Received: 26 July 2017 | Revised: 13 October 2017 | Accepted: 26 October 2017

DOI: 10.1111/geb.12693

DATA PAPER

WILEY Global Ecology and Biogeography A Journal of Macroecology

# Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling

Jorge Assis[1] | Lennert Tyberghein[2] | Samuel Bosch[2,3] | Heroen Verbruggen[4] | Ester A. Serrão[1] | Olivier De Clerck[3]

[1]Centre for Marine Sciences, CCMAR-CIMAR, University of Algarve, Faro, Portugal

[2]Flanders Marine Institute (VLIZ), InnovOcean site, Ostend, Belgium

[3]Phycology Research Group, Biology Department, Ghent University, Ghent, Belgium

[4]School of BioSciences, University of Melbourne, Melbourne, Victoria, Australia

**Correspondence**
Jorge Assis, Centre for Marine Sciences, CCMAR-CIMAR, University of Algarve, Campus Gambelas, 8005-139 Faro, Portugal.
Email: jorgemfa@gmail.com

**Funding information**
Pew Foundation; Foundation for Science and Technology (FCT), Grant/Award Number: SFRH/BPD/111003/2015, PTDC/MAR-EST/6053/2014, BIODIVERSA/004/2015 and CCMAR/Multi/04326/2013; EU FP7 ERANET, Grant/Award Number: SEAS-ERA/INVASIVES SD/ER/010; Australian Research Council, Grant/Award Number: FT110100585

Editor: Derek Tittensor

## Abstract

**Motivation:** The availability of user-friendly, high-resolution global environmental datasets is crucial for bioclimatic modelling. For terrestrial environments, WorldClim has served this purpose since 2005, but equivalent marine data only became available in 2012, with pioneer initiatives like Bio-ORACLE providing data layers for several ecologically relevant variables. Currently, the available marine data packages have not yet been updated to the most recent Intergovernmental Panel on Climate Change (IPCC) predictions nor to present times, and are mostly restricted to the top surface layer of the oceans, precluding the modelling of a large fraction of the benthic diversity that inhabits deeper habitats. To address this gap, we present a significant update of Bio-ORACLE for new future climate scenarios, present-day conditions and benthic layers (near sea bottom). The reliability of data layers was assessed using a cross-validation framework against in situ quality-controlled data. This test showed a generally good agreement between our data layers and the global climatic patterns. We also provide a package of functions in the R software environment (*sdmpredictors*) to facilitate listing, extraction and management of data layers and allow easy integration with the available pipelines for bioclimatic modelling.

**Main types of variable contained:** Surface and benthic layers for water temperature, salinity, nutrients, chlorophyll, sea ice, current velocity, phytoplankton, primary productivity, iron and light at bottom.

**Spatial location and grain:** Global at 5 arcmin (*c*. 0.08° or 9.2 km at the equator).

**Time period and grain:** Present (2000–2014) and future (2040–2050 and 2090–2100) environmental conditions based on monthly averages.

**Major taxa and level of measurement:** Marine biodiversity associated with sea surface and epibenthic habitats.

**Software format:** ASCII and TIFF grid formats for geographical information systems and a package of functions developed for R software.

KEYWORDS
Bio-ORACLE, bioclimatic modelling, environmental data, global, kriging, macroecology, marine, species distribution modelling

## 1 | INTRODUCTION

Early attempts to model the relationship between the occurrence or abundance of species and their natural environment relied heavily on environmental variables measured in situ (Sutherst & Maywald, 1985) and often involved complex software pipelines specifically developed to extract, organize and visualize data (Kemp, Loon, Shamoun-Baranes, & Bouten, 2012; Lima-Ribeiro et al., 2015). The spatial and temporal resolution of environmental data also showed high variability, precluding smooth integration and comparison of bioclimatic analyses (Lima-Ribeiro et al., 2015). Alongside the development of geographical information systems (GIS), the advent of cutting-edge spatial interpolation resulted in data layers representing global environmental conditions, conformal in extent and resolution. Pioneer initiatives, such as the Climatic Research Unit Terrestrial Climatology (New, Hulme, & Jones, 1999) and WorldClim (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005a), significantly pushed the application of bioclimatic modelling in ecology, biogeography, conservation biology and evolution. Yet, these gridded datasets were tailored for terrestrial climates only, and the availability of marine data layers lagged significantly behind (Robinson et al., 2011). National Oceanic and Atmospheric Administration's World Ocean Atlas (Levitus, 2001), AquaMaps (Kaschner et al., 2008) and Hexacoral (Fautin & Buddemeier, 2008) were only recently enhanced by the more comprehensive and higher resolution datasets Bio-ORACLE (Tyberghein et al., 2012) and Marspec (Sbrocco & Barber, 2013).

The increased accessibility of marine data layers allowed an emerging body of research to describe the global distribution of species (e.g., Chaudhary, Saeedi, & Costello, 2017; Chefaoui, Assis, Duarte, & Serrão, 2015; Hill & Terblanche, 2014; Parravicini et al., 2013; Stuart-Smith, Edgar, Barrett, Kininmonth, & Bates, 2015), address niche-based questions (e.g., Assis et al., 2015; Lee-Yaw et al., 2016; Verbruggen et al., 2009), support biodiversity conservation (e.g., Boavida, Assis, Silva, & Serrão, 2016; Guisan & Thuiller, 2005; Guisan et al., 2013) and ecosystem-based management (Hobday, Hartog, Timmiss, & Fielding, 2010) and infer the likely anthropogenic pressures leading to population turnover and extinction (e.g., Scherner et al., 2013). The establishment of standard protocols (e.g., Coupled Model Intercomparison Project; CMIP) delivering the outputs of atmosphere–ocean general circulation models (AOGCMs) for past and future climate scenarios (Moss et al., 2010; Otto-Bliesner et al., 2009) further expanded the applications for marine data layers, for instance, to predict range shifts through time (e.g., Assis, Berecibar et al., 2017; Burrows et al., 2014; Neiva et al., 2015; Thomas et al., 2004) or test relevant evolutionary hypotheses such as the location of marine biodiversity hotspots free from past bottlenecks and extinctions (i.e., climatic refugia; Assis, Serrão, Claro, Perrin, & Pearson, 2014; Waltari et al., 2007).

