**Methods**

*Estimating FAO stock status*

We constructed time series of stock status (B/BMSY) for 1,740 FAO fish stocks using a superensemble model that estimates B/BMSY from the B/BMSY predictions of four individual catch-only models (**Table 1**) and two spectral properties of the catch time series. The superensemble model is adapted from the Anderson et al. (2017) superensemble model, which produces better estimates of stock status than other catch-only models (Anderson et al. 2017; Free et al. in prep) and has been used to estimate the terminal year status of 785 FAO fish stocks (Rosenberg et al. 2017). We extended the analysis of Rosenberg et al. (2017) to estimate status from 1950-2015 for the 1,740 FAO fish stocks (FAO area-country-species triples) meeting the following criteria: marine wild capture fisheries for finfish and invertebrates with taxonomic identification resolved to the species-level and with catch time series ≥20 yrs and ≥250 mt of median annual catch. We also excluded highly migratory species whose population dynamics cannot be described by catch within a single country’s exclusive economic zone.

The superensemble model uses boosted regression trees to estimate B/BMSY in year *i* using: (1) the 0.20 and 0.05 spectral densities of the scaled catch time series (catch divided by maximum catch) from year *0* to year *i* and (2) B/BMSY predictions for year *i* from four individual catch-only models (**Table 1**) applied to the full catch time series. We calculated the 0.20 and 0.05 spectral densities, which correspond to 5- and 20-year cycles respectively, using the *spec.ar* function in R. We implemented the individual catch-only models using the *datalimited* (Anderson et al. 2016) and *datalimited2* (Free 2018) packages in R. In addition to catch time series, mPRM requires the classification of species into 17 life history categories (**Supp. Table 1**), cMSY-13 and COM-SIR require estimates of resilience (i.e., the capacity to withstand exploitation; **Supp. Table 2**), and OCOM requires estimates of natural mortality. We classified life history based on taxonomy and derived resilience and natural morality estimates from a mixture of FishBase (Froese & Pauly 2017), SeaLifeBase (Palomares & Pauly 2017), and FishLife (Thorson et al. 2018) life history information.

All analyses were conducted in R v.3.4.2 (R Core Team 2017) and the code is available here: <https://github.com/cfree14/trade_and_collapse>. See the supplemental methodsfor more details on the development and testing of the superensemble model.

*Exploring sequential exploitation*

We examined patterns in sequential exploitation by mapping the year of first development and year of first overexploitation for the 32 species with ≥10 stocks. We defined the years of first development and overexploitation as the years with the first B/BMSY mean (5-yr left-aligned rolling mean) less than 1.0 and 0.5, respectively.

**Supplemental methods**

*FAO stock selection*

We analyzed the 1,740 FAO fish stocks (FAO area-country-species triples) meeting the following criteria: marine wild capture fisheries for finfish and invertebrates with taxonomic identification resolved to the species-level and with catch time series ≥20 yrs and ≥250 mt of median annual catch after trimming years of zero catch from the beginning of the time series. We also excluded: (1) stocks of species that could not be placed into life history categories consistent with the mPRM model (e.g., barnacles, corals, sea cucumbers, sea urchins, starfish, sponges, etc.); (2) stocks of highly migratory species whose population dynamics cannot be described by catch within a single country’s exclusive economic zone; and (3) stocks targeted by a distant water fleet whose catch time series is unlikely to be representative of total removals from that population (i.e., stocks whose FAO area and EEZ don’t overlap were excluded).

*Building the superensemble model*

We developed the superensemble model using simulated fish stocks from Rosenberg et al. (2014) and tested the models on a set of simulated stocks withheld from model training and on real fish stocks in the RAM Legacy Stock Assessment Database (RAMLDB v. 2.95; Ricard et al. 2012). The Rosenberg et al. (2014) simulated stocks represent a fully factorial set of 5760 simulated fisheries comprised of three fish life histories, three levels of initial biomass depletion, four exploitation scenarios, two levels of recruitment variability, two levels of recruitment autocorrelation, and two levels of measurement error, with each combination of parameters run through ten stochastic iterations (**Supp. Table 3**). The RAMLDB is a global database of catch data and stock assessment output, including reference points and time series of biomass and fishing mortality.

