Acta Adriatica 65 (2024): 61-74
https://doi.org/10.32582/aa.65.1.8
ORIGINAL ARTICLE

www.izor.hr

Offshore distribution of yelkouan shearwaters in the north-western Adriatic Sea: insight from machine learning

Silvia Bonizzoni[1,2*], Blair D. Sterba-Boatwright[3], Sarah Piwetz[1] and Giovanni Bearzi[1,2,4]

[1] *Dolphin Biology and Conservation, Cordenons, Italy*
[2] *OceanCare, Wädenswil, Switzerland*
[3] *Texas A&M University, Corpus Christi, TX, USA*
[4] *ISMAR Institute of Marine Sciences, CNR National Research Council, Venice, Italy*

Abstract: The yelkouan shearwater *Puffinus yelkouan* is endemic to the Mediterranean and Black Seas, and classified as Vulnerable in the IUCN Red List. Information on the species' distribution and habitat use in the eastern Mediterranean is scant, and only a few studies were based on direct visual observations in offshore waters. Here, we provide information on 1) the occurrence of yelkouan shearwaters within a 3000 km^2 study area off the region of Veneto, Italy, in the north-western Adriatic Sea, based on visual surveys conducted from small boats between April and October 2018–2022 (effort: 169 days, 23,836 km), and 2) the geographic, bathymetric and oceanographic variables likely to drive the species' offshore distribution. Yelkouan shearwaters (238 sightings, 916 individuals) were observed in waters 9–33 m deep, between 2 and 24 km from the coast. Individual counts ranged between 1 and 100, with 95% of the encounters having less than 10 individuals. An Explainable Boosting Machine model – a machine learning technique based on generalized additive models – selected chlorophyll *a* as the most important variable to explain the species' occurrence, followed by distance from the coast, and bottom depth. The model indicated a higher occurrence in waters with chlorophyll *a* less than ~2.3 mg/m^3, farther than ~15 km from the coast, and deeper than ~22 m. The effects of SST, salinity, and day of the year were less clear. This study provides insight into the offshore distribution of yelkouan shearwaters, within one of the Mediterranean areas most exposed to cumulative human threats.

Keywords: distribution; Explainable Boosting Machine; machine learning; Mediterranean Sea; occurrence; Procellariidae; shearwaters; visual survey

Sažetak: RASPROSTRANJENOST GREGULE U OTVORENIM VODAMA SJEVEROZAPADNOG JADRANA: SPOZNAJE NA TEMELJU STROJNOG UČENJA. Gregula (*Puffinus yelkouan*) je endemska vrsta Sredozemnog i Crnog mora koja se nalazi na IUCN crvenom popisu u kategoriji ranjivih vrsta. Informacije o rasprostranjenosti ove vrste i staništima u istočnom Sredozemnom moru su oskudne, a samo se nekoliko studija temeljilo na izravnim vizualnim opažanjima na otvorenom moru. Ovdje donosimo informacije o 1) pojavi gregule na 3000 km^2 istraživanog područja ispred obale regije Veneto u Italiji, u sjeverozapadnom Jadranskom moru, temeljem vizualnih opažanja provedenih s malih brodova između travnja i listopada 2018. i 2022. (napor: 169 dana, 23 836 km) i 2) geografskim, batimetrijskim i oceanografskim varijablama koje vjerojatno utječu na distribuciju ove vrste na otvorenom moru. Gregule (238 opažanja, 916 jedinki) su uočene na morskom području gdje se dubine kreću od 9 do 33 m, između 2 i 24 km od obale. U pojedinačnim opažanjima izbrojeno je između jedne i 100 jedinki, a u 95% slučajeva manje od 10 jedinki. Primjenom EBM (*Explainable Boosting Machine*) modela, tehnike strojnog učenja zasnovane na generaliziranim aditivnim modelima, odabran je klorofil *a* kao najvažnija varijabla za objašnjenje pojave vrste, nakon čega slijedi udaljenost od obale i dubina. Model je ukazao na veću pojavu vrste u vodama s koncentracijom klorofila *a* manjom od ~2,3 mg/m^3, udaljenijim od ~15 km od obale i dubljim od ~22 m. Učinci SST-a, slanosti i dana u godini bili su manje jasni. Ova studija daje uvid u rasprostranjenost gregule na otvorenom moru, u jednom od Sredozemnih područja koja su najpodložnija kumulativnim prijetnjama ljudskih aktivnosti.

Ključne riječi: rasprostranjenost; *Explainable Boosting Machine*; strojno učenje; Sredozemno more; pojava; Procellariidae; gregule; vizualno opažanje

INTRODUCTION

The yelkouan shearwater *Puffinus yelkouan* (Fig. 1) is a pelagic seabird from the family Procellariidae, endemic to the central and eastern Mediterranean Sea, and Black Sea (Bourgeois and Vidal, 2008; Genovart *et al.*, 2012; Pérez-Ortega and İsfendiyaroğlu, 2017; BirdLife International, 2021a; Pezzo *et al.*, 2021). The breeding population in Europe is estimated to include between 47,000 and 81,800 mature individuals (BirdLife International, 2021a). This figure should be taken with caution as it was derived from a combination of estimates based on different time frames and methodologies (BirdLife International, 2021b). The

Corresponding author: silvia.bonizzoni@gmail.com

Received: 29 January 2024, accepted: 22 April 2024
ISSN: 0001-5113, eISSN: 1846-0453

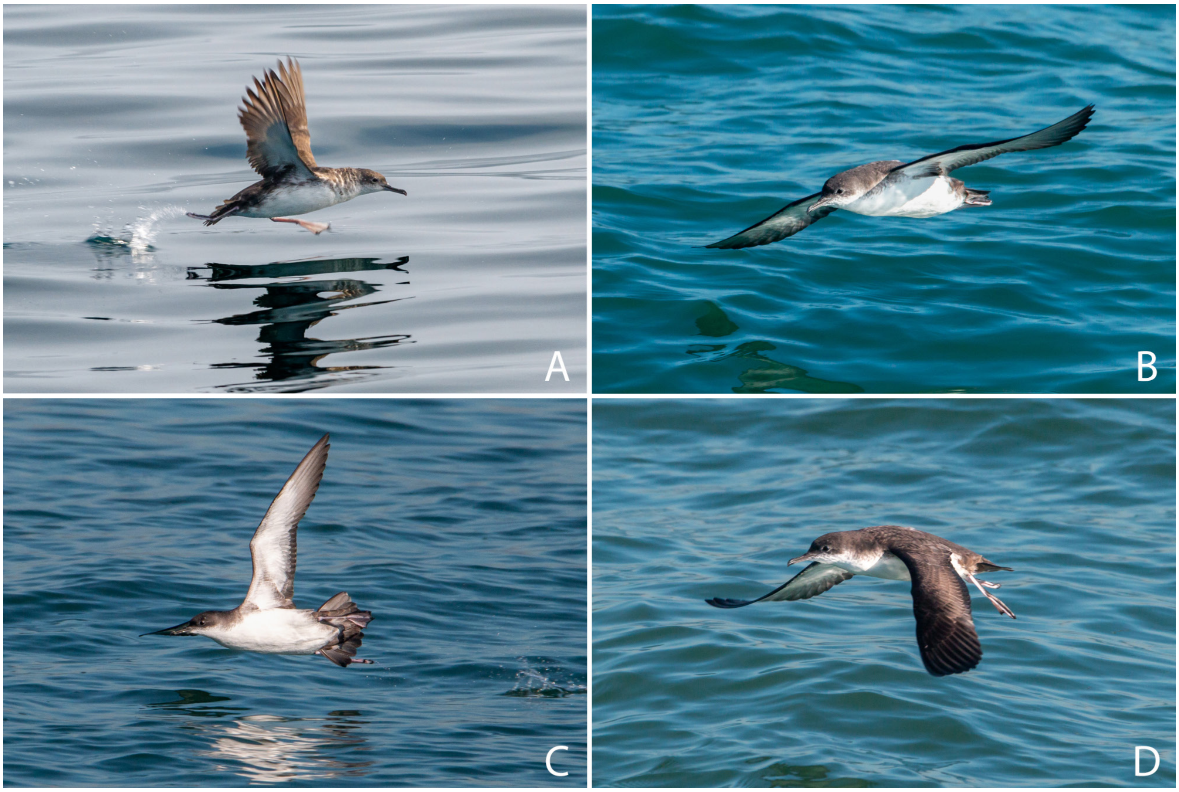

Fig. 1. Yelkouan shearwaters photographed off the region of Veneto, north-western Adriatic Sea. Photos by Silvia Bonizzoni (**A**) and Federico Fanesi (**B, C, D**).

species is currently classified as Vulnerable in the IUCN Red List (BirdLife International, 2021a), due to population decline and low breeding success caused by predation (largely by introduced mammals; Tranchant *et al*., 2003; UNEP/MAP-RAC/SPA, 2003; Bonnaud *et al*., 2007; Gaudard, 2018) and low survival rates (due to incidental mortality in fishing gear, and to predation; Bourgeois and Vidal, 2008; Oppel *et al*., 2011; BirdLife International, 2021a). The species is also included in Annex II of the Specially Protected Areas and Biological Diversity (SPA/BD) Protocol of the Barcelona Convention ("List of endangered or threatened species"; EC, 1999), in Annex II of the Bern Convention ("Strictly protected fauna species"; EC, 1979), in Annex I of the Birds Directive ("Subjects of special conservation measures concerning their habitat in order to ensure their survival and reproduction in their area of distribution"; EC, 2010). In addition, an "Action Plan for the conservation of marine and coastal bird species listed in Annex II of the Protocol concerning Specially Protected Areas and Biological Diversity in the Mediterranean" (UNEP/MAP-RAC/SPA, 2017) lists twelve actions to protect yelkouan shearwaters, including the identification of important at-sea areas. The "International single species action plan for the yelkouan shearwater *Puffinus yelkouan*" (Gaudard, 2018) identified three main conservation goals: 1) increase adult survival to 92% or more and breeding success to 75% or more, 2) improve the quality of breeding habitat, and 3) fill knowledge gaps about species' distribution and population estimates.

