OPEN

Coralligenous and maërl habitats: predictive modelling to identify their spatial distributions across the Mediterranean Sea

C. S. Martin[1,2], M. Giannoulaki[1], F. De Leo[3], M. Scardi[4], M. Salomidi[5], L. Knitweiss[6], M. L. Pace[6], G. Garofalo[7], M. Gristina[7], E. Ballesteros[8], G. Bavestrello[9], A. Belluscio[10], E. Cebrian[8,11], V. Gerakaris[5], G. Pergent[12], C. Pergent-Martini[12], P. J. Schembri[13], K. Terribile[13], L. Rizzo[3], J. Ben Souissi[14], M. Bonacorsi[12], G. Guarnieri[3], M. Krzelj[15], V. Macic[16], E. Punzo[17], V. Valavanis[1] & S. Fraschetti[3]

[1]Institute of Marine Biological Resources and Inland Waters, Hellenic Centre for Marine Research, PO box 2214, 71003 Heraklion (Greece), [2]UNEP World Conservation Monitoring Centre, 219 Huntingdon Road, Cambridge CB3 0DL (UK), [3]Laboratory of Ecology and Marine Biology, University of Salento, CoNISMa, 73100 Lecce (Italy), [4]Department of Biology, University of Rome ''Tor Vergata'', Via della Ricerca Scientifica, 00133 Rome (Italy), [5]Institute of Oceanography, Hellenic Centre for Marine Research, PO box 712, 19013 Anavyssos (Greece), [6]Laboratory of Ecological Assessment & Monitoring Department of Biology, Faculty of Science University of Malta Msida, MSD 2080, Malta, [7]Istituto per l'Ambiente Marino Costiero, National Research Council, 91026 Mazara del Vallo (TP) (Italy), [8]Centre d'Estudis Avançats de Blanes-CSIC, 17300 Blanes (Spain), [9]Dipartimento di Scienze della Terra, dell'Ambiente e della Vita, University of Genoa, Genoa (Italy), [10]Department of Environmental Biology, University of Rome ''La Sapienza'', Viale dell'Università, 32, 00185 Rome (Italy), [11]Departament de Ciències Ambientals, Facultat de Ciències, Universitat de Girona, 17071 Girona (Spain), [12]FRES 3041, EqEL, University of Corsica, Faculty of Sciences, BP 52, 20250 Corte (France), [13]Department of Biology, University of Malta, Msida MSD2080 (Malta), [14]Institut National Agronomique de Tunisie, 43 Avenue Charles Nicolle, 1082 Tunis (Tunisia), [15]University Department of Marine Studies, University of Split, Livanjska 5/III, 21000 Split (Croatia), [16]Institute of marine biology, University of Montenegro, 85330 Kotor (Montenegro), [17]Institute of Marine Sciences, National Research Council, Largo Fiera della Pesca, 60125 Ancona (Italy).

Bioconstructions such as coralligenous outcrops and maërl beds are typical Mediterranean underwater seascapes. Fine-scale knowledge on the distribution of these sensitive habitats is crucial for their effective management and conservation. In the present study, a thorough review of existing spatial datasets showing the distribution of coralligenous and maërl habitats across the Mediterranean Sea was undertaken, highlighting current gaps in knowledge. Predictive modelling was then carried out, based on environmental predictors, to produce the first continuous maps of these two habitats across the entire basin. These predicted occurrence maps for coralligenous outcrops and maërl beds provide critical information about where the two habitats are most likely to occur. The collated occurrence data and derived distribution model outputs can help addressing the challenge of developing basin-wide spatial plans and to guide cost-effective future surveys and monitoring efforts towards areas that are presently poorly-sampled.

F ine-scale knowledge on the distribution of species and habitats is crucial for effective management and conservation of marine resources[1–4]. Conservation prioritisation exercises require good quality information on the spatial distribution of vulnerable species and their associated habitats, including different life history stages[5]. Such spatial information is also critical to decision-makers and managers, so that marine resources are sustainably exploited, and other human activities (e.g. extractive industries, maritime transport, fisheries and aquaculture) seek to minimise negative impacts[6]. At present, data available to address these issues typically consist of sparse geo-referenced information on species and habitat occurrences. At best, absence records for non mobile species and benthic habitats are usually only available for a limited number of sites since the absence of a species or habitat is only ascertained when a given site has been systematically surveyed. Mapping marine biodiversity

remains operationally complicated and expensive, with the result that fine-scale knowledge of species and habitat distributions is unavailable for most marine areas[7–9].

In the Mediterranean Sea, there have been several attempts at assessing the distribution patterns of species and habitats across the entire basin, based on literature reviews[10–14]. Recently, Giakoumi et al.[15] assessed potential spatial priorities for the conservation of three Mediterranean habitats (*Posidonia oceanica* meadows, coralligenous formations, and marine caves) by considering their eco-regional representation, as well as the opportunity costs for fisheries and aquaculture. These efforts are partly driven by existing, as well as emerging policies at the national and international levels. Member States (MS) of the European Union (EU) have, for instance, committed to collating knowledge on the distribution of species and habitats, and assessing their ecological status, as part of an ecosystem approach to marine management. Based on this information, targets and associated indicators will be used to guide progress towards achieving 'Good Environmental Status' (GES) in MS' marine waters by 2020[16]. A similar process is emerging at the scale of the Mediterranean basin for the Contracting Parties to the Barcelona Convention (1976). To date, these efforts have documented important gaps in knowledge, especially on species and habitats that are considered of critical importance for the Mediterranean Sea and its conservation[15].

Bioconstructions such as coralligenous outcrops and maërl beds are typical Mediterranean underwater seascapes, comprising coralline algal frameworks that grow in dim light conditions[17]. They are the result of the building activities of algal and animal constructors, counterbalanced by physical, as well as biological, eroding processes. Because of their extent, biodiversity and production, coralligenous and maërl habitats rank among the most important ecosystems in the Mediterranean Sea[17–21], and they are considered of great significance both for fisheries[22] and carbon regulation[23,24].

Mechanical disturbance and re-suspension of nearby sediments, particularly by bottom trawling, is probably the most destructive human activity currently affecting coralligenous outcrops and maërl beds[17,25,26]. Other threats include pollution (e.g. wastewater discharge, aquaculture), which results in increased turbidity and sedimentation, but also direct habitat destruction through artisanal and recreational fishing (e.g. fishnets, long-lines), coastal or offshore construction activities (including submarine cables), and unregulated diving activities and anchoring[17,19,27]. Climate change is also known to affect several key species that are part of coralligenous habitats, by increasing the incidence of thermal anomalies (e.g.[28–30]) and storms[31]. Some invasive algal species (*Womersleyella setacea*, *Acrothamnion preissii*, *Caulerpa racemosa* v. *cylindracea* and *C. taxifolia*) can also pose a severe threat to these communities, either by forming dense carpets (i.e. physical barriers) or by increasing sedimentation[26,32–35]. Such a pervasive range of impacts, coupled with the slow growth rates and long recovery periods of these systems, have driven efforts aimed at conserving them.

