the catch rule, inter-annual limits on the relative variation in catches were evaluated. The constraints were defined as the maximum change in the advised catch compared to the last catch. Eight lower limits (0, i.e. no constraint, 0.5, 0.6, 0.7, 0.75, 0.8, 0.85 and 0.9), seven upper limits (1.1, 1.15, 1.2, 1.25, 1.3, 1.5 and no limit) and all combinations of these (56 combinations in total) were implemented.

By default, the management simulated here followed the ICES assessment cycle for data-limited stocks (ICES, 2012a, 2018a). This meant that the catch rule was applied in an intermediate (assessment) year

y based on data up to the previous year $(y-1)$ and the TAC was set biennially for the following two years $y+1$ and $y+2$. The data used in the catch rule were from the years up to the year before the intermediate year $(y-1)$, i.e. $y-1$ for the catch data for components C_{y-1} and f, $y-1$ for the index for b, and years $y-5 \ldots y-1$ for r. The effect of time-lags on management was explored by including data in the catch rule up to the intermediate year y. The survey index was calculated based on the stock at the beginning of a given year and could therefore be extended two years up to the advice year $y+1$. Additionally, setting the TAC annually instead of biennially was explored.

Perfect information scenario

Finally, to check whether the catch rule worked when all the information available to the catch rule was available without error, an additional scenario was run for all the simulated stocks and fishing histories. For these scenarios, only recruitment variability was implemented. The survey index was replaced with the SSB from the operating model to remove the impact of survey selectivity, $I_{trigger}$ was set to exactly $B_{trigger}$, which, in agreement with ICES data limited guidelines (ICES, 2018b), was set to $0.5B_{MSY}$. This modification meant that the biomass threshold was set irrespective of the historical exploitation and comparable for all stocks. The reference length for the f component of the catch rule was defined as the equilibrium length obtained in the operating model when fished at F_{MSY}.

RESULTS

Figure 1 shows the median trajectories for the 29 simulated stocks when the catch rule was implemented for the two fishing history scenarios. In the one-way scenario, seven (anchovy, brill, herring, John Dory, lemon-sole, sandeels, European pilchard and whiting) out of the 29 stocks collapsed by the end of the 100-year simulation period. In the roller-coaster scenario, two additional stocks (tub gurnard and black seabream) collapsed. The remaining stocks survived and displayed stock-specific long-term oscillations. One stock, megrim, approached virgin SSB and the other stocks reached terminal biomass values between 7% - 63% of virgin SSB.

In general, the catch rule was influenced most by component r (equation 3). Figure 2 shows the time series of the individual components for two example stocks. In the beginning, after the implementation of the catch rule, component b (equation 6) acted and reduced the catch, however, this effect lasted only for a few years. Component f (equations 4-5) gave some information throughout the entire simulation period but at a markedly lower magnitude compared to r.

Penalized regression

Performing a lasso regression with the primary input parameters resulted in a model fit that selected only the von Bertalanffy growth parameter k to explain the six performance statistics for the one-way fishing scenario (Figure 3). Allowing elastic-net regularization in the penalized regression model led to minor improvements in the model fit (the mean squared error was reduced from 0.45 to 0.41), but came at the cost of adding complexity to the model by returning non-zero coefficients for all supplied input parameters. Consequently, k was selected as the single most important factor for the performance of the catch rule for the simulated stocks. This was particularly evident for the risk and catch. Higher values of k were linked to higher risks (both collapse risk and B_{lim} risk) and lower long-term catch. Stocks with very low collapse risks were clustered at $k \leq 0.32$ whereas collapse risks above 20% were only observed for stocks with $k > 0.32$. When also using ~~also~~ derived input parameters as predictor variables, the lasso regression selected only the conditional population growth rate at MSY (g_c) and using the elastic-net regularization resulted in minor improvements but retained all provided input parameters.

Clustering

Clustering was performed on the median of the SSB/B_{MSY} time-series for the 29 simulated stocks. Figure 4 shows the results from the hierarchical clustering for up to four clusters for the one-way fishing history. Hierarchical clustering does not compute centroids for the clusters; for plotting purposes (Figure 4B), centroids for the clusters were calculated post-hoc as the annual average of the SSB/B_{MSY} values of all stocks within a cluster. If all stocks were kept in a single cluster, the centroid SSB/B_{MSY} trend showed a recovery after the start of the MSE simulation and equilibrated at a level slightly above

294 1. The first separation in the hierarchical cluster distinguished between two distinct patterns (second row in Figure 4B); the first cluster was composed of stocks that experienced early peaks but collapsed by the end of the simulation period, whereas the stocks in the second cluster survived. This split corresponds well to the von Bertalanffy k values for these stocks (Figure 4C). The first cluster (collapsed) is comprised of stocks with $k > 0.32$. On the other hand, the stocks with lower k ($k \leq 0.32$) survived.

Following the dendrogram further, the next two splits occurred within the cluster of surviving stocks. Firstly, there is a separation of stocks that stay around B_{MSY} in the long term and ~~the~~ ones that end up markedly above B_{MSY} (third row of Figure 4B). These stocks are mainly characterised by k values around the median of the simulated range, although two of the stocks inside this cluster have k values at the lower end of the total range (megrim and redfish). For megrim, the catch was reduced substantially at the beginning of the simulation and approaches zero; consequently, the stock moves towards virgin biomass. Redfish displayed a similar behaviour; however, the catch recovered later in the simulation and the stock declined again from very high levels. Secondly, the stocks reaching levels above B_{MSY} are divided further into one cluster where the SSB converged at around $2B_{MSY}$ and one cluster where the SSB approaches levels close to $3B_{MSY}$ (fourth row of Figure 4B). In terms of k, these stocks overlap and no clear distinction is evident. Moving further along the dendrogram, these clusters are divided further; however, clusters increasingly represent individual stocks instead of general trends, because stocks are singled out as the number of clusters grows. The clusters in Figure 4 are colour-coded and this colour-code is maintained throughout the study. Results in this figure are for the one-way trip scenario, but results for the roller-coaster scenario are almost identical when considering four clusters.

Modifications to the catch rule

Adding a multiplier (x in equation 7) of less than one to the catch rule reduced the risk (both collapse risk and B_{lim} risk) for all stocks and for both fishing scenarios (Figure 5). This risk reduction was a result of higher terminal SSB values, and the smaller the multiplier, the higher the SSB values, capped at the top at the virgin biomass level. For the stocks where the median SSB collapsed during the simulation period (cluster 1), adding the multiplier delayed this collapse, and by reducing the multiplier further, the collapse was avoided altogether. This behaviour of the SSB trajectory was stock specific. For example, in the default catch rule, the median SSB of anchovy in the one-way fishing scenario reached zero roughly 40 years after the start of the simulation and adding a multiplier of only 0.95 avoided this collapse. On the other hand, pilchard and John Dory collapsed in the roller-coaster fishing scenario within approximately five years, and this collapse could only be averted by implementing a multiplier of 0.8 or below.

The performance of the catch rule for these cluster 1 stocks was highly sensitive to small changes of the multiplier. Once a threshold multiplier was reached, the long-term stock levels increased rapidly and overshot B_{MSY}, thereby foregoing catch. Stocks in cluster 2 were kept around B_{MSY} in the long term when the catch rule was applied without a multiplier. Introducing the multiplier for these stocks reduced their risks but moved them above B_{MSY}. Stock levels for stocks from clusters 3 and 4 where shifted further above B_{MSY} when the multiplier was added. For 16 of the 29 stocks tested, adding the multiplier reduced the catch. For the remaining 13 stocks, the maximum was achieved within a range 0.9–0.95. When considering all stocks together, there does not seem to be a multiplier that increases risk performance for all stocks without jeopardizing catch for some.

Implementing an upper catch constraint reduced the risks for all stocks, and more restrictive constraints led to lower risks (Figure 6A). For most stocks, the catch was reduced when upper catch constraints were used. An exception was for the stocks from cluster 1. For the one-way fishing history, the catch peaked at upper constraints between 1.15 and 1.3, and for the roller-coaster scenario, the catch increased up to the most restrictive constraint (1.1). For most of the remaining stocks, the catch is relatively stable for constraints at or above 1.2, and this value seems to be a reasonable compromise between risk reduction and maximising catch. In general, including a lower constraint on the catch increased the risk of stock collapse and resulted in subsequent reductions in catch. If the lower constraint was implemented in combination with an upper constraint, for some stocks a small peak in catch was observed at lower constraint levels above 0 and below 1. Figure 6B shows the effect of including lower catch constraints on the performance of the catch rule in combination with an upper constraint of 1.2.

More restrictive lower constraints (i.e. restricting catch reductions) caused a large increase in risks and a large decrease in catch, with this behaviour being particularly pronounced at constraint levels above 0.7. Below 0.7, the risks and catches were relatively stable.

For the stocks surviving the default implementation of the catch rule ($k \leq 0.32$), using more recent data and setting the TAC more frequently improved performance by reducing oscillations and reaching final biomass values earlier (Figure 7). The lowest fluctuations were observed when the TAC was set annually, the catch data provided up to the intermediate year, and the survey data up to the beginning of the advice year. The terminal biomass values were similar irrespective of the timing. Some of the high k stocks (cluster 1, $k > 0.32$) could be saved; however, two stocks (John Dory and herring) still collapsed even if the TAC was set annually and the most recent data was used.

Perfect information scenario

When the catch rule was implemented with perfect information ~~and knowledge~~ (i.e. the SSB from the operating model was used as the index and $I_{trigger}$ set to $0.5B_{MSY}$ from the operating model), the performance of the catch rule was substantially improved and most stocks converged towards B_{MSY}, indicating that the catch rule did work under these unrealistically perfect conditions (Figure 8). However, stocks with higher k values generally displayed a stronger oscillatory behaviour. Among the high k stocks ($k > 0.32$), three stocks survived (brill, whiting and lemon sole) in the one-way fishing history scenario but the remaining stocks, and all high k stocks in the roller-coaster fishing scenario, showed poor performance with collapses. The highest k stock, sandeel, showed a recovery to very high biomass levels, but this behaviour could be attributed to the fact that this stock was at the brink of stock collapse, with catches reduced to very low levels, and consequently the stock could recover with almost no fishing activity.

DISCUSSION

This study simulation tested a simple catch rule making use of proxy MSY reference points for a range of data-limited fish stocks. The main result was that the performance of the catch rule was stock specific and could broadly be linked to life-history characteristics, with the von Bertalanffy growth parameter k emerging as the most important one from a penalized regression model. When also using derived input parameters, the conditional growth rate at MSY (g_c) appeared as crucial factor determining the performance of the catch rule. However, this parameter might not always be readily available and is computed from the primary input parameters. Therefore, we focused the analysis on k.

It was clear from a visual inspection of the results that the response of stocks to the application of the catch rule could be organised into different groups, and therefore a time series clustering approach using dynamic time warping was adopted. The relative stock status SSB/B_{MSY} was selected as a time series index because it provided the overall best indicator of the performance of the catch rule over time. Biomass was used in relative terms because the catch rule's long-term target is MSY, and consequently both undershooting (overfishing) and overshooting (loosing yield through fishing below MSY) of B_{MSY} could be identified and was comparable across all simulated stocks. Both the clustering analysis and the penalized regression approach indicated that there is a clear relationship between the life histories of the simulated stocks, and the performance of the catch rule. The most important finding is the separation of the simulation trajectories into two groups: one where the stocks collapsed during the simulation, and the other where the stocks survived and ended up at or above B_{MSY}. The split corresponded well to the von Bertalanffy growth parameter k and the catch rule seemed to perform reasonably for stocks with $k \leq 0.32$, but very poorly for stocks with $k > 0.32$. The $k \leq 0.32$ stocks reached levels of between B_{MSY} and $3B_{MSY}$, i.e. stock collapses were avoided, but frequently there was a loss of~~n~~ yield compared to the yield achieved when fishing at F_{MSY}.

The result that the catch rule performed worse for more productive stocks ($k \geq 0.32$) compared to less productive stocks ($k \leq 0.32$) might at first glance appear counter-intuitive. However, the performance of the catch rule, as measured by the summary statistics, is an emergent property of the interaction between the operating model and the catch rule. We showed that the advised catch was mostly influenced by the r component of the rule (the stock trend) and stocks with higher k are inherently more variable, which led to higher fluctuations in the catch. When subjected to the catch rule, the higher k stocks collapsed on average early during the simulation. This behaviour can be attributed to an initial rapid recovery, also causing the catch to increase, but once the stocks started to decline again, the catch was not reduced quickly enough to avoid stock collapse. This undesirable feature is caused by the design of the catch rule, which bases the newly advised catch on the previous catch and observed data with a time-lag. The less productive stocks were also less variable, and the catch rule was sufficiently reactive to avoid stock collapse.

Previous studies have tested simple empirical data-limited catch rules with various simulated stocks (e.g. Jardim *et al.*, 2015), or based operating models on knowledge from fully ~~analytical~~ quantitative stock assessments (e.g. Geromont and Butterworth, 2015a; Carruthers *et al.*, 2016). In the simulation exercise of Carruthers *et al.* (2016), various data-limited methods have been tested, but only three stocks (Pacific herring, Atlantic bluefin tuna and Pacific canary rockfish) were simulated, and therefore, possible inferences from life-histories were limited. Jardim *et al.* (2015) tested a simplified version of the catch rule tested here, including only a single component at a time (either r using survey data or f using length frequency data). The results from their simulation study are in agreement with the current work, showing a wide range of stock trajectories and yields often below MSY. The basis for the simulation of the stocks in Jardim *et al.* (2015) were averaged life-history parameters to generate a variety of life-history traits. For the work presented here, we went one step further and used life-history parameters from real stock units; by doing so, we were able to link the performance of the catch rule back to the original life-history parameters.

Modifications to the catch rule (multipliers, catch constraints, using more recent data) were able to improve the performance of the catch rule. However, the improvement was stock specific and a trade-off between yield and risk was evident. Although application of the multiplier always reduced the risk, the stocks frequently ended up above B_{MSY}, and the catch rule was overly reactive to minor changes of the multiplier for higher k stocks, not a good feature in a situation of high uncertainty. For stocks for which the catch rule kept the stock at or above B_{MSY} in the long term, the multiplier moved the stock level further away from B_{MSY} and reduced yield. Stocks which collapsed when the default catch rule was applied (the higher-k stocks) could be saved, but only at the cost of moving the stocks far above B_{MSY} and losing yield.

Regarding the catch constraint, an upper limit of 1.2 was deemed appropriate because the long-term yield hardly changed for most stocks if less restrictive constraints were implemented; furthermore, this value provides an important reduction in risk compared to the application of the catch rule without any constraints. For this level of upper constraint, a lower constraint of 0.7 seemed to be a suitable choice, because implementing more restrictive lower constraints would cause a large increase in risk and a drop in yield. Less restrictive lower constraints did not have much impact on either yield or risk.

As could be expected, more recent data did improve the performance of the catch rule, mainly by reducing oscillations, but this approach did not prove successful for the high-k stocks.

Challenges remain for the catch rule tested here. For example, the components of the catch rule make use of different commercial and scientific data and are designed to account for stock dynamics. However, if just one of the components of the catch rules fails or produces very low (close to zero) or high values, it will inevitably overrule the other components and dominate the final catch advice; in such circumstances, the use of the catch constraints become important. The analysis into the components of the catch rule showed that the rule is mainly dominated by the trend in the index, frequently masking information from the other components. The biomass safeguard is important to recover the stock above a threshold, but depending on how this level is set, it may not be effective enough (e.g. if the threshold is set too low). The problem of dominant components of the rule could be dealt with through variable weighting of the different components and is a subject of future work.

If there is perfect information available, (catch data, survey index, mean length in the catch) and reference points were set correctly according to MSY, then the catch rule performed well and approached the desired MSY target for low-to-medium-k stocks. The results from these perfect information scenarios showed the importance of setting reference points appropriately, because, for example, setting the index trigger value dependant on the fishing history based on the lowest ever observed value governed where the biomass ended up. The lower-k stocks were less depleted relative to B_{MSY}, and therefore, the trigger point in the b-component of the catch rule was higher, which in turn resulted in a higher terminal biomass when the stocks were subjected to the catch rule. In a real-life application of the catch rule to data-limited stocks, reference values are uncertain, possibly impeding the performance of the rule.

One concern raised during this work was the appropriateness of using a constant steepness of 0.75 in the recruitment model for all stocks. We addressed this issue by conducting additional sensitivity runs (detailed in the supplementary online material) and found that changing steepness to 0.9 or imposing a positive linear relationship between steepness and k did not affect the conclusions.

The starting point for the simulations in this study represented highly depleted stocks and might be considered as a worst case. We used this condition to examine whether the catch rule was able to correctly identify the depletion and recover stocks. If a catch rule works in such harsh conditions, the rule is likely to work well in less depleted states. Additionally, due to the long simulation period (100 years), all stocks moved away from their initial state during the simulation and this provided insight into whether a long-term equilibrium was reached.

Trends and fluctuations in populations are determined by complex interactions between extrinsic forcing and intrinsic dynamics. For example, stochastic recruitment can induce low-frequency variability, i.e. 'cohort resonance', which can induce apparent trends in abundance and may be common in age-structured populations; such low-frequency fluctuations can potentially mimic or cloak critical variation in abundance linked to environmental change, over-exploitation or other types of anthropogenic forcing (Bjørnstad *et al.*, 2004). Although important, these effects can be difficult to disentangle. The simulations so far show that life histories are important and should be used to help condition operating models to ensure robust feedback-control rules. MSE is important to help develop these robust feedback control rules and to help identify appropriate observational systems.

Although the performance of the HCR depended on the life-history characteristic, it was not in the way initially expected, i.e. the outcomes could not be grouped solely by whether the operating models represented fast growing vs. late maturing species or demersal vs. pelagic stocks. What was important was the nature of the dynamics, i.e. how variable was the stock between years; for example, a stock could exhibit high interannual variability if natural mortality and recruitment variability was high, regardless of the values of k, L_{inf}, L_{50}. The nature of the indices is also important; for example, even if a stock had low interannual variability, an index could be highly variable if it was based on juveniles or there were large changes in spatial distribution between years. It is therefore necessary to look at the robustness of management strategies to the nature of the time-series of the stock (as represented by the operating model), and to the characteristics of the data collected from it. This will require tuning by constructing a reference set of operating models and then tuning the management strategy to secure the desired trade-offs. The work so far can be considered as focusing first on developing management strategies that perform satisfactorily for a reference set; the next step is to develop case-specific strategies.

Finally, simple empirical management procedures are usually considered in a data-limited context; such simple rules can sometimes achieve similar performance compared to management procedures based on fully ~~analytical~~ quantitative assessments (e.g. shown by Carruthers *et al.*, 2014; Geromont and Butterworth, 2015b) or even out-perform them, particularly if operational effort is included.

SUPPLEMENTARY MATERIAL

The following supplementary material is available at ICESJMS online: A document describing the operating models, including input parameters and equations and an analysis of the sensitivity of the study to the stock recruitment model steepness and variability.

ACKNOWLEDGEMENTS

The authors would like to thank the FLR developers at the Joint Research Centre of the European Commission for ongoing effort into FLR and to provide an easily adaptable MSE framework, the ICES WKLIFE (Workshop on the Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks) workshop series which played an important part for this work. Part of this work was carried out within

Defra project MF1253. LK's involvement was funded through the MyDas project under Marine Biodiversity Scheme, which is financed by the Irish government and the European Maritime & Fisheries Fund as part of the EMFF Operational Programme for 2014-2020.

REFERENCES

Aghabozorgi, S., Shirkhorshidi, A. S., and Wah, T. Y. 2015. Time-series clustering – A decade review. Information Systems, 53: 16–38.

Bentley, N. 2015. Data and time poverty in fisheries estimation: potential approaches and solutions. ICES Journal of Marine Science, 72: 186–193.

Berndt, D. J., and Clifford, J. 1994. Using Dynamic Time Warping to Find Patterns in Time Series. *In* KDD-94 Workshop on Knowledge Discovery in Databases, pp. 359–370. Seattle.

Beverton, R. J. H., and Holt, S. J. 1957. On the Dynamics of Exploited Fish Populations. HMSO for Ministry of Agriculture, Fisheries and Food, London.

Bjørnstad, O. N., Nisbet, R. M., and Fromentin, J.-M. 2004. Trends and cohort resonant effects in age-structured populations. Journal of Animal Ecology, 73: 1157–1167.

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, 119–120: 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.

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, 24: 25–38.

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, 338: 517–520.

Dowling, N., Wilson, J., Rudd, M., Babcock, E., Caillaux, M., Cope, J., Dougherty, D., *et al.* 2016. FishPath: A Decision Support System for Assessing and Managing Data- and Capacity- Limited Fisheries. *In* Assessing and Managing Data-Limited Fish Stocks. Alaska Sea Grant, University of Alaska Fairbansk.

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.

Dowling, N. A., Smith, A. D. M., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury, K., Wilson, J. R., *et al.* 2019. Generic solutions for data-limited fishery assessments are not so simple. Fish and Fisheries, 20: 174–188.

Fitzgerald, S. P., Wilson, J. R., and Lenihan, H. S. 2018. Detecting a need for improved management in a data-limited crab fishery. Fisheries Research, 208: 133–144. Elsevier.

Friedman, J., Hastie, T., and Tibshirani, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33: 1–20.

Garcia, S. M. 1995. The precautionary approach to fisheries and its implications for fishery research, technology and management: an updated review. *In* Precautionary approach to fisheries Part 2: Scientific papers. FAO FISHERIES TECHNICAL PAPER 350/2.

Geromont, H. F., and Butterworth, D. S. 2015a. Generic management procedures for data-poor fisheries: forecasting with few data. ICES Journal of Marine Science, 72: 251–261.

Geromont, H. F., and Butterworth, D. S. 2015b. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. ICES Journal of Marine Science, 72: 262–274.

Gislason, H., Daan, N., Rice, J. C., and Pope, J. G. 2010. Size, growth, temperature and the natural mortality of marine fish. Fish and Fisheries, 11: 149–158.

Hillary, R. M., Preece, A. L., Davies, C. R., Kurota, H., Sakai, O., Itoh, T., Parma, A. M., *et al.* 2016. A scientific alternative to moratoria for rebuilding depleted international tuna stocks. Fish and Fisheries, 17: 469–482.

Hoerl, A., and Kennard, R. 1988. Ridge Regression. *In* Encyclopedia of Statistical Sciences. Volume 8. Regressograms - St. Petersburg Paradox, The. Ed. by S. Kotz, N. L. Johnson, and C. B. Read. Wiley.

ICES. 2012a. ICES Implementation of Advice for Data-limited Stocks in 2012 in its 2012 Advice. ICES CM 2012/ACOM 68: 42 pp.

ICES. 2012b. Report of the Workshop 3 on Implementing the ICES Fmsy Framework , 9-13 January 2012, ICES, Headquarters. ICES CM 2012/ACOM:39: 33 pp.

ICES. 2013a. General Context of ICES Advice. *In* Report of the ICES Advisory Committee 2013. ICES Advice, 2013. Book 1, pp. 3–24. Copenhagen.

ICES. 2013b. Report of the Workshop on the Development of Quantitative Assessment Methodologies based on LIFE-history traits, exploitation characteristics, and other key parameters for Data-limited Stocks (WKLIFE III), 28 October–1 November 2013, Copenhagen, Denmark. ICES CM 2013/ACOM:35: 98 pp.

ICES. 2017a. Report of the ICES Workshop on the Development of Quantitative Assessment Methodologies based on Life-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks in categories 3-6 (WKLIFE VI), 3-7 October 2016, Lisb. ICES CM 2016/ACOM:59: 106 pp.

ICES. 2017b. Report of the ICES Workshop on the Development of Quantitative Assessment Methodologies based on Life-history traits, exploitation characteristics, and other relevant parameters for data-limited stocks in categories 3-6 (WKLIFE VII), 2-6 October 2017, Lis. ICES CM 2017/ACOM:43: 221 pp.

ICES. 2017c. Report of the Workshop on the Development of the ICES approach to providing MSY advice for category 3 and 4 stocks (WKMSYCat34), 6–10 March 2017, Copenhagen, Denmark. ICES CM 2017/ACOM:47: 53 pp.

ICES. 2018a. Advice basis: General context of ICES advice. *In* Report of the ICES Advisory Committee 2018. ICES Advice 2018. Copenhagen.

ICES. 2018b. ICES reference points for stocks in categories 3 and 4. ICES Technical Guidelines. Copenhagen.

