



Multimodal deep learning for cetacean distribution modeling of fin whales (*Balaenoptera physalus*) in the western Mediterranean Sea

D. Cazau¹ · P. Nguyen Hong Duc² · J.-N. Druon³ · S. Matwins⁴ · R. Fablet⁵

Received: 24 June 2020 / Revised: 18 June 2021 / Accepted: 22 June 2021 /

Published online: 26 July 2021

© The Author(s), under exclusive licence to Springer Science+Business Media LLC, part of Springer Nature 2021

Abstract

Cetacean Distribution Modeling (CDM) is used to quantify mobile marine species distributions and densities. It is essential to better understand and protect whales and their relatives. Current CDM approaches often fail in capturing general species-environment relationships, which would be valid within a broader range of environmental conditions that characterize the surveyed regions. This paper aims at investigating the usefulness of deep learning based schemes, namely multi-task and transfer learning, in CDM. Co-training of a stochastic presence-background model on a classification task and a deterministic rule-based model on a regression task was performed. Whale presence-only records were used for the first task, and index outputs of a feeding habitat occurrence model for the second one. This new approach has been experimented through the study case of fin whales in the western Mediterranean Sea. To evaluate our approach, a new metric called True Positive rate per unit of Surface Habitat (TPSH) and an original multimodal fully-connected neural networks were developed. A Generalized Additive Model (GAM)—a standard CDM method—was also used as a reference for performance. Results show that our multi-task learning model improves both the feeding habitat model by 10.8% and data-driven models such as GAM by 16.5% on our TPSH metric in relative terms, revealing a higher accuracy of our approach in estimating whale presence. Such trends in results have been further supported by the use of two other independent datasets that forced models to generalize beyond their training dataset of species-environment relationships. Performance could be further improved by adopting more optimal thresholds as observed from Receiver Operating Characteristic curves, e.g. the multi-task learning model could reach absolute gains up to 10% in the median of the True Positive Rate while maintaining its habitat spatial spreading. Globally, our work confirmed our working hypothesis that expert information on whale behaviour represent a good knowledge base for model generalization. This result can be further improved by a concurrent learning of more local species-environment relationships from in-situ presence data.

Editors: Dino Ienco, Thomas Corpetti, Minh-Tan Pham, Sebastien Lefevre, Roberto Interdonato.

Extended author information available on the last page of the article

1 Introduction

Cetacean Distribution Models (CDM) have become a prominent technique in conservation biogeography and are increasingly used as prediction tools for environmental change forecasts (Franklin, 2010). The goal of CDM is to infer the spatial distribution of a given cetacean species based on a function that takes a matrix of ocean variables (e.g. temperature, currents, salinity, and chlorophyll levels) of a given location as input and outputs an estimate of the abundance or a presence index of the species. An accurate knowledge of the spatial distribution of species is actually of crucial importance for many concrete scenarios including the ocean management, the preservation of rare and/or endangered species, the surveillance of alien invasive species, the measurement of human impact or climate change on species, etc. Especially, a current need has raised for larger datasets of CDM data at broader scales to achieve for example Good Environmental Status within the Marine Strategy Framework Directive (European Commission, 2008, 2017), and implement spatially explicit conservation measures for Ecologically and Biologically Significant Marine Areas (Dunn, 2014). A common modeling approach in CDM relies on the niche theory (Druon et al. 2012), which links environment with ecology to select a limited set of predictors and improve model interpretability. In other words, CDM relies on behavioural assumptions that cetaceans, and in particular fin whales, have regular and unusual seasonal movements subject to complex set of factors affecting their habitat use, notably in response to environmental fluctuations (Notarbartolo di Sciara et al., 2016).

The Mediterranean Sea, our area of interest, is unique among large sea basins because it constitutes a miniature ocean with contrasted physical, climatic and biological characteristics (Bethoux et al., 1999), and supports a highly diverse marine fauna including large mobile animals such as cetaceans (Coll 2010; Notarbartolo di Sciara et al., 2016). The wide range of cetaceans across the Mediterranean Sea (e.g., fin whale and sperm whale (Aïssi et al., 2008; Panigada et al., 2017)), combined with their vulnerability to the multiple anthropogenic pressures (e.g., ship strike, noise disturbance, pollution exposure) (Pesante et al., 2006; David et al., 2011; Castellote et al., 2012; Cañadas and Vazquez 2014), stress the need to develop predictive models to map their densities throughout this Sea. One particular endangered species is the fin whale (*Balaenoptera physalus* L. 1758), which is the only commonly observed mysticete in the Mediterranean Sea. Although its presence has been documented since ancient times (Notarbartolo di Sciara et al., 2003), this species is among the least known of all cetaceans in the Mediterranean Sea, mainly because of its pelagic distribution (e.g. unknown population size, unknown winter breeding grounds, unknown presence in the eastern basin, relatively unknown behavioural response to maritime traffic and noise) (Notarbartolo di Sciara et al., 2016). The Mediterranean fin whales' population size is small and possibly decreasing. Their survival rate is low and the pressure from many threats deriving from human activities such as vessel traffic, noise, chemical pollution and likely climate change, is high, as reviewed in Notarbartolo di Sciara et al. (2016). This recent review also recalls that the status of the Mediterranean fin whales' population raises considerable concern and conservation measures should be urgently implemented. Nowadays, the degree of protection of fin whales has raised, with a vulnerable ACCOBAMS/IUCN global conservation status¹ in the Mediterranean sea, and endangered

¹ <https://accobams.org/species/garampus/>.

one worldwide. The Mediterranean fin whale subpopulation is estimated to be of few thousands (Notarbartolo di Sciara et al., 2016).

Also, fin whales are among the fastest cetaceans in the world; they can sustain speeds between 37 km/h (23 mph) and 41 km/h (25 mph) and bursts up to 46 km/h (29 mph),² and can cover > 150 km daily, making its behaviour sensitive to meso-scale oceanographic features. Whales appear to be fairly widespread during winter, whereas in summer favourable feeding habitat is dramatically reduced, concentrating mostly in the western Ligurian Sea and Gulf of Lions (Druon et al., 2012; Notarbartolo di Sciara et al., 2016; Panigada et al., 2017). Analyses of feeding habitat therefore represent a key element of protection measures. Furthermore, because fin whales have a complex behaviour, notably involving aggregations for reproduction and long-distance communication, it is of utmost importance that these habitat models be sufficiently elaborated to account for or to detect some of this complexity. This is why a data-driven machine learning based approach is particularly relevant to be compared to more model-based approaches used in habitat modeling.

Gap analyses of observation data on mobile marine species have led to identify of geographic areas and seasons with important knowledge gaps (Mannocci et al., 2018). For example, a global gap analysis of line transect surveys used to derive abundance estimates and habitat-based density models of cetaceans has revealed large geographic gaps in the Southern Hemisphere and seasonal gaps in non-summer months (Kaschner et al., 2012). Especially, Mediterranean waters characterised by comparatively warmer temperatures, lower productivity and higher eddy activity were poorly surveyed for cetaceans. This raised the prospect that cetacean density models fitted to environmental covariates would have to be extrapolated in order to provide predictions for the entire Mediterranean Sea. This problem is all the more critical as a fundamental assumption in CDM calibration is the unbiased sampling of available conditions in the environmental space (Pearce & Boyce, 2006), whereas opportunistic data, not collected following strict random or systematic sampling methods (Yackulic et al., 2012), are generally used. Presence-only CDMs are particularly vulnerable to sampling bias, when compared to presence–absence CDMs, since background points are generally randomly drawn across the area of interest instead of representing the environmental conditions of the sampled area (Phillips 2009; Yackulic et al., 2012). The resulting predictions from these naïve CDMs reflect the joint distributions of species probability of presence and sampling effort (Elith et al., 2011; Guillera-Arroita, 2015). Note however that absence for fin whales in particular can include substantial missed presence, and that species–environment relationships are unknown in unsampled and/or extreme environments (Elith and Leathwick 2009; Elith et al., 2010). Hence, extrapolation in environmental space using presence–absence can lead to highly uncertain predictions. Furthermore, the selected model strongly influences predictions, particularly when predicting outside the sampled range of the covariate data (Pearson et al., 2006; Mainali et al., 2015). Overall, current CDM methods generally fail in capturing the general species–environment relationships characterizing cetacean habitats (Mannocci et al., 2015), which would be valid within a broader range of environmental conditions that characterize the surveyed regions.

