Status of Global Unassessed Fisheries will Remain Highly Uncertain without Better Data

Daniel Ovando

Ray Hilborn

Cole Monnahan

Merrill Rudd

Rishi Sharma

James Thorson

Yannick Rousseau

Yimin Ye

2021-01-16

Assessments of the global state of fish populations play an important role in tracking the implementation of the United Nations Sustainable Development Goals. While we have reliable estimates of stock status for fisheries accounting for approximately half of recent global catch, our knowledge of the state of the majority of the word’s ‘unassessed’ fish stocks remains highly uncertain. Numerous high-profile publications have produced estimates of the global status of these unassessed fisheries, but limited quantity and quality of data along with methodological differences have produced counterintuitive and conflicting results. Here, we show that despite numerous efforts, our understanding of the status of global fish stocks remains incomplete, even when new sources of broadly available data are added. Obtaining accurate estimates of stock status for the world’s fisheries depends on prioritizing the collection of high-priority data around the globe, rather than the development of new modeling methods alone.

data limited are not just messier versions of data rich

80% in trouble

# Introduction

The United Nations Sustainable Development Goal 14 (SDG 14), focusing on “Life under water”, calls for the global community to “Conserve and sustainably use the oceans, seas and marine resources for sustainable development”. Marine fisheries represent one of the largest anthropenic impacts on the oceans, and when managed sustainably are a critical source of economic prosperity and food security around the globe. As such, meeting the SDG 14 targets depends in part on our ability to effectively measure the status of global marine fish populations and fisheries. While our understanding of the state of world fisheries has improved over the last decade the majority of the world’s fish populations, making up roughly 50% of marine landings, lack formal statistical assessments (stock assessments) of their biological status [1,2]. This is a major impediment to ensuring the sustainable development of the world’s oceans. In this paper, we assess why the assessment of global fisheries remains a challenge, and chart a path towards a better understanding of the state of the world’s marine resources.

The Food and Agriculture Organization of the United Nations’ (FAO) is the custodian agency of the SDG 14 Indicator on fisheries sustainability, and the FAO’s State of World Fisheries and Aquaculture (SOFIA) report is the most widely used primary source for tracking the global state of fisheries. The latest SOFIA report, covering 70% of the landings of all fisheries in the world, estimates that as of 2017 59.6% of marine fish stocks are maximally sustainably fished, 6.2% are underfished, and 34.2% are overfished [2]. Where possible, these statements about the status of individual fish stocks are made on the basis of formal stock assessments reported in the RAM Legacy Stock Assessment Database (RAM) [3]. The “assessed” fisheries in RAM represent roughly 50% of global capture as of 2020 [1], and represent our best estimates of the biological state of “assessed” fish populations around the globe.

That leaves roughly 50% of global landings, and the majority of global fisheries, as currently “unassessed”. The SOFIA report includes unassessed stocks and bases its estimates for these fisheries on data-limited methods or qualitative expert opinion for each region where these stocks are distributed [2]. While these methods combined with local knowledge can provide good insight as to the status of unassessed fish stocks, the SOFIA assessment was designed in the 1970s based on the then available data and methods. With the surge in data availability, such as global assessments of management strength, trawl footprints, and fishing effort, the SOFIA global assessment methods for unassessed fish populations need to be updated to meet the demand for tracking progress of global fish populations towards the SDG goals.

We evaluate the ability of alternative assessment models to improve our estimate of the population status of unassessed fish stocks around the world. While these unassessed stocks are generally individually smaller and less economically valuable than the fish populations in the assessed category, collectively they are a vital source of food, employment, cultural value, and ecosystem services around the world. Numerous studies in recent years have put forward versions of “data-limited” models that have attempted to provide estimates for the global status of unassessed fish stocks lacking the data needed for formal stock assessment [4–8]. Due to data limitations, all of these global assessment efforts use forms of “catch-only” data-limited models (Free et al. 2020 [9] and references therein). These models seek to infer stock status, for example biomass (*B*) relative to the biomass at maximum sustainable yield (*BMSY*), from characteristics of a fishery’s catch history, and have been the foundation of many subsequent analyses of the future of global fisheries (e.g. [**cabral2020?**], [**gaines2018?**]).

However, Free et al. 2020 [9] demonstrated that these catch-only models can often produce both imprecise and biased estimates of current stock status in terms of *B/BMSY*. These issues become apparent when we consider some of the macro-level predictions made by these models. While the estimates *B/BMSY* and other reference points reported in RAM are themselves model outputs subject to their own non-trivial errors and biases, a simple benchmark is to compare the estimates of fishery status from RAM to those predicted by potentially less reliable catch-only models intended for use when insufficient data are available for a full stock assessment model.

Costello et al. 2016 [5] found similar rankings of regions in terms of stock status as RAM, but their estimate of state of fisheries in the Mediterranean/Black Sea regions and Southeast Asia seem to be overly-optimistic, and the Northeast Pacific should be better by comparison. The superensemble method [10] used in Rosenberg et al. 2018 [7] demonstrates the same problem, with stocks in Southeast Asia estimated as doing better than the Northeast Atlantic or Northeast Pacific. The Pauly 2007 [6] catch-based approach finds the stocks of Southeast Asia in much better condition than those in the Northeast Pacific or Northeast Atlantic (Table.1). The non-RAM results in Table.1 include both formally assessed and unassessed stocks, and as such we would expect them to differ broadly in their estimates of regional stock status, particularly since we might expect unassessed stocks to be less rigorously managed and by extension have poorer stock status. However, the lack of consistency across heavily assessed regions such as the Northeast Pacific, and the lack of contrast in stock status between heavily and lightly managed regions is concerning.

Table 1: Estimates of B/BMSY by FAO, RAM Legacy Stock Assessment Database through Hilborn et al. 2020, Costello et al. 2016, Rosenberg et al. 2018, and Pauly 2007

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| FAO Area | FAO % Overfished | B/BMSY Costello | B/BMSY Rosenberg | Pauly - Catch/ Max(catch) | B/BMSY RAM |
| Pacific, Northeast | 0% | 1.18 | 1.02 | 0.31 | 1.61 |
| Atlantic, Northeast | 15% | 0.91 | 0.97 | 0.24 | 1.27 |
| Indian Ocean, Eastern | 18% | 0.94 | 1.04 | 0.75 | 1.01 |
| Pacific, Western Central | 22% | 0.89 | 1.06 | 0.69 | 1.46 |
| Mediterranean and Black Sea | 48% | 0.88 | 0.96 | 0.28 | 0.52 |

While each of these prior efforts at assessing global fisheries made important advances in our understanding of the global oceans, none have proven to be a consistently reliable means of estimating stock status of the worlds unassessed fisheries. In this paper we ask, can augmenting the types of catch-only models employed in global assessments to date with new broadly available data improve our understanding of the state of global fish populations? We use a flexible stock assessment software package, [sraplus](file:///Users/danovan/projects/assessing-global-fisheries/documents/www.github.com/danovando/sraplus) to demonstrate how different data sources can be used to augment catch-only models at a global scale, and to evaluate how our perception of global stock status would vary depending on which sources of data we include. We show that improving our understanding of the world’s unassessed fisheries depends on a redoubled effort at global data collection and synthesis.