Jardim, E., Azevedo, M., and Brites, N. M. 2015. Harvest control rules for data limited stocks using length-based reference points and survey biomass indices. Fisheries Research, 171: 12–19. Elsevier B.V.

Jardim, E., Scott, F., Mosqueira, I., Citores, L., Devine, J., Fischer, S., Ibaibarriaga, L., *et al.* 2017. Assessment for All initiative (a4a) Workshop on development of MSE algorithms with R/FLR/a4a. EUR 28705 EN, Publications Office of the European Union. Luxembourg.

Kell, L. T., Pilling, G. M., Kirkwood, G. P., Pastoors, M., Mesnil, B., Korsbrekke, K., Abaunza, P., *et al.* 2005. An evaluation of the implicit management procedure used for some ICES roundfish stocks. ICES Journal of Marine Science, 62: 750–759.

Kell, L. T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., *et al.* 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES Journal of Marine Science, 64: 640–646.

Punt, A. E., Butterworth, D. S., de Moor, C. L., De Oliveira, J. A. A., and Haddon, M. 2016. Management strategy evaluation: best practices. Fish and Fisheries, 17: 303–334.

Rosenberg, A. A., Kleisner, K. M., Afflerbach, J., Anderson, S. C., Dickey-Collas, M., Cooper, A. B., Fogarty, M. J., *et al.* 2018. Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries around the World. Conservation Letters, 11: 1–9.

Smith, A. D. M. 1994. The light on the hill. Population Dynamics for Fisheries Management: 249–253.

Tibshirani, R. 1996. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58: 267–288.

Von Bertalanffy, L. 1950. An Outline of General System Theory. The British Journal for the Philosophy of Science, 1: 134–165.

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.

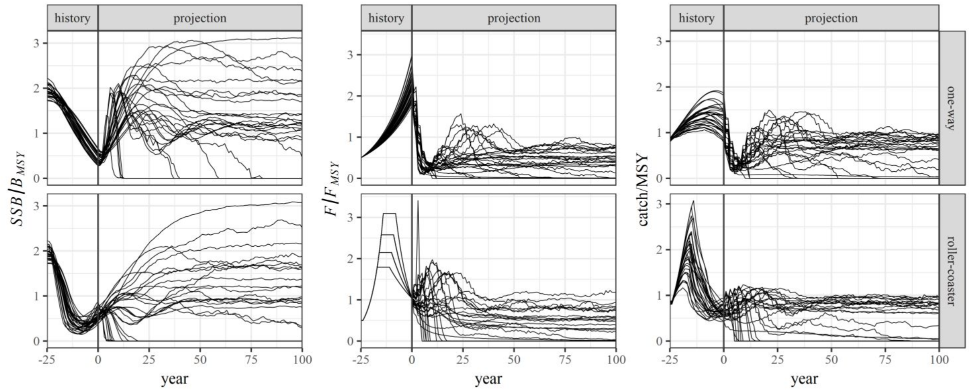
Zou, H., and Hastie, T. 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67: 301–320.

TABLES

Table 1. The 29 stocks on which the operating models are based. Given are the scientific and common names, a unique stock ID and the von Bertalanffy growth parameter k.

Scientific name	Common name	ID	*k*
Ammodytes spp.	sandeels	san	1
Anarchias lupus	Atlantic wolffish	wlf	0.11
Argentina silus	greater argentine	arg	0.23
Chelidonichtys lucerna	tub gurnard	gut	0.32
Clupea harengus	herring	her	0.606
Engraulis encrasicolus	anchovy	ane	0.44
Lepidorhombus whiffiagonis	megrim	meg	0.12
Lophius budegassa	blackbellied angler	ang3	0.08
Lophius piscatorius	angler	ang	0.18
Lophius piscatorius	angler	ang2	0.18
Melanogrammus aeglefinus	haddock	had	0.2
Merlangius merlangus	whiting	whg	0.38
Microstomus kitt	lemon sole	lem	0.42
Molva molva	ling	lin	0.14
Mullus surmuletus	striped red mullet	mut	0.21
Mustelus asterias	starry smooth-hound	sdv	0.15
Nephrops	Norway lobster	nep	0.2
Pleuronectes platessa	European plaice	ple	0.23
Pollachius pollachius	pollack	pol	0.19
Raja clavata	thornback ray	rjc2	0.14
Raja clavata	thornback ray	rjc	0.09
Sardina pilchardus	European pilchard	sar	0.6
Scophthalmus rhombus	brill	bll	0.38
Scopthalmus maximus	turbot	tur	0.32
Scyliorhinus canicula	lesser spotted dogfish	syc	0.15
Scyliorhinus canicula	lesser spotted dogfish	syc2	0.23
Sebastes norvegicus	golden redfish	smn	0.11
Spondyliosoma cantharus	black seabream	sbb	0.22
Zeus faber	John Dory	jnd	0.47

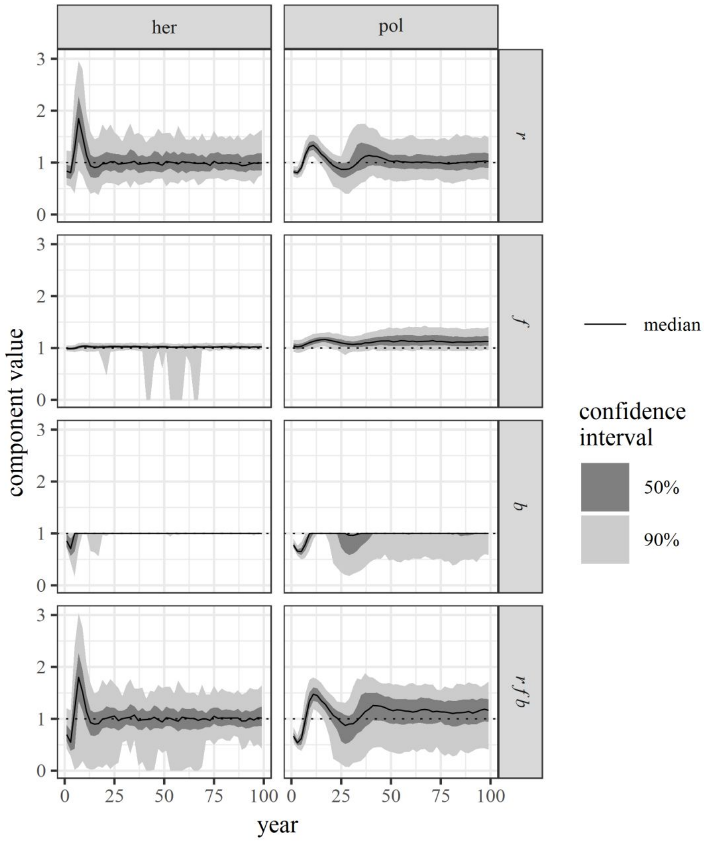
619 **FIGURES**



620

621 **Figure 1.** Median trajectories for spawning stock biomass (SSB), mean fishing mortality and catch
622 relative to MSY reference points for the 29 simulated stocks. Shown are the historical fishing period
623 ("history", years -25-0) and the results of subsequently applying the catch rule (years 1-100). The top
624 row shows the one-way fishing history and the bottom row the roller-coaster fishing history.

625



626

627 **Figure 2.** Components of the catch rule (r, f and b) and their product ($r\ f\ b$, which scales the recent
628 catch) for two example stocks: herring (her) and pollack (pol). The higher the deviation of a component
629 from one (up or down), the higher is its contribution in the catch rule. Please note that for herring, the
630 stock collapsed in most simulated replicates (the median SSB collapsed after 13 years) and in the
631 distributions shown for the components, these collapsed replicates were excluded because they did not
632 provide any stock status information.

633

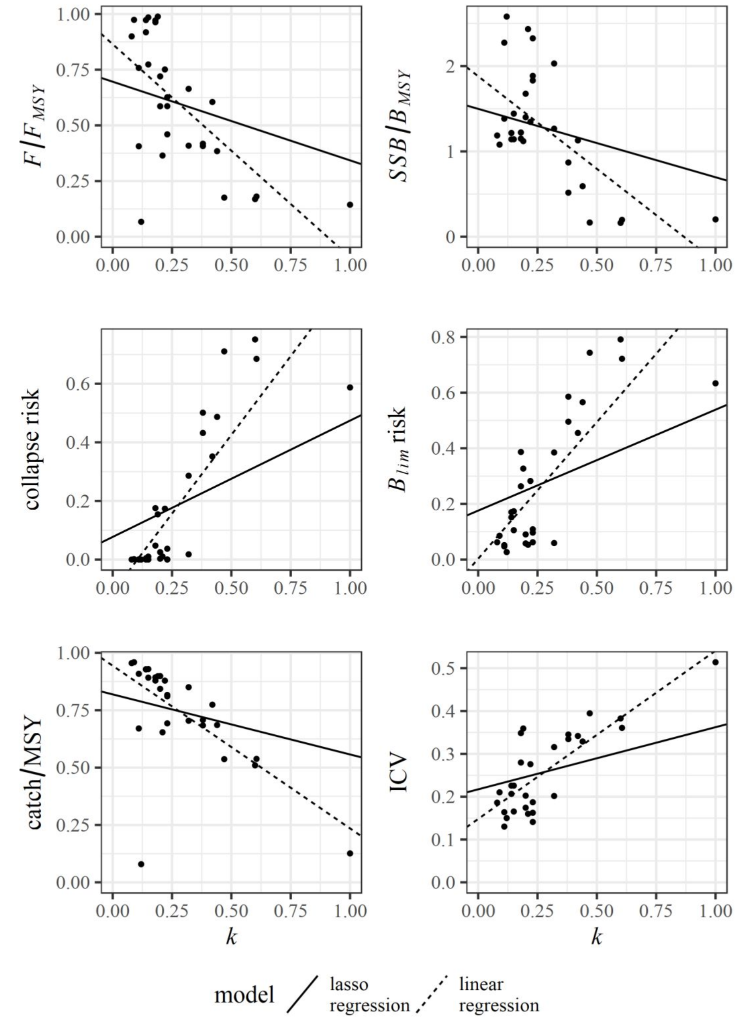


Figure 3. Six performance statistics versus the von Bertalanffy growth model parameter k for the tested catch rule and the one-way fishing history for all 29 stocks. The solid lines show the fit from the lasso regression model, and the dotted lines a linear regression of the data.

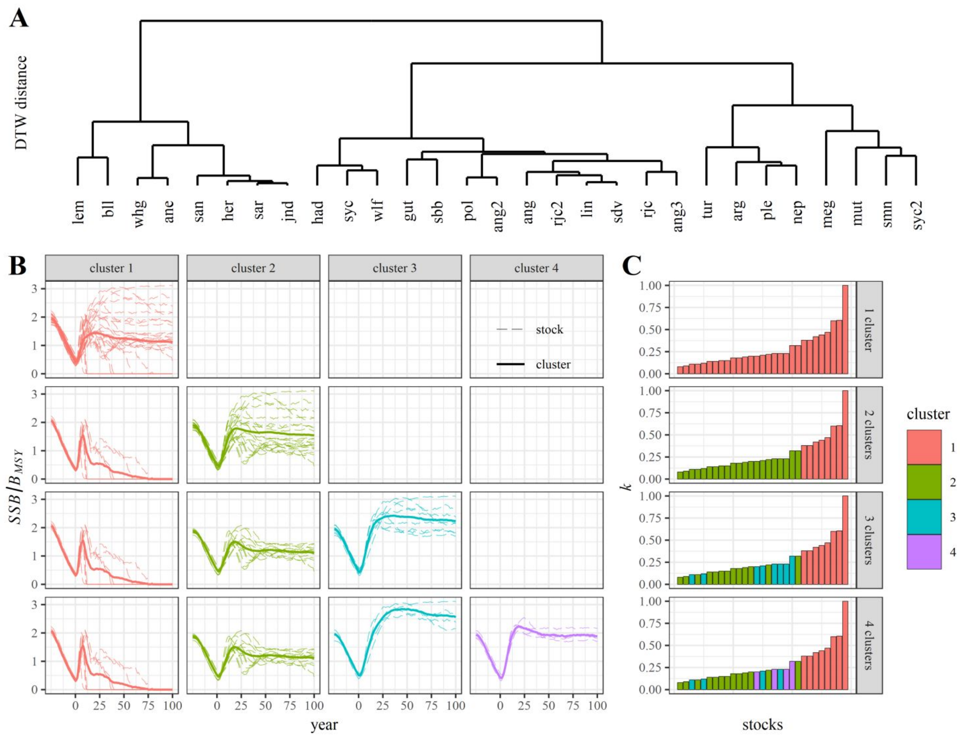


Figure 4. Results of the hierarchical clustering approach of relative SSB for the one-way fishing history. A shows a dendrogram of the time series for the 29 simulated stocks, the names correspond to the stock IDs defined in Table 1. The y-axis corresponds to the dynamic time warping (DTW) distance between the time series. B represents the median SSB/B_{MSY} times series for all stocks (dashed lines) and the centroids (solid bold line). Rows represent the number of clusters and each column is one cluster. C shows von Bertalanffy growth model parameter *k* for all stocks, sorted in ascending order and colour-coded for the clusters shown in B.

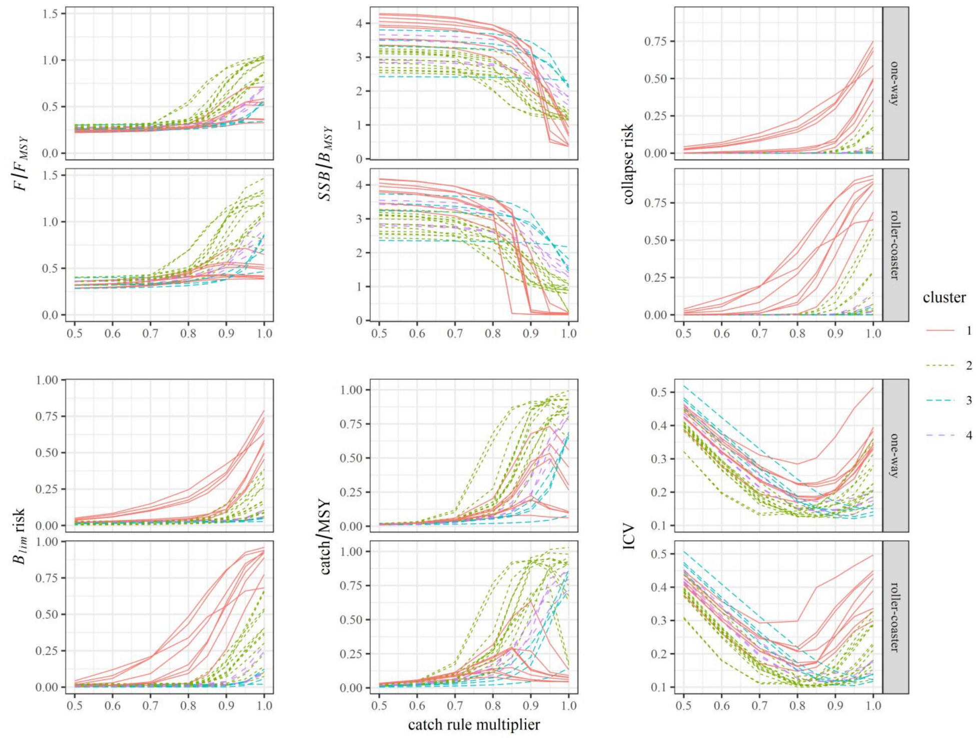


Figure 5. Effect of implementing a multiplier to the catch rule on the six performance statistics for the 29 simulated stocks and both fishing histories. The clusters correspond to the ones defined in Figure 4.

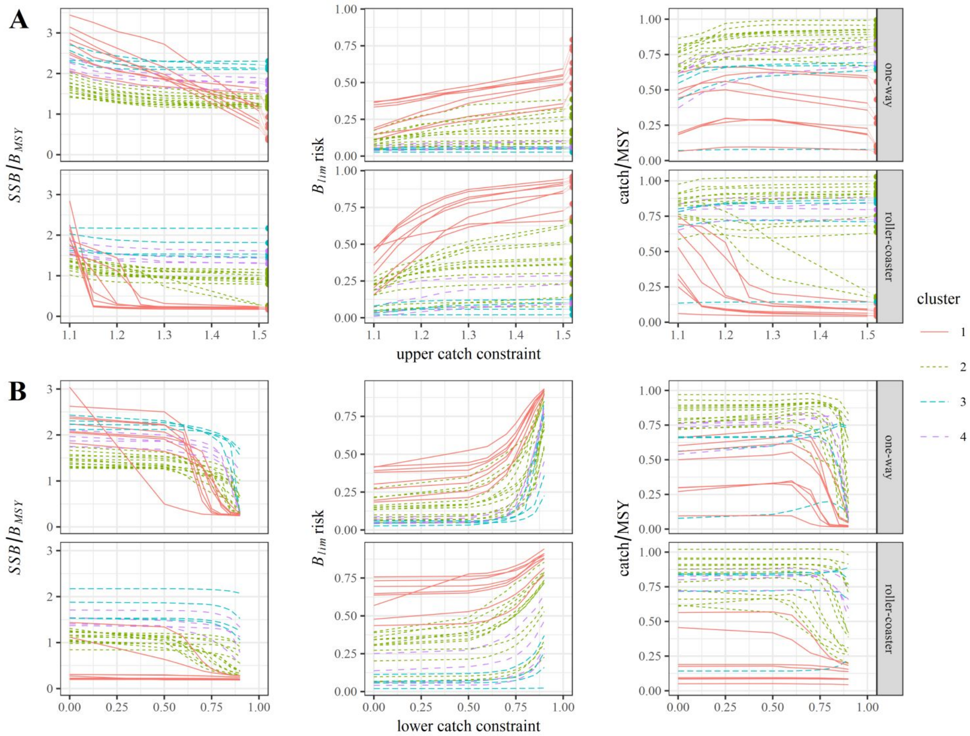


Figure 6. Effect of catch constraints on three performance statistics. **A** shows the effect of upper catch constraints. The points above an upper catch constraint of 1.5, connected with thin lines, indicate the performance when no upper catch constraint was implemented. **B** shows the effect of lower catch constraints in combination with an upper catch constraint of 1.2. The clusters correspond to the ones defined in Figure 4.

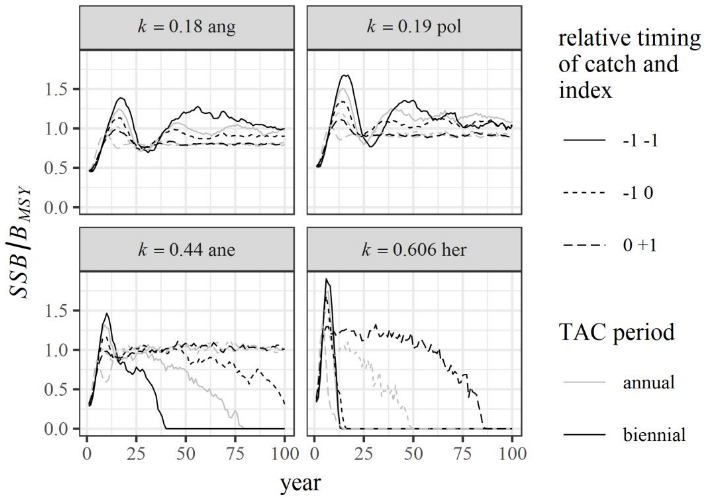


Figure 7. Effect of time-lags for the data used in the catch rule and periodicity of TAC-setting (annual vs. biennial) for four example stocks (sorted by von Bertalanffy *k*) in the one-way fishing scenario. The timing is relative to the intermediate year (0) and -1 refers to the year before, +1 to year after the intermediate year.

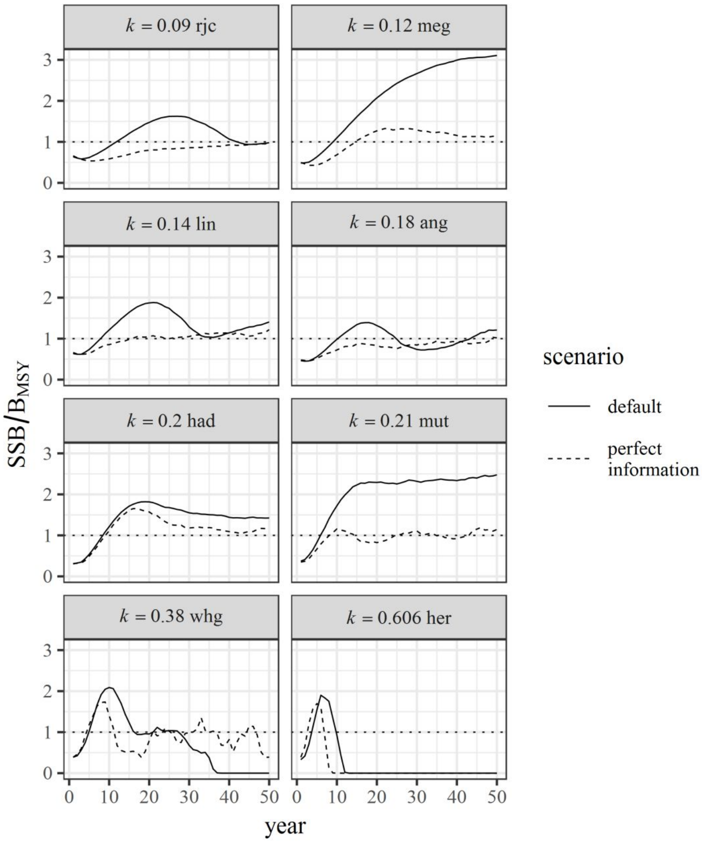


Figure 8. Application of the catch rule with and without perfect information for eight example stocks (as defined in Table 1) for the one-way fishing history. In the perfect information scenario, no uncertainty, apart from recruitment variability, has been implemented; the survey is an exact representation of the spawning stock biomass and $I_{trigger} = 0.5B_{MSY}$.

Canadian Science Publishing | Canadian Journal of Fisheries and Aquatic Sciences

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

Journal:	*Canadian Journal of Fisheries and Aquatic Sciences*
Manuscript ID	cjfas-2019-0276
Manuscript Type:	Article
Date Submitted by the Author:	16-Aug-2019
Complete List of Authors:	Pons, Maite; University of Washington, Aquatic and Fishery Sciences Cope, Jason; National Marine Fisheries Service - NOAA, Northwest Fisheries Science Center Kell, Laurie; Comision Internacional para la Conservacion del Atun Atlantico
Keyword:	data-limited assessment methods, depletion, life-history, harvest rates
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

Performance of catch-based and length-based stock assessment 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. Email: mpons@uw.edu

[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. Email: jason.cope@noaa.gov

[3] Centre for Environmental Policy, Imperial College London, London SW7 1NE, UK. Email: laurie@kell.es

[*] Corresponding author: School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, Washington 98195-5020, USA. Email: mpons@uw.edu. Telephone number: +1 206 883 5102.

Abstract

The quality of data from many small-scale fisheries is insufficient to allow for the application of conventional assessment methods. Even though in many countries they are moving to close-loop simulations to assess the performance of different management procedures in data limited situations, managers in most developing countries are still demanding information on stock status. In this study we use the common metric of harvest rate to evaluate and compare the performance of the following catch-only and length-only assessment models: Catch-Maximum Sustainable Yield (Catch-MSY), State-Space Catch-Only Model (SSCOM), Depletion Based Stock Reduction Analysis (DBSRA), Simple Stock Synthesis (SSS), an extension of Catch-MSY (CMSY), Length Based Spawning Potential Ratio (LBSPR), Length-Based Integrated Mixed Effects (LIME), and Length-Based Bayesian (LBB). In general, results were more biased for slightly depleted than for highly depleted stocks, and for long-lived than for short-lived species. Length-based models, such as LIME, performed as well as catch-based methods in many scenarios and, among the catch-base models the one with the best performance was SSS.