In this current work, we wish to address this so-called extrapolation problem with an original approach based on modern deep learning based schemes. Especially, multi-task and transfer learning schemes were used to co-train two different models on two

² From https://animaldiversity.org/site/accounts/information/Balaenoptera_physalus.html.

different tasks: (1) a stochastic presence-background model on a classification task, using whale presence-only records and pseudo-absences, and (2) a deterministic rule-based model on a regression task, using the index outputs of a feeding habitat occurrence model. Using this hybrid modeling approach, an improved fusion of knowledge extracted from these two tasks is expected, in order to improve accuracy of current models in performing whale presence estimation. Our model (1) is based on multimodal deep learning architectures, while our model (2) is the Feeding Habitat Occurrence (FHO) model by Druon et al. (2012). They will be described in Sect. 3. Experimental results and discussion have been put in Sect. 5.

Within the CDM literature based on productivity fronts as feeding proxy, our study will draw from two broad families of methods that are rarely processed together. A first one is the niche model, where observation data are processed as a “whole” to provide summary statistics over large spatio-temporal scales used to calibrate parameters of explicit “behavioral rules”. For example, it is known that chlorophyll-a fronts are productive oceanic features that attract large predators like fin whales, a knowledge that is at the core of Druon et al. (2012)’s favourable feeding habitat of fin whale (more details will be provided at Sect. 3.2). But note that the favourable feeding habitat is centred on the productivity fronts and not strictly on the local environment surrounding each observation. Indeed, being active long enough (from weeks to months) to allow the development of mesozooplankton populations (Druon, 2019), productivity fronts were shown to attract a large variety of top predators (Druon et al., 2012, 2017, 2021a; Olson et al., 1994; Polovina et al., 2001). Chlorophyll-a fronts are thus thought to be explanatory (or proxy) variables for fin whale distribution (Panigada et al., 2017). A second family of CDM methods involve pure data-driven approaches, where different statistical algorithms (e.g. Derville et al., 2018 benchmarked GAM, SVM and MaxEnt methods) are first used to learn specific patterns in the environmental variables related to the presence of a whale or to its “supposed” absence. Then, such an algorithm can then predict a presence versus background probability at new temporal and spatial points. Our work aims at proposing a new possible CDM framework that consists in better combining knowledge both from rule-based expert models and data-driven learning.

Conceptually, our work connects to the literature of multimodal machine learning (Baltrusaitis et al., 2017). This discipline aims to build models that can process and relate information from multiple modalities. A modality refers to the way in which something happens or is experienced and a research problem is characterized as multimodal when it includes multiple such modalities. Multimodal machine learning models, generally trained end-to-end from large amount of data with multiple modalities, aim to extract a joint representation integrating all modalities. Thanks to their representation power, models based on deep neural networks have been intensively explored through different applications such as audio-visual emotion classification (Wöllmer et al., 2010), gesture recognition (Neverova et al., 2016), affect analysis (Kahou et al., 2016), and video description generation (Jin & Liang, 2016). For Earth observation, first works using multimodal deep learning have mostly focused on the question of fusing multi-scale and multi-resolution data. Most works have dealt with various types of satellite images, e.g. High Spatial Resolution from the Sentinel-2 program and Very High Spatial Resolution SPOT6/7 satellite images in Benedetti et al. (2018). Closer to our experimental setup, Rao et al. (2017) have used multimodal deep learning to build a semantic mapping of unseen or unfamiliar benthic habitats using scarce in-situ high-resolution images from an autonomous underwater vehicles (AUV) and large-scale low-resolution bathymetric maps from shipborne multibeam sonar.

With the recent revival popularity of Deep Neural Networks (DNN) (Hinton et al., 2006), a few studies have begun to explore the use of DNNs for species distribution modeling. One layered-neural network had already been tested in the context of species distribution modeling (Thuiller, 2003). More recently, Chen et al. (2017) jointly modeled the distribution of multiple bird species while also simultaneously learning the shared habitat preferences among species. Their Deep Multi-Species Embedding framework considers two heterogeneous contextual information feature sets (environmental features and inter-species relationships). It uses a deep neural network to learn a representation from environmental covariates where species interactions and shared habitat preferences among species could be described. They show that their model better deals with the fusion of heterogeneous multi-scale features, as the environmental features used in the model describe habitat characteristics at a much coarser spatial resolution than that of the inter-species interactions, this model formulation can be seen as a multi-scale approach that shares information at coarse scales while simultaneously allowing fine-scale variabilities between species. Botella et al. (2018) applied DNNs on plant distribution modeling. They show how mutualizing model features for many species prevent deep NN to overfit and finally allow them to reach a better predictive performance than the MAXENT baseline. Our current work proposes the first use of such approaches based on multimodal deep learning to CDM, so as to validate its applicability to this discipline, before developing more complex modeling features such as multi-scale fusion.

2 Observation data

2.1 Fin whale presence data

The present study compiled one of the most extensive databases of geo-located presence data of fin whales in the Mediterranean Sea, as detailed in Supplementary Table 1. Presence data consist of samples of locations with known presences (at-sea sightings and GPS-precision electronic tagging). These data are mainly derived from scientific surveys with dedicated sighting protocols along fixed-line transects using ferries as platforms of observation or aboard vessels implementing the line-transect protocol (Buckland et al., 2001) and mostly travelling at speeds of 5–6 knots (i.e. 9–11 km h⁻¹). Other presence data were collected using less systematic data collection methods (Panigada et al., 2005) or on random routes, at about the same speed. Several opportunistic surveys were also used (e.g. French Customs). In most cases, an observational route was planned in order to search in all types of habitat (coastal, continental shelf and slope, and oceanic waters) and to accommodate for changeable weather conditions.

Overall, our database has $n = 1479$ samples of fin whale presence data in the Mediterranean Sea from 2000 to 2011, mostly summer and autumn, and with two years containing a higher number of observations in 2012 and 2015 (Fig. 1). These data cover most areas of the north-western Mediterranean Sea and northern Alboran Sea; however, the large majority are located in the Liguro-Provençal Basin (Fig. 2).

Since our objective is to model foraging habitat, presence data corresponding to a migratory behaviour were excluded. Such a behaviour can be identified on presence data coming from e-tagging (the campaigns with the mention “e-tagging” in their names, see Supplementary Table 1), as only this type of data gave the opportunity to estimate the animal mean speed based on the elapsed time between successive locations. We thus excluded

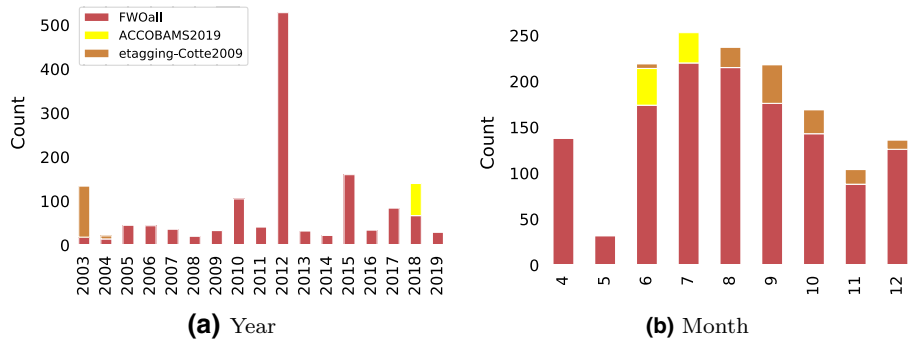


Fig. 1 Temporal distributions of fin whale presence data over years and months. The legend refers to the various datasets used (see Supplementary Table 1)