The superensemble model uses boosted regression trees (BRT) to estimate stock status (B/BMSY) from the B/BMSY estimates of four individual catch-only assessment models (**Table 1**) and two spectral properties of the catch time series. Boosted regression trees combine regression and machine learning, offer predictive power superior to other modeling methods (Elith et al. 2008), and produced the best superensemble model in Anderson et al. (2017). We excluded SSCOM, one of the individual catch-only models included in the original Anderson et al. (2017) because of its enormous run-time (<8 stocks / day), and included OCOM, which was developed after the Anderson et al. (2017) model was published. We included the 0.05 and 0.20 spectral densities of the scaled catch time series (catch divided by maximum catch) because they were shown to improve predictive performance in Anderson et al. (2017). Because B/BMSY is a ratio bounded at zero, we fit the BRT models using the log of B/BMSY and exponentiated predictions from the model. Thus, each of the superensemble models has the following conceptual structure:

where represents the superensemble estimate of B/BMSY, ’s represent the individual model estimates of B/BMSY, and *SD*’s are the spectral densities of the scaled catch time series.

We divided the simulated stocks for model training (90% of data) and testing (10% of data) by withholding the 10th iteration of each simulation scenario. The training stocks were used to fit the BRT model while the test stocks were used to independently evaluate the model’s predictive ability. A grid search for the BRT model parameters that minimize the RMSE using repeated 10-fold cross validation found the following optimal parameters: learning rate=0.005, interaction depth=10, and number of trees=7500. The BRT models were fit using the *caret* (Kuhn 2016) and *gbm* (Ridgeway 2016) packages in R v.3.4.2 (R Core Team 2017).

*Assigning life history traits to the FAO species*

In addition to catch time series, the individual catch-only models require information on life history category, resilience, and natural mortality. We collected this information using a combination of FishBase (Froese & Pauly 2017), SeaLifeBase (Palomares & Pauly 2017), and FishLife (Thorson et al. 2018) life history information.

We used the *rfishbase* package in R (Boettiger et al. 2012) to correct the taxonomy of species in the FAO landings data and download their habitat types, Von Bertalanffy growth parameters, maximum size, and vulnerability and resilience from FishBase (FB, for finfish; Froese & Pauly 2017) and SeaLifeBase (SLB, for invertebrates; Palomares & Pauly 2017). We also used the *FishLife* package in R (Thorson et al. 2017) to estimate natural mortality and Von Bertalanffy growth parameter for all finfish species. *FishLife* uses a multivariate model trained on FishBase to predict eight life history traits for >32,000 fish (Thorson et al. 2017).

We classified species into the 17 life history categories used by the mPRM catch-only model based on taxonomy using **Supp. Table 1**. We classified species into resilience categories (**Supp. Table 2**) using, in order of preference, resilience values: (1) reported on FB/SLB; (2) derived from the FishLife Von Bertalanffy K parameter; (3) derived from the FB/SLB Von Bertalanffy K parameter; (4) derived from the FB/SLB vulnerability metric; (5) derived from the FB/SLB Von Bertalanffy maximum age; (6) derived from the genus mode; or (7) derived from the family mode (**Supp. Table 2**). We used natural mortality estimates in the following order of preference: (1) FishLife values; (2) FB/SLB values; or (3) derived using the tmax- and growth-based estimators recommended by Then et al. (2014). Resilience and natural mortality estimates remained unavailable for only 37 and 64 invertebrate species, respectively (of >1500 species).

**References**

Anderson, S.C., J. Afflerbach, A.B. Cooper, M. Dickey-Collas, O.P. Jensen, K.M. Kleisner, C. Longo, G.C. Osio, D. Ovando, C. Minte-Vera, C. Minto, I. Mosqueira, A.A. Rosenberg, E.R. Selig, J.T. Thorson, and J.C. Walsh. 2016. datalimited: Stock assessment methods for data-limited fisheries. R package version 0.0.2. Available at:<https://github.com/datalimited/datalimited>

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

Costello, C., D. Ovando, R. Hilborn, S.D. Gaines, O. Deschenes, and S.E. Lester. 2012. Status and solutions for the world’s unassessed fisheries. Science338:517-520.

Jensen, O.P., Free, C.M. (2017) Testing and comparison of data-limited assessment models for estimating global and regional stock status. UN Food & Agriculture Organization.

Kuhn, M., 2016. Classification and Regression Training. R Package Version 6 (0–71).

Martell, S., and R. Froese. 2013. A simple method for estimating MSY from catch and resilience. Fish & Fisheries 14:504-514.

Ridgeway, G., 2016. Gbm: Generalized Boosted Classification Models. R Package Version 2.1.1.

Rosenberg, A.A., M.J. Fogarty, A.B. Cooper, M. Dickey-Collas, E.A. Fulton, N.L. Gutiérrez, K.J.W. Hyde, K.M. Kleisner, C. Longo, C.V. Minte-Vera, C. Minto, I. Mosqueira, G.C. Osio, D. Ovando, E.R. Selig, J.T. Thorson, and Y. Ye. 2014. Developing new approaches to global stock status assessment and fishery production potential of the seas. FAO Fisheries and Aquaculture Circular, Rome, Italy.

