

Harnessing citizen science for marine conservation in Malta: a comparative analysis of GAM and MaxEnt models in bottlenose dolphin habitat mapping

Francesca Soster^{1,*}, Tim Awbery^{2,*}, Nina Vérité–Taulet³, Timothy Zammit³ and Kimberly Terribile⁴

- ¹ Applied Research and Innovation Centre, Malta College of Arts Science & Technology (MCAST), Paola, Malta
- ² Marine Mammal Research Team, Scottish Association for Marine Science, Oban, United Kingdom
- ³ Discover the Blu, San Pawl il-Baħar, Malta
- ⁴ Centre for Agriculture, Aquatics and Animal Sciences, Malta College of Arts, Science and Technology (MCAST), Paola, Malta
- * These authors contributed equally to this work.

ABSTRACT

Background. Species distribution models (SDMs) are powerful tools for informing conservation, particularly for highly mobile marine species such as common bottlenose dolphins (*Tursiops truncatus*). In Maltese waters, the limited availability of data on this species has constrained the effectiveness of conservation efforts. Despite the designation of offshore Special Areas of Conservation (SACs), key coastal regions need more detailed spatial studies to support evidence-based management.

Methods. In this study, we analyzed and compared the outputs of a generalized additive model (GAM) and a maximum entropy (MaxEnt) model to assess summer habitat suitability for bottlenose dolphins within a coastal SAC in Malta. The models were informed by presence-only data collected through systematic surveys and a citizen science campaign, integrated with environmental and anthropogenic predictors including chlorophyll-a concentration, sea surface temperature anomaly, slope, and distance to aquaculture sites.

Results. Both modeling approaches identified high habitat suitability in shallow, nearshore regions, with chlorophyll-a concentration and proximity to aquaculture sites emerging as the most important predictors. Slope and sea surface temperature anomaly contributed less substantially. The two models showed spatial agreement in highlighting these nearshore areas as core habitats, though GAM predicted a broader extent of suitable habitat, whereas MaxEnt results were more spatially restricted. Both models demonstrated strong predictive performance (AUC > 0.85), reinforcing the ecological relevance of the identified drivers.

Conclusion. This study demonstrates the potential of integrating opportunistic data with SDMs to support habitat assessments in data-limited contexts. The use of complementary modeling approaches provides robust insights into species—environment relationships. These results aim to guide spatial planning and future assessments of conservation priorities in Maltese coastal waters.

Submitted 7 May 2025 Accepted 7 July 2025 Published 5 August 2025

Corresponding author Francesca Soster, francesca.soster@mcast.edu.mt

Academic editor Stuart Pimm

Additional Information and Declarations can be found on page 21

DOI 10.7717/peerj.19804

© Copyright 2025 Soster et al.

Distributed under Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Conservation Biology, Ecology, Marine Biology

Keywords *Tursiops truncatus*, Conservation planning, Coastal research, Species distribution model, Maximum entropy (MaxEnt), Citizen science, Generalized Additive Model (GAM), Habitat suitability modeling

INTRODUCTION

Understanding species distribution and habitat use is fundamental for developing effective conservation strategies (*Rodríguez et al.*, 2007; *Guisan et al.*, 2013). However, assessing these aspects can be particularly challenging for highly mobile species like cetaceans (*Fernandez*, *Sillero & Yesson*, 2022). In this context, species distribution models (SDMs) offer a solid foundation for gaining valuable ecological insights, particularly in data-deficient regions (*Redfern et al.*, 2006; *Fiedler et al.*, 2018; *Fernandez et al.*, 2021). By evaluating the relationships between cetacean populations and the environment, as well as anthropogenic factors, SDMs contribute to a more comprehensive understanding of species ecology and potential conservation priorities (*Rodríguez et al.*, 2007; *Guisan et al.*, 2013; *Marshall*, *Glegg & Howell*, 2014; *Pace*, *Tizzi & Mussi*, 2015; *Giralt Paradell*, *Díaz López & Methion*, 2019).

The ability of SDMs to assess species distributional ranges *via* predictive modeling (*Anderson*, *Lew & Peterson*, 2003; *Pitchford et al.*, 2016) or to identify the environmental drivers of distribution through explanation modeling (*Azzellino et al.*, 2008; *La Manna et al.*, 2023b) has led to an increase in their use in recent years for cetacean species (*Pasanisi et al.*, 2024). When used together, predictive and explanation modeling provides information on the distribution of habitats that are suitable for species survival (*Hirzel & Le Lay*, 2008). Within this framework, habitat suitability models serve as practical tools to estimate the ecological conditions and areas within which a species is most likely to occur (*Guisan & Thuiller*, 2005).

Among the most widely adopted modeling techniques are machine learning algorithms such as maximum entropy (MaxEnt) models (*Melo-Merino, Reyes-Bonilla & Lira-Noriega, 2020*), which estimate the habitat suitability from presence-only data using the principle of maximum entropy (*Phillips, Anderson & Schapire, 2006*), and statistical approaches such as generalized additive models (GAMs), which model non-linear relationships between species occurrence (*e.g.*, presence/absence) and environmental factors (*Hastie & Tibshirani, 1986*). Both methods enable researchers to identify critical conservation areas and mitigate potential threats by analyzing species' habitat preferences (*Rodríguez et al., 2007*; *Guisan et al., 2013*). Furthermore, they help predict how marine species respond to environmental changes, contributing to the development of targeted conservation strategies (*Giralt Paradell, Díaz López & Methion, 2019*; *Díaz López & Methion, 2024*).

The performance and reliability of these models are inherently dependent on several factors, including model parametrization (*Elith et al.*, 2006), the ecological relevance and statistical independence of environmental predictors (*Guisan & Zimmermann*, 2000; *Dormann et al.*, 2013), the spatial and temporal scale at which models are implemented (*Elith & Leathwick*, 2009) and the overall quality, completeness, and bias of species occurrence records (*Hernandez et al.*, 2006; *Araújo & Guisan*, 2006; *Merow*, *Smith & Silander*, 2013).

Dedicated surveys, although considered robust, are often resource-intensive and limited by logistical and financial constraints, resulting in incomplete spatial coverage and potential underrepresentation of species distribution (*Evans & Hammond, 2004*; *Meyer et al., 2015*). Recent studies have demonstrated that non-traditional data sources, such as opportunistic and citizen science data (*Giovos et al., 2016*; *Pace et al., 2019*; *Robbins, Babey & Embling, 2020*), offer a cost-effective approach to addressing specific research challenges, facilitating the generation of estimates of cetacean distribution (*Fernandez et al., 2021*). In the study of species inhabiting dynamic environments, non-traditional data sources offer several advantages over conventional survey methodologies. For instance, they can enhance spatial and temporal coverage and reduce logistical constraints, as they often do not require specialized equipment or dedicated field effort (*Robbins, Babey & Embling, 2020*; *Corr et al., 2024*). Nevertheless, the inherent biases and variability associated with opportunistic data necessitate careful quality assessment and data integration techniques to ensure their effective use in large-scale modeling efforts (*Isaac et al., 2020*; *Martino et al., 2021*).

Given these challenges, the integration of SDMs with non-traditional data sources presents a promising approach for studying cetaceans in regions where systematic monitoring is limited. In Maltese waters, the common bottlenose dolphin (*Tursiops truncatus*) represents one of the most frequently observed cetacean species (*Notarbartolo Di Sciara*, 2002) and plays an important ecological role as a top predator within the marine ecosystem. Bottlenose dolphins are protected under a framework of national, regional and international legal frameworks. At the national level, all cetacean species are protected through the Flora, Fauna and Natural Habitats Protection Regulations (S.L. 549.44), which transpose the European Union (EU) Habitats Directive 92/43/EEC into Maltese law. In particular, the bottlenose dolphin is listed in Annex II of the Directive, which requires the designation of Special Areas of Conservation (SACs) forming part of the Natura 2000 network, and in Annex IV as a species of community interest in need of strict protection.

Despite its protected status, the limited availability of systematic data on bottlenose dolphins in Maltese waters presents challenges for fully supporting conservation planning. Previous studies in the region have suggested a coastal distribution and spatial association between dolphin presence and anthropogenic features such as aquaculture sites, which may influence foraging behavior and local prey availability (Laspina, Terribile & Said, 2022; Soster et al., 2025). Recent modeling efforts have highlighted potential gaps in the spatial alignment between currently designated Natura 2000 sites and areas predicted to be of high habitat suitability for bottlenose dolphins around the Maltese archipelago (Soster et al., 2025). Specifically, although SACs have been designated for bottlenose dolphin conservation in Malta, these mainly encompass offshore regions, while key coastal regions remain outside protected boundaries. This highlights the need for further in-depth investigation of bottlenose dolphin spatial ecology and habitat use in these areas, particularly through temporally and spatially specific studies that can build upon existing baseline knowledge. Moreover, dolphins in the region are exposed to a range of human pressures, including maritime traffic, aquaculture, and unregulated recreational activities (Said et al., 2017; Filletti et al., 2023; Mizzi et al., 2024; Soster et al., 2025). They are additionally subject to environmental pressures, such as marine heatwaves (Garrabou et al., 2022), which require

further investigation to be adequately addressed under existing management frameworks. These conservation gaps highlight the urgent need for spatially explicit data on dolphin distribution and habitat use to inform evidence-based conservation strategies in Malta's dynamic and heavily used coastal waters.

