

# ECOGRAPHY

## Investigating uncertainties in zooplankton composition shifts under climate change scenarios in the Mediterranean Sea

Fabio Benedetti, François Guilhaumon, Fanny Adloff and Sakina-Dorothee Ayata

*F. Benedetti (fabio.benedetti@obs-vlfr.fr) and S.-D. Ayata, Sorbonne Univ., UPMC Université Paris 06, CNRS, Laboratoire d'océanographie de Villefranche (LOV), Observatoire Océanologique, FR-06230 Villefranche-sur-Mer, France. – F. Guilhaumon, IRD UMR 9190 MARBEC, IRD-CNRS-IFREMER-UM, Univ. de Montpellier, Montpellier, France. – F. Adloff, CNRM UMR 3589, Météo-France/CNRS, Toulouse, France.*

### Ecography

41: 345–360, 2018

doi: 10.1111/ecog.02434

Subject Editor: Christine Meynard

Editor-in-Chief: Miguel Araújo

Accepted 16 December 2016

Ensemble niche modelling has become a common framework to predict changes in assemblages composition under climate change scenarios. The amount of uncertainty generated by the different components of this framework has rarely been assessed. In the marine realm forecasts have usually focused on taxa representing the top of the marine food-web, thus overlooking their basal component: the plankton. Calibrating environmental niche models at the global scale, we modelled the habitat suitability of 106 copepod species and estimated the dissimilarity between present and future zooplanktonic assemblages in the surface Mediterranean Sea. We identified the patterns (species replacement versus nestedness) driving the predicted dissimilarity, and quantified the relative contributions of different uncertainty sources: environmental niche models, greenhouse gas emission scenarios, circulation model configurations and species prevalence. Our results confirm that the choice of the niche modelling method is the greatest source of uncertainty in habitat suitability projections. Presence-only and presence-absence methods provided different visions of the niches, which subsequently lead to different future scenarios of biodiversity changes. Nestedness with decline in species richness is the pattern driving dissimilarity between present and future copepod assemblages. Our projections contrast with those reported for higher trophic levels, suggesting that different components of the pelagic food-web may respond discordantly to future climatic changes.

---

### Introduction

The last decade witnessed the emergence of environmental niche models as a popular tool for studying biogeography and macroecology (Zimmermann et al. 2010, Brotons 2014). Environmental niche models (ENMs) refer to a wide array of statistical methods that enable scientists to empirically relate the distribution of a species to a suite of abiotic predictors in order to approximate its environmental niche (sensu Hutchinson 1957). The constructed models can be used to infer areas of potential habitat suitability for the species within, or outside, its known spatial range (Guisan and Zimmermann 2000, Owens et al. 2012). In a context of climate change, ENMs



www.ecography.org

---

© 2016 The Authors. Ecography © 2016 Nordic Society Oikos

have been increasingly coupled with climate models to forecast future changes in species habitat suitability and distributions (Pearson and Dawson 2003, Guisan and Thuiller 2005, Brotons 2014). ENMs can be used to identify regions with potential diversity losses or gains, and can therefore help in determining priority areas for biodiversity management (Bellard et al. 2012).

The niche modelling procedure is sprinkled with decisional steps, from the initial conditions to the projections. Each step requires a decision among alternative options, and each option has a consequence on the final inference, thus generating variability (or uncertainty) around the mean projection (Araújo and New 2007, Beaumont et al. 2008). Uncertainty maps are often not provided with ENMs predictions (Rocchini et al. 2011). Studies quantifying and mapping uncertainties associated with niche models predictions are scarce and are generally based on terrestrial taxa such as New World birds (Diniz-Filho et al. 2009), freshwater fishes in France (Buisson et al. 2010), African vertebrates (Garcia et al. 2012), and European trees (Goberville et al. 2015). While these studies have identified the choice of the statistical technique to be the main driver of uncertainty, there is still no consensus on the best model. Ensemble forecasting has therefore been suggested as a solution to handle the variability of predictions based on different methods (Araújo and New 2007). But distinct modelling strategies often rely on different theoretical assumptions and data properties (e.g. presence-only versus presence-absence data, or distance-based versus regression-based methods), and therefore may not approximate the same facet of a species niche (Soberón and Nakamura 2009, Sillero 2011). Consequently, more research is needed to summarise the relative importance of the different choices made within the niche modelling framework (i.e. aims of the study, species' ecological properties, ENM type and complexity, data characteristics etc.).

Oceans play a key role in regulating Earth's climate (Sabine et al. 2004) and provide energetic and food resources for people throughout the world. Since climate change is threatening marine biodiversity and the associated ecosystem services (Hoegh-Guldberg and Bruno 2010, Gattuso et al. 2015), it is crucial to better understand how species may respond to environmental fluctuations and how the later may impact ecosystem functioning. Additionally, marine taxa live in habitats with different characteristics compared to their terrestrial peers: there are fewer barriers to dispersal (Steele 1991), and temperature gradients are declined over much larger spatial scales. Moreover, marine taxa present original life cycles and ecological characteristics (ontogenetic shifts, spatial aggregation, mobility), as well as peculiar data properties (and biases) which may pose new challenges for niche modelling (Robinson et al. 2011). Occurrence-based niche modelling has been relatively ignored in marine ecology (Robinson et al. 2011). Rather, marine ecologists have focused on the available abundance records to model the population dynamics of high trophic level taxa and their impact on food supplies (Cheung et al. 2010), or to relate

their distribution shifts to climate fluctuations (Perry et al. 2005, Pinsky et al. 2013).

Amid all oceanic basins, the Mediterranean Sea occupies a particular position: this semi-enclosed basin is a hotspot for biodiversity (Myers et al. 2000, Bianchi and Morri 2000) that undergoes intense human pressures (The MerMex Group 2011, Coll et al. 2012) and strong rates of warming due to anthropogenic climate change (Giorgi 2006). Ongoing and future warming of the basin will likely trigger biodiversity losses across different benthic communities (Danovaro et al. 2004, Coma et al. 2009, Garrabou et al. 2009). Pelagic communities are also likely to experience alterations in species composition due to the northwards migrations of species tracking their optimal thermal niche (Sabatès et al. 2006). Northwards migration would promote the replacement of cold-water species by warm-water competitors (Ben Rais Lasram et al. 2010, Albouy et al. 2012), thus leading to alterations in the food-web structure and modifications of the functional and phylogenetic components of biodiversity (Albouy et al. 2014, 2015). However, previous projections only concerned coastal fish assemblages. Potential community shifts in the lower trophic levels (i.e. the plankton) are still lacking despite their major ecological importance. Mediterranean fishes primarily feed on mesozooplankton (Stergiou and Karpouzi 2002, Costalago et al. 2014), a group that is dominated (in terms of abundance and diversity) by copepods (Siokou-Frangou et al. 2010, Mazzocchi et al. 2014). Copepods represent a key group in pelagic trophic webs and are major contributors to the marine carbon cycle (Mauchline 1998, Beaugrand et al. 2010). Climate variability has been shown to modify copepod biogeography and phenology (Beaugrand et al. 2002, Mackas et al. 2012), therefore altering upper trophic levels through bottom-up processes (Beaugrand and Kirby 2010). Copepods have been identified as “beacons of climate change” (Richardson 2008) meaning they are an ideal group for monitoring changes in environmental conditions, and the ensuing ecosystem modifications. For these reasons, ENMs are appropriate and necessary tools to estimate climate change impacts on copepod assemblages (Sunday et al. 2012).

Plankton belong to the taxa that are the least studied through occurrence-based niche modelling (Robinson et al. 2011). To our knowledge, only four studies have used ENMs to model species niches for oceanic phytoplankton (Irwin et al. 2012, Pinkernell and Beszteri 2014, Brun et al. 2015, Barton et al. 2016). Zooplankton has been subjected to more niche-based studies, yet nearly all ENMs were developed for the North Atlantic and they implied a limited set of model algorithms and species (Reygondeau and Beaugrand 2011, Beaugrand et al. 2013, Chust et al. 2013, Villarino et al. 2015, Brun et al. 2016). These studies have focused on predicting shifts in habitat suitability for a few species and on inferring changes in local biodiversity ( $\alpha$  diversity). To account for changes in species assemblages composition, one can use indices of  $\beta$  diversity (representing the variation of species assemblages composition; Anderson et al. 2011), and combine them with  $\alpha$  diversity estimates to identify the

patterns driving dissimilarity between present and future species assemblages (Dobrovolski et al. 2012).

The aims of our study are to: 1) estimate potential shifts in zooplankton surface assemblages composition in the Mediterranean Sea under several climate change scenarios; and 2) quantify the relative importance of uncertainty sources. Shifts in assemblages composition are derived from a set of ten ENMs, covering the majority of the commonly-used algorithms, and six different configurations of a regional circulation model. These configurations allow investigating the effects of greenhouse gas (GHG) emission scenarios and boundary forcings (BF), the consequences of the latter having never been explored before though it has been shown to be at least of the same order of magnitude as the one related to the GHG emission scenario (Adloff et al. 2015). In addition, five different levels in species prevalences are tested to investigate this potential source of substantial uncertainty.

## Material and methods

### Species data

To select the copepod species to be modelled, we filtered those with more than 50 occurrences in the Mediterranean Sea, based on a regional dataset (Supplementary material Appendix 1 Table A1) and, since these are not endemic, we selected only those represented globally in the Ocean Biogeographic Information System (OBIS; <www.iobis.org>), which lead to a total of 106 copepod species. Models were therefore calibrated using global data to avoid truncated distributions (Thuiller et al. 2004). The species list represents nearly 20% of the total number of copepod species reported in the Mediterranean Sea ( $n = 560$ , Razouls et al. 2005–2016). Most of the missing species being rare, our species list does represent the most commonly sampled species in the basin (Siokou-Frangou et al. 2010, Mazzocchi et al. 2014).

