# Supplementary information: Global trade network patterns coupled to marine fisheries sustainability

## SI 1. FAO stock status estimates

### 1.1 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).

### 1.2 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 (**Table S4**). 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 S1**) 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:

(3)

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).

### 1.3 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 and Pauly, 2017), SeaLifeBase (Palomares and Pauly, 2017), and FishLife (Thorson et al., 2017, p. 201) 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 and 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 **Table S2**. We classified species into resilience categories (**Table S3)** 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 (SI Table S3). 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 >1,500 species).

#### Table S1. Individual catch-only models used in the superensemble model.

### 

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Method | References | Data input/output | Brief description |
| 1 | cMSY-2013  Catch-MSY | Martell and 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 and 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., 2018 | 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 |

### 

### 

#### Table S2. 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 |

### 

### 

#### Table S3. 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] | ------ | ------ | ------ | ------ |

### 

### 

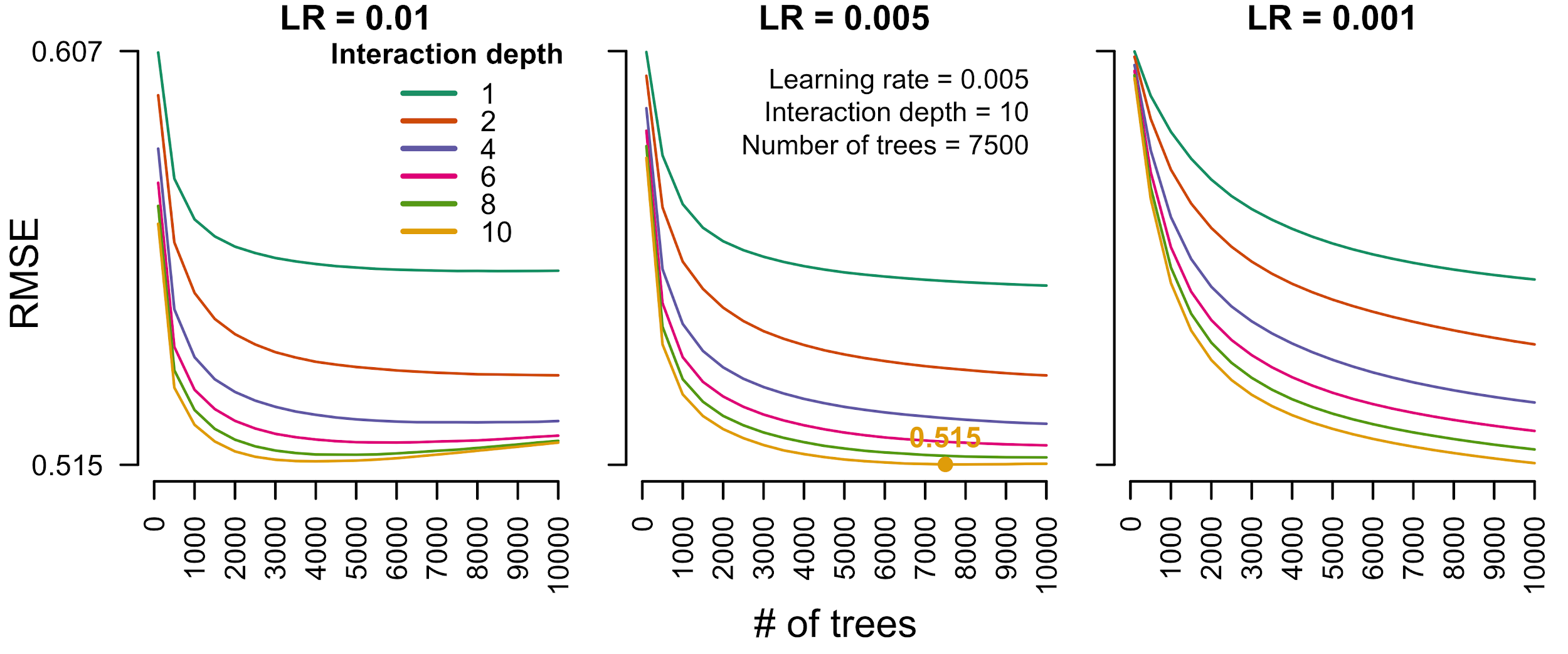
#### Table S4. Factorial design of the Rosenberg et al. (2014) simulated stocks.

