https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

A multi-decadal, depth-resolved Area-of-Habitat dataset for global marine species

Yixuan Li[1, #], Shijun Zheng[1, #], Le Yu[1,2, *], Zhenrong Du[3], Fanghua Xu[1], Mingjie Li[4], Zhibing Jiang[5], Yibo Liao[5], Yining Chen[5], Bangyi Tao[6]

[1] Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing 100084, China

[2] Ministry of Education Ecological Field Station for East Asian Migratory Birds, Beijing 100084, China

[3] School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China

[4] Technology Innovation Center for South China Sea Remote Sensing, Surveying and Mapping Collaborative Application, South China Sea Development Research Institute, Ministry of Natural Resources, Guangzhou 510300, China

[5] Key Laboratory of Marine Ecosystem Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China

[6] State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China

# These authors contributed equally.

* Corresponding author: Le Yu (leyu@tsinghua.edu.cn)

Abstract

Marine biodiversity is being rapidly redistributed in response to ocean change, yet globally consistent datasets that resolve suitable habitat across both depth and time remain scarce. Here, we present a multi-decadal, depth-resolved Area-of-Habitat (AOH) dataset for a globally distributed subset of marine species assessed by the IUCN Red List. We integrated approximately 45.5 million occurrence records with IUCN expert range maps and physical and biogeochemical ocean reanalysis data spanning 1993–2025. Species-specific MaxEnt models were successfully fitted for 5,619 species primarily fishes (4,815 species), together with reef-forming corals, molluscs, lobsters, marine mammals, marine reptiles, mangroves and seagrasses. The fitted models were

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

projected across 50 standard depth layers for 1995, 2000, 2005, 2010, 2015, 2020 and 2025. Phytoplankton, zooplankton and macroalgae were not included. Model predictions were constrained to the original IUCN ranges to generate continuous, range-constrained estimates of environmental suitability, operationally defined here as marine AOH. The resulting dataset comprises species-level layers and aggregated products resolved by depth and year. Internal consistency assessments showed that predicted suitability was generally higher at species occurrence locations across the evaluated species and depth layers. Independent comparisons with global coral-reef distributions and FishGlob scientific survey observations further supported the ecological plausibility and spatial realism of the resulting AOH estimates.

Aggregated analyses revealed pronounced vertical differentiation in marine AOH, with shallow and deeper waters exhibiting distinct spatial concentrations and temporal trajectories. By integrating geographic space, ocean depth and multi-decadal environmental change within a consistent modelling framework, this dataset provides a standardized resource for investigating marine habitat redistribution, identifying persistent areas of high suitability and potential climate refugia, and evaluating the correspondence between suitable habitats and marine protected areas in a changing ocean. The dataset is publicly available at https://doi.org/10.5281/zenodo.21933262 (Yu and Li, 2026).

1. Introduction

Marine biodiversity is declining under the combined pressures of climate change, overexploitation, pollution and habitat degradation, with broad consequences for ecosystem functioning and the benefits that marine ecosystems provide to society (IPBES, 2019; Pörtner et al., 2023). In response, the Kunming–Montreal Global Biodiversity Framework calls for the effective conservation and management of at least 30% of coastal and marine areas by 2030 (CBD, 2022). Achieving this target requires spatially explicit, repeatable and comparable information on species and their habitats, enabling biodiversity patterns, conservation priorities and protected-area coverage to

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

be assessed consistently across space and time. Area of Habitat (AOH) mapping provides a standardized framework for delineating environmentally suitable habitat within species' geographic ranges and for aggregating such information across species to support biodiversity assessment (Brooks et al., 2019; Lumbierres et al., 2022). When generated consistently across multiple time points, AOH maps can further characterize changes in the extent and spatial distribution of suitable habitat, thereby supporting biodiversity monitoring and conservation planning (Yang et al., 2025; Li et al., 2025). Extending this framework to marine systems therefore offers an important basis for assessing large-scale patterns of marine habitat suitability and tracking their changes through time.

Substantial progress has been made in developing standardized AOH datasets for terrestrial biodiversity, including multispecies products for mammals, birds and amphibians at regional to global scales (Rondinini et al., 2011; Ficetola et al., 2015; Palacio et al., 2021; Lumbierres et al., 2022). More recent datasets have further incorporated historical time series or future scenario projections, enabling dynamic assessments of habitat change (Yang et al., 2025; Li et al., 2025). The need for dynamic assessment also extends to marine systems, including deeper waters where species adapted to long-term environmental stability are nevertheless expected to face increasing climate velocities even under ambitious mitigation pathways (Brito-Morales et al., 2020; UNEP, 2026). Comparable global marine mapping has advanced primarily through habitat-suitability and species-distribution frameworks such as AquaMaps and AquaX, which combine occurrence records, geographic constraints and environmental information to estimate suitable areas for thousands of marine species (Kaschner et al., 2006; Ready et al., 2010; Reygondeau et al., 2026). However, existing global marine products remain limited in two important respects. First, most represent either climatological baselines or a small number of widely separated future periods, rather than a consistent multi-decadal sequence suitable for examining historical changes in marine habitat. Such limited temporal coverage constrains the ability to distinguish

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

persistent redistribution from short-term variability and to identify when suitable habitats have expanded, contracted or shifted. Second, existing products generally represent marine distributions in two-dimensional geographic space or distinguish only broadly between surface and bottom environments. Yet the ocean is strongly vertically stratified: temperature, oxygen, salinity, light, productivity and pressure vary substantially through the water column, and both species composition and habitat suitability can differ markedly among depth layers (Duffy and Chown, 2017; Woolley et al., 2016; Woods et al., 2023; Duhamet et al., 2023). Models that associate observations collected at depth with surface conditions, or that collapse the water column into a small number of broad depth categories, may therefore mischaracterize species' environmental niches and obscure vertically structured habitat redistribution (Duffy and Chown, 2017; Robinson et al., 2017; Klaassen et al., 2025). Together, limited temporal coverage and coarse vertical representation constrain comprehensive assessments of marine habitat change across space, depth and time.

To address these limitations, we developed a multi-decadal, depth-resolved AOH dataset for global marine species. We integrated approximately 45.5 million occurrence records from the Global Marine Biodiversity Dataset (Dou, 2024) with physical and biogeochemical ocean reanalysis data spanning 1993–2025 from two Copernicus Marine Service products (Copernicus Marine Service, 2024a, 2024b). Environmental conditions were harmonized across 50 standard depth layers and used to fit species-specific MaxEnt models for 5,619 marine species. The fitted models were projected for seven target years (1995, 2000, 2005, 2010, 2015, 2020 and 2025), selected at regular five-year intervals to provide temporally consistent snapshots and to include 2020 as a policy-relevant reference point marking the endpoint of the Aichi Biodiversity Targets and the transition to the post-2020 global biodiversity framework. These projections generated species-level estimates of range-constrained environmental suitability across geographic space, depth and time. We assessed the resulting dataset using complementary internal consistency checks at the species and depth-layer levels

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

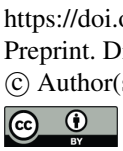
together with independent comparisons against global coral-reef distributions and scientific bottom-trawl observations. By integrating long-term temporal coverage with explicit vertical resolution, this dataset provides a standardized resource for investigating marine habitat redistribution, examining depth-dependent changes in habitat suitability, identifying persistent areas of high suitability and potential climate refugia, and evaluating the correspondence between suitable habitats and marine protected areas in a changing ocean.

2. Methods

2.1 Overall workflow

We developed an integrated workflow to generate a multi-decadal, depth-resolved Area-of-Habitat (AOH) dataset for global marine species (Fig. 1). The workflow combined species occurrence records, IUCN expert range maps, and depth-resolved physical and biogeochemical ocean variables to estimate species-specific environmental suitability across geographic space, depth and time. It consisted of three main stages: preparation and harmonization of biological and environmental inputs; species-specific distribution modelling and spatiotemporal projection; and construction, aggregation and evaluation of the resulting AOH products. First, species occurrence records, IUCN range maps and depth-resolved environmental variables were compiled, screened and harmonized to a common spatial, temporal and vertical framework. Second, independent MaxEnt models were fitted for individual species using occurrence records together with depth-specific, sampling-bias-adjusted background points, and the fitted models were projected across seven target years and the depth layers represented by each species. Third, the continuous suitability predictions were constrained to the original IUCN ranges to generate species-level marine AOH layers. These layers were subsequently aggregated using area-weighted metrics to characterize spatial, vertical and temporal patterns, and their plausibility was assessed through internal consistency checks and independent external comparisons.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.



Open Access Earth System Science Data Discussions

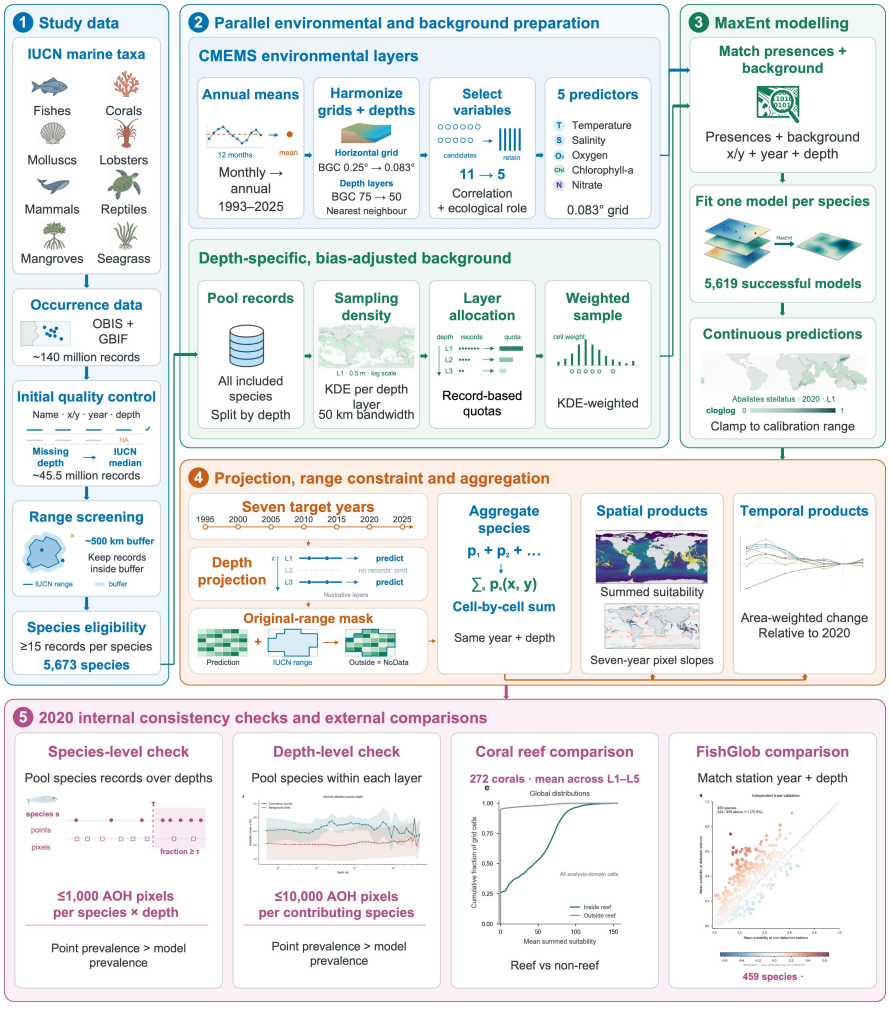


Fig. 1. Workflow for generating the multi-decadal, depth-resolved marine Area-of-Habitat dataset.

