DEEP ~~LEARNING AND TRAJECTORY REPRESENTATION FOR THE PREDICTION OF SEABIRD DIVING BEHAVIOUR~~ INFERENCE OF SEABIRD DIVES FROM GPS-ONLY RECORDS: PERFORMANCE AND GENERALIZATION PROPERTIES

Amédée Roy
Institut de Recherche pour le Développement (IRD),
MARBEC (Univ. Montpellier, Ifremer, CNRS, IRD)
Avenue Jean Monnet, 34200, Sète, France
amedee.roy@ird.fr

Sophie ~~**Lanco-Bertrand**~~ **Lanco Bertrand**
Institut de Recherche pour le Développement (IRD),
MARBEC (Univ. Montpellier, Ifremer, CNRS, IRD)
Avenue Jean Monnet, 34200, Sète, France
sophie.bertrand@ird.fr

Ronan Fablet
IMT Atlantique,
UMR CNRS Lab-STICC
Brest, France
ronan.fablet@imt-atlantique.fr

November 15, 2021

ABSTRACT

~~Seabirds are considered as suitable indicators for the study of marine ecosystems, since their foraging strategies provide a real-time response to complex ecosystem dynamics . By deploying GPS sensors on seabirds it is possible to obtain their trajectories , and deep learning have recently~~ At-sea behaviour of seabirds have received significant attention in ecology over the last decades as it is a key process in the ecology and fate of these populations. It is also, through the position of top predator that these species often occupy, a relevant and integrative indicator of the dynamics of the marine ecosystems they rely on. Seabird trajectories are recorded through the deployment of GPS, and a variety of statistical approaches have been tested to infer probable behaviours from these location data. Recently, deep learning tools have shown promising results for the segmentation and classification of animal behaviour from trajectory data. Yet~~there is still lot of investigation needed in terms of network architectures, data representation but also to demonstrate the generalization properties of these approaches~~, these approaches have not been widely used and investigation is still needed to identify optimal network architecture and to demonstrate their generalization properties. From a database of about ~~250~~ 300 foraging trajectories derived from GPS data deployed simultaneously with pressure sensors for the identification of dives, this work has ~~consisted in training deep~~

~~networks~~ benchmarked deep neural network architectures trained in a supervised manner for the prediction of dives from trajectory data. ~~In this study, we confirm~~ It first confirms that deep learning allows better dive prediction than usual methods such as Hidden Markov Models~~for two distinct seabirds species. We propose a novel deep learning model for trajectory data. It combines the computational efficiency of convolutional neural networks to distance-matrix-based representations of trajectory data. Our model considerably increases the ability of deep networks to infer behaviours, as well as their stability to different data inputs. The considered trajectory data representation might enable deep networks to better capture spatial information than from longitude and latitude time-series considered in previous works~~. It also demonstrates the generalization properties of the trained networks for inferring dives distribution for seabirds from other colonies and ecosystems. In particular, convolutional networks trained on Peruvian boobies from a specific colony show great ability to predict dives of boobies from other colonies and from distinct ecosystems. We further investigate accross-species generalization using a transfer learning strategy known as 'fine-tuning'. Starting from a convolutional network pre-trained on Guanay cormorant data reduced by two the size of the dataset needed to accurately predict dives in a tropical booby from Brazil. We believe that the networks trained in this study will provide relevant starting point for future fine-tuning works for seabird trajectory segmentation.

1 Introduction

Marine ~~megafauna (i.e. species that depend on marine resources for their food and located at the top of the trophic food webs)~~ top predators have received significant attention in marine ecology over the last decades [1]. ~~These species~~ They are known to use vast areas for feeding, thus requiring specific adaptive foraging strategies in order to localize their preys, especially in the pelagic environments which are highly variable. [2]. They offer a unique perspective into ocean processes and dynamics, given that they can amplify information on the structure of the seascape across multiple ~~spatiotemporal~~ spatio-temporal scales due to their relatively high mobility and longevity. Often considered as sentinels of the environmental variability and bio-indicators for ecosystem structure and dynamics, their study has been particularly contextualized into ecosystem-based management and conservation issues [3], [4].

~~Among marine megafauna species, seabird are considered as suitable indicators because they are both sensitive to variations in food supply and relatively easy to observe [5], [6]. During breeding season, seabird must return regularly on land to brood and feed chicks between foraging trips. Feeding at sea while breeding at land (i.e. central place foraging) implies that seabirds had to develop specific morphological, physiological and behavioural abilities in order to navigate efficiently to foraging zones eventually far from the breeding site, to capture preys and to go back to the nest in relatively short times [7]. At this period seabirds have no constraints related to predator avoidance, partner or site selection, foraging trips are thus essentially dedicated to food acquisition. For these reasons, their foraging movements have been suggested to reflect prey abundance and distribution at sea [8].~~

~~Over the last decades, the study of animal movements has been revolutionized~~ Numerous studies have focused on the variability of seabirds' foraging strategies and in particular of dive distributions. Assessing consistency or shifts in foraging locations [9]–[11], and in the resource spatial partitioning [12], [13] provide indeed crucial information for understanding marine ecosystems. This has been particularly enabled by great technical advances in the miniaturization and autonomy of ~~devices **monaco_bio-logging_2016**, [14]~~biologging devices [14], [15]. GPS loggers have been at the forefront of this breakthrough, and can now provide precise and accurate data on the ~~movements~~ foraging trajectories of many free-ranging species, such as seabirds [6], [16]. ~~The derived data is particularly useful in studying a variety of behavioral aspects including habitat selection, migration or dispersion patterns, and foraging strategies [17].~~ Detailed information on the ~~foraging behaviour has~~ diving behaviour has also been gained through the ~~combined use of GPS and~~ additional use of pressure sensors, such as Time Depth Recorders (TDR) devices ~~. TDR devices capture dive profiles of marine animals and many studies have used pressure as a proxy to seabirds' foraging behaviour [18]–[20].~~ [18]–[20]. Yet, for ~~practical~~historical, financial and ethical reasons, the deployment of several sensors ~~is not always possible and efforts have been made to develop accurate methods to infer foraging locations from GPS data. Many~~ has not always been possible and a substantial amount of tracking datasets consist in GPS tracks only. The development of tools dedicated to animal trajectories segmentation (i.e. for dive identification) is therefore needed to extract more out of historical seabird foraging trajectories [21].

Among existing approaches to dive identification from GPS tracks, many individual-based studies aim to infer behavioral state directly by applying thresholds ~~from~~ to various ecological metrics of movement data, such as speed, direction and tortuosity [22], [23]. A common example is the so-called First-Passage Time ~~(FPT) method~~method (hereafter, FPT), which is defined as the time taken for an individual to cross a virtual circle of given radius [24]–[26]. Here foraging behaviour is assumed to occur when birds fly at very low speeds [27]. ~~Modelling~~ Statistical methods have also been used to predict diving behaviour ~~taking the sequence of trajectory data into account through~~ with clustering schemes such as the Expectation Maximization binary clustering technique [28], [29] or using hidden Markov models (hereafter, HMM) typically with 2 or 3 distinct behavioural modes ~~[30]–[33], and more occasionally~~through gaussian ~~mixtures models guilford_gps_2008, [28], or~~ to explicit account for time-related priors [30]–[33]. More occasionally, supervised machine learning approaches such as artificial neural networks, support vector ~~machine,~~ machines and random forests have also been used [34], [35]. We may refer the reader to [36] for a more detailed review of these methods.

Recently, deep learning methods have been suggested to be a potentially useful tool for behavioural pattern segmentation [37]. Deep learning refers to a neural network with multiple layers of processing units [38]. By decomposing the data through these multiple layers, deep neural networks may learn complex features for representing the data with a high level of abstraction at multiple scales. The trajectory of an animal being the result of complex processes at multiple ~~spatiotemporal~~ spatio-temporal scales [17], deep learning might be able to extract relevant representations of trajectories for performing tasks such as classification, segmentation or simulation. Deep learning has become the state-of-the-art framework for a wide range of problems in text, speech, audio and image processing and applications in ecology have mainly addressed image analysis and computer vision case-studies [39], [40]. Fewer studies have explored deep learning for animal trajectory data. Recurrent neural networks ~~(RNNs)~~ have been used for movement prediction [41], [42], and for the identification of representative movement patterns [43]. Very recently, an attention network has also been proposed for comparative analysis of animal trajectories [44]. Related to our study, a fully-connected network (hereafter, FCN) has been used to predict seabirds' diving in European shags, common guillemots and razorbills [21]. With a very simple ~~fully-connected network~~ FCN with 4 layers comprising hundreds of hidden nodes, this study demonstrated the improved accuracy of this approach over commonly-used behavioural classification methods. These promising results support new investigations to further explore the potential of deep learning schemes for movement ecology studies.

As in [21], this work addresses the inference of seabird diving behaviour from GPS data using Deep Learning methods. ~~From a database of about 250~~ Besides, their FCN architecture, we investigated Convolutional Neural Networks and U-Networks [45], which are state-of-the-art architectures for time series and image data processing and shall better account for the time structure of trajectory data. As case-studies, we considered two tropical seabird genus with distinct diving behaviour (Boobies vs Cormorants). The associated datasets comprised 297 foraging trajectories derived from GPS data deployed simultaneously with pressure sensors for the identification of dives~~, we trained different deep networks within a supervised setting. Two network architectures were compared (Fully Connected Network vs~~

U-Network) , and different trajectory representations were used (Time-series vs Distance Matrix) . These approaches were applied to two tropical seabird species with distinct diving behaviour (Boobies vs Cormorants) and compared to usual methods for dive prediction (Hidden Markov Model and First-Time-Passage) . We also evaluated the potential impact on the estimation of dives distribution. Our specific objectives were therefore (a) to confirm the performance of deep networks over state-of-the-art tools for dives identification, (b) to demonstrate generalization properties of trained network to predict dives of seabirds from other colonies and (c) to evaluate the benefits of a transfer learning strategy known as 'fine-tuning' for accross-species generalization.