Broad areas currently identified as important for yelkouan shearwaters include the Alborán Sea, the Tyrrhenian Sea, the Gulf of Lions, the waters of Algeria, Tunisia, south-eastern Sicily (Italy) and Israel, the northern Adriatic Sea, the Aegean Sea and the Black Sea (Bourgeois and Vidal, 2008). A total of 75 islands in the Mediterranean and Black Seas were identified as certain or possible breeding sites (Bourgeois and Vidal, 2008). The southernmost breeding site may be the island of Crete, in Greece (Genovart *et al.,* 2012). Most of the information on the distribution and habitat use of yelkouan shearwaters comes from telemetry tracking of single individuals (Péron *et al.,* 2013; Raine *et al.,* 2013; Pezzo *et al.,* 2021; Zec *et al*. 2023) and land-based studies at nesting sites (Borg *et al.,* 2002; Bourgeois *et al.,* 2011; Oppel *et al.,* 2011), whereas few studies have been based on direct visual observations in offshore waters. Studies of this kind were conducted in the Gulf of Lion, southern France (Péron *et al.,* 2013), off southern Sicily, Italy (Ranù *et al.,* 2022), off the Adriatic Sea coast of Croatia (Stipčević and Lukač, 2001; Zec *et al.* 2023), in the northern Aegean Sea, Greece (Zakkak *et al.,* 2013), and in the central-western Black Sea (Pérez Ortega and İsfendiyaroğlu, 2017).

The northern Adriatic Sea represents an important area for yelkouan shearwater moulting, feeding and passage during the interbreeding period (Stipčević and Lukač, 2001; Bourgeois and Vidal, 2008). A review of sighting data collected in the eastern Adriatic between

1888 and 1998 suggested a main occurrence between May and October, with a few sightings in April, November and December, and no sightings in January and February (Stipčević and Lukač, 2001). Similarly, in the northern Adriatic, the period between June and October was identified as the one with the highest occurrence (Bourgeois and Vidal, 2008). Based on recent information from the checklist of the birds of Croatia between 2005 and 2012, yelkouan shearwaters were considered a "regular", "breeding", "migratory" and "wintering" species (Kralj and Barišić, 2013). However, information on distribution, habitat use and behaviour is scant in the Adriatic Sea and eastern Mediterranean, where quantitative investigations are lacking (UNEP/MAP-RAC/SPA, 2007; Bourgeois and Vidal, 2008).

In the north-western Adriatic Sea, this species was regularly observed in the Italian region of Friuli-Venezia Giulia, but rarely in the adjacent region of Veneto (Brichetti, 1979). More recent observations appear consistent with this pattern. For instance, in the Gulf of Trieste, Brichetti and Fracasso (2018) reported hundreds of yelkouan shearwaters between June and October, with up to 1000 individuals observed in September (though the authors failed to specify the year of observation). Conversely, in Veneto, ornithological reports between 2018 and 2022 included only one sighting of three individuals in 2018, two sightings of three and 15 individuals in 2020, and one sighting of eight individuals in 2022 (Sighele *et al.,* 2019, 2020, 2021, 2022, 2023).

Individuals observed along the coasts of Friuli-Venezia Giulia and Veneto were inferred to come from colonies in the north-eastern Adriatic Sea, along the coast of Dalmatia and in the Kvarner Gulf, Croatia, where the species was reported to occur year-round, tolerating seasonal changes and generally completing the reproductive cycle (Brichetti, 1979). More recent studies confirmed the presence of insular breeding sites in Croatia (Stipčević and Lukač, 2001; Crnković, 2005; Tutiš *et al.,* 2013; Zec *et al*., 2023), but the presumption of breeding colonies in the Kvarner Gulf could not be confirmed (Stipčević and Lukač, 2001). The Special Protected Area of Lastovo archipelago is thought to represent the main breeding area in Croatia, with 250–300 nesting pairs out of 300–400 estimated nationally (Tutiš *et al.,* 2013). Another known breeding area, which includes the islands of Vis, Palagruža, Biševo, Svetac, Brusnik, Kamik and Jabuka, lies in the Special Protected Area Pučinski otoci (Zec *et al.,* 2023). The coast of Slovenia and the Gulf of Trieste were considered important feeding areas, where the species was considered a "very abundant summer visitor". Large flocks between 500 and 1000 individuals were reported from the Gulf of Trieste, Kvarner region and northern Dalmatia (Stipčević and Lukač, 2001; Vrezec, 2006). However, in recent years the species was considered "rare" in Slovenia, with only one sighting per year between 2015 and 2018 (Hanžel, 2016, 2017; Hanžel and Denac, 2018; Denac *et al.,* 2019). These observations were limited to four separate flocks: 43 individuals in October 2015 (Hanžel, 2016), 257 individuals in June 2016 (Hanžel, 2017), 75 individuals in June 2017 (Hanžel and Denac, 2018), and 78 individuals in June 2018 (Denac *et al.,* 2019).

Here, we provide quantitative information on the occurrence of yelkouan shearwaters within a north-western Adriatic Sea area situated off the region of Veneto, Italy, based on five years of visual surveys conducted at sea. We also examine environmental variables likely to drive the species' distribution in offshore waters.

METHODS

Study area

The study area (Fig. 2) is situated off the region of Veneto, Italy, in the north-western Adriatic Sea. It includes waters up to approximately 30 km from the coast, encompassing approximately 3000 km^2 of sea surface. In this area, waters reach a maximum depth of 33 m, with a gradual topographic slope and a strong runoff from large rivers (Russo and Artegiani, 1996) resulting in high primary production (Zavatarelli *et al.,* 1998). Of several rivers flowing into this area, the most ecologically important is the nutrient-rich river Po (Europe's fifth largest river), which contributes one third of the total riverine freshwater input in the Adriatic Sea (Raicich, 1996).

The Adriatic Sea is one of the Mediterranean areas most exposed to cumulative human stressors, resulting in biodiversity decline and poor ecosystem health (Fortibuoni *et al*., 2017; Ramírez *et al*., 2018). This sea is one of the most intensively fished areas, globally (Eigaard *et al*., 2017; Amoroso *et al*., 2018; Russo *et al*., 2019; FAO, 2020). For instance, a recent assessment on the effects of bottom trawling in 24 areas around the world described the Adriatic Sea as the most intensively trawled, and the one with the worst seabed status (Pitcher *et al*., 2022). Fisheries have long been an important driver of ecosystem change (Coll *et al*., 2009; Fortibuoni *et al*., 2010; Lotze *et al.,* 2011; Fortibuoni *et al*., 2017) as well as a potential driver of demographic, ecological, and/or behavioural change for seabirds (e.g. Oro and Ruiz, 1997; Karris *et al.,* 2018). Largely due to destructive and excessive fishing impacts, today's Adriatic ecosystems are trophically simpler and less resilient than a few decades ago (Sguotti *et al*., 2022). The relative composition of species has changed, several marine communities have suffered sharp declines (Eigaard *et al*., 2017; Amoroso *et al*., 2018; Russo *et al*., 2019), and some long-lived vulnerable species have virtually disappeared. For instance, elasmobranchs declined by more than 94% across 60 years (Ferretti *et al*., 2013), whereas common dolphins *Delphinus delphis*, historically abundant in this area, have virtually disappeared (Bearzi *et al.,* 2004, 2024). However, the historic and ongoing effects of human activities and ecosystem changes on the Adriatic Sea population of yelkouan shearwaters are unknown.

