#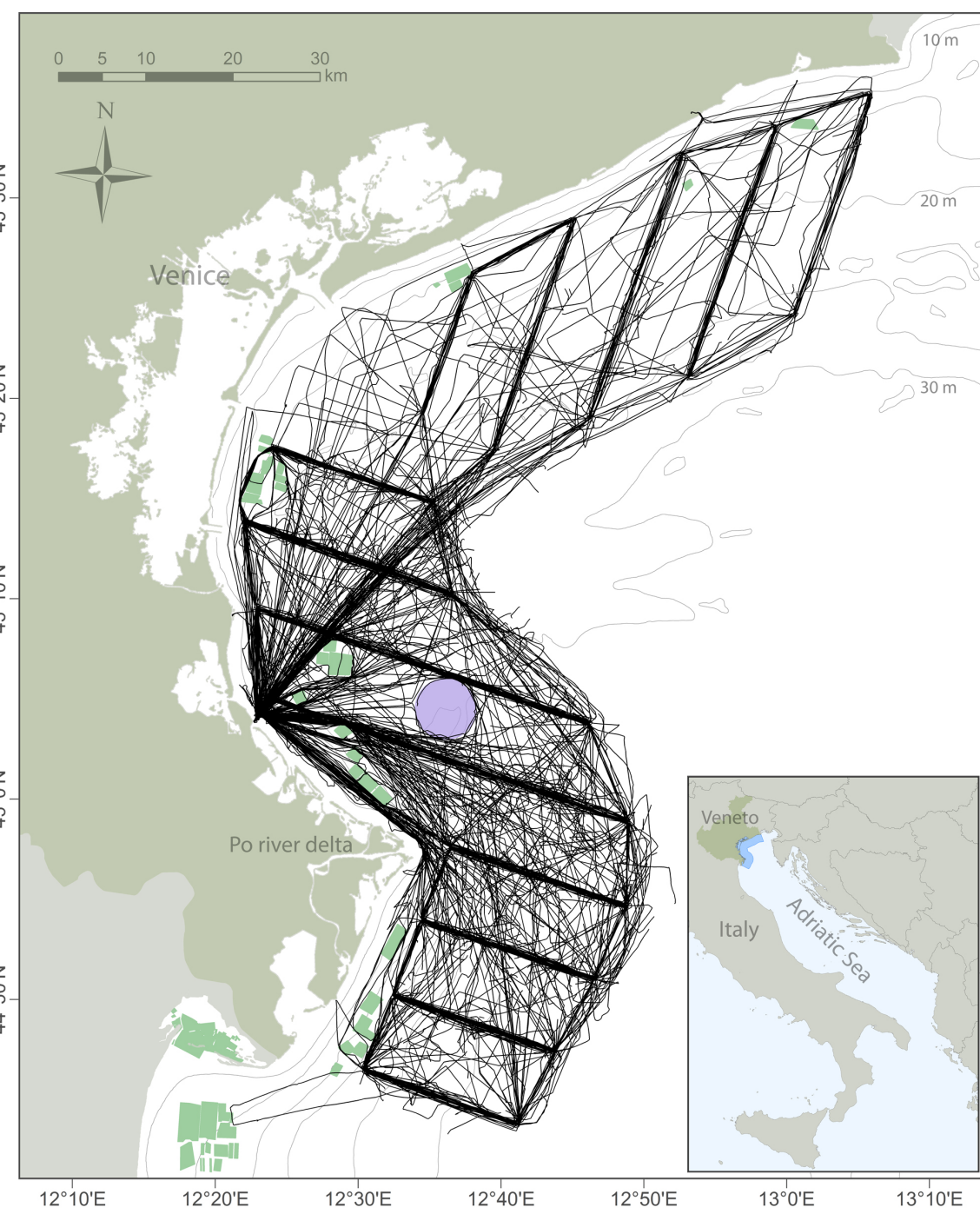 Factors affecting the offshore distribution of yelkouan shearwaters in the northwestern Adriatic Sea: insight from machine learning

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

1) Dolphin Biology and Conservation, Italy; 2) OceanCare, Switzerland; 3) Texas A&M University, USA ; 4) Institute of Marine Sciences, CNR, Italy

