Received: 15 July 2022 | Revised: 14 February 2023 | Accepted: 17 February 2023

DOI: 10.1111/geb.13657

DATA ARTICLE

Global Ecology and Biogeography | WILEY

The Reef Environment Centralized InFormation System (RECIFS): An integrated geo-environmental database for coral reef research and conservation

Oliver Selmoni[1] | Gaël Lecellier[2,3] | Véronique Berteaux-Lecellier[2] | Stéphane Joost[4]

[1]Laboratory of Geographic Information Systems (LASIG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[2]UMR250/9220 ENTROPIE IRD-CNRS-Ifremer-UNC-UR, Nouméa, New Caledonia

[3]Institut des Sciences Exactes et Appliquées (ISEA) EA7484, Université de la Nouvelle-Calédonie, Nouméa, New Caledonia

[4]Geospatial Molecular Epidemiology group (GEOME), Laboratory for Biological Geochemistry (LGB), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Correspondence
Stéphane Joost, Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.
Email: stephane.joost@epfl.ch

Handling Editor: Shane Blowes

Abstract

Motivation: Host to intricate networks of marine species, coral reefs are among the most biologically diverse ecosystems on Earth. Over the past few decades, major degradations of coral reefs have been observed worldwide, which is largely attributed to the effects of climate change and local stressors related to human activities. Now more than ever, characterizing how the environment shapes the dynamics of the reef ecosystem (e.g., shifts in species abundance, community changes, emergence of locally adapted populations) is key to uncovering the environmental drivers of reef degradation, and developing efficient conservation strategies in response. To achieve these objectives, it is pivotal that environmental data describing the processes driving such ecosystem dynamics, which occur across specific spatial and temporal scales, are easily accessible to coral reef researchers and conservation stakeholders alike.

Main types of variable contained: Multiple environmental variables characterizing various facets of the reef environment, including water chemistry and physics (e.g., temperature, pH, chlorophyll concentration), local anthropogenic pressures (e.g., boat traffic, distance from agricultural or urban areas) and sea currents patterns.

Spatial location and grain: Worldwide reef cells of 5 by 5 km.

Time period and grain: Last 3–4 decades, monthly and yearly resolution.

Major taxa and level of measurement: Environmental data important for coral reefs and associated biodiversity.

Software format: Interactive web application available at https://recifs.epfl.ch.

KEYWORDS
anthropogenic stress, climate change, conservation, coral reef, environmental database, remote sensing

1 | INTRODUCTION

Hosting a quarter of marine species, coral reefs are among the most productive and biologically diverse ecosystems on Earth (Bouchet, 2006; Knowlton et al., 2010; Reaka-Kudla, 1997). These biodiversity hotspots are currently facing critical threats imposed by both climate change and human-induced local disturbances (Hughes et al., 2017). During the 2009–2018 decade, for example,

14% of hard corals were lost worldwide, a decline that was mainly attributed to anomalous heat waves causing coral bleaching (Souter et al., 2021). Additionally, the resilience of corals to thermal stress can be hampered by local water conditions, such as turbidity levels or nutrient loads (MacNeil et al., 2019). As hard corals shape the physical reef habitat, their loss can lead declines in reef fish abundance (Jones et al., 2004), and such declines can be exacerbated by local stressors such as overfishing (Hughes et al., 2017) and pollution (e.g., agricultural or industrial runoff; Wenger et al., 2015). Now more than ever, it is of paramount importance to characterize how the environment and human disturbances shape the dynamics of the reef ecosystem, such as shifts in species abundance, community changes, or the emergence of locally adapted populations.

Here, we present the Reef Environment Centralized Information System (RECIFS), an online repository of datasets describing the reef environment worldwide over the past few decades. The datasets provided through RECIFS originate from the public domain, and characterize various facets of the reef environment including water chemistry (e.g., chlorophyll concentration, salinity, etc.) and physics (e.g., temperature, heat, velocity, etc.), as well as local anthropogenic impacts (e.g., population density, land use of agricultural and urban areas).

In comparison to existing repositories of marine environmental data (Bio-ORACLE: "Ocean Rasters for Analysis of Climate and Environment" Assis et al., 2018; Tyberghein et al., 2012; MARSPEC: "Ocean Climate Layers for Marine Spatial Ecology", Sbrocco & Barber, 2013), RECIFS has distinctive features to promote the study of coral reef ecosystem dynamics, and to facilitate the uptake of environmental information in reef conservation. These features are:

1. *Reef specificity*: RECIFS is the first database centred on the reef environment worldwide. Data provided through RECIFS characterize key environmental constraints of the reef ecosystem, and such characterization is missing in existing repositories of marine environmental data. Among these environmental constraints are conditions that trigger coral bleaching events (e.g., degree heating week) and that represent human-related disturbances (e.g., boat traffic, coastal land use, human population density).
2. *Spatial–temporal flexibility*: RECIFS data are available under customizable spatial (i.e., environmental measures from 5 to 100 km around a reef of interest) and temporal scales (i.e., environmental measures for a given period of the year, and/or for a given range of years). This is key, since understanding specific dynamics of the reef ecosystem requires environmental data at distinct spatial and temporal resolutions to be available (Fernandez et al., 2017; Leempoel et al., 2017; Mannocci et al., 2017; Melo-Merino et al., 2020; Murray et al., 2018). For instance, the abundance of sessile taxa might be driven by fine-scale changes in water condition (e.g., environmental variation measured at 5-km spatial resolution), whereas for mobile taxa (e.g., sharks) large-scale changes might be more relevant (e.g., 50-km resolution). Similarly, the reef community composition or adaptive processes might be shaped by long-term environmental constraints (e.g., overall average temperature over past few decades), or by seasonal or episodic events (e.g., heat waves during the hot seasons of the past 5 years). In existing repositories of marine environmental data such customization is not possible, since datasets are typically provided at a fixed spatial resolution, and as overall temporal trends (bioclimatic variables, climatologies) of contemporary seascape conditions.
3. *Web application*: RECIFS is available in open access through an intuitive interface requiring no prior knowledge on use of geographic information systems (GIS). While existing repositories of marine environmental data provide online tools for the visualization of the environmental datasets, none of them allows users to subset and process environmental data in real time, nor to customize, annotate and print the map visualization. These characteristics make remotely sensed environmental data accessible to non-specialists of GIS, including to practitioners of coral reef conservation (Selmoni, Lecellier, Ainley, et al., 2020). RECIFS can therefore provide a common syntax to describe the reef environment shared between scientists and conservation stakeholders. For instance, if a study reports that a given environmental variable is associated with a particular shift in a reef ecosystem dynamic (see 'Examples of use' for practical examples), any reef conservation stakeholder worldwide could then use the RECIFS interface to assess how the environmental variable is distributed in their reef system of interest. This mediatory role of RECIFS facilitates the access to environmental information in the conservation domain, as advocated by previous reviews in the field (Foo & Asner, 2019, 2021; Guisan et al., 2013; Hedley et al., 2016).