The marine datasets currently available, however, are almost exclusively restricted to the top surface layer of the oceans (e.g., Bio-ORACLE), and those including benthic layers adjacent to the seabed are particularly coarse in resolution (Hexacoral to *c*. 56 $km^2$ and World Ocean Atlas to *c*. 112 $km^2$, at the equator) or limited to biophysical features extracted from bathymetric profiles (MARSPEC; Sbrocco & Barber, 2013). These constraints significantly limit the potential for modelling benthic species (Boavida et al., 2016; Davies & Guinotte, 2011; Reiss et al., 2014), which include a large proportion of marine biodiversity. For instance, the exploration of deep cryptic refugia for marine species is suboptimal when using surface data only (Assis et al., 2016; Graham, Kinlan, Druehl, Garske, & Banks, 2007; Perry, Low, Ellis, & Reynolds, 2005).

To address this gap, we present a significant extension of the marine data layers available in Bio-ORACLE. New ecologically relevant surface and benthic layers tailored for mechanistic and correlative modelling (Kearney & Porter, 2009; Peterson et al., 2011) are provided for present conditions and the new generation of climate change scenarios (Moss et al., 2010). Besides the extension of Bio-ORACLE to include benthic layers for temperature, salinity, nutrients and chlorophyll (Table 1), we also provide new data on sea ice, current velocity, phytoplankton, primary productivity, iron and light at the bottom for a better understanding of marine macroecological processes. We also determine the reliability of data layers (as stressed by Hall & Hall, 2014) using a cross-validation framework against in situ quality-controlled data. We provide a package of functions in the R software environment (R Development Core Team, 2016) for easy integration with the available pipelines for bioclimatic modelling (e.g., Naimi & Araújo, 2016; Thuiller, Lafourcade, Engler, & Araújo, 2009).

## 2 | MARINE DATA LAYERS

Marine data layers for present conditions were produced with climate data describing monthly averages for the period 2000–2014, obtained from pre-processed global ocean re-analyses combining satellite and in situ observations at regular two- and three-dimensional spatial grids (Table 1). Future layers were produced for 2040–2050 and 2090–2100 by averaging data from distinct AOGCMs provided by the CMIP 5 (Table 2). The available data (temperature, salinity, current velocity and sea ice thickness) were obtained for the new representative concentration pathway scenarios (RCP): the RCP26, a peak-and-decline scenario ending in very low greenhouse gas concentration levels by the end of the 21th century; the RCP45 and RCP60, in which levels stabilize; and the RCP85, a scenario of increasing emissions over time, leading to high greenhouse gas concentration levels (reviewed by Moss et al., 2010).

The monthly averages for the present and future were used to produce six distinct predictors per variable for bioclimatic modelling: the long-term average; the minimum and maximum records; the long-term average of the minimum and maximum records per year (e.g., temperature of the warmest month, on average); and range, given by the average of the absolute difference between the minimum and maximum records per year. These predictors were statistically downscaled (i.e., from coarse- to fine-scale resolution) to a common spatial resolution of 5 arcmin (*c*. 0.08° or 9.2 km at the equator) by fitting a kriging model based on the 12 nearest values of each focal cell (e.g., Hofstra, Haylock, New, Jones, & Frei, 2008; Lima-Ribeiro et al., 2015). The choice of kriging over other interpolation methods was based on studies showing higher performance for this method (e.g., Hofstra et al., 2008; Lima-Ribeiro et al., 2015) and also on a priori testing performed against inverse distance weighting (IDW; e.g., Assis et al., 2014; Kemp et al., 2012). This test cross-validated the interpolation of $1 \times 10^4$ random records with both methods for different variables, and showed lower root mean square error (RMSE) and mean absolute error (MAE) for kriging, despite the lack of differences in the mean value of all

**TABLE 1** Marine data layers, units, correspondence with the first version of Bio-ORACLE, range of values (determined for benthic layers, at their average depth), accuracy assessed with quality-control data, number of quality-control records (*n*) and source of climate data