# Methods

We compared estimates of stock status, defined by *B/BMSY* obtained from sraplus using various forms of broadly available data to the best-available estimates of *B/BMSY* available for those fish populations reported in the RAM database. All analysis were conducted in the R programming language [11]. Model fitting was conducted using Rcpp [12] and Stan [13] implemented through Template Model Builder [14] by the tmbstan package [15]. The sraplus package is publicly available at [github.com/danovando/sraplus](https://github.com/DanOvando/sraplus), and all materials needed to fully reproduce this manuscript are available at [github.com/DanOvando/assessing-global-fisheries](https://github.com/DanOvando/assessing-global-fisheries). We encourage readers to explore the documentation available at the package website at www.github.com/danovando/sraplus. Here we describe the structure of the population model underpinning sraplus, the estimation models used, and the construction of priors used in this paper.

## Data Sources

We restricted ourselves to databases with wide coverage around the world, representing the types of information that can be leveraged in efforts to assess the status of global fisheries (Table.2). While fishery-independent surveys are becoming increasingly available [16], they are not yet sufficiently distributed or accessible to serve as a foundation for global assessments of unassessed fisheries.

Table 2: Data sources included across model fits.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Source | Short Name | Data Use | Caveats |
| Catch data [2] | catches | Priors on stock status, scaling of population size, exploitation history | Heuristics or regressions used to translate shape of catch history into priors on stock status |
| Fisheries Management Index [17] | FMI | Priors on most recent F/FMSY values | Priors produced by regression trained on data from RAM Legacy Stock Assessment Database |
| Swept Area Ratio [18] | SAR | Priors on most recent F/FMSY values | Priors produced by regression trained on data from RAM Legacy Stock Assessment Database |
| Reconstructed effort data [19] | effort | Combined with catch data to create an index of abundance | Total reconstructed effort across all sectors. Assumed rate of technology creep reported in individual sections |

## Experiment Structure

We assessed the ability of Fisheries Management Index (FMI), swept area ratio (SAR), and effort data to improve estimates of the state of global unassessed fish populations. We based this test on 393 stocks from RAM, covering 19 broad taxonomic groups, with estimates of *B/BMSY* and greater than 25 years of continuous catch history. *B/BMSY* values from RAM are themselves estimates, not data, but they are the best available information on global stock status. We then paired the catch histories for these RAM stocks with regional-level SAR, FMI, and effort data. This process approximates a global-level assessment exercise, where data are available at regional levels, but not for specific fisheries.

We also estimated *B/BMSY* values of our candidate RAM stocks by using sraplus to fit to an abundance index drawn directly from RAM. We then fit a range of models utilizing different combinations FMI, SAR, and effort data, along with the CMSY method described in Froese et al. 2017 [20] (See Table.??). We assessed performance using three metrics: median percent error (MPE, a measure of bias), median absolute percent error (MAPE, a measure of accuracy), and classification accuracy. Classification accuracy was calculated as the proportion of times that use of a given combination of data resulted in a stock being classified into the correct FAO status classification (one of underfished, maximally sustainably fished, and overfished).

## Population Model

The core of sraplus is a Pella-Tomlinson [21] production model constructed in the manner of Winker *et al.* 2018 [22]. While models of these kinds abstract away many important details of fish biology and fleet behavior, they are the highest resolution model that the data types evaluated here will support. The purpose of sraplus is not to make substantial innovations in the fitting of surplus production models, but to provide a flexible tool for exploring the impacts of adding different kinds of data and priors on estimates of fishery status

The population growth equation is

Where is biomass at time *t*, *K* is carrying capacity ,*r* is the intrinsic growth rate, *m* is the scaling parameter that allows for the ratio of *BMSY/K* to shift. When *m* is two *BMSY / K* = 0.5, lower values of *m* shift the production function left, higher values right. is a vector catches, and is vector of process errors. Growth rates can become unrealistically large when the population reaches low sizes under the Pella-Tomlinson model. We deal with this problem by following the methods described in [22]. to reduce the production of the population when it falls below a threshold of 25% of carrying capacity.

We allow for process error *p* (in the manner of the stochastic stock reduction analysis suggested by Walters *et al.* 2006 [23]). This allows the population dynamics to deviate from the exact values given by the Pella-Tomlinson operating model, while still conforming to the assumptions of this model on average. Incorporation of process errors is useful for two reasons: (1) when you have an abundance index, process errors can reduce bias arising from lack of fit in a deterministic SRA whenever dynamics are poorly explained by catch-history alone, and (2) with or without an abundance index (or other info), the stochastic portion is necessary to get good uncertainty intervals (i.e., with close to nominal coverage, see [24]).

Process error *p* is assumed to be log-normally distributed, such that

## Estimation Model

All of our estimates are Bayesian in nature. sraplus can be run in two forms: either as a stock reduction analysis [SRA, 23], or fit to an index of abundance (fishery dependent or independent). Unless there is an abundance index to fit to the model runs as a stock reduction analysis. A stock reduction analysis works by specifying prior distributions on population parameters and critically the recent state of the fishery. sraplus allows users to specify the most recent status in units of depletion, B/BMSY, F, or F/FMSY. We then sample from the prior distributions of the population model parameters and apply those to the production model, along with the catch history. Any run that results in the collapse of the population (catch greater than biomass in any time step) is immediately rejected. The remaining viable draws from the prior distributions are sampled in proportion to the supplied prior on recent stock status. Any model results based on data sources listed in Table.?? that do not contain “cpue” or “RAM data” are estimated through stock reduction analysis. All SRA style runs in our paper sampled 2,000 draws of the prior-predictive distribution from a total of 1e6 candidate draws.

The full list of estimable parameters are listed in Table.3. *r* and *K* are the only two parameters that are always estimated. Estimation of every other parameter can be turned on or off. When estimation is turned off estimable parameters are fixed at their initial values, which can either be set to model defaults or specified by the user. For our main sets of results (everything excluding the value of information analysis), the estimated parameters are *r*, *K*, , , and *B0*. *q* is also estimated when needed.

Table 3: Name, abbreviations, and priors distribution for parameters potentially estimated by sraplus in this manuscript.

|  |  |  |
| --- | --- | --- |
| Parameter | Abbreviation | Default Prior |
| Carrying Capacity |  | Prior predictive tuning |
| Growth rate |  | Thorson, 2020 [25] updated by prior predictive tuning |
| Shape parameter |  | Drawn from Thorson et al. 2012 ([8]) |
| Catchability |  |  |
| Observation Error |  |  |
| Ratio of process to observation error |  |  |
| Initial State |  | Posterior probability dist. of catch-based regressions |

When an index of abundance is available the model estimates the posterior probability distributions of the estimated and transformed parameters using Hamiltonian Monte Carlo implemented in stan [26] accessed through the tmbstan interface [15]. By default the model uses 2000 draws with a 1000 step warm-up and one chain. Any detailed fit for a particular fishery would likely use more draws and chains, but we verified that this sampling routine produced an acceptable tradeoff of speed and convergence criteria. The model fits to a direct estimate of abundance (e.g. a fishery independent survey or a standardized catch-per-unit-effort index), the likelihood is

where *f* is the Pella-Tomlinson production model (Equation.(1)). When an effort index is available, sraplus constructs an index of abundance based on the catch and effort data. [27] measure an index of abundance as catch divided by their effort index, either nominal or effective (assuming the 2.6% annualized technology creep). This rate of technology creep assumes that every unit increase in effort is log-linearly greater than the unit of effort before it. When effort increases dramatically above historic levels, this can create a CPUE index that decreases faster than the true population. This is due to the fact that the marginal fishing mortality produced by increasing unit of efforts increases decreases as effort approaches infinity (since fishing mortality is bounded between 0 and 1). To accommodate this, we generate a catch per effective harvest rate index of abundance, as

Where can has a technology creep component

We then fit to the index of abundance per

### CMSY

In addition to the results from sraplus, we include a set of results produced by the default settings of the CMSY method [20]. For computational efficiency, we used a ported version of the CMSY model available at <https://github.com/DanOvando/portedcmsy>. The only modification made is to convert the underlying population model to C++ for faster computation. For each stock we used all the default options and priors provided and generated by CMSY, except for resilience, which was pulled from the vulnerability scores from FishBase accessed through rfishbase [28]. Vulnerability scores greater than 66 were scored as low resilience, between 33 and 66 medium resilience, and lower than 33 high resilience.