2 | DATA AND METHODS

2.1 | Data structure: Reef environment

The RECIFS repository is based on three types of datasets describing the reef environment retrospectively. The first type describes data at monthly temporal resolution, the second at a yearly temporal resolution, while the third type is time-invariant (Table 1). These datasets were accessed at a global level as raw data, all sharing the same spatial projection (WGS 84, EPSG:4326), and were then processed and synthesized as outlined here below (and as summarized in Figure 1).

Raw datasets with a monthly temporal resolution were sets of worldwide maps (called stacks of raster images), where every map (called a raster image) described a monthly statistic (usually the average) of a given oceanic environmental variable globally. RECIFS includes the raw monthly datasets for the following environmental variables: chlorophyll concentration, degree heating week, iron concentration, oxygen concentration, pH, nitrate concentration, phosphate concentration, sea current velocity, suspended matter concentration, salinity, and temperature. These variables represent data at near-surface depth, with spatial resolutions ranging between 5–25 km,

TABLE 1 Datasets included in Reef Environment Centralized Information System (RECIFS).

Name	Description	Time window (resolution)	Spatial res.	% NA	Source	Original product id
Chlorophyll concentration (mg/m^3)	Monthly averages of mass concentration of chlorophyll-a in sea water	1997–2021 (monthly)	4 km	12.06	CMEMS	OCEANCOLOUR_GLO_CHL_L4_REP_OBSERVATIONS_009_082
Degree heating week (°C-week)	Monthly maxima of degree heating week (DHW). DHW is calculated as the accumulation of thermal stress (i.e., temperature > 1 °C above the monthly maximal mean temperature) over the previous 12 weeks	1985–2021 (monthly)	5 km	3.14	NOAA CRW	ct5km_dhw-max_v3.1
Iron concentration (mmol/m^3)	Monthly averages of mole concentration of dissolved iron in sea water	1993–2019 (monthly)	0.25°	20.62	CMEMS	GLOBAL_MULTIYEAR_BGC_001_029
O2 concentration (mmol/m^3)	Monthly averages of mole concentration of dissolved molecular oxygen in sea water	1993–2019 (monthly)	0.25°	20.62	CMEMS	GLOBAL_MULTIYEAR_BGC_001_029
Sea water pH	Monthly averages of sea water pH reported on total scale	1993–2019 (monthly)	0.25°	20.62	CMEMS	GLOBAL_MULTIYEAR_BGC_001_029
Nitrate concentration (mmol/m^3)	Monthly averages of mole concentration of nitrate in sea water	1993–2019 (monthly)	0.25°	20.62	CMEMS	GLOBAL_MULTIYEAR_BGC_001_029
Phosphate concentration (mmol/m^3)	Monthly averages of mole concentration of phosphate in sea water	1993–2019 (monthly)	0.25°	20.62	CMEMS	GLOBAL_MULTIYEAR_BGC_001_029
Sea water velocity (m/s)	Monthly averages of sea water surface velocity, computed as the Euclidean norm of eastward and northward velocity	1993–2019 (monthly)	0.083°	16.01	CMEMS	GLOBAL_REANALYSIS_PHY_001_030
Suspended matter concentration (g/m^3)	Monthly averages of mass concentration of suspended matter in sea water	1997–2021 (monthly)	4 km	0	CMEMS	OCEANCOLOUR_GLO_OPTICS_L4_REP_OBSERVATIONS_009_081
Sea surface salinity (1e^{-3})	Monthly averages of sea surface salinity	1993–2019 (monthly)	0.083°	15.99	CMEMS	GLOBAL_REANALYSIS_PHY_001_030
Sea surface temperature (°C)	Monthly average of sea surface temperature	1985–2021 (monthly)	5 km	3.14	NOAA CRW	ct5km_sst-mean_v3.1
Density of built-up (%)	Percentage of ground cover of built-up areas per pixel	2015–2019 (yearly)	100 m	0	CGLS	LandCover100m:collection_3:epoch2019
Density of cropland (%)	Percentage of ground cover of cropland per pixel	2015–2019 (yearly)	100 m	0	CGLS	LandCover100m:collection_3:epoch2019
Boat detection (%)	Percentage of boat detections per satellite overpass coverage	2017–2020 (yearly)	0.04°	0	EOG	VBD_npp_global-saa_pc_v23
Human population density (hab/km^2)	Human population density	2000–2020 (yearly - by 5 years)	0.04°	0	SEDAC	gpw-v4- rev11
Depth (m)	Depth	-	0.017°	0.03	GMRT	GMRTv3_9

TABLE 1 (Continued)

Name	Description	Time window (resolution)	Spatial res.	% NA	Source	Original product id
Surface land	Binary map, indicating whether pixel is on land or sea	-	0.017°	0.03	GMRT	GMRTv3_9
Data sources						
Source id	**Source name**			**Reference**		
CMEMS	Copernicus Marine Service			EU Copernicus Marine Service (2022)		
NOAA CRW	National Oceanic and Atmospheric Administration, Coral Reef Watch			NOAA Coral Reef Watch (2018); Skirving et al. (2020)		
CGLS	Copernicus Global Land Service			Copernicus Global Land Service (2022)		
EOG	Earth Observation Group – Payne Institute for Public Policy			Elvidge et al. (2015, 2018); Hsu et al. (2019)		
SEDAC	Socioeconomic Data and Applications Center, Columbia University			Center for International Earth Science Information Network (2022)		
GMRT	Global Multi-Resolution Topography Data Synthesis			Ryan et al. (2009)		

Note: For each dataset, the table shows the name and description of the variable of interest, the temporal window covered by the dataset (with the temporal resolution in parenthesis), the spatial resolution of the dataset, the % of missing values (% NA) across all the reef cells, the source repository of the dataset (with the abbreviations defined at the bottom) and the identifier of the original product from the source repository.

and across temporal windows generally covering the past 3–4 decades (see Table 1 for details). These raw monthly datasets are sourced from the repositories from the Coral Reef Watch of the National Oceanic and Atmospheric Administration (NOAA Coral Reef Watch, 2018) and the Copernicus Marine Service (EU Copernicus Marine Service, 2022).

Raw datasets at yearly temporal resolution were also raster stacks, but here each image was a yearly measure of a given variable worldwide. The datasets of this type included in RECIFS were: land cover of built-up and cropland areas (covering the 2015–2019 period, at 100-m resolution; Copernicus Global Land Service, 2022), boat detection (based on night-time light detection on sea; 2017–2020, at 500-m resolution; Elvidge et al., 2015, 2018; Hsu et al., 2019), and human population density (available for the years 2000, 2005, 2010, 2015, 2020, at 5-km resolution, Center for International Earth Science Information Network, 2022).

Two time-invariant raw datasets were included in RECIFS, both derived from the Global Multi-Resolution Topography Data Synthesis (v.4.1, accessed in January 2021), which is a single raster image describing contemporary topography (on land) and bathymetry (on sea) worldwide at ~1-km spatial resolution (Ryan et al., 2009). The first derived dataset is a depth map where the pixels represent the topography of below sea surface level, and the second is a binary map where pixels were either assigned to land (topography > 0) or to sea (topography < 0).