Observations were aggregated within the  $0.25^\circ \times 0.25^\circ$  grid cells of the World Ocean Atlas 2013 (WOA13, Levitus et al. 2013). Ultimately, only the presences recorded across a defined 30-year baseline period (1965–1994) were kept. See Supplementary material Appendix 2 Table A2, for the final list of copepod species names and their corresponding numbers of global and regional occurrences after re-sampling on WOA13's grid.

### Present and future climatic data

Sea surface temperature (SST) and sea surface salinity (SSS) were used as environmental predictors. These variables are commonly used when modelling copepod distributions because copepods are: poikilothermic, passively dispersed, and not exploited by human activities (Richardson 2008, Reygondeau and Beaugrand 2011, Chust et al. 2013).

Since the use of 30-year climatologies is often advocated for calibrating ENMs when predicting species distributions under climate change scenarios (Roubicek et al. 2010, Harris

et al. 2014), 30-year climatologies for the baseline period (1965–1994) were constructed as follows. First, four global seasonal (spring/summer/fall/winter) SST and SSS in situ climatologies were retrieved from the WOA13 (Locarnini et al. 2013, Zweng et al. 2013, available at: <www.nodc.noaa.gov/cgi-bin/OC5/woa13/woa13.pl>) at a  $1/4^\circ$  resolution, for each of the three decades constituting the chosen baseline periods: 1965–1974, 1975–1984, 1985–1994. For each variable, global decadal climatologies were computed by averaging the four initial seasonal climatologies. In addition, the standard deviation of SST was computed to obtain decadal climatologies of the seasonal variation of SST ( $\sigma$ SST) for each of the three periods. The newly-defined decadal climatologies were used to compute the final estimates of average SST, average SSS, and average  $\sigma$ SST for the 1965–1994 period, which were used to calibrate the ENMs.

Future predictions of SST, SSS and  $\sigma$ SST over the Mediterranean Sea were obtained from the regional ocean general circulation model NEMOMED8 (Beuvier et al. 2010) under multiple forcing configurations (Adloff et al. 2015). Its horizontal resolution is  $1/8^\circ$  ( $\sim 9$  to 12 km grid cells depending on latitude) and it has 43 vertical levels. NEMOMED8 has been previously used to project fish distributions under climate change scenarios (Ben Rais Lasram et al. 2010, Albouy et al. 2012, Hattab et al. 2014).

This ocean model presents three main sources of boundary forcing: the Atlantic hydrography, the river runoff and the atmospheric surface fluxes. To assess projection uncertainty related to the choice of the GHG emission scenario and to the choice of the different boundary forcings, we used the six-member ensemble scenario simulations of Adloff et al. (2015). In their numerical experiments, the origin of the model boundary forcings (surface flux, river runoff and Atlantic hydrography) was alternatively changed, and three different scenarios of GHG emission were considered. The GHG emission scenarios used in their study are based on the Special Report on Emission Scenarios (SRES) of the Intergovernmental Panel for Climate Change (IPCC 2007). The authors' annotations for the model runs were kept (A2, A2-F, A2-RF, A2-ARF, A1B-ARF, B1-ARF) with A2, A1B and B1 indicating the GHG emission scenario, F the updated surface fluxes conditions, R the updated river runoff conditions, and A the updated Atlantic hydrography conditions. The sensitivity to each of the ocean model boundary forcings can be assessed through the 'one to one' comparison among the simulations A2, A2-F, A2-RF and A2-ARF. The comparison between A2-ARF, A1B-ARF and B1-ARF allows to assess the uncertainty related to the choice of the GHG emission scenario. The latest IPCC report provides more recent GHG concentration scenarios (Representative Concentration Pathways RCPs; IPCC 2013), but to date there is no regionalized climate change model for the Mediterranean Sea under RCP scenarios.

Monthly outputs of the ocean model were used to calculate seasonal and decadal climatologies of mean SST, mean SSS, and mean  $\sigma$ SST following the WOA13 mode (Levitus et al. 2013) for two future 30-year periods: 2020–2049

and 2069–2098. Similarly, monthly outputs for the baseline period were used to construct additional climatologies of mean SST, SSS and  $\sigma$ SST. These were used to compute the modelled climatological anomalies for the three environmental predictors. The modelled climatological anomalies were added to the baseline in situ climatologies to obtain future fields of SST, SSS and  $\sigma$ SST that are corrected for the bias between NEMOMED8 outputs and the observations. Therefore, the ENMs that were calibrated on observational climatologies were not directly projected on modelled data (Hattab et al. 2014).

The final climatologies were then used to project the habitat suitabilities of the 106 species in 2020–2049 and 2069–2098 using the ENMs (Fig. 1). See Supplementary material Appendix 3 Table A3, for a comparison of NEMOMED8 outputs against WOA13 observations, and a summary of the climatic anomalies used for predicting species habitat suitabilities in 2020–2049 and in 2069–2098.

### **Environmental niche modelling and uncertainty due to species prevalences**

To investigate the uncertainty due to the choice of the ENM, we used ten algorithms that cover the complexity range of the commonly-used statistical methods (Merow et al. 2014): three regression-based models: generalized linear model (GLM), generalized additive model (GAM), multivariate adaptive regression splines (MARS); one classification-based model: flexible discriminant analysis (FDA); two tree-based models: random forest (RF), classification tree analysis (CTA); and three machine learning models: boosted regression trees (BRT), artificial neural networks (ANN), and maximum entropy (MAXENT; Phillips et al. 2006). These nine models require either presence/absence (P/A) data or presence–background (P/B) data (for MAXENT, Yackulic et al. 2013, Guillera-Aroita et al. 2014). An additional presence-only (P/O) ENM was also used: surface range envelope (SRE, equivalent to the Bioclim model; Busby 1991).

For the nine algorithms requiring P/A (or P/B) data, pseudo-absences (psA) were randomly generated after defining both environmental and spatial weighting (Engler et al. 2004, Hengl et al. 2009), to place them in regions of lowest environmental suitability and far from the known presences (Hattab et al. 2014). The method employed to generate pseudo-absences strongly impacts ENMs outputs, and should be chosen in light of the species' ecological characteristics (Chefaoui and Lobo 2008, VanDerWal et al. 2009, Barbet-Massin et al. 2012). Zooplankton is composed of ectotherms whose population dynamics are tightly coupled to climate (Hays et al. 2005, Richardson 2008, Beaugrand et al. 2013), and whose individuals are passively dispersed over very large spatial scales in relatively short time periods (Jönsson and Watson 2016). Building on the results of Chefaoui and Lobo (2008), we chose to employ the following method to randomly draw psA.

Firstly, the reverse environmental weighting was based on a P/O ecological niche factor analysis (ENFA; Hirzel et al. 2002). This multivariate ordination technique allocates a

degree of similarity (ranging between 0 and 100) to each cell (at the global scale for ENM calibration) by comparing the species environmental envelope to the environmental conditions (using the three selected predictors). It therefore provides a habitat suitability index (HSI) that was used together with the distance to presences to produce the following probability distribution  $\tau$  (Hengl et al. 2009):

$$\tau_x = \left[ \frac{d_x + (100 - HSI_x)}{2} \right]^2$$

with  $d_x$  being the distance to presences normalized by the maximum distance, and pseudo-absences are increasingly drawn at the edge of low HSI values because of the squared term.  $\tau$  was used as a probability density function to randomly simulate pseudo-absences in unsuitable habitats and further away from known presences. We argue that this method allows to draw the psA in the environment that is theoretically reachable for the studied species (as recommended by Barve et al. 2011) because at decadal time scales, planktonic communities are well-connected anywhere in the ocean, as recently demonstrated by Jönsson and Watson (2016).

The chosen number of psA also impacts ENMs projections (Barbet-Massin et al. 2012, Meynard and Kaplan 2013). Considering that the regional occurrence data do not allow to approximate each species' prevalences in the Mediterranean Sea (because of the low surface coverage of scientific cruises), we chose to draw a varying number of psA for each species. This allowed to investigate the relative amount of uncertainty related to species prevalence. Different levels of species prevalences were considered by increasing the ratio of drawn pseudo-absences over the number of presences (which was kept constant at the number of observations): 1 (npsA = nP), 0.67 (npsA = 1.5 × n), 0.5 (npsA = 2 × nPres), 0.1 (npsA = 10 × nP), and 0.02 (npsA = 50 × nP). These five prevalence levels were finally used in the variance analysis as an additional uncertainty factor, together with ENM choice, SRES, BF and the ensuing interaction terms.

To account for the stochasticity in the psA generation process, 10 different psA realisations were carried out for each species (so each species presents 50 P-psA datasets). For every species and every ENMs algorithm, the 50 P/psA datasets were split into a calibration set (80%) and a testing set (20%). Models were evaluated according to the true skill statistic (TSS) criterion (Allouche et al. 2006) with a three-fold cross-validation. See Supplementary material Appendix 4 and 5 Fig. A4–5, for the species and the ENMs evaluation scores.