Rosenberg, A.A., Kleisner, K.M., Afflerbach, J., Anderson, S.C., Dickey‐Collas, M., Cooper, A.B., Fogarty, M.J., Fulton, E.A., Gutiérrez, N.L., Hyde, K.J.W., Jardim, E., Jensen, O.P., Kristiansen, T., Longo, C., Minte-Vera, C.V., Minto, C., Mosqueira, I., Chato Osio, G., Ovando, D., Selig, E.R., Thorson, J.T., Walsh, J.C., Ye, Y. (2017) Applying a new ensemble approach to estimating stock status of marine fisheries around the world. *Conservation Letters*: doi: 10.1111/conl.12363

Then, A.Y., Hoenig, J.M., Hall, N.G., Hewitt, D.A. (2014) Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES Journal of Marine Science* 72(1): 82-92.

Thorson, J.T., C. Minto, C.V. Minte-Vera, K.M. Kleisner, C. Longo. 2013. A new role for effort dynamics in the theory of harvested populations and data-poor stock assessment. Canadian Journal of Fisheries & Aquatic Sciences 70(12):1829-1844.

Thorson, J.T., S. B. Munch, J. M. Cope, J. Gao. (2017) Predicting life history parameters for all fishes worldwide. *Ecological Applications* 27(8): 2262–2276.

Vasconcellos, M., and K. Cochrane. 2005. Overview of world status of data-limited fisheries: inferences from landings statistics. In: Kruse, G.H., V.F. Gallucci, D.E. Hay, R.I. Perry, R.M. Peterman, T.C. Shirley, P.D. Spencer, B. Wilson, and R. Woodby (Eds.): Fisheries Assessment and Management in Data-Limited Situations. Alaska Sea Grant College Program, University of Alaska Fairbanks, Fairbanks, AK, pp. 1–20.