To process these raw environmental datasets, we extracted values from the raster stacks using custom R scripts (v. 4.1; R Core Team, 2022), featuring the raster (v. 3.5; Hijmans, 2021) and the rgdal (v. 1.5; Bivand et al., 2021) libraries. First, polygons representing coral reefs worldwide were retrieved from the Global Distribution of Coral Reefs dataset (v. 4.1; UNEP-WCMC et al., 2021), and were reported to 5 by 5 km cells. The resulting grid was composed of 61,038 reef cells. For each reef cell, we extracted the environmental values from each raw dataset using a buffer to average over different radiuses (2.5 km, referred to as 'on reef' extraction, and 10, 25 and 50 km). Of note, the extraction of environmental values was performed on the raw datasets at their native resolution (i.e., the one shown in Table 1), without any down- or up-scaling preprocessing. The result was a set of tables (called 'Extended tables'), with one table per environmental variable at a given buffer size. For each table, the rows represented the worldwide reef cells (such that there were 61,038 rows in each table) and columns represented the temporal dimension (months, years or a single column for time-invariant variables).

Finally, the 'Extended tables; were synthesized together into a single table (called the 'Summary table'). For the variables at a monthly resolution, each reef cell had overall temporal average and standard deviation 'on reef' (i.e., using the 2.5-km buffer) for each environmental variable. For the annual and non-temporal variables, each reef cell had overall averages measured using two of the buffer sizes: 10 and 50 km.

Environmental variables had an average proportion of missing values of 9% of the reef cells, but with substantial variation between variables (standard deviation of 9%; Table 1). Indeed,

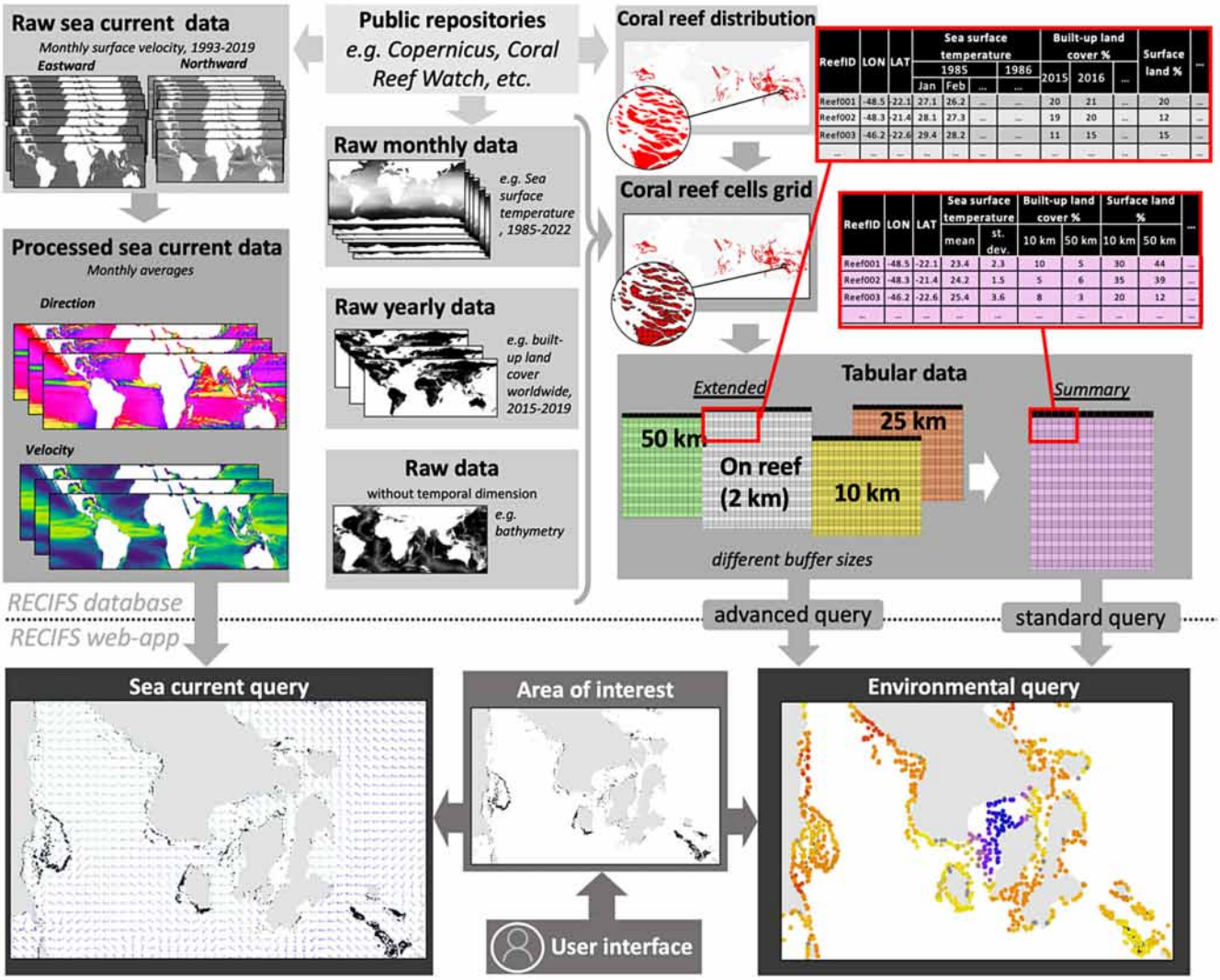


FIGURE 1 Data workflow. Raw environmental data used in Reef Environment Centralized Information System (RECIFS) come from publicly available repositories, and have distinct temporal resolutions (monthly, annual or time invariant). Shapes of coral reefs worldwide were summarized into a regular grid (reef cell size: 5 by 5 km), and raw environmental values were extracted for each reef cell. These extractions were performed using four different buffer sizes (2.5 km – on reef, 10, 25 or 50 km) around the centres of reef cells. The result is a set of tables ('Extended tables') describing variation across reef cells for each environmental variable, at different buffer sizes and through different time periods (months or years). These tables are synthesized into a 'Summary table', reporting overall temporal average and standard deviation of each environmental variable, for every reef cell. RECIFS also includes data on sea surface current velocity and direction, obtained from processed raw data describing eastward and northward surface water velocity. These processed data are accessible through the RECIFS web application: the user selects an area of interest through an interactive map and runs either an environmental query to visualize environmental variation (standard query to access the Summary table, advanced query to access the Extended tables), or a sea current query to visualize sea currents.

the highest missing proportions (~21% of the reef cells) were observed for variables with native spatial resolution of 25 km (pH, iron, nitrate, phosphate, and oxygen concentration), whereas land cover-derived variables had no missing values. It should be noted that missing values distributions did not vary over time (i.e., for a given environmental variable, the same reef cells were always missing over time). The proportion of missing values appeared to vary between oceanic basins worldwide (Supporting Information Figure S1a), and this variation seemed to reflect differences in the seascape configuration (Supporting Information Figure S1b). Indeed, missing proportions were higher for reef cells located next to surface land (e.g., fringing reefs next to Southeast Asia islands), than at reef cells far from land (e.g., outshore platform reefs from the Great Barrier Reef). This difference is probably due to the fact that pixels located in coastal waters typically have higher uncertainty in the environmental variable estimation via satellite imagery (Gilerson et al., 2022). Of note, when employing buffer radiuses above the size of the reef cell (i.e., 10, 25 and 50 km) to average environmental values, missing values were excluded from calculations.

