The Reef Check Med Dataset on Key Mediterranean Marine Species 2001–2020

Eva Turicchia[1,2,3,4,5†‡], Massimo Ponti[2,3,4,5†‡], Gianfranco Rossi[4†] and Carlo Cerrano[4,5,6,7,8†]*

[1] *Dipartimento di Beni Culturali, Università di Bologna, Ravenna, Italy,* [2] *Centro Interdipartimentale di Ricerca Industriale Fonti Rinnovabili, Ambiente, Mare ed Energia, Università di Bologna, Ravenna, Italy,* [3] *Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, Ravenna, Italy,* [4] *Reef Check Italia Onlus, Ancona, Italy,* [5] *Consorzio Nazionale Interuniversitario per le Scienze del Mare (CoNISMa), Rome, Italy,* [6] *Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona, Italy,* [7] *Stazione Zoologica Anton Dohrn, Naples, Italy,* [8] *Fano Marine Center, Fano, Italy*

BACKGROUND

Mediterranean marine coastal habitats have been and continue to be threatened by human-related pressures, such as resource over-exploitation, pollution, habitat loss and fragmentation, and the invasion of non-native species (Airoldi and Beck, 2007; Micheli et al., 2013). These pressures are exacerbated by disturbances associated with global climate change, which have led to major shifts in marine ecosystems, impacting their resilience and ability to provide goods and services (Ponti et al., 2014; Garrabou et al., 2019). Ecological shifts in marine benthic communities are difficult to recognize because of the scarcity of: findable, accessible, interoperable, and reusable quantitative data (FAIR data principles; Wilkinson et al., 2016), which could serve as a baseline. The situation is impaired further by a lack of long-term monitoring capability at a regional scale. To be successful, marine coastal habitat conservation requires ecosystem-based management approaches that give ample consideration to the spatial and temporal distribution of key species over broad scales (Foley et al., 2010). It is evident that, easily accessible, reliable, and accurate data are essential to successfully monitor marine ecosystem health providing the knowledge needed to address the threats to coastal marine habitats, develop policies and regulations to protect vulnerable areas, understand trends, and forecast future changes (Martín Míguez et al., 2019). However, data obtained from scientific investigations and institutional monitoring programs, albeit very accurate, are generally too scarce and fragmented to be used effectively for spatial planning (Hochachka et al., 2012). This is particularly true for subtidal marine environments, as making sufficient repeated observations and measurements requires a large effort. As a solution, volunteers—citizen scientists—trained in the use of specifically developed monitoring protocols can help fill the gap in high-quality data acquisition, by performing monitoring over broad spatial and temporal scales.

Since 2001, volunteer certified trained snorkelers, freedivers, and scuba divers (hereafter EcoDivers) have collected data for selected key marine species, recording their occurrence, distribution, abundance, and bathymetric range along the Mediterranean Sea coasts, using the Reef Check Mediterranean Underwater Coastal Environment Monitoring (RCMed U-CEM) protocol (Turicchia et al., 2021b). Here, we describe the resulting dataset, the "Reef Check Med dataset on key Mediterranean marine species 2001–2020" (RCMed_2001–2020; Ponti et al., 2021), which is hosted by the European Marine Observation and Data Network (EMODnet; Martín Míguez et al., 2019) open repository. The organization and consistency of the data, the standards adopted, and how they can be accessed and used are also reported. The dataset is maintained by the non-profit organization Reef Check

Italia (RCI), collaborating with the other European Reef Check organizations, members of the worldwide Reef Check Foundation, and within the Reef Check Mediterranean Sea network[1].