Keywords: data-limited assessment methods, depletion, life-history, harvest rates

INTRODUCTION

Major commercial fish species usually have substantial sets of data that can be integrated by complex stock assessment models (e.g., Methot and Wetzel 2013); these data may include time series of total removals, catch-at-length or -age, relative abundance indices, fishing effort, tag recoveries and information on life-history parameters. These datasets required for such stock assessments, however, are unavailable for most of the small-scale fisheries and by-catch species around the world. Fisheries and stocks lacking comprehensive datasets 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 increasing demand for science-based fisheries management of unassessed fisheries where data and 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 and life-history data started many years ago with the development of Stock Reduction Analysis (SRA; Kimura and Tagart 1982; Kimura et al. 1984). Since then, SRA 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 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-Maximum Sustainable Yield, Catch-MSY; Martell and Froese 2013), bayesian approches (CMSY, Froese et al. 2017), and age-structured population dynamics (Simple Stock Synthesis, SSS; Cope 2013). Despite these differences, this family of catch-only models share a common dependence on prior assumptions about final stock depletion. Simulation testing has previously indicated that these methods perform well only when assumptions regarding final relative abundance are met (Wetzel and Punt 2015). Unsurprisingly, because final stock depletion is a prior assumption, the methods perform differently under different stock depletion levels (i.e., highly depleted or slightly depleted stocks, Walters et al. 2006) or under different harvest history or catch trends.

For many small-scale fisheries, obtaining reliable time series on historical total catch is difficult, whereas sampling lengths from 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. This basic method has been extended by length-based spawning potential ratio (LBSPR, Hordyk et al. 2015a), length-based Integrated Mixed Effects (LIME, Rudd and Thorson 2017) and Length-Based Bayesian approach (LBB, Froese et al. 2018) models, among others. These allow for the estimation of instantaneous fishing mortality (F) and spawning potential ratio (SPR) when basic biological parameters are known. In contrast to LBSPR and LBB, LIME does not assume equilibrium conditions. The mixed-effects aspect of LIME extends length-based methods by estimating changes in recruitment and separating them from fishing mortality over time (Rudd and Thorson 2018).

It is good practice to simulation test the performance of assessment methods before applying them in practice (Cope 2008). This can be done using a variety of approaches, though it is most often accomplished using an Operating Model (OM) to generating pseudo-data with error to fit an assessment model (Punt et al. 2016). Simulation can either be an open loop or a closed loop with feedback. Carruthers (2016), using closed loop simulations, 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 (Carruthers et al. 2014). In many data poor fisheries, measuring total removals is difficult, as is enforcing catch limits. Recently, 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. They found that the LBSPR assessment model with an iterative effort-based harvest control rule can be used to rebuild an overfished stock to sustainable levels or fish down a stock to the target SPR. What is needed however, is a comparison of the performance of both length-based and catch-based methods to estimate stock status. Unfortunately, finding a common metric between catch-based and length-based stock status metrics is difficult; the former measures overfishing and the latter stock depletion.

Even though in many countries they are moving to close-loop simulations or Management Strategy Evaluation (MSE) to assess the performance of different management procedures for data-limited fisheries, managers in most of the developing countries are still demanding information on stock status in order to manage their fisheries. Furthermore, Dowling et al. (2019) noted the dangers in the indiscriminate use of generic methods and frameworks and emphasized the importance of using care in acknowledging and interpreting uncertainties. Therefore, this study used open-loop simulations to better estimate relative bias and precision of a range of data-limited methods. A key challenge is to maintain a balance between the opposing risks of inappropriate management "action" due to assessment inaccuracy and inappropriate management "inaction" due to assessment uncertainty.

We used OMs to represent the main sources of uncertainty and to generate data for use in data-limited stock assessment methods and to evaluate how well the methods perform when the data and knowledge requirements are met. A common metric is then used for comparison across models, namely exploitation or harvest rates, to evaluate the performance of catch-based and length-based models in a simulation context. We evaluate performance considering three fish stocks with contrasting life-history strategies, under three different harvest trends, and three different levels of final stock depletion.

METHODS

Twenty-seven different OMs were created using a factorial design comprising 3 harvest rates, 3 life-history types, and 3 depletion scenarios. The different harvest rates scenarios, each considered a 20-year time series of fishing, correspond to fishing mortality histories commonly seen in many fisheries. In harvest rate scenario 1, fishing mortality increases until it reaches a maximum and starts declining afterwards. This is commonly seen once management measures are implemented to reduce fishing pressure. Harvest rate scenario 2 assumes that fishing mortality increases and remains constant after reaching a maximum. This harvest rate profile could result from the implementation of catch or effort management limits. Harvest rate scenario 3 has constantly increasing fishing mortality, which would occur for fisheries that are still developing (Figure 1a).

Three life history types of varying longevity and somatic growth rate were simulated, namely (i) a short-lived fast-growing species, Pacific chub mackerel (*Scomber japonicus)*, (ii) a medium-lived medium-growing fish, albacore tuna (*Thunnus alalunga*, and (iii) a longer-lived slow-growing species, canary rockfish (*Sebastes pinniger)* (Table 1). Finally, the following three depletion levels were considered: (i) heavily fished (depletion = 0.2), (ii) sustainably fished (depletion = 0.4) and (iii) slightly fished (depletion = 0.6).

Operating model specifications

The OM was developed to simulate resource dynamics under the different fishing scenarios, life histories and final depletion levels. The OMs consist of an age-structured population with numbers at age over time modelled as follows:

$$N_{a,t} = \begin{cases} R_t, & a = 0 \text{ and } t = 0 \\ N_{a-1,t}\, e^{(-M - F_t S_{a-1})}, & 0 < a < A \text{ and } t = 1 \\ \dfrac{N_{a-1,t}\, e^{(-M - F_t S_{a-1})}}{1 - e^{(-M - F_t S_{a-1})}}, & a = A \text{ and } t = 1 \\ N_{a-1,t-1}\, e^{(-M - F_t S_{a-1})}, & 0 < a < A \text{ and } t > 1 \\ (N_{a-1,t-1} + N_{a,t-1})\, e^{(-M - F_{t-1} S_{a-1})}, & a = A \text{ and } t > 1 \end{cases}$$

R_t is the number of age-0 animals at the start of year t, $N_{a,t}$ is the number of fish of age a at the start of the year t, S_a is the selectivity at age, F_t is the instantaneous fishing mortality rate during year t, M is the instantaneous rate of natural mortality, and A is the age of the plus group. Fishing mortality deviations were included as $F_t \sim lognormal\,(F_{t-1}, \sigma_F^2)$. A Beverton–Holt spawner–recruitment function (Beverton and Holt 1957) and annual normally distributed recruitment deviations $N(0, \sigma_R)$ were assumed (Table 1).

The biomass in each year t was calculated as $B_t = \sum_{a=1}^{A} N_{a,t} w_a$ where $w_a = \alpha L_a^{\beta}$ (parameters in Table 1). In addition, the predicted total catch by year (C_t) was calculated as $C_t = \sum_{a=0}^{A} C_{a,t}$ with:

$$C_{a,t} = \frac{F_t S_a}{M + F_t S_a} N_{a,t} \left(1 - e^{(-M - F_t S_a)}\right)$$

For each OM (N=27), 100 time series of fishing mortality were simulated, and the harvest rate per year (U_t) as C_t/B_t calculated (Figures S1 to S3). Each simulated population began at the unfished biomass level and all fishing trend scenarios terminate at the specified depletion level (Appendix Figures A1 to A27).

To condition the OM, published life-history parameter values (Table 1) reported in formal stock assessments were used (Crone and Hill 2015 for the short-lived Pacific chub mackerel; ICCAT 2014 for the medium-lived albacore tuna; and Thorson and Wetzel 2015 for long-lived canary rockfish). Each population was assumed to be targeted in a single area, by one fleet with a selectivity pattern (Table 1) that was logistic and constant through time.

Length bins were defined as they were in their respective assessments; every 2 cm (Crone and Hill 2015; ICCAT 2014; Thorson and Wetzel 2015). To obtain the catch length frequency, the probability (p) of being in a length bin (j) given age (a) was calculated as:

$$p_{j,a} = \begin{cases} \emptyset\left(\dfrac{j - L_a}{L_a CV_L}\right), & j = 1 \\ \emptyset\left(\dfrac{j - L_a}{L_a CV_L}\right) - \emptyset\left(\dfrac{j - 1 - L_a}{L_a CV_L}\right), & 1 < j < J \\ 1 - \emptyset\left(\dfrac{j - 1 - L_a}{L_a CV_L}\right), & j = J \end{cases}$$

With the predicted probability of harvest by length bin being:

$$\pi_j = p_{j,a}\frac{\sum_{a=0}^{A} N_{a,t} S_a}{N_t}$$

One thousand fish per year were drawn using a multinomial distribution with a π_j probability (Rudd and Thorson 2018).

Comparing methods outputs

One of the challenges when comparing catch-based and length-based methods is that they produce different model outputs. Catch-only models calculate total and/or spawning stock biomass and sustainable catches, whereas length-based models estimate exploitation and transient SPR, which can be used to infer stock status. Multiple studies have shown that catch-based methods might be appropriate to predict sustainable catch at the end of the time series, but not to reconstruct a biomass time series (Carruthers et al. 2012; Wetzel and Punt 2015). These are fundamentally different measures of the population status. As a result, our performance metric is defined as the error relative (RE) to the OM, where RE = (U_{Method} - U_{OM}) / U_{OM}. This allows for a measure of uncertainty, in both bias and precision, for all 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 OM. Imprecision is related to the variability (variance) around the central tendency.

We used U as a common measure for comparisons between each data limited method and the OM. For catch-only approaches it is defined as the ratio catch/biomass; while for the length-based models, the estimated F was transformed to an exploitation rate via $U = 1 - exp(- F)$. In addition, we present the average RE across the last five years of the time series, not along the entire time series of data, because we are interested in the estimation of the current exploitation rates.

Estimation models

All simulations, and data-limited model calculations, were conducted using the open-source statistical software R (R Core Team 2018). Each catch-based and length-based method evaluated here are summarized below.

Catch-based data-limited methods

Catch-MSY (Martell and Froese 2013) is a SRA approach that assume a Schaefer biomass dynamic model. Inputs are a time series of removals, priors for the population rate of increase at low population size (r), carrying capacity (K), and a range of stock depletion in the final year (Table 1). Values of r and K are randomly sampled using 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). 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. Here we use the modified version of CMSY (Rosenberg et al. 2017) to extract biomass trends from all viable *r*-*K* pairs using the R package *datalimited* version 0.1.0 (Anderson et al. 2016). The biomass trajectory is calculated as the median of all viable biomass trajectories generated under the Monte Carlo process.

CMSY (Froese et al. 2017) extends Catch-MSY by using a Monte-Carlo filter (instead of the SIR algorithm) that fixes systematic biases in the Catch-MSY method. It also, explicitly incorporates process error and estimates target reference points (MSY, F_{MSY}, B_{MSY}) as well as relative stock size (B/B_{MSY}) and exploitation (F/F_{MSY}) from catch data and priors for *r* and depletion at the beginning and the end of the time series. CMSY has an inbuilt piecewise "hockey-stick" to prevent over-estimating of rebuilding potential at very low abundance $B<0.25B_0$. The CMSY package implements a Bayesian state-space implementation of the Schaefer surplus production model (Winker 2019).

SSCOM 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 trends in fishing mortality follow 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. The package *datalimited* version 0.1.0 (Anderson et al. 2016) was used and the code was extended 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). Using this modified version of SSCOM, the required inputs are priors for *r*, *K*, and final stock depletion (Table 1).

DBSRA (Dick and MacCall 2011) modifies the SRA approach by using Monte Carlo draws from four parameter distributions (M, F_{MSY}/M, B_{MSY}/B_0, and *depletion*) and age at maturity (A_{mat}) to separate the total 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:

$$B_{t+1} = B_t + P(B_{t-Amat}) - C_t$$

where B_t is the biomass at the start of year t, $P(B_{t-Amat})$ is the latent annual production based on a function of adult biomass in year t-$Amat$ and C_t is the catch in year t. Biomass in the first year (B_0) is assumed to be equal to K. The package *fishmethods* version 1.10-3 was used to perform this analysis (Nelson 2017). For DBSRA we used the $Amat$ and M as fixed inputs and three priors: final stock depletion, F_{MSY}/M, and B_{MSY}/B_0 (distributions in Table 1). Each of these is assigned a distribution from which the Monte Carlo draws are taken.

SSS is based on the Stock Synthesis age-structured stock assessment model (Methot and Wetzel 2013). SSS fix all parameters in the Stock Synthesis model except for initial recruitment ($\ln R_0$). It also sets up an artificial index of abundance that represents the relative stock biomass. The first value of the index is always 1, and the value in the final year represents the percent of the population left in that year. The values of steepness (h) and the final year of the abundance survey are all randomly drawn from a specified distribution using a Monte Carlo approach (Cope 2013) and $\ln R_0$ is then estimated. Benefits of this approach are that it retains the same modelling framework as the data-rich stock assessments, but still allows for flexibility in a variety of parameter and model specifications, if desired. The input priors used for SSS were relative stock status and steepness and selectivity was matched to the OM (Cope 2019).

Length-based data-limited methods

SPR is the proportion of the unfished reproductive potential per recruit under a given level of fishing pressure (Goodyear 1993). In **LBSPR,** SPR in an exploited population is calculated as a function of the ratio of fishing mortality to natural mortality (*F/M*), selectivity, and the two life-history ratios *M/k* and L_m/L_∞; *k* is the von Bertalanffy growth coefficient, L_m is the size of maturity and L_∞ is asymptotic size (Hordyk et al. 2015a). The inputs of LBSPR are *M/k*, L_∞, the variability of length-at-age (CVL_∞), which is normally assumed to be around 10%, and the length at maturity specified in terms of L_{50} and L_{95} (the size at which 50% and 95% of a population matures). Given the assumed values for *M/k* and L_∞ and that length composition data come from an exploited stock, the LBSPR model uses maximum likelihood methods to estimate the selectivity ogive, which is assumed to be of a logistic form defined by the selectivity-at-length parameters S_{50} and S_{95} (the size at which 50% and 95% of a population is retained by the fishing gear), and *F/M*. The selectivity ogive and relative fishing mortality are then used to calculate SPR (Hordyk et al. 2015a, 2015b). Estimates of SPR are primarily determined by the length of fish relative to L_{50} and L_∞, but it also depends on life history parameters such as fecundity-at-age/length and selectivity. 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 the LBSPR package, version 0.1.3 in R (Hordyk 2018). We used the Rauch-Tung-Striebel smoother function available in the LBSPR package to smooth out the multi-year estimates of *F*.

LIME uses length composition of the catch 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 separately 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). The inputs to LIME (Rudd 2019) are: M, k, L_∞, t_0, CVL_∞, L_{50}, L_{95}, h, and the parameters of the length-weight relationship a and b (Table 1).

LBB is a simple and fast method for estimating relative stock size that uses a Bayesian Monte Carlo Markov Chain (MCMC) approach (Froese et al. 2018). In contrast to other length-based methods, LBB uses pre-specified priors on parameters, and thus, technically does not require further inputs in addition to length frequency data, if the user is willing to accept the life history defaults. However, it provides the user the option to specify priors for the inputs L_∞, length at first capture (Lc), and relative natural mortality (M/k). We specified the true M/k value and we let the model calculate Lc (length where 50% of the individuals are retained by the gear) and L_∞, which is approximated by the maximum observed length L_{max}. In addition, F/M is estimated as means over the age range represented in the length-frequency sample.

During our simulation testing, we assumed that length models had the correct values for the von Bertalanffy length-at-age relationship, L_∞, k and t_0, length-weight parameters α and β, M, and the parameters L_{50} and L_{95} from a logistic maturity-at-length curve. We did not evaluated misspecifications in life history parameters inputs. The objective is only to compare the performance of different data-limited methods under different scenarios, not to evaluate each of them under different parameters misspecifications, which has already been done in the original publications of each method.

RESULTS

Large variability is seen in model performance across harvest scenarios, life-history types, and depletion levels (Figure 1b-d, Figure 2). Figures 1 and 2 can be used to search for the best model, i.e. to identify the least biased and least imprecise method based on the life history of the species, harvest trends, and knowledge of final depletion. The imprecision of each method is shown in the variability around those estimations in Figure 1 and the specific values in Figure 2. A robust method would show low bias and high precision for all stocks and harvest scenarios; for example, among all catch-based methods, SSS appears to be the most and SSCOM the least robust method (Figure 2).

In general, catch-based methods tended to be more biased at relatively higher stock sizes (D=0.6) and for long-lived species (Figures 1 and 2). Catch-MSY and SSCOM tended to perform poorly across a range of scenarios. Among the length-based models, all were less biased and more precise for the medium-lived species (Figure 2) without a clear different performance among harvest scenarios. The most biased length-based method was LBSPR although lesser imprecise than LIME and LBB (Figure 2). Overall, LIME and SSS were the most robust methods.

Catch-only methods

It was to be expected that the various catch-based models considered in this study would perform differently among them because they have different model structure and assumptions. SSS performed best in most cases estimating unbiased exploitation rates across different scenarios of harvest trends, final stock depletion and life histories. However, it tended to underestimate harvest rates by 30% for medium and short-lived species at relatively high stock sizes (D=0.6). For long-lived species the estimations were slightly overestimated for sustainable relative stock sizes (D=0.4), but SSS was always the model that was less biased. DBSRA was the most precise but it underestimated harvest rates in general, and in contrast to the other catch-based methods, it was less biased for stocks at relatively high stock sizes (Figure 2). Catch-MSY was the most biased of the catch-based models tested, overestimating harvest rates in particular for stocks at relatively high stock sizes (D=0.6). It was less biased for medium-lived species and produced non-biased estimates of MSY for highly depleted (D=0.2) medium-lived or short-lived species at sustainable relative stock sizes (D=0.4) (Table A1). SSCOM was less biased than Catch-MSY in most scenarios, but less precise than any other catch-based model, showing the broadest range of RE, particularly when stocks were at relatively high stock sizes (D=0.6) or where catch decreased at the end of the time series (i.e. harvest scenario 1). SSCOM was also highly and positively biased for harvest scenario 1 (Figure 1 and 2). SSCOM showed multiple modes across scenarios, suggesting that the estimates are unstable. CMSY was less biased in the relatively medium to high stock sizes (D=0.4 and D=0.6) and for medium-lived species. In general, catch-based models were less biased and more precise when stocks were at relatively low stock sizes (i.e. using a prior centered around 0.2, Figure 1). In general, estimation models in harvest scenario 1 (ramp) produced the most variable estimations in harvest rates (Figure 1). The catch-based methods performed the best for the medium-lived species.

Length-based methods

In many cases, length-based models gave a less biased estimation of *U* than catch-based models (Figure 1). LIME was the least biased length-based method. However, in general, LIME did not converge around 10% of the time. The three length-based methods used here (LIME, LBSPR and LBB) produced more variable estimations in harvest scenario 1, where the fishing intensity decreased at the end of the time series (Figure 1) but no clear pattern was observed in terms of harvest scenarios in bias (Figure 2). LBSPR in general underestimated harvest rates. Compared to itself, LBB performed better in scenarios where the stocks have relatively low to medium stock sizes (D=0.2 and D=0.4) and for medium-lived species (Figure 2). Both, LBB and LIME were highly variable for long-lived species and harvest scenario 1 (Figure 1). Multiple modes for harvest scenario 1 may suggest poor convergence. In general, length-based models were less biased for the medium-lived species.

Short-lived species

For the short-lived life-history strategy, SSS was the least biased and most precise among the catch-based methods. The second most precise method was CMSY, but this method was positively biased for relatively high stock sizes (D=0.6). Among the three length-based models, LIME had better overall performance than LBSPR and LBB. However, LBSPR was less biased in harvest scenario 1. LBB was the one model that presented the most variability in harvest-rate estimations for the short-lived species.

Medium-lived species

Both catch-based and length-based models gained more precision as the stocks were more depleted. Among the catch-based methods, CMSY was the less biased followed by SSS, although SSS was more precise. DBSRA was highly negatively biased particularly for relatively low stock sizes (D=0.2). Catch-MSY was highly positively biased but less so for highly depleted stocks (Figure 1c). Between the length-based models, LIME showed very good performance with regards to bias for harvest scenarios 2 and 3 under the three depletion levels, although it was slightly biased for harvest scenario 1. LBSPR and LBB showed similar performance but LBSPR was more precise (Figure 1c).

Long-lived species

Both catch-based and length-based methods were less precise (more variability in RE) as the long-lived stocks had relatively higher stock sizes (D=0.6), as they did for the other life-history strategies. Among the catch-based methods, SSS was the most precise and least biased method except in harvest scenario 1 and D=0.2 where CMSY was the least biased. However, SSS underestimated harvest rates in scenarios 2 and 3, where the catch history is constant or increases at the end of the time series. Among assessment methods, a less biased and more precise estimation was observed for sustainable depleted stocks (D=0.4) than for other depletion levels. SSCOM was less biased in harvest scenario 3 and D=0.4 (Figure 2) but highly imprecise (Figure 1). LBSPR and DBSRA was negatively biased in all cases (Figure 1d). LIME was highly imprecise in harvest scenario 1, but the least biased among the other length-based methods (Figure 1c, Figure 2).

Discussion

Simulation studies commonly make different operating model assumptions from those of the methods being tested to allow the evaluation of robustness; in some cases, however, the same population model is used for both simulation and estimation, i.e. self-testing (Deroba et al., 2015). Using the same model structure for simulation and estimation can result in more optimistic results that might not be true under many scenarios (Francis, 2012). Since it is not possible to explore the impact of model assumptions when the model used for simulation and estimation is the same. If a method performs poorly, however, when the assumptions in the OM are the same as the assessment model, it is unlikely to perform well in practice. To evaluate robustness to model structure our approach evaluated multiple data-limited assessment methods with a range of assumptions about population and fishery dynamics using an operating model decoupled from the tested methods.

It is to be expected that the various methods would perform differently. Rosenberg et al. (2017) used four catch-based data-limited models and found that models frequently disagreed about population status estimations, with no model showing overall good performance, i.e. high precision and low bias across all case studies. When scenarios represent specific resource dynamics or particular stocks or fisheries, it may be difficult to draw any overall conclusions. Therefore, we chose scenarios that represented different fishing intensity trends, depletion levels, and life histories. We found that model performance is highly dependent on all these factors. More imprecision was found, where fishing pressure decreases at the end of the time series (harvest scenario 1), in comparison with harvest scenarios where F was either stable or increasing (harvest scenarios 2 and 3) where both performed similarly in terms of bias and imprecision. In addition, most bias was found for scenarios when stock abundance is high and/or for slow-grow long-lived species.

In particular, catch-based models performed better (i.e. were less biased and more precise) for stocks that were medium to highly depleted than for slightly depleted stocks. Walters et al. (2006) suggested that for SRA, stocks that have experienced extensive historical depletion gain precision due to a high rate of rejected parameter draws. Moreover, SSS which is an age structure model, performed better than the models that are based on surplus production functions such as Catch-MSY, CMSY, DBSRA, and SSCOM, even when priors for depletion were centered on the true values for all methods. Although SSS seems to be the least biased catch-based model, unlike other catch-based models, more detailed life-history information (e.g., age and growth estimates) are required by SSS to define age structure and remove catch according to age-/size-based selectivity patterns (Cope, 2013).

High imprecision was observed in the estimates from SSCOM. To increase precision, additional and different priors could be specified as well as trying different effort-dynamics. In this study we used, as a default, the effort dynamics specified in Anderson et al. (2017). Other effort dynamics could be assumed. Anderson et al. (2017) found that SSCOM was the most imprecise and that the representation of effort dynamics was more suitable at low biomasses. In addition, SSCOM might be more appropriate for stocks with longer time series of catch. Pons (2018) found better performance of SSCOM for the same species using longer time series (~ 80 years of catch data).

Catch-MSY performed poorly in all scenarios, overestimating harvest rates even when given a prior for depletion close to the true value. A key point of Catch-MSY 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 100 times the maximum catch as the upper bound for K based on Martell and Froese (2013). Other K values could be explored to see if this improves the outcome, but it remains a difficult parameter to specify. Rosenberg et al. (2017) and Free et al. (2017) found that Catch-MSY was the model that performed second best and better than SSCOM in their scenarios. One of the differences between their study and with ours is that they considered a uniform prior for depletion in SSCOM and we considered a lognormal prior centered on the true value.

CMSY on the other hand performed particularly well with respect to bias and precision for medium-lived species, even better than SSS for medium to low depleted stocks (Figure 2). Also, CMSY was more accurate than the original Catch-MSY method (Martell and Froese, 2013). The difference is that Catch-MSY was designed to select the most probable r-K pair as the geometric mean of this distribution, but CMSY searches not in the center of the distribution but rather near the right tip of viable pairs. According to Froese et al. (2017) since r is defined as the maximum net productivity, the right tip of the distribution of *r*-*K* pairs is where these parameters should be found. So, between Catch-MSY and CMSY, CMSY is preferred.

