Species-area uncertainties impact the setting of habitat conservation targets and propagate across conservation solutions

Sabrine Drira, Frïda Ben Rais Lasram, Amel Ben Rejeb Jenhani, Yunne Jai Shin, François Guilhaumon

To cite this version:
Sabrine Drira, Frïda Ben Rais Lasram, Amel Ben Rejeb Jenhani, Yunne Jai Shin, François Guilhaumon. Species-area uncertainties impact the setting of habitat conservation targets and propagate across conservation solutions. Biological Conservation, Elsevier, 2019, 235, pp.279–289. 10.1016/j.biocon.2019.05.012. hal-02434136

HAL Id: hal-02434136
https://hal.archives-ouvertes.fr/hal-02434136
Submitted on 25 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Title: Species-area uncertainties impact the setting of habitat conservation targets and propagate across conservation solutions

Authors: Sabrine Drira¹,², Frida Ben Rais Lasram³, Amel Ben Rejeb Jenhani¹, Yunne Jai Shin² and François Guilhaumon²

¹Ecosystemes et Ressources Aquatiques, Institut National Agronomique de Tunisie, Université de Carthage, 43 Avenue Charles-Nicoles, 1082 Tunis-Mahrajène, Tunisia

²MARBEC, IRD-CNRS-IFREMER-UM, Université de Montpellier, 34095 Montpellier, France

³Univ. Littoral Cote d’Opale, Univ. Lille, CNRS, UMR 8187, LOG, Laboratoire d’Océanologie et de Géosciences, 62930 Wimereux, France

Corresponding author: Sabrine Drira

E-mail: sabrinedrira@gmail.com

Address: Marbec, 093 Place Eugène Bataillon, 34090 Montpellier

Acknowledgements: Sabrine Drira is funded by an ARTS PhD grant from the Institut de Recherche pour le Développement (IRD, France). We are very grateful to Michael S. Corson for English editing. We also thank Tarek Hattab for helpful suggestions.
1. Introduction

Human domination of the natural world is resulting in biodiversity losses similar to those during mass-extinction events recorded over geological time (Ceballos et al., 2017). In the marine realm, biodiversity is threatened by numerous anthropogenic stressors, such as overfishing (Pauly et al., 1998), climate change (Cheung et al., 2009) and biological invasions (Simberloff et al., 2013). In this context, marine protected areas (MPAs) are promoted as tools to help conserve biodiversity heritage (Wood et al., 2008), maintain ecosystem processes (Roberts et al., 2005), and favour a sustainable exploitation of living resources (Agardy et al., 2003; Spalding et al., 2008).

Systematic Conservation Planning (SCP; Margules & Pressey, 2000) is a widely-used approach to develop networks of protected areas. Based on the principle of complementarity, this approach selects protected area systems that optimize the representation of each biodiversity feature occurring in a region (e.g. species, habitats, ecosystem services; Possingham et al., 2000; Carwardine et al., 2009). A crucial step in the SCP process is to set conservation targets for biodiversity features. These targets are explicit goals that quantify the minimum amount of each biodiversity feature to be covered by the protected areas (Ball et al., 2009).

At the international level, the parties to the Convention on Biological Diversity (CBD) agreed in 2010 to protect 10% of their marine exclusive economic zone and 17% of their terrestrial lands and inland waters by 2020 (Aichi Biodiversity Target n°11; CBD, 2010). These targets, expressed as percentages of broad geographical regions are often translated in the SCP process as uniform conservation targets such as 10% for each habitat type occurring in a region (Barr et al. 2011; Pressey et al. 2003). Setting such uniform targets assumes that all habitats need to be equally protected, and ignores biological variations across regions and habitats (Agardy et al., 2003; Pressey et al., 2003). Moreover, the broadscale Aichi targets have been negotiated to encompass countries with remarkably different economic, political and cultural backgrounds (Amengual & Alvarez-Berastegui, 2018), and have been thus criticized for being justified in terms of political expediency and for their lack of scientific background (Carwardine et al., 2009; Noss et al., 2012; Svancara et al., 2005; Müller et al., 2018). In the marine realm particularly, a meta-analysis of published MPAs networks concludes that the 10% marine target proves insufficient to meet many conservation goals, such as the protection of biodiversity, has socio-economic limitations and does not account for the connectivity within MPAs networks (O’Leary et al., 2016). Considering environmental and biological variations by determining habitat-specific conservation targets has thus been advocated to provide ecologically realistic and relevant conservation solutions (Wood, 2011; Woodley et al., 2012).
The methodological strategy used for setting conservation targets varies according to the biodiversity feature considered (Metcalfe et al., 2013). When the biodiversity features are different habitats occurring in the planning region, an extensively used approach to set habitat-specific conservation targets is based on the Species-Area Relationship (SAR; Fig. 1; Rosenzweig, 1995; Triantis et al., 2012). Building on the relationship between species richness and area (SAR), habitat-specific conservation targets are set by estimating the proportion of a habitat required to represent a given proportion of the total number of species occurring within it (step 2 Fig. 1 and for instance Desmet & Cowling, 2004; Reyers et al., 2007; Metcalfe et al., 2013; Davis et al., 2017). These habitat-specific conservation targets are then used to constrain the SCP reserve selection algorithm when designing protected area networks (step 3 Fig. 1).