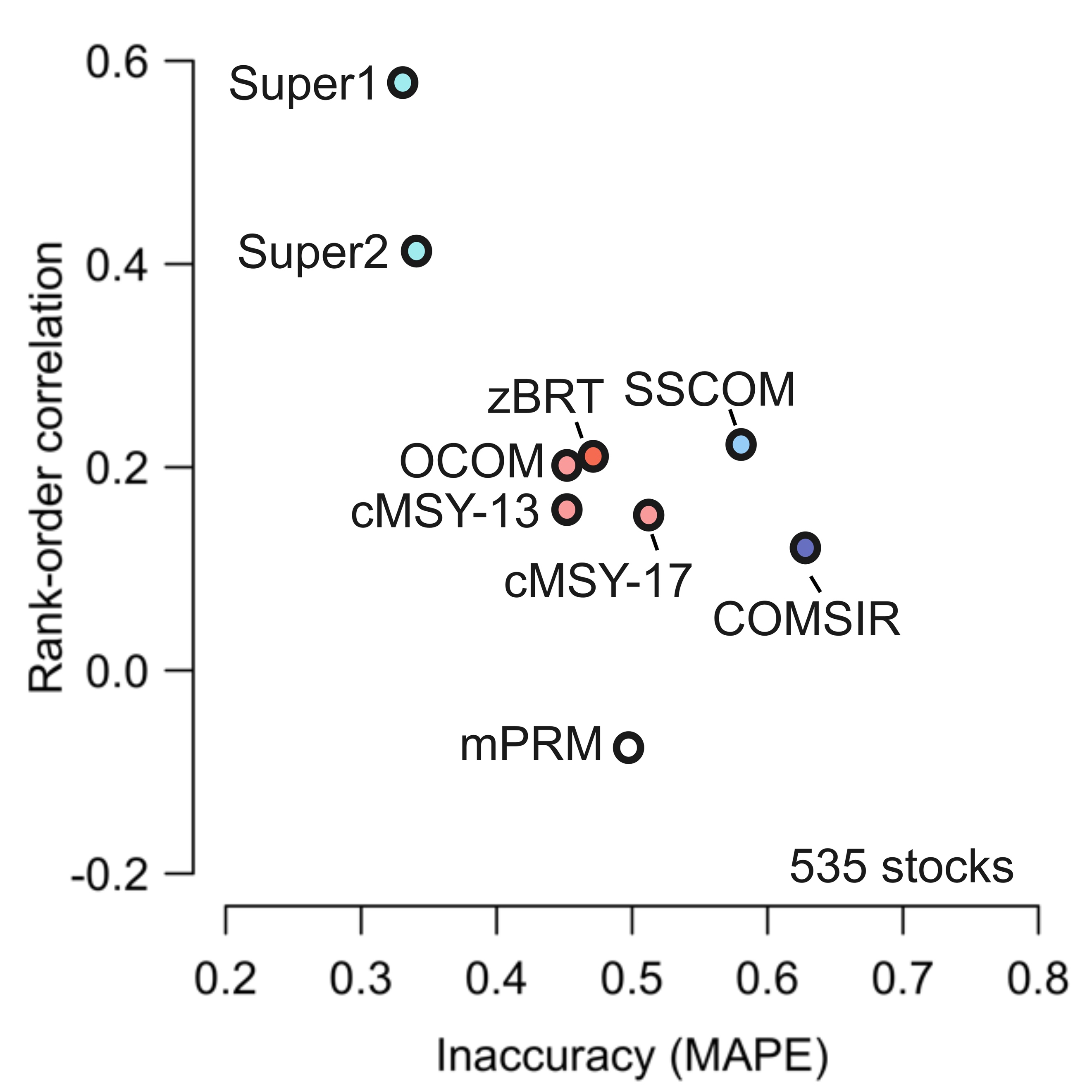
### 

|  |  |  |
| --- | --- | --- |
| 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 |  |