In light of these gaps, this case study aims to improve understanding of bottlenose dolphin habitat suitability using presence-only data from opportunistic research platforms and citizen science. We apply GAM and MaxEnt approaches within a heavily impacted coastal SAC that, despite its designation, does not explicitly include bottlenose dolphins as a target species. While this study is conservation-focused and not a methodological comparison of modeling techniques, the inclusion of both models allows for a more nuanced interpretation of species—habitat relationships and enhances stakeholder confidence in the results. By analyzing and comparing the outputs, we identify summer habitat suitability patterns and assess the influence of environmental and anthropogenic drivers. The ultimate long-term goal of this work is to support evidence-based spatial planning and inform the potential inclusion of bottlenose dolphins within the site's conservation framework.

MATERIALS & METHODS

Study area

The study area covers 159 km² and is situated on the northeastern side of the Maltese archipelago, extending from central Malta to the northern coast of Gozo, with a maximum depth of 100 m (Fig. 1). The area includes the Natura 2000 site Zona fil-Baħar bejn Il-Ponta ta 'San Dimitri (Għawdex)u Il-Qaliet, designated as a SAC for the protection of the Maltese top-shell (*Steromphala nivosa*) and the loggerhead turtle (*Caretta caretta*). Although bottlenose dolphins are known to occur in the area, they are not currently listed among the protected species for this site (*EUNIS*, 2025). This coastal SAC hosts a wide range of important habitats, such as *Posidonia oceanica* meadows, maërl beds, and *Cymodocea nodosa* meadows. However, this site is affected by numerous human activities, including four main finfish aquaculture sites and one bunkering area for large vessels (Fig. 1). Furthermore, pleasure boating is especially prevalent, driven by the high concentration of marinas and the region's appeal as a tourist destination.

Occurrence data

Occurrence data of bottlenose dolphins were collected using two complementary approaches: (i) dedicated boat surveys were conducted between July and September 2024 aboard a six-meter rigid-hull boat equipped with a 225 hp engine, following a standardized observational protocol during whale watching activities, permitted by the Environment and Resources Authority (EP 0249/24; Supplementary Material 1A); (ii) opportunistic data were gathered through a citizen science campaign, which encouraged sea users to report dolphin sightings by submitting information on coordinates, group size and composition, behavior, and photographs (Supplementary Material 1B). Outreach included the distribution of flyers to marinas, diving centers, charter companies, and tour operators across the study area. In addition, opportunistic reports shared *via* social

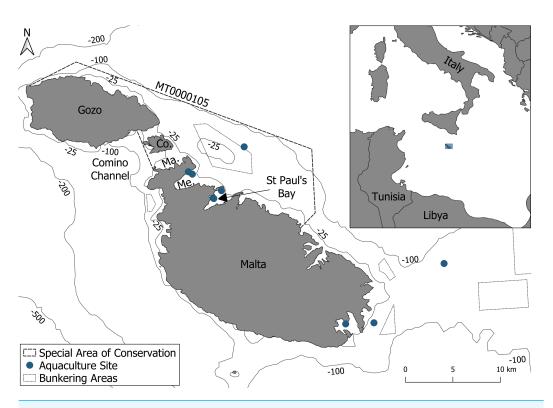


Figure 1 The survey area consisting of the Special Area of Conservation (MT0000105). Within the Special Area of Conservation (SAC), delineated by a dashed line, are aquaculture sites and a bunkering area delineated by blue points and a dotted black line respectively. Abbreviations: Co., Comino; Me., Mellieha Bay; Ma., Marfa Bay.

media platforms were also collected and screened (*Pace et al.*, 2019). All records were validated through expert verification before the inclusion in the database, a method widely utilized for its high level of accuracy (*Yu, Wong & Hutchinson, 2010; Baker et al., 2021*). The validation process considered the reliability of the observer, the availability of supporting visual evidence (*e.g.*, videos or photographs), and existing knowledge of the local bottlenose dolphin population (*Bonter & Cooper, 2012*).

Data collection was limited to the summer season due to resource availability and its alignment with the regional peak in dolphin sightings and more favorable sea conditions, thereby enhancing both detection probability and data quality. An effort-based GAM was initially generated, but the low number of presence points with associated effort meant that the results were not reliable (*i.e.*, deviance explained >99.9%). Thus, presence-only data were considered for two modeling approaches using a binomial GAM and a MaxEnt model.

Environmental data

A number of static environmental variables were included in the models, including depth, slope, aspect, and distance from shore. Dynamic variables comprised sea surface temperature (SST), chlorophyll-a concentration, mixed layer depth, and salinity. The selection of these variables was informed by previous studies on bottlenose dolphins

(Pitchford et al., 2016; Carlucci et al., 2016; La Manna, Ronchetti & Sarà, 2016; Fontanesi et al., 2024). In addition, SST anomaly, chlorophyll-a anomaly, and distance from aquaculture sites were incorporated to account for the potential influence of direct and indirect anthropogenic impacts on habitat suitability (Maricato et al., 2022). Details of the downloaded datasets, spatial resolutions and sources are given in Table 1. All dynamic variables were downloaded at a daily temporal scale, mean averaged across the survey period and resampled to the resolution of the depth layer (Fig. S1). A number of the variables had a high level of correlation (Pearson's correlation coefficient |r|>0.7) and were removed from the analysis (Dormann et al., 2013). In order to decide which of a pair of correlated variables was removed, two single-variable GAMs were run, and the variable in the model with the lowest AIC value was selected. Additionally, following the fitting of the best-performing GAM, a check of concurvity (a non-linear equivalent of a collinearity test) was made. This diagnostic is specific to models like GAMs that include smooth functions and allow assessment of non-linear dependencies among predictors. This step is not applicable to MaxEnt models, as they do not estimate smooth functions and lack a direct analogue to concurvity diagnostics. The final set of environmental predictors retained after this process was used consistently in both the GAM and MaxEnt models to ensure comparability of outputs.

Pseudoabsences and background points selection

Whilst both GAM and MaxEnt are used regularly in species distribution modeling, they differ notably in how they handle absence information and the selection of pseudo-absence or background points. GAMs require both presence and absence (or pseudo-absence) data, and model performance can be influenced by the ratio between these points. A common recommendation is to use a presence-to-pseudo-absence ratio between 1:1 and 1:5, although using a large number of pseudo-absences (e.g., 10,000) with equal weighting can also enhance model accuracy, particularly in large or heterogeneous study areas (Barbet-Massin et al., 2012). In contrast, MaxEnt operates on presence-only data and uses background points to represent the available environmental space, without assuming true absences. The default setting in MaxEnt is 10,000 background points, which is generally suitable for most applications, though this number can be adjusted depending on the extent and resolution of the study area (Phillips & Dudík, 2008). These methodological differences influence model outputs and will be considered when comparing results.

To account for sampling bias and improve model accuracy, a bias raster was generated using a kernel density estimation based on the spatial distribution of dolphin sightings (Fig. S2). A focal (neighborhood-based) smoothing operation was applied to create a continuous surface representing relative sampling effort, constrained within a two-kilometer buffer around both dolphin sightings and the boundaries of the SAC. The resulting raster was normalized and used to guide the selection of pseudo-absence/background points. A total of 30,000 points were initially generated (Fig. S3) using the randomPoints function in the dismo package in R, version 1.3-16 (*Hijmans et al.*, 2024).

Det ille and description		
Predictor	Details and description	
Depth	Water depth was obtained from the GEBCO 2023 global bathymetric grid (15 arc-seconds, approximately 460 ×380 m at Malta's latitude spatial resolution) (https://www.gebco.net/data_and_products/gridded_bathymetry_data/). The bathymetry raster was clipped to the study area extent and projected to UTM Zone 33N to ensure accurate distance calculations.	
Slope and Aspect	Slope and aspect were derived from the GEBCO bathymetry raster using the R package $terra$. Slope represents the rate of change in elevation, while aspect indicates the direction of the steepest slope. These rasters were generated at the same resolution as the input bathymetry grid (approximately 460×380 m).	
Distance from Shore and Distance from Aquaculture Sites	Euclidean distance from each raster cell to the nearest coastline or aquaculture facility, calculated using the "Distance" tool in the <i>terra</i> package in R. The coastline was obtained from official Maltese marine boundaries, and aquaculture sites were manually digitised from government sources (https://msdi.data.gov.mt/geoportal.html).	
Sea Surface Temperature (SST)	Daily mean SST values were obtained from the STREAM App (https://stream-srf.com/products/), providing high-resolution satellite-derived SST data (~1 km grid resolution) based on Copernicus Marine Service inputs. SST values were averaged over the study period to generate a composite summer surface.	
Chlorophyll-a	Daily chlorophyll-a concentration data were sourced from the STREAM App. The product integrates Copernicus Ocean Colour multi-sensor datasets at \sim 1 km spatial resolution, using the MedOC4 algorithm for offshore waters and AD4 for coastal waters. Values were averaged over the summer period.	
Mixed Layer Depth	Mixed layer depth was obtained from the Mediterranean Sea Physics Analysis and Forecast product (Copernicus Marine Service: https://marine.copernicus.eu/). The product provides daily gridded data at 1/24° (~4 km) spatial resolution.	
Salinity	Daily surface salinity data were also extracted from the Mediterranean Sea Physics Analysis and Forecast product (Copernicus Marine Service, 1/24° spatial resolution). Mean salinity values were calculated for the summer study period.	
SST Anomaly	SST anomaly, representing the deviation of daily SST from the 30-year climatological mean (1989–2019), was obtained from the STREAM App. This allows identification of thermal anomalies across the study area at \sim 1 km resolution.	
Chlorophyll-a Anomaly	Chlorophyll-a anomaly data were sourced from the STREAM App, calculated as the difference between daily chlorophyll-a concentrations and the 26-year climatological mean (Sept 1997–Aug 2023). The product uses multi-sensor satellite observations and gap-free data at \sim 1 km resolution.	