### **Mapping future shifts in species assemblages**

P/A distribution maps were generated from the habitat suitabilities over the Mediterranean Sea for each species (Fig. 1), and for each combination of ENM (n = 10), cross-validation runs (n = 3), prevalence level (n = 5), pseudo-absence realisation (n = 10) and hydrodynamical model's boundary forcings (n = 6). This was done for the two future time periods. Species assemblages (i.e. the sum of species modelled

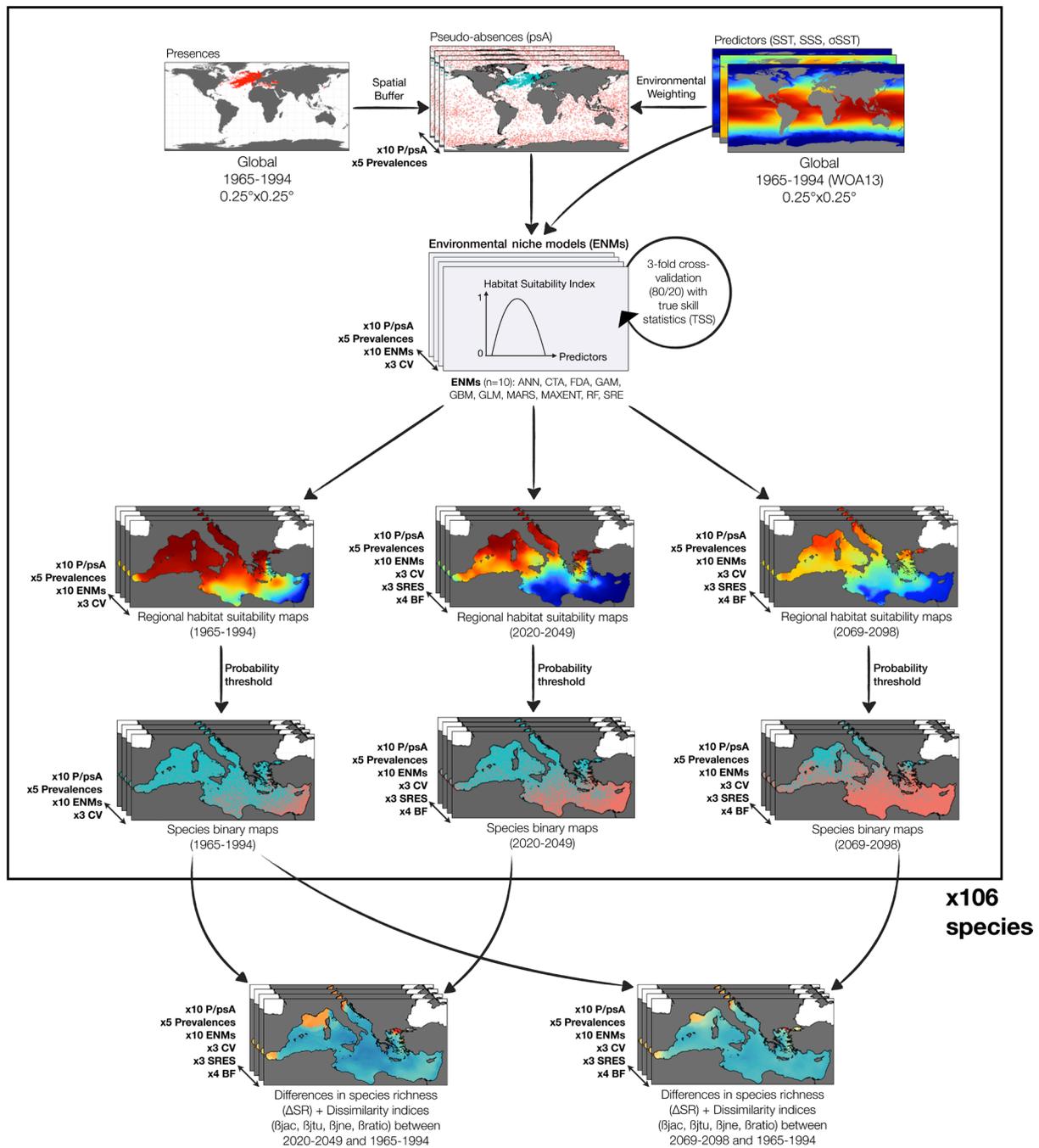


Figure 1. Schematic summary of the analytical framework of the study. For each 106 species, environmental weighting and spatial buffering are used to generate 10 P/psA datasets and each is used to calibrate 10 different ENMs. The niche models are used to project species habitat suitabilities in present (1965–1994) and future (2020–2049, 2069–2098) time periods at the regional scale. Different configurations (SRES and BF) of the regional circulation model are used to investigate their relative contribution to uncertainties. Habitat suitability maps are transformed into binary (P/A) maps (maximizing TSS threshold criterion) which are used to generate the species assemblages. Differences in species richness ( $\Delta SR$ ) and composition ( $\beta$ div indices) are computed by comparing the assemblages of 2020–2049 and 2068–2098 to the 1965–1994 assemblages.

as present or absent in each cell grid) were built by stacking all the species P/A maps (according to the identity of the above-mentioned parameters). Species assemblages for the baseline period were simulated in the same manner, but for a single set of environmental conditions (WOA13 baseline climatologies).

By comparing the present to the future assemblages, indices of community shifts related to  $\alpha$  and  $\beta$  diversity were computed within each grid cell: difference in species richness ( $\Delta SR$ ), and Jaccard's dissimilarity index ( $\beta_{jac}$ ).

For each period, the sum of the species modelled as present was used to estimate species richness.  $\Delta SR$  was computed

as the difference between future species richness and the baseline species richness, and was used to assess whether climate changes would promote or weaken copepod  $\alpha$  diversity.

Pairwise Jaccard's dissimilarity index (ranging between 0 and 1) is given by:

$$\beta_{jac} = \frac{b+c}{a+b+c}$$

where  $a$  is the number of species present at both time periods,  $b$  is the number of species present in the baseline period only, and  $c$  is the number of species present in the future time period only. It was used to assess the temporal changes in species assemblages composition. In addition, by applying the framework of Baselga (2010),  $\beta_{jac}$ -values were decomposed into its two additive components: nestedness ( $\beta_{jne}$ ) and turnover ( $\beta_{jtu}$ ). The latter expresses species replacement without the influence of  $\Delta SR$  between time steps as follows:

$$\beta_{jtu} = \frac{2 \min(b,c)}{\alpha + 2 \min(b,c)}$$

The difference between  $\beta_{jac}$  and  $\beta_{jtu}$  expresses the nestedness component  $\beta_{jne}$  that accounts for the amount of dissimilarity that is due to differences in richness ( $\beta_{jne} = \beta_{jac} - \beta_{jtu}$ ). It is formulated as follows:

$$\beta_{jne} = \frac{\max(b,c) - \min(b,c)}{a+b+c} \times \frac{a}{a+2 \min(b,c)}$$

Moreover, the ratio between  $\beta_{jne}$  and  $\beta_{jtu}$  ( $\beta_{ratio} = \beta_{jne}/\beta_{jtu}$ ) was computed and related to  $\Delta SR$  in order to understand which component has the highest contribution to future changes in the species assemblages (Dobrovolski et al. 2012, Albouy et al. 2012). For instance, a  $\beta_{ratio}$ -value greater than 0.5 indicates that the observed dissimilarity is driven by nestedness, which can occur under both increases or decreases in richness. Alternatively, a value lower than 0.5 indicates the shift is driven by species replacement. When  $\Delta SR$  is positive and turnover drives the dissimilarity in the assemblage, it means that climate change promotes diversity by creating favourable conditions for species that were not present previously. When  $\Delta SR$  is positive and nestedness drives the dissimilarity, climate change promotes richness while not changing the initial assemblage composition.

### Processing novel climate conditions

To identify where niches are projected into novel combinations of environmental predictors (Zurell et al. 2012, Mesgaran et al. 2014), cells where ENMs extrapolation occurs were determined according to the species multivariate environmental similarity surface (MESS; Elith et al. 2010). It enables to evaluate how dissimilar the environment used for projecting the ENMs is from the species native range (i.e. reference envelope used for ENMs calibration). The MESS maps present both positive and negative values, the later indicating the cells where novel climate conditions occur. Since the maps depend on the calibration dataset, MESS values were computed for each species and for every combination of prevalence level and future circulation model forcing conditions

(i.e. SRES and BF, thus 30 maps per species). The psA realisations or the ENMs' cross-evaluation runs do not significantly affect the MESS, so they were not taken into account when identifying novel climate conditions.

Within each cell, the species presenting negative MESS values were discarded from the assemblage's species list, and changes in  $\alpha$  and  $\beta$  diversity were computed without them. The ratio of species being discarded was computed and mapped for each prevalence and NEMOMED8 configuration (data not shown) in order to assess where non analog climates have the strongest impact. In the cases where novel conditions do not allow to predict changes in copepod diversity (because all species had to be discarded from the assemblages), the corresponding cells were left blank in the consensus projections and were ignored in the subsequent variance analysis. All analyses were also carried out while ignoring MESS outputs, but as it did not alter the main results, only those obtained when accounting for non analog climates are presented. Density distributions in future Mediterranean environmental ranges were visually compared to the current ones to identify the combination of predictors that may lead to the appearance of novel climate conditions.

### Consensus projections and partitioning sources of uncertainties

To identify the dominant patterns in assemblages shifts, consensus maps were drawn for each of the calculated indices by averaging their values within each cell and across every model run. The associated standard deviation was used to assess variability between runs as well as its spatial distribution over the basin.

Three-way ANOVA was used to assess the contributions of the different uncertainty sources to the overall variability in  $\Delta SR$  (Diniz-Filho et al. 2009, Garcia et al. 2012): ENMs, emission scenario (SRES), circulation model boundary forcing (BF), species prevalence, and the associated interaction terms. BF and SRES were treated separately in the variance analyses because not all BF have been coupled with every SRES. The relative amount of uncertainty attributable to the sources was estimated as the proportion of sum of squares with respect to the total sum (Diniz-Filho et al. 2009, Garcia et al. 2012).

The three-way ANOVA was performed in a linear mixed-effect model framework that allowed to account for the variation in effect size produced when iterating three cross-validation runs within 10 psA realisations.

To further examine how projections differ under combinations of ENMs and ocean model configurations, values of  $\Delta SR$  were averaged to obtain consensus projections for each combination of BF/SRES and ENMs (e.g. SRE-A2, MAXENT-B1ARF, GLM-A2F etc.). Similarity between these projections were assessed by analysing their loadings on the first principal component of an unscaled principal component analysis (PCA; Legendre and Legendre 2012).

All statistical analyses were conducted under the R environment (<www.r-project.org>) using the biomod2 package

(Thuiller et al. 2013) for ENMs and the betapart package (Baselga and Orme 2012) for biodiversity indices' estimates.

## Data deposition

Data available from the Dryad Digital Repository: <<http://dx.doi.org/10.5061/dryad.63134>> (Benedetti et al. 2017).