SARs are both universal and versatile. Positive relationships between species-richness and area have been described for a wide variety of taxa, ranging from bacteria (Horner-Devine et al., 2004) to vertebrates (Newmark et al., 2017), including invertebrates and plants (Triantis et al., 2012), as well as for a variety of ecosystems and geographic locations (Adams et al., 2017; Deane et al., 2017). Despite the universality of the pattern, systematic variations in the shape of SARs have been reported across taxa and ecosystems. For example, Qian et al. (2007) identified a strong decrease in the slope of SARs with increasing latitude for vascular plants in North America; and SARs have been shown to vary in shape among taxa (Guilhaumon et al., 2008). The shape of SARs varies also with scale, changing from exponential for small sampling areas to Power-law for intermediate sampling areas, and sigmoid when representing entire biogeographical regions (He & Legendre, 1996); or depends on the method used to construct them (Scheiner, 2003). Reflecting the variety of observed SAR shapes and the large body of theoretical work that interprets the patterns, over 20 mathematical models have been developed to describe SARs (Tjørve, 2003, 2009). Although no agreement has been reached on a single and universal model (Connor and McCoy 1979; Veech 2000; Fattorini 2007; Guilhaumon et al. 2008), the Power-law model (Arrhenius, 1921) is often assumed to describe SARs well and is therefore the most widely used in ecology and conservation science (Smith, 2010). For analytical convenience, simple linear regression models obtained by log-transforming the Power-law equation are generally used in conservation applications (see equations 1-3 in methods). When setting habitat conservation targets using SARs, most studies do not even fit the log-transformed Power-law model to data, but instead estimate the Power model parameters (and derive corresponding conservation targets) for a given habitat on the basis of two data points ($P_1, P_2$): $P_1 (x_1 = \log$ of average area of sampling units, $y_1 = \log$ of average number of species per sampling unit) and $P_2 (x_2 = \log$ of the total habitat area, $y_2 = \log$ of the total number of species) (see methods for details, Desmet & Cowling, 2004; Rouget et al., 2004;
Metcalfe et al., 2013; Davis et al., 2017). Even though this approach (called hereafter the ‘log-transformed model’) is more ecologically relevant than setting uniform conservation targets (e.g. the 10% target adopted internationally) and may be useful in data-poor situations, it is not suited to take into account the wide range of SAR shapes encountered across ecological conditions and habitats. Consequently, when the a priori chosen model (e.g. log-transformed model) is inappropriate for fitting the SAR, the estimation of conservation targets may potentially be inaccurate and lead to inefficient reserve networks, in particular when the derived targets are less ambitious than those estimated using the data-driven best fit. Conversely, when the a priori chosen model estimates too ambitious targets, scarce conservation effort could be mismanaged, while it could be rather used to protect more depleted habitats. Also, several SAR models could provide equivalent good fits to empirical data sets, resulting in model-selection uncertainty. In such situations, the choice of a particular SAR model to calculate conservation targets is delicate, especially when the estimated habitat targets vary among good fits. Several studies have reported an uncertainty when selecting the best SAR model for a given data set (Stiles & Scheiner, 2007; Guilhaumon et al., 2008), with significant influence on the outcomes of a variety of ecological applications, such as detecting richness hotspots (Guilhaumon et al., 2008) or estimating species extinction rates (Matthews et al., 2014). For coping with SAR uncertainties, multi-model inference has been recommended as a solution to provide a robust final inference, weighting several models fits with respect to their relative probabilities to best fit the SAR (Step 2 Fig. 1; Grace et al., 2010; Guilhaumon et al., 2008, 2010; Smith, 2010; Scheiner et al., 2011).

In this study, we investigate whether the uncertainty in the selection of SAR models propagates throughout the entire SCP process and leads to different habitat targets and conservation solutions. Using fishes of the coastal Mediterranean Sea as a case study, we aimed to answer the following questions: (1) What are the uncertainties underlying the selection of SAR models for different Mediterranean coastal marine habitats? We constructed empirical SARs for six Mediterranean coastal marine habitats and compared in a model selection framework 20 SAR models (hereafter SAR scenarios) to fit these empirical SARs. (2) Do SAR scenarios lead to different habitat conservation targets, and do they propagate in the SCP process, resulting in distinct conservation solutions? For each SAR scenario, we calculated conservation targets for the different habitats and identified corresponding conservation solution using a reserve-selection algorithm. We then tested whether dissimilarities between conservation targets estimated by different SAR models were correlated to dissimilarities between the corresponding conservation solutions. (3) Finally, we asked whether conservation targets and conservation solutions generated by taking into account SAR uncertainties (i.e. using a multi-model inference) were different from those based on the log-
transformed model? We estimated conservation targets and associated conservation solutions using both methods and analysed differences in conservation outcomes.

2. Material and methods

2.1. Data

Spatial distribution data for 635 Mediterranean fish species, which represent the entire Mediterranean fish assemblage, were obtained as rasters at 0.1° resolution from the FishMed database (Albouy et al., 2015). Raster maps for 6 Mediterranean Sea marine habitats, mainly coastal (i.e. rocky reef, shallow soft, soft slope, soft shelf, seagrass and rocky intertidal; Fig. A1, Table 1), based on the distribution of substrate types (e.g. soft, hard) and three depth ranges (shallow for 0-60 m, shelf for 60-200 m and slope for 200-2000 m), were obtained from Micheli et al. (2013) at 1km² resolution (see Halpern et al, 2008 and Micheli et al, 2013 for further details), and resampled on the species distribution grid.

2.2. Species-area sampling

Using gridded species distributions, we constructed empirical SAR for each habitat by tallying both the number of species and the habitat surface area occurring at increasing levels of cells number (step 1 Fig.1). For each level of \( n \) cells number (\( n \) ranging from 1 to \( N-1 \) grid cell, with \( N \) the total number of grid cells observed in the focal habitat, and using a step of one grid cell), we sampled randomly 100 sets of cells and averaged species richness and habitat surface area over these 100 sets, constructing empirical species-area curves of type IIB as defined by Scheiner (2003). For the last level of \( N \) cells we used the total number of species observed in the focal habitat, without replications.

When constructing empirical SARs, species are usually counted at the first encounter and considered present at a given level of area even if they are represented by a single individual (He & Hubbell, 2011). For conservation purposes, this approach would hardly promote the maintenance of populations that are sufficiently viable to ensure species persistence. To address this limitation, we considered that a species was present at a given level of area only if at least 10% of the species’ total range in the habitat considered was encompassed. This 10% coverage threshold was chosen conservatively to allow for species persistence (Rodrigues et al., 2004; Venter et al., 2014; Guilhaumon et al., 2015; Mouillot et al., 2016).

2.3. Species-area modeling

Many functional forms have been proposed to fit SARs; they differ in shape (e.g. convex, sigmoid, asymptotic, non-asymptotic) and in the complexity of the corresponding mathematical model (Triantis et al., 2012). For each habitat, we fitted 20 SAR models\( \hat{S}_i = g_i(A) \), \( 1 \leq i \leq 20 \) where \( \hat{S}_i \) is
the species richness inferred by model $i$ for the habitat area $A$ and $g_i$ is the functional form, including a linear function, nine convex functions and ten sigmoid functions, thus encompassing the various SAR shapes gathered in the literature (Table 2). To compare the goodness-of-fit of SAR models, we used the version of Akaike's Information Criterion (AICc; Akaike, 1987) corrected for small samples. Model selection using AICc considers both fit and complexity when comparing several models. We used Akaike weights ($\omega_i$) derived from AICc values to reflect the probability of each model to be the best in explaining the data (Burnham & Anderson, 2003). When the best-fitted model has an Akaike weight greater than 0.9, robust inferences can be made using only this model; however if none of the models has a weight greater than 0.9, no conclusion can be drawn about the best model, reflecting a substantial uncertainty surrounding the model selection (Burnham & Anderson, 2003). If so, model averaging has been recommended to make robust SAR predictions (Guilhaumon et al., 2008). Here we performed multi-model inference by averaging SAR inferences across models and weighting according to Akaike-weights: $\bar{S} = \sum_i \hat{S}_i \omega_i$.