## Priors

The shape parameter is usually not reliably estimable given available data for surplus production models, however, Thorson et al. 2012 [8] provides estimate the ratio of *BMSY* to *K* for many fish taxa. While we estimate *m* by default throughout the results presented here, we use highly informative priors for the shape parameter based on Thorson et al. 2012 [8] for the genus of the species in question.

We address two critical features of prior use in sraplus below: tuning of the prior-predictive distribution and translation of outside data into priors usable by sraplus.

### Prior Predictive Tuning

Suppose that the only thing we observe from a fishery is a catch history. Assuming only Pella-Tomlinson population dynamics, the only thing we can learn from this catch history alone is the set of model parameters that ensure that the population still exists and never collapsed in the past. In the absence of any data to fit to, the a stock reduction algorithm (SRA) works by assuming that we know current stock status, and then finds feasible parameters to satisfy that belief given a catch history and model structure. This creates a problem for the Bayesian nature of our analysis. Consider a production model with two parameters, a growth rate *r* and a carrying capacity *K*. Once we specify prior distributions on *r* and *K*, and then apply these distributions to our model (the shape of the production function along with the catch histories), we have implicitly provided a prior on the status of the stock in all time periods, since each unique combination of *r* and *K* together with the model and the catch history produces a deterministic stock status in each time step. Doing so places two priors on recent stock status: one implicit prior through the population parameter priors, and one explicit through the users perception of recent stock status, creating a problem termed Borel’s Paradox (See Poole and Raftery 2000 [29] and references therein for a discussion of Borel’s Paradox in a fisheries context).

Borel’s Paradox poses a particular problem for the SRA version of sraplus. Due to the fact that there are more ways for a fishery to be relatively unexploited than for a fishery to be close to collapse but not collapsed, in this context Borel’s Paradox causes the posterior distribution of stock status to be positively biased relative to the supplied prior (although other modeling choices can result in a net negative bias in stock status, [9]). This process can also make it easy for users to accidentally supply very informative priors on stock status, without realizing that choices relating to population biology priors that may appear independent of stock status are in fact dictating the the posterior distributions of stock status resulting from the SRA algorithm.

We use an approximate solution to this problem here, similar in spirit to Bayesian melding [29]. Our solution amounts to a two-step sample-importance-resampling (SIR) algorithm. We first run the standard SIR algorithm as described above. We then break the resulting draws into bins based on terminal stock status, and calculate the mean sampling probability of each bin. The net result of this is that it allows users to place explicit prior on stock status, and then adjust their priors on life history parameters to reflect this prior, rather than creating a complicated and biased prior on stock status based on a mixture of explicit and implicit priors. See XX for a detailed explanation of this problem and our solution.

### Priors Informed by Outside Data

Along with allowing users to supply their own priors, the sraplus package contains three built-in methods for converting information on stock status from additional outside data into a form usable as a stock status prior by sraplus. We paired data on catch histories, swept area ratio (SAR), and fisheries management index (FMI) with estimates of stock status from the RAM legacy stock assessment database. We then trained a model of the general form for each of these three data types. Given values of these variables for a new fishery, sraplus uses the fitted model to generate posterior predictive distributions of stock status based on these data, which can then be used as priors on stock status by sraplus for new fisheries. For example, given data on SAR or FMI scores, together with a catch history, sraplus uses these regressions to convert those SAR and FMI values into priors on B/BMSY or F/FMSY in the most recent year of the fishery usable by sraplus. See XX for a detailed explanation of this process.

All prior regression models where tested by out-of-sample predictive power, and where competing models were considered the final model was chosen by leave-on-out validation [30]. The final models are intended as a reasonably robust means of translating available data (catch histories, FMI, and SAR values) into a form usable by sraplus. Given the scope of this analysis, we do not claim that the regressions used here are the best possible model relating these data with the fishery status indicators of interest. Rather, each regression was tested to ensure that it is unlikely that, given the same data, an alternative model would perform substantially better than those presented here.

## Value of Information Calculations

We performed a value-of-information (VOI) assessment to determine what types of data may be most beneficial to acquire at a global scale if we are to improve our knowledge of the state of global fisheries. The VOI analysis was performed by using sraplus to generate estimates of *B/BMSY* for stocks in RAM, and comparing the estimated values to the values reported in RAM. We generate fits for 3000 combinations of a RAM stock and available data. For any one draw, we randomly sample a RAM stock and a list of available data and data quality. For example, we might sample stock *A* with information on recent fishing mortality rates for the first iteration, and stock *A* again for the second iteration but now with information on recent fishing mortality rates and a recent index of abundance. The result is a set of model performance estimates where the characteristics of the stock and the data made available to the model are randomized.

# Results

## Case Study

We first present a case study demonstrating how different kinds of data can lead to different conclusions about stock status. From there, we assess the performance of models fit using different kinds of broadly available data: combinations of catches, effort, Fisheries Management Index scores, and swept area ratio (SAR) values (Table.2).

As a case study, we selected 26 stocks for which we have stock specific FMI and SAR scores. We then paired effort data at the resolution of year, country, and FAO statistical area from Rousseau et al. 2019 [27] to each stock. As a benchmark, we first estimated stock status for these case study fisheries using the CMSY [20] method, as this has become one of the most widely used catch-only models currently available. We then used stock-specific data on SAR and FMI to generate priors on F/FMSY for each of the stocks, which were then passed to sraplus. Lastly, we used the reconstructed effort data [27] to create an index of abundance for each stock, and estimated stock status by fitting to this index while using priors on fishing mortality rates informed by each stock’s FMI and SAR values. While CMSY systemically overestimated fishing mortality rates and underestimated stock status, use of the SAR, FMI, and effort data produced substantially more accurate results for both B/BMSY and F/FMSY (Fig.1).

