



## RESEARCH PAPER

Journal of  
Biogeography

WILEY

# A treasure from the past: Former sperm whale distribution in Indonesian waters unveiled using distribution models and historical whaling data

Achmad Sahri<sup>1</sup> | Mochamad I. H. Putra<sup>2</sup> | Putu L. K. Mustika<sup>3,4</sup> |  
Albertinka J. Murk<sup>1</sup>

<sup>1</sup>Marine Animal Ecology Group, Wageningen University and Research, Wageningen, The Netherlands

<sup>2</sup>Marine Megafauna Research Group, Misool Foundation, Savu Sea Program, East Flores, Indonesia

<sup>3</sup>Cetacean Sirenian Indonesia (CETASI), Jakarta Utara, Indonesia

<sup>4</sup>College of Business, Law and Governance, James Cook University, Townsville, QLD, Australia

**Correspondence**

Achmad Sahri, Marine Animal Ecology Group, Wageningen University and Research, Droevendaalsesteeg 1 6708 PB Wageningen, The Netherlands.  
Email: achmad.sahri@wur.nl

**Funding information**

Indonesia Endowment Fund for Education (LPDP) Scholarship from the Ministry of Finance of the Republic of Indonesia, Grant/Award Number: PRJ-482/LPDP.3/2017

**Handling Editor:** Christine N. Meynard

**Abstract**

**Aim:** This study is the first in Indonesia to assess historical sperm whale (*Physeter macrocephalus*) seasonal distributions by combining historical whaling data with environmental factors associated with sperm whale habitat preferences. As current records of whale occurrence covering the whole of Indonesian waters are incomplete, we used historical whaling data summarized by Charles Haskins Townsend in 1935 to model its potential distribution for each season.

**Location:** Indonesian waters (92-143E, 9N-14S).

**Taxa:** Sperm whale (*P. macrocephalus*).

**Methods:** We used a presence-only habitat model – Maximum Entropy (Maxent) –, and a presence–pseudo-absence method – generalized additive model (GAM) – with nine submerged topographic variables to predict historical seasonal distributions.

**Results:** Both Maxent and GAM predict similar potential distribution which align closely with the whaling data. The results indicate that in four areas in the eastern part of Indonesia, no seasonal differences occurred in sperm whale distribution, while noticeable seasonal differences were indicated in other areas. The key parameters that characterize sperm whale habitat in both models were distance to coast, distance to –1,000 and –5,000 m isobaths, and submarine key features such as trough and trench.

**Main conclusions:** The historical catch data of this species can be used to describe the historical species distribution and provide a baseline to assess present distribution, prioritize current research and monitoring and recommend future data collection. Our models also predict distributions that are significantly larger than the ones occupied by sperm whales nowadays, another example of a shifting baseline. Our study demonstrates the benefits of incorporating historical whaling data into habitat models for ecological investigation and to inform conservation efforts for cetaceans.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Journal of Biogeography* published by John Wiley & Sons Ltd

## KEYWORDS

cetacean, conservation management, generalized additive model, historical whaling data, Maxent, species distribution model, sperm whale

## 1 | INTRODUCTION

Comprehensive knowledge on where marine animals occur, and which characteristics determine their habitat preferences, is important in understanding the ecology of the species and guiding appropriate conservation and management efforts (Kanaji, Okazaki, Kishiro, & Miyashita, 2015; Redfern et al., 2006). However, animal occurrence data are deficient for many marine species (Redfern et al., 2017) in many regions (Moura, Sillero, & Rodrigues, 2012). For instance, records of cetacean occurrence in Indonesian waters are lacking. This can be attributed to the high logistical costs of conducting expansive surveys (Johnson et al., 2016; Redfern et al., 2006; Thorne et al., 2012). Historical data are therefore an interesting alternative to consider (Johnson et al., 2016; Reeves, Smith, Josephson, Clapham, & Woolmer, 2004; Torres et al., 2013). It could even serve to demonstrate shifting baselines in species densities caused by historical factors such as intensive exploitation (Jackson, 2001).

Sperm whales (*Physeter macrocephalus*) were hunted extensively across all oceans for two centuries during the Yankee whaling era (Whitehead, 2002), when females were the focus on tropical grounds (Best, 1979; Whitehead, 2003). Males were preferred by modern whalers for their larger body size (Johnson et al., 2016). This difference in targeting led to a sex ratio imbalance that may have limited population recovery ever since (Carroll, Hedley, Bannister, Ensor, & Harcourt, 2014; Whitehead, 2002). Sperm whale recovery was also hindered by whaling throughout the last century, some of it illegal and poorly documented (Smith, Reeves, Josephson, & Lund, 2012). This species is now listed as 'vulnerable' under the IUCN Red List of Endangered Species (Taylor et al., 2008).

The long history of the exploitation of sperm whales provides rough information on their past distribution. In 1935, Charles Haskins Townsend published four global charts from the information recorded in American whaling logbooks and journals between 1761 and 1920. The charts show where and in what month American sail whalers captured a total of 36,909 sperm whales globally. Of the five species for which Townsend published charts, only the sperm whale was reported to be captured in large numbers in equatorial Indonesian waters (Figure 1), with a ratio of 79:1 compared to humpback whales. The extensive global analysis published by Townsend (1935) reveals three priority whaling grounds in the Indonesian archipelago: the Molucca Passage, the Celebes Sea and the Sulu Sea Grounds (Figure 1).

Townsend's hand plotting of catch points, however, creates difficulties for interpretation (Smith, Reeves, et al., 2012), and efforts are needed to relate the fragmented catch points to environmental factors in order to assess the continuous surface distribution of the species. In Indonesian waters, Townsend mentions three whaling grounds across the archipelago, with all sperm whales caught outside them referred

to as stragglers. The West Banda Sea is not mentioned, even though abundant sperm whale captures were reported there (Figure 1). Townsend did report a seasonal oscillation of sperm whales between the northern and southern latitude. For a better interpretation of the historical data and to assess the full former distribution of sperm whales, habitat modelling can be applied. This approach could also help to indicate seasonal migration between habitats in the past by analysing whale presence across the different seasons.

Species distribution models (SDMs) can provide quantitative predictions of the geographic distribution of species based on environmental variables. These results can then be used for conservation, policy and spatial planning. SDMs have been constructed based on historical whaling data before, and this has helped scientists to identify the core habitats of certain species, which is valuable for understanding species ecology and for conservation management (Elith et al., 2011; Gregr, 2011; Pearce & Boyce, 2006; Torres et al., 2013) and provided improvements in the usability of the available data (Kanaji et al., 2015). Identifying important habitats can help minimize adverse human and cetacean interactions (Breen, Brown, Reid, & Rogan, 2016) by implementing spatially explicit conservation measures. However, this attempt has never been done for Indonesian waters, even though it would provide important information about suitable sperm whale habitats.

The objective of this study was to identify important habitat areas for sperm whales within the Indonesian archipelago during the historical whaling era and to understand their past distribution patterns and seasonal differences. The relevance of the information gained from this study for current sperm whale management is also discussed.

## 2 | MATERIALS AND METHODS

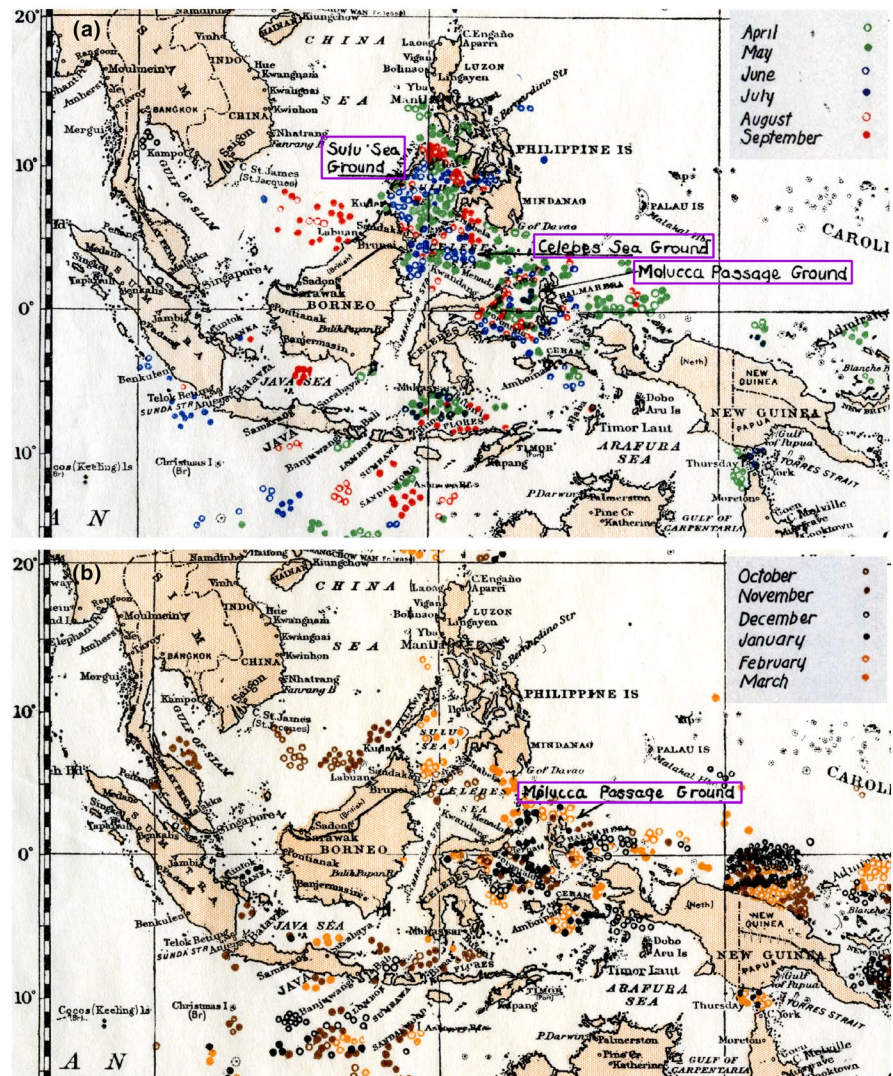
### 2.1 | Study area

The study area covers Indonesian waters from 92° to 143°E and from 9°N to 14°S (Figure 2). The area spans c. 14.4 million km<sup>2</sup> and contains a diversity of habitat types, from shallow-sloping to deep-steep submerged waters (Figure A2a,b in Appendix S1).

