Can We Avoid Tacit Trade-Offs between Flexibility and Efficiency in Systematic Conservation Planning? The Mediterranean Sea as a Case Study
Sabrine Drira, Frida Ben Rais Lasram, Tarek Hattab, Yunne-Jai Shin, Amel Ben Rejeb Jenhani, Francois Guilhaumon

To cite this version:
Sabrine Drira, Frida Ben Rais Lasram, Tarek Hattab, Yunne-Jai Shin, Amel Ben Rejeb Jenhani, et al.. Can We Avoid Tacit Trade-Offs between Flexibility and Efficiency in Systematic Conservation Planning? The Mediterranean Sea as a Case Study. DIVERSITY-BASEL, 2021, 14 (1), pp.9. 10.3390/d14010009. hal-03622232

HAL Id: hal-03622232
https://hal.umontpellier.fr/hal-03622232
Submitted on 2 Jun 2022

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 4.0 International License
Can We Avoid Tacit Trade-Offs between Flexibility and Efficiency in Systematic Conservation Planning? The Mediterranean Sea as a Case Study

Sabrine Drira 1,2,*, Frida Ben Rais Lasram 3, Tarek Hattab 2,*, Yunne-Jai Shin 2, Amel Ben Rejeb Jenhani 1 and François Guilhaumon 2,4,8

1 Laboratoire de Recherche Ecosystèmes et Ressources Aquatiques (LR21AGR01), Institut National Agronomique de Tunisie, Université de Carthage, 43 Avenue Charles-Nicoles, 1082 Tunis-Mahrajene, Tunisia; sabrinedrira@gmail.com (S.D.); jenhani.amel@gmail.com (A.B.R.J.)
2 MARBEC, IRD, IFREMER, CNRS, Université de Montpellier, 34095 Montpellier, France; Tarek.Hattab@ifremer.fr (T.H.); yunne-jai.shin@ird.fr (Y.-J.S.)
3 Université Littoral Côte d’Opale, Université Lille, CNRS, IRD, UMR 8187, LOG, Laboratoire d’Océanologie et de Géosciences, 62930 Wimereux, France; frida.lasram@univ-littoral.fr
4 UMR 9220 EN TroPIE, IRD, Université de la Réunion, Université de la Nouvelle-Calédonie, IFREMER, CNRS, 97744 La Réunion, France
* Correspondence: francois.guilhaumon@ird.fr

Abstract: Species distribution models (SDMs) provide robust inferences about species-specific site suitability and are increasingly used in systematic conservation planning (SCP). SDMs are subjected to intrinsic uncertainties, and conservation studies have generally overlooked these. The integration of SDM uncertainties in conservation solutions requires the development of a suitable optimization algorithm. Exact optimization algorithms grant efficiency to conservation solutions, but most of their implementations generate a single binary and indivisible solution. Therefore, without variation in their parameterization, they provide low flexibility in the implementation of conservation solutions by stakeholders. Contrarily, heuristic algorithms provide such flexibility, by generating large amounts of sub-optimal solutions. As a consequence, efficiency and flexibility are implicitly linked in conservation applications: mathematically efficient solutions provide less flexibility, and the flexible solutions provided by heuristics are sub-optimal. To avoid this trade-off between flexibility and efficiency in SCP, we propose a reserve-selection framework, based on exact optimization combined with a post-selection of SDM outputs. This reserve-selection framework provides flexibility and addresses the efficiency and representativeness of conservation solutions. To exemplify the approach, we analyzed an experimental design, crossing pre- and post-selection of SDM outputs versus heuristics and exact mathematical optimizations. We used the Mediterranean Sea as a biogeographical template for our analyses, integrating the outputs of eight SDM techniques for 438 fish species.

Keywords: systematic conservation planning; species distribution models; exact optimization; heuristics; efficiency; flexibility; Mediterranean Sea; fish