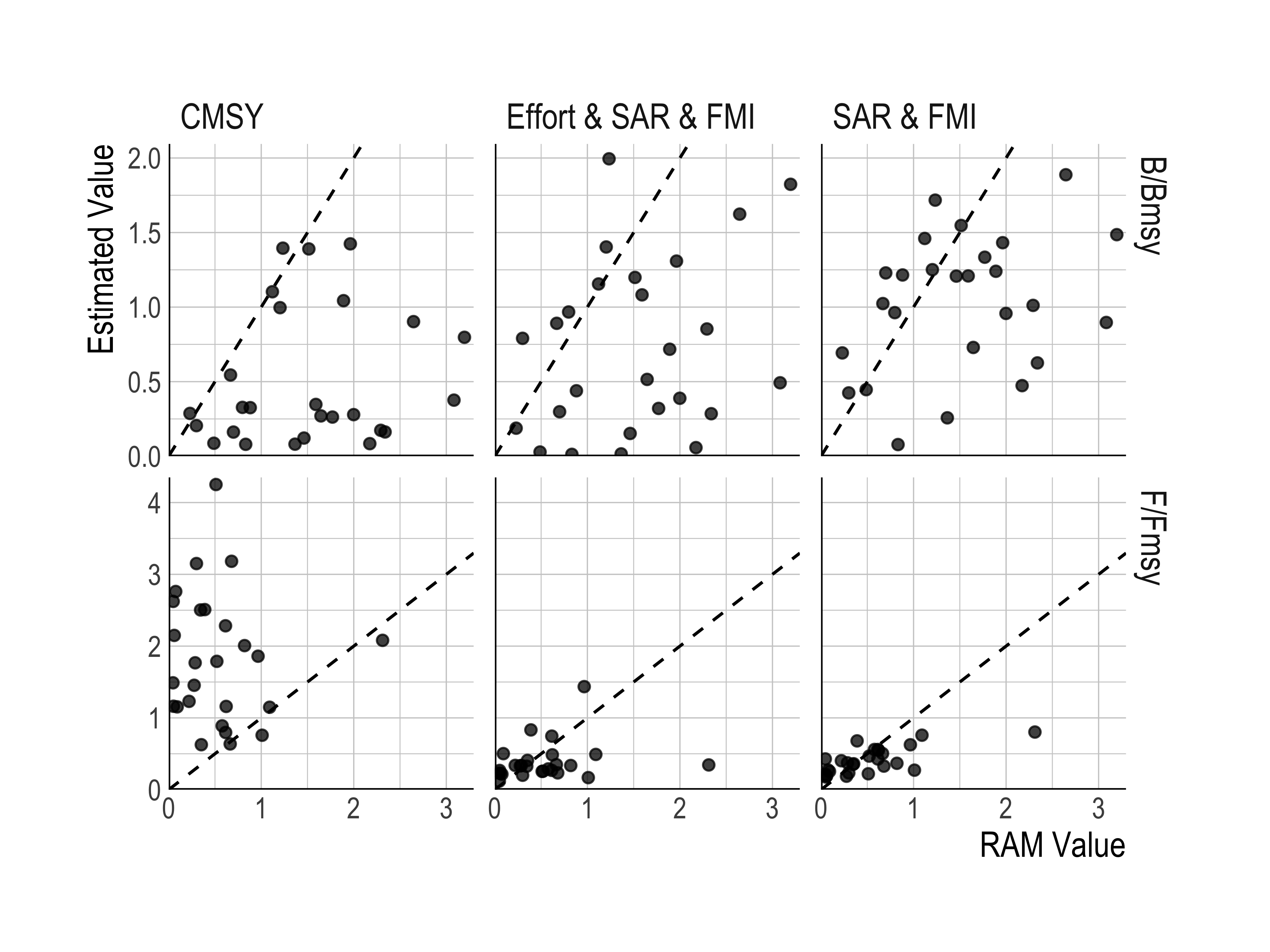


Figure 1: RAM values of B/BMSY and F/FMSY (x-axes) for case study fisheries plotted against estimated values (y-axes) using CMSY [20], priors informed by stock-specific Fisheries Management Index (FMI) and swept area ratio (SAR) scores, and an abundance index based on reconstructed effort trends assuming a rate of technological increase of 2.6%. Each point is a stock. Black dashed line shows the 1:1 relationship.

## Global Performance

We next assessed the ability of FMI, SAR, and effort data to improve estimates of global stock status. Overall the sraplus estimates of B/BMSY resulting from using the RAM data were relatively accurate and unbiased at a macro level (MPE ,MAPE , accuracy = , Table.4, Fig.2-4). This exercise tells us that given sufficiently high quality data, a surplus production model such as sraplus is reasonably capable of reproducing the global state of fisheries as understood from formally assessed fisheries.

Performance limitations in assessments of global fishery status then are likely to arise less from model misspecification than from the quality of the data themselves. This becomes clearer once we consider the performance of sraplus models fit to combinations of broadly available datasets. Many of the datasets used produced similar levels of bias as the RAM data (Table.4). However, this is somewhat an artifact of the data. The status of most stocks in RAM is also relatively good, with recent B/BMSY values generally near one. This means that a model that more or less reproduces the global average of stock status will be relatively unbiased on average, but imprecise. Focusing on MAPE instead, the error of the models jumps dramatically as soon as data other than RAM are used, to a minimum value of 47% and a maximum of 68%. The mean accuracy of the sraplus models across all non-RAM data fits was only 44%. Note that there are only three bins in the FAO stock status classifications, and a “model” that randomly assigns a stock to a status category would have a mean accuracy of .

Looking geographically we see a similar pattern of a rapid decrease in performance for models using non-RAM data intended to simulate a global assessment process. Across the models, performance was not consistent in space: use of different data performed best or worst for different FAO regions. For example, models fit to nominal CPUE data substantially overestimate stock status in the Mediterranean, while models based on data using effective CPUE perform better in that region (but worse in others) Fig.3. We find similarly inconsistent performance for both bias (Fig.2) and accuracy (Fig.4). Overall, while some data sources performed slightly better than others by some metrics in some places, no models using any non-RAM data were able to capture the overall state or geographic distribution of stock status represented in RAM in a consistently satisfactory manner.

Table 4: Global performance statistics in the most recent year available of models using different sources of data. mpe = median percent error (bias), mape = median absolute percent error (error), accuracy = percent of times that stocks were classified to the correct FAO status bin (underfished, maximally sustainably fished, overfished). Performance is judged relative to reported values in RAM Legacy Stock Assessment Database. See Table.5 for details of data types.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Used | mpe | mape | accuracy |
| RAM Index | 0.14 | 0.29 | 0.69 |
| FMI | -0.09 | 0.47 | 0.42 |
| SAR | -0.04 | 0.50 | 0.38 |
| Effective CPUE+ | -0.30 | 0.52 | 0.43 |
| Nominal CPUE+ | -0.01 | 0.52 | 0.46 |
| RAM U/Umsy | -0.09 | 0.58 | 0.40 |
| Guess | -0.16 | 0.60 | 0.32 |
| CMSY | -0.54 | 0.60 | 0.41 |
| Nominal CPUE | 0.05 | 0.63 | 0.48 |
| Effective CPUE | -0.36 | 0.68 | 0.41 |