2 Materials and Methods

2.1 Dataset

GPS and TDR devices were fitted to Peruvian boobies (jointly fitted to breeding tropical seabirds both in Peru (92 Peruvian boobies, 106 Guanay cormorants) and Brazil (37 masked boobies). Peruvian boobies (*Sula Variegata*) and Guanay Cormorant cormorants (*Leucocarbo Bougainvilli*) breeding off the coast of Peru were captured at Isla Pescadores (11.775S, 77.265 11°46'30.34"S, 77°15'51.22"W) every year in December from 2008 to 2013. Similar devices were also deployed on Peruvian boobies from 2013 and at Isla Guanuape (8.566S, 78.966 8°32'18.92"S, 78°57'42.72"W) in December 2007. 2007, while masked boobies (*Sula dactylatra*) were captured at Fernando de Noronha archipelago (3°49'9.71"S, 32°23'36.11"W) every year in April from 2017 to 2019. The GPS were attached with Tesa tape on the tail feathers for boobies and on the back feathers for cormorants for 1 to 2 days and the TDR were fixed on the bird's leg with a metal band. In total, GPS devices (Gipsy GPS, 25–30 g, Technosmart, Rome, Italy; i-gotU GPS GT 600, 25–30 g, Mobile action Tecnology Technology, NewTaipei City, Taiwan; MiniGPSlog 30 g, Earth and Ocean GPS, Kiel, Germany; Axy-trek 14g, Technosmart, Rome, Italy) and time-depth recorders (TDRs, 3 g; resolution, 4 cm; G5 CEFAS Technology, Lowesoft, UK) were fitted to 80 Peruvian boobies and 68 Guanay cormorants. The GPS recorded locations at 1-s intervals and were attached with Tesa tape on the tail feathers for boobies and on the back feathers for cormorants for 1 to 235 seabirds.

After recovery, each GPS track was split into foraging trips by selecting locations further than a given distance to the colony and longer than a given time. Foraging trips were linearly interpolated to the TDR sampling resolution (i.e. 1s) and the coverage ratio was computed as in [21]. It is defined as the ratio between the number of recorded fixes and the number of fixes that should have been recorded with a perfectly regular sampling in a fixed temporal window. Amount of missing data is detailed in Table 1. True dives were defined by depth measured by TDR higher than 2 days. The TDR recorded depth at 1-s intervals and were fixed on the bird's leg with a metal band. Each GPS track was then splitted into foraging trips meters. Each GPS position was thus associated with a boolean value detailing the 'dive' status. This dataset consists therefore in a total of 234 297 foraging trips of seabirds with doubled-deployment GPS and TDR (see Table ??1).

Table 1: **Dataset Overview** - General statistics on the four linearly-interpolated datasets used in this study. (m ± s) is for respectively mean and standard deviation. Dives refer to the proportion of positions labeled as 'dive' (TDR-derived depth higher than 2 meters). Gaps consist in the proportion of missing fixes that have been linearly interpolated. Resting has been defined as the proportion of time with speeds inferior to 1 m.s-1 associated to non-diving behaviour.

Species	Colony Location	Nb of trips	Trip Duration (min)	Dives (%)	Dives Duration (s)	Gaps (%)	Resting (%)
Sula variegata	Pescadores Island, Peru	132	64 ± 37	1.3 %	2.5 ± 1.3	2.2 %	4.4 %
Leucocarbo bougainvilli	Pescadores Island, Peru	79	143 ± 69	9.4 %	12.9 ± 14.1	25.5 %	36.6 %
Sula variegata	Guañape Island, Peru	22	162 ± 75	0.7 %	3.3 ± 2.5	1.5 %	6.6 %
Sula dactylatra	Fernando de Noronha, Brazil	64	491 ± 377	0.2 %	2.2 ± 1.4	6.1 %	33 %

Missing fixes in GPS data were linearly interpolated. The coverage ratio was computed as in [21]. It is defined as the ratio between the number of recorded fixes and the number of fixes that should have been recorded with a perfectly regular sampling in a fixed temporal window. True dives were defined by depth measured by TDR larger than 2 meters. To assess the impact of the sampling rate of the trajectory data, we resampled each trip every 5, 15 s and 30s. Temporal windows containing at least one dive were classified as dives.

The two datasets from Pescadores Island, corresponding to the two study species, were splitted into training, validation and test datasets with respective size of 70% , 20% and 10%. The dataset from Guañape was only used as test dataset in order to evaluate the the generalization performance of trained networks, i.e.their ability to deal with datasets that have not been used during the training process.

2.2 Deep Neural Network Architectures

In this work, we investigated deep neural networks. As baseline architecture, we considered the fully-connected network (FCNet) proposed in [21]. Besides, as the considered problem may be stated as a segmentation issue, an adaption of a U-Net architecture naturally arised as a state-of-the-art solution described in Figure 1 we considered a fully convolutional neural network (CNNet) and a U-shape network particularly adapted to segmentation problems (UNet) [45]. We describe below these two three architectures and the associated supervised training procedure. We refer the reader to [40] for an introduction to deep neural networks dedicated to ecologists.

2.2.1 Fully-Connected Network (FCN FCNet)

The first architecture implemented was similar to the fully-connected network presented by [21]. As input vector, we used the concatenation of longitude, latitude step speed, turning angle and coverage time series for over a 20-second window. This input vector is fed to a layer of 100 nodes followed by 3 layers of 500 nodes. Each node applies a linear transformation to the incoming data and a non-linear activation chosen as a Rectified Linear Unit (ReLU - $\text{RELU}(x) = \max(0, x)$). The last layer applied a softmax binary function so that the output of the vector is a time series of values between 0 and 1, which can be interpreted as binary classification probabilities. This architecture is a classical example of a so-called multilayer perceptron, with a rectified linear activation which is the default activation in deep learning architectures. Overall, this architecture involves 5.10^5 500k parameters.
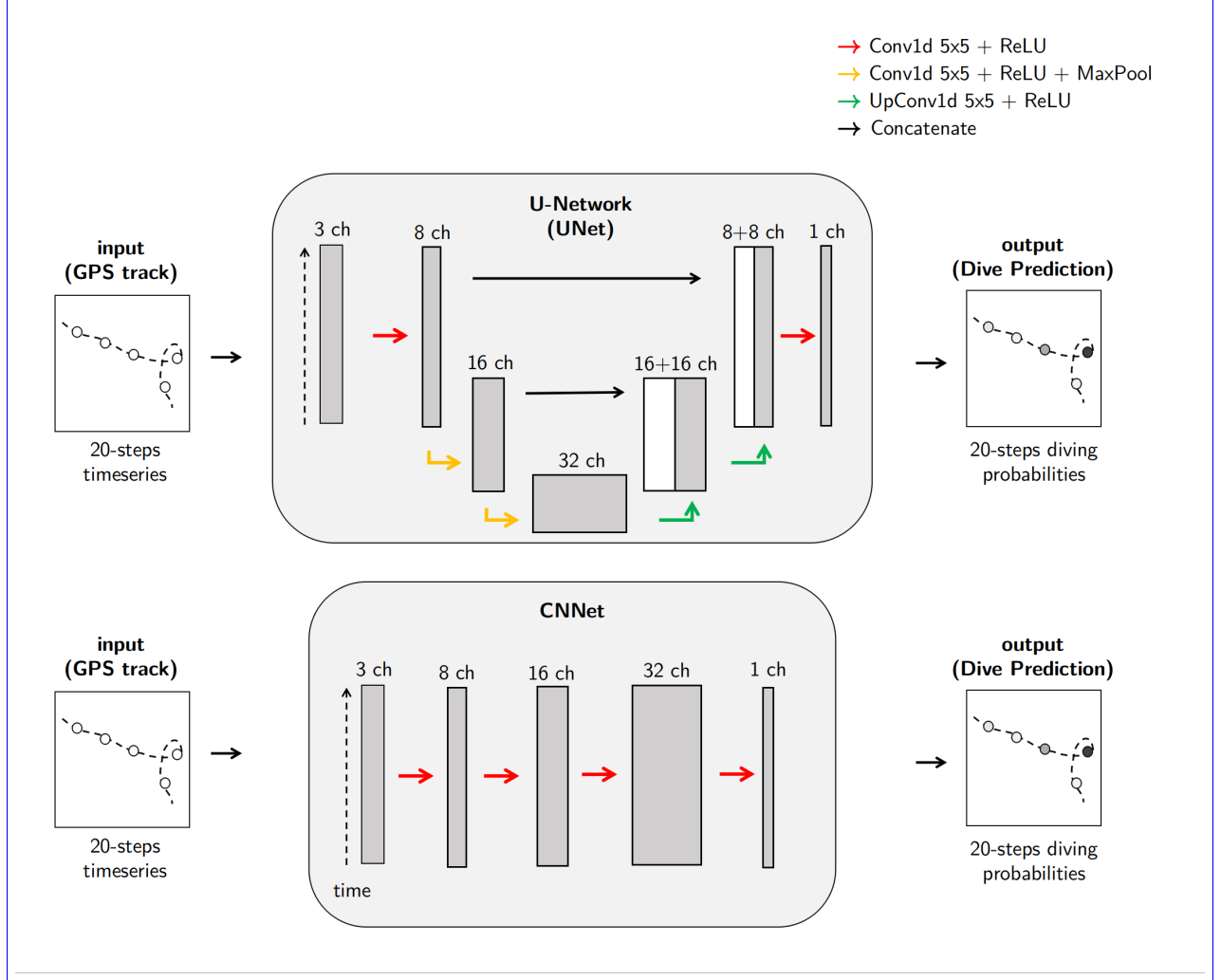
Figure 1: **Network Architectures** - CNNet refers to a fully convolutional neural network. UNet refers to a U-shape network. A channel refer to deep learning terminology and describes a representation of the input data as output of some computation layer. Conv1d, MaxPool, and UpConv1d are abbreviations for usual deep learning operations. Details can be found on pytorch's documentation [46]

2.2.2 ~~U-Network with Distance Matrix Encoder~~ Fully Convolutional Networks (~~DME-UNet~~CNN)

~~As described in Figure **??**, we also introduced a novel architecture based on a U-Network (UNet) [45]. This architecture exploits a Distance Matrix Encoder (DME) to represent geometrical features along a trajectory. Similarly to the FCNet architecture~~Convolutional networks exploit convolutional layers and are the state-of-art architectures for a wide range of applications, especially for signal and image processing tasks [47]. Thus, we investigated a basic neural network fully composed of convolutional layers. Similar to FCNet, its input vector is the concatenation of ~~longitude, latitude~~ step speed, turning angle and coverage time series over a 20-second window ~~and it outputs~~ but its output is a vector of diving probability of the same length. Overall, this architecture U-Net involves 5k parameters.

