PNAS

Supporting Information Appendix for

A diverse portfolio of marine protected areas can better advance global conservation and equity

David A. Gill, Sarah E. Lester, Christopher M. Free, Alexander Pfaff, Edwin Iversen, Brian J Reich, Shu Yang, Gabby Ahmadia, Dominic A. Andradi-Brown, Emily Darling, Graham J. Edgar, Helen E. Fox, Jonas Geldmann, Duong Trung Le, Roosevelt Mesa-Gutiérrez, Michael B. Mascia, Peter J. Mumby, Laura Veverka, Laura M. Warmuth

Corresponding author: david.gill@duke.edu

The PDF file includes:

1 Study design

1.1 MPA impacts on fish populations (mechanisms)

The IUCN defines protected areas (PAs) as "a clearly defined geographical space, recognized, dedicated and managed, through legal or other effective means, to achieve the long-term conservation of nature with associated ecosystem services and cultural values" (1). Marine protected areas (MPAs) refer to "any area of intertidal or subtidal terrain, together with its overlying water and associated flora, fauna, historical and cultural features, which has been reserved by law or other effective means to protect part or all of the enclosed environment (2)." MPAs can lead to increased fish biomass by imposing restrictions that reduce extractive and destructive activities (primarily fishing) within the MPA boundaries, thus the magnitude and rate of fish population recovery (i.e., MPA impact) is highly dependent on the level of restrictions, the adequacy and appropriateness of management to facilitate sustainable use, and the pre-existing context (e.g., historic exploitation rates, species life history, habitat quality) (3–8).

1.2 Factors affecting MPA placement (selection criteria)

MPAs are not randomly placed in the ocean; many MPAs are situated in areas with low benefits from exploitation (9) or to protect areas with high biodiversity or tourism value (10, 11). In these cases, MPA fish populations would be expectedly higher than those in non-MPA locations regardless of protection, as non-MPA locations might have higher historical exploitation rates or lower habitat quality (e.g., lower fish abundance from low reef complexity (12)) than MPA locations. This may also be the case for MPA types, where protected areas that prohibit all extraction (no-take) are systematically placed in different locations than other protected area types (e.g., multiple-use MPAs) (13). Such differences were observed in terrestrial PAs, with more strict PAs located in areas where land conversion was less likely (14). Given that MPAs that prohibit all fishing (i.e., no-take MPAs) and MPAs that allow some fishing (multiple-use MPAs) incur different opportunity and social costs for extractive uses, we hypothesize that choice of MPA type can also be influenced by these opportunity costs and the associated fishing history. For example, no-take MPAs may be established (intentionally or unintentionally) in areas in better ecological condition (15) (e.g., higher coral cover in Florida Keys no-take zones (11)) or areas with fewer human stressors (e.g., areas less accessible or desirable for fishing). If so, no-take MPAs that are placed in relatively undisturbed areas where fish stocks are less threatened and in better condition may have greater standing biomass but a smaller impact compared to multiple-use MPAs placed in more heavily used areas (13).

1.3 Factors affecting outcomes

In addition to biases in no-take and multiple-use MPA/zone placement, any spatial comparison is likely affected by time-invariant and time-varying socio-environmental factors that differ between the treated and control locations. Social (e.g., traditional/cultural uses) (16), political (e.g., regulatory framework), economic (e.g., capital invested in marine resource activities) (17), and environmental (e.g., habitat types, ocean conditions, pollution levels) (18–20) conditions and trends can differ between sites, and account for much of the differences in fish populations between protected and unprotected sites (21, 22).

An ideal study design for evaluating MPA impacts uses comparable no-take, multiple-use, and non-MPA sites where long-term monitoring data on fish populations and socio-environmental conditions are collected for a significant duration before and after establishment (10, 22). However, such datasets are scarce or potentially non-existent, given the paucity of baseline and long-term MPA monitoring data for specific sites much less across regional or global scales (23, 24). It also requires considerable foresight and resources to collect data at control sites which account for biases in MPA placement (10, 21). In the absence of such data and recognizing its limitations (Section 6), we used our previously assembled global dataset of over 15,000 no-take, multiple-use, and non-MPA sites within and around over 287 MPAs in 43 countries and territories to estimate the relative impacts of MPAs that prohibit (no-take) and restrict (multiple-use) fishing. We also compiled detailed data on the social, environmental, management, and regulatory conditions at each study site, using matching and regression analyses to account for confounding factors that affect MPA placement and fish biomass outcomes (21, 25). We use MPAs as our unit of analysis or, in the case of MPAs with multiple-use and no-take zones, the zone type within the MPA.

A detailed methods plan for the initial iteration of this study was posted and publicly shared on the Open Science Framework platform in 2019 (26).

2 Step 1: Data compilation

Fig. S1 describes the analysis steps 1-3, and the corresponding sensitivity and diagnostic tests.

2.1 Fish biomass outcomes

We built on a global synthesized dataset of fish populations from Gill et al (6), with the full dataset comprising species or family level observations from 15,978 underwater visual census surveys conducted in and around 287 MPAs in 58 countries (see Gill et al (6) for data sources and description). Each survey site represents a single survey from a specific location at a particular period in time. We averaged transect level observations to get total biomass (grams) to units per 100m^2 for each site, based on the survey area and methodology, as the number of transects at each site varied by survey methodology. If no fish biomass data were provided by the data providers, we calculated fish biomass using the individual body lengths and allometric length-weight data obtained either from the data provider or from FishBase (27).

We used total biomass of all fish >10cm at each survey site (g/100m^2) as an indicator of conservation outcomes given its sensitivity to recovery from fishing, and its strong relationship to other conservation objectives such as species diversity (23, 28, 29). We only considered fish >10 cm to avoid recruitment effects, high variability associated with observing small cryptic species, and to focus the analysis on fishable species (30, 31). To reduce the effect of outliers, we use the natural log of the ratio of fish biomass of matched treatment and control sites as an indicator of MPA performance (i.e., treatment effects).

Given that not every site had fish >10cm (resulting in total biomass of 0 for that site; 997 survey sites), we followed Thiault et al (32) and Cresswell et al (24) by adding a small constant to biomass for all sites to avoid calculating log ratios with zero values (10). This constant represents the smallest total biomass value observed in the dataset (1.09 g/100m^2). To ensure that the choice of constant did not affect the results, we ran sensitivity tests with constants of different magnitudes (0.1 -10x constant). Such substantial changes in magnitude of the constant only marginally changed the estimated site-level average treatment effect (Fig. S6).

2.1.1 Unclear or ambiguous observations and sites

We removed species-level records with biomass values >200,000 g/100m^2. These were predominantly observations of large schools of mobile pelagic fish (e.g., *Scombrids, Sphyraenids*) or large transient species (e.g*., Manta birostris*) as it would be difficult to attribute these observations to protection at that site, particularly for small MPAs. We tested for the effect of schooling species by re-estimating the treatment effects after removing observations of fish with > 250 individuals. Removing these observations did not result in a significant change in the average estimated site-level treatment effects (Fig. S6). While our approach seeks to incorporate as much available data as possible, future studies could explore the sensitivity of results to wholly excluding species known to school.

Of the 15,978 survey sites in the dataset, we removed surveys where we were unable to calculate biomass due to missing data (n=65 sites). Habitat is an important confounder when assessing protection effects (19, 32–34), therefore, we removed survey sites without habitat data (n= 36 survey sites) and those conducted in rare habitats (e.g., pinnacles, ridges, back reefs; n=13 survey sites) given that these sites were unlikely to have an appropriate match. We also excluded sites where protection-level was unclear (n=158) or unclear how MPA protection might affect fin-fish (e.g., most whale sanctuaries). To reduce the likelihood of spillover affecting the results, we removed all survey sites within 1 km of the MPA boundary (n = 1,116 sites). Finally, to allow a lag time for fish response to protection, we removed sites in MPAs less than 3 years old (n = 850 sites). The final dataset used in the analysis included 14,044 sites (89.9% of original dataset) from within 335 zones in 216 MPAs (Fig. 1a) [1].

2.2 MPA regulations

In this study, we defined no-take as areas that do not allow any forms of fishing (subsistence, commercial, recreational) at any period of time. To identify the fishing regulations at each survey site, we compiled spatial and regulatory information on MPA and MPA zone boundaries. We extracted spatial and attribute data for these MPAs from the spatial datasets described in the "*MPA spatial data"* section. To identify fishing regulations, we compiled over one thousand documents and maps that described the activities permitted or prohibited in each zone of each MPA. Here we extracted information on zone names, area, and regulations, with a focus on fishing regulations to determine whether sites were located within a no-take or a multiple-use zone. We ranked each document or map based on 1) credibility of the source and 2) publication date. Here we prioritized resources from state or other management agencies (e.g., MPA gazettement document, local or national fishing regulations) or scientific publications and documents created within three years from the time when the fish surveys were conducted given the dynamic nature of MPA regulations. To further ensure reliability, we attempted to source at least two highly credible and relevant sources for each MPA, cross-validating information with other independent sources (e.g., World Database of Protected Areas (WDPA)), and reached out to local or regional experts for MPAs where we did not have high confidence in the provided information. While low confidence data does not mean that the information is incorrect (e.g., only one non-government source found, regulations not within three years of fish survey), as a sensitivity test, we re-estimated treatment effects after removing sites where we had low confidence in the fishing regulations (n=2,318 or 15.8% of final dataset) and this did not result in a significant change the average site-level treatment effects (Fig. S6).

MPA spatial data: We compiled spatial and regulatory information on MPA and MPA zone boundaries to identify the fishing regulations at each survey site, using MPA and regulatory data close in time to when surveys were conducted (see section above). We extracted spatial and attribute data for these MPAs from a larger spatial dataset of over 17,000 MPA and zone boundary shapefiles compiled by the authors and

---

[1] The sum of number of disqualified sites is larger than total number of sites removed as some sites were excluded for multiple reasons.

other research partners (see Acknowledgements). For this larger dataset, we first used the boundary files for MPAs from the November 2017 version of the World Database of Protected Areas (35), and validated or filled missing data using other national, regional, and global datasets. These datasets provided polygon information for MPAs in the Caribbean ((36), (37), (38)), Australia (39), United States mainland and overseas territories (37), and other locations around the world (e.g., MPA Atlas (40)). We focused on nationally designated marine and coastal MPAs and excluded large sanctuaries that spanned national waters or seas (e.g., Southern Ocean Whale Sanctuary), as: 1) most were established after the fish surveys were conducted; 2) given that regulations usually focused on non-fish species, their effect on fish populations were likely to be minimal or uniform across MPA and non-MPA survey sites. The final dataset comprised 17,122 MPA and zone boundary polygons.

We also used our compiled documents (see MPA regulations section) and other online sources to acquire spatial information on zone boundaries for our study MPAs (n=287 MPAs) that were not available in the above datasets. We used Google Earth and ESRI ArcGIS to georeference MPA zoning maps from the most credible and relevant sources available and create polygons for each MPA zone. We then imported the resulting 809 polygons into R and added the relevant regulation data for each zone. We plotted and compared each digitized zone polygon to sourced maps to ensure that they retained the correct spatial information after processing. This resulted in an additional 1,031 polygons containing zone and (outer) MPA boundaries for our study MPAs.

2.3 MPA management and governance

We used a dataset compiled by Gill et al (6) that describes the adequacy and appropriateness of management within 433 MPAs in 70 countries. These indicators, drawn from surveys completed by MPA management staff and/or other stakeholders, provide insight on the effectiveness (i.e., adequacy of management activities and capacities) and equity (i.e., fairness or justness of management) of management processes. Although these types of assessments are vulnerable to strategic biases (e.g., when conducted by a single MPA manager who may want to demonstrate high or low performance), other studies have shown that these data are effective at representing realities on the ground (41).

Assessment data took the form of ordinal responses to each question along with descriptive text responses. Gill et al (6) sourced and compiled these data from three management assessment tools: the Management Effectiveness Tracking Tool (METT) (42), the World Bank MPA Score Card (43), and the NOAA Coral Reef Conservation Program's (CRCP) MPA Management Assessment Checklist (44, 45), rescaling the ordinal response data to ensure construct validity across assessments (Table S4). In addition to the rescaled scores, following Gill et al (6), we developed binary thresholds for effective management for each indicator based on the scoring criteria and alignment with social theory (46–48). See Gill et al (6) for more information on data sourcing and processing and Table S4 for descriptions of response categories and thresholds.

3 Step 2: Site-level bias-adjusted treatment effects

3.1 Estimating fish biomass impacts

Estimating the relative effectiveness of these two policies requires considering several policy (treatment) scenarios. For example, to assess no-take MPA impacts, we will need to compare no-take fish biomass outcomes with outcomes if instead they were unprotected (here denoted by NT:0), as well as with outcomes if these same sites were inside a MU MPA instead (NT:MU|NT). Similar to Ferraro et al (13) and

Rasolofoson et al (49), we estimate relative impacts between no-take and multiple-use MPAs by comparing each MPA type to two counterfactual outcomes: no MPA and the alternative MPA type, resulting in four treatment effects (Table S6).

Here we define treatment as no-take or multiple-use MPA establishment. We also define MPA impacts as the expected difference in total fish biomass (Y) observed in treated $(Y_1 \mid T_i = 1)$ and counterfactual units $(Y_0 \mid T_i = 1)$ in treated locations (i.e. average treatment effect on the treated (ATT)), conditional on observed conditions X (Equation (1)) (50). See Table S6 for the treatment and counterfactual group for each comparison. The ATT for survey site *i* is:

$$ATT_i = E(Y_{i1} - Y_{i0} \mid T_i = 1, X_i) \quad (1)$$

where Y_i is total fish biomass at site *i*, T_i is treatment (0-control; 1- treated) and X_i are covariates.

3.2 Matching

We estimated the average treatment effect on the treated (ATT) for each comparison (Table S6) using statistical matching and regression adjustments. Here we identify comparable untreated sites to pair with treated sites based on the covariates (confounding factors: *X*) that affect both treatment and outcomes (25, 50) (Table S7). The assumption is, conditional on these observed factors, that treated and matched control units are interchangeable, and will on average have the same outcomes in the absence of treatment, making them an appropriate counterfactual (51, 52).

3.2.1 Matching covariates

Table S7 describes the covariates we used to match treated and control sites. We used exact matching for habitat type, country, ecoregion, and data source, ensuring that no treated site is paired with a control site from a different ecosystem, political or regulatory system, or sampled using a different methodology. This controlled for unobserved factors associated with these covariates (e.g., national fisheries policies, oceanographic conditions unique to particular ecoregions, etc.). For the remaining (continuous) covariates, we used nearest-neighbor matching based on Mahalanobis distances to identify control units for each treatment group, as this matching approach provided greater covariate balance with fewer dropped observations than other attempted approaches (see Section 5.1 for more details).

We used the Matching package v4.10-2 (53) in the R statistical software v4.2.0 (R Core Team, 2019) to match two control sites to each treated site (2:1 matching), reducing potential biases introduced from extreme observations when matching to a single control site (i.e., 1:1 matching). To improve match quality with the limited number of control sites, we matched with replacement, and permitted ties between matches (i.e., >1 control site with the same multivariate distance from a treatment site) where the weights of tied matches summed to one (53, 54). Given that human population density, distance to population center, and shore distance have been shown to strongly affect fish biomass outcomes and likely to affect MPA placement (due to their association with extractive uses such as fishing; (55, 56)), we applied calipers during the matching process to ensure that matched treatment and control sites did not differ in these covariates beyond one standard deviation. See Table S7 for more details on covariates and calipers.

3.3 Post-matching bias adjustment

Post-matching balance statistics suggest that matching significantly reduced the differences between treatment and control sites (Fig. S7 and Table S8), with the standardized mean differences of each covariate well below the recommended level of 0.25 or 25% (57). Nonetheless, some bias in treatment effects can remain due to residual differences in covariates between treated and control sites (58). We therefore used random forest models to identify and remove the post-matching differences in outcomes that can be explained by differences in the social and environmental conditions (i.e., covariates) between treated and control sites (59). Random forest models estimate the predicted values based on higher order interactions and non-parametric recursive partitioning over the n-dimensional covariate space (60). We chose random forests over linear approaches given their ability to predict in cases where there are heterogeneous, unscaled predictors with various non-linear functional forms and potential higher order interactions (61–64) making them better suited to model relationships within complex social-ecological systems.

