The Colombian Caribbean Sea: a tropical habitat for the Vulnerable sperm whale *Physeter macrocephalus*?

ISABEL CRISTINA AVILA, NOHELIA FARÍAS-CURTIDOR, LUISA CASTELLANOS-MORA
KARINA BOHRER DO AMARAL, DALIA C. BARRAGÁN-BARRERA
CARLOS ANDRÉS OROZCO, JORGE LEÓN and VLADIMIR PUENTES

Abstract We studied the sperm whale *Physeter macrocephalus* in the Colombian Caribbean by combining data from our offshore surveys of behaviour, encounter rate, group structure and density with data from the literature. We describe for the first time the potential distribution of sperm whales in the Colombian Caribbean, using sighting and acoustic data obtained during our surveys, published information, and opportunistic encounters during 1988–2020. We conducted surveys on seismic vessels over 703 days during 2011–2016, covering an area of 68,904 km². We recorded 98 individuals in a total of 50 groups, a density of 1.42 individuals per 1,000 km². To determine the potential distribution of the species, we built *Maxent* models with uncorrelated environmental variables at five depths (from the surface to c. 2,000 m). The model for 1,000 m depth had the best performance, with areas of high probability of occurrence of sperm whales in the south and north-east Colombian Caribbean over the shelf break to waters up to c. 3,000 m deep, at a median distance of 107 km from the coast, and near the Archipelago of San Andrés, Old Providence and Saint Catherine in the north-west. This area may be an important tropical habitat for sperm whales, in which they socialize, rest, breed and feed. Our study underlines the importance of monitoring marine mammals offshore and describes the potential distribution of sperm whales in the Colombian Caribbean, supporting conservation actions for this Vulnerable species, which is currently facing several threats in this region.

Keywords Atlantic, Colombia, depth levels, environmental variables, *Maxent* model, *Physeter macrocephalus*, species distribution, sperm whale

Supplementary material for this article is available at doi.org/10.1017/S0030605321001113

ISABEL C. AVILA (Corresponding author, orcid.org/0000-0003-1389-8908, isabel_c_avila@yahoo.com) Grupo de Ecología Animal, Universidad del Valle, Cali, Colombia, and Institute for Terrestrial and Aquatic Wildlife Research, University of Veterinary Medicine Hannover, Foundation, Büsum, Germany

NOHELIA FARÍAS-CURTIDOR (orcid.org/0000-0002-2617-8988) Fundación Macuáticos Colombia, Medellín, Colombia

LUISA CASTELLANOS-MORA (orcid.org/0000-0001-6247-7021) Independent researcher, Bogotá, Colombia

KARINA BOHRER DO AMARAL (orcid.org/0000-0002-5800-8887) Laboratório de Sistemática e Ecologia de Aves e Mamíferos Marinhos, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

DALIA C. BARRAGÁN-BARRERA (orcid.org/0000-0003-4023-9908) Fundación Macuáticos Colombia, Medellín, Colombia, and Centro de Investigaciones Oceanográficas e Hidrográficas del Caribe, Cartagena de Indias, Colombia

CARLOS ANDRÉS OROZCO (orcid.org/0000-0001-6711-7930) Independent researcher, Archipelago of San Andrés, Old Providence and Saint Catherine, Colombia

JORGE LEÓN and VLADIMIR PUENTES (orcid.org/0000-0002-8796-1810) Anadarko Colombia Company, Bogotá, Colombia

Received 7 April 2021. Revision requested 30 July 2021.
Accepted 5 August 2021.

Introduction

The sperm whale *Physeter macrocephalus* is a deep diving, top marine predator distributed worldwide, commonly in offshore and deep waters (usually > 1,000 m), concentrated in areas known as 'grounds' (Whitehead, 2003). The distribution of sperm whales is associated with areas of upwelling, temperature gradients, seafloor relief and with processes supporting food webs that include mesopelagic or demersal cephalopods, on which sperm whales feed (Baumgartner et al., 2001; Evans & Hindell, 2004). However, females with their young are usually restricted to temperate and tropical waters at low latitudes, where sea surface temperatures are > 15 °C. Males leave their mothers at c. 10 years of age, moving to colder waters at higher latitudes, returning in their late 20s to the tropical and subtropical habitat of females, to mate (Whitehead, 2003; Whitehead et al., 2012).

Sperm whales are categorized as Vulnerable on the IUCN Red List (Taylor et al., 2019) and nationally in Colombia (Rodríguez-Mahecha et al., 2006), having been hunted for 2 centuries, until the 1990s, across all oceans (Whitehead, 2002). Although commercial hunting has ceased, sperm whales face threats from incidental catch, interactions with fishing gear, collisions with boats, and pollution (Avila et al., 2018). The global population trend of the species is unknown (Taylor et al., 2019), but the sperm whale population in the eastern Caribbean declined during 2005–2015 (Gero & Whitehead, 2016). The global offshore distribution of sperm whales and their habitat use are poorly known (Whitehead, 2003), and there have been few studies of the species in Colombia despite its occurrence in both Pacific and Caribbean waters (Trujillo et al., 2013).

The Colombian Caribbean region is important for cruise tourism (Aguilera et al., 2006) and fishing activities (Suárez & Rehder, 2009), but there is little information on the ecology of this region to support management decisions. Studies of the occurrence of marine mammals in the Colombian Caribbean Sea have been limited to coastal areas, with few studies offshore (e.g. Pardo et al., 2009; Farías-Curtidor et al., 2017). Information about sperm whales in Colombian waters is scarce, and the few studies on sperm whales in this region have been in The Bahamas (Ward et al., 2012) and close to Dominica and surrounding islands (e.g. Gordon et al., 1998; Gero et al., 2007, 2014; Gero & Whitehead, 2016), where small aggregations of up to 14 adult females and subadults of unknown sex have been reported (Ward et al., 2012; Gero et al., 2014). It has been suggested that the eastern Caribbean is an ecological trap for the species (Whitehead & Gero, 2015). Given the importance of the eastern Caribbean region for sperm whales and the documented population decline there, which potentially suggests migration to surrounding areas with better conditions (Whitehead & Gero, 2015; Gero & Whitehead, 2016), such as The Bahamas, it is important to determine the environmental conditions that affect sperm whale distribution in other areas in the Caribbean, such as Colombian waters.

Here we report our research on the behaviour, encounter rate, group structure and density of sperm whales in the Colombian Caribbean through the compilation and analysis of data from offshore surveys. To describe the potential distribution of the species in this region, we investigate how environmental conditions at various depths influences its distribution. We identify areas of the Colombian Caribbean where sperm whales are present, and demonstrate the importance of offshore monitoring of marine mammals to generate data for management plans in this region.

Study area

The study area is the Colombian Caribbean, which comprises a total area of 132,288 km² (Fig. 1). This region is characterized by a mean depth of 2,700 m, with a maximum of 4,500 m, and a wide continental shelf (70–150 km) that extends to 200 m depth (Tabares et al., 1996). Temperature at the surface is 27–29 °C, and varies from 18 °C at 200 m to 6 °C at 800 m; salinity is 35–37 ppt (Andrade et al., 2015).