Hordyk et al. (2015a) explained how LBSPR relies on detecting the signal of fishing mortality in the right-hand side of the length composition. Consequently, fishing is not likely to have a visible impact on the length composition until fishing mortality is very high and stocks are highly depleted. This is why LBSPR was less biased for more depleted populations and in fishing scenario 1.

Our study found that LIME was highly imprecise for long-lived species. Rudd and Thorson (2017) also showed that LIME is more imprecise for long lived species. The model is trying to track cohorts through the length data to estimate recruitment deviations and this is likely difficult for long-lived species when time series of length data are short or much of the population is found near the asymptotic size (Rudd and Thorson 2017).

Hordyk et al. (2019) suggested that LBB has not been sufficiently simulated tested and it can produce biased estimates of fishing mortality. We found that LBB was the most biased and imprecise length-based method although for the less depleted stocks it generally performed better than LBSPR. One of Hordyk et al. (2019) critics to LBB is that it assumes that M/k = 1.5, however here, we specified the true M/k value. So, the main bias was associated to the estimations of L_∞ due to the approximation used to the maximum observed length L_{max}. In addition to L_∞, L_c was always overestimated by the LBB assessment model (see Table A2).

In general, all catch-based and length-based methods seem to perform worse for long-lived life-history types, where there is likely to be less contrast in the dynamics over time, than for medium and short-lived. The length of the time series for the long-lived canary rockfish is therefore probably too short in comparison to the age they reach the maximum length (64 years), to capture the true dynamics of the population and the response to different harvest rates.

The present study did not look at parameter misspecification but correctly specified (unbiased) the life-history parameters and catch histories. As we mentioned before, parameters misspecification testing was performed in the original publications for each method. With accurate prior information, length-based models such as LIME showed better performance in many cases than some catch-based models, as the latter were more sensitive to the catch history scenarios and depletion levels. LIME was not sensitive to catch trends because the model integrates the catch scenarios into the length compositions, but in most cases, LIME was more imprecise than other length-based assessments. In addition, LIME has lack of convergence in many runs compared to LBSPR and LBB.

Bias and precision are both important factors to consider when assessing fish stocks, bias reflects how close an estimate is to an accepted value and precision reflects reproducibility of the estimate. For example, if an assessment is to be re-conducted every year to monitor the impact of a management measure, a precise but biased method would be able to detect a trend better than an unbiased but imprecise method. As with scientific instruments this trade-off require calibration, which in the case of fish stock assessment can be performed using MSE, where the choice of parameters and reference points in a management procedure are tuned, i.e. calibrated, to meet the desired management objectives as represented by the OM. Therefore, a biased method (e.g., DBSRA) may be preferable to one that is less biased, but more imprecise (e.g., LIME). Alternatively, imprecision can be addressed through the choice of the percentile (e.g., median being the 50% percentile value) for the derived model output used by management (e.g., catch or SPR); assuming that the true value is contained within the parameter distribution. For example, instead of taking the median value, one could instead use the derived model output associated with the 40th

percentile. Such an approach (Ralston et al. 2011) is used in fisheries management systems to directly incorporate scientific uncertainty (both bias and imprecision), and can also be tuned using MSE.

Recommendations

To provide estimates of stock status for unassessed fisheries where data are limited, but reconstructing time series of catch is possible, SSS is recommended. The performance of SSS hinges on the correct specification of the input parameters such as stock depletion, productivity, maturity, and growth parameters. Knowledge about these parameters is likely to be poor for data-limited stocks, resulting in misspecification of these parameters. Meta-analyses may offer some starting values for certain parameters (e.g., Myers 2001, Thorson *et al.* 2012, and Zhou *et al.* 2012, 2017), however other inputs remain difficult to specify (Chrysafi and Cope *in review*).

For fisheries where the time series of catch are unavailable, using length-composition data can provide good approximations of the status of the stock, in particular for medium-lived species. It has been shown here that in some cases, length-based models such as LIME can provide the same or less biased estimates of exploitation status than catch-based models. However, growth parameters are even more important for length-based than for any catch-based method, so it is important to have good estimates of those parameters before using any length-based assessment.

Making recommendations on which models should be applied to estimate exploitation intensity in different fisheries is challenging because model choice is dependent on data availability, trends in fishing intensity, and the biology of the species. If possible, simulation studies testing different data-limited methods with OMs based on the focus species and the dynamics of the fishery can greatly inform which method is most appropriate. Likewise, decision support tools such as FishPath (Dowling et al. 2015) can also help one weigh the input requirements and assumptions to identify the most appropriate methods given data and life history. Based on the OMs used in this study, we conclude that when only catch data is available, SSS should be considered. When only length data is available, LIME appeared to be less biased being able to capture changes in recruitment and fishing mortality better than LBSPR and LBB. However, Pons et al. (2019) found that neither LBSPR or LIME are good in all situations, and thus, both should be considered and compared. LIME sometimes undergoes convergence issues and has difficulties separating changes in recruitment from changes in fishing mortality (Pons 2018; Pons et al. 2019).

For long-lived species it is necessary to have longer time series of data to draw more conclusions. However, Pons (2018) recommended SSS and LBSPR when long time series (i.e. 80 years) of data are available for a species that lives more than 60 years to evaluate changes in fishing intensity.

If both catch and length data are available, models that integrate both data types should be considered. LIME, although primarily length based, allows for the inclusion of catch data as well as an index of abundance if one is available. Moreover, integrated assessment models (that use catch as well as length information) like Stock Synthesis could also be considered (Methot and Wetzel 2013). Length information can therefore be added to the SSS data file, with the possibility of freeing up the stock status assumption input, and running the model more like a traditional statistical-catch-at-age model (Cope 2013).

For the scenarios analyzed here, including the specific life-histories considered, we do not recommend Catch-MSY for estimating exploitation rates, even with a good estimate of stock depletion. This method will however, as it was originally created to do, produce unbiased estimates of MSY, in particular for short and medium-lived and highly depleted species (Table A1).

Future directions

Dowling et al. (2019) in a review of data limited methods, noted the dangers in the indiscriminate use of generic methods and recommended obtaining better data, using care in acknowledging and interpreting uncertainties, developing harvest strategies that are robust to these higher levels of uncertainty and tailoring them to the species and fisheries specific data and context. Therefore, methods should be tested using a management strategy evaluation (MSE) to specify Management Procedures (MP) 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. This must include evaluation of the robustness of the methods to misspecification of input parameters and the benefits of improving knowledge on them.

This study provides a way of conditioning OMs and generating pseudo data for use by the MP. 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 an unbiased but imprecise estimator. Additionally, if a method only provides estimates of exploitation level or MSY, then management controls may be different i.e. have to be based on a total allowable catch or effort. Traditional stock assessment and advice based upon it mainly considers measurement and process error only. MSE allows for the consideration of additional uncertainty, such as uncertainty in the actual dynamics, which has a larger impact on achieving management objectives (Punt 2008).

Acknowledgments

MP was funded by the Walton Family Foundation (#2018-51) and International Seafood Sustainability Foundation (ISSF). LK’s involvement was funded through the MyDas project under the Marine Biodiversity Scheme which is financed by the Irish government and the European Maritime & Fisheries Fund (EMFF) as part of the EMFF Operational Programme for 2014-2020. We want to acknowledge Trevor Branch, Victor Restrepo, Dave Fluharty, Jim Ianelli and Ray Hilborn for their significant comments in the early stages of this study. We also want to acknowledge Merrill Rudd, Chantel Wetzel, Ian Taylor, and James Thorson for their help with SS, r4ss, and other packages and models used in this study and Nicole Baker and Kelli Johnson for proofreading the final version of the manuscript.

References

Anderson, S.C., Afflerbach, Jamie Cooper, A.B., Dickey-Collas, M., Jensen, O.P., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Minte-Vera, C., Minto, C., Mosqueira, I., Rosenberg, A.A., Selig, E., Thorson, J.T., and Walsh, J.C. 2016. datalimited: Stock Assessment Methods for Data-limited Fisheries. R package version 0.1.0. Available from https://github.com/datalimited/datalimited.

Anderson, S.C., Cooper, A.B., Jensen, O.P., Minto, C., Thorson, J.T., Walsh, J.C., Afflerbach, J., Dickey-Collas, M., Kleisner, K.M., Longo, C., Osio, G.C., Ovando, D., Mosqueira, I., Rosenberg, A.A., and Selig, E.R. 2017. Improving estimates of population status and trend with superensemble models. Fish Fish. **18**(4): 732–741. doi:10.1111/faf.12200.

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.

Carruthers, T.R., Walters, C.J., and McAllister, M.K. 2012. Evaluating methods that classify fisheries stock status using only fisheries catch data. Fish. Res. **119**: 66–79. doi:10.1016/j.fishres.2011.12.011.

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. Fish. Res. **153**: 48–68. Elsevier B.V. doi: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 J. Mar. Sci. **73**(2): 464–482. doi: 10.1093/icesjms/fsv212.

Chrysafi, A., and Kuparinen, A. 2016. Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. Environ. Rev. **1**: 1–44. doi:10.1017/CBO9781107415324.004.

Chrysafi, A., and J.M. Cope. in review. Testing methods of determining relative stock abundance priors when setting catch recommendations using data-limited approaches. Fish. Res. under review.

Cope, J.M. 2008. Issues and Advances in Data-Limited Stock Assessment: Experimentation through Simulation. Diss. - Dr. Philos. 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. Fish. Res. **142**: 3–14. Elsevier B.V. doi:10.1016/j.fishres.2012.03.006.

Cope, J. M. 2019. Simple Stock Synthesis code and examples. Available at https://github.com/shcaba/SSS.

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 **338**(6106): 517–20. doi:10.1126/science.1223389.

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

Deroba, J.J., Butterworth, D.S., Methot Jr, R.D., De Oliveira, J.A.A., Fernandez, C., Nielsen, A., Cadrin, S.X., Dickey-Collas, M., Legault, C.M., Ianelli, J. and Valero, J.L., 2014. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES J. Mar. Sci. **72**(1): 19–30. doi: 10.1093/icesjms/fst237.

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. Fish. Res. **110**(2): 331–341. Elsevier B.V. doi: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. Fish. Res. **171**: 141–153. Elsevier B.V. doi:10.1016/j.fishres.2014.11.005.

Dowling, N., Wilson, J., Rudd, M., Babcock, E., Caillaux, M., Cope, J., Dougherty, D., Fujita, R., Gedamke, T., Gleason, M., Guttierrez, M., Hordyk, A., Maina, G., Mous, P., Ovando, D., Parma, A., Prince, J., Revenga, C., Rude, J., Szuwalski, C., Valencia, S. and Victor, S., 2016. FishPath: A Decision Support System for Assessing and Managing Data- and Capacity- Limited Fisheries, in: Quinn II, T., Armstrong, J., Baker, M., Heifetz, J., Witherell, D. (Eds.), Assessing and Managing Data-Limited Fish Stocks. Alaska Sea Grant, University of Alaska Fairbansk. https://doi.org/10.4027/amdlfs.2016.03

Dowling, N. A., Smith, A. D. M., Smith, D. C., Parma, A. M., Dichmont, C. M., Sainsbury K., Wilson, J. R., Dougherty, D. T. and Cope, J. M. 2019. Generic solutions for data-limited fishery assessments are not so simple. Fish Fish. **20**(1): 174–188. doi: 10.1111/faf.12329.

Francis, C. 2012. Reply to "The reliability of estimates of natural mortality from stock assessment models." Fish. Res. **119–120**: 133–134. Elsevier B.V.

doi:10.1016/j.fishres.2012.02.015.

Free, C.M, Jensen, O.P., Wiedenmann, J., and Deroba, J.J. 2017. The refined ORCS approach: A catch-based method for estimating stock status and catch limits for data-poor fish stocks. Fish. Res. **193**: 60–70. doi: 10.1016/j.fishres.2017.03.017

Froese, R., Demirel, N., Coro, G., Kleisner, K.M., and Winker, H. 2017. Estimating fisheries reference points from catch and resilience. Fish Fish. **18**(3): 506–526. doi:10.1111/faf.12190.

Froese, R., Winker, H., Coro, G., Demirel, N., et al. 2018. A new approach for estimating stock status from length frequency data, ICES J. Mar. Sci. **75**(6). 2004–2015. doi: 10.1093/icesjms/fsy078.

Goodyear, C.P. 1993. Spawning stock biomass per recruit in fisheries management: foundation and current use. In Risk evaluation and biological reference points for fisheries management. Edited by S.J. Smith, J.J. Hunt, and D. Rivard. Can. Spec. Publ. Fish. Aquat. Sci. No. 120.

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**(1): 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. Fish. Res. **171**: 20–32. Elsevier B.V. doi:10.1016/j.fishres.2014.12.018.

Hordyk A. R. 2018. LBSPR: Length-Based Spawning Potential Ratio. R package version 0.1.3. https://github.com/AdrianHordyk/LBSPR.

Hordyk, A. R., Prince, J. d., Carruthers, T. R. and Walters, C. J. 2019. Comment on “A new approach for estimating stock status from length frequency data” by Froese et al. (2018).

ICES J. Mar. Sci. **76**(2): 457–460. doi: 10.1093/icesjms/fsy168.

ICCAT. 2014. Report of the 2013 ICCAT North and South Atlantic Albacore stock assessment meeting. Collect. Vol. Sci. Pap. ICCAT. **70**(3): 830–995.

Kimura, D.K., and Tagart, J.V. 1982. Stock reduction analysis, another solution to the catch equations. Can. J. Fish. Aquat. Sci. **39**: 1467–1472. doi: 10.1139/f82-198.

Kimura, D.K., Balsiger, J.W., and Ito, D.H. 1984. Generalized stock reduction analysis. Can. J. Fish. Aquat. Sci. **41**: 1325–1333. doi: 10.1139/f84-162.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell B.M. 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Software. **70**(5), 1-21. doi:10.18637/jss.v070.i05.

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

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. **142**: 86–99. doi:10.1016/j.fishres.2012.10.012.

Myers, R.A., 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES J. Mar. Sci. **58**(5): 937–951 doi:10.1006/jmsc.2001.1109.

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.

Pons, M. 2018. Stock Status and Management in Tuna Fisheries: From Data-rich to Data-poor. University of Washington, ProQuest Dissertations Publishing, 2018: 10846859. 186 pp. https://search.proquest.com/docview/2126689812? accountid¼14784 (last accessed 15 November 2018).

Pons, M., Kell, L., Rudd, M. B., Cope, J. M., and Lucena Fredou, F. 2019. Performance of length-based data-limited methods in a multifleet context: application to small tunas, mackerels, and bonitos in the Atlantic Ocean. ICES J. Mar. Sci. **76**(4): 960–973, doi:10.1093/icesjms/fsz004.

Punt, A. 2008. Refocusing stock assessment in support of policy evaluation. Fish. Glob. Welf. Environ. 5th World Fish. Congr. (January 2008): 139–152. Available from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Refocusing+Stock+Assessment+in+Support+of+Policy+Evaluation#0%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Refocusing+stock+assessment+in+support+of+policy+evaluation#0.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A., Haddon, M., 2016. Management strategy evaluation: best practices. Fish Fish. **17**, 303–334. doi: 10.1111/faf.12104.

R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org.

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J.D., Conser, R.J. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fish. Bull. **109**(2): 217–232.

Rosenberg, A.A., Kleisner, K.M., Afflerbach, J., Anderson, S.C., Dickey-Collas, M., Cooper, A.B., Fogarty, M.J., Fulton, E.A., Gutiérrez, N.L., Hyde, K.J.W., Jardim, E., Jensen, O.P., Kristiansen, T., Longo, C., Minte-Vera, C. V., Minto, C., Mosqueira, I., Osio, G.C., Ovando, D., Selig, E.R., Thorson, J.T., Walsh, J.C., and Ye, Y. 2017. Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries Around the World. Conserv. Lett. **00**(April): 1–9. doi: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. Can. J. Fish.

Aquat. Sci. **75**(7): 1019–1035. doi:10.1139/cjfas-2017-0143.

Rudd, M. B. 2019. Length-based integrated mixed effects model code. Available at: https://github.com/merrillrudd/LIME.

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. Can. J. Fish. Aquat. Sci. **69**(9): 1556–1568. doi: 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. Can. J. Fish. Aquat. Sci. **70**(12): 1829–1844. doi:10.1139/cjfas-2013-0280.

Thorson, J.T., and Wetzel, C. 2015. The status of canary rockfish (Sebastes pinniger) in the California Current in 2015. Pacific Fish. Manag. Counc. (June): 22.

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

Walters, C.J., Martell, S.J.D., and Korman, J. 2006. A stochastic approach to stock reduction analysis. Can. J. Fish. Aquat. Sci. **63**(1): 212–223. doi: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. Fish. Res. **110**(2): 342–355. doi: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. Fish. Res. **171**: 170–187. Elsevier B.V. doi:10.1016/j.fishres.2015.06.005.

Winker, Henning. 2019. The CMSY method for data-limited stock assessment. Package available at https://github.com/Henning-Winker/cmsy.

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. Can. J. Fish. Aquat. Sci. **69**(8): 1292–1301. doi:10.1139/f2012-060.

Zhou, S., Punt, A.E., Ye, Y., Ellis, N., Dichmont, C.M., Haddon, M., Smith, D.C. and Smith, A.D. 2017. Estimating stock depletion level from patterns of catch history. Fish and Fisheries. **18**(4): 742–751. doi: 10.1111/faf.12201.

Table 1. Life-history information and priors for the three species used in the study. Notation is as follows: *Lognormal* (μ, σ^2); Uniform *U* (a, b). Priors for *K* were Uniform between the maximum catch in the time series and 100 times the maximum catch. For the Catch-MSY method, the depletion priors were Uniform centered on the true value with a minimum of *true - 0.1* and a maximum of *true + 0.1*.

Operating model inputs	Symbol	Short-lived	Medium-lived	Long-lived
Maximum age	Age_{max}	10	15	64
Age at 50% maturity (years)	A_{mat}	3	5	16
Length where 50% of the fish are mature (FL cm)	L_{50}	29	90	55
Length where 95% of the fish are mature (FL cm)	L_{95}	34	100	57
Length-weight scaling parameter	α	2.73x10^{-6}	1.34x10^{-5}	1.80x10^{-5}
Length-weight allometric parameter	β	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	t_0	-0.6	-1.3	-1.9
Coefficient of variation of length at age for all ages	CVL	0.1	0.1	0.1
Natural mortality (1/years)	M	0.60	0.30	0.05
Steepness	h	0.5	0.9	0.8
Selectivity at 50% (cm)	S_{50}	25	60	45
Selectivity at 95% (cm)	S_{95}	30	75	50
Recruitment deviations	σ_R	0.3	0.4	0.5
Fishing mortality deviations	σ_F	0.2	0.2	0.2
Estimation models prior distributions				
Depletion (used for all catch-based models)	XB_0	*Lognormal* (true, 0.1)	*Lognormal* (true, 0.1)	*Lognormal* (true, 0.1)
Carrying capacity (used for Catch-MSY, CMSY and SSCOM)	K	*U* (max(catch), max(catch)x100)	*U* (max(catch), max(catch)x100)	*U* (max(catch), max(catch)x100)
Population rate of increase (used for Catch-MSY, CMSY and SSCOM)	r	*U* (0.8, 1.2)	*U* (0.2, 0.6)	*U* (0.05, 0.4)
Steepness (used for SSS)	h	*Normal* (0.5,0.1)	*Normal* (0.9,0.1)	*Normal* (0.8,0.1)
Vulnerability (used for DBSRA)	F_{MSY}/M	*U* (0, 2)	*U* (0, 2)	*U* (0, 2)
Compensation (used for DBSRA)	B_{MSY}/B_0	*U* (0, 1)	*U* (0, 1)	*U* (0, 1)

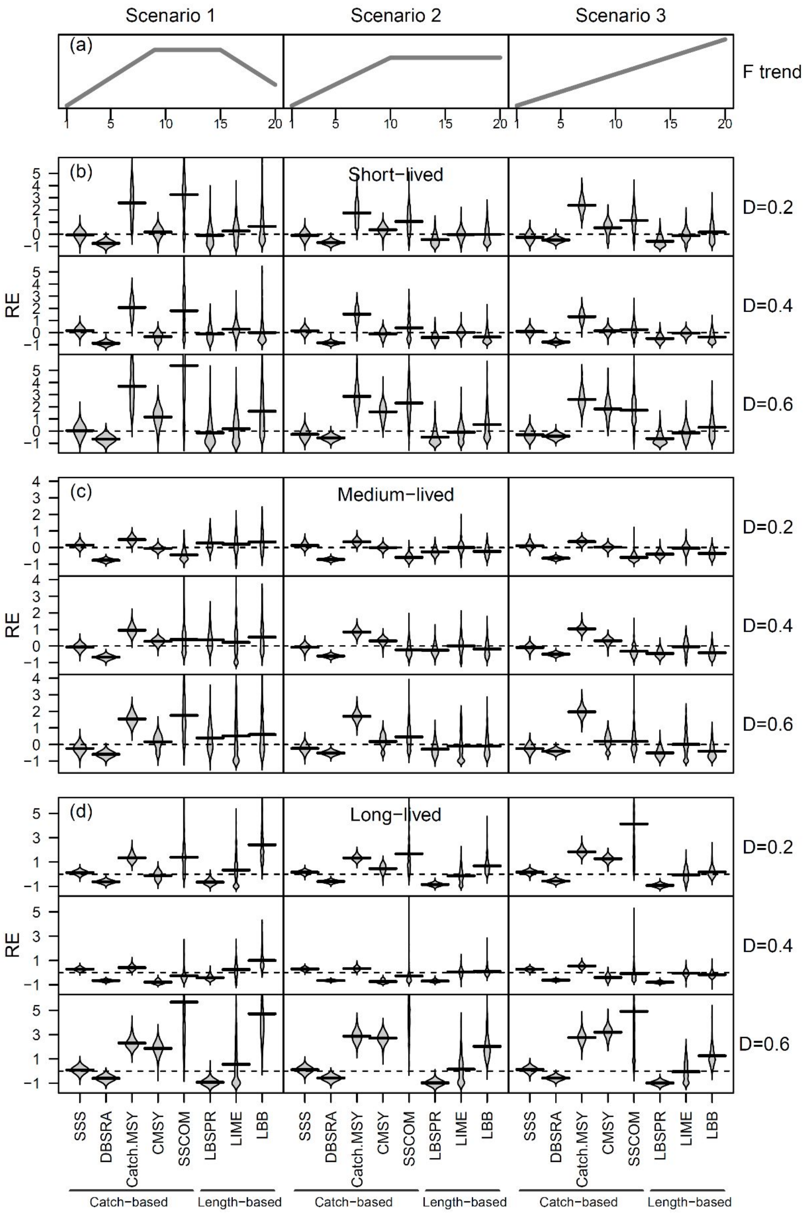


Figure 1. Relative error (RE) in exploitation rate for all the catch-based and length-based models considered under the different harvest scenarios (a) and depletion scenarios for differing life histories, (b) short-lived, (c) medium-lived, and (d) long-lived.