Although not legally binding, the Barcelona Convention's 'Action plan adopted in 2008 for the conservation of coralligenous outcrops and other calcareous bio-concretions in the Mediterranean Sea' asserts that "coralligenous/maërl assemblages should be granted legal protection at the same level as *Posidonia oceanica* meadows"[26]. Coralligenous outcrops also appear in the EU's Habitats Directive[36] (under *1170 Reefs*), and in the Bern Convention[37]. Two of the most common maërl-forming Mediterranean species, *Lithothamnion corallioides* and *Phymatolithon calcareum*, are included in Annex V of the Habitats Directive. Finally under European law[38], destructive fishing is prohibited over Mediterranean coralligenous and maërl bottoms. The substantial lack of relevant geospatial data, however, significantly hinders the effective implementation of these policies[11,19]. Giakoumi et al.[15] recently produced a basin-scale distribution map integrating all benthic assemblages thriving on hard substrata of biogenic origin and

under low irradiance levels, along with rhodolith beds in coastal detritic bottoms, selected deep-sea habitats (e.g. seamount peaks, offshore rocky banks), and some deep coral communities. There was no attempt at discriminating between these very different systems and spatial planning analyses were carried at coarse spatial resolution (presence/absence in a grid of 10 km cell size). Continuous spatial information hence remains unavailable, hampering the development of effective spatial measures to protect coralligenous outcrops and maërl beds.

A number of modelling techniques can be used to fill gaps in the knowledge of the spatial distribution of species and habitats by predicting the location of areas that are likely to be suitable for a species or a community to live[39–41]. Models are usually based on physical and environmental variables (e.g. water temperature, salinity, depth, nutrient concentrations, seabed types, etc), which are typically easier to record and map across vast expanses (i.e. regional, global scale) in contrast to species and habitat data[42–44]. Despite inherent limitations and associated uncertainties, predictive modelling is a cost-effective alternative to field surveys as it can help identifying and mapping where sensitive marine ecosystems may occur.

In the present study, a thorough review of existing spatial datasets showing the distribution of coralligenous and maërl habitats across the Mediterranean Sea was undertaken, with particular attention given to the basin's eastern and southern parts, where data have traditionally been limited in regional reviews (see for instance[11]). Based on the collated spatial datasets (parts of which are, to date, unpublished), predictive modelling was carried out to produce the first continuous maps of these two habitats across the Mediterranean Sea. We anticipate that our results will be critical (i) for the development of basin-wide spatial planning initiatives (including representative networks of marine protected areas) based on realistic information on habitat distribution, and (ii) to guide cost-effective future surveys and monitoring efforts towards areas that are presently poorly-sampled and under-represented in current conservation planning exercises.

Results

Occurrence datasets for coralligenous outcrops and maërl beds. Datasets on coralligenous outcrops and maërl beds came from a total of 17 countries (Supplementary Table S2), and in a wide variety of formats: from shapefiles and lossless rasters, to image maps in paper format, or electronic format with information loss through compression. The datasets were found to be heterogeneous, with scales from 1:4,000 to 1:250,000, and un-standardised legends, even within the same country. The collated coralligenous outcrops dataset was composed of 4,293 points, 12 lines and 23,632 polygons (Figure 1a). That of maërl beds had 416 points and 748 polygons (Figure 1b). Together, the surface areas corresponding to the polygons only amounted to 2,763.4 km² (coralligenous outcrops) and 1,654.5 km² (maërl beds). Point and line data do not have associated surface areas. Thus, they were used in the modelling but not for surface areas estimates.

We estimated that, combining all the collected information in terms of points/lines/polygons, and only considering the length of the coast where data had been retrieved, datasets provided information for approximately 30% of the coasts of the Mediterranean basin. For the remaining coastline (70%), no further information could be found and/or accessed. *In situ* depth of occurrences collected from the publications (see the Supplementary References) revealed that records were located between 10 and 140 m (Figure 2), peaking in the shallower sector of this range. Based on this, modelling was restricted to the 0–200 m depth zone. Published information was found to be insufficient for deriving a value of sampling effort across depth bins.

Scientific information on these two habitats remained unevenly distributed, essentially because the majority of systematic studies have taken place in the western Mediterranean. Areas for which
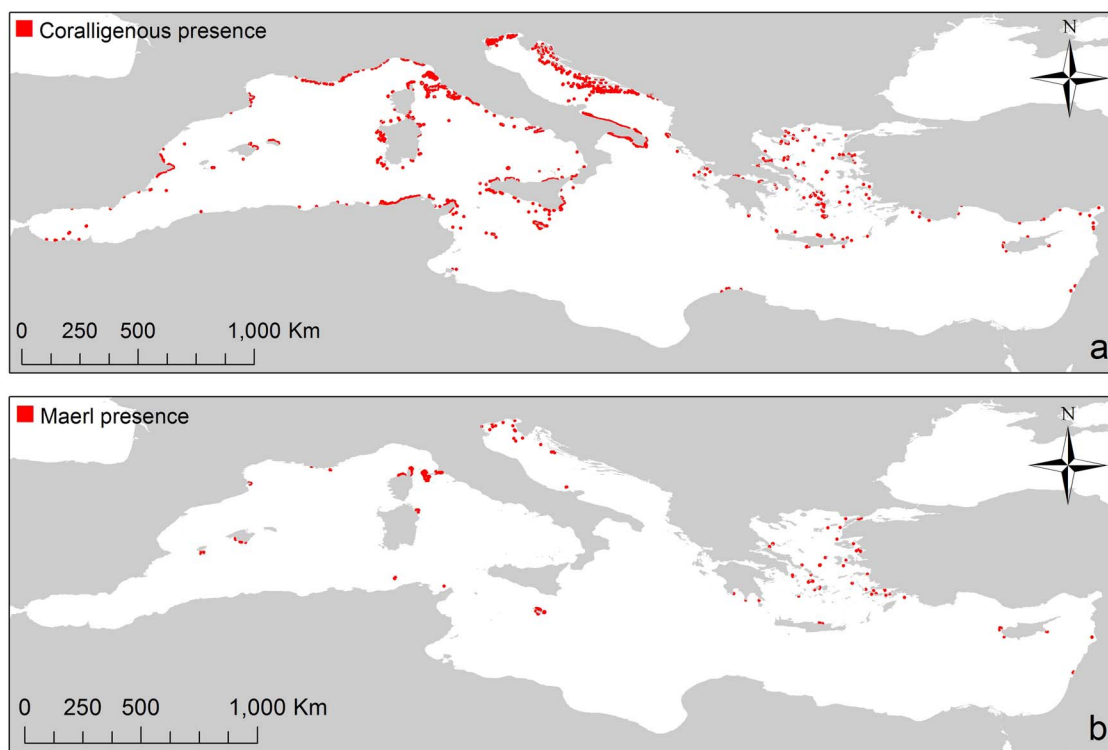
Figure 1 | Occurrences of (a) coralligenous outcrops and (b) maërl beds across the Mediterranean Sea, as extracted from the review work, and used in the distribution models (see text for details). Data sources are listed in the Supplementary References. Boundaries of point/line/polygon features of the data layers have artificially been enhanced so that very small-scale occurrences are visible on the illustrative maps shown here. As a result, surface areas covered by these habitats appear much larger than they are in reality (e.g. around Malta). These data layers are more accurately viewed within a Geographic Information System. Maps were created using ArcGIS® software by Esri (Environmental Systems Resource Institute, ArcMap 10.1, (www.esri.com).