### 2.2 | Historical whaling data

The presence data of sperm whales ( $n = 793$ ) were obtained from historical whaling (Townsend, 1935, Figure A2j in Appendix S1). This information was available from sperm whalers' log books from the 18th to the early 20th century. Their voyages typically lasted 2–4 years, encompassing tropical and temperate seas throughout all seasons, only docking at port for supplies or repairs (Townsend, 1935). Age

**FIGURE 1** Townsend's (1935) sperm whale charts within Indonesian archipelago during Yankee whaling period 1761–1920: (a) April–September and (b) October–March. Reproduced with permission from the New Bedford Whaling Museum. The colours used to indicate the locations of sperm whale catches indicate the catch months. The whaling ground names are shown in purple boxes



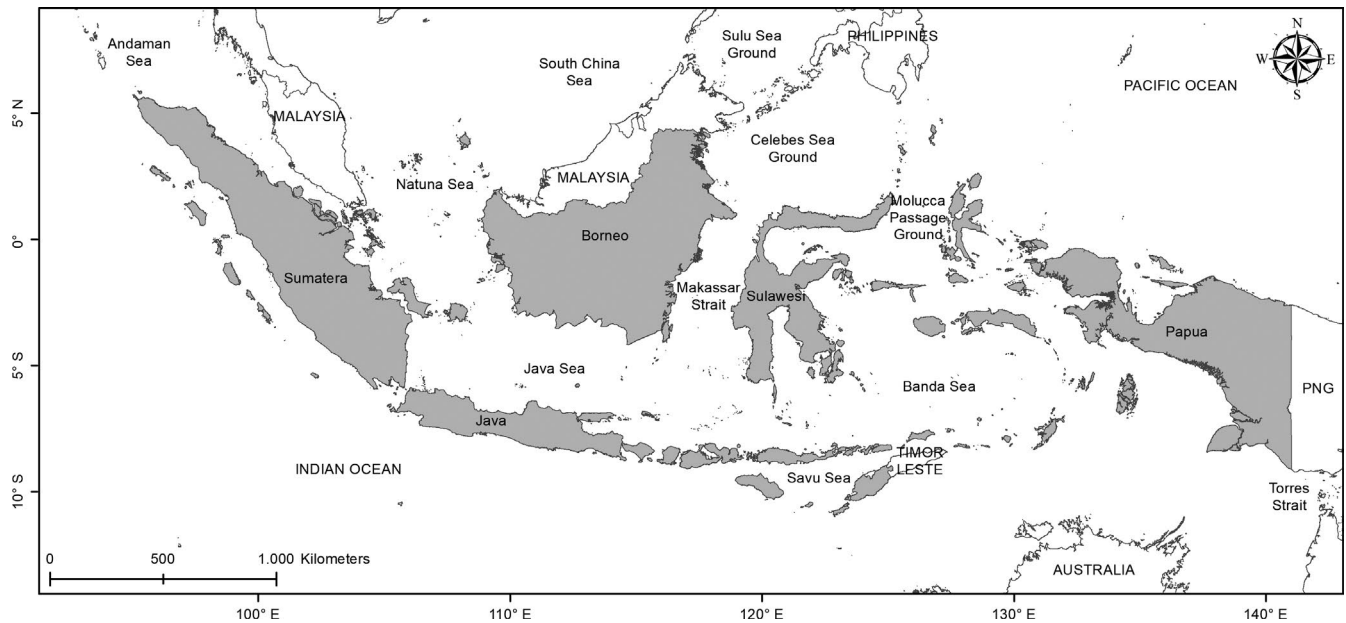
(juvenile, adult) and sex (male, female) segregation were not available from the dataset. However, as Yankee whaling focused on female social groups on tropical grounds (Best, 1979; Whitehead, 2003), the sperm whales were assumed to be mainly adult females. We combined the whaling data into one all year dataset, ignoring seasons (hereafter, 'all year') and four independent groups based on the monsoonal seasons, i.e. Transition 1 (T1 season, March–May,  $n = 188$ ), SE monsoon (June–August,  $n = 146$ ), Transition 2 (T2 season, September–November,  $n = 207$ ) and SW monsoon (December–February,  $n = 252$ ).

The absence data were available from daily locations of vessels extracted from logbooks for voyages between 1780 and 1920 (Smith, Reeves, et al., 2012). However, the amount of absence data was very limited and unevenly distributed over the study area (mostly in the southern part). As our preliminary work showed that this results in biased distribution models and poor model predictions (data not shown), we randomly generated 10,000 pseudo-absence data for the models as an alternative approach. It has been reported before that the ranks of habitat model calculated using presence

and absence data were similar to that calculated using presence and pseudo-absence data (Redfern et al., 2017).

### 2.3 | Environmental variables

Twelve submerged topographic variables were selected based on the expected ecological relevance to sperm whale habitats (Azzellino et al., 2012; Fiedler et al., 2018; Johnson et al., 2016; Redfern et al., 2017; Schlacher, Rowden, Dower, & Consalvey, 2010) in the study area. Due to unavailability of data during the historical whaling period, and the uncertainty of their stability over time, dynamic oceanographic variables (e.g. chlorophyll-a concentration, sea surface temperature and sea surface salinity) were not included in this work. We checked the collinearity among variables and used only nine (Table 1) with Pearson's correlation values  $<0.75$  (Figure A1 in Appendix S1) in modelling. These variables are bathymetry, slope as well as distance to- coast,  $-1,000$  and  $-5,000$  m isobaths, seamount, shelf, trench and trough (Figure A2 in Appendix S1).



**FIGURE 2** Study area in the Indonesian archipelago. Names of seas and whaling grounds are used in the text

Bathymetry data were obtained from the General Bathymetric Chart of the Ocean (GEBCO, <https://www.gebco.net/>), providing a 1-km grid of bathymetric surface for the study area. Slope was derived from the GEBCO using Spatial Analyst extension in ArcGIS 10.6.1 (Environmental Systems Research Institute, Inc.). Coastal lines were obtained from the Indonesian Geospatial Information Agency, while undersea topographic features were obtained from the Seafloor Geomorphic Features Map (Harris, Macmillan-Lawler, Rupp, & Baker, 2014). Distance to coast, the two isobaths and topographic features were generated using the Euclidean Distance Tool-Spatial Analyst extension in ArcGIS 10.6.1. All variables were in ASCII raster format. The selection of spatial resolutions for environmental variables was primarily based on data availability. As bathymetry and slope were already in a  $1 \times 1 \text{ km}^2$  grid, the other variables were aggregated to match the same grid size and cover the same area. In total, there are 13,659,947 grids within the study area.

## 2.4 | Species distribution modelling

We used two SDMs in this study: the maximum entropy model (Maxent) and the generalized additive model (GAM). The use of different models makes it possible to compare and evaluate their abilities and outputs, and thereby determine which models (and variables) describe the species distribution and habitat preferences with the greatest confidence.

Maxent is a presence-only model previously applied for cetacean distribution studies (Edrén, Wisz, Teilmann, Dietz, & Söderkvist, 2010; Smith, Grantham, et al., 2012; Thorne et al., 2012). Maxent is advantageous for this study because it accounts for spatial bias in presence data and can identify areas that fall beyond the

range of occupied environmental condition (Phillips, Anderson, & Schapire, 2006). Maxent predicts a focal species' most uniform distribution across the study area (the distribution with maximum entropy), in relation to the environmental conditions at the locations of the presence data. The model is then extrapolated to other unsampled areas within the study area to give the habitat suitability of the species based on environmental characteristics of the sites (Phillips & Dudík, 2008; Phillips et al., 2006). We used the Maximum Entropy Species Distribution Modelling software v. 3.4.1 ([https://biodiversityinformatics.amnh.org/open\\_source/maxent](https://biodiversityinformatics.amnh.org/open_source/maxent)).

The use of GAM is common in SDM because it allows the data to identify nonlinearity in species-habitat relationships rather than imposing parametric fits through polynomial terms in a linear regression (Chambers & Hastie, 2017; Hastie & Tibshirani, 1990). GAM is a presence-absence model that has been applied for cetacean studies (Fiedler et al., 2018; Virgili, Racine, Authier, Monestiez, & Ridoux, 2017). The GAM model attempts to differentiate the environmental conditions where the species is observed (presence points) and where it is absent (i.e. in pseudo-absence points). The preferred niche of the species was then projected onto the geographic space with its characteristics to depict its potential distribution. We used R package 'BIOMOD2' v. 3.3-7.1 (<https://cran.r-project.org/web/packages/biomod2/index.html>).