2.2 Species and environmental data

We first compiled accepted scientific names and species-specific depth limits for marine species. Because the IUCN Red List covers marine species that have undergone formal conservation assessment, the resulting species pool represents a taxonomically incomplete subset of described marine biodiversity rather than a comprehensive inventory of all marine species. In this study, “global marine species” therefore refers

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

to the worldwide spatial coverage of the IUCN-assessed marine taxa included in the analysis, subject to the availability of usable depth, occurrence and range information. Upper and lower depth limits were available for 16,445 species.

We then obtained approximately 140 million raw marine species occurrence records from the Global Marine Biodiversity Dataset (Dou, 2024), released by the Oceanographic Data Center, Chinese Academy of Sciences. This point-based dataset integrates occurrence records originally sourced from the Ocean Biodiversity Information System (OBIS) and the Global Biodiversity Information Facility (GBIF), provides environmental attributes matched to the coordinates, depths and collection times of individual records, and includes species names matched to the accepted names provided by IUCN. Records were screened for valid scientific names, geographic coordinates, depths and collection years. Where an observed depth was unavailable, it was assigned using the median of the species' IUCN depth range; records lacking both observed and IUCN-derived depth information were excluded. Following this initial screening and matching with the IUCN species list, approximately 45.5 million occurrence records representing 15,128 species remained, while 1,317 IUCN species without usable occurrence records were excluded. To reduce the risk of unstable models under our depth-stratified, density-weighted background-sampling design, we required a minimum of 15 valid occurrence records per species for species-level MaxEnt model fitting, consistent with the threshold adopted in comparable marine MaxEnt applications (La Manna et al., 2024). Applying this criterion retained 5,673 species for model training (Section 2.3).

Pairwise correlations among the 11 candidate predictors were evaluated using the harmonized, spatially continuous CMEMS environmental grids (Copernicus Marine Service, 2024a, 2024b). These predictors represent environmental variables commonly used to characterize marine species distributions, including temperature, salinity, ocean currents, oxygen conditions, nutrient availability, acidity and primary productivity (Tyberghein et al., 2012; Assis et al., 2018; Pickens et al., 2021). The 11 variables

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

included seawater temperature, salinity, eastward current velocity, northward current velocity, chlorophyll-a concentration, nitrate concentration, phosphate concentration, silicate concentration, dissolved oxygen concentration, pH and net primary production. Pairwise Pearson correlations were calculated to assess collinearity, with $|r| > 0.8$ considered indicative of strong correlation. Based on the observed correlation structure across representative depth layers, together with ecological relevance, we retained seawater temperature, salinity, chlorophyll-a concentration, nitrate concentration and dissolved oxygen concentration. Although temperature and dissolved oxygen were strongly correlated in some upper-ocean layers, both variables were retained because they represent distinct thermal and aerobic constraints on marine species distributions, and their correlation varied substantially with depth. Their selection was also supported by their established relevance to marine species distributions. Temperature is a major environmental driver shaping the distribution and diversity patterns of marine organisms (Tittensor et al., 2010), while dissolved oxygen constrains their aerobic habitats (Deutsch et al., 2020). Salinity and nitrate reflect hydrographic and resource gradients and have been identified as important predictors of marine species distributions, although their relative importance varies among taxonomic groups and depth zones (Egorova et al., 2025). Chlorophyll-a provides an indicator of phytoplankton biomass and associated resource availability and has contributed substantially to explaining the spatiotemporal distributions of pelagic fishes (Wang et al., 2018).

The selected predictor layers were obtained from the Copernicus Marine Service (CMEMS) for model calibration and spatiotemporal projection (Copernicus Marine Service, 2024a, 2024b). Monthly data covering 1993–2025 were aggregated into annual means. The physical variables were provided at a native horizontal resolution of approximately 0.083° (1/12°), whereas the biogeochemical variables were provided at 0.25°. To establish a common horizontal grid, the biogeochemical variables were resampled to the 0.083° physical grid using nearest-neighbour interpolation. Seawater

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

temperature and salinity were provided across 50 standard depth layers, whereas chlorophyll-a concentration, nitrate concentration and dissolved oxygen concentration were originally provided across 75 depth layers. The biogeochemical variables were also aligned with the 50 physical depth layers using nearest-neighbour interpolation. The processed environmental layers were used to extract environmental values for occurrence and background records and to generate the temporal and depth-resolved model projections described in Section 2.3. Species range polygons were obtained from the IUCN Red List spatial database and used to screen occurrence records and constrain the spatial extent of the species-level AOH estimates. We retained polygons with presence classified as Extant, Probably Extant, or Possibly Extant and origin classified as Native, Reintroduced, or Introduced. Multiple polygons belonging to the same species were merged by scientific name. To exclude geographically distant occurrence records while retaining observations near mapped range boundaries, each species' range was buffered outward by approximately 500 km. Occurrence records were intersected with the corresponding buffered ranges, and only records located within these areas were retained for model fitting.

The labels L1–L50 denote the 50 physical-product depth coordinates, from approximately 0.5 to 5,727.9 m, rather than 50 equal-thickness depth intervals. This retains the native vertical framework of the CMEMS physical reanalysis (Copernicus Marine Service, 2024b). For Figures 2–4, eight illustrative levels were used: L1, L2, L5, L10, L15, L20, L30 and L40 (0.5, 1.5, 5.1, 13.5, 29.4, 65.8, 380.2 and 1,941.9 m). This subset provides a compact view of densely resolved shallow waters and progressively deeper levels; it is a visualization subset, not a separate ecological zonation or a statistically optimized selection. Modelling retained the full 50-level framework, subject to species-level data availability.

The approximately 500-km range buffer was a pragmatic tolerance for screening occurrences near expert-mapped range boundaries, while excluding geographically remote records. A 500-km buffering scale has also been used in marine SDMs, for

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

example around occurrence-derived domains in clownfishes (García-Jiménez et al., 2025). That application differs from buffering IUCN polygons and provides a methodological precedent rather than evidence that 500 km is optimal here. The buffer does not represent a common dispersal distance across taxa, and it does not expand the final AOH maps, which remain restricted to the original IUCN polygons. Alternative buffer distances were not evaluated in this study.

2.3 Species distribution modelling and temporal projection

We fitted an independent MaxEnt model for each species using the maxnet package in R (Phillips et al., 2017). MaxEnt estimates environmental suitability by contrasting the environmental conditions associated with species occurrences against those available at a set of background locations, making it suitable for presence-only occurrence data for which systematically collected absence records are unavailable (Phillips et al., 2006). Model calibration was based on the five environmental predictors retained in Section 2.2. Separate models were fitted for individual species so that each model characterized the environmental conditions associated with that species rather than pooling environmental responses across taxa. Feature classes were selected using the default sample-size-dependent settings of maxnet, and models were fitted using the default regularization multiplier (regmult = 1) (Phillips, 2021).

To reduce the influence of spatial sampling bias, we adapted the target-group background approach to depth-resolved marine occurrence data. Spatially biased occurrence records combined with uniformly sampled background points can cause presence–background models to reproduce geographic variation in sampling intensity rather than the environmental conditions associated with species occurrences. Drawing background points according to the sampling distribution of the broader target group can reduce this effect (Phillips et al., 2009; Barber et al., 2022). We therefore pooled the occurrence records of all included species and constructed a kernel-density surface for each of the 50 standard depth layers using a bandwidth of 50 km and a spatial resolution of 0.083°. These depth-specific surfaces represented the spatial distribution

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

of marine sampling within each layer and were used as probability weights for background selection. For each species, background points were drawn only from the depth layers represented by its occurrence records, and the total number of background points was distributed among these layers in proportion to the number of occurrence records in each layer. Locations were then sampled using the corresponding kernel-density values as weights. A maximum of 20,000 background points was used when more than 40,000 valid candidate cells were available; otherwise, half of the valid candidate cells were sampled. Candidate points lacking complete values for any of the five predictors were excluded.

The 50-km bandwidth is a fixed smoothing scale for the pooled sampling-effort proxy, rather than a species-specific movement or habitat scale. It was used to reduce highly localized sampling peaks while retaining broad sampling gradients and was not optimized separately for each taxon or depth. The density surfaces were generated on the same 0.083° grid used for the harmonized environmental predictors.