All SAR analyses were conducted using an updated version of the R package “mmSAR” (Guilhaumon et al., 2010), which estimates non-linear model parameters for numerous SAR functional forms (Triantis et al., 2012). Of the 20 SAR models initially fitted, those with negative $R^2$ (i.e. the model does not follow the trend of the data and fits worse than an intercept-only linear regression) and those yielding negative predictions or a decrease in species richness with increasing area were discarded. The remaining models differed by habitat, but for comparison purposes, only the 13 models validated for all habitats were used to evaluate the sensitivity of the SCP process to the selection of the SAR model (Table 2). We used an $R^2$ measure that compares the deviance of the non linear SAR model with the deviance of a linear intercept-only model (Kvålseth, 1985) and used these $R^2$ values as indicators of the proportion of the total variation in species richness that is explained (accounted for) by the non linear SAR models.

### 2.4. Conservation targets

For a given habitat, targets were expressed as the proportion of that habitat area required to represent a given proportion of the total number of species occurring within it. The proportion of species is generally set at 70-80% (Desmet & Cowling, 2004; Metcalfe et al., 2013; Davis et al., 2017). In this study, habitat targets were calculated to protect 80% of the fish fauna of each habitat. For the six habitats, 13 fitting SARs models ($\hat{S} = g(A)$; Table 2) were used to derive the habitat-target area $A_t$ corresponding to 80% of the species richness ($S_t$) inferred for the total area of the habitat.
focal habitat (\( S^c_{\text{tot}} \)). We also used multi-model inference for each habitat to estimate corresponding conservation targets.

In addition, we calculated habitat targets following the *log-transformed model* approach developed in Desmet & Cowling (2004), based on simplifications of the logarithmic transformation of the Power-law model (Eq.1 and Eq.2; Arrhenius, 1921):

\[
\hat{S} = cA^z \quad (1)
\]

\[
\log(\hat{S}) = c + z\log(A) \quad (2)
\]

where \( \hat{S} \) = species richness, \( A \) = habitat area, \( z = \) the rate at which new species are encountered in a system (Lomolino, 2000) and \( c \) is a scaling factor that relates to the area of an individual sampling unit, which can be ignored when comparing percentages of species and area (Rondinini & Chiozza, 2010). The habitat-specific \( z \)-parameter was calculated using Eq. (3) as the slope of the log-transformed model (Eq. 2).

\[
z = (y_2 - y_1)/(x_2 - x_1) \quad (3)
\]

where \( y_2 = \log \) (total number of species in a habitat class), \( y_1 = \log \) (mean number of species per sampling point), \( x_2 = \log \) (total area of the habitat class) and \( x_1 = \log \) (mean area of sampling points). Habitat-target area (\( A_t \)) was then calculated to protect \( S_t = 80\% \) of predicted species richness using equation 4 (Desmet & Cowling, 2004; Metcalfe et al., 2013).

\[
\log(A_t) = \log(S_t)/z \quad (4)
\]

### 2.5. Systematic conservation planning

We used the reserve-selection algorithm in the Marxan decision support tool (Ball et al., 2009) to assess the implications of considering habitat-targets from different SAR scenarios on spatial conservation solutions. Marxan uses a heuristic algorithm to minimize the opportunity cost (surface area or socio-economic cost) of the set of Planning Units (PUs, here grid cells) selected as a protected area system while ensuring predefined representativeness for each conservation feature (here SARs habitat-targets). Near-optimal conservation solutions are generated by minimizing an *objective function*, linear combination of three components: (1) total cost of the solution (sum of the opportunity costs of the PUs selected in a given solution); (2) penalty associated with the spatial compactness level of the solution, calibrated using a Boundary Length Modifier (BLM) and (3) sum of Features Penalty Factor (FPF), which represents a penalty cost for each unmet target.

Because the value of the FPF and the BLM depend on the magnitude of conservation targets, it is necessary to adjust, for each SAR scenario independently, the weights in the *objective function* components (i.e. BLM and FPF) to be of similar magnitude, thus preventing any component from outweighing the others (Stewart & Possingham, 2005; Ardron et al., 2008). Here we conducted
analyses without using BLM to allow comparison of conservation solutions based on different SARs habitat-targets. FPF was incrementally increased for each habitat separately until all the habitats would meet their SAR targets. We allocated an equal cost (set to 1) to the 15,539 PUs of the planning region, as our aim was to examine differences in the selection of priority areas among the SAR scenarios and not to design a cost-effective MPA network. We ran Marxan to identify a set of 100 potential MPAs network solutions for each of the 15 SAR scenarios (13 fitting non-linear SARs models, the log-transformed model and the SAR multi-model inference) using 1,000,000 iterations that were necessary for the heuristic algorithm to achieve near-optimal solutions. We then calculated the “irreplaceability” value for each PU as the percentage of solutions that contained the PU. The irreplaceability values emphasize the importance of each PU for achieving conservation targets (Carwardine et al., 2007).

To emphasize the sensitivity of conservation planning to various SAR scenarios, we used two different metrics. We first investigated the spatial overlap between the best conservation solutions arising from the 13 different SAR scenarios (i.e. fitting non-linear SARs models). As such, PUs were identified as priority areas of the least investment risk when selected among all models, or as unsure for conservation plans when their selection depended on the SAR scenario. Second, we derived pairwise comparisons between the 13 SAR scenarios by calculating, for each PU, the difference in irreplaceability. We additionally defined the PUs selected across more than 50% of conservation solutions as "highly irreplaceable" (and essential to effectively meet conservation objectives (Ardron et al., 2008)), and compared them across SAR scenarios. The same analyses were conducted to compare match and mismatches between the multi-model inference and the log-transformed model, the relationship between respective PUs irreplaceability values was projected onto an RGB (Red-Green-Blue) two-dimensional color space and mapped geographically.

### 2.6. Uncertainty propagation analysis

We used the Mantel test (Mantel, 1967) to test for a relationship between dissimilarities in conservation targets and dissimilarities in conservation solutions arising from different SAR models. In the Mantel test, a coefficient of correlation is computed between the two observed dissimilarity matrices and its significance is evaluated by re-calculating the correlation after many random permutations of the rows and columns of one dissimilarity matrix. We calculated the dissimilarity between all pairs of habitat targets for the 13 fitting SAR models using Euclidean distances. A second dissimilarity matrix between the same models was calculated on the basis of PU irreplaceability values, using Euclidean distances as well. A strong and significant Mantel correlation indicates that the dissimilarity in habitat-targets between SARs models is tightly linked
to the dissimilarity in conservation solutions, revealing that SAR uncertainties spread across the full SCP process.