~~**DME-UNet Architecture** - This network is composed of two blocks entitled Distance Matrix Encoder (DME) and U-Network (UNet). It takes as input a GPS of 20 successive positions and outputs a diving probability to each of these positions.~~

7

A channel refer to deep learning terminology and describes a representation of the input data as output of some computation layer. Conv1d, Conv2dn MaxPool, and UpConv1d are abbreviations for usual deep learning operations in convolutional networks. Details can be found on pytorch's documentation [46]

The DME takes as input a distance matrix. This distance matrix is simply obtained by computing the orthodromic distance between each pair of positions in the input track. Its lines and rows are ordered in increasing time. The idea behind this representation of trajectory data is that longitude and latitude time-series may not be the best representations for neural-network-based methods since movement trajectories inherently contain 2D spatial information which might not be easily perceivable in simple time-series. The performance of machine learning methods is heavily dependent on the choice of data representation (or features) on which they are applied [48] . We expect the distance matrix representation to better encode 2D spatial patterns along a trajectory. Given a matrix distance, the DME applies a 2-dimensional convolutionnal layer with 8 channels (i.e. 8 output features) combined with rectifier linear unit activations. As a feature extraction step, we then sum the resulting output over the horizontal direction to extract a 8-dimensional feature vector at each position of the considered trajectory segment. In short, this network can be regarded as a trainable function that would compute residence time defined as in **barraquand_animal_2008**, but with circle of different radius which values will be adjusted automatically by the network.

We adapt

2.2.3 U-Network (UNet)

As the considered problem can be seen as a segmentation task, a U-Net architecture , which is the naturally arises as a state-of-the-art neural architecture for segmentation tasks [45], to a multivariate 1-dimensional setting. As inputs, we provide the U-Net with a 11-dimensional time series which concatenate the raw longitude, latitude and coverage ratio time series and the output of the DME. solution [45]. The key feature of the U-Net this architecture is to combine the information extracted by convolutional blocks applied at different temporal scales. To achieve this multi-scale analysis, the U-Net applies pooling layer to coarsen the time resolution and interpolation layers (UpConv1d layers) to increase the time resolution as sketched in Fig.??1. At each scale, we apply a specific convolution block. We concatenate its output with the interpolated output of the coarser scale to a convolutional block, whose output is interpolated to the finer resolution. Overall, we may notice that the output of the U-Net architecture is a time series with the same time resolution as the input time series. Similarly to the fully-connected architectureFCNet and CNNet, the last layer applies a sigmoïd sigmoid activation to transform the output into a time series of diving probabilities. Overall, this architecture (DME + U-Net) involves 2.10^4 parameterswhich is approximately 30 times less than for the FCNetinvolves 20k parameters.

2.2.4 Network Training and Validation

Given a selected neural network architecture, the training procedure relies on a supervised learning scheme using a weighted binary cross entropy as loss function. This function evaluates the performance of a prediction by comparing

the dive prediction (output of the model) with the true dives defined by TDR data. We consider a weighted version of the binary cross entropy because of the unbalanced presence of dive and no-dive behaviour in the studied trajectories (see Table ??1). The objective is to penalize more for mistakes on the smaller class (diving behaviour) than for false positive, thus ensuring for convergence. In the reported experiments, the weight was empirically set to 5 for cormorants datasets and 30 for boobies dataset.

The minimisation of the training loss exploits the Adam stochastic optimizer [49]. Networks were evaluated on training and validation datasets every epoch (defined as one pass through the entire train dataset). We consider an early-stopping criterion such that the training procedure was stopped as soon as the validation loss started increasing. Overall, given a trajectory the diving probability at a given location was assessed by computing the mean probability of all predictions derived from all 20 positions windows. These models were implemented, trained and tested with python using pytorch library [46]. Our pytorch Code is available on our github repository: https://github.com/AmedeeRoy/BirdDL.

2.3 Benchmarked methods

Two classical methods for dive prediction First-Passage Time (FPT), and Hidden Markov Models (HMM) were evaluated for intercomparison purposes. FPT was computed following [50], by selecting the radius that maximizes the variance of passage times. Time passage values were converted into a probability of dives with min-max normalization. Regarding HMMs, we applied the momentuHMM R package [32]. We implemented HMMs with 3 ~~behavioural modes~~ (resp. 4) behavioural modes for boobies (resp. cormorants) associated to traveling, searching~~and diving~~, diving and resting behaviours. This approach represents trajectories as a sequence of steps and angles. It models steps as random variables following a gamma marginal distribution and angle following a von mises marginal distribution. ~~In this HMM setting, the coverage data was used as covariate in the transition probability matrix, assuming that low coverage ratio might provide a proxy of the likelihood of the diving behaviour and vice-versa.~~ We may point out that the HMMs direclty provide as outputs a likelihood value of the diving behaviour.

2.4 Evaluation scheme

We describe below the evaluation scheme we implemented to assess the performance of the proposed neural network approaches. We first ~~focuses~~ focus on the benchmarking of the performance of the considered approaches in terms of dive prediction accuracy for different data input. For the proposed neural network architectures, we further analyze ~~the impact on the~~ their generalization properties. The methodological framework is exposed in Figure 2.

(a) Network training We assessed the dive prediction performance of ~~different data types used as inputs as well as their generalization performance~~ the 5 benchmarked methods (FPT, HMM, FCNet, CNNet and UNet) considering trajectory data derived from the two dataset from Pescadores Island (see Tab. 1). To test for the effect of temporal resolution, the two datasets have been downsampled every 5, 15 and 30s. When downsampling, temporal windows containing at least one dive were classified as dives. Each dataset were then splitted into training, validation and test
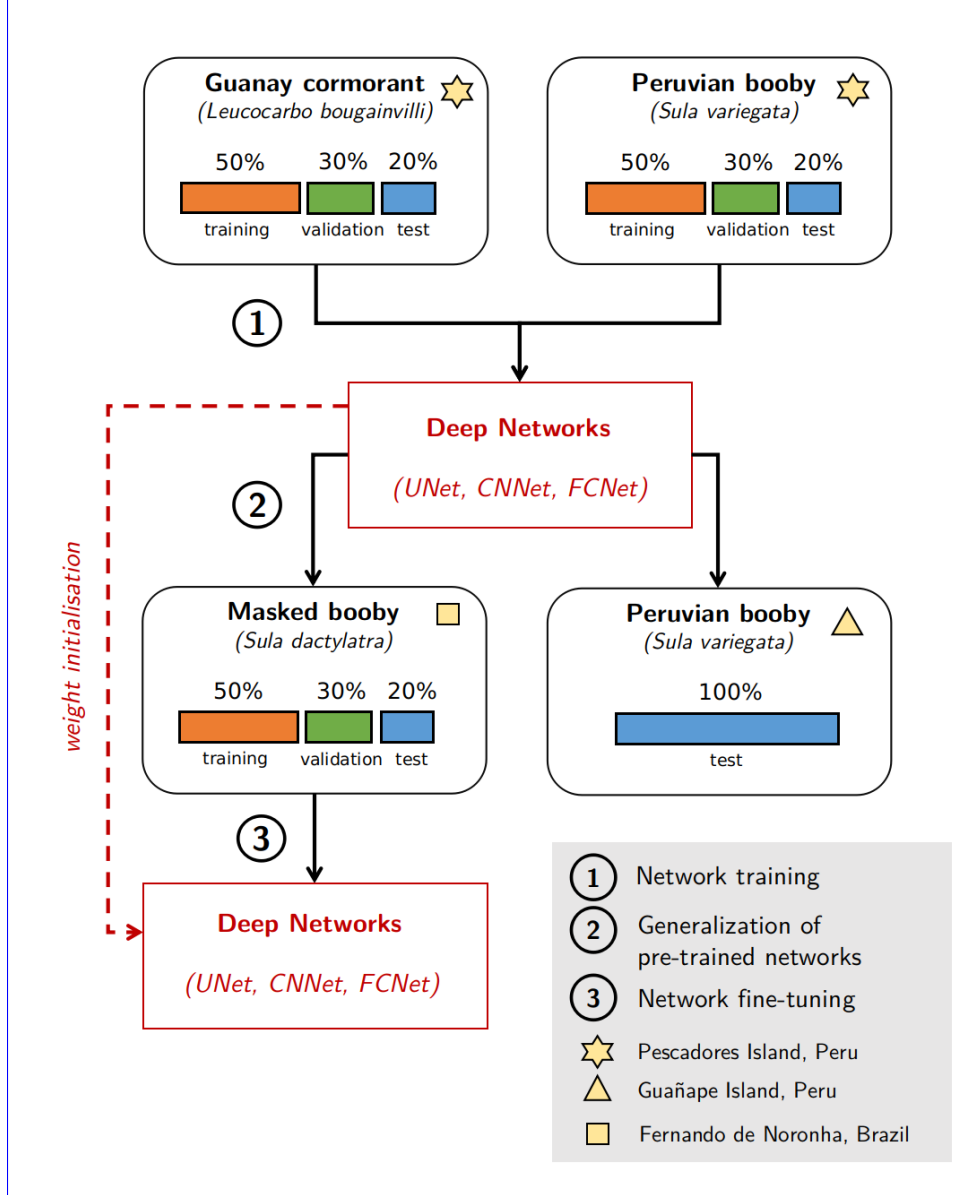
Figure 2: **Evaluation scheme** - (1) The datasets from Pescadores Island (see Tab. 1) have been used to train, validate and test deep networks. UNet, CNN and FCNet refer to the deep network architectures used in this study (see Fig. 1). (2) The trained networks have been directly used to predict dives on two other datasets without any additional training (Datasets from Guañape Island and Fernando de Noronha). (3) The dataset from Brazil have been used to train, validate and test deep networks. However, the deep networks previously obtained at step (1) have been used for weight initialization. This is known as Fine-tuning.

datasets with respective size of 50%, 30% and 20%. Deep networks were trained and selected based on the training and validation datasets. All approaches were finally compared on the testing dataset. Overall, this led to the quantitative comparison of the performance of 5 models on 6 datasets all listed in Table 2.