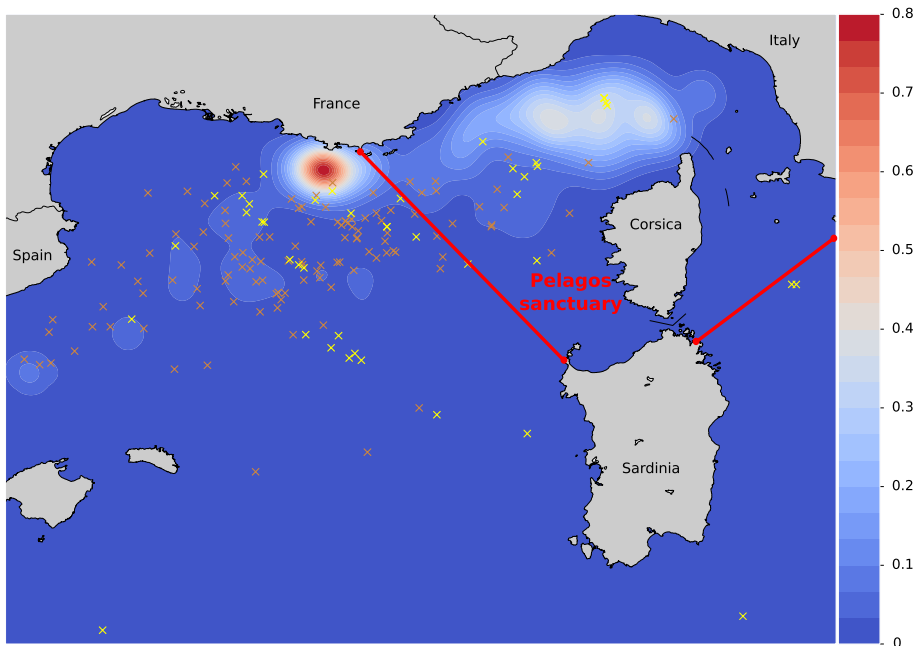


Fig. 2 Regional seas in the western Mediterranean Sea showing the geo-located presence data ($n = 1312$) fitted with a Gaussian kernel density model. Colorbar represents log(density) evaluations, normalized to be probability densities. Data from the two campaigns *etaggingCotte2009* ($n = 123$) and *ACCOBAMS2019* ($n = 44$) (see Supplementary Table 1) have been drawn separately as orange and yellow crossed, respectively, and will be used to evaluate generalization performance of models. The Pelagos sanctuary is also represented within the red region

observation samples corresponding to an average daily speed above 8 km/h, as these levels can be considered as a migratory behaviour rather than a foraging behaviour (Lydersen et al., 2020). The proportion of potentially migrating e-tag observations is around 10% (see histogram of average daily whale speed from the *etaggingCotte2009* dataset in Supplementary Fig.1). A more elaborated method presented in Panigada et al. (2017) and not used

Table 1 Candidate environmental predictors of cetacean habitat models

Environmental predictor	Variable name	Source
Bathymetry		
Depth (m)	<i>Bathy</i>	General Bathymetric Chart of the Oceans ^a
Biological		
Chlorophyll-a concentration (mg m ⁻³)	<i>CHL</i>	MODIS-Aqua sensors ^b
Chlorophyll-a gradient (mg m ⁻³ km ⁻¹)	<i>gradCHL</i>	Processed from CHL (see Eq. 1)

^a<http://www.gebco.net/>.

^b<https://oceandata.sci.gsfc.nasa.gov/>.

here is the area-restricted search behaviour which is characterised by higher turning angles and lower autocorrelation values in direction and speed to maximise searching effort in the most profitable areas.

2.2 Pseudo-absences

As we do not have sampling effort from all campaigns, we used a background sampling method to generate pseudo-absences that do not need them, namely random sampling. It consists in generating a first set of naive background points (i.e. unbiased) taken at random within the whole study area and over all the months from the year 2013, except at the occurrence localities of the target species. We followed past studies (Barbet-Massin et al., 2012; Cerasoli et al., 2017) in taking a number of random pseudo-absences equal to the number of presence records, which have already been shown for example to provide a good accuracy in boosted regression tree models. The number of pseudo-absences was set to $n \times \alpha_{pa}$, where α_{pa} is a multiplicative factor set to 1 or 3 in our experiments. This latter value is used to unbalance the number of pseudo-absences comparatively to presence ones, as classically done in presence/absence model (Barbet-Massin et al., 2012).

2.3 Environmental data

To describe the characteristics of pelagic habitats, we gathered and compiled geographic rasters containing the pixel values of three different environmental descriptors, as described in Table 1.

2.3.1 Bathymetry

The water depth was taken from the General Bathymetric Chart of the Oceans (GEBCO) of the British Oceanographic Data Centre, with a spatial resolution of 1 arc-minute grid (ca. 1.85 km at the latitude of interest; GEBCO³) and interpolated to a 1/24° grid resolution (≈ 4 km). Note that in contrast to other variables, bathymetric features (e.g., coastal areas, continental slopes, submarine canyons and seamounts) are all static.

³ www.bodc.ac.uk/products/bodc_products/gebco/.

2.3.2 Biological data

Several studies have shown that feeding fin whales and other fin whales are often located in areas of Chlorophyll-a fronts (Cotté & Guinet, 2009), where many zooplankton species are abundant. The concentration of small and large zooplankton in convergence areas, such as chlorophyll-a fronts, is, in turn, known to attract higher trophic level predators, leading to the assemblage of a complete pelagic food web (Olson et al., 1994; Polovina et al., 2001; Briscoe et al., 2017). The size of chlorophyll-a fronts were more specifically shown to be linked to mesozooplankton biomass composed of an assemblage of 131 most common species of abiotic niches ranging from subtropical to polar conditions in the North Atlantic (Druon, 2019). The chlorophyll-a gradient and associated mesozooplankton feeding habitat provides appropriate information on the useful fraction [10–20%, e.g. Raymont (1980); Libralato et al. (2008)] of the primary productivity transferred along marine food chains and primarily in the pelagic ecosystems (Druon et al., 2021a, 2021b). Fin whale mainly feeding on northern krill in the Mediterranean Sea (*M. norvegica*, Notarbartolo di Sciarra et al., 2016), productivity fronts were therefore used as a direct proxy for potential feeding (Panigada et al., 2017). It results that the vicinity of chlorophyll fronts together with medium chlorophyll-a content (CHL in mgChl m^{-3}) are assumed to describe the fin whale foraging niches in the FHO model (Druon et al., 2012; Panigada et al., 2017). The geo-projected product CHL was selected from MODIS-Aqua sensor, mapped daily (at the scale of the processes involved, i.e. within 12 h) with a medium resolution of a $1/24^\circ$ (≈ 4 km). The horizontal chlorophyll-a gradient ($gradCHL$ in $\text{mgChl m}^{-3} \text{ km}^{-1}$) was derived from the daily chlorophyll-a data, using a bi-directional gradient norm over a three-by-three grid-cell window as follows (Druon, 2019):

$$gradCHL = \sqrt{D_x^2 + D_y^2} \quad (1)$$

with D_x , D_y , the longitudinal and latitudinal chlorophyll-a horizontal gradient, respectively, corrected by the pixel size in km. Small and large chlorophyll-a fronts refer to variable absolute levels of chlorophyll-a gradient values.

3 Modeling frameworks

In this section, we detail the models used in our study, starting with the Feeding Habitat Occurrence (FHO) model from Druon et al. (2012).

3.1 Principle and formalism

Our models \mathcal{F} aim to predict a probability presence P_k , which can be either continuous or binary, from an input vector X_k for each data sample indexed by k , that is $P_k = \mathcal{F}(X^k)$. Here, a data sample refers to either a presence or a pseudo-absence of fin whales. The input vector X_k was composed of the three environmental variables (described above in Sect. 2.3) : *Bathy* (bathymetry in m), CHL (chlorophyll-a in mgChl m^{-3}) and $gradCHL$ (gradient of chlorophyll-a in $\text{mgChl m}^{-3} \text{ km}^{-1}$). After aggregating rasters of these variables to match the $1/24^\circ$ grid resolution (≈ 4 km), they were sampled at each data sample location and concatenated to build the three-dimensional input vector X_k . 5×5 (i.e. $23 \text{ km} \times 23 \text{ km}$)

patches around these locations were also extracted, each patch thus containing 25 independent input vectors X_k . In order to smooth out too local effects, we defined our probability presence P_k as the overall median of the output probabilities obtained from the 25 input vectors X_k within each 5×5 patch. In the following, index k will be dropped for conciseness.

All input patches have been post-processed with the following operations. Temporal interpolation was first performed, which consisted in applying a pixel-wise 5-days interpolation on values of *CHL* and *gradCHL* variables, under the hypothesis that they are persistent over this time duration. *Bathy* values corresponding to land have been set to NaN as are the missing values. Patches containing more than half of NaN values were discarded of our analysis.

3.2 FHO (feeding habitat occurrence) model

Druon's favourable feeding habitat of fin whale (simply referred to FHO in the following) is a daily indicator relating whale feeding behavior to the occurrence of chlorophyll-a fronts that are detected by satellite sensors of ocean colour. Chlorophyll-a fronts are mesoscale features such as ocean eddies or meandering currents, generating upwelling (divergence) and downwelling (convergence), and that persist long enough (i.e. weeks to months) to potentially sustain zooplankton production (Labat, 2009; Druon, 2019) prior to becoming hotspots for feeding at higher trophic levels. These productive frontal features have then been shown to attract large predators like fin whales (D'Amico et al., 2003; Notarbartolo di Sciara et al., 2016; Panigada et al., 2017), but also e.g. Atlantic bluefin tuna (Druon et al., 2016) and skipjack tuna (Druon et al., 2017).