**Tables & Figures**

**Figure 1.** FAO stock sample size.

**Figure 2.** FAO stock status over time (from superensemble model).

**Figure 3.** FAO stock area centroids.

**Supp. Figure 1.** Superensemble tuning.

**Supp. Figure 2.** Superensemble performance.

**Table 1.** Catch-only methods used in superensemble model.

**Supp. Table 1.** Life history categorizations for mPRM.

**Supp. Table 2.** Resilience life history traits and r priors.

**Supp. Table 3.** Simulated stocks factorial design.

**Appendix A.** Catch time series.

**Appendix B.** B/BMSY time series.

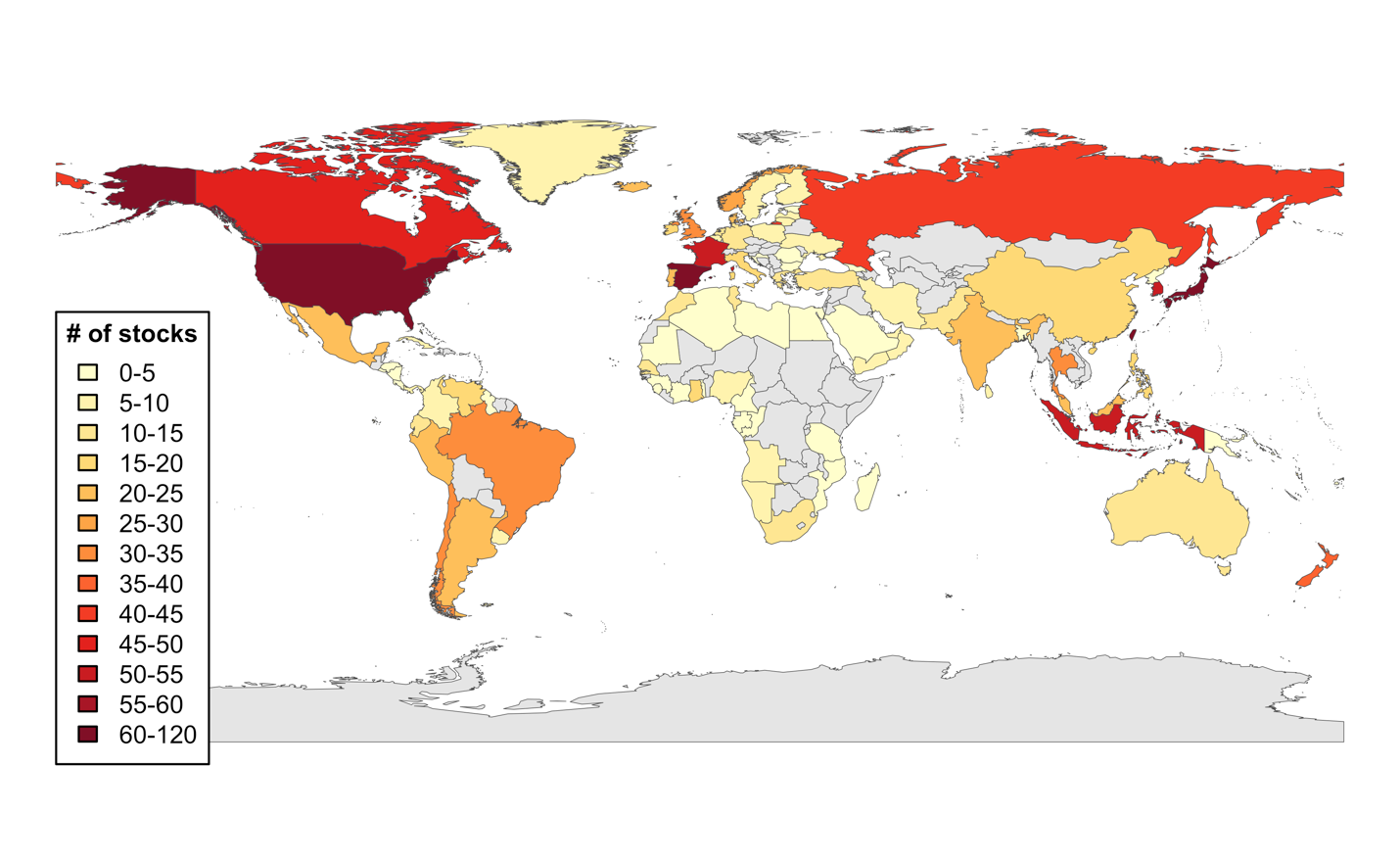
**Appendix C.** Maps showing year of first development and first collapse.