2.2 | Data structure: Surface currents

The raw environmental dataset used to characterize sea currents was sourced from the Copernicus Marine Service (dataset id: GLOBAL_REANALYSIS_PHY_001_030; EU Copernicus Marine Service, 2022). This dataset is obtained from a stack of raster images displaying global monthly averages of eastward and northward sea surface velocity, from 1993 to 2019. Using custom R scripts featuring the raster and the circular (v. 0.4; Agostinelli & Lund, 2017) libraries, we combined eastward and northward velocities of each pixel into a vector of sea currents, and then computed (a) sea current direction (the angle of the vector) and (b) sea current velocity (the norm of the vector) for each monthly measure. Finally, we summarized trends of surface circulation by computing raster images representing the overall yearly average and 12 by-month averages for both sea current direction and velocities.

2.3 | Web application and queries

The web-based application of RECIFS is built on a NodeJS server (v. 10.13; OpenJS Foundation). On the front-end, the server features an Openlayers (v. 5.3; Open Source Geospatial Foundation) map interface displaying the reef cells, while the back-end of the server stores environmental data in a tabular format and the raster images summarizing sea surface currents.

When the front-end user defines an 'Area of Interest' through the interactive map, the server queries the RECIFS database through R scripts (featuring the raster and the sp libraries; v. 1.4; Bivand et al., 2013) processing, sub-setting and finally returning the data stored on the back-end. Three types of query are possible:

1. *Standard environmental query*: The user selects an 'Area of Interest', and the server returns precomputed environmental values from the 'Summary table' for the reef cells from this area.
2. *Advanced environmental query*: The user selects an 'Area of Interest', an environmental variable of interest, a spatial buffer size, a temporal window (years and/or months), and a statistic to summarize environmental variation over time (mean, standard deviation, median, minimum or maximum), and the server returns the corresponding values from the 'Extended tables' for the reef cells in the area.
3. *Sea current query*: The user selects an 'Area of Interest' and specifies a month of interest, and the server returns a visual display of arrows indicating the corresponding strength and direction of sea surface currents.

The results of the queries are displayed on the interactive map. The front-end interface features different functionalities allowing the user to customize map visualization, such as defining the colour scale used to display environmental data, setting the transparency of layers or modifying the background layer. The map rendered on the web interface can be downloaded either in a PDF (reef cell environmental values and sea surface current arrows) or a tabular format (reef cell environmental values only). Furthermore, the user can define 'Points of interest' on the map (e.g., sampling locations, reefs of interest) and download the environmental data for the closest reef cells to such points. All these functionalities are available in complete open access, without any registration required. An interactive tutorial is also available at https://recifs.epfl.ch/tutorial to walk first-time users through the different functionalities described above.

2.4 | Examples of use

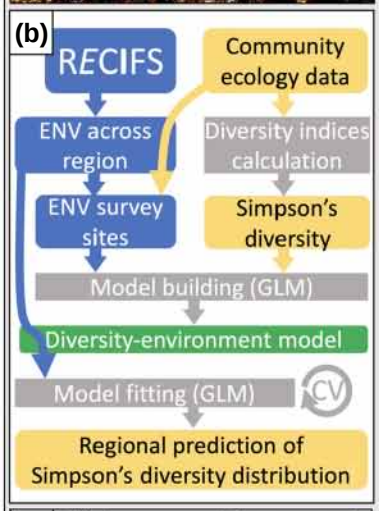
One common method to identify potential drivers of ecosystem dynamics is to investigate the association between remote sensed environmental data, and in-situ measurements replicated at multiple locations across reef systems. These in-situ measurements can include various ecological surveys, including surveys of coral abundance (e.g., Sully et al., 2022), coral diversity (e.g., Kusumoto et al., 2020), coral bleaching severity (e.g., McClanahan et al., 2020), fish biomass (e.g., Cinner et al., 2016), or the presence/absence of a given reef species (e.g., Förderer et al., 2018; Ottimofiore et al., 2017; Principe et al., 2021). In-situ measurements also include molecular data that represent genetic diversity in populations, which can be used in genotype–environment association (GEA) studies to uncover genetic variants potentially underpinning local adaptation processes (Fuller et al., 2020; Lundgren et al., 2013; Selmoni, Lecellier, Magalon, et al., 2021 ; Selmoni, Rochat, et al., 2020; Sherman et al., 2020; Thomas et al., 2017).

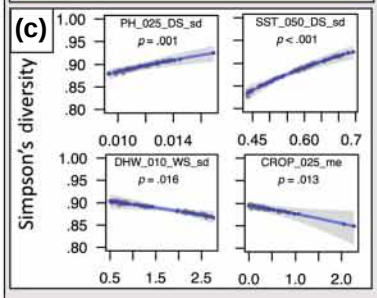
In Boxes 1 and 2 we show possible implementations of the RECIFS data to (a) characterize coral diversity in the Caribbean, and (b) study local adaption of a reef fish population along the north-western coast of Australia.

2.5 | Perspectives

In the years to come, the development of RECIFS will focus on three domains: data, functionalities and integration. The datasets already available in the repository will be updated on an annual basis to provide the most recent records for the different environment variables. Furthermore, new variables that describe environmental factors shaping the dynamics of coral reef species will be added to the database as these variables become available and/or are required; we encourage coral reef researchers and conservation actors to contact us (the authors) to propose environmental descriptors to be added to RECIFS. With the future improvements of remote sensed data and environmental modelling techniques, we also aim to increase the spatial and temporal resolution of the reef cells, so that the environmental characterization can be more pertinent to fine-scale dynamics of the reef ecosystem (Murray et al., 2018).

Concerning functionalities, the main future goal is to develop an objective quantification of reef connectivity calculated from sea current data. Such quantification could combine transition matrices,





Box 1. Coral taxonomic diversity in the Caribbean.

The first case study is based on photographic surveys performed across the reefs of the Caribbean (**Figure a**: a coral reef from the Caribbean) by the Catlin Seaview Survey (CSS) project (González-Rivero *et al.*, 2014, 2016). The CSS project used machine learning to annotate the presence of 17 hard coral taxa along 183 surveys performed in 12 territories of the region. We used these taxonomic annotations to calculate the Simpson's diversity index (SDI) for each survey (high values indicate high diversity of coral taxa).

Following the analytical pipeline shown in **Figure b**, we investigated the association between spatial variation of SDI and 302 environmental variables (ENV) obtained from RECIFS. The 302 variables characterized the reef environment of the region combining different environmental conditions, at varying spatial scale (different buffer sizes), different temporal resolution (wet season, dry season, and overall) and using different statistics to summarize environmental trends over time (mean and standard deviation).