To visualize the dissimilarities between conservation solutions arising from different SAR scenarios we used Nonmetric Multi-Dimensional Scaling (NMDS). We used the function “metaMDS” of the R package “vegan” (Oksanen et al., 2008). Besides, a k-means partitioning was performed on the irreplaceability values for all SAR scenarios, and the resulting grouping was projected in the NMDS space to highlight groupings of SAR scenarios that generated similar conservation solutions.

3. Results

3.1. SAR modeling

Coefficients of determination ($R^2$) for the 13 SAR models ranged widely across habitats (i.e. from 0 to 1 for the gompertz model for the soft slope and rocky intertidal habitats, respectively) and explained on average c. 90% of the total variation in species richness (mean ± standard deviation (SD) = 0.88 ± 0.18) (Table 3). More complex sigmoid models (i.e. weibull3 and mmf) generally explained more variation in species richness than simpler non-asymptotic convex models. Akaike weights varied greatly among models and habitats, and the data empirically supported between 1 and 6 models (i.e. Akaike weights > 0), depending on the habitat (Table 3). For two habitats (i.e. shallow soft and rocky intertidal), model selection probabilities spread almost equally between the mmf and heleg models. For the seagrass habitat, data provided empirical support for the Power-law model and five sigmoid asymptotic models (i.e. mmf, weibull3, weibull4, gompertz and heleg), indicating uncertainty in the selection of the best fitting model. As for the other habitats (i.e. rocky reef, soft slope, and soft shelf), the weibull4 model was selected without uncertainty (i.e. Akaike weights > 0.9) as the best-fitting model.

3.2. Conservation targets

For a given habitat, conservation targets differed substantially among SAR models (Table 4). The difference between the maximum and minimum conservation targets predicted by different SAR models exceeded 30% for all habitats. For example, conservation targets for the soft shelf habitat varied from c. 8% for the gompertz model to c. 38% for the power model. Similarly, a given model provided varying conservation targets among habitats. For example, the monod model identified targets ranging from 18% to 87%, depending on the habitat. Calculating the mean target for a given SAR model across habitats revealed a relationship between target sizes and model characteristics: non-asymptotic models generated the largest targets (i.e. 54% and 59% for the koba and power, respectively), while convex models generated larger targets than sigmoid models (Table 4).
Multi-model inference, weighting all 13 SAR models, determined a consensual conservation target ranging from 9-97% depending on the habitat (38% ± 35.5) whereas the log-transformed model yielded relatively less ambitious targets, ranging between 12-40% among habitats (21% ± 9.7). The differences between the conservation targets estimated by these two methods can exceed 40% when the multi-model inference generated the most ambitious targets (i.e. seagrass and rocky reef habitats). Otherwise, targets estimated by the log-transformed model exceeded those estimated by the multi-model by c. 8% depending on the habitat (Table 3).

3.3 Conservation solutions

Although SAR scenarios that provided more ambitious targets ultimately led to more spatially extensive conservation networks, conservation solutions for different SAR scenarios showed different spatial distributions beyond their difference in size. While the less ambitious SAR scenario (i.e. gompertz) provided an average target of c. 26% across habitats, consensual priority areas among all SAR scenarios did not exceed 1% of the study area (Fig. 3 and Fig. A2): the PUs selected by models estimating moderately ambitious targets were not necessarily selected by models that estimate larger targets.

The spatial distribution of irreplaceability differed among SAR scenarios (Fig. 4). PUs that are “highly irreplaceable” for achieving targets for a given SAR scenario are not likely to be so for another SAR scenario. In all pairwise SAR scenarios comparisons, between c. 2%-50% of “Highly irreplaceable” PUs differed, while only c. 2% presented maximum irreplaceability across all SAR scenarios (Fig. A3; Table A4).

The comparison between irreplaceability values for the multi-model SAR and the log-transformed model revealed substantial differences over 10% of the study area (Fig. 5). Several patches of "highly irreplaceable” PUs under the multi-model SAR scenario were spread across the Balearic Islands (Spain), Sicily (Italy) and Corsica (France), covering mainly seagrass and rocky reef habitats; irreplaceability values for the corresponding PUs under the log-transformed model were systematically low.

3.4. SARs uncertainties propagation

A Mantel r-statistic of 0.97 (p-value=0.01) suggested a strong, positive and significant correlation between SARs habitat-targets dissimilarities and conservation solutions dissimilarities, revealing that differences in SARs inferences propagate through the SCP process and lead to different conservation solutions. The two-dimensional NMDS (Fig. 6) produced high congruence between the observed and ordinated distances (non-metric fit $R^2 = 0.99$) and a stress value of 0.02 distinguished clearly conservation solutions of the log-transformed model from other SAR scenarios. A k-means clustering distinguishing five groups of SAR models partitioned the models
as follows. A first group of three non-asymptotic models (power, koba and P2). A second group
gathering four sigmoid asymptotic models (weibull3, weibull4, heleg and mnf) and the multi-model
inference. A third group including both convex asymptotic and sigmoid models with 2-3 parameters
(i.e. monod, epm2, negexpo and chapman) while the gompertz and betap models were aggregated
into a fourth group. The log-transformed model was isolated in a fifth group. Note that all models of
the second group were most frequently selected as the best models among habitats and that models
of the third group had null Akaike weights across all habitats (Table 3).

4. Discussion

As expected from the literature evidence, the shapes of coastal Mediterranean SARs varied greatly
among habitats, including convex, exponential and sigmoid tendencies (Fig. 2). The best fitting
SAR model differed by habitat and although the mean $R^2$ was relatively high, the goodness-of-fit of
the 13 models varied greatly among the six marine habitat types we considered. Hence, assuming
any of the numerous models described in the literature to provide an optimal fit for all the habitat
data sets analysed here would have led to spurious inferences (Smith, 2010). Indeed, our results
confirm that models that fit well for one habitat may not do so for other habitats (Connor & McCoy,
1979; Guilhaumon et al., 2008; Stiles & Scheiner, 2007; Triantis et al., 2012).

Furthermore, based on AICc weights, we detected a substantial uncertainty in the selection of the
best SAR model for three habitats, with several models (mostly sigmoid asymptotic) supported by
empirical SARs. For the other habitats, a sigmoid asymptotic model was selected as best-fitting
SAR (Table 3). Our findings of substantial support for sigmoid asymptotic SARs models for
Mediterranean habitats, would not be considered in classical analyses assuming a priori the Power-
law SAR or its log-transformed model. As pertaining from the versatility of the SAR, the
justification for particular SAR models is problematic in the literature, and when the data allows,
testing among existing SAR models is advocated (Smith et al., 2010; Scheiner et al., 2011; Davis et
al., 2017). Furthermore, when numerous models show equivalent support for empirical SARs,
testing a range of SARs models is prospectively essential to allow a multi-model inference robust to
the assumption of a particular SAR model (Burnham and Anderson, 2003).