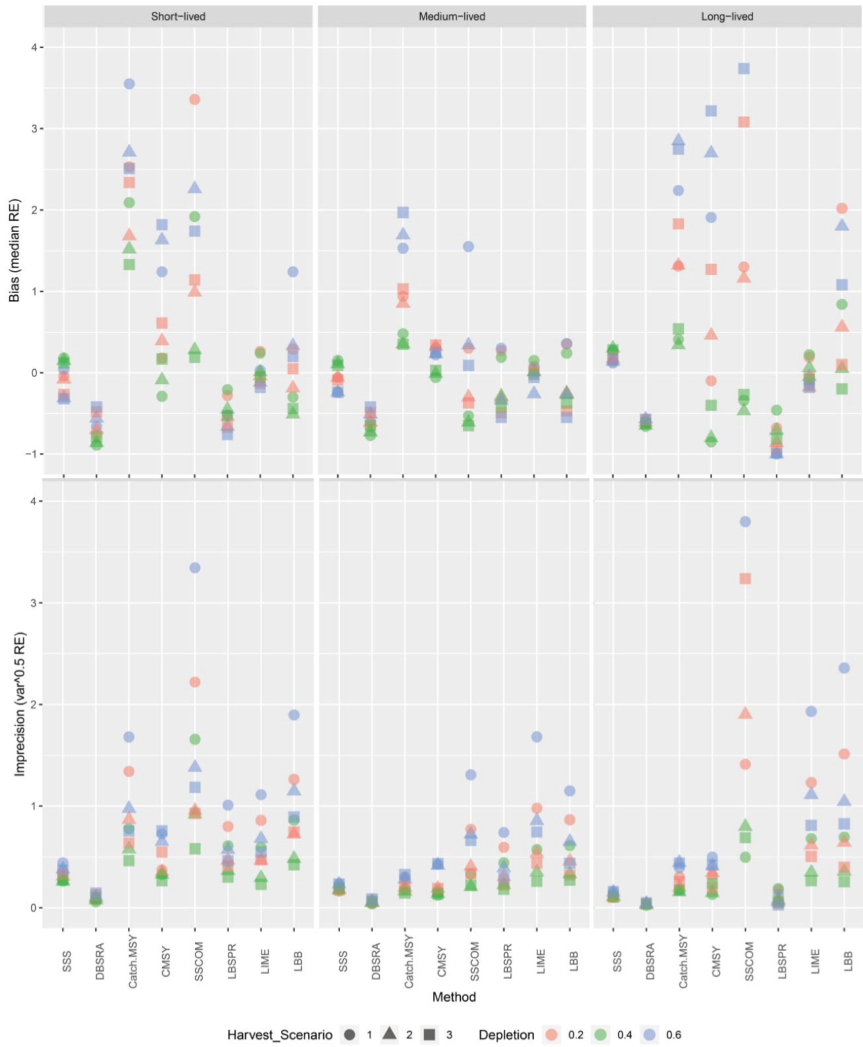


Figure 2. Bias (top panels) and imprecision (bottom panels) for all the catch-based and length-based models considered under the different harvest scenarios and life histories. The y-axes were truncated at 4, so 3 and 2 scenarios are not visualized in the upper and lower panels, respectively.

Appendix

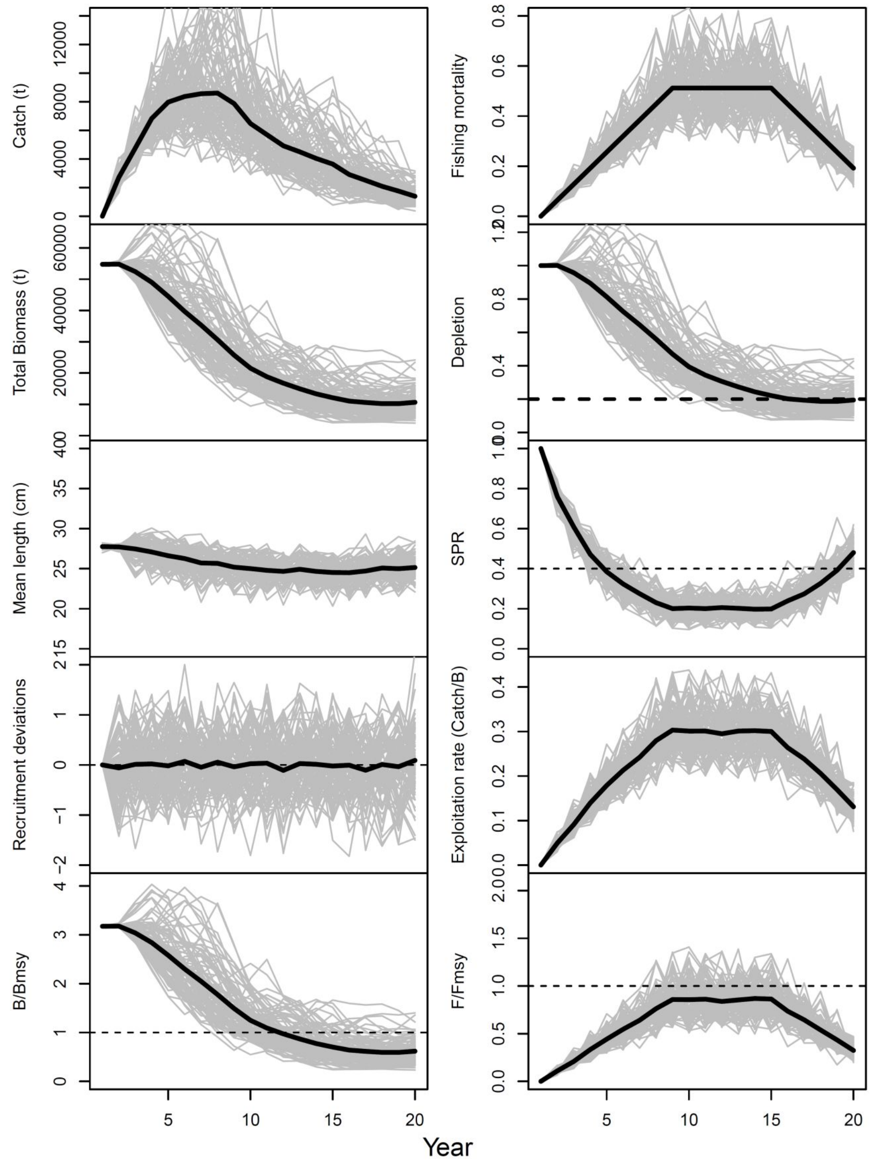


Figure A1. Time series for each simulated fast-grow chub mackerel population for harvest scenario 1 and depletion = 0.2. The black solid lines represent the mean value for all runs.

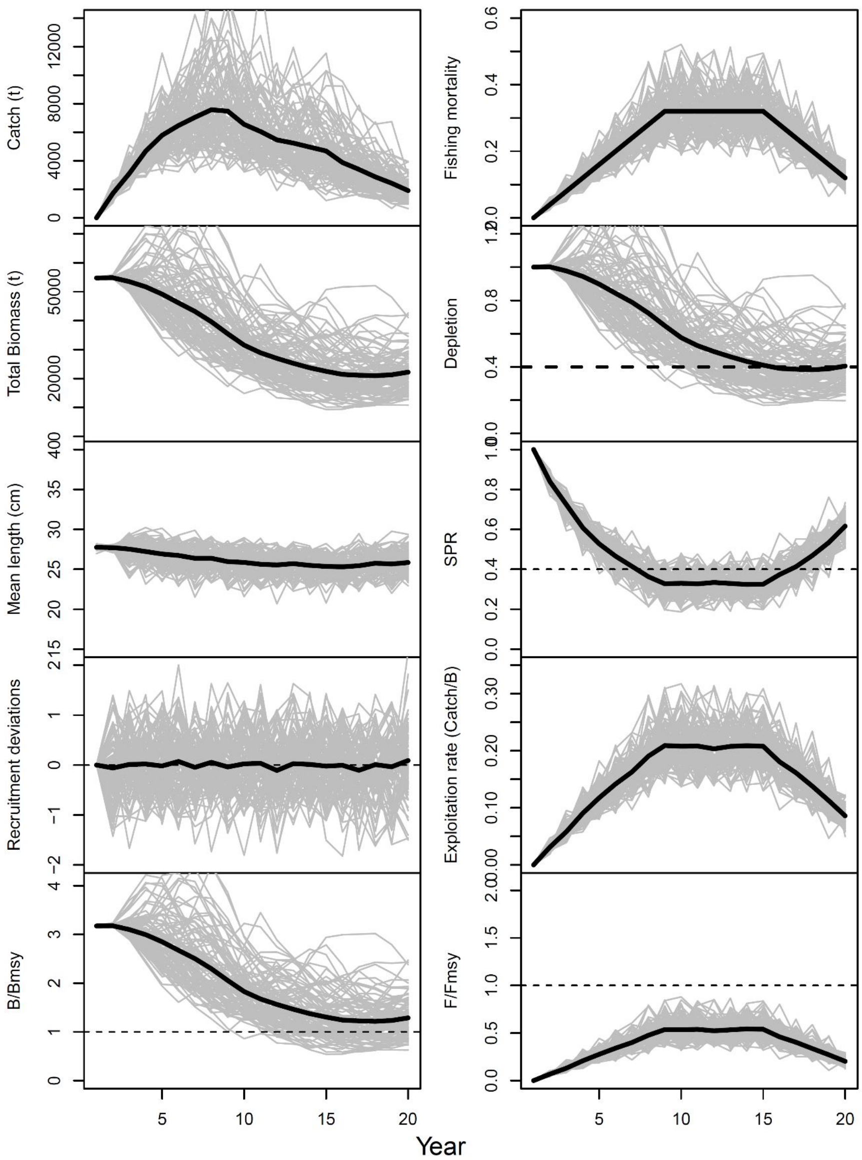


Figure A2. Time series for each simulated fast-grow chub mackerel population for harvest scenario 1 and depletion = 0.4. The black solid lines represent the mean value for all runs.

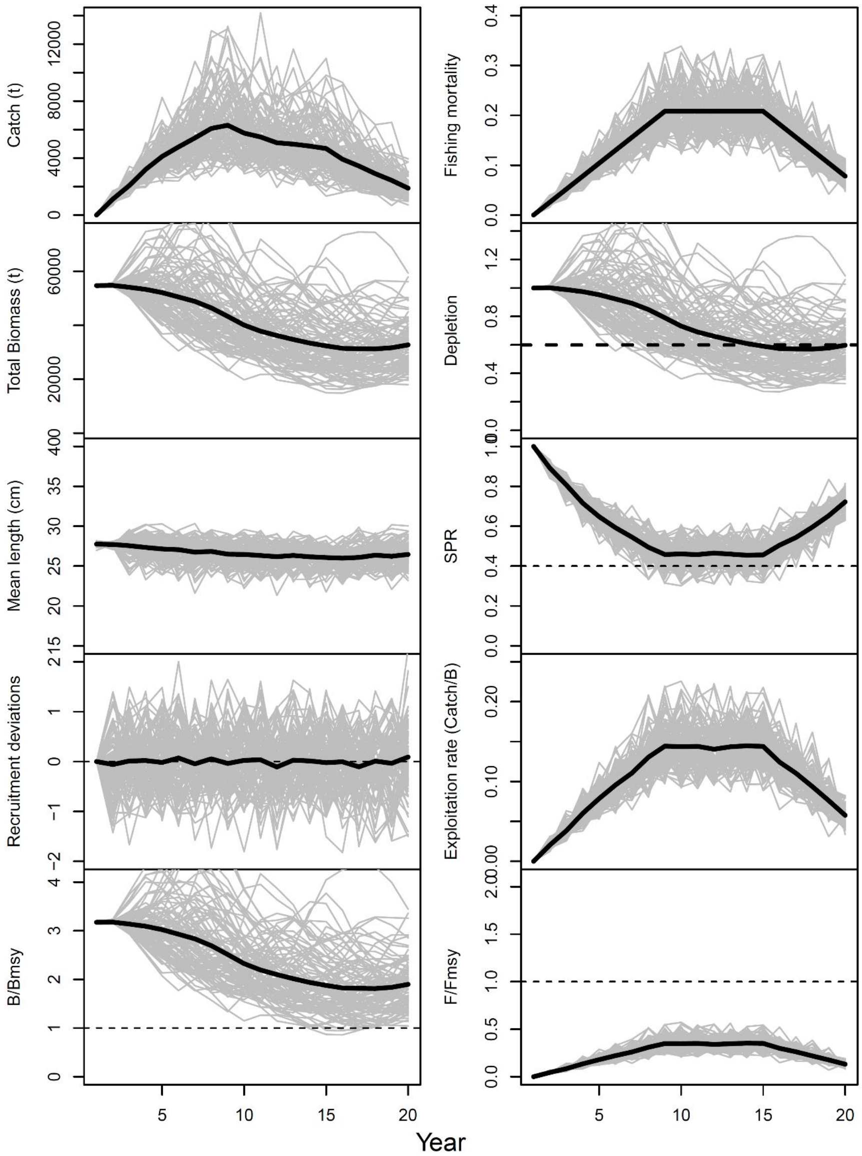


Figure A3. Time series for each simulated fast-grow chub mackerel population for harvest scenario 1 and depletion = 0.6. The black solid lines represent the mean value for all runs.

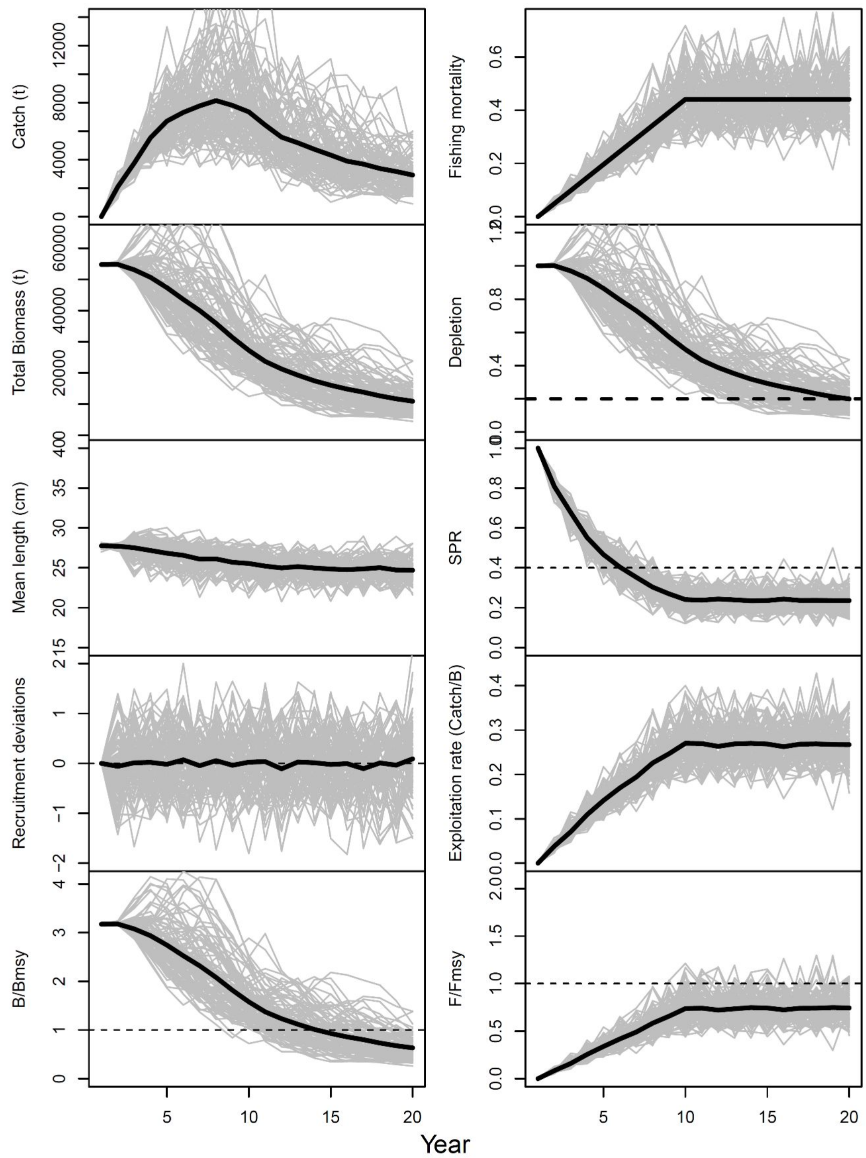


Figure A4. Time series for each simulated fast-grow chub mackerel population for harvest scenario 2 and depletion = 0.2. The black solid lines represent the mean value for all runs.

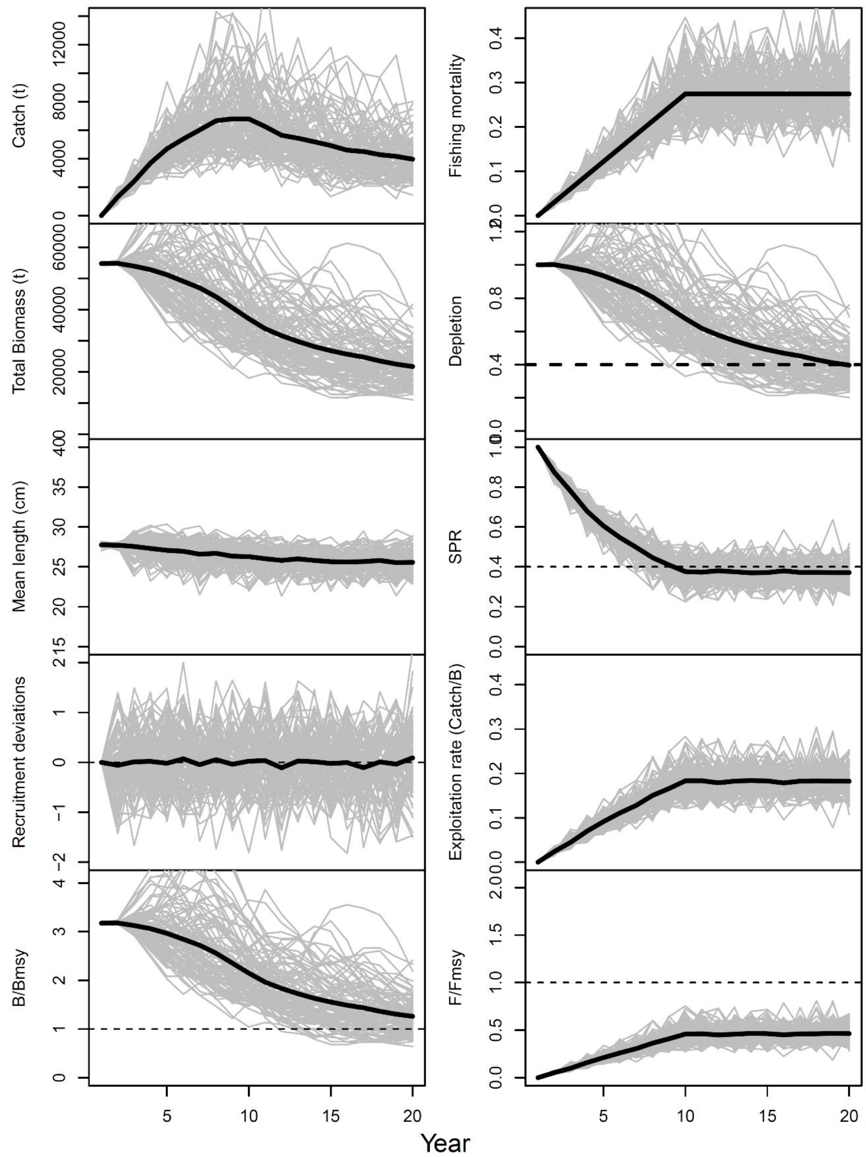


Figure A5. Time series for each simulated fast-grow chub mackerel population for harvest scenario 2 and depletion = 0.4. The black solid lines represent the mean value for all runs.

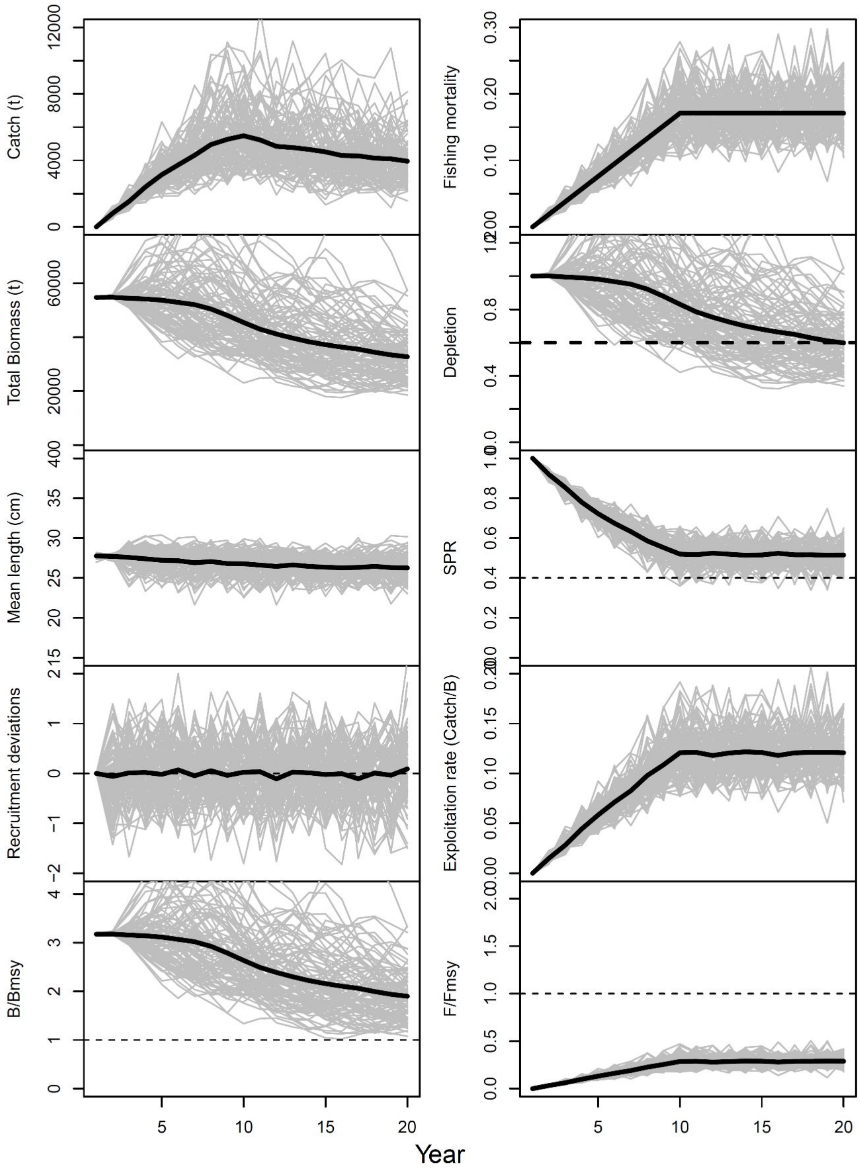


Figure A6. Time series for each simulated fast-grow chub mackerel population for harvest scenario 2 and depletion = 0.6. The black solid lines represent the mean value for all runs.

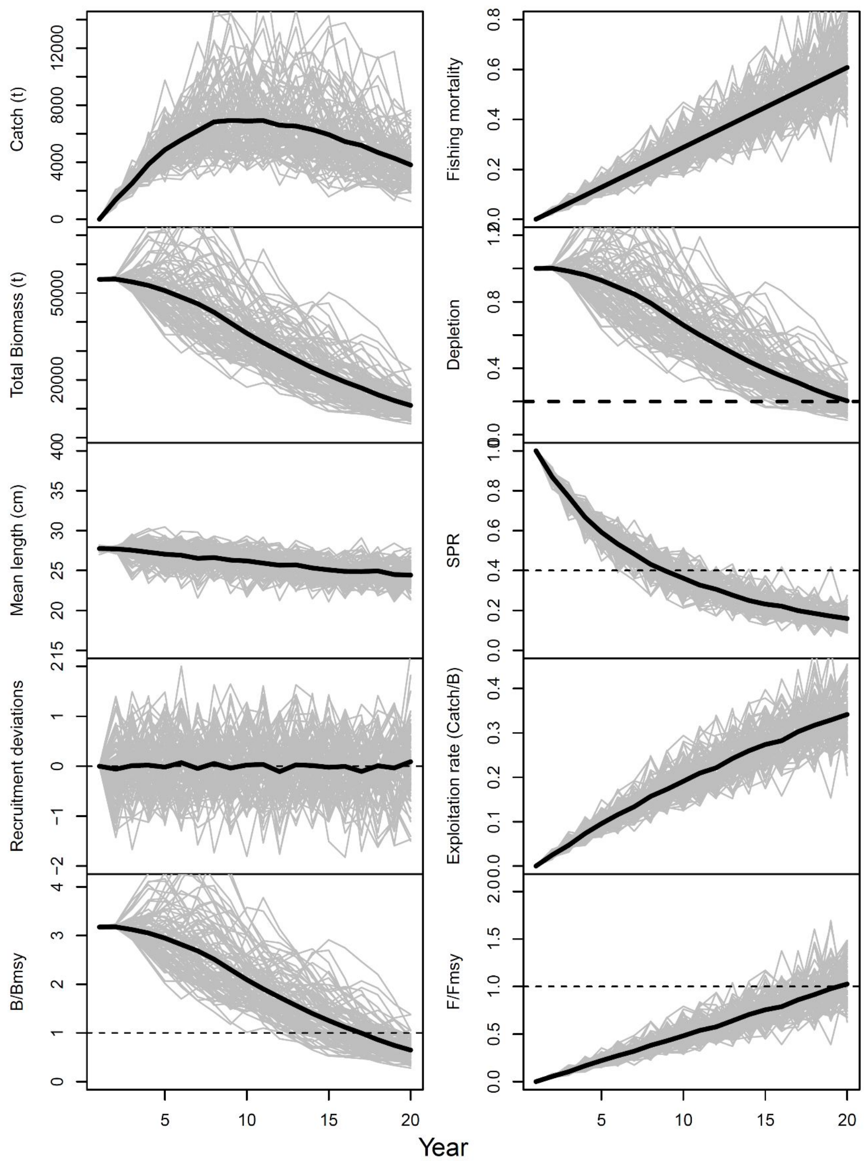


Figure A7. Time series for each simulated fast-grow chub mackerel population for harvest scenario 3 and depletion = 0.2. The black solid lines represent the mean value for all runs.