We carried out three major steps in this approach. First, we used the *regression_forest* function in the *grf* package v2.1.0 (65) to model the relationships between the covariates and biomass outcomes in untreated (non-MPA) locations (to avoid influence of treatment effect), and used this model to predict fish biomass outcomes for all (treated and control) sites. Second, for each match set (i.e., comparisons in Table S6), we then calculated the difference between predicted biomass between treated and matched control units $\left(\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\right)$. This estimates the difference in outcomes that can be explained by the observed covariates. Lastly, we removed this predicted difference from our treatment effects to get the bias-adjusted values (Equation (2)). For the random forest models, we included a cluster level factor to account for similarity between sites of the same MPA or zone, and generated 10,000 trees in each forest to ensure stability. Given the sensitivity of random forest models to the choice of tuning parameters, we used the "*tune.parameters''* argument which uses cross-validation to identify the most appropriate parameters based on the data (100 tuning trees) (65). We also used "out of bag" prediction to improve accuracy (66).

$$ATT_{bias-adjuste\quad,i} = E(Y_{i1} - Y_{i0} | T_i = 1, X_i) - \left(\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\right) \qquad (2)$$

4 Step 3: MPA-level treatment effects

We used Bayesian linear hierarchical models to estimate the MPA-level average treatment effects for each estimand (Table S6). In each model, we used the site-level bias-adjusted treatment effects ($ATT_{bias-adjust\quad,i}$; Equation (2)) as the response variable. We used hierarchical models as they give inference on the mean treatment effect based on population of MPAs and not just on our sample (67) while allowing us to account for correlations induced by sampling multiple sites per MPA and re-use of control sites in the matching process. This also helped to account for the uneven distribution of sites per MPA which varied greatly (n=1 to 1,619 sites per MPA). For all models (Equations (5)-(8)), we report the model intercept α which represents the MPA-average treatment effect, and the probabilities of a positive absolute (NT:0, MU:0) or relative (NT:MU|NT, NT:MU|MU) impacts (Table S9). We also report percent biomass differences using the exponent of the model intercept α (Equation (3).

$$\text{percent biomass difference} = (\exp(\alpha) \times 100) - 100 \qquad (3)$$

We fitted the Bayesian models using JAGS software v4.3.0 (68) and R package Rjags v4-13 (69). To ensure model stability, we ran 50,000 iterations for two chains, with sufficient burn in (10,000 iterations) and then thinned (every 10th iteration) to reduce auto-correlation.

Our model evaluation efforts identified features in the data that required extensions of the standard linear mixed model formulation. First, both the random effects estimates and models' residuals had more extreme upper and lower quantiles than allowed for by normal distributions. We found that t-distributions provided a better fit and therefore adjusted the model to include t-distributed random effects and error terms to make the model more robust against outliers. In our most general model, we assumed response Y_i followed a Student t-distribution with location m_{i}, scale σ_{i} and d_1 degrees of freedom:

$$\mathrm{Y_i} \sim \mathrm{t}(m_{\mathrm{i}}, \sigma_{\mathrm{i}}, \mathrm{d}_1) \tag{4}$$

We considered four different models for the location that include different types of random effects:

$$m_{\mathrm{i}} = \alpha + \gamma_{\mathrm{MPA_i},i} + \left(\delta_{CON_{i1},i} + \delta_{CON_{i2},i}\right)/2 \tag{5}$$

or

$$m_{\mathrm{i}} = \alpha + \gamma_{\mathrm{MPA_i},i} + \left(\delta_{CON_{i1},i} + \delta_{CON_{i2},i}\right)/2 + \left(\theta_{CON\_MPA_{i1},i} + \theta_{CON\_MPA_{i2},i}\right)/2 \tag{6}$$

or

$$m_{\mathrm{i}} = \alpha + \beta_i + \gamma_{\mathrm{MP\ \ i},i} + \left(\delta_{CON_{i1},i} + \delta_{CON_{i2},i}\right)/2 \tag{7}$$

or

$$m_{\mathrm{i}} = \alpha + \beta_i + \gamma_{\mathrm{MPA_i},i} + \left(\delta_{CON_{i1},i} + \delta_{CON_{i2},i}\right)/2 + \left(\theta_{CON\_MPA_{i1},i} + \theta_{CON\_MPA_{i2},i}\right)/2 \tag{8}$$

The scale was modeled as

$$\log(\sigma_i) = a_1 + a_2 log(N_i). \tag{9}$$

α represents the MPA treatment effect of interest. β is the binary fixed effect term for the population center distance, capacity, and regulations models (Equations (7) and (8); see Table S3 and Table S4 for distance, capacity, and regulations definitions and other details), MPA_i is the index of the MPA for observation i, CON_{i1} and CON_{i2} are the indices of the two controls matched to treated observation i and N_i is the number of samples drawn from observation $i's$ MPA (Equation (9)). The random effects γ_j and δ_k were included to account for dependence between observations that share an MPA and control respectively, and are modeled as $\gamma_{\mathrm{j}} \sim \mathrm{t}(0, \tau_1, \mathrm{d}_2)$ and $\delta_k \sim t(0, \tau_2, d_2)$. For the direct no-take to multiple-use NT:MU comparisons, we also included a random intercept θ_{m} for control sites from the same MPA (Equations (6) and (8)). While other random effects structures were possible (e.g., random intercepts for each MPA zone, country, or ecoregion, nested random effects, etc.), after running various iterations of these models, Equations (5)-(8) represented the most parsimonious structure that best represented the data.

We used uninformative prior distributions for the remaining parameters, d_1, d_2 ~ Uniform(1,30), α ~ Normal(0,100000), a_1, a_2~ Normal(0,1), and τ_1, τ_2 ~ Half-Cauchy(0,1.8). The Student t distributions are used to allow for extremely large and small observations, and the log scale modeled in terms of MPA sample size (Equation (9)) explains observed heteroskedasticity given that the error variance increased with the number of sites per MPA.

5 Sensitivity tests

5.1 Matching sensitivity tests

Match quality can vary greatly depending on multiple factors: covariate specification, number of matched controls per treated unit, choice of calipers, and more (53, 54). To identify the best approach for our data, we assessed the match quality of multiple specifications using post-matching balance statistics to identify the specification that had better covariate balance (e.g., lower standardized mean difference and quantile-quantile differences, etc.; Table S8) and fewer dropped "unmatched" treated observations. Tested specifications included nearest-neighbor matching on 1) all covariates, 2) propensity scores, 3) propensity scores and covariates strongly correlated with treatment and outcomes (i.e., human population density, shore distance, market distance), and 4) propensity scores and all covariates. We also assessed the post-matching statistics with two and three controls matched to each treated site, more and less strict caliper widths on influential covariates (e.g., maximum shore distance difference of 0.5 or 1 SD), and transformations on covariates with skewed distributions. Of the specifications attempted, nearest-neighbor matching on all untransformed covariates with two control sites per treated unit achieved the best covariate balance and fewest dropped observations. To account for potential interactions between covariates, we also added a quadratic term for some covariates and estimated counterfactual conditions using random forest models (which account for higher order interactions (61)) instead of matching. These approaches did not improve match quality.

5.2 Unobserved bias

While matching reduces the biases between treated and control units based on observed covariates, MPA placement could be determined by other unaccounted or unobserved factors. For example, if implementors systematically place no-take or multiple-use MPAs in specific locations for reasons other than those included in the list of covariates (Table S7), social-ecological conditions and thus fish biomass outcomes may differ greatly between no-take, multiple-use, and unprotected sites even without protection. However, given that these factors are mostly unknown, they are impossible to identify or directly measure (25). Here we employed a confounding sensitivity test, adapted from Blackwell (70), to ascertain the sensitivity of our results to unknown and unaccounted confounding factors that might explain differences in MPA placement. This test has advantages over other sensitivity tests given that: 1) it provides information on the magnitude and direction of the confounding needed to overturn inference (e.g., assuming positive impacts when they are actually negative); and 2) it is independent of the number and types of unknown confounders or estimation strategy (70). In this test, we examined the sensitivity of our results to unknown confounding by examining how the magnitude and direction of each treatment effect varies when exposed to varying amounts of hypothetical confounding. Here we created scenarios where unknown factors (U) that affect MPA placement have half up to two ($U = \alpha = 0.5 - 2$) times the effect on fish biomass than all the observed covariates (e.g., shore distance, depth, etc.). We did this in four steps. First, we calculated the bias in treatment effects explained by the observed covariates. We did this by using regression forest models to predict fish biomass outcomes based on observed covariates for treated ($\hat{\mu}_1(X_i)$) and control ($\hat{\mu}_0(X_i)$) units separately. We used regression forests to account for non-linear functional forms and potential higher order interactions between variables. Here $obs.effect_i$ (Equation (10)) represents the difference in fish biomass outcomes between control and treatment sites that can be explained by the remaining post-matching differences in covariates.

$$obs.effect_i = \hat{\mu}_1(X_i) - \hat{\mu}_0(X_i). \tag{10}$$

Second, we computed a range of hypothetical confounding estimates based on the $obs.effect_i$ ($\alpha \times obs.effect_i$; where $0.5 < \alpha < 2$) and added these values to control outcomes (Equation (11)). This simulated a potential confounding that increases or decreases counterfactual outcomes where negative α values (i.e., biasing control outcomes downward) suggest that treated sites are placed in locations that inherently have greater fish biomass than control locations.

$$Y_i^q = Y_i + (1 - T_i)\alpha \times obs.effect_i P[T_i = 1 \mid X_i] \tag{11}$$

where $T_i = 1$ in treated and 0 in control sites. See Figure S8 for sensitivity test results.

Third, we re-estimated the bias-adjusted ATT treatment effects using the confounded control outcomes Y_i^q. As expected, our hypothetical results suggested that if such confounding factors exist, our estimates of MPA impacts would decrease with increasing amounts of (positive) confounding, especially the estimates of the absolute impacts of no-take MPAs (NT:0; Fig. S8). In other words, of all the estimands, the estimate of no-take MPA impacts were most sensitive to selection bias (e.g., positive impacts disappear if the effect of unknown factors on biomass outcomes are ~75% of the effect from observed covariates). As mentioned before, whether such biases exist, or would have such a strong effect on outcomes that were close to or beyond those of the observed covariates is unknown and unmeasurable. This is unlikely, as such unknown placement biases would have to be systemic, with a consistently strong effect on numerous MPAs globally. Furthermore, by using exact matching, we account for unobserved biases introduced by social, political, and ecological factors (e.g., selecting control sites from the same country only; Table S7), and the included observed covariates are some of the most well-known factors in the literature to frequently affect MPA placement and fish biomass (e.g., neighboring human population density (19, 28, 55, 56); Table S7).

6 Study limitations

We recognize that our sample does not represent a random selection of the global pool of MPAs given that we relied on open-source data or those shared by data managers. As a result, our sample mostly comprised data from tropical coral and temperate reef locations (Fig 1a). Limitations in management data methodologies and availability also confined some of our analyses to specific locations (e.g., offshore territories or countries receiving development funding) and thus inference (71, 72). As mentioned in Section 1, the absence of baseline and trend data due to data limitations can result in confounded estimates if all major observed and unobserved factors that contribute to site selection are not accounted for. Also, while exact matching and calipers reduced numerous observed and unobserved biases, the matching process also resulted in a considerable number of dropped observations (n=14-46% of treatment sites depending on the estimand). For example, there were countries/regions where only MPA sites were surveyed (e.g., Galápagos), and thus excluded from estimates of absolute impacts (NT:0 and MU:0) as no appropriate non-MPA match could be found. We examined whether the above limitations might have had a strong influence on our results (e.g., if post-matching subsamples were biased towards specific social, environment, or geographic contexts or whether treatment effects varied significantly between MPAs with and without management data). As expected, sub-sampling resulted in some geographic and contextual differences. For example, MPAs used to assess relative impacts were larger, younger, and further away from shore and population centers than those in the absolute impacts sample. However, given that relative impacts analysis directly matches sites by MPA type (e.g., matching remote no-take with remote multiple-use MPA sites) and not MPA and non-MPA sites, and that these covariates have disparate effects on outcomes (4, 17, 73), how these covariate differences influences the results is not clear. Additionally, differences in sample sizes (and thus variation) made it not possible to determine if treatment effects differed between samples with or without management data. While sample selection bias is an issue in all observational studies (74) and our dataset is one of the largest of its kind, we do not claim that our results are representative of all MPAs, but for the MPAs in our sample. These MPAs are located in regions with high biodiversity, subject to multiple local and global stressors and thus represent high-priority areas for conservation (75–77). Further, we applied modeling approaches appropriate for unbalanced data (Section 4), accounting for confounding factors (Section 3), and for population-level inference based on samples (i.e., random-effects models; (Section 4) (67)).

7 References

1. J. Day, *et al.*, "Guidelines for applying the IUCN protected area management categories to marine protected areas. Second edition." (IUCN, 2019).

2. IUCN-WCPA, "Establishing Marine Protected Area Networks–Making It Happen" (IUCN World Commission on Protected Areas (IUCN-WCPA), National Oceanic and Atmospheric Administration and The Nature Conservancy, 2008).

3. N. J. Bennett, P. Dearden, Why local people do not support conservation: Community perceptions of marine protected area livelihood impacts, governance and management in Thailand. *Marine Policy* **44**, 107–116 (2014).

4. G. J. Edgar, *et al.*, Global conservation outcomes depend on marine protected areas with five key features. *Nature* **506**, 216–220 (2014).

5. W. J. Fletcher, R. E. Kearney, B. S. Wise, W. J. Nash, Large-scale expansion of no-take closures within the Great Barrier Reef has not enhanced fishery production. *Ecological Applications* **25**, 1187–1196 (2015).

6. D. A. Gill, *et al.*, Capacity shortfalls hinder the performance of marine protected areas globally. *Nature* **543**, 665–669 (2017).

7. K. Grorud-Colvert, *et al.*, The MPA Guide: A framework to achieve global goals for the ocean. *Science* **373**, eabf0861 (2021).

8. M. J. Rees, *et al.*, Accounting for habitat structural complexity improves the assessment of performance in no-take marine reserves. *Biological Conservation* **224**, 100–110 (2018).

9. R. Devillers, *et al.*, Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquatic Conservation: Marine and Freshwater Ecosystems* **25**, 480–504 (2015).

10. C. Osenberg, J. Shima, S. Miller, A. Stier, "Assessing effects of marine protected areas: confounding in space and possible solutions" in *Marine Protected Areas: A Multidisciplinary Approach*, J. Claudet, Ed. (Cambridge University Press, 2011), pp. 143–167.

11. L. T. Toth, *et al.*, Do no-take reserves benefit Florida's corals? 14 years of change and stasis in the Florida Keys National Marine Sanctuary. *Coral Reefs* **33**, 565–577 (2014).

12. S. K. Wilson, *et al.*, Relative and combined effects of habitat and fishing on reef fish communities across a limited fishing gradient at Ningaloo. *Marine environmental research* **81**, 1–11 (2012).

13. P. J. Ferraro, *et al.*, More strictly protected areas are not necessarily more protective: evidence from Bolivia, Costa Rica, Indonesia, and Thailand. *Environmental Research Letters* **8**, 025011 (2013).

14. L. N. Joppa, A. Pfaff, High and Far: Biases in the Location of Protected Areas. *PLoS ONE* **4**, e8273 (2009).

15. K. a. Karr, *et al.*, Thresholds in Caribbean coral reefs: implications for ecosystem-based fishery management. *Journal of Applied Ecology* **52**, 402–412 (2015).

16. J.-B. Jouffray, *et al.*, Parsing human and biophysical drivers of coral reef regimes. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20182544 (2019).

17. J. E. Cinner, *et al.*, Gravity of human impacts mediates coral reef conservation gains. *Proceedings of the National Academy of Sciences* **115**, E6116–E6125 (2018).

18. G. P. Fioravanti, J. F. Valentine, Impacts of reef architecture and fishing on food webs in back reef environments of the upper Florida Keys National Marine Sanctuary. *Florida Scientist* **70**, 18 (2007).

19. A. M. Friedlander, E. K. Brown, P. L. Jokiel, W. R. Smith, K. S. Rodgers, Effects of habitat, wave exposure, and marine protected area status on coral reef fish assemblages in the Hawaiian archipelago. *Coral Reefs* **22**, 291–305 (2003).

20. I. D. Williams, *et al.*, Differences in Reef Fish Assemblages between Populated and Remote Reefs Spanning Multiple Archipelagos Across the Central and Western Pacific. *Journal of Marine Biology* **2011**, 1–14 (2011).

21. G. N. Ahmadia, *et al.*, Integrating impact evaluation in the design and implementation of monitoring marine protected areas. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140275 (2015).

22. P. J. Ferraro, J. N. Sanchirico, M. D. Smith, Causal inference in coupled human and natural systems. *Proceedings of the National Academy of Sciences* **116**, 201805563 (2018).

23. R. A. Abesamis, A. L. Green, G. R. Russ, C. R. L. Jadloc, The intrinsic vulnerability to fishing of coral reef fishes and their differential recovery in fishery closures. *Reviews in Fish Biology and Fisheries* **24**, 1033–1063 (2014).

24. A. K. Cresswell, *et al.*, Disentangling the response of fishes to recreational fishing over 30 years within a fringing coral reef reserve network. *Biological Conservation* **237**, 514–524 (2019).

25. P. R. Rosenbaum, *Design of Observational Studies* (Springer, 2010).

26. Gill, David A, OSF | Relative impacts of NT and fished MPAs; https://osf.io/pxrf8/. *Open Science Framework* (2019) (May 1, 2019).

27. R. Froese, D. Pauly, Fishbase. *World Wide Web electronic publication* (2010) (December 1, 2010).

28. T. R. McClanahan, Biogeography versus resource management: how do they compare when prioritizing the management of coral reef fish in the south-western Indian Ocean? *Journal of Biogeography* **42**, 2414–2426 (2015).

29. K. L. Nash, N. A. J. Graham, Ecological indicators for coral reef fisheries management. *Fish and Fisheries* **17**, 1029–1054 (2016).

30. T. R. McClanahan, Community biomass and life history benchmarks for coral reef fisheries. *Fish and Fisheries* **19**, 471–488 (2018).

31. S. K. Wilson, *et al.*, Habitat degradation and fishing effects on the size structure of coral reef fish communities. *Ecological Applications* **20**, 442–451 (2010).