The comparison of conservation habitat-targets revealed substantial differences among SAR
scenarios. For example, while the best-fitting model for soft shelf habitat (i.e. weibull4) predicted a
target of 13%, other models predicted markedly different ranges of targets (e.g. from 7% to 38% for
gompertz and power models, respectively). In general, for the threshold of 80% of species richness
coverage employed here, non-asymptotic models predicted the largest targets, and convex models
predicted larger targets than sigmoid models. The model used, whether convex or sigmoid or with an asymptote, significantly influenced the amount of each habitat to be secured in protected areas. Using multi-model inference resulted in a mean conservation target of \( c. 38\% \) over the entire study area. This conservation target, devised to protect 80\% of the species within each habitat, as recommended by international conservation institutions (Desmet & Cowling, 2004; Metcalfe et al., 2013; Davis et al., 2017), and combined with the 10\% coverage threshold to construct SARs, promotes both species representation and persistence within each habitat. Accordingly, in the Mediterranean Sea, the conservation target of 10\% of the marine area set by the CBD (2010) would leave some key habitats under-protected to represent 80\% of the species (i.e. seagrass and rocky reefs). Additionally, this result supports the recommendations of the 2014 World Parks Congress, which called for protection of at least 30\% of marine areas, as in previous studies (O’Leary et al., 2016). On the other hand, the use of the log-transformed model is justified on the assumption that the power model is the most appropriate at fitting all SAR datasets. However, our results show that a power model, fitted as a non-linear regression by minimizing the residual sum of squares, was never selected as the best fitting one, except for the seagrass habitat, where the power model holds a share of the statistical information in explaining empirical SAR. If employed, habitat-conservation targets estimated with the log-transformed model would provide insufficient protection level for some key habitats in the life cycle of Mediterranean fish species (i.e. seagrass and rocky reef habitats). On the flip side, when targets estimated with the log-transformed model exceeded those estimated by the multi-model SAR, the protection of the corresponding habitats may be overstated, leading to waste of scarce conservation resources that could be used to protect more needful habitats. Overall, our investigations revealed strong differences in SAR scenarios on the setting of habitat-specific conservation targets, with the log-transformed model providing conservation targets diverging markedly from data-driven SAR inferences. As such, beyond providing ecologically relevant conservation targets, considering habitat specificities by modeling empirical SARs is relevant for a parsimonious allocation of conservation funds.

Importantly, beyond an effect on the magnitude of conservation targets, our study revealed that the spatial distribution of selected priority areas depends on the SAR scenario used (Fig. 3, Fig. 4). As shown here, conservation solutions for less ambitious targets are not always subsets of conservation solutions for more ambitious targets (Fig. 3), confirming that the spatial distribution of conservation solutions is target-specific (Margules & Pressey, 2000). Pairwise comparison between irreplaceability values under different SAR scenarios revealed a maximal congruence of 50\% of "highly irreplaceable" PUs (Fig.A3; Table A4). Our analyses highlighted that dissimilarities in habitat-targets associated with alternative SAR scenarios propagate throughout the SCP process and
lead to correlated dissimilarities in conservation outcomes. This is of special concern given the wide range of magnitude for targets estimated under different SAR scenarios (exceeding 30% for e.g. rocky reefs or soft slopes; Table 4). The clustering of SAR scenarios in the NMDS based on irreplaceability values (Fig. 6) shows that models with similar shapes were likely to generate similar conservation solutions (e.g. the group of sigmoid models vs the group of convex models in Fig. 6). This suggests the necessity to consider a wide range of SAR shapes when estimating habitat conservation targets and solutions. On another note, the overlap among the 13 best conservation solutions, each based on a different SAR scenario, revealed that priority areas of the least investment risk (i.e. consensual areas among all SAR models) occupied only 1% of the total study area. Most of these consensual PUs are already located in established MPAs (e.g. Cabrera National Park (Spain), reserve of Lavezzi islands (France), Egadi MPA (Italy), MPA of “Penisola del Sinis - Isola di Mal di Ventre” (Italy) and MPA of Ustica Island (Italy); Fig.5), indicating that the current Mediterranean MPAs system provides a valuable starting point to establish a relevant conservation network for coastal habitat conservation (Guilhaumon et al. 2015).

Comparing the multi-model SAR and the log-transformed model, SCP outcomes showed strong effects of the method on the spatial distribution of irreplaceability values. Particular conservation implications of using both methods can be highlighted by scrutinizing areas with larger differences in irreplaceability values (Fig. 5). Most of these areas, such as the Balearic Islands in Spain, the Gulf of Gabes in Tunisia, and the coasts of Sicily in Italy and Corsica in France, are predicted to be irreplaceable under the multi-model SAR scenario but not by the log-transformed model. Some of these priority areas overlap with established MPAs, while most are located in unprotected areas, such as the Gulf of Gabes, known to be a shelter for one of the most extensive Posidonia oceanica meadows in the Mediterranean sea (Batisse & Jeudy de Grissac, 1998) and a fish functional diversity hotspot (Mouillot et al., 2011).

In conclusion, our findings highlight the potential consequences of the choice of a particular SAR model on the different stages of habitat-based SCP. SAR uncertainties propagate through the SCP process, confirming our theoretical anticipation that ignoring SARs specificities when setting habitat conservation targets may lead either to a waste of resources limited by socio-economic constraints or to the under-representation of biological conservation features. When the data allows, a shift from an a priori defined SAR model or uniform targets towards a data-based approach would provide stakeholders with better guidance on the appropriate strategies for habitats conservation. Failing to do so may well lead to uncertain conservation outcomes, as exemplified herein the identification of conservation plans among coastal Mediterranean habitats.


Ceballos, G., Ehrlich, P.R., Dirzo, R., 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the National Academy of Sciences 114, E6089–E6096.


Deane, D.C., Fordham, D.A., He, F., Bradshaw, C.J., 2017. Future extinction risk of wetland plants is higher from individual patch loss than total area reduction. Biological Conservation 209, 27–33.


Fattorini, S. 2007. To fit or not to fit? A poorly fitting procedure produces inconsistent results when the species–area relationship is used to locate hotspots. Biodiversity and Conservation 16, 2531-2538.


analysis: priority regions for expanding the global protected-area network. AIBS Bulletin 54, 1092–1100.