Methods

Surveys

During February–November in 2011 and 2013–2016 we recorded occurrences of sperm whales and collected environmental data in the Colombian Exclusive Economic Zone in the Caribbean Sea from aboard seismic vessels in an area being explored for oil and gas (Supplementary Fig. 1). Visual surveys were during 06.00–18.30 by two biologists trained in observation of marine fauna. Observations, using 10 × 50 binoculars, were made from the highest platform of five survey vessels (Veritas Viking, Osprey Explorer, Polar Duke, Oceanic Sirius and Oceanic Vega), with a mean observation height of 18 m and at a mean speed of 4.2 knots. When possible, observed sperm whales were photographed. Sperm whales were located via their lateral blowing out and by their dorsal fins, tails or body (Farías-Curtidor et al., 2020). For each sighting, date, time, location (with a GPS), number of individuals, presence of any juveniles, follow-up time, depth of bottom, sea conditions (Beaufort), cloudiness and visibility were recorded. Behaviour was recorded using ad libitum sampling. Vessels halted any seismic activity when a marine mammal was close to the vessel (< 500 m), to avoid or mitigate any potential negative impact of this activity on them (JNCC, 2017), and therefore we did not evaluate the behaviour of whales in relation to seismic activity.

For seismic surveys during 2013–2016, acoustic data were collected by a passive acoustic monitoring operator during 18.31–05.59). Sperm whales were detected acoustically from their wideband clicks, which can be distinguished from other marine sounds (Mellinger et al., 2003).

Sperm whale occurrence and density data analyses

Encounter rate and group size statistics were estimated for sighting surveys, the former as the number of individuals and groups sighted per 100 h of observation effort and the latter as the number of individuals and groups observed per 1,000 km².

Modelling sperm whale distribution

We used the maximum entropy algorithm, in *Maxent* (Phillips et al., 2017), to model the potential distribution of sperm whales in the Colombian Caribbean Exclusive Economic Zone using our observations combined with published data. *Maxent* estimates the geographical range of a species by finding the distribution that has the maximum entropy constrained by the environmental conditions recorded at occurrence locations. Models were performed using the *maxnet* function in *Maxent* with a complementary log–log transformation, which appears to be most appropriate for estimating probability of presence (Phillips & Dudík, 2008; Phillips et al., 2017).

As our data were not collected along survey track lines, real absences were not available, and therefore to represent pseudo-absences we randomly selected locations lacking presence data (Phillips & Dudík, 2008; Merow et al., 2013).

  doi:10.1017/S0030605321001113

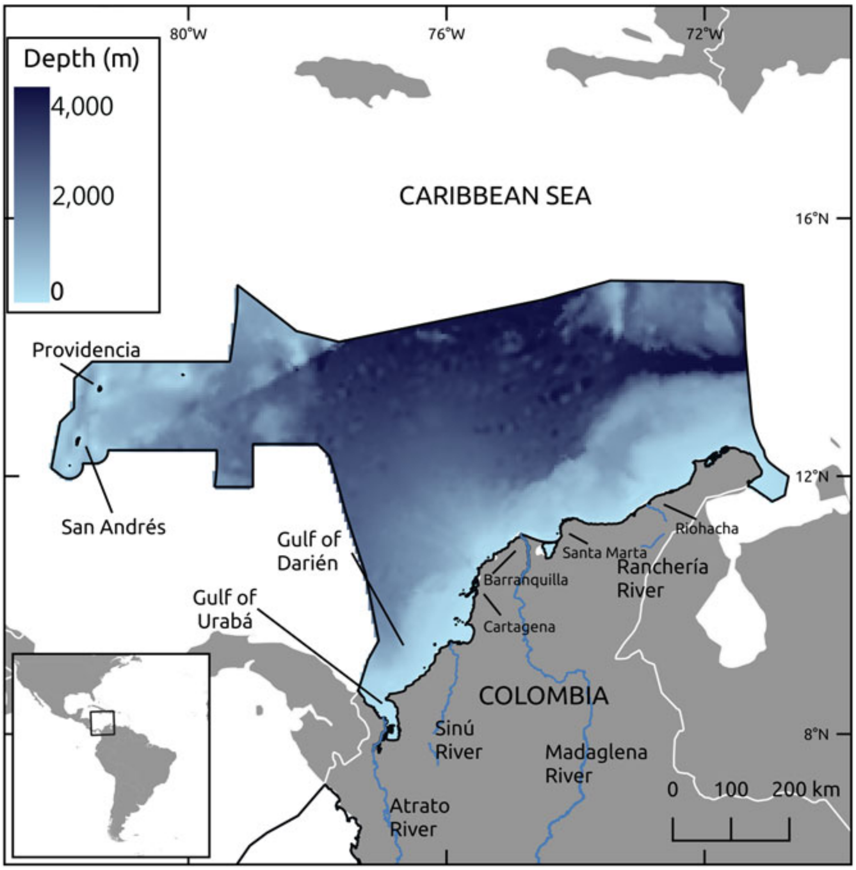

FIG. 1 The Colombian Caribbean Sea, showing the bathymetry of the study area and the locations of the Archipelago of San Andrés, Old Providence and Saint Catherine (the latter two islands labelled Providencia), the Gulf of Urabá, the Gulf of Darién, and the rivers Atrato, Sinú, Magdalena and Ranchería.

We created one random point per 4.5 × 4.5 km grid cell for the surface model and one random point per 9.2 × 9.2 km grid cell for the models at various depth levels, using the *randomPoints* function in the *Dismo* package (Hijmans, 2017) in *R 3.6.3* (R Core Team, 2020) over the study area defined for each model (Supplementary Material 1).

Environmental variables used in the modelling were selected based on information in Baumgartner et al. (2001), Tobeña et al. (2016) and Barragán-Barrera et al. (2019). As the sperm whale is a deep diving species, we included environmental variables from different levels of the water column: on the surface, at c. 0.5 m depth (Level 1), c. 500 m (Level 2), c. 1,000 m (Level 3), c. 1,500 m (Level 4) and c. 2,000 m (Level 5) (Supplementary Fig. 2). Source, types of environmental data available, spatial resolution and time span of sea surface data differ from those of data for various depths, and therefore we chose data that we considered to be equivalent. One exception was the inclusion of the ocean mixed layer thickness data in the models at different depth levels, which was not included as a surface layer. We analysed variable importance and selected uncorrelated environmental variables following the method proposed by Dormann et al. (2013), implemented by Zurell et al. (2020). Firstly, we examined the importance of variables for the surface and different depths using a simple generalized linear model for each potential predictor, and ranked variable importance with Akaike's information criterion. We then inspected correlations between environmental layers to identify all pairs of variables that had a Spearman correlation coefficient > 0.7, removing the less important variable from further analyses (Supplementary Material 1, Supplementary Tables 1–3). From a total of 111 environmental layer candidates, we selected 22, including dynamic (ocean mixed layer thickness, salinity, temperature, total chlorophyll a and phytoplankton) and static (bathymetry, distance to shore, seafloor aspect and slope) variables. We extracted predictor values for each occurrence, and background points, using the function *Fun_Extract* (Derville et al., 2018), which returns the closest values for empty cells. To summarize the environmental conditions at the surface and depth levels, and to describe the environmental heterogeneity of the water column, we conducted a principal component analysis (PCA) for the 22 selected layers, using the function *rasterPCA* in the *Rstoolbox* package (Leutner & Horning, 2016) in *R* (Supplementary Table 4).

