

Global Assessments of Fishery Status need Better Data more than Better Models

Daniel Ovando Ray Hilborn Cole Monnahan Merrill Rudd Rishi Sharma
James Thorson Yannick Rousseau Yimin Ye

2020-10-06

Abstract

Assessments of the global state of fisheries 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 49% of global catch, our knowledge of the state of the remaining 51%, the worlds ‘unassessed’ fisheries, is poor. 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 fisheries remains poor, 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 at a global scale, not on the development of new modeling methods alone.

Contents

1	Publication Requirements	2
2	Introduction	2
3	Results	3
4	Discussion	6
5	Methods	11
5.1	Population Model	11
5.2	Estimation Model	12
5.3	Priors	14
5.4	Value of Information Calculations	17
6	References	18
7	Supplemental Material	19

1 Publication Requirements

Target Journal: [Nature Sustainability](#)

3500 words main text

50 references (main text)

3000 words (max) methods

2 Introduction

The United Nations Sustainable Development Goal 14 (SDG 14), related to “Life under water”, calls for the global community to “Conserve and sustainably use the oceans, seas and marine resources for sustainable development”. Meeting these targets depends in part on our ability to effectively track the status of global marine fish stocks. The Food and Agriculture Organization of the United Nations’ (FAO) State of World Fisheries and Aquaculture (SOFIA) report is the standard source for tracking the global state of fisheries. As of the most recent report, the FAO estimates that as of 2017 59.6% of marine fish stocks are maximally sustainably fished, 6.2% are underfished, and 34.2% are overfished.¹ While foundational, the SOFIA assessment was designed in the 1970s based on the then available data and methods. With the surge in data availability and models designed for data-limited stocks, new global assessment methods are needed to meet the new demand for estimation and tracking progress towards the SDG goals.

Forty nine percent of landings of marine fishes reported by the FAO are represented in the RAM Legacy Stock Assessment Database, a repository of “gold standard” estimates of fishery status.² The SOFIA process uses these formal assessments for their determinations of stock status wherever possible. However, that leaves 51% of global fisheries landings lacking in formal stock assessments. While these “unassessed” stocks are generally individually smaller than the typically larger and more valuable stocks in the assessed category, collectively they are a vital source of food, employment, cultural value, and ecosystem services around the world.

The SOFIA report bases its estimates of included unassessed stocks mostly on expert opinion. While local experts can be well informed as to the status of their fish stocks, a more quantitative and reproducible process would be desirable. Numerous studies in recent years have put forward versions of “data-limited” models that have attempted to provide estimates for the global status of these unassessed stocks.³⁻⁷ Due to data limitations, all of these global assessment efforts use forms of “catch-only” stock assessment models (Free et al. 2020⁸ and references therein). These models seek to infer stock status, for example in terms of biomass B relative to the biomass at maximum sustainable yield B_{MSY} , from characteristics of a fishery’s catch history, for example the ratio of catch to maximum catch.⁹

However, Free et al. 2020⁸ demonstrated that these catch-only models can often produce both imprecise and biased estimates of current stock status in terms of B/B_{MSY} . These issues become apparent when we consider some of the macro-level predictions made by these models. The RAM Legacy Stock Assessment Database¹⁰ contains the best available estimates of B/B_{MSY} and other fishery reference points for hundreds of fisheries.² While it must be noted that the estimates in RAM are themselves model outputs subject to their own non-trivial errors and biases, a simple benchmark is to compare the best available estimates of fishery status from RAM to those predicted by potentially less reliable methods intended for use when insufficient data are available for a full stock assessment model.

Costello et al. 2016⁴ finds 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 over-optimistic, and the Northeast Pacific should be better by comparison. Rosenberg et al. 2018⁶ demonstrates the same problem, with stocks in Southeast Asia estimated as doing better than the Northeast Atlantic or Northeast Pacific. The Pauly 2007⁵ catch based approach finds the stocks of Southeast Asia in much better condition than the Northeast Pacific or Northeast Atlantic (Table.1). The methods besides RAM 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/B_{MSY} by FAO, RAM through Hilborn et al. 2020, Costello et al. 2016, Rosenberg et al. 2018, and Pauly 2007

FAO Area	FAO % Overfished	B/B_{MSY} Costello	B/B_{MSY} Rosenberg	Pauly - Catch/ Max(catch)	B/B_{MSY} 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

In this paper we ask, can combining the FAO’s catch statistics with other broadly available data improve our understanding of the state of global fisheries? We use a flexible surplus-production based stock assessment package, **srplus** to demonstrate how different sources of data 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 our understanding of global fishery status is poor, and that improvements depend on an redoubled effort at global data collection.

3 Results

Table 2: Broadly available data used to inform estimates of global stock status

Data Source	Short Name	Data Use	Caveats
Catch data ¹	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 ¹¹	FMI	Priors on most recent F/F_{MSY} values	Priors produced by regression trained on data from RAM Legacy Stock Assessment Database
Swept Area Ratio ¹²	SAR	Priors on most recent F/F_{MSY} values	Priors produced by regression trained on data from RAM Legacy Stock Assessment Database
Reconstructed effort data ¹³	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

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 (FMI) scores (a measure of management capacity), and swept area ratio (SAR) values (a measure of trawl fishing intensity) (Table.2).

For our 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. 2019¹⁴ to each stock. We first used the catch history heuristics internal to CMSY¹⁵ to estimate stock status. We then used stock-specific data on SAR and FMI to generate priors on F/F_{MSY} for each of the stocks, which were then passed to `sraplus`. Lastly, we used the reconstructed effort data from¹⁴ 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 (Fig.1).