To estimate the optimal number of background points for GAMs, 30 values were randomly sampled from the generated dataframe using base R (v4.4.3; *R Core Team, 2025*). Presence and background data were split into 90% training and 10% testing sets. GAMs were fitted to the training data and evaluated on the test set using the area under the curve (AUC) and the true skill statistic (TSS). AUC is threshold-independent and widely used for its simplicity and interpretability (*Liu, White & Newell, 2011*), while TSS provides a threshold-dependent metric less sensitive to prevalence (*Shabani, Kumar & Ahmadi, 2018*). This process was repeated 50 times using k-values (wiggliness parameter) of 3, 5,

and 10, and background point values ranging from 30 to 250, exceeding the commonly recommended 1:1 to 1:5 presence-to-pseudo-absence ratio (*Barbet-Massin et al.*, 2012). For the GAM, AUC stabilized beyond 50 background points (Fig. S4A), while TSS peaked between 50 and 90 (Fig. S4B). Models with k = 3 or 5 consistently outperformed those with k = 10. A 1:1 presence-to-background ratio and k = 5 were selected for final modeling.

The same procedure was applied to MaxEnt, with background values from 1,000 to 12,000 and beta (regularization multiplier) values from 1 to 5. For MaxEnt, AUC and TSS remained stable across most background values, with AUC between 0.87 and 0.89, and TSS between 0.50 and 0.55 (Figs. S5A, S5B). Given the small size (159 km²) and homogeneity of the study area, 8,000 background points were selected. A beta value of 1 was retained, following *Phillips & Dudík* (2008) for small samples, and to avoid over-regularization, which may reduce ecological interpretability (*Merow, Smith & Silander, 2013*).

Model specifications and evaluation

The GAM was fitted using the binomial family with a logit link function in the mgcv package in R (version 1.9-1). Variable selection was enabled using shrinkage smoothers, which apply penalties to smooth terms that do not contribute meaningfully to the model, effectively shrinking them towards zero and allowing their influence to be removed.

The MaxEnt model was fitted using specific parameters implemented through the dismo package in R. Linear, quadratic, and hinge features were included, while product features were excluded, as they are only enabled when the number of presence points exceeds 80 (*Phillips & Dudík*, 2008). Threshold features were also omitted to minimize model complexity (*Elith et al.*, 2011). The model was run using 500 iterations and 8,000 background points.

For both models, a 10-fold cross-validation procedure (90% train/10% test) was conducted to evaluate predictive performance as previously done when dataset sizes are constrained (Fielding & Bell, 1997; Breiner et al., 2015; Deneu et al., 2021). This approach was selected to provide a robust estimate of model accuracy while minimizing potential biases related to overfitting and data imbalance. The dataset was stratified into ten folds, ensuring a balanced distribution of presence and pseudo-absence observations for the GAM, and presence and background observations for the MaxEnt model. Each fold was used once as a testing set, while the remaining nine folds served as the training set. This process was repeated iteratively so that each observation contributed to both training and validation. For each iteration, a model was fitted to the environmental predictors, and predictions of species presence probability were generated for the testing set. Model performance was evaluated using receiver operating characteristic (ROC) curve analysis (Fielding & Bell, 1997; Phillips, Anderson & Schapire, 2006; Elith et al., 2006). The AUC was calculated for each fold as a measure of predictive accuracy, with higher AUC values indicating better model performance (Phillips, Anderson & Schapire, 2006; Elith et al., 2006; Merow, Smith & Silander, 2013).

Following model fitting, uncertainty was quantified for both GAM and MaxEnt predictions to assess the reliability of habitat suitability outputs. For GAMs, standard errors were extracted directly from the model's built-in prediction outputs (*Wood*, 2017).

For MaxEnt, a bootstrap resampling approach was used, and the resulting raster stack of 500 replicate predictions was analyzed (Elith & Leathwick, 2009). In both cases, a mean habitat suitability map was generated to represent the central tendency of model outputs (Phillips, Anderson & Schapire, 2006). To further quantify uncertainty, the standard error was calculated for each raster cell, reflecting the absolute variability in predicted values, while the coefficient of variation (CV) was computed to express relative uncertainty as a percentage of the mean prediction (Chen, Dimitrov & Meyers, 2019). These uncertainty metrics were used to identify areas of high prediction confidence and regions where model outputs were more variable. Given the two different methodologies of extracting the coefficient of variation, these are not directly comparable between the two modeling techniques but provide a useful understanding of spatial uncertainty around the model. Similarly, the output predictions of GAMs and MaxEnt are different; GAMs use logistic regression (binomial family with a logit link) to predict the probability of a presence whereas MaxEnt returns a predicted suitability (Guisan, Edwards & Hastie, 2002; Phillips, Anderson & Schapire, 2006). Whilst not directly comparable, both methods can be used to determine the area which is most suitable for the species of interest.

Contribution and importance of variables

Both GAMs and MaxEnt offer visual tools to help interpret how environmental variables influence species presence. In GAMs, the shape of the partial response curves illustrates the effect of each predictor on the log-odds of species presence (*Guisan, Edwards & Hastie*, 2002). In MaxEnt, response curves indicate the species' relative preference across the range of each environmental variable (*Phillips, Anderson & Schapire*, 2006). Given the differences, the two should not be directly compared. They can however, both be useful in interpreting the effects of variables and thus are both given in the results.

Similarly, MaxEnt also provides two metrics of variable importance: percent contribution, based on the increase in model gain during training, and permutation importance, which measures the decrease in model performance when a variable's values are randomly permuted (*Phillips, Anderson & Schapire, 2006*; *Elith et al., 2011*). While GAMs do not include built-in measures of variable importance, a comparable estimate can be derived by examining deviance explained. This involves removing one predictor at a time and comparing the deviance of the reduced model to the full model (*Wood, 2017*). As with the visual response curves, these importance measures are not directly comparable across modeling approaches, but each helps to understand the role of environmental drivers and has been presented accordingly.

Investigating the additional value of including citizen science data

To assess the influence of citizen science data on model performance, both the GAM and MaxEnt models were run without the inclusion of citizen science records. Model parameters were kept identical to those used in the full models: a 1:1 ratio of presence to background points and a k-value of five for the GAM, and 8,000 background points with a beta value of one for MaxEnt.

RESULTS

Occurrence

A total of 17 boat-based surveys covered 484.8 km of on-effort trackline, with a mean vessel speed of 9.7 kt (SE \pm 0.18). The survey effort was unevenly distributed across the study area, with the majority of effort concentrated in the northern coastal waters of Malta (Fig. 2). A total of 20 bottlenose dolphin sightings were recorded during these surveys, with most sightings occurring in shallow to moderately deep waters (<50 m) to the north of Malta. The encounter rate for boat-based surveys was 4.12 groups and 19.8 individuals per 100 km surveyed. The citizen science campaign contributed an additional 49 sightings within the study area, out of a total of 52 validated records submitted during the campaign period. Together, the combined dataset resulted in 69 sightings (Fig. 2).

Collinearity tests

The collinearity test resulted in the retention of slope, aspect, chlorophyll-a, SST anomaly, and distance to aquaculture. The subsequent concurvity assessment for the GAM led to the removal of aspect due to high non-linear dependency with other predictors. As a result, the final models were fitted using chlorophyll-a, SST anomaly, slope and distance to aquaculture.

Habitat suitability analysis with GAM

The GAM showed strong predictive performance, with a mean AUC of 0.86 (SE \pm 0.02) under 10-fold cross-validation, indicating a high level of discriminatory power (*Swets*, 1988). All four explanatory variables included in the model were found to have statistically significant relationships with bottlenose dolphin occurrence. In terms of variable importance, the proportion of deviance explained was highest for chlorophyll-a (0.79) and distance to aquaculture (0.65), suggesting these were the strongest drivers of dolphin habitat suitability (Table 2). Although slope and SST temperature anomaly were statistically significant, they contributed less to the overall deviance explained (0.20 and 0.16, respectively), indicating a weaker influence on habitat suitability patterns.

The model predicted an increased probability of occurrence with increasing values of slope and chlorophyll-a and a decreased probability of occurrence with increasing values of SST anomaly and distance to aquaculture facilities (Fig. 3).

The mean prediction map generated across 1,000 iterations showed that the highest habitat suitability for bottlenose dolphins was concentrated around Mellieha Bay and St Paul's Bay. Additional areas of moderate suitability were identified in the Comino Channel, in proximity to the bunkering area and along the north-eastern coast of Gozo (Fig. 4). The uncertainty associated with the spatial predictions, represented by the coefficient of variation across, was generally low within the borders of the SAC but higher at the outer edges of the survey region.

Habitat suitability analysis with MaxEnt

The MaxEnt model demonstrated strong predictive performance (*Swets*, *1988*), with a mean AUC of 0.89 (SE \pm 0.0001) under 10-fold cross-validation. The regularized training

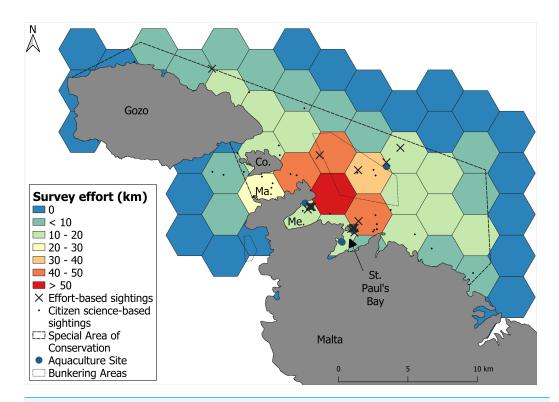


Figure 2 Survey effort calculated as the amount of on-effort track line across the survey area. Warmer colours represent areas of higher survey effort and cooler colours represent areas of lower survey effort. Sightings of dolphins are delineated by black crosses.