information was previously unavailable were, however, much better covered by the present study, particularly the eastern Mediterranean Sea. Important new information was gained from Malta, Italy, France (Corsica), Spain, Croatia, Greece, Albania, Algeria, Tunisia and Morocco, making the present datasets the most comprehensive to date. In Malta and Italy, knowledge was particularly extensive. Distribution maps of bioconstructions were available in shapefiles for several portions of the Italian coastline, covering continuous stretches of coasts of hundreds of kilometres (i.e. Ligurian Sea, Tyrrhenian Sea, Apulia, Sicily). Still, there were areas of the Mediterranean Sea where data remained extremely scarce (Albania, Algeria,

Cyprus, Israel, Libya, Montenegro, Morocco, Syria, Tunisia and Turkey) or totally absent (Bosnia and Herzegovina, Egypt, Lebanon and Slovenia). Knowledge on maërl beds was somewhat limited compared to what was available for coralligenous outcrops; a significant update was nevertheless achieved. Previously unknown spatial information on maërl distribution was brought to light for Greece, France (Corsica), Cyprus, Turkey, Spain and Italy. Malta and Corsica, in particular, had significant datasets for this habitat as highlighted by fine-scale surveys in targeted areas.

Coralligenous outcrops occurrence model. A total of 11,174 presence points (i.e. the training set derived from the combined polygon, line and point occurrence dataset) were used to model the occurrence of coralligenous outcrops across the Mediterranean Sea. The final model based on this training set retained six variables from the starting subset of 12, and the AUCs were 0.80 for the training set and 0.77 (standard deviation 0.003) for the geographically independent test set of 5,581 points (Figure 3a). Bathymetry, slope of the seafloor and nutrient input were the three main contributors to the model (combined contribution of 84.1%; Table 1), whilst the remaining three predictors (euphotic depth, phosphate concentration and geostrophic velocity of sea surface current) had a combined contribution of 16%. Variable response curves suggested a unimodal response, in support of grouping these species together, at the spatial scale considered.

Based on a jackknife test of variable importance (for the test gain; Supplementary Figure S5a), the predictor variable with the highest gain when used in isolation was nutrient input, which therefore appeared to have the most useful information by itself. The predictor variable that decreased the gain the most when it was omitted was euphotic depth, which therefore had the most information that was not present in the other predictor variables. The jackknife test on the
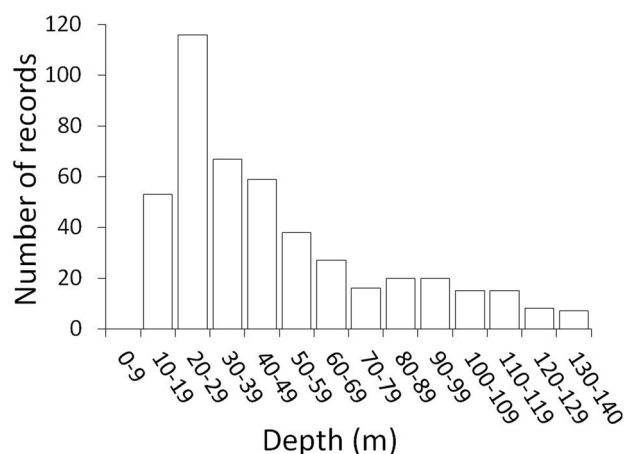

Figure 2 | *In situ* depths of occurrences for coralligenous outcrops and mäerl beds, as extracted from the review work.
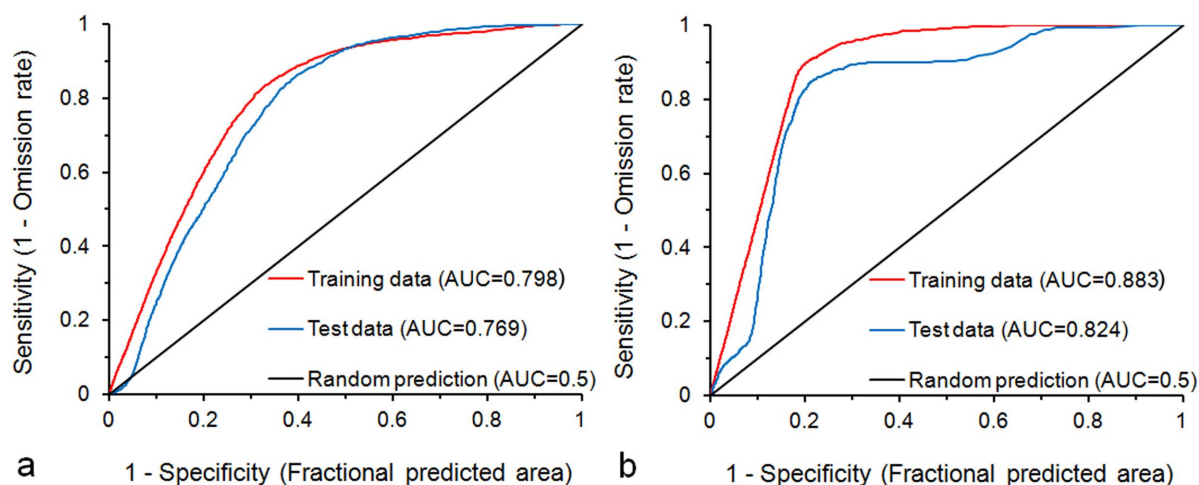
3

Figure 3 | ROC (Receiver Operating Characteristic) curves for the training and test sets of (a) coralligenous outcrops and (b) maërl beds. AUC: Area Under the Curve.

test set's AUC (Supplementary Figure S5b) confirmed that bathymetry, slope of the seafloor and nutrient input were the main contributors to the model and highlighted the role of sea surface current in predicting the occurrence of coralligenous outcrops.

Areas predicted to have suitable conditions for the occurrence of coralligenous outcrops are shown in Figure 4a and were generally consistent with known presence areas. Predicted occurrences for the North African coast highlighted suitable areas for which there were no occurrence data. This suggested that the measures taken to (i) address the geographic sampling bias (target group background) and (ii) to prevent overfitting (hinge feature; regularisation multiplier) had been efficient.