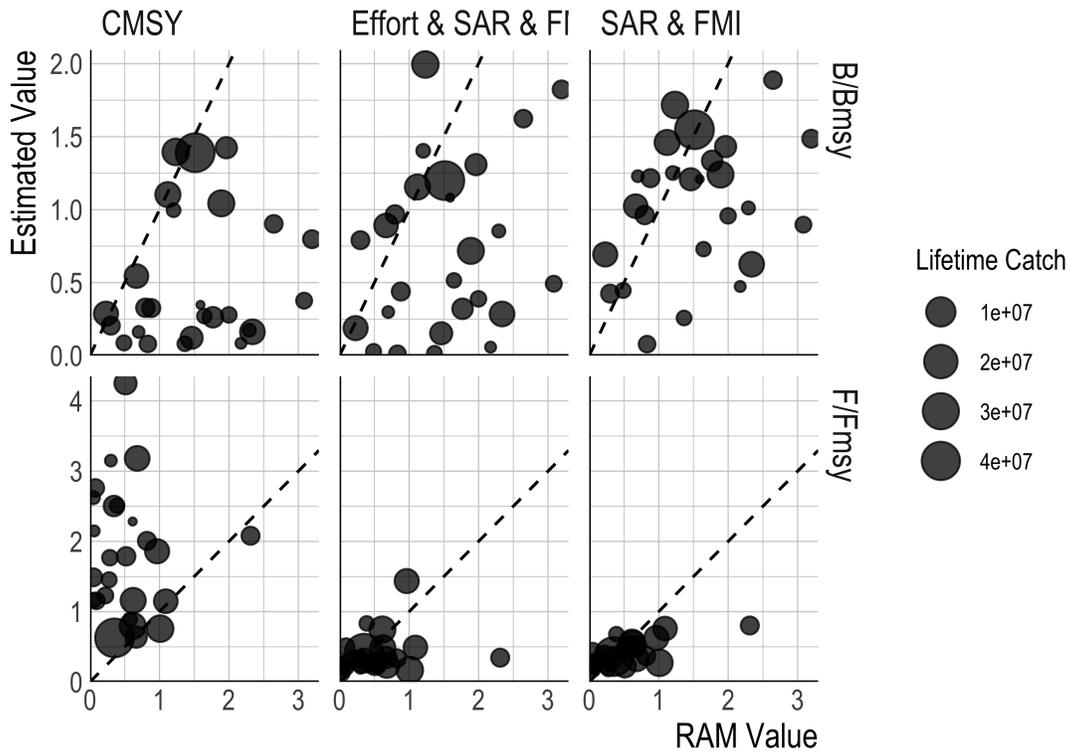


Figure 1: RAM values of B/B_{MSY} and F/F_{MSY} (x-axes) for case study fisheries plotted against estimated values (y-axes) using CMSY,¹⁵ priors informed by stock-specific Fisheries Management Index (FMI) and swept area ratio (SAR) scores, and an abundance index based on reconstructed effort trends assuming a rate of technological increase of 2.6%. Each point is a stock, point size is a function of stock size. Black dashed line shows the 1:1 relationship.

We next assessed the ability of FMI, SAR, and effort data to improve estimates of global stock status. We based this test on 393 fisheries from RAM, covering 19 broad taxonomic groups, with estimates of B/B_{MSY} and greater than 25 years of continuous catch history. B/B_{MSY} 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.

As a proof of concept, we also estimated B/B_{MSY} 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 catch-only method described in Froese et al. 2017¹⁵ (See Table.S1). 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 is 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).

Overall the `sraplus` estimates of B/B_{MSY} resulting from using the RAM data are reasonably good (median absolute percent error 29%, accuracy = 69%, Table.3, 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 then are likely to arise less from model misspecification than from the quality of the data themselves. These becomes clearer once we consider the performance of `sraplus` models fit to combinations of our broadly available datasets. Many of the datasets used produced similar levels of bias as the RAM data (Table.3). However, this is somewhat an artifact of the data. The status of most stocks in RAM is also relatively good, with recent B/B_{MSY} 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 72%. The mean accuracy across all non-RAM data fits was only 41%. 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 has a mean accuracy of 38%.

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 3: 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.

data	mpe	mape	accuracy
ram-data	0.14	0.29	0.69
fmi	-0.09	0.47	0.42
sar	-0.04	0.50	0.38
cpue-plus	-0.30	0.52	0.43
nominal-cpue-plus	-0.01	0.52	0.46
Guess	-0.10	0.53	0.33
u_umsy	-0.09	0.58	0.40
cmsy	-0.54	0.60	0.41
nominal-cpue	0.05	0.63	0.48
cpue	-0.36	0.68	0.41
guess	-0.09	0.72	0.38

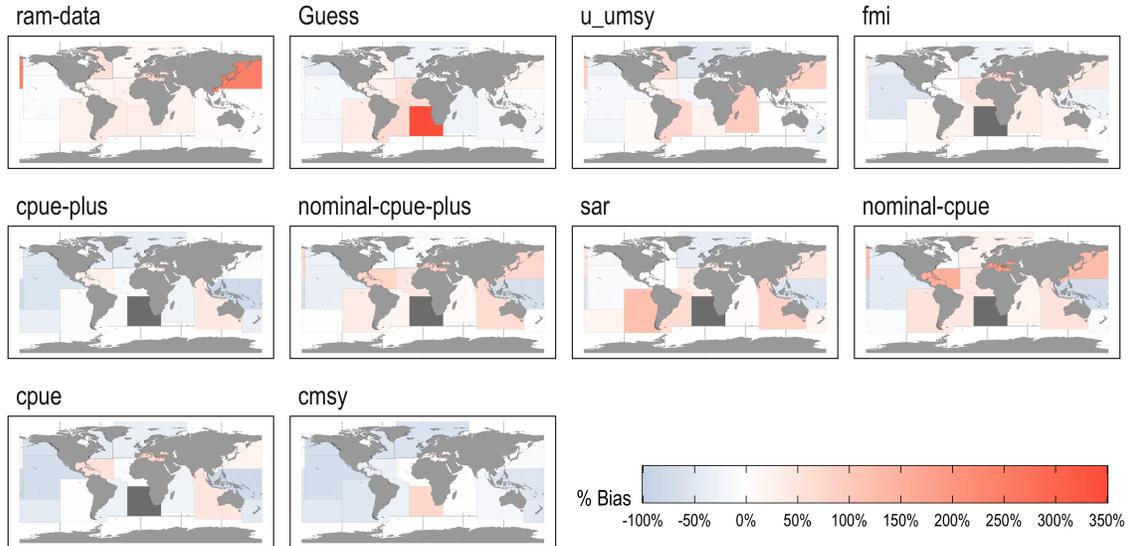


Figure 2: Median percent error in most recent B/B_{MSY} by FAO statistical area from different data sources. ram-data refers to catch and abundance index drawn from RAM. CPUE refers to an index of abundance based on reconstructed effort data. cpue-plus uses CPUE along with Fisheries Management Index (FMI) and/or wept area ratio (SAR) data. For both CPUE series ‘nominal’ assumes a 0% technology rate, otherwise a 2.6% technology rate 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. 2017.¹⁵

4 Discussion

Global-level assessment are critical for guiding management agendas for the world’s oceans, and tracking critical 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 fisheries remains murky in many parts of the world. 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, produced biased and imprecise estimates of stock status, frequently performing worse than a simple guess (Table.3).