To be able to make a comparison, the same settings were applied within both models: 30% random test percentage, 10 replicates with bootstrap replicated run type, regularization multiplier 1 and maximum number of randomly selected background points 10,000 (the last two parameters were set as defaults, see Elith et al., 2011; Phillips et al., 2006). Both SDMs were built for each of the 'all year' and four monsoonal seasons (T1 season, SE monsoon, T2 season and SW monsoon). Maxent generates logistical spatial predictions of distribution with values between 0 (not suitable) and 1 (highly suitable),

**TABLE 1** Environmental variables used for species distribution modelling

Variables	Unit	Sources	Rationale
Bathymetry	m	General Bathymetric Chart of the Ocean (GEBCO; <a href="https://www.gebco.net/">https://www.gebco.net/</a> )	Shallow water is associated with high primary production. Top predators respond to bathymetric features; shallow topography may provide favourable foraging opportunities (Yen, Sydeman, & Hyrenbach, 2004)
Slope	%	GEBCO ( <a href="https://www.gebco.net/">https://www.gebco.net/</a> ) and ArcGIS derived	Associated with currents, steep benthic relief promotes water movements. High slope induces prey aggregation and/or increases primary production (Yen et al., 2004)
Distance to isobaths			
-200 m <sup>a</sup>	km	GEBCO ( <a href="https://www.gebco.net/">https://www.gebco.net/</a> ) and ArcGIS derived	These distinct depths are associated with the deep diving habits of species (Johnson et al., 2016; Lambert, Mannocci, Lehodey, & Ridoux, 2014)
-1,000 m (d_1000)	km	GEBCO ( <a href="https://www.gebco.net/">https://www.gebco.net/</a> ) and ArcGIS derived	
-2,500 m <sup>a</sup>	km	GEBCO ( <a href="https://www.gebco.net/">https://www.gebco.net/</a> ) and ArcGIS derived	
-5,000 m (d_5000)	km	GEBCO ( <a href="https://www.gebco.net/">https://www.gebco.net/</a> ) and ArcGIS derived	
Distance to			
Coast (d_coast)	km	Indonesian Geospatial Information Agency and ArcGIS derived	Distance provides an indication of preference for near or offshore habitats (Dalla Rosa, Ford, & Trites, 2012)
Ridges <sup>a</sup>	km	Seafloor Geomorphic Features Map (Harris et al., 2014) and ArcGIS derived	
Seamount (d_seamount)	km	Seafloor Geomorphic Features Map (Harris et al., 2014) and ArcGIS derived	
Shelf (d_shelf)	km	Seafloor Geomorphic Features Map (Harris et al., 2014) and ArcGIS derived	
Trench (d_trench)	km	Seafloor Geomorphic Features Map (Harris et al., 2014) and ArcGIS derived	
Trough (d_trough)	km	Seafloor Geomorphic Features Map (Harris et al., 2014) and ArcGIS derived	

<sup>a</sup>Variables that were eliminated due to multicollinearity. The variable names in parentheses in the first column are as named in our models and in Table 3.

while that of GAM using Biomod2 ranges from 0 (not suitable) to 1,000 (highly suitable); therefore, the output values were standardized to be comparable. The correlation between the suitability values predicted by Maxent and GAM per cell values was calculated as a Pearson's correlation coefficient.

To distinguish suitable and unsuitable habitats, we applied the '10th percentile training presence threshold' to the predicted distribution maps (raster format) in ArcGIS 10.6.1. This threshold was chosen because it is the best threshold when true absence data are not available (Brito, Acosta, Álvares, & Cuzin, 2009), and it selects values above which 90% of the training presence are correctly classified, thus giving conservative estimates of prediction (Padalia, Srivastava, & Kushwaha, 2014). The mean threshold of 10 replicates of each seasonal model was used as a binary threshold for presence/absence of sperm whales, above which a suitable habitat is considered to occur. To reveal seasonal similarities and differences in habitat suitability, a seasonal overlap map was generated with the best performing SDM (either Maxent or GAM, see below). The map

was created by combining (stacking) the individual binary seasonal maps.

## 2.5 | Model performance assessment

Species distribution models are commonly assessed by the Area Under the receiver-operating-characteristics (ROC) Curve, or AUC (Phillips et al., 2006). An AUC curve is a graphical plot which illustrates the probability that a random presence data point will be ranked above a random absence (or pseudo-absence) one (Hanley & McNeil, 1982). The ROC curve shows 'sensitivity' over '1 minus specificity' at a range of threshold probability values, where the sensitivity is the proportion of observed presences correctly predicted by the model and the specificity is the proportion of observed absences (or pseudo-absences) correctly predicted by the model (Raes & ter Steege, 2007). An AUC value above 0.5 indicates that the model performs better than random (Elith et al., 2006; Phillips

& Dudík, 2008). A potentially useful model will have an AUC above 0.75 (Elith, Burgman, & Regan, 2002).

We also applied two additional model evaluation metrics called True Skill Statistic or TSS (Allouche, Tsoar, & Kadmon, 2006) and Sensitivity. A TSS above 0.4 reflects a potentially useful model and above 0.6 is considered to denote good to excellent performance (Landis & Koch, 1977; Tobeña, Prieto, Machete, & Silva, 2016). A model with a sensitivity value above 0.7 is considered a good model (Ochoa-Ochoa, Flores-Villela, & Bezaury-Creel, 2016). We used several evaluation metrics of model performance because none are perfect when true absence data are not available. Significant differences for AUC, TSS and Sensitivity values between Maxent and GAM in 'all year' and each seasonal model were tested using a t-test.

The Jackknife test was used to determine the contribution of each environmental variable to the predicted output, by running the model with and without that variable. The percentage contribution to model gain and the permutation importance of each environmental variable used to build the model could then be assessed (Phillips et al., 2006). The environmental variables with the

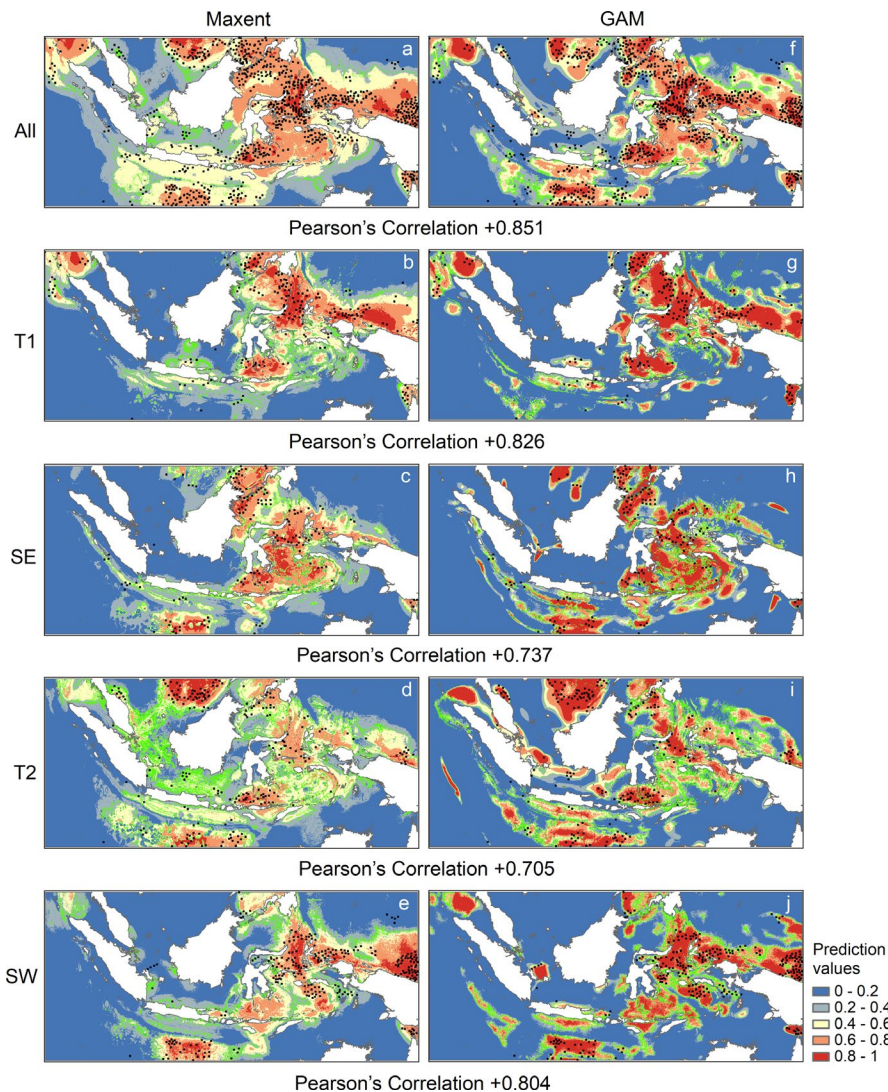
highest training gain are considered to have the highest contribution to the model.

### 3 | RESULTS

#### 3.1 | Historical distribution of sperm whale

The overall predictions per season are visually similar for both models (Figure 3), and align closely with the whaling data. The patchiness of sperm whale distribution is evident in GAM, while the distributions modelled with Maxent were more continuous. Some obvious examples of the patchiness in GAM and the continuity in Maxent can be seen in the northern part of Papua island, and the south-western part of the Sumatera (Figure 3).

The similarity of the spatial distributions between both models is also apparent from the high Pearson's correlation coefficients (Fiedler et al., 2018) between predicted grid values for sperm whale habitat suitability. The Pearson's correlations were high in all models, ranging from 0.705 to 0.851 (Figure 3).