## Results

Consensus patterns of  $\alpha$  and  $\beta$  diversity show that species nestedness with decreases in richness is the main pattern driving the dissimilarity between present and future copepod assemblages ( $\Delta SR < 0$  and  $\beta ratio < 0.5$ ; Fig. 2). Indeed, 91.63% and 95.85% of the cells exhibit a decline in species richness by 2020–2049 and 2069–2098 respectively. Mean  $\Delta SR$  is  $-2.13$  for the mid-century and  $-5.13$  for the end-of-century period (mean SR for the baseline period being 75.83). The projected patterns are spatially structured. The largest decreases in richness are observed in the eastern Mediterranean (Fig. 2c–d). The few cells that show positive  $\Delta SR$  (8.37% by 2020–2049; 4.15% by 2069–2098) are located in the northwestern area (Gulf of Lions, Ligurian Sea), the Alboran Sea, and the northernmost parts of the Adriatic and the Aegean Seas (Fig. 2c–d), which are characterised by milder temperature and lower salinity waters, both now and in the future. Gains and losses in richness are mainly associated with nestedness. Most of Mediterranean cells show mean  $\beta ratio$  values higher than 0.5 for 2020–2049 (total mean  $\beta ratio = 0.63$ ). By 2069–2098, 98.31% of the cells exhibit a higher contribution of nestedness than turnover in assemblages' dissimilarity (total mean  $\beta ratio = 0.77$ ). Gains in species richness associated with turnover only appear in the Alboran and Marmara Seas by 2069–2098 (Fig. 2d).

As shown on Fig. 2c–d, non analog conditions occur in the central Aegean Sea and the easternmost part of the Levantine basin by 2069–2098, due to non analog SSS values (future SST and  $\sigma_{\theta}$  were always within the range of the calibration data). The entire Levantine basin is actually affected by novel salinity conditions, as nearly 50% of the species had to be discarded from the assemblages, depending on the considered model forcings and prevalences (data not shown).

Standard deviations of  $\Delta SR$  were computed within each cell to assess its variability across runs (Fig. 3). The amplitude of the predicted losses in species richness scales with its variability ( $R^2 = 0.54$ ,  $p$ -value  $< 2.2 \times 10^{-16}$ ), which is also true for the cells that exhibit positive  $\Delta SR$  value ( $R^2 = 0.59$ ,  $p$ -value  $< 2.2 \times 10^{-16}$ ). The strongest differences in richness and assemblage dissimilarity rates occur in the eastern part of the Mediterranean basin. Uncertainties across model runs are therefore larger for these regions.

Three-way ANOVA based on linear mixed-effects models helped disentangling the relative contribution of different sources of uncertainty. For both future time periods, the

choice of the ENMs explains most of the variability in projections of  $\Delta SR$  (Fig. 4a–b). The same result was obtained for the dissimilarity indices, and are therefore not presented. On average, the relative contribution of ENMs to the total sum of squares increases from 72.68% in 2020–2049 to 74.14% in 2069–2098, when accounting for the choice of BF. When accounting for SRES rather than BF, ENMs' average contribution decreases from 82.25% in 2020–2049 to 68.66% in 2069–2098. The second most important contributing factors are the interaction terms, indicating divergent ENMs responses according to different BF and SRES configurations (23.16% in 2020–2049 and 17.96% in 2069–2098 when accounting for BF; 14.44% in 2020–2049 and 19.81% in 2069–2098 with SRES). The choice of the BF, of the SRES, or of the prevalence level, always show mean relative contributions lower than 6%.

Predicted shifts in species assemblages are mainly driven by the ENMs and their interactions with either the BF or the SRES. Therefore similarities across the ENMs' average  $\Delta SR$  projections were examined through a PCA. The first two principal components (PC1 and PC2) explain 87.73% of total variance. However the SRE predictions' relative contribution to PC1 and PC2 reaches 48.85% (Fig. 5a). Since all PCs are orthogonal to one another and the P/A ENMs are contributing to non-significant components, SRE projections should not be averaged together with the other methods in an ensemble forecasting framework. When averaging SRE projections, higher consensus values of negative  $\Delta SR$  and nestedness are obtained (Fig. 5b). Mean  $\Delta SR$  in the Mediterranean Sea for 2069–2098 decreases to  $-11.21$  while the mean  $\beta ratio$  increases up to 0.95. The proportion of cells exhibiting strictly positive  $\Delta SR$  slightly increases from 4.15% to 7.85%.

A second PCA was performed with the P/psA models' projections only. Again, the loading coordinates are used to explore their degree of similarity in their  $\Delta SR$  forecasts (Fig. 6). Examining the maps of mean  $\Delta SR$  for each P/psA ENM revealed that the loadings along PC1 (64.00% of total variance) are to be interpreted as an increasing gradient in the predicted species loss. All nine p/psA ENMs present the same spatial pattern in species richness increases, while the range of their predicted decreases vary. ENMs with higher loadings on PC1 are those that predict greater declines in richness (Fig. 6). MARS projections are the most pessimistic regarding  $\Delta SR$ -values, with all its configurations having higher loadings than the other ENMs. The more pessimistic models include MARS, CTA and ANN (in order of pessimism). The least pessimistic forecasts are produced by the remaining models: MAXENT, RF, FDA, GAM, GBM and GLM.

The ENMs predicting the most negative values of  $\Delta SR$  are the ones with the most variability in loadings along PC1 (MARS followed by CTA and ANN). A significant correlation was found between the ENMs average loadings and the range of loadings (the distance between the least conservative configuration (A2ARF) and the most conservative one (B1ARF) ( $R^2 = 0.82$ ,  $p$ -value  $< 0.001$ ). This indicates the

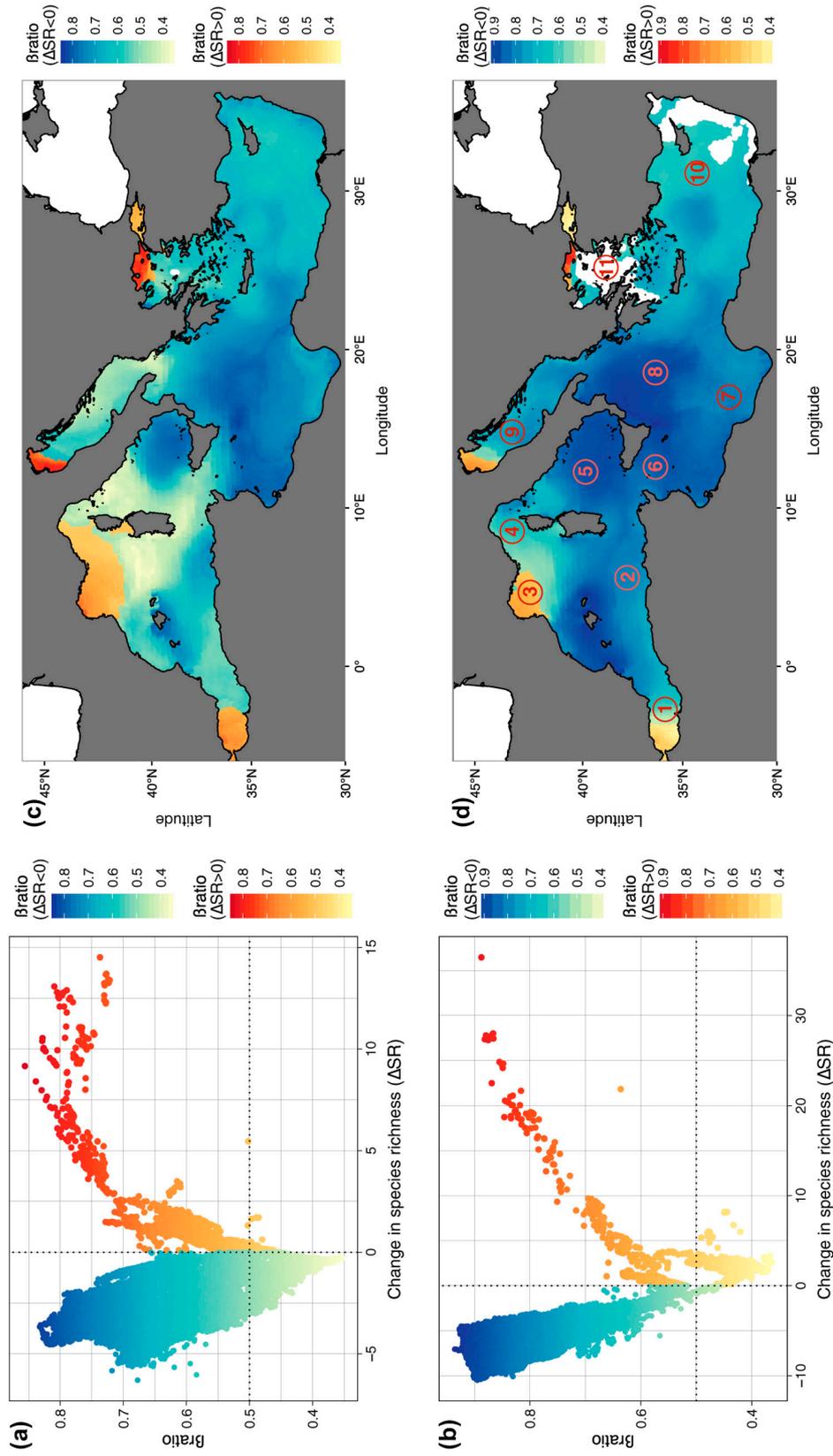


Figure 2. Changes in species richness and composition between the baseline period (1965–1994) and 2020–2049 (a, c) and 2069–2098 (b, d) for the copepod assemblages of the surface Mediterranean Sea. Changes in species composition are quantified using Jaccard's dissimilarity index  $\beta_{jac}$  (a, b) and the Bratio index (c, d). Numbers in (d) indicate the main Mediterranean sub-basins: 1) Alboran Sea 2) Algerian and Tunisian waters 3) Gulf of Lions 4) Ligurian Sea 5) Tyrrhenian Sea 6) Sicilian Strait 7) Tunisian and Libyan shelf waters 8) Ionian Sea 9) Adriatic Sea 10) Levantine Sea 11) Aegean Sea.