Dive prediction performance   As evaluation metrics for dive prediction, we evaluated the receiver operating characteristics curve (ROC), the area under the curve (AUC) as well as the binary cross entropy (BCE) for the test datasets.

10

For binary classification, the ROC curve plots the which describes the performance of a binary classifier. It consists in plotting the true positive rate (i.e. true predicted dives) against the false positive rate (i.e. false predicted dives). We obtain this curve by varying the probability threshold defining dive/no dive behaviours. Moreover, we evaluate the area under the curve (AUC) as well as the binary cross entropy (BCE) for the test datasets. Regarding the AUC, it was estimated by integrating the ROC curve along the x axis using the composite trapezoidal rule. For neural network approaches, we also analyzed the value of the training loss for the training and test datasets.

Data inputs(b) Generalization performance of pre-trained networks Both studied species breed in the most productive upwelling system, the Humboldt Current System (HCS) and feed on the same preys, i. e. Peruvian anchovies [51]. However, they are known to have distinct foraging strategies: boobies are plunge divers reaching in average about 2 m depth and spending most of the time in fly, while cormorants dive deeper and longer in average, reach up to In this section, we evaluated the generalization performance through the application of models trained on Pescadores dataset to data that have not been used during the training process. In particular, we evaluate previously fitted deep networks performance on the two datasets from Guañape Island and from Fernando de Noronha Archipelago, composed of Peruvian boobies and masked boobies trajectories, respectively. In this experiment, we compared the dive prediction performance of FPT and HMM methods to the best FCNet, CNN and UNet models. Beyond AUC and BCE performance metrics, we also evaluated the relevance of the estimated maps of dive distributions. The later were computed using a weighted Kernel Density Estimator (KDE) using dive probabilities as weighing factor. As groundtruth, we considered the map of dive distributions estimated from true dive locations defined by TDR data. Kernel densities were estimated using a $0.01\times0.01°$grid and a bandwith of $0.25°$. From these maps, we evaluated an means square error (MSE) as an integrated performance metrics for the different approaches [52].

(c) Network Fine-Tuning Transfer learning refers to the fact of using knowledge that was gained from solving one problem and applying it to a new but related problem. For this purpose, a solution known as 'fine-tuning' consists in using a trained model as the initialization of the training scheme rather than training a new model from scratch. Thus, we evaluated the benefits of fine-tuning for predicting dives of the 15s-resampled Brazilian dataset (Tab. 1) based on the deep networks fitted on the dataset from Pescadores. The Brazilian dataset was therefore split into training, validation and test datasets with respective size of 50%, 30m depth, and spend up to 40% of the time in water [53]. We assessed the dive prediction performance of the benchmarked methods when considering trajectory data derived from these two distinct seabird and associated with different time resolutions, more precisely trajectory data sampled at % and 20%. We then trained models from scratch and using fine-tuning for the three studied network architectures and following the learning procedure presented before. We also evaluate the impact of the training dataset size by randomly selecting respectively 1, 5, 15 and 30s. We specifically evaluated the relative importance of coverage data and of the Distance Matrix Encoder by training neural network approaches with and without this information. As mentioned above, we also applied the HMM with and without coverage data used as covariate in the transition probability matrix.

273 ~~Overall, this led to the quantitative comparison of the performance of 17 models and 6 datasets all listed in Table ??~~30

274 foraging trips for the training step. All models were finally compared to HMM and FPT methods on the testing dataset

275 and using the AUC evaluation metric.

Table 2: **Deep Networks Training** - Performance metrics for all trained deep networks on the trajectories of Pescadores along with benchmarked methods used for comparison. AUC means the Area Under the ROC curve, BCE is for binary cross entropy computed on the testing trajectories. Train and Validation Loss correspond to the loss computed after model training on respectively training and validation datasets. SV is for Peruvian boobies (*Sula variegata*), LB is for Guanay cormorants (*Leucocarbo bougainvilli*)

Dataset	Resolution	Model	AUC	BCE	F-score	Train Loss	Validation Loss	Reference Name
SV (Pescadores)	5s	FPT	0.62	0.70	0.55	-	-	-
		HMM	0.86	1.07	0.69	-	-	-
		FCNet	0.89	0.38	0.81	0.61	0.61	SV_FCNet_5s
		CNNet	0.94	0.29	0.88	0.48	0.49	SV_CNNet_5s
		UNet	**0.96**	**0.23**	**0.91**	**0.48**	**0.45**	**SV_UNet_5s**
	15s	FPT	0.71	0.79	0.66	-	-	-
		HMM	0.87	2.39	0.84	-	-	-
		FCNet	0.82	0.81	0.80	1.35	1.16	SV_FCNet_15s
		CNNet	0.91	0.58	0.85	0.89	0.86	SV_CNNet_15s
		UNet	**0.93**	**0.57**	**0.86**	**0.87**	**0.79**	**SV_UNet_15s**
	30s	FPT	0.73	0.97	0.70	-	-	-
		HMM	0.84	1.22	0.68	-	-	-
		FCNet	0.82	1.10	0.79	1.69	1.74	SV_FCNet_30s
		CNNet	0.85	0.98	0.80	1.55	1.47	SV_CNNet_30s
		UNet	**0.91**	**0.73**	**0.86**	**1.12**	**1.10**	**SV_UNet_30s**
LB (Pescadores)	5s	FPT	0.61	1.59	0.57	-	-	-
		HMM	0.78	1.42	0.72	-	-	-
		FCNet	0.87	0.40	0.79	0.55	0.67	LB_FCNet_5s
		CNNet	0.92	0.30	0.84	0.48	0.57	LB_CNNet_5s
		UNet	**0.93**	**0.28**	**0.85**	**0.46**	**0.54**	**LB_UNet_5s**
	15s	FPT	0.58	1.73	0.62	-	-	-
		HMM	0.76	3.35	0.72	-	-	-
		FCNet	0.67	0.77	0.75	0.85	0.94	LB_FCNet_15s
		CNNet	0.89	0.43	0.85	0.60	0.73	LB_CNNet_15s
		UNet	**0.90**	**0.37**	**0.83**	**0.52**	**0.76**	**LB_UNet_15s**
	30s	FPT	0.56	1.81	0.62	-	-	-
		HMM	0.75	2.90	0.74	-	-	-
		FCNet	0.65	0.92	0.75	0.96	1.10	LB_FCNet_30s
		CNNet	0.74	0.74	0.76	0.86	1.01	LB_CNNet_30s
		UNet	**0.88**	**0.37**	**0.83**	**0.97**	**1.05**	**LB_UNet_30s**

276 **Generalization properties** ~~When considering neural network approaches, training models which may apply beyond~~

277 ~~the considered training framework is a key feature, generally referred to as the generalization performance of the~~

278 ~~trained neural networks. Here, we evaluated this generalization performance through the application of models trained~~

279 ~~on Pescadores dataset to the Guañape dataset. This dataset is composed of Peruvian boobies trajectories but from a~~

280 ~~different breeding colony. Beyond the different in the case-study region, we expect peruvian boobies behaviour to~~

281 ~~share common features across colonies, so that models fitted to the trajectory data of a given colony might still be~~

282 ~~meaningful when applied to another colony. In this experiment, we compared the dive prediction performance of FPT~~

283 ~~and HMM methods to the best FCNet and DME-UNet models. Beyond AUC and BCE performance metrics, we also~~

evaluated the relevance of the estimated map of dive distributions. The later were computed using a weighted Kernel Density Estimator (KDE) using dive probabilities as weighing factor. As groundtruth, we considered the map of dive distributions estimated from true dive locations defined by TDR data. From these maps, we evaluated an Hellinger distance as an integrated performance metrics for the different approaches [52].

3 Results

We ~~detailed~~ detail below the numerical experiments performed in this study to assess the relevance of the proposed neural network approaches to predict dive behaviour of ~~Peruvian~~ boobies and cormorants from trajectory data.

~~Overall performance~~(a) Network Training ~~For all datasets,~~ On the Pescadores Island dataset used for network training, we reported a contrasted performance of the different methods~~obtained contrasted performance results~~, with AUC going from ~~0.71 to 0.97~~ 0.61 to 0.96 (see Table ~~??~~2), which corresponds in the best cases to correct prediction rates of diving and non-diving behaviour of approximately ~~90~~95% and of ~~70~~60% in the worst cases (see Figure ~~??~~3). Overall, ~~neural networks-based approaches obtained better performance than the state-of-the art methods, with highest AUC (around 0.9) and lowest binary cross entropy (around 0.5). First-Passage Time obtained quasi systematically the lowest AUC, and the Hidden Markov Models had most of the time AUC around~~ all methods performed better at predicting the dives of boobies than those of cormorants. The UNet obtained systematically the best prediction performance with averaged AUC of 0.93 (resp. 0.90) for boobies and cormorants respectively. The CNN also achieved very good predictions, consistently performing at least as well as state-of-the-art tools with averaged AUC of 0.9 (resp. 0.85~~but the highest BCE. State-of-the-art methods were outperformed by neural networks specifically for the 5s resolution datasets, with AUC improvements up to 15% with neural networks over HMM. In particular, the DME-UNet was the most consistent method, being able to get better performance on most datasets, and to improve substantially the performance of the FCNet proposed by [21]. As illustration, on~~). The lowest performance was reported for the FPT approach, which never predicted dives with AUC higher than 0.73. The HMM obtained relatively good performance on the ~~boobies dataset, BCEwas approximately two times lower with the DME-UNet than with the FCNet, and the DME-UNet obtained systematically higher AUC .~~ boobies dataset with AUC indices around 0.85, yet it did not get AUC higher than 0.76 on the cormorants dataset. It also obtained the highest BCE, approximately 2 to 10 times higher than the UNet. Regarding the FCNet, the AUC index ranged from 0.65 to 0.89, showing a much greater variability than for CNN and UNet architectures. The best neural network predictions for Pescadores dataset are illustrated in Fig. 4.