1. Introduction

In response to declining wildlife populations and concerns about the anthropogenic footprint on biodiversity [1,2], systematic conservation planning (SCP) has rapidly emerged as a viable approach for identifying reserve networks (sets of sites where human activities are limited, also called conservation solutions) [3]. To efficiently meet transparently defined objectives for biodiversity conservation, SCP builds on principles, in order to set adequate conservation objectives (ensuring the protected area system protects a sample of all biodiversity present), identify the most cost-effective conservation solutions, and provide stakeholders with flexibility (the ability to choose among alternative sites to meet conservation objectives) when engaging conservation actions on the ground [4].
A fundamental step in SCP is setting quantitative targets to adequately represent and protect biodiversity [3,4]. Conservation targets represent ‘the minimum amount of a particular biodiversity feature that we would like to conserve through one or several conservation actions’. This step is generally addressed by setting quantitative targets for the different conservation features concerned (conservation features are most often species, but can be functional types [5] or habitats [6,7]). Ideally, conservation targets should relate directly to the probability of species persistence [8]. As these kinds of data are usually unavailable or incomplete [9], an indirect approach is to secure a variable proportion of species range sizes, such that species with restricted ranges (who present higher extinction risk because of lower local abundance and higher demographic stochasticity) have more ambitious targets [10]. However, our knowledge about species distribution is limited [12–15]. To overcome this shortfall, species distribution models (SDMs, also called ecological niche models) have been proposed as a way to provide a robust inference about species distributions [16]. These models use various correlative statistical methods to associate observed species occurrences with environmental predictor variables and infer species probabilities of occurrence over the whole study area (hereafter ‘distribution scenario’) [16]. As the different statistical methods used to build SDMs are differently sensitive to geographical range properties [17], including species prevalence [18], they yield variable predictions about species distributions (e.g., [19–21]). In turn, reserve selection, affected by SDM commission (false species presences) and omission (false species absences) errors, can become uncertain [22,23]. For this reason, as the use of SDMs is growing in SCP applications (16, Appendix A), the adequacy of conservation solutions, and our ability to evaluate them, increasingly depends on the accuracy and the variability of SDM outputs [22–25].

An efficient reserve network is one that meets its objectives at the lowest possible cost (or smallest area of the conservation solution [3,26]). The efficiency of conservation solutions depends on the particular reserve-selection algorithm used to solve the conservation problem, or on its parameterization. A large body of research has shown that exact optimization algorithms (implemented via. mathematical programming) give very precise control over the gap to optimality of conservation solutions (i.e., degree of solution efficiency), allowing avoiding a waste of scarce conservation resources [27–29]. However, the current implementations of optimal reserve-selection algorithms generally yield a single binary and indivisible solution (e.g., [30,31]) and provide no information regarding the relative importance of selected sites, nor about the opportunity that sites outside the optimal solution represent for achieving conservation targets. Therefore, optimal algorithms provide low flexibility in the implementation of conservation solutions by stakeholders [32]. However, flexibility, the opportunity to choose alternative sites to achieve conservation objectives, is central to SCP, empowering stakeholders to schedule conservation actions and negotiate the inclusion of sites having particular ecological, social, or political interests [25]. In the pursuit of flexibility, random-based heuristics (integrating a stochastic component into their search algorithms, e.g., Marxan [33]) providing many sub-optimal solutions (containing more sites than necessary or capturing less conservation features than feasible with the same amount of resources) have been widely supported [34,35]. Indeed, heuristics quantify the potential contribution of all possible sites to the achievement of conservation objectives. This potential contribution, called irreplaceability, is calculated as the selection frequency of each planning unit across all solutions. Hence, the flexibility associated with heuristics-based conservation solutions arises from the availability of many alternative and sub-optimal conservation solutions. As such, in the current context, the two principles of efficiency and flexibility are implicitly linked in most current conservation applications, with the most mathematically efficient solutions being inflexible and the flexible solutions provided by random-based heuristics being sub-optimal.