What quality of assessment is needed and what constitutes a meaningful improvement in assessment quality

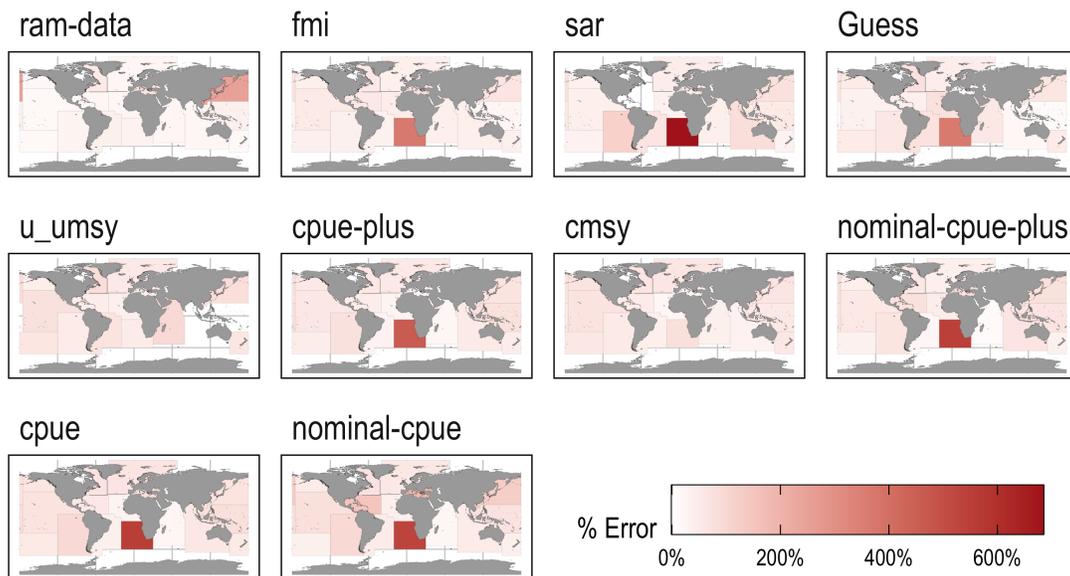


Figure 3: Median absolute percent error in most recent B/B_{MSY} by FAO statistical area from different data sources. ram-data refers to catch and abundance index drawn from RAM. CPUE refers to an index of abundance based on reconstructed effort data. cpue-plus uses CPUE along with Fisheries Management Index (FMI) and/or wept area ratio (SAR) data. For both CPUE series ‘nominal’ assumes a 0% technology rate, otherwise a 2.6% technology rate 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. 2017.¹⁵

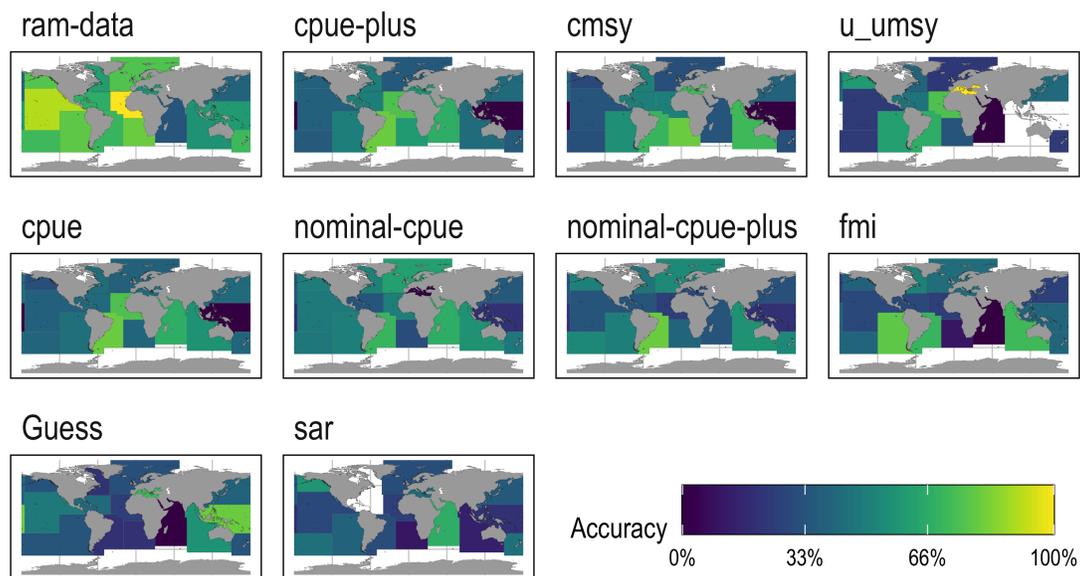


Figure 4: Mean classification accuracy (assignment to FAO stock status category) by FAO statistical area arising from different data sources. ram-data refers to catch and abundance index drawn from RAM. CPUE refers to an index of abundance based on reconstructed effort data. cpue-plus uses CPUE along with Fisheries Management Index (FMI) and/or wept area ratio (SAR) data. For both CPUE series ‘nominal’ assumes a 0% technology rate, otherwise a 2.6% technology rate 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. 2017.¹⁵

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 methods; the difficulty comes in attempting to apply data types uniformly across the globe. While it is unreasonable to expect any global-scale data to be able to perform as well as data pulled from RAM assessments themselves, 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 trends in stock status in time and space. That 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 data themselves.