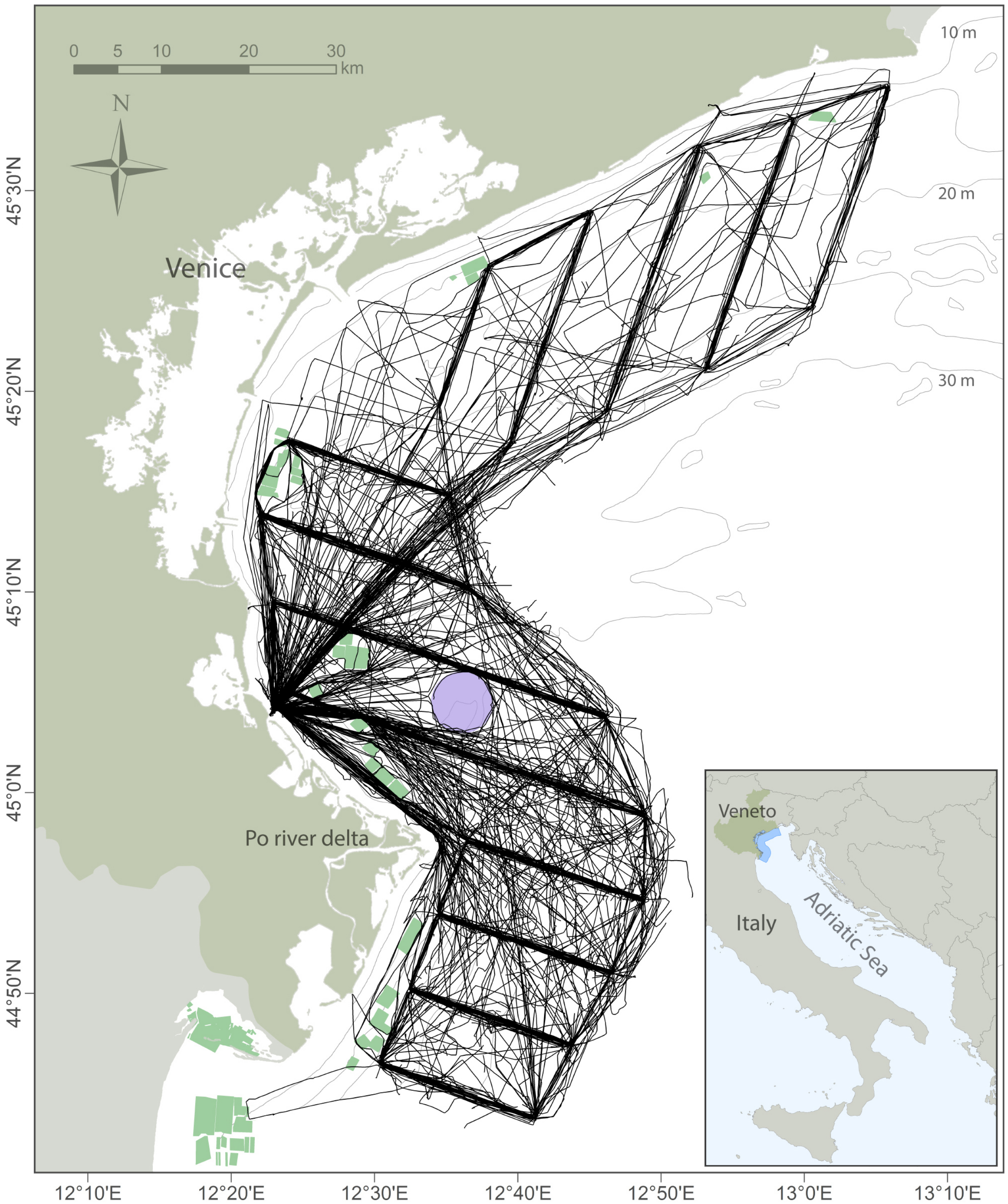

Fig. 2. Overall survey effort in years 2018–2022, totalling 169 days at sea and 23,836 km of navigation. The map also shows the bathymetry, the position of mussel farms (green polygons) and the position of a large gas terminal where navigation is not allowed (purple circle). The inset shows the position of the study area (blue polygon) relative to the Adriatic Sea.

Boat-based surveys

All data were collected in the context of a marine megafauna monitoring project, based on visual observations from a dedicated 6 m long boat equipped with a 150 HP outboard engine. Surveys were conducted between April and October 2018–2022 (Table 1), totalling 169 days at sea and 23,836 km of navigation (Fig. 2). Navigation was carried out at speeds of 28–35 km h^{-1}, in daylight with no fog or rain, with sea state $\leq$ 2 (Douglas sea scale), and with very low to no swell. Depending on weather and light conditions, surveys started in the morning (as early as 05:00) and ended in the afternoon/evening (as late as 21:00). A total of 13 survey line transects, covered monthly, helped obtain a more homogeneous coverage of the study area. The

Table 1. Number of surveys at sea, hours of survey effort, and encounters with yelkouan shearwaters by year and month. The table also reports annual and monthly encounter rates (encounters per survey hour; see the Methods section).

Year	April			May			June			July			August			September			October			Total			
	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Survey days	Survey hours	Encounters	Encounter rate
2018	2	15	0	5	22	0	6	38	0	5	29	3	7	36	3	5	30	7	2	13	18	32	184	31	0.17
2019	3	17	0	4	24	1	5	31	1	5	31	2	5	33	6	5	31	7	4	27	35	31	193	52	0.27
2020	0	0	0	0	0	0	6	37	5	6	38	3	4	29	12	7	39	2	4	22	1	27	166	23	0.14
2021	0	0	0	3	20	1	6	42	7	5	36	5	5	41	43	6	45	1	3	21	1	28	204	58	0.28
2022	0	0	0	3	21	6	3	18	9	3	16	16	9	56	21	6	33	1	10	67	21	34	211	74	0.35
Total	5	32	0	15	87	8	26	166	22	24	150	29	30	195	85	29	178	18	23	150	76	152	958	238	0.25
Encounter rate	0.00			0.09			0.13			0.19			0.44			0.10			0.51			0.25			

boat's position was recorded automatically via GPS at 1-min intervals, throughout the surveys.

The presence of yelkouan shearwaters was assessed visually by two experienced observers throughout navigation. Recorded data included date, time, GPS position, bird activity when first seen (either "flying" or "sitting on water") and number of individuals (based on visual counts by the two researchers). When yelkouan shearwaters were sighted, boat speed and route were not altered and navigation continued, resulting in instantaneous sampling that reduced the chance of re-sampling the same birds.

All data were analysed using ArcMap (ESRI, Redlands, CA, USA). Data collected under non-standard navigation conditions were removed from the analyses, and only days with a minimum survey effort of 2.5 hours were considered, to reduce excessive sampling of waters near the home port. Such stratification resulted in a survey effort of 152 days, 22,379 km and 958 hours of observation, used for the analyses.

Data for statistical modelling

Data considered in the models were collected between May and October, with April removed from the analyses due to insufficient survey effort (Table 1). Information from other studies indicated that May corresponded to the hatching period, June–August to fledging, and October to returning to insular breeding sites (Bourgeois and Vidal, 2008; Bourgeois *et al.,* 2011; Oppel *et al.,* 2011).

The study area was divided into a frame of 4x4 km grid cells (a resolution consistent with remote sensing data). To avoid potential bias due to outliers, the following cells were removed from the analyses: cell portions along the coastline with surface areas of less than 6 km^2, any cell with less than 6 km of total navigation (i.e. less than a cell's diagonal), and cells adjacent to the port where the research boat was based. To account for a different probability of encountering yelkouan shearwaters depending on survey effort (Buckland *et al.,* 2001), a sampling effort index was calculated based on the number of 1-min GPS points of effort within each grid cell, divided by the sea area in that cell (following Bonizzoni *et al.,* 2023). All GPS points were linked with information obtained 1) at sea (sea state, presence/absence and number of yelkouan shearwaters), 2) calculated within ArcMap (minimum distance from coast), or 3) from online datasets (bottom depth, sea surface temperature, chlorophyll *a*, salinity). Bottom depth was obtained from EMODNET (emodnet-hydrography.eu). Sea surface temperature (SST, °C) data were obtained from Physical Oceanography Distributed Active Archive Center (podaac.jpl.nasa.gov), with a temporal resolution of one day and a spatial resolution of 0.01 x 0.01 (latitude x longitude) degrees. Chlorophyll *a* (Chl a, expressed as mg/m^3) data were obtained from NASA OceanColor (oceancolor.gsfc.nasa.gov) as monthly-averaged MODIS-SMI products at a spatial resolution of 4 x 4 km. Salinity data were obtained from Copernicus Marine Environmental Monitoring Service (resources.marine.copernicus.eu) as monthly averages (S, PSU) at a spatial resolution of 0.042 x 0.042 degrees.

Statistical methods

The model was selected by using an Explainable Boosting Machine (EBM; Nori *et al.,* 2019), a new type of machine learning technique. Explainable Boosting Machines are based on generalized additive models (GAMs), but are specifically designed to produce interpretable models for high dimensional datasets. While the dataset involved in this paper is not high dimensional, an EBM still outperformed more classical techniques such as random forests. An EBM constructs GAMs of the form:

$$g(E[y]) = \sum f_j(x_j) + \sum f_{ij}(x_i, x_j)$$

where x_i and x_j are input variables and $g(.)$ is a link function. The interpretability of the model comes from the fact that each f_j can be plotted as a function of x_j, and each f_{ij} can be plotted as a heat map based on (x_i, x_j).

The functions of the form $f_j(x_j)$ are defined using *gradient boosting* (Friedman, 2001) of models defined as a small forest of simple regression trees (Lou *et al.,* 2012). EBMs also incorporate an algorithm called FAST (Lou *et al.,* 2013) that efficiently determines a relatively small set of interactions $f_{ij}(x_i, x_j)$ of input variables that are best to include in the overall additive model. The chosen interactions are added to the set of inputs for gradient boosting, and the small forests for those interactions have trees with either one or two inputs used. One result of FAST is that model outputs are defined on partitions of the data, rather than at individual points.

EBMs combine the results of multiple GAM models using *bagging* (Breiman, 1996). Thus, they take bootstrap samples of the data, fit a gradient boosted model as described above to each, and return the mean values across all bootstrap samples. For each input variable, the EBM produces a plot of the function $f_j(x_j)$. For a binary presence/absence classification model such as the one presented here, the plots use a score of 0 to represent the overall log odds of presence. Thus, scores above 0 represent the additional log odds of presence, while scores below 0 represent the decreased log odds of presence. Error bounds of ± one standard deviation from the bagged models are added. For the interaction models $f_{ij}(x_i, x_j)$, the EBM produces a heat map of the scores.

The EBM also returns a feature importance graph to help understand which inputs were most important in the overall model. The inputs are ranked in order of the weighted mean absolute scores from these plots; the inputs with the largest average scores have the highest influence on the predicted values.