**FIGURE 3** Historical sperm whale distributions as predicted by Maxent and generalized additive model (GAM) for all year and per season: T1 = Transition 1 (March–May), SE = Southeast monsoon (June–August), T2 = Transition 2 (September–November), SW = Southwest monsoon (December–February). The prediction values range between 0 and 1 and are classified into five classes as shown in the bottom right corner, and the actual whale catch data are indicated by black squares (■). The Pearson's correlations between Maxent and GAM prediction values are given. Green lines indicate 10th percentile training presence threshold

### 3.2 | Performance of the models

The performance metrics, AUC, TSS and Sensitivity for Maxent and GAM are given in Table 2. Most model outputs predicted by both Maxent and GAM yielded good discrimination power, with AUC and Sensitivity values >0.75 and TSS values >0.4 indicating model robustness (Table 2). Maxent performed better than GAM, indicated by significantly higher AUC values for all modelled seasons. Sensitivity values of three Maxent outputs also were significantly higher than those of GAM. However, based on TSS values, four of five model outputs did not significantly differ between Maxent and GAM. GAM never performed significantly better than Maxent (Table 2).

As can be seen in Figure 3, the Maxent and GAM distributions fit the sperm whale catch data well. Both models predicted a high suitability of sperm whale habitat in several areas where catches were not reported. Some prominent examples were in the East Banda Sea during the SE monsoon, and in the West Banda Sea during the SW monsoon (Figure 3).

### 3.3 | Seasonal differences in distribution

The season-specific habitat suitability maps were generated with Maxent only because Maxent performed better than GAM (see Section 3.2). By regrouping the whale occurrences per quarter based

**TABLE 2** Performance metrics for Maxent and GAM

Season	Metric	Maxent	Sig <sup>a</sup>	GAM
All	AUC	0.806 ± 0.009	>	0.793 ± 0.011
	TSS	0.461 ± 0.023	ns	0.475 ± 0.020
	Sensitivity	0.888 ± 0.028	>	0.820 ± 0.041
T1	AUC	0.884 ± 0.010	>	0.815 ± 0.020
	TSS	0.575 ± 0.042	>	0.521 ± 0.041
	Sensitivity	0.825 ± 0.054	>	0.769 ± 0.071
SE	AUC	0.899 ± 0.014	>	0.810 ± 0.023
	TSS	0.511 ± 0.046	ns	0.514 ± 0.042
	Sensitivity	0.860 ± 0.045	>	0.802 ± 0.064
T2	AUC	0.874 ± 0.008	>	0.804 ± 0.013
	TSS	0.451 ± 0.036	ns	0.455 ± 0.029
	Sensitivity	0.810 ± 0.056	ns	0.798 ± 0.044
SW	AUC	0.895 ± 0.007	>	0.842 ± 0.020
	TSS	0.609 ± 0.048	ns	0.585 ± 0.035
	Sensitivity	0.839 ± 0.051	ns	0.846 ± 0.040

Abbreviations: All, all year; AUC, area under the curve; GAM, generalized additive model; SE, Southeast monsoon (June–August); SW, Southwest monsoon (December–February); T1, Transition 1 (March–May); T2, Transition 2 (September–November); TSS, True Skill Statistic.

<sup>a</sup>Significant differences in performance metrics (AUC, TSS, Sensitivity) between Maxent and GAM results were tested using t-test. '>' indicates Maxent performed significantly better than GAM, 'ns' indicates no significant difference. GAM never performed significantly better than Maxent.

on Indonesian monsoonal seasons, different seasonal positions of suitable habitat were revealed (Figure 4).

A wide distribution of sperm whales was predicted in the South China Sea during the T2 season, which indicates possible seasonal migrations from surrounding Southeast Asian waters (Figure 4). Limited or no sperm whale distribution was predicted in the Celebes Sea Ground during the SW monsoon, in the Indian Ocean south of Java island during the T1 season, in the Andaman Sea during the SE monsoon and in the Torres Strait during the T2 season (Figure 4). A narrow distribution was also predicted to the north of Papua during the SE monsoon.

The seasonal overlap map (Figure 4f) unveils some of the seasonal variability in sperm whale presence. In the four marked areas, no seasonal differences occur in sperm whale distribution, i.e. in the West Banda Sea, the Molucca Passage Ground, the North Papua and the Sulu Sea Ground, while noticeable seasonal differences in distribution were observed in other areas.

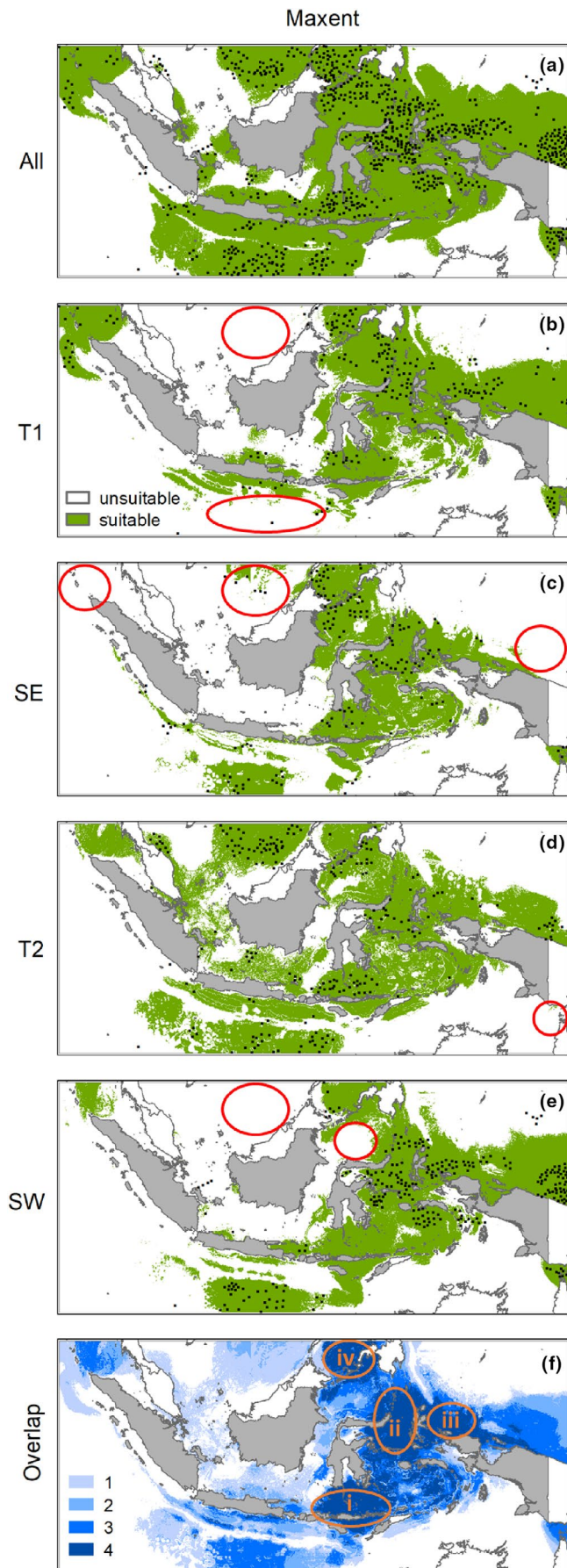
### 3.4 | The importance of environmental variables for sperm whale distribution

Five of nine variables appeared to be important in defining environmental niches for the species. Only distance to coast was retained as the most important variable in all model outputs. The next variable most commonly retained was distance to trough (retained in 5 model outputs), followed by distance to -1,000 and -5,000 m isobaths and trench (each in 4 model outputs). The remaining variables, except bathymetry and distance to seamount, were retained in only 1 or 2 model outputs (Table 3). In general, distance to coast, -1,000 and -5,000 m isobaths, trough and trench were the most important variables that determined sperm whale habitat suitability (Table 3).

Three categories of variables determine sperm whale occurrence in habitat modelling (Figure 5). First, variables that do significantly determine different whale distribution across seasons and available habitat. These variables were distance to -5,000 m isobath, coast, trench and trough (Figure 5f–i). Whales occurred mainly in areas with a distance of 118–151 km from the -5,000 m isobath in most seasons, but farther during the T2 season (median 176 km, the farthest is c. 700 km); closer to the coast (200 km) during the T1 season compared to available habitat (up to 375 km); closer to a trench during the SW monsoon and closer to a trough during the SE monsoon and the T2 seasons.

Second, variables that show the same between all four seasons: distance to -1,000 m isobath, distance to seamount and slope (Figure 5c–e). Sperm whales mostly occurred close to the -1,000 m isobath (44–54 km), close to seamount (115–155 km compared with 180 km of available habitat) and were prevalent in areas with steeper slope (2.5%–3.2% compared to only 1.7% of available habitat).

Third, variables that do not determine sperm whale distribution in any season: bathymetry and distance to shelf (Figure 5a,b). Sperm whales occurred in areas with bathymetry values ranging from -2,151 to -3,047 m (the median in the available habitat is -2,422 m),



**FIGURE 4** Historical seasonal habitat suitability maps of sperm whale in Indonesia generated with the Maxent model (the first five figures). Suitable habitat is indicated in green and is season dependent: (a) All = all year (1761–1920), (b) T1 = Transition 1 (March–May), (c) SE = Southeast monsoon (June–August), (d) T2 = Transition 2 (September–November), (e) SW = Southwest monsoon (December–February). Red circles indicate the areas that differ most among seasons. The last figure (f) shows the degree of seasonal overlap of the distributions in increasingly dark blue colour, and the numbers 1–4 show the number of season(s) with overlapping distribution. Orange circles indicate areas without differences in distribution among seasons. (i) West Banda Sea, (ii) Molucca Passage Ground, (iii) Northern Papua, (iv) Sulu Sea Ground

and were located 48–60 km from shelf (the median in the available habitat is 49 km) effectively occupying the full range of available values for these two variables.