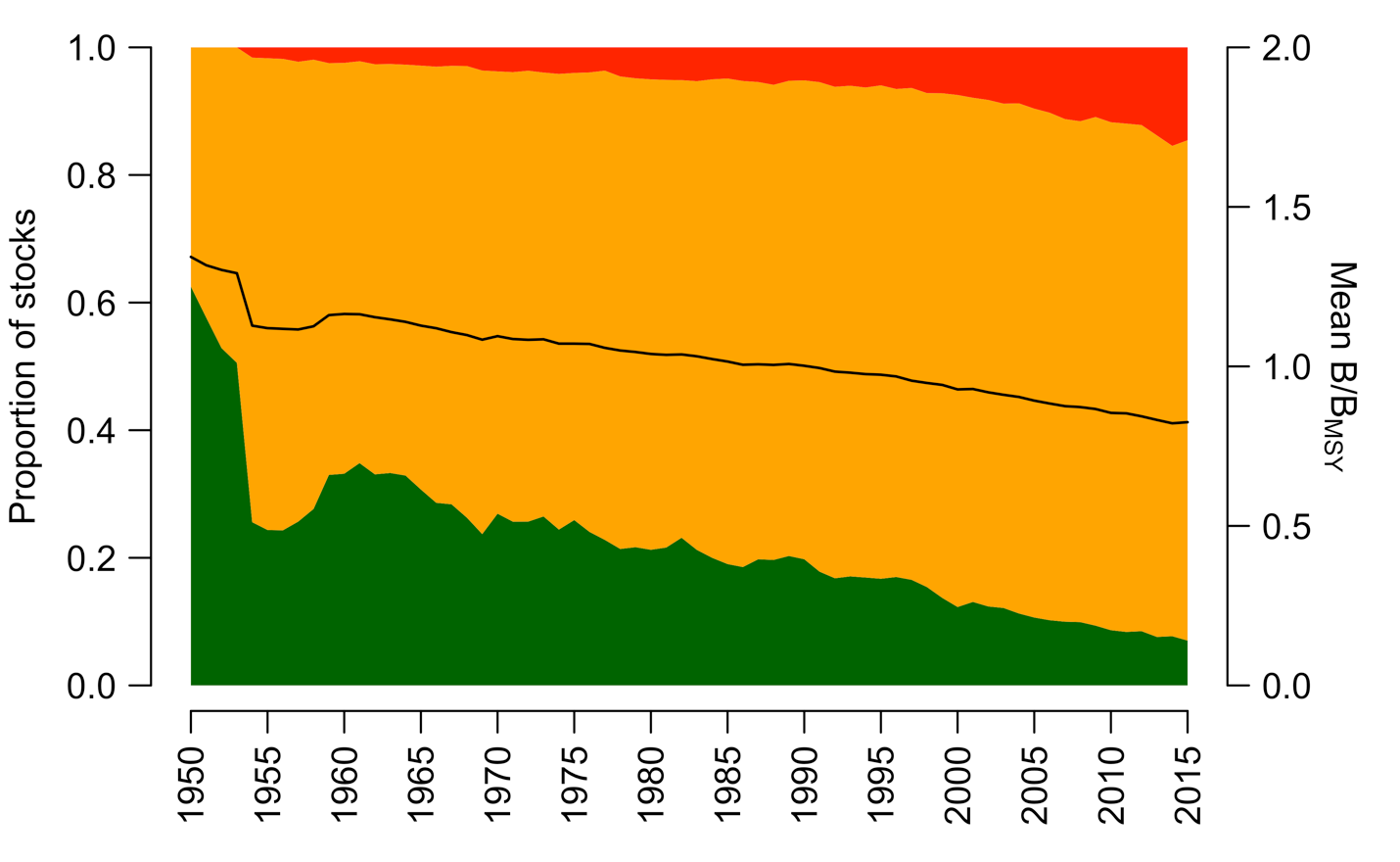
**Appendix captions**

**Appendix A.** Catch time series for FAO stocks used in the analysis(where each stock is a FAO area-country-species triple). For all stocks, the x-axis span 1950-2010 but are unlabeled to minimize clutter; y-axis show catch in metric tons but are also unlabeled.

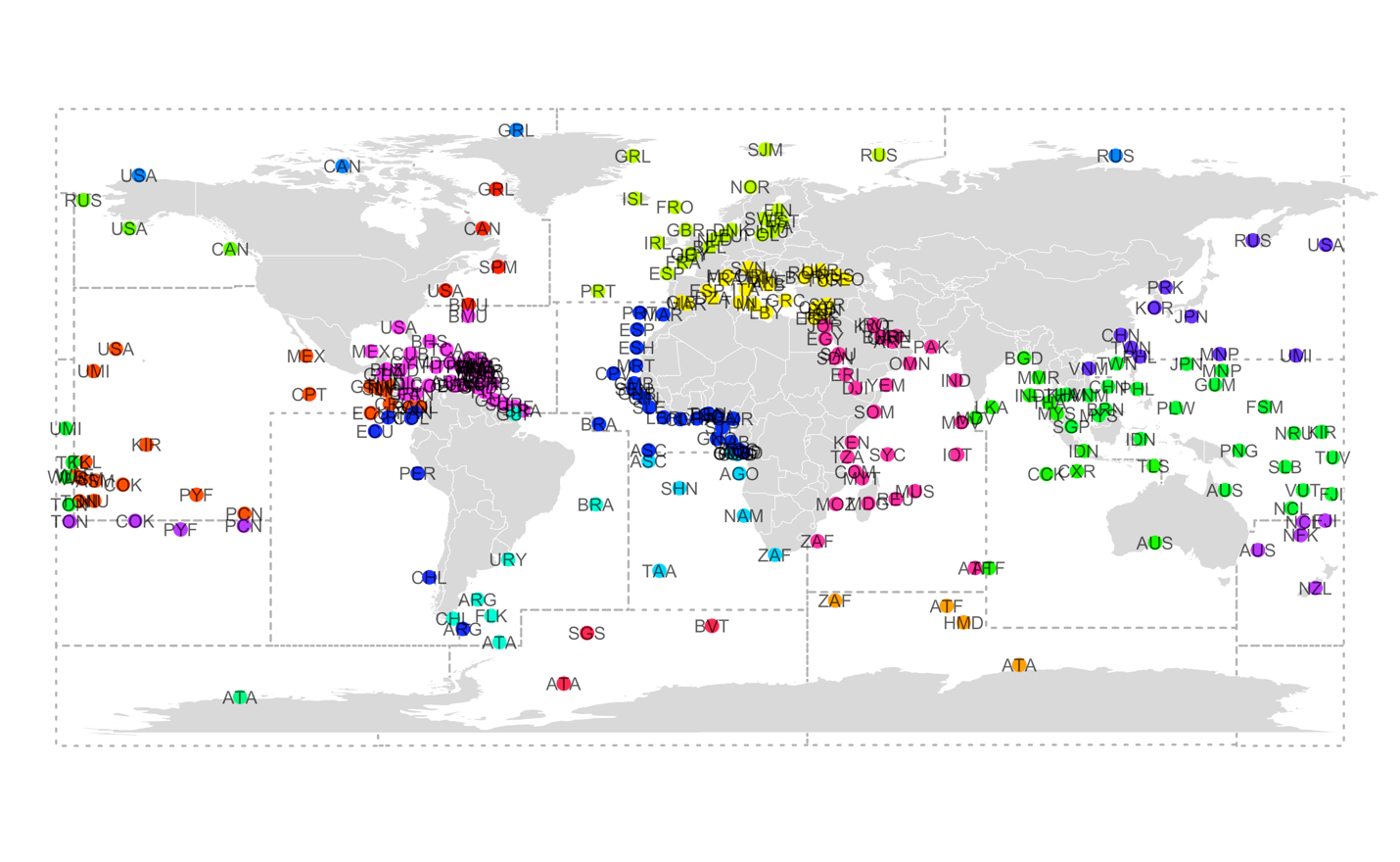
**Appendix B.** B/BMSY time series estimated by: mPRM (red), cMSY-13, OCOM (blue), COM-SIR (green), and the superensemble model (black). For all stocks, the x-axis span 1950-2010 but are unlabeled to minimize clutter.