Maxent model settings were defined through the *ENMevaluate* function of the *ENMeval 0.3.0* package (Muscarella et al., 2014) in *R*, which provides species-specific settings such as feature classes and regularization multipliers to generate models (see Supplementary Material 2 for details). Model performance and cross-validation predictions were estimated using a series of adapted functions

TABLE 1 Number of groups and individuals of sperm whales *Physeter macrocephalus* recorded during the daytime and night-time during 2011 and 2013–2016 in the Colombian Caribbean (Fig. 1), with hours of survey effort and area surveyed. Number of individuals by age was only recorded during the daytime.

	Effort (h)		Area (km^2)		Number of groups		Number of individuals		Number of individuals by age	
Year	Daytime	Night-time	Daytime	Night-time	Daytime	Night-time	Daytime	Night-time	Adults	Juveniles
2011	573.0	0.0	2,876.0	0.0	5	0	15	0	15	0
2013	1,739.2	872.9	22,980.7	3,097.0	16	9	37	15	36	1
2014	333.3	0.0	4,410.0	0.0	6	0	13	0	12	1
2015	2,757.1	2,330.3	10,878.6	9,105.4	3	5	4	5	3	1
2016	2,547.7	1,916.3	8,878.7	6,678.3	5	1	8	1	7	1
Total	7,950.3	5,119.5	50,024.0	18,880.7	35	15	77	21	73	4

(Zurell et al., 2020). The functions partition the data into k folds ($k = 15$), determine the model algorithm, update the model for the new training data, and make predictions for the hold-out data. The values of the area under the curve (AUC) and the true skill statistic were used as indicators of the predictive ability of the models. The best model is that with an AUC value closest to 1 (Phillips et al., 2006). For the true skill statistic, a value > 0.5 indicates a good prediction (Tobeña et al., 2016). Two types of model output are commonly used to describe the potential distribution of a species: continuous results in which sites are assigned a probability of being part of a species' distribution, and binary results in which sites are classified as either part of the distribution of the species or outside their distribution (Liu et al., 2009). For the former, the final output maps derived from the cross-validation predictions for each type of model were exported in raster format with values in the range 0–1, and were interpreted as an estimate of occurrence probability. For the latter, binary maps were constructed to indicate where sperm whales could be present. To do this, the maps of occurrence probability were transformed by calculating an optimal threshold (a value that maximizes the sum of model sensitivity plus specificity) using the *PresenceAbsence* package (Freeman & Moisen, 2008) in *R*, implemented in the script of Zurell et al. (2020). As we used several record types of sperm whale occurrence (published and opportunistic acoustic and sighting data), data were first cleaned, removing duplicate records and retaining only one occurrence record per grid cell.

Results

Monitoring

During 2011 and 2013–2016 we surveyed for a total of 13,069.8 h in 703 days over an area of 68,904.7 km^2 (Table 1). We covered 52.1% of the offshore areas of the Colombian Caribbean Sea (Supplementary Fig. 1). We recorded a total of 98 individual sperm whales in 50 groups, of which 77 individuals (73 adults and four juveniles) in 35 groups were during daytime in 7,950.3 h of surveys over 50,024 km^2, and 21 individuals in 15 groups were during night-time in 5,119.5 h of acoustic surveys over 18,880.7 km^2. Group sizes were 1–10 (mean 2.0 ± SD 1.5) and each group was recorded for a mean of 18.9 ± SD 22.1 minutes. All observations of sperm whales were during seismic surveys and were outside the mitigation zone (a radius of 500 m around the sound source; JNCC, 2017); 62 and 38% of these observations were when the seismic airgun was active and inactive, respectively. Sperm whales exhibited slow and fast swimming, exposure of pectoral and caudal fins, resting, spyhopping (putting head out of water and looking around) and breaching behaviours (Plate 1).

Mean encounter rate was 0.8 individuals and 0.4 groups per 100 h, with 0.7 individuals and 0.5 groups per 100 hours during daytime, and 0.8 individuals and 0.3 groups per 100 h during night-time. The estimated density of sperm whales was 1.42 individuals and 0.7 groups per 1,000 km^2, with 1.6 individuals and 0.7 groups per 1,000 km^2 during daytime and 1.1 individuals and 0.8 groups per 1,000 km^2 during night-time.

Distribution models

Data for a total of 66 groups of sperm whales and at least 124 individuals were recorded in our surveys and opportunistic sightings, and published data, combined, during 1988–2020, of which eight records were juveniles (0.8 per year). Sperm whales were recorded at a mean distance of 72.5 km from the coast (6–237 km), in a mean water depth of 1,700 m (244–4,191 m; Table 2, Fig. 2).

The PCA indicated environmental heterogeneity across the water column, with the surface the most differentiated compared to the five water depth levels (Supplementary Fig. 3). The generalized linear models, ranked by Akaike's information criterion, indicated the most important variables that explain sperm whale distribution were distance to shore, and range and standard deviation of ocean mixed layer thickness (Supplementary Table 4).

  doi:10.1017/S0030605321001113

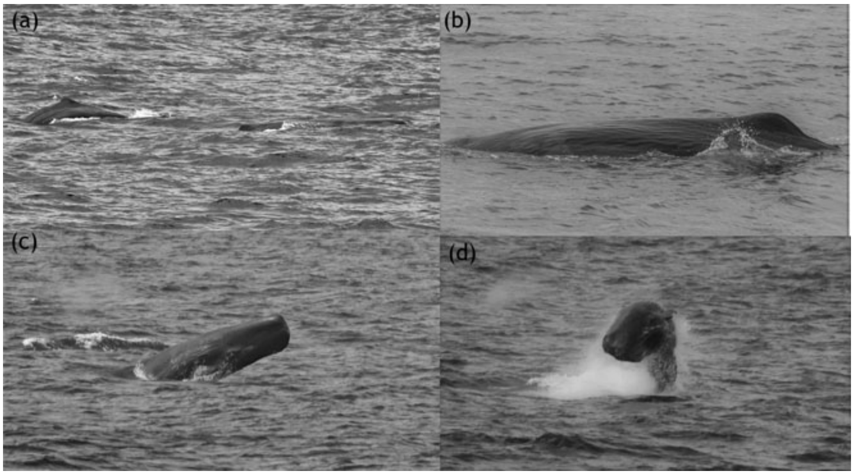

PLATE 1 Some of the behaviours of sperm whales *Physeter macrocephalus* that we recorded in the Colombian Caribbean: (a) two adults swimming slowly (photo: Nohelia Farías-Curtidor); (b) an adult resting (photo: Javier Alarcón); (c–d) a juvenile breaching (photo: Nohelia Farías-Curtidor).