Here, after screening the literature for SCP applications that used SDMs as inputs, we identified three different strategies used to deal with both SDM uncertainties and reserve-selection efficiency (Appendix A). Each of these strategies results from a particular treatment of SDM uncertainties (whether they are taken into account or not) and the
targeted reserve selection efficiency (i.e., sub-optimal vs. optimal solutions) (Figure 1). Most SCP applications have overlooked the implications of SDM uncertainties and have focused on a single and a priori chosen statistical method (Appendix A, [36]). Others studies have considered a range of models to predict species distributions and used only the best model identified in terms of predictive performance (e.g., [37–39]). This risks the possibility that a conservation solution based on one particular distribution scenario might not meet the required target identified under a different good model [40]. Proceeding further, some studies used ensemble approaches to summarize the information across the multiple model predictions (Figure 1B,D), this takes the uncertainty between different models into account, leading to predictions representing a central tendency (e.g., using median or averaged predicted values) (e.g., [41]). However, building reserve networks using ensemble SDM predictions (called hereafter the ‘pre-selection approach’) masks the variability in conservation outcomes induced by different SDMs [42]. To overcome this limitation, an alternative ‘post-selection’ approach has been introduced [42], where several scenarios for the distribution of biodiversity are constructed by sampling one particular model prediction for each species, and which are used as the various different inputs for SCP algorithms (Figure 1A,C). This latter approach provides more reliable conservation outcomes, notably, showing a better representation of rare species than using the pre-selection approach [42]. Reserve-selection efficiency has been considered less when implying SDM in the SCP process. Most conservation applications used random-based heuristics (Figure 1A,B), while only 4% of the reviewed publications implemented mathematical programming to provide optimal conservation solutions. All of these studies used a pre-selection approach of SDMs (Figure 1D, Appendix A). Though the approach has never been implemented, we believe that a post-selection of SDM outputs combined with mathematical programming optimization (Figure 1C), would lead to a reserve-selection framework that might provide flexibility, while simultaneously addressing solution efficiency, thus avoiding the current tacit trade-off existing between the two concepts of flexibility and efficiency.

Here, we used the fishes of the Mediterranean Sea as an ecological template to exemplify this fourth approach. To compare the different approaches we analyzed an experimental design, crossing pre- and post-selection of SDM outputs versus random-based heuristics and exact mathematical optimizations. In the Mediterranean, species richness exhibits a decreasing gradient from West to East, with the highest species number found in coastal areas around the island of Sicily [43]. Diversity patterns in the Mediterranean are also strongly correlated with the distance to the coast, with higher species numbers over the continental shelf than in pelagic environments [43–45]. We show that accounting for uncertainties in species distributions by varying initial conditions for the SCP process, and obtaining optimal solutions for each of these initial conditions, offers flexibility in the implementation of conservation solutions, while optimizing the allocation of conservation resources.
Figure 1. Schematic representation of the different approaches for the definition of spatial priorities when accounting for species distribution model uncertainties. For each species (represented by different colors), the spatial distribution is modeled using various statistical methods (each represented by a different shape), from this pool of species/method predictions (central gray squares), four approaches are possible. The first two approaches use random-based heuristics, to identify conservation solutions. (A) A post-selection approach, where a set of sub-optimal solutions are generated for different distribution scenarios, each constructed by selecting randomly a model among all SDMs for each species. Priority rankings corresponding to different distribution scenarios are then summarized by averaging. (B) A pre-selection approach, where species distribution models are summarized into a unique ensemble model before reserve selection. The priority ranking is estimated from several sub-optimal solutions. The two latter approaches use exact optimization. (C) A post-selection approach, where priority ranking values are estimated for a set of optimal solutions, each corresponding to a different distribution scenario. (D) A pre-selection approach of SDM outputs (ensembling), leading to a unique optimal solution. The way uncertainties related to both model accuracy and solution efficiency spread across the SCP process is illustrated by gray shades of different heights in each panel.

2. Materials and Methods

2.1. Species Distribution Modeling

The occurrences of Mediterranean fish species (i.e., geographic locations) used for distribution modeling were obtained from multiple databases: OBIS (Ocean Biogeographic Information System), Global Biodiversity Information Facility (GBIF), iNaturalist (A Community for Naturalists), VertNet (vertebrate biodiversity networks), Ecoengine (UC Berkeley’s Natural History Data), and Fishmed [46]. We filtered the data by removing potentially erroneous occurrences (e.g., located on continents) and keeping only the observations made since 1975 (to match the available environmental data). For the calibration of SDMs, temperature and salinity climatologies [44,45] were acquired from the global World Ocean Database 2013 Version 2, at a spatial resolution of 1/4°. These climatologies represent decadal averages of temperatures and salinities for 1975–1984, 1985–1994, 1995–2004, and 2005–2012, distributed over 40 standard depth layers. We constructed a regular grid at a spatial resolution of 1/12th of a degree (5 arcmin) over the Mediterranean Sea, of which
each cell was used as a planning unit (PU) for the SCP process. The occurrences and environmental data were interpolated on the grid and aggregated vertically by calculating mean values in the first 50 and 200 m depth, for the calibration of pelagic and benthopelagic species, respectively; and the last 50 m depth for benthic and demersal species.