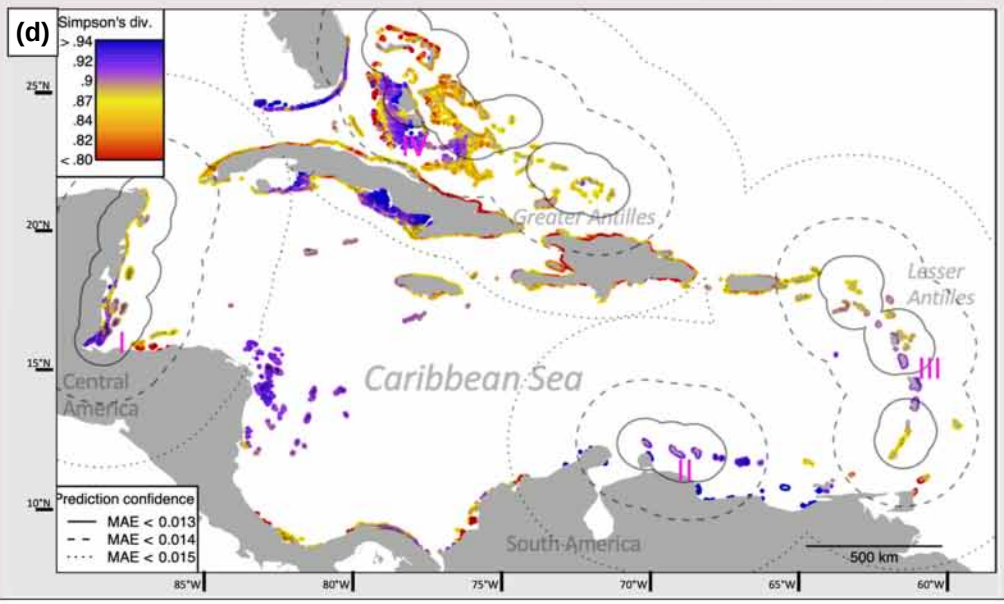
We built a stepwise generalized linear model (GLM) to find which of these 302 environmental variables were associated with spatial variation of the SDI. The resulting "Diversity-environment model" included 11 of environmental variables that significantly ($p<0.05$) explained variation of the SDI ($R^2=0.56$). These variables describe environmental conditions that are well known to mediate changes in coral abundance/mortality, such as thermal stress (standard deviation of Degree Heating Week during the wet season), ocean acidification (standard deviation of pH), or run-off from agricultural activities along the coastline (land use of cropland; Cornwall et al., 2021; D'Angelo & Wiedenmann, 2014; De'ath & Fabricius, 2010; R. Jones et al., 2020; McClanahan et al., 2020; Sully et al., 2019). In **Figure c**, we show the association between SDI and four of these variables: standard deviation of PH (PH_025_DS_sd), sea surface temperature (SST_050_DS_sd), degree heating week (DHW_010_WS_sd), and density of cropland (CROP_025_me).

Based on the associations described above, the RECIFS environmental data across the region were used to predict SDI for each reef of the Caribbean (**Figure d**). This spatial prediction highlighted regions expected to host reefs with high coral diversity, for instance in Belize (highlighted by the pink "I" on map); Bonaire, Aruba and Curaçao ("II"); Dominica and Martinique ("III"); and the Andros Island in Bahamas ("IV").

Of note, we ran a spatially explicit cross validation (CV; Pohjankukka et al., 2017) to assess the uncertainty around the SDI prediction. The mean absolute error (MAE) of the prediction was ~11% of the range of SDI measured in real surveys and did not appear to increase with distance from survey locations. Yet, care must be taken when interpreting this prediction as the survey data was systematically collected at 10 meters of depth, such that these diversity estimations might not be relevant at other depth levels.

Some of the biodiversity patterns highlighted here (*e.g.,* hotspots in Belize, the south of Cuba, northeast of Puerto Rico) mirror recent Caribbean-wide estimations of corals morpho-functional diversity (Melo-Merino *et al.*, 2022), and could indicate priority target for regional conservation initiatives.

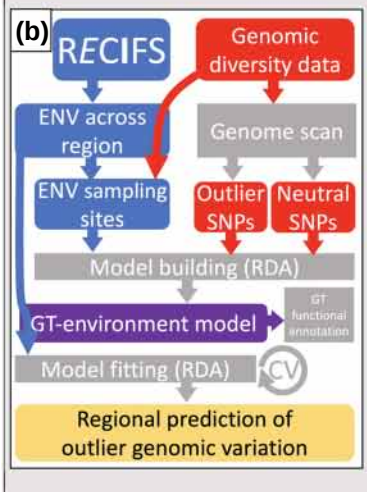
The extended methods and results of this case study are available as supplementary material.

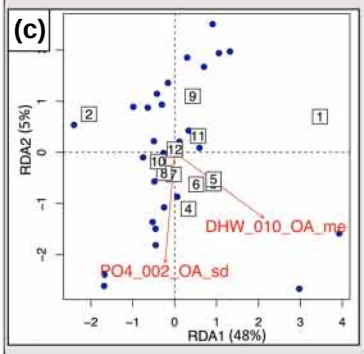


graph theory (van Etten, 2018), and oceanographic models to estimate the probability of transition from one reef to another. Overall patterns of regional transitions could then be synthesized in connectivity indices, summarizing how each reef of a given region is interconnected to those located downstream or upstream via oceanic currents (Matz et al., 2020; Selmoni, Rochat, Lecellier, et al., 2020).

Going forwards, it will also be important to integrate RECIFS with other projects and complementary datasets that characterize alternative facets of reef environments. Importantly, the cross-link between repositories should also be extended to various types of data (e.g., field surveys on species abundance/diversity, geo-referenced genomic data) that describe the reef ecosystem (Hedley et al., 2016;

Picture credits: Ian Shaw





Box 2. Stripey snapper adaptation in NW Australia.

For the second case study we used pre-existing population genomics data from 1,016 stripey snapper individuals (*Lutjanus carponotatus*; shown in **Figure a**) from 51 reefs along the northwestern coast of Australia (DiBattista *et al.*, 2017). Each sample in this dataset was genotyped at 17,007 single nucleotide polymorphisms (SNPs). Using genome scans, we identified 62 outlier SNPs that did not follow the neutral genetic structure of the population (*i.e.,* SNPs potentially involved in local adaptation). We then ran a principal component analysis (PCA) to summarize the allelic frequencies of outlier SNPs into 12 axes of outlier genomic variation (outlier-PCs).

Using the analytical pipeline shown in **Figure b**, we investigated the association between the spatial distribution of outlier-PCs and 302 environmental variables (ENV) from RECIFS. The 302 variables characterized the reef environment of the region combining different environmental conditions, at varying spatial (different buffer sizes), different temporal resolution (wet season, dry season, and overall) and using different statistics to summarize environmental trends over time (mean and standard deviation).