## 4 | DISCUSSION

This study presents the first historical seasonal distribution modelling of sperm whales in Indonesian waters in relation to their habitat preferences. First, we evaluated the model outputs and their performance and caveats. Both Maxent and GAM predict a similar area of sperm whale distribution; and both distributions agree well with the whaling data. This work resulted in three main findings: identifying historical sperm whale seasonal distribution patterns, finding the important variables characterizing sperm whale habitat preferences and confirming the application of SDM based on historical whaling data to support conservation management.

### 4.1 | Performance and caveats of models

The model evaluation metrics indicate that, overall, Maxent and GAM predicted reasonably well the distribution of the sperm whales based on environmental characteristics for ‘all year’ and each season (Figure 3). The majority of the models showed moderate to high discrimination power based on AUC, TSS and sensitivity values. The standard deviation of the average AUC, TSS and sensitivity values was mostly small (<0.05) for each season, either in Maxent or GAM (Table 2). This indicates consistent and reliable model outputs, and appropriate robustness (Tobea et al., 2016). For extended analysis of the model performance and caveats see Texts A1 in Appendix S1.

### 4.2 | Historical sperm whale seasonal distribution patterns

The modelled sperm whale distributions clearly show that important sperm whale habitats were located mainly in the eastern part of the equatorial Indonesian waters. Four areas were identified as important habitats for sperm whales in any season, while they were



**TABLE 3** Relative importance of variables in Maxent and GAM model outputs

Variables	Maxent					GAM				
	All	T1	SE	T2	SW	All	T1	SE	T2	SW
Bathymetry	0.081	0.042	0.076	0.093	0.106	0.021	0.098	0.067	0.037	0.017
d_1000	0.063	<b>0.127</b>	0.085	0.051	0.007	<b>0.164</b>	0.105	<b>0.188</b>	<b>0.187</b>	0.124
d_5000	0.075	0.052	0.073	0.087	<b>0.141</b>	<b>0.163</b>	<b>0.138</b>	0.101	<b>0.157</b>	0.124
d_coast	<b>0.390</b>	<b>0.354</b>	<b>0.214</b>	<b>0.202</b>	<b>0.251</b>	<b>0.324</b>	<b>0.268</b>	<b>0.177</b>	<b>0.190</b>	<b>0.252</b>
d_seamount	0.080	0.051	0.099	0.089	0.056	0.065	0.132	0.114	0.026	0.100
d_shelf	0.069	0.112	0.048	0.052	0.121	0.128	0.047	0.117	0.149	<b>0.154</b>
d_trench	0.049	0.100	0.079	<b>0.231</b>	<b>0.159</b>	0.069	<b>0.136</b>	0.098	0.111	<b>0.153</b>
d_trough	<b>0.091</b>	<b>0.119</b>	<b>0.197</b>	<b>0.143</b>	0.103	0.059	0.069	<b>0.120</b>	0.110	0.055
slope	<b>0.103</b>	0.044	<b>0.129</b>	0.052	0.057	0.007	0.007	0.018	0.033	0.020

Note: Bold shaded numbers are the three most important variables that determine sperm whale habitat. The variable names in the first column are the names given in Table 1.

Abbreviations: All, all year; GAM, generalized additive model; SE, Southeast monsoon (June–August); SW, Southwest monsoon (December–February); T1, Transition 1 (March–May); T2, Transition 2 (September–November).

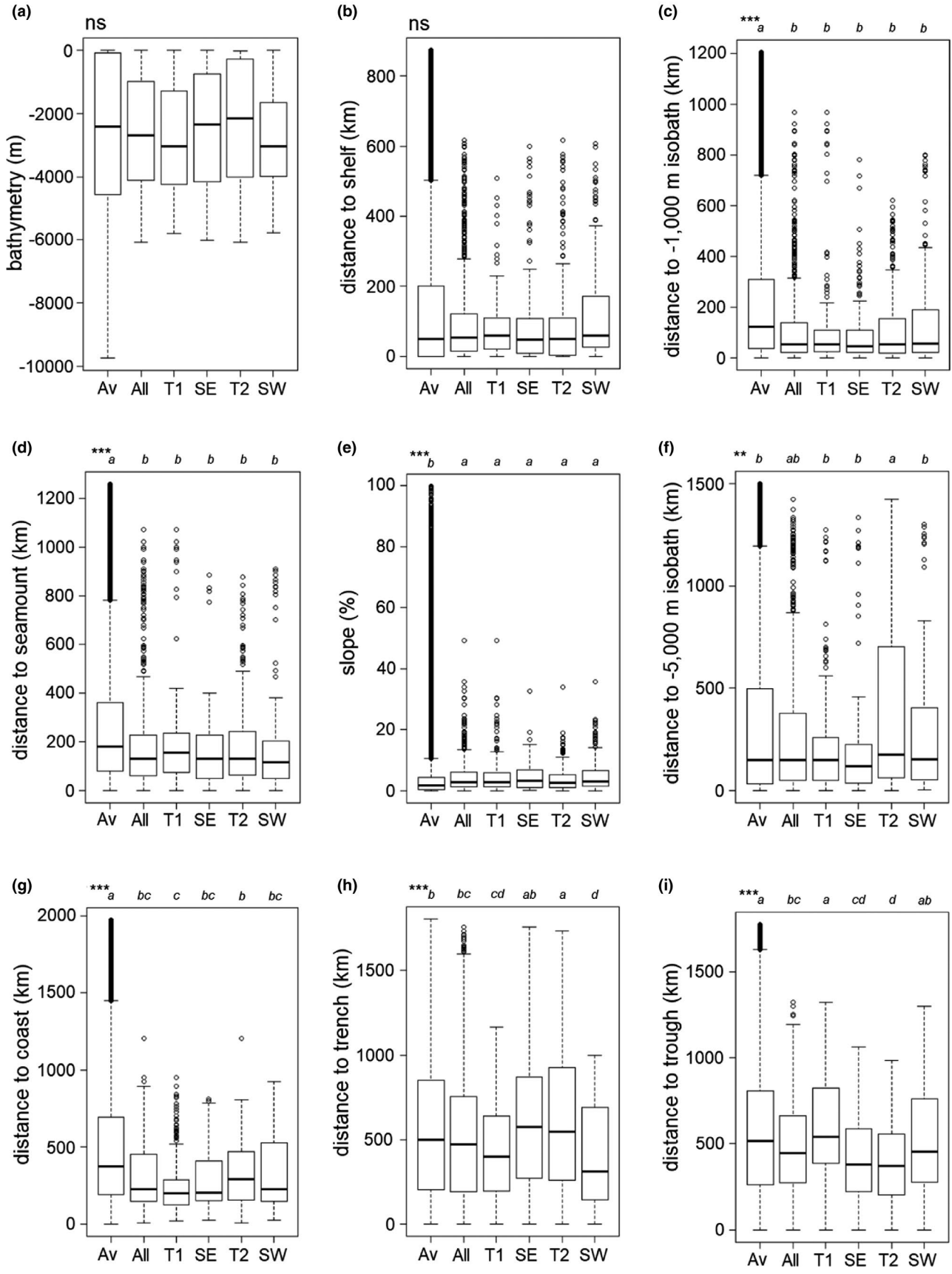
seasonally absent in other locations (Figure 4). Several studies have reported historical year-round presence of sperm whales in the low-latitude areas (Jaquet, Whitehead, & Lewis, 1996; Smith, Reeves, et al., 2012) which is corroborated by our results. Townsend (1935) reported sperm whale occurrences in the form of ‘points’, while this study provides a continuous surface distribution based on the SDMs. Because Townsend plotted the data in two-half-year periods (Figure 1), it seems sperm whales occurred in the north of Borneo island year-round. However, our plotting and modelling based on Indonesian monsoonal seasons reveals they mainly occur there during the T2 season (inter-monsoonal season, September–November; Figures 3d and 4d).

Migrations of the sperm whale are poorly understood compared to other large whales such as baleen whales. While male sperm whales are reported to move to higher latitudes in summer, females (as in our study) remain in equatorial and lower latitudes throughout the year (Evans, 1997) and their seasonal migration routes are unknown (Evans, 1997; Whitehead, 2003, 2009). Availability of different types of prey is suggested to drive these sex-specific movement patterns (Flinn, Trites, Gregr, & Perry, 2002). In the Galapagos, a comparable equatorial area, upwelling events mainly driven by increasing sea surface temperature increased the feeding success of migrating sperm whales (Flinn et al., 2002; Whitehead, 1996). A study on historical trends of sperm whale stranding events in the North Sea suggested that they mainly occurred around migration events and were associated with increased temperature anomalies (Pierce, Santos, Smeenk, Saveliev, & Zuur, 2007). For our study area, during the T2 season the water temperatures at northern latitudes start to drop and the sperm whales were assumed to migrate from other parts of Southeast Asia to the north of Borneo. However, no sperm whales were found in this region during the SW monsoon coinciding with winter in northern latitudes. It therefore seems that sperm whales continue dispersing to other locations. Additionally, the monsoonal regime with seasonally reversing currents leading

to water exchange (Drushka, Sprintall, Gille, & Brodjonegoro, 2010) potentially drives the seasonal upwelling in these waters which enhances local productivity that may attract the whales. This should be corroborated by more targeted future research.