32. L. Thiault, *et al.*, Ecological evaluation of a marine protected area network: a progressive-change BACIPS approach. *Ecosphere* **10**, e02576 (2019).

33. J. a. García-Charton, *et al.*, Multi-scale spatial heterogeneity, habitat structure, and the effect of marine reserves on Western Mediterranean rocky reef fish assemblages. *Marine Biology* **144**, 161–182 (2004).

34. V. Stelzenmüller, F. Maynou, P. Martín, Spatial assessment of benefits of a coastal Mediterranean Marine Protected Area. *Biological Conservation* **136**, 571–583 (2007).

35. IUCN and UNEP-WCMC, The World Database on Protected Areas (WDPA). Accessed April 2019 (2019) (April 21, 2019).

36. TNC, CAMPAM, CaribbeanMPA: A database of the Wider Caribbean's marine protected areas (2014).

37. NOAA, The Marine Protected Areas Inventory (2014).

38. Healthy Reefs, Healthy Reefs Eco-Audit data and maps (2014).

39. Dept. of Environment, Collaborative Australian Protected Areas Database (CAPAD) (2016).

40. Marine Conservation Institute, MPAtlas (2015) (November 17, 2015).

41. J. Geldmann, *et al.*, Changes in protected area management effectiveness over time: A global analysis. *Biological Conservation* **191**, 692–699 (2015).

42. S. Stolton, *et al.*, "Management Effectiveness Tracking Tool: Reporting progress in Protected areas sites; second edition" (World Bank/WWF Forest Alliance and WWF, 2007).

43. F. Staub, M. E. Hatziolos, "Score Card to Assess Progress in Achieving Management effectiveness goals for Marine Protected Areas" (Prepared for the World Bank, 2004).

44. NOAA, "NOAA Coral Reef Conservation Program MPA Management Assessment Checklist" (National Oceanic and Atmospheric Administration (NOAA), 2010).

45. M. Gombos, *et al.*, "A Management Capacity Assessment of Selected Coral Reef Marine Protected Areas in the Caribbean" (Commissioned by the National Oceanic and Atmospheric Administration (NOAA) Coral Reef Conservation Program (CRCP), the Gulf and Caribbean Fisheries Institute

(GCFI) and by the UNEP-CEP Caribbean Marine Protected Area Management Network and Forum (CaMPAM), 2011).

46. M. B. Mascia, “Social Dimensions of Marine Reserves” in *Marine Reserves: A Guide to Science, Design, and Use*, J. Sobel, C. Dahlgren, Eds. (Island Press, 2004), pp. 164–186.

47. E. Ostrom, *Governing the commons: the evolution of institutions for collective action* (Cambridge University Press, 1990).

48. C. Scianna, F. Niccolini, S. D. Gaines, P. Guidetti, ‘Organization Science’: A new prospective to assess marine protected areas effectiveness. *Ocean & Coastal Management* **116**, 443–448 (2015).

49. R. A. Rasolofoson, P. J. Ferraro, C. N. Jenkins, J. P. G. Jones, Effectiveness of Community Forest Management at reducing deforestation in Madagascar. *Biological Conservation* **184**, 271–277 (2015).

50. J. S. Sekhon, Multivariate and Propensity Score Matching. *Journal of Statistical Software* **42**, 1–52 (2011).

51. P. J. Ferraro, M. M. Hanauer, Advances in Measuring the Environmental and Social Impacts of Environmental Programs. *Annual Review of Environment and Resources* **39**, 495–517 (2014).

52. J. S. Sekhon, Opiates for the Matches: Matching Methods for Causal Inference. *Annual Review of Political Science* **12**, 487–508 (2009).

53. J. S. Sekhon, Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R (version 4.9-6). *Journal of Statistical Software* **42**, 1–52 (2011).

54. M. Caliendo, S. Kopeinig, Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* **22**, 31–72 (2005).

55. T. D. Brewer, J. E. Cinner, a Green, R. L. Pressey, Effects of human population density and proximity to markets on coral reef fishes vulnerable to extinction by fishing. *Conservation biology : the journal of the Society for Conservation Biology* **27**, 443–52 (2013).

56. J. E. Cinner, N. A. J. Graham, C. Huchery, M. A. Macneil, Global effects of local human population density and distance to markets on the condition of coral reef fisheries. *Conservation Biology* **27**, 453–8 (2013).

57. E. A. Stuart, Matching Methods for Causal Inference: A Review and a Look Forward. *Statist. Sci.* **25**, 1–21 (2010).

58. A. Abadie, G. W. Imbens, Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics* **29**, 1–11 (2011).

59. R. Y. Fidler, *et al.*, Participation, not penalties: Community involvement and equitable governance contribute to more effective multiuse protected areas. *Science Advances* **8**, eabl8929 (2022).

60. C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, A. Zeileis, Conditional Variable Importance for Random Forests. *BMC Bioinformatics* **9**, 307 (2008).

61. C. Strobl, J. Malley, G. Tutz, An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological methods* **14**, 323–348 (2009).

62. T. Hothorn, K. Hornik, A. Zeileis, Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* **15**, 651–674 (2006).

63. D. R. Cutler, *et al.*, Random Forests for Classification In Ecology. *Ecology* **88**, 2783–2792 (2007).

64. A. Hapfelmeier, T. Hothorn, K. Ulm, C. Strobl, A new variable importance measure for random forests with missing data. *Statistics and Computing* **24**, 21–34 (2012).

65. J. Tibshirani, S. Athey, S. Wager, grf: Generalized Random Forests (2020).

66. M. Lu, S. Sadiq, D. J. Feaster, H. Ishwaran, Estimating Individual Treatment Effect in Observational Data Using Random Forest Methods. *Journal of Computational and Graphical Statistics* **27**, 209–219 (2018).

67. J. C. Pinheiro, D. M. Bates, *Mixed effects models in S and S-PLUS* (Springer Verlag, 2000).

68. M. Plummer, JAGS: A Program for Analysis of Bayesian Graphical Models using Gibbs Sampling. *3rd International Workshop on Distributed Statistical Computing (DSC 2003); Vienna, Austria* **124** (2003).

69. M. Plummer, A. Stukalov, M. Denwood, rjags: Bayesian Graphical Models using MCMC (2023) (July 13, 2023).

70. M. Blackwell, A Selection Bias Approach to Sensitivity Analysis for Causal Effects. *Polit. anal.* **22**, 169–182 (2014).

71. C. N. Cook, R. W. (Bill) Carter, M. Hockings, Measuring the accuracy of management effectiveness evaluations of protected areas. *Journal of Environmental Management* **139**, 164–171 (2014).

72. C. N. Cook, M. Hockings, Opportunities for improving the rigor of management effectiveness evaluations in protected areas. *Conservation Letters* **4**, 372–382 (2011).

73. S. J. Campbell, *et al.*, Fishing restrictions and remoteness deliver conservation outcomes for Indonesia’s coral reef fisheries. *Conservation Letters* **13**, e12698 (2020).

74. D. E. Ho, K. Imai, G. King, E. a. Stuart, Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* **15**, 199–236 (2007).

75. L. Burke, M. Spalding, K. Reytar, A. Perry, “Reefs at risk revisited” (World Resources Institute, 2011).

76. T. P. Hughes, *et al.*, Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science* **359**, 80–83 (2018).

77. E. R. Selig, *et al.*, Global priorities for marine biodiversity conservation. *PloS one* **9**, e82898 (2014).

78. Flanders Marine Institute, MarineRegions.org (2021) (October 17, 2017).

79. NCCOS, Benthic Habitat Mapping of Coral Reefs in Florida, Caribbean, and the Pacific; National Oceanic and Atmospheric Administration (NOAA) National Centers for Coastal Ocean Science (NCCOS) (2014) (April 20, 2006).

80. UNEP-WCMC, WorldFish Center, WRI, TNC, "Global distribution of coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project. Version 1.3." (Includes contributions from IMaRS-USF and IRD (2005), IMaRSUSF (2005) and Spalding et al. (2001). Cambridge (UK): UNEP World Conservation Monitoring Centre., 2010).

81. L. Tyberghein, *et al.*, Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* **21**, 272–281 (2012).

82. SOEST, 50 Meter Bathymetry Grids (GMT and Arc). School of Ocean and Earth Science and Technology, University of Hawaii (2014) (September 5, 2014).

83. NOAA, U.S. Virgin Islands, VI 1 arc-second MHW DEM. National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) (2014) (September 5, 2014).

84. NOAA, US Coastal Relief Model - Puerto Rico. National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) (2014) (September 5, 2014).

85. H. L. Tolman, "User manual and system documentation of WAVEWATCH III version 3.14, Technical Note" (U.S. Department of Commerce, National Oceanographic and Atmospheric Administration, National Weather Service, National Centers for Environmental Predictions, 2009).

86. P. Marchand, D. A. Gill, waver. Calculate fetch and wave exposure (R package) (2016).

87. M. D. Spalding, *et al.*, Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. *BioScience* **57**, 573–583 (2007).

88. P. Wessel, W. H. F. Smith, A global, self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research* **101**, 8741 (1996).

89. CIESIN, CIAT, Center for International Earth Science Information Network - CIESIN - Columbia University, and Centro Internacional de Agricultura Tropical - CIAT. 2005. Gridded Population of the World, Version 3 (GPWv3): Population Density Grid. Palisades, NY: NASA Soci (2014) https:/doi.org/dx.doi.org/10.7927/H4XK8CG2 (August 9, 2014).

90. ESRI, World Cities. Esri Geographic Information System company (2015) (March 6, 2015).

91. M. D. Spalding, I. Meliane, A. Milam, C. Fitzgerald, L. Z. Hale, Protecting Marine Spaces: Global Targets and Changing Approaches. *Ocean Yearbook 27*, 213–248 (2013).

92. G. P. Jones, C. Syms, Disturbance, habitat structure and the ecology of fishes on coral reefs. *Austral Ecology* **23**, 287–297 (1998).

93. D. Pelletier, D. Mallet, A. Powel, W. Roman, "Disentangling the Effects of Habitat and Protection on Coral Reef Fish Communities in Long-Established Marine Reserves" in *Formal Methods in Macro-Biology*, (2014), pp. 156–158.

94. S. Williams, P. Mumby, I. Chollett, J. Cortés, Importance of differentiating Orbicella reefs from gorgonian plains for ecological assessments of Caribbean reefs. *Marine Ecology Progress Series* **530**, 93–101 (2015).

95. M. Dallimer, N. Strange, Why socio-political borders and boundaries matter in conservation. *Trends in Ecology & Evolution* **30**, 1–8 (2015).

96. M. O. Nadon, *et al.*, Re-Creating Missing Population Baselines for Pacific Reef Sharks. *Conservation Biology* **26**, 493–503 (2012).

97. J. M. Gove, *et al.*, Quantifying Climatological Ranges and Anomalies for Pacific Coral Reef Ecosystems. *PLoS ONE* **8**, e61974 (2013).

98. J. D. Bell, Effects of Depth and Marine Reserve Fishing Restrictions on the Structure of a Rocky Reef Fish Assemblage in the North-Western Mediterranean Sea. *Journal of Applied Ecology* **20**, 357–369 (1983).

99. I. Chollett, P. J. Mumby, Predicting the distribution of Montastraea reefs using wave exposure. *Coral Reefs* **31**, 493–503 (2012).

100. N. T. Shears, R. C. Babcock, A. K. Salomon, Context-Dependent Effects of Fishing: Variation in Trophic Cascades across Environmental Gradients. *Ecological Applications* **18**, 1860–1873 (2008).

101. I. Chollett, S. W. J. Canty, S. J. Box, P. J. Mumby, Adapting to the impacts of global change on an artisanal coral reef fishery. *Ecological Economics* **102**, 118–125 (2014).

102. F. Bastardie, J. R. Nielsen, B. S. Andersen, O. R. Eigaard, Integrating individual trip planning in energy efficiency - Building decision tree models for Danish fisheries. *Fisheries Research* **143**, 119–130 (2013).

103. J. E. Wilen, M. D. Smith, D. Lockwood, L. W. Botsford, Avoiding surprises: incorporating fisherman behavior into management models. *Bulletin of Marine Science* **70**, 553–575 (2002).

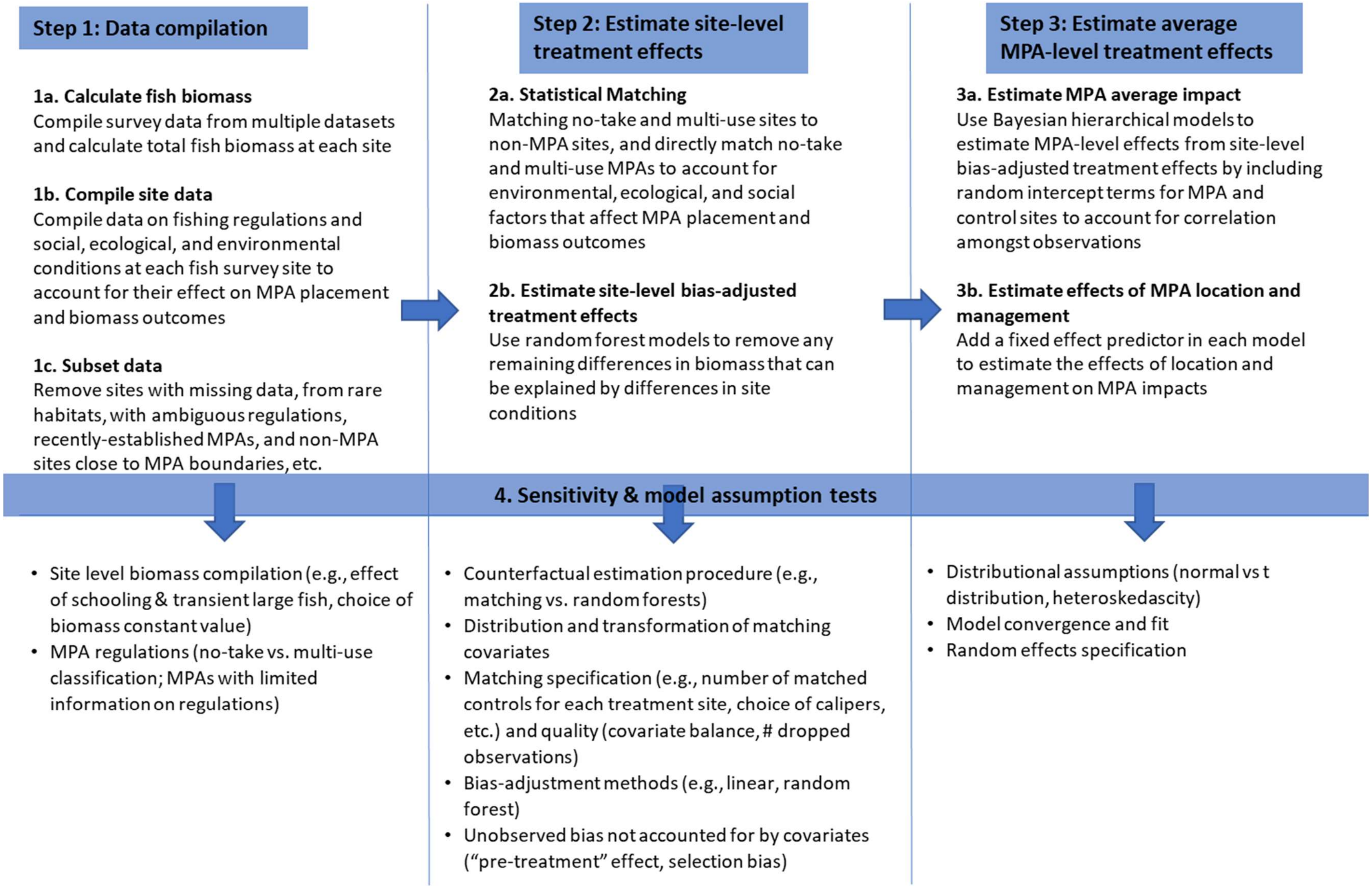
104. A. C. Haynie, L. Pfeiffer, Why economics matters for understanding the effects of climate change on fisheries. *ICES Journal of Marine Science* **69**, 1160–1167 (2012).

105. K. Newton, I. M. Côté, G. M. Pilling, S. Jennings, N. K. Dulvy, Current and future sustainability of island coral reef fisheries. *Current Biology* **17**, 655–658 (2007).

106. S. Advani, L. N. Rix, D. M. Aherne, M. A. Alwany, D. M. Bailey, Distance from a fishing community explains fish abundance in a no-take zone with weak compliance. *PLoS ONE* **10**, e0126098 (2015).

107. A. R. Harborne, *et al.*, Modelling and mapping regional-scale patterns of fishing impact and fish stocks to support coral-reef management in Micronesia. *Diversity and Distributions* **24**, 1729–1743 (2018).

108. J. Claudet, *et al.*, Marine reserves: size and age do matter. *Ecology Letters* **11**, 481–489 (2008).



Step 1: Data compilation

1a. Calculate fish biomass
Compile survey data from multiple datasets and calculate total fish biomass at each site

1b. Compile site data
Compile data on fishing regulations and social, ecological, and environmental conditions at each fish survey site to account for their effect on MPA placement and biomass outcomes

1c. Subset data
Remove sites with missing data, from rare habitats, with ambiguous regulations, recently-established MPAs, and non-MPA sites close to MPA boundaries, etc.