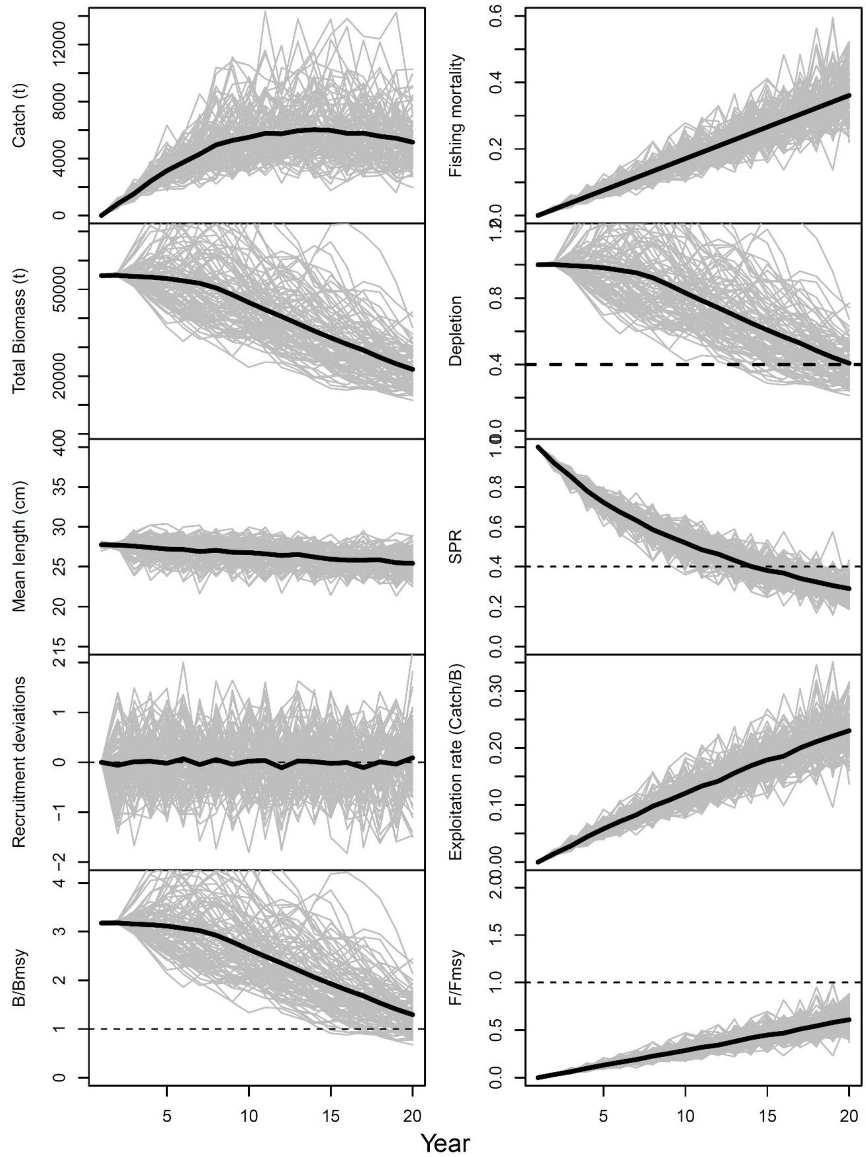


Figure A8. Time series for each simulated fast-grow chub mackerel population for harvest scenario 3 and depletion = 0.4. The black solid lines represent the mean value for all runs.

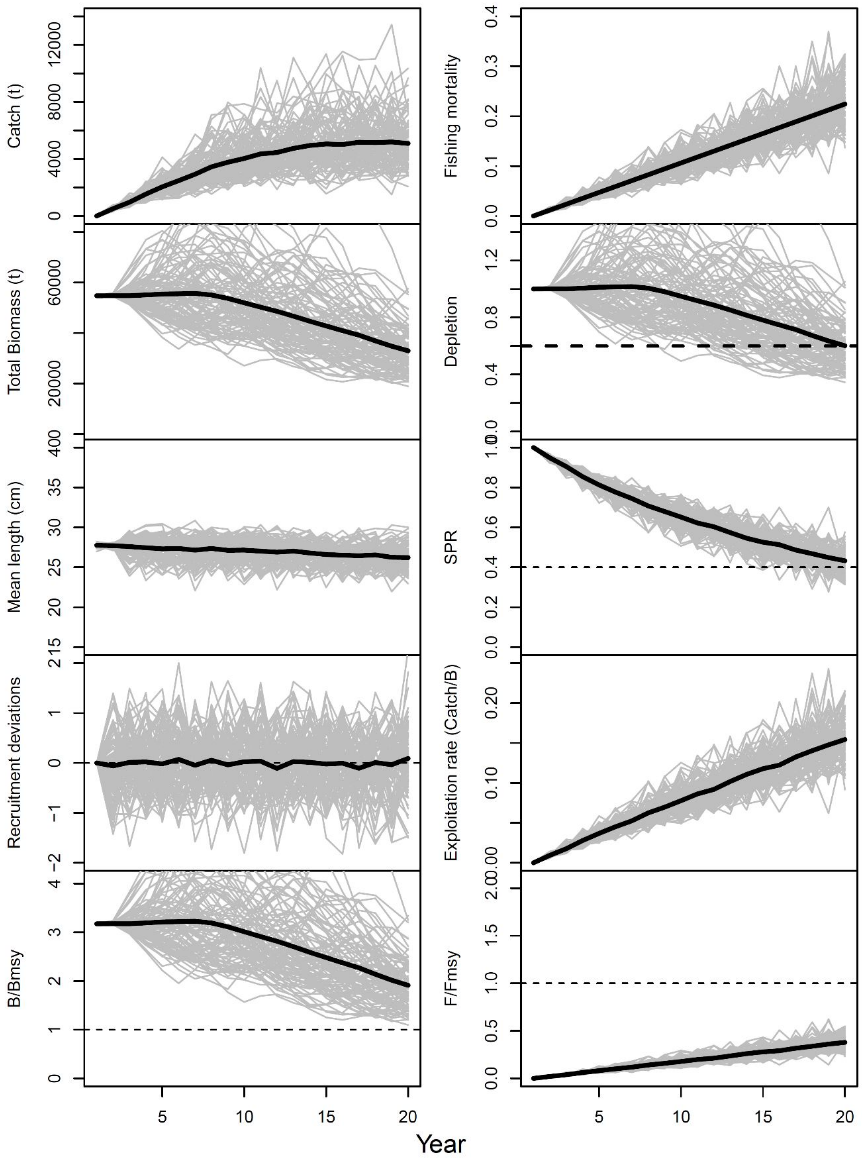


Figure A9. Time series for each simulated fast-grow chub mackerel population for harvest scenario 3 and depletion = 0.6. The black solid lines represent the mean value for all runs.

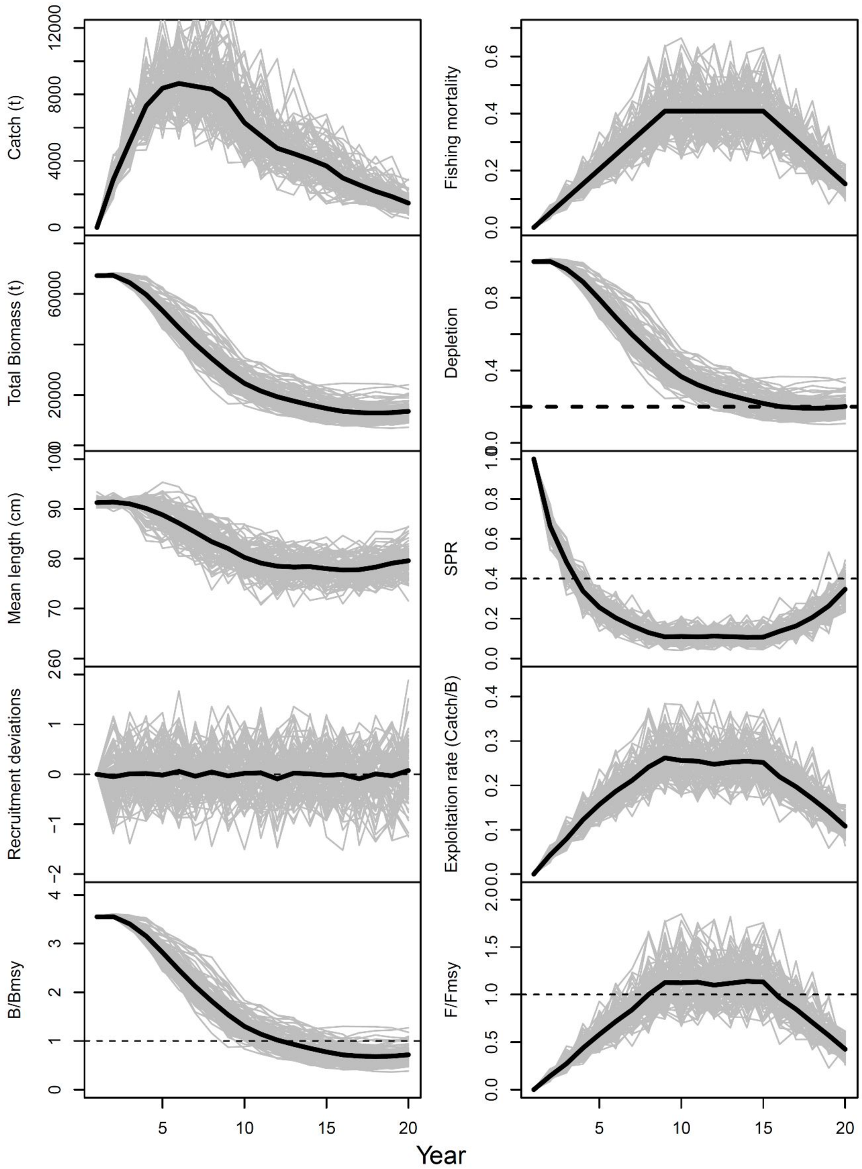


Figure A10. Time series for each simulated medium-grow albacore tuna population for harvest scenario 1 and depletion = 0.2. The black solid lines represent the mean value for all runs.

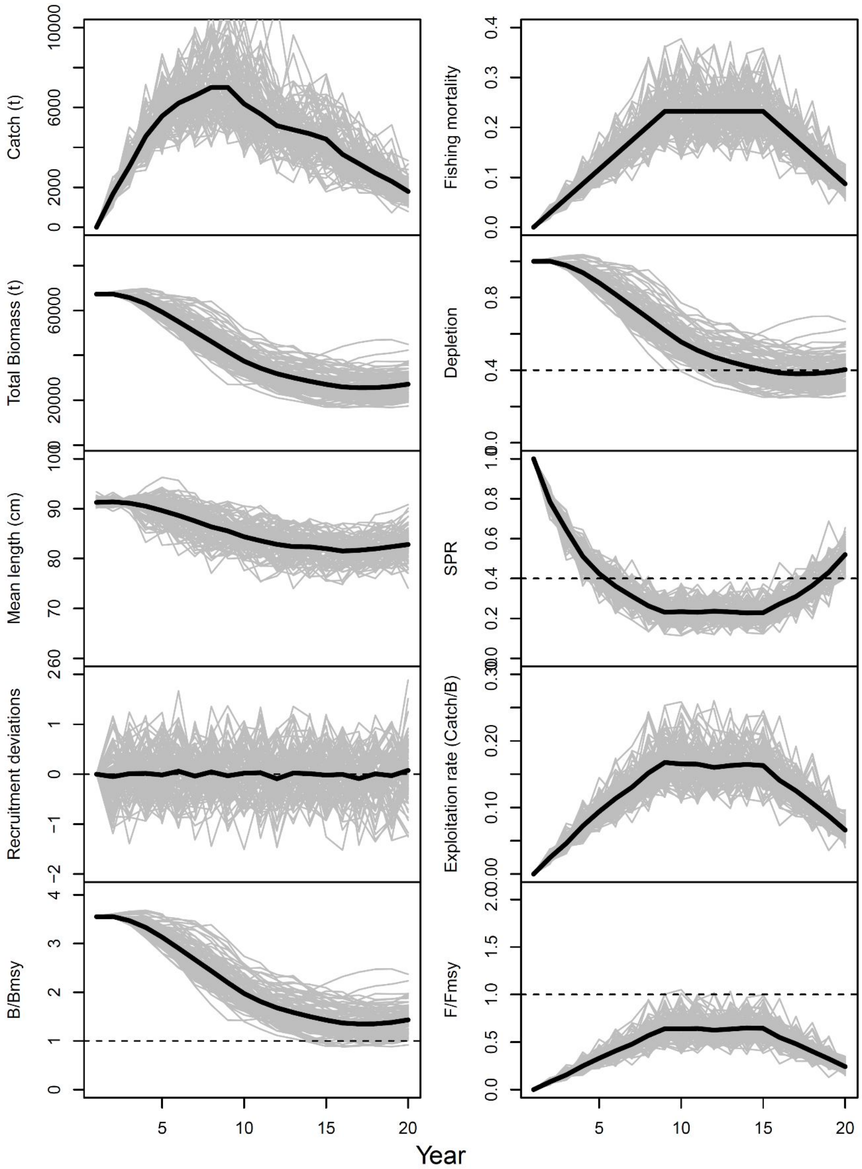


Figure A11. Time series for each simulated medium-grow albacore tuna population for harvest scenario 1 and depletion = 0.4. The black solid lines represent the mean value for all runs.

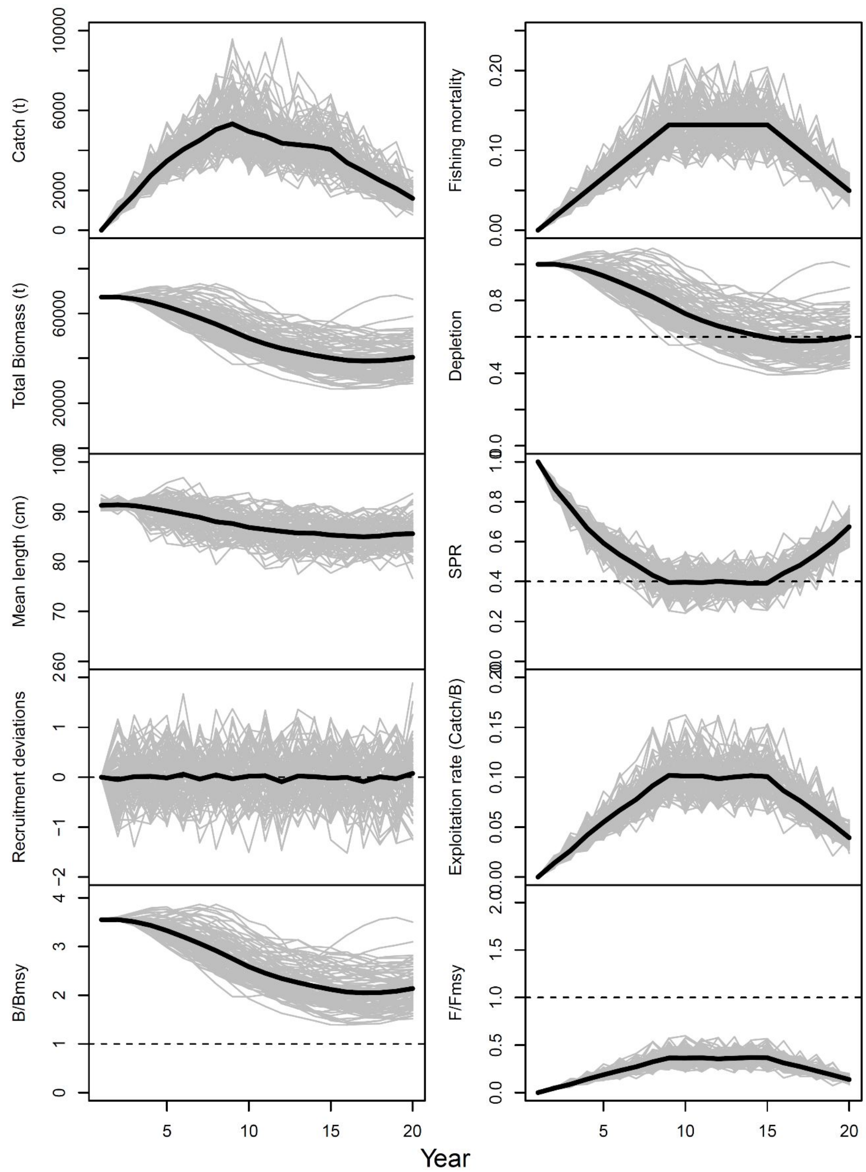


Figure A12. Time series for each simulated medium-grow albacore tuna population for harvest scenario 1 and depletion = 0.6. The black solid lines represent the mean value for all runs.

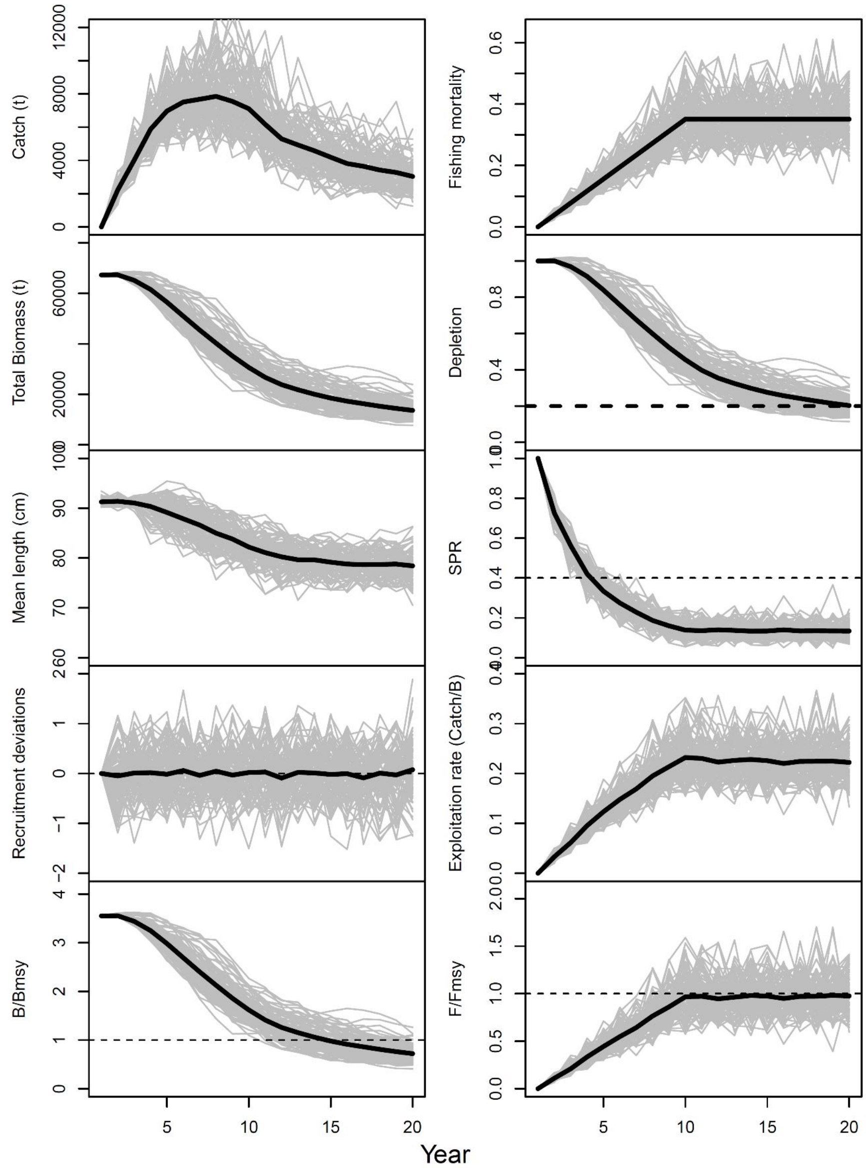


Figure A13. Time series for each simulated medium-grow albacore tuna population for harvest scenario 2 and depletion = 0.2. The black solid lines represent the mean value for all runs.

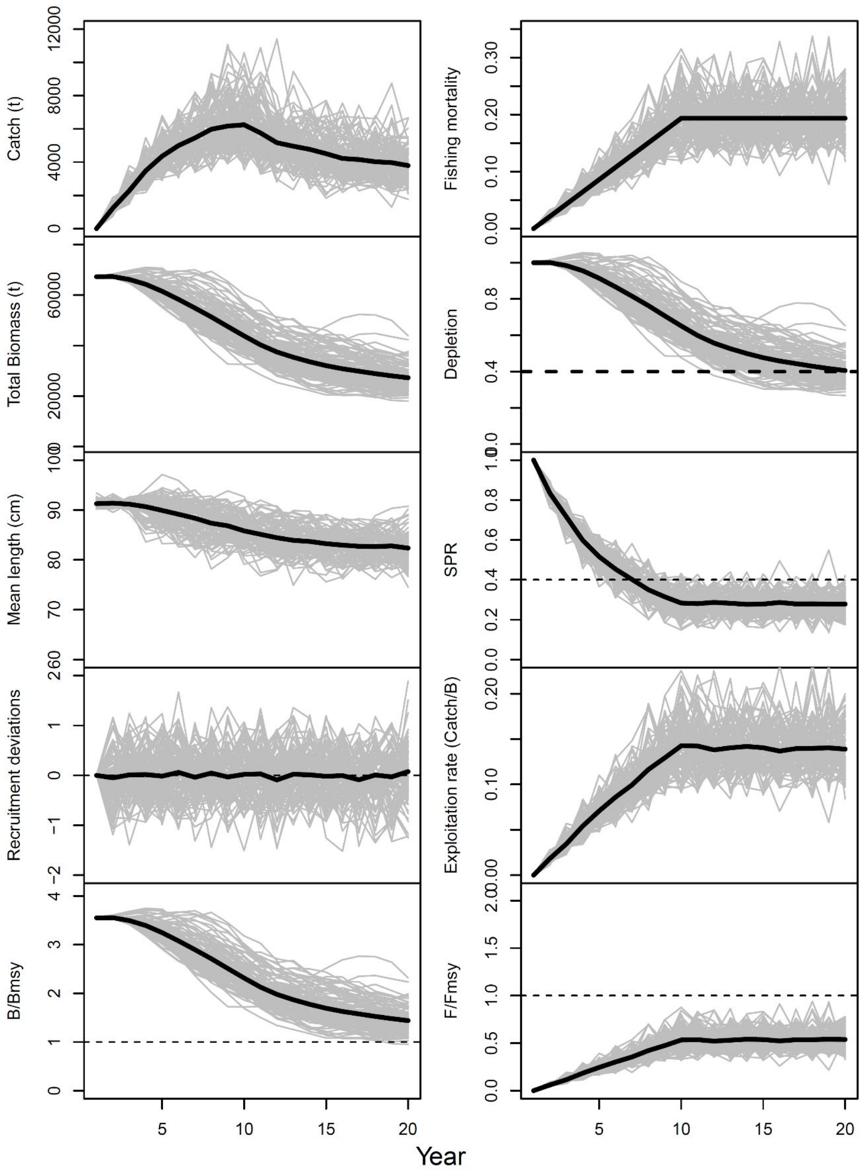


Figure A14. Time series for each simulated medium-grow albacore tuna population for harvest scenario 2 and depletion = 0.4. The black solid lines represent the mean value for all runs.