Table 2 Relative importance of explanatory variables in the generalised additive model (GAM). The sum of the proportion of the total deviance explained exceeds 1 because the method penalises more complex models.

Variable	Deviance explained	Proportion of total deviance explained
Chlorophyll-a	31.4%	0.79
Distance to aquaculture	26%	0.65
Slope	7.89%	0.20
SST anomaly	6.36%	0.16

Notes.

*The sum of the proportion of the total deviance explained exceeds 1 because the method penalises more complex models.

gain was 1.23, compared to an unregularized gain of 1.18, indicating that the application of regularization successfully reduced overfitting to a small extent.

Chlorophyll-a was the most influential predictor in the model, contributing 61.2% of the total model gain and 25.8% of the permutation importance (Table 3). Distance to aquaculture was the second most influential variable, contributing 20.1% to model gain and 53.3% to permutation importance.

The model showed a marked decrease in habitat suitability with increasing distance from aquaculture facilities, particularly within the first 5,000 m. Habitat suitability increased

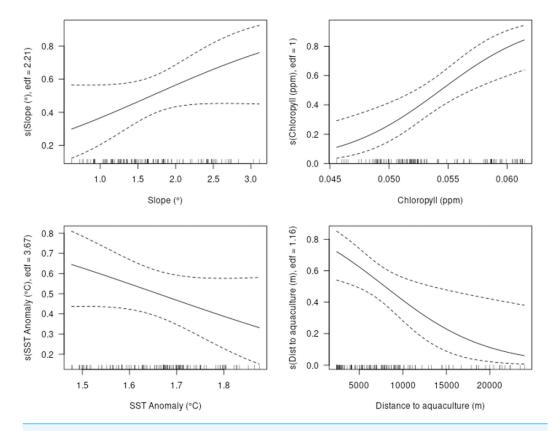


Figure 3 Partial plots of explanatory variables for the GAM. Slope, chlorophyll-a, sea surface temperature (SST) anomaly and distance to aquaculture.

exponentially with higher chlorophyll-a concentrations (Fig. 5). Slope and SST anomaly contributed less than 15% cumulatively to the model's predictive power. However, slope showed a higher permutation importance (20.4%), suggesting a moderate influence on model sensitivity. The response curve indicated that suitability peaked at slope values between 2.5° and 3°, declining on either side. Sea surface temperature anomaly exhibited a nearly linear negative relationship with suitability, but its overall contribution to the model was minimal (Fig. 5).

As with the GAM, the most suitable areas for bottlenose dolphins were the waters surrounding Mellieha Bay and Saint Paul's Bay. The MaxEnt model predicted high suitability only along the Maltese coast of the Comino Channel and a moderate suitability in the shallow bunkering area (Fig. 6). The coefficient of variation was highest near the edges of the survey area, and higher uncertainty was also observed within the SAC, especially along the central coastal waters of Gozo.

Investigating the additional value of including citizen science data

When using the GAM with effort-based presences only (n = 21), model performance was notably lower than when research and citizen science data were combined (n = 69), with a mean AUC of 0.69 compared to 0.86, and a comparable TSS of 0.984 *versus* 0.951.

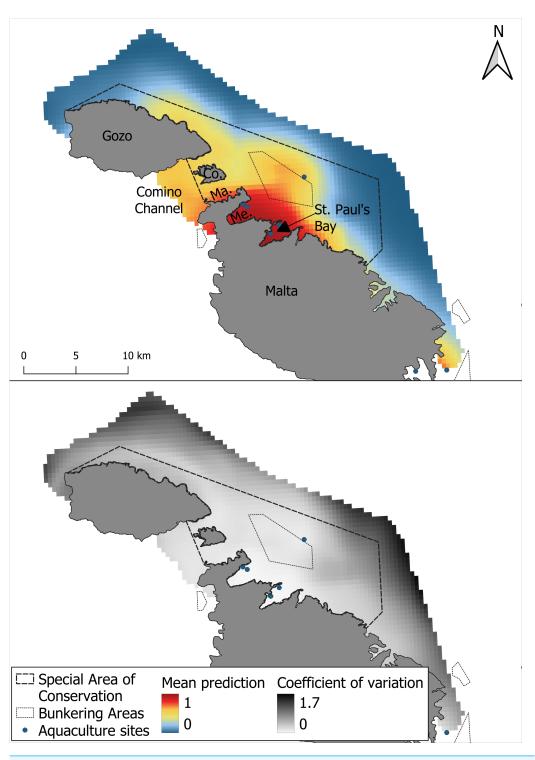


Figure 4 Mean prediction for bottlenose dolphin habitat suitability and the coefficient of variation across 1,000 iterations of the GAM modelling process.

Table 3 Percent contribution and permutation importance for the Maximum Entropy Model (Max-Ent).

Predictor	Percent contribution of predictors	Permutation importance
Chlorophyll-a	61.2	25.8
Distance to aquaculture	20.1	53.3
Slope	10.4	20.4
SST anomaly	2.3	0.5

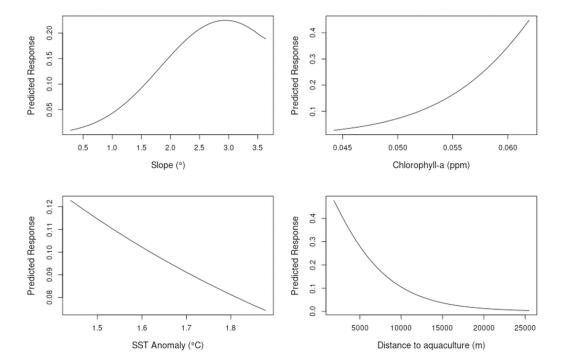


Figure 5 Single variable response for the MaxEnt model. Slope, chlorophyll-a, SST anomaly and distance to aquaculture.

Full-size DOI: 10.7717/peerj.19804/fig-5

Prediction uncertainty was also higher, with a coefficient of variation of 154.9% across the study area, particularly in the northern region (Figs. S6D–S6F). Furthermore, the limited number of data points led to an underestimation of habitat suitability in areas where dolphins were observed in the citizen science data, such as the Comino Channel (Fig. 2) and the southeast of Malta (Figs. S6A–S6C).

Although the MaxEnt model using only effort-based presences produced higher mean AUC (0.94) and TSS (0.89) compared to the model based on the full dataset (AUC = 0.89; TSS = 0.64), it showed increased prediction uncertainty, with the coefficient of variation rising by 76.5% across the survey area, particularly in the eastern region (Figs. S7D–S7F). Additionally, the model failed to predict suitable habitat in the Comino Channel and the southeastern part of the survey area (Figs. S7A–S7C).

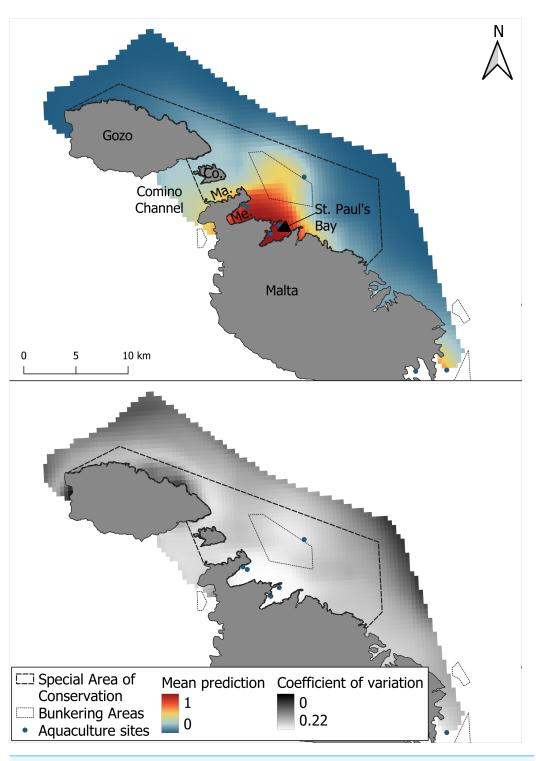


Figure 6 Mean prediction for bottlenose dolphin habitat suitability using the MaxEnt model and the coefficient of variation across 1,000 iterations of the modelling process.

DISCUSSION

Effective conservation planning relies on identifying areas that are important for species survival and on understanding how environmental factors influence their distribution. SDMs are widely used tools for obtaining this information, particularly in data-limited contexts (*Elith & Leathwick*, 2009; *Guisan et al.*, 2013). This case study identifies core suitable habitats for a top marine predator by linking occurrence records, obtained with the support of citizen science, to key environmental variables. Through the application of a GAM and a MaxEnt model, summer habitat suitability maps were generated for the bottlenose dolphin in a relatively understudied area of Malta, offering a practical tool to support decision-makers in the formulation of conservation strategies.

This study examined the bottlenose dolphin summer habitat suitability, revealing that the northeastern coastal waters of the archipelago represent an important summer habitat for the species. In addition to being in line with previous broader research in the region (Soster et al., 2025), this study provides a finer-scale perspective by adding spatial detail and seasonal context to earlier habitat suitability assessments. Both the GAM and the MaxEnt models used in this research identified the area between Mellieha Bay and St Paul's Bay as a region of high suitability, suggesting that this portion of the study area may represent a key habitat for the species in summer. This area coincides with shallow, productive waters and proximity to inshore aquaculture sites, factors previously associated with higher dolphin presence in other Mediterranean regions, where similar environmental and anthropogenic features have been linked to increased dolphin presence (Díaz López, 2012; Pace et al., 2019; Gnone et al., 2022; La Manna et al., 2023a; Bellingeri et al., 2025). Moderate suitability was also found across the broader study area, with particular interest in the Comino Channel, but with some degree of spatial variability between the two models.