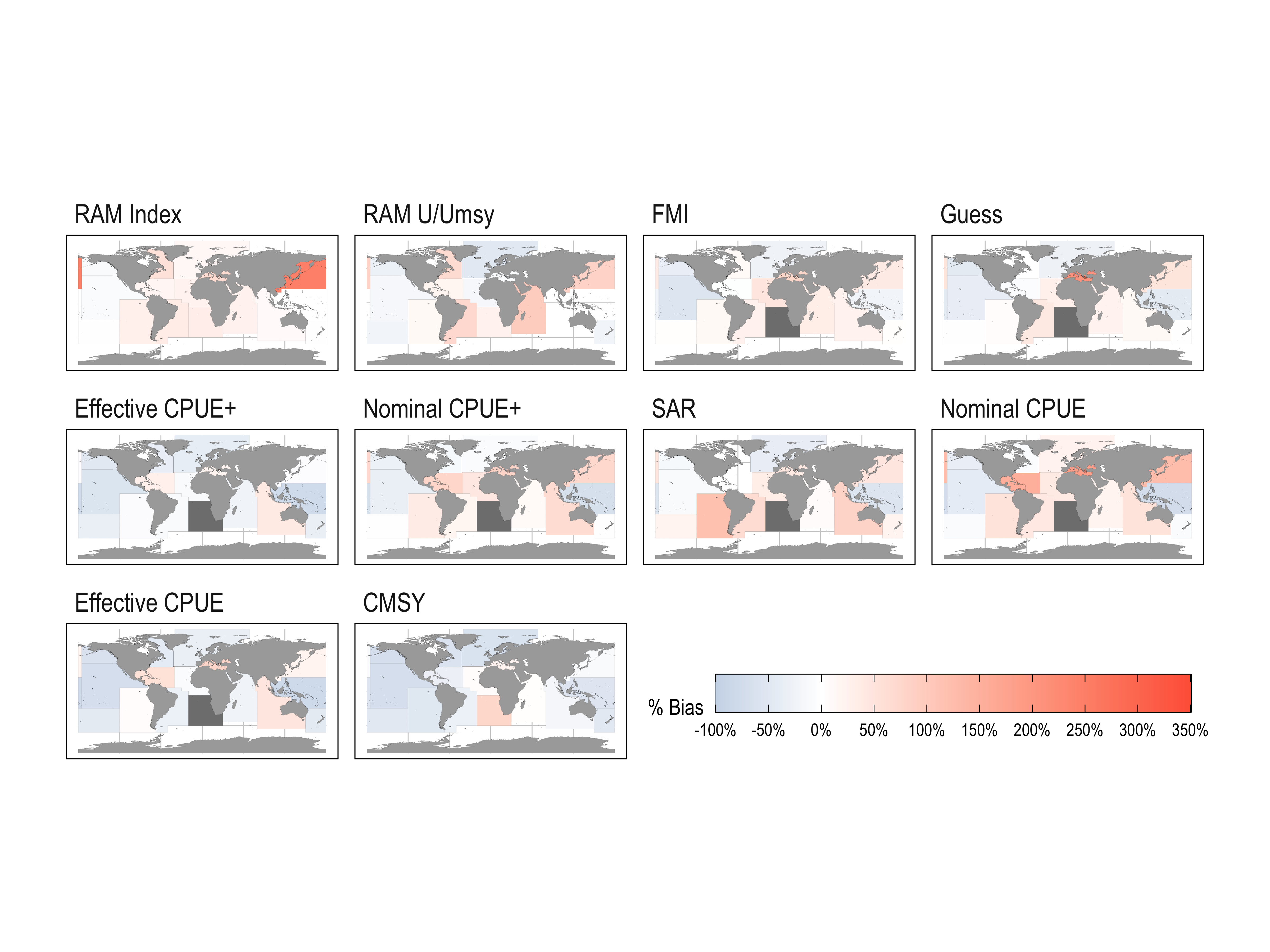


Figure 2: Median percent error in most recent B/BMSY by FAO statistical area from different data sources. RAM Index refers to catch and abundance index drawn from RAM. Effective CPUE refers to an index of abundance based on reconstructed effort data. Effective CPUE+ uses CPUE along with Fisheries Management Index (FMI) and/or swept area ratio (SAR) data. For both CPUE series ‘nominal’ assumes a 0% technology creep, for effective a 2.6% technology creep is assumed. RAM U/UMSY assumes all fisheries in the region share a common U/UMSY series with formally assessed fisheries in the region. FMI uses FMI scores to develop a prior on recent fishing mortality rates, SAR does the same but based on swept area ratio. CMSY uses the methods from Froese et al. 2017 [20]. Guess assignes a random recent B/BMSY of 0.4,1, or 1.6.

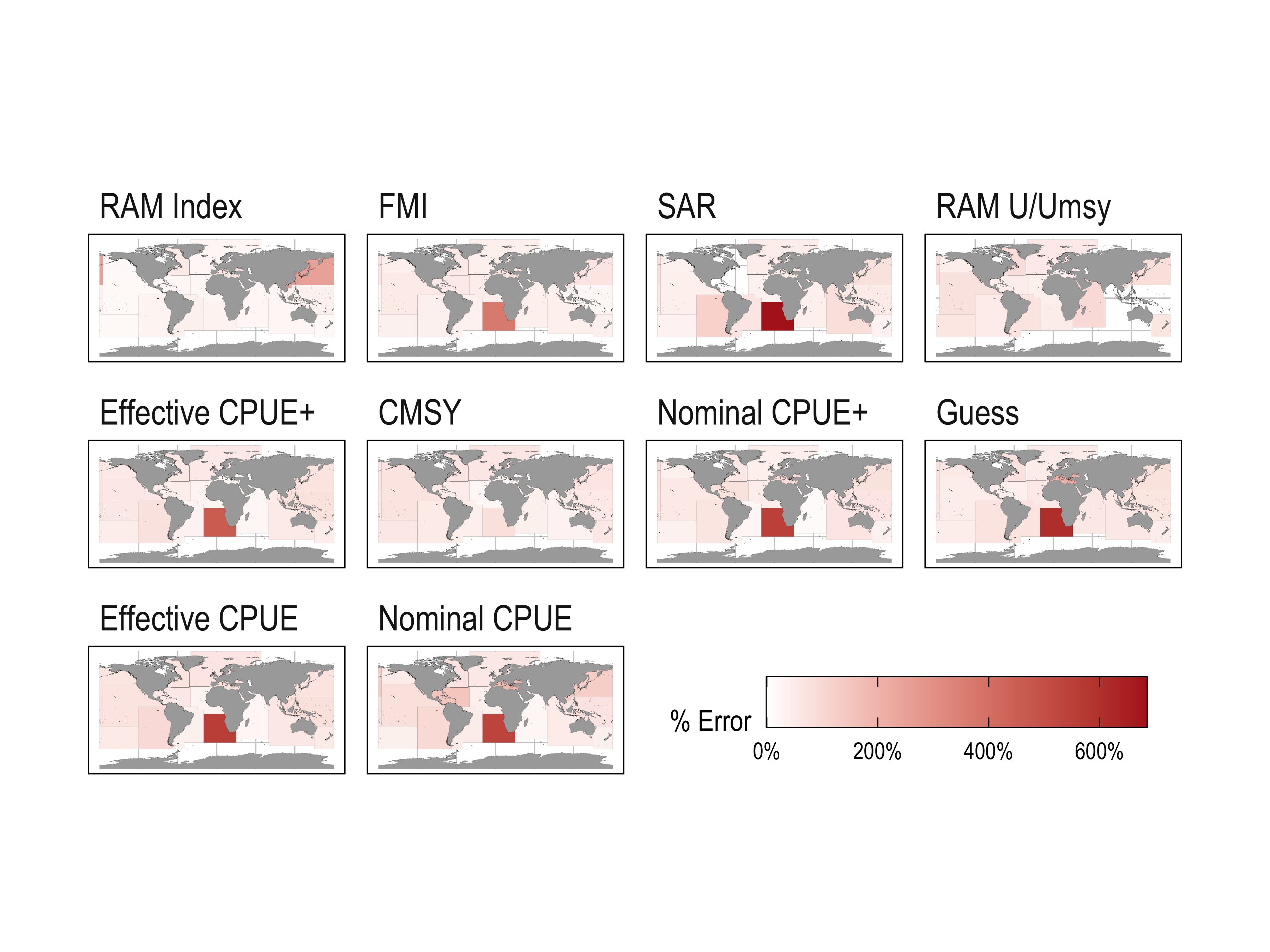


Figure 3: Median absolute percent error in most recent B/BMSY by FAO statistical area from different data sources. RAM Index refers to catch and abundance index drawn from RAM. Effective CPUE refers to an index of abundance based on reconstructed effort data. Effective CPUE+ uses CPUE along with Fisheries Management Index (FMI) and/or swept area ratio (SAR) data. For both CPUE series ‘nominal’ assumes a 0% technology creep, for effective a 2.6% technology creep is assumed. RAM U/UMSY assumes all fisheries in the region share a common U/UMSY series with formally assessed fisheries in the region. FMI uses FMI scores to develop a prior on recent fishing mortality rates, SAR does the same but based on swept area ratio. CMSY uses the methods from Froese et al. 2017 [20]. Guess assignes a random recent B/BMSY of 0.4,1, or 1.6.

