# Methods

## 1. Underwater visual census database

We used a published global database of reef fish abundance and sizes collected along belt transects1. This database encompasses 9118 transects across 585 sites (98 localities) in the Central Indo-Pacific, Central Pacific, Eastern Pacific, Western Indian, Eastern Atlantic, Western Atlantic. The database only includes sites with a hard reef bottom at the outer reef slope. Transects were carried out at a constant depth, parallel to the reef crest. We selected the species inside families for which we have body stoichiometric data, that were at least 7cm to minimize the bias related to the indentification of small individuals, and finally we discarded rare species, for which less than 20 individuals were ever recorded across all transects. The dataset then included 1110 species that belong to 25 families (Acanthuridae, Balistidae, Bothidae, Chaetodontidae, Cirrhitidae, Fistulariidae, Haemulidae, Holocentridae, Kyphosidae, Labridae, Lethrinidae, Lutjanidae, Monacanthidae, Mugilidae, Mullidae, Ostraciidae, Pempheridae, Pomacanthidae, Pomacentridae, Sciaenidae, Scorpaenidae, Serranidae, Siganidae, Tetraodontidae, Zanclidae).

## 2. Bioenergetic modeling

We applied a series of bioenergetic models for each existing combination of species identity, body size, and sea surface temperature (n = 30668). The bioenergetic model framework integrates elements of metabolic theory, stoichiometry, and flexible elemental limitation2. As such, this model is well suited for the application of a wide variety of fish species. The model relies on a set of input parameters including elements of metabolism, growth, and diet and body stoichiometry. We quantified the input parameters for all species through the integration of empirical data, data synthesis, and Bayesian phylogenetic models (See supplementary methods). The bioenergetic model then allows for the prediction of elemental fluxes mediated by individuals, including ingestion rate, excretion rates of N and P, and growth rate. Here, we focused on 5 key processes: N excretion rate (g N / day), P excretion rate (g P/day), production of body mass through growth (g C/day), herbivory, i.e. ingestion rate of macrophytes (g/day), and piscivory, i.e. ingestion rate of fishes (g/day). We summed the contribution of each individual to each process on the transect-level, to get community-level estimates, then dividing it by the surface area that was surveyed. Each process is thus expressed in dry mass per day per square meter. We note that N excretion, P excretion, and production include contributions of all fishes, whereas herbivory and piscivory are only carried out by a subset of the community, with respect to their trophic guild (3).

## 3. Relationship between functions and biomass

To assess the effect of biomass and sea surface temperature sst), we performed a Bayesian mixed effect regression of each log-transformed function for transect-scale observations ():

We then assessed the covariation between functions, independent of biomass and sst. To do so, we first extracted the median residuals for each function per transect. In some transects, there were no piscivores or herbivores observed. In those cases, we did not include these transects in the analysis. To assess the covariation, we then quantified the correlations that exist among the different functions. To visualize the residuals per locality on a world map, we first grouped all transects per site by taking the median value for each process across multiple transect. We then grouped the sites again into localities by taking the median.

Finally, we quantified baseline values for each function and assessed the occurrence of communities that reach baseline values for each function. Based on the biomass reference for a healthy reef of ~100g/m2 (4), we predicted the baseline for each function for the full range of sst’s. Then, for all communities, we quantified whether or not the community had higher values for all functions compared to their baselines. Using this logical variable, we fitted a Bernoulli model with a logit-link to predict the probability of having all five functions above their threshold in relation to biomass and sea surface temperature, similar to the model specified above.

## 4. Effect of community structure on ecosystem functions

To investigate the effect of the community structure accounting for the effect of standing biomass and sea surface temperature, we quantified a set of variables that characterize the community. These variables discribe the size, age, and trophic distribution of the community, as these may all affect functions2. Specifically, we took the 2.5%, 50% and 97.5% quantiles of the total length, immaturity, and trophic level of all individuals per transect. The total length is based on the visual estimation by divers. The immaturity is quantified using the following formula:

where is the species-specific growth rate parameter and is the species-specific asymptotic adult length, and is the total length of individual i. Essentially, this is the derivative of the Von Bertalanffy growth model for a certain length, and the higher this value is, the younger the individual. Finally, the trophic level was extracted from fishbase5. Additionally, we quantified the transect-level species richness. For each log-transformed function we then fitted a Bayesian mixed-effect model with all 12 above-mentioned variables:

To compare effects across functions and assess the relative importance of each variable, we standardized all variables prior to model fitting. We fitted all 5 models by using 4 cores, that each had 2000 iterations with a warm-up of 1000 iterations. We used weakly-informative priors6.