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. 2019¹⁴ can help distinguish between regions with similar catch histories but different 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/F_{MSY} of any of the datasets we explore on an individual stock basis, a Bayesian R^2 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 results in meaningfully incorrect estimates of stock status for regions whose nature does not match the assumptions needed to apply these data sources.

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 average reduction in error resulting from having access to different kinds of data (Fig.5). 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

Our value-of-information analysis also shows though the high utility of having access to even a recent snapshot of F/F_{MSY} (Fig.5). Swept area ratios, Fisheries Management Index scores, or other similar metrics can be used to construct 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.¹⁷⁻¹⁹ While length-based assessments come with a host of assumptions and pitfalls, properly implemented in some fisheries this may provide an overlooked source of fisheries data at a global scale, at least as an improvement over relying on catch-alone or regional proxies. Such a database could be used to construct stock or stock complex specific priors on fishing mortality for particular regions around the globe, which when paired with catch data and where possible indices of abundance could meaningfully improve our understanding of global fisheries.²⁰

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](#). 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 previously had 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](#) in combination with the reconstruction approaches of¹⁴ might allow us to construct and use timeseries of fishing effort specific to particular areas, fleets, and species complexes

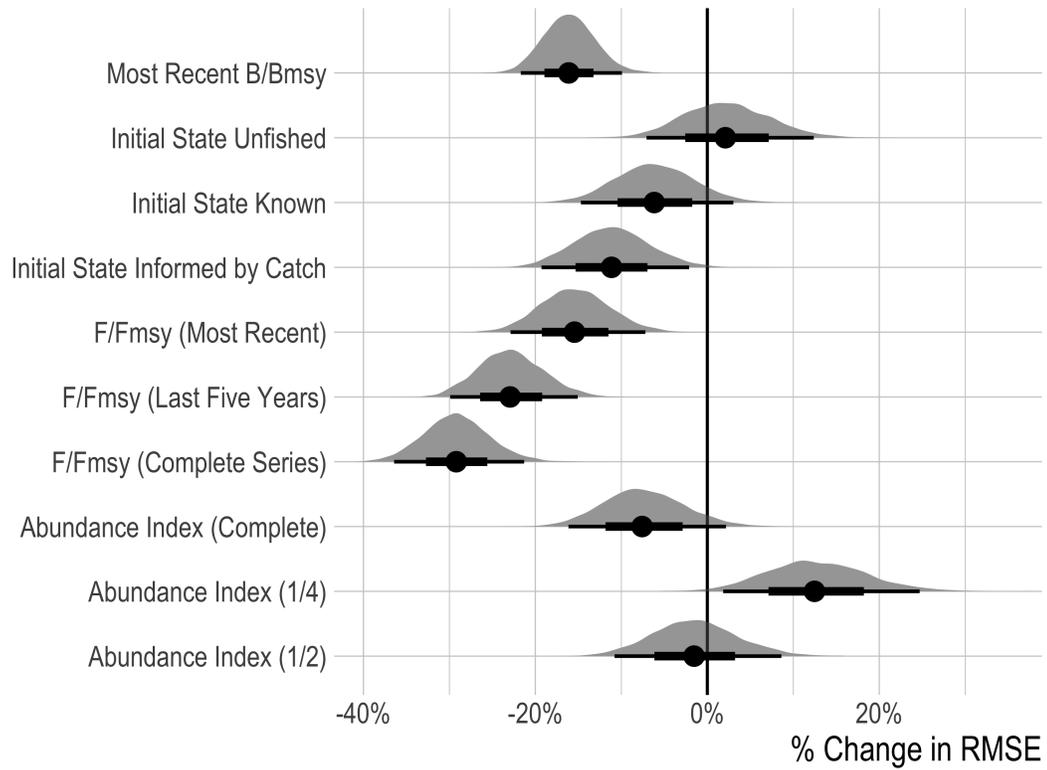


Figure 5: Posterior probability distributions of estimated effect of different data types on root mean squared error of B/B_{MSY} in the most recent 5 years of data available for each model fit. Distribution is full posterior probability distribution. Point is median, thicker black section inner 66th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

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 the data-rich world 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 both improve the health and management of local stocks and provide a means for ensuring that global estimates of stock status are as accurate as they can be.

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 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. Improving estimates of global stock status depends on investing in an improved and expanded global network of fishery data and fisheries scientists.

5 Methods

All analysis were conducted in the R programming language.²¹ Model fitting was conducted using Rcpp [eddelbuettel2011] and stan,²² implemented through Template Model Builder²³ by the tmbstan package.²⁴ The `srplus` package is publicly available at github.com/danovando/srplus, and all materials needed to fully reproduce this manuscript are available at github.com/DanOvando/assessing-global-fisheries.

`srplus` contains too many options to cover within this methods section. We encourage readers to explore the documentation available at the package website at www.github.com/danovando/srplus. Below we describe the structure of the population model underpinning `srplus`, the estimation models used, and the construction of priors used in this paper.

5.1 Population Model

The core of `srplus` is a Pella-Tomlinson²⁵ production model constructed in the manner of.²⁶ While models of these kinds abstract away many important details of fish biology and fleet behavior, they are the highest resolution model that the potential data evaluated here will support. The purpose of `srplus` is not to make substantial improvements 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

$$f(x) = \begin{cases} B_{t+1} = \left(B_t + B_t \frac{r}{m-1} \left(1 - \left(\frac{B_t}{K} \right)^{m-1} \right) - \hat{c}_t \right) p_t, & \text{if } B_t > 0.25 \times K. \\ B_{t+1} = \left(B_t + \frac{B_t}{0.25 \times K} \left(B_t \frac{r}{m-1} \left(1 - \left(\frac{B_t}{K} \right)^{m-1} \right) - \hat{c}_t \right) \right) p_t, & \text{otherwise.} \end{cases} \quad (1)$$

Where B_t is biomass at time t , r is the intrinsic growth rate, m is the scaling parameter that allows for the ratio of B_{MSY}/K to shift. When m is two $B_{MSY} / K = 0.5$, lower values of m shift the production function left, higher values right. \mathbf{c} is a vector catches, and \mathbf{p} 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.²⁶ 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 error suggested by²⁷). 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, see²⁸).