After data cleaning, 61 occurrences of sperm whale groups were available for *Maxent* modelling, but with differing numbers of records at different depth levels (Table 2). All models generated had an AUC > 0.77 and true skill statistic > 0.47, indicating a good performance in general. However, when we considered the optimal threshold, models of the surface and Level 1 failed to assign presence to 59.3% and 17.2% of real sperm whales occurrences used as training data, respectively, whereas the other models failed to assign < 12%. The resulting maps indicated that the probability of occurrence of sperm whales differed between the surface model and the models for Levels 1–5 (Supplementary Figs 4 & 5). The Level 3 model, at c. 1,000 m depth, best represented the potential distribution of sperm whales in the Colombian Caribbean, with the highest AUC (0.84) and true skill statistic (0.55), and assigned presence to > 88% of real sperm whale occurrences (Table 3). The map resulting from this model indicated there is high probability of occurrence of sperm whales in the south and north-east Colombian Caribbean over the shelf break to waters up to c. 3,000 m deep, and near the Archipelago of San Andrés, Old Providence and Saint Catherine in the north-west (Fig. 3). The area of high probability of sperm whale occurrence is characterized by a distance to shore of 9.8–220.5 km (median 106.9 km) and an ocean mixed layer thickness of 5.2–50.3 m (median 29.7 m). The area of occurrence resulting from the model for Level 3 (Fig. 3) had similar features to the real occurrence data (Supplementary Table 5).

Discussion

This study provides the first assessment of the occurrence of sperm whales in the Colombian Caribbean, which appears to be an important habitat for this species. The mean encounter rates of 0.8 individuals and 0.4 groups per 100 h are similar to rates reported in the Gulf of Mexico (1.3 individuals and 0.6 groups per 100 h), where the sperm whale is considered the most abundant large cetacean (Barkaszi et al., 2012). The sperm whale density we recorded in the Colombian Caribbean (1.42 individuals per 1,000 km^2) is similar to that reported worldwide (1.4 individuals per 1,000 km^2; Whitehead, 2002) and to that of other American tropical regions of the Pacific Ocean, but lower than reported for the Colombian Pacific (3.8 individuals per 1,000 km^2; Gerrodette & Palacios, 1996; Supplementary Table 6). Previous studies in the Caribbean had reported lower encounter rates for sperm whales (e.g. 0.35 individuals per 1,000 km^2; Mullin & Fulling, 2004). However, most of these data were from coastal research platforms. Our findings highlight the value of marine mammal occurrence data obtained during seismic surveys, which cover offshore areas that researchers may not usually be able to survey.

Whitehead et al. (2012) found that sperm whale social units in the North Atlantic are based around 6–12 often matrilineally related individuals that move together, raise their calves communally, and probably share important knowledge among themselves. Our findings for the Colombian Caribbean are similar, with groups of up to 10 individuals, and similar to the Gulf of Mexico where groups have up to 16 individuals (Barkaszi et al., 2012). Given our documentation of juveniles and the stranding of a juvenile in 2009 in the Urabá Gulf on the Colombian Caribbean coast (Trujillo et al., 2013), it is possible that the waters of the Colombian Caribbean are a breeding area for Atlantic sperm whales. Further studies are required to examine this possibility. Genetic and photo-identification studies are also required, to assess whether sperm whales sighted in the Colombian Caribbean belong to either of the better-known populations of the eastern Caribbean (Gordon et al., 1998; Gero et al., 2007) or the Gulf of Mexico (Weller et al., 2000), or whether the Colombian Caribbean is an area of connectivity between the eastern and western Atlantic populations. It has been proposed that the eastern Caribbean is a sink with favourable conditions for sperm whales (Whitehead & Gero, 2015), but current threats related to

  doi:10.1017/S0030605321001113

Table 2 Records of sperm whales *Physeter macrocephalus* in the Colombian Caribbean during 1988–2020, with corresponding depth range and distances to the coast.

Year	Month	Colombian Caribbean location[1]	Number of groups	Total number of individuals	Number of juveniles	Location[2]	Record type & source	Depth/depth range (m)	Distance/distance range from coast (km)	Source
1988	May	SE	1	1*	0	Off Barranquilla	Acoustic survey	1,036	26	Pardo et al. (2009)
1990	July	NW	1	10	0	San Andrés	Acoustic survey	441	6	Pardo et al. (2009)
1992	June	NW	4	4*	0	San Andrés	Opportunistic acoustic record	1,043–2,349	19–24	H. Whitehead, unpubl. data
2008	Aug.	NE	2	2*	0	Off Riohacha	Opportunistic acoustic record	561–3,403	34–202	H. Whitehead, unpubl. data
2009	June	NW	1	1*	0	San Andrés	Opportunistic acoustic record	2,047	55	H. Whitehead, unpubl. data
2011	Nov.	NE	5	15	0	East of Riohacha	Sighting survey	774–1,592	41–73	This study
2012	Nov.	SW	1	1	0	Off Cartagena	Sighting survey	900	30	Fundación Omacha & Consultoría y Monitoreo Ambiental (2012)
2013	July–Nov.	SW	24	50	1	Off mouth of Sinú River	Sighting & acoustic surveys	385–2,995	36–99	This study
2013	Sep.–Nov.	SC, NE	1	2	0	Off Barranquilla, Santa Marta & Riohacha	Sighting survey	2,551	38	This study
2014	Oct.–Nov.	NE	6	13	1	East of Riohacha	Sighting survey	1,687–2,277	72–93	This study
2015	June–Nov.	NE	8	9	1	Off Barranquilla & Santa Marta	Sighting & acoustic surveys	2,273–3,889	80–226	This study
2016	Feb.–Sep.	NE	6	9	1	Off Santa Marta & Riohacha	Sighting & acoustic surveys	3,012–4,191	163–237	This study
2020	Aug.	NW	2	2	2	San Andrés	Opportunistic sighting	1,222–1,684	11–20	This study
2020	Sep.	NW	1	2	1	San Andrés	Opportunistic sighting	244	54	This study
2020	Oct.	NW	2	2	1	San Andrés	Opportunistic sighting	446–658	24–44	This study
2020	Nov.	SW	1	1	0	Off mouth of Atrato River in Gulf of Urabá	Opportunistic sighting	20	13	This study

[1]S, south (≤ 12°N); N, north (> 12°N); W, west (< 75°W); E, east (≥ 75°W); C, central.
[2]San Andrés indicates Archipelago of San Andrés, Old Providence and Saint Catherine.
*Minimum number of individuals (total number not confirmed).