To better understand the environmental drivers behind these patterns, the relative importance of predictor variables for both models was examined. While GAM and MaxEnt differ in their analytical frameworks, both models' outputs revealed a comparable pattern in the importance of environmental variables shaping bottlenose dolphin habitat preference. Particularly, both models identified chlorophyll-a and distance to aquaculture as the most important predictors of habitat suitability, while slope and SST anomaly contributed less significantly. However, the relative contribution of these variables differed between the two approaches. In the GAM, chlorophyll-a explained the highest proportion of deviance, followed closely by distance to aquaculture. In contrast, MaxEnt revealed distance to aquaculture as the most influential predictor based on its permutation importance, with chlorophyll-a as the second most important variable. These differences reflect how each model handles variable interactions (*Guisan & Zimmermann, 2000*; *Elith et al., 2011*), but the consistency in identifying chlorophyll-a and distance to aquaculture as the top predictors highlights their ecological relevance to this population.

To further clarify these ecological patterns, the response curves offer information on how particular environmental factors affect habitat suitability. Higher chlorophyll-a concentrations showed a positive relationship with habitat suitability in both models, likely reflecting higher primary productivity and, by extension, greater prey availability (*Methion*

et al., 2023). Chlorophyll-a has frequently been used as a proxy for such favorable foraging conditions, given its correlation with the abundance of zooplankton and small pelagic fish (Torres, Read & Halpin, 2008). Indeed, its importance as a predictor of dolphin occurrence and habitat use has been demonstrated in other regions of the Mediterranean (*La Manna*, Ronchetti & Sarà, 2016; Giannoulaki et al., 2017; Karamitros et al., 2020). In this study, chlorophyll-a levels were likely influenced by both natural productivity and anthropogenic input from aquaculture (Díaz López & Methion, 2017), as chlorophyll-a concentrations were consistently higher in areas surrounding aquaculture sites during the summer months (Fig. S1), contributing to their predictive strength in both models. Additionally, bottlenose dolphins are highly opportunistic predators with a well-documented behavioral plasticity (Reynolds, Wells & Eide, 2013), often concentrating their foraging activity in areas where prey is predictably abundant. Suitability also declined with increasing distance from aquaculture sites, suggesting that these areas may provide foraging opportunities by promoting local prey aggregations (Díaz López, 2012; Díaz López, 2017; Díaz López & Methion, 2017). However, this does not imply that the area would be unsuitable in the absence of aquaculture; rather, the presence of these sites appears to increase habitat use within an ecologically favorable region. Finally, while the observed negative correlation between bottlenose dolphin presence and distance to aquaculture facilities improves our ability to predict their habitat preferences, this aggregation near fish farms may pose conservation concerns, as previous studies have associated such clustering with increased risks of vessel collisions, entanglement, habitat degradation, and exposure to elevated noise and light pollution (Guisan & Zimmermann, 2000; Martino et al., 2021).

While chlorophyll-a and distance to aquaculture were the strongest predictors in both models, slope and SST anomaly also contributed to explaining dolphin habitat suitability, albeit to a lesser extent. In the GAM, slope showed a significant but weaker relationship, with habitat suitability peaking at intermediate values. Similarly, the MaxEnt response curve indicated that suitability decreased on either side of an optimal slope range, suggesting a preference for specific seabed structures. Both models indicated a preference for areas characterized by gentle slopes (~2.5–3°), indicating that bottlenose dolphins in the study area are associated with mildly inclined rather than very flat or steep seabed. This is consistent with previous findings suggesting that moderate bottom inclinations may support preferred foraging or transit zones, particularly in coastal environments where bathymetric features influence prey availability (*Cañadas, Sagarminaga & Garci'a-Tiscar, 2002; Gnone et al., 2022*).

Although SST anomaly was the weakest predictor in both models, the response curves indicated a consistent negative relationship between increasing temperature anomalies and habitat suitability. This suggests that bottlenose dolphins may tend to avoid areas experiencing greater thermal deviation, potentially due to indirect effects on prey availability or ecosystem stability (*Wernberg et al.*, 2013). In the Mediterranean, SST anomalies have been associated with shifts in the abundance and distribution of demersal and pelagic fish species (*Sabatés et al.*, 2006; *Hidalgo et al.*, 2011), which could in turn reduce the attractiveness of the areas affected to predators. Given that the study only covered summer months, the role of thermal variability might be underestimated, and further investigation

across seasons would be needed to better understand its influence. This is of particular importance, in light of the potential consequences of extreme weather events such as marine heatwaves that are becoming more frequent in the Mediterranean (*Darmaraki et al.*, 2019; Wild et al., 2019).

One important finding of this study is the spatial agreement shown by the two models in predicting areas of high habitat suitability, particularly in the waters surrounding Mellieha Bay and St Paul's Bay. However, a certain difference between the two modeling approaches was observed. Particularly, the GAM predicted a broader extent of high suitability, including the Comino Channel, as well as moderate suitability across most of the study area. In contrast, the MaxEnt model identified a more spatially restricted area of high suitability in the channel, primarily concentrated in Marfa Bay, with moderate suitability extending across the bunkering area. Additionally, while the coefficient of variation was higher around the SAC boundaries for both models, the MaxEnt predictions exhibited greater uncertainty within the SAC itself, particularly around the northern coastal waters of Gozo relative to the overall coefficient of variation. These differences likely reflect the methodological dissimilarities between the models and their respective sensitivities to the spatial distribution and structure of the input data. GAM's use of pseudo-absence data and smooth functions may have allowed it to detect localized patterns of suitability, particularly if species-environment relationships in this area are non-linear or influenced by smaller gradients (Guisan, Edwards & Hastie, 2002; Wood, 2017). In contrast, MaxEnt's reliance on background sampling and global feature fitting may have reduced sensitivity to weak or spatially clustered signals in regions with lower observation density (Elith et al., 2011; Merow, Smith & Silander, 2013). Interestingly, while MaxEnt's outputs have been found to be only similar to those of GAMs when background points are drawn from observed absences (Fiedler et al., 2018), these results show spatial agreement despite the use of pseudo-absences and background data, suggesting a robust ecological indicator captured by both modeling approaches. These model-specific differences highlight the value of employing complementary modeling techniques, as they can reveal different aspects of species-habitat relationships and strengthen the ecological interpretation of suitability maps (Guisan & Zimmermann, 2000; Elith & Graham, 2009).

Given its significance in conservation applications of SDMs (*Guisan et al., 2013*), it is important to highlight the low degree of uncertainty found in the central core area identified between Mellieha Bay and St. Paul's Bay. While some uncertainty is inherently irreducible (*Regan et al., 2005*), the low uncertainty and strong spatial agreement of highly suitable areas, together with the high predictive performance of both models (AUC > 0.85), reinforce confidence in the identified ecological patterns. Making the uncertainty explicit also helps stakeholders understand the level of confidence they can place in the prediction (*Regan et al., 2005*). The highly suitable area found by both models should therefore be treated as a conservation priority, while the broader GAM-predicted moderately suitable areas may inform future monitoring efforts. In data-limited contexts, this complementarity can provide a more holistic and robust foundation for decision-making.

Incorporating citizen science data was a key aspect of this study, enabling a nearly threefold increase in presence records as well as expanding the number of sightings in areas

with less survey coverage. In contrast, models based solely on research data predicted low habitat suitability in the Comino Channel, despite frequent dolphin sightings reported by citizen scientists (Fig. 2). The reduced coefficient of variation in both the GAM and MaxEnt models using the full dataset further underscored the value of including opportunistic observations, especially when systematic survey effort is limited. These findings align with previous research showing that appropriately incorporating citizen science and opportunistic data can significantly enhance species distribution models (*Tiago, Pereira & Capinha, 2017; Coxen et al., 2017; Matutini et al., 2021*). Our results contribute to the growing recognition of non-traditional data sources as valuable tools for contributing to cost-effective long-term monitoring and the generation of reliable habitat suitability predictions. These predictions help govern conservation, particularly in marine environments where systematic data collection is often constrained (*Giovos et al., 2016*; *Pace et al., 2019*).

Protecting bottlenose dolphins in Maltese waters effectively calls for the inclusion of species—habitat knowledge into site-based and pressure-based management approaches (*Reisinger, Johnson & Friedlaender, 2022*). This study adopts a conservation-oriented perspective and identifies key summer habitats which, although located within an existing SAC, are not currently designated for the protection of bottlenose dolphins. The strong agreement between habitat suitability maps highlights consistent ecological patterns and supports spatial planning, local mitigation, and the potential inclusion of bottlenose dolphins in the site's conservation objectives. Furthermore, the integration of citizen science data demonstrates how participatory monitoring expands spatial coverage and enhances model reliability in under-surveyed areas. Employing complementary modeling approaches, such as MaxEnt and GAMs, further strengthens the ecological interpretation of habitat use and provides evidence for adaptive management in data-limited marine environments.

Study limitations

Despite providing valuable insights into the summer habitat suitability of bottlenose dolphins in Maltese waters, this work is subject to certain limitations that should be acknowledged. First, the study was limited to a single season, potentially overlooking important temporal dynamics and habitat preferences that vary across seasons. Moreover, data were collected during a single summer period, which may not fully capture interannual variability in dolphin occurrence and habitat use. As such, conclusions drawn are primarily reflective of summer habitat suitability patterns and may not represent consistent seasonal patterns across years or year-round habitat use. Secondly, although sampling bias was addressed in both modeling exercises, the integration of citizen science and opportunistic data may remain susceptible to uneven effort and reporting. These limitations may result in underrepresentation of certain areas, where lower predicted suitability could reflect limited observation effort rather than true absence. Third, the moderate spatial disagreement between the GAM and MaxEnt predictions, particularly in areas such as the Comino Channel, suggests that methodological sensitivity to data structure may affect spatial outputs. While both models showed strong AUC values, the reliance on pseudo-absence or

background data remains a known limitation in interpreting habitat suitability, particularly in data-poor or spatially heterogeneous systems (*Lobo*, *Jiménez-Valverde & Real*, 2008; *Phillips et al.*, 2009).