Maërl beds occurrence model. A total of 4,612 presence points (i.e. the training set derived from the combined polygon and point occurrence dataset) were used to model the occurrence of maërl beds across the Mediterranean. The final model based on this training set retained seven variables from the starting subset of 12, and the AUCs were 0.88 for the training set, and 0.82 (standard deviation 0.004) for the geographically independent test set of 2,204 points (Figure 3b). Phosphate concentration, geostrophic velocity of sea surface current, silicate concentration and bathymetry were the four main contributors to the model (combined contribution of 83.6%; Table 2), whilst the remaining three predictors (bottom salinity, euphotic depth and slope of the seafloor) had a combined contribution of 16.4%. As for coralligenous outcrops, unimodal variable response curves supported grouping these species together, at the spatial scale considered.

Based on a jackknife test on the test gain (Supplementary Figure S6a), phosphate concentration appeared to have the most useful information by itself, and also the most information that was not present in the other predictor variables. The jackknife test on the test set's AUC (Supplementary Figure S6b) confirmed that phosphate

and silicate concentrations and sea surface current were the strongest contributors to the model.

Predicted areas with suitable conditions for the occurrence of maërl beds are shown in Figure 4b, and were mostly consistent with known presence areas, with some exceptions such as the Po river estuary in the north of Italy. Given the paucity of occurrence data for this habitat across the Mediterranean, and especially the North African coast, the model output was relatively informative in highlighting several suitable areas where no occurrence data were available to train the model. Again, this suggested that measures taken to prevent issues of geographic sampling bias and overfitting worked. The model predicted high suitability (probability of occurrence >0.8) in one area having no known record of maërl beds: the southern Evoikos Gulf (Greece). This area happens to have a relatively high phosphate concentration, and groundtruthing would be necessary to confirm the actual presence of maërl beds, beyond the predicted suitability of the area.

Discussion

Spatial data on coralligenous and maërl habitats have become increasingly available during the last twenty years[11,15,45], indicating that these bioconstructions occur widely across the Mediterranean basin[17]. The present study has provided (i) the most comprehensive update for the distributions of coralligenous outcrops and maërl beds across the Mediterranean Sea, going much further than previous studies, in particular for the eastern Mediterranean basin, and (ii) the first basin-wide and continuous distribution maps based on predictive modelling. Knowledge acquisition was particularly acute for maërl beds, for which data on spatial occurrence had remained comparatively scarce before this review and associated modelling exercise.

Surface areas reported here for coralligenous outcrops (2,763 km²) and maërl beds (1,654 km²) were based on ("raw") polygon data resulting from *in situ* observations (i.e. not from the model outputs), predominantly from small-scale studies, limited to the 0 to 200 m depth band. These figures do not include surface areas associated with vertical cliffs, where coralligenous outcrops are commonly found. Point and line data were not used in the surface area estimations, as they do not have associated surface areas. Hence, the figures given here clearly underestimate the real spatial extent of coralligenous and maërl habitats in the Mediterranean Sea. In this region, spatial data on species and habitat distributions remain very patchy and in many locations the information ranges from low quality to completely unavailable. It remains challenging for the Mediterranean but also elsewhere, to integrate information that has often

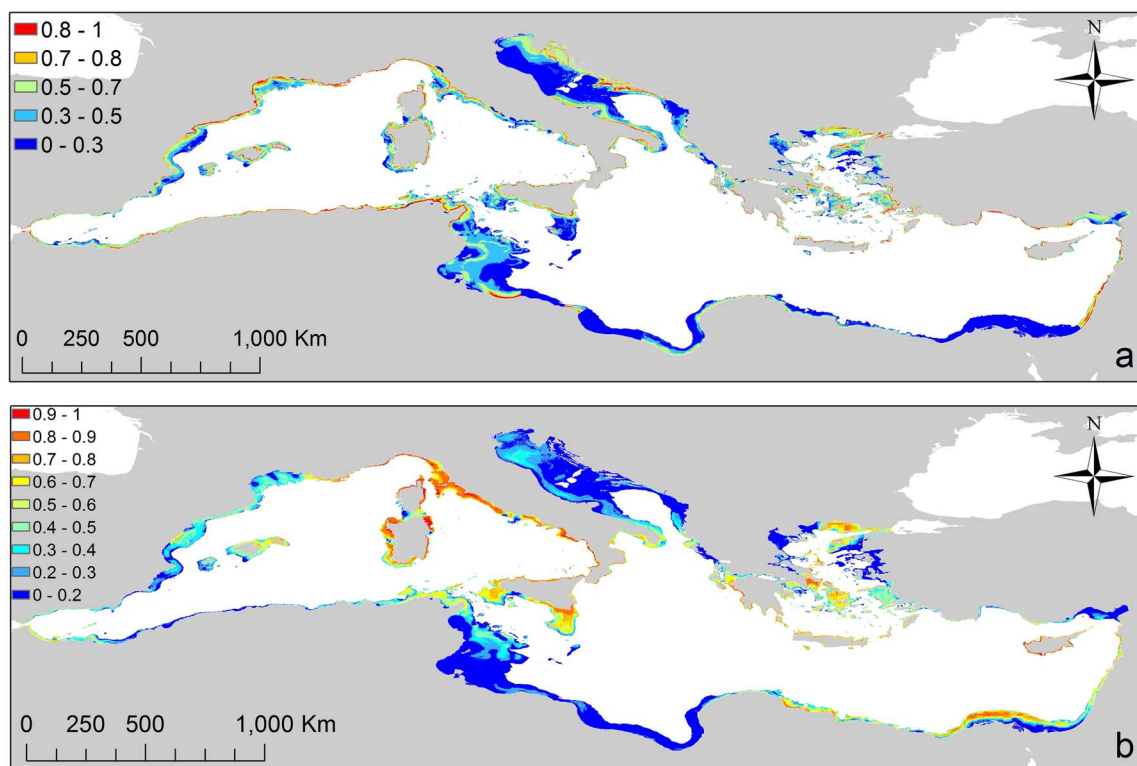| Table 1 | Relative contributions of each predictor variable to the coralligenous outcrops distribution model | |
|---|---|
| Predictor variable | Contribution (%) |
| Bathymetry | 37.4 |
| Slope of the seafloor | 31.9 |
| Nutrient input | 14.8 |
| Euphotic depth | 6.5 |
| Phosphate concentration | 4.8 |
| Sea surface current | 4.7 |

Figure 4 | Spatial distributions showing occurrence probabilities for (a) coralligenous outcrops and (b) maërl beds across the Mediterranean Sea, as predicted using distribution modelling. Maps were created using ArcGIS® software by Esri (Environmental Systems Resource Institute, ArcMap 10.1, (www.esri.com).

been collected using differing approaches and that is stored in various repositories. Considering the available polygon data and their associated surface areas, we roughly estimate that as much as 95% of coralligenous habitat may still need to be mapped across the Mediterranean basin, especially in deeper areas. The value is probably even higher for maërl beds. Given their high biodiversity value, the systematic mapping of these two habitats across the Mediterranean should be a priority, especially as they can be used more widely to track anthropogenic disturbances, for instance as part of the EU's Marine Strategy Framework Directive[16] and the Barcelona Convention.