METHODS

Abundance data for target species were collected according to the RCMed U-CEM protocol developed by RCI for a Citizen Science (CS) initiative that aims to monitor the ecological status of the Mediterranean marine coastal habitats. For this protocol, 43 taxa were selected based on two or more criteria, including ease of identification, being included in the international lists of protected species, being sensitive to human impacts, and being key indicators of the shift that Mediterranean coastal habitats can undergo under local pressures and climate change. Morphologically and ecologically similar species have been included at the genus or higher taxa level (Cerrano et al., 2017). Before going diving or snorkeling, each trained EcoDiver chooses one or more taxa, among the 43 included in the protocol (**Table 1**), to actively search for, according to the type of habitat typology, survey depth, and personal interests. EcoDivers make independent observations along random swims (as defined in Hill and Wilkinson, 2004) and upload their records to the online database using the specific smartphone app or the Internet form. Not encountered but actively searched taxa are reported as absent. No data is provided for not searched taxa. New data are made publicly available following quality assurance and control (QA/QC) procedures. Data that do not meet the standards of the QA/QC procedures are discarded. The detailed monitoring protocol and methodology used to collect and record the data, including species selection, participant training and QA/QC procedures, is described in Turicchia et al. (2021b). EcoDiver personal data are managed in accordance with the European general data protection regulation (GDPR), which allows sharing the collected data on their behalf but leaves each one responsible for the quality of the data they provided. No ethical approval was obtained regarding plants and animals because the protocol does not provide for collecting or manipulating organisms, but only visual observations into the wild.

DATASET STRUCTURE

The RCMed_2001–2020 dataset is fully compliant with the EMODnet biology standards (Martín Míguez et al., 2019). The taxonomic guideline used is based on the World Register of Marine Species (WoRMS; Vandepitte et al., 2015), the authoritative and comprehensive global list of marine organisms' names. Biotic and abiotic measurements are reported using the controlled thesaurus from the Natural Environment Research Council (NERC; http://vocab.nerc.ac.uk) Vocabulary Server maintained by the British Oceanographic Data Center (BODC), and the Darwin Core Archive (DwC-A), an internationally recognized biodiversity informatics standardized data system intended to facilitate information sharing on biological diversity.

This ensures interoperability and maximizes reusability, by providing a core standard (Wieczorek et al., 2012).

Following the EMODnet biology standards, data are organized in three tables: the DwC Event Core table stores information on the survey events, the DwC Occurrence extension table stores occurrence details, while the DwC Measurement or Facts extension (eMoF) table contains quantitative and additional information collected during survey events and species occurrences. The metadata records are based on ISO19115 standards.

DwC Event Core Table

Individual survey events correspond to single dives or swims, carried out independently by single EcoDivers inspecting the seabed at a specific time and place to collect data on single or multiple species. Several EcoDivers can investigate the same place simultaneously, but each provides an independent survey event. Each survey has a unique ID (*eventID*, including a progressive number, automatically attributed at the time of data entry) and is characterized by: the survey date (*eventDate*, in the format YYYY-MM-DD, conforms to the ISO 8601 1:2019); geographical coordinates in decimal degrees (*decimalLatitude*, *decimalLongitude*) based consistently on the same geodetic datum (*geodeticDatum* = WGS84; i.e., EPSG:4326); and with an accuracy (*coordinateUncertaintyInMeters*) of 200 m, as provided by the adopted protocol. Minimum (*minimumDepthInMeters*) and maximum (*maximumDepthInMeters*) depths represent the bathymetrical range of the survey and are expressed in meters. The *verbatimLocality* field contains textual information on the survey site (i.e., the site's local name and municipality).

The prevailing *habitat* surveyed is identified according to the following categories (when available, the corresponding European Nature Information System marine habitat classification, EUNIS v2019[2], is shown in parentheses):

- Coastal rocks (MA153, MA154, MA255, MB151, MB251, MC151, MC251, MD15, MD25)
- Offshore rocks (MC151, MC251, MD15, MD25)
- Rocky cliff (MC151, MC251, MD15, MD25)
- Posidonia (MB252)
- Posidonia and sand (MB252)
- Posidonia and rock
- Cave (MC152)
- Metal wreck
- Sand (MA452, MB254, MB551, MB552, MB553, MC35, MC55, MD35)
- Mud (MB651, MC451, MC65)
- Breakwaters and ports
- River mouth (MA353, MA553)
- Coastal lagoon (MB152, MB253, MB554)

All records also report the codified institution name providing data (*institutionCode* = RCI), the name of the dataset (*datasetName* = Reef Check Med - key Mediterranean marine

---

[2]Permalink to this version https://www.eea.europa.eu/ds_resolveuid/6d0484fd0078483ca73bec230574b34e.