A set of eight statistical methods based on presence/absence occurrences were used. The eight models belong to four main categories: multiple regressions (generalized linear model, generalized additive model, multiple adaptive regression splines), regression trees (boosted regression tree, random forest, classification tree analysis), flexible discriminant analysis, and artificial neural network. These algorithms were implemented using the BIOMOD2 multi-model platform [47]. All models were evaluated using a cross-validation procedure, with random partitioning of 3-fold cross validation occurrences. For each subset of data, 75% of occurrences were used for calibration and the remaining 25% for validation. In total, we obtained 24 suitability maps for each species, which were transformed into presence/absence maps using the probability threshold that maximizes the models’ true skill statistic (TSS; [48]). This criterion evaluates the predictive power of the models, taking into account both omission and commission errors, and indicates a very good discrimination power when close to one, and a performance no better than random when close to zero or less.

We measured the uncertainty in SDM predictions at two levels: species and assemblages. First, for each species and within each planning unit, we calculated the average of binary predictions (i.e., models committee averaging; [47]). A committee averaging score close to 0 or 1 means that all models agree to predict 0 and 1, respectively. A score of 0.5 means that half the models predict 1 and the other half 0, and reflects a maximal prediction uncertainty. We modified these scores to measure the uncertainty in model predictions, such that 0 represents a total agreement among models on the absence or the presence of the species, and 1 represents the situation where the same number of models predict an absence and a presence, using the following Equation (1):

$$\frac{0.5 - |\text{score} - 0.5|}{0.5}$$

(1)

To provide a measure of SDM uncertainties at the assemblage level we averaged the scores of all species present in a PU.

2.2. Species Distribution Scenarios

In the pre-selection approach, SDM uncertainties are accounted for by summarizing the full range of predictions into an ensemble model, a priori to reserve selection. Here, for each species, several distribution maps (only SDMs with good predictive performance among the 24; i.e., TSS > 0.6) were combined into one final ensemble, by averaging the selected SDMs occurrence probabilities. In the post-selection approach, different ‘distribution scenarios’ are created, by randomly choosing one distribution among the available distribution maps for each species, and we produced 100 of these distribution scenarios.

2.3. Conservation Targets, Reserve Selection Algorithms, and Priority Rankings

We defined conservation targets for each data set, as follows: a target of 100% representation was set for species with restricted distribution (range <1000 km²) and a target of 10% was used for extended species (those with a geographic extent exceeding the two-thirds range of the species). For intermediate-sized species, the target was interpolated as a linear function of the log-transformed range area. In addition, we modified area-based targets for critically endangered species, as determined by the IUCN Red List categories, and targeted protecting 100% of their spatial range.

We identified conservation solutions that ensured species representation, while minimizing the total area of selected PUs, following the minimum set problem formulation [49,50]. We used an exact optimization algorithm to solve an integer linear program-
ming problem (ILP; Figure 1C,D; [51]); and the Marxan decision support tool, implementing a random-based heuristic algorithm (Figure 1A,B).

The ILP problem was formulated to minimize an objective function, giving the cost of the solution, while respecting a set of linear constraints:

\[
\begin{align*}
\text{Minimize} & \quad \sum_{i=1}^{M} s_i x_i \\
\text{subject to} & \quad \sum_{i=1}^{M} x_i a_{ij} \geq t_j \forall j \in E \\
& \quad x_i \in \{0, 1\}
\end{align*}
\]

where \( M \) is the number of PUs and \( E \) is the set of species. Let \( a_{ij} = 1 \) if species \((j \in E)\) is present in PU \((i \in S)\) and zero otherwise. The Boolean variable \( x_i = 1 \) if the PU \((i \in S)\) is selected and zero otherwise. Each PU \((i \in S)\) is described by a surface area, \( s_i \) and \( t_j \) is the minimum amount of each species range to be included in the solution (i.e., the conservation target of species \((j \in E)\)). The ILP problems were solved using the GUROBI software, which implements the ‘branch and bound’ exact algorithm [52].

We used Marxan [33], which implements a simulated annealing algorithm (a meta-heuristic algorithm) to identity sub-optimal systems of priority areas. Marxan iteratively minimizes an objective function that sums the total cost of PUs in the solution and penalties for species targets not being met (species’ penalty factor, SPF, species specific weights here set such as to ensure the representation of all species). We set the boundary length modifier to 0 (a parameter that measures the trade-off between cost and compactness of the solution), as our aim was to examine differences in the selection of priority areas among the strategies and not to design an MPA network with a desirable level of compactness. In our study, Marxan was run 100 times and consisted of 1,000,000 iterations per run. In addition, the cost of each site was equal to its area, favoring the selection of sites with high ecological importance.