The consistent occurrence of sperm whales in certain areas during specific seasons (Figure 4) suggests site fidelity of the species. This is in accordance with observations reported for other areas (Correia, Tepsich, Rosso, Caldeira, & Sousa-Pinto, 2015). Site fidelity is characteristic of sperm whales (Antunes et al., 2011) and has been supported by genetic studies (Engelhaupt et al., 2009). Our modelling showed that the animals were permanently present in four core areas, the West Banda Sea, the Molucca Passage Ground, northern Papua and the Sulu Sea Ground, while modelled and observed sperm whale distribution varied greatly between seasons in other places. This habitat segregation over the seasons (Figure 4) includes the high density of animals in the north of Borneo Island exclusively during the T2 season and the absence of sperm whales in the Andaman Sea during the SE monsoon and in the Torres Strait during the T2 season. Sperm whales have been known to exhibit seasonal changes in their distribution associated with migratory movements (Smith, Reeves, et al., 2012). Of course, it is important to realize that historical whaling data not only reflect the presence of the animals but could also be influenced by seasonal differences in catch effort (see Whitehead & Jaquet, 1996) and differences in the amount of presence data from Townsend (1935) by season.

The absence of sperm whales at several locations in the modelled distributions (Figure 4, red circles) does not necessarily indicate unsuitable habitat; it could also be due to the absence of (successful) whaling at that location during that season. The same absence could occur nowadays through non-occupancy during times of observation effort (Torres et al., 2013). As the animals may also follow prey abundance, additional local physical and biological processes that cause prey to aggregate (Benson et al., 2011) may account for part of the unexplained deviances of the sperm whale reports from the



**FIGURE 5** Boxplots showing the habitat ranges of sperm whale clustered based on the variables used in the habitat modelling: (a) bathymetry (m); (b) distance to shelf (km); (c) distance to -1,000 m isobath (km); (d) distance to seamount (km); (e) slope (%); (f) distance to -5,000 m isobaths (km); (g) distance to coast (km); (h) distance to trench (km); (i) distance to trough (km). Av = value of a variable in available habitat (whole study area). All = All year, T1 = Transition 1 (March–May), SE = Southeast monsoon (June–August), T2 = Transition 2 (September–November), SW = Southwest monsoon (December–February). Significant differences among seasonal habitat ranges were checked using Kruskal–Wallis (\*\* $p < 0.01$ , \*\*\* $p < 0.001$ , ns = not significant). Wilcoxon post hoc test was applied for multiple comparisons after Kruskal–Wallis testing indicated a significant difference. The same letters indicate the boxplots do not differ significantly

predicted distribution. Unfortunately, variables that directly quantify the productivity (e.g. chlorophyll-a concentrations and sea surface temperature) during the historical whaling period were not available. Exploring the potential use of current dynamic variables for historical sperm whale distribution modelling could be the focus of future research.

### 4.3 | Important variables characterizing sperm whale habitat preferences

The habitat preferences of sperm whales over the seasons are also reflected in the set of important environmental variables for each modelled distribution. A key environmental factor in explaining sperm whale distribution for both Maxent and GAM in all seasons was distance to coast. More productive areas are closer to the coast (Fiedler et al., 2018), which the whales seem to take advantage of. The fact that commercial whaling activity in Western Australia also occurred primarily relatively close to land (Johnson et al., 2016) corroborates our current study results that distance to coast considerably determines their habitat. Prey abundance has indeed been reported to be generally higher in coastal waters (Kanaji et al., 2015), possibly explaining the more suitable habitat closer to the coast, as highlighted for all seasons by both Maxent and GAM.

Distance to trough and trench were the next most important variables in predicting sperm whale distribution (Table 3). These variables reflect increased topographic complexity that may be important in creating physical processes that support enough productivity and associated cephalopods to attract sperm whales (Bouchet, Meeuwig, Salgado Kent, Letessier, & Jenner, 2015; Whitehead, 2009). Prey availability, however, could not be included in the current study as data for this from the historical whaling era are sparsely available. Although oceanographic and biological variables such as sea surface temperature and chlorophyll-a concentration can serve as proxies of prey species density, the availability of such data is also very limited for this period. Unravelling the parameters that determine the species' ecological niche requires fine-scale seasonally dependent habitat use studies, investigating among other things prey distribution, associated cetacean foraging behaviour as well as predator avoidance strategies. Such ecological information can provide a more comprehensive selection of input variables to fine-tune the distribution models.

Distance to -1,000 and -5,000 m isobaths were the next most important variables, particularly for GAM model outputs (Table 3). During the historical and commercial whaling era, most whales were killed offshore in very deep waters with a median depth of

c. -5,000 m (Johnson et al., 2016). The high presence in very deep waters shown in our habitat modelling aligns with the deep diving habits of the sperm whales to catch their primary food source, cephalopods (Whitehead, 2009). In the study area, -5,000 m isobath (Figure A2d in Appendix S1) also coincides with complex topographic features like seamounts (Figure A2f in Appendix S1) and steep slope seabed (Figure A2b in Appendix S1), although seamounts did not significantly influence our models. Distance to the -1,000 m isobath has also been indicated as suitable habitat for foraging sperm whales in other regions for other types of prey (Chua, Lane, Ooi, Tay, & Kubodera, 2019; Hooker, Whitehead, & Gowans, 1999; Kawakami, 1980; Pirotta, Matthiopoulos, MacKenzie, Scott-Hayward, & Rendell, 2011). In Western Australia, sperm whales also occurred at this depth, as commercial whaling locations were reported to have a median depth of around -1,000 m (Johnson et al., 2016). Another toothed whale, the short-finned pilot whale, has also been reported to forage on mesopelagic cephalopods at a depth of c. -1,000 m (Aguilar Soto et al., 2008).

Previous studies have indicated that slope is an environmental factor influencing sperm whale distribution (Hooker et al., 1999; Pirotta et al., 2011), and sperm whales are more prevalent in areas with a steeper slope (Mannocci, Monestiez, Spitz, & Ridoux, 2015). However, we only found a relationship between slope and the modelled whale distributions for the Maxent outputs in 'all year' and SE monsoon. Sperm whales often occur close to the shelf edge (Whitehead, 2009) due to the presence of upwelling-modified waters (Redfern et al., 2017).

Both Maxent and GAM failed to identify bathymetry and seamount as important variables in predicting sperm whale distribution, therefore the two variables were not retained in the model outputs. Other studies, however, did find a relationship between bathymetry and presence of whales in general (Hooker et al., 1999; Pirotta et al., 2011). This is often the case for the distribution of whale species that feed on benthic prey following seafloor structures (Anderwald et al., 2012; Praca, Gannier, Das, & Laran, 2009), instead of deep-diving pelagic species like sperm whale. Our results are also in agreement with Morato et al. (2008) and Tobeña et al. (2016) who found that current sperm whale sighting frequencies were not influenced by distance to seamount.

### 4.4 | How SDM based on historical data can help conservation management

Our models provide spatial maps of the former distribution of sperm whales and their seasonal difference patterns for a vast

area of the equatorial Indonesian marine ecosystem. Our predicted distributions can help to clarify the history of the exploitation of this species by exploring the surviving historical records in the light of these results. Understanding the past distribution of a species is important in order to recognize the past population structure (Knowles, Carstens, & Keat, 2007), the migration paths (Ruegg, Hijmans, & Moritz, 2006) and the niches that are valuable to the species (Nogués-Bravo, 2009). As the current populations of marine animals could capture a fraction of their historical occurrences (Josephson, Smith, & Reeves, 2008), the historical distribution from our study can be used as a guideline to assess present species occurrence and eventually direct ecological investigation for future management support. For instance, marine protected areas and marine spatial planning in Indonesia should take into account species seasonal variability. The historical data may also demonstrate that areas now considered to be abundant in sperm whales only represent a fragment of former densities that used to occur there, as intensive exploitation in the past may have led to a shifting baseline in species densities (Jackson, 2001). It is not certain, however, whether or not contemporary sperm whale populations occupy the same habitats and seasonal migration paths as former populations. To understand this issue, further habitat modelling of modern sperm whale distribution using current data is needed.

The habitat preference assessment performed in this study can help determine suitable habitats for sperm whales in Indonesian waters, where little or no published data on its distribution exist, to inform future marine management and policy. The results of this study can be used to prioritize current sperm whale research and monitoring efforts in the areas predicted to be core sperm whale habitats: West Banda Sea, Molucca Passage Ground and Bird Head Seascape-northern Papua in all seasons (Figure 4). This is especially important as these habitats overlap with some of the busiest shipping routes in the world (Tournadre, 2014), and therefore may require mitigation measures. This will help us to gain a better understanding of the interactions between the animals and their environment, and how the population may respond to human disturbances and changes in their habitat. Our results also show that inter-island areas are very important for sperm whales and thus deserve special attention regarding the management of human activities that may threaten the cetaceans.

This study shows that sperm whales were very likely to be found year-round in equatorial Indonesian waters during the whaling era, with important habitats in the four aforementioned areas. In addition, our results show that sperm whale distribution varied substantially between seasons in other places. The use of historical whaling data in habitat models appears to be a promising approach to accelerate our knowledge on equatorial marine mammal distribution. Although the historical whaling data were not collected in a scientifically systematic way, it still provides important information in a cost-effective manner for a region for which knowledge is sparse. The habitat models synthesized which environmental factors are the primary determinants of habitat suitability for sperm whales from the historical whaling data. Incorporating such an unconventional

dataset into habitat models yields benefits for scientific understanding of sperm whale habitat preferences as well as former sperm whale occurrences. This information can support management decisions on conservation measures for endangered large whales such as sperm whales.

## ACKNOWLEDGEMENTS

This work was financially supported by the Indonesia Endowment Fund for Education (LPDP) Scholarship from the Ministry of Finance of the Republic of Indonesia for the first author (contract number: PRJ-482/LPDP.3/2017). The authors would like to thank Neftalí Sillero for his valuable input on habitat modelling, and the New Bedford Whaling Museum for permitting the historical whaling data and maps to be used in this study.