CONCLUSIONS

This study applied a comparative modeling approach using generalized additive models and MaxEnt to assess summer habitat suitability for bottlenose dolphins in a coastal Special Area of Conservation (SAC) in Malta. The models integrated citizen science and opportunistic data with environmental and anthropogenic variables to identify key drivers of dolphin habitat suitability and highlight important habitats within a data-limited context.

Both modeling approaches consistently identified shallow, nearshore regions, particularly around Mellieha Bay and St Paul's Bay, as key summer habitats for bottlenose dolphins. Chlorophyll-a concentration and proximity to aquaculture sites emerged as the strongest predictors, while slope and sea surface temperature anomaly played a secondary role. Despite some methodological differences between the models, the observed spatial agreement strengthens confidence in the ecological relevance of these findings.

By combining traditional ecological knowledge with citizen science contributions, this work demonstrates a scalable and cost-effective approach for habitat assessment, especially valuable in regions where systematic monitoring is limited. The results contribute to filling critical knowledge gaps on dolphin spatial ecology in Malta's coastal waters and offer science-based support for potential updates to conservation planning frameworks, including the consideration of bottlenose dolphins as a qualifying species within the SAC.

However, this study is not without limitations. The research focused on a single summer season and was based on presence-only data, limiting its ability to capture interannual variability and broader seasonal dynamics. Additionally, while sampling bias was addressed, the use of opportunistic data remains susceptible to uneven spatial effort and reporting.

Future research should prioritize increasing the temporal resolution of data collection across multiple seasons and years, improving systematic survey coverage, and further investigating the ecological linkages between aquaculture, productivity hotspots, and dolphin habitat use. Exploring these dynamics under different climatic scenarios, including marine heatwaves, may also provide critical insights for adaptive conservation management.

In summary, this case study contributes to the growing body of evidence supporting the integration of SDMs and participatory data collection into marine conservation planning. The approach presented here offers a valuable tool for informing spatial management decisions aimed at balancing species protection with human activities.

ACKNOWLEDGEMENTS

This study was carried out as part of the project "Coastal Opportunities for Climate Change Adaptation through Sustainable Tourism, Research and Integrated Marine Ecosystem Management" (COASTWISE). Visual surveys were conducted during marine life observation tours with Discover the Blu. The staff of the Malta College of Arts, Science & Technology (MCAST) is gratefully acknowledged for their support and guidance during

this project. In particular, Professor Aldo Drago for providing oceanographic datasets through the STREAM platform (https://app.stream-srf.com/auth/welcome). The authors extend their gratitude to the research assistants Leah Camilleri and Joseph Cassar and to all the volunteers, participants and citizens who took part in the data collection, particularly Benjamin Metzger, Lucien Cambie, Sam Bard, Marie Delattre, Gabriela Górska and Justine Previ. The authors are especially grateful to Henry Copperstone and Daniele Baraggioli for providing observation platforms that complemented vessel-based surveys and protocol development. Finally, the authors are grateful to Martina Cutajar and Shona Brincat from the Ċirkewwa Marine Park for their collaboration in facilitating the citizen science campaign.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by Malta's Ministry for Education, Sport, Youth, Research and Innovation (MEYR). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors: Malta's Ministry for Education, Sport, Youth, Research and Innovation (MEYR).

Competing Interests

Timothy Zammit is the owner of Discover the Blu. Nina Vérité-Taulet is employed by Discover the Blu.

Author Contributions

- Francesca Soster conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Tim Awbery conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Nina Vérité–Taulet performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Timothy Zammit performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Kimberly Terribile conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (*i.e.*, approving body and any reference numbers):

The animal study protocol was approved by Malta's Environment and Resources Authority (ERA), Malta (permit number EP 0249/24, issued on 12 July 2024).

Data Availability

The following information was supplied regarding data availability:

The bottlenose dolphin sighting data are available in the Supplementary Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.19804#supplemental-information.

REFERENCES

- Anderson RP, Lew D, Peterson AT. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* **162**:211–232 DOI 10.1016/S0304-3800(02)00349-6.
- **Araújo MB, Guisan A. 2006.** Five (or so) challenges for species distribution modelling. *Journal of Biogeography* **33**:1677–1688 DOI 10.1111/j.1365-2699.2006.01584.x.
- **Azzellino A, Gaspari S, Airoldi S, Nani B. 2008.** Habitat use and preferences of cetaceans along the continental slope and the adjacent pelagic waters in the western Ligurian Sea. *Deep Sea Research Part I: Oceanographic Research Papers* **55**:296–323 DOI 10.1016/j.dsr.2007.11.006.
- Baker E, Drury JP, Judge J, Roy DB, Smith GC, Stephens PA. 2021. The verification of ecological citizen science data: current approaches and future possibilities. *Citizen Science: Theory and Practice* 6(1):1–14 DOI 10.5334/cstp.351.
- Barbet-Massin M, Jiguet F, Albert CH, Thuiller W. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3:327–338 DOI 10.1111/j.2041-210X.2011.00172.x.
- Bellingeri M, Nebuloni A, La Manna G, Akkaya A, Alessi J, Arcangeli A, Ascheri D, Awbery T, Azzinari G, Azzinari C, Azzolin M, Bacchetta M, Amer IB, Buscaino G, Calogero G, Campana I, Canadas AM, Carlucci R, Castelli A, Chicote CA, Corrias V, David L, Dede A, Degollada E, De Santis V, Dhermain F, Diaz Lopez B, Di Meglio N, Fontanesi E, Genov T, Giacoma C, Giardina F, Giralt Paradell O, Gonzalvo J, Labach H, Lanfredi C, Mancusi C, Mandich A, Menniti MA, Mevorach Y, Monaco C, Mussi B, Nuti S, Pace DS, Papale E, Paraboschi M, Pedrazzi G, Pellegrino G, Ronchetti F, Rutigliano R, Santacesaria FC, Santoni M-C, Stanzani L, Scheinin A, Tepsich P, Tintore B, Tomasi N, Tonay AM, Tort B, Tringali LM, Verga A, Vighi M, Violi B, Železnik J, Gnone G. 2025. Investigating the presence of different bottlenose dolphin ecotypes in the Mediterranean Sea. *Scientific Reports* 15:5526 DOI 10.1038/s41598-025-88605-0.
- **Bonter DN, Cooper CB. 2012.** Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* **10**:305–307 DOI 10.1890/110273.
- Breiner FT, Guisan A, Bergamini A, Nobis MP. 2015. Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution* **6**:1210–1218 DOI 10.1111/2041-210X.12403.

- Cañadas A, Sagarminaga R, Garcı'a-Tiscar S. 2002. Cetacean distribution related with depth and slope in the Mediterranean waters off southern Spain. *Deep Sea Research Part I: Oceanographic Research Papers* 49:2053–2073

 DOI 10.1016/S0967-0637(02)00123-1.
- Carlucci R, Fanizza C, Cipriano G, Paoli C, Russo T, Vassallo P. 2016. Modeling the spatial distribution of the striped dolphin (*Stenella coeruleoalba*) and common bottlenose dolphin (*Tursiops truncatus*) in the Gulf of Taranto (Northern Ionian Sea, Central-eastern Mediterranean Sea). *Ecological Indicators* 69:707–721 DOI 10.1016/j.ecolind.2016.05.035.
- Chen X, Dimitrov NB, Meyers LA. 2019. Uncertainty analysis of species distribution models. *PLOS ONE* 14:e0214190 DOI 10.1371/journal.pone.0214190.
- Corr S, Dudley R, Brereton T, Clear N, Crosby A, Duncan S, Evans PGH, Jones D, Sayer S, Taylor T, Tregenza N, Williams R, Witt MJ, Ingram SN. 2024. Using citizen science data to assess the vulnerability of bottlenose dolphins to human impacts along England's South Coast. *Animal Conservation* 27:461–477 DOI 10.1111/acv.12921.
- Coxen CL, Frey JK, Carleton SA, Collins DP. 2017. Species distribution models for a migratory bird based on citizen science and satellite tracking data. *Global Ecology and Conservation* 11:298–311 DOI 10.1016/j.gecco.2017.08.001.
- Darmaraki S, Somot S, Sevault F, Nabat P, Cabos Narvaez WD, Cavicchia L, Djurdjevic V, Li L, Sannino G, Sein DV. 2019. Future evolution of marine heatwaves in the Mediterranean Sea. *Climate Dynamics* 53:1371–1392 DOI 10.1007/s00382-019-04661-z.
- **Deneu B, Servajean M, Bonnet P, Botella C, Munoz F, Joly A. 2021.** Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. *PLOS Computational Biology* **17**:e1008856 DOI 10.1371/journal.pcbi.1008856.
- **Díaz López B. 2012.** Bottlenose dolphins and aquaculture: interaction and site fidelity on the north-eastern coast of Sardinia (Italy). *Marine Biology* **159**:2161–2172 DOI 10.1007/s00227-012-2002-x.
- **Díaz López B. 2017.** Temporal variability in predator presence around a fin fish farm in the Northwestern Mediterranean Sea. *Marine Ecology* **38**:e12378 DOI 10.1111/maec.12378.
- **Díaz López B, Methion S. 2017.** The impact of shellfish farming on common bottlenose dolphins' use of habitat: running head: impact of mussel farming on bottlenose dolphins. *Marine Biology* **164**:83 DOI 10.1007/s00227-017-3125-x.
- **Díaz López B, Methion S. 2024.** Killer whales habitat suitability in the Iberian Peninsula and the Gulf of Biscay: implications for conservation. *Ocean & Coastal Management* **255**:107245 DOI 10.1016/j.ocecoaman.2024.107245.
- Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36:27–46 DOI 10.1111/j.1600-0587.2012.07348.x.