A priority ranking for planning units was calculated as the selection frequency of each site across a number of conservation solutions, sites selected in more than 90% of the solutions were considered as high-priority conservation areas [53]. For the post-selection approach coupled with heuristic (Figure 1A), the priority rankings obtained for the different distribution scenarios were averaged among the 100 scenarios and 100 Marxan runs. For the pre-selection approach coupled with heuristic optimization (Figure 1B), we calculated the PU selection frequency across the 100 sub-optimal solutions derived from Marxan outputs. For the post-selection approach coupled with exact optimization (Figure 1C), we calculated the selection frequency over the 100 optimal solutions, each based on a different ‘distribution scenario’. For the pre-selection approach coupled with exact optimization (Figure 1D) there is only a single solution.

2.4. Comparative Analysis

First, we compared the distributions of priority rankings (10-bin histograms) derived from the pre-selection approach coupled with a heuristic reserve selection algorithm (the most common strategy, Figure 1B) and the post-selection approach coupled with exact reserve selection algorithm (the unexplored strategy, Figure 1C). To highlight the spatial differences and similarities between priority rankings, we mapped the differences in PUs ranking and measured the overlap in high-priority conservation areas (PUs with irreplaceability >0.9) between the two approaches. We additionally used the Wilcoxon test, known as a \( t \)-test for paired samples, to determine statistically significant differences between the rankings obtained. These rankings were compared to the uncertainties associated with the modeling of species distributions using Spearman’s correlation test, to determine whether these uncertainties were better represented by one approach or another.

Additionally, we compared the total protected area for each conservation solution obtained, based on the different approaches, and quantified the efficiency of heuristics outcomes for each ‘distribution scenario’ as the difference in total area with respect to the optimal solution.
3. Results

Overall, the SDMs showed variable performances in predicting observed species distributions, with TSS values ranging from 0 to 0.98 (0.73 ± 0.1; mean ± standard deviation). To avoid spurious conclusions, based on unfair predictions, only models with TSS greater than 0.6 were used for the remaining analyses. For all 438 species considered, several statistical methods reached the performance threshold for predicting observed distributions, which prevented a single ‘better’ model from being distinguished.

In addition, significant variability was observed in the modeled distribution ranges of species (Figure A2). The inferences about species distributions and, consequently, conservation targets are decisively dependent on the SDM modeling approach, with substantial uncertainties associated with the choice of a single best statistical modeling technique. This species-level variability resulted in spatially-structured uncertainties at the assemblage level (Equation (1)), with congruence areas mainly located in pelagic environments; and areas of disagreement at the edge of the species distribution areas, along the margins of the continental shelf, notably along the shores of the Aegean and Ionian seas (Figure 2).

Figure 2. Map of assemblage-level uncertainties in SDM predictions. Uncertainty is calculated at the species level (Equation (1)) and averaged across species present in each planning unit. Green areas indicate cells for which model prediction uncertainties are the highest.

The uncertainties of the SDMs have most often been taken into account in the literature by opting for the pre-selection approach, based on a consensual overall model as input data for SCP (Appendix A). The conservation results based on the pre-selection approach, with the ensemble model as input data for Marxan, revealed that the sub-optimal solutions contained 2.85 to 3.34% more PUs than the optimal solution obtained with ILP. We identified the 32.6% of PUs that were consistently selected across the 100 sub-optimal solutions (Figure A3), while the minimum proportion of PUs required to achieve species targets (i.e., full optimal solution) was 36.7% (Figure A3). Indeed, the distribution of PU selection frequency was strongly bimodal using both exact and heuristics algorithms, with most of the PUs belonging to the first and last 20% quantiles (0–20 and 80–100; Figure 3), reflecting the similarity of the sub-optimal solutions identified using Marxan’s heuristic algorithm.
Figure 3. Frequency distributions of priority ranking values derived from the two approaches used for integrating species distribution uncertainties. Bars of full color represent the result of the post-selection approach coupled with exact optimization, hatched bars represent the result from the pre-selection approach coupled with heuristics. The curve represents the frequency differences across different levels of priority between the two approaches.