Oryx, Page 6 of 11 © The Author(s), 2022. Published by Cambridge University Press on behalf of Fauna & Flora International doi:10.1017/S0030605321001113

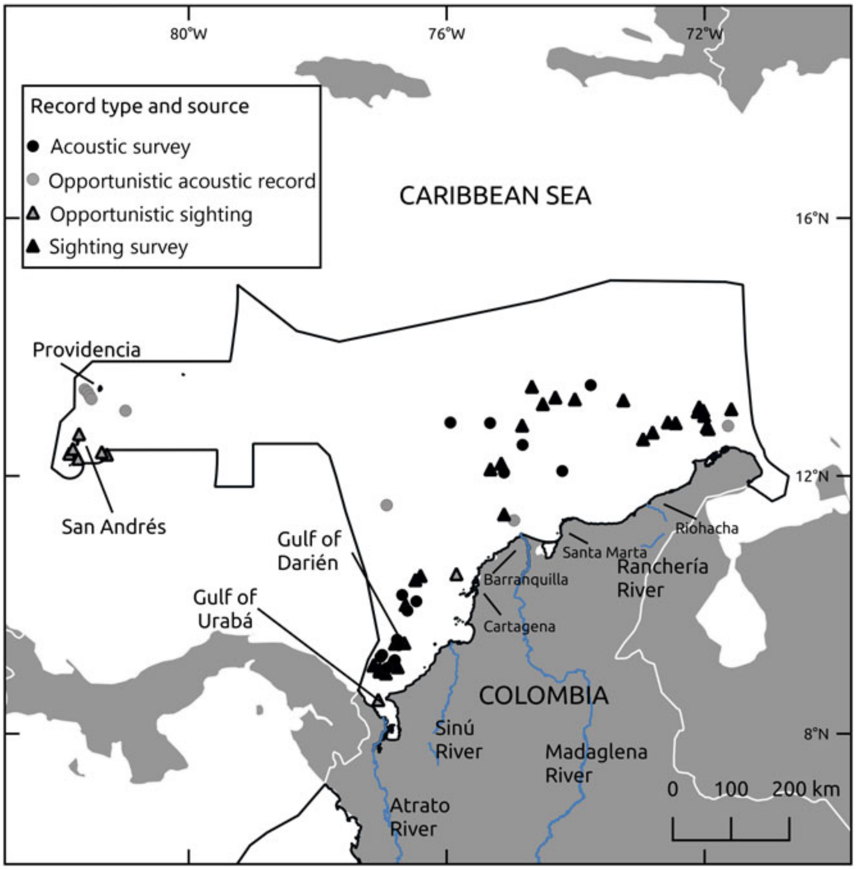

FIG. 2 Records of sperm whales *Physeter macrocephalus* in the Colombian Caribbean during 1988–2020.

human activities in this area (e.g. around the Lesser Antilles) are probably affecting the species (Whitehead & Gero, 2015; Gero & Whitehead, 2016). Although the sperm whale does not appear to conduct long migrations in equatorial waters, when feeding and survival conditions are poor sperm whales tend to roam widely (Whitehead, 2018). Therefore, considering the Caribbean is a relatively small basin, if Colombian waters offer suitable conditions, it is not surprising to find sperm whales from the eastern Caribbean there.

This is the first study to describe the potential distribution of sperm whales in the Colombian Caribbean, using high resolution spatio-temporal variables that are likely to influence sperm whale distribution. Some studies of cetacean distribution in Colombian waters have used line transect surveys (e.g. Palacios et al., 2012), which is the most widely used method to estimate cetacean occurrence. However, this method has high costs and logistical challenges, and low detectability for many cetacean species (Kaschner et al., 2012). Species distribution models, as used here, are useful to estimate the potential distribution of species, particularly in areas where there have not been any line transect surveys, such as in the offshore Colombian Caribbean. In our study, the model built with conditions at 1,000 m depth had the best performance; the surface model had a relatively poor performance, failing to predict probability of occurrence in areas where the species was recorded over the shelf break. This suggests that the analysis of sea surface conditions alone is insufficient to describe the distribution of sperm whales. This is not unexpected, as the sperm whale dives deeply to feed (Whitehead, 2003; Evans & Hindell, 2004). Our results indicate that of the environmental variables tested, the most important were distance to shore and ocean mixed layer thickness. The model for 1,000 m depth identified that the area with a high probability of sperm whale occurrence is close to the shore (median = 106.9 km), with an average range in ocean mixed layer thickness of 29.7 m. These variables may be related to the presence of sperm whale prey. Areas close to the continental shoreline are influenced by rivers and their nutrients, which favour the presence of prey. The ocean mixed layer, which has homogeneous density, temperature and salinity, varies greatly in time and space (e.g. in subpolar latitudes it can be < 20 m in summer but > 200 m in winter; de Boyer Montégut et al., 2004), and plays an important role in phytoplankton and food chain dynamics (Carvalho et al., 2017).

Our modelling indicates that the potential distribution area of sperm whales includes the south and north-east Colombian Caribbean over the shelf break to waters up to c. 3,000 m deep, and near the Archipelago of San Andrés, Old Providence and Saint Catherine in the north-west (Fig. 3). The south and north-east Colombian Caribbean

TABLE 3 Definition of the six models, from the surface to a depth of c. 2,000 m, generated using *Maxent*, with the number of presences used as training data, number of presences with missing data, number of background points, modelling settings, and metrics of cross-validation model performance (area under the curve, true skill statistic, and optimal threshold). The best model is in bold. For additional details, see Supplementary Table 4 and Supplementary Material 2.

Model	Water column level (m)	Number of presences	Number of presences with missing data	Background points	Settings[1]	AUC[2]	TSS[3]	Optimal threshold	Number of real presences assigned as absences
Surface	0	59	2	10,000	LQPT 1	0.80	0.47	0.48	35
Level 1	c. 0.5	58	3	10,000	H 4	0.82	0.48	0.42	10
Level 2	c. 500	54	0	4,245	L 1	0.83	0.49	0.16	5
Level 3	**c. 1,000**	**44**	**1**	**3,971**	**L 3**	**0.84**	**0.55**	**0.27**	**5**
Level 4	c. 1,500	35	0	3,676	H 1	0.81	0.51	0.19	0
Level 5	c. 2,000	21	5	3,422	H 1	0.77	0.48	0.14	0

[1]L, linear; Q, quadratic; P, product; T, threshold; H, hinge.
[2]Area under the curve.
[3]True skill statistic.