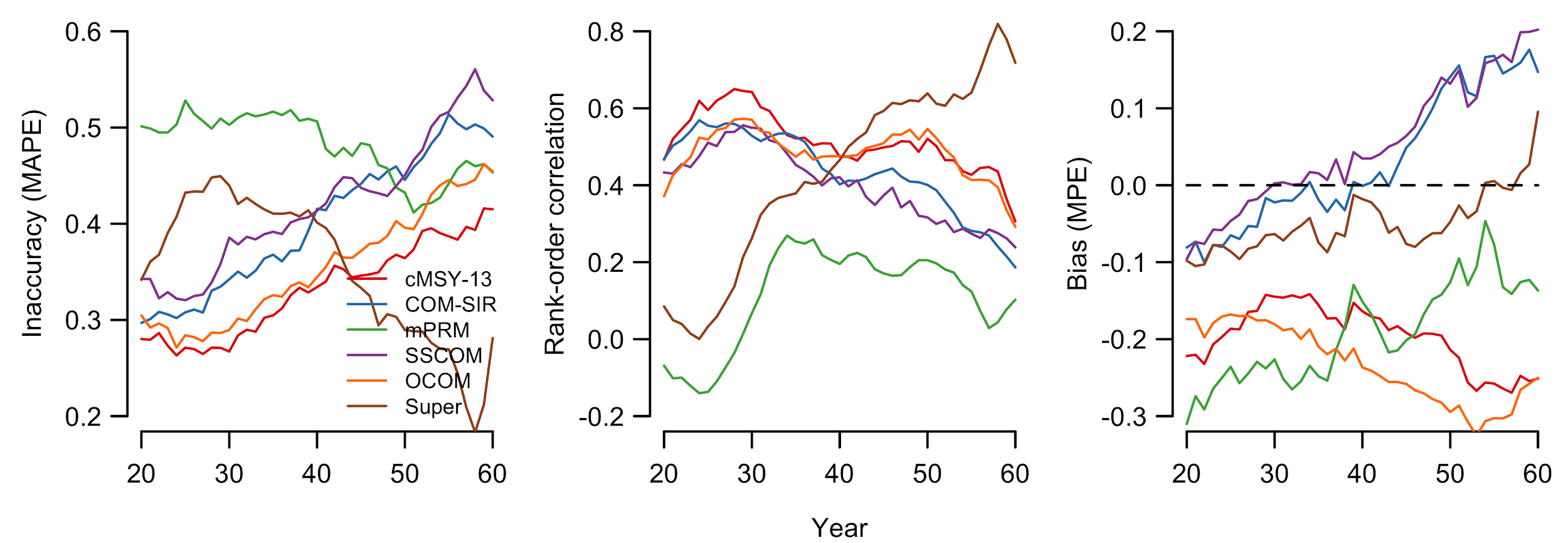
### 

### 

**Figure S1.** 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.

****

**Figure S2.** 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).

**Figure S3.** Status estimation performance of six catch-only stock status models. The performance of the superensemble model improves over time while the performance of the other catch-only models degrades over time.

## 

## SI 2. Species groups

#### Table S5. Species groups included in the analysis (listed alphabetically).

|  |
| --- |
| **Species groups** |

Anchovies

Coalfish

Cod

Crab

Cuttlefish

Eel

Flatfish

Haddock

Hake

Halibut

Herring

Homarus

Lobster

Mackerel

Mussel

Octopus

Oysters

Plaice

Rock lobster

Sardines

Scallop

Seabass

Shark

Sole

## 

## SI 3. Indicators of speed and scale

#### Table S6. Indicators of speed and scale in seafood trade.

|  |  |  |  |
| --- | --- | --- | --- |
| **Network characteristic** | **Calculation** | **Reference** | **Data resolution** |
| Clustering coefficient | where *v* is all closed network triangular trade connections and *L* is the number of triangular trade connections that are possible in the network. | [Watts and Strogatz, 1998](https://www.zotero.org/google-docs/?Gxx1MV) | At the network level (i.e. per year and species group) |
| Average node degree | where *u* is the number of trade connections of a specific importer or exporter and *n* is the number of countries in the network. | [Diestel, 2017](https://www.zotero.org/google-docs/?5l0d0F) | At the network level (i.e. per year and species group) |
| Trade duration | where *D* is the number of consecutively traded years between a specific importer and exporter pair and *l* is the number of importers trading with the specified exporter. |  | At the level of the exporter (i.e. per year, species group and exporter) |
| Turnover of trade connections | *Mt* is the total number of unique importer-exporter combinations recorded in year *t*,  and *k* is the unique importer exporter combinations in common between the network-years. | [Magurran, 1988](https://www.zotero.org/google-docs/?VeEbXj) | At the network level (i.e. per year and species group) |

## SI 4. Fixed effects estimations

The model setting is static and the network effects on fishery stocks are assumed to be contemporaneous (**Model 4, Table S7**).