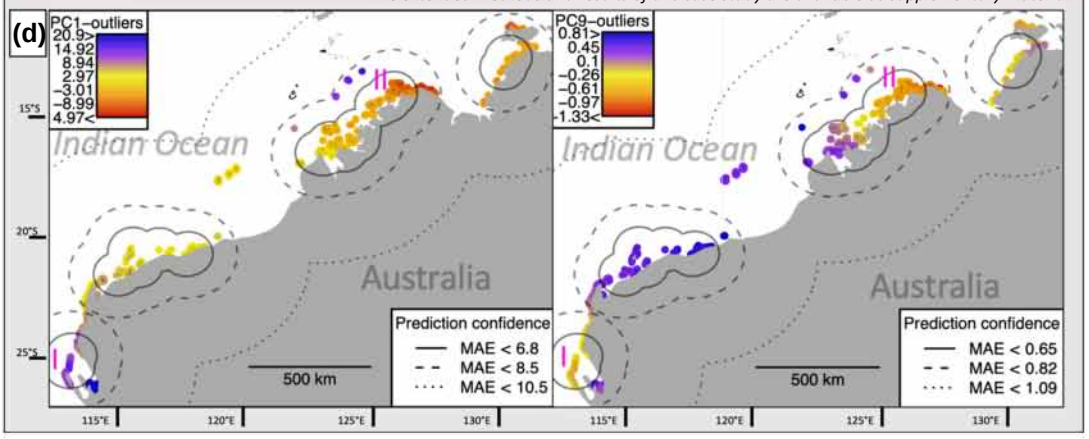
We ran a redundancy analysis (RDA) to uncover which of the 302 environmental variables were associated with axes of outlier genomic variation ("GT-environment model"). In **Figure c**, the tri-plot displays the redundancy between variation of outlier-PCs (white boxes), sampling sites (blue dotes) and environmental conditions (red arrows). We found that more than half of the outlier-PCs variation (R^2=0.52) could be explained by two environmental variables. The first variable was average Degree Heating Week (DHW_010_OA me), which showed the strongest correlation with the first axis of genomic variation (outlier-PC1). The second environmental variable was the standard deviation of Phosphate Concentration (PO4_002_OA_sd), which was correlated with outlier-PC1 and outlier-PC9.

It should be noted that RDA is an exploratory analysis. As such, RDA only uncovers correlations between axes of genetic variation and environmental gradients, without necessarily implying a causal (*i.e.,* adaptive) relationship. One way to find support for the putative adaptive role of an axis of genomic variation is to investigate the functional genomic annotations ("GT functional annotation"). For instance, we found that outlier-PC1 summarized the variation of SNPs located in genes potentially involved in the "detection of temperature stimulus", which corroborates the hypothesis of an adaptive role against heat stress.

Finally, we employed the association models from the RDA to predict the spatial distribution of the genetic axes potentially involved in local adaptation (**Figure d**: left side: outlier-PC1; right side: outlier-PC9). High values of outlier-PC1, indicating a higher frequency of alleles putatively implicated in heat adaptation, were observed in the Shark Bay area (highlighted by the purple "I" on maps). As for outlier-PC9, we assumed that low values corresponded to an adaptation to high variability in Phosphate concentration and observed reef cells with such values in the Shark Bay area and the Kimberly region ("II" on map).

We ran a spatially explicit cross validation (CV; Pohjankukka et al., 2017) to assess the uncertainty around the genetic variation prediction. The cross-validation suggests that predictions must interpreted with care, as the mean absolute error (MAE) was non-negligible (~36% and 14% of the range of outlier-PC1 and outlier-PC9, respectively) and increased with distance from sampling locations. Yet, these results suggest that reefs of the Shark Bay area might host a stripey snapper population with exceptional adaptive capacities to both thermal stress and variability in phosphate concentration, compared with the reefs in the northeastern regions. In the original work that produced the genomic dataset used in this case study, DiBattista and colleagues highlighted a connectivity break between the reefs in Shark Bay and those located in northern areas (DiBattista *et al.*, 2017). For these reasons, stripey snapper in the Shark Bay area might be of particular interest for conservation efforts.

The extended methods and results of this case study are available as supplementary material.



Lecours et al., 2022). A hub centralizing different data sources on the state of the reef ecosystem will be of paramount importance in the future, as it will facilitate the establishment of links between all the different information layers. In the long term, the goal of RECIFS is to serve as the foundation from which to build such a hub.

ACKNOWLEDGMENTS

We thank the environmental data providers for sharing the datasets included in RECIFS. We thank the Catlin Seaview Project and Joseph DiBattista for sharing the data used for the case studies. We thank Annie Guillaume and the anonymous referees for the comments and

suggestions provided during the writing of this paper. Open access funding provided by Ecole Polytechnique Federale de Lausanne.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The R scripts used to process raw environmental data and the R scripts of the analyses from the two case studies (Boxes 1 and 2) are available on Dryad (https://doi.org/10.5061/dryad.12jm63z2f). Codes from the RECIFS web application are publicly available on GitHub (https://github.com/Oselmoni/RECIFS).

ORCID

Oliver Selmoni https://orcid.org/0000-0003-0904-5486

REFERENCES

Agostinelli, C., & Lund, U. (2017). R package circular: Circular statistics.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., & de Clerck, O. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, *27*, 277–284.

Bivand, R., Keitt, T., & Rowlingson, B. (2021). rgdal: Bindings for the geospatial data abstraction library.

Bivand, R., Pebesma, E., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). Springer.

Bouchet, P. (2006). The magnitude of marine biodiversity. In C. Duarte (Ed.), *The exploration of marine biodiversity: Scientific and technological challenges* (pp. 31–64). Fundación BBVA.

Center for International Earth Science Information Network. (2022). Gridded population of the world, version 4.11 (GPWv4): Population density, revision 11.

Cinner, J. E., Huchery, C., MacNeil, M. A., Graham, N. A. J., McClanahan, T. R., Maina, J., Maire, E., Kittinger, J. N., Hicks, C. C., Mora, C., Allison, E. H., D'Agata, S., Hoey, A., Feary, D. A., Crowder, L., Williams, I. D., Kulbicki, M., Vigliola, L., Wantiez, L., … Mouillot, D. (2016). Bright spots among the world's coral reefs. *Nature*, *535*, 416–419.

Copernicus Global Land Service. (2022). Land Cover 100m: version 3 Globe 2015–2019.

Cornwall, C. E., Comeau, S., Kornder, N. A., Perry, C. T., van Hooidonk, R., DeCarlo, T. M., Pratchett, M. S., Anderson, K. D., Browne, N., Carpenter, R., Diaz-Pulido, G., D'Olivo, J. P., Doo, S. S., Figueiredo, J., Fortunato, S. A. V., Kennedy, E., Lantz, C. A., McCulloch, M. T., González-Rivero, M., … Lowe, R. J. (2021). Global declines in coral reef calcium carbonate production under ocean acidification and warming. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, e2015265118.

D'Angelo, C., & Wiedenmann, J. (2014). Impacts of nutrient enrichment on coral reefs: New perspectives and implications for coastal management and reef survival. *Current Opinion in Environmental Sustainability*, *7*, 82–93.

De'ath, G., & Fabricius, K. (2010). Water quality as a regional driver of coral biodiversity and macroalgae on the Great Barrier Reef. *Ecological Applications*, *20*, 840–850.

DiBattista, J. D., Travers, M. J., Moore, G. I., Evans, R. D., Newman, S. J., Feng, M., Moyle, S. D., Gorton, R. J., Saunders, T., & Berry, O. (2017). Seascape genomics reveals fine-scale patterns of dispersal for a reef fish along the ecologically divergent coast of Northwestern Australia. *Molecular Ecology*, *26*, 6206–6223.