The present study adopted a presence-only modelling approach because of the paucity of known absence areas for coralligenous outcrops, and even more for maërl beds. The "species" datasets were strongly spatially biased towards northwest marine regions. Despite measures taken to minimise this bias and also overfitting, some areas of known absence were predicted to be suitable for coralligenous outcrops (e.g. Nile delta, north-eastern coast of Italy): these false predicted presences were manually removed from the final maps. In contrast, areas of known presence, especially those based on point data, did not necessarily show very high suitability levels. Model

outputs (Figure 4) were hence presented herein in combination with the collated observed occurrence data (Figure 1). This implies that spatial management measures for fisheries that are aimed at protecting coralligenous outcrops and maërl beds, should not be based solely on the model outputs presented here; targeted groundtruthing should be carried out so that informed decisions are taken.

Due to data limitations on species lists across the various component datasets, coralligenous outcrops and maërl beds were each modelled as a whole, instead of modelling multispecific assemblages with distinct habitat preferences. Inspection of (unimodal) response curves did suggest that the approach taken was appropriate at the spatial scale used. This may not necessarily be the case at the local or site scale. It thus remains important to encourage systematic assessments of species composition of the two habitats across the region, beyond simply recording the presence of the habitat as a whole, so that species and assemblages with different habitat preferences can be modelled and mapped separately.

While better coralligenous outcrops and maërl beds data would certainly improve model outputs, so would better alternatives for predictor variables, especially if they have finer spatial resolutions. Predictor resolution, from the global/continental scales to the site/micro scales, influences the importance of different variables in controlling species distributions across varying spatial scales[46]. For instance, very good site-scale predictors of coralligenous outcrops occurrence are hard substrata and steep underwater cliffs (although concretions over flat rocky surfaces and platform coralligenous assemblages are also extremely common), combined with strong currents, light between 0.05% and 3% of surface irradiance, and nutrient-poor waters[17]. Coralligenous outcrops would be unlikely to occur in high sedimentary zones without hard substrata, in enclosed estuarine systems, and in sandy areas with low salinities such as river mouths, although some exceptions exist[17]. For maërl beds, flat and coarse grained areas would tend to be suitable habitats, as well as straits with strong bottom currents that reduce sedimenta-

Table 2 | Relative contributions of each predictor variable to the maërl beds distribution model

Predictor variable	Contribution (%)
Phosphate concentration	44.6
Sea surface current	16.3
Silicate concentration	12.9
Bathymetry	9.8
Bottom salinity	8.0
Euphotic depth	5.7
Slope of the seafloor	2.7

tion[23,47]. The present study did not manage to unearth such fine-resolution predictor variables in mapped formats that would cover the entire Mediterranean Sea; besides, the spatial resolutions of the variables that were available were coarser than desired. As fine-resolution (i.e. local, site and micro-scales) spatial data on e.g. bottom types or salinity were not available at the scale of the Mediterranean basin, coarser-resolution (i.e. regional and landscape scales) surrogates were used, thereby constraining model behaviour and the derived local-scale interpretation.

The main drivers of the coralligenous outcrops model were bathymetry, slope of the seafloor and nutrient input. Those of the maërl beds model were phosphate concentration, sea surface current, silicate concentration and bathymetry. For bathymetry, a link of causality with presence of bioconstructions is possible. In contrast, predictor variables that are measured *in situ* and interpolated (e.g. phosphate and silicate concentrations), or even modelled (e.g. bottom salinity), harbour more uncertainty, and may play a role in the model because of their "shape" (i.e. spatial gradients and patterns), not because of causality or correlation.

The map of predicted coralligenous outcrops occurrence agrees in parts with the multi-criteria evaluation approach of Cameron and Askew[48] for the western Mediterranean: there, data layers such as bottom substratum, current and bathymetry were combined within a Geographic Information System (GIS) using various thresholds. The two approaches give comparable results for the North Algerian coast, parts of the south-eastern French coastline, the Spanish coast, western Corsican and Sardinian coasts, and the Balearic Islands. In other areas, discrepancies are significant: for instance, most of the western Italian coast is missed by the multi-criteria approach, and also the eastern Corsican and Sardinian coasts, as well as the Tunisian coast.

Disentangling the environmental variables driving the distribution of coralligenous and maërl habitats across the Mediterranean Sea is clearly a challenge and requires detailed knowledge on the ecology and biology of these complex habitats. Specific experimental and observational studies are required to address this issue. However, the predicted occurrence maps for coralligenous outcrops and maërl beds can be of critical importance to guide more-cost-effective survey and monitoring efforts targeting poorly-surveyed areas (e.g. in non-EU countries) and areas where these bioconstructions are putatively likely to occur. In turn, the newly collected data, preferably less spatially biased, could then be used to improve distribution models, since a systematic survey of the whole Mediterranean basin is not a realistic option. Model performance will also improve with finer resolution, or more relevant, predictor variables, resulting in better predicted occurrence maps. To date, however, this predictive modelling exercise remains unique in this regional sea, having provided continuous predicted occurrence maps for two of its most important habitats in terms of biodiversity.

Human impacts on coralligenous and maërl habitats can be substantial, and these effects may grow further in the future as a result of the interlinked effects of climate change and rising anthropogenic pressure. In light of the importance of the processes produced by these habitats, increasing our understanding of their distribution is critical in helping to protect their associated biodiversity. Presently, coralligenous and maërl habitats are considered priority habitats at the European and regional levels, with specific conservation and management measures. The occurrence and predictive maps presented here can be fed into the development of basin-wide conservation plans (e.g. for establishing networks of marine protected areas) or other forms of marine spatial planning, and also in policy development that, at present, are often largely limited by the scarce spatial information on both the distribution and extent of such marine habitats.

Methods

Compiling occurrence datasets for coralligenous outcrops and maërl beds. Georeferenced occurrence records for coralligenous outcrops and maërl beds across the Mediterranean basin were compiled as part of two international research projects:

'MEDISEH'[41] (*Mediterranean Sensitive Habitats*), which was financed by the European Commission under the MAREA Framework, and CoCoNET (Towards COast to COast NETworks of marine protected areas), financed by the EU's 7th Framework Programme. Data sources included peer-reviewed articles, and national, regional and international reports ('grey literature'). This resulted in a total of 771 scientific documents (see Supplementary References), a subset of which had associated spatial information (i.e. maps), information on *in situ* depth of occurrence, and/or species lists for the communities encountered. Spatial information also came from unpublished *in situ* observations by experts and divers. Where digital spatial information (e.g. shapefiles) was not available, shapefiles were created manually by digitising image maps, or by manually extracting spatial information from textual descriptions, based on expert knowledge. All the GIS work (including the maps) was carried out using ArcGIS® software by Esri (Environmental Systems Resource Institute, ArcMap 10.1, www.esri.com).