| Layer | Units | BO1 | Range | MAE | RMSE | Cor | *n* | Source |
|---|---|---|---|---|---|---|---|---|
| Temperature | °C | Yes | [−1.94;39.22] | 0.39 | 0.75 | 0.99 | 445,248 | ARMOR |
| Salinity | PSS | Yes | [4.75;41.96] | 0.13 | 0.52 | 0.93 | 444,925 | ARMOR |
| Sea ice concentration | Fraction | No | [0;1] | – | – | – | – | ORAP |
| Sea ice thickness | m | No | [0;10.94] | – | – | – | – | ORAP |
| Current velocity | $m \cdot s^{-1}$ | No | [0;2.42] | – | – | – | – | ORAP |
| Nitrate | $mmol \cdot m^{-3}$ | Yes | [0;164.51] | 1.62 | 2.47 | 0.98 | 93,201 | PISCES |
| Phosphate | $mmol \cdot m^{-3}$ | Yes | [0;3.55] | 0.15 | 0.23 | 0.97 | 349,074 | PISCES |
| Silicate | $mmol \cdot m^{-3}$ | Yes | [0.46;316.67] | 5.98 | 9.05 | 0.99 | 367,629 | PISCES |
| Dissolved molecular oxygen | $mmol \cdot m^{-3}$ | Yes | [0;789.94] | 15.27 | 23.50 | 0.96 | 417,790 | PISCES |
| Dissolved iron | $mmol \cdot m^{-3}$ | No | [0;0.03] | – | – | – | – | PISCES |
| Chlorophyll | $mg \cdot m^{-3}$ | Yes | [0;17.46] | 0.1 | 0.15 | 0.56 | 5,401 | PISCES |
| Phytoplankton* | $mmol \cdot m^{-3}$ | No | [0;44.94] | – | – | – | – | PISCES |
| Primary productivity* | $g \cdot m^{-3} \cdot day^{-1}$ | No | [0;0.95] | – | – | – | – | PISCES |
| Light at the bottom | $E \cdot m^{-2} \cdot yr^{-1}$ | No | [0;69.21] | – | – | – | – | GlobColour |

ARMOR = Global Observed Ocean Physics Reprocessing (resolution: 0.25°/33 vertical levels); BO1 = first version of Bio-ORACLE; Cor = Pearson's correlation coefficient; GlobColour merging MERIS/MODIS/SeaWiFS (resolution: 0.05°/surface); MAE = mean average error; ORAP = Global Ocean Physics Reanalysis ECMWF (resolution: 0.25°/75 vertical levels); PISCES = Global Ocean Biogeochemistry Non-assimilative Hindcast (resolution: 0.25°/75 vertical levels); RMSE = root mean square error. *Expressed as carbon in sea water
[Correction added on 10 January 2018, after first online publication: Units have been corrected in this version]

variables (nonparametric Kruskal–Wallis *p*-values > .05; see Supporting Information Table S1.1).

The downscaling process for benthic layers considered the geographical position and depth of cells (e.g., Assis et al., 2016; Boavida et al., 2016) as inferred from the general bathymetric chart of the oceans (GEBCO, 2015). Given that focal cells included a range of depth values, the benthic layers were produced for the minimum, average and maximum depths. The future layers were downscaled using the change-factor approach (Lima-Ribeiro et al., 2015; Varela, Lima-Ribeiro, & Terribile, 2015; Wilby et al., 2004). This technique is based on applying the predicted magnitude of climate change to the data layers produced for the present. For this purpose, data for the period 2000–2014 were also obtained from the AOGCMs to determine the difference (change-factor) between the present conditions and the future scenarios of change, at the native resolution of each AOGCM (Table 2). Next, the change-factor was downscaled to 0.08° resolution with kriging (as previously described) and applied to the corresponding baseline layer for the present conditions.

The layers for current velocity and light at the bottom were further post-processed. The current velocity was determined with the Pythagoras theorem on the meridional (along the longitude circle) and zonal (along the latitude circle) components of ocean currents, whereas light at the bottom used a standard depth-dependent exponential function (Assis et al., 2016; Graham et al., 2007) based on photosynthetically active radiation (PAR) and diffuse attenuation coefficient (Kd490):

$$\text{Light at bottom} = \text{PAR} \times \exp(-\text{Kd490} \times z),$$

where $z$ is depth inferred from the general bathymetric chart of the oceans (GEBCO, 2015).

**TABLE 2** Average and range of climatic anomalies for the future (2090–2100) under different scenarios of change, and source of data (coupled atmosphere–ocean general circulation models)

| Layer | RCP26 | RCP45 | RCP60 | RCP85 | Source (AOGCM) |
|---|---|---|---|---|---|
| Temperature | 0.21 [−2.06;4.45] | 0.29 [−2.01;6.72] | 0.33 [−1.24;6.89] | 0.51 [−2.62;10.98] | CCSM4, HadGEM2-ES, MIROC5 |
| Salinity | −0.03 [−6.03;1.37] | −0.04 [−4.97;1.39] | −0.05 [−4.48;1.34] | −0.07 [−4.33;1.83] | CCSM4, HadGEM2-ES, MIROC5 |
| Current velocity | 0.01 [−0.17;0.18] | 0.01 [−0.11;0.28] | 0.01 [−0.13;0.16] | 0.01 [−0.18;0.24] | CCSM4, HadGEM2-ES, MIROC5 |
| Sea ice thickness | −0.12 [−8.33;1.43] | −0.18 [−8.78;1.98] | −0.22 [−8.95;1.54] | −0.27 [−9.01;1.88] | CCSM4, HadGEM2-ES, MIROC5 |

*Note*. Climatic anomalies were inferred by determining the difference between the benthic layers (average depth) produced for the different representative concentration pathway (RCP) scenarios and those for the present. CCSM4 = The Community Climate System Model 4 (resolution: 1.13° × 0.47°/60 vertical levels); HadGEM2-ES = Hadley Centre Global Environmental Model 2 (resolution: 1.00° × 0.84°/40 vertical levels); MIROC5 = Model for Interdisciplinary Research on Climate 5 (resolution: 1.41° × 0.81°/50 vertical levels).

All layers were exported to ASCII (ESRI) and TIFF grid formats for easy downloading and integration in modern GIS technologies (e.g., ESRI and QGIS). To minimize the possible spatial gap with 'no data' between the data layers provided and the available vector shorelines, the global self-consistent, hierarchical, high-resolution geography database (Wessel & Smith, 1996) is recommended.