## 5. Species dominance and contributions to functions

To investigate the role of each species for each function, we quantified the contribution of each species to each function for all transects as followed:

where i is a certain species, j is a transect, F is the value of function f.

Then, we quantified the degree of species dominance per function for each transect. We did this by first ranking species according to their contribution to function, followed by quantifying the cumulative contributions of species to functions. Then, we used the area under the species accumulation curve as a measure for the degree of dominance. Specifically, the degree of dominance (DD) was calculated as followed:

where is the area under the curve, is the theoretical area under the curve where each species has an equal contribution to a certain function, is the theoretical area under the curve where one species performs the entire function. They are quantified as:

where is the contribution of a certain species and R is the species richness. The degree of dominance thus ranges between 0 and 1, where 0 means that each species contributes equally and 1 means that one species performs the entire function.

Finally, to know how often species are contributing more than average for a certain function, we quantified the frequency of dominance, i.e. the number of times a species is dominant divided by the total number of transects in which that species is observed. A species is considered dominant for a certain function in a given transect if their contribution is higher than 1/R, i.e. they contribute more than the situation in which each species contributes equally to a certain function.

## 6. Species vulnerability to fishing and clilate change

For each species, we quantified two measures of vulnerability: vulnerability to climate change and vulnerability to fishing pressure7. For species’ vulnerability to climate change, we solely focus on their vulnerability to the loss of live corals. Vulnerability to climate change is related to diet specialization, habitat specialization, recruitment specialization for live coral and body size7. Graham et al. (2011)7 developed a score for climate change vulnerability for 134 species. We used these scores to fit a Bayesian mixed effect predictive model that relates the vulnerability with the log-transformed maximum size of fish (extracted from Fishbase 5), the dependence on coral for food (3 categories: not dependent, facultative corallivore, and obligate corallivore), and dependence on coral for habitat (2 categories: dependent vs. not dependent). We also included a random effect for family. This model had a Bayesian R2 of 0.97. We then used this model to extrapolate the vulnerability measure to all 1110 species in our dataset. For species’ vulnerability to fishing, we extracted the index from Cheung et al. (2005) 8.

## References

1. Barneche, D. R. *et al.* Body size, reef area and temperature predict global reef-fish species richness across spatial scales. *Global Ecology and Biogeography* **28**, 315–327 (2019).

2. Schiettekatte, N. M. D. *et al.* Nutrient limitation, bioenergetics, and stoichiometry: a new model to predict elemental fluxes mediated by fishes. *Functional Ecology* (2020) doi:[10.1111/1365-2435.13618](https://doi.org/10.1111/1365-2435.13618).

3. Parravicini, V., Casey, J. M., Schiettekatte, N. M. D. & Brandl, S. J. Global gut content data synthesis and phylogeny delineate reef fish trophic guilds. *bioRxiv* 0–3 (2020) doi:[10.1101/2020.03.04.977116](https://doi.org/10.1101/2020.03.04.977116).

4. MacNeil, M. A. *et al.* Recovery potential of the world’s coral reef fishes. *Nature* **520**, 341–344 (2015).

5. Froese, R. & Pauly, D. FishBase. *World Wide Web electronic publication.* (2018).

6. Bürkner, P.-C. brms : An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software* **80**, 1–28 (2017).

7. Graham, N. A. *et al.* Extinction vulnerability of coral reef fishes. *Ecology Letters* **14**, 341–348 (2011).

8. Cheung, W. W., Pitcher, T. J. & Pauly, D. A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. *Biological Conservation* **124**, 97–111 (2005).