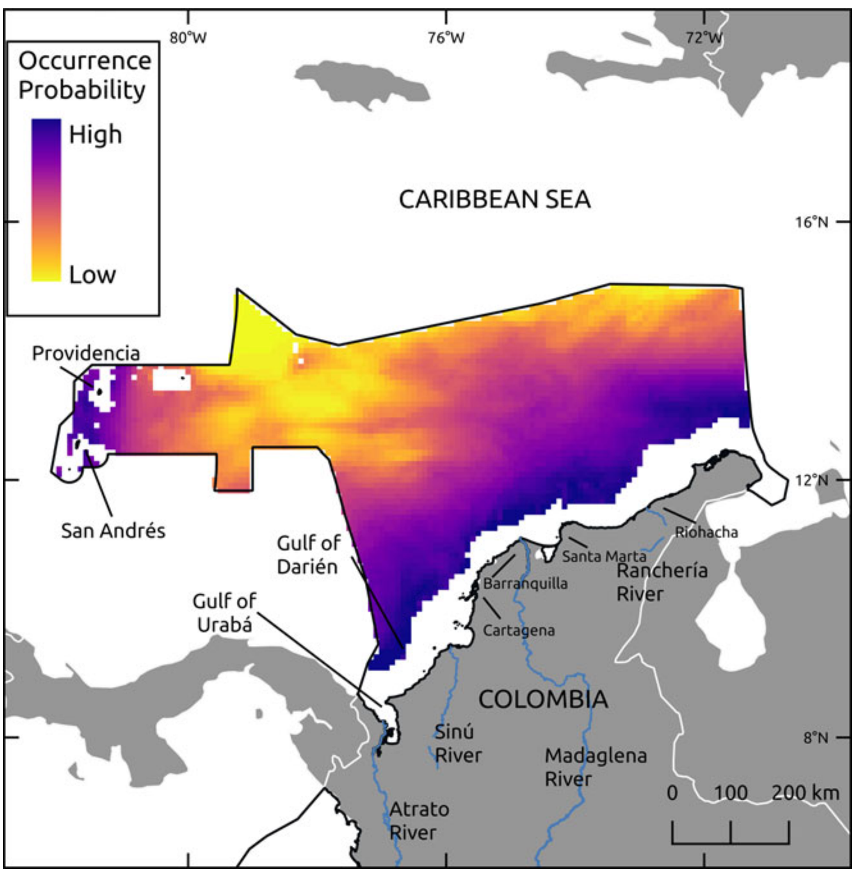

FIG. 3 Potential distribution of the sperm whale in the Colombian Caribbean Sea. Occurrence probability is based on the Level 3 model for environmental conditions at c. 1,000 m depth (Table 3).

are influenced by freshwater discharge from the Atrato, Sinú, Magdalena and Riohacha Rivers and their nutrients, and by upwelling off the southern Caribbean coast, which is probably the main nutrient source supporting biological productivity in this sea (Correa-Ramírez et al., 2020). Rivers provide a nutrient source for the oceans, and their plumes can be zones of high biological productivity, supporting phytoplankton and zooplankton (Montoya et al., 2016). Additionally, such plumes can generate an alluvial fan and an offshore canyon (a geological process that occurs at the mouth of rivers), which could provide habitat suitable for deep-sea species such as the sperm whale. The Archipelago of San Andrés, Old Providence and Saint Catherine, designated a Seaflower Biosphere Reserve in 2000, has extensive coral reefs, seagrass beds and mangroves (CORALINA–INVEMAR, 2012), productive ecosystems that could provide prey for sperm whales. In addition, the Archipelago is close to the 2,350 m Roncador seamount

(Idárraga-García & León, 2019). This could explain the high probability of occurrence of sperm whales in this area, as in the Azores where seamount complexes are prefered habitat for the sperm whale (Tobeña et al., 2016). Sperm whales are deep divers and their main diet is cephalopods, usually subtropical and muscular cephalopod species of the Onychoteuthidae and Histioteuthidae families (Evans & Hindell, 2004), which range throughout the water column to at least 3,000 m depth. In the Colombian Caribbean, 48 cephalopod species have been documented, including two species of Onychoteuthidae (Guerrero-Kommritz, 2021).

Our findings indicate that the Caribbean waters of Colombia may be an important tropical feeding and breeding habitat for sperm whales. In the Caribbean Region there are increasing pressures on marine mammals from coastal development, fishing, boat traffic, river sediment loading, alien species and climate change (Miloslavich et al., 2010; SPAW-RAC, 2020; Avila & Giraldo, 2022). Sperm whales in the region are affected by incidental catch in fishing nets and collision with boats (SPAW-RAC, 2020). In the Colombian Caribbean Sea commercial fisheries are concentrated in the south and north-east (Kroodsma et al., 2018), areas where there is a high probability of sperm whale presence. The Colombian Caribbean region is also an international and national tourist destination (during 2000–2019 an average of 145 tourist cruise ships $>$ 150 m long arrived annually in San Andrés, Cartagena and Santa Marta; CITUR, 2021). The transit of such large ships puts whales at risk of collisions (Laist et al., 2001). In 2009 the tourist cruise ship Summer Flower (169 m long) on the route from Santa Marta (Colombia) to Antwerp (Belgium) collided with a fin whale *Balaenoptera physalus*, and arrived in Belgium with the dead whale on its bow (Haelters et al., 2018).

Our analyses suggest that the Vulnerable sperm whale may be both feeding and breeding in the Colombian Caribbean Sea, and provides information that will be useful for management of this cetacean species. Efforts for the conservation and sustainable use of the distribution area of this species, identified here, need to be implemented in this region.

Acknowledgements We thank Anadarko Colombia Company (since 2019, Occidental Petroleum Corporation), Repsol Exploración Colombia and Shell for allowing the use of data. OSC Limited, EPI Group, CGG & Vision Project Services provided the services of marine fauna observers and passive acoustic monitoring operators and we acknowledge all observers and operators who supported data collection (Salomé Dussán, Juan Manuel Salazar, Alexandra Gärtner, Chad Leedy, Israel Ribeiro, James Doom, Marilia Olio, Leif Halvors and Nicholas Engelmann). We thank fishers from San Andrés and Luis F. Batista from Urabá Gulf for providing information on whales, Javier Alarcón for provision of a photograph, Hal Whitehead for providing access to his data, Ana Marroquim, Paula Gómez, Andrés Palacios and Octavio Ortega for comments on the text, Greg Jenssen for editing help, and two anonymous reviewers and the Editor for their critiques. ICA and DCB-B were supported in 2020–2021 by the Universidad del Valle and Centro de Investigaciones Oceanográficas e Hidrográficas del Caribe, respectively, and the Colombian Sciences Ministry (Call no. 848 of 2019).

Author contributions Study design: ICA; data collection: ICA, NF-C, LC-M; data analysis: ICA, NF-C, KBdA; writing: ICA, NF-C; revision, editing: all authors.

Conflicts of interest None.

Ethical standards This research abided by the *Oryx* guidelines on ethical standards.

References

Aguilera, M.M., Bernal, C. & Quintero, P. (2006) *Turismo y Desarrollo en el Caribe Colombiano*. Documentos de Trabajo Sobre Economía Regional y Urbana, No. 79. Banco de la República–Sucursal, Cartagena, Colombia.

Andrade, C.A., Rangel, O.E. & Herrera, E. (2015) *Atlas de los Datos Oceanográficos de Colombia 1922–2013 Temperatura, Salinidad, Densidad, Velocidad Geostrófica*. Dirección General Marítima-Ecopetrol S.A. Ed. Dimar, Bogotá, Colombia.

Avila, I.C. & Giraldo, A. (2022) Áreas en riesgo para los mamíferos marinos en Colombia. *Revista de Biología Tropical*, 70, 96–113.