## 3 | RELIABILITY OF MARINE DATA LAYERS

The reliability of downscaled data layers was inferred with in situ quality-controlled data provided by the Global Ocean Data Analysis Project (GLODAP; Olsen et al., 2016) for temperature, salinity, phosphate, nitrate, silicate, dissolved oxygen and chlorophyll. This was performed by cross-validating the outputs of downscaling the raw data used to produce data layers at the locations (geographical position and depth) of each sample available in GLODAP, with the actual data provided by this dataset (e.g., Assis, Berecibar et al., 2017; Boavida et al., 2016; Davies & Guinotte, 2011). The paired relationships between the interpolated and in situ data were statistically analysed with mean absolute error, root mean square error and Pearson's correlation. These tests showed the layers mirroring most climatic patterns present in quality-controlled data. All variables retrieved low error rates (Table 1) and high correlation coefficients (Pearson's correlation: $>.93$; Table 1) with a unique exception for chlorophyll, which showed a much lower correlation (Pearson's correlation: .56; Table 1; see Supporting Information Figure S1.1). This exception is likely to result from coupling the high temporal variability known for chlorophyll on a daily basis (e.g., Iriarte, González, Liu, Rivas, & Valenzuela, 2007; Wang, Le Borgne, Murtugudde, Busalacchi, & Behrenfeld, 2009; Wang Hladik et al., 2010) with the monthly averages used to produce the layers (as discussed by Boavida et al., 2016; Davies & Guinotte, 2011). Moreover, chlorophyll was the variable with the lowest number of in situ observations (5,401 for the global ocean; Table 1), a fact that might have precluded a proper assessment of reliability in cross-validation.

The spatial distribution error of the layers was illustrated by mapping the difference between the interpolated and in situ data onto a 2.5° grid (e.g., Davies & Guinotte, 2011) and by plotting this difference against depth. In general, the distribution of errors also showed high accuracy for all layers. Temperature, phosphate, nitrate and dissolved molecular oxygen displayed only specific anomalies, highly restricted to discrete regions of the global ocean (e.g., east Siberian Sea and southern North Sea; Supporting Information Figures S1.4, S1.8, S1.10 and S1.14), and with no relationship with depth (Supporting Information Figures S1.3, S1.7, S1.9 and S1.13). The errors for silicate were mostly in the Southern Ocean (Supporting Information Figure S1.12) and those for salinity were in the top layers (surface waters with higher positive anomaly; Supporting Information Figure S1.5) of the Canadian Arctic and east Siberian Sea (Supporting Information Figure S1.6), particularly for values below 30 PSS (Supporting Information Figure S1.5). The spatial distribution of errors also showed that dissolved molecular oxygen, phosphate, salinity, silicate and temperature have good coverage of in situ samples (GLODAP dataset), whereas nitrate and chlorophyll are mostly uncovered throughout the globe.

## 4 | R-PACKAGE TOOL

In addition to providing the layers for downloading in ASCII and TIFF formats, we also developed the *sdmpredictors* package of functions in R (R Development Core Team, 2016) to facilitate listing, extraction and management of data layers. This package, whose functions are detailed in Table 3, also integrates the layers from the first version of Bio-ORACLE, as well as those of MARSPEC (Sbrocco & Barber, 2013) and BioClim (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005b). The source code and related help files are available via the CRAN repository and

**TABLE 3** List of functions available in *sdmpredictors* package, description and arguments that users may specify

| Function | Description | Arguments |
|---|---|---|
| list_datasets() | Explore datasets available in the package | Terrestrial (logical) only terrestrial data are returned; marine (logical) only marine data are returned |
| list_layers() | Explore layers in a dataset | Datasets (character); terrestrial (logical); marine (logical); monthly (logical) only annual and monthly layers are returned (default) |
| list_layers_future() | Explore future layers in the package | Datasets (character); scenario (character) for climate change scenario (e.g., RCP85); year (integer) for the climate change prediction (e.g., 2100); terrestrial (logical); marine (logical); monthly (logical) |
| get_future_layers() | Get the name of a future climate layer(s) based on the current climate layer(s) | Current_layer_codes (character) with the code(s) of the layers either as a vector or dataframe provided by list_layers); scenario (character); year (integer) |
| load_layers() | Download specific layers to the current directory | Layercodes (character) with the codes of the layers to be loaded; equalarea (logical) for Behrmann cylindrical equal-area projection; rasterstack (logical) to stack the layers in a unique object; datadir (character) for the directory to store data |
| layer_stats() | Layer statistics | Layercodes (character) |
| layers_correlation() | Pearson's correlation coefficient between layers | Layercodes (character) |