Step 2: Estimate site-level treatment effects

2a. Statistical Matching
Matching no-take and multi-use sites to non-MPA sites, and directly match no-take and multi-use MPAs to account for environmental, ecological, and social factors that affect MPA placement and biomass outcomes

2b. Estimate site-level bias-adjusted treatment effects
Use random forest models to remove any remaining differences in biomass that can be explained by differences in site conditions

Step 3: Estimate average MPA-level treatment effects

3a. Estimate MPA average impact
Use Bayesian hierarchical models to estimate MPA-level effects from site-level bias-adjusted treatment effects by including random intercept terms for MPA and control sites to account for correlation amongst observations

3b. Estimate effects of MPA location and management
Add a fixed effect predictor in each model to estimate the effects of location and management on MPA impacts

4. Sensitivity & model assumption tests

- Site level biomass compilation (e.g., effect of schooling & transient large fish, choice of biomass constant value)
- MPA regulations (no-take vs. multi-use classification; MPAs with limited information on regulations)

- Counterfactual estimation procedure (e.g., matching vs. random forests)
- Distribution and transformation of matching covariates
- Matching specification (e.g., number of matched controls for each treatment site, choice of calipers, etc.) and quality (covariate balance, # dropped observations)
- Bias-adjustment methods (e.g., linear, random forest)
- Unobserved bias not accounted for by covariates ("pre-treatment" effect, selection bias)

- Distributional assumptions (normal vs t distribution, heteroskedascity)
- Model convergence and fit
- Random effects specification

Fig. S1 | Major analytical steps and diagnostic tests used in this study. Step 4 (lower panels) show the sensitivity and model assumptions tests conducted in association with each of the three steps.

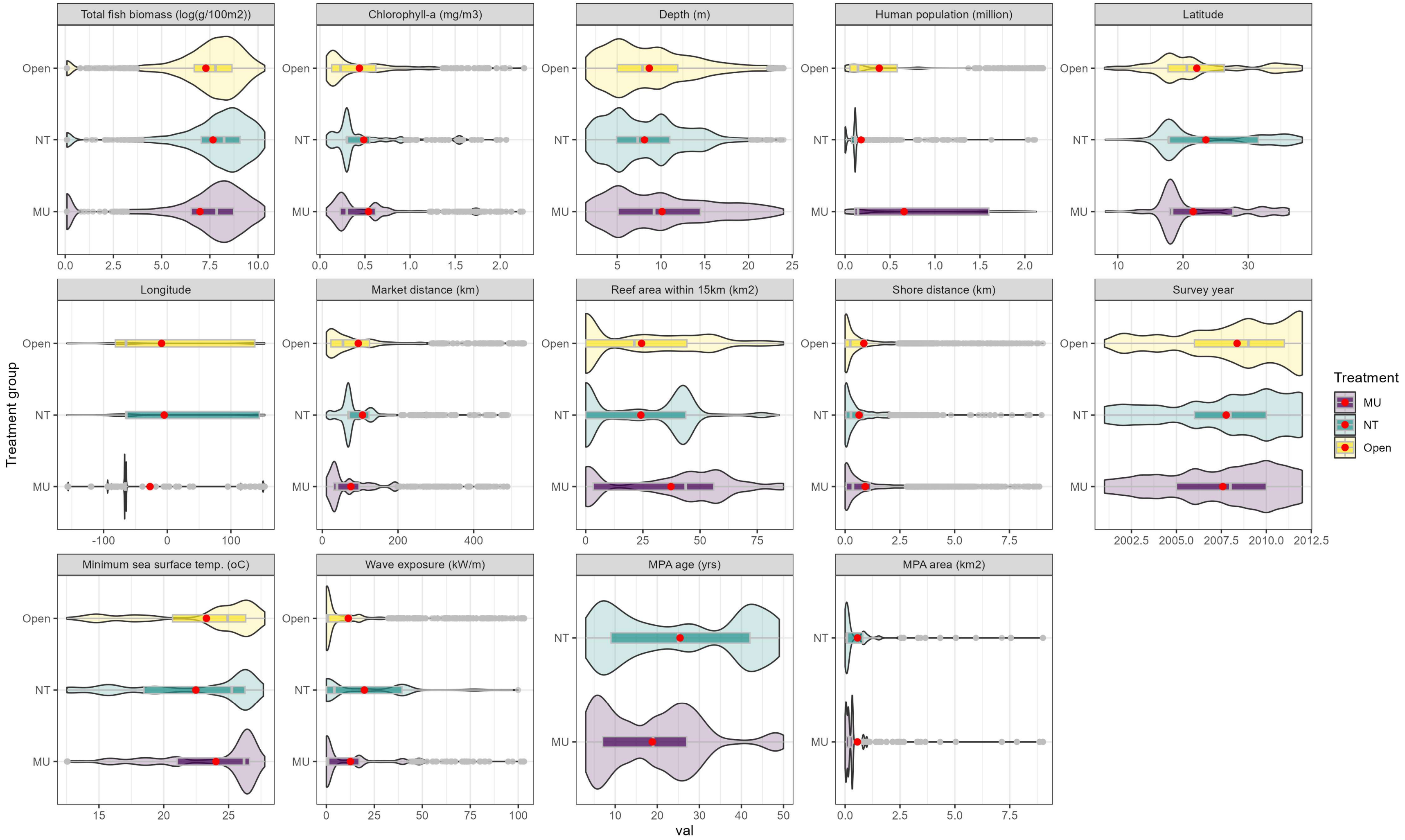


Fig. S2 | Covariate distribution by treatment group: MU: multiple-use, NT: no-take, and Open: non-MPA sites (n= 14,044 sites). Box represents the lower (25%) and upper (75%) quantile. Mean (red dot), median (vertical line in box), and 1.5 times the interquartile range (horizontal line) also shown. For visualization purposes, plots exclude observations below the 5th and above the 95th percentiles. See Table S3 variable descriptions.

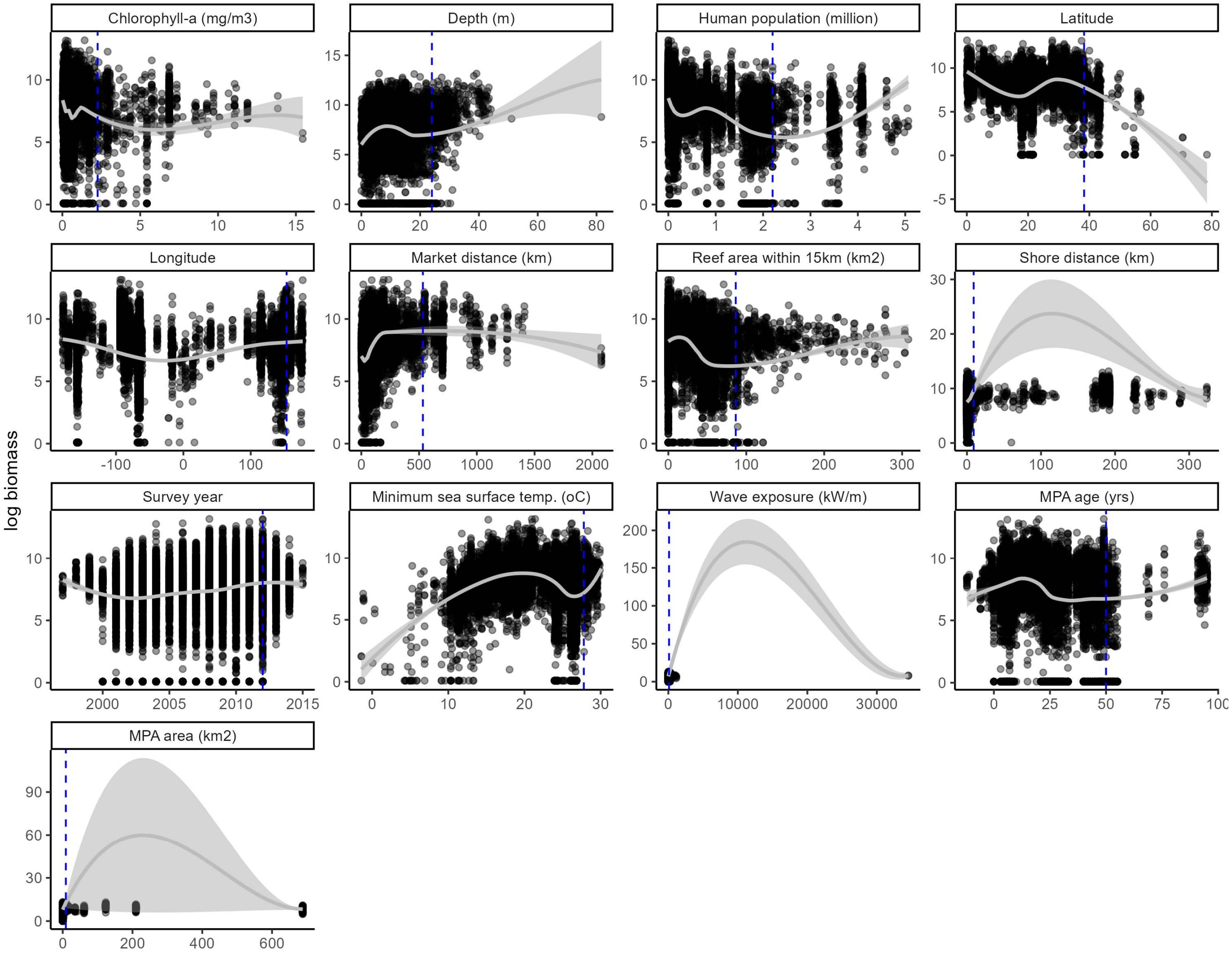


Fig. S3 | Relationship between fish biomass (log) and matching covariates (n= 14,044 sites). Included are smoothed LOESS lines (grey line) along with the standard error regions (shaded grey area). Dashed vertical blue lines represent the 95th percentile. See Table S3 variable descriptions.

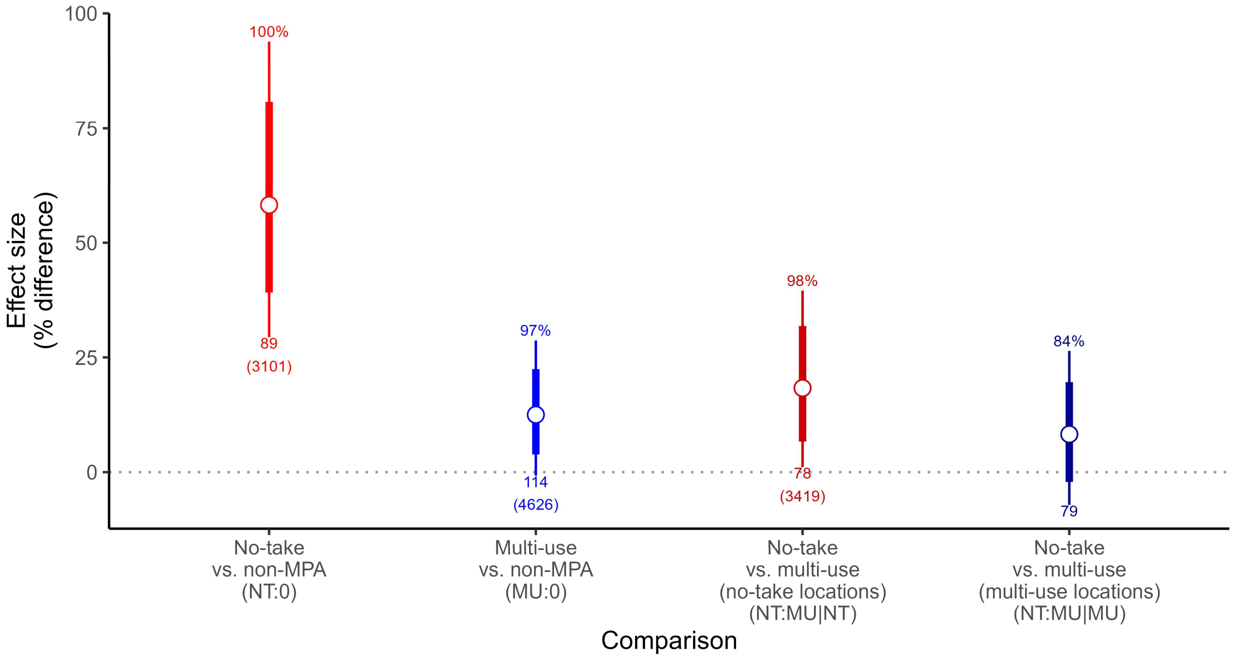


Fig. S4 | Absolute and relative no-take and multiple-use MPA impacts. Effect sizes of the expected difference in total fish biomass in percent differences, comparing biomass in no-take (left, red) and multiple-use (second left, blue) MPA sites to unprotected sites, no-take to counterfactual multiple-use MPA sites in no-take locations (second right, dark-red), and multiple-use to counterfactual no-take MPA sites in multiple-use locations (right, dark-blue) (Table S6). Greater values in the relative impact estimates (second right and right) represent larger expected biomass increases from converting to no-take restrictions. Thick and thin lines show the 80% and 95% credible intervals, respectively, around the median effect size (white dot). Probability of positive effects are shown above the estimates and number of MPAs (and number of sites, in parentheses) are shown below estimates. Percent differences were calculated using modelled response ratios (Table S9; Equation 3). NT:MU|MU model estimates (right; dark blue) generated as MU-NT but converted to NT-MU by multiplying by -1 for easier interpretation.

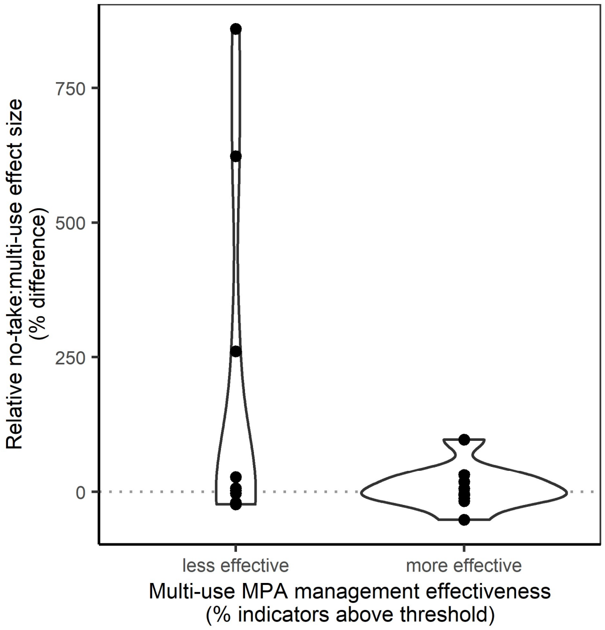


Fig. S5 | Relative no-take:multiple-use MPA impacts versus management effectiveness of multiple-use MPAs within high pressure locations (<100 km from population center). Effect sizes represent the expected difference in total fish biomass from converting existing multiple-use to no-take MPAs where greater values represent larger expected biomass increases from converting to no-take restrictions (n=19 MPAs). Management effectiveness values (x-axis) represent multiple-use MPAs where fewer (less effective) or more (more effective) than half of ten management effectiveness indicators exceed the indicator thresholds for adequate and appropriate management. Management effectiveness indicators were staff capacity, sustainable-use regulations, budget capacity, management plan implementation, monitoring activities, enforcement capacity, inclusive decision making, devolution of management authority, level of legislative support, and MPA boundary delineation. See Gill et al (6) and Table S4 for more details on indicators and indicator thresholds.

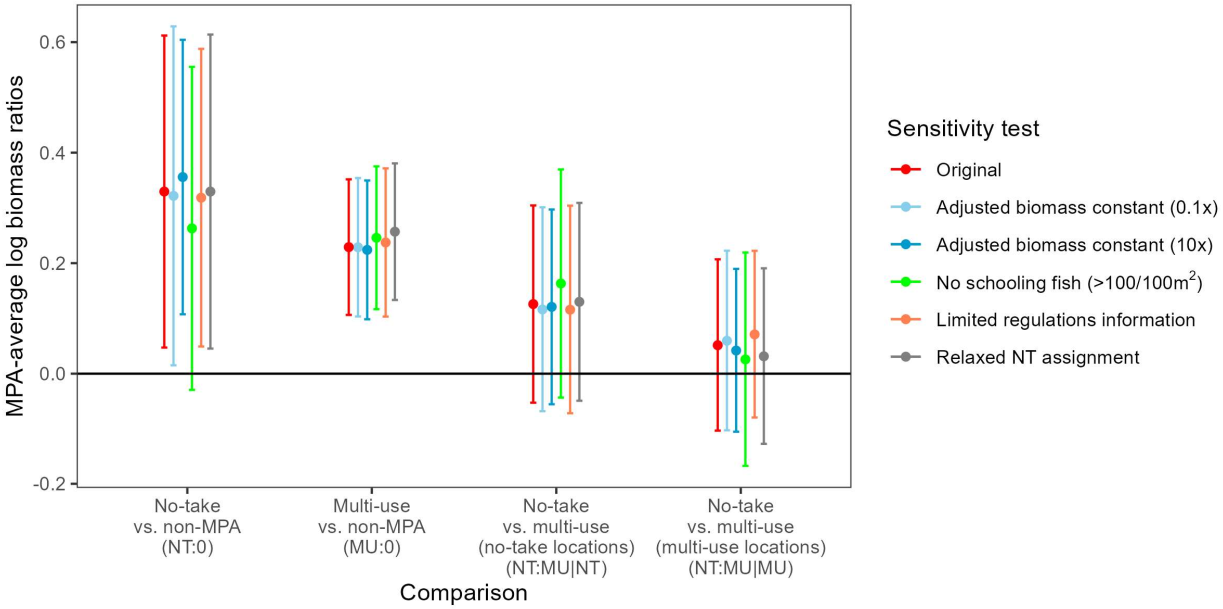


Fig. S6 | Fish biomass compilation sensitivity tests. Plot shows (unmodelled) MPA average logged biomass ratios based on original data (red); Adjusted biomass constant: changing the magnitude of the constant biomass value (i.e., minimum fish biomass observed in dataset) added to each site's total biomass to avoid calculating log ratios with zero values (0.1 and 10 times minimum fish biomass value; light and dark blue respectively); No schooling fish (green): removing all records with schools of fish (i.e., >100 individuals per 100m^2); Limited regulations information (orange): removing sites where two sources of regulations information could not be obtained; Relaxed NT assignment (grey): relaxing stringent requirement that no-take sites (as designated by the data provider) be re-categorized as multiple-use if some low-level fishing allowed. See Table S6 for comparison descriptions.