The FHO model provides a daily unitary index of the fin whale's preferred feeding habitat. This index (orange line, Supplementary Fig.2) is based on the distribution of horizontal chlorophyll-a gradients at the whale locations (green histogram, Supplementary Fig.2). These chlorophyll-a gradients represent small and large productive front features, as detected by MODIS-Aqua sensor. The minimum chlorophyll-a gradient value that defined an index value of zero (lower value of the orange line, Supplementary Fig. 2) corresponds to the percentile 5th of species observations, while the slope between the zero and one index values was defined by the maximum slope of the cumulative distribution of these gradients (green dashed line, Supplementary Fig. 2). The FHO model used in our study is a recalibrated version of the one used in Druon (2017, Fig. 4), using 1479 fin whale presence data samples instead of 1999. Numerical values of the five parameters of this model are as follow:

- $CHL_{min} = 0.111$ (in mgChl m^{-3}), equal to the 3rd percentile;
- $CHL_{max} = 0.4196$ (in mgChl m^{-3}), equal to the 97th percentile;
- $gradCHL_{min} = 0.000628$ (in $\text{mgChl m}^{-3} \text{ km}^{-1}$), equal to the 25th percentile;
- $gradCHL_{int} = 0.0041346$ (in $\text{mgChl m}^{-3} \text{ km}^{-1}$), extracting from a linear fit of the cumulated distribution of *gradCHL* values located at CS (see Supplementary Fig. 2 for details);
- $bathy_{min} = 90$ (in m), equal to the 3rd percentile value since sightings in the inner shelf are unusual.

where the variables *gradCHL* and *CHL* correspond to chlorophyll-a fronts and chlorophyll-a, respectively. Note that a minimum water depth of 90 m is used to

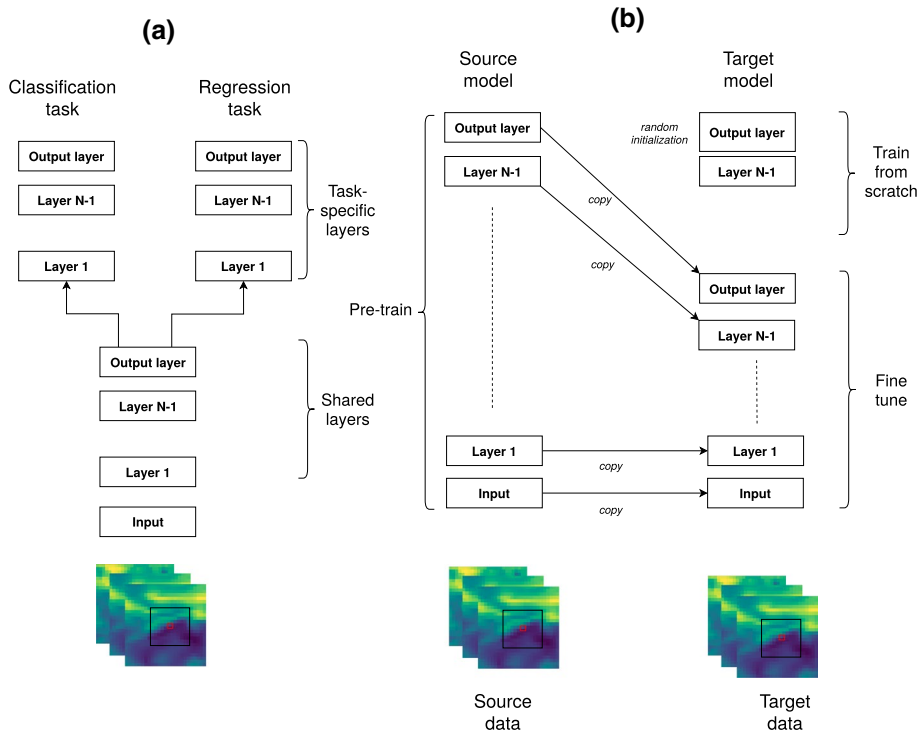


Fig. 3 Schematic overview of **a** multi-task learning and **b** transfer learning approaches

exclude the inner continental shelf. For $gradCHL_{min} \leq gradCHL \leq gradCHL_{int}$ and $CHL_{min} \leq CHL \leq CHL_{max}$, we formulate P_k^{FHO} as

$$P_k^{FHO} = 1 + \frac{\ln(gradCHL) - \ln(gradCHL_{int})}{\ln(gradCHL_{int}) - \ln(gradCHL_{min})} \tag{2}$$

Note that, Druon (2017, Druon et al., 2021a) arbitrarily set P_k^{FHO} to zero when it is inferior to 0.3, as a way to highlight the high mobility of fin whale to select productivity fronts of effective interest for feeding. This outcome leads to an absence of feeding habitat between 0 and 0.3, conditions for which energy costs for feeding would exceed energy gains by feeding on low productivity fronts (i.e. sub-optimal feeding conditions).

3.3 Proposed models

In our proposed models, $\mathcal{F}(X)$ will be a nonlinear function approximator modelled as one or several neural network(s). The considered models use different learning schemes that can be broadly divided into:

1. *data-driven* (DD), where models are trained performing a binary classification from whale presence data and pseudo-absences;

2. *multi-task learning* (MTL) (a.k.a parallel transfer learning), where a multi-output model performs a joint learning of both fin whale presence / pseudo absence data and P_k^{FHO} processed respectively as binary classification and regression tasks;
3. *transfer learning* (TL) (a.k.a sequential transfer learning), which applies a two-step learning procedure: (1) starting by learning a knowledge base from P_k^{FHO} only, and then (2) fine-tuning the model with fin whale presence / pseudo absence data;

Different multi-task learning and transfer learning schemes were developed (Fig. 3). Note that the difference between them lies in the way they transfer the knowledge: the shared information among all tasks is learned jointly in multi-task learning, while the knowledge in transfer learning is transferred from one to the other.

The three proposed models are listed in Table 2, named as *learnStrat_archi*, where *learnStrat* refers to one of the three learning schemes just described, and *archi* corresponds to the core architecture of the model.

DD_MLP This model is a simple Multi Layer Perceptron (i.e. a feed-forward artificial neural network) that performs a binary classification task to predict a fin whale presence of pseudo-absence. It consists of a simple stack of 3 fully connected layers with relu activations, each having 16 hidden units, and followed by a fourth single-unit layer with a softmax activation and a binary cross-entropy as loss function defined as

$$L_{\text{cla}} = -\frac{1}{N} \sum_{k=1}^K (y_k \log(P_k) + (1 - y_k) \log(1 - P_k)) \quad (3)$$

where $K = n \times (1 + \alpha_{\text{pa}})$ is the number of samples, y_k is the ground truth of fin whale presence or absence (i.e. 0 or 1) in patch k , and P_k is its predicted probability.

MTL_MLP//MLP This model performs multi-task learning by jointly running a regression task and a binary classification task (graph (A) of Fig. 3). The model follows a single-input multi-output architecture, using the first three layers of the DD_MLP model as the shared network branch, on which we added two parallel fourth layers that are trained jointly through a combination of their loss functions, as follows:

$$L_{\text{mtl}} = w_{\text{reg}} * L_{\text{reg}} + w_{\text{cla}} * L_{\text{cla}} \quad (4)$$

where L_{reg} is the function loss of the regression task, defined as a mean absolute error:

$$L_{\text{reg}} = -\frac{1}{k} \sum_{k=1}^K |P_k - P_k^{\text{FHO}}| \quad (5)$$

and w_{reg} and w_{cla} are two weights to balance the two task-specific losses, P_k is the predicted value of P_k^{FHO} . From preliminary numerical experiments, we set these weights to 0.6 and 0.4, respectively for w_{reg} and w_{cla} , which best satisfied both accuracy performance in our results and model optimization.

TL_MLP+MLP This last model is a case of (sequential) transfer learning where we first built a knowledge base pre-trained on a P_k^{FHO} regression task. To do so, the exact same MLP architecture as DD_MLP was used with a sigmoid (as P_k^{FHO} values are unitary) activation in its last layer and a mean absolute error loss function, performing a fivefold cross validation to predict the P_k^{FHO} values. This knowledge base was then transferred to our binary classification task by first adding a fully connected classifier on top

Table 2 Details of proposed models

Model	Learning strategy	Architecture	Tasks	
			Regression	Binary classification
DD_MLP	Data-driven based on presence data and pseudo-absences	MultiLayer perceptron	X	✓
MTL_MLP//MLP	Multi-task learning jointly regression and classification tasks	Two multiLayer perceptron run in parallel	✓	✓
TL_MLP+MLP	Transfer learning from a regression task on FHO to a classification task on presence data and pseudo-absences with fine-tuning	Two multiLayer perceptron run sequentially with fine tuning	✓	✓

Table 3 Details of test / training sets for the different experimental datasets. LOGO stands for Leave-One-Group-Out

Dataset	Train samples	Test samples	Splitting procedure
Complete	1331	148	10-fold cross-validation
Gene_1	1331	123	10-fold LOGO using etaggingCotte2009 as test group
Gene_2	1331	44	10-fold LOGO using ACCOBAMS2019 as test group

of it, and secondly by fine tuning the last two layers of MD_MLP, i.e. jointly training the new added classifier and these two layers while freezing the other layers.