Elvidge, C. D., Ghosh, T., Baugh, K., Zhizhin, M., Hsu, F. C., Katada, N. S., Penalosa, W., & Hung, B. Q. (2018). Rating the effectiveness of fishery closures with visible infrared imaging radiometer suite boat detection data. *Frontiers in Marine Science*, *5*, 132.

Elvidge, C. D., Zhizhin, M., Baugh, K., & Hsu, F. C. (2015). Automatic boat identification system for VIIRS low light imaging data. *Remote Sensing*, *7*, 3020–3036.

EU Copernicus Marine Service. (2022). Global ocean—In-situ-near-real-time observations.

Fernandez, M., Yesson, C., Gannier, A., Miller, P. I., & Azevedo, J. M. N. (2017). The importance of temporal resolution for niche modelling in dynamic marine environments. *Journal of Biogeography*, *44*, 2816–2827.

Foo, S. A., & Asner, G. P. (2019). Scaling up coral reef restoration using remote sensing technology. *Frontiers in Marine Science*, *6*, 79.

Foo, S. A., & Asner, G. P. (2021). Impacts of remotely sensed environmental drivers on coral outplant survival. *Restoration Ecology*, *29*, e13309.

Förderer, M., Rödder, D., & Langer, M. R. (2018). Patterns of species richness and the center of diversity in modern Indo-Pacific larger foraminifera. *Scientific Reports*, *8*, 1–9.

Fuller, Z. L., Mocellin, V. J. L., Morris, L. A., Cantin, N., Shepherd, J., Sarre, L., Peng, J., Liao, Y., Pickrell, J., Andolfatto, P., Matz, M., Bay, L. K., & Przeworski, M. (2020). Population genetics of the coroal Acropora millepora: Toward genomic prediction of bleaching. *Science*, *369*, eaba4674.

Gilerson, A., Herrera-Estrella, E., Foster, R., Agagliate, J., Hu, C., Ibrahim, A., & Franz, B. (2022). Determining the primary sources of uncertainty in retrieval of marine remote sensing reflectance from satellite ocean color sensors. *Frontiers in Remote Sensing*, *3*, 25.

González-Rivero, M., Beijbom, O., Rodriguez-Ramirez, A., Holtrop, T., González-Marrero, Y., Ganase, A., Roelfsema, C., Phinn, S., & Hoegh-Guldberg, O. (2016). Scaling up ecological measurements of coral reefs using semi-automated field image collection and analysis. *Remote Sensing*, *8*, 30.

González-Rivero, M., Bongaerts, P., Beijbom, O., Pizarro, O., Friedman, A., Rodriguez-Ramirez, A., Upcroft, B., Laffoley, D., Kline, D., Bailhache, C., Vevers, R., & Hoegh-Guldberg, O. (2014). The Catlin Seaview Survey—Kilometre-scale seascape assessment, and monitoring of coral reef ecosystems. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *24*, 184–198.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., Mcdonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., … Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*, 1424–1435.

Hedley, J. D., Roelfsema, C. M., Chollett, I., Harborne, A. R., Heron, S. F., Weeks, S. J., Skirving, W. J., Strong, A. E., Mark Eakin, C., Christensen, T. R. L., Ticzon, V., Bejarano, S., & Mumby, P. J. (2016). Remote sensing of coral reefs for monitoring and management: A review. *Remote Sensing*, *8*, 118.

Hijmans, R. J. (2021). raster: Geographic Data Analysis and Modeling.

Hsu, F. C., Elvidge, C. D., Baugh, K., Zhizhin, M., Ghosh, T., Kroodsma, D., Susanto, A., Budy, W., Riyanto, M., Nurzeha, R., & Sudarja, Y. (2019). Cross-matching VIIRS boat detections with vessel monitoring system tracks in Indonesia. *Remote Sensing*, *11*, 995.

Hughes, T. P., Barnes, M. L., Bellwood, D. R., Cinner, J. E., Cumming, G. S., Jackson, J. B. C., Kleypas, J., van de Leemput, I. A., Lough, J. M., Morrison, T. H., Palumbi, S. R., van Nes, E. H., & Scheffer, M. (2017). Coral reefs in the Anthropocene. *Nature*, *546*, 82–90.

Jones, G. P., McCormick, M. I., Srinivasan, M., & Eagle, J. V. (2004). Coral decline threatens fish biodiversity in marine reserves. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8251–8253.

Jones, R., Giofre, N., Luter, H. M., Neoh, T. L., Fisher, R., & Duckworth, A. (2020). Responses of corals to chronic turbidity. *Scientific Reports*, *10*, 1–13.

Knowlton, N., Brainard, R., Fisher, R., Moews, M., Plaisance, L., & Caley, J. (2010). Coral reef biodiversity. In A. McIntyre (Ed.), *Life in the World's ocean: Diversity, distribution and abundance* (pp. 65–78). Blackwell Publishing Ltd.

Kusumoto, B., Costello, M. J., Kubota, Y., Shiono, T., Wei, C. L., Yasuhara, M., & Chao, A. (2020). Global distribution of coral diversity: Biodiversity knowledge gradients related to spatial resolution. *Ecological Research*, *35*, 315–326.

Lecours, V., Disney, M., He, K., Pettorelli, N., Rowcliffe, J. M., Sankey, T., & Scales, K. (2022). Remote sensing and the UN Ocean Decade: High expectations, big opportunities. *Remote Sensing in Ecology and Conservation*, *8*, 267–271.

Leempoel, K., Duruz, S., Rochat, E., Widmer, I., Orozco-terWengel, P., & Joost, S. (2017). Simple rules for an efficient use of geographic information systems in molecular ecology. *Frontiers in Ecology and Evolution*, *5*, 33.

Lundgren, P., Vera, J. C., Peplow, L., Manel, S., & van Oppen, M. J. H. (2013). Genotype—Environment correlations in corals from the Great Barrier Reef. *BMC Genetics*, *14*, 9.

MacNeil, M. A., Mellin, C., Matthews, S., Wolff, N. H., McClanahan, T. R., Devlin, M., Drovandi, C., Mengersen, K., & Graham, N. A. J. (2019). Water quality mediates resilience on the Great Barrier Reef. *Nature Ecology & Evolution*, *3*, 620–627.

Mannocci, L., Boustany, A. M., Roberts, J. J., Palacios, D. M., Dunn, D. C., Halpin, P. N., Viehman, S., Moxley, J., Cleary, J., Bailey, H., Bograd, S. J., Becker, E. A., Gardner, B., Hartog, J. R., Hazen, E. L., Ferguson, M. C., Forney, K. A., Kinlan, B. P., Oliver, M. J., … Winship, A. J. (2017). Temporal resolutions in species distribution models of highly mobile marine animals: Recommendations for ecologists and managers. *Diversity and Distributions*, *23*, 1098–1109.

Matz, M. V., Treml, E. A., & Haller, B. C. (2020). Estimating the potential for coral adaptation to global warming across the Indo-West Pacific. *Global Change Biology*, *26*, 3473–3481.