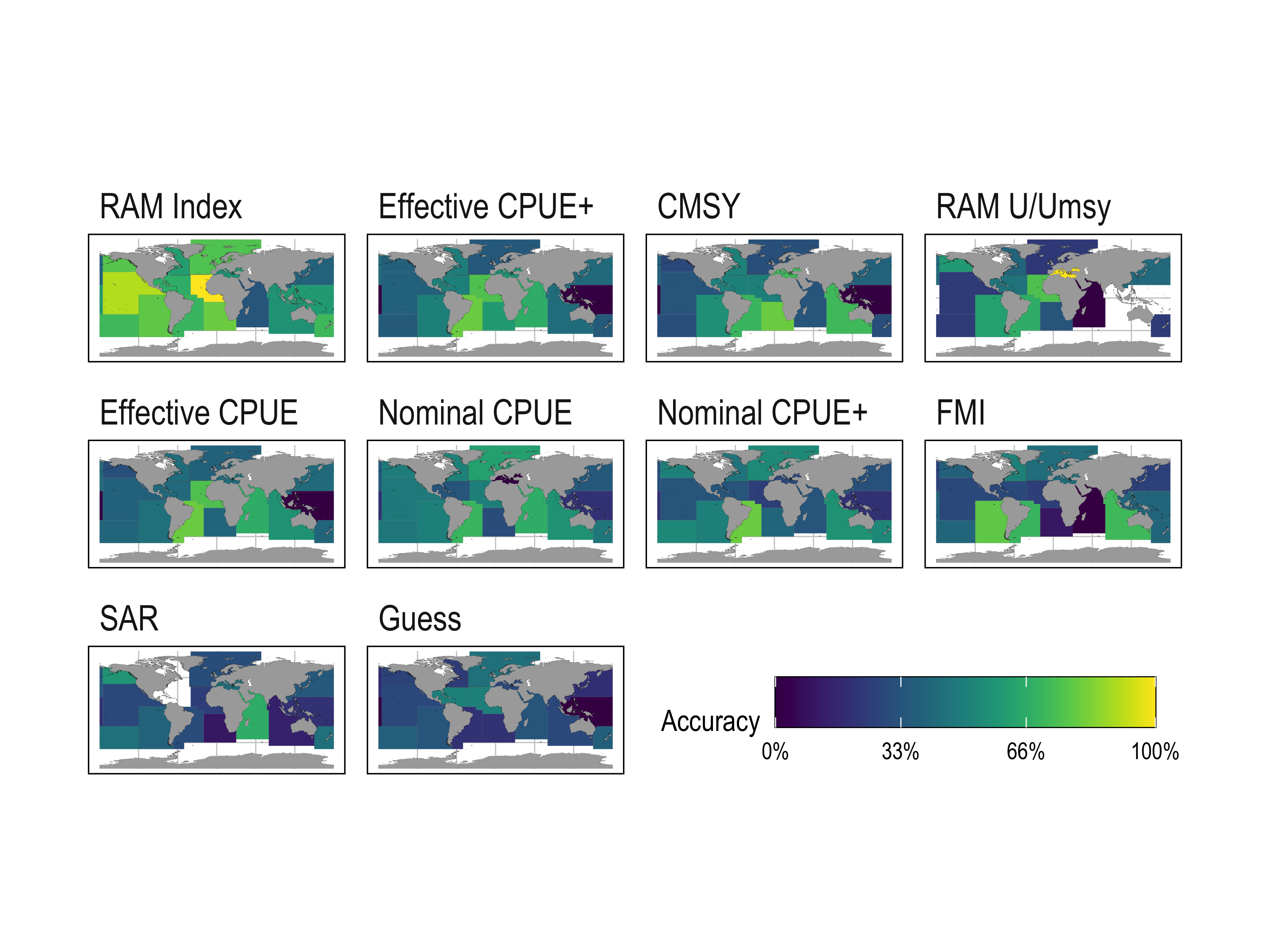


Figure 4: Mean classification accuracy (assignment to FAO stock status category) by FAO statistical area arising from different data sources. RAM Index refers to catch and abundance index drawn from RAM. Effective CPUE refers to an index of abundance based on reconstructed effort data. Effective CPUE+ uses CPUE along with Fisheries Management Index (FMI) and/or swept area ratio (SAR) data. For both CPUE series ‘nominal’ assumes a 0% technology creep, for effective a 2.6% technology creep is assumed. RAM U/UMSY assumes all fisheries in the region share a common U/UMSY series with formally assessed fisheries in the region. FMI uses FMI scores to develop a prior on recent fishing mortality rates, SAR does the same but based on swept area ratio. CMSY uses the methods from Froese et al. 2017 [20]. Guess assignes a random recent B/BMSY of 0.4,1, or 1.6.

The poor global performance of models fit to currently accessible broadly available datasets, relative to models fit to data from RAM, indicates that we must prioritize collection and aggregation of new or underutilized data sources if we are to improve our estimates of global fishery status. What sources of data might provide the greatest value in improving our estimates of global stock status? We used sraplus together with the RAM database to estimate the value of information, measured as the average reduction in root mean squared error, resulting from use of different kinds of data (Fig.5). Having access to estimates of *F/FMSY* reduced model error in proportion to the number of years for which *F/FMSY* values are available. Interestingly though, having access to only an accurate estimate of F/FMSY in the most recent year was extremely informative, reducing error on average by 15%, on par with an estimate of recent *B/BMSY* itself. While having access to complete index of abundance, such as a fishery independent survey, was on average able to reduce error relative to a baseline catch-only heuristic, using only the most recent quarter of the available abundance index actually increased error on average. We may have to wait many years for new surveys to provide substantial improvements in status estimates, or work to expand access to long-running existing surveys that have yet to be fully utilized in fisheries assessment [16].



Figure 5: Posterior probability distributions of estimated effect of different data types on root mean squared error of B/BMSY in the most recent 5 years of data available for each model fit. Distribution is full posterior probability distribution. Point is median, thicker black section inner 66th quantile of the posterior, the thinner black line the 95th. Change is relative to the mean performance of a catch-only heuristic model.

# Discussion

Global-level assessments of fish populations are critical for guiding management agendas for the world’s oceans, and tracking indicators such as the United Nations Sustainable Development Goals. Despite this need, and despite advances in stock assessment methods and available data, we show that our understanding of the world’s fish stocks remains uncertain in many parts of the world. The hope of efforts using catch-only models to estimate the status of global unassessed fisheries is fundamentally that characteristics of a fishery’s catch history and life history are informative as to the status its population. Our results show that estimates of stock status based on catch histories alone stem almost entirely from underlying assumptions, not the information content of the catches, producing results that are inconsistent with best-available knowledge of the state of assessed fish populations in RAM. While in some cases addition of globally available data, such as quality of fisheries management or effort reconstructions, provided value above and beyond catch histories alone (Fig.1), at the global level models fit using each of the available datasets, besides the RAM-derived indices, generally produced biased and imprecise estimates of fish stock status (Table.4).