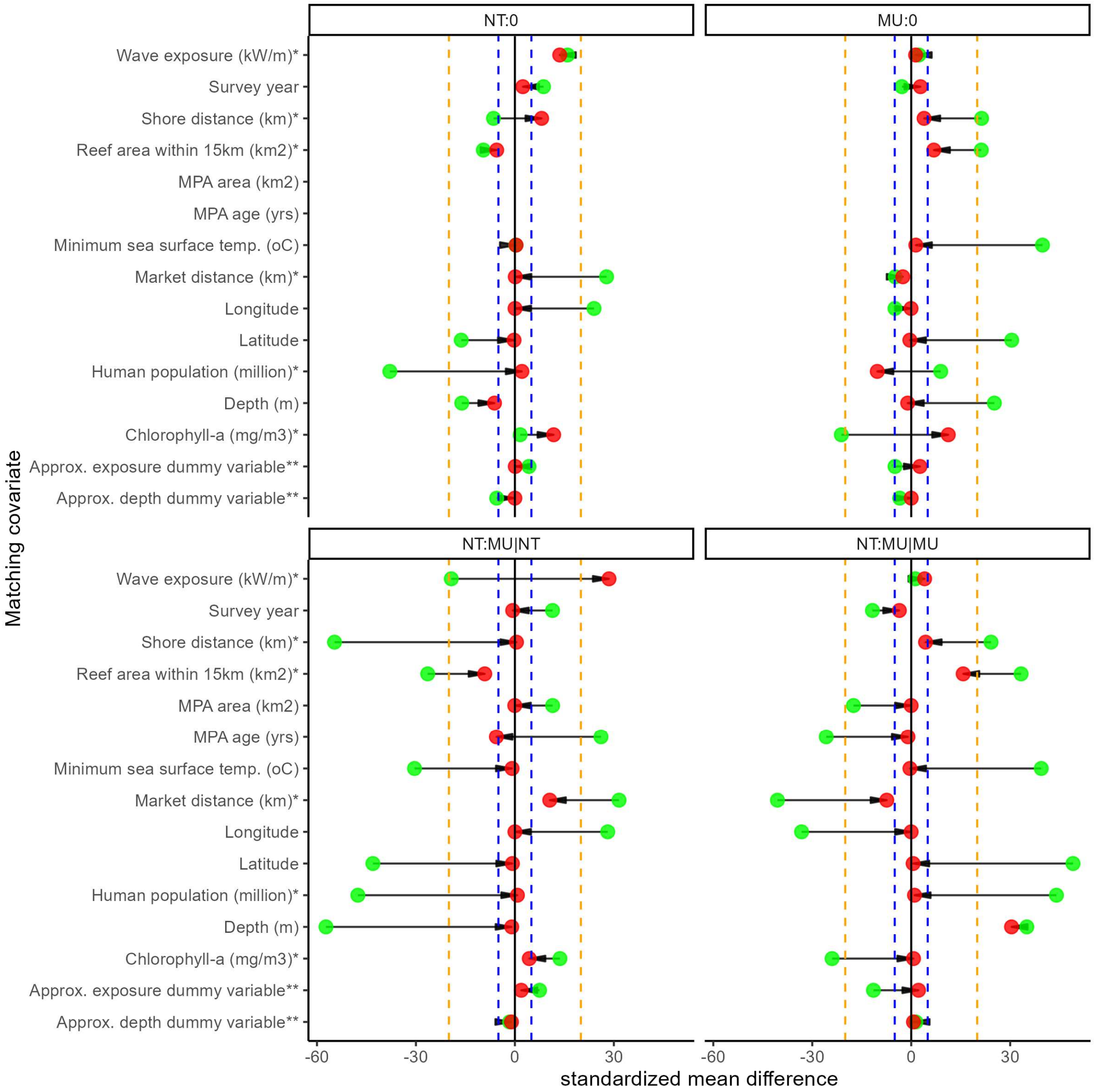


Fig. S7 | Standardized mean difference between treated and control values before (green) and after (red) matching for each matching covariate and each of the four comparisons. Blue and orange vertical dashed lines represent a standardized mean difference of 5% and 20% respectively. See Tables S3 and S7 for covariate descriptions and Table S6 for comparison descriptions.

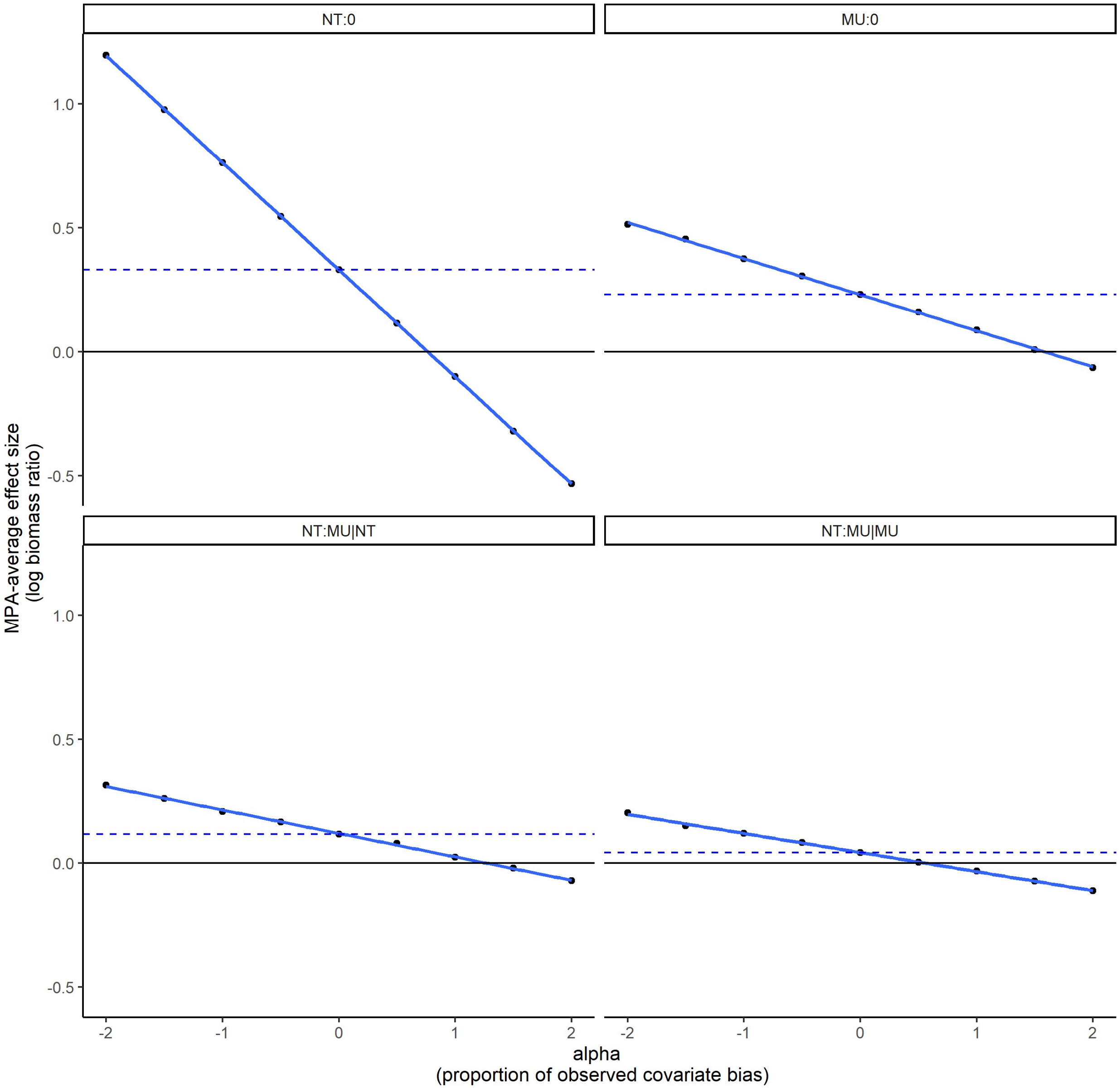


Fig. S8 | Sensitivity of MPA treatment effects to increasing and decreasing amounts of hypothetical confounding (alpha) on counterfactual control outcomes. Alpha values represent multiples of the difference in fish biomass outcomes between control and treatment sites that can be explained by the remaining post-matching differences in covariates. Here alpha values were added to the observed counterfactual biomass. Vertical axes represent the (unmodelled) MPA average treatment effects (i.e., average bias-adjusted logged biomass ratios) and the horizontal dashed blue lines show the estimates without hypothetical confounding (i.e., alpha = 0). Negative alpha values suggest that treated sites are placed in locations that inherently have greater fish biomass than control locations and visa versa. See Table S6 for comparison descriptions.

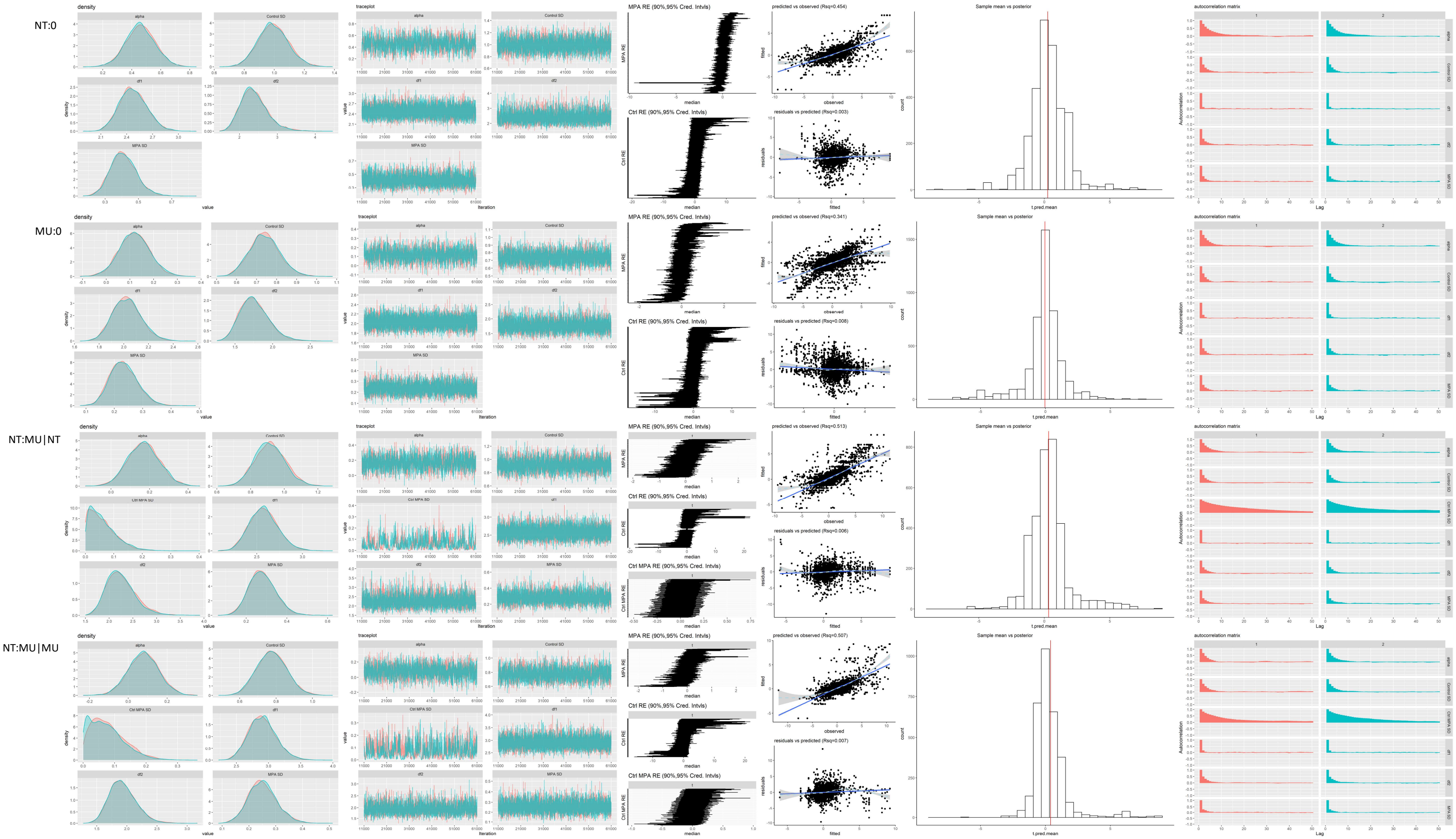


Fig. S9 | Diagnostic plots for Bayesian linear models of MPA-level impacts on fish biomass (log response ratios) for models without fixed effect covariate predictors. Plots (from left to right) include density and trace plots for each model parameter; caterpillar plots of MPA, matched control (Ctrl) site, and matched control site MPA random effects (RE); predicted vs. observed and residual vs. predicted scatter plots; sample mean (red vertical line) vs. posterior distribution plots; and auto-correlation plots for each model parameter (rows) and chain (columns). The four rows represent the results for the NT:0, MU:0, NT:MU|NT, and NT:MU|MU models respectively (Table S6).

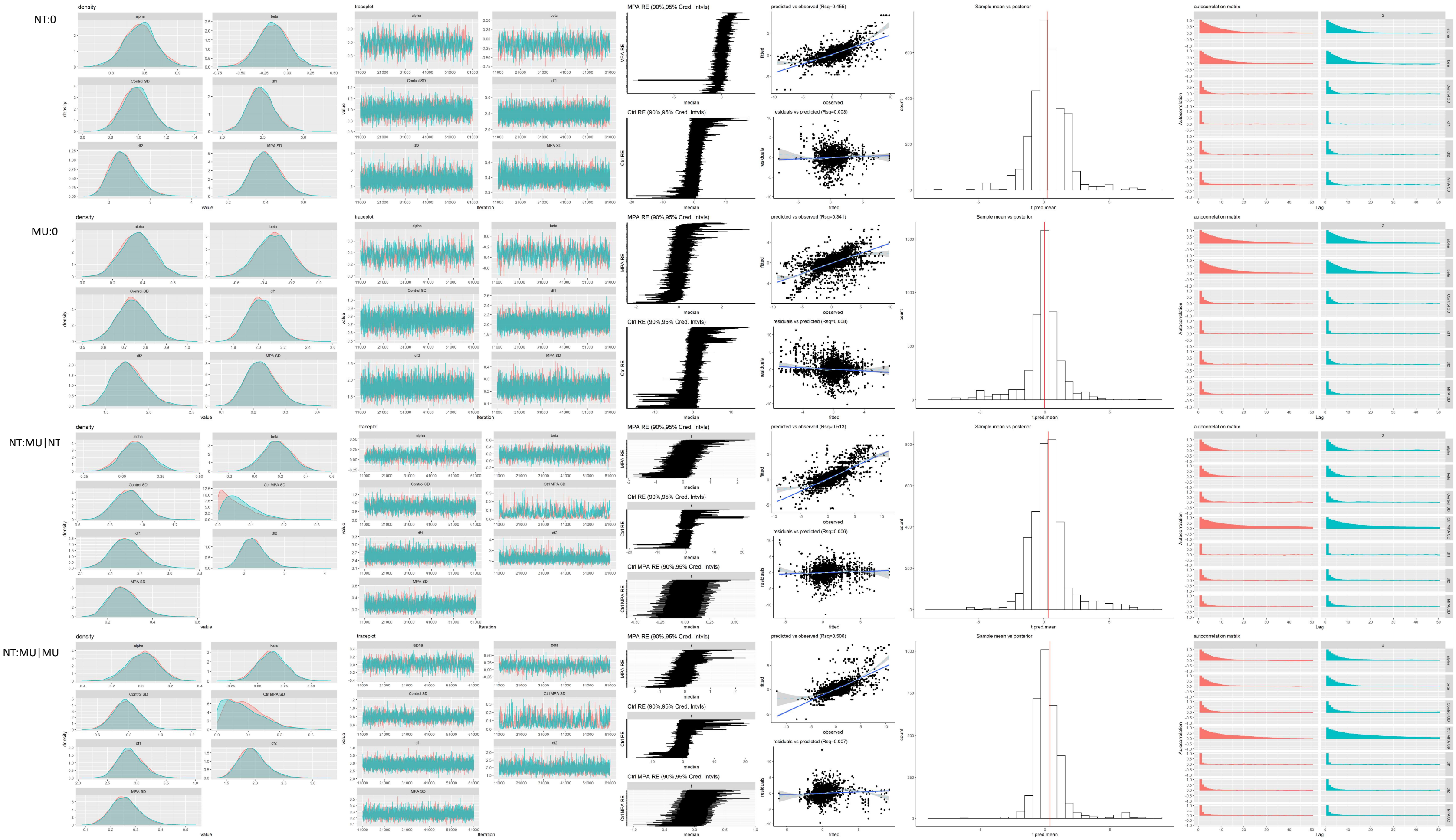


Fig. S10 | Diagnostic plots for Bayesian linear models of MPA-level impacts on fish biomass (log response ratios) by population center distance. Plots (from left to right) include density and trace plots for each model parameter; caterpillar plots of MPA, matched control (Ctrl) site, and matched control site MPA random effects (RE); predicted vs. observed and residual vs. predicted scatter plots; sample mean (red vertical line) vs. posterior distribution plots; and auto-correlation plots for each model parameter (rows) and chain (columns). The four rows represent the results for the NT:0, MU:0, NT:MU|NT, and NT:MU|MU models respectively (Table S6).