~~Comparison of algorithms~~ ~~ROC curves obtained from the prediction of 4 algorithms, First-Time Passage (FPT), 3-states Hidden Markov Models (HMM), Fully-Connected Network (FCN) and U-Network with a Distance Matrix Encoder (DME-UNet) on 4 distinct test datasets derived from two seabirds species breeding in Pescadores Island from 2008 to 2013~~

Impact of the temporal resolution ~~Interestingly, on the cormorant dataset the sampling resolution~~ Interestingly, deep networks were relatively sensitive to the sampling resolution, whereas it did not affect much the performance of the ~~neural network approaches (DME-UNet with AUC of 0.95) whereas state-of-the-art methods FPT~~ /HMM ~~had increasing AUC with larger resolution going from respectively 0.73/0.78 to 0.79/0.86 (see Table ??). At the opposite, on the peruvian datasets the~~ FPT /HMM ~~methods had constant performance disregarding the sampling resolution (AUC around 0.76/0.88), while deep learning tools had decreasing AUC with larger resolution. In particular, the dataset of Peruvian boobies with the larger resolution (i. e.~~ and HMM approaches. For both species, the higher the resolution, the better the performance for UNets, CNNs and FCNets. For instance, we reported a mean AUC of 0.92 for a 5s resolution vs. 0.81 for a 30s ~~) is the only one where neural networks did not do any better than~~ resolution. This was particularly true for the CNN and mostly for the FCNet which did not performed better than HMM on the 30s-resoluted datasets, whereas they were able to outperform state-of-the-art ~~. Yet, if DME-UNet's performance are pretty much equivalent to the highest AUC obtained by HMM (AUC of 0.87) , on this dataset the FCNet had lowest prediction performance (AUC of 0.76)(see Figure ??).~~ approaches on the 5s-resoluted datasets.

Impact of data inputs ~~Deep networks had substantial increase of accuracy through the coverage ratio data, all three networks FCNet, UNet and DME-UNet having very high prediction accuracy with AUC around 0.95 (see Table ??). Yet, without coverage information, accuracy was way poorer with AUC around 0.7, except for the network using a Distance Matrix Encoder (DME-UNet)that maintains an AUC at 0.92. At the opposite, the removal of coverage data did not change significantly the performance indexes of the HMM with a relatively constant AUC of 0.87 (see Figure ??).~~

Maps of diving probabilities ~~We further analysed the relevance of the dive predictions through the estimated dive distributions maps reported in Figure ?? and ??. For both species, we can observe neural networks and HMM approaches have contrasted outputs (either very high or low diving probabilities), while the FPT approach is the one that discriminate the least dive from non-diving behaviours. Moreover, these maps also point out that from all methods false positive are predicted with high probabilities in areas where no true dives occurred (i.e. isolated blue circle with high radius). Yet, it is particularly visible from FCNet and HMM output and specifically in the vicinity of the colony.~~

Maps of Guanay cormorants predicted dives ~~Maps of the testing trajectories. Red points represent true dive derived from TDR data. Blue points represents diving probabilities of each location with radius increasing for higher probabilities. These probabilities are the results of four methods on the same Dataset (LB 5s), with our proposed network DME-UNet (top left), the fully connected network used by [21] FCNet (top right), the 3 states Hidden Markov Models HMM (bottom left), and the First Time Passage apporach FPT (bottom right)~~

Application to Guañape dataset ~~The use of deep networks trained on the dataset of Peruvian boobies from Pescadores Island on trajectories from birds from another colony (i.e. Guañape) revealed once again the higher~~

14

~~**Comparison of data inputs** - ROC curves obtained from the prediction of 3 deep networks, Fully-Connected Network (FCN) and U-Network with/without a Distance Matrix Encoder (resp. UNet / DME-UNet) on the test dataset derived from Peruvian boobies sampled at 5s breeding in Pescadores Island from 2008 to 2013~~
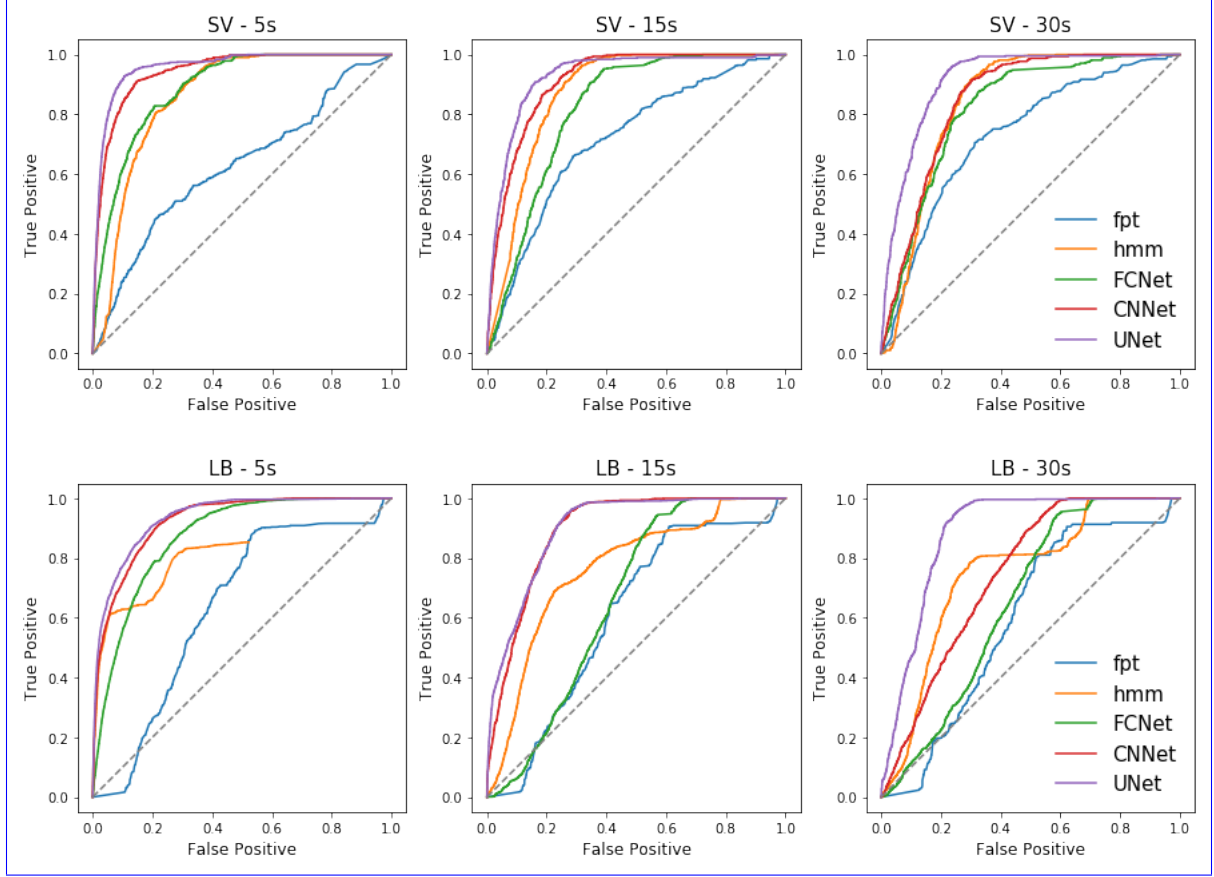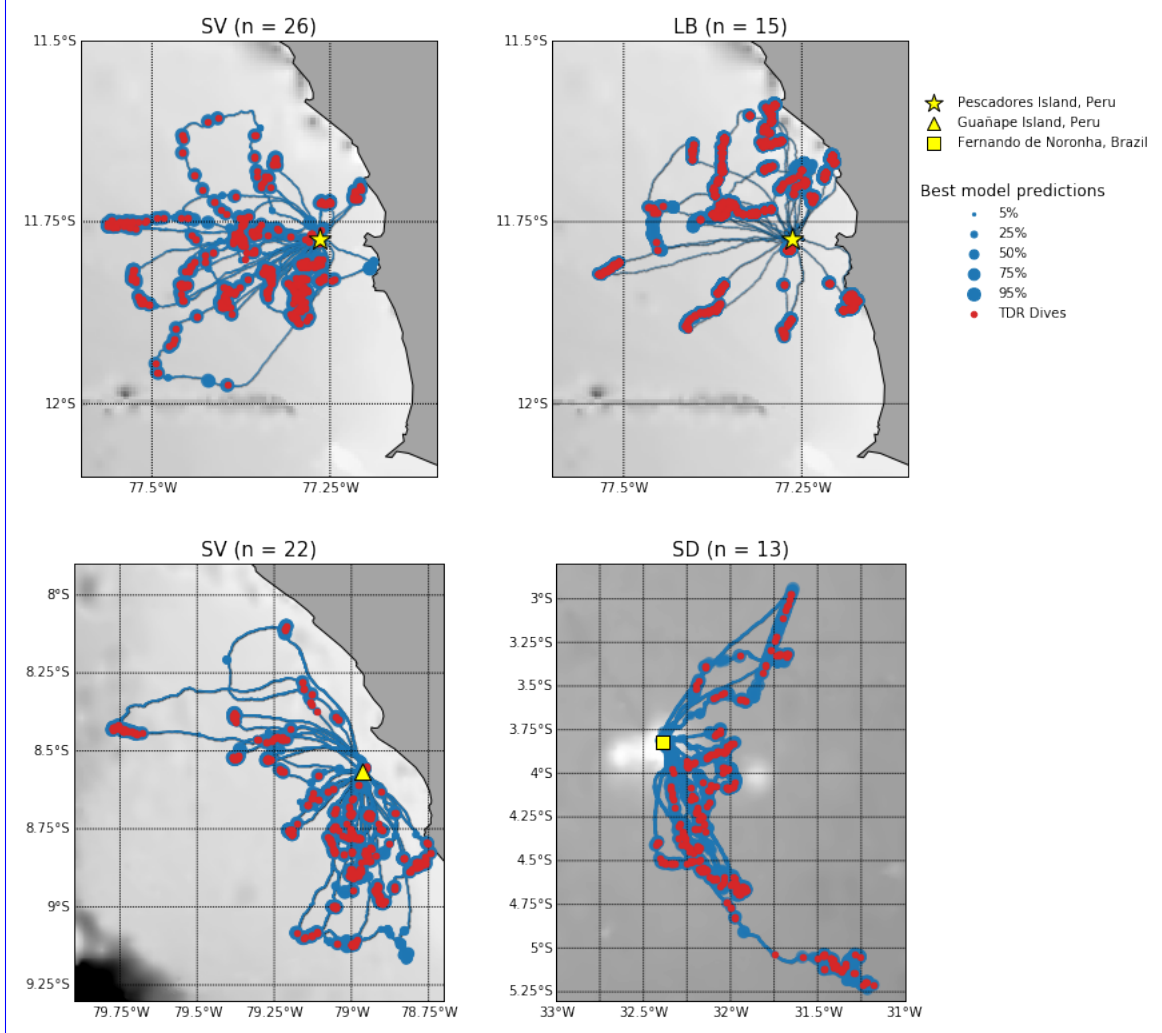

Figure 3: **Performance of Deep Networks on Pescadores Dataset** - ROC curves obtained from the prediction of 5 algorithms, First-Time Passage (FPT), Hidden Markov Models (HMM), Fully-Connected Network (FCNet), Fully-Convolutional Network (CNN) and U-Network (UNet) on 2 distinct test datasets resampled at 3 different resolutions (5, 15 and 30s) derived from two seabirds species breeding in Pescadores Island from 2008 to 2013. SV stands for Peruvian boobies (*Sula variegata*), and LB stands for Guanay cormorants (*Leucocarbo bougainvilli*).