Survey effort in years 2018–2022, totalling 169 days at sea and 23,836 km of navigation.
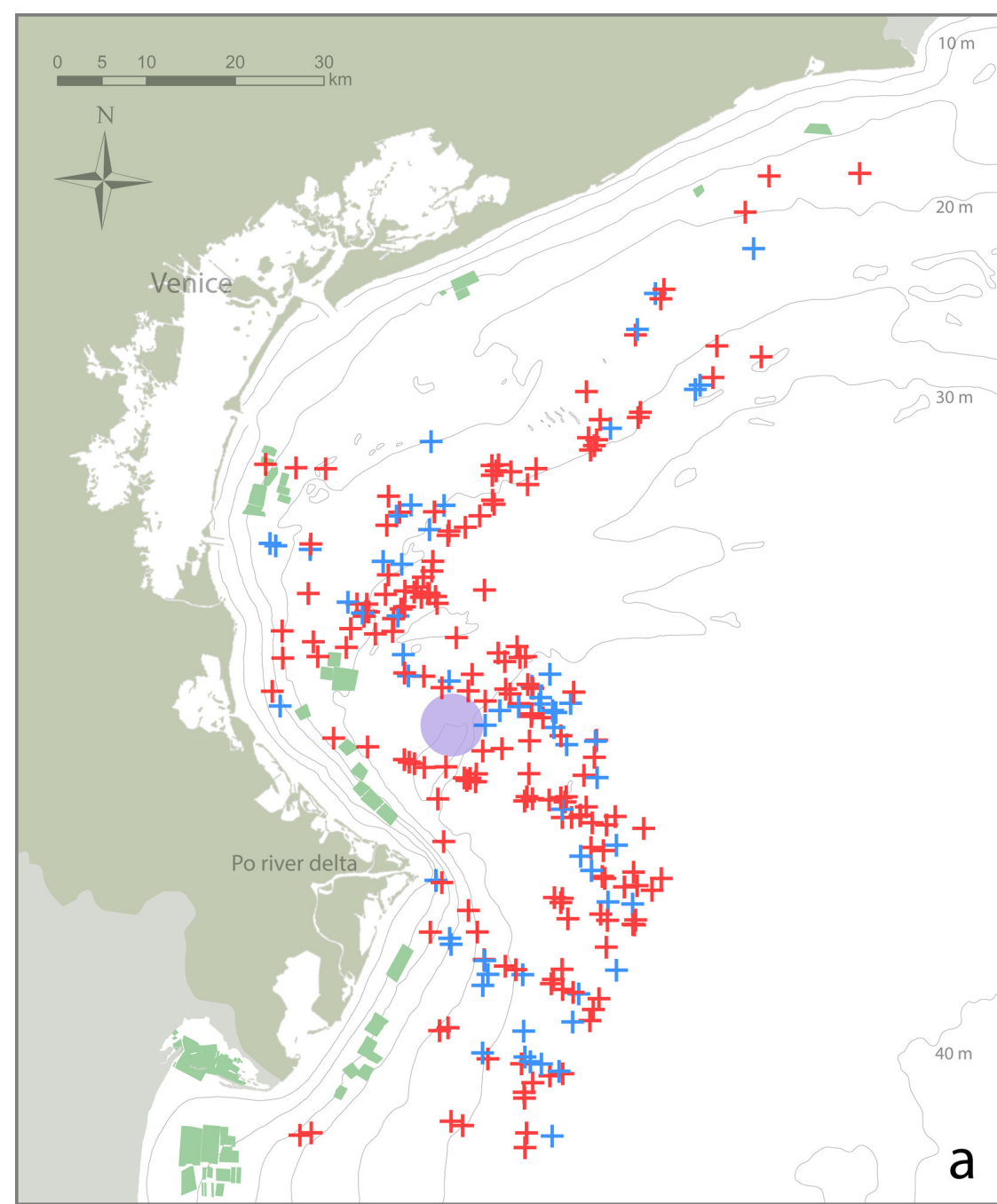
Position of 238 sightings of yelkouan shearwaters in 2018–2022 (**a**), and number of birds observed in each sighting (**b**). Bird activity recorded as "flying" is indicated in red, "sitting on water" in blue.
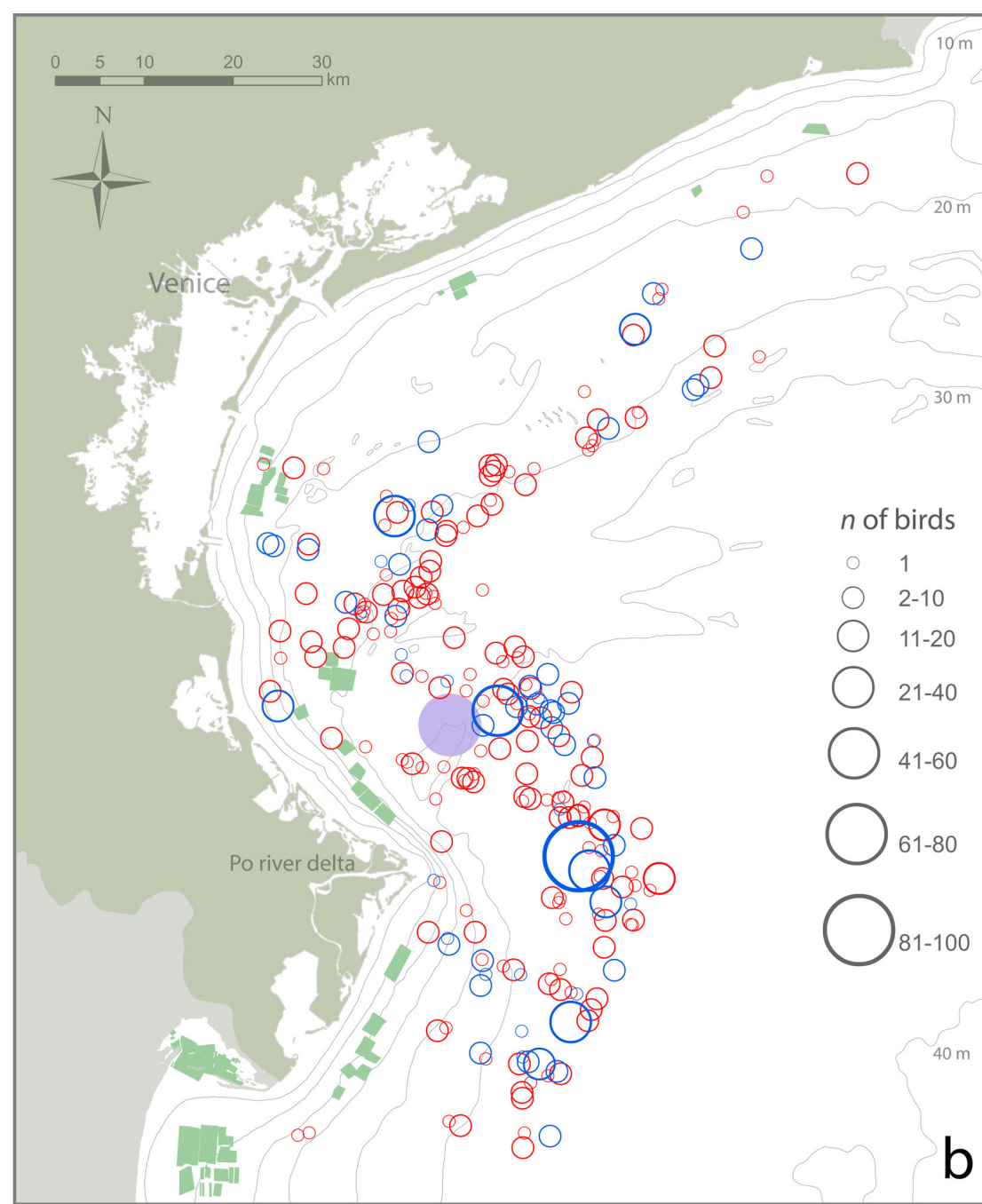
Results