(5)

where status (B/BMSY) of species group *g* exported by exporter *i* in year *t* is predicted by the network scale metrics (clustering and degree) and network speed metrics (turnover and duration) of species group *g* in year *t*. The error term is composed of two terms: first, represents unobserved exporter-species time-invariant fixed effects, capturing factors such as unchanged institutional and geographical characteristics, and, second, represents a stochastic error term to introduce variation in other variables that could potentially affect stock status but are not included in our model. The second model specification is represented by (**Model 5, Table S7**):

(6)

Above equation includes *(t – 1)* year dummies (), which captured the year-specific factors common to all exporter-species pairs. Year dummies, as in the GMM estimator (**Model 2, Table 1**) are used to remove time trends in dependent and independent variables but can be negligible if there are no common trends for each species group or each exporter' stock.

The term might be correlated with other covariates and thus leads to biased estimation. The fixed effect estimator (i.e. within-estimator) can eliminate the fixed effects and thus address part of the endogeneity problem. The endogeneity resulting from can be addressed using the GMM estimator. We tested model fit for the fixed effect estimator using within r-squared and evaluated model significance using the p-value with a < 0.05 alpha value.

## 

## Table S7. Fixed-effects model estimation for the contemporaneous model (Within-Estimator). Standard errors in parentheses. Significance levels reported as \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001.

|  |  |  |
| --- | --- | --- |
| Network Characteristics | Model 4 | Model 5 |
|  | B/BMSY | B/BMSY |
| clustering | 0.682\*\*\* | -0.0845 |
|  | (0.132) | (0.144) |
| degree | -0.0203\*\*\* | -0.00284 |
|  | (0.00342) | (0.00378) |
| turnover | -0.00000238 | 0.000113 |
|  | (0.0000846) | (0.0000957) |
| trade duration | -0.0576\*\*\* | -0.0440\*\* |
|  | (0.0156) | (0.0150) |
| Year dummies | No | Yes |
| *N* | 5423 | 5423 |
| within- R2 | 0.039 | 0.077 |

## 

## 

## SI 5. Trade duration and traded volume

## 

**Figure S4.** Shows traded volume and trade duration by exporting country. Traded volume sums all reported trades per year, species group and exporter-importer pair. To calculate maximum trade duration, each consecutive year of trade between an exporter-importer pair is summed.

## SI References

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

Boettiger, C., Lang, D.T., Wainwright, P.C., 2012. rfishbase: exploring, manipulating and visualizing FishBase data from R. J. Fish Biol. 81, 2030–2039. https://doi.org/10.1111/j.1095-8649.2012.03464.x

Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., Lester, S.E., 2012. Status and Solutions for the World’s Unassessed Fisheries. Science 338, 517–520. https://doi.org/10.1126/science.1223389

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Froese, R., Pauly, D., 2017. FishBase 2017, version (march, 2017). URL http://www fishbase.org

Kuhn, M., 2016. Classification and Regression Training.

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

Palomares, M.L.D., Pauly, D., 2017. SeaLifeBase.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Ricard, D., Minto, C., Jensen, O.P., Baum, J.K., 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish Fish. 13, 380–398. https://doi.org/10.1111/j.1467-2979.2011.00435.x

Ridgeway, G., 2016. Gbm: Generalized Boosted Classification Models.

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

Thorson, J.T., Munch, S.B., Cope, J.M., Gao, J., 2017. Predicting life history parameters for all fishes worldwide. Ecol. Appl. 27, 2262–2276. https://doi.org/10.1002/eap.1606

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

Zhou, S., Punt, A.E., Smith, A.D.M., Ye, Y., Haddon, M., Dichmont, C.M., Smith, D.C., 2018. An optimized catch-only assessment method for data poor fisheries. ICES J. Mar. Sci. 75, 964–976. https://doi.org/10.1093/icesjms/fsx226