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

$$p_t \sim e^{\text{normal}(-\sigma_{proc}^2/2, \sigma_{proc})} \quad (2)$$

5.2 Estimation Model

All of our estimates are Bayesian in nature. We can break the use of `sraplus` into two distinct categories: with data and without. By “data”, we refer to measurements which are used to confront model estimates within a likelihood function. In our context, these include fishery-independent survey data, or a CPUE index. When there are no data, the model amounts to filtering priors through the model (the combination of the Pella-Tomlinson operating model and the catches for the stock in question, along with any fixed parameters). Under this mode, the model is essentially a stock-reduction analysis model, in the manner of,²⁷ in which we ask, which combinations of prior probability distributions of parameters do not crash the population (i.e. results in biomass less than catches in any time step in the fishery’s history), given the constraints of the population model and the catches. This step updates the prior distribution of population parameters by eliminating combinations of priors that are impossible for a given catch history and a specified functional form.

The full list of estimable parameters are listed in Table.4. 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 , σ_{obs} , γ , and $B0$. q is also estimated when needed.

Table 4: Name, abbreviations, and priors distribution for parameters potentially estimated by `sraplus` in this manuscript.

Parameter	Abbreviation	Default Prior
Carrying Capacity	K	Prior predictive tuning
Growth rate	r	Thorson, 2020 ²⁹ updated by prior predictive tuning
Shape parameter	m	Drawn from Thorson et al. 2012 (7)
Catchability	q	$\text{logn}(1e^{-3}, 1)$
Observation Error	σ_{obs}	$\text{logn}(.05, 1)$
Ratio of process to observation error	γ	$\text{logn}(.5, 0.25)$
Initial State	$B0$	Posterior probability dist. of catch-based regressions

`sraplus` can be run in two forms: either as a stock reduction analysis,²⁷ 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/B_{MSY} , F , or F/F_{MSY} . 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.5 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.

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 stan³¹ accessed through the `tmbstan` interface.²⁴ 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

$$\log(a_t) \sim \text{normal}(f(r, K, m, B0, \mathbf{p}, \mathbf{c}) \times q, \text{sigma}_{obs})$$

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.¹⁴ measure an index of abundance as catch divided by their effort index, either nominal or effective (assuming the 2.6% annualized technology rate). 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

$$cpue_t = \frac{catch_t}{(1 - e^{-f_t})}$$

$$f_t = q_t E_t$$

Where q_t can has a technology rate component τ

$$q_t = q_{t-1} \times (1 + \tau)$$

We then fit to the index of abundance per

$$\log(cpue_t) \sim \text{normal}(f(r, K, m, B0, \mathbf{p}, \mathbf{C}), \sigma_{obs})$$

5.2.1 CMSY

In addition to the results from `sraplus`, we include a set of results produced by the CMSY method.¹⁵ 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 provided by CMSY, except for resilience, which was pulled from the vulnerability scores from FishBase accessed through `rfishbase` [boettiger2012]. Vulnerability scores greater than 66 were scored as low resilience, between 33 and 66 medium resilience, and lower than 33 high resilience.

5.3 Priors

Priors for all estimable parameters can be left at default values or set by the user. The shape parameter is usually not reliably estimable given available data for surplus production models, however, Thorson et al. 2012⁷ provides estimate the ratio of B_{MSY} 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. 2012⁷ 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`.

5.3.1 Prior Predictive Tuning

Suppose that the only thing we observe from a fishery is a catch history. Assuming 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. We can think of this as a binomial process in which we fit a model of the population, conditional on catches, to the fact that we know that population existed in each time step of the catch history. Beyond that though, barring additional information or model assumptions we have no way of knowing whether these catches represent a substantial proportion of a small population or a minuscule fraction of a massive population; all we know is that the current population must be greater than 0.

In the absence of any data to fit to, the SRA algorithm works by assuming that we know current stock status, and then finds feasible parameters to satisfy that belief. This creates a problem for the Bayesian nature of our analysis though. 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 essentially 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. This may seem like an academic concern, and indeed in our experience when the data are sufficiently informative the Bayesian version of our model subject to Borel's paradox produces effectively identical results to those produce by the same model fit by maximum likelihood. However, particularly for the SRA version of `sraplus`, Borel's Paradox poses a particular problem.

The SRA algorithm works in two steps. First, the algorithm rejects any draws that resulted in the collapse of the population (biomass less than catch in a given timestep). From there a standard SRA would sample from the priors in proportion to the stated prior on recent stock status. If the bulk of the prior on terminal stock status was concentrated at 50% of K , combinations of r and K that produce terminal stock status near 50% of K are sampled proportionally more frequently. However, lower values of terminal stock status have fewer candidate values of r and K , since it becomes harder and harder to find viable pairs that come close to but do not crash the population at any time step. Conversely, in the absence of constraints higher values of stock status have infinite combinations of plausible r and K combinations: since under this model the population cannot be greater than carrying capacity, as for example K approaches infinity terminal stock status asymptotes at close to 100% of K . The net result of this is that even though individual combinations of r and K that produce higher stock status than the mean of the prior on recent stock status individually have lower probability of being sampled, there are many more opportunities for the lower-probability events that produce higher stock status to be sampled. As a result, the post-model-pre-data prior on terminal depletion will always be higher under this method than the supplied prior on stock status.

We use an approximation to this problem here, similar in spirit to Bayesian melding.³² Our solution amounts to a two-step 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.

$$p(bin_i) = \frac{1}{N_i} \sum_{n=1}^{N_i} p(b_{n,i})$$

We then divide the sampling probability of of bin i evenly among each of the draws within that bin n

$$p(n_i) = \frac{p(bin_i)}{N_i}$$

And we then perform a second SIR algorithm but now sampling each observation n_i in proportion to $p(n_i)$. The net result of this is a post-model-pre-data distribution of parameters r and K that produce a distribution of recent stock status that roughly matches the supplied prior on recent stock status. In effect, this process answers the question “given the model, what combinations of parameters produce my prior on recent stock status”. This is only an approximate solution, but it helps ensure that the post-model-pre-data distribution of stock status much more closely matches the stated prior on recent stock status, and reduced the positive bias resulting from use of the raw SRA algorithm (Fig.6, Fig.S1).