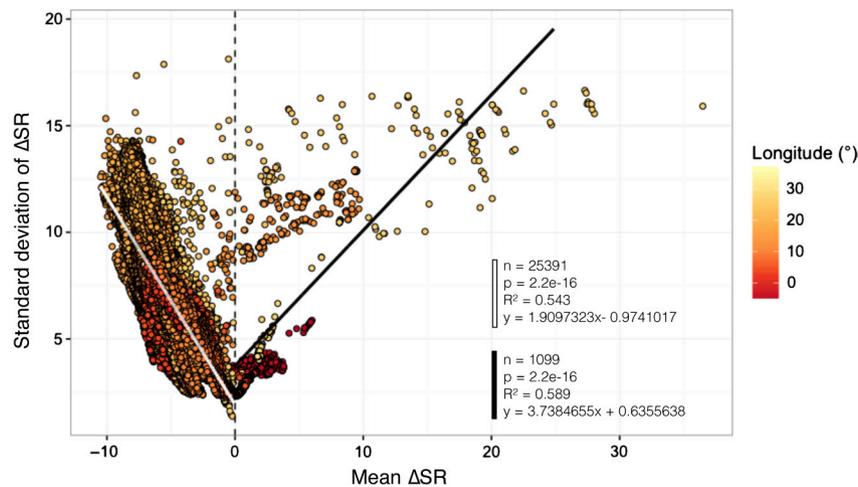


Figure 3. Linear relationships between the average value of predicted  $\Delta$ SR and the associated standard deviation, for the Mediterranean cells exhibiting both losses and increases in richness. Mean  $\Delta$ SR-values come from all ENMs predictions for the 2069–2098 period. The two linear model statistics are given in the figure.

most pessimistic ENMs are also the most sensitive to the choice of the BF and SRES.

## Discussion

In the present study, we explored the uncertainties in future composition changes in Mediterranean copepod assemblages. Different combinations of species prevalences, ENMs, circulation model boundary forcings, and emission scenarios were used to forecast shifts in species assemblages for the 21st century. Our results have implications for studies aiming to forecast changes in habitat suitability for planktonic species with ENMs, from both technical and theoretical perspectives.

### Main sources of uncertainties

Our results are in agreement with previous studies (Diniz-Filho et al. 2009, Buisson et al. 2010, Garcia et al. 2012) that documented the variability in forecasts related to the differences between ENMs outputs (Fig. 4). The main divergence occurs between the SRE model, the sole P/O ENM considered here, and the other ENMs that were all based on P/psA (or P/B for MAXENT) data (Fig. 5).

Dissimilarity between present and future assemblages was much greater when predicted by the SRE than with the other ENMs, but nestedness in species losses remained the dominant pattern driving the dissimilarity. Previous studies comparing P/O to P/A (or P/psA and P/B) methodologies consistently found SRE models to underestimate species ranges, which translated into more pessimistic forecasts of changes in richness (Pearson et al. 2006, Hijmans and Graham 2006). The SRE uses percentile distribution to draw a rectangular ‘box’ (a range envelope) around the presence data in environmental space (Busby 1991). SRE projections depend on the overlap between the defined envelope and the future conditions,

whereas P/A models (e.g. GLM-derived response curves) allow to recognise favourable areas even beyond the range of the observed presences. The large discrepancies between the SRE and the P/psA models are also due to the methodology chosen to generate psA. Environmental and spatial weightings were applied such that psA were assigned in unsuitable areas as far as possible from known presences (Hengl et al. 2009). Therefore the environmental range captured by the SRE is much narrower than the one captured by the other models.

The similarity between correlative ENMs along the first component of a PCA can be related to the similarity between the models’ algorithms (Fig. 6). Regression-based methods forecasts are quite similar to one another (GLM, GAM, FDA). The same is evidenced for complex classification-trees (GBM, RF). GBM predictions are similar to regression models because the algorithm used here is equivalent to boosted regression trees (Ridgeway 1999, Friedman 2001). MAXENT projections are similar to both GBM’s and regression-based models’, as its core algorithm contains a machine-learning piece (Elith et al. 2011), and it may be close to GLMs depending on the tuning of its parameters (Guillera-Arroita et al. 2014, Halvorsen et al. 2015). MARS-based projections forecasted the greatest rates of species loss among P/psA models. It may seem surprising that MARS projections were not closer to regression-based models since they rely on a non-parametric regression procedure that is often seen as an extension of GLMs and GAMs (Friedman 1991, Franklin 2009). The greater species losses predicted by MARS models could be due to the first-order interactions that were enabled between the predictive variables (they were disabled for the other models).

The interaction terms between ENMs and BF, and/or the choice of the SRES, were identified as the second uncertainty-generating factor (Fig. 4). The BF can have an important but very local impact, meaning it is restricted to the few

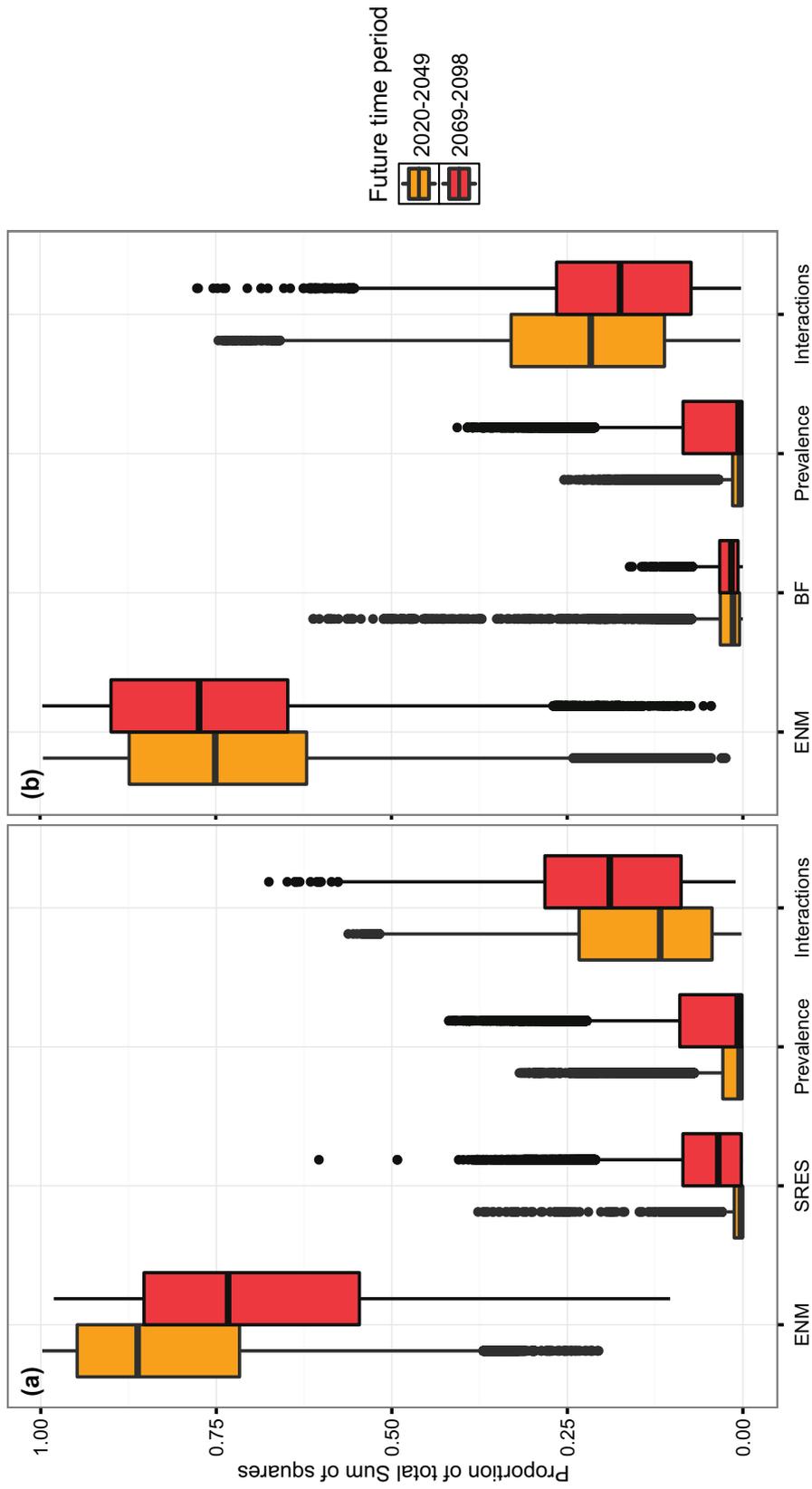


Figure 4. Proportion of the total sum of square attributed to the following source of uncertainties: (a) ENM, GHG emission scenario (SRES), species prevalence, and associated interaction terms, and (b) ENM, boundary forcings, species prevalence and associated interaction terms, for the two future time periods (2020–2049, 2069–2098).