Yelkouan shearwaters were observed in all months (except in April, when effort was lower). They were encountered on 66 days, for a total of 238 sightings including 916 individuals, in waters 9–33 m deep, between 2 and 24 km from the coast.

Individual counts ranged between 1 and 100 (mean 3.85, SD 8.33, mode 1), with 95% of the encounters having <10 individuals. When the birds were spotted, behaviour was recorded as "flying" (75%) or "sitting on water" (25%).

Introduction

The yelkouan shearwater *Puffinus yelkouan* is endemic to the Mediterranean and Black Seas and is classified as Vulnerable in the IUCN Red List.

Most information on its distribution and habitat use comes from telemetry tracking of single individuals and land-based studies at nesting sites, whereas few studies have been based on direct visual observations conducted offshore.

Information on the species' distribution and habitat use in the eastern Mediterranean is scant. Land-based studies suggest that the northern Adriatic Sea represents an important area for moulting, feeding and passage during the interbreeding period; however, the species' distribution at sea is unknown.

An **Explainable Boosting Machine model**——a machine learning technique based on generalized additive models, specifically designed to produce interpretable models for high dimensional datasets——**selected chlorophyll a as the most important variable** to explain the species' occurrence, **followed by distance from the coast, and bottom depth.**

EBM variable importance: chlorophyll a, distance from coast, bottom depth, sea surface temperature, day of the year, salinity, sampling effort index.
Blue: single variables.
Green: interactions between variables.

The model indicated a **higher occurrence in waters with chlorophyll <2.3 mg/m³, farther than ~15 km from the coast, and deeper than ~22 m.**