Following the post-selection approach, 100 ‘distribution scenarios’ were generated by randomly sampling one model for each species among the SDMs selected as having a good predictive performance, thus considering the full range of good predictions (Figure 1A–C). Then, we identified conservation solutions and investigated the variability in their size and in the spatial distribution of the selected PUs. The optimal solutions, each based on a different distribution scenario, varied in size, ranging from c. 27% to c. 43% of the Mediterranean Sea (Figures A3 and A4). Among these optimal solutions, 85% of PUs were selected at least once, with 17% of PUs redundant in 20% to 80% of the solutions. Irreplaceability values derived from the post-selection approach revealed more flexibility than with the pre-selection approach, allocating moderate values (20 to 80) to a greater number of PUs than the pre-selection approach (Figure 3). In this respect, the number of high priority PUs (selected more than 90 times out of the 100 solutions) were about 7% lower than for the pre-selection approach (Figure 3). Moreover, the Wilcoxon tests, used to compare the priority rankings derived from the two pre-selection and post-selection approaches, were significant (p-value < 0.05), revealing that the selection frequency of the PUs depends on the approach followed for considering species distribution uncertainties.

Furthermore, we examined the spatial mismatch of selection frequencies between the two approaches (Figure 4). Only 14% of areas were identified as totally irreplaceable by both approaches (i.e., a selection frequency of 100%; Figure 4). In contrast, the frequency selection of 9.83% of planning units showed differences of 20% and more between the two approaches. Areas with higher priority following the post-selection approach, are mainly located along the Aegean Sea and Ionian Sea coasts; while areas presenting higher priority following the pre-selection approach are spread in small and isolated areas along all Mediterranean coasts (Figure 4). The correlation between SDM uncertainties and the selection frequencies from the post-selection approach was greater than that of the pre-selection approach (Spearman rs 0.77 and 0.57 respectively). Hence, the conservation solutions identified with the post-selection approach were more representative of the uncertainty map of predictions than the pre-selection approach.
Figure 4. Comparison of priority ranking values based on the two approaches, including uncertainty associated with distribution modeling. Map of difference in planning unit (PU) selection frequency. Scatter plot showing the PU selection frequency under the different approaches. Histogram of selection frequency differences.

The post-selection approach can also be based on sub-optimal solutions, derived from heuristics reserve selection algorithms (Figure 1A). In this study, the statistics on the optimality gaps, calculated between the optimal and sub-optimal solutions for each ‘distribution scenario’, reveal that half of the solutions identified using heuristics are c. 3 to 8% more costly than those found by the exact optimization (Figure 5).

Figure 5. Left: scatter plot showing the area of conservation solutions (expressed as percentage of the sea) for random-based heuristics and optimal conservation outcomes for both the pre-selection and post-selection approaches; each point represents the percentage of protected areas necessary to achieve species target. Right: boxplot of optimality gap between sub-optimal and optimal conservation outcomes for each ‘distribution scenario’. The gap between the best Marxan solution and the optimal solution based on the pre-selection approach is shown by a red point.

4. Discussion

Systematic conservation planning is a data-intensive process, requiring species occurrence data that fully cover the region of interest [54]. Such knowledge about species distributions is usually so scarce and geographically variable that conservation applications rely increasingly on modeling techniques to predict the distribution of species (Appendix A). The literature review revealed that the choice of the statistical method used to model distributions is likely to be opportunistic in conservation applications, as the great majority of studies identified conservation solutions on the basis of species distributions modeled with only one statistical method. Using a single and a priori chosen model (opportunistically), particularly for rare species, may yield spurious predictions, affecting the selection of
MPAs and jeopardizing the adequacy of conservation targets and the representativeness of conservation solutions [55,56]. Before any conservation exercise, it is a good practice to evaluate the predictive performance of SDMs using an evaluation metric (e.g., TSS, AUC). However, while this criterion provides guidance on the relative performance of SDMs, it does not identify the ‘best’ statistical method for distribution modeling. Indeed, the TSS or equivalent measures evaluate the performance of the internal model and, therefore, do not allow deciding on the ecological validity of the predictions [57]. In this study, various models yielded equivalent predictive performances, which prevented distinguishing a single ‘best’ model. It is, thus, good practice for any conservation exercise to start by using different statistical methods for species distribution modeling, and to evaluate the predictive performance of the models obtained [58].