Implementation details All models output a presence probability through their softmax activation in last layers. They have been developed using Keras functional API and TensorFlow backend.⁴ Our network is trained end-to-end using RMSProp optimizer (Tieleman & Hinton, 2012). We decreased the learning rate accross iterations, typically from $10 e^{-3}$ to $10 e^{-6}$. Note that all models involving fin whale presence data are relatively shallow networks with low capacity, in order to prevent overfitting the small amount of these data. All the experiments were run under workstations with a single GPU (Nvidia GTX 1080 and GTX 1080 Ti), resulting in computational times for model learning up to 10 min for the models with the highest capacity.

4 Experimental setup

4.1 GAM as a baseline model

As a baseline model, we implemented a Generalized Additive Model (GAM), as it has been widely used for whale occurrence prediction (e.g. Derville et al., (2018)). We used the pyGAM implementation,⁵ mainly written based on the R mgcv package⁶ by Wood (2006). As we deal with a binary classification task, logisticGAM is used, featuring a Binomial error distribution and a logit link. Like for our NN-based models, we did not perform hyperparameter tuning of these models as it is generally done in CDM literature (e.g. (Derville et al., 2018)), except for quick optimization on the kernel selection.

4.2 Training and test sets

Table 3 details our evaluation protocols in terms of train/test splitting. Cross-validation has been systematically used to evaluate the stability of models, e.g. by observing no overfitting between the loss on the validation versus test set during cross-validation, and to report standard deviation along the median values, but it was not used to validate choices of hyper-parameters. The complete dataset gathers all of our fin whale presence

⁴ <https://keras.io/>, <https://www.tensorflow.org/>.

⁵ <https://pygam.readthedocs.io/en/latest/>.

⁶ <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.

data except those from the campaigns etaggingCotte2009 and ACCOBAMS2019. Data from these campaigns have been used as two independent datasets to test generalization capacity of our models, namely the Gene_1 and Gene_2 datasets. For the latter datasets, Leave-One-Group-Out has been employed, with the campaign names of etagging-Cotte2009 and ACCOBAMS2019 as group information to encode domain specific pre-defined cross-validation folds, and keeping the same number of training samples as in the complete dataset.

For evaluation of our models, we report the True Positive rate per unit of favourable Surface Habitat (TPSH). This metric is defined for the k th test sample as the model output probability P_k weighted by the Spreading of Habitat Surface (SHS) around this test patch. More precisely, this latter weighting metric counts the number of true positive samples estimated by the model in a 49×49 (i.e. $\approx 225 \text{ km} \times 225 \text{ km}$) area around the test sample, and penalize predictions of true positive samples associated with a too large habitat spreading. Put it another way, TPSH increases inverse proportionally to its number of pseudo false positive predictions over a 49×49 pixel patch around the test sample. This metric thus includes an implicit evaluation of false positive rates, yet it does not consider an explicit count of absences correctly predicted or not as in the True Skill Statistic metric used in presence/absence model, and defined as the sum of sensitivity and specificity minus one, the sensitivity being the proportion of presences correctly predicted, and while the specificity is the proportion of absences correctly predicted (Barbet-Massin et al., 2012).

Mathematically, TPSH is thus defined as

$$\text{TPSH} = \frac{1}{K} \sum_{k=1}^K (\mathbb{1}_\lambda(P_k) \times (1 - \text{SHS}_k)) \quad [\text{in}\%] \tag{6}$$

where $\mathbb{1}_\lambda(P_k)$ is the indicator function (i.e. $\mathbb{1}_\lambda(x)=1$ for $x \geq \lambda$ and 0 otherwise), P_k is the model output probability for patch k and K the number of test samples. The term SHS_k is the Spreading of Habitat Surface of pixels in patch k , reduced to the unitary interval, that is

$$\text{SHS}_k = \sum_{n=1}^{N_k} \frac{\mathbb{1}_\lambda(P_k(n))}{(N_k - N_k^{\text{nan}})} \quad [\text{in}\%] \tag{7}$$

where $N_k = 49^2 = 2401$ the total number of pixels per patch, N_k^{nan} the number of NaN values in the patch, and $P_k(n)$ is the model output probability for each pixel n .

The threshold λ is set to 0.5 by default, but we also report in our study Receiver Operating Characteristic (ROC) curves by exploring the value range [0.35 ; 0.65] with a 0.02 step. From such curves, the euclidean distance to the top-left corner of the ROC curve for each λ can be computed, where lower values will show better performance than larger ones.

The performance of a habitat model is always relative to the proportion of favourable habitat compared to the entire domain. For instance for the FHO model, the surface area of suitable habitat (defined as above 50% favourable daily occurrence) ranged from about 7% of the western Mediterranean Sea in summer to about 26% in winter while 40% of the observations (of which 95% were in summer months) were within this favourable habitat and 80% were within 8 km from it (Panigada et al., (2017), $n = 1,287$). Since top predators spend most of their time searching for their effective environment for feeding (in our context), they are likely often close to it but not necessarily often in it (effectively feeding). The distance to the closest favourable habitat is therefore a powerful metric for fin whales as their movement capacity is large. However, we choose to use as a metric the True Positive rate per Unit of favourable Surface Habitat (TPSH), with relatively low absolute levels,

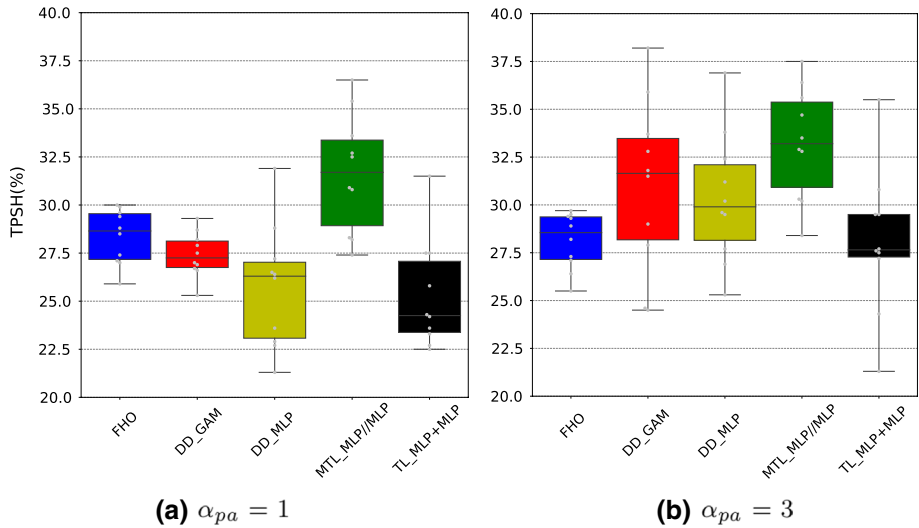


Fig. 4 Boxplots of TPSH metric (in %) to assess model performance on a presence–absence classification task on the complete dataset. Remember that coefficient α_{pa} is used to unbalance the number of pseudo-absences w.r.t to presence ones, as detailed in the background sampling strategy in Sect. 2.2. A value of $\lambda = 0.5$ has been used

since the main aim of the paper was to compare the different model performances using a simple metric rather than to evaluate their practical usefulness in management.