TABLE 1 | List of taxa recorded in the dataset and corresponding Life Sciences Identifier (LSID).

scientificName	identificationQualifier	scientificNameID
Aplidium conicum		urn:lsid:marinespecies.org:taxname:103641
Aplidium tabarquensis		urn:lsid:marinespecies.org:taxname:103667
Aplysina	spp.	urn:lsid:marinespecies.org:taxname:131975
Arca noae		urn:lsid:marinespecies.org:taxname:138788
Astroides calycularis		urn:lsid:marinespecies.org:taxname:135178
Axinella	spp.	urn:lsid:marinespecies.org:taxname:131774
Balanophyllia europaea		urn:lsid:marinespecies.org:taxname:1451203
Caulerpa cylindracea		urn:lsid:marinespecies.org:taxname:660621
Caulerpa taxifolia		urn:lsid:marinespecies.org:taxname:144476
Centrostephanus longispinus		urn:lsid:marinespecies.org:taxname:124331
Chromis chromis		urn:lsid:marinespecies.org:taxname:127000
Cladocora caespitosa		urn:lsid:marinespecies.org:taxname:135146
Conger conger		urn:lsid:marinespecies.org:taxname:126285
Corallium rubrum		urn:lsid:marinespecies.org:taxname:125416
Diplodus	spp.	urn:lsid:marinespecies.org:taxname:126076
Epizoanthus	spp.	urn:lsid:marinespecies.org:taxname:100790
Eunicella cavolini		urn:lsid:marinespecies.org:taxname:125361
Eunicella singularis		urn:lsid:marinespecies.org:taxname:125365
Eunicella verrucosa		urn:lsid:marinespecies.org:taxname:125366
Geodia cydonium		urn:lsid:marinespecies.org:taxname:134025
Hippocampus	spp.	urn:lsid:marinespecies.org:taxname:126224
Homarus gammarus		urn:lsid:marinespecies.org:taxname:107253
Keratosa		urn:lsid:marinespecies.org:taxname:366651
Leptopsammia pruvoti		urn:lsid:marinespecies.org:taxname:135193
Maasella edwardsii		urn:lsid:marinespecies.org:taxname:125420
Ophidiaster ophidianus		urn:lsid:marinespecies.org:taxname:124101
Palinurus elephas		urn:lsid:marinespecies.org:taxname:107703
Paracentrotus lividus		urn:lsid:marinespecies.org:taxname:124316
Paramuricea clavata		urn:lsid:marinespecies.org:taxname:125387
Parazoanthus axinellae		urn:lsid:marinespecies.org:taxname:101055
Patella ferruginea		urn:lsid:marinespecies.org:taxname:140679
Pecten jacobaeus		urn:lsid:marinespecies.org:taxname:394429
Pectinidae		urn:lsid:marinespecies.org:taxname:213
Pinna nobilis		urn:lsid:marinespecies.org:taxname:140780
Polycitor adriaticus		urn:lsid:marinespecies.org:taxname:103625
Pyuridae		urn:lsid:marinespecies.org:taxname:103449
Rapana venosa		urn:lsid:marinespecies.org:taxname:140416
Savalia savaglia		urn:lsid:marinespecies.org:taxname:383014
Sciaena umbra		urn:lsid:marinespecies.org:taxname:127010
Scyllarides latus		urn:lsid:marinespecies.org:taxname:107708
Stolonifera		urn:lsid:marinespecies.org:taxname:1368
Tethya	spp.	urn:lsid:marinespecies.org:taxname:132077
Trisopterus minutus		urn:lsid:marinespecies.org:taxname:126446

species 2001–2020), and the protocol name (*samplingProtocol* = RCMed U-CEM protocol).