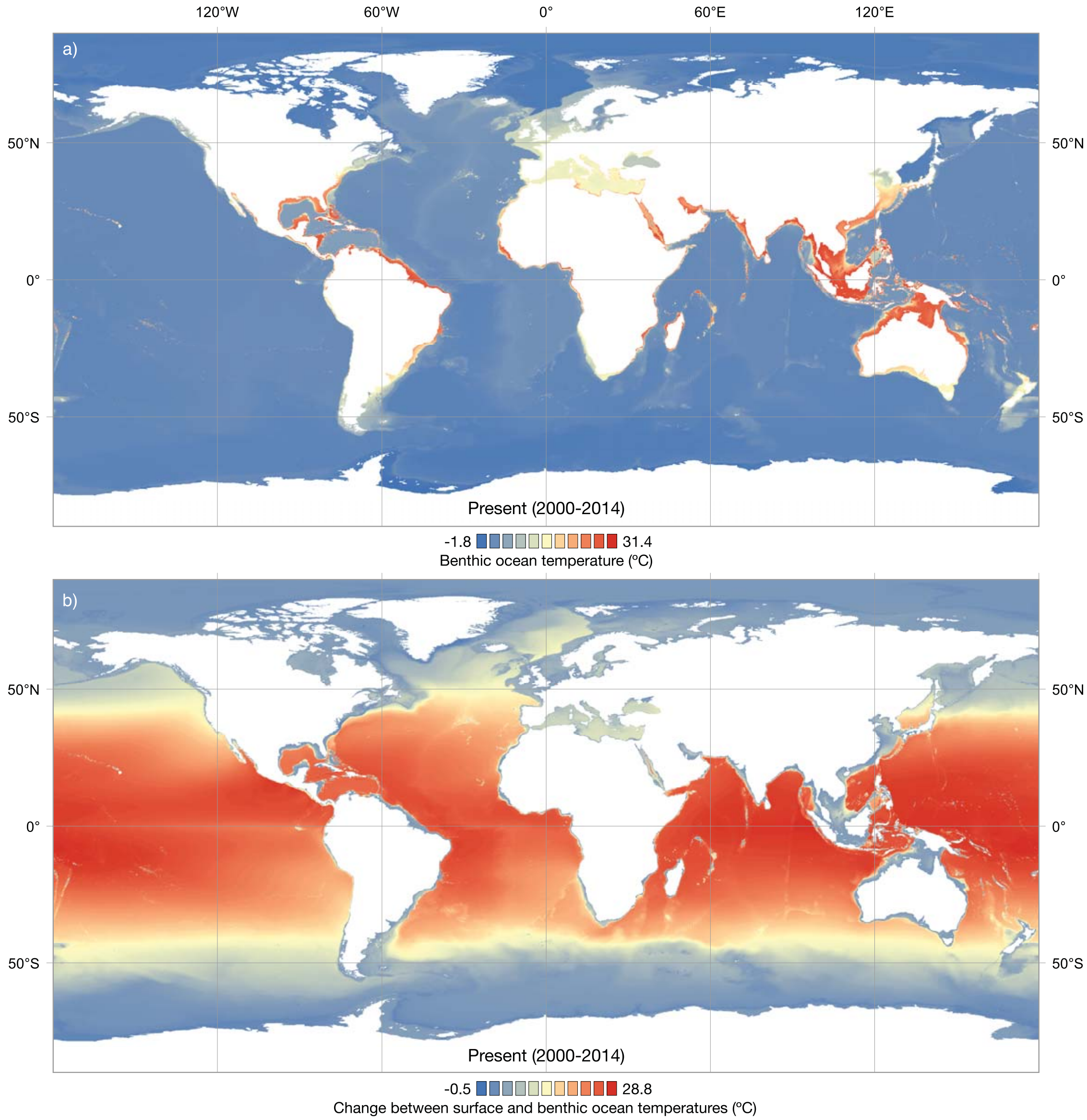


**FIGURE 1** (a) Mean benthic ocean temperature for the present (period 2000–2014) and (b) change (difference) between surface and benthic ocean temperatures

can be installed easily by entering the following lines of code into the R command prompt:

1. install.packages('sdmpredictors').
2. library(sdmpredictors).

## 5 | CONCLUSION

A comprehensive set of new ecologically relevant data layers are provided for bioclimatic modelling in a user-friendly format, with global coverage and comparable grid system. The current update expands the potential of Bio-ORACLE by (a) covering the benthic realm with climate and environmental data, (b) adding new variables to present-day conditions (i.e., current velocity, iron, light, phytoplankton, primary productivity and sea ice), (c) adding the new generation of climate change scenarios, (d) providing means of data reliability and uncertainty, and (e) providing a package of functions in the R software environment.

The relevance of new data aiming for species associated with sea benthic features (e.g., Assis et al., 2016; Boavida et al., 2016) is clearly underlined by the disparity in ocean temperatures between surface and benthic layers, which can amount to up to 28.8 °C in the deeper regions of lower latitudes (Figure 1). The data provided for the new generation of climate change scenarios further diversifies the range of scientific questions that can be addressed using Bio-ORACLE. One such case is the possibility to explore climate-induced depth range shifts (e.g., Assis et al., 2016; Assis, Araújo, & Serrão, 2017), the marine equivalent to elevation range shifts for terrestrial species (Chen, Hill, Ohlemüller, Roy, & Thomas, 2011; Galbreath, Hafner, & Zamudio, 2009). The reliability of new layers has been assessed with in situ quality-controlled data. This information is relevant for marine