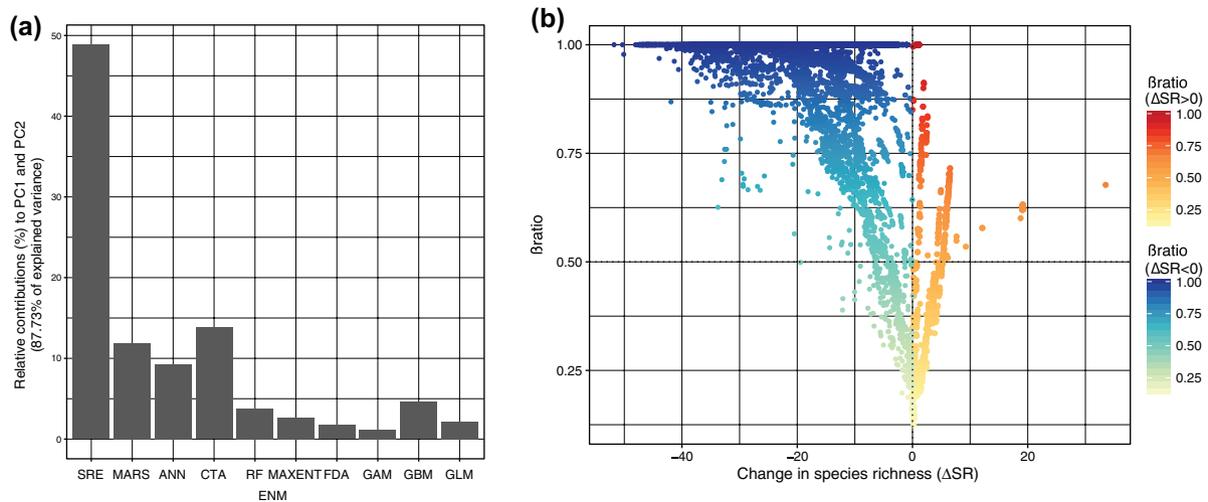


Figure 5. (a) Relative contribution of the 10 ENMs to the first two principal components of a PCA, SRE projections are incomparable to presence-absence and presence-background ENMs' forecasts. (b) Consensus changes in species richness and surface assemblages composition between the baseline (1965–1994) and the future (2069–2098) time periods, based on SRE models only, quantified using the Jaccard's dissimilarity index  $\beta_{jac}$ .

cells located near the forcing fluxes (e.g. the Alboran Sea for the Atlantic hydrography, and the Northern Aegean sea for the river runoff, since the Black Sea is treated as a river in this circulation model). The sensitivity to the choice of the SRES slightly increases between the two future periods

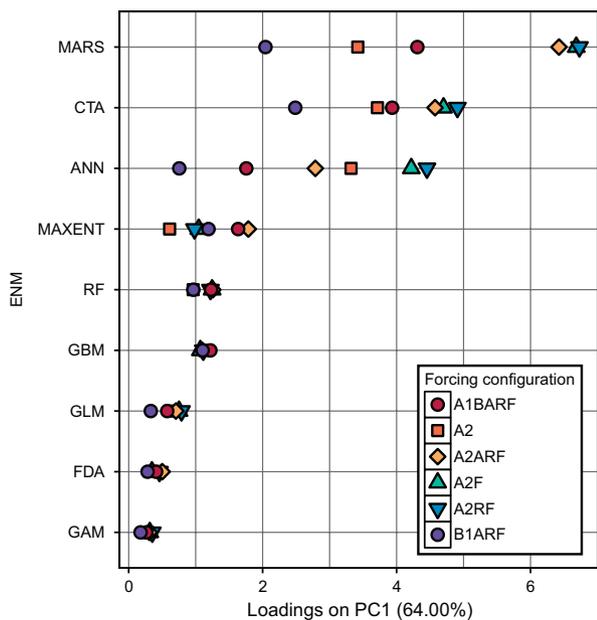


Figure 6. ENMs loadings (excluding SRE) on the first principal component (72.36% of explained variance) of an un-scaled PCA based on their mean projections of  $\Delta SR$  for the 2069–2098 period. Colors and shapes describe the forcing configuration of the circulation model (BF and SRES). PC1 represents a gradient in predicted species losses (negative  $\Delta SR$ ), ENMs with higher loadings are the ones predicting greater rates of species loss. Within-ENM spreading of the loadings along PC1 indicates the method sensitivity to the choices of the SRES/BF (i.e. the strength of interactions).

(Fig. 4b) consistently with the response of the physical variables in climate scenarios. Due to the long lifetime of anthropogenic  $CO_2$  in the atmosphere, the magnitude of oceanic response is smaller in the first half of the 21st century.

Our results have important implications for interpreting previous studies that have predicted shifts in habitat suitability for fishes over the Mediterranean continental shelf (Ben Rais Lasram et al. 2010, Albouy et al. 2012). Their results were all based on species habitat suitabilities that were estimated through a weighted average consensus across seven ENMs, including the SRE. Combining the SRE in ensemble predictions will lead to less conservative changes (i.e. higher predicted rates of species loss) than excluding it based on the evaluation criterion score (which is the most common criterion for weight attribution). From the present results, we argue that P/O and P/psA (or P/B) models outputs should be compared (e.g. with a PCA) before being mixed together as they rely on different data and assumptions, and are likely to model different components of the species niche (Brotons et al. 2004, Sillero 2011). Indeed, the PCA identified these two types of methods as two different 'visions' of the species niches, both leading to two different scenarios of biodiversity change.

## Guidelines for modelling zooplankton with ENMs

Modelling habitat suitabilities through P/O or P/A methods holds different implications depending on the ecological properties of the species (Hernandez et al. 2006, Tsoar et al. 2007, Jiménez-Valverde et al. 2008). One group of methods might be better suited than another according to the taxa of interest. Brotons et al. (2004) suggested that P/O models may be more accurate for species that are far from equilibrium with their environment. For several reasons, we argue

that correlative models, based on environmentally-weighted psA, are appropriate for modelling zooplankton taxa. First, they are short-lived ectotherms whose physiology and population dynamics are tightly coupled with climate variability (Hays et al. 2005, Richardson 2008). Sunday et al. (2011, 2012) showed that the spatial ranges of marine ectotherms closely match their thermal tolerance limits. Most of the zooplankton are not commercially exploited, so in absence of direct human harvesting they are likely to be near equilibrium with the environment, and their geographical distribution is a good indicator of their abiotic preferences.

Copepods exhibit very broad latitudinal ranges (Razouls et al. 2005–2016) which result from both wide environmental preferences and huge dispersal potential due to turbulent oceanic circulation (Jönsson and Watson 2016). However, the potentially worldwide distributions of these organisms (Finlay 2002, Cermeño and Falkowski 2009, de Vargas et al. 2015), combined with spatially (and temporally) biased data sets, limit the ENMs' capacity to link species occurrences to environmental predictors properly. In consideration of the datasets attributes (large spatial autocorrelation due to sampling biases coarse resolution etc.), future niche modelling studies should not focus on using complex ENM algorithms. Since they are likely to fit spurious relationships, or natural stochasticity (visible through noisy response curves), and thus to be less transferable in time and space (Jiménez-Valverde et al. 2008, Heinänen et al. 2012, Merow et al. 2014).

We advocate that P/psA methods applied to zooplankton should be coupled with environmentally-weighted simulations of pseudo-absences because 1) absence data are impossible to ascertain in the plankton realm, and 2) marine ectotherms are at equilibrium with their environments. Multiple methods of model evaluation and comparison should be considered (Brun et al. 2016), such as niche transferability tests in space and time (niche hindcasting), or comparison with mechanistic models outputs and/or response curves from laboratory experiments. Habitat suitability estimates along environmental gradients will be of great use for marine ecologists as they can easily be coupled with functional traits data (Benedetti et al. 2016) to better explore trait biogeography, and their link with ecosystem functioning (Albouy et al. 2015).

### Future shifts in zooplankton surface assemblages

Previous studies modelling climate change impacts on zooplankton have generally focused on changes of habitat suitability or species richness (Reygondeau and Beaugrand 2011, Beaugrand et al. 2015). Here, we extended the use of planktonic niche models for measuring  $\beta$  diversity. We predicted that climate change might lead to a loss of copepod diversity throughout most of the surface of the Mediterranean Sea (although some northern regions exhibit increases in species richness), with nestedness as the main pattern driving

the dissimilarity between present and future assemblages (Fig. 2). Our results imply that future copepod assemblages in most of the surface Mediterranean Sea will be composed of less species, all remnants being present in the initial assemblages. Areas of potential future increases in copepod diversity are restricted to the coldest regions: the Gulf of Lions, the Alboran Sea and the northern Adriatic and Aegean Seas. Again, our predictions imply that climate change might make these areas suitable for new species, without removing the ones present in the initial assemblages.

This pattern may be explained by the northward shifts of temperature and salinity sensitive species towards the Gulf of Lions, and the northern Adriatic and Aegean Seas. These results are in agreement with the rate of climate change estimated by Burrows et al. (2014), who tracked SST isotherms modelled by a global ocean model that was forced by the RCP 8.5 emission scenario.

Nestedness was also the dominant pattern in forecasted changes in Mediterranean coastal fish assemblages (Albouy et al. 2012). However, the rates of changes in species richness we found for copepods are arguably much lower. In addition, Albouy et al. (2012) predicted higher proportions of cells displaying increases in species richness, and in more diverse locations throughout the basin. They predicted increases in fish diversity in the central Adriatic, the central and northern Aegean, and the coastal Levantine. Meanwhile, we predict higher diversity losses in these regions, that will experience the strongest rates of warming and saltening (Adloff et al. 2015).

The fact that planktonic copepods present broader environmental niches, compared to coastal fishes, may explain this discrepancy. The considered fishes include endemic species with much narrower thermal amplitudes (Ben Rais Lasram et al. 2010) than planktonic species characterized by global scale distributions. This is in line with the results from Mediterranean coastal time series that demonstrated the strong resilience of copepod communities to highly variable conditions, over pluri-decadal scales (Siokou-Frangou et al. 2010, Mazzocchi et al. 2011).

The comparison between our results and previous studies conducted on other components of Mediterranean marine ecosystems implies that different components of the pelagic food web may not respond to climate changes in unique ways. Consequently, predicting climate-induced shifts in ecosystems requires to account for multiple trophic levels.