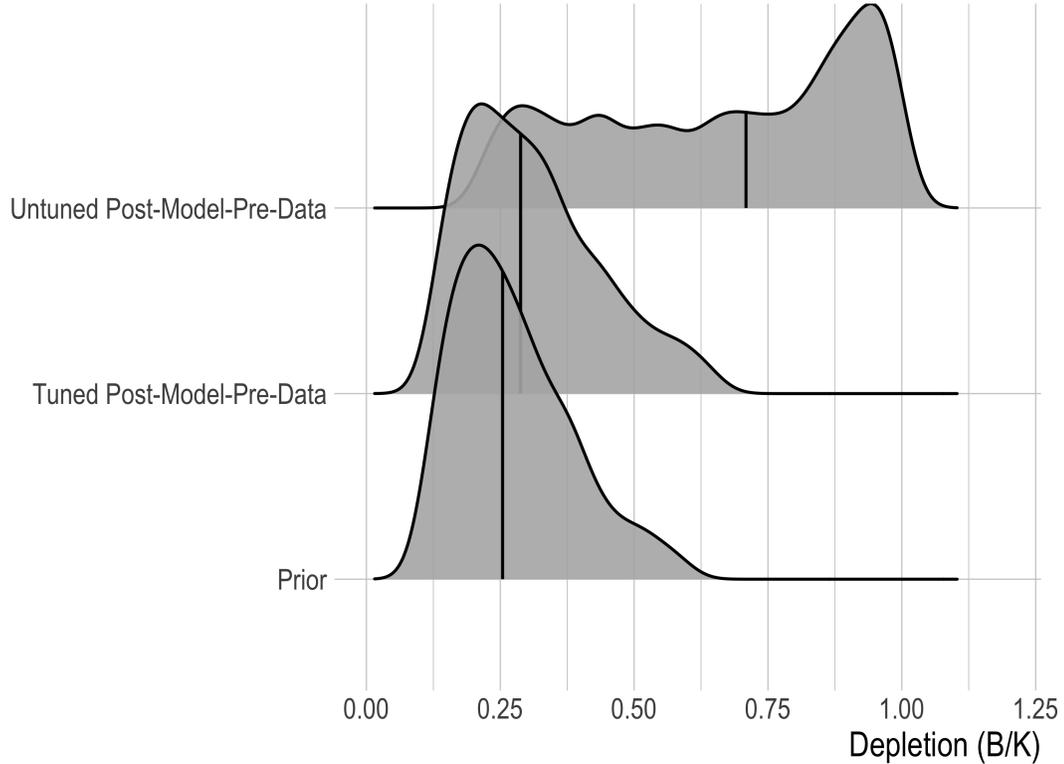


Figure 6: Post-model-pre-data distribution of depletion (biomass relative to carrying capacity) from raw SRA algorithm (untuned, top row), from SRA algorithm with approximate tuning applied (tuned, middle row), compared to the supplied prior on depletion (bottom row). Black vertical line indicates median value.

5.3.2 Priors Informed by Outside Data

Along with allowing users to supply their own priors, the `sraplus` package contains three built-in methods used in this manuscript for converting information on stock status from additional outside data into a form usable as a 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 $\log(status) \sim normal(variable, \sigma)$ for each of these three data types. Given values of these variables for a new fishery then, `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`.

All prior regression models were tested by out-of-sample predictive power, and where competing models were considered the final model was chosen by leave-on-out validation.³³ 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 presented regressions 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.

5.3.2.1 Catch-Only Priors Many of the current methods for estimating global stock status of unassessed stocks are based on predicting stock status from characteristics of the catch history.³⁻⁶ While these catch-only methods have been shown to have serious shortcomings,⁸ we include them as a point of reference given their ubiquity in the global assessment literature.

We used data from the RAM Legacy Stock Assessment Database to estimate a regression of stock status as a function of catch history characteristics. To facilitate the process, we first fit a spectral clustering algorithm to the scaled catch histories of fisheries in RAM, in order to identify four possible clusters of catch history types within the data. Emergent clusters show for example one built around a downward “one way trip” style catch histories, others with a boom and bust pattern, others with stable but fluctuating catches.

We then trained a classification algorithm to predict which catch cluster a given fishery would fall into based on the shape of its catch history. This algorithm was then used to assign fisheries to one of the four identified catch history types, and the catch history type was then used as a hierarchical term within our catch-based regressions (where s refers to a smoothing term). For the first regression, we restrict the data to the first year of data available for each fishery i , in order to estimate initial stock status

$$\log(value_i) \sim normal\left(s\left(\frac{first(catch)}{max(catch)}\right)|cluster_i\right) + s(\log(length_i)|cluster_i) + 1, \sigma$$

For the second regression, we included data for all available years y for fishery i . The model is then used to construct a prior on fishery status in the terminal year of the data

$$\log(value_{i,y}) \sim normal\left(s(fyear|cluster_i) + s\left(\frac{catch_{i,y}}{max(catch_i)}\right)|cluster_i\right) + cluster_i, \sigma$$

where $fyear$ is the year of the fishery, starting from 0.

5.3.2.2 Fisheries Management Index Priors The Fisheries Management Index (FMI), as presented in,¹¹ utilizes surveys filled out by regional experts to score a fishery against a set of 46 specific questions for individual species about what elements of fisheries management were in place. These questions are then aggregated into broader categories of science, enforcement, management, and socioeconomic. The higher the score, the better the expert judges that a given metric is met in that fishery. Importantly, FMI surveys can be filled out in the absence of stock assessments. This allows us to explore how FMI values map onto

stock status, and explore the ability then to use FMI scores to produce priors on stock status for unassessed fisheries (in a manner similar to³⁴ and³⁵).