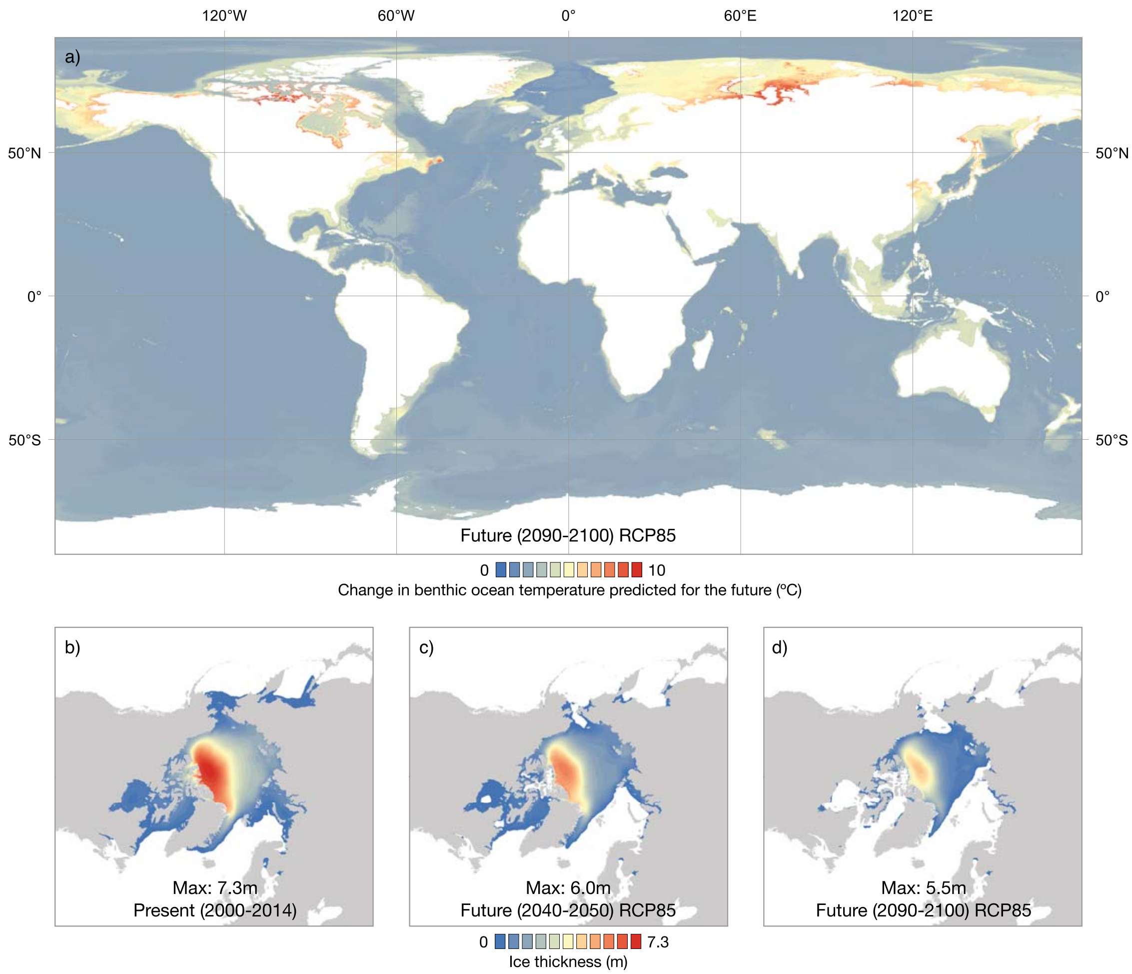


**FIGURE 2** (a) Change in the maximal benthic ocean temperature predicted for the period 2090–2100 with RCP85. (b–d) Arctic mean ice thickness for the periods (b) 2000–2014, (c) 2040–2050 and (d) 2090–2100 predicted with RCP85

modellers and showed a good agreement between data layers and the global climatic patterns. The future data using the magnitude of climate changes (change-factor) of different AOGCMs provide the highest confidence level currently attainable (Hall & Hall, 2014). Despite their inherent uncertainties, the AOGCMs used represent the present scientific understanding linking greenhouse gas emissions with global climate changes.

The familiar data structure of Bio-ORACLE (rasters for GIS) and its integration with R computing language should allow easy acquisition, exploration and manipulation of data, as well as smooth integration with the available statistical tools (e.g., Naimi & Araújo, 2016; Thuiller et al., 2009). The new data layers represent a valuable addition to the spatial information about climate available for the global ocean. Their key features can be used to improve bioclimatic modelling and provide valuable insights into the current and future states of marine biodiversity and indirectly into the services it provides to society. When requested by decision-makers, these outcomes may guide important climate change-integrated conservation strategies (Hannah, Midgley, & Millar, 2002; Hobday et al., 2010) and feed baseline assessments, such as those required for the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES).

## ACKNOWLEDGMENTS

This study was supported by the Pew Foundation (E.A.S.) and the Foundation for Science and Technology (FCT) of Portugal through a fellowship to Jorge Assis (SFRH/BPD/111003/2015) and projects PTDC/MAR-EST/6053/2014, BIODIVERSA/004/2015 and CCMAR/Multi/04326/2013. S.B. and O.D.C. are indebted to EU FP7 ERANET (Project SEAS-ERA/INVASIVES SD/ER/010). H.V. was supported by the Australian Research Council (FT110100585). The datasets ARMOR, ORAP and PISCES were made available by the E.U. Copernicus Marine Service Information (http://marine.copernicus.eu). The GlobColour dataset (http://globcolour.info) used in this study has been developed, validated and distributed by ACRI-ST, France. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups (listed in Table 2 of this paper) for producing and making available

their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

## DATA ACCESSIBILITY

The Bio-ORACLE layers are accessible online at http://www.bio-oracle.org for download (as ASCII and TIFF files) and preview.

## ORCID

*Jorge Assis* http://orcid.org/0000-0002-6624-4820

## REFERENCES

Assis, J., Araújo, M. B., & Serrão, E. A. (2017). Projected climate changes threaten ancient refugia of kelp forests in the North Atlantic. *Global Change Biology* . https://doi.org/10.1111/gcb.13818.

Assis, J., Berecibar, E., Claro, B., Alberto, F., Reed, D., Raimondi, P., & Serrão, E. A. (2017). Major shifts at the range edge of marine forests: The combined effects of climate changes and limited dispersal. *Scientific Reports*, *7* , 44348.