Avila, I.C., Kaschner, K. & Dormann, C.F. (2018) Current global risks to marine mammals: taking stock of the threats. *Biological Conservation*, 221, 44–58.

Barkaszi, M.J., Butler, M., Compton, R., Unietis, A. & Bennet, B. (2012) *Seismic Survey Mitigation Measures and Marine Mammal Observer Reports*. Unpublished Report. U.S. Dept. of the Interior, Bureau of Ocean Energy Management, Gulf of Mexico OCS Region, New Orleans, USA.

Barragán-Barrera, D.C., do Amaral, K.B., Chávez-Carreño, P.A., Farías-Curtidor, N., Lancheros-Neva, R., Botero-Acosta, N. et al. (2019) Ecological niche modeling of three species of *Stenella* dolphins in the Caribbean Basin, with application to the Seaflower biosphere reserve. *Frontiers in Marine Science*, 6, 1–17.

Baumgartner, M.F., Mullin, K.D., May, L.N. & Leming, T.D. (2001) Cetacean habitats in the northern Gulf of Mexico. *Fishery Bulletin*, 99, 219–239.

Carvalho, F., Kohut, J., Oliver, M.J. & Schofield, O. (2017) Defining the ecologically relevant mixed-layer depth for Antarctica's coastal seas. *Geophysical Research Letters*, 44, 338–345.

Citur (Centro de Información Turística) (2021) *Estadísticas Nacionales – Turismo Receptor – Llegada de Cruceros*. Ministerio de Comercio, Industria y Turismo, Bogotá, Colombia, citur.gov.co/estadisticas/df_cruceros/barcos/8#gsc.tab=0 [accessed 12 February 2021].

CORALINA–INVEMAR (2012) *Atlas de la Reserva de Biósfera Seaflower. Archipiélago de San Andrés, Providencia y Santa Catalina*. Instituto de Investigaciones Marinas y Costeras "José Benito Vives De Andréis" y Corporación para el Desarrollo Sostenible del Archipiélago de San Andrés, Providencia y Santa Catalina. Serie de Publicaciones Especiales de INVEMAR no. 28, Santa Marta, Colombia.

Correa-Ramírez, M., Rodriguez-Santana, A., Ricaurte-Villota, C. & Páramo, J. (2020) The Southern Caribbean upwelling system off Colombia: water masses and mixing processes. *Deep Sea Research Part I: Oceanographic Research Papers*, 155, 103145.

de Boyer Montégut, C., Madec, G., Fischer, A.S., Lazar, A. & Iudicone, D. (2004) Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, 109, C12003.

Derville, S., Torres, L.G., Iovan, C. & Garrigue, C. (2018) Finding the right fit: comparative cetacean distribution models using multiple data sources and statistical approaches. *Diversity and Distribution*, 24, 1657–1673.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carre, G., Garcia-Marquez, J.R. et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36, 27–46.

Evans, K. & Hindell, M.A. (2004) The diet of sperm whales (*Physeter macrocephalus*) in southern Australian waters. *ICES Journal of Marine Science*, 61, 1313–1329.

Farías-Curtidor, N., Barragán-Barrera, D.C., Chávez-Carreño, P.A., Jiménez-Pinedo, C., Palacios, D., Caicedo, D. et al. (2017) Range extension for the common dolphin (*Delphinus* sp.) to the Colombian Caribbean, with taxonomic implications from genetic barcoding analysis. *PLOS ONE*, 12, e0171000.

Farías-Curtidor, N., López, P., Alarcón, J., Duque, L., Barragán-Barrera, D.C., Avila, I.C. et al. (2020) Marine mammals sighted offshore in the Colombian Caribbean. *In Visual Guide for Birds, Marine Mammals and Other Organisms Sighted Offshore in the Colombian Caribbean* (eds V. Puentes & J. León), pp. 78–97. Anadarko Colombia Company, Bogotá, Colombia.

Freeman, E.A. & Moisen, G. (2008) *PresenceAbsence*: An *R* package for presence-absence model analysis. *Journal of Statistical Software*, 23, 1–31.

Fundación Omacha & Consultoría y Monitoreo Ambiental (2012) *Observación de Fauna Marina, Proyecto Off Shore Pacífico y Caribe Colombiano "Primera y Segunda Grilla"*. Unpublished Report. Fundación Omacha, Bogotá, Colombia.

Gero, S., Gordon, J., Carlson, C., Evans, P. & Whitehead, H. (2007) Population estimate and inter-island movement of sperm whales, *Physeter macrocephalus*, in the eastern Caribbean Sea. *Journal of Cetacean Research and Management*, 9, 143–150.

Gero, S., Milligan, M., Rinaldi, C., Francis, P., Gordon, J., Carlson, C. et al. (2014) Behavior and social structure of the sperm whales of Dominica, West Indies. *Marine Mammal Science*, 30, 905–922.

Gero, S. & Whitehead, H. (2016) Critical decline of the eastern Caribbean sperm whale population. *PLOS ONE*, 11, e0162019.

Gerrodette, T. & Palacios, D.M. (1996) *Estimates of Cetacean Abundance in EEZ Waters of the Eastern Tropical Pacific*. Southwest Fisheries Science Center, Administrative Report JL-96-10. La Jolla, USA.

Gordon, J., Moscrop, A., Caroson, C., Ingram, S., Leaper, R., Matthews, J. & Young, K. (1998) Distribution, movements and residency of sperm whales off the commonwealth of Dominica, eastern Caribbean: implications for the development and regulation of the local whalewatching industry. *Reports of the International Whaling Commission*, 48, 551–557.

Guerrero-Kommritz, J. (2021) Cephalopoda (Mollusca) of the Colombian Caribbean Sea. *Boletín de Investigaciones Marinas y Costeras-INVEMAR*, 50, 191–196.

Haelters, J., Kerckhof, F. & Jauniaux, T. (2018) Strandings of cetaceans in Belgium from 1995 to 2017. *Lutra*, 61, 107–126.

Hijmans, R. (2017) *Raster: Geographic Data Analysis and Modeling. R Package Version* 2.6-7. CRAN.R-project.org/package=raster [accessed 15 November 2020].

Idárraga-García, J. & León, H. (2019) Unraveling the underwater morphological features of Roncador bank, archipelago of San Andrés, Providencia and Santa Catalina (Colombian Caribbean). *Frontiers in Marine Science*, 6, 77.

JNCC (Joint Nature Conservation Committee) (2017) *JNCC Guidelines for Minimising the Risk of Injury to Marine Mammals from Geophysical Surveys*. Joint Nature Conservation Committee, Aberdeen, UK.

Kaschner, K., Quick, N.J., Jewell, R., Williams, R. & Harris, C.M. (2012) Global coverage of cetacean line-transect surveys: status Quo, data gaps and future challenges. *PLOS ONE*, 7, e44075.

Kroodsma, D.A., Mayorga, J., Hochberg, T., Miller, N.A., Boerder, K., Ferretti, F. et al. (2018) Tracking the global footprint of fisheries. *Science*, 359, 904–908.