The final selected model relating FMI variable to stock status metrics was a generalized additive model (GAM) of the form

$$\log(value_i) \sim N(s(research_i)+s(management_i)+s(enforcement_i)+s(socioeconomics_i)+\frac{catch_i}{max(catch)_i}+1, \sigma_{SAR})$$

5.3.2.3 Swept Area Ratio Priors ¹² provides an extensive database of trawling footprints throughout the world, including both regions heavily covered by stock assessments and largely unassessed areas. This makes the trawl footprint data an ideal candidate for supporting global stock assessment efforts. As illustrated in,¹² there is an evident positive relationship between the swept area ratio (SAR, the total annual area trawled divided by the total area of the region) and U/U_{MSY} . Note that SAR can be greater than 1 since the same area can be trawled multiple times in a year, e.g. if all trawl-able areas are trawled twice a year then the SAR will be 2. Also note the skewed distribution of SAR values with most concentrated well below 1 and only a handful above 1.

The final selected model relating SAR to stock status metrics was

$$\log(value_i) \sim normal(s(SAR_i) + s(\frac{catch_i}{max(catch)_i}) + 1, \sigma_{SAR})$$

5.4 Value of Information Calculations

What sources of data might provide the greatest value in improving our estimates of global stock status? We fit 3,000 `sraplus` models to randomly sampled fisheries from RAM, each time varying the kind and quality of data made available to the model, and what parameters the model estimated. We then calculated the root-mean-squared-error (RMSE) between the observed and predicted B/B_{MSY} over the most recent five years of each fit, and then fit a regression to these data to estimate the posterior probability of the effect of different data types and model states on RMSE (Fig.5).

The value-of-information (VOI) calculations presented in Fig.5 help illustrate 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 stock status (B/B_{MSY}) for stocks in the RAM legacy stock assessment, 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.

Using this set of fits, we assess performance as the root-mean-squared-error of B/B_{MSY} over the most recent 5 years of the fishery, in order to evaluate the ability of the model to capture the recent trends in stock status and not just the most recent year. We evaluate the contributing of each data type to RMSE using a Gamma GLM with a log link of the form

$$rmse \sim Gamma(\beta\mathbf{X} + (1|stock), shape, scale)$$

Where β is the vector of coefficients associated with the matrix of dummy variables marking the use of different data types in the vector \mathbf{X}

6 References

1. FAO. *STATE OF WORLD FISHERIES AND AQUACULTURE 2020: Sustainability in action*. (FOOD & AGRICULTURE ORG, 2020).
2. Hilborn, R. *et al.* Effective fisheries management instrumental in improving fish stock status. *Proceedings of the National Academy of Sciences* (2020) doi:10.1073/pnas.1909726116.
3. Costello, C. *et al.* Status and Solutions for the World's Unassessed Fisheries. *Science* **338**, 517–520 (2012).
4. Costello, C. *et al.* Global fishery prospects under contrasting management regimes. *Proceedings of the National Academy of Sciences* **113**, 5125–5129 (2016).
5. Pauly, D. The Sea Around Us Project: Documenting and Communicating Global Fisheries Impacts on Marine Ecosystems. *AMBIO: A Journal of the Human Environment* **36**, 290–295 (2007).
6. Rosenberg, A. A. *et al.* Applying a New Ensemble Approach to Estimating Stock Status of Marine Fisheries around the World. *Conservation Letters* **11**, e12363 (2018).
7. Thorson, J. T., Cope, J. M., Branch, T. A. & Jensen, O. P. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. *Canadian Journal of Fisheries and Aquatic Sciences* **69**, 1556–1568 (2012).
8. Free, C. M. *et al.* Blood from a stone: Performance of catch-only methods in estimating stock biomass status. *Fisheries Research* **223**, 105452 (2020).
9. Martell, S. & Froese, R. A simple method for estimating MSY from catch and resilience. *Fish and Fisheries* **14**, 504–514 (2013).
10. Ricard, D., Minto, C., Jensen, O. P. & Baum, J. K. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish and Fisheries* **13**, 380–398 (2012).
11. Melnychuk, M. C., Peterson, E., Elliott, M. & Hilborn, R. Fisheries management impacts on target species status. *Proceedings of the National Academy of Sciences* **114**, 178–183 (2017).
12. Amoroso, R. O. *et al.* Bottom trawl fishing footprints on the world's continental shelves. *Proceedings of the National Academy of Sciences* 201802379 (2018) doi:10.1073/pnas.1802379115.
13. Rousseau, Y., Watson, R., Blanchard, J. & Fulton, E. Evolution of global marine fishing fleets and the response of fished resources. *Proceedings of the National Academy of Sciences* 201820344 (2019) doi:10.1073/pnas.1820344116.
14. Rousseau, Y., Watson, R. A., Blanchard, J. L. & Fulton, E. A. Evolution of global marine fishing fleets and the response of fished resources. *Proceedings of the National Academy of Sciences* 201820344 (2019) doi:10.1073/pnas.1820344116.
15. Froese, R., Demirel, N., Coro, G., Kleisner, K. M. & Winker, H. Estimating fisheries reference points from catch and resilience. *Fish and Fisheries* **18**, 506–526 (2017).
16. Maureaud, A. *et al.* Are we ready to track climate-driven shifts in marine species across international boundaries? - A global survey of scientific bottom trawl data. (2020) doi:10.1101/2020.06.18.125930.
17. Hordyk, A., Ono, K., Prince, J. D. & Walters, C. J. 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* (2016) doi:10.1139/cjfas-2015-0422.
18. Prince, J. & Hordyk, A. What to do when you have almost nothing: A simple quantitative prescription for managing extremely data-poor fisheries. *Fish and Fisheries* **20**, 224–238 (2019).
19. Rudd, M. B. & Thorson, J. T. 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 (2017) doi:10.1139/cjfas-2017-0143.