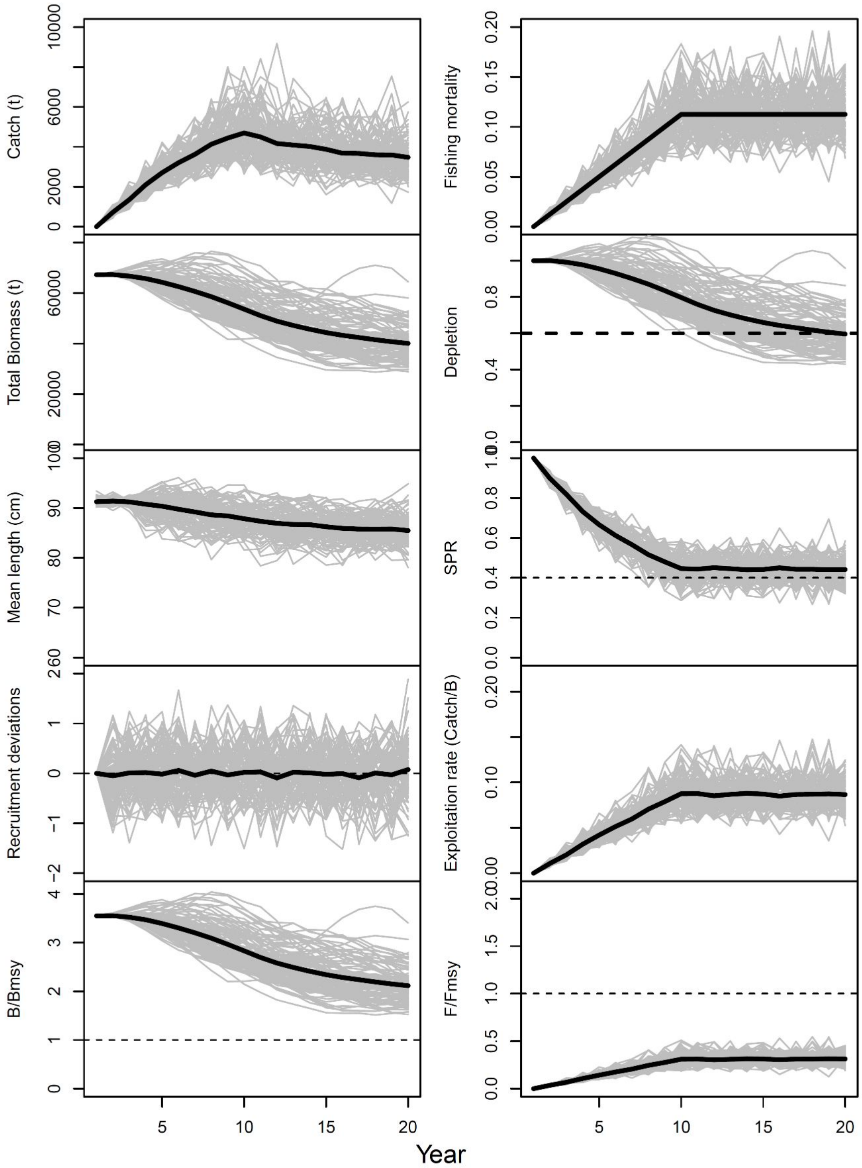


Figure A15. Time series for each simulated medium-grow albacore tuna population for harvest scenario 2 and depletion = 0.6. The black solid lines represent the mean value for all runs.

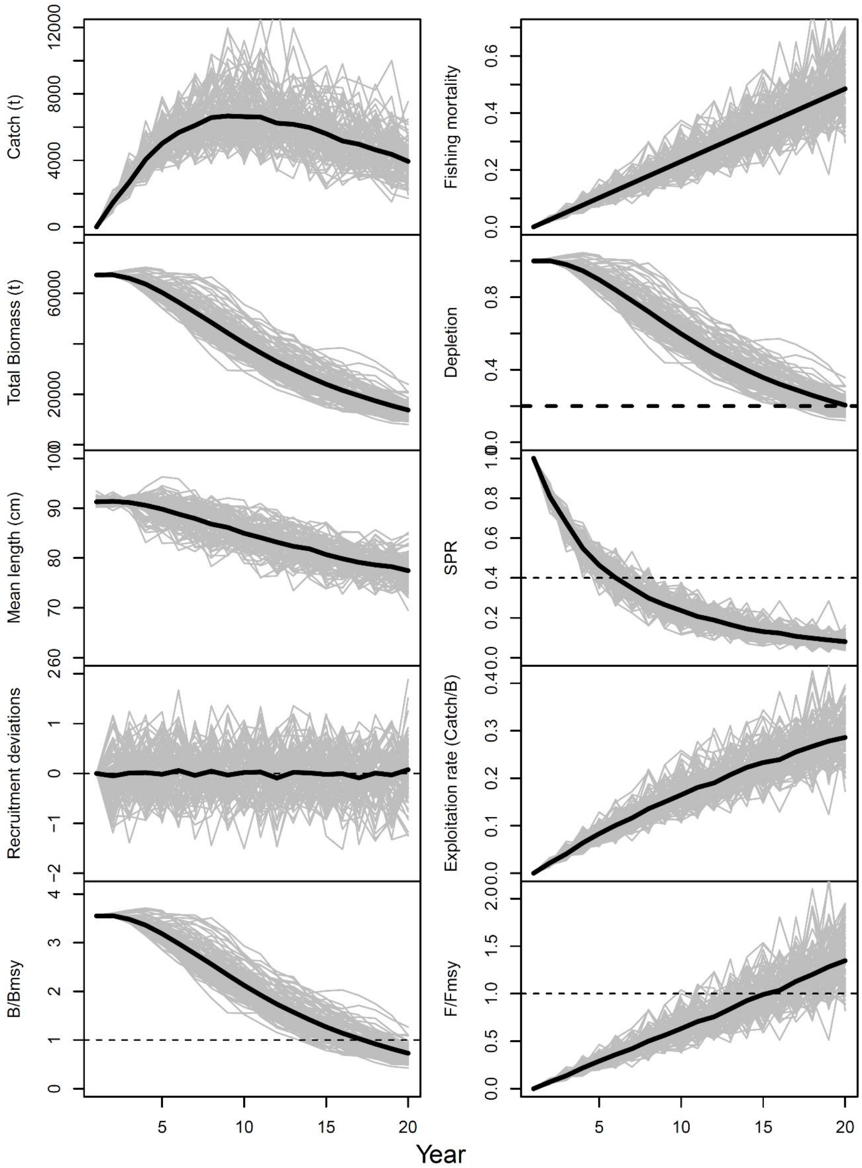


Figure A16. Time series for each simulated medium-grow albacore tuna population for harvest scenario 3 and depletion = 0.2. The black solid lines represent the mean value for all runs.

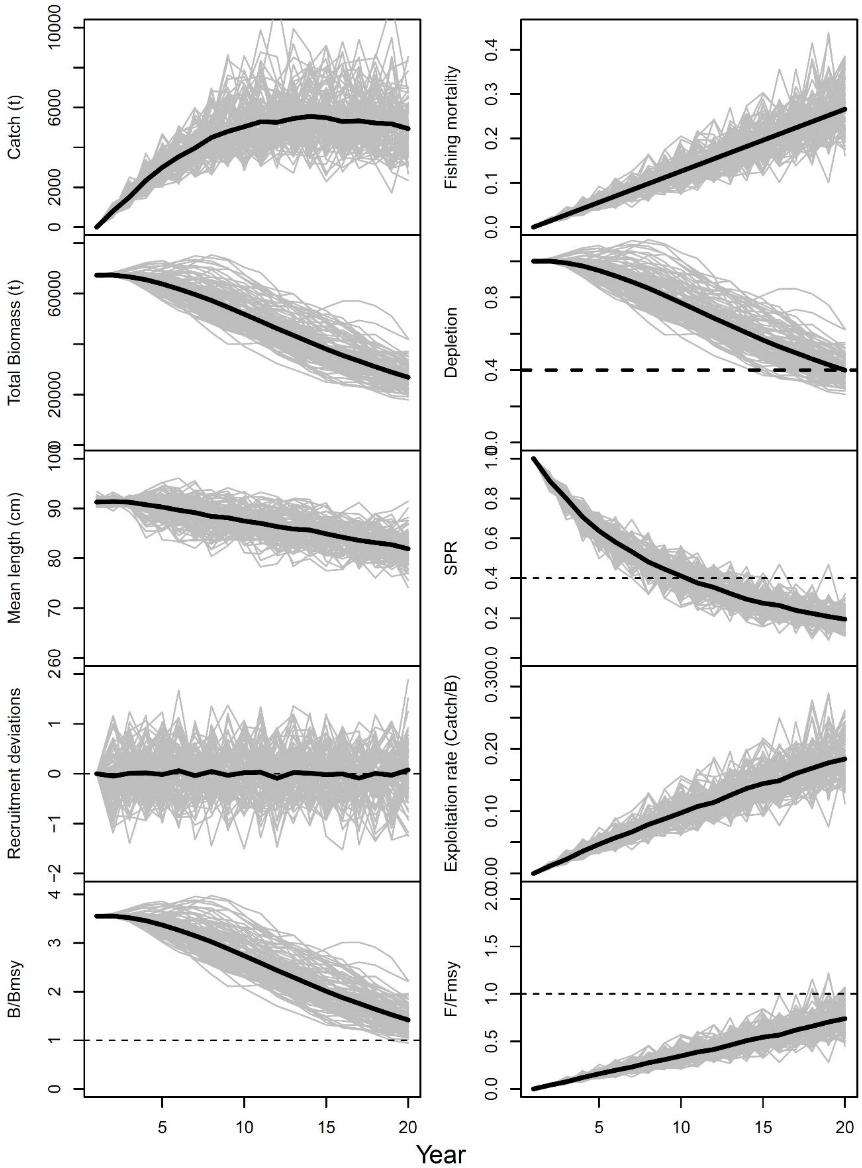


Figure A17. Time series for each simulated medium-grow albacore tuna population for harvest scenario 3 and depletion = 0.4. The black solid lines represent the mean value for all runs.

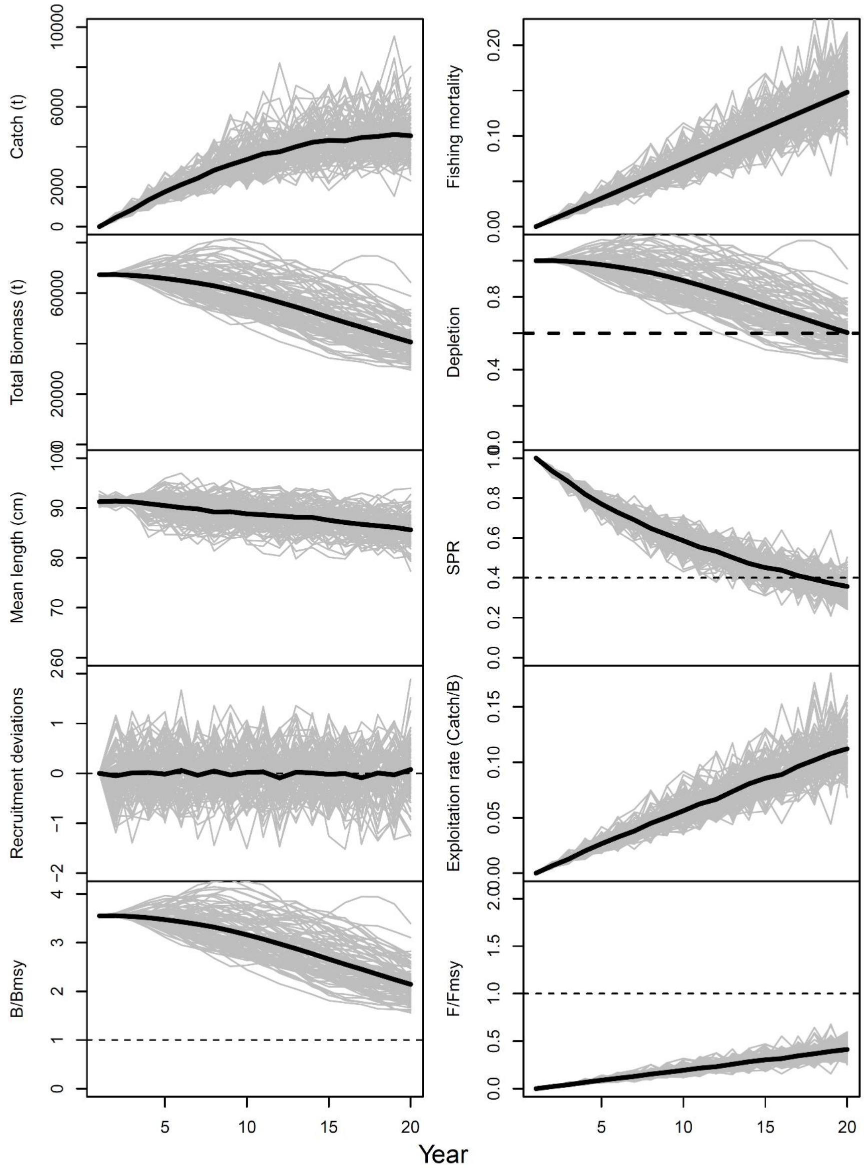


Figure A18. Time series for each simulated medium-grow albacore tuna population for harvest scenario 3 and depletion = 0.6. The black solid lines represent the mean value for all runs.

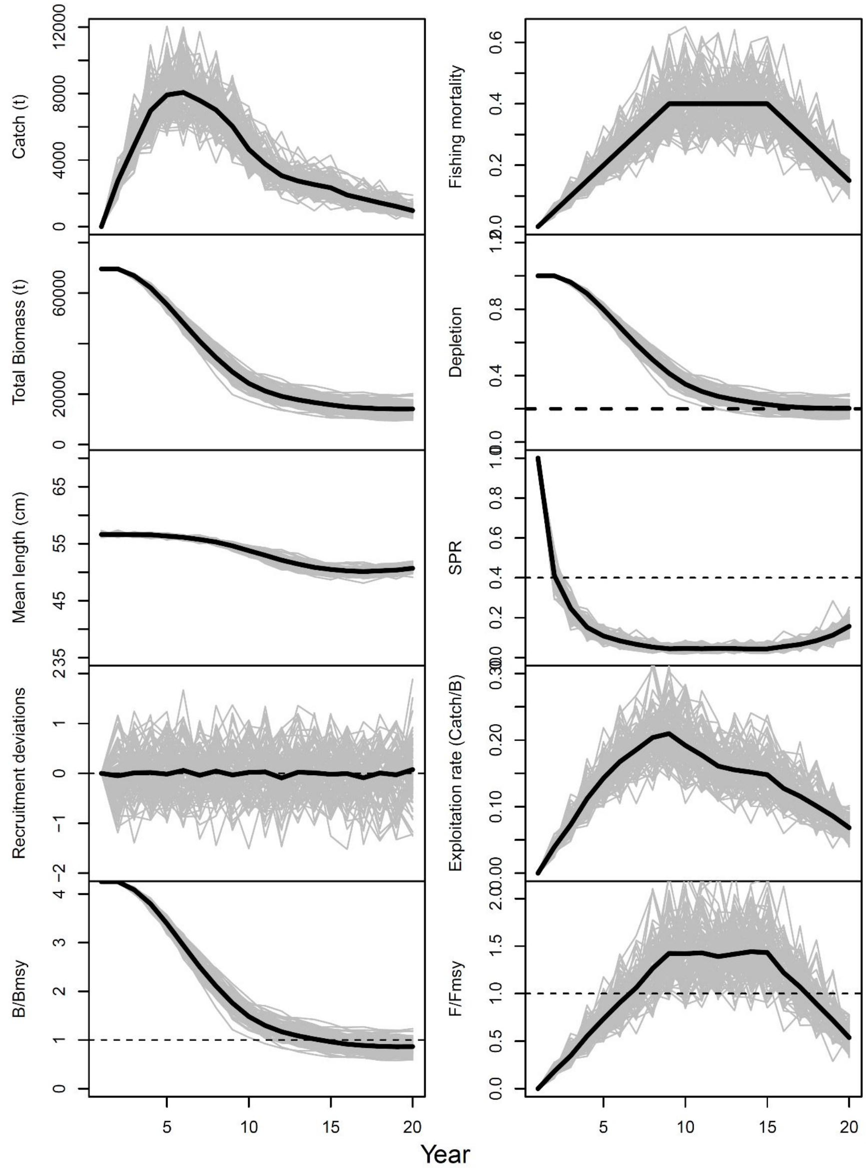


Figure A19. Time series for each simulated slow-grow canary rockfish population for harvest scenario 1 and depletion = 0.2. The black solid lines represent the mean value for all runs.

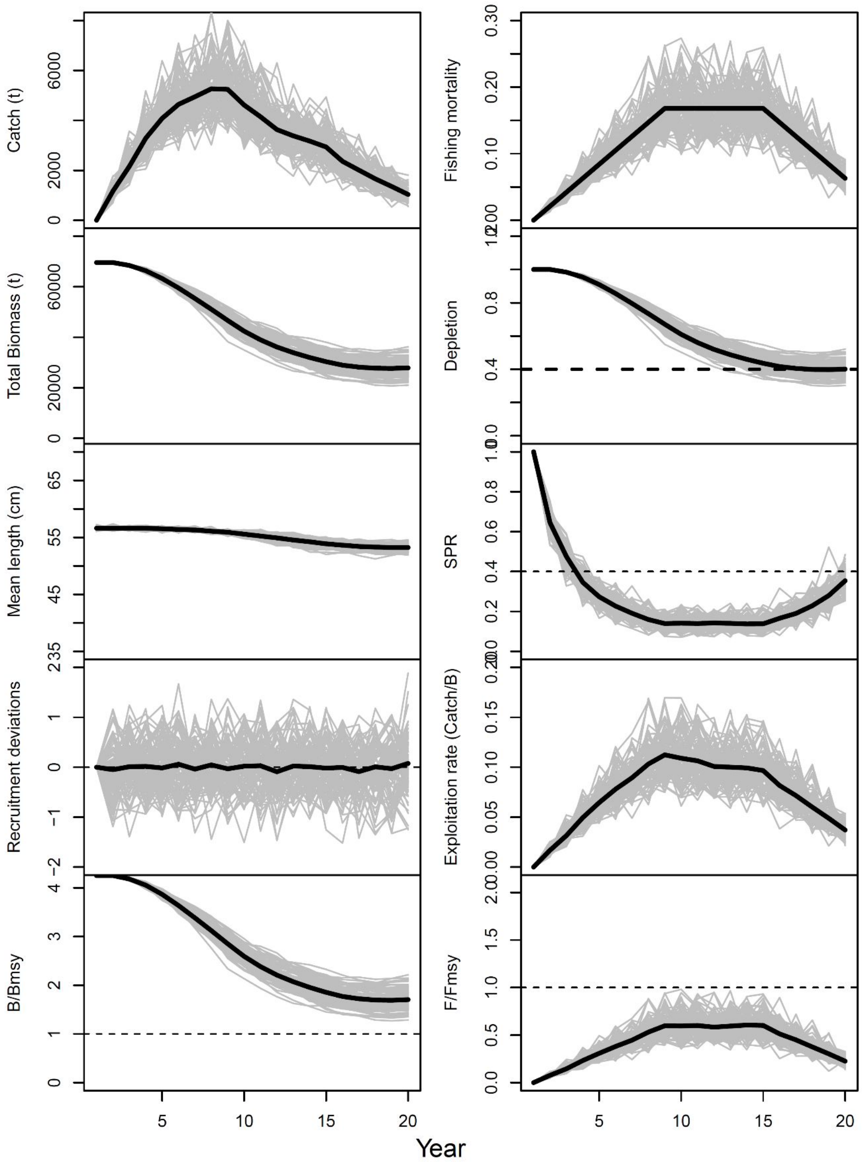


Figure A20. Time series for each simulated slow-grow canary rockfish population for harvest scenario 1 and depletion = 0.4. The black solid lines represent the mean value for all runs.

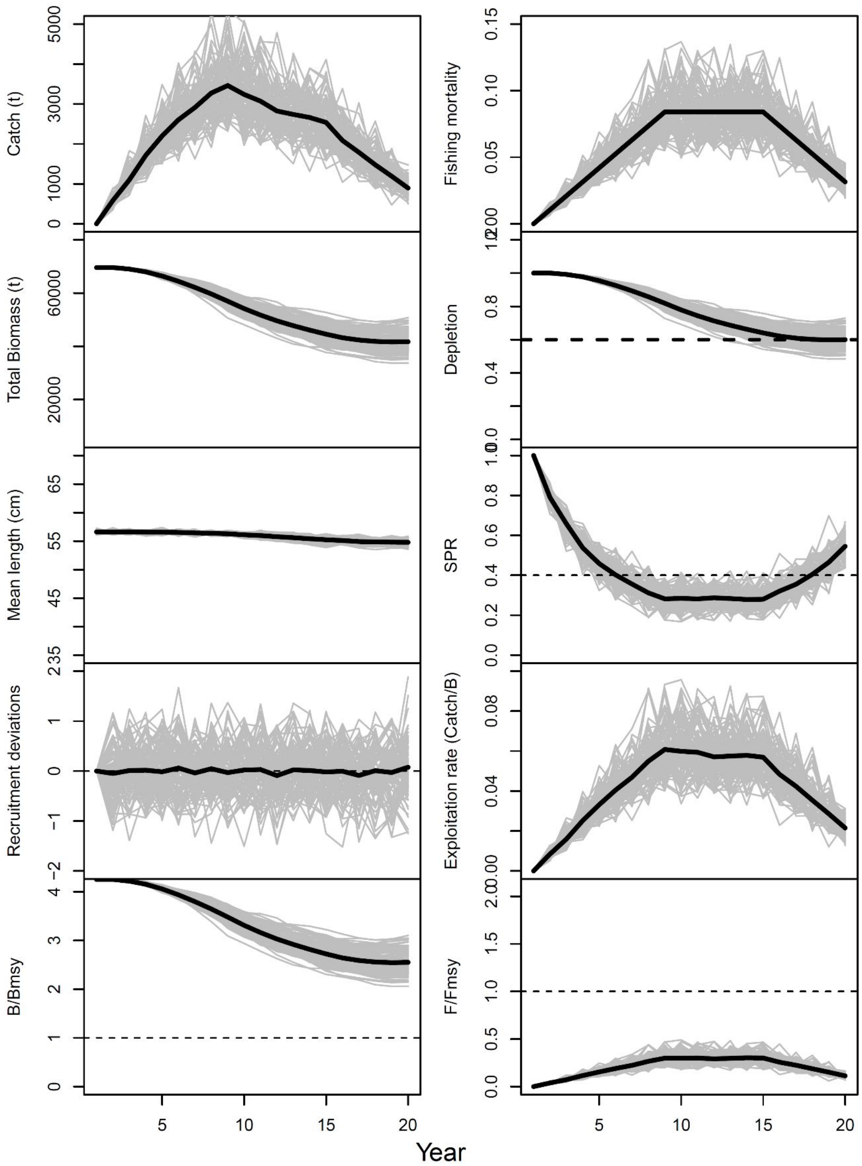


Figure A21. Time series for each simulated slow-grow canary rockfish population for harvest scenario 1 and depletion = 0.6. The black solid lines represent the mean value for all runs.

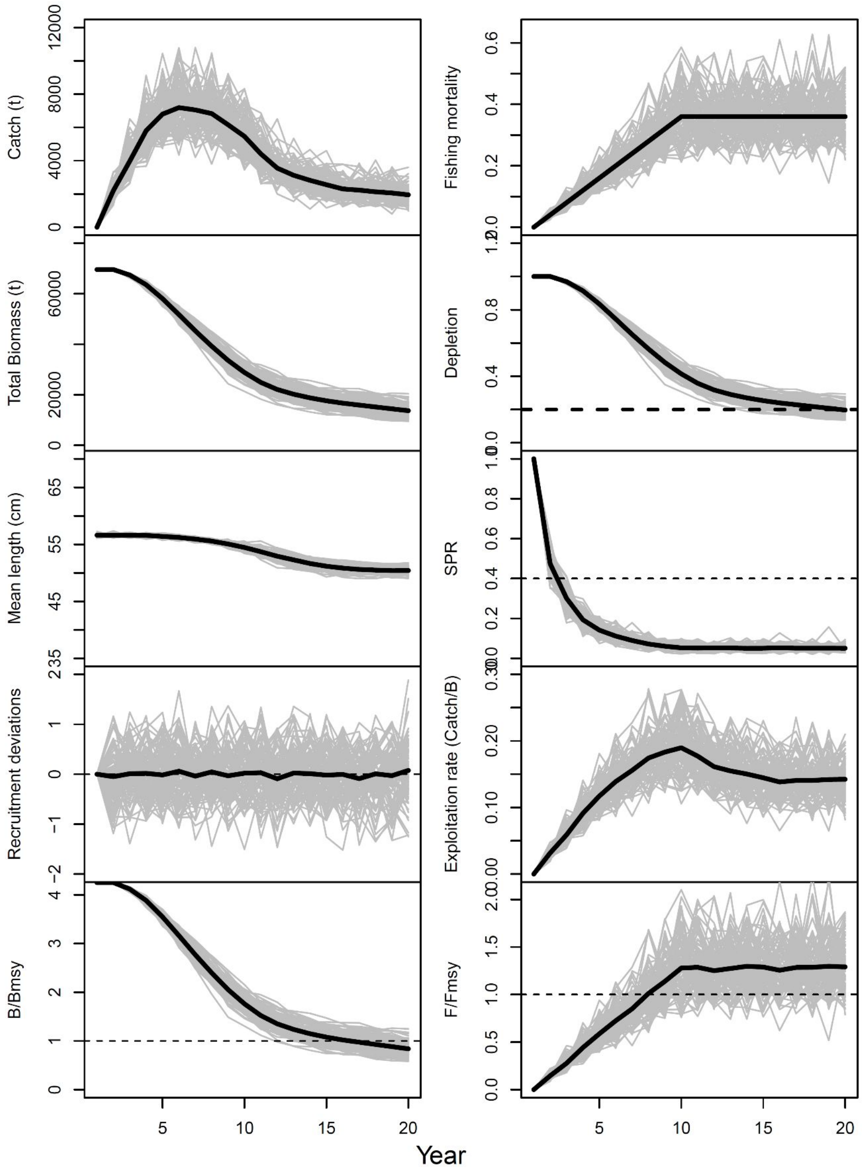


Figure A22. Time series for each simulated slow-grow canary rockfish population for harvest scenario 2 and depletion = 0.2. The black solid lines represent the mean value for all runs.