Modeling uncertainties predominate in the literature, and most studies opt for the pre-selection approach, in order to address them (Appendix A; e.g., [41]). Several studies used, to complement SDM ensembles as inputs for SCP, a measure of the variability in species predictions in each cell as a cost. Including uncertainties in SDMs as a constraint for the SCP algorithm implies that sites for which there is low concordance among model predictions are not prioritized [59]. This approach limits the risk of misallocating scarce conservation resources where the uncertainty about the occurrence of species is high. However, as SDMs tend to disagree, particularly for rare species [55,56], it could be that, without further constraints on the representation of rare species (e.g., by adding a strong penalty for the under-representation of rare species in the objective function), places exhibiting a low SDM congruence are the places where most of rare species do occur, potentially decreasing the effective representativeness of conservation solutions for rare species.

The use of an exact reserve selection algorithm to solve an ILP problem based on predictions from the ensemble models (Figure 1D), provided a unique, optimal, and efficient conservation solution covering 36.7% of the Mediterranean Sea and most of the continental shelf (Figure A2). Marxan’s heuristic algorithm produced a set of sub-optimal solutions, making it possible to estimate the priority ranking of PUs (Figures 1B and A3). However, the ranking obtained was strongly bimodal (Figure 3) and geographically structured (Figure A3), with 32.6% of the PUs located over the continental shelf being completely irreplaceable to achieve the conservation objectives. This limits the flexibility offered by heuristics for the implementation of conservation solutions. With this limited flexibility being obtained at the ‘cost’ of efficiency (here 2.85 to 3.34% more sites than the optimal solution), better strategies could be devised to offer flexibility in the implementation of conservation solutions, while optimizing the allocation of conservation resources.

Flexibility can also be provided by applying algorithms repeatedly, under different initial conditions. Following the post-selection approach, different ‘distributions scenarios’ are considered by randomly sampling model predictions for each species among the full range of relevant predictions (Figure 1A–C). Here, the post-selection approach generated priority ranking values with more replaceable PUs and greater flexibility for decision-makers than the pre-selection approach. Furthermore, these rankings were shown to be more representative of species distribution uncertainties than the ones obtained with the pre-selection approach. As such, the post-selection approach coupled with exact optimization represents an alternative to provide both flexibility and efficiency in the SCP process, while acknowledging our limited knowledge of biodiversity distribution.

In practice, this approach can be based on optimal solutions, but also the best sub-optimal solutions from heuristics. Several studies have commented on the sub-optimality of conservation solutions [34,59,60]. Vanderkam et al. (2007) [35] found that heuristics yielded reserve networks that were 2–70% larger than the network identified by an optimal algorithm (e.g., mathematical programming), depending on the formulation of optimization problem and the heuristic algorithm used. Heuristic algorithms cannot inform the user about the degree of sub-optimality of the solutions. Our results show that in the Mediterranean the sub-optimality of Marxan’s solutions can reach up to 8% of the study area (Figure 5). Moilanen (2008) [34] concluded that a 5–10% efficiency loss out of limited
resources can be considered meaningful. Combined with the fact that over the 10% of marine areas targeted for protection by the Convention on Biological Diversity only 7.14% of the Mediterranean sea is under a legal conservation designation [60], these results reinforce the need to opt for an algorithm that saves the scarce resources allocated to conservation in this region.

In conclusion, our study highlighted that the modeling technique choice is an important factor to consider when using SDMs as inputs for SCP. The way modeling uncertainties are handled can lead to distinct perspectives on the prioritization of conservation actions and jeopardize the adequacy and representativeness of conservation outcomes. While the pre-selection approach estimates site priorities by varying solutions efficiency, the post-selection approach (coupled with the exact algorithm) provides decision-makers with the greatest information related to the uncertainties in species distributions across different modeling techniques. This information, communicated as priority values, where the most irreplaceable areas are selected across different potential ‘distribution scenarios’ could help to better implement the fundamental principles of SCP, aiming for adaptive and effective management of resources.


Funding: The authors would like to acknowledge valuable financial support from the Institut de Recherche pour le Développement (IRD, France) in the framework of the GAMBAS project. Sabrine Drira was fully funded by IRD ARTS.

Institutional Review Board Statement: Not applicable.


Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

In order to assay the main trends in the use of SDMs through spatial conservation planning, a literature search was conducted using Web of Science (WoS). The intention of the review was to assess the ways in which SDMs have been integrated into conservation planning by evaluating two questions: (1) What types of models have been included? and (2) How uncertainty between different models has been accounted for in conservation planning?