20. Thorson, J. T. & Cope, J. M. Catch curve stock-reduction analysis: An alternative solution to the catch equations. *Fisheries Research* **171**, 33–41 (2015).
21. R Core Team. R: A Language and Environment for Statistical Computing. (2019).
22. Carpenter, B. *et al.* Stan : A Probabilistic Programming Language. *Journal of Statistical Software* **76**, (2017).
23. Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H. & Bell, B. M. **TMB** : Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* **70**, (2016).
24. Monnahan, C. & Kristensen, K. No-u-turn sampling for fast bayesian inference in admB and tmb: Introducing the admB and tmbstan R packages. *PloS one* **13**, (2018).
25. Pella, J. J. & Tomlinson, P. K. A generalized stock production model. *Inter-American Tropical Tuna Commission Bulletin* **13**, 416–497 (1969).
26. Winker, H., Carvalho, F. & Kapur, M. JABBA: Just Another Bayesian Biomass Assessment. *Fisheries Research* **204**, 275–288 (2018).
27. Walters, C. J., Martell, S. J. D. & Korman, J. A stochastic approach to stock reduction analysis. *Canadian Journal of Fisheries and Aquatic Sciences* **63**, 212–223 (2006).
28. Thorson, J. T., Rudd, M. B. & Winker, H. The case for estimating recruitment variation in data-moderate and data-poor age-structured models. *Fisheries Research* (2018) doi:10.1016/j.fishres.2018.07.007.
29. Thorson, J. T. 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 (2020).
30. Thorson, J. T., Cope, J. M., Branch, T. A. & Jensen, O. P. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. *Canadian Journal of Fisheries and Aquatic Sciences* **69**, 1556–1568 (2012).
31. Stan Development Team. `{}R`: The `{}R` interface to `{}Stan`. (2018).
32. Poole, D. & Raftery, A. E. Inference for Deterministic Simulation Models: The Bayesian Melding Approach. *Journal of the American Statistical Association* **95**, 1244–1255 (2000).
33. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**, 1413–1432 (2017).
34. Osio, G. C., Orío, A. & Millar, C. P. Assessing the vulnerability of Mediterranean demersal stocks and predicting exploitation status of un-assessed stocks. *Fisheries Research* **171**, 110–121 (2015).
35. Cope, J. M., Thorson, J. T., Wetzel, C. R. & DeVore, J. Evaluating a prior on relative stock status using simplified age-structured models. *Fisheries Research* **171**, 101–109 (2015).

7 Supplemental Material

Table 5: Data sources used for terminal stock status estimate

Data Name	Description
RAM-Data	Fit to abundance index from RAM
sar	Prior on terminal F/Fmsy set by regional swept area ratio
fmi	Prior on terminal F/Fmsy set by regional fisheries management index scores
cpue	Fit to CPUE index created from RAM catch and regional effort index. 2.5% tech. creep
cpue-plus	Fit to CPUE index created from RAM catch and regional effort index with priors informed by SAR and FMI. 2.5% tech. creep
nominal-cpue	Fit to CPUE index created from RAM catch and regional effort index. 0% tech. creep

Data Name	Description
nominal-cpue-plus	Fit to CPUE index created from RAM catch and regional effort index with priors informed by SAR and FMI. 0% tech. creep
catch-only	priors on terminal B/Bmsy informed by catch history
heuristic	priors on initial and terminal B/Bmsy informed by CMSY heuristic

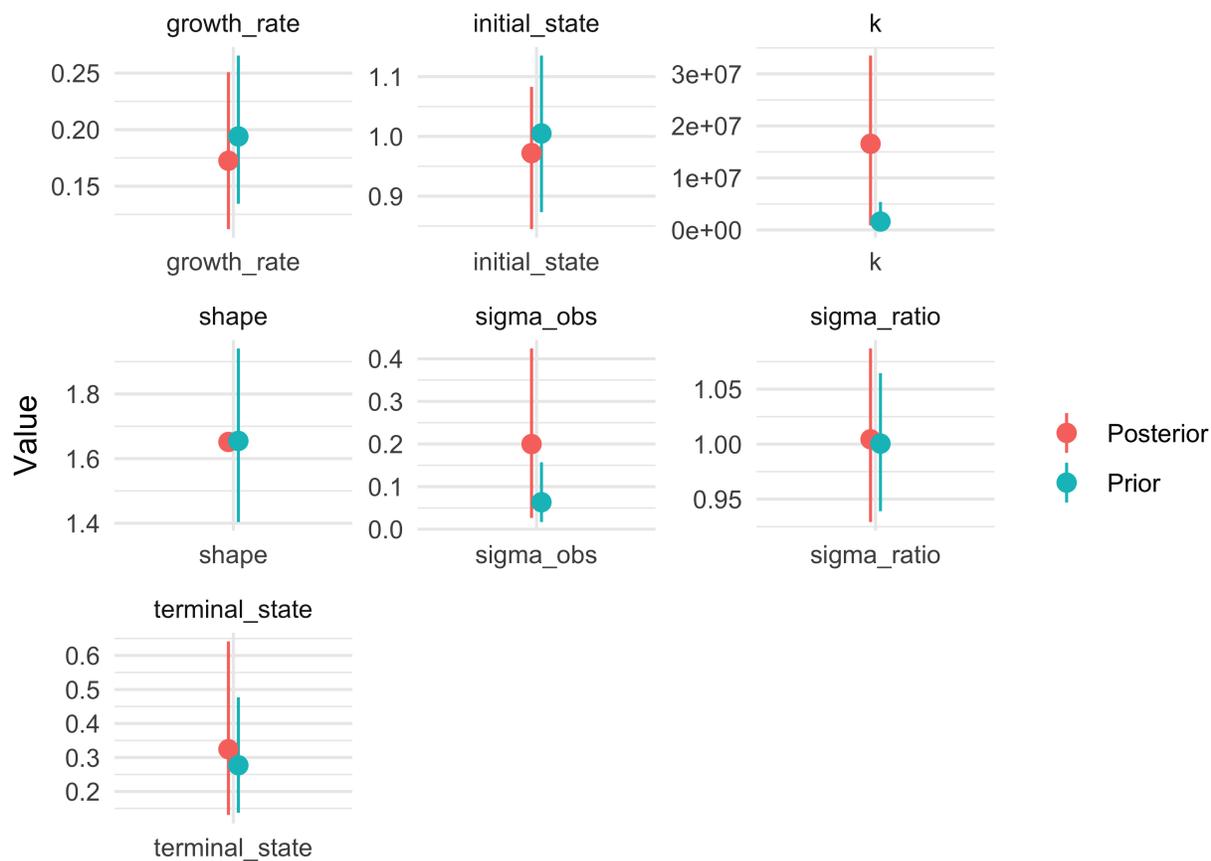


Figure 7: Prior-posterior plots of fits for case study fishery in Fig.6