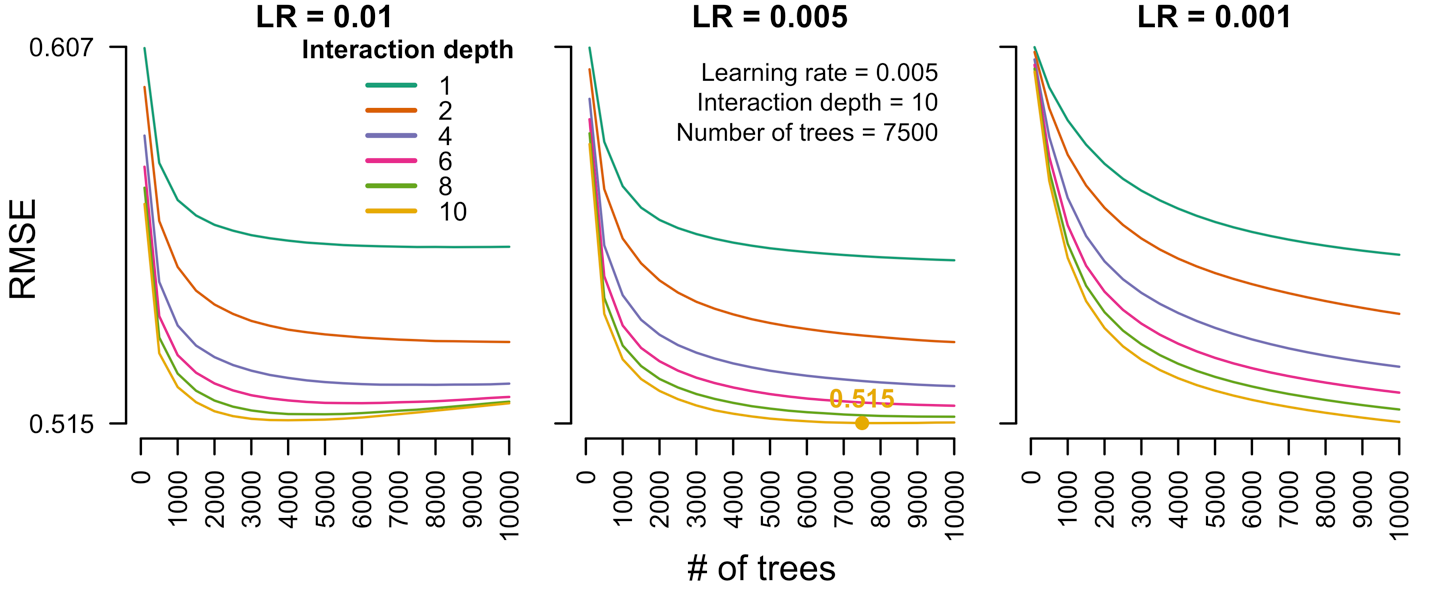
**Appendix C.** Maps showing **(A)** the year of first development; **(B)** the year of first overexploitation; and **(C)** the year of first collapse for species with ≥10 stocks.