5 Experiments

The models were first benchmarked on the task of presence–absence prediction of whale sightings using the complete dataset. For $\alpha_{pa} = 1$ and $\lambda = 0.5$, the proposed model MTL_MLP//MLP performs the best in predicting fin whale presence, improving FHO, DD_GAM and DD_MLP of 3.1%, 4.5% and 5.4% on the median of TPSH, respectively (subfigure (a) of Fig. 4). In relative terms, these improvements from FHO and GAM correspond to 10.8% and 16.5%, respectively (obtained by dividing the absolute gains of 3.1% and 4.5% by 28.7% and 27.2%, i.e. the median values of FHO and GAM models from subfigure (a) of Fig. 4). MTL_MLP//MLP and FHO also exhibited a smaller variance than more data-driven approach like DD_GAM and DD_MLP models. Thus, incorporating expert knowledge in models is also an efficient way to reduce variability of estimations. Also, increasing the number of pseudo-absences improves all models, with greater benefits for the two data-driven models, e.g. up to 5% of TPSH for DD_MLP, and to a lesser extent the MTL_MLP//MLP and TL_MLP+MLP models, e.g. with a gain of 1% for MTL_MLP//MLP (subfigure (b) of Fig. 4). This also globally increases dispersion of output models. This result is in line with CDM literature, as it is known that a random selection of pseudo-absences is recommended when high specificity is valued over high sensitivity (e.g. for reserve planning) (Barbet-Massin et al., 2012). Furthermore, when comparing our two different learning strategies, clearly the knowledge from FHO is better exploited in the multi-task learning MTL_MLP//MLP than in the sequential transfer learning TL_MLP+MLP models, where

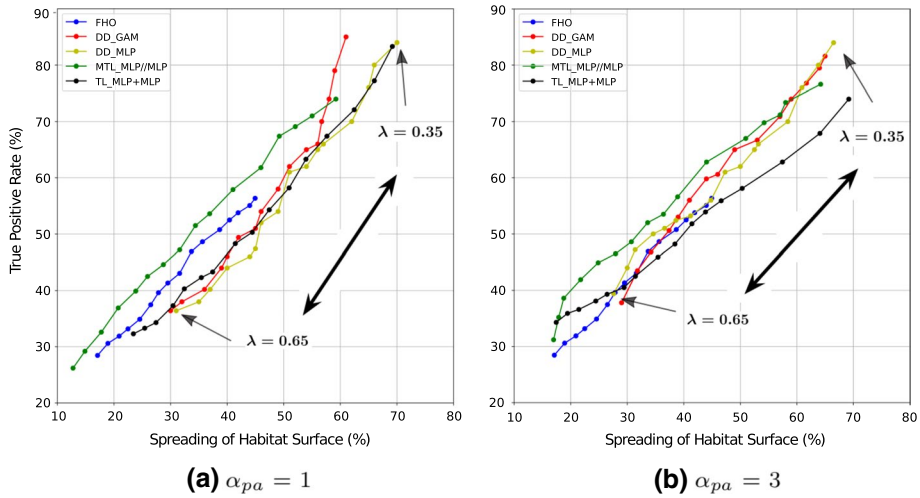


Fig. 5 ROC curves to assess model performance on a presence–absence classification task on the complete dataset. Each point of the curves corresponds to a minimum favourable habitat occurrence (lambda value) linearly sampled from 0.35 to 0.65 (with an increment of 0.02). Note that the most interesting cases in terms of use are when the habitat surface is below about 35%, corresponding to a relatively highly discriminant habitat

in this latter the classification fine-tuning seems to overwhelm the FHO knowledge base as this model does not even reach FHO performance.

Further investigations were conducted on the trade-off between True Positive Rate (in %) and SHS (in %) inside the TPSH metric. ROC curves were computed by thresholding from 0.35 to 0.65 the output probability of each model, i.e. each point of the curves corresponds to a minimum favourable habitat occurrence (lambda value) linearly sampled from 0.35 to 0.65 (Fig. 5). Naturally, similar trends in results as with the TPSH metric (Fig. 4) were obtained, but we can now better visualize how data-driven models are quite limited in decreasing the spatial extent of their habitat estimation (subfigure (a) of Fig. 5), demonstrating their inability in properly isolating locally-defined environmental features related to fin whale occurrences. Benefits from increasing the α_{pa} parameter can be simply described as translating all ROC curves on the left, plus forming a plateau-shaped slope on certain models like MTL_MLP//MLP and DD_MLP for reasonably small SHS values ($\leq 35\%$) (subfigure (b) of Fig. 5). This means that this mechanism of unbalancing presence–absence classes during training allows to reduce habitat spatial spreading while boosting performance accuracy of models (gains up to 5% for data-driven models), which are promising properties towards better CDM models. This augmentation also makes the crossing between the ROC curves of MTL_MLP//MLP and data driven models occurring for smaller SHS values.

It is now interesting to visualize how the proposed models deviate from original FHO maps. Using for each model the training set from the best performing fold, i.e. with the highest TPSH metric, maps of whale presence–absence over 49×49 pixel patches (i.e. $\times 225 \text{ km} \times 225 \text{ km}$) were computed (Fig. 6). The MTL_MLP//MLP model (map 4 in subfigures (a)–(c) of Fig. 6) basically tends to slightly spread around FHO patterns. However, it captures in finer details salient mesoscales structures than the data-driven models, while being also able to slightly expand its habitat surface over new areas. Such property

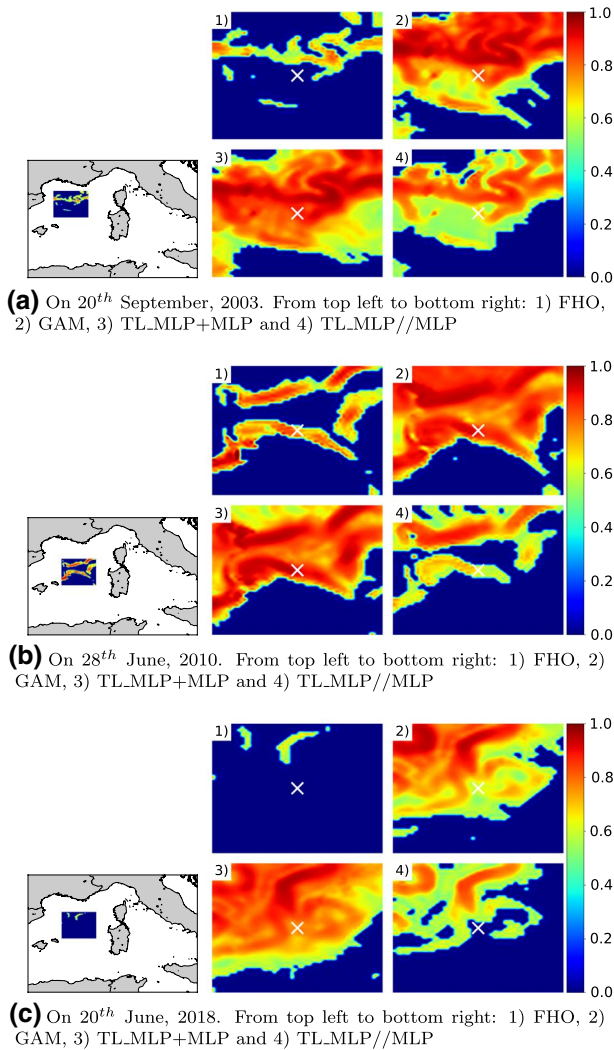


Fig. 6 a–c Examples of predicted presence probability maps of size 49×49 pixel (i.e. $\approx 225 \text{ km} \times 225 \text{ km}$) for the different models, from top left to bottom right: (1) FHO, (2) GAM, (3) TL_MLP + MLP and (4) TL_MLP//MLP. The white cross points to the location of a whale presence observation

most likely results from a deeper use of the learnt local relationships between presence data and the surrounding environment. It is also interesting to visualize how models extrapolate beyond FHO patterns so as to better reach visual sighting locations based on the learnt relationships from data. Although salient structures from these extrapolations are quite similar, once again the TL_MLP//MLP model performs the best in terms of narrowing its spatial prediction of habitats.

Our last experiment consisted in evaluating generalization capacity of our models using two independent campaigns of whale sightings, i.e. none of which have been seen during model training. Note that in the previous experiment, some presence data from a same