For this study, EBM models were fit using 5-fold cross validation; more specifically, an EBM was fit on 80% of the data (the training set) and the resulting model was evaluated on the remaining 20% (the test set). The entire dataset was partitioned randomly into five disjoint subset sets; each subset was used as a test set, with the remaining 80% as the training set. The results of the EBMs on the five test sets generated this way were combined to form a model for the entire dataset.

Model quality was evaluated using two measures: area under the receiver operating characteristic curve (AUC; Fielding and Bell, 1997) and the True Skill Statistic (TSS; Allouche *et al.,* 2006). Hosmer *et al.* (2013) suggested that an AUC above 0.8 is “excellent”. While the TSS is widely used in ecological modelling, there does not appear to be a generally accepted scale for it. As a result, a random forest model (Breiman, 2001) was also fit, allowing the comparison of the EBM model with a standard, widely used model type. To improve the performance of the random forest model, data augmentation was performed before fitting the model using the R package ROSE (Lunardon *et al*., 2014) to create a dataset roughly balanced between presences and absences. Such augmentation did not improve EBM results. To verify that the EBM model was not overfitting, random inputs were included in the dataset. If these random inputs were deemed to be important by the EBM, that would be a signal that the models were overfit.

EBMs are implemented in the open-source python interpret package (github.com/interpretml), and are designed to be used with jupyter notebooks (jupyter.org). EBMs for this study were fitted with the default settings, except that “outer_bags” and “inner_bags” were set to 25, based on the recommendation of the program developers.

Model output was exported to R (R Core Team, 2023) for additional computation and generation of graphs. Graphs were created using the ggplot2 package (Wickham, 2016), and calculations used the AUC package (Ballings and Van den Poel, 2022). Random forests were generated using the package randomForest (Liaw and Wiener, 2002).

RESULTS

Encounters

Yelkouan shearwaters were encountered on 66 (43.4%) of the 152 days considered, between 05:25 and 19:08, for a total of 238 sightings and 916 individuals. The birds were observed in all months during the study period, except in April, probably due to low effort in this month. Table 1 shows the temporal distribution of survey effort by year and by month, as well as the number of yelkouan shearwater sightings. Fig. 3 shows sighting locations by activity (“flying” or “sitting on water”) and flock size ranges.

When considering the number of daily sightings by year, and the daily presence/absence by year, Shapiro-Wilk normality tests highlighted a non-normal distribution of the data (daily sightings: $W = 0.644$, $p < 0.001$; daily presence/absence: $W = 0.825$, $p < 0.001$). Kruskal-Wallis tests suggested no significant differences in the number of daily sightings across the five years of the study ($H = 8.388$, $df = 4$, $p = 0.08$), nor in the daily presence across the five years of the study ($H = 9.176$, $df = 4$, $p = 0.06$)

When yelkouan shearwaters were sighted, the number of individuals ranged between 1 and 100 (Fig. 3), but 95% ($n = 227$) of the sightings had less than 10 individuals. The activity was recorded as “flying” in 75% of the cases ($n = 178$), and “sitting on water” in 25% (n

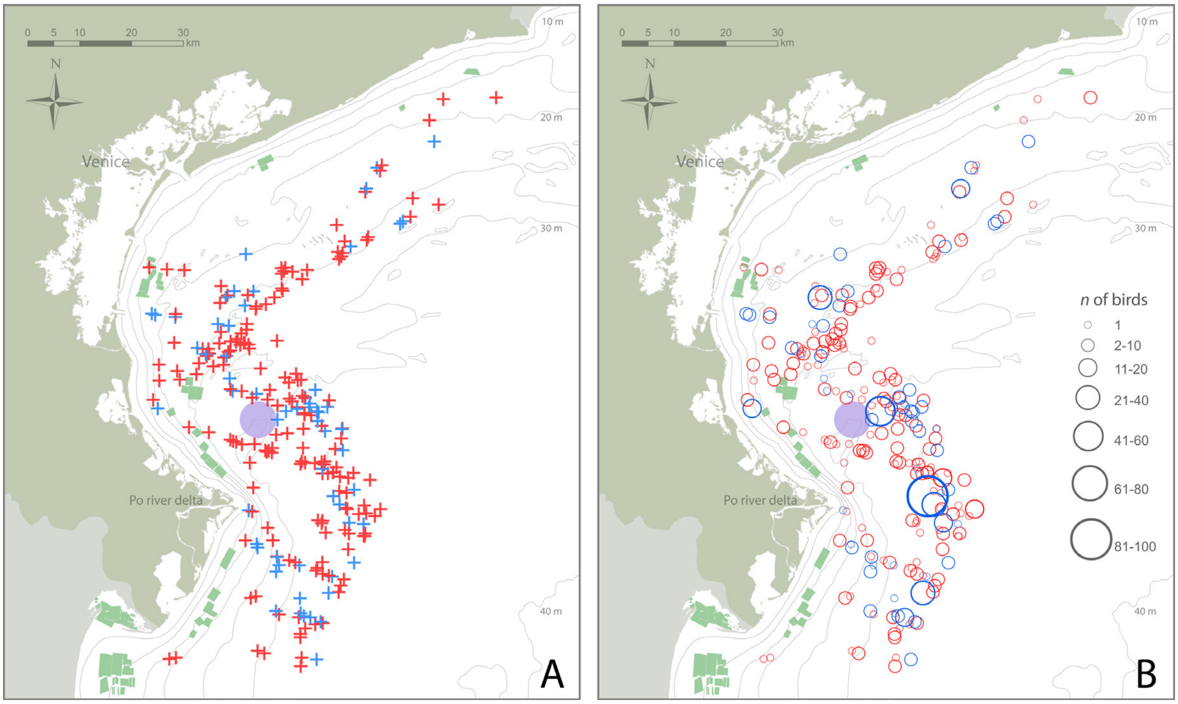

Fig. 3. Position of 238 sightings of yelkouan shearwaters in 2018–2022 (**A**), and number of birds observed in each sighting (**B**). Bird activity recorded as "flying" is indicated in red, "sitting on water" in blue.

= 60). Yelkouan shearwaters were observed in waters 9–33 m deep, between 2 and 24 km from the coast.

Comparison of EBM and random forest results

Ten different EBM models were fit using 5-fold cross-validation, with consistent results. The average AUC was 0.819, with a standard deviation of 0.005. The True Skill Statistic was maximized at 0.488, resulting from a sensitivity (true presence rate) of 0.684 and a specificity (true absence rate) of 0.804. Similarly, ten different random forest models were fit using 5-fold cross-validation, again with consistent results. The average AUC was 0.784, with a standard deviation of 0.007. The True Skill Statistic was maximized at 0.428, resulting from a sensitivity (true presence rate) of 0.680 and a specificity (true absence rate) of 0.748.

Effect on EBM of including random predictors

When random predictors were included in the cross-validated EBM models, they were always ranked below all main factors in importance: an indication that the EBMs were unlikely to be overfitting the data. Further, interaction terms always or nearly always ranked below the random predictors in importance, indicating that they were of little value to the model.

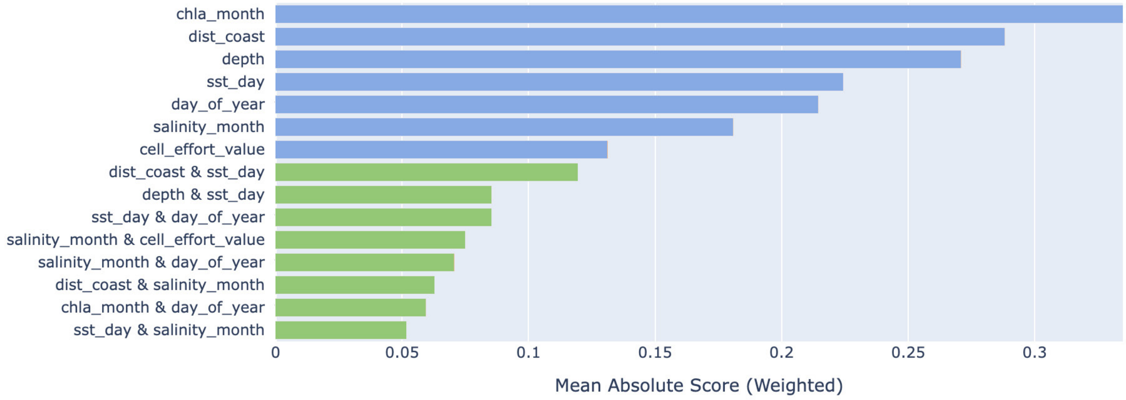

Fig. 4. Barplot of the EBM variable importance measure for the yelkouan shearwater data: chlorophyll *a* (chla_month), distance from coast (dist_coast), bottom depth (depth), sea surface temperature (sst_day), day of the year (day_of_year), salinity (salinity_month), sampling effort index (cell_effort_value). The first seven bars, in blue, represent the main effects of single variables; the remainder, in green, are interactions between variables. The horizontal length for each main effects bar is the weighted mean of the score for that variable. The interaction bars are calculated similarly for two-way plots.

Importance of variables

The train/test model (Fig. 4) shows the ranking of variable importance, including two-way interactions: the longer the horizontal bar, the more important the explanatory variable. The results were consistent from one run to the next, as well as for the whole dataset. Chlorophyll *a* (chla_month) was the most important variable selected by the model, followed by distance from coast (dist_coast), bottom depth (depth), sea surface temperature (sst_day), day of the year (day_of_year), salinity (salinity_month) and sampling effort index (cell_effort_value). The AUC for the overall model varied randomly from run to run, but centred at about 0.82.