Laist, D.W., Knowlton, A.R., Mead, J.G., Collet, A.S. & Podesta, M. (2001) Collisions between ships and whales. *Marine Mammal Science*, 17, 35–75.

Leutner, B. & Horning, N. (2016) *Tools for Remote Sensing Data Analysis*. The Comprehensive R Archive Network. cran.r-project.org/web/packages/RStoolbox/RStoolbox.pdf [accessed 12 February 2021].

Liu, C., White, M. & Newell, G. (2009) Measuring the accuracy of species distribution models: a review. In *Proceedings 18th World IMACs/MODSIM Congress, July 2009* (eds R.S. Anderssen, R.D. Braddock & L.T.H. Newham), pp. 4241–4247. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Cairns, Australia.

Mellinger, D.K., Thode, A.M. & Martinez, A. (2003) Passive acoustic monitoring of sperm whales in the Gulf of Mexico, with a model of acoustic detection distance. In *Proceedings: Twenty-First Annual Gulf of Mexico Information Transfer Meeting, January 2002* (eds M. McKay & J. Nides), pp. 493–501. U.S. Department of the Interior, Minerals Management Service, Gulf of Mexico OCS Region, New Orleans, USA.

Merow, C., Smith, M.J. & Silander, Jr, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings mater. *Ecography*, 36, 1058–1069.

Miloslavich, P., Díaz, J.M., Klein, E., Alvarado, J.J., Díaz, C., Gobin, J. et al. (2010) Marine biodiversity in the Caribbean: regional estimates and distribution patterns. *PLOS ONE*, 5, e11916.

Montoya, L.J., Toro-Botero, F.M. & Gómez-Giraldo, A. (2016) Study of Atrato river plume in a tropical estuary: effects of the wind and tidal regime on the Gulf of Urabá, Colombia. *DYNA*, 84, 367–375.

Mullin, K.D. & Fulling, G.L. (2004) Abundance of cetaceans in the oceanic northern Gulf of Mexico. *Marine Mammal Science*, 20, 787–807.

Muscarella, R., Galante, P.J., Soley-Guardia, M., Boria, R.A., Kass, J.M., Uriarte, M. & Anderson, R. (2014) *ENMeval*: an *R* package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, 5, 1198–1205.

Palacios, D.M., Herrera, J.C., Gerrodette, T., García, C., Soler, G.A., Avila, I.C. et al. (2012) Cetacean distribution and relative abundance in Colombia's Pacific EEZ from survey cruises and platforms of opportunity. *Journal of Cetacean Research Management*, 12, 45–60.

Pardo, M.A., Mejía-Fajardo, A., Beltrán-Pedreros, S., Trujillo, F., Kerr, I. & Palacios, D.M. (2009) Odontocete sightings collected during offshore cruises in the western and southwestern Caribbean Sea. *Latin American Journal of Aquatic Mammals*, 7, 57–62.

Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E. & Blair, M.E. (2017) Opening the black box: an open-source release of Maxent. *Ecography*, 40, 887–893.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190, 231–259.

Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with *Maxent*: new extensions and a comprehensive evaluation. *Ecography*, 31, 161–175.

R Core Team (2020) *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. R-project.org [accessed 20 November 2020].

Rodríguez-Mahecha, J.V., Alberico, M., Trujillo, F. & Jorgenson, J. (eds) (2006) *Libro Rojo de los Mamíferos de Colombia*. Serie libros rojos de especies amenazadas de Colombia. Conservación Internacional Colombia y Ministerio de Ambiente, Vivienda y Desarrollo Territorial, Bogotá, Colombia.

Spaw-Rac. (2020) *Implementation of the Action Plan for Conservation of Marine Mammals (MMAP) in the Wider Caribbean Region: A Scientific and Technical Analysis*. UN Environment Programme, Caribbean Environment Programme, Specially Protected Areas and Wildlife Regional Activity Centre, Kingston, Jamaica.

Suárez, A.M. & Rehder, J. (2009) *Actualización del Estado de la Flota Pesquera Comercial Industrial en Colombia*. Instituto Colombiano Agropecuario. Produmedios, Bogotá, Colombia.

Tabares, N., Soltau, J.M. & Díaz, J. (1996) Caracterización geomorfológica del sector suroccidental del mar Caribe. *Boletín Científico CIOH*, 17, 3–16.

Taylor, B.L., Baird, R., Barlow, J., Dawson, S.M., Ford, J., Mead, J.G. et al. (2019) *Physeter macrocephalus* (amended version of 2008 assessment). In *The IUCN Red List of Threatened Species* 2019. dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T41755A160983555.en [accessed 2 March 2022].

Tobeña, M., Prieto, R., Machete, M. & Silva, M.A. (2016) Modeling the potential distribution and richness of cetaceans in the Azores from fisheries observer program data. *Frontiers in Marine Science*, 3, 202.

Trujillo, F., Gärtner, A., Caicedo, D. & Diazgranados, M.C. (eds) (2013) *Diagnóstico del Estado de Conocimiento y Conservación de los Mamíferos Acuáticos en Colombia*. Ministerio de Ambiente y Desarrollo Sostenible, Fundación Omacha, Conservation International and WWF, Bogotá, Colombia.

Ward, J.A., Thomas, L., Jarvis, S., DiMarzio, N., Moretti, D., Marques, T.A. et al. (2012) Passive acoustic density estimation of sperm whales in the tongue of the ocean, Bahamas. *Marine Mammal Science*, 28, E444–E455.

Weller, D.W., Würsig, B., Lynn, S.K. & Schiro, A.J. (2000) Preliminary findings on the occurrence and site fidelity of photo-identified sperm whales (*Physeter macrocephalus*) in the northern Gulf of Mexico. *Gulf of Mexico Science*, 18, 35–39.

Whitehead, H. (2002) Estimates of the current global population size and historical trajectory for sperm whales. *Marine Ecology Progress Series*, 242, 295–304.

Whitehead, H. (2003) *Sperm Whales: Social Evolution in the Ocean*. University of Chicago Press, Chicago, USA.

Whitehead, H. (2018) Sperm whale: *Physeter macrocephalus*. In *Encyclopedia of Marine Mammals* (eds B. Würsig, J.G.M. Thewissen & K.M. Kovacs), pp. 919–925. Academic Press, Cambridge, USA.

Whitehead, H., Antunes, R., Gero, S., Wong, S.N.P., Engelhaupt, D. & Rendell, L. (2012) Multilevel societies of female sperm whales (*Physeter macrocephalus*) in the Atlantic and Pacific: why are they so different? *International Journal of Primatology*, 33, 1142–1164.

Whitehead, H. & Gero, S. (2015) Conflicting rates of increase in the sperm whale population of the eastern Caribbean: positive observed rates do not reflect a healthy population. *Endangered Species Research*, 27, 207–218.

Zurell, D., Zimmermann, N.E., Gross, H., Baltensweiler, A., Sattler, T. & Wüest, R.O. (2020) Testing species assemblage predictions from stacked and joint species distribution models. *Journal of Biogeography*, 47, 101–113.