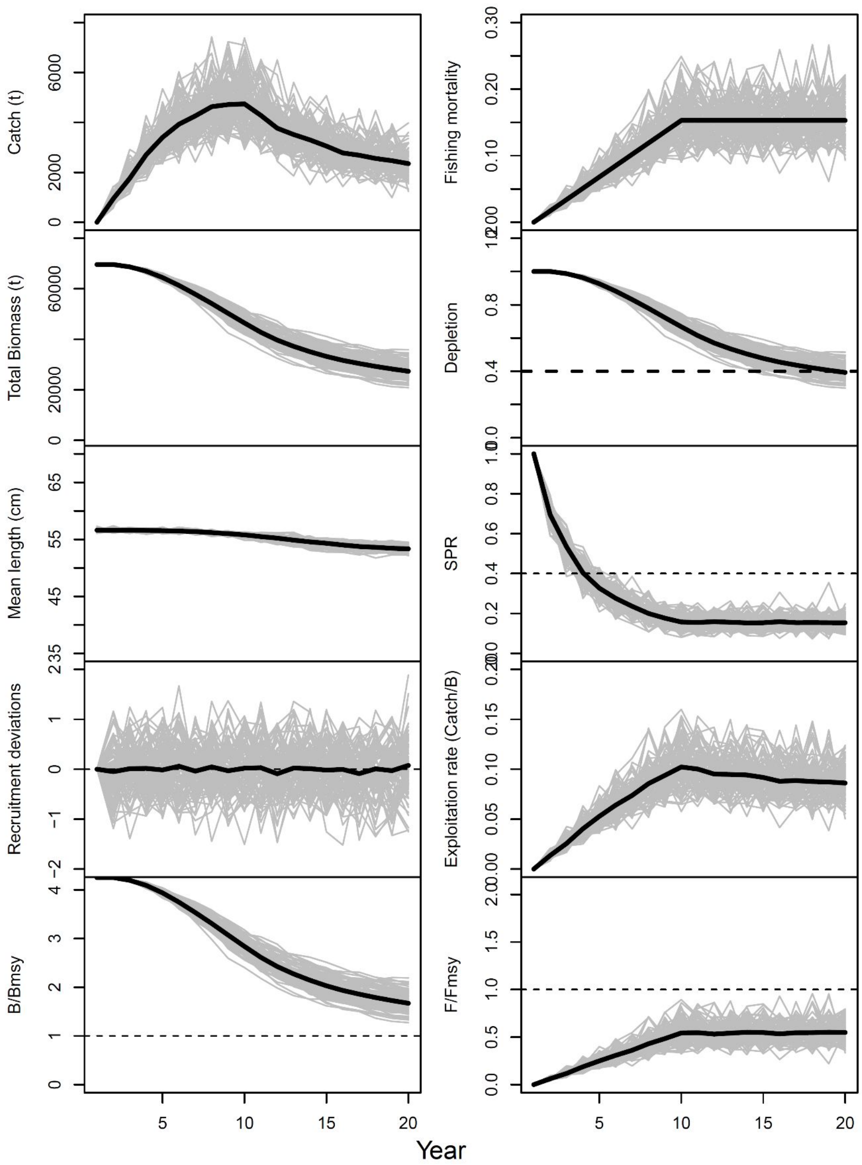


Figure A23. Time series for each simulated slow-grow canary rockfish population for harvest scenario 2 and depletion = 0.4. The black solid lines represent the mean value for all runs.

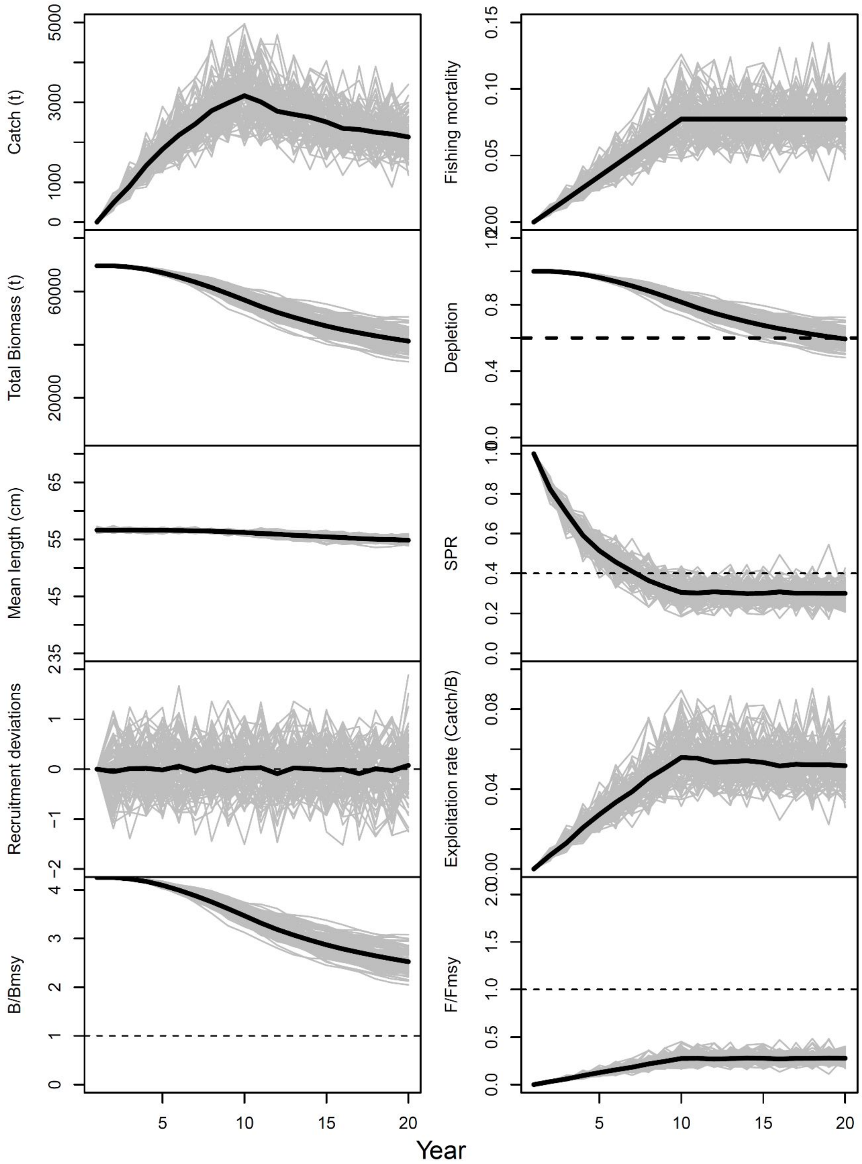


Figure A24. Time series for each simulated slow-grow canary rockfish population for harvest scenario 2 and depletion = 0.6. The black solid lines represent the mean value for all runs.

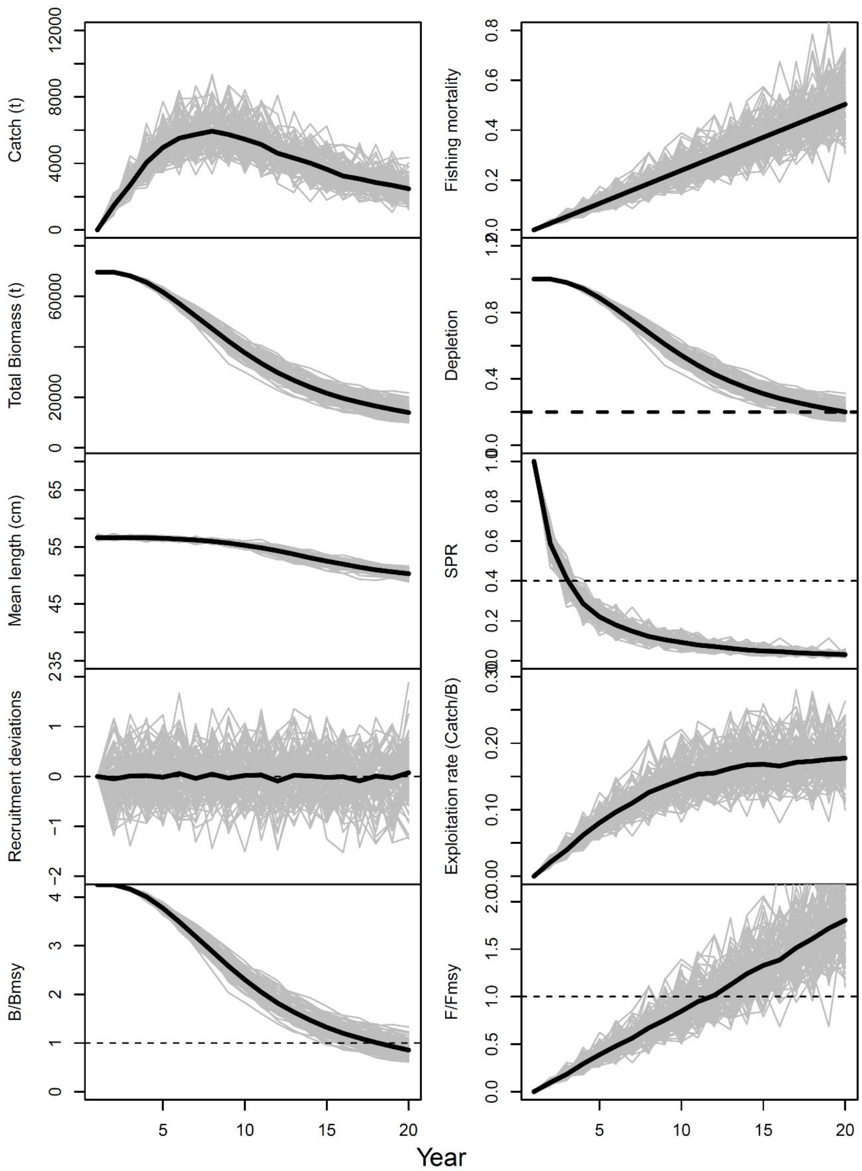


Figure A25. Time series for each simulated slow-grow canary rockfish population for harvest scenario 1 and depletion = 0.2. The black solid lines represent the mean value for all runs.

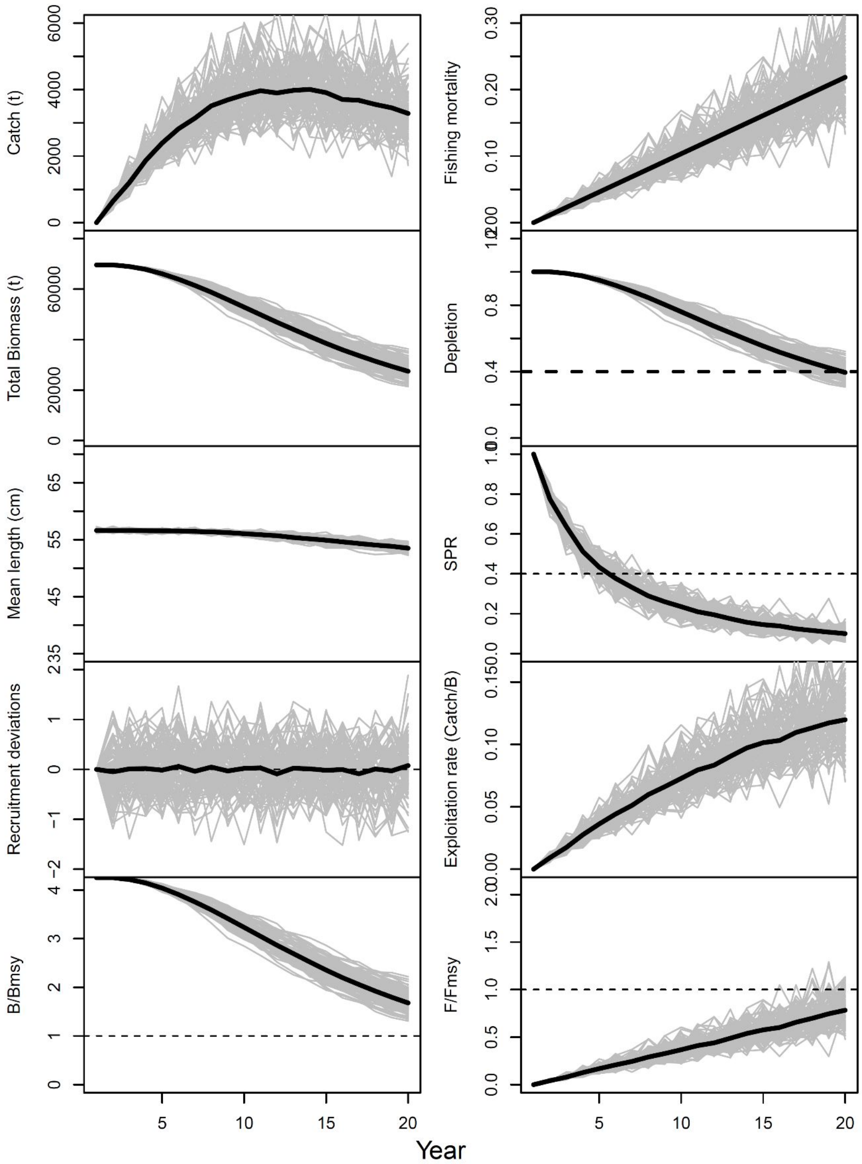


Figure A26. Time series for each simulated slow-grow canary rockfish population for harvest scenario 3 and depletion = 0.4. The black solid lines represent the mean value for all runs.

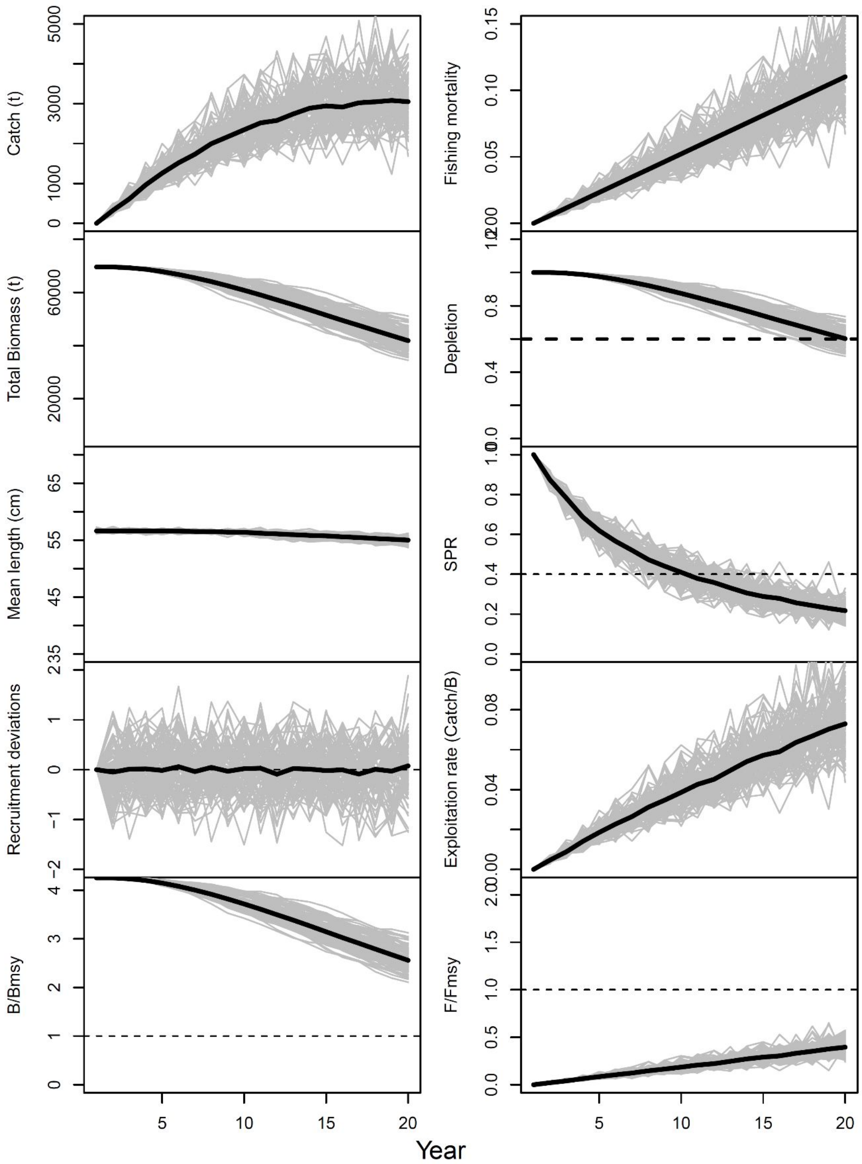


Figure A27. Time series for each simulated slow-grow canary rockfish population for harvest scenario 3 and depletion = 0.6. The black solid lines represent the mean value for all runs.

Table A1. Mean and standard deviation of relative error (RE) between the maximum sustainable yield (MSY) estimated by the operating model (OM) and the MSY estimated by the CMSY method. Values are proportions. Positive values mean that the MSY was overestimated and negative values that it was underestimated.

Species	Harvest Scenario	Mean ± Standard deviation		
		Depletion level		
		0.2	0.4	0.6
Mackerel	Scenario 1	0.37 ± 0.08	0.18 ± 0.09	-0.13 ± 0.15
	Scenario 2	0.28 ± 0.08	0.03 ± 0.09	-0.22 ± 0.13
	Scenario 3	0.33 ± 0.11	0.05 ± 0.11	-0.26 ± 0.14
Albacore	Scenario 1	-0.07 ± 0.02	-0.18 ± 0.03	-0.46 ± 0.02
	Scenario 2	-0.25 ± 0.02	-0.42 ± 0.02	-0.42 ± 0.02
	Scenario 3	-0.12 ± 0.02	-0.24 ± 0.02	-0.43 ± 0.02
Rockfish	Scenario 1	1.61 ± 0.27	1.13 ± 0.31	0.80 ± 0.32
	Scenario 2	1.67 ± 0.26	1.27 ± 0.33	0.86 ± 0.34
	Scenario 3	1.77 ± 0.33	1.29 ± 0.45	0.94 ± 0.45

Table A2. True OM and LBB estimated values for L_∞ and S_{50} (~Lc). LCL is the lower confidence limit and UCL the upper confidence limit for the estimated values.

Scenarios	Life-history	Harvest trend	Final depletion	True L_∞	True S_{50}	Estimated L_∞	Estimated LCL L_∞	Estimated UCL L_∞	Estimated Lc	Estimated LCL Lc	Estimated UCL Lc
1	Short-lived	Scenario 1	0.2	38.2	25	44.5	44.0	45.1	16.9	16.5	17.4
2			0.4	38.2	29	43.6	43.3	44.0	15.6	15.5	15.8
3			0.6	38.2	29	45.1	44.6	45.6	20.1	19.5	20.7
4		Scenario 2	0.2	38.2	29	44.7	44.2	45.2	18.0	17.4	18.6
5			0.4	38.2	29	43.7	43.4	44.1	15.7	15.5	15.8
6			0.6	38.2	29	45.3	45.0	45.8	20.3	19.7	20.9
7		Scenario 3	0.2	38.2	29	45.5	45.2	46.0	20.3	19.8	21.0
8			0.4	38.2	29	44.5	44.0	45.1	15.9	15.7	16.1
9			0.6	38.2	29	45.7	45.4	46.2	20.7	20.1	21.3
10	Medium-lived	Scenario 1	0.2	122.2	60	142.5	141.4	143.6	62.4	61.7	63.2
11			0.4	122.2	60	141.4	140.1	142.5	60.4	59.7	61.0
12			0.6	122.2	60	143.2	142.7	143.9	64.1	63.2	64.9
13		Scenario 2	0.2	122.2	60	143.0	141.8	143.9	62.8	62.0	63.6
14			0.4	122.2	60	141.9	141.1	143.3	60.7	60.0	61.4
15			0.6	122.2	60	143.3	143.0	144.0	64.6	63.8	65.4
16		Scenario 3	0.2	122.2	60	143.3	143.0	144.0	63.3	62.5	64.1
17			0.4	122.2	60	142.5	141.4	143.7	61.1	60.4	61.8
18			0.6	122.2	60	143.5	143.1	144.3	65.4	64.5	66.3
19	Long-lived	Scenario 1	0.2	60.0	60	76.7	76.0	77.6	50.4	49.9	51.0
20			0.4	60.0	45	75.3	74.5	76.3	47.0	46.5	47.4
21			0.6	60.0	45	76.9	76.2	77.8	51.2	50.7	51.8
22		Scenario 2	0.2	60.0	45	76.5	75.8	77.4	50.5	50.0	51.1
23			0.4	60.0	45	75.4	74.6	76.4	47.8	47.2	48.3
24			0.6	60.0	45	77.0	76.3	78.0	51.4	50.8	52.0
25		Scenario 3	0.2	60.0	45	76.8	76.1	77.7	51.0	50.4	51.6
26			0.4	60.0	45	76.0	75.3	76.9	49.5	49.0	50.1
27			0.6	60.0	45	77.1	76.3	78.0	51.7	51.1	52.2

5.03 ROC

An Evaluation of the Robustness of Length Based Indicators using Receiver Operating Characteristics.

Laurence T. Kell

Centre for Environmental Policy, Imperial College London, London SW7 1NE

e-mail Lauriel@seapluslus.co.uk

Coílín Minto

GMIT

Hans Gerritsen

Marine Institute

October 31, 2019

Abstract

Keywords:

An Example of Conducting Management Strategy Evaluation Using Machine Learning to Evaluate Trade-offs Between Multiple Management Objectives

Laurence T. Kell

Coílín Minto

Hans Gerritsen

Alexander J. M. Kell

October 14, 2019

Abstract

An example of conducting Management Strategy Evaluation (MSE) using machine learning to evaluate trade-offs between management objectives using an Operating Model conditioned on life history characteristics to evaluate an empirical Harvest Control Rule. The specific aims of the study are to

- Develop a risk based framework, where is risk is an uncertainty that matters and what matters are management objectives.
- Develop a way of tuning Management Procedures so that case specific management strategies can be developed efficiently.
- Allow stakeholders to more easily agree management objectives and to evaluate the trade-offs between them.

Keywords:

FLife

Laurie Kell (Sea++); Ernesto Jardim (EC JRC); Coilin Minto (GMIT); Iago Mosqueira (EC JRC); Alex Tidd (GMIT)

27 January, 2016

Outline

- Empirical relationships between life history invariants
  - **Fishbase** dataset with 5 habitat groups showing relationships between growth parameters and maturity
  - **Filling in holes** using *lhPar*
- Functional forms
  - **Biology** growth and maturity
  - **Gislason** relationship between M and k and length
  - **Lorenzen** relationship between M and mass-at-age
  - **Fishing Mortality** selectivity
- Equilibrium Dynamics
  - **lhEql** Production function
  - **Stock** recruitment relationship
- **leslie** Leslie matrix
- Elasticity Analysis
- Dynamics
  - **fwd**
  - **noise**
  - **Resonant cohort effects**
- Density Dependence
  - **SRR**
  - **M**
  - **Growth**
- Indicators
  - **lopt** ...
- Estimators
  - **priors**
- Feedback control
  - **hcr P**
  - **hcr D**

Contents