Occurrence and background records were matched to the environmental data according to their geographic coordinates, collection year and assigned depth. The nearest available annual environmental layer and standard depth layer were identified for each record, from which values of the five retained predictors were extracted. Occurrence locations were designated as presences and the density-weighted sampled locations as background records, forming a separate calibration dataset for each species. Records with incomplete predictor values were removed before model fitting. Of the 5,673 species meeting the minimum occurrence threshold, independent MaxEnt models were fitted successfully for 5,619 species; species for which model fitting failed were excluded from subsequent processing. Predictions were generated using the cloglog output format implemented in maxnet, producing continuous suitability values ranging from 0 to 1 (Phillips et al., 2017). These values represent relative environmental suitability under the fitted model and should not be interpreted directly as occurrence probability or observed occupancy.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Each fitted model was projected onto environmental conditions for seven target years: 1995, 2000, 2005, 2010, 2015, 2020 and 2025. Predictions were generated separately for each standard depth layer represented by the species, thereby preserving variation in environmental suitability across both time and depth. Predictor values outside the ranges encountered during model calibration were clamped to the corresponding calibration ranges using the default prediction settings of maxnet, thereby reducing unsupported extrapolation into environmental conditions not represented by the training data (Stohlgren et al., 2011). For each species, the environmental grids and resulting suitability estimates were spatially constrained using the original, unbuffered IUCN range polygons; predictions outside the mapped range were excluded. The resulting continuous suitability estimates were stored separately for each species, target year and depth layer and were subsequently used to construct the marine AOH dataset described in Section 2.4.

2.4 Construction and aggregation of the marine AOH dataset

Conventional terrestrial AOH mapping commonly refines a species' range using habitat and elevation associations. Because the present analysis did not apply a separate marine habitat-type layer, we operationally defined marine AOH as the environmentally suitable portion of the original IUCN range. The resulting values should therefore be interpreted as range-constrained environmental suitability rather than as direct estimates produced by conventional terrestrial habitat and depth filtering.

$$AOH_{s,i,l,t} = R_{s,i}\, S_{s,i,l,t} \quad (1)$$

Here, $S_{s,i,l,t}$ is the predicted suitability of species s in grid cell i, depth layer l and year t, and $R_{s,i}$ is the species-specific IUCN range mask, taking a value of 1 inside the original IUCN range and 0 outside it. Applying this mask retained the continuous suitability estimates within the species' range and excluded predictions outside it. Buffered ranges were used only to screen occurrence records and define the modelling domain; the final suitability estimates were constrained using the original IUCN polygons.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Species-level AOH layers were stored separately for each target year and depth layer. To summarize the joint distribution of modelled species, continuous AOH values were summed cell by cell:

$$SS_{i,l,t} = \sum_{s \in S_l} AOH_{s,i,l,t} \quad (2)$$

where $SS_{i,l,t}$ is the summed suitability in grid cell i, depth layer l and year t, and S_l is the set of modelled species represented in that layer. Continuous values were retained during aggregation to avoid introducing an arbitrary binary suitability threshold (Calabrese et al., 2014; Grenié et al., 2020). Accordingly, summed suitability represents the cumulative modelled suitability of the included species and should not be interpreted as directly observed species richness.

Because grid-cell area varies with latitude, layer-level totals were calculated using cell-area weights:

$$T_{l,t} = \sum_{i \in \Omega_{l,t}} A_i \, SS_{i,l,t} \quad (3)$$

where $T_{l,t}$ is the area-weighted summed suitability for depth layer l and year t, A_i is the area of grid cell i, and $\Omega_{l,t}$ is the set of cells with valid values in that layer and year. We subsequently summarized spatial and temporal variation using the percentage change relative to 2020, the proportion of the layer-level total occurring in tropical waters, the contribution of the highest-valued cells occupying 5% of valid area, the total valid area and the maximum cell value. The temporal trajectories in Fig. 4 use area-weighted means over valid cells, expressed relative to 2020.

Temporal change was evaluated across the seven target years. For each depth layer, an ordinary least-squares regression was fitted to the area-weighted totals. Cell-level trends were estimated only for grid cells with complete values across all target years, and the area-weighted proportion of cells with positive slopes was calculated for each depth layer. The outcomes of these analyses and the assessments of dataset plausibility are reported in the Results section. For interactive exploration, the aggregated layers

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

can be visualized by target year and depth layer through the Marine AOH Viewer provided in the Data availability section.

2.5 Internal consistency and external validation

Because globally consistent independent observations are scarce for a dynamic, depth-resolved marine AOH dataset, we assessed the plausibility of the 2020 predictions through two complementary internal consistency checks and two comparisons with independent external datasets. The internal checks followed the point-prevalence-versus-model-prevalence framework proposed by Dahal et al. (2022), which evaluates whether predicted suitable conditions are more strongly concentrated at species occurrence locations than across the corresponding AOH rasters. Because the occurrence records used in these assessments were also used during model calibration, these analyses were interpreted as internal consistency checks rather than independent tests of predictive performance.

Framework A was conducted at the species level. For each species, occurrence points were pooled across all represented depth layers, and the corresponding cloglog suitability values were extracted from the 2020 predictions. To characterize suitability across the modelled range, we randomly sampled up to 1,000 valid AOH pixels from each represented depth layer of the species' IUCN-range-masked rasters. Framework B was conducted at the depth-layer level. For each of the 50 standard depth layers, suitability values were pooled across the occurrence points of all species represented in that layer, together with up to 10,000 valid AOH pixels sampled from each contributing species' raster. In both frameworks, AOH pixels were restricted to non-missing cells within the original IUCN range of the corresponding species.

We evaluated five fixed suitability thresholds, τ, ranging from 0.1 to 0.5 at intervals of 0.1. At each threshold, point prevalence was calculated as

$$P_{point}(\tau) = \frac{n(occurrence\ points, \hat{S} \geq \tau)}{n(occurrence\ points)} \quad (4)$$

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

where n (occurrence points, $\hat{S} \geq \tau$) is the number of occurrence points with predicted suitability $\hat{S}$ at or above threshold τ, and n(occurrence points) is the total number of occurrence points evaluated. Model prevalence was calculated as

$$P_{model}(\tau) = \frac{n(AOH\ pixels, \hat{S} \geq \tau)}{n(AOH\ pixels)} \quad (5)$$

where n (AOH pixels, $\hat{S} \geq \tau$) is the number of sampled valid AOH pixels with predicted suitability at or above the threshold, and n (AOH pixels) is the total number of sampled valid AOH pixels evaluated within the species' original IUCN range. A species or depth layer was considered to pass the threshold-based assessment when $P_point(\tau) > P_model(\tau)$, indicating that occurrence locations exceeded the threshold more frequently than expected from the corresponding proportion of the modelled AOH. We additionally compared mean suitability at occurrence points with mean suitability across sampled AOH pixels to evaluate the same relationship without converting the continuous predictions into binary suitable and unsuitable areas. Extracted values were retained so that all thresholds and continuous summaries could be calculated from the same occurrence-point and AOH-pixel samples.

The two external datasets were selected to evaluate complementary aspects of model plausibility. FishGlob provides species-level, station-based scientific bottom-trawl observations with associated year and depth information (Maureaud et al., 2024), enabling comparison of predicted suitability between detections and non-detections. In contrast, the UNEP-WCMC global coral-reef dataset provides an independently mapped distribution of reef habitat (UNEP-WCMC et al., 2021), allowing broad-scale spatial correspondence of shallow-water predictions to be assessed.

The first external comparison evaluated whether predicted suitability for reef-building corals corresponded with independently mapped coral reef locations. Of 372 species identified from the IUCN REEF_FORMING_CORALS group, 272 had 2020 predictions for all five shallow layers (L1–L5; representative depths 0.5, 1.5, 2.6, 3.8 and 5.1 m). Continuous predictions were summed across these species within each layer

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

and then averaged equally across the five layers, without applying a species-level binary suitability threshold. The common analysis domain retained ocean cells with consistent valid-prediction support across layers for each species; cells outside all species' prediction support retained the original aggregate-zero background convention. Grid cells intersecting UNEP-WCMC reef polygons (UNEP-WCMC et al., 2021) were classified as reef cells. Reef and non-reef values were compared using a one-sided Mann–Whitney U test, sampling at most 200,000 cells per group (seed 42); descriptive metrics used all eligible cells. Spatial agreement was summarized using the inside-to-outside mean ratio and reef enrichment among cells at or above the 75th percentile of positive mean summed suitability, relative to the reef-cell fraction in the corresponding domain. These comparisons were repeated within the Great Barrier Reef, Coral Triangle and Caribbean, with percentile cutoffs calculated separately for each domain. The analysis evaluated horizontal spatial correspondence without matching reef locations to actual bathymetric depths.

The second external comparison used FishGlob data (Maureaud et al., 2024), which integrates 29 scientific bottom-trawl surveys comprising 216,548 stations across the continental shelves and slopes of North America and Europe. Species captured at a station were treated as presences, whereas surveyed stations at which a species was not captured were treated as non-detections. Matching FishGlob data with the 5,619 successfully modelled species yielded 1,053 candidate species. For each candidate species, stations within its surveyed range were assigned to presence and non-detection groups. When more than 3,000 non-detection stations were available, 3,000 stations were randomly selected. Station coordinates, survey years and sampling depths were matched to the nearest target year and standard prediction depth layer. Stations outside the original IUCN range or lacking a valid prediction were excluded, and only species with at least 10 valid stations in both groups were retained. For each retained species, we compared mean continuous suitability between presence and non-detection stations. A comparison was considered directionally consistent when mean predicted suitability

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

was higher at presence stations than at non-detection stations. Non-detections can reflect imperfect catchability as well as unsuitable conditions.

2.6 Assumptions and constraints

First, species–environment relationships fitted to the pooled calibration records are assumed to be transferable among the seven target years and represented depth levels. Annual means, nearest-year/depth matching and imputed depths simplify seasonal variability, vertical movements and observation uncertainty. Predictions therefore describe environmental suitability under these assumptions, rather than observed occupancy or abundance.

Second, static IUCN polygons delimit all final predictions, so suitable conditions outside these boundaries cannot be recovered by this product. Density-weighted backgrounds approximate sampling effort but cannot fully correct taxonomic, geographic or temporal biases. The fixed bandwidth and buffer settings were not optimized across taxa. Environmental predictors do not explicitly represent substrate or biotic interactions.