349 ~~accuracy of deep learning based algorithms (see Figure 3) . Both FCNet and DME-UNet outperformed HMM and~~

350 ~~FPT in terms of AUC~~

351 **(b) Generalization performance of pre-trained networks** Overall, all networks trained with data from Pescadores

352 reported a AUC performance higher than 0.78 (resp. 0.56) when applied to Guañape (resp. Fernando de Noronha)

353 dataset (Fig. 5). AUC performance averaged 0.85 when using deep networks trained with the boobies dataset and

354 0.72 with cormorants data (see Tab. 3). On both datasets, the best models were UNet and CNN models trained from

355 boobies data with respectively AUC scores of 0.98 and 0.87. In particular, they outperformed the HMM that were

356 specifically fitted to Guañape and Fernando de Noronha data. By contrast, the FCNet used by [21] that obtained better

357 results than HMM on the Pescadores dataset (AUC of 0.89 vs 0.86 for HMM) did not predict better than HMM when

358 used at Guañape (e.g. ~~, with AUC of around 0.97 for neural networks vs . 0.88~~ AUC of 0.89 vs 0.92 for HMM). The

359 ~~Hellinger distance of~~ MSE for the estimated dive distribution maps stressed the greater relevance of ~~DME-UNet~~ UNet

15

Figure 4: **Maps of predicted dives for all 'test' datasets** - Red points represent true dive derived from TDR data. Blue points represent diving probabilities at each location with radius increasing for higher probabilities. These probabilities are the outputs of the best deep networks for each dataset: Peruvian boobies from Pescadores (top left), and from Guañape Island (bottom left), Guanay cormorants from Pescadores Island (top right), and masked boobies from Fernando de Noronha archipelago (bottom right). SV stands for Peruvian boobies (*Sula variegata*), LB for Guanay cormorants (*Leucocarbo bougainvilli*) and SD for masked boobies (*Sula dactylatra*). Bathymetry is shown in grey.

360 predictions with a ~~distance~~ MSE value 1.6 times smaller than the one derived from ~~FCNet estimations and 2~~ CNN

361 estimations and 1.9 times smaller than the one derived from HMM ~~estimation.~~ estimations (Tab. 3). As illustrated in

362 Fig. 6, only the Unet did not overestimate the number of dives in the vicinity of the colony as well as an other foraging

363 area southward from the colony.

16

Table 3: **Deep Network Testing** - The deep networks fitted on the dataset from Pescadores have been used for dive prediction in Guañape and in Fernando de Noronha. Deep networks are described by their reference name (see Table 2). AUC is for area under the roc curve. BCE is the binary cross entropy. MSE corresponds to the mean square error of the diving distribution maps estimated with kernel density estimations and plotted in Figure 5 to the correct diving distribution. SV is for Peruvian boobies (*Sula variegata)*

Dataset	Resolution	Model	AUC	BCE	F-score	MSE
SV (Guañape)	5s	FPT	0.65	0.57	0.43	10.9
		HMM	0.92	2.46	0.88	7.7
		SV_FCNet_5s	0.89	0.31	0.80	7.7
		SV_CNNet_5s	0.97	0.20	0.91	6.5
		SV_UNet_5s	**0.98**	**0.10**	**0.91**	**4.0**
		LB_FCNet_5s	0.78	0.09	0.05	13.8
		LB_CNNet_5s	0.87	0.07	0.07	7.8
		LB_UNet_5s	0.87	0.08	0.09	14.5
SD (FdN)	15s	FPT	0.50	0.75	0.22	7.5
		HMM	0.84	1.86	0.81	3.4
		SV_FCNet_15s	0.63	1.17	0.71	4.8
		SV_CNNet_15s	**0.87**	**0.59**	**0.83**	**3.8**
		SV_UNet_15s	0.73	0.58	0.55	5.8
		LB_FCNet_15s	0.56	0.61	0.48	6.5
		LB_CNNet_15s	0.62	0.27	0.08	12.9
		LB_UNet_15s	0.63	0.18	0.08	13.2

(c) Network Fine-Tuning In this section, we evaluated the benefits of a fine-tuning strategy for the prediction of masked boobies dives. As expected, all deep networks initialized using previous models converged more quickly than deep networks trained from scratch. In particular, a dataset of 15 foraging trips (i.e. around 30k GPS positions) was enough for convolutional networks to obtain AUC of 0.9 using fine-tuning, whereas deep networks trained from scratch needed twice as many trips for the same predictive performance (Figure 7). The improvement issued from a fine-tuning was notably important for small-to-medium datasets (5-10 foraging trips, i.e. 10k to 20k GPS positions), and for the CNN. It decreased as the size of the dataset increased. From our experiments, at least 5 trips were necessary to fine-tune a relevant network compared with the HMM baseline (see for instance bottom-center in Fig. 7). The best neural network predictions for the Brazilian dataset are illustrated in Fig. 4.

4 Discussion

This study aimed at predicting seabirds dives from GPS data only ~~by training a deep network in a supersized~~ using deep neural networks trained in a supervised manner based on TDR data to define the ~~true~~ groundtruthed dives. In line with [21], this study further supports the relevance of deep learning approach over classical methods for dive predictions. ~~Moreover, we introduce a new deep network, the so-called DME-UNet, which combines a state-of-the-art deep network architecture with a meaningful geometrical representation of trajectory data. We~~ Using convolutional architectures rather fully-connected ones, we reported even better results with higher stability to the different data inputs, as well as better generalization abilities.
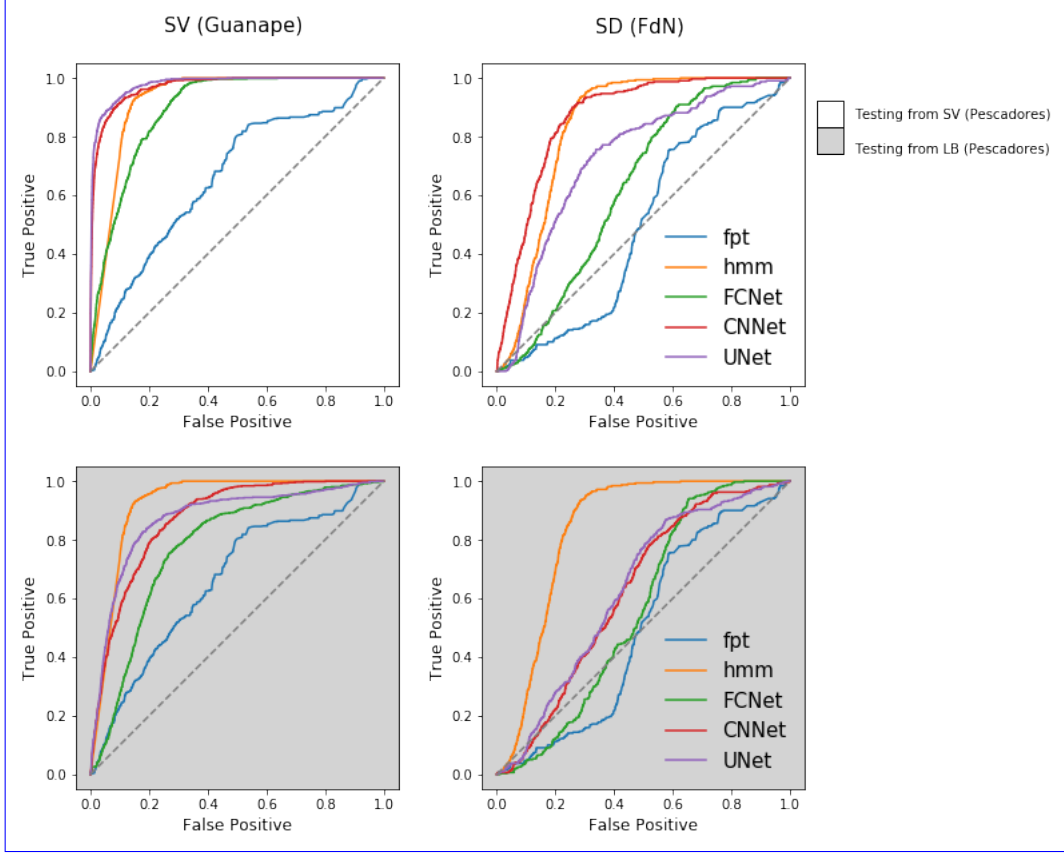
Figure 5: ~~Maps of dive distributions of Peruvian Boobies from Guañape Island~~ **Performance of Tested Deep Networks** - ~~These density maps were~~ ROC curves obtained ~~through Kernel Density Estimation. The left map has been computed~~ from ~~true dives derived from TDR data. The four other maps are estimations~~ the prediction of ~~this map5 algorithms,~~ ~~using all points of the trajectories with weights associated to diving probabilities estimated by the four studied approaches, with our proposed network DME-UNet~~ First-Time Passage (~~top left~~FPT), ~~the fully connected network used by [21] FCNet (top right), the 3 states~~ Hidden Markov Models (HMM), Fully-Connected Network (~~bottom left~~FCNet), Fully-Convolutional Network (CNNet) and U-Network (UNet) on 2 distinct test datasets. SV stands for Peruvian boobies (*Sula variegata*), and LB stands for Guanay cormorants (*Leucocarbo bougainvilli*). The deep networks used in this figure have been trained on the ~~First Time Passage apporach FPT~~ dataset from Pescadores (~~bottom right~~see Fig. 3 and Tab. 2). ~~Hellinger distance~~ They have been used for dive prediction of ~~the 4 estimations to the ground truth defined as the TRUE DIVES map are~~ Peruvian boobies in ~~Table ??~~Guañape (left column) and for masked boobies in Fernando de Noronha (right column).