### Limitations

A notable limitation of our study is that we were unable to test the relative contribution of the choice of the regional circulation model because of data availability. Previous studies have shown this factor to be the second-most important in explaining variability across predictions, persistently ranking in front of interaction terms (Diniz-Filho et al. 2009,

Buisson et al. 2010, Garcia et al. 2012). Consequently, it is reasonable to believe that it could represent a second-order uncertainty factor in our case as well. But it is not likely to overstep ENMs as the major source of uncertainty (Garcia et al. 2012). It is noteworthy that different regional ocean circulation models generally agree on the future impacts of climate change on the overall Mediterranean circulation and physical conditions (Dubois et al. 2012, Gualdi et al. 2013). So we are confident our consensus patterns of shifts might not drastically change when switching to another regional circulation model.

Additionally, zooplankton is known to perform diel vertical migrations that can span several hundreds of meters depending on the species (Roe 1974, Ohman 1990). So it is crucial to note that our habitat suitability predictions are only valid for the surface waters of the Mediterranean Sea. Changes in the species surface habitat suitability could lead to horizontal spatial range shifts as the species track optimal growth conditions (Sunday et al. 2012, Poloczanska et al. 2013), but it could also trigger a deepening of their distribution (Dulvy et al. 2008). The exact depth of the species occurrences is difficult to establish for each observation which limits the development of three-dimensional niche models (Bentlage et al. 2013). Still, we point out that the majority of the occurrence data used here comes from surface layers (0–200 m depth), and that only a few of the studied copepod species do perform large vertical migrations in the Mediterranean basin (Scotto di Carlo et al. 1984, Benedetti et al. 2016).

*Acknowledgments* – The authors are grateful to Jean-Olivier Irisson and Guilhem Marre for their numerical inputs all along the study. We are most grateful to Tarek Hattab for his help regarding the pseudo-absence generation algorithm. We also thank Samuel Somot. Special thanks are given to Traci Erin Cox for her corrections on the text. We are also grateful to Christine Meynard for comments on the manuscript.

*Funding* – Financial support was provided by the EC FP7 PERSEUS Project (grant. Agr. 287600), the MerMex (Marine Ecosystems Response in the Mediterranean Experiment)/MISTRALS French National Program, through the PlankMed action, and by the Climate-KIC initiative of the European Institute of Innovation & Technology (EIT) through a PhD grant given to FB. This work is part of the Med-CORDEX initiative (<www.medcordex.eu>) supported by the HyMeX programme (<www.hymex.org>).

## References

- Adloff, F. et al. 2015. Mediterranean Sea response to climate change in an ensemble of twenty first century scenarios. – *Climate Dynamics* 45: 2775–2802.
- Albouy, C. et al. 2012. Combining projected changes in species richness and composition reveals climate change impacts on coastal Mediterranean fish assemblages. – *Global Change Biol.* 18: 2995–3003.
- Albouy, C. et al. 2014. From projected species distribution to food-web structure under climate change. – *Global Change Biol.* 20: 730–741.
- Albouy, C. et al. 2015. Projected impacts of climate warming on the functional and phylogenetic components of coastal Mediterranean fish biodiversity. – *Ecography* 38: 681–689.
- Allouche, O. et al. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). – *J. Appl. Ecol.* 43: 1223–1232.
- Anderson, M. J. et al. 2011. Navigating the multiple meanings of  $\beta$  diversity: a roadmap for the practicing ecologist. – *Ecol. Lett.* 14: 19–28.
- Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. – *Trends Ecol. Evol.* 22: 42–47.
- Barbet-Massin, M. et al. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? – *Meth. Ecol. Evol.* 3: 327–338.
- Barton, A. D. et al. 2016. Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. – *Proc. Natl Acad. Sci. USA* 113: 2964–2969.
- Barve, N. et al. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. – *Ecol. Modell.* 222: 1810–1819.
- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. – *Global Ecol. Biogeogr.* 19: 134–143.
- Baselga, A. and Orme, C. D. L. 2012. betapart: an R package for the study of beta diversity. – *Meth. Ecol. Evol.* 3: 808–812.
- Beaugrand, G. and Kirby, R. R. 2010. Climate, plankton and cod. – *Global Change Biol.* 16: 1268–1280.
- Beaugrand, G. et al. 2002. Diversity of calanoid copepods in the North Atlantic and adjacent seas: species associations and biogeography. – *Mar. Ecol. Progr. Ser.* 232: 179–195.
- Beaugrand, G. et al. 2010. Marine biodiversity, ecosystem functioning and carbon cycles. – *Proc. Natl Acad. Sci. USA* 107: 10120–10124.
- Beaugrand, G. et al. 2013. Applying the concept of the ecological niche and a macroecological approach to understand how climate influences zooplankton: advantages, assumptions, limitations and requirements. – *Progr. Oceanogr.* 111: 75–90.
- Beaugrand, G. et al. 2015. Future vulnerability of marine biodiversity compared with contemporary and past changes. – *Nat. Climate Change* 5: 695–701.
- Beaumont, L. J. et al. 2008. Why is the choice of future climate scenarios for species distribution modelling important? – *Ecol. Lett.* 11: 1135–1146.
- Bellard, C. et al. 2012. Impacts of climate change on the future of biodiversity. – *Ecol. Lett.* 15: 365–377.
- Ben Rais Lasram, F. et al. 2010. The Mediterranean Sea as a “cul de sac” for endemic fishes facing climate change. – *Global Change Biol.* 16: 3233–3245.
- Benedetti, F. et al. 2016. Identifying copepod functional groups from species functional traits. – *J. Plankton Res.* 38: 159–166.
- Benedetti, F. et al. 2017. Data from: Investigating uncertainties in zooplankton composition shifts under climate change scenarios in the Mediterranean Sea. – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.63134>.
- Bentlage, B. et al. 2013. Plumbing the depths: extending ecological niche modelling and species distribution modelling in three dimensions. – *Global Ecol. Biogeogr.* 22: 952–961.

- Beuvier, J. et al. 2010. Modeling the Mediterranean Sea interannual variability during 1961–2000: focus on the eastern Mediterranean transient. – *J. Geophys. Res. Oceans* (1978–2012) 115:C8.
- Bianchi, C. N. and Morri, C. 2000. Marine biodiversity of the Mediterranean Sea: situation, problems and prospects for future research. – *Mar. Pollut. Bull.* 40: 367–376.
- Brotons, L. 2014. Species distribution models and impact factor growth in environmental journals: methodological fashion or the attraction of global change science. – *PLoS ONE* 9: e111996.
- Brotons, L. et al. 2004. Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. – *Ecography* 27: 437–448.
- Brun, P. et al. 2015. Ecological niches of open ocean phytoplankton taxa. – *Limnol. Oceanogr.* 60: 1020–1038.
- Brun, P. et al. 2016. The predictive skill of species distribution models for plankton in a changing climate. – *Global Change Biol.* 22: 3170–3181.
- Buisson, L. et al. 2010. Uncertainty in ensemble forecasting of species distribution. – *Global Change Biol.* 16: 1145–1157.
- Burrows, M. T. et al. 2014. Geographical limits to species-range shifts are suggested by climate velocity. – *Nature* 507: 492–495.
- Busby, J. R. 1991. BIOCLIM—a bioclimate analysis and prediction system. – *Plant Protection Quarterly (Australia)* 6: 8–9.
- Cermeño, P. and Falkowski, P. G. 2009. Controls on diatom biogeography in the ocean. – *Science* 325: 1539–1541.
- Chefaoui, R. M. and Lobo, J. M. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. – *Ecol. Modell.* 210: 478–486.
- Cheung, W. W. et al. 2010. Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. – *Global Change Biol.* 16: 24–35.
- Chust, G. et al. 2013. Are *Calanus* spp. shifting poleward in the North Atlantic? A habitat modelling approach. – *ICES J. Mar. Sci.* 2013: fst147.
- Coll, M. et al. 2012. The Mediterranean Sea under siege: spatial overlap between marine biodiversity, cumulative threats and marine reserves. – *Global Ecol. Biogeogr.* 21: 465–480.
- Coma, R. et al. 2009. Global warming-enhanced stratification and mass mortality events in the Mediterranean. – *Proc. Natl Acad. Sci. USA* 106: 6176–6181.
- Costalago, D. et al. 2014. Seasonal comparison of the diets of juvenile European anchovy *Engraulis encrasicolus* and sardine *Sardina pilchardus* in the Gulf of Lions. – *J. Sea Res.* 89: 64–72.
- Danovaro, R. et al. 2004. Biodiversity response to climate change in a warm deep sea. – *Ecol. Lett.* 7: 821–828.
- de Vargas, C. et al. 2015. Eukaryotic plankton diversity in the sunlit ocean. – *Science* 348: 1261605.
- Diniz-Filho, J. A. F. et al. 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. – *Ecography* 32: 897–906.
- Dobrovolski, R. et al. 2012. Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. – *Global Ecol. Biogeogr.* 21: 191–197.
- Dubois, C. et al. 2012. Future projections of the surface heat and water budgets of the Mediterranean Sea in an ensemble of coupled atmosphere–ocean regional climate models. – *Climate Dynamics* 39: 1859–1884.
- Dulvy, N. K. et al. 2008. Climate change and deepening of the North Sea fish assemblage: a biotic indicator of warming seas. – *J. Appl. Ecol.* 45: 1029–1039.
- Elith, J. et al. 2010. The art of modelling range-shifting species. – *Meth. Ecol. Evol.* 1: 330–342.
- Elith, J. et al. 2011. A statistical explanation of MaxEnt for ecologists. – *Divers. Distrib.* 17: 43–57.
- Engler, R. et al. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. – *J. Appl. Ecol.* 41: 263–274.
- Finlay, B. J. 2002. Global dispersal of free-living microbial eukaryote species. – *Science* 296: 1061–1063.
- Franklin, J. 2009. Mapping species distributions: spatial inference and prediction. – Cambridge Univ. Press.
- Friedman, J. H. 1991. Multivariate adaptive regression splines. – *Ann. Stat.* 19: 1–67.
- Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. – *Ann. Stat.* 5: 1189–1232.
- García, R. A. et al. 2012. Exploring consensus in 21st century projections of climatically suitable areas for African vertebrates. – *Global Change Biol.* 18: 1253–1269.
- Garrabou, J. et al. 2009. Mass mortality in northwestern Mediterranean rocky benthic communities: effects of the 2003 heat wave. – *Global Change Biol.* 15: 1090–1103.
- Gattuso, J.-P. et al. 2015. Contrasting futures for ocean and society from different anthropogenic CO<sub>2</sub> emissions scenarios. – *Science* 349: aac4722.
- Giorgi, F. 2006. Climate change hot-spots. – *Geophys. Res. Lett.* 33: L08707.
- Goberville, E. et al. 2015. Uncertainties in the projection of species distributions related to general circulation models. – *Ecol. Evol.* 5: 1100–1116.
- Gualdi, S. et al. 2013. The CIRCE simulations: regional climate change projections with realistic representation of the Mediterranean Sea. – *Bull. Am. Meteorol. Soc.* 94: 65–81.
- Guillera-Aroita, G. et al. 2014. Maxent is not a presence–absence method: a comment on Thibaud et al. – *Meth. Ecol. Evol.* 5: 1192–1197.
- Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Modell.* 135: 147–186.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Halvorsen, R. et al. 2015. Opportunities for improved distribution modelling practice via a strict maximum likelihood interpretation of MaxEnt. – *Ecography* 38: 172–183.
- Harris, R. M. B. et al. 2014. Climate projections for ecologists. – *Climate Change* 5: 621–637.
- Hattab, T. et al. 2014. Towards a better understanding of potential impacts of climate change on marine species distribution: a multiscale modelling approach. – *Global Ecol. Biogeogr.* 23: 1417–1429.
- Hays, G. C. et al. 2005. Climate change and marine plankton. – *Trends Ecol. Evol.* 20: 337–344.
- Heinänen, S. et al. 2012. High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. – *Landscape Ecol.* 27: 545–555.
- Hengl, T. et al. 2009. Spatial prediction of species’ distributions from occurrence-only records: combining point pattern analysis, ENFA and regression-kriging. – *Ecol. Modell.* 220: 3499–3511.