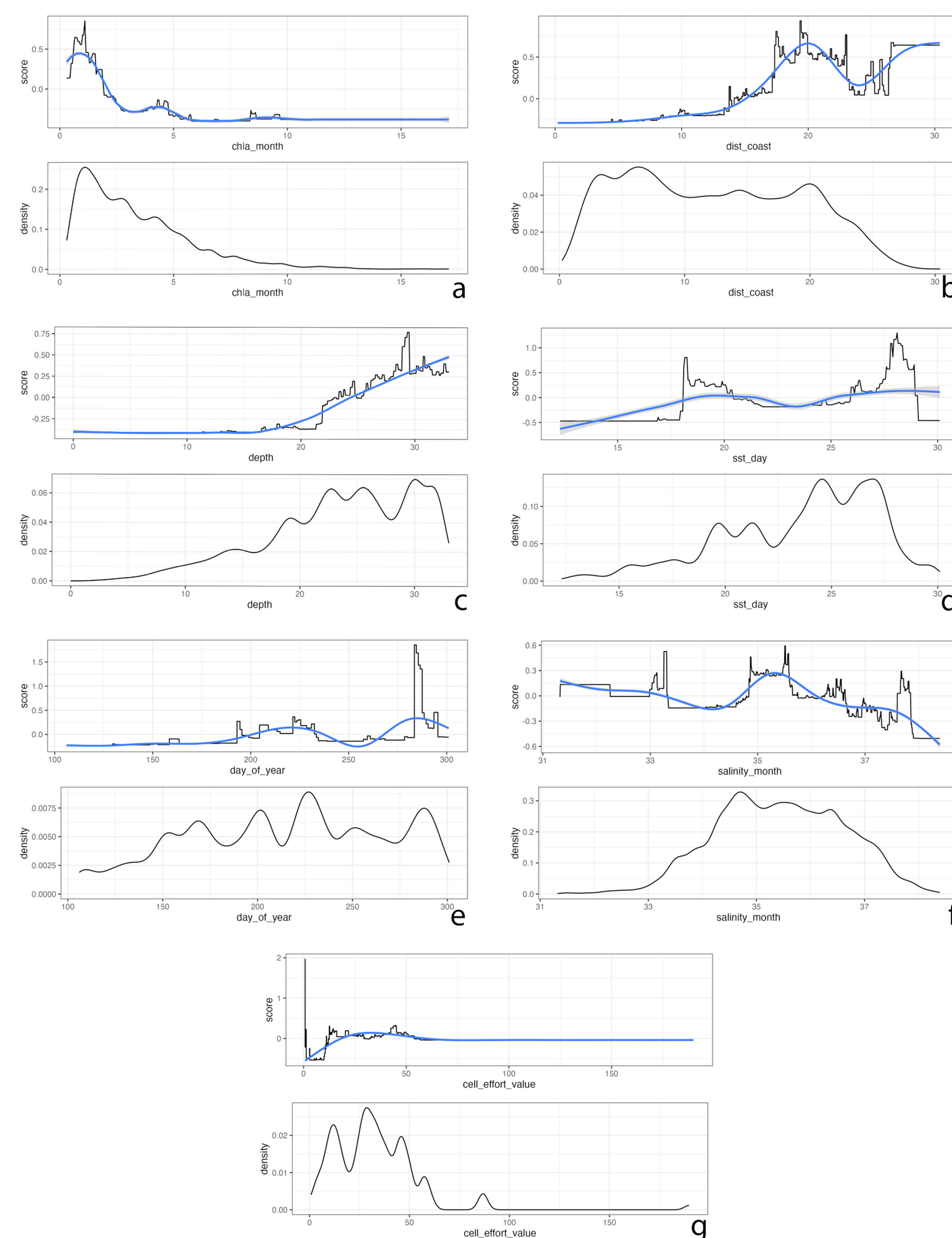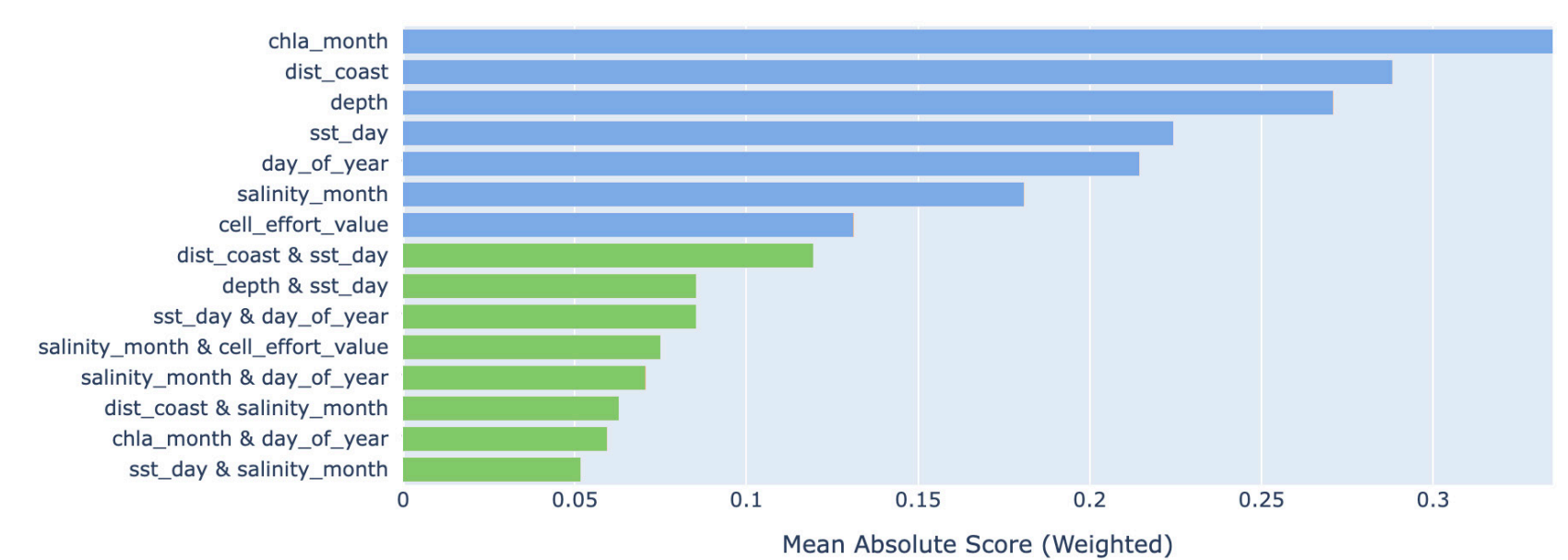The effects of sea surface temperature, salinity, and day of the year were less clear.

Methods

Based on visual surveys conducted from small boats between April and October 2018–2022 (effort: 169 days and 23,836 km), we provide information on **1)** the occurrence of yelkouan shearwaters within a study area of 3000 km² off the region of Veneto, Italy, in the northwestern Adriatic Sea; and **2)** the geographic, bathymetric and oceanographic variables likely to drive the species' offshore distribution in this area.

This study provides insight into the offshore distribution of yelkouan shearwaters, within one of the Mediterranean areas most exposed to cumulative human threats.

Relationship between explanatory variables and occurrence of yelkouan shearwaters: chlorophyll a (a), distance from coast (b), bottom depth (c), sea surface temperature (d), day of the year (e), salinity (f), sampling effort index (g). Top graphs represent the effect of each variable on the prediction of bird presence/absence, bottom graphs are density plots for each variable.