McClanahan, T. R., Darling, E., Maina, J., Muthiga, N., D'agata, S., Leblond, J., Arthur, R., Jupiter, S., Wilson, S., Mangubhai, S., Ussi, A., Guillaume, M., Humphries, A., Patankar, V., Shedrawi, G., Pagu, J., & Grimsditch, G. (2020). Highly variable taxa-specific coral bleaching responses to thermal stresses. *Marine Ecology Progress Series*, *648*, 135–151.

Melo-Merino, S. M., Lira-Noriega, A., González-Barrios, F. J., Reyes-Bonilla, H., & Álvarez-Filip, L. (2022). Functional divergence from ecological baselines on Caribbean coral reefs. *Ecography*, *2022*, e05811.

Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, *415*, 108837.

Murray, N. J., Keith, D. A., Bland, L. M., Ferrari, R., Lyons, M. B., Lucas, R., Pettorelli, N., & Nicholson, E. (2018). The role of satellite remote sensing in structured ecosystem risk assessments. *Science of the Total Environment*, *619–620*, 249–257.

NOAA Coral Reef Watch. (2018), updated daily. *NOAA Coral Reef Watch Version 3.1 Daily Global 5km Satellite Coral Bleaching Degree Heating Week Product*, Jun. 3, 2013-Jun. 2, 2014. College Park, Maryland, USA: NOAA Coral Reef Watch. https://ftp.star.nesdis.noaa.gov/pub/sod/mecb/crw/data/5km/v3.1/nc/v1.0/daily/dhw/

Ottimofiore, E., Albouy, C., Leprieur, F., Descombes, P., Kulbicki, M., Mouillot, D., Parravicini, V., & Pellissier, L. (2017). Responses of coral reef fishes to past climate changes are related to life-history traits. *Ecology and Evolution*, *7*, 1996–2005.

Pohjankukka, J., Pahikkala, T., Nevalainen, P., & Heikkonen, J. (2017). Estimating the prediction performance of spatial models via spatial k-fold cross validation. *International Journal of Geographical Information Science*, *31*, 2001–2019.

Principe, S. C., Acosta, A. L., Andrade, J. E., & Lotufo, T. M. C. (2021). Predicted shifts in the distributions of Atlantic reef-building corals in the face of climate change. *Frontiers in Marine Science*, *8*, 912.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reaka-Kudla, M. (1997). The global biodiversity of coral reefs: A comparison with rain forests. In M. Reaka-Kudla, D. Wilson, & E. Wilson (Eds.), *Biodiversity II: Understanding and protecting our biological resources* (pp. 83–108). Joseph Henry Press.

Ryan, W. B. F., Carbotte, S. M., Coplan, J. O., O'Hara, S., Melkonian, A., Arko, R., Weissel, R. A., Ferrini, V., Goodwillie, A., Nitsche, F., Bonczkowski, J., & Zemsky, R. (2009). Global multi-resolution topography synthesis. *Geochemistry, Geophysics, Geosystems*, *10*, 1–9.

Sbrocco, E. J., & Barber, P. H. (2013). MARSPEC: Ocean climate layers for marine spatial ecology. *Ecology*, *94*, 979.

Selmoni, O., Lecellier, G., Ainley, L., Collin, A., Doucet, R., Dubousquet, V., Feremaito, H., Ito Waia, E., Kininmonth, S., Magalon, H., Malimali, S., Maugateau, A., Meibom, A., Mosese, S., René-Trouillefou, M., Satoh, N., van Oppen, M. J. H., Xozamé, A., Yékawene, M., … Berteaux-Lecellier, V. (2020). Using modern conservation tools for innovative Management of Coral Reefs: The MANACO consortium. *Frontiers in Marine Science*, *7*, 609.

Selmoni, O., Lecellier, G., Magalon, H., Vigliola, L., Oury, N., Benzoni, F., Peignon, C., Joost, S., & Berteaux-Lecellier, V. (2021). Seascape genomics reveals candidate molecular targets of heat stress adaptation in three coral species. *Molecular Ecology*, *30*(8), 1892–1906. https://doi.org/10.1111/mec.15857

Selmoni, O., Rochat, E., Lecellier, G., Berteaux-Lecellier, V., & Joost, S. (2020). Seascape genomics as a new tool to empower coral reef conservation strategies: An example on North-Western Pacific Acropora digitifera. *Evolutionary Applications*, *13*, 1923–1938.

Sherman, K. D., Paris, J. R., King, R. A., Moore, K. A., Dahlgren, C. P., Knowles, L. C., Stump, K., Tyler, C. R., & Stevens, J. R. (2020). RAD-Seq analysis and in situ monitoring of Nassau grouper reveal fine-scale population structure and origins of aggregating fish. *Frontiers in Marine Science*, *7*, 157.

Skirving, W., Marsh, B., De La Cour, J., Liu, G., Harris, A., Maturi, E., Geiger, E., & Eakin, C. M. (2020). CoralTemp and the Coral Reef Watch Coral Bleaching Heat Stress Product Suite Version 3.1. *Remote Sensing*, *12*(23), 3856. https://doi.org/10.3390/rs12233856

Souter, D., Planes, S., Wicquart, J., Logan, M., Obura, D., & Staub, F. (2021). Status of Coral Reefs of the World: 2020, Australian Government, United Nation Environment Program (UNEP).

Sully, S., Burkepile, D. E., Donovan, M. K., Hodgson, G., & van Woesik, R. (2019). A global analysis of coral bleaching over the past two decades. *Nature Communications*, *10*, 1–5.

Sully, S., Hodgson, G., & van Woesik, R. (2022). Present and future bright and dark spots for coral reefs through climate change. *Global Change Biology*, *28*, 4509–4522.

Thomas, L., Kennington, W. J., Evans, R. D., Kendrick, G. A., & Stat, M. (2017). Restricted gene flow and local adaptation highlight the vulnerability of high-latitude reefs to rapid environmental change. *Global Change Biology*, *23*, 2197–2205.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & de Clerck, O. (2012). Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, *21*, 272–281.

UNEP-WCMC, WorldFish Centre, World Resources Institute & The Nature Conservancy. (2021). Global distribution of warm-water coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project.

van Etten, J. (2018). gdistance: Distances and routes on geographical grids.

Wenger, A. S., Fabricius, K., Jones, G. P., & Brodie, J. E. (2015). Effects of sedimentation, eutrophication, and chemical pollution on coral reef fishes. In C. Mora (Ed.), *Ecology of fishes and coral reefs* (pp. 145–153). Cambridge University Press.

BIOSKETCH

The research teams are the Laboratory of Geographic Information System (LASIG) at EPFL (Switzerland) and the UMR-250/9220 ENTROPIE at the Centre National de la Recherche Scientifique (CNRS) and the Institute of Research for Development (IRD; France). The two teams have been collaborating since 2017 to investigate coral adaptation to environmental changes combining remote sensing and genomics analyses. In 2019, the two teams founded ManaCo, an international network aiming to facilitate the collaboration between coral researchers and stakeholders of reef conservation.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Selmoni, O., Lecellier, G., Berteaux-Lecellier, V., & Joost, S. (2023). The Reef Environment Centralized InFormation System (RECIFS): An integrated geo-environmental database for coral reef research and conservation. *Global Ecology and Biogeography*, *32*, 622–632. https://doi.org/10.1111/geb.13657