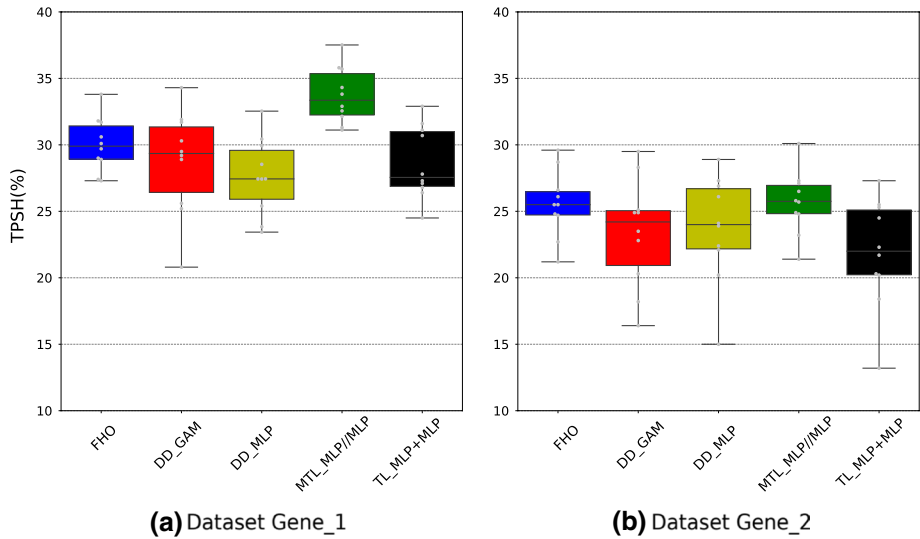


Fig. 7 Boxplots of TPSH metric (in %) to assess model performance on a presence–absence classification task on the **a** Gene_1 (i.e. the campaign etaggingCotte2009) and **b** Gene_2 (i.e. the campaign ACCO-BAMS2019) datasets with $\alpha_{pa} = 1$

campaign could appear both in training and test sets, limiting the evaluation of model generalization. Using $\lambda = 0.5$ and $\alpha_{pa} = 3$, performance results on the first evaluation dataset Gene_1 (subfigure (a) of Fig. 7) are quite similar to those obtained for the complete dataset (Fig. 4), and degrade significantly on the second Gene_2 dataset (subfigure (b) of Fig. 7), showing also higher dispersion values on all models. Explanations of such a result can be initiated with the respective multivariate density distributions of these two datasets in comparison to the complete dataset (Supplementary Fig. 3), which reveal stronger similarities between distributions from the complete and Gene_1 datasets rather than with the Gene_2 dataset, especially through the *gradientCHL* variable. Note that this was expected considering the time periods of both campaigns, where a higher primary productivity is expected in early winter (Gene_1 time period) than in spring and summer (Gene_2 time period) which are more transitional seasons. When comparing model results in more details, FHO and MTL_MLP//MLP models exhibit the best performance on the TPSH metric, with performance gains up to 5% in comparison to data-driven averaged performance (Fig. 7), confirming our initial hypothesis that expert information on whale behaviour represent a good knowledge base for model generalization, which can be further improved by a concurrent learning of more local species–environment relationships from in-situ presence data.

6 Conclusion and perspectives

In this study we have proposed a deep learning based strategy, i.e. a neural network model plus a learning scheme, able to perform concurrently pixel-based classification and regression tasks over both whale sightings and pre-computed expert maps. Such an approach has revealed a promising capacity in enhancing models w.r.t the extrapolation problem. It allows to obtain more refined and delimited likely occurrence areas of whales in contrary

to data-driven models. This is a crucial requirement for operational management applications dedicated to marine mammal protection and conservation.

While a deep learning approach may lead to better results using supplementary physical variables (such as surface temperature or currents), we selected the same environmental variables than the FHO model so as to favour a strict comparison of model performance, as being the main objective of this current paper. Future directions of our work will investigate the use of more explanatory variables. We will also investigate more complex spatio-temporal CDM modeling. Indeed, although this is not the case of the feeding habitat model used in our study, current environmental covariables in CDM are mostly restricted to localized, single-pixel values alone (or average at fixed spatio-temporal resolutions), preventing them to capture larger contextual information that may be of importance for whale habitat modelling. Indeed, the precise time and spatial scales of raster data relevant to whale habitat modeling is still an open question. Note however that the FHO model already explicitly captures these larger relevant information as being centered on the productivity fronts. The integration of contextual information with multiscale features can be done conveniently in some deep learning architectures, e.g. by feeding convolutional neural networks with high-resolution spatial patches of raster data around each spotted whale. In particular, end-to-end training of networks should be able to fuse multi-scale features from both large-scale contextual information and fine-grained details, and discover themselves the most optimal space scales for habitat modelling, which are difficult to set a priori.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10994-021-06029-z>.

Acknowledgements We would like to thank our main sponsors in this work: CominLabs (<https://www.cominlabs.u-bretagne.fr/>) through the innovation action Tech4Whales, DREC Agence Française de la Biodiversité (<https://www.afbiodiversite.fr/>) and ISblue (<https://www.isblue.fr/about-us/>). We are also very grateful for all providers of cetacean visual sightings and e-tagging data, namely: Cotté, C. and Guinet, C., the associations Souffleurs d'Ecume and GECEM (<https://www.gecem.org>, now MIRACETI), Pelagis Observatoire and Tethys Research Institute.

References

- Aïssi, M., Celona, A., Comparetto, G., & Mangano, R. (2008). *Large-scale seasonal distribution of fin whales (*balaenoptera physalus*) in the central mediterranean sea*. U. K.: J. Mar. Biol. Assoc.
- Balrusaitis, T., Ahuja, C., & Morency, L. P. (2017). Multimodal machine learning: A survey and taxonomy. *CoRR*, *abs/1705.09406*.
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, *3*, 327–338.
- Benedetti, P., Ienco, D., Gaetano, R., Ose, K., & Pensa, R. (2018). M³ fusion: A deep learning architecture for multiscale multimodal multitemporal satellite data fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *IEEE*, *11*(12), 4939–4949.
- Bethoux, J. P., Gentili, B., Morin, P., Nicolas, E., Pierre, C., & Ruiz-Pino, D. (1999). The mediterranean sea: A miniature ocean for climatic and environmental studies and a key for the climatic functioning of the North Atlantic. *Progress in Oceanography*, *44*, 131–146.
- Botella, C., Joly, A., Bonnet, P., Monestiez, P., & Munoz, F. (2018). A deep learning approach to species distribution modelling. In A. Joly, S. Vrochidis, K. Karatzas, A. Karppinen, & P. Bonnet (Eds.), *Multimedia Tools and Applications for Environmental, Biodiversity Informatics* (pp. 169–199). Springer.
- Briscoe, D. K., Hobday, A. J., Carlisle, A., Scales, K., Eveson, J. P., Arrizabalaga, H., & Fromentin, J. M. (2017). Ecological bridges and barriers in pelagic ecosystems. *Deep Sea Research Part II: Topical Studies in Oceanography*, *140*, 182–192.
- Buckland, S., Anderson, D., Burnham, K., Laake, J., & Borchers, D. (2001). *Introduction to distance sampling: Estimating abundance of biological populations*. Oxford: Oxford University Press.


- Cañadas, A., & Vazquez, J. A. (2014). Conserving Cuvier's beaked whales in the alboran sea (sw mediterranean): Identification of high density areas to be avoided by intense man-made sound. *Biological Conservation*, *178*, 155–162.
- Castellote, M., Clark, C. W., & Lammers, M. O. (2012). Acoustic and behavioural changes by fin whales (*balaenoptera physalus*) in response to shipping and airgun noise. *Biological Conservation*, *147*, 115–122.
- Cerasoli, F., Iannella, M., D'Alessandro, P., & Biondi, M. (2017). Comparing pseudo-absences generation techniques in Boosted Regression Trees models for conservation purposes: A case study on amphibians in a protected area. *PLoS ONE*, *12*, 1–23.
- Chen, D., Xhue, Y., Chen, S., Fink, D., & Gomes, C. (2017). Deep multi-species embedding.
- Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Ben Rais Lasram, F., Aguzzi, J., et al. (2010) The biodiversity of the mediterranean sea: Estimates, patterns, and threats. *PLoS ONE*, *5*(8):e11842, 1–36. <https://doi.org/10.1371/journal.pone.0011842>
- Cotté, C., & Guinet, C. (2009). Scale-dependent habitat use by a large free-ranging predator the mediterranean fin whale. *Deep Sea Research I*, *56*, 801–811.
- David, L., Alleaume, S., & Guinet, C. (2011). High risk areas of collision between fin whales and ferries in the north-western mediterranean sea. *Journal of Marine Animals and Their Environment*, *4*, 17–28.
- 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 Distributions*, *2018*(24), 1657–1673.
- Druon, J. N. (2017). *Ocean Productivity index for Fish in the Arctic: First assessment from satellite-derived plankton-to-fish favourable habitats*. Luxembourg: Publications Office of the European Union.
- Druon, J. N., Chassot, E., Murua, H., & Lopez, J. (2017). Skipjack tuna availability for purse seine fisheries is driven by suitable feeding habitat dynamics in the Atlantic and Indian oceans. *Frontiers in Marine Science*, *4*, 315. <https://doi.org/10.3389/fmars.2017.00315>.
- Druon, J. N., Fromentin, J. M., & Hanke, A. R. E. A. (2016). Habitat suitability of the Atlantic bluefin tuna by size class: An ecological niche approach. *Progress in Oceanography*, *142*, 30–46.
- Druon, J. N., Hélaouët, P., Fromentin, J. M., Palialexis, A., Hoepfner, N. (2019). Satellite-based indicator of zooplankton distribution for global monitoring. *Science Report*, *9*, 4732. <https://doi.org/10.1038/s41598-019-41212-2>.
- Druon, J. N., Panigada, S., David, L., Gannier, A., Mayol, P., Arcangeli, A., Cañadas, A., Laran, S., Di Méglion, N., & Gauffier, P. (2012). Potential feeding habitat of fin whales in the western mediterranean sea: An environmental niche model. *Marine Ecology Progress Series*, *464*, 289–306.
- Druon, J. N., Gascuel, D., Gibin, M., Zanzi, A., Fromentin, J. M., Colloca, et al. (2021a). Mesoscale productivity fronts and local fishing opportunities in the European Seas. *Fish and Fisheries*, 1–21. <https://doi.org/10.1111/faf.12585>
- Druon, J. N., Mangin, A., Hélaouët, P., & Palialexis, A. (2021b). The chlorophyll-a gradient as primary Earth observation index of marine ecosystem feeding capacity. Copernicus Ocean State Report 5. <https://doi.org/10.1080/1755876X.2021.1946240>
- Dunn, D. C. (2014). The convention on biological diversity's ecologically or biologically significant areas: Origins, development, and current status. *Marine Policy*, *49*(2014), 137–145.
- D'Amico, A., Bergamasco, A., Zanasca, P., Carniel, S., Nacini, E., Portunato, N., Teloni, V., Mori, C., & Barbanti, R. (2003). Qualitative correlation of marine mammals with physical and biological parameters in the Ligurian sea. *IEEE Journal of Oceanic Engineering*, *28*, 29–43.
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*, 330–342.
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, *40*, 677–697.
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, E. Y., & Yates, C. J. (2011). A statistical explanation of maxent for ecologists. *Diversity and Distributions*, *17*, 43–57.
- European Commission, E. (2008). The role of cfp in implementing an ecosystem approach to marine management. Technical report to the council and the European parliament.
- European Commission, E. (2017). Commission Decision (EU) 2017/848 of 17 may 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing Decision 2010/477/EU (Text with EEA relevance), pp. 43–74.
- Franklin, J. (2010). Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, *16*, 321–330.
- Guillera-Arroita, G. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, *24*, 276–292.