Predictor variables used as input to the models. An initial screening phase for data layers relevant to predictive modelling of coralligenous and maërl habitats identified 17 datasets that were also spatially continuous at the scale of the Mediterranean basin (Supplementary Table S1). Predictors under consideration ranged from physical (e.g. bathymetry), environmental (e.g. salinity) and anthropogenic (e.g. nutrient input) variables, to calculations (e.g. distance to ports), and *in situ* (e.g. silicate concentration) or remotely-sensed (e.g. euphotic depth) measurements.

All 17 layers were standardised to raster format, having the same geographic extent (the Mediterranean basin), coordinate system (WGS 1984 datum; cylindrical equal-area projection), and resolution (cell size 400 m). This choice of working resolution, albeit artificial, allowed for a better fit along the coastline (extracted from the GSHHS Database[49], version 2.2.1), i.e. with minimal gaps between the 'end' of the predictor layers and the land boundaries. For predictor variables with coarser native resolutions, the resolution was artificially made finer without re-interpolating the data, so that the variables would retain their native resolutions. Gaps in spatial coverage were retained and coded as such.

Habitat modelling. Coralligenous outcrop is a collective term that refers to a very complex biogenic structure mainly created by the outgrowth of encrusting calcareous algae on hard substrata in dim light conditions[17]. Some fleshy and turf algae as well as several groups of sessile invertebrates (e.g. sponges, ascidians, cnidarians, bryozoans, serpulid polychaetes, molluscs) contribute to create the final coralligenous habitat[17,29].

Maërl is also a collective term for a biogenic structure composed of one or more species of free-living (unattached) calcareous red algae (mostly Corallinaceae but also Peyssonneliaceae), dwelling on sedimentary bottoms. These algae can display a branching or a laminar appearance. They sometimes grow as nodules known as rhodoliths that cover all the sea floor, or accumulate within the sand and gravel ripple marks[11,23,50]. Some Authors distinguish between dense accumulations of interlocking rhodoliths within the ripples of muddy and sandy substrates (maërl beds) and rhodoliths dispersed among sediments (rhodolith bottoms). However, in the literature, the terms maërl and rhodolith are also used as synonyms. In the present study, both were combined under the umbrella term maërl beds.

Therefore, coralligenous outcrops and maërl beds are two complex habitats featured by a suite of different species, which will vary locally and regionally[51,52]. The coralligenous habitat can even be considered to be a submarine seascape, or community mosaic, rather than a single community[17]. Although habitat modelling should be better performed on distinct sub-communities of each, e.g. as defined by a prior multivariate analysis[53], this, however, could not be done here due to the scarcity of species lists across the various component datasets. As a result, coralligenous and maërl occurrences were each modelled as a whole, without distinguishing between their component sub-communities.

Point, line and polygon (i.e. boundary) occurrence data were used to develop and test the distribution models. Polygons and lines were converted to sets of point data: this involved first converting them to raster format (using the same grid resolution as that of the predictor variables), and then converting the raster to a point shapefile (using the default filtering to one point per pixel). The resulting point shapefile was then merged with the other point dataset for the same habitat. Excluded from the coralligenous model were occurrence data located in the Po estuary (Italy), due to their unique and unusual habitat preferences. These are indeed a specific type of coralligenous outcrops called *tegnue*, defined as submerged rocky substrates of biogenic concretions, irregularly scattered in the sandy or muddy seabed, and containing extraordinary zoobenthic assemblages[22].

Model development. Data exploration was carried out on the coralligenous, maërl and predictor variable datasets, using R software (R Development Core Team; www.r-project.org), so as to detect potential outliers. Modelling techniques are often sensitive to multicollinearity among the predictor variables used. Available predictor variables (17) were hence iteratively tested for multicollinearity based on a combination of variance inflation factor (VIF < 2.5) and Spearman's rank correlation (r_s < 0.6). This resulted in a subset of 12 mostly uncorrelated predictor variables[54] (Table 3; Supplementary Figures S1 and S2), which were used as initial input to the models.

Maximum entropy, a well-known approach in machine-learning, is widely used to model species geographic distributions (i.e. their occurrence) in the terrestrial and marine environments, using, for instance, museum collections that only record occurrence localities[55,56]. The software Maxent[55,56] (version 3.3.3 k) was used to build models for coralligenous outcrops and maërl beds, starting with the subset of 12 predictor environmental variables. The algorithm used in Maxent aimed to find the

Table 3 | Predictor variables used in the models. dd: decimal degrees; psu: practical salinity units; n/a: not applicable. Maps are shown in Supplementary Figure S1

Name	Description	Source	Unit	Native resolution	Abbreviated name
Bathymetry	Depth of the seafloor	EMODNet Hydrography Portal	m	0.00416 dd	emodnet
Bottom salinity	Water salinity near the seafloor. Model output showing annual mean	World Ocean Database 2009[60]	psu	0.25 dd	botsalin
Bottom temperature	Water temperature near the seafloor. Model output showing annual mean	World Ocean Database 2009[60]	°C	0.25 dd	bottemp
Bottom type	Classed as 'hard' or 'soft'	[9]	n/a	1 km^2	bottype
Distance to ports	Euclidian distance (GIS calculation by the present study).	World Port Index 2011	dd	n/a	distport
Distance to (major) river mouths	Euclidian distance (GIS calculation by the present study).	River mouth location from ESRI Data and Maps (2012) and [49]	dd	n/a	distriver
Euphotic depth	Lee's euphotic depth. Annual mean from monthly climatologies (satellite telemetry). Prepared using MGET (Duke University)	Ocean Color Web	m	0.0416 dd	zeumean
Nutrient input	Based on mean annual use of fertilisers (from FAO statistics), and distributed across landscapes with dasymetric techniques. Transformed (0-1)	[9]	n/a	1 km^2	nutrient
Phosphate concentration	Collected *in situ* at the sea surface	[44]	µmol.L^{-1}	0.0833 dd	phosphate
Sea surface current	Absolute gridded geostrophic velocities of sea surface current showing annual mean from monthly climatologies (satellite telemetry)	Aviso SSALTO/DUACS	cm.sec^{-1}	0.0129 dd	sscmean
Silicate concentration	Collected *in situ* at the sea surface. Proxy for diatom phytoplankton growth	[44]	µmol.L^{-1}	0.0833 dd	silicate
Slope of the seafloor	GIS calculation by the present study	Based on bathymetry layer of EMODNet Hydrography Portal	%	n/a	gebslo

largest spread, or maximum entropy, in the geographic dataset composed of occurrence records of coralligenous outcrops or maërl beds, in relation to the 12 predictor variables. For each of the two models being developed, Maxent started with a uniform distribution of occurrence probability values for coralligenous outcrops or maërl beds over the entire Mediterranean basin, and conducted an optimisation routine that iteratively improved model fit, measured as the loss of entropy (i.e. the "gain" of information).