Assis, J., Coelho, N. C., Lamy, T., Valero, M., Alberto, F., & Serrão, E. A. (2016). Deep reefs are climatic refugia for genetic diversity of marine forests. *Journal of Biogeography*, *43*, 833–844.

Assis, J., Serrão, E. A., Claro, B., Perrin, C., & Pearson, G. A. (2014). Climate-driven range shifts explain the distribution of extant gene pools and predict future loss of unique lineages in a marine brown alga. *Molecular Ecology*, *23*, 2797–2810.

Assis, J., Zupan, M., Nicastro, K. R., Zardi, G. I., McQuaid, C. D., & Serrão, E. A. (2015). Oceanographic conditions limit the spread of a marine invader along Southern African shores. *PLoS One*, *10*, e0128124.

Boavida, J., Assis, J., Silva, I., & Serrão, E. A. (2016). Overlooked habitat of a vulnerable gorgonian revealed in the Mediterranean and Eastern Atlantic by ecological niche modelling. *Scientific Reports*, *6*, 36460.

Burrows, M. T., Schoeman, D. S., Richardson, A. J., Molinos, J. G., Hoffmann, A., Buckley, L. B., . . . Poloczanska, E. S. (2014). Geographical limits to species-range shifts are suggested by climate velocity. *Nature*, *507*, 492–495.

Chaudhary, C., Saeedi, H., & Costello, M. J. (2017). Bimodality of latitudinal gradients in marine species richness. *Trends in Ecology and Evolution*, *31*, 670–676.

Chefaoui, R. M., Assis, J., Duarte, C. M., & Serrão, E. A. (2015). Large-scale prediction of seagrass distribution integrating landscape metrics and environmental factors: The case of *Cymodocea nodosa* (Mediterranean–Atlantic). *Estuaries and Coasts*, 123–137.

Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, *333*, 1024–1026.

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

Fautin, D. G., & Buddemeier, R. W. (2008). *Biogeoinformatics of the hexacorals*. Retrieved from http://www.kgs.ku.edu/Hexacoral/

Galbreath, K. E., Hafner, D. J., & Zamudio, K. R. (2009). When cold is better: Climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution*, *63*, 2848–2863.

GEBCO. (2015). *General bathymetric chart of the oceans*. In B. O. D. Centre (Ed.), Liverpool, UK: British Oceanographic Data Centre. Retrieved from http://www.gebco.net

Graham, M. H., Kinlan, B. P., Druehl, L. D., Garske, L. E., & Banks, S. (2007). Deep-water kelp refugia as potential hotspots of tropical marine diversity and productivity. *Proceedings of the National Academy of Sciences USA*, *104*, 16576–16580.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, *8*, 993–1009.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., . . . Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*, 1424–1435.

Hall, B. A., & Hall, A. (2014). Projecting regional change. *Science*, *346*, 1461–1462.

Hannah, L., Midgley, G. F., & Millar, D. (2002). Climate change-integrated conservation strategies. *Global Ecology and Biogeography*, *11*, 485–495.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005a). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005b). WORLDCLIM – a set of global climate layers (climate grids). *International Journal of Climatology*, *25*, 1965–1978.

Hill, M. P., & Terblanche, J. S. (2014). Niche overlap of congeneric invaders supports a single-species hypothesis and provides insight into future invasion risk: Implications for global management of the *Bactrocera dorsalis* complex. *PLos One*, *9*, e90121.

Hobday, A. J., Hartog, J. R., Timmiss, T., & Fielding, J. (2010). Dynamic spatial zoning to manage southern bluefin tuna (*Thunnus maccoyii*) capture in a multi-species longline fishery. *Fisheries Oceanography*, *19*, 243–253.

Hofstra, N., Haylock, M., New, M., Jones, P., & Frei, C. (2008). Comparison of six methods for the interpolation of daily, European climate data. *Journal of Geophysical Research Atmospheres*, *113*, 1–19.

Iriarte, J. L., González, H. E., Liu, K. K., Rivas, C., & Valenzuela, C. (2007). Spatial and temporal variability of chlorophyll and primary productivity in surface waters of southern Chile (41.5–43° S). *Estuarine, Coastal and Shelf Science*, *74*, 471–480.

Kaschner, K., Kesner-Reyes, K., Garilao, C., Rius-Barile, J., Rees, T., & Froese, R. (2008). *AquaMaps: Predicted range maps for aquatic species*. Retrieved from http://www.aquamaps.org

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology Letters*, *12*, 334–350.

Kemp, M. U., Loon, E. E. V., Shamoun-Baranes, J., & Bouten, W. (2012). RNCEP: Global weather and climate data at your fingertips. *Methods in Ecology and Evolution*, *3*, 65–70.

Lee-Yaw, J. A., Kharouba, H. M., Bontrager, M., Mahony, C., Csergõ, A. M., Noreen, A. M. E., . . . Angert, A. L. (2016). A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecology Letters*, *19*, 710–722.

Levitus, S. (2001). *World ocean atlas*. Silver Spring, MD: NOAA.

Lima-Ribeiro, M., Varela, S., González-Hernández, J., Oliveira, G., Diniz-Filho, J., & Terribile, L. (2015). ecoClimate: A database of climate data from multiple models for past, present, and future for macroecologists and biogeographers. *Biodiversity Informatics*, *10*, 1–21.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., . . . Wilbanks, T. J. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, *463*, 747–756.

Naimi, B., & Araújo, M. B. (2016). sdm: A reproducible and extensible R platform for species distribution modelling. *Ecography*, *39*, 368–375.