Fig. 5 shows the effects of explanatory variables on the log-odds of finding yelkouan shearwaters, together with density plots of data distribution. The graphs are indicative of higher chances of finding yelkouan shearwaters in waters with chlorophyll *a* values of less than approximately 2.3 mg/m^3 (Fig. 5a), farther than approximately 15 km from the coast (Fig. 5b), and deeper than approximately 22 m (Fig. 5c). The effects of SST, day of the year, and salinity (Fig. 5d, e, f) were less clear. Areas with low survey effort (index less than approximately 12) had low occurrence of yelkouan shearwaters, whereas an intermediate survey effort resulted in higher yelkouan shearwater occurrence; however, the curve flattened at effort indices above 60 (Fig. 5g).

DISCUSSION

Boat-based visual surveys of seabirds have limitations (such as limited spatial and temporal coverage) and may be affected by bias (such as bird attraction to, or avoidance of, survey vessels). However, the importance of boat-based surveys to monitor yelkouan shearwaters in the Adriatic Sea was emphasized by several authors, e.g. as a means to identify at-sea distribution hotspots and shed light on the distribution of wintering or non-breeding birds (Brichetti, 1979; Vrezec, 2006; Bourgeois and Vidal, 2008; Derhé, 2012).

In the present Adriatic Sea study, the proportion of sightings beyond approximately 15 km from shore, and EBM model results highlighting a preference for deep offshore waters, attest to the importance of data collection from boats. Despite the relatively low number of encounters with yelkouan shearwaters (238 sightings, 916 birds), as compared to approximately 24,000 km of survey effort, the present study suggests that monitoring this species at sea is feasible, and can help us understand its offshore diurnal distribution and the factors affecting it. Offshore studies of this type may be particularly cost-effective when conducted in the context of wider monitoring efforts.

In the study area, the EBM model selected chlorophyll *a* as the most important variable to explain the species' occurrence, followed by distance from coast, and bottom depth. As the study area represents only a portion of yelkouan shearwaters' habitat in the Adriatic Sea, these preferences cannot be generalized. However, information provided by this study can be compared with information obtained in other areas. For instance, in the North Aegean Sea, Greece, Zakkak *et al.* (2013) estimated the abundance of marine birds from vessels, and suggested that the density of yelkouan shearwaters was linked to chlorophyll concentration and distance from shore, as well as latitude and longitude, without providing information on how the density of birds varied accordingly to these variables. Telemetry data from yelkouan shearwaters tagged in Tavolara Island (Sardinia, Italy), indicated a preference for productive coastal waters within the continental shelf. Specifically, during the incubation period, birds avoided areas farther than 400 km from the colony, as well as waters deeper than 1000 m; during chick rearing, birds could be found in areas farther from the colony (Pezzo *et al*., 2021). Aerial surveys in the north-western Mediterranean Sea indicated that "small-sized shearwaters" (including Balearic and yelkouan shearwaters) preferred coastal waters with high chlorophyll *a* concentration (Lambert *et al.,* 2017). To our knowledge, there are no other studies that describe the factors affecting the offshore distribution of yelkouan shearwaters. However, additional valuable information comes from studies of other species within the genus *Puffinus*. For instance, aerial surveys off the coast of Portugal (Araújo *et al.,* 2017) suggested that the distribution of Balearic shearwaters *P. mauretanicus* was best predicted by chlorophyll concentration, with other important variables including bathymetry, sea surface temperature and bottom slope, depending on the year. More specifically, occurrence peaked with 1) chlorophyll concentrations around 2.5 mg/m^3 (but higher values did not always imply higher occurrence), 2) bottom depths between 5 and 79 m, 3) sea surface temperatures between 15.9°C and 19.3°C, and 4) smooth bottom slopes (Araújo *et al.,* 2017). Aerial surveys in the Bay of Biscay and English Channel indicated that Balearic shearwaters and Manx shearwaters *P. puffinus* preferred coastal and shelf areas, while great shearwaters *P. gravis* and sooty shearwaters *P. griseus* preferred oceanic and continental slope areas (Pettex *et al.,* 2017). In the western Mediterranean, chlorophyll *a* was an important variable affecting the distribution of Balearic shearwaters (Louzao *et al.,* 2006, 2012). Vessel-based survey data indicated that Balearic shearwater occurrence was linked to areas with high chlorophyll *a* concentration, close to the coast and close to breeding colonies (Louzao *et al.,* 2006). Satellite transmitter data from six tagged adult Balearic shearwaters showed that bottom depth was the best predictor of distribution, followed by distance to colony, chlorophyll *a* concentration, and spatial gradients of chlorophyll *a*; presence probability increased with decreasing bottom depth, at 200 km from the colony, with lower chlorophyll *a* concentration, and at 20% and 1% of spatial gradients of chlorophyll *a* (Louzao *et al.,* 2012). Occurrence was higher in waters with chlorophyll *a* less than ~1 mg/m^3 (see Fig. 4 in Louzao *et al.,* 2012). In the Gulf of Maine, tagging data from 66 great shearwaters indicated correlations between foraging habitat and chlorophyll *a* (as well as depth and SST), but these relationships were not

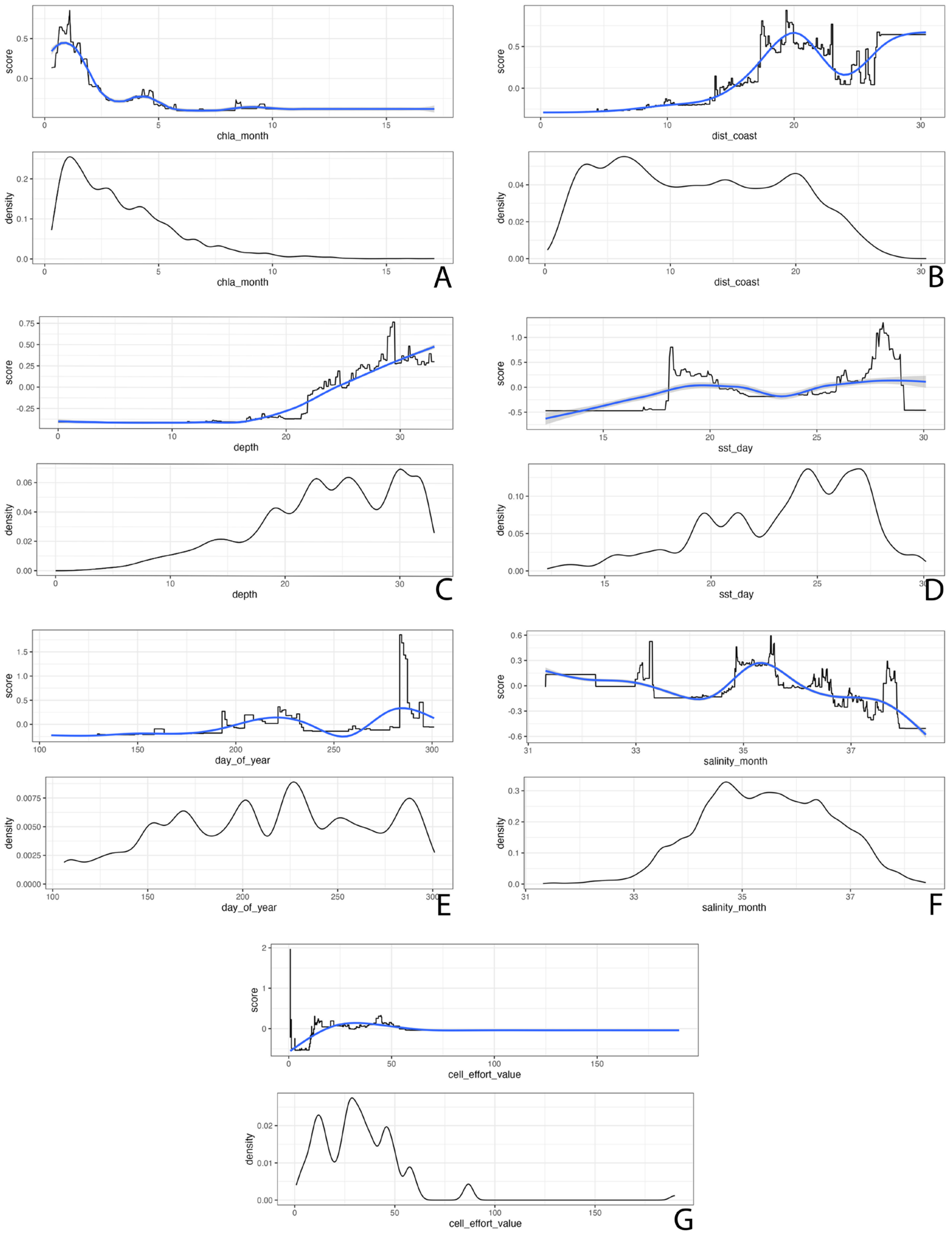

Fig. 5. Relationship between the explanatory variables and occurrence of yelkouan shearwaters: chlorophyll *a* (**A**), distance from the coast (**B**), bottom depth (**C**), sea surface temperature (**D**), day of the year (**E**), salinity (**F**), and sampling effort index (**G**). Plots for (**A**)-(**G**) top graphs represent the effect of each main variable on the prediction of the presence/absence of shearwaters. The vertical scale is log-odds of presence. The histogram lines are the direct EBM output; the smooth blue curves were added to clarify trends. Plots for (**A**)-(**G**) bottom graphs are density plots for each variable.

consistent across birds from different tagging sites, suggesting a flexible foraging strategy based on local habitat conditions and high mobility (Powers *et al.,* 2017).