381 ~~The proposed network allows not only to better predict dive behaviour but also results~~ Peruvian boobies and Guanay
382 cormorants tracked in Peru breed in a highly productive upwelling system, the Humboldt Current System (HCS) and
383 feed on the same preys, i.e. Peruvian anchovies [51]. However, they are known to have distinct foraging strategies:
384 boobies are plunge divers reaching in average about 2 m depth and spending most of the time in fly, while cormorants
385 dive deeper and longer on average, reach up to 30 m depth, and spend up to 40% of the time resting on the water
386 surface [53] (Tab. 1). By contrast masked boobies breeding at Fernando de Noronha are plunge divers similarly to
387 Peruvian boobies, yet they forage mainly in oligotrophic waters [54] and feed mainly on flying fish and flying squids
388 [55], [56]. Their foraging strategies then differ from Peruvian boobies as they perform longest trips and spend more
389 time resting at sea surface (Tab. 1). We demonstrated that for these three species, our best deep network models
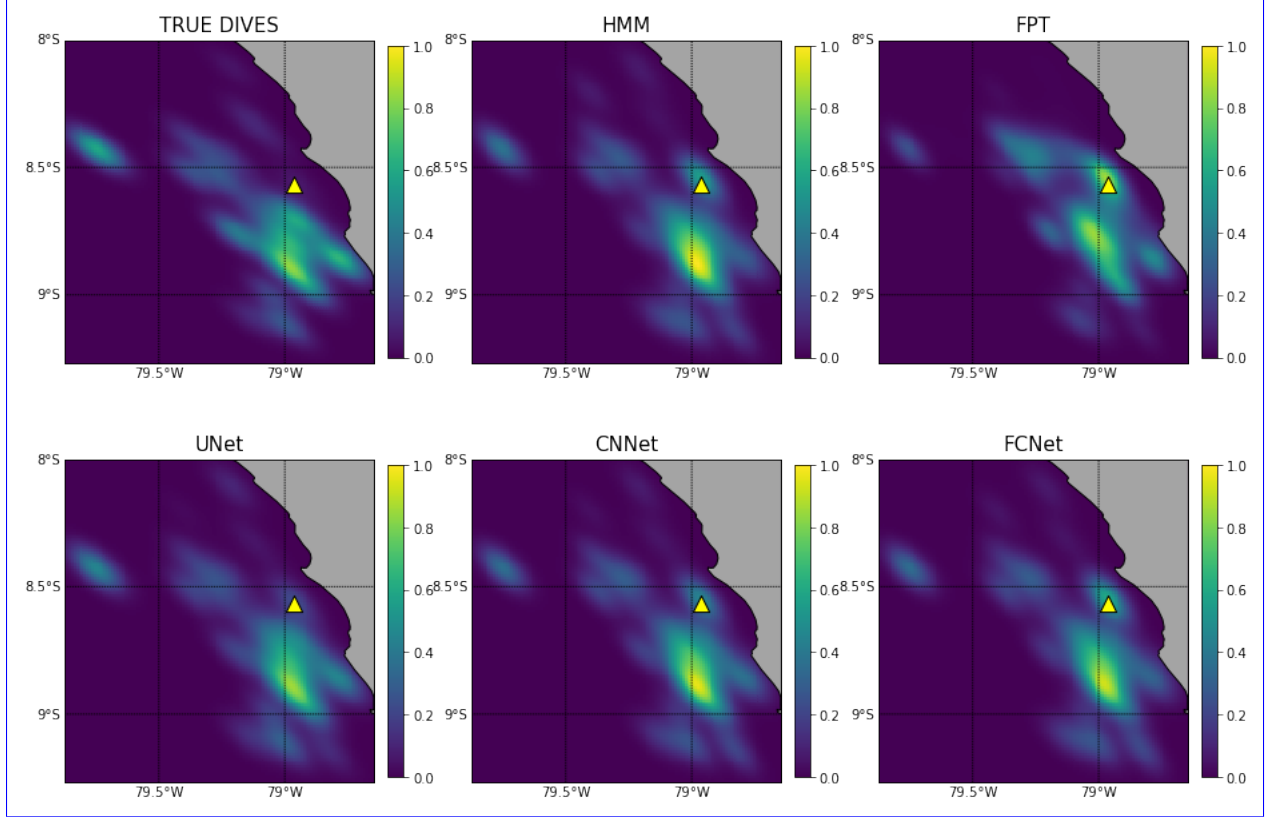
18

Figure 6: **Maps of dive distributions of Peruvian Boobies from Guañape Island** - These density maps were obtained through Kernel Density Estimation. The top left map has been computed from true dives derived from TDR data. The five other maps are estimations of this map, using all points of the trajectories with weights associated to diving probabilities estimated by the studied approaches: First-Passage Time (FPT), Hidden Markov Model (HMM), Fully Connected Network used by [21] FCNet (top right), fully-Convolutional Network (CNNet), and the U-Network (UNet). Dive map mean square error (MSE) between estimated and reference distribution are in Table 3

were able to accurately predict around 95% of dives and outperformed HMM that predicted around 85% of dives. In particular, the proposed U-shape deep network (UNet) demonstrated a greater robustness to different data inputs, as it obtained the best results whatever the sampling resolution (Tab.3).

Additionally, UNet also resulted in better seabird dive distribution maps ~(Fig. 6). Recently numerous studies used seabirds dive as a proxy for prey distribution, and such distribution are usually computed by applying KDE on dive predictions derived from HMMs [57]–[59]. Here, we show that the error in the estimation of dive distributions maps can be ~~divised~~ divided by two by using deep learning tools rather than HMM tools. In our specific study, HMMs ~~have~~ over-estimated the frequency of dives at ~~proximity of breeding locations~~ specific locations (including the vicinity of the colony). Sulids and cormorants spend time bathing near their breeding territories involving vigorous splashing and beating the water with the wings ~~,~~[55]. Such behaviours associated to low speed might be erroneously classified as diving behaviour by ~~HMMs~~ state-of-the-art tools which could explain the observed bias. ~~For the same reason, we might explain that~~ This might also explain why HMM are better ~~on the boobiesdataset than with cormorants~~ at predicting boobies' than cormorants' dives because these birds spend more time resting at the surface, which corresponds to
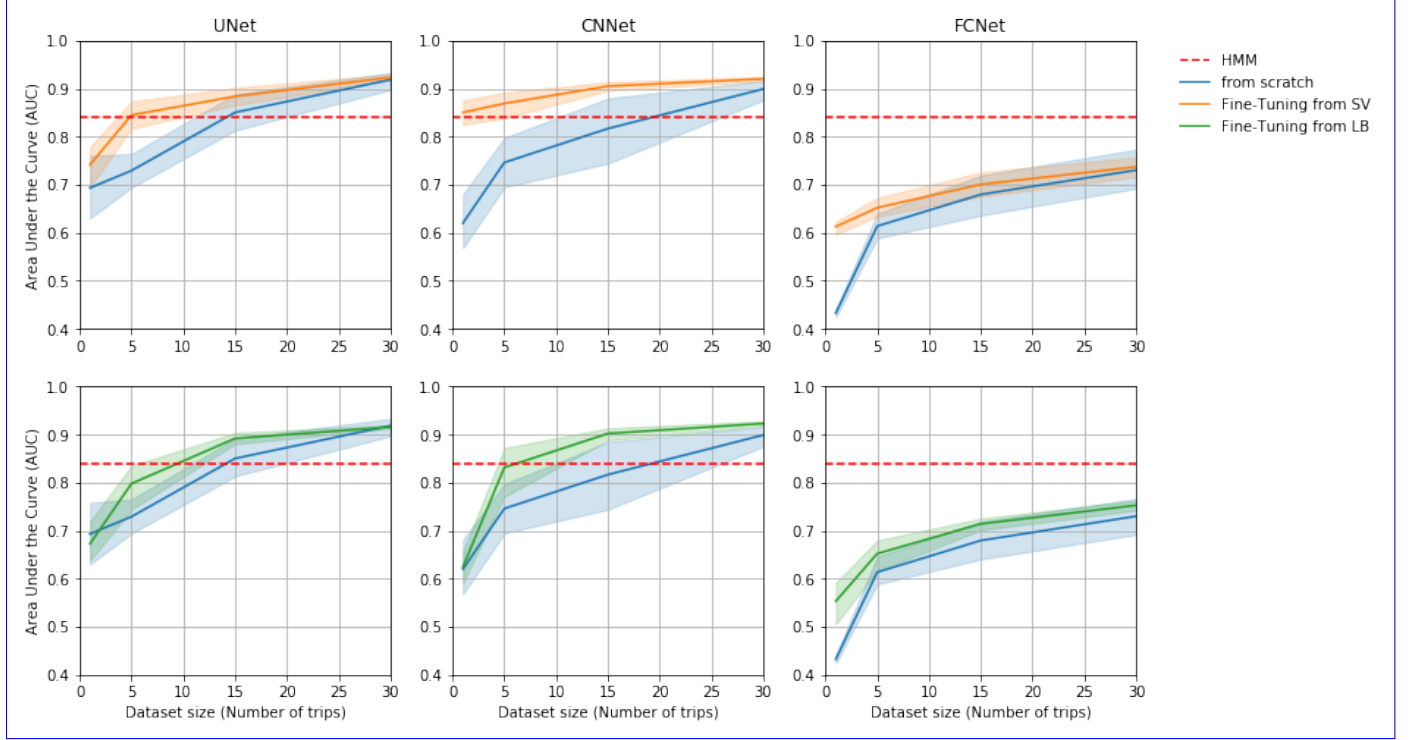
19

Figure 7: **Fine-tuning** - AUC indices of trained deep network function of dataset size.