DwC Occurrence Extension Table

The DwC Occurrence extension table stores details on species occurrence linked to the individual survey events (*eventID*). Each record has a unique numeric identifier (*occurenceID*), attributed in post-processing after the QA/QC procedures, and is related to a single taxon that was searched for during the survey. Taxa are identified by their scientific name at the lowest possible taxonomic level (*scientificName*), with the indication of multiple species (spp.) belonging to the same genus when appropriate (*identificationQualifier*), and the corresponding Life Sciences Identifier (LSID), a consistent globally unique identifier

based on the AphiaID (Vandepitte et al., 2015) from the World Register of Marine Species (stored in the field *scientificNameID*). Each record reports whether the species searched for during the survey was found or not (*occurrenceStatus* = present or absent). As explicitly indicated, all records are based on an on-site visual census (*basisOfRecord* = HumanObservation) carried out by an EcoDiver identified by name and unique certification number (*identifiedBy*).

DwC Measurement or Facts Extension (eMoF) Table

The DwC eMoF table contains additional quantitative information on species occurrences and events. Records are linked to every single occurrence (*occurenceID*) and to the individual survey events (*eventID*) to which they belong. Four types of measurement (*measurementType*) are stored:

- "Abundance category of a biological entity specified elsewhere" for each occurrence;
- "Depth minimum of biological entity specified elsewhere on the bed by epibenthic sampling" (in meters) for each occurrence;
- "Depth maximum of biological entity specified elsewhere on the bed by epibenthic sampling" (in meters) for each occurrence;
- "Sample duration" (in minutes) for each survey event.

Measurement type and units refer to the NERC vocabulary, as indicated in the appropriate fields (*measurementTypeID*, *measurementUnitID*). Corresponding values are stored in the *measurementValue* and *measurementUnit* fields. Abundance categories are identified by a ranking number, using numerical or descriptive classes according to the countability of the taxa/organisms searched for (in brackets the corresponding descriptive categories):
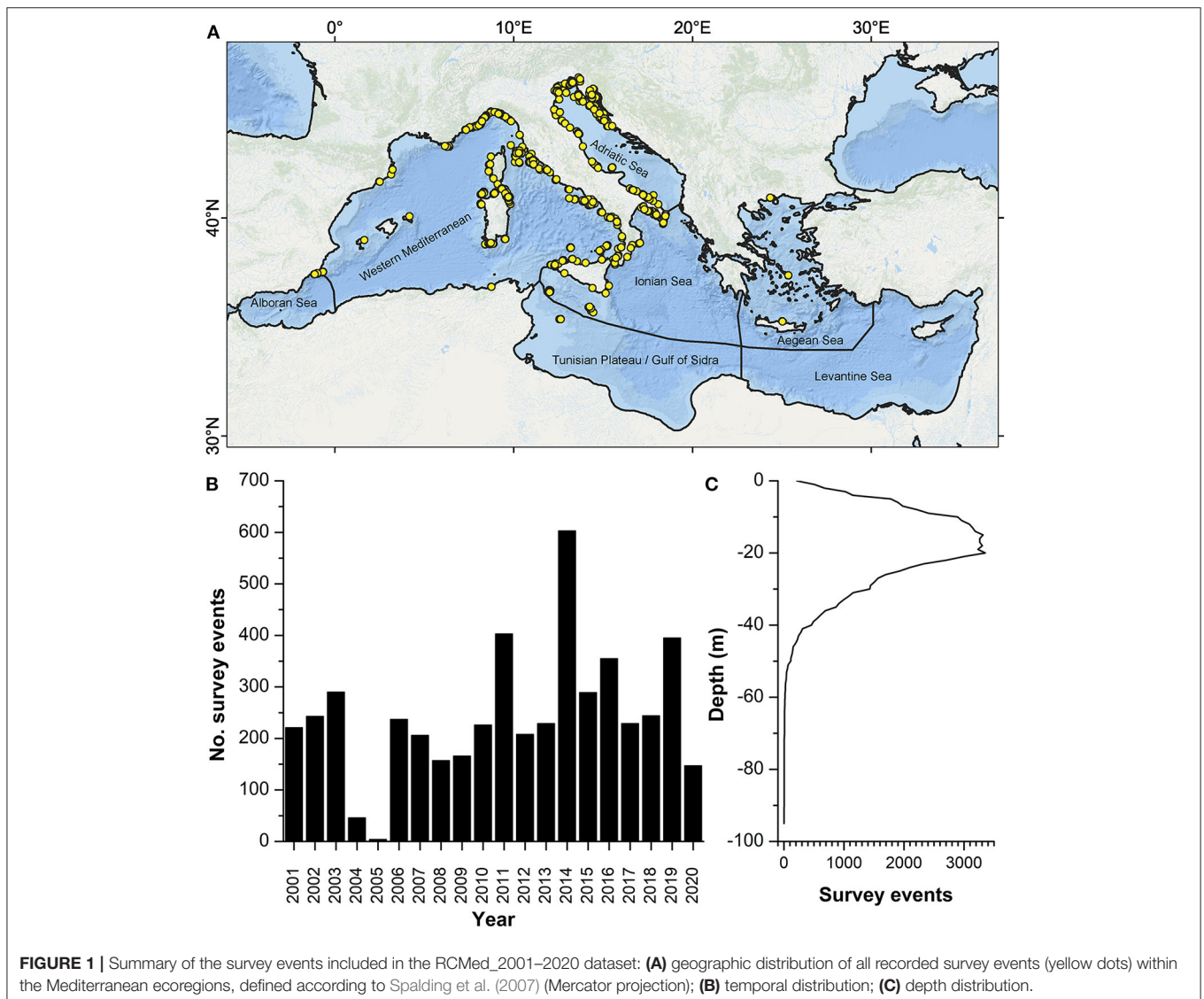