What quality of assessment is needed and what constitutes a meaningful improvement in assessment quality depends on the needs of those using the assessment outputs. It may be that for particular regions, species, or uses the results presented here or in other past global analyses are sufficient. In some instances using the data presented here did provide some improvement over use of catch-only style assessment methods; the difficulty comes in attempting to apply data types uniformly across the globe. While it is unreasonable to expect models based solely on global-scale data to be able to perform as well as detailed stock assessments reported in RAM, or that data-limited methods would perform well for every individual stock, our hope would be that a data-limited approach based on globally available data sources would be able to correctly capture general patterns in stock status in time and space. That use of none of the datasets collected here can achieve that, and that our test on the RAM data suggest that model misspecification is not the primary culprit, tells us that improvements in estimates of global stock status must come from improvements in the quality and use of data themselves.

We chose to test the performance of methods against estimated values reported in RAM. A reasonable critique of this choice is that unassessed stocks, on which these methods would actually be used, are likely to have vastly different dynamics than the heavily managed fish populations represented in RAM. The clear and systemically poor performance of the methods tested here when applied to RAM stocks would require though that these methods have massively lower bias and higher accuracy for unassessed fisheries than RAM stocks if they are to be expected to provide reasonable picture of global fish populations. Free et al. 2020’s [9] simulation testing of the types of methods included here suggests that this is unlikely to be true.

Our results do not imply that the kinds of broadly available data presented here are not valuable under the right conditions. The FMI and SAR based priors are an improvement over catch-only models in applicable situations (i.e. those that sufficiently resemble the data on which the regressions were trained, Fig.1). Effort data such as those reconstructed by Rousseau et al. 2019 [27] can help distinguish between regions with similar catch histories but different large-scale effort trajectories, and may be quite useful as indices of abundance for areas with good knowledge of rates of evolution of fishing technology and a broadly selective fishing fleet. Despite not adding a great deal in terms of performance at the global scale, swept-area-ratio was the strongest predictor of F/FMSY of any of the datasets we explore on an individual stock basis, with a Bayesian R2 value of 0.43.

But, we must simultaneously consider data quality and resolution: applying one SAR value to all stocks in a region, even if that SAR value can provide valuable information for a subset of fisheries, causes inaccurate estimates of stock status when applied too broadly. Our analysis does not show that the data considered here are without value, but that attempting to indiscriminately apply these data to all areas of the globe results in meaningfully incorrect estimates of stock status for regions whose nature does not match the assumptions needed to apply these data sources.

Our value-of-information analysis also shows though the high utility of having access to even a recent snapshot of F/FMSY (Fig.5). Swept area ratios, Fisheries Management Index scores, or other similar metrics can be used to construct fishery-specific priors on fishing mortality rates, though care must be taken in applying them at the appropriate spatial resolution. Another avenue would be to prioritize the development of a global repository for length and age composition data. Given appropriate conditions, these length measurements can be used to estimate local fishing mortality rates [31–33]. While length-based assessments come with a host of assumptions and pitfalls, properly implemented in some fisheries with appropriate life histories these methods may provide an overlooked source of information on fisheries at a global scale, at least as an improvement over relying on catch-only or regional proxies alone. Such a database could be used to construct stock or stock complex specific priors on fishing mortality for particular regions around the globe, which could meaningfully improve our understanding of global fisheries, particularly when paired with catch data and where possible indices of abundance [33,34].

We must also prioritize collection and curation of fish population survey data worldwide. Repositories of fishery-independent survey data would be immensely beneficial, such as those maintained by [FishStat](https://james-thorson.github.io//). Recent research confirms that there are bottom trawl data to support analysis of biomass-trends since 2001 and potentially earlier in many regions [16], and survey data are available for more stocks than have stock assessments. Effort reconstructions such as those utilized here may help create fishery-dependent abundance indices in some instances, and going forward datasets such as those compiled by [Global Fishing Watch](https://globalfishingwatch.org/) in combination with the reconstruction approaches of [27] might allow us to construct and use timeseries of fishing effort specific to particular areas, fleets, and species complexes

Expanded training of fisheries scientists around the globe is another critical need. Even were we to dramatically expand the amount and types of data available for global assessment, individual fisheries and regions will need to make informed decisions about which sources of data may be applicable and which not, and to critically evaluate the results of any model based on local expertise. This is why stock assessments even in data-rich fisheries are not an automated process; the real challenge is often not in fitting a model to data but in understanding how best to use the data and the quality and limitations of the model used. Empowering a global network of fisheries scientists through training and peer-support would help local experts make the most of available data, ensure the reliability of newly collected data, and improve the interpretation of assessment results.

# Conclusions

The coming decades are a critical time for the future of fisheries and ocean health. Achieving the United Nations Sustainable Development Goal 14 for the conservation and sustainable use of the world’s oceans depends on our ability to effectively assess the status of fish stocks around the world. The RAM Legacy Stock Assessment Database combined with the FAO’s expert elicitation of status for select stocks have dramatically improved our understanding of global fisheries in recent years. However, this process still leaves a substantial number of fisheries and proportion of global catch unassessed. Numerous catch-based data-limited approaches have attempted to fill that gap, and while these efforts have advanced our knowledge and interest in unassessed fisheries, none have yet been able to provide a clear solution to this problem.

The lack of strong information on stock status within catch histories demonstrated here means that differences in models and assumptions between catch-based assessment efforts can produce starkly contrasting conclusions on global stock status, leading to debates that are inconclusive as they are inherently driven by assumptions. The FAO is leading efforts to increase technical capacity and monitoring and evaluation infrastructure to improve fisheries management in places with limited data. Such projects stand to provide a better picture of fishery status at global and local scales, furthering our ability to meet the UN SDG targets. Our results emphasize the urgency and rationale for building the infrastructure and capacity that can lead to better marine resource management globally. Catch-only methods may be a useful starting point when applied properly, but achieving meaningful improvements in the assessment and management of global unassessed fisheries will depend on expanded collection of targeted data types, active management, and local capacity building.

# References

1. Hilborn R *et al.* 2020 Effective fisheries management instrumental in improving fish stock status. *Proceedings of the National Academy of Sciences* (doi:[10.1073/pnas.1909726116](https://doi.org/10.1073/pnas.1909726116))

2. FAO. 2020 *STATE OF WORLD FISHERIES AND AQUACULTURE 2020: Sustainability in action.* S.l.: FOOD & AGRICULTURE ORG.

3. Ricard D, Minto C, Jensen OP, Baum JK. 2012 Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish and Fisheries* **13**, 380–398. (doi:[10.1111/j.1467-2979.2011.00435.x](https://doi.org/10.1111/j.1467-2979.2011.00435.x))

4. Costello C, Ovando D, Hilborn R, Gaines SD, Deschenes O, Lester SE. 2012 Status and Solutions for the World’s Unassessed Fisheries. *Science* **338**, 517–520. (doi:[10.1126/science.1223389](https://doi.org/10.1126/science.1223389))

5. Costello C *et al.* 2016 Global fishery prospects under contrasting management regimes. *Proceedings of the National Academy of Sciences* **113**, 5125–5129. (doi:[10.1073/pnas.1520420113](https://doi.org/10.1073/pnas.1520420113))