Figure captions:

**Figure 1**: Description of the different steps of the workflow for habitats conservation planning using Species-Area Relationships (SARs). Gray points indicate the empirical SAR and SAR scenarios are represented by solid lines. On each plot in *Step 2*, horizontal and vertical dashed lines indicate respectively the threshold of 80% of total predicted species richness and the corresponding habitat area to protect. In *Step 3*, the colors of the selected planning units correspond to the color of the scenario in *Step 2*.

**Figure 2**: Empirical (gray points) and fitted (lines) SAR for six Mediterranean coastal habitats: (a) rocky reef, (b) shallow soft, (c) soft slope, (d) soft shelf, (e) seagrass and (f) rocky intertidal. Model codes as in Table 1.

**Figure 3**: The overlap between the best conservation solutions under 13 SAR scenarios. The color of each planning unit indicate the number of SAR scenarios among which it was part of the best conservation solution. Locations of five marine protected areas: 1, Cabrera National Park; 2, reserve of Lavezzi Islands; 3, Egadi MPA; 4, MPA of Ustica Island and 5, MPA of “Penisola del Sinis - Isola di Mal di Ventre”.

**Figure 4**: Irreplaceability values of planning units in the Mediterranean Sea under different SAR scenarios: (a) power, (b) epm2, (c) P2, (d) koba, (e) mmf, (f) monod, (g) negexpo, (h) chapman, (i) weibull3, (j) gompertz, (k) weibull4, (l) betap, (m) heleg, (n) the log-transformed model and (o) the multi-model inference. Model codes as in Table 1.

**Figure 5**: Matches and mismatches between the multi-model inference and the *log-transformed model* SAR scenarios. Planning units colors are produced by projecting irreplaceability values within a two-dimensional RGB (Red-Green-Blue) color space as indicated in the inset.

**Figure 6**: Non-metric multi-dimensional scaling ordination of the irreplaceability values of SAR scenarios. Colors represent groups identified using a k-means clustering method. The point “multi-model” represents the multi-model inference, while “log-transform” indicates the *log-transformed model*.
Table 1: Characteristics of six coastal habitat types in the Mediterranean Sea used in this study: total area, species richness, mean habitat area per planning unit (PU) and mean number of species per planning unit.

<table>
<thead>
<tr>
<th>Habitat type</th>
<th>Area (km²)</th>
<th>No. of species</th>
<th>No. of PUs</th>
<th>Mean habitat area per PU (km²)</th>
<th>Mean no. of species per PU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rocky reef</td>
<td>37689.248</td>
<td>576</td>
<td>837</td>
<td>45.02</td>
<td>207</td>
</tr>
<tr>
<td>Shallow soft</td>
<td>190334.881</td>
<td>595</td>
<td>4462</td>
<td>42.65</td>
<td>203</td>
</tr>
<tr>
<td>Soft slope</td>
<td>87948.256</td>
<td>589</td>
<td>11349</td>
<td>77.49</td>
<td>62</td>
</tr>
<tr>
<td>Soft shelf</td>
<td>225536.802</td>
<td>593</td>
<td>5304</td>
<td>42.52</td>
<td>168</td>
</tr>
<tr>
<td>Seagrass</td>
<td>3158.366</td>
<td>520</td>
<td>286</td>
<td>11.04</td>
<td>285</td>
</tr>
<tr>
<td>Rocky intertidal</td>
<td>18522.185</td>
<td>594</td>
<td>2844</td>
<td>6.51</td>
<td>229</td>
</tr>
</tbody>
</table>

Table 2: The analytical formula, shape, presence/absence of an asymptote and number of parameters of the 20 models analyzed. S is species richness, A is the area, and c, d, f, and z are fitted parameters. Bold font identifies the 13 models selected for conservation planning analyses.