Third, comparisons among depths depend on the species and valid cells represented in each layer. Summed suitability is not observed species richness, and an area-weighted mean differs from an area-integrated total when the valid area changes. Internal checks reuse calibration occurrences; reef comparisons assess horizontal habitat correspondence; and trawl non-detections do not establish biological absence. External data provenance alone does not establish that every validation record was independent of the OBIS/GBIF-derived calibration pool because no record-level overlap audit was performed. These assumptions and constraints should be considered when interpreting the results and validation outcomes.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

3. Results

3.1 Global depth-stratified patterns of AOH

Aggregated AOH showed pronounced vertical differentiation across the eight representative depth layers examined for 2020 (selected from L1–L40, approximately 0.5–1,942 m; Fig. 2). Shallow-water layers were characterized by strong spatial concentration and a predominance of AOH in tropical waters, whereas deeper layers exhibited progressively more diffuse spatial patterns. From L2 to L15 (approximately 1.5–29 m), more than half of the area-weighted global AOH total occurred within the tropical band (23.5°S–23.5°N), and the highest-valued grid cells occupying 5% of valid area accounted for well over half of the corresponding layer-level total. From L20 downward (approximately 66–1,942 m), both tropical dominance and spatial concentration declined markedly and subsequently remained comparatively stable. The area-weighted total AOH also decreased substantially with depth, declining by approximately 79% from L1 to L40. In contrast, the proportion of grid cells retaining valid predictions decreased more gradually, from just over 70% at L1 to approximately 59% at L40. The shallowest layer, L1, differed from the immediately deeper shallow layers by exhibiting lower tropical dominance and weaker spatial concentration. Together, these patterns reveal a clear transition from highly concentrated shallow-water AOH to more spatially diffuse distributions at greater depths.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

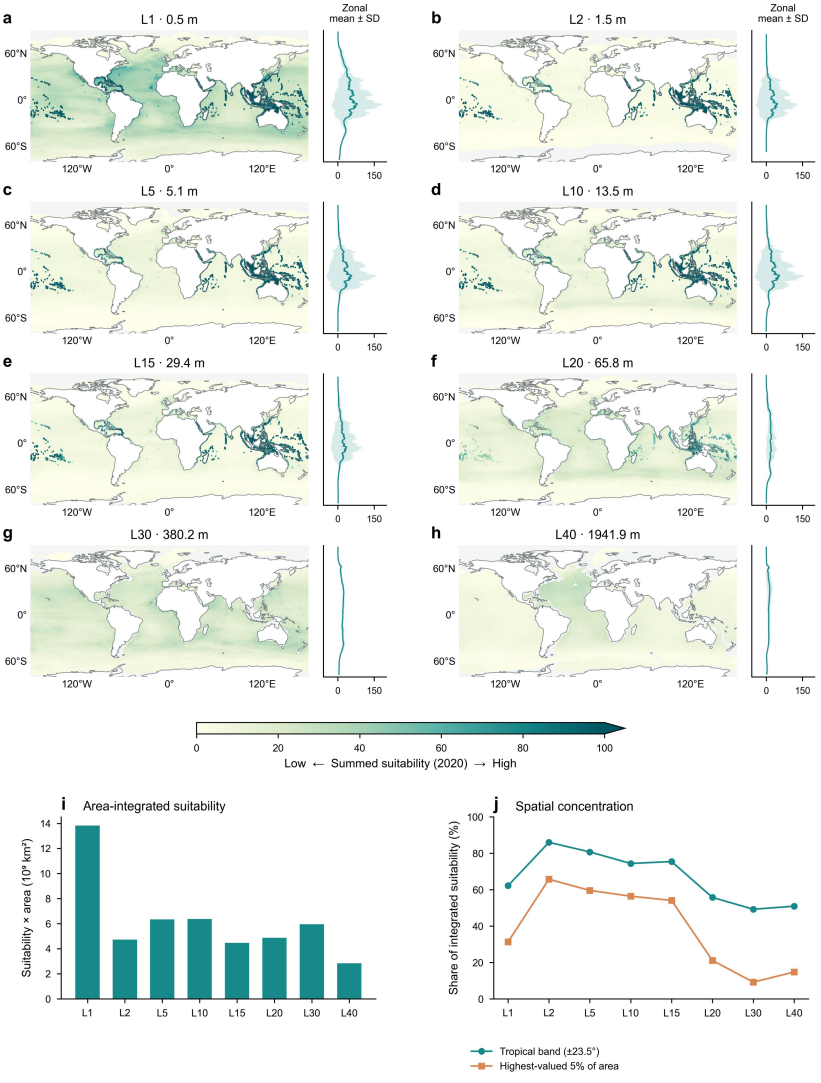


Fig. 2. Global depth-resolved patterns of aggregated marine AOH in 2020. *The accompanying latitudinal profiles show the mean ± SD of aggregated AOH across longitudes at each latitude. Aggregated AOH represents summed modelled environmental suitability across species. (a–h) Maps and latitude profiles for the eight representative layers. (i) Area-integrated summed suitability. (j) Contributions of the tropical band and the highest-valued cells occupying 5% of valid area to the corresponding layer total.*

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

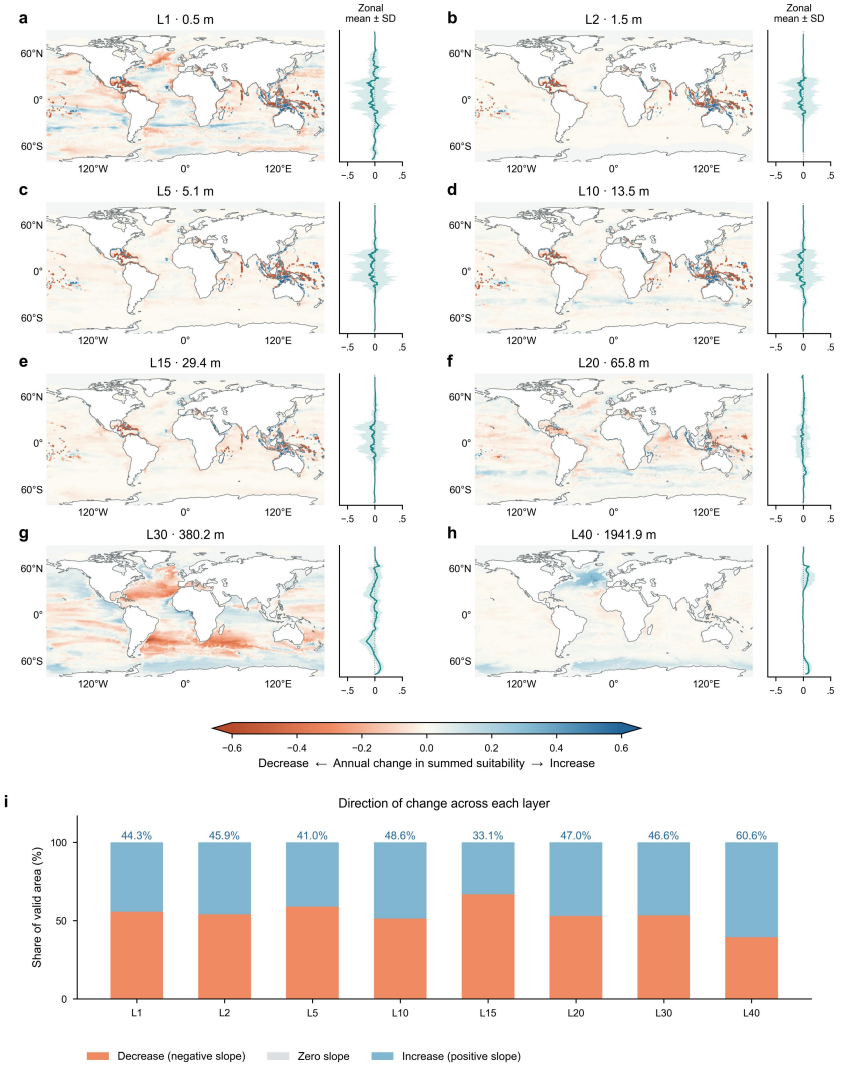
3.2 Temporal changes in AOH during 1995-2025

Temporal trends in aggregated AOH varied markedly among depth layers between 1995 and 2025 (Fig. 3). Pixel-wise linear trends were predominantly negative from L1 to L30, with the strongest overall decline occurring at L5 and a comparatively weaker negative trend at L20. L40 was the only representative depth layer showing a positive net trend, with increasing AOH over more than half of its valid area. At L30, negative trends were concentrated primarily in the mid-latitudes of both hemispheres, particularly in the Southern Hemisphere, whereas trends were close to zero within approximately 23.5° of the equator and became positive poleward of 50°. Prediction coverage at L50 was insufficient to support reliable trend estimation.

Layer-level time series showed similarly depth-dependent temporal trajectories (Fig. 4). The five shallowest representative layers (L1, L2, L5, L10 and L15; approximately 0.5–29 m) generally reached their highest area-weighted mean AOH around 2000 and subsequently declined toward 2020, although the magnitude of change varied among layers. L20 fluctuated within a comparatively narrow range throughout the study period. In contrast, L30 exhibited a persistent decline across the seven target years and remained below its earlier levels in 2025, whereas L40 increased consistently from 1995 to 2025. These contrasting trajectories indicate that temporal changes in marine AOH were strongly depth dependent, with shallow and intermediate waters generally declining while deeper layers showed distinct or opposing patterns.