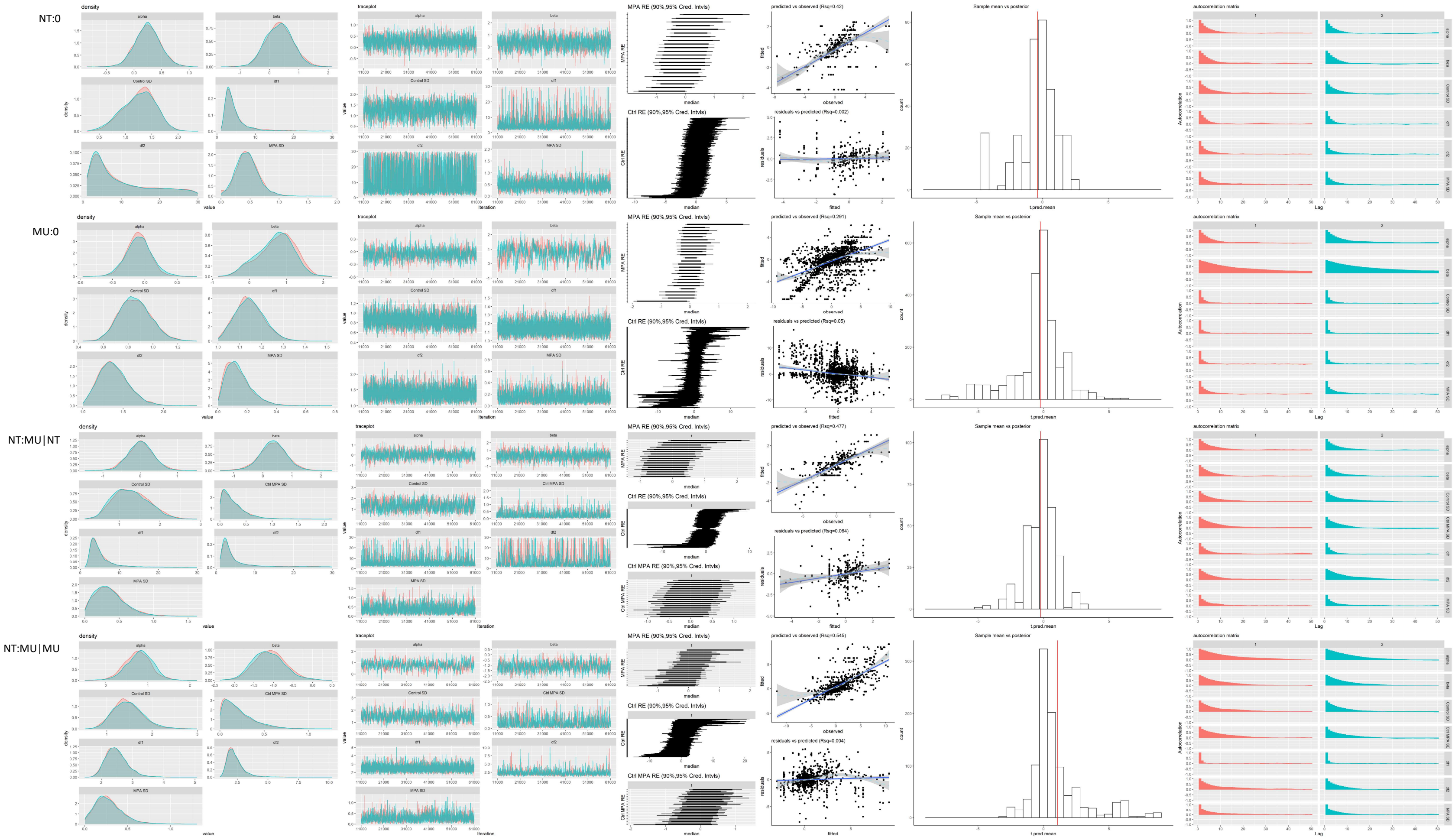


Fig. S11 | Diagnostic plots for Bayesian linear models of MPA-level impacts on fish biomass (log response ratios) by staff capacity. Plots (from left to right) include density and trace plots for each model parameter; caterpillar plots of MPA, matched control (Ctrl) site, and matched control site MPA random effects (RE); predicted vs. observed and residual vs. predicted scatter plots; sample mean (red vertical line) vs. posterior distribution plots; and auto-correlation plots for each model parameter (rows) and chain (columns). The four rows represent the results for the NT:0, MU:0, NT:MU|NT, and NT:MU|MU models respectively (Table S6).

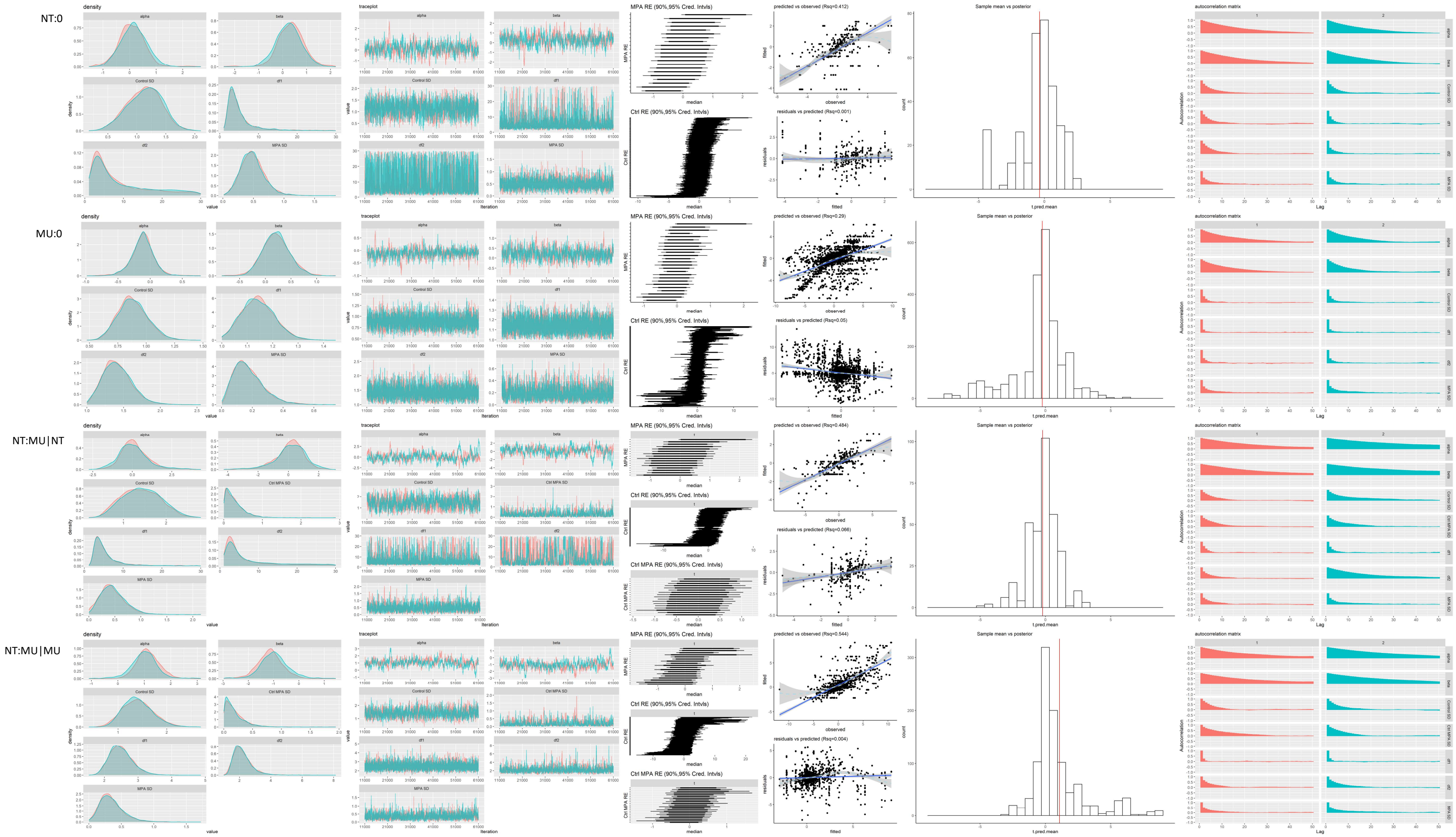


Fig. S12 | Diagnostic plots for Bayesian linear models of MPA-level impacts on fish biomass (log response ratios) by resource-use regulations. Plots (from left to right) include density and trace plots for each model parameter; caterpillar plots of MPA, matched control (Ctrl) site, and matched control site MPA random effects (RE); predicted vs. observed and residual vs. predicted scatter plots; sample mean (red vertical line) vs. posterior distribution plots; and auto-correlation plots for each model parameter (rows) and chain (columns). The four rows represent the results for the NT:0, MU:0, NT:MU|NT, and NT:MU|MU models respectively (Table S6).

Table S1 | Sources and description of MPA management assessment and fish population data. Adapted from Gill et al (6).

Dataset/ Management Assessment Tool	Geographic Coverage/Habitat	Date Range	Number of Assessments/ Survey Sites	Data Type	Data Source
Management assessment tool					
Management Effectiveness Tracking Tool (METT)	Global; mostly developing countries	2000-2014	533*	Likert-scaled management assessments	Global Database for Protected Area Management Effectiveness; Conservation International
World Bank MPA Scorecard	Global; mostly in developing countries	2011-2015	166*	Likert-scaled management assessments	Conservation International; WWF Birds Head Seascape project
NOAA Coral Reef Conservation Program (CRCP) MPA Management Assessment Checklist (NOAA CRCP MPA Checklist) & Caribbean MPA Capacity Assessment Tool ‡	US Caribbean and Pacific as well as other Caribbean MPAs	2011	51*	Likert-scaled management assessments	NOAA CRCP; Caribbean MPA Management Capacity Assessment
Fish population data					
Atlantic Gulf Rapid Reef Assessment (AGRRA)**	Wider Caribbean; coral reefs	1997-2012	1,394	Underwater Visual Census (UVC) surveys (ecologically and commercially important species)	www.agrra.org
NOAA National Coral Reef Monitoring Program (NOAA NCRMP)***	US Caribbean and Pacific (Hawaii, Guam, Tutuila); coral reefs and associated ecosystems	2000-2014	8,534	UVC surveys	NOAA NCRMP
Reef Life Surveys (RLS)	Global; rocky and coral reefs	2006-2013	5,760	UVC surveys	Reef Life Surveys

Dataset/ Management Assessment Tool	**Geographic Coverage/Habitat**	**Date Range**	**Number of Assessments/ Survey Sites**	**Data Type**	**Data Source**
Wildlife Conservation Society (WCS)	East Africa coral reefs (Madagascar and Mozambique)	2007-2015	103*	UVC surveys (fishable biomass: >10 cm excluding non-target species)	Wildlife Conservation Society
WWF Bird's Head Seascape Ecological Impact Evaluation program	Indonesia coral reefs	2011-2014	200	UVC surveys (major fish families)	World Wildlife Fund

*MPA level survey data. **AGRRA data are derived from multiple data providers, including data provided by the Healthy Reef Initiative, Living Oceans Foundation, Bahamas National Trust, Perry Institute for Marine Science, and the Kerzner Foundation. ***NOAA NCRMP comprised data from the online NCRMP dataset as well as data made available from the NOAA Pacific Islands Fisheries Science Center, Coral Reef Ecosystem Program (CREP). Survey sites refer to spatially explicit sampling events. ǂ The Caribbean MPA Capacity Assessment Tool was commissioned and conducted by the NOAA CRCP, the Gulf and Caribbean Fisheries Institute (GCFI), and the UNEP-CEP Caribbean Marine Protected Area Management Network and Forum.

Table S2 | Data sources for MPA and zone spatial data.

Spatial dataset	Data	Release Date
World Database of Protected Areas	MPA and zone boundary polygons (or buffered points where no polygon data were available; n=140)	Nov 2017
NOAA MPA Inventory	MPA boundary polygons	2014
Collaborative Australian Protected Areas Database (CAPAD) Australia	MPA and zone boundary polygons	2016
The Nature Conservancy (TNC) Caribbean MPAs	MPA and zone boundary polygons	2014
Healthy Reef Initiative	MPA and zone boundary polygons	2014
MPA Atlas	MPA and zone boundary polygons	2015
Digitized polygons	MPA and zone boundary polygons	Dependent on source

Table S3 | Description and sources of variables used in study. Adapted from Gill et al (6)**.** Variable types include the response variable (fish biomass), covariates used in the matching procedures, and management and contextual predictors. Data types were continuous (Cont), ordinal (Ord) or binary (Bin). Spatial scale refers to the scale applied in the matching or analysis. See Table S1 for more details on the management and ecological data sources, Table S4 for management indicator scoring levels, and Table S5 for summary statistics.

Variable	**Data Type**	**Spatial Scale**	**Variable Description**	**Data Source**
Response variable				
Fish biomass	Cont	Site	Total fish biomass (g/100m^2) of all recorded species >10cm (see Section 2 in methods for exceptions), Calculated from transect/site level data.	Underwater visual census datasets (Table S1)
Matching covariates				
MPA age	Cont	Site	MPA age at the time of fish survey (years)	Calculated from MPA establishment data from official government/NGO sources and/or WDPA (35) (October 2015 release)
MPA size	Cont	MPA	MPA size (km^2)	Based on data from official government/NGO sources and/or WDPA (October 2015 release); Some values calculated from spatial data
Latitude/ longitude	Cont	Site	Location of fish survey site	Fish survey data
Country	Cat	Site	Location of fish survey site	Fish survey data; EEZ Maritime Boundaries (78)
Habitat	Cat	Site	Marine habitat at fish survey site (e.g., coral reefs, rocky reefs, seagrass, mangroves) at the highest resolution available (e.g., fringing vs patch coral reefs)	Fish survey data; benthic NOAA habitat maps (79); WCMC Global Distribution of Coral Reefs (80)
Minimum sea surface temperature	Cont	Site	Minimum sea surface temperature (2002-2009; ^{o}C)	Bio-ORACLE (81)
Chlorophyll-a	Cont	Site	Proxy for primary productivity at study site (Chlorophyll-a (2002-2009; mg/m^3))	Bio-ORACLE (81)
Depth	Cont	Site	Depth at survey site (m)	Fish survey data; NOAA bathymetric raster maps (82–84)
Exposure	Cont	Site	Wave energy at fish survey site (kW/m)	Calculated using wind/wave data from WAVEWATCH III (WW3) (85) and fetch using the 'waver' R package (86)
Marine ecoregions	Cat	Site	Marine biogeographic region	WWF Marine Ecosystems of the World GIS layer (87)
Distance to shoreline	Cont	Site	Distance to nearest coastline (km)	Calculated using the high resolution shoreline layer from the Global Self-consistent, Hierarchical, High-resolution Geography (GSHHG) dataset (88)

Variable	Data Type	Spatial Scale	Variable Description	Data Source
Human population density	Cont	Site	Coastal population within 100 km radius of fish survey site (# individuals)	Calculated using the Socioeconomic Data and Application Centre (SEDAC) Gridded Population Of The World dataset (89)
Distance to population center	Cont; Bin	Site	Distance to capital or population center, used as a proxy for distance to markets (km) and fishing pressure. Converted to binary to separate sites near (<100 km) vs far (>100 km) from population centers in Bayesian models	Calculated using the World Cities base map layer provided by ESRI (Version 10.1)(90)
Management variables				
Staff capacity	Ord	MPA	Adequacy of (on-site) staff capacity/numbers to carry out critical management activities (including designated community members)	MPA Management assessment datasets (Table S1)
Sustainable-use regulations	Ord	MPA	Appropriate regulations to control use/unsustainable activities are defined and in place	MPA Management assessment datasets (Table S1)

Table S4 | Indicators, thresholds, scores, and descriptions from the three management assessments used in this study. Adapted from Gill et al (6). Blue dotted line indicates the threshold levels for each indicator. See Gill et al (6) for more details. Original scores were adjusted to permit alignment between indicator values across assessments.