- **Elith J, Graham CH. 2009.** Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* **32**:66–77 DOI 10.1111/j.1600-0587.2008.05505.x.
- Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JMM, Townsend Peterson A, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151 DOI 10.1111/j.2006.0906-7590.04596.x.
- **Elith J, Leathwick JR. 2009.** Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* **40**:677–697 DOI 10.1146/annurev.ecolsys.110308.120159.
- Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17:43–57 DOI 10.1111/j.1472-4642.2010.00725.x.
- EUNIS. EUNIS Site factsheet for Zona fil-Bahar bejn Il-Ponta ta' San Dimitri (Ghawdex) u Il-Qaliet. *Available at https://eunis.eea.europa.eu/sites/MT0000105* (accessed on April 21 2025).
- **Evans PGH, Hammond PS. 2004.** Monitoring cetaceans in European waters. *Mammal Review* **34**:131–156 DOI 10.1046/j.0305-1838.2003.00027.x.
- Fernandez M, Alves F, Ferreira R, Fischer J-C, Thake P, Nunes N, Caldeira R, Dinis A. 2021. Modeling fine-scale cetaceans' distributions in oceanic Islands: Madeira Archipelago as a case study. *Frontiers in Marine Science* **8**:688248 DOI 10.3389/fmars.2021.688248.
- **Fernandez M, Sillero N, Yesson C. 2022.** To be or not to be: the role of absences in niche modelling for highly mobile species in dynamic marine environments. *Ecological Modelling* **471**:110040 DOI 10.1016/j.ecolmodel.2022.110040.
- Fiedler PC, Redfern JV, Forney KA, Palacios DM, Sheredy C, Rasmussen K, García-Godos I, Santillán L, Tetley MJ, Félix F, Ballance LT. 2018. Prediction of large whale distributions: a comparison of presence–absence and presence-only modeling techniques. *Frontiers in Marine Science* 5:419 DOI 10.3389/fmars.2018.00419.
- **Fielding AH, Bell JF. 1997.** A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24**:38–49 DOI 10.1017/S0376892997000088.
- **Filletti JM, Gauci A, Deidun A, Riccobene G, Viola S. 2023.** A preliminary snapshot investigation of the marine soundscape for malta: a steppingstone towards achieving good ecological status. *Journal of Marine Science and Engineering* **11**:2163 DOI 10.3390/jmse11112163.
- Fontanesi E, Ascheri D, Bertulli CG, Salvioli F, McGinty N. 2024. Bottlenose dolphin (*Tursiops truncatus*) habitat partitioning in relation to age classes in the northwest Mediterranean Sea. *Marine Biology* 171:137 DOI 10.1007/s00227-024-04457-8.

- Garrabou J, Gómez-Gras D, Medrano A, Cerrano C, Ponti M, Schlegel R, Bensoussan N, Turicchia E, Sini M, Gerovasileiou V, Teixido N, Mirasole A, Tamburello L, Cebrian E, Rilov G, Ledoux J-B, Souissi JB, Khamassi F, Ghanem R, Benabdi M, Grimes S, Ocaña O, Bazairi H, Hereu B, Linares C, Kersting DK, Rovira Gla, Ortega J, Casals D, Pagès-Escolà M, Margarit N, Capdevila P, Verdura J, Ramos A, Izquierdo A, Barbera C, Rubio-Portillo E, Anton I, López-Sendino P, Díaz D, Vázquez-Luis M, Duarte C, Marbà N, Aspillaga E, Espinosa F, Grech D, Guala I, Azzurro E, Farina S, Cristina Gambi M, Chimienti G, Montefalcone M, Azzola A, Mantas TP, Fraschetti S, Ceccherelli G, Kipson S, Bakran-Petricioli T, Petricioli D, Jimenez C, Katsanevakis S, Kizilkaya IT, Kizilkaya Z, Sartoretto S, Elodie R, Ruitton S, Comeau S, Gattuso J-P, Harmelin J-G. 2022. Marine heatwaves drive recurrent mass mortalities in the Mediterranean Sea. *Global Change Biology* 28:5708–5725 DOI 10.1111/gcb.16301.
- **Giannoulaki M, Markoglou E, Valavanis VD, Alexiadou P, Cucknell A, Frantzis A. 2017.** Linking small pelagic fish and cetacean distribution to model suitable habitat for coastal dolphin species, *Delphinus delphis* and *Tursiops truncatus*, in the Greek Seas (Eastern Mediterranean). *Aquatic Conservation: Marine and Freshwater Ecosystems* **27**:436–451 DOI 10.1002/aqc.2669.
- **Giovos I, Ganias K, Garagouni M, Gonzalvo J. 2016.** Social media in the service of conservation: a case study of dolphins in the Hellenic seas. *Aquatic Mammals* **42**:12–20 DOI 10.1578/AM.42.1.2016.12.
- Giralt Paradell O, Díaz López B, Methion S. 2019. Modelling common dolphin (*Delphinus delphis*) coastal distribution and habitat use: insights for conservation. *Ocean & Coastal Management* 179:104836 DOI 10.1016/j.ocecoaman.2019.104836.
- Gnone G, Bellingeri M, Molinari Y, Dhermain F, Labach H, Díaz López B, David L, Di Meglio N, Azzinari G, Azzinari C, Airoldi S, Lanfredi C, Gonzalvo J, De Santis V, Nuti S, Álvarez Chicote C, Gazo M, Mandich A, Alessi J, Azzellino A, Tomasi N, Santoni M-C, Mancusi C, Falabrino M, Cañadas AM. 2022. The Seabed Makes the Dolphins: physiographic features shape the size and structure of the bottlenose dolphin geographical units. *Journal of Marine Science and Engineering* 10:1036 DOI 10.3390/jmse10081036.
- **Guisan A, Edwards TC, Hastie T. 2002.** Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* **157**:89–100 DOI 10.1016/S0304-3800(02)00204-1.
- **Guisan A, Thuiller W. 2005.** Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**:993–1009 DOI 10.1111/j.1461-0248.2005.00792.x.
- Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AIT, Regan TJ, Brotons L, Mcdonald-Madden E, Mantyka-Pringle C, Martin TG, Rhodes JR, Maggini R, Setterfield SA, Elith J, Schwartz MW, Wintle BA, Broennimann O, Austin M, Ferrier S, Kearney MR, Possingham HP, Buckley YM. 2013. Predicting species distributions for conservation decisions. *Ecology Letters* 16:1424–1435 DOI 10.1111/ele.12189.

- **Guisan A, Zimmermann NE. 2000.** Predictive habitat distribution models in ecology. *Ecological Modelling* **135**:147–186 DOI 10.1016/S0304-3800(00)00354-9.
- **Hastie T, Tibshirani R. 1986.** Generalized additive models. *Statistical Science* 1:297–310 DOI 10.1214/ss/1177013604.
- Hernandez PA, Graham CH, Master LL, Albert DL. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773–785 DOI 10.1111/j.0906-7590.2006.04700.x.
- Hidalgo M, Rouyer T, Molinero JC, Massutí E, Moranta J, Guijarro B, Stenseth NC. 2011. Synergistic effects of fishing-induced demographic changes and climate variation on fish population dynamics. *Marine Ecology Progress Series* 426:1–12 DOI 10.3354/meps09077.
- **Hijmans RJ, Phillips S, Leathwick J, Elith J. 2024.** dismo: species distribution modeling. R package version 1.3-16. *Available at https://rspatial.org/raster/sdm/*.
- **Hirzel AH, Le Lay G. 2008.** Habitat suitability modelling and niche theory. *Journal of Applied Ecology* **45**:1372–1381 DOI 10.1111/j.1365-2664.2008.01524.x.
- Isaac NJB, Jarzyna MA, Keil P, Dambly LI, Boersch-Supan PH, Browning E, Freeman SN, Golding N, Guillera-Arroita G, Henrys PA, Jarvis S, Lahoz-Monfort J, Pagel J, Pescott OL, Schmucki R, Simmonds EG, O'Hara RB. 2020. Data integration for large-scale models of species distributions. *Trends in Ecology & Evolution* 35:56–67 DOI 10.1016/j.tree.2019.08.006.
- **Karamitros G, Gkafas GA, Giantsis IA, Martsikalis P, Kavouras M, Exadacty-los A. 2020.** Model-based distribution and abundance of three delphinidae in the Mediterranean. *Animals: An Open Access Journal from MDPI* **10**:260 DOI 10.3390/ani10020260.
- **La Manna G, Ronchetti F, Perretti F, Ceccherelli G. 2023a.** Determinants of group size in the common bottlenose dolphins: the role of water temperature and noise. *Behavioral Ecology and Sociobiology* **77**:45 DOI 10.1007/s00265-023-03322-w.
- **La Manna G, Ronchetti F, Perretti F, Ceccherelli G. 2023b.** Not only wide range shifts: marine warming and heat waves influence spatial traits of a Mediterranean common bottlenose dolphin population. *Estuarine, Coastal and Shelf Science* **285**:108320 DOI 10.1016/j.ecss.2023.108320.
- La Manna G, Ronchetti F, Sarà G. 2016. Predicting common bottlenose dolphin habitat preference to dynamically adapt management measures from a Marine Spatial Planning perspective. *Ocean & Coastal Management* 130:317–327 DOI 10.1016/j.ocecoaman.2016.07.004.
- Laspina M, Terribile K, Said A. 2022. Interactions between cetaceans and small-scale fisheries around the central Mediterranean Maltese islands. *MCAST Journal of Applied Research & Practice* 6:113–126 DOI 10.5604/01.3001.0015.9134.
- **Liu C, White M, Newell G. 2011.** Measuring and comparing the accuracy of species distribution models with presence–absence data. *Ecography* **34**:232–243 DOI 10.1111/j.1600-0587.2010.06354.x.