The increasing trajectory at L40 represents a modelled change in environmental suitability at approximately 1,941.9 m, rather than a measured increase in deep-sea abundance or direct evidence of downward migration. The post-2020 comparison contains only the 2020 and 2025 target years and therefore represents a five-year endpoint difference rather than a continuously observed annual trend.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

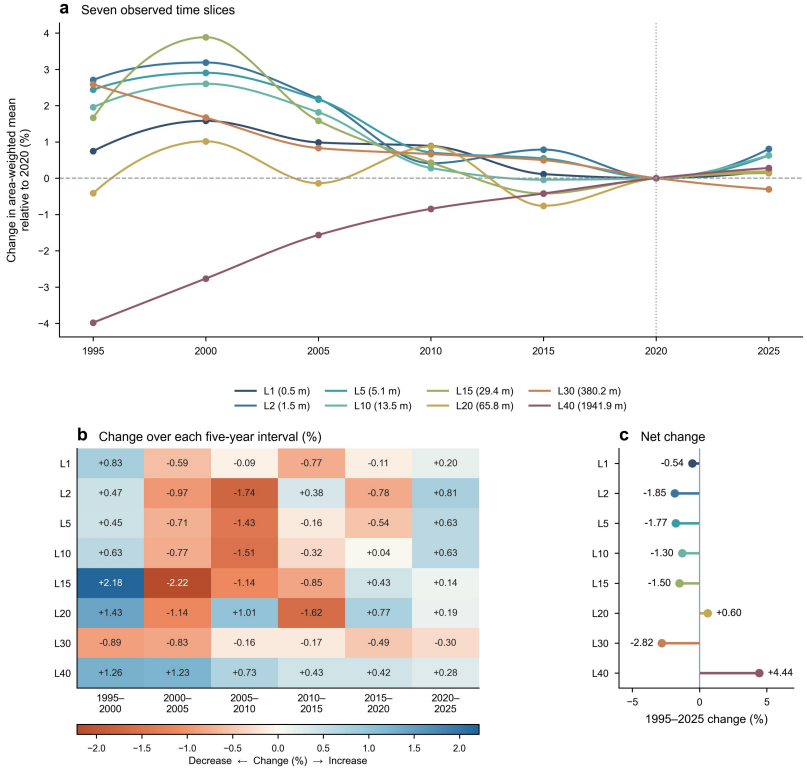


492

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

493 ***Fig. 3. Pixel-wise linear trends in aggregated marine AOH across eight representative depth layers during***
494 ***1995–2025.*** *Linear trends were estimated for each grid cell using aggregated AOH values from seven target*
495 *years (1995, 2000, 2005, 2010, 2015, 2020 and 2025), and only grid cells with valid predictions for all seven*
496 *years were included. The accompanying latitudinal profiles show the mean ± SD of pixel-wise slopes across*
497 *longitudes at each latitude. Panels (a–h) show the maps and profiles; panel (i) shows the area-weighted*
498 *proportions of grid cells with decreasing, zero or increasing trends across the eight depth layers, with labels*
499 *indicating the proportion with positive slopes. Slope direction does not imply statistical significance.*



500

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

501 ***Fig. 4. Temporal trajectories of area-weighted mean marine AOH across eight representative depth***
502 ***layers during 1995–2025.*** *For each depth layer, continuous species-level AOH values were summed across*
503 *species and then averaged over valid grid cells using grid-cell-area weights to obtain the layer-level mean for*
504 *each of the seven target years (1995, 2000, 2005, 2010, 2015, 2020 and 2025). In panel a, values are expressed*
505 *as percentage change relative to the corresponding 2020 mean, allowing temporal trajectories to be compared*
506 *among depth layers despite differences in absolute AOH magnitude. (b) Percentage change over each*
507 *consecutive five-year interval, relative to the earlier value. (c) Net percentage change from 1995 to 2025,*
508 *relative to 1995.*
509

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

3.3 Validation

We evaluated the 2020 predictions at complementary biological and spatial scales. Following the point-prevalence-versus-model-prevalence framework of Dahal et al. (2022), internal consistency was examined first at the species level and then across depth layers. The predictions were subsequently compared with two independent external datasets representing reef-associated spatial patterns and station-level species observations.

3.3.1 Species-level internal consistency

Among the 4,996 species evaluated for 2020, occurrence-point prevalence exceeded model prevalence for the majority of species across all five suitability thresholds (τ = 0.1–0.5; Fig. 5). The species-level pass rate declined progressively as the threshold became more restrictive, from 94.70% (4,731/4,996) at τ = 0.1 to 94.38%, 92.55%, 90.27% and 85.69% (4,281/4,996) at τ = 0.2, 0.3, 0.4 and 0.5, respectively. At the fixed threshold of τ = 0.5, most species lay above the 1:1 line, indicating that suitability values exceeding the threshold occurred proportionally more often at known occurrence locations than across sampled pixels within the corresponding range-constrained AOH. The threshold-free comparison of continuous suitability showed the same overall pattern, with mean predicted suitability generally higher at occurrence points than across sampled AOH pixels. Only 132 species (2.6%) failed the prevalence comparison at all five thresholds, indicating that the observed relationship was robust to the choice of suitability threshold. Together, these results demonstrate strong internal consistency between species occurrences and the corresponding modelled suitability patterns. However, because the occurrence records used in these comparisons also contributed to model calibration, the analysis should be interpreted as an internal consistency assessment rather than as an independent evaluation of predictive performance.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

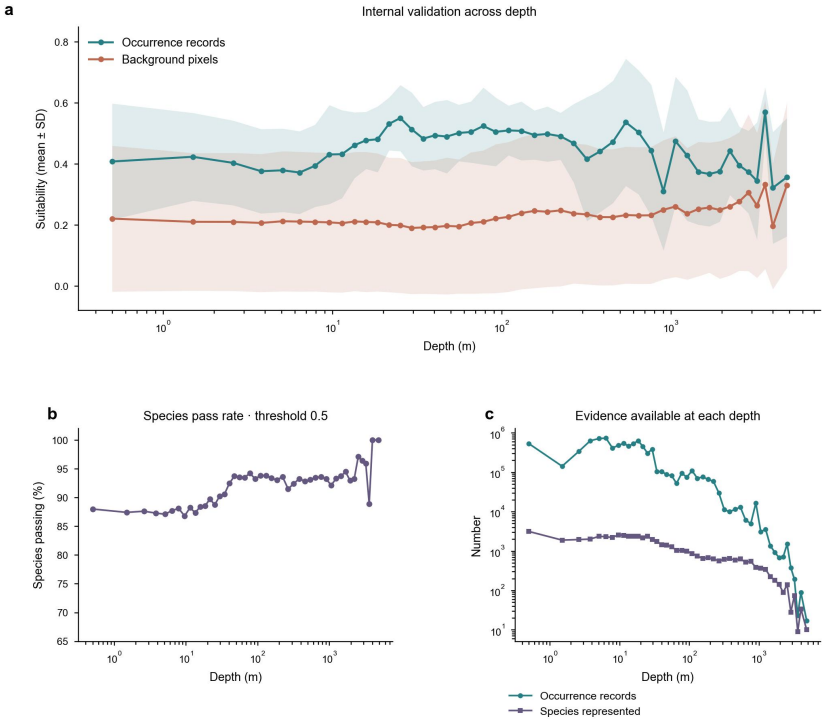


Fig. 5. Depth-resolved internal consistency of marine AOH predictions in 2020. *Mean predicted suitability (cloglog; mean ± SD, panel a) is shown for occurrence points and sampled AOH pixels across the 47 standard depth layers with sufficient data for evaluation. Occurrence-point suitability was extracted from the 2020 species-level predictions, whereas AOH-pixel suitability was sampled from valid cells within the corresponding original IUCN ranges. Panel b shows the species pass rate for each depth layer, defined as the proportion of constituent species for which occurrence-point prevalence exceeded model prevalence at the fixed suitability threshold of 0.5. Because the occurrence records used in this assessment also contributed to model calibration, the comparison represents an internal consistency check rather than an independent validation of predictive performance. Panel c shows the numbers of occurrence records and species represented at each depth.*

3.3.2 Depth-layer-level internal consistency

Of the 50 standard depth layers, 47 contained sufficient data for evaluation. Across all 47 layers, mean predicted suitability at occurrence points was higher than that across

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

sampled AOH pixels, and this pattern persisted from shallow waters to the deepest evaluated layers (Fig. 5). Although both the magnitude and variability of predicted suitability differed among depths, the layer-specific species pass rate at the fixed threshold of $\tau = 0.5$ remained consistently high, averaging 91.95% ± 3.28% and ranging from 86.78% to 100.00%. The lowest pass rate occurred at approximately 9.6 m, whereas the two deepest evaluated layers reached 100%. These results indicate that the concentration of higher predicted suitability at known occurrence locations was maintained across the water column and was not restricted to species-rich surface layers. As with the species-level assessment, however, the occurrence records used here also contributed to model calibration. The analysis should therefore be interpreted as a depth-resolved internal consistency assessment rather than as an independent validation of predictive performance.