<table>
<thead>
<tr>
<th>Model name</th>
<th>Code</th>
<th>Formula</th>
<th>Parameters</th>
<th>Shape</th>
<th>Asymptote?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>linear</td>
<td>( S = c + zA )</td>
<td>2</td>
<td>Linear</td>
<td>x</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td>power</td>
<td>( S = cA^2 )</td>
<td>2</td>
<td>Convex</td>
<td>x</td>
</tr>
<tr>
<td>Power Rosenzweig</td>
<td>power_R</td>
<td>( S = k + cA^z )</td>
<td>3</td>
<td>Convex</td>
<td>x</td>
</tr>
<tr>
<td>Extended power-law</td>
<td>epm1</td>
<td>( S = cA^{2-d} )</td>
<td>3</td>
<td>Sigmoid</td>
<td>x</td>
</tr>
<tr>
<td>Extended</td>
<td>epm2</td>
<td>( S = cA^{z-(c/A)} )</td>
<td>3</td>
<td>Sigmoid</td>
<td>x</td>
</tr>
<tr>
<td>Persistence function 1</td>
<td>P1</td>
<td>( S = cA^ze^{-dA} )</td>
<td>3</td>
<td>Convex</td>
<td>x</td>
</tr>
<tr>
<td>Persistence</td>
<td>P2</td>
<td>( S = cA^ze^{-dA} )</td>
<td>3</td>
<td>Sigmoid</td>
<td>x</td>
</tr>
<tr>
<td>function 2</td>
<td>( S = c + z \log A )</td>
<td>( S = \log(1 + A/z) )</td>
<td>( S = d/(1 + cA^{-1}) )</td>
<td>( S = d/(1 + cA^{-z}) )</td>
<td>( S = c/(f + A^{-z}) )</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>Exponential</td>
<td>expo</td>
<td>2</td>
<td>Convex</td>
<td>Convex</td>
<td>Convex</td>
</tr>
<tr>
<td>Kobayashi logarithmic</td>
<td>koba</td>
<td>2</td>
<td>Convex</td>
<td>( \times )</td>
<td>( \times )</td>
</tr>
<tr>
<td>Monod</td>
<td>monod</td>
<td>2</td>
<td>Convex</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Morgan-Mercer-Foldin</td>
<td>mmf</td>
<td>3</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Logistic</td>
<td>heleg</td>
<td>3</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Negative exponential</td>
<td>negexpo</td>
<td>2</td>
<td>Convex</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Chapman–Richards</td>
<td>chapman</td>
<td>3</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Weibull 3</td>
<td>weibull3</td>
<td>3</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Weibull4</td>
<td>weibull4</td>
<td>4</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Asymptotic regression</td>
<td>asymp</td>
<td>3</td>
<td>Convex</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Rational</td>
<td>ratio</td>
<td>3</td>
<td>Convex</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Gompertz</td>
<td>gompertz</td>
<td>3</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
<tr>
<td>Beta-P</td>
<td>betap</td>
<td>4</td>
<td>Sigmoid</td>
<td>( \checkmark )</td>
<td>( \checkmark )</td>
</tr>
</tbody>
</table>
Table 3: Model selection and goodness-of-fit for 13 SAR models in six Mediterranean coastal habitats. $R^2$ is the coefficient of determination, AICc is the Akaike Information Criterion corrected for small samples and $\omega$ is Akaike model weight, which reflects the probability that a model is the best at fitting the data. Model codes as in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Rocky reef</th>
<th>Shallow soft</th>
<th>Soft slope</th>
<th>Soft shelf</th>
<th>Seagrass</th>
<th>Rocky intertidal</th>
</tr>
</thead>
<tbody>
<tr>
<td>power</td>
<td>0.94</td>
<td>0.54</td>
<td>0.9</td>
<td>0.68</td>
<td>0.98</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>6 904</td>
<td>54 349</td>
<td>89 174</td>
<td>48 726</td>
<td>1 834</td>
<td>31 112</td>
</tr>
<tr>
<td>epm2</td>
<td>0.99</td>
<td>0.83</td>
<td>0.97</td>
<td>0.93</td>
<td>0.97</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>5 639</td>
<td>41 365</td>
<td>76 515</td>
<td>40 224</td>
<td>1 962</td>
<td>28 650</td>
</tr>
<tr>
<td>P2</td>
<td>0.99</td>
<td>0.85</td>
<td>0.94</td>
<td>0.86</td>
<td>0.98</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>5 453</td>
<td>40 712</td>
<td>89 178</td>
<td>46 349</td>
<td>1 849</td>
<td>30 407</td>
</tr>
<tr>
<td>koba</td>
<td>0.94</td>
<td>0.63</td>
<td>0.95</td>
<td>0.86</td>
<td>0.98</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>6 919</td>
<td>45 245</td>
<td>82 590</td>
<td>46 380</td>
<td>1 836</td>
<td>29 282</td>
</tr>
<tr>
<td>mmf</td>
<td>1</td>
<td>14 900</td>
<td>0.48</td>
<td>0.89</td>
<td>0.98</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>4 547</td>
<td>14 000</td>
<td>26 153</td>
<td>25 292</td>
<td>1 836</td>
<td>28 207</td>
</tr>
<tr>
<td>monod</td>
<td>0.94</td>
<td>0.78</td>
<td>0.97</td>
<td>0.86</td>
<td>0.98</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>6 919</td>
<td>42 548</td>
<td>76 452</td>
<td>46 728</td>
<td>1 836</td>
<td>28 207</td>
</tr>
<tr>
<td>negexpo</td>
<td>0.94</td>
<td>0.87</td>
<td>0.99</td>
<td>0.87</td>
<td>0.98</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>6 919</td>
<td>40 025</td>
<td>67 774</td>
<td>37 477</td>
<td>1 836</td>
<td>28 207</td>
</tr>
<tr>
<td>chapman</td>
<td>0.94</td>
<td>0.87</td>
<td>0.98</td>
<td>0.86</td>
<td>0.98</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>6 921</td>
<td>40 027</td>
<td>72 379</td>
<td>37 479</td>
<td>1 836</td>
<td>28 207</td>
</tr>
<tr>
<td>weibull3</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>4 961</td>
<td>19 295</td>
<td>64 386</td>
<td>20 227</td>
<td>1 836</td>
<td>15 020</td>
</tr>
<tr>
<td>gompertz</td>
<td>0.81</td>
<td>0.06</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>8 000</td>
<td>49 994</td>
<td>115 925</td>
<td>47 084</td>
<td>1 836</td>
<td>15 387</td>
</tr>
<tr>
<td>weibull4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>4 466</td>
<td>16 235</td>
<td>40 626</td>
<td>20 182</td>
<td>1 838</td>
<td>15 271</td>
</tr>
<tr>
<td>betap</td>
<td>0.99</td>
<td>0.99</td>
<td>0.51</td>
<td>0.9</td>
<td>0.98</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>4 963</td>
<td>23 323</td>
<td>107 688</td>
<td>46 424</td>
<td>2 425</td>
<td>14 999</td>
</tr>
<tr>
<td>heleg</td>
<td>1</td>
<td>1</td>
<td>0.52</td>
<td>0.99</td>
<td>0.98</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>4 547</td>
<td>14 899</td>
<td>71 464</td>
<td>26 153</td>
<td>1 836</td>
<td>14 581</td>
</tr>
</tbody>
</table>
Table 4: Conservation targets, expressed as the percentage of each habitat required to protect 80% of species according to 13 non-linear SAR models, the log-transformed model and the multi-model SAR. Mean habitat targets for a given SAR model across habitats; Mean habitat targets for a given habitat across SAR scenarios; ranges of targets estimated across SAR scenarios for each habitat (maximum minus minimum target).

<table>
<thead>
<tr>
<th>Habitat</th>
<th>Power</th>
<th>EPM2</th>
<th>P2</th>
<th>Koba</th>
<th>MMF</th>
<th>Monod</th>
<th>Neg expo</th>
<th>Chapman</th>
<th>Weibull3</th>
<th>Gompertz</th>
<th>Weibull4</th>
<th>Betap</th>
<th>Heleg</th>
<th>Multi-model</th>
<th>Log-transformed model</th>
<th>Mean % target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rocky reef</td>
<td>83.02</td>
<td>74.34</td>
<td>73.06</td>
<td>81.45</td>
<td>62.84</td>
<td>81.32</td>
<td>81.45</td>
<td>81.32</td>
<td>61.36</td>
<td>36.62</td>
<td>64.13</td>
<td>61.36</td>
<td>62.84</td>
<td>64.13</td>
<td>22.95</td>
<td>66.14</td>
</tr>
<tr>
<td>Shallow soft</td>
<td>34.15</td>
<td>18.23</td>
<td>17.18</td>
<td>31.95</td>
<td>9.31</td>
<td>13.46</td>
<td>18.47</td>
<td>13.46</td>
<td>9.34</td>
<td>0.41</td>
<td>9.37</td>
<td>9.52</td>
<td>9.31</td>
<td>9.31</td>
<td>17.42</td>
<td>14.72</td>
</tr>
<tr>
<td>Soft slope</td>
<td>57.69</td>
<td>38.28</td>
<td>57.69</td>
<td>51</td>
<td>37.21</td>
<td>43.63</td>
<td>43.39</td>
<td>33.39</td>
<td>34.17</td>
<td>0.06</td>
<td>33.35</td>
<td>13.07</td>
<td>37.21</td>
<td>33.35</td>
<td>39.61</td>
<td>36.44</td>
</tr>
<tr>
<td>Seagrass</td>
<td>97.42</td>
<td>95.95</td>
<td>97.27</td>
<td>87.57</td>
<td>97.42</td>
<td>87.8</td>
<td>87.57</td>
<td>87.57</td>
<td>97.42</td>
<td>98.98</td>
<td>97.42</td>
<td>85.8</td>
<td>97.42</td>
<td>97.42</td>
<td>12.18</td>
<td>88.34</td>
</tr>
<tr>
<td>Rocky intertidal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean % target</td>
<td>45.67</td>
<td>26.55</td>
<td>24.85</td>
<td>41.01</td>
<td>11.94</td>
<td>19.44</td>
<td>26.92</td>
<td>19.44</td>
<td>11.77</td>
<td>11.94</td>
<td>11.94</td>
<td>11.77</td>
<td>11.94</td>
<td>11.94</td>
<td>21.56</td>
<td>20.16</td>
</tr>
</tbody>
</table>
SUPPORTING INFORMATION