To identify relevant literature, we carried out a topic search examining the title, keywords and abstracts among papers, using the following query: TS = ((“species distribution model” OR “habitat model” OR “ecological niche model” OR “habitat suitability ind” OR “habitat suitability mod” OR “occupancy mod” OR maxent OR “presence-absence mod” OR “presence-only mod” OR “niche model” OR “climate change” OR “changing climate”) AND (“Spatial planning” OR “Spatial optimization” OR “Marine reserve” OR “spatial prioritization” OR “reserve selection” OR “area-selection algorithm” OR “spatial conservation prioritization” OR “conservation plan” OR “land use plan” OR “regional
This query results initially in 137 papers. Articles without study cases for both distribution modeling and conservation planning application were excluded. Our literature review also exclude national journals, reviews, case reports, letters, editorials, and conference abstracts. Furthermore, we browsed the most frequently cited papers in the gathered corpus, to have a final corpus composed of 76 articles in total. Nevertheless, our review may be incomplete; but it remains representative. We then extracted the following information from all studies: (1) year of publication, (2) number of techniques used for species distribution modeling, (3) projection in future under climate change scenarios (Boolean), (4) reserve-selection algorithm used for conservation planning (scoring; heuristics or mathematical programming), (5) uncertainty of feature data (deal with it or not; Boolean).

The use of SDMs is increasingly favored and common in SCP applications (Figure A1). Most conservation applications have neglected the uncertainties associated with different modeling techniques, by using a single modeling method, chosen a priori, to predict the distribution of species (n = 65); while others have examined a wide range of techniques and have used only the ‘best’ model identified in terms of predictive performance (e.g., Leach et al., 2013; Passoni et al., 2017; Walther and Pirsig, 2017). However, when such particular solutions are evaluated using other statistical methods, the representation of species in the conservation solution might not meet the required target level (Loiselle et al., 2003).

Only c. 23% of the reviewed publications explicitly incorporated SDM uncertainties into the conservation planning process, especially for the modeling of future projections (Figure A1). To do this, all studies adopted ensemble prediction approaches (called ensemble or multi-model models or model averaging), representing a consensus model with median or mean prediction values, in order to combine different SDMs, and significantly improve the precision of the predicted distributions (e.g., Alagador et al., 2016, Faleiro and Loyola, 2013). This approach, hereafter referred to as the pre-selection approach, aims to reduce uncertainty a priori to conservation planning, so that the distributions predicted by the ensemble models are used as inputs for the optimization algorithms. Meller et al. (2014) introduced another approach, hereafter referred to as the post-selection approach, in which different ‘distribution scenarios’ are constructed by randomly sampling one model from across the range of SDMs; and used a heuristic algorithm to identify conservation priorities.

The effectiveness of conservation solutions has been less taken into account when SDMs are involved in the SCP process. Most conservation applications (about 80% of the reviewed publications) used heuristic algorithms, with slightly more than half using
Zonation (n = 42; Moilanen et al., 2009), and almost one quarter using Marxan (n = 19; Ball and Possingham, 2000). Only 4% (n = 3) of the reviewed publications implemented mathematical programming to provide optimal conservation solutions. All three studies had a pre-selection approach for managing SDMs uncertainties.

Figure A2. Percentage of predicted species range across distribution models with good predictive performance. The points represent the predicted range from the ensemble model.

Figure A3. Conservation outcomes following the pre-selection approach. Top: optimal solution from exact optimization; Bottom: priority ranking values across sub-optimal solutions.
Figure A4. Conservation outcomes following the post-selection approach. Map of priority ranking values across optimal solutions built for each ‘distribution scenario’. Boxplot of the percentage of protected areas across those optimal solutions. The grey point represents the percentage of protected areas necessary to achieve the species target based on the ensemble model.

References


22. Lentin, P.E.; Wintle, B.A. Spatial conservation priorities are highly sensitive to choice of biodiversity surrogates and species distribution model type. *Ecography*** **2015**, *38*, 1101–1111. [CrossRef]


34. Moilanen, A. Two paths to a suboptimal solution—one more about optimality in reserve selection. *Biol. Conserv.*** **2008**, *141*, 379–385. [CrossRef]


52. Gurobi Optimization; Gurobi Optimization. Inc.: Houston, TX, USA, 2012.