**Figure 1.** Distribution of FAO stocks included in analysis (where each stock is a FAO area-country-species triple).

****

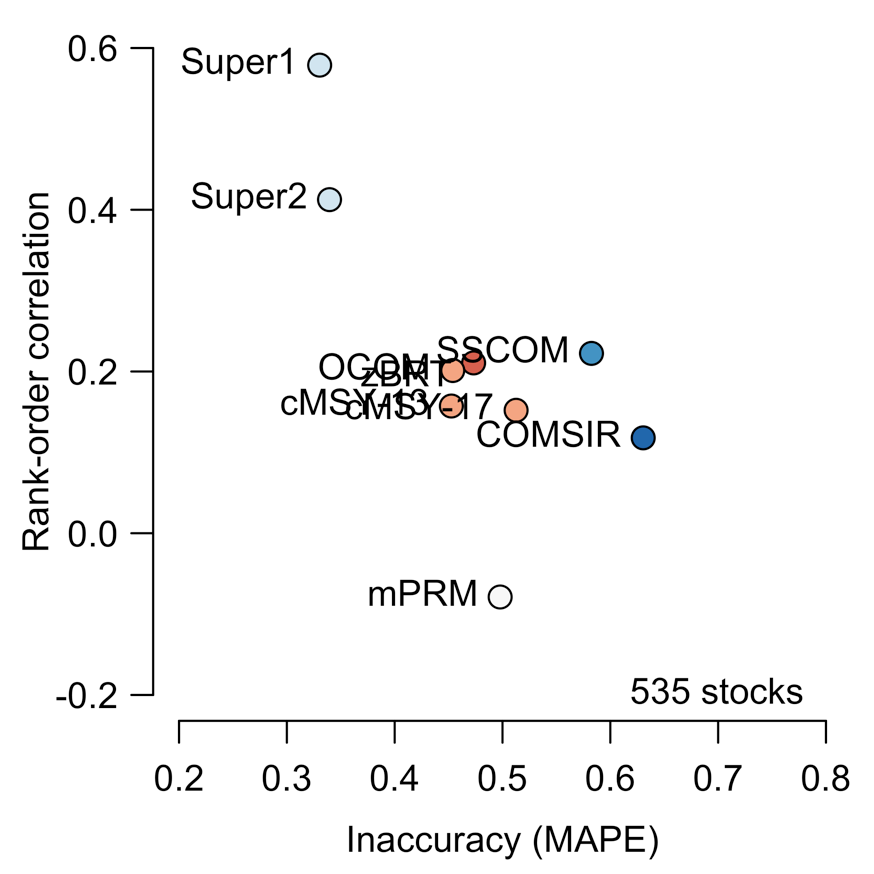
**Figure 2.** Status of FAO stocks over time. Colors indicate the proportion of stock classified as: underexploited (B/BMSY > 1.25, green), fully exploited (0.5 < B/BMSY < 1.25, orange), and overexploited (B/BMSY < 0.5, red). The black line shows mean B/BMSY over time.

**Figure 3.** Centroids of stock areas (i.e., FAO area-country intersects) represented in the FAO analysis. Points are colored by FAO area (also shown in dashed lines).

****

**Supp. Figure 1.** Model tuning curves showing the average root mean square error (RMSE) for each combination of candidate BRT model parameters (learning rate, interaction depth, # of trees) for the superensemble model. The optimal combination of model parameters (marked and labeled) is the combination that minimizes the RMSE.

When retraining model, add more trees (15,000? 20,000?) and reduce number of interaction depths (1,2,4,8,12).

****

**Supp. Figure 2.** The performance of COMs evaluated on the simulated stocks withheld from the BRT model training (n=535). The best performing methods are indicated by high rank-order correlation and low inaccuracy (top-left corner).

Add categorical performance. Add performance on RAMLDB stocks. Add scatterplots of true and superensemble predicted B/BMSY for both simulated stocks and RAMLDB stocks.

**Tables**

**Table 1.** Individual catch-only models used in the superensemble model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Method** | **References** | **Data input/output** | **Brief description** |
| 1 | **cMSY-2013** Catch-MSY | Martell & Froese 2013 Rosenberg et al. 2014 | In: Catch, resilience Out: B/BMSY, MSY, B, BMSY | Uses a stock reduction analysis with priors for r, k, and initial/final year depletion derived from resilience to estimate status |
| 2 | **COM-SIR** Catch-only-model with sampling-importance-resampling | Vasconcellos & Cochrane 2005 Rosenberg et al. 2014 | In: Catch, resilience Out: B/BMSY | Uses a coupled harvest-dynamics model fit using a sampling-importance-resampling algorithm to estimate status |
| 3 | **OCOM** Optimized catch-only model | Zhou et al. 2017 | In: Catch, natural mortality (M) Out: Saturation, MSY | Uses a stock reduction analysis with priors for r and final year depletion derived from M and saturation from Zhou-BRT to estimate status |
| 4 | **mPRM** Modified panel regression model | Costello et al. 2012 Anderson et al. 2017 | In: Catch, taxonomic group Out: B/BMSY | Uses a panel regression model trained on the RAMLDB to predict status from characteristics of the catch time series and taxonomic group |

Consider adding cMSY-2017 to superensemble model. Consider training on RAMLDB?