An ability to either seek or avoid areas with high chlorophyll *a* concentration implies that yelkouan shearwaters can discriminate among areas of different productivity. Yelkouan shearwaters are Procellariiformes, a taxon equipped with a well-developed olfactory system with the ability to detect dimethyl sulfide (DMS) – an organosulfur compound often associated with high phytoplankton concentration and zooplankton feeding (Nevitt, 2008; Dell'Ariccia *et al.,* 2014). The sensitivity and attraction to DMS of Procellariiformes is not restricted to Antarctic and oceanic waters, but was also observed in the Mediterranean Sea (Dell'Ariccia *et al.,* 2014). While no studies of this type were conducted specifically on yelkouan shearwaters, these birds may be "early detectors" responding to DMS cues to exploit small prey patches, before "late detectors" – usually more aggressive species – arrive to take over (Nevitt, 2006).

In the north-western Adriatic Sea, the apparent preference for areas with lower chlorophyll *a* concentration contrasts with what is generally observed in other shearwater species, i.e. a preference for areas with *higher* chlorophyll *a* concentration (though such preference may vary across subareas; Powers *et al.,* 2017). One hypothesis may be that yelkouan shearwaters are attracted by productive waters in seas that are relatively oligotrophic, but avoid the most productive waters in highly eutrophic seas. Productivity in the coastal waters off the region of Veneto is extremely high, due to massive runoff of nutrients from large rivers (Russo and Artegiani, 1996; McKinney, 2007), which also results in highly murky and turbid river plumes that extend several kilometres offshore. In the study area, yelkouan shearwaters might simply tend to avoid the most eutrophic waters, and forage preferentially away from river plumes.

Telemetry data of yelkouan shearwaters (chick-rearing adults) tagged in Croatian breeding sites (Zec *et al.,* 2023) suggested that individuals in the north-western Adriatic Sea came from the Lastovo archipelago and Pučinski otoci – areas located in the eastern-central Adriatic and characterized by lower chlorophyll *a* concentrations (Barale *et al*., 2005; Mélin *et al*., 2011). Combining this information with evidence provided in the present study, it may be speculated that yelkouan shearwaters from breeding sites in the eastern Adriatic forage in productive north-western waters, while possibly avoiding the most eutrophic and murky waters off the coast of Veneto.

In addition to chlorophyll *a*, the EBM model in this study selected distance from coast and bottom depth among the most significant explanatory variables, and indicated a preference for deeper waters away from the coast. This information may contribute to the identification of suitable areas for the protection of this vulnerable species.

The occurrence and distribution of yelkouan shearwaters might also be affected by fisheries, as observed for Balearic shearwaters in the western Mediterranean Sea (Arcos and Oro, 2002). Balearic shearwaters follow operating trawlers and obtain up to 40% of their energetic requirements by feeding on discards – primarily clupeoids (European anchovy *Engraulis encrasicolus* and European pilchard *Sardina pilchardus*), gadoids, and flatfish (Arcos and Oro, 2002). In our study in the north-western Adriatic Sea, the potential effect of fisheries (particularly the large fleet of trawlers operating in this area; OSEPA, 2022) was not specifically investigated, largely due to inappropriate sample size. However, the hypothesis of yelkouan shearwater's spatio-temporal distribution being influenced at least in part by trawl fisheries (e.g. because of a potential reliance on scavenging or opportunistic foraging) deserves future investigation.

In summary, this study provides basic information on yelkouan shearwater distribution and its potential drivers in the north-western Adriatic Sea. Such information can help identify critical habitat for the species, possibly contributing to the identification and creation of new Important Bird and Biodiversity Areas (Donald *et al*., 2019), such as those proposed between Korčula island and Lastovo Archipelago, and between Po river Delta and Istria (Zec *et al.,* 2023).

ACKNOWLEDGEMENTS

This study was funded by OceanCare. Additional support was received from the Regional Environmental Protection Agency of Veneto (June–September 2018, August–September 2022), and from the University of Padua (June–September 2020). We are grateful to ESRI for supporting us with a GIS software grant. Special thanks to Mauro Bastianini for helping us in the field and in other important ways. We are grateful to Federico Fanesi for photos of yelkouan shearwaters and feedback. Two anonymous reviewers offered valuable comments that helped improve an earlier draft. We thank Nicola Baccetti and Camilla Gotti for providing important literature.

SUPPLEMENTARY MATERIALS

Supplement 1. Annotated jupyter notebook for the Explainable Boosting Machine model.

Supplement 2. Annotated R script for the random forest model.

AUTHOR CONTRIBUTIONS

Silvia Bonizzoni: Data curation; Formal analysis; Investigation; Methodology; Visualization; Writing - original draft. Blair D. Sterba-Boatwright: Formal analysis; Writing - original draft. Sarah Piwetz: Writing - review & editing. Giovanni Bearzi: Supervision; Conceptualization; Investigation; Writing - original draft; Visualization.

REFERENCES

Allouche, O., Tsoar, A., Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology, 43(6), 1223-1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Amoroso, R.O., Pitcher, C.R., Rijnsdorp, A.D., McConnaughey, R.A., Parma, A.M., Suuronen, P., Eigaard, O.R. *et al*. 2018. Bottom trawl fishing footprints on the world's continental shelves. Proceedings of the National Academy of Sciences, 115, E10275-E10282. https://doi.org/10.1073/pnas.1802379115

Araújo, H., Bastos-Santos, J., Rodrigues, P.C., Ferreira M., Pereira A., Henriques A.C., Monteiro, S.S. *et al.* 2017. The importance of Portuguese continental shelf waters to Balearic shearwaters revealed by aerial census. Marine Biology, 164, 55. https://doi.org/10.1007/s00227-017-3089-x

Arcos, J.M., Oro, D. 2002. Significance of fisheries discards for a threatened Mediterranean seabird, the Balearic shearwater *Puffinus mauretanicus*. Marine Ecology Progress Series, 239, 209-220. https://doi.org/10.3354/meps239209

Ballings, M., Van den Poel, D. 2022. AUC: threshold independent performance measures for probabilistic classifiers.

Barale, V., Schiller, C., Tacchi, R., Marechal, C. 2005. Trends and interactions of physical and bio-geo-chemical features in the Adriatic Sea as derived from satellite observations. Science of the Total Environment, 353(1-3), 68-81. https://doi.org/10.1016/j.scitotenv.2005.09.060

Bearzi, G., Holcer, D., Notarbartolo di Sciara, G. 2004. The role of historical dolphin takes and habitat degradation in shaping the present status of northern Adriatic cetaceans. Aquatic Conservation: Marine and Freshwater Ecosystems, 14(4), 363-379. https://doi.org/10.1002/aqc.626

Bearzi, G., Bonizzoni, S., Genov, T., Notarbartolo di Sciara, G. 2024. Whales and dolphins of the Adriatic Sea: present knowledge, threats and conservation. Acta Adriatica, 65(1). https://doi.org/10.32582/aa.65.1.1

BirdLife International. 2021a. *Puffinus yelkouan*. The IUCN Red List of Threatened Species 2021, e.T22698230A166330721. https://doi.org/10.2305/IUCN.UK.2021-3.RLTS.T22698230A166330721.en

BirdLife International. 2021b. *Puffinus yelkouan*. European Red List of Birds. Luxembourg: Publications Office of the European Union. https://www.iucnredlist.org/species/pdf/166330721/attachment (accessed 14 April 2024)

Bonizzoni, S., Gramolini, R., Furey, N.B., Bearzi, G. 2023. Bottlenose dolphin distribution in a Mediterranean area exposed to intensive trawling. Marine Environmental Research, 188, 105993. https://doi.org/10.1016/j.marenvres.2023.105993

Bonnaud, E., Bourgeois, K., Vidal, E., Kayser, Y., Legrand, J. 2007. Feeding ecology of a feral cat population on a small Mediterranean Island. Journal of Mammalogy 88(4), 1074-1081. https://doi.org/10.1644/06-MAMM-A-031R2.1

Borg, J.J., Sultana, J., Coleiro, C. 2002. Pre-breeding movements and early nest visits by the Levantine shearwater *Puffinus yelkouan* in Malta. II-Merill 30, 20-23. https://www.um.edu.mt/library/oar/handle/123456789/36587 (accessed 14 April 2024)

Bourgeois, K., Vidal, E. 2008. The endemic Mediterranean yelkouan shearwater *Puffinus yelkouan*: distribution, threats and a plea for more data. Oryx, 42(2), 187-194. https://doi.org/10.1017/S0030605308006467

Bourgeois, K., Vorenger, J., Faulquier, L., Legrand, J., Vidal, E. 2011. Diet and contamination of the yelkouan shearwater *Puffinus yelkouan* in the Hyères archipelago, Mediterranean basin, France. Journal of Ornithology, 152, 947-953. https://doi.org/10.1007/s10336-011-0677-2

Breiman, L. 1996. Bagging predictors. Machine Learning 24(2), 123-40. https://doi.org/10.1007/BF00058655

Breiman, L. 2001. Random Forests. Machine Learning, 45, 5-32. https://doi.org/10.1023/A:1010933404324

Brichetti, P. 1979. Distribuzione geografica degli uccelli nidificanti in Italia, Corsica e isole Maltesi. Natura Bresciana. 77 pp.