3.3.3 External spatial validation using the UNEP-WCMC global coral reef distribution

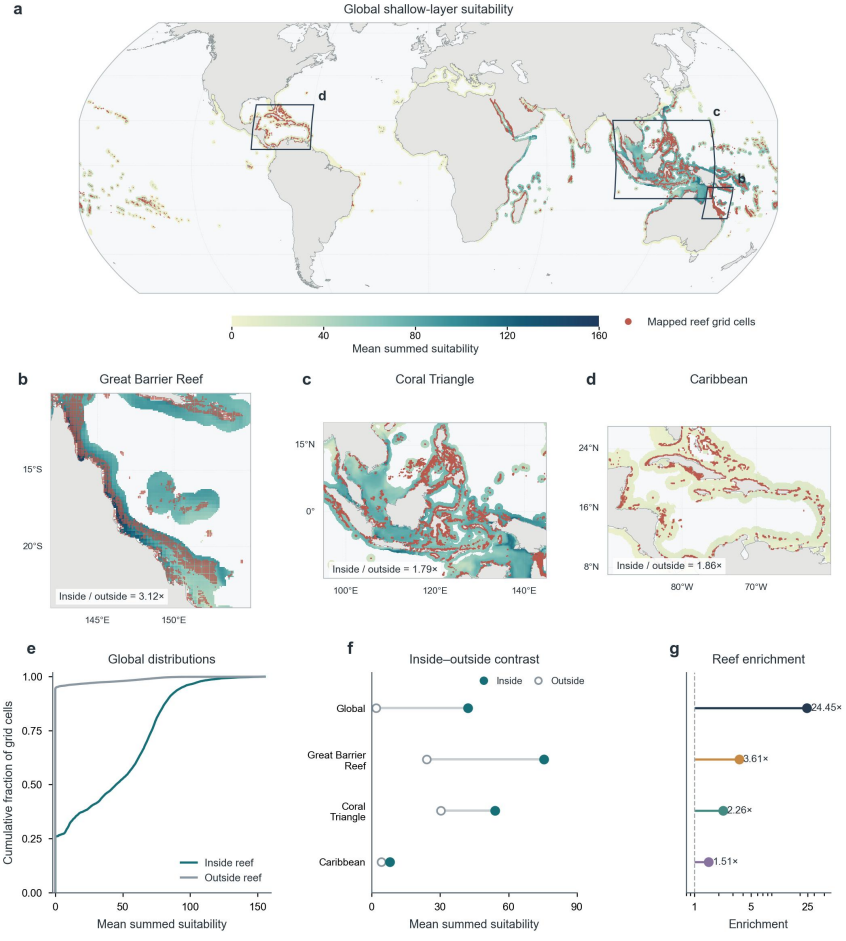
For the 272 reef-building coral species evaluated using the 2020 L1–L5 average (representative depths 0.5–5.1 m), aggregated predicted suitability showed strong spatial correspondence with the independent UNEP-WCMC global coral reef distribution (Fig. 6). Mean shallow-layer summed suitability was 42.17 within mapped reef boundaries and 1.98 outside them, corresponding to a 21.33-fold difference. Areas of high predicted suitability were also disproportionately associated with mapped reefs: 9.60% of grid cells at or above the 75th percentile of positive mean summed suitability overlapped reef boundaries, compared with a global spatial baseline of 0.3927%, representing a 24.45-fold enrichment. The same directional pattern was observed across three major coral-reef regions. Mean summed suitability was consistently higher within than outside mapped reef boundaries in the Great Barrier Reef, Coral Triangle and Caribbean, with inside-to-outside ratios of 3.12, 1.79 and 1.86, respectively. The agreement observed at both global and regional scales indicates that the aggregated

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

shallow-layer predictions recovered the spatial distribution of a well-established marine habitat independently of the data used for model calibration.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.



579

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Fig. 6. External spatial validation of shallow-layer coral suitability against mapped reefs. Continuous 2020 predictions for 2*72 reef-building coral species were summed within layers and averaged equally across L1–L5 (representative depths 0.5–5.1 m), without species-level binarization.* ***(a)*** *Global mean summed suitability and UNEP-WCMC reef cells. (b–d) Great Barrier Reef, Coral Triangle and Caribbean, using the same colour scale. Reef symbols aid visibility without enlarging the quantitative mask. (e) Cumulative suitability distributions inside and outside reef cells, including aggregate-zero ocean background. (f) Corresponding mean suitability. (g) Reef enrichment among cells at or above the 75th percentile of positive suitability, relative to each domain's reef-cell baseline. The common global domain comprised 24,114 reef and 6,115,775 non-reef cells. Grid cells are spatially dependent; this comparison does not validate actual vertical occupancy.*

3.3.4 External station-level validation using FishGlob trawl observations

FishGlob provided an independent station-level evaluation of species-specific AOH predictions using scientific bottom-trawl observations from continental shelf and slope ecosystems. Of the 1,053 FishGlob species shared with the modelled dataset, 459 met the requirement of at least 10 valid presence stations and 10 valid non-detection stations after matching observations to the corresponding prediction year, depth layer and original IUCN range (Fig. 7). Species not retained generally lacked sufficient matched stations or valid predictions within the required spatiotemporal and geographic constraints. Across the 459 evaluated species, mean predicted suitability was higher at presence stations (0.315) than at non-detection stations (0.245). A total of 324 species (70.6%) lay above the 1:1 line, indicating higher mean predicted suitability at presence than at non-detection stations, whereas 135 species showed equal or lower suitability at presence stations. This predominance of species above the 1:1 line indicates that the model predictions generally discriminated observed presences from non-detections in an independent survey dataset. These species-level station contrasts cannot be ranked directly against reef-level enrichment, because the two comparisons differ in response variable, aggregation, spatial domain and observation process. Nevertheless, the overall directional consistency provides independent support for the ability of the species-level

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

Open Access Earth System Science Data Discussions

predictions to distinguish between locations where species were observed and surveyed locations where they were not detected.

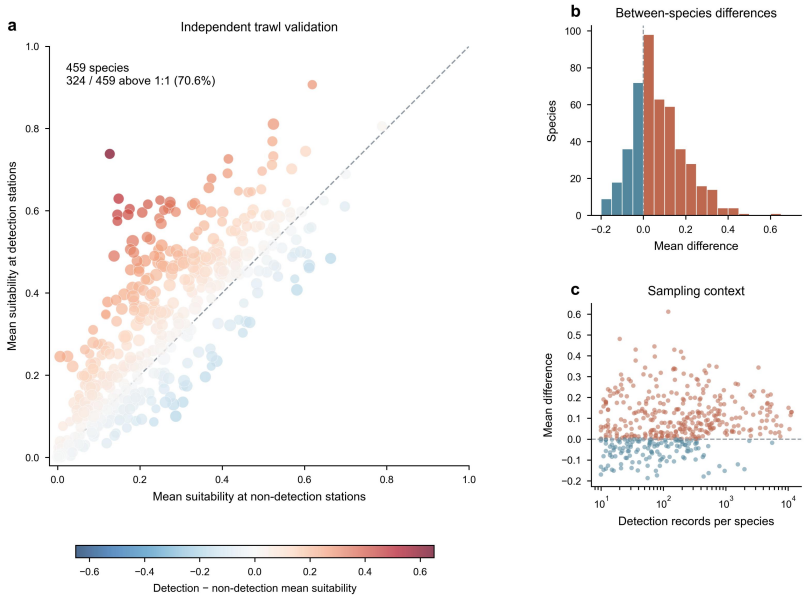


Fig. 7. External validation of species-level AOH predictions using FishGlob scientific bottom-trawl observations. *The comparison included 459 species for which at least 10 valid presence stations and 10 valid non-detection stations could be matched to the corresponding prediction year, depth layer and original IUCN range. In panel a, for each species, the x-axis shows mean predicted suitability at non-detection stations and the y-axis shows mean predicted suitability at presence stations. The diagonal 1:1 line indicates equal mean suitability between the two groups; points above the line therefore represent species for which predicted suitability was higher at presence than at non-detection stations. A total of 324 species (70.6%) occurred above the 1:1 line. Point colour represents the difference between mean suitability at presence and non-detection stations, and point area increases with the logarithm of the total number of stations included for each species. (b) Distribution of species-level presence-minus-non-detection mean differences. (c) Mean differences in relation to the number of presence records per species. Non-detection does not establish biological absence.*

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

4. Discussion

The principal contribution of this dataset is its representation of marine AOH across horizontal space, depth and multiple decades. By integrating species occurrence records, expert range constraints and depth-resolved environmental variables, this framework extends AOH-based biodiversity assessment into the marine environment and provides a consistent basis for examining habitat redistribution through three dimensions. The observed patterns highlight that marine AOH is strongly structured along the vertical axis, with different depth layers showing distinct spatial concentrations and temporal trajectories. Such depth-dependent patterns are consistent with the importance of explicitly representing depth in marine species-distribution modelling (Duffy and Chown, 2017), the pronounced bathymetric structuring and turnover documented for marine fishes (Duhamet et al., 2023), and evidence that climate velocities—and hence the expected pressures for species redistribution—differ substantially among ocean depth zones (Brito-Morales et al., 2020). However, these results represent changes in modelled environmental suitability rather than direct changes in species richness, occupancy, abundance or movement; therefore, changes in AOH should not be interpreted as direct evidence of population redistribution.

The validation analyses provide complementary evidence that the dataset captures meaningful ecological patterns. Internal assessments showed that predicted suitability was generally higher at occurrence locations than across the corresponding range-constrained prediction areas, although these evaluations represent consistency checks because the same occurrence records were used for model calibration. External comparisons further supported the spatial realism of the predictions: predicted suitability for reef-building corals was strongly concentrated within mapped reef areas, and most FishGlob-tested species showed higher mean suitability at observed-presence stations than at observed-absence stations. These two external assessments evaluate different properties of the predictions: the coral-reef comparison examines broad-scale spatial correspondence with an independently mapped habitat, whereas the FishGlob

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

comparison tests discrimination between observed-presence and observed-absence stations. Their results are therefore not expected to be directly equivalent, because different evaluation approaches may target different properties of model performance and differ in their degree of independence from the calibration data (Araújo et al., 2019). The FishGlob comparison may additionally be influenced by variation among surveys in sampling design, spatiotemporal coverage, and gear- and species-specific catchability because the integrated dataset combines multiple scientific bottom-trawl surveys and primarily samples fish communities associated with the seafloor (Maureaud et al., 2024). Together, these assessments support the comparative use of the dataset while highlighting uncertainty associated with taxa, depth layers and observation methods.

Reef-building corals are sessile after settlement, so mapped reef habitat provides a relatively persistent spatial reference compared with mobile fish detections. Nevertheless, reef extent and condition can change, and mapped reef habitat does not demonstrate the presence of each coral species.

The dataset should be interpreted according to its operational definition of marine AOH. Conventional terrestrial AOH is defined as suitable habitat within a species' geographic range and is commonly mapped by refining IUCN range maps using habitat associations and elevational limits (Brooks et al., 2019; Lumbierres et al., 2022). In this dataset, this concept is operationally extended to the marine environment by defining species-level marine AOH as continuous environmental suitability constrained by the original IUCN range. Species-level layers therefore represent range-constrained suitability estimates rather than binary occupied areas, and aggregated summed suitability should not be interpreted as equivalent to observed species richness. Taxonomic coverage was restricted to marine taxa assessed by the IUCN Red List and was further limited by the availability of usable depth, occurrence and range data and successful model fitting. The resulting dataset should therefore not be interpreted as a taxonomically complete representation of global marine biodiversity. Uneven

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

occurrence sampling and the omission of taxon-specific ecological predictors are widely recognized sources of uncertainty in marine SDMs (Robinson et al., 2017; Pickens et al., 2021; Klaassen et al., 2025). Additional dataset-specific uncertainty arises from assigning missing observation depths using available species-level depth information, applying static expert range polygons, using annually averaged environmental conditions and excluding some ecological constraints, such as substrate characteristics, prey availability and biotic interactions. Depth-specific target-group background sampling can reduce the influence of spatially uneven sampling effort (Phillips et al., 2009; Barber et al., 2022), while geographic range constraints and prediction clamping limit geographically implausible projections and unsupported environmental extrapolation (Elith et al., 2010; Stohlgren et al., 2011). Nevertheless, these procedures cannot completely eliminate uncertainty associated with the occurrence data, environmental predictors and modelling assumptions.