Management Indicator	**Threshold (dashed blue line)**	**Adjusted Score**	**Indicator Score Descriptions**		
			Management Effectiveness Tracking Tool (42)	**World Bank MPA Scorecard (and variants)** (43)	**NOAA CRCP MPA Checklist** (44) **& Caribbean MPA Capacity Assessment Tool** (45)
Staff capacity and/or presence	Adequate staff capacity/presence	1	There are no staff	There are no staff	No management personnel assigned to site and/or little or no formalized community oversight
		2	Staff numbers are inadequate or below optimum	Staff numbers are inadequate for critical management activities Staff numbers are below optimum level for critical management activities	Some management personnel assigned to site or some formalized community oversight
		3	Staff numbers are adequate	Staff numbers are adequate for the management needs of the site	Full-time site manager and programmatic personnel assigned to site or local community-based management leader in place that has been formally designated and accepted and is able to dedicate sufficient time to the management of the site
Appropriateness of regulations controlling use	Appropriate MPA regulations in place controlling use	1	There are no regulations or regulations with major weaknesses		Site has been legally established or is under equivalent customary tenure or other form of community-based protection status, but there are few or no official or community-based rules and regulations in place supporting the MPA and its management plan
		2	Regulations with some weaknesses or gaps		Laws or customary instruments for the establishment of the MPA are in place, and official or community-based rules or regulations governing some specific activities within the MPA are also in place
		3	Regulations provide an excellent basis for management		Clearly defined laws or customary instruments and official or community-based rules and regulations governing all specific activities included in the objectives of the site management plan are in place

Table S5. | Summary statistics for outcome, MPA attribute, contextual, and management variables. SD: standard deviation; SE: standard error.

Variable	n (MPAs)	Mean	SD	SE	Median	Min	Max
Outcome variable							
Total fish biomass (log(g/100m^2))	217	8.04	1.25	0.08	8.30	1.38	10.26
MPA attribute and contextual variables							
Chlorophyll-a (mg/m^3)	217	0.83	1.47	0.10	0.36	0.04	10.71
Depth (m)	217	8.31	3.83	0.26	7.87	0.55	27.99
Human population (million)	217	0.65	1.05	0.07	0.15	0.00	4.86
Latitude	217	21.81	10.86	0.74	19.66	0.18	55.90
Longitude	217	-10.02	116.23	7.89	-71.11	-178.17	177.13
Market distance (km)	217	123.72	171.70	11.66	64.82	1.05	1151.22
Reef area within 15km (km^2)	217	35.35	41.74	2.83	28.00	0.00	266.01
Shore distance (km)	217	3.11	18.89	1.28	0.39	0.00	200.81
Survey year	217	2007.75	4.44	0.30	2009.77	1998.00	2014.00
Minimum sea surface temp. (^{o}C)	217	22.97	5.78	0.39	25.85	-1.18	29.58
Wave exposure (kW/m)*	217	185.22	2343.35	159.08	8.05	0.00	34532.05
MPA age (yrs)	217	18.57	14.34	0.98	14.56	3.00	95.00
MPA area (km^2)	217	6.20	49.68	3.37	0.04	0.00	687.99
Management variables							
Adequate staff capacity	217	0.19	0.40	0.05	0.00	0.00	1.00
Sustainable use regulations	217	0.76	0.43	0.06	1.00	0.00	1.00

* Extreme outliers present, likely due to an error from wind time series from the Mediterranean (not used in the analysis); median wave exposure value is more representative. Statistics for the MPA attribute and contextual variables represent the mean values from all survey sites inside the MPA.

Table S6 | Comparisons examined in the study, with descriptions of the estimands and treatment and counterfactual groups used to estimate the absolute and relative impacts of no-take and multiple-use MPAs (Adapted from Rasolofoson et al (49)).

Comparison		Abbreviation	Estimand	Treatment	Counterfactual
Absolute impacts	No-take MPA to no MPA in no-take MPA locations	NT:0	expected difference between biomass in protected no-take MPA/zones compared to the same sites if they were not protected $E(Y_{NT} - Y_0 \mid T_{NT} = 1)$	No-take MPA	No MPA
	Multiple-use MPA to no MPA in multiple-use MPA locations	MU:0	expected difference between biomass in multiple-use MPA/zones compared to the same sites if they were not protected $E(Y_{NT} - Y_0 \mid T_{MU} = 1)$	Multiple-use MPA	No MPA
Relative impacts	No-take MPA to multiple-use MPA in no-take MPA locations	NT:MU|NT	expected difference between biomass in no-take MPA/zones compared to the same sites if they were multiple-use MPA sites instead $E(Y_{NT} - Y_{MU} \mid T_{NT} = 1)$	No-take MPA	Multiple-use MPA
	No-take MPA to multiple-use MPA in multiple-use MPA locations*	NT:MU|MU	expected difference between biomass in multiple-use MPA/zones compared to the same sites if they were no-take MPA sites instead $E(Y_{NT} - Y_{MU} \mid T_{MU} = 1)$	Multiple-use MPA	No-take MPA

*Model estimates generated as MU-NT but converted to NT-MU by multiplying by -1 for easier interpretation

Table S7 | Covariates used in the matching process to identify appropriate control fish survey sites to pair with treated sites based on their influence on site selection and fish biomass outcomes. Calipers were used for covariates known to strongly affect fish biomass outcomes. Adapted from Gill et al (6). See Table S8 for post-matching statistics.

Matching Covariates	Rationale	Restrictions in Treatment-Control Matched Pairs (calipers)
Exact matching variables		
Sampling protocol	Control for differences in sampling methodologies (24).	Same methodology only
Habitat type	Control for habitat selection bias in MPA placement (91) and natural variation in fish communities by habitat (8, 19, 24, 33, 92–94).	Similar habitat type only
Country	Control for variation in national policies and/or resource use patterns between countries (95).	Same country only
Marine ecoregions	Control for large-scale biogeographic variation (87).	Same ecoregion only
Nearest neighbor matching variables		
Minimum sea surface temperature (2002-2009; °C)	Temperature affects fish community structure. Low temperatures can act as spatial boundaries for warm water fish species (96), which make up the majority of the sample.	Minimize mean difference
Chlorophyll-a (2002-2009; mg/m^3)	Control for variations in available primary productivity which could affect community composition (73, 97).	Minimize mean difference
Depth (m)	Control for natural variation in community composition by depth (98).	Minimize mean difference
Wave exposure (kW/m)	Wave energy explains some of the variation in marine community composition (19, 99, 100). Adverse sea conditions can also be a deterrent for small fishing vessels resulting in lower fishing pressure (101, 102).	Minimize mean difference
Distance to shoreline (km)	All else equal, fishing intensity is usually negatively correlated with shore distance (103, 104). Also accounts for other land-based human stressors (e.g., pollution, destructive nearshore activities) from neighboring coastal populations.	Minimize mean difference (maximum difference of 1 standard deviation)
Coastal population (individuals within 100km^2)	Control for human impacts (e.g. pollution, destructive nearshore activities) from neighboring coastal populations (55, 105).	Minimize mean difference (maximum difference of 1 standard deviation)
Distance to provincial capital (market) (km)	Distance to capital used as a proxy for distance to major markets, which is commonly negatively correlated with fishing intensity (55, 56, 106).	Minimize mean difference (maximum difference of 1 standard deviation)
Reef area within 15 km (km^2)	Neighboring reefs that can act as sources for larvae and other forms of connectivity amongst meta-populations (107).	Minimize mean difference
Sample date (years)	Control for unobserved temporal variation caused by factors such as exogenous shocks (e.g., storm events, algal blooms) and other changes between survey periods (32).	Minimize mean difference (maximum difference of 4 years)
Latitude/ longitude	Control for unobserved spatial variation caused by time-invariant factors (e.g., differences in local social-ecological conditions) or localized shocks (e.g., storm history) that may vary between sites. Also reduces latitudinal effects (i.e.,	Minimize mean difference (maximum difference of 2.5° latitude)

Matching Covariates	Rationale	Restrictions in Treatment-Control Matched Pairs (calipers)
	varying environmental conditions with increasing distance from equator).	
MPA age*	Control for the effects of MPA age on fish biomass (e.g., older MPAs allow for greater population recovery) (4, 23, 108).	Minimize mean difference (maximum difference of 1 standard deviation)
MPA area*	Control for the effects of MPA size on biomass (e.g., larger MPAs protect a greater portion of fish species ranges) (4, 108).	Minimize mean difference

* Covariates used for estimating relative impacts (i.e., directly matching NT and MU MPA sites: NT:MU|NT; NT:MU|MU; Table S6).

Table S8 | Covariate balance statistics for the unmatched (before) and matched (after) fish survey site data. Table shows treatment and control means, standardized mean differences, raw mean (Mean eQQ diff.) and maximum (max eQQ diff.) differences from empirical quantile-quantile plots between the groups for each comparison (Table S6) and each covariate (Tables S3 and S7).

Comparison	**Treatment Mean**		**Control Mean**		**Standardized Mean Difference**		**Raw eQQ diff (Mean)**		**Raw eQQ diff (Max)**	
	Before	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**
Wave exposure (kW/m)										
NT:0	26.16	20.17	19.49	16.26	15.97	13.58	8.83	5.23	341	166
MU:0	34.22	31.23	19.49	21.68	2.33	1.33	14.71	24.9	33874	33874
NT:MU|NT	26.16	24.36	34.22	14.18	-19.29	28.53	17.96	10.66	34215	41
NT:MU|MU	34.22	24.03	26.16	21.52	1.27	4.06	17.96	9.01	34215	384
Survey year*										
NT:0	2008	2007	2007	2007	8.68	2.41	0.39	0.47	3	3
MU:0	2007	2007	2007	2007	-2.84	2.81	0.49	0.15	2	1
NT:MU|NT	2008	2007	2007	2007	11.41	-0.67	0.45	0.15	2	1
NT:MU|MU	2007	2008	2008	2008	-11.76	-3.6	0.45	0.13	2	1
Shore distance (km)*										
NT:0	2.96	0.65	4.35	0.58	-6.48	8.1	1.45	0.18	106.60	3.54
MU:0	14.64	1	4.35	0.94	21.33	3.9	11.66	0.15	184.24	8.91
NT:MU|NT	2.97	1.04	14.64	1.01	-54.68	0.49	12.26	0.42	188.06	21.51
NT:MU|MU	14.64	1.09	2.97	0.92	24.19	4.34	12.26	0.27	188.06	20.98
Reef area within 15km (km^2)										
NT:0	29.79	27.88	33.38	29.43	-9.56	-5.52	6.48	3.2	62.25	55.50
MU:0	39.72	42.96	33.38	41.02	21.25	6.85	10.92	3.05	93.75	43.50
NT:MU|NT	29.78	31.99	39.72	35.24	-26.43	-9.14	14.42	9.17	116.00	105.50
NT:MU|MU	39.72	38.41	29.78	33.26	33.31	15.75	14.42	7.63	116.00	76.25
MPA area (km^2)										
NT:MU|NT	22.72	24.19	10.14	24.24	11.45	-0.04	12.71	0.06	626.80	1.25
NT:MU|MU	10.14	15.75	22.72	15.78	-17.49	-0.03	12.71	0.06	626.80	26.11
MPA age (yrs)*										
NT:MU|NT	26.62	27.43	22.05	28.43	26.1	-5.61	6.35	2.35	42.00	9.00
NT:MU|MU	22.05	21.55	26.62	21.77	-25.76	-0.98	6.35	1.92	42.00	10.00
Minimum sea surface temp. (^{o}C)										
NT:0	22.39	22.53	22.38	22.51	0.27	0.35	0.81	0.21	10.52	2.14
MU:0	24	24.48	22.38	24.43	39.79	1.4	1.74	0.15	8.01	1.07
NT:MU|NT	22.4	22.94	24	22.98	-30.43	-0.85	1.7	0.22	7.28	2.29
NT:MU|MU	24	23.17	22.4	23.18	39.43	-0.39	1.7	0.09	7.28	0.85
Market distance (km)*										
NT:0	162.18	97.7	106.74	97.6	27.79	0.11	64.12	14.32	967.21	132.64
MU:0	99.21	63.4	106.74	65.68	-4.85	-2.55	18.78	8.18	947.05	130.92
NT:MU|NT	162.2	149.07	99.21	129.62	31.58	10.57	63.17	22.04	405.44	88.55
NT:MU|MU	99.21	114.06	162.2	127.97	-40.56	-7.41	63.17	15.79	405.44	101.92
Longitude										
NT:0	7.46	-1.9	-20.52	-1.95	24	0.04	28.48	0.18	190.77	3.28
MU:0	-25.38	-28.12	-20.52	-28.07	-4.92	-0.05	31.26	0.19	198.60	1.68
NT:MU|NT	7.43	1.84	-25.38	1.84	28.14	0	35.28	0.15	211.86	1.37
NT:MU|MU	-25.38	4.86	7.43	4.88	-33.24	-0.02	35.28	0.08	211.86	1.57
Latitude*										
NT:0	-3.14	-0.6	0.84	-0.53	-16.27	-0.26	4.65	0.26	30.97	2.28
MU:0	7.4	7.25	0.84	7.34	30.47	-0.41	7.41	0.19	36.22	1.57

Comparison	**Treatment Mean**		**Control Mean**		**Standardized Mean Difference**		**Raw eQQ diff (Mean)**		**Raw eQQ diff (Max)**	
	Before	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**
NT:MU|NT	-3.15	-0.72	7.4	-0.54	-43.04	-0.75	10.54	0.34	47.86	1.73
NT:MU|MU	7.4	-0.36	-3.15	-0.49	49.05	0.57	10.54	0.22	47.86	1.51
Human population (million)*										
NT:0	0.35	0.38	0.65	0.36	-37.91	2.12	0.3	0.03	2.35	0.51
MU:0	0.72	0.76	0.65	0.84	8.94	-10.3	0.32	0.12	2.02	0.89
NT:MU|NT	0.35	0.22	0.72	0.21	-47.6	0.76	0.48	0.02	1.49	0.30
NT:MU|MU	0.72	0.29	0.35	0.29	44.02	0.99	0.48	0.01	1.49	0.30
Depth (m)										
NT:0	8.22	8.15	9.02	8.46	-16.1	-6.18	0.95	0.4	45.88	10.30
MU:0	11.07	10.76	9.02	10.84	25.21	-1.08	2.14	0.68	30.54	9.46
NT:MU|NT	8.22	8.21	11.07	8.26	-57.27	-0.96	3.03	0.27	15.35	11.20
NT:MU|MU	11.07	10.24	8.22	8.21	35.04	30.4	3.03	2.06	15.35	15.35
Chlorophyll-a (mg/m^3)										
NT:0	0.81	0.85	0.78	0.67	1.59	11.75	0.16	0.22	3.72	4.60
MU:0	0.61	0.6	0.78	0.5	-21.18	11.24	0.27	0.14	6.87	4.99
NT:MU|NT	0.81	0.78	0.61	0.71	13.66	4.36	0.25	0.08	5.11	5.03
NT:MU|MU	0.61	0.56	0.81	0.56	-24	0.7	0.25	0.05	5.11	2.84
Exposure dummy variable**										
NT:0	0.02	0.02	0.01	0.02	4.27	0.11	0.01	0	1	1
MU:0	0.01	0.01	0.01	0.01	-4.91	2.66	0	0	1	1
NT:MU|NT	0.02	0.02	0.01	0.02	7.48	1.91	0.01	0	1	1
NT:MU|MU	0.01	0.01	0.02	0.01	-11.45	2.21	0.01	0	1	1
Depth dummy variable**										
NT:0	0	0	0.01	0	-5.53	0	0	0	1	0
MU:0	0	0	0.01	0	-3.48	0	0	0	1	0
NT:MU|NT	0	0	0	0	-1.59	-1.03	0	0	1	1
NT:MU|MU	0	0	0	0	1.41	0.68	0	0	1	1

*Calipers used in matching procedure (see Table S7). **Dummy variable used to indicate where mean values were used to fill missing data (<1% of data). Lower post-matching differences indicate good matching performance for that covariate (25). See Table S3 and S7 for covariate descriptions and caliper restrictions.

Table S9 | Summary of posterior statistics for Bayesian models estimating MPA impacts. Intercept (alpha) represents the MPA treatment effect of interest; beta is the binary fixed effect term for the population center distance (alpha:far >100km, alpha+beta:near <100km), capacity (alpha:inadequate; alpha+beta:adequate), and regulations (alpha:weak; alpha+beta:strong) models. Mean, SD (standard deviation) and distribution percentiles are also shown. Percent biomass differences were calculated using the log biomass ratios and Equation 3. NT:MU|MU model estimates generated as MU-NT but converted to NT-MU by multiplying by -1 for easier interpretation. P(positive): probability of a positive effect. See Table S6 for comparison descriptions.