Brichetti, P., Fracasso, G. 2018. The birds of Italy: Anatidae-Alcidae. Edizioni Belvedere, Latina, 512 pp.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, New York. 432 pp. https://doi.org/10.1093/oso/9780198506492.001.0001

Coll, M., Santojanni, A., Palomera, I., Arneri, E. 2009. Food-web changes in the Adriatic Sea over the last three decades. Marine Ecology Progress Series, 381, 17-37. https://doi.org/10.3354/meps07944

Crnković, R. 2005. Gniježdenje gregule *Puffinus yelkouan* na Hrvatskoj obali Jadrana (Breeding of the yelkouan shearwater *Puffinus yelkouan* in Croatian part of Adriatic Sea). Larus, 49, 52-53.

Dell'Ariccia, G., Célérier, A., Gabirot, M., Palmas, P., Massa, B., Bonadonna, F. 2014. Olfactory foraging in temperate waters: sensitivity to dimethylsulphide of shearwaters in the Atlantic Ocean and Mediterranean Sea. Journal of Experimental Biology, 217(10), 1701-1709. https://doi.org/10.1242/jeb.097931

Denac, M., Hanžel, J., Bordjan, D., Božič, L., Vogrin, M. 2019. Redke vrste ptic v Sloveniji v letu 2018 - Poročilo Nacionalne komisije za redkosti (Rare birds in Slovenia in 2018 - Slovenian Rarities Committee). Acrocephalus, 40(182-183), 45-60. https://doi.org/10.1515/acro-2019-0012

Derhé, M. 2012. Developing a population assessment for yelkouan shearwaters *Puffinus yelkouan*. *In* Ecology and Conservation of Mediterranean Seabirds and other bird species under the Barcelona Convention (eds. P. Yésou, N. Baccetti, J. Sultana). Proceedings of the 13th Medmaravis Pan-Mediterranean Symposium. Alghero 14-17 October 2011. pp. 66-73.

Donald, P.F., Fishpool, L.D., Ajagbe, A., Bennun, L.A., Bunting, G., Burfield, I.J., Butchart S.H.M. *et al.* 2019. Important Bird and Biodiversity Areas (IBAs): the development and characteristics of a global inventory of key sites for biodiversity. Bird Conservation International, 29(2), 177-198. https://doi.org/10.1017/S0959270918000102

EC. 1979. Convention on the Conservation of European Wildlife and Natural Habitats. Council of Europe, European Treaty Series, 104, 1-10.

EC. 1999. Protocol concerning specially protected areas and biological diversity in the Mediterranean. Official Journal of the European Communities OJ L322, 3-17.

EC. 2010. Directive 2009/147/EC of the European Parliament and of the Council of 30 November 2009 on the conservation of wild birds. Official Journal of the European Union, OJ L20, 7-25.

Eigaard, O.R., Bastardie, F., Hintzen, N.T., Buhl-Mortensen, L., Buhl-Mortensen, P., Catarino, R., Dinesen, G.E. *et al*. 2017. The footprint of bottom trawling in European waters: distribution, intensity, and seabed integrity. ICES Journal of Marine Science, 74(3), 847-865. https://doi.org/10.1093/icesjms/fsw194

FAO. 2020. The state of Mediterranean and Black Sea fisheries 2020. General Fisheries Commission for the Mediterranean. FAO, Rome, 172 pp. https://doi.org/10.4060/cb2429en

Ferretti, F., Osio, G.C., Jenkins, C.J., Rosenberg, A.A., Lotze, H.K. 2013. Long-term change in a meso-predator community in response to prolonged and heterogeneous human impact. Scientific Reports 3, 1057. https://doi.org/10.1038/srep01057

Fielding, A.H, Bell J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation, 24(1), 38-49. https://doi.org/10.1017/S0376892997000088

Fortibuoni, T., Libralato, S., Raicevich, S., Giovanardi, O., Solidoro, C. 2010. Coding early naturalists' accounts into long-term fish community changes in the Adriatic Sea (1800–2000). PLoS ONE, 5, e15502. https://doi.org/10.1371/journal.pone.0015502

Fortibuoni, T., Giovanardi, O., Pranovi, F., Raicevich, S., Solidoro, C., Libralato, S. 2017. Analysis of long-term changes in a Mediterranean marine ecosystem based on fishery landings. Frontiers in Marine Science, 4, 1-16. https://doi.org/10.3389/fmars.2017.00033

Friedman, J.H. 2001. Greedy function approximation: a gradient boosting machine. The Annals of Statistics 29(5), 1189-1232.

Gaudard, C. 2018. International single species action plan for the yelkouan shearwater *Puffinus yelkouan*. Project LIFE 14 PRE/UK/000002. Coordinated Efforts for International Species Recovery EuroSAP. LPO/BirdLife France. Rochefort. 45 pp.

Genovart, M., Juste, J., Contreras-Díaz, H., Oro, D. 2012. Genetic and phenotypic differentiation between the critically endangered Balearic shearwater and neighboring colonies of its sibling species. Journal of Heredity, 103(3), 330-341. https://doi.org/10.1093/jhered/ess010

Hanžel, J. 2016. Redke vrste ptic v Sloveniji v letu 2015 - Poročilo Nacionalne komisije za redkosti (Rare birds of Slovenia in 2015 - Slovenian Rarities Committee's Report). Acrocephalus, 37(168-169), 69-78. https://doi.org/10.1515/acro-2016-0004

Hanžel, J. 2017. Redke vrste ptic v Sloveniji v letu 2016; Poročilo Nacionalne komisije za redkosti / Rare birds of Slovenia in 2016 - Slovenian Rarities Committee Report. Acrocephalus, 38(172-173), 21-30. https://doi.org/10.1515/acro-2017-0002

Hanžel, J., Denac, M. 2018. Redke vrste ptic v Sloveniji v letu 2017 - Poročilo Nacionalne komisije za redkosti (Rare birds of Slovenia in 2017 - Slovenian Rarities Committee's Report). Acrocephalus, 39(178-179), 177-184. https://doi.org/10.1515/acro-2018-0013

Hosmer, Jr., D.W., Lemeshow, S., Sturdivant, R.X. 2013. Applied logistic regression, 3rd ed. John Wiley & Sons, Hoboken. 528 pp.

Karris, G., Ketsilis-Rinis, V., Kalogeropoulou, A., Xirouchakis, S., Machias, A., Maina, I., Kavadas, S. 2018. The use of demersal trawling discards as a food source for two scavenging seabird species: a case study of an eastern Mediterranean oligotrophic marine ecosystem. Avian Research, 9, 1-14. https://doi.org/10.1186/s40657-018-0118-5

Kralj, J., Barišić, S. 2013. Rare birds in Croatia: third report of the Croatian Rarities Committee. Natura Croatica, 22(2), 375-396.

Lambert, C., Laran, S., David, L., Dorémus, G., Pettex, E., Van Canneyt, O., Ridoux, V. 2017. How does ocean seasonality drive habitat preferences of highly mobile top predators? Part I: the north-western Mediterranean Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 141, 115-132. https://doi.org/10.1016/j.dsr2.2016.06.012

Liaw, A., Wiener, M. 2002. Classification and regression by randomForest. R News, 2(3), 18-22.

Lotze, H.K., Coll, M., Dunne, J.A. 2011. Historical changes in marine resources, food-web structure and ecosystem functioning in the Adriatic Sea, Mediterranean. Ecosystems, 14, 198-222. https://doi.org/10.1007/s10021-010-9404-8

Lou, Y., Caruana, R., Gehrke, J. 2012. Intelligible models for classification and regression. *In* Proceedings of the 18th ACM SIGKDD International Conference on Knowledge discovery and data mining. pp. 150-158. https://doi.org/10.1145/2339530.2339556

Lou, Y., Caruana, R., Gehrke, J., Hooker, G. 2013. Accurate intelligible models with pairwise interactions. *In* Proceedings of the 19th ACM SIGKDD International Conference on Knowledge discovery and data mining. pp. 623-631. https://doi.org/10.1145/2487575.2487579

Louzao, M., Hyrenbach, K.D., Arcos, J.M., Abelló, P., Gil de Sola, L., Oro, D. 2006. Oceanographic habitat of a endangered Mediterranean procellariiform: implications for marine protected areas. Ecological Applications, 16(5), 1683-1695. https://doi.org/10.1890/1051-0761(2006)016[1683:OHOAEM]2.0.CO;2

Louzao, M., Delord, K., García, D., Boué, A., Weimerskirch, H. 2012. Protecting persistent dynamic oceanographic features: transboundary conservation efforts are needed for the critically endangered Balearic shearwater. PloS One, 7(5), e35728. https://doi.org/10.1371/journal.pone.0035728

Lunardon, N., Menardi, G., Torelli, N. 2014. ROSE: a package for binary imbalanced learning. R Journal, 6(1), 79-89. https://doi.org/10.32614/RJ-2014-008

McKinney, F.K. 2007. The northern Adriatic ecosystem: deep time in a shallow sea. Columbia University Press, New York. 328 pp.

Mélin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J.F., Canuti, E. 2011. Multi-sensor satellite time series of optical properties and chlorophyll-a concentration in the Adriatic Sea. Progress in Oceanography, 91(3), 229-244. https://doi.org/10.1016/j.pocean.2010.12.001

Nevitt, G.A. 2006. Olfactory foraging strategies of procellariiform seabirds. Acta Zoologica Sinica, 52, 510-513.