Within these limitations, this dataset provides a standardized framework for comparing marine habitat suitability across species, depth layers and time periods. Species-level predictions can support assessments of changes in suitable environmental conditions within mapped ranges, while aggregated products enable analyses of vertically differentiated habitat patterns, persistent suitability hotspots and potential climate refugia. The dataset may also support spatial comparisons between suitable marine habitats, potential climate refugia and existing marine protected areas (Brito-Morales et al., 2020; Zhuang et al., 2025). Its depth-resolved structure further provides an opportunity to evaluate whether protected areas encompass suitable environmental conditions across the water column rather than only at the surface. Future applications should integrate species-specific ecological knowledge, local observations and explicit uncertainty assessments to improve the interpretation of marine biodiversity change.

5. Conclusion

We developed a multi-decadal, depth-resolved marine Area-of-Habitat (AOH) dataset for 5,619 species across 50 standard depth layers. Using environmental

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

reanalysis data spanning 1993–2025, we generated AOH for seven discrete target years (1995, 2000, 2005, 2010, 2015, 2020, and 2025) to capture temporal variation at 5-year intervals. By integrating species occurrence records, IUCN expert range maps, and vertically resolved physical and biogeochemical ocean variables within a consistent modelling framework, the dataset provides range-constrained estimates of environmental suitability across geographic space, ocean depth, and time. Its principal advance lies in explicitly incorporating the vertical dimension into long-term, global marine habitat mapping, thereby extending existing approaches beyond predominantly static or surface-oriented representations. The contrasting spatial patterns and temporal trajectories observed among depth layers further demonstrate the importance of accounting for depth when assessing large-scale changes in marine habitat suitability. Complementary internal consistency assessments and independent external comparisons support the ecological plausibility of the resulting predictions.

The dataset provides a standardized basis for comparing suitable environmental conditions among species, regions, depth zones, and time periods. Species-level layers enable changes in habitat suitability to be examined for individual species across years and depths, whereas aggregated products facilitate analyses of broader spatial patterns, vertically structured suitability hotspots, and potential climate refugia. The dataset may also support assessments of how suitable marine habitats overlap with existing marine protected areas across the water column. As improved occurrence records, ocean reanalysis products, and species-specific ecological information become available, the framework can be updated and refined accordingly. The dataset should nevertheless be interpreted as continuous, range-constrained environmental suitability rather than as direct observations of species occupancy, abundance, or richness.

6. Acknowledgements

This work was supported by the National Key R&D Program of China (2024YFF1307600), the National Natural Science Foundation of China (42401314), and the Technology Innovation Center for South China Sea Remote Sensing, Surveying

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

and Mapping Collaborative Application, Ministry of Natural Resources (RSSMCA-2025-B006).

7. Data availability

The data for this manuscript are available at https://doi.org/10.5281/zenodo.21933262 (Yu and Li, 2026). An interactive visualization of the aggregated marine AOH layers is available through the Marine AOH Viewer at https://zzssjj.projects.earthengine.app/view/marine-aoh-viewer, where users can explore the dataset by target year and standard depth layer. The application provides a visualization interface, whereas the complete species-level and aggregated data files are archived on Zenodo.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

References

Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R. B., Zimmermann, N. E., and Rahbek, C.: Standards for distribution models in biodiversity assessments, Science Advances, 5, https://doi.org/10.1126/sciadv.aat4858, 2019.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., and De Clerck, O.: Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling, Global Ecology and Biogeography, 27, 277–284, https://doi.org/10.1111/geb.12693, 2018.

Barber, R. A., Ball, S. G., Morris, R. K. A., and Gilbert, F.: Target-group backgrounds prove effective at correcting sampling bias in Maxent models, Diversity and Distributions, 28, 128–141, https://doi.org/10.1111/ddi.13442, 2022.

Brito-Morales, I., Schoeman, D. S., Molinos, J. G., Burrows, M. T., Klein, C. J., Arafeh-Dalmau, N., Kaschner, K., Garilao, C., Kesner-Reyes, K., and Richardson, A. J.: Climate velocity reveals increasing exposure of deep-ocean biodiversity to future warming, Nature Climate Change, 10, 576–581, https://doi.org/10.1038/s41558-020-0773-5, 2020.

Brooks, T. M., Pimm, S. L., Akçakaya, H. R., Buchanan, G. M., Butchart, S. H. M., Foden, W., Hilton-Taylor, C., Hoffmann, M., Jenkins, C. N., Joppa, L., Li, B. V., Menon, V., Ocampo-Peñuela, N., and Rondinini, C.: Measuring terrestrial area of habitat (AOH) and its utility for the IUCN red list, Trends in Ecology & Evolution, 34, 977–986, https://doi.org/10.1016/j.tree.2019.06.009, 2019.

Calabrese, J. M., Certain, G., Kraan, C., and Dormann, C. F.: Stacking species distribution models and adjusting bias by linking them to macroecological models, Global Ecology and Biogeography, 23, 99–112, https://doi.org/10.1111/geb.12102, 2014.

Comte, L. and Olden, J. D.: Climatic vulnerability of the world's freshwater and marine fishes, Nature Climate Change, 7, 718–722, https://doi.org/10.1038/nclimate3382, 2017.

Convention on Biological Diversity: Kunming-Montreal Global Biodiversity Framework, 2022.

Copernicus Marine Service: Global Ocean Biogeochemistry Hindcast, https://data.marine.copernicus.eu/product/GLOBAL_MULTIYEAR_BGC_001_029/description, last access: 21 August 2026, 2024a.

Copernicus Marine Service: Global Ocean Physics Reanalysis, https://data.marine.copernicus.eu/product/GLOBAL_MULTIYEAR_PHY_001_030/description, last access: 21 August 2026, 2024b.

Dahal, P. R., Lumbierres, M., Butchart, S. H. M., Donald, P. F., and Rondinini, C.: A validation standard for area of habitat maps for terrestrial birds and mammals,

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

787 Geoscientific Model Development, 15, 5093–5105, https://doi.org/10.5194/gmd-
788 15-5093-2022, 2022.

789 Deutsch, C., Penn, J. L., and Seibel, B.: Metabolic trait diversity shapes marine
790 biogeography, Nature, 585, 557–562, https://doi.org/10.1038/s41586-020-2721-y,
791 2020.

792 Dou, F.: Global Marine Biodiversity Dataset with Environmental Elements,
793 Oceanographic Data Center, Chinese Academy of Sciences [data set],
794 https://msdc.qdio.ac.cn/data/metadata-special-
795 detail?id=1745022461423996930&otherId=1745022461554020354 (last access:
796 21 August 2026), 2024.

797 Duffy, G. and Chown, S.: Explicitly integrating a third dimension in marine species
798 distribution modelling, Marine Ecology Progress Series, 564, 1–8,
799 https://doi.org/10.3354/meps12011, 2017.

800 Duhamet, A., Albouy, C., Marques, V., Manel, S., and Mouillot, D.: The global depth
801 range of marine fishes and their genetic coverage for environmental DNA
802 metabarcoding, Ecology and Evolution, 13, https://doi.org/10.1002/ece3.9672,
803 2023.

804 Durán, A. P., Green, J. M. H., West, C. D., Visconti, P., Burgess, N. D., Virah-Sawmy,
805 M., and Balmford, A.: A practical approach to measuring the biodiversity impacts
806 of land conversion, Methods in Ecology and Evolution, 11, 910–921,
807 https://doi.org/10.1111/2041-210x.13427, 2020.

808 Egorova, Y., Reygondeau, G., Cheung, W. W. L., and Pakhomov, E. A.: Species
809 Distribution Models for Mesopelagic Mesozooplankton Community, Journal of
810 Biogeography, 52, 13–26, https://doi.org/10.1111/jbi.15011, 2025.

811 Elith, J., Kearney, M., and Phillips, S.: The art of modelling range-shifting species,
812 Methods in Ecology and Evolution, 1, 330–342, https://doi.org/10.1111/j.2041-
813 210x.2010.00036.x, 2010.

814 Ficetola, G. F., Rondinini, C., Bonardi, A., Baisero, D., and Padoa-Schioppa, E.:
815 Habitat availability for amphibians and extinction threat: a global analysis,
816 Diversity and Distributions, 21, 302–311, https://doi.org/10.1111/ddi.12296, 2015.

817 García-Jiménez, A., Talbi, M., Fitzgerald, L. M., Heim, A., Marcionetti, A., Schmid,
818 S., Bertrand, J., Shaughnessy, A., Santiago, C., Rangseethampanya, P.,
819 Ruttanachuchote, P., Aunkhongthong, W., Pengsakun, S., Sutthacheep, M.,
820 Malinsky, M., Frédérich, B., Cortesi, F., Kochzius, M., Yeemin, T., Gaboriau, T.,
821 and Salamin, N.: Habitat Specialisation Impacts Clownfish Demographic
822 Resilience to Pleistocene Sea-Level Fluctuations, Molecular Ecology, 34, e70134,
823 https://doi.org/10.1111/mec.70134, 2025.

824 Grenié, M., Violle, C., and Munoz, F.: Is prediction of species richness from stacked
825 species distribution models biased by habitat saturation?, Ecological Indicators,
826 111, 105970, https://doi.org/10.1016/j.ecolind.2019.105970, 2020.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

827 IPBES: Global Assessment Report on Biodiversity and Ecosystem Services of the
828 Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem
829 Services, IPBES, https://doi.org/10.5281/zenodo.6417333, 2019.