Predictor	**Variable**	**Comparison**	**Log Biomass Ratio**							**% Biomass Difference**					
			Mean	**SD**	**2.5%**	**10%**	**50%**	**90%**	**97.5%**	**2.5%**	**10%**	**50%**	**90%**	**97.5%**	**P(positive)**
No predictor	alpha	MU:0	0.12	0.07	-0.01	0.04	0.12	0.20	0.25	-0.76	3.85	12.58	22.46	28.67	0.97
		NT:0	0.46	0.10	0.26	0.33	0.46	0.59	0.66	29.40	39.17	58.21	80.80	93.91	1.00
		NT:MU|MU	0.08	0.08	-0.07	-0.02	0.08	0.18	0.23	-7.12	-2.13	8.25	19.65	26.40	0.84
		NT:MU|NT	0.17	0.08	0.01	0.06	0.17	0.28	0.33	1.11	6.66	18.36	31.79	39.58	0.98
Near population center	alpha	MU:0	0.37	0.11	0.16	0.22	0.36	0.51	0.59	17.61	25.06	43.95	65.86	80.27	1.00
	alpha+beta		0.04	0.07	-0.10	-0.05	0.04	0.13	0.18	-9.65	-5.07	3.97	14.03	19.89	0.71
	alpha	NT:0	0.56	0.15	0.27	0.37	0.56	0.75	0.86	30.81	45.38	75.38	111.28	135.31	1.00
	alpha+beta		0.41	0.12	0.18	0.25	0.41	0.56	0.64	19.54	29.03	50.22	74.66	89.04	1.00
	alpha	NT:MU|MU	0.01	0.11	-0.19	-0.12	0.02	0.15	0.22	-17.67	-11.51	1.79	16.00	24.35	0.56
	alpha+beta		0.15	0.10	-0.05	0.02	0.15	0.28	0.36	-4.93	1.59	15.66	32.51	42.91	0.92
	alpha	NT:MU|NT	0.09	0.09	-0.09	-0.03	0.09	0.21	0.27	-9.05	-3.11	9.16	23.26	31.61	0.83
	alpha+beta		0.26	0.10	0.06	0.13	0.26	0.39	0.46	6.22	13.94	30.19	48.03	58.82	0.99
Adequate staff capacity	alpha	MU:0	-0.06	0.12	-0.29	-0.20	-0.06	0.09	0.18	-24.93	-18.46	-5.71	9.17	19.56	0.29
	alpha+beta		0.66	0.47	-0.30	-0.02	0.71	1.23	1.48	-26.27	-1.84	103.79	243.82	340.84	0.89
	alpha	NT:0	0.25	0.23	-0.21	-0.05	0.25	0.55	0.72	-18.89	-4.82	27.94	72.77	104.86	0.86
	alpha+beta		0.58	0.42	-0.20	0.05	0.57	1.13	1.45	-18.33	5.14	77.54	209.17	325.54	0.92
	alpha	NT:MU|MU	0.78	0.32	0.15	0.38	0.79	1.17	1.40	16.20	46.71	120.81	222.22	303.66	0.99
	alpha+beta		-0.32	0.31	-0.95	-0.72	-0.32	0.07	0.30	-61.24	-51.14	-27.44	6.93	34.41	0.14
	alpha	NT:MU|NT	0.01	0.35	-0.68	-0.44	0.01	0.45	0.71	-49.24	-35.45	1.08	56.74	103.96	0.51
	alpha+beta		0.36	0.37	-0.35	-0.09	0.35	0.83	1.15	-29.39	-8.21	41.93	129.72	214.52	0.85
Sustainable use regulations	alpha	MU:0	-0.12	0.17	-0.49	-0.33	-0.11	0.09	0.22	-38.59	-28.17	-10.73	8.91	25.01	0.22
	alpha+beta		0.09	0.19	-0.25	-0.14	0.09	0.34	0.48	-22.23	-13.31	8.94	40.71	61.79	0.68
	alpha	NT:0	0.12	0.50	-0.84	-0.50	0.11	0.73	1.09	-56.72	-39.17	11.79	107.54	196.49	0.59
	alpha+beta		0.41	0.24	-0.03	0.12	0.41	0.72	0.90	-2.74	12.97	50.62	104.96	146.38	0.97
	alpha	NT:MU|MU	1.07	0.47	0.14	0.51	1.06	1.64	2.01	14.81	66.03	189.73	416.66	646.12	0.98
	alpha+beta		0.01	0.28	-0.52	-0.33	0.00	0.36	0.61	-40.59	-27.91	0.00	43.45	83.39	0.50
	alpha	NT:MU|NT	0.09	0.90	-1.53	-0.91	0.01	1.21	2.19	-78.43	-59.83	1.12	235.29	795.95	0.51
	alpha+beta		0.22	0.32	-0.43	-0.17	0.23	0.62	0.86	-35.17	-15.41	25.39	85.96	135.99	0.77

Table S10 | Diagnostic statistics for parameters in the Bayesian models without fixed effect predictors. Intercept (alpha) represents the MPA treatment effect of interest. SD: standard deviation; df: degrees of freedom. See Table S6 for comparison descriptions.

Predictor	Variable	Comparison	Log Biomass Ratio		# Sites	# MPAs	Rhat	Effective Sample Size	Total Iterations	Raftery & Lewis Min # Iterations	Geweke Z Scores	
			Mean	SD							Chain 1	Chain 2
No predictor	alpha	MU:0	0.12	0.07	4626	114	1	1656	4000	3746	-0.79	0.29
	Control SD		0.74	0.08	4626	114	1	2923	4000	3746	0.36	-0.36
	df1		2.03	0.12	4626	114	1	6544	4000	3746	-1.04	-0.47
	df2		1.76	0.19	4626	114	1	3623	4000	3746	0.61	0.59
	MPA SD		0.23	0.05	4626	114	1	3520	4000	3746	0.42	-1.31
	alpha	NT:0	0.46	0.1	3101	89	1	1294	4000	3746	-0.33	1.35
	Control SD		0.99	0.1	3101	89	1.01	2869	4000	3746	-2.18	0.94
	df1		2.48	0.16	3101	89	1	7105	4000	3746	1.03	-1.81
	df2		2.39	0.36	3101	89	1	3633	4000	3746	-2.71	1.04
	MPA SD		0.41	0.08	3101	89	1	3726	4000	3746	0.14	1.42
	alpha	NT:MU|MU	0.08	0.08	3261	79	1	1856	4000	3746	1.02	0.27
	Control SD		0.78	0.08	3261	79	1	2120	4000	3746	-0.74	-0.93
	Ctrl MPA SD		0.08	0.06	3261	79	1	466	4000	3746	0.58	-2.11
	df1		2.91	0.22	3261	79	1	7255	4000	3746	-0.07	-0.69
	df2		1.93	0.23	3261	79	1	2791	4000	3746	-0.7	-0.65
	MPA SD		0.25	0.05	3261	79	1	4363	4000	3746	0.07	1.19
	alpha	NT:MU|NT	0.17	0.08	3419	78	1	1685	4000	3746	-0.72	-1.06
	Control SD		0.92	0.09	3419	78	1	2466	4000	3746	-0.02	0.62
	Ctrl MPA SD		0.06	0.05	3419	78	1	332	4000	3746	-0.39	-2.25
	df1		2.58	0.16	3419	78	1	8616	4000	3746	-0.13	-2.22
	df2		2.25	0.31	3419	78	1	3021	4000	3746	-0.54	1.14
	MPA SD		0.28	0.07	3419	78	1	3014	4000	3746	-0.93	0.07

Table S11 | Diagnostic statistics for parameters in the Bayesian models with a binary predictor (beta) to identify sites in MPAs near (alpha.beta) or far from (alpha) population centers. Intercept (alpha) represents the MPA treatment effect of interest. SD: standard deviation; df: degrees of freedom. alpha.beta parameters are the sum of the intercept (alpha) and predictor (beta) coefficients. See Table S6 for comparison descriptions.

			Log Biomass Ratio							Raftery & Lewis	Geweke Z Scores	
Predictor	**Variable**	**Comparison**	**Mean**	**SD**	**# Sites**	**# MPAs**	**Rhat**	**Effective Sample Size**	**Total Iterations**	**Min # Iterations**	**Chain 1**	**Chain 2**
Near population center	alpha	MU:0	0.37	0.11	742	41	1	580	4000	3746	-0.04	-1.44
	alpha.beta		0.04	0.07	3884	76						
	beta		-0.33	0.12	3884	76	1	590	4000	3746	0.16	1.35
	Control SD		0.74	0.08	4626	117	1	2787	4000	3746	-0.85	-0.19
	df1		2.03	0.12	4626	117	1	6170	4000	3746	2.67	-2.69
	df2		1.76	0.19	4626	117	1	3512	4000	3746	-0.49	-0.03
	MPA SD		0.22	0.05	4626	117	1	3522	4000	3746	-0.87	-0.15
	alpha	NT:0	0.56	0.15	846	35	1.01	634	4000	3746	-1.1	0.09
	alpha.beta		0.41	0.12	2255	56						
	beta		-0.15	0.17	2255	56	1	694	4000	3746	-0.39	-0.12
	Control SD		0.99	0.1	3101	91	1	2987	4000	3746	-1.91	-2.5
	df1		2.48	0.16	3101	91	1	7020	4000	3746	-0.62	-0.81
	df2		2.37	0.35	3101	91	1	3687	4000	3746	-1.76	-0.73
	MPA SD		0.4	0.08	3101	91	1	3600	4000	3746	-1.12	3.3
	alpha	NT:MU|MU	0.01	0.11	971	33	1.01	1008	4000	3746	-0.54	-0.27
	alpha.beta		0.15	0.1	2290	49						
	beta		0.13	0.13	2290	49	1	1004	4000	3746	1.27	0.41
	Control SD		0.79	0.09	3261	82	1	2029	4000	3746	0.08	-0.78
	Ctrl MPA SD		0.09	0.06	3261	82	1.01	409	4000	3746	0.16	-0.63
	df1		2.89	0.22	3261	82	1	7389	4000	3746	-0.15	0.03
	df2		1.96	0.24	3261	82	1	2729	4000	3746	0.8	-1.56
	MPA SD		0.26	0.06	3261	82	1	3630	4000	3746	0.5	-0.55
	alpha	NT:MU|NT	0.09	0.09	1300	40	1.01	1273	4000	3746	1.24	-1.95
	alpha.beta		0.26	0.1	2119	44						
	beta		0.17	0.11	2119	44	1	1649	4000	3746	-1.33	0.94
	Control SD		0.92	0.09	3419	84	1	2255	4000	3746	-0.04	-0.44
	Ctrl MPA SD		0.06	0.05	3419	84	1.08	426	4000	3746	-0.13	-0.9
	df1		2.57	0.16	3419	84	1	8246	4000	3746	0.41	1.37
	df2		2.25	0.3	3419	84	1	2855	4000	3746	-0.22	-0.45
	MPA SD		0.27	0.07	3419	84	1	2871	4000	3746	0.66	0.75

Table S12 | Diagnostic statistics for parameters in the Bayesian models with a binary predictor (beta) to identify sites in MPA with (alpha.beta) or without (alpha) adequate staff capacity. Intercept (alpha) represents the MPA treatment effect of interest. SD: standard deviation; df: degrees of freedom. alpha.beta parameters are the sum of the intercept (alpha) and predictor (beta) coefficients. See Table S6 for comparison descriptions.

			Log Biomass Ratio								Geweke Z Scores	
Predictor	**Variable**	**Comparison**	**Mean**	**SD**	**# Sites**	**# MPAs**	**Rhat**	**Effective Sample Size**	**Total Iterations**	**Raftery & Lewis Min # Iterations**	**Chain 1**	**Chain 2**
Adequate staff capacity	alpha	MU:0	-0.06	0.12	1896	20	1.01	1006	4000	3746	0.54	-0.6
	alpha.beta		0.66	0.47	525	6						
	beta		0.72	0.48	525	6	1.01	280	4000	3746	-1.69	1.83
	Control SD		0.86	0.13	2421	26	1	3113	4000	3746	2.35	1.3
	df1		1.15	0.07	2421	26	1	4700	4000	3746	-0.73	0.34
	df2		1.39	0.19	2421	26	1	3853	4000	3746	3.03	0.82
	MPA SD		0.14	0.09	2421	26	1	1867	4000	3746	1.99	0.31
	alpha	NT:0	0.25	0.23	138	15	1	1370	4000	3746	0.84	0.22
	alpha.beta		0.58	0.42	201	7						
	beta		0.34	0.47	201	7	1	1095	4000	3746	-1.72	-0.73
	Control SD		1.26	0.31	339	22	1	1210	4000	3746	-0.14	-0.43
	df1		4.99	3.95	339	22	1	2030	4000	3746	-1.32	0.61
	df2		10.02	7.92	339	22	1	1949	4000	3746	-0.2	0.04
	MPA SD		0.45	0.2	339	22	1	1733	4000	3746	0.03	2.5
	alpha	NT:MU|MU	0.78	0.32	544	13	1.03	406	4000	3746	1.34	0.32
	alpha.beta		-0.32	0.31	505	6						
	beta		-1.11	0.41	505	6	1.02	581	4000	3746	-1.45	-0.73
	Control SD		1.47	0.33	1049	19	1	790	4000	3746	1.89	1.1
	Ctrl MPA SD		0.21	0.18	1049	19	1	668	4000	3746	-0.26	-0.9
	df1		2.49	0.35	1049	19	1	5902	4000	3746	-0.22	-0.61
	df2		2.34	0.78	1049	19	1	779	4000	3746	0.28	0.81
	MPA SD		0.28	0.17	1049	19	1	1305	4000	3746	-1.59	-2.8
	alpha	NT:MU|NT	0.01	0.35	133	9	1	961	4000	3746	-1.34	0.48
	alpha.beta		0.36	0.37	203	8						
	beta		0.35	0.48	203	8	1	1088	4000	3746	1.79	-0.74
	Control SD		1.29	0.42	336	17	1	793	4000	3746	1.1	0.57
	Ctrl MPA SD		0.25	0.22	336	17	1	769	4000	3746	1.67	0.48
	df1		5.61	4.14	336	17	1	1976	4000	3746	1.43	0.95
	df2		6.37	6.72	336	17	1.01	722	4000	3746	-0.13	0.55
	MPA SD		0.37	0.22	336	17	1	1370	4000	3746	1.13	-0.1

Table S13 | Diagnostic statistics for parameters in the Bayesian models with a binary predictor (beta) to identify sites in MPA with (alpha.beta) or without (alpha) strong sustainable use regulations. Intercept (alpha) represents the MPA treatment effect of interest. SD: standard deviation; df: degrees of freedom. alpha.beta parameters are the sum of the intercept (alpha) and predictor (beta) coefficients. See Table S6 for comparison descriptions.

			Log Biomass Ratio								Geweke Z Scores	
Predictor	**Variable**	**Comparison**	**Mean**	**SD**	**# Sites**	**# MPAs**	**Rhat**	**Effective Sample Size**	**Total Iterations**	**Raftery & Lewis Min # Iterations**	**Chain 1**	**Chain 2**
Sustainable use regulations	alpha	MU:0	-0.12	0.17	1788	7	1	406	4000	3746	0.08	-2.21
	alpha.beta		0.09	0.19	629	18						
	beta		0.21	0.26	629	18	1	542	4000	3746	-0.14	1.27
	Control SD		0.88	0.13	2417	25	1	3098	4000	3746	1.87	2.27
	df1		1.14	0.07	2417	25	1	4866	4000	3746	-1.12	-1.35
	df2		1.42	0.2	2417	25	1	3221	4000	3746	1.97	2.16
	MPA SD		0.18	0.1	2417	25	1	1523	4000	3746	1.99	1.61
	alpha	NT:0	0.12	0.5	9	4	1.02	285	4000	3746	-0.46	-0.65
	alpha.beta		0.41	0.24	325	17						
	beta		0.3	0.55	325	17	1.03	281	4000	3746	0.33	0.93
	Control SD		1.13	0.3	334	21	1	1295	4000	3746	-0.12	0.7
	df1		5.5	4.51	334	21	1.01	2036	4000	3746	-0.83	1.1
	df2		8.9	7.45	334	21	1	1834	4000	3746	-0.28	1.37
	MPA SD		0.48	0.19	334	21	1	1965	4000	3746	-0.54	0.9
	alpha	NT:MU|MU	1.07	0.47	470	4	1.08	208	4000	3746	-0.86	-0.07
	alpha.beta		0.01	0.28	575	14						
	beta		-1.06	0.54	575	14	1.11	205	4000	3746	1.1	-0.22
	Control SD		1.41	0.31	1045	18	1.01	1062	4000	3746	-1.35	0.17
	Ctrl MPA SD		0.19	0.17	1045	18	1	642	4000	3746	-0.29	-1.57
	df1		2.5	0.36	1045	18	1	5710	4000	3746	-0.72	-0.09
	df2		2.22	0.65	1045	18	1	1113	4000	3746	-1.82	-0.08
	MPA SD		0.35	0.18	1045	18	1	1225	4000	3746	-0.43	-0.08
	alpha	NT:MU|NT	0.09	0.9	7	2	1.05	164	4000	3746	0.59	-1.39
	alpha.beta		0.22	0.32	324	14						
	beta		0.13	0.95	324	14	1.04	171	4000	3746	-0.48	1.42
	Control SD		1.42	0.45	331	16	1.03	755	4000	3746	1.66	-0.45
	Ctrl MPA SD		0.28	0.25	331	16	1.01	750	4000	3746	1.2	-0.86
	df1		6.01	4.69	331	16	1	2493	4000	3746	0.23	0.56
	df2		7.64	7.43	331	16	1.02	865	4000	3746	1.22	-0.02
	MPA SD		0.47	0.25	331	16	1.01	1473	4000	3746	1.48	-1.06