Nevitt, G.A. 2008. Sensory ecology on the high seas: the odor world of the procellariiform seabirds. The Journal of Experimental Biology, 211(11), 1706-1713. https://doi.org/10.1242/jeb.015412

Nori, H., Jenkins, S., Koch, P., Caruana, R. 2019. InterpretML: a unified framework for machine learning interpretability. arXiv:1909.09223v1, 8 pp. https://doi.org/10.48550/arXiv.1909.09223

Oppel, S., Raine, A.F., Borg, J.J., Raine, H., Bonnaud, E., Bourgeois, K., Breton, A.R. 2011. Is the yelkouan shearwater *Puffinus yelkouan* threatened by low adult survival probabilities? Biological Conservation, 144(9), 2255-2263. https://doi.org/10.1016/j.biocon.2011.05.017

Oro, D., Ruiz, X. 1997. Exploitation of trawler discards by breeding seabirds in the north-western Mediterranean: differences between the Ebro Delta and the Balearic Islands areas. ICES Journal of Marine Science, 54(4), 695-707. https://doi.org/10.1006/jmsc.1997.0246

OSEPA, 2022. La pesca in Veneto: flotta, imprese, produzione e commercio, anno 2021 (Fisheries in Veneto: fleet, businesses, production and trade, year 2021). Osservatorio Socio Economico della Pesca e dell'Aquacoltura-Veneto, 35 pp.

Pérez-Ortega, M., İsfendiyaroğlu, S. 2017. Predicting foraging hotspots for yelkouan shearwater in the Black Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 141, 237-247. http://doi.org/10.1016/j.dsr2.2016.07.007

Péron, C., Grémillet, D., Prudor, A., Pettex, E., Saraux, C., Soriano-Redondo, A., Authier, M. *et al*. 2013. Importance of coastal Marine Protected Areas for the conservation of pelagic seabirds: the case of Vulnerable yelkouan shearwaters in the Mediterranean Sea. Biological Conservation, 168, 210-221. https://doi.org/10.1016/j.biocon.2013.09.006

Pettex, E., Laran, S., Authier, M., Blanck, A., Dorémus, G., Falchetto, H., Lambert, C. *et al*. 2017. Using large scale surveys to investigate seasonal variations in seabird distribution and abundance. Part II: The Bay of Biscay and the English Channel. Deep Sea Research Part II: Topical Studies in Oceanography, 141, 86-101. https://doi.org/10.1016/j.dsr2.2016.11.012

Pezzo, F., Zenatello, M., Cerritelli, G., Navone, A., Giunchi, D., Spano, G., Pollonara, E. *et al*. 2021. Productivity changes in the Mediterranean Sea drive foraging movements of yelkouan shearwater *Puffinus yelkouan* from the core of its global breeding range. Marine Ecology, 42(4), e12668. https://doi.org/10.1111/maec.12668

Pitcher, C.R., Hiddink, J.G., Jennings, S., Collie, J., Parma, A.M., Amoroso, R., Mazor, T. *et al*. 2022. Trawl impacts on the relative status of biotic communities of seabed sedimentary habitats in 24 regions worldwide. Proceedings of the National Academy of Sciences, 119, e2109449119. https://doi.org/10.1073/pnas.2109449119

Powers, K.D., Wiley, D.N., Allyn, A.J., Welch, L.J., Ronconi, R.A. 2017. Movements and foraging habitats of great shearwaters *Puffinus gravis* in the Gulf of Maine. Marine Ecology Progress Series, 574, 211-226. https://doi.org/10.3354/meps12168

R Core Team. 2023. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/ (accessed 14 April 2024)

Raicich, F. 1996. On the fresh water balance of the Adriatic Sea. Journal of Marine Systems, 9, 305-319. https://doi.org/10.1016/S0924-7963(96)00042-5

Raine, A.F., Borg, J.J., Raine, H., Phillips, R.A. 2013. Migration strategies of the yelkouan shearwater *Puffinus yelkouan*. Journal of Ornithology, 154, 411-422. https://doi.org/10.1007/s10336-012-0905-4

Ramírez, F., Coll, M., Navarro, J., Bustamante, J., Green, A.J. 2018. Spatial congruence between multiple stressors in the Mediterranean Sea may reduce its resilience to climate impacts. Scientific Reports, 8(1), 1-8. https://doi.org/10.1038/s41598-018-33237-w

Ranù, M., Vanacore, A., Mandich, A., Alessi, J. 2022. Bottlenose dolphins and seabirds distribution analysis for the identification of a marine biodiversity hotspot in Agrigento waters. Journal of Marine Science and Engineering, 10(3), 345. https://doi.org/10.3390/jmse10030345

Russo, A., Artegiani, A. 1996. Adriatic Sea hydrography. Scientia Marina, 60, 33-43.

Russo, T., Franceschini, S., D'Andrea, L., Scardi, M., Parisi, A., Cataudella, S. 2019. Predicting fishing footprint of trawlers from environmental and fleet data: an application of artificial neural networks. Frontiers in Marine Science, 6, 670. https://doi.org/10.3389/fmars.2019.00670

Sguotti, C., Bischoff, A., Conversi, A., Mazzoldi, C., Möllmann, C., Barausse, A. 2022. Stable landings mask irreversible community reorganizations in an overexploited Mediterranean ecosystem. Journal of Animal Ecology, 91, 2465-2479. https://doi.org/10.1111/1365-2656.13831

Sighele, M., Verza, E., Stival, E., Cassol, M., Bon, M. 2019. Rapporto ornitologico per la regione Veneto, anno 2018 (Ornithological report for the region of Veneto, year 2018). Bollettino del Museo di Storia Naturale di Venezia 70, 45-70.

Sighele, M., Verza, E., Stival, E., Cassol, M., Bon, M. 2020. Rapporto ornitologico per la regione Veneto, anno 2019 (Ornithological report for the region of Veneto, year 2019). Bollettino del Museo di Storia Naturale di Venezia 71, 89-111.

Sighele, M., Stival, E., Verza, E., Cassol, M., Bon, M. 2021. Rapporto ornitologico per la regione Veneto, anno 2020 (Ornithological report for the region of Veneto, year 2020). Bollettino del Museo di Storia Naturale di Venezia 72, 33-58.

Sighele, M., Stival, E., Verza, E., Sartori, A., Bon, M. 2022. Rapporto ornitologico per la regione Veneto, anno 2021 (Ornithological report for the region of Veneto, year 2021). Bollettino del Museo di Storia Naturale di Venezia 73, 45-68.

Sighele, M., Stival, E., Corvino, R., Fagotto, N., Fanelli, V., Feltrin, G., Moretto, F. *et al*. 2023. Rapporto ornitologico per la regione Veneto, anno 2022 (Ornithological report for the region of Veneto, year 2022). Bollettino del Museo di Storia Naturale di Venezia, 74, 31-53.

Stipčević, M., Lukač, G. 2001. Status of tubenose seabirds Procellariiformes breeding in the eastern Adriatic. Acrocephalus, 22(104/105), 9-21.

Tranchant, A., Vidal, E., Kayser, Y. 2003. Premières données sur le régime alimentaire du chat haret *Felis catus* en situation micro-insulaire Méditerranéenne. Revue d'Ecologie, Terre et Vie, 58(4), 411-418.

Tutiš, V., Kralj, J., Radović, D., Ćiković, D., Barišić, S. 2013. Crvena Knjiga ptica Hrvatske (Red Book of Croatian birds). Ministry of Environmental and Nature Protection, State Institute for Nature Protection, Zagreb. 258 pp.

UNEP/MAP-RAC/SPA. 2003. Action Plan for the Conservation of bird species listed in Annex II of the Protocol concerning Specially Protected Areas (SPAs), and Biological Diversity in the Mediterranean. Ed. RAC/SPA, Tunis, 80 pp

UNEP/MAP-RAC/SPA. 2007. Guidelines for management and monitoring threatened populations of marine and coastal bird species and their important areas in the Mediterranean (by J. Sultana). Ed. RAC/SPA, Tunis, 10 pp.

UNEP/MAP-RAC/SPA. 2017. Action Plan for the conservation of marine and coastal bird species listed in Annex II of the Protocol concerning Specially Protected Areas and Biological Diversity in the Mediterranean. SPA/RAC, Tunis, 63 pp.

Vrezec, A. 2006. Marine and coastal birds of Slovenia: status, population size and conservation of Mediterranean action plan species. *In* Proceedings of the first symposium on the Mediterranean action plan for the conservation of marine and coastal birds (ed. N. Aransay). Vilanova i la Geltrú, 17-19 November 2005. UNEP/MAP - SPA/RAC, Tunis, pp. 81−85.

Wickham, H. 2016. Ggplot2: elegant graphics for data analysis. Springer, New York, 260 pp. https://doi.org/10.1007/978-3-319-24277-4

Zakkak, S., Panagiotopoulou, M., Halley, J.M. 2013. Estimating the abundance of sheawters and gulls in the north Aegean Sea. Marine Ornithology, 41, 141-148.

Zavatarelli, M., Raicich, F., Bregant, D., Russo, A., Artegiani, A. 1998. Climatological biogeochemical characteristics of the Adriatic Sea. Journal of Marine Systems, 18(1-3), 227-263. https://doi.org/10.1016/S0924-7963(98)00014-1

Zec, M., Engelen, D., Handley, J., Rajković, Ž. 2023. IBA identification for seabirds in Croatia: supporting designation of marine SPAs. LIFE Artina (LIFE 17 NAT/HR/000594) report for actions C5 and E1. Association Biom (BirdLife Croatia), 21 pp.