6. Pauly D. 2007 The Sea Around Us Project: Documenting and Communicating Global Fisheries Impacts on Marine Ecosystems. *AMBIO: A Journal of the Human Environment* **36**, 290–295. (doi:[10.1579/0044-7447(2007)36[290:TSAUPD]2.0.CO;2](https://doi.org/10.1579/0044-7447(2007)36%5b290:TSAUPD%5d2.0.CO;2))

7. Rosenberg AA *et al.* 2018 Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries around the World. *Conservation Letters* **11**, e12363. (doi:[10.1111/conl.12363](https://doi.org/10.1111/conl.12363))

8. Thorson JT, Cope JM, Branch TA, Jensen OP. 2012 Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. *Canadian Journal of Fisheries and Aquatic Sciences* **69**, 1556–1568. (doi:[10.1139/f2012-077](https://doi.org/10.1139/f2012-077))

9. Free CM, Jensen OP, Anderson SC, Gutierrez NL, Kleisner KM, Longo C, Minto C, Osio GC, Walsh JC. 2020 Blood from a stone: Performance of catch-only methods in estimating stock biomass status. *Fisheries Research* **223**, 105452. (doi:[10.1016/j.fishres.2019.105452](https://doi.org/10.1016/j.fishres.2019.105452))

10. Anderson SC *et al.* 2017 Improving estimates of population status and trend with superensemble models. *Fish and Fisheries* **18**, 732–741. (doi:[10.1111/faf.12200](https://doi.org/10.1111/faf.12200))

11. R Core Team. 2019 R: A Language and Environment for Statistical Computing.

12. Eddelbuettel D, François R. 2011 Rcpp: Seamless R and C++ integration. *Journal of Statistical Software* **40**, 1–18. (doi:[10.18637/jss.v040.i08](https://doi.org/10.18637/jss.v040.i08))

13. Carpenter B *et al.* 2017 *Stan* : A Probabilistic Programming Language. *Journal of Statistical Software* **76**. (doi:[10.18637/jss.v076.i01](https://doi.org/10.18637/jss.v076.i01))

14. Kristensen K, Nielsen A, Berg CW, Skaug H, Bell BM. 2016 **TMB** : Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* **70**. (doi:[10.18637/jss.v070.i05](https://doi.org/10.18637/jss.v070.i05))

15. Monnahan C, Kristensen K. 2018 No-u-turn sampling for fast bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan r packages. *PloS one* **13**. (doi:[10.1371/journal.pone.0197954](https://doi.org/10.1371/journal.pone.0197954))

16. Maureaud A *et al.* 2020 Are we ready to track climate-driven shifts in marine species across international boundaries? - A global survey of scientific bottom trawl data. (doi:[10.1101/2020.06.18.125930](https://doi.org/10.1101/2020.06.18.125930))

17. Melnychuk MC, Peterson E, Elliott M, Hilborn R. 2017 Fisheries management impacts on target species status. *Proceedings of the National Academy of Sciences* **114**, 178–183.

18. Amoroso RO *et al.* 2018 Bottom trawl fishing footprints on the world’s continental shelves. *Proceedings of the National Academy of Sciences*, 201802379. (doi:[10.1073/pnas.1802379115](https://doi.org/10.1073/pnas.1802379115))

19. Rousseau Y, Watson RA, Blanchard JL, Fulton EA. 2019 Evolution of global marine fishing fleets and the response of fished resources. *Proceedings of the National Academy of Sciences*, 201820344. (doi:[10.1073/pnas.1820344116](https://doi.org/10.1073/pnas.1820344116))

20. Froese R, Demirel N, Coro G, Kleisner KM, Winker H. 2017 Estimating fisheries reference points from catch and resilience. *Fish and Fisheries* **18**, 506–526. (doi:[10.1111/faf.12190](https://doi.org/10.1111/faf.12190))

21. Pella JJ, Tomlinson PK. 1969 A generalized stock production model. *Inter-American Tropical Tuna Commission Bulletin* **13**, 416–497.

22. Winker H, Carvalho F, Kapur M. 2018 JABBA: Just Another Bayesian Biomass Assessment. *Fisheries Research* **204**, 275–288. (doi:[10.1016/j.fishres.2018.03.010](https://doi.org/10.1016/j.fishres.2018.03.010))

23. Walters CJ, Martell SJD, Korman J. 2006 A stochastic approach to stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences* **63**, 212–223. (doi:[10.1139/f05-213](https://doi.org/10.1139/f05-213))

24. Thorson JT, Rudd MB, Winker H. 2018 The case for estimating recruitment variation in data-moderate and data-poor age-structured models. *Fisheries Research* (doi:[10.1016/j.fishres.2018.07.007](https://doi.org/10.1016/j.fishres.2018.07.007))

25. Thorson JT. 2020 Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. *Fish and Fisheries* **21**, 237–251. (doi:[10.1111/faf.12427](https://doi.org/10.1111/faf.12427))

26. Stan Development Team. 2018 {{{}: The {}R{} interface to {}Stan{}{}.

27. Rousseau Y, Watson RA, Blanchard JL, Fulton EA. 2019 Evolution of global marine fishing fleets and the response of fished resources. *Proceedings of the National Academy of Sciences*, 201820344. (doi:[10.1073/pnas.1820344116](https://doi.org/10.1073/pnas.1820344116))

28. Boettiger C, Temple Lang D, Wainwright P. 2012 Rfishbase: Exploring, manipulating and visualizing FishBase data from r. *Journal of Fish Biology*

29. Poole D, Raftery AE. 2000 Inference for Deterministic Simulation Models: The Bayesian Melding Approach. *Journal of the American Statistical Association* **95**, 1244–1255. (doi:[10.1080/01621459.2000.10474324](https://doi.org/10.1080/01621459.2000.10474324))

30. Vehtari A, Gelman A, Gabry J. 2017 Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**, 1413–1432. (doi:[10.1007/s11222-016-9696-4](https://doi.org/10.1007/s11222-016-9696-4))

31. Hordyk A, Ono K, Prince JD, Walters CJ. 2016 A simple length-structured model based on life history ratios and incorporating size-dependent selectivity: Application to spawning potential ratios for data-poor stocks. *Canadian Journal of Fisheries and Aquatic Sciences* (doi:[10.1139/cjfas-2015-0422](https://doi.org/10.1139/cjfas-2015-0422))

32. Prince J, Hordyk A. 2019 What to do when you have almost nothing: A simple quantitative prescription for managing extremely data-poor fisheries. *Fish and Fisheries* **20**, 224–238. (doi:[10.1111/faf.12335](https://doi.org/10.1111/faf.12335))

33. Rudd MB, Thorson JT. 2017 Accounting for variable recruitment and fishing mortality in length-based stock assessments for data-limited fisheries. *Canadian Journal of Fisheries and Aquatic Sciences*, 1–17. (doi:[10.1139/cjfas-2017-0143](https://doi.org/10.1139/cjfas-2017-0143))

34. Thorson JT, Cope JM. 2015 Catch curve stock-reduction analysis: An alternative solution to the catch equations. *Fisheries Research* **171**, 33–41. (doi:[10.1016/j.fishres.2014.03.024](https://doi.org/10.1016/j.fishres.2014.03.024))

35. Thorson JT, Branch TA, Jensen OP. 2012 Using model-based inference to evaluate global fisheries status from landings, location, and life history data. *Canadian Journal of Fisheries and Aquatic Sciences* **69**, 645–655. (doi:[10.1139/f2012-016](https://doi.org/10.1139/f2012-016))