**Supp. Table 1.** Classification of FAO stocks into life history categories consistent with the mPRM catch-only model.

|  |  |  |
| --- | --- | --- |
| **Type** | **Category** | **Taxonomic groups** |
| Finfish | Cods, hakes, haddocks | Order: Gadiformes |
| Finfish | Flounders, halibuts, soles | Order: Pleuronectiformes |
| Finfish | Herrings, sardines, anchovies | Order: Clupeiformes (except the shads) |
| Finfish | Shads | Subfamily: Alosinae; Genera: Alosa, Brevoortia, Ethmalosa, Ethmidium, Gudusia, Hilsa, Tenualosa |
| Finfish | Tunas, bonitos, billfishes | Scombridae (tunas, bonitos, mackerel), Istiophoridae (marlin), Xiphiidae (swordfish) |
| Finfish | Sharks, rays, chimaeras | Classes: Elasmobranchii (sharks, rays, skates, sawfish), Holocephali (chimaeras) |
| Finfish | Miscellaneous coastal fishes | Unclassified finfish/cephalopods with w/ a reef-associated, benthopelagic, pelagic-neritic, or pelagic habitat type |
| Finfish | Miscellaneous demersal fishes | Unclassified finfish/cephalopods with w/ a demersal, bathydemersal, or bathypelagic habitat type |
| Finfish | Miscellaneous diadromous fishes | Families: Salmonidae (salmon), Moronidae (temperate basses) |
| Finfish | Miscellaneous pelagic fishes | Unclassified finfish/cephalopods with w/ a pelagic-oceanic habitat type |
| Molluscs | Abalones, winkles, conchs | Class: Gastropoda |
| Molluscs | Clams, cockles, arkshells | Orders: Veneroida (clams/cockles), Arcoida (ark shells) plus Myoida (other clams), Mytiloida (mussels) |
| Molluscs | Scallops, pectens | Order: Ostreoida |
| Crustaceans | Crabs, sea-spiders | All crab/sea spider families plus geryon families |
| Crustaceans | King crabs, squat-lobsters | Families: Lithodidae (king crabs), Galatheidae (squat lobsters) |
| Crustaceans | Lobsters, spiny-rock lobsters | Families: Nephropidae (true lobsters), Palinuridae (spiny lobsters), Scyllaridae (slipper lobsters) |
| Crustaceans | Shrimps, prawns | All shrimp/prawn families plus krill/seabobs/mantis shrimp families |

**Supp. Table 2.** AFS and FishBase guidelines for using life history traits to classify the resilience of fish stocks to exploitation and the r priors used by COM-SIR and cMSY-13 for each resilience category.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Resilience** | **r prior** | **Von B**  **K (1/yr)** | **Age at**  **maturity (yr)** | **Maximum**  **age (yr)** | **Fecundity**  **(1/yr)** |
| High | [0.6, 1.5] | >0.3 | <1 | 1-3 | >10,000 |
| Medium | [0.2, 1.0] | 0.16-0.30 | 2-4 | 4-10 | 100-1000 |
| Low | [0.05, 0.5] | 0.05-0.15 | 5-10 | 11-30 | 10-100 |
| Very low | [0.015, 0.1] | <0.05 | >10 | > 30 | <10 |
| Unknown | [0.2, 1.0] | ------ | ------ | ------ | ------ |

**Supp. Table 3.** Factorial design of the Rosenberg et al. (2014) simulated stocks.

|  |  |  |
| --- | --- | --- |
| **Factor** | **# of levels** | **Levels** |
| Life history | 3 | Demersal, small pelagic, or large pelagic |
| Initial biomass depletion | 3 | 100%, 70%, or 40% of carrying capacity |
| Exploitation dynamics | 4 | Constant, biomass-coupled, increasing, or roller coaster rates |
| Recruitment variability | 2 | Low or high variability |
| Recruitment autocorrelation | 2 | With or without autocorrelation |
| Catch measurement error | 2 | With or without catch measurement error |
| Time series length | 2 | 20 or 60 years |
| Iterations | 10 | Iterations for each combination of the above parameters |
| **Total # of stocks:** | 5760 |  |