**Figure A1:** Habitat types in the Mediterranean Sea: (a) rocky reef, (b) shallow soft, (c) soft slope, (d) soft shelf, (e) seagrass and (f) rocky intertidal.

**Figure A2:** Best solutions for each SAR scenario: (a) power, (b) epm2, (c) P2, (d) koba, (e) mmf, (f) monod, (g) negexpo, (h) chapman, (i) weibull3, (j) gompertz, (k) weibull4, (l) betap, (m) heleg, (n) the log-transformed model and (o) the multi-model inference. Model codes as in Table 1.

**Figure A3:** Pairwise comparison of irreplaceability values of planning units under 13 SAR scenarios. Upper triangle illustrates differences in irreplaceability values between two given SAR scenarios. Colors of planning units in scatter plots (lower triangle) and maps indicate the differences in irreplaceability values between two SAR scenarios.
**Table A4**: Pairwise comparison of matches and mismatches of “highly irreplaceable” planning units under 13 SAR scenarios: The percentage of “highly irreplaceable” planning units selected under the SAR scenario and not under the other are given in upper triangle; the percentage of “highly irreplaceable” planning units selected under both SAR scenarios are represented in lower triangle.

<table>
<thead>
<tr>
<th>Model</th>
<th>power</th>
<th>epm2</th>
<th>P2</th>
<th>koba</th>
<th>mmf</th>
<th>monod</th>
<th>negexpo</th>
<th>chapman</th>
<th>weibull3</th>
<th>gompertz</th>
</tr>
</thead>
<tbody>
<tr>
<td>power</td>
<td>45.82</td>
<td>6.7</td>
<td>13.59</td>
<td>46.97</td>
<td>33.95</td>
<td>47.31</td>
<td>50.26</td>
<td>49.95</td>
<td>53.14</td>
<td></td>
</tr>
<tr>
<td>epm2</td>
<td>8.04</td>
<td>42.56</td>
<td>32.59</td>
<td>5.67</td>
<td>12.53</td>
<td>4.96</td>
<td>4.2</td>
<td>4.64</td>
<td>6.28</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>49.16</td>
<td>18.33</td>
<td>14.32</td>
<td>43.45</td>
<td>30.96</td>
<td>44.06</td>
<td>47</td>
<td>46.43</td>
<td>49.62</td>
<td></td>
</tr>
<tr>
<td>koba</td>
<td>7.92</td>
<td>17.29</td>
<td>4.27</td>
<td>33.77</td>
<td>22.87</td>
<td>33.66</td>
<td>36.42</td>
<td>36.41</td>
<td>39.49</td>
<td></td>
</tr>
<tr>
<td>mmf</td>
<td>39.78</td>
<td>5.68</td>
<td>4.37</td>
<td>4.07</td>
<td>13.55</td>
<td>5.01</td>
<td>4.25</td>
<td>3.78</td>
<td>5.22</td>
<td></td>
</tr>
<tr>
<td>monod</td>
<td>7.74</td>
<td>6.72</td>
<td>3.52</td>
<td>3.64</td>
<td>1.64</td>
<td>13</td>
<td>14.42</td>
<td>14.8</td>
<td>17.25</td>
<td></td>
</tr>
<tr>
<td>negexpo</td>
<td>37.75</td>
<td>4.73</td>
<td>4.25</td>
<td>3.31</td>
<td>1.63</td>
<td>3.9</td>
<td>3.1</td>
<td>3.84</td>
<td>5.22</td>
<td></td>
</tr>
<tr>
<td>chapman</td>
<td>6.98</td>
<td>6.6</td>
<td>3.95</td>
<td>1.79</td>
<td>1.87</td>
<td>3.55</td>
<td>3.13</td>
<td>2.28</td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td>weibull3</td>
<td>4.49</td>
<td>6.63</td>
<td>4.56</td>
<td>1.78</td>
<td>4.25</td>
<td>3.34</td>
<td>3.19</td>
<td>3.08</td>
<td>2.79</td>
<td></td>
</tr>
<tr>
<td>gompertz</td>
<td>7</td>
<td>4.18</td>
<td>3.86</td>
<td>1.8</td>
<td>3.79</td>
<td>3.54</td>
<td>3.08</td>
<td>1.61</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>weibull4</td>
<td>6.64</td>
<td>5.9</td>
<td>4.56</td>
<td>1.67</td>
<td>4.26</td>
<td>1.78</td>
<td>3.11</td>
<td>3.09</td>
<td>6.76</td>
<td></td>
</tr>
<tr>
<td>betap</td>
<td>18.64</td>
<td>4.4</td>
<td>4.38</td>
<td>1.78</td>
<td>4.08</td>
<td>3.24</td>
<td>3.1</td>
<td>7.1</td>
<td>4.38</td>
<td></td>
</tr>
<tr>
<td>heleg</td>
<td>6.73</td>
<td>4.05</td>
<td>3.82</td>
<td>1.65</td>
<td>3.69</td>
<td>3.13</td>
<td>3.08</td>
<td>4.49</td>
<td>5.71</td>
<td></td>
</tr>
</tbody>
</table>
Step 1: Empirical SARs
Species richness and habitat area are tallied for increasing surface level (number of planning units) for each habitat of study area.

Input data
Species distributions
Habitats map

Step 2: SAR models and conservation targets
For each habitat, the empirical SAR is fitted according to different SAR scenarios (here 3 SAR models and a multi-model inference). The corresponding conservation targets are then inferred to protect a given percentage (here 80%) of the total species richness of each habitat.

SAR Scenario 1
SAR Scenario 2
SAR Scenario 3
multi-model SAR scenario

Step 3: Conservation planning
A reserve-selection algorithm (here Marxan) is used to generate a conservation solution based on each SAR scenario.
Irreplaceability of log-transformation method

Irreplaceability of multi-model inference