Neiva, J., Assis, J., Coelho, N. C., Fernandes, F., Pearson, G. A., & Serrão, E. A. (2015). Genes left behind: Climate change threatens cryptic genetic diversity in the canopy-forming seaweed *Bifurcaria bifurcata*. *PLoS One*, *10*, e0131530.

New, M., Hulme, M., & Jones, P. (1999). Representing twentieth-century space – Time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate*, *12*, 829–856.

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., . . . Suzuki, T. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2) – An internally consistent data product for the world ocean. *Earth System Science Data*, *8*, 297–323.

Otto-Bliesner, B. L., Schneider, R., Brady, E. C., Kucera, M., Abe-Ouchi, A., Bard, E., . . . Yu, Y. (2009). A comparison of PMIP2 model simulations and the MARGO proxy reconstruction for tropical sea surface temperatures at last glacial maximum. *Climate Dynamics*, *32*, 799–815.

Parravicini, V., Kulbicki, M., Bellwood, D. R., Friedlander, A. M., Chabanet, P., Floeter, S. R., . . . Mouillot, D. (2013). Global patterns and predictors of tropical reef fish species richness. *Ecography*, *36*, 1254–1262.

Perry, A. L., Low, P. J., Ellis, J. R., & Reynolds, J. D. (2005). Climate change and distribution shifts in marine fishes. *Science*, *308*, 1912–1915.

Peterson, A. T. T., Soberón, J., Pearson, R. G. R. G., Anderson, R. P. R. P., Martínez-Meyer, E., Nakamura, M., & Bastos Araújo, M. (2011). Ecological niches and geographic distributions. In S.E. Simon, A. Levin, & S. Henry (Eds.), *Horn monographs* (pp. 328). Princeton, NJ: Princeton University Press.

R Development Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reiss, H., Birchenough, S., Borja, A., Buhl-Mortensen, L., Craeymeersch, J., Dannheim, J., . . . Degraer, S. (2014). Benthos distribution modelling and its relevance for marine ecosystem management. *ICES Journal of Marine Science*, *72*, 297–315.

Robinson, L. M., Elith, J., Hobday, A. J., Pearson, R. G., Kendall, B. E., Possingham, H. P., & Richardson, A. J. (2011). Pushing the limits in marine species distribution modelling: Lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, *20*, 789–802.

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

Scherner, F., Antunes, P., Cabral, E., Oliveira, D., Carlos, J., Hall-Spencer, J. M., . . . Pereira, S. M. B. (2013). Coastal urbanization leads to remarkable seaweed species loss and community shifts along the SW Atlantic. *Marine Pollution Bulletin*, *76*, 106–115.

Stuart-Smith, R. D., Edgar, G. J., Barrett, N. S., Kininmonth, S. J., & Bates, A. E. (2015). Thermal biases and vulnerability to warming in the world's marine fauna. *Nature*, *528*, 88–92.

Sutherst, R. W., & Maywald, G. F. (1985). A computerised system for matching climates in ecology. *Agriculture, Ecosystems and Environment*, *13*, 281–299.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., . . . Williams, S. E. (2004). Extinction risk from climate change. *Nature*, *427*, 145–148.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – A platform for ensemble forecasting of species distributions. *Ecography*, *32*, 369–373.

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.

Varela, S., Lima-Ribeiro, M. S., & Terribile, L. C. (2015). A short guide to the climatic variables of the last glacial maximum for biogeographers. *PLoS One*, *10*, e0129037.

Verbruggen, H., Tyberghein, L., Pauly, K., Vlaeminck, C., Van Nieuwenhuyze, K., Kooistra, W. H. C. F., . . . De Clerck, O. (2009). Macroecology meets macroevolution: Evolutionary niche dynamics in the seaweed *Halimeda*. *Global Ecology and Biogeography*, *18*, 393–405.

Waltari, E., Hijmans, R. J., Peterson, A. T., Nyári, Á. S., Perkins, S. L., & Guralnick, R. P. (2007). Locating pleistocene refugia: Comparing phylogeographic and ecological niche model predictions. *PLoS One*, *2*, e563.

Wang, X., Le Borgne, R., Murtugudde, R., Busalacchi, A. J., & Behrenfeld, M. (2009). Spatial and temporal variability of the phytoplankton carbon to chlorophyll ratio in the equatorial Pacific: A basin-scale modeling study. *Journal of Geophysical Research: Oceans*, *114*, C07008. https://doi.org/10.1029/2008JC004942.

Wang Hladik, C., Huang, W., Milla, K., Edmiston, L., Harwell, M., & Schalles, J. (2010). Detecting the spatial and temporal variability of chlorophyll-*a* concentration and total suspended solids in Apalachicola Bay, Florida using MODIS imagery. *International Journal of Remote Sensing*, *31*, 439–453.

Wessel, P., & Smith, W. H. F. (1996). A global, self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research: Solid Earth*, *101*, 8741–8743.

Wilby, R. L., Charles, S. P., Zorita, E., Timbal, B., Whetton, P., & Mearns, L. O. (2004). *Guidelines for use of climate scenarios developed from statistical downscaling methods. IPCC Task Group on data and scenario support for impact and climate analysis (TGICA)*. Geneva, Switzerland: IPCC Data Distribution Centre.

## BIOSKETCH

**Jorge Assis** is a post-doctoral researcher at CCMAR-CIMAR, University of Algarve. His research is focused on ecological niche modelling, past and future climate-driven range shifts and landscape genetics at multiple temporal and spatial scales.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article .

**How to cite this article:** Assis J, Tyberghein L, Bosch S, Verbruggen H, Serrão EA, De Clerck O. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecol Biogeogr*. 2018;27:277–284. https://doi.org/10.1111/geb.12693