- Hinton, G. E., Osindero, S., & Teh, Y. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18, 1527–1554.
- Jin, Q., & Liang, J. (2016). Video description generation using audio and visual cues. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval* (pp. 239–242). ACM.
- Kahou, S. E., Bouthillier, X., Lamblin, P., Gulcehre, C., Michalski, V., Konda, K., Jean, S., Froumenty, P., Dauphin, Y., Boulanger-Lewandowski, N., & Ferrari, R. C. (2016). Emonets: Multimodal deep learning approaches for emotion recognition in video. *Journal on Multimodal User Interfaces*, 10(2), 99–111.
- 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.
- Labat, J. P. (2009). Mesoscale distribution of zooplankton biomass in the northeast Atlantic ocean determined with an optical plankton counter: Relationships with environmental structures. *Deep-Sea Research Part I: Oceanographic Research Papers*, 56, 1742–1756.
- Libralato, S., Coll, M., Tudela, S., Palomera, I., & Pranovi, F. (2008). Novel index for quantification of ecosystem effects of fishing as removal of secondary production. *Marine Ecology Progress Series*, 355, 107–129.
- Lydersen, C., Vacquié-Garcia, J., Heide-Jorgensen, M., Øien, N., Guinet, C., & Kovacs, K. M., (2020). Autumn movements of fin whales (*balaenoptera physalus*) from Svalbard, Norway, revealed by satellite tracking. *Scientific Reports*, 10, 16966, 2045–2322. <https://doi.org/10.1038/s41598-020-73996-z>
- Mainali, K., Warren, D., Dhileepan, K., McConnachie, A., Strathie, L., Hassan, G., Karki, D., Shrestha, B. B., & Parmesan, C. (2015). Projecting future expansion of invasive species: Comparing and improving methodologies for species distribution modeling. *Global Change Biology*, 21, 4464–4480.
- 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, 1267–1280.
- Mannocci, L., Roberts, J., Halpin, P., Authier, M., Boisseau, O., Bradai, M. N., Cañadas, A., Chicote, C., David, L., Di-Méglio, N., & Fortuna, C. M. (2018). Assessing cetacean surveys throughout the mediterranean sea: A gap analysis in environmental space. *Scientific Reports*, 8, 3126.
- Neverova, N., Wolf, C., Taylor, G., & Nebout, F. (2016). Moddrop: Adaptive multi-modal gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(8), 1692–1706.
- Notarbartolo di Sciarra, G., Castellote, M., Druon, J. N., & Panigada, S. (2016). Chapter three-Fin whales, *balaenoptera physalus*: At home in a changing mediterranean sea? *Advances in Marine Biology*, 2016(75), 75–101. <https://doi.org/10.1016/bs.amb.2016.08.002>.
- Notarbartolo di Sciarra, G., Zanardelli, M., Jahoda, M., Panigada, S., & Airoldi, S. (2003). The fin whale *balaenoptera physalus* (l. 1758) in the mediterranean sea. *Mammal Review*, 33, 105–150.
- Olson, D. B., Hitchcock, G. L., Mariano, A. J., Ashjian, C. J., Peng, G., Nero, R. W., & Podesta, G. P. (1994). Life on the edge: Marine life and fronts. *Oceanography*, 7(2), 52–60.
- Panigada, S., Donovan, G., Druon, J., Lauriano, G., Pierantonio, N., Pirodda, E., Zanardelli, M., Zerbini, A. N., & di Sciarra, G. N. (2017). Satellite tagging of mediterranean fin whales: Working towards the identification of critical habitats and the focussing of mitigation measures. *Scientific Reports*, 7, 3365.
- Panigada, S., Notarbartolo di Sciarra, G., Zanardelli, M., Airoldi, S., Borsani, J., & Jahoda, M. (2005). Fin whales summering in the Ligurian sea: Distribution, encounter rate, mean group size and relation to physiographic variables. *Journal of Cetacean Research and Management*, 7, 137–145.
- Pearce, J. L., & Boyce, M. S. (2006). Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, 43, 405–412.
- Pearson, R., Thuiller, W., Araujo, M., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T. P. & Lees, D. C. (2006). Model-based uncertainty in species range prediction. *Journal of Biogeography*, 33, 1704–1711.
- Pesante, G., Zanardelli, M., Capoulade, F., Gannier, A., & Weinrich, M. T. (2006). Mediterranean fin whales at risk from fatal ship strikes. *Marine Pollution Bulletin*, 52, 1287–1298.
- Phillips, S. J. E. A. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*, 19(1), 181–197.
- Polovina, J. J., Howell, E., Kobayashi, D. R., & Seki, M. P. (2001). The transition zone chlorophyll front, a dynamic global feature defining migration and forage habitat for marine resources. *Progress in Oceanography*, 49(1–4), 469–483.
- Rao, D., De Deuge, M., Nourani-Vatani, N., Williams, S. B., & Pizarro, O. (2017). Multimodal learning and inference from visual and remotely sensed data. *The International Journal of Robotics Research*, 36(1), 24–43.
- Raymont, J. E. G. (1980). *Plankton and productivity in the oceans: Volume 1: Phytoplankton* (2nd Edn.). Elsevier.

- Thuiller, W. (2003). Biomod-optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, 9(10), 1353–1362.
- Tieleman, T., & Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural Networks for Machine Learning 4(2)*:26–31.
- Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC, 496 p
- Wöllmer, M., Metallinou, A., Eyben, F., Schuller, B., & Narayanan, S. S. (2010). Context-sensitive multimodal emotion recognition from speech and facial expression using bidirectional lstm modeling. In *INTERSPEECH 2010* (pp. 2362–2365).
- Yackulic, C., Yackulic, C., Chandler, R., Zipkin, E. F., Andrew Royle, J., Nichols, J., Campbell Grant, E. H., & Veran, S. (2012). Presence-only modelling using maxent: When can we trust the inferences? *Methods in Ecology and Evolution*, 4, 236–243.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

D. Cazau¹  · P. Nguyen Hong Duc² · J.-N. Druon³ · S. Matwins⁴ · R. Fablet⁵

✉ D. Cazau
dorian.cazau@ensta-bretagne.fr

¹ ENSTA Bretagne, UMR CNRS Lab-STICC, Brest, France

² Institut Jean Le Rond d'Alembert, Sorbonne Université, Paris, France

³ European Commission - Joint Research Centre, Ispra, Italy

⁴ Faculty of Computer Science, Dalhousie University, Halifax, Canada

⁵ IMT Atlantique, UMR CNRS Lab-STICC, Brest, France