## CONFLICT OF INTEREST

The authors declare no conflict of interest regarding the results presented in this study.

## DATA AVAILABILITY STATEMENT

The original data for this study are available from Townsend (1935) with the permission from the New Bedford Whaling Museum. The processed data in raster format are available at Dryad (<https://doi.org/10.5061/dryad.18931zct6>).

## ORCID

Achmad Sahri  <https://orcid.org/0000-0002-3778-7244>

Mochamad I. H. Putra  <https://orcid.org/0000-0001-9202-7857>

Putu L. K. Mustika  <https://orcid.org/0000-0001-5157-4635>

Albertinka J. Murk  <https://orcid.org/0000-0002-4881-4236>

## REFERENCES

- Aguilar Soto, N., Johnson, M. P., Madsen, P. T., Díaz, F., Domínguez, I., Brito, A., & Tyack, P. (2008). Cheetahs of the deep sea: Deep foraging sprints in short-finned pilot whales off Tenerife (Canary Islands). *Journal of Animal Ecology*, 77(5), 936–947. <https://doi.org/10.1111/j.1365-2656.2008.01393.x>
- 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>
- Anderwald, P., Evans, P. G. H., Dyer, R., Dale, A., Wright, P. J., & Hoelzel, A. R. (2012). Spatial scale and environmental determinants in minke whale habitat use and foraging. *Marine Ecology Progress Series*, 450, 259–274. <https://doi.org/10.3354/meps09573>
- Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J., & Rendell, L. (2011). Individually distinctive acoustic features in sperm whale codas. *Animal Behaviour*, 81(4), 723–730. <https://doi.org/10.1016/j.anbehav.2010.12.019>
- Azzellino, A., Panigada, S., Lanfredi, C., Zanardelli, M., Airoidi, S., & Notarbartolo di Sciara, G. (2012). Predictive habitat models for managing marine areas: Spatial and temporal distribution of marine mammals within the Pelagos Sanctuary (Northwestern Mediterranean sea). *Ocean and Coastal Management*, 67, 63–74. <https://doi.org/10.1016/j.ocecoaman.2012.05.024>
- Benson, S. R., Eguchi, T., Foley, D. G., Forney, K. A., Bailey, H., Hitipeuw, C., ... Dutton, P. H. (2011). Large-scale movements and high-use



- areas of western Pacific leatherback turtles, *Dermochelys coriacea*. *Ecosphere*, 2(7), 1–27. <https://doi.org/10.1890/ES11-00053.1>
- Best, P. (1979). Chapter 7: Social organization in sperm whales, *Physeter macrocephalus*. In H. Winn & B. Olla (Eds), *Behavior of marine animals*. Vol. 3: *Cetaceans*. New York, NY; London, UK: Plenum Press.
- Bouchet, P. J., Meeuwig, J. J., Salgado Kent, C. P., Letessier, T. B., & Jenner, C. K. (2015). Topographic determinants of mobile vertebrate predator hotspots: Current knowledge and future directions. *Biological Reviews*, 90(3), 699–728. <https://doi.org/10.1111/brv.12130>
- Breen, P., Brown, S., Reid, D., & Rogan, E. (2016). Modelling cetacean distribution and mapping overlap with fisheries in the northeast Atlantic. *Ocean and Coastal Management*, 134, 140–149. <https://doi.org/10.1016/j.ocecoaman.2016.09.004>
- Brito, J. C., Acosta, A. L., Álvares, F., & Cuzin, F. (2009). Biogeography and conservation of taxa from remote regions: An application of ecological-niche based models and GIS to North-African canids. *Biological Conservation*, 142(12), 3020–3029. <https://doi.org/10.1016/j.biocon.2009.08.001>
- Carroll, G., Hedley, S., Bannister, J., Ensor, P., & Harcourt, R. (2014). No evidence for recovery in the population of sperm whale bulls off Western Australia, 30 years post-whaling. *Endangered Species Research*, 24(1), 33–43. <https://doi.org/10.3354/esr00584>
- Chambers, J. M., & Hastie, T. J. (2017). *Statistical models in S*. New York, NY: CRC Press. <https://doi.org/10.1201/9780203738535>
- Chua, M. A. H., Lane, D. J. W., Ooi, S. K., Tay, S. H. X., & Kubodera, T. (2019). Diet and mitochondrial DNA haplotype of a sperm whale (*Physeter macrocephalus*) found dead off Jurong Island, Singapore. *PeerJ*, 7, e6705. <https://doi.org/10.7717/peerj.6705>
- Correia, A. M., Tepsich, P., Rosso, M., Caldeira, R., & Sousa-Pinto, I. (2015). Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems*, 143, 73–85. <https://doi.org/10.1016/j.jmarsys.2014.10.016>
- Dalla Rosa, L., Ford, J. K. B., & Trites, A. W. (2012). Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. *Continental Shelf Research*, 36, 89–104. <https://doi.org/10.1016/j.csr.2012.01.017>
- Drushka, K., Sprintall, J., Gille, S. T., & Brodjonegoro, I. (2010). Vertical structure of Kelvin waves in the Indonesian throughflow exit passages. *Journal of Physical Oceanography*, 40(9), 1965–1987. <https://doi.org/10.1175/2010JPO4380.1>
- Edrén, S. M. C., Wisz, M. S., Teilmann, J., Dietz, R., & Söderkvist, J. (2010). Modelling spatial patterns in harbour porpoise satellite telemetry data using maximum entropy. *Ecography*, 33(4), 698–708. <https://doi.org/10.1111/j.1600-0587.2009.05901.x>
- Elith, J., Burgman, M. A., & Regan, H. M. (2002). Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological Modelling*, 157(2–3), 313–329. [https://doi.org/10.1016/S0304-3800\(02\)00202-8](https://doi.org/10.1016/S0304-3800(02)00202-8)
- Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Engelhaupt, D., Rus hoelzel, A., Nicholson, C., Frantzis, A., Mesnick, S., Gero, S., ... Mignucci-giannoni, A. A. (2009). Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (*Physeter macrocephalus*). *Molecular Ecology*, 18(20), 4193–4205. <https://doi.org/10.1111/j.1365-294X.2009.04355.x>
- Evans, P. G. H. (1997). Ecology of sperm whales (*Physeter macrocephalus*) in the eastern North Atlantic, with special reference to sightings & strandings records from the British Isles. *Bulletin De l'institut Royal Des Sciences Naturelles De Belgique Biologie*, 67, 37–46.
- Fiedler, P. C., Redfern, J. V., Forney, K. A., Palacios, D. M., Sheredy, C., Rasmussen, K., ... Ballance, L. T. (2018). Prediction of large whale distributions: A comparison of presence-absence and presence-only modeling techniques. *Frontiers in Marine Science*, 5, 1–15. <https://doi.org/10.3389/fmars.2018.00419>
- Flinn, R. D., Trites, A. W., Gregr, E. J., & Perry, R. I. (2002). Diets of fin, sei, and sperm whales in British Columbia: An analysis of commercial whaling records, 1963–1967. *Marine Mammal Science*, 18(3), 663–679. <https://doi.org/10.1111/j.1748-7692.2002.tb01065.x>
- Gregr, E. J. (2011). Insights into North Pacific right whale *Eubalaena japonica* habitat from historic whaling records. *Endangered Species Research*, 15(3), 223–239. <https://doi.org/10.3354/esr00381>
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29–36. <https://doi.org/10.1148/radiology.143.1.7063747>
- Harris, P. T., Macmillan-Lawler, M., Rupp, J., & Baker, E. K. (2014). Geomorphology of the oceans. *Marine Geology*, 352, 4–24. <https://doi.org/10.1016/j.margeo.2014.01.011>
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models*. London, UK: Chapman and Hall. <https://doi.org/10.1201/9780203753781>
- Hooker, S. K., Whitehead, H., & Gowans, S. (1999). Marine protected area design and the spatial and temporal distribution of cetaceans in a submarine canyon. *Conservation Biology*, 13(3), 592–602. <https://doi.org/10.1046/j.1523-1739.1999.98099.x>
- Jackson, J. B. C. (2001). What was natural in the coastal oceans? *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), 5411–5418. <https://doi.org/10.1073/pnas.091092898>
- Jaquet, N., Whitehead, H., & Lewis, M. (1996). Coherence between 19th century sperm whale distributions and satellite-derived pigments in the tropical Pacific. *Marine Ecology Progress Series*, 145(1–3), 1–10. <https://doi.org/10.3354/meps145001>
- Johnson, C. M., Beckley, L. E., Kobryn, H., Johnson, G. E., Kerr, I., & Payne, R. (2016). Crowdsourcing modern and historical data identifies sperm whale (*Physeter macrocephalus*) habitat offshore of South-Western Australia. *Frontiers in Marine Science*, 3, 1–16. <https://doi.org/10.3389/fmars.2016.00167>
- Josephson, E., Smith, T. D., & Reeves, R. R. (2008). Historical distribution of right whales in the North Pacific. *Fish and Fisheries*, 9(2), 155–168. <https://doi.org/10.1111/j.1467-2979.2008.00275.x>
- Kanaji, Y., Okazaki, M., Kishiro, T., & Miyashita, T. (2015). Estimation of habitat suitability for the southern form of the short-finned pilot whale (*Globicephala macrorhynchus*) in the North Pacific. *Fisheries Oceanography*, 24(1), 14–25. <https://doi.org/10.1111/fog.12074>
- Kawakami, T. (1980). A review of sperm whale food. *Scientific Reports of the Whales Research Institute*, 32(32), 199–218.
- Knowles, L. L., Carstens, B. C., & Keat, M. L. L. (2007). Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology*, 17(11), 940–946. <https://doi.org/10.1016/j.cub.2007.04.033>
- Lambert, C., Mannocci, L., Lehodey, P., & Ridoux, V. (2014). Predicting cetacean habitats from their energetic needs and the distribution of their prey in two contrasted tropical regions. *PLoS One*, 9(8), e105958. <https://doi.org/10.1371/journal.pone.0105958>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Mannocci, L., Monestiez, P., Spitz, J., & Ridoux, V. (2015). Extrapolating cetacean densities beyond surveyed regions: Habitat-based