- Hernandez, P. A. et al. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. – *Ecography* 29: 773–785.
- Hijmans, R. J. and Graham, C. H. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. – *Global Change Biol.* 12: 2272–2281.
- Hirzel, A. et al. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? – *Ecology* 83: 2027–2036.
- Hoegh-Guldberg, O. and Bruno, J. F. 2010. The impact of climate change on the world's marine ecosystems. – *Science* 328: 1523–1528.
- Hutchinson, G. E. 1957. Concluding remarks. – *Cold Spring Harbor Symp. Quantitative Biol.* Cold Spring Harbor Laboratory Press, pp. 415–427.
- IPCC 2007. Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. – Cambridge Univ. Press.
- IPCC 2013. Climate change 2013: the physical science basis. Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change. – Cambridge Univ. Press.
- Irwin, A. J. et al. 2012. Phytoplankton niches estimated from field data. – *Limnol. Oceanogr.* 57: 787–797.
- Jiménez-Valverde, A. et al. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. – *Divers. Distrib.* 14: 885–890.
- Jönsson, B.F. and Watson, J.R. 2016. The timescales of global-surface ocean connectivity. – *Nat. Commun.* 7: 11239.
- Legendre, P. and Legendre, L. 2012. Numerical ecology. – Elsevier.
- Levitus, S. et al. 2013. The World ocean database. – *Data Sci. J.* 12: WDS229–WDS234.
- Locarnini, R. et al. 2013. World ocean atlas 2013. Vol. 1: Temperature. – NOAA Atlas NESDIS 73: 40.
- Mackas, D. et al. 2012. Changing zooplankton seasonality in a changing ocean: comparing time series of zooplankton phenology. – *Progr. Oceanogr.* 97: 31–62.
- Mauchline, J. 1998. The biology of calanoid copepods. – *Adv. Mar. Biol.* 33: 1–710.
- Mazzocchi, M. G. et al. 2011. Zooplankton associations in a Mediterranean long-term time-series. – *J. Plankton Res.* 33: 1163–1181.
- Mazzocchi, M. et al. 2014. Regional and seasonal characteristics of epipelagic mesozooplankton in the Mediterranean Sea based on an artificial neural network analysis. – *J. Mar. Syst.* 135: 64–80.
- Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? – *Ecography* 37: 1267–1281.
- Mesgaran, M. B. et al. 2014. Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. – *Divers. Distrib.* 20: 1147–1159.
- Meynard, C. N. and Kaplan, D. M. 2013. Using virtual species to study species distributions and model performance. – *J. Biogeogr.* 40: 1–8.
- Myers, N. et al. 2000. Biodiversity hotspots for conservation priorities. – *Nature* 403: 853–858.
- Ohman, M. D. 1990. The demographic benefits of diel vertical migration by zooplankton. – *Ecol. Monogr.* 60: 257–281.
- Owens, H. L. et al. 2012. Predicting suitable environments and potential occurrences for coelacanths (*Latimeria* spp.). – *Biodiv. Conserv.* 21: 577–587.
- Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.
- Pearson, R. G. et al. 2006. Model-based uncertainty in species range prediction. – *J. Biogeogr.* 33: 1704–1711.
- Perry, A. L. et al. 2005. Climate change and distribution shifts in marine fishes. – *Science* 308: 1912–1915.
- Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – *Ecol. Modell.* 190: 231–259.
- Pinkernell, S. and Beszteri, B. 2014. Potential effects of climate change on the distribution range of the main silicate sinker of the Southern Ocean. – *Ecol. Evol.* 4: 3147–3161.
- Pinsky, M. L. et al. 2013. Marine taxa track local climate velocities. – *Science* 341: 1239–1242.
- Poloczanska, E. S. et al. 2013. Global imprint of climate change on marine life. – *Nat. Climate Change* 3: 919–925.
- Razouls, C. et al. 2005–2016. Diversity and geographic distribution of marine planktonic copepods. – <<http://copepodes.obs-banyuls.fr/en>>.
- Reygondeau, G. and Beaugrand, G. 2011. Future climate-driven shifts in distribution of *Calanus finmarchicus*. – *Global Change Biol.* 17: 756–766.
- Richardson, A. J. 2008. In hot water: zooplankton and climate change. – *ICES J. Mar. Sci.* 65: 279–295.
- Ridgeway, G. 1999. The state of boosting. – *Comput. Sci. Stat.* 31: 172–181.
- Robinson, L. et al. 2011. Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. – *Global Ecol. Biogeogr.* 20: 789–802.
- Rocchini, D. et al. 2011. Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. – *Progr. Phys. Geogr.* 35: 211–226.
- Roe, H. 1974. Observations on the diurnal vertical migrations of an oceanic animal community. – *Mar. Biol.* 28: 99–113.
- Roubicek, A. et al. 2010. Does the choice of climate baseline matter in ecological niche modelling? – *Ecol. Modell.* 221: 2280–2286.
- Sabatès, A. et al. 2006. Sea warming and fish distribution: the case of the small pelagic fish, *Sardinella aurita*, in the western Mediterranean. – *Global Change Biol.* 12: 2209–2219.
- Sabine, C. L. et al. 2004. The oceanic sink for anthropogenic CO<sub>2</sub>. – *Science* 305: 367–371.
- Scotto di Carlo, B. et al. 1984. Vertical zonation patterns for Mediterranean copepods from the surface to 3000 m at a fixed station in the Tyrrhenian Sea. – *J. Plankton Res.* 6: 1031–1056.
- Sillero, N. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. – *Ecol. Modell.* 222: 1343–1346.
- Siokou-Frangou, I. et al. 2010. Plankton in the open Mediterranean Sea: a review. – *Biogeosciences* 7: 1543–1586.
- Soberón, J. and Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. – *Proc. Natl Acad. Sci. USA* 106: 19644–19650.
- Steele, J. H. 1991. Can ecological theory cross the land-sea boundary? – *J. Theor. Biol.* 153: 425–436.
- Stergiou, K. I. and Karpouzi, V. S. 2002. Feeding habits and trophic levels of Mediterranean fish. – *Rev. Fish Biol. Fish.* 11: 217–254.

- Sunday, J. M. et al. 2011. Global analysis of thermal tolerance and latitude in ectotherms. – *Proc. R. Soc. B* 278: 1823–1830.
- Sunday, J. M. et al. 2012. Thermal tolerance and the global redistribution of animals. – *Nat. Climate Change* 2: 686–690.
- The MerMex Group 2011. Marine ecosystems' responses to climatic and anthropogenic forcings in the Mediterranean. – *Progr. Oceanogr.* 91: 97–166.
- Thuiller, W. et al. 2004. Effects of restricting environmental range of data to project current and future species distributions. – *Ecography* 27: 165–172.
- Thuiller, W. et al. 2013. biomod2: Ensemble platform for species distribution modeling. – R package ver. 2: r560.
- Tsoar, A. et al. 2007. A comparative evaluation of presence-only methods for modelling species distribution. – *Divers. Distrib.* 13: 397–405.
- VanDerWal, J. et al. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? – *Ecol. Model.* 220: 589–594.
- Villarino, E. et al. 2015. Modelling the future biogeography of North Atlantic zooplankton communities in response to climate change. – *Mar. Ecol. Progr. Ser.* 531: 121–142.
- Yackulic, C. B. et al. 2013. Presence-only modelling using MAXENT: when can we trust the inferences? – *Meth. Ecol. Evol.* 4: 236–243.
- Zimmermann, N. E. et al. 2010. New trends in species distribution modelling. – *Ecography* 33: 985–989.
- Zurell, D. et al. 2012. Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. – *Divers. Distrib.* 18: 628–634.
- Zweng, M. et al. 2013. *World Ocean Atlas 2013*. Vol. 2: Salinity. – NOAA Atlas NESDIS 74: 39.

Supplementary material (Appendix ecog-02434 at <[www.ecography.org/appendix/ecog-02434](http://www.ecography.org/appendix/ecog-02434)>). Appendix 1–5.