- **Lobo JM, Jiménez-Valverde A, Real R. 2008.** AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17(2)**:145–151 DOI 10.1111/j.1466-8238.2007.00358.x.
- Maricato G, Tardin R, Lodi L, Wedekin LL, Daura-Jorge FG, Maciel I, Maria TF, Alves MAS. 2022. Identifying suitable areas for common bottlenose dolphin in anthropized waters. *Marine Biology* **169**:110 DOI 10.1007/s00227-022-04095-y.
- Marshall CE, Glegg GA, Howell KL. 2014. Species distribution modelling to support marine conservation planning: the next steps. *Marine Policy* 45:330–332 DOI 10.1016/j.marpol.2013.09.003.
- Martino S, Pace DS, Moro S, Casoli E, Ventura D, Frachea A, Silvestri M, Arcangeli A, Giacomini G, Ardizzone G, Jona Lasinio G. 2021. Integration of presence-only data from several sources: a case study on dolphins' spatial distribution. *Ecography* 44:1533–1543 DOI 10.1111/ecog.05843.
- Matutini F, Baudry J, Pain G, Sineau M, Pithon J. 2021. How citizen science could improve species distribution models and their independent assessment. *Ecology and Evolution* 11:3028–3039 DOI 10.1002/ece3.7210.
- Melo-Merino SM, Reyes-Bonilla H, Lira-Noriega A. 2020. Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. *Ecological Modelling* 415:108837 DOI 10.1016/j.ecolmodel.2019.108837.
- Merow C, Smith MJ, Silander JA. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36:1058–1069 DOI 10.1111/j.1600-0587.2013.07872.x.
- Methion S, Giralt Paradell O, Padín XA, Corrège T, Díaz López B. 2023. Group size varies with climate and oceanographic conditions in bottlenose dolphins. *Marine Biology* 170:7 DOI 10.1007/s00227-022-04154-4.
- Meyer CFJ, Aguiar LMS, Aguirre LF, Baumgarten J, Clarke FM, Cosson J-F, Estrada Villegas S, Fahr J, Faria D, Furey N, Henry M, Jenkins RKB, Kunz TH, Cristina MacSwiney González M, Moya I, Pons J-M, Racey PA, Rex K, Sampaio EM, Stoner KE, Voigt CC, von Staden D, Weise CD, Kalko EKV. 2015. Species undersampling in tropical bat surveys: effects on emerging biodiversity patterns. *Journal of Animal Ecology* 84:113–123 DOI 10.1111/1365-2656.12261.
- **Mizzi M, Deidun A, Gauci A, Gauci R. 2024.** The impact of anchoring on seafloor integrity: an integrated assessment within a major bunkering area of the Maltese Islands. *Geographies* 4:612–629 DOI 10.3390/geographies4040033.
- Notarbartolo Di Sciara G. 2002. Action plan for the conservation of cetaceans in Maltese waters. Tunis, Tunisia: Regional Activity Centre for Specially Protected Areas (RAC/SPA), UNEP/MAP. Prepared under the SAP/BIO framework. *Available at https://www.disciara.org/documents/NotarbartolodiSciara_Mifsud_2002.pdf* (accessed on 20 January 2025).
- Pace DS, Giacomini G, Campana I, Paraboschi M, Pellegrino G, Silvestri M, Alessi J, Angeletti D, Cafaro V, Pavan G, Ardizzone G, Arcangeli A. 2019. An integrated approach for cetacean knowledge and conservation in the central Mediterranean

- Sea using research and social media data sources. *Aquatic Conservation: Marine and Freshwater Ecosystems* **29**:1302–1323 DOI 10.1002/aqc.3117.
- Pace DS, Tizzi R, Mussi B. 2015. Cetaceans value and conservation in the Mediterranean Sea. *Journal of Biodiversity & Endangered Species* 1(S1):S1-004 DOI 10.4172/2332-2543.S1-004.
- Pasanisi E, Pace DS, Orasi A, Vitale M, Arcangeli A. 2024. A global systematic review of species distribution modelling approaches for cetaceans and sea turtles. *Ecological Informatics* 82:102700 DOI 10.1016/j.ecoinf.2024.102700.
- Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231–259 DOI 10.1016/j.ecolmodel.2005.03.026.
- Phillips SJ, Dudík M. 2008. Modeling of species distributions with MaxEnt: new extensions and a comprehensive evaluation. *Ecography* 31:161–175 DOI 10.1111/j.0906-7590.2008.5203.x.
- Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* 19:181–197 DOI 10.1890/07-2153.1.
- Pitchford JL, Howard VA, Shelley JK, Serafin BJS, Coleman AT, Solangi M. 2016.

 Predictive spatial modelling of seasonal bottlenose dolphin (*Tursiops truncatus*) distributions in the Mississippi Sound. *Aquatic Conservation: Marine and Freshwater Ecosystems* 26:289–306 DOI 10.1002/aqc.2547.
- **R Core Team. 2025.** R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. *Available at https://www.r-project.org/*.
- Redfern J, Ferguson M, Becker E, Hyrenbach K, Good C, Barlow J, Kaschner K, Baumgartner M, Forney K, Ballance L, Fauchald P, Halpin P, Hamazaki T, Pershing A, Qian S, Read A, Reilly S, Torres L, Werne F. 2006. Techniques for cetacean habitat modeling. *Marine Ecology Progress Series* 310:271–295 DOI 10.3354/meps310271.
- Regan HM, Ben-Haim Y, Langford B, Wilson WG, Lundberg P, Andelman SJ, Burgman MA. 2005. Robust decision-making under severe uncertainty for conservation management. *Ecological Applications* 15:1471–1477 DOI 10.1890/03-5419.
- **Reisinger RR, Johnson C, Friedlaender AS. 2022.** Marine Mammal Movement Ecology in a Conservation and Management Context. In: Notarbartolo diSciara G, Würsig B, eds. *Marine mammals: the evolving human factor*. Cham: Springer International Publishing, 149–192 DOI 10.1007/978-3-030-98100-6_5.
- **Reynolds JE, Wells RS, Eide SD. 2013.** *The bottlenose dolphin: biology and conservation.* Gainesville: University Press of Florida.
- **Robbins JR, Babey L, Embling CB. 2020.** Citizen science in the marine environment: estimating common dolphin densities in the north-east Atlantic. *PeerJ* **8**:e8335 DOI 10.7717/peerj.8335.
- **Rodríguez JP, Brotons L, Bustamante J, Seoane J. 2007.** The application of predictive modelling of species distribution to biodiversity conservation. *Diversity & Distributions* **13**:243–251 DOI 10.1111/j.1472-4642.2007.00356.x.

- **Sabatés A, Martín P, Lloret J, Raya V. 2006.** Sea warming and fish distribution: the case of the small pelagic fish, *Sardinella aurita*, in the western Mediterranean. *Global Change Biology* **12**:2209–2219 DOI 10.1111/j.1365-2486.2006.01246.x.
- Said A, MacMillan D, Schembri M, Tzanopoulos J. 2017. Fishing in a congested sea: what do marine protected areas imply for the future of the Maltese artisanal fleet? *Applied Geography* 87:245–255 DOI 10.1016/j.apgeog.2017.08.013.
- Shabani F, Kumar L, Ahmadi M. 2018. Assessing accuracy methods of species distribution models: AUC, specificity, sensitivity and the true skill statistic. *Global Journal of Human-Social Science: B Geography, Geo-Sciences, Environmental Science & Disaster Management* 18(1).
- Soster F, Galdies C, Awbery T, Gauci A, Metzger B, Barbara N, Díaz López B. 2025.

 Common bottlenose dolphin habitat suitability in Malta's coastal waters as a humanaltered environment. SSRN Preprint DOI 10.2139/ssrn.5220799.
- **Swets JA. 1988.** Measuring the accuracy of diagnostic systems. *Science* **240**:1285–1293 DOI 10.1126/science.3287615.
- **Tiago P, Pereira HM, Capinha C. 2017.** Using citizen science data to estimate climatic niches and species distributions. *Basic and Applied Ecology* **20**:75–85 DOI 10.1016/j.baae.2017.04.001.
- **Torres LG, Read AJ, Halpin P. 2008.** Fine-scale habitat modeling of a top marine predator: do prey data improve predictive capacity? *Ecological Applications* **18**:1702–1717 DOI 10.1890/07-1455.1.
- Wernberg T, Smale DA, Tuya F, Thomsen MS, Langlois TJ, De Bettignies T, Bennett S, Rousseaux CS. 2013. An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nature Climate Change* 3:78–82 DOI 10.1038/nclimate1627.
- Wild S, Krützen M, Rankin RW, Hoppitt WJE, Gerber L, Allen SJ. 2019. Long-term decline in survival and reproduction of dolphins following a marine heatwave. *Current Biology* **29**:R239–R240 DOI 10.1016/j.cub.2019.02.047.
- **Wood SN. 2017.** *Generalized additive models: an introduction with R, second edition.* New York: Chapman and Hall/CRC DOI 10.1201/9781315370279.
- Yu J, Wong W-K, Hutchinson RA. 2010. Modeling experts and novices in citizen science data for species distribution modeling. In: 2010 IEEE international conference on data mining. Piscataway: IEEE, 1157–1162 DOI 10.1109/ICDM.2010.103.