- predictions in the circumtropical belt. *Journal of Biogeography*, 42(7), 1267–1280. <https://doi.org/10.1111/jbi.12530>
- Morato, T., Varkey, D. A., Damaso, C., Machete, M., Santos, M., Prieto, R., ... Santos, R. S. (2008). Evidence of a seamount effect on aggregating visitors. *Marine Ecology Progress Series*, 357, 23–32. <https://doi.org/10.3354/meps07269>
- Moura, A. E., Sillero, N., & Rodrigues, A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecologica*, 38, 24–32. <https://doi.org/10.1016/j.actao.2011.08.006>
- Nogués-Bravo, D. (2009). Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, 18(5), 521–531. <https://doi.org/10.1111/j.1466-8238.2009.00476.x>
- Ochoa-Ochoa, L. M., Flores-Villela, O. A., & Bezaury-Creel, J. E. (2016). Using one vs. many, sensitivity and uncertainty analyses of species distribution models with focus on conservation area networks. *Ecological Modelling*, 320, 372–382. <https://doi.org/10.1016/j.ecolmodel.2015.10.031>
- Padalia, H., Srivastava, V., & Kushwaha, S. P. S. (2014). Modeling potential invasion range of alien invasive species, *Hyptis suaveolens* (L.) Poit. in India: Comparison of MaxEnt and GARP. *Ecological Informatics*, 22, 36–43. <https://doi.org/10.1016/j.ecoinf.2014.04.002>
- Pearce, J. L., & Boyce, M. S. (2006). Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, 43(3), 405–412. <https://doi.org/10.1111/j.1365-2664.2005.01112.x>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(2), 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Pierce, G. J., Santos, M. B., Smeenk, C., Saveliev, A., & Zuur, A. F. (2007). Historical trends in the incidence of strandings of sperm whales (*Physeter macrocephalus*) on North Sea coasts: An association with positive temperature anomalies. *Fisheries Research*, 87(2–3), 219–228. <https://doi.org/10.1016/j.fishres.2007.06.001>
- Pirotta, E., Matthiopoulos, J., MacKenzie, M., Scott-Hayward, L., & Rendell, L. (2011). Modelling sperm whale habitat preference: A novel approach combining transect and follow data. *Marine Ecology Progress Series*, 436, 257–272. <https://doi.org/10.3354/meps09236>
- Praca, E., Gannier, A., Das, K., & Laran, S. (2009). Modelling the habitat suitability of cetaceans: Example of the sperm whale in the northwestern Mediterranean Sea. *Deep-Sea Research Part I: Oceanographic Research Papers*, 56(4), 648–657. <https://doi.org/10.1016/j.dsr.2008.11.001>
- Raes, N., & ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, 30(5), 727–736. <https://doi.org/10.1111/j.2007.0906-7590.05041.x>
- Redfern, J. V., Ferguson, M. C., Becker, E. A., Hyrenbach, K. D., Good, C., Barlow, J., ... Werne, F. (2006). Techniques for cetacean – Habitat modeling. *Marine Ecology Progress Series*, 310, 271–295. <https://doi.org/10.3354/meps310271>
- Redfern, J. V., Moore, T. J., Fiedler, P. C., de Vos, A., Brownell, R. L., Forney, K. A., ... Ballance, L. T. (2017). Predicting cetacean distributions in data-poor marine ecosystems. *Diversity and Distributions*, 23(4), 394–408. <https://doi.org/10.1111/ddi.12537>
- Reeves, R. R., Smith, T. D., Josephson, E. A., Clapham, P. J., & Woolmer, G. (2004). Historical observations of humpback and blue whales in the North Atlantic Ocean: Clues to migratory routes and possibly additional feeding grounds. *Marine Mammal Science*, 20(4), 774–786. <https://doi.org/10.1111/j.1748-7692.2004.tb01192.x>
- Ruegg, K. C., Hijmans, R. J., & Moritz, C. (2006). Climate change and the origin of migratory pathways in the Swainson's thrush, *Catharus ustulatus*. *Journal of Biogeography*, 33(7), 1172–1182. <https://doi.org/10.1111/j.1365-2699.2006.01517.x>
- Schlacher, T. A., Rowden, A. A., Dower, J. F., & Consalvey, M. (2010). Seamount science scales undersea mountains: New research and outlook. *Marine Ecology*, 31, 1–13. <https://doi.org/10.1111/j.1439-0485.2010.00396.x>
- Smith, J. N., Grantham, H. S., Gales, N., Double, M. C., Noad, M. J., & Paton, D. (2012). Identification of humpback whale breeding and calving habitat in the Great Barrier Reef. *Marine Ecology Progress Series*, 447, 259–272. <https://doi.org/10.3354/meps09462>
- Smith, T. D., Reeves, R. R., Josephson, E. A., & Lund, J. N. (2012). Spatial and seasonal distribution of American whaling and whales in the age of sail. *PLoS One*, 7(4), e34905. <https://doi.org/10.1371/journal.pone.0034905>
- Taylor, B. L., Baird, R., Barlow, J., Dawson, S. M., Ford, J., Mead, J. G., Pitman, R. L. (2008). *Physeter macrocephalus*. The IUCN Red List of Threatened Species 2008: e.T41755A10554884. Retrieved from <http://www.iucnredlist.org/details/41755/0>
- Thorne, L. H., Johnston, D. W., Urban, D. L., Tyne, J., Bejder, L., Baird, R. W., ... Chapla Hill, M. (2012). Predictive modeling of spinner dolphin (*Stenella longirostris*) resting habitat in the main Hawaiian Islands. *PLoS One*, 7(8), e43167. <https://doi.org/10.1371/journal.pone.0043167>
- 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. Retrieved from <https://www.frontiersin.org/article/10.3389/fmars.2016.00202>
- Torres, L. G., Smith, T. D., Sutton, P., Macdiarmid, A., Bannister, J., & Miyashita, T. (2013). From exploitation to conservation: Habitat models using whaling data predict distribution patterns and threat exposure of an endangered whale. *Diversity and Distributions*, 19(9), 1138–1152. <https://doi.org/10.1111/ddi.12069>
- Tournadre, J. (2014). Anthropogenic pressure on the open ocean: The growth of ship traffic revealed by altimeter data analysis. *Geophysical Research Letters*, 41(22), 7924–7932. <https://doi.org/10.1002/2014GL061786>
- Townsend, C. H. (1935). The distribution of certain whales as shown by logbook records of American whaleships. *Zoologica*, 19, 1–50 (plus maps).
- Virgili, A., Racine, M., Authier, M., Monestiez, P., & Ridoux, V. (2017). Comparison of habitat models for scarcely detected species. *Ecological Modelling*, 346, 88–98. <https://doi.org/10.1016/j.ecolmodel.2016.12.013>
- Whitehead, H. (1996). Variation in the feeding success of sperm whales: Temporal scale, spatial scale and relationship to migrations. *The Journal of Animal Ecology*, 65(4), 429–438. <https://doi.org/10.2307/5778>
- Whitehead, H. (2002). Estimates of the current global population size and historical trajectory for sperm whales. *Marine Ecology Progress Series*, 242, 295–304. <https://doi.org/10.3354/meps242295>
- Whitehead, H. (2003). *Sperm whales: Social evolution in the ocean*. Chicago, IL: The University of Chicago Press.
- Whitehead, H. (2009). Sperm whale: *Physeter macrocephalus*. In W. F. Perrin, B. Würsig, & J. G. M. Thewissen (Eds), *Encyclopedia of marine mammals* (pp. 1091–1097). <https://doi.org/10.1016/B978-0-12-373553-9.00248-0>
- Whitehead, H., & Jaquet, N. (1996). Are the charts of Maury and Townsend good indicators of sperm whale distribution and seasonality? Forty-Sixth Report of the International Whaling Commission, pp. 643–647.
- Yen, P. P. W., Sydeman, W. J., & Hyrenbach, K. D. (2004). Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: Implications for trophic transfer and conservation. *Journal of Marine Systems*, 50(1–2), 79–99. <https://doi.org/10.1016/j.jmarsys.2003.09.015>



## BIOSKETCH

**Achmad Sahri** is a marine ecologist and habitat modeller with research focus on the distribution, habitat preference and habitat use patterns of cetaceans for better understanding of their biology and ecology for conservation purposes. He is currently a researcher in Marine Animal Ecology Group, Wageningen University, The Netherlands. The group studies how marine animals adapt in response to a changing environment and the consequences for the ecosystem, its sustainable use and its management. Within this group, his works mainly fall under the research theme of Marine Conservation Management and Planning.

Author contributions: All authors conceived the ideas; A.S and M.I.H.P collected the data; A.S. analysed the data; A.S. and A.J.M. led the writing; All authors discussed the results and commented on the manuscript.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Sahri A, Putra MIH, Mustika PLK, Murk AJ. A treasure from the past: Former sperm whale distribution in Indonesian waters unveiled using distribution models and historical whaling data. *J Biogeogr.* 2020;00:1–15. <https://doi.org/10.1111/jbi.13931>