Available occurrence points for each habitat were split between 'training' and 'test' sets, the latter accounting for approximately a third of occurrences. The test set for each habitat model was geographically independent (Supplementary Figure S3) so as to avoid spatial autocorrelation between the test and training sets (which would occur if test points were selected randomly by Maxent). Test areas were selected so as to encompass, as much as possible, a variety of environmental conditions. The test set was not used for model development, but kept aside and fed separately to Maxent so as to assess model performance across the region.

Of the several feature types available in Maxent, hinge features were used so as to obtain smoother models and to help prevent overfitting to the training data[57]. In addition, Maxent's 'regularisation multiplier'[58] was tuned to 2.5 (the default value being 1), so as to reduce overfitting further and control model complexity.

So as to reduce the effect of the geographic sampling bias in the occurrence datasets for coralligenous and maerl habitats, 'target group background'[59] was used. Areas of the Mediterranean were attributed a relative value of sampling effort, based on expert knowledge (Supplementary Figure S4). This information was fed to Maxent in raster (ASCII) format.

Based on estimated relative contributions to the model by the 12 predictor variables, the ones contributing the least to the model were removed (i.e. usually if they contributed less than 5% and based on expert judgement), and the final model was re-run without them. The importance of each retained predictor variable was then measured through a jackknife (also called 'leave-one-out') test of variable importance, by training with each predictor variable first omitted, and then used in isolation. The model output was spatialised in the form of raster showing the logistic probability (ranging from 0 to 1) of occurrence for the habitat considered.

The Receiver Operating Characteristic (ROC) curve[39] was used to investigate the trade off between prediction sensitivity and specificity. The associated Area Under the Curve (AUC) is 0.5 in the case of random prediction, and higher values (to a maximum of 1) correspond to better performing models.

1. Rushton, S. P., Ormerod, S. J. & Kerby, G. New paradigms for modelling species distributions? *J. Appl. Ecol.* **41**, 193–200 (2004).
2. Guisan, A. & Thuiller, W. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009 (2005).
3. Cogan, C. B., Todd, B. J., Lawton, P. & Noji, T. T. The role of marine habitat mapping in ecosystem-based management. *ICES J. Mar. Sci.* **66**, 2033–2042 (2009).
4. Fraschetti, S. *et al.* Design of marine protected areas in a human-dominated seascape. *Mar. Ecol. Prog. Ser.* **375**, 13–24 (2009).
5. Moilanen, A., Wilson, K. & Possingham, H. *Spatial conservation prioritization: quantitative methods and computational tools.* (Oxford University Press, 2009).
6. Foley, M. M. *et al.* Guiding ecological principles for marine spatial planning. *Mar. Policy* **34**, 955–966 (2010).
7. Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography (Cop.).* **29**, 129–151 (2006).
8. Fraschetti, S., Terlizzi, A. & Boero, F. How many habitats are there in the sea (and where)? *J. Exp. Mar. Bio. Ecol.* **366**, 109–115 (2008).
9. Halpern, B. S. *et al.* A global map of human impact on marine ecosystems. *Science* **319**, 948–52 (2008).
10. Bianchi, C. N. & Morri, C. Marine Biodiversity of the Mediterranean Sea: Situation, Problems and Prospects for Future Research. *Mar. Pollut. Bull.* **40**, 367–376 (2000).
11. UNEP-MAP-RAC/SPA. Proceedings of the 1st Mediterranean symposium on the conservation of the coralligenous and other calcareous bio-concretions. (Tabarka, 15-16 January 2009. in Pergent-Martini, C. & Brichet, M.) (RAC/SPA, 2009).
12. UNEP-MAP-RAC/SPA. Proceedings of the 9th Meeting of the Focal Points for SPAs on the State of knowledge on the geographical distribution of marine magnoliophyta meadows in the Mediterranean. Floriana, Malta, 3-6 June 2009. in (Leonardini, R., Pergent, G. & Boudouresque, C.F., MED WG 331/Inf 5) (UNEP, 2009).
13. Coll, M. *et al.* The biodiversity of the Mediterranean Sea: estimates, patterns, and threats. *PLoS One* **5**, e11842 (2010).
14. Danovaro, R. *et al.* Deep-sea biodiversity in the Mediterranean Sea: the known, the unknown, and the unknowable. *PLoS One* **5**, e11832 (2010).
15. Giakoumi, S. *et al.* Ecoregion-Based Conservation Planning in the Mediterranean: Dealing with Large-Scale Heterogeneity. *PLoS One* **8**, e76449 (2013).
16. European Parliament and Council. Marine Strategy Framework Directive, 2008/56/EC. *Off. J. Eur. Union* **L164/19** (2008).
17. Ballesteros, E. Mediterranean coralligenous assemblages: a synthesis of present knowledge. *Oceanogr. Mar. Biol. Annu. Rev.* **44**, 123–195 (2006).

18. Bordehore, C., Ramos-Espla, A. A. & Riosmena-Rodriguez, R. Comparative study of two maerl beds with different otter trawling history, southeast Iberian Peninsula. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **13**, S43–S54 (2003).

19. Salomidi, M. *et al.* Assessment of goods and services, vulnerability, and conservation status of European seabed biotopes: a stepping stone towards ecosystem-based marine spatial management. *Mediterr. Mar. Sci.* **13**, 49–88 (2012).

20. Steller, D. L., Hernandez-Ayon, J. M., Riosmena-Rodriguez, R. & Cabello-Pasini, A. Effect of temperature on photosynthesis, growth and calcification rates of the free-living coralline alga *Lithophyllum margaritae*. *Ciencias Mar.* **33**, 441–456 (2007).

21. Piazzi, L., Gennaro, P. & Balata, D. Threats to macroalgal coralligenous assemblages in the Mediterranean Sea. *Mar. Pollut. Bull.* **64**, 2623–9 (2012).

22. Casellato, S. & Stefanon, A. Coralligenous habitat in the northern Adriatic Sea: an overview. *Mar. Ecol.* **29**, 321–341 (2008).

23. Savini, A., Basso, D., Alice Bracchi, V., Corselli, C. & Pennetta, M. Maërl-bed mapping and carbonate quantification on submerged terraces offshore the Cilento peninsula (Tyrrhenian Sea, Italy). *Geodiversitas* **34**, 77–98 (2012).

24. Martin, S., Charnoz, A. & Gattuso, J.-P. Photosynthesis, respiration and calcification in the Mediterranean crustose coralline alga *Lithophyllum cabiochae* (Corallinales, Rhodophyta). *Eur. J. Phycol.* **48**, 163–172 (2013).

25. Barbera, C. *et al.* Conservation and management of northeast Atlantic and Mediterranean maërl beds. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **13**, S65–S76 (2003).

26. UNEP-MAP-RAC/SPA. *Action plan for the conservation of the coralligenous and other calcareous bio-concretions in the Mediterranean Sea.* (Regional Activity Centre for Specially Protected Areas, ed.) 57 pp., (Tunis, Cedex, Tunisia, 2008).

27. Coma, R., Pola, E., Ribes, M. & Zabala, M. Long-term variation of temperate octocoral mortality patterns, protected vs. unprotected areas. *Ecol. Appl.* **14**, 1466–1478 (2004).