low ~~speeds~~ speed patterns without being dives (see Table ~~**??**). The high BCE obtained by HMM also suggests that~~ ~~in these situations HMM predicts erroneously dive with high confidence. In opposition, deep learning seem better~~ ~~for discriminating such~~ 1). We may also stress that Cormorants trajectories are characterized by relatively long gaps in the regularly sampled sequence of locations, since these devices do not receive a satellite signal while submerged [30], [60]. This may in turn make more complex the analysis of Cormorants trajectories. In this respect, UNet showed a greater ability to discriminate the resting/bathing behaviours from dives~~. It would be nice to see further why and~~ ~~how deep networks succeed in outperforming HMMs, but the interpretation is one of the inherent difficulty of deep~~ ~~learning. Not that it is not possible, but that it would require further analysis.~~Yet, having such results on deep networks~~ ~~performance over HMM is a first step in the demonstration that there is some information that the HMMs misses and~~ ~~that it can be captured by deep networks.~~

, and a greater robustness to the presence of linearly-interpolated segments. Whereas HMM are mostly driven by fine-scale features (w.r.t. the considered time resolution), UNets exploit a multi-scale analysis of trajectory data and can extract relevant multi-scale information to retrieve dive. Future work could investigate further the key features extracted by UNets. As shown in Figure ~~**??**, the better performance reported with DME-UNet~~ 3, the performance of the deep networks was closely related to the temporal resolution of the sampled dataset. ~~In particular, from~~ ~~Peruvian boobies' trajectories resampled at 15s and 30s, similar results were obtained through DME-UNet and HMM~~ ~~. However, we observed that with the 5s resolution dataset, deep networks had outperformed HMM. This suggests~~ ~~that HMMs are great approaches to trajectory segmentation with relatively large temporal resolution, and that they are~~

20

~~unable to interpret movement process at very fine spatio-temporal scales~~ Whereas HMM did not succeed in exploiting higher-resolution data, UNets led to better performance when the resolution increased. This supports a greater ability of UNets both to deal with potential aliasing effects as well as ~~deep networks do. Knowing that Peruvian boobies dive during very short periods, approximately 2 seconds (see Table ??), this suggests that the deep networks would be able to capture diving behavioural movements from a GPS track sampled with a resolution in the same order of magnitude than dive durations. For the Guanay cormorants dataset, DME-UNet had better results than HMM for all tested sampling resolution. However, cormorants dives last about 20s, which could explain why even with the 30s sampled dataset deep networks had better results than HMMs. We might infer that a decrease of the performance of deep networks would occur for sampling resolution of 1 minute or more.~~ to exploit fine-scale features. With technological advances in sensor technology, ecologists are able to collect larger amount of data than ever before. We might expect GPS with lower consumption ~~,~~ and higher resolution in the future. Such an expected trend would make more critical the exploitation of ~~development of~~ the proposed deep learning approaches to make the most of the collected high-resolution animal trajectories [16], [61], [62].

~~The GPS tracks were characterized by short gaps in the regularly sampled sequence of locations, since these devices do not receive a satellite signal while submerged [30], [60]. These gaps are therefore sometimes directly considered indicative of diving behaviour [53]. Surprisingly, HMM did not take benefits from such variable for inferring dives. Nevertheless, the use of this coverage information as input helped substantially deep networks to predict dives. Indeed, without coverage data, the FCNet proposed by [21] predicted dives with relatively low accuracy, with in particular AUC of 0.71 which is low compared to HMMs' performance (AUC of 0.87), while it obtained AUC of 0.94 using this external information (see Figure ??). In their study, Browning et al. also used in some cases altitude data to get better performance than HMM. One could therefore argue that deep networks are not relevant with longitude and latitude data only. However, the DME-UNet proposed in this study demonstrated the ability of deep networks to classify animal behaviours in such situation. Indeed, without the coverage ratio variable it obtained still an AUC of 0.92 which was about 5 % more than HMM. This demonstrates that this approach is not only relevant for seabirds where coverage ratio is an important external variable, but also for any trajectory dataset.~~

~~A key breakthrough of our approach was thus to use what we called a Distance Matrix Encoder (DME), which is a simple neural network block which aims to extract features for a better geometrical description of a trajectory. The performance of machine learning methods is heavily dependent on the choice of data representation (or features) on which they are applied [48]. Most studies dealing with trajectory data represented movement with longitude and latitude time-series. Yet, some studies argued that such representation might not be relevant for neural-network based methods since movement trajectories inherently contain 2D spatial information which might not be easily perceivable in simple time-series. For instance, previous works investigated one-hot representations initially introduced for text data [63] as well as trajectory image [64]. Here, we explored distance matrix representations as the distance matrix of positions has been shown to capture relevant spatial information for protein structure prediction [65], [66]. Our study demonstrates its relevance when considered as input to a neural network architecture to complement longitude~~

and latitude time series. Future work may further investigate how distance matrix representation could be of interest for other deep learning approaches to trajectory data, including among others Recurrent Neural Network (RNN), and Generative Adversarial Networks (GAN) for trajectory prediction and simulation [41], [42], [67], [68].

When considering learning-based schemes, the assessment of the generalization performance is of key importance When considering neural network approaches, training models which may apply beyond the considered training framework is a key feature, generally referred to as the generalization performance of the trained neural networks. Beyond the evaluation of dive prediction performance on a trajectory dataset, which is independent from the training dataset, the question whether a model trained on a given dataset, e.g. for a given species, colony and time period, may apply to other species, colonies and/or time periods, naturally arises as a key question. Numerous studies in the deep learning literature [69], [70] have highlighted that some neural architectures show relevant generalization properties whereas others may not. Here, we evaluated the generalization performance of the benchmarked methods for two colonies of Peruvian boobies. three benchmarked deep networks.

Thus we demonstrate the ability of deep networks trained at a colony for one species to also apply to an another colony (of the same ecosystem) for the same species. In our example, Peruvian boobies from Guañape Guañape Island did have different foraging strategies from their counterparts from Pescadores island, with trips two times longer and dives slightly longer (see Table ??). It turned out that the two deep networks were very stable on this testing dataset predicting dives with high accuracy (AUC of 0.971). However, the UNet reached similar dive prediction performance when applied to Guañape data. This suggests that dive patterns are highly similar between Peruvian boobies from both colonies. We also show the great ability of the CNN to generalize dive prediction to a seabird of same genus but from a totally distinct ecosystem. When applied to masked boobies trajectories from a Brazilian colony the CNN trained from Peruvian boobies data obtained an AUC of 0.87 despite the important difference in foraging strategies (Tab. 3). The improvement of neural networks over the HMM fitted to the dataset was about 10% same model trained on masked boobies data reached an AUC of 0.93 (Fig. 7), suggesting that diving characteristics are slightly different. Masked boobies from the Brazilian colony feed indeed on different preys, and spend way more time resting at the surface (Tab.1). As deep networks trained on cormorants unsurprisingly led to less accurate prediction when used to predict boobies dives, we suggest that the CNN may capture genus-specific features. These results then support the relevance of deep learning schemes as 'ready-to-use' tools which could be used by ecologists to predict seabirds dives on new datasets(small) datasets, including when these datasets do not include groundtruthed dive data for a supervised training. To make easier such applications, we share online the different models we pre-trained trained on the considered dataset datasets (https://github.com/AmedeeRoy/BirdDL/results). models).

Beyond such a direct application, trained models may also be are also of key interest to explore transfer learning strategies, which refer to the ability of exploiting some previously trained models to address a new task or dataset rather than training a new model from scratch. Fine tuning is a typical example of transfer learning procedure. It uses a previously trained model as the initialization of the training scheme. Such a strategy may be We illustrated how fine-tuned CNN and UNet models could outperform HMM with smaller training datasets. For instance, the fine-tuned

CNN for the prediction of masked boobies' dive was able to converge and outperform HMM with a dataset twice as small as the dataset required to reach same performance without fine-tuning (Fig. 7). Such a result was even possible by initializing neural networks with the model trained with cormorant data. This further supports the ability of deep networks to generalize their prediction from deep diving seabirds (e.g. cormorants) to plunge divers (e.g. boobies). Fine-tuning is thus particularly relevant when the training dataset may not be sufficiently large to train a model from scratch. ~~We then expect the models trained in this work to be potentially of interest for the~~ While the need of large dataset is often presented as a drawback for supervized techniques, we demonstrated that relatively small datasets (5-10 foraging trips, i.e. 10k to 20k GPS data) may be enough to fine-tune deep networks and outperform stat-of-the-art approach to data segmentation. Thus, we expect that our models will be of interest for future work on seabird trajectory segmentation, as they could be used as initializations for fine-tuning ~~of these models on datasets referring to other colonies and species of same genus~~procedures.

Acknowledgements

Funding statement

Ethics statement

Foraging trip routes and dive depth data were obtained from electronic devices attached to Peruvian ~~Boobies and Guanay Cormorants~~ boobies and Guanay cormorants tagged at the Pescadores ~~Guañape~~ and Guañape Islands, Peru, from 2007 to ~~2013. Capture, restraint, holding and release of birds for attachment and subsequent removal of electronic devices were performed under the auspices of institutional animal use and care a~~ 2013, and from masked boobies tagged at the Fernando de Noronha, Brazil, from 2017 to 2019. This work was conducted with the approval of the Peruvian federal agency, Programa de Desarrollo Productivo Agrario Rural, commonly known as "Agrorural". Headquarters of Agrorural are located at Av. Salaverry 1388, Lima, Peru~~. Protocols for care and use of subject birds were reviewed and approved~~

~~by Agrorural officials prior to every fieldwork sessions~~, and of the Brazilian Ministry of Environment—Instituto Chico Mendes de Conservação da Biodiversidade (Authorization No 52583-5).

References

[1] M. Authier, J. Spitz, A. Blanck, and V. Ridoux, "Conservation science for marine megafauna in Europe: Historical perspectives and future directions," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 141, pp. 1–7, 2017, ISSN: 09670645. DOI: 10.1016/j.dsr2.2017.05.002.

[2] E. L. Hazen, B. Abrahms, S. Brodie, G. Carroll, M. G. Jacox, M. S. Savoca, K. L. Scales, W. J. Sydeman, and S. J. Bograd, "Marine top predators as climate and ecosystem sentinels," *Frontiers in Ecology and the Environment*, vol. 17, no. 10