FIGURE 1 | Summary of the survey events included in the RCMed_2001–2020 dataset: (A) geographic distribution of all recorded survey events (yellow dots) within the Mediterranean ecoregions, defined according to Spalding et al. (2007) (Mercator projection); (B) temporal distribution; (C) depth distribution.

- Category 0: 0 specimens (absent)
- Category 1: 1 specimen (isolated specimen)
- Category 2: 2 specimens (some scattered specimens)
- Category 3: 3–5 specimens (several scattered specimens)
- Category 4: 6–10 specimens (a crowded area)
- Category 5: 11–50 specimens (some crowded areas)
- Category 6: > 50 specimens (several crowded areas)

DATA SEARCH, UPDATES, AND USE

The RCMed_2001–2020 dataset is distributed under the international Creative Common license (CC BY 4.0), which guarantees transparency on the origin of the data and allows for free sharing and adaptation, giving appropriate credit to the Reef Check Mediterranean network. It can be directly accessed from the EMODnet Biology Portal[3] that offers different services, including the data catalog, a data download toolbox with a step-wise filtering approach, a map viewer, the atlas of marine life data, and a web feature service (WFS), compliant with the Open Geospatial Consortium (OGC) standards for direct integration in geographic information systems (Martín Míguez et al., 2019). Thanks to the interoperability of the network (Tanhua et al., 2019), the dataset is redistributed under the Ocean Biodiversity Information System (OBIS) networks (including EurOBIS, MedOBIS; Costello and Vanden Berghe, 2006 and references therein), the European infrastructure on biodiversity and ecosystem research (LifeWatch; Basset and Los, 2012), and the Global Biodiversity Information Facility (GBIF; Flemons et al., 2007). Periodic submissions of newly acquired data to EMODnet are expected.

DATASET CONTENTS AND APPLICATIONS

The RCMed_2001–2020 dataset consists of 50,255 observations unevenly distributed among 43 key taxa in the Mediterranean Sea recorded in 4,898 individual survey events, carried out by 692 EcoDivers from 2001 to 2020. The data comes from Croatia, France, Greece, Italy, Spain, and Tunisia, covering parts of the following ecoregions (*sensu* Spalding et al., 2007): Western Mediterranean (52.3% of the surveys), Adriatic Sea (42.2%), Ionian Sea (4.9%), Alboran Sea (0.2%), Aegean Sea (0.2%), and Tunisian Plateau/Gulf of Sidra (0.1%; **Figure 1A**). After an initial period of protocol development in the Adriatic Sea (2001–2003, originally called "Adriatic Underwater Watching Project") with 200–300 surveys carried out per year, there was a reduction in the number of surveys the following two years. After this, the number of surveys per year has varied from 150 to 600, with the minimum value in 2020. This is likely related to the COVID-19 pandemic lockdown (**Figure 1B**). While ∼ 97% of observations took place in the recreational diving depth range (0–40 m), the maximum depth reached during surveys was 95 m (**Figure 1C**). The spatial and temporal distribution of the data is affected by the volunteers' willingness, habits and preferences applying the RCMed U-CEM protocol. However, spatial and temporal biases are recognized as major issues in CS projects and biodiversity databases, remaining

intrinsically unavoidable for this and most other CS initiatives (Beck et al., 2014).

The United Nations Decade of Ocean Science for Sustainable Development (2021–2030) asks for an urgent improvement in marine conservation actions worldwide. Similarly, the EU Biodiversity Strategy for 2030 includes among its main tasks an enhanced focus on Natura 2000 species and habitats and a Nature Restoration Plan of degraded ecosystems across the EU, addressing the key drivers of biodiversity loss. Without a detailed census and mapping of the distribution and abundances of target species, it is impossible to address these objectives effectively. Marine Citizen Science is a promising and powerful tool to enhance engagement in marine conservation worldwide and increase ocean observation capability ensuring long-term monitoring whenever appropriate protocols are applied. In these regards, the application of the RCMed_2001–2020 dataset ranges from: monitoring the ecological status of Mediterranean coastal environments to assessing the effects of human impacts and management interventions (Turicchia et al., 2021a); raising public awareness; and involving people in marine conservation (Lucrezi et al., 2018 and references therein). Moreover, the dataset has been used to complement scientific papers on species distribution and abundance, distribution modeling, and comparing historical data series (Cerrano et al., 2017; Ponti et al., 2018; Turicchia et al., 2018). A list of applications and publications obtained by applying the protocol and using this data is kept up to date on the Reef Check Med website, and authors are encouraged to report their outcomes.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: EMODnet Biology data portal, https://doi.org/10.14284/468.

AUTHOR CONTRIBUTIONS

MP conceived the database of Reef Check Med Dataset—key Mediterranean marine species 2001–2020. ET and MP wrote the first draft of the manuscript. All authors contributed to and approved the final version of the manuscript.

FUNDING

ACKNOWLEDGMENTS

---

[3]https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6454

REFERENCES

Airoldi, L., and Beck, M. W. (2007). Loss, status and trends for coastal marine habitats of Europe. *Oceanogr. Mar. Biol., Annu. Rev.* 45, 345–405. doi: 10.1201/9781420050943.ch7

Basset, A., and Los, W. (2012). Biodiversity e-Science: LifeWatch, the European infrastructure on biodiversity and ecosystem research. *Plant Biosyst.* 146, 780–782. doi: 10.1080/11263504.2012.740091

Beck, J., Boller, M., Erhardt, A., and Schwanghart, W. (2014). Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecol. Inform.* 19, 10–15. doi: 10.1016/j.ecoinf.2013.11.002

Cerrano, C., Milanese, M., and Ponti, M. (2017). Diving for science – science for diving: volunteer scuba divers support science and conservation in the Mediterranean Sea. *Aquat. Conserv.* 27, 303–323. doi: 10.1002/aqc.2663

Costello, M. J., and Vanden Berghe, E. (2006). 'Ocean biodiversity informatics': a new era in marine biology research and management. *Mar. Ecol. Prog. Ser.* 316, 203–214. doi: 10.3354/meps316203

Flemons, P., Guralnick, R., Krieger, J., Ranipeta, A., and Neufeld, D. (2007). A web-based GIS tool for exploring the world's biodiversity: The Global Biodiversity Information Facility Mapping and Analysis Portal Application (GBIF-MAPA). *Ecol. Inform.* 2, 49–60. doi: 10.1016/j.ecoinf.2007.03.004

Foley, M. M., Halpern, B. S., Micheli, F., Armsby, M. H., Caldwell, M. R., Crain, C. M., et al. (2010). Guiding ecological principles for marine spatial planning. *Mar Policy* 34, 955–966. doi: 10.1016/j.marpol.2010.02.001

Garrabou, J., Gómez-Gras, D., Ledoux, J.-B., Linares, C., Bensoussan, N., López-Sendino, P., et al. (2019). Collaborative database to track mass mortality events in the Mediterranean Sea. *Front. Mar. Sci.* 6:707. doi: 10.3389/fmars.2019.00707

Hill, J., and Wilkinson, C. (2004). *Methods for Ecological Monitoring of Coral Reefs. A Resource for Managers.* Townsville, AU: Australian Institute for Marine Science.

Hochachka, W. M., Fink, D., Hutchinson, R. A., Sheldon, D., Wong, W. K., and Kelling, S. (2012). Data-intensive science applied to broad-scale citizen science. *Trends Ecol. Evol.* 27, 130–137. doi: 10.1016/j.tree.2011.11.006

Lucrezi, S., Milanese, M., Palma, M., and Cerrano, C. (2018). Stirring the strategic direction of scuba diving marine Citizen Science: A survey of active and potential participants. *PLoS ONE* 13:e0202484. doi: 10.1371/journal.pone.0202484

Martín Míguez, B., Novellino, A., Vinci, M., Claus, S., Calewaert, J.-B., Vallius, H., et al. (2019). The European Marine Observation and Data network (EMODnet): Visions and roles of the gateway to marine data in Europe. *Front. Mar. Sci.* 6:313. doi: 10.3389/fmars.2019.00313

Micheli, F., Halpern, B. S., Walbridge, S., Ciriaco, S., Ferretti, F., Fraschetti, S., et al. (2013). Cumulative human impacts on Mediterranean and Black Sea marine ecosystems: assessing current pressures and opportunities. *PLoS ONE* 8:e79889. doi: 10.1371/journal.pone.0079889

Ponti, M., Perlini, R. A., Ventra, V., Grech, D., Abbiati, M., and Cerrano, C. (2014). Ecological shifts in Mediterranean coralligenous assemblages related to gorgonian forest loss. *PLoS ONE* 9:e102782. doi: 10.1371/journal.pone.0102782

Ponti, M., Turicchia, E., Ferro, F., Cerrano, C., and Abbiati, M. (2018). The understorey of gorgonian forests in mesophotic temperate reefs. *Aquat. Conserv.* 28, 1153–1166. doi: 10.1002/aqc.2928

Ponti, M., Turicchia, E., Rossi, G., and Cerrano, C. (2021). Reef check med – key Mediterranean marine species 2001-2020. Dataset maintained by Reef Check Italia onlus, EMODnet Biology data portal. [Dataset]

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M. A. X., et al. (2007). Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* 57, 573–583. doi: 10.1641/B570707

Tanhua, T., Pouliquen, S., Hausman, J., O'Brien, K., Bricher, P., de Bruin, T., et al. (2019). Ocean FAIR data services. *Front. Mar. Sci.* 6:440. doi: 10.3389/fmars.2019.00440

Turicchia, E., Abbiati, M., Sweet, M., and Ponti, M. (2018). Mass mortality hits gorgonian forests at Montecristo Island. *Dis. Aquat. Org.* 131, 79–85. doi: 10.3354/dao03284

Turicchia, E., Cerrano, C., Ghetta, M., Abbiati, M., and Ponti, M. (2021a). *MedSens* index: The bridge between marine citizen science and coastal management. *Ecol. Indic.* 122:107296. doi: 10.1016/j.ecolind.2020.107296

Turicchia, E., Ponti, M., Rossi, G., Milanese, M., Di Camillo, C. G., and Cerrano, C. (2021b). The reef check Mediterranean underwater coastal environment monitoring protocol. *Front. Mar. Sci.* 8:620368. doi: 10.3389/fmars.2021.620368

Vandepitte, L., Vanhoorne, B., Decock, W., Dekeyzer, S., Trias Verbeeck, A., Bovit, L., et al. (2015). How Aphia—the platform behind several online and taxonomically oriented databases—can serve both the taxonomic community and the field of biodiversity informatics. *J. Mar. Sci. Eng.* 3, 1448–1473. doi: 10.3390/jmse3041448

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., et al. (2012). Darwin core: an evolving community-developed biodiversity data standard. *PLoS ONE* 7:e29715. doi: 10.1371/journal.pone.0029715

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* 3:160018. doi: 10.1038/sdata.2016.18