28. Cerrano, C. *et al.* A catastrophic mass-mortality episode of gorgonians and other organisms in the Ligurian Sea (North-western Mediterranean), summer 1999. *Ecol. Lett.* **3**, 284–293 (2000).

29. Pérès, J. M. & Picard, J. Nouveau manuel de bionomie benthique de la Mer Méditerranée. *Recl. Trav. Stn. Mar. Endoume* **31**, 1–137 (1964).

30. Garrabou, J. *et al.* Mass mortality in Northwestern Mediterranean rocky benthic communities: effects of the 2003 heat wave. *Glob. Chang. Biol.* **15**, 1090–1103 (2009).

31. Teixidó, N., Casas, E., Cebrian, E., Linares, C. & Garrabou, J. Impacts on coralligenous outcrop biodiversity of a dramatic coastal storm. *PLoS One* **8**, e53742 (2013).

32. Piazzi, L., Balata, D. & Cinelli, F. Invasions of alien macroalgae in Mediterranean coralligenous assemblages. *Cryptogam. Algol.* **28**, 289–301 (2007).

33. Klein, J. C. & Verlaque, M. Macroalgal assemblages of disturbed coastal detritic bottoms subject to invasive species. *Estuar. Coast. Shelf Sci.* **82**, 461–468 (2009).

34. Cebrian, E., Linares, C., Marschal, C. & Garrabou, J. Exploring the effects of invasive algae on the persistence of gorgonian populations. *Biol. Invasions* **14**, 2647–2656 (2012).

35. De Caralt, S. & Cebrian, E. Impact of an invasive alga (*Womersleyella setacea*) on sponge assemblages: compromising the viability of future populations. *Biol. Invasions* **15**, 1591–1600 (2013).

36. European Council. Council Directive 92/43/EEC on the conservation of natural habitats and of wild fauna and flora. *Off. J. Eur. Union* **1992L0043**, (2007).

37. Council of Europe. Resolution No. 4 (1996) listing endangered natural habitats requiring specific conservation measures. Revised (2010) Annex I of Resolution 4 (1996) of the Bern Convention on endangered natural habitat types using the EUNIS habitat classification. (1996). at https://wcd.coe.int/ViewDoc.jsp?id=1475213&Site= (date of access 4/03/2014).

38. European Council. Council Regulation (EC) No 1967/2006 concerning management measures for the sustainable exploitation of fishery resources in the Mediterranean Sea. *Off. J. Eur. Union2* **L404/11**, (2006).

39. Guisan, A. & Zimmermann, N. E. Predictive habitat distribution models in ecology. *Ecol. Model.* **135**, 147–186 (2000).

40. Franklin, J. *Mapping species distributions.* (Cambridge University Press, ed.) 320 pp. (The Edinburgh Building, Cambridge, 2009).

41. Giannoulaki, M. *et al. Mediterranean Sensitive Habitats (MEDISEH), final project report. DG MARE Specific Contract SI2.600741.* 557 pp. (Hellenic Centre for Marine Research, 2013).

42. Tittensor, D. P. *et al.* Predicting global habitat suitability for stony corals on seamounts. *J. Biogeogr.* **36**, 1111–1128 (2009).

43. Davies, A. J. & Guinotte, J. M. Global habitat suitability for framework-forming cold-water corals. *PLoS One* **6**, e18483 (2011).

44. Tyberghein, L. *et al.* Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Glob. Ecol. Biogeogr.* **21**, 272–281 (2012).

45. Bonacorsi, M. *et al.* Cartography of main coastal ecosystems in a NATURA 2000 site. Paper presented at *Proceeding on the Tenth International Conference The Mediterranean Coastal Environment, Medcoast 11, 25-29 Oct. 2011, Rhodes, Greece* 253–262 (Medcoast, 2011).

46. Pearson, R. G. & Dawson, T. P. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* **12**, 361 (2003).

47. Picard, J. Recherches qualitatives sur les biocoenoses marines des substrats meubles dragables de la région marseillaise. *Recl. Trav. Stn. Mar. Endoume* **36**, 1–160 (1965).

48. Cameron, A. & Askew, N. *EUSeaMap - Preparatory action for development and assessment of a European broad-scale seabed habitat map final report. EC contract no. MARE/2008/07.* 240 pp. (JNCC, 2011).

49. Wessel, P. & Smith, W. H. F. A global, self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res.* **101**, 8741–8743 (1996).

50. Ballesteros, E. The deep-water *Peyssonnelia* beds from the Balearic Islands (Western Mediterranean). *Mar. Ecol.* **15**, 233–253 (1994).

51. Kipson, S. *et al.* Rapid biodiversity assessment and monitoring method for highly diverse benthic communities: a case study of mediterranean coralligenous outcrops. *PLoS One* **6**, e27103 (2011).

52. Gatti, G. *et al.* Seafloor integrity down the harbor waterfront: the coralligenous shoals off Vado Ligure (NW Mediterranean). *Adv. Oceanogr. Limnol.* **3**, 51–67 (2012).

53. Vaz, S., Carpentier, A. & Coppin, F. Eastern English Channel fish assemblages: measuring the structuring effect of habitats on distinct sub-communities. *ICES J. Mar. Sci.* **64**, 1–17 (2007).

54. Zuur, A. F., Ieno, E. N. & Smith, G. M. *Analysing ecological data.* (Springer, ed.) 672 pp. (New York, NY 10013, USA, 2007).

55. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modelling of species geographic distributions. *Ecol. Modell.* **190**, 231–259 (2006).

56. Phillips, S. J. & Dudı, M. Modelling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography (Cop.).* **31**, 161–175 (2008).

57. Elith, J. *et al.* A statistical explanation of MaxEnt for ecologists. *Divers. Distrib.* **17**, 43–57 (2011).

58. Elith, J., Kearney, M. & Phillips, S. The art of modelling range-shifting species. *Methods Ecol. Evol.* **1**, 330–342 (2010).

59. Phillips, S. J. *et al.* Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol. Appl.* **19**, 181–197 (2009).

60. Boyer, T. *et al.* Objective analyses of annual, seasonal, and monthly temperature and salinity for the World Ocean on a 0.25° grid. *Int. J. Climatol.* **25**, 931–945 (2005).

Acknowledgments

Author contributions

C.S.M. and S.F. led the writing of the manuscript, with significant contributions from M.Gr., G.Ga., M.Sc., M.Sa., L.K., M.L.P., A.B., V.G., G.M., E.B., E.C., P.J.S., C.P.M., G.P., G.B. and V.V. S.F. coordinated the collaborative effort of collecting spatial data on the two habitats. S.F. and F.D.L. reviewed data sources and combined all spatial datasets into one. C.S.M., M.Sc. and M.Gi. carried out the predictive modelling. K.T., L.R., J.B.S., M.B., G.Gu., M.K., V.M. and E.P. reviewed the manuscript.

Additional information