830 The IUCN Red List of Threatened Species: https://www.iucnredlist.org, last access: 21
831 August 2026.

832 Kaschner, K., Watson, R., Trites, A., and Pauly, D.: Mapping world-wide distributions
833 of marine mammal species using a relative environmental suitability (RES) model,
834 Marine Ecology Progress Series, 316, 285–310,
835 https://doi.org/10.3354/meps316285, 2006.

836 Klaassen, M., Marques, T. A., Alves, F., and Fernandez, M.: Trends in marine species
837 distribution models: a review of methodological advances and future challenges,
838 Ecography, https://doi.org/10.1002/ecog.07702, 2025.

839 La Manna, G., Ronchetti, F., Moro Merella, M., Vargiu, R., Perretti, F., and Ceccherelli,
840 G.: Using spatial distribution modeling of commercial species to inform
841 management of small-scale fisheries in a Mediterranean marine protected area,
842 Ecology and Evolution, 14, https://doi.org/10.1002/ece3.70169, 2024.

843 Li, B., Cheng, C., Zhang, T., Mu, N., Li, Z., Yang, S., and Wu, X.: Global 1-km habitat
844 distribution for endangered species and its spatial changes under future warming
845 scenarios, Scientific Data, 12, https://doi.org/10.1038/s41597-025-05898-6, 2025.

846 Lumbierres, M., Dahal, P. R., Soria, C. D., Di Marco, M., Butchart, S. H. M., Donald,
847 P. F., and Rondinini, C.: Area of habitat maps for the World's terrestrial birds and
848 mammals, Scientific Data, 9, https://doi.org/10.1038/s41597-022-01838-w, 2022.

849 Maureaud, A. A., Palacios-Abrantes, J., Kitchel, Z., Mannocci, L., Pinsky, M. L.,
850 Fredston, A., Beukhof, E., Forrest, D. L., Frelat, R., Palomares, M. L. D., Pecuchet,
851 L., Thorson, J. T., van Denderen, P. D., and Mérigot, B.: FISHGLOB_data: an
852 integrated dataset of fish biodiversity sampled with scientific bottom-trawl
853 surveys, Scientific Data, 11, 24, https://doi.org/10.1038/s41597-023-02866-w,
854 2024.

855 Palacio, R. D., Negret, P. J., Velásquez-Tibatá, J., and Jacobson, A. P.: A data-driven
856 geospatial workflow to map species distributions for conservation assessments,
857 Diversity and Distributions, 27, 2559–2570, https://doi.org/10.1111/ddi.13424,
858 2021.

859 Pecl, G. T., Araújo, M. B., Bell, J. D., Blanchard, J., Bonebrake, T. C., Chen, I.-C.,
860 Clark, T. D., Colwell, R. K., Danielsen, F., Evengård, B., Falconi, L., Ferrier, S.,
861 Frusher, S., Garcia, R. A., Griffis, R. B., Hobday, A. J., Janion-Scheepers, C.,
862 Jarzyna, M. A., Jennings, S., Lenoir, J., Linnetved, H. I., Martin, V. Y.,
863 McCormack, P. C., McDonald, J., Mitchell, N. J., Mustonen, T., Pandolfi, J. M.,
864 Pettorelli, N., Popova, E., Robinson, S. A., Scheffers, B. R., Shaw, J. D., Sorte, C.
865 J. B., Strugnell, J. M., Sunday, J. M., Tuanmu, M.-N., Vergés, A., Villanueva, C.,
866 Wernberg, T., Wapstra, E., and Williams, S. E.: Biodiversity Redistribution under

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

867 Climate change: Impacts on Ecosystems and Human well-being, Science, 355,
868 eaai9214, https://doi.org/10.1126/science.aai9214, 2017.
869 Phillips, S. J., Anderson, R. P., and Schapire, R. E.: Maximum entropy modeling of
870 species geographic distributions, Ecological Modelling, 190, 231–259,
871 https://doi.org/10.1016/j.ecolmodel.2005.03.026, 2006.
872 Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., and
873 Ferrier, S.: Sample selection bias and presence-only distribution models:
874 implications for background and pseudo-absence data, Ecological Applications,
875 19, 181–197, https://doi.org/10.1890/07-2153.1, 2009.
876 Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., and Blair, M. E.: Opening
877 the black box: an open-source release of Maxent, Ecography, 40, 887–893,
878 https://doi.org/10.1111/ecog.03049, 2017.
879 Phillips, S.: maxnet: Fitting 'Maxent' Species Distribution Models with 'glmnet', R
880 package version 0.1.4, https://CRAN.R-project.org/package=maxnet, 2021.
881 Pickens, B. A., Carroll, R., Schirripa, M. J., Forrestal, F. C., Friedland, K. D, and Taylor,
882 J. C.: A systematic review of spatial habitat associations and modeling of marine
883 fish distribution: A guide to predictors, methods, and knowledge gaps, PLOS ONE,
884 16, e0251818-e0251818, https://doi.org/10.1371/journal.pone.0251818, 2021.
885 Pörtner, H.-O., Scholes, R. J., Arneth, A., Barnes, D. K. A., Burrows, M. T., Diamond,
886 S. E., Duarte, C. M., Kiessling, W., Leadley, P., Managi, S., McElwee, P., Midgley,
887 G., Ngo, H. T., Obura, D., Pascual, U., Sankaran, M., Shin, Y. J., and Val, A. L.:
888 Overcoming the coupled climate and biodiversity crises and their societal impacts,
889 Science, 380, 2023.
890 Ready, J., Kaschner, K., South, A. B., Eastwood, P. D., Rees, T., Rius, J., Agbayani,
891 E., Kullander, S., and Froese, R.: Predicting the distributions of marine organisms
892 at the global scale, Ecological Modelling, 221, 467–478,
893 https://doi.org/10.1016/j.ecolmodel.2009.10.025, 2010.
894 Reygondeau, G., Egorova, Y., Boerder, K., Tittensor, D. P., Kaschner, K., Kesner-
895 Reyes, K., Bailly, N., and Cheung, W. W. L.: AquaX: An enhanced and revised
896 AquaMaps framework to model marine species distributions and biodiversity,
897 PLOS One, 21, e0335823, https://doi.org/10.1371/journal.pone.0335823, 2026.
898 Robinson, N. M., Nelson, W. A., Costello, M. J., Sutherland, J. E., and Lundquist, C.
899 J.: A Systematic Review of Marine-Based Species Distribution Models (SDMs)
900 with Recommendations for Best Practice, Frontiers in Marine Science, 4,
901 https://doi.org/10.3389/fmars.2017.00421, 2017.
902 Rondinini, C., Di Marco, M., Chiozza, F., Santulli, G., Baisero, D., Visconti, P.,
903 Hoffmann, M., Schipper, J., Stuart, S. N., Tognelli, M. F., Amori, G., Falcucci, A.,
904 Maiorano, L., and Boitani, L.: Global habitat suitability models of terrestrial
905 mammals, Philosophical Transactions of the Royal Society B: Biological Sciences,
906 366, 2633–2641, https://doi.org/10.1098/rstb.2011.0113, 2011.

https://doi.org/10.5194/essd-2026-785
Preprint. Discussion started: 29 October 2026
© Author(s) 2026. CC BY 4.0 License.

907 Stohlgren, T. J., Jarnevich, C. S., Esaias, W. E., and Morisette, J. T.: Bounding species
908 distribution models, Current Zoology, 57, 642–647,
909 https://doi.org/10.1093/czoolo/57.5.642, 2011.

910 Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., and Worm,
911 B.: Global patterns and predictors of marine biodiversity across taxa, Nature, 466,
912 1098–1101, https://doi.org/10.1038/nature09329, 2010.

913 Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., and De Clerck, O.:
914 Bio-ORACLE: a global environmental dataset for marine species distribution
915 modelling, Global Ecology and Biogeography, 21, 272–281,
916 https://doi.org/10.1111/j.1466-8238.2011.00656.x, 2012.

917 UNEP-WCMC, WorldFish Centre, WRI, and TNC: Global distribution of coral reefs,
918 Version 4.1, https://doi.org/10.34892/t2wk-5t34, 2021.

919 UNEP: Limiting Overshoot - Navigating exceedance of 1.5°C and pathways towards
920 return, UNEP - UN Environment Programme, 2026.

921 Wang, L., Kerr, L. A., Record, N. R., Bridger, E., Tupper, B., Mills, K. E., Armstrong,
922 E. M., and Pershing, A. J.: Modeling marine pelagic fish species spatiotemporal
923 distributions utilizing a maximum entropy approach, Fisheries Oceanography, 27,
924 571–586, https://doi.org/10.1111/fog.12279, 2018.

925 Woods, B. L., Van, A. P., Hindell, M. A., Raymond, B., Saunders, R. A., Walters, A.,
926 and Rowan Trebilco: Species distribution models describe spatial variability in
927 mesopelagic fish abundance in the Southern Ocean, Frontiers in Marine Science,
928 9, https://doi.org/10.3389/fmars.2022.981434, 2023.

929 Woolley, S. N. C., Tittensor, D. P., Dunstan, P. K., Guillera-Arroita, G., Lahoz-Monfort,
930 J. J., Wintle, B. A., Worm, B., and O'Hara, T. D.: Deep-sea diversity patterns are
931 shaped by energy availability, Nature, 533, 393–396,
932 https://doi.org/10.1038/nature17937, 2016.

933 Yang, Y., Yang, Y., Li, X., Yu, T., Pei, K., Zhong, Y., and Xie, Y.: Area of Habitat
934 maps for national key protected wildlife in China, Scientific Data, 12,
935 https://doi.org/10.1038/s41597-025-05489-5, 2025.

936 Yu, Le and Li, Y.: A dynamic, depth-resolved Area-of-Habitat dataset for global marine
937 species, Zenodo, https://doi.org/10.5281/zenodo.21933262, 2026.

938 Zhuang, H., Zhao, L., Wang, Z., Zhang, Z., and Yang, L.: Identifying global marine
939 climate refugia through a conservative approach to ocean biodiversity preservation,
940 Nature Communications, 16, https://doi.org/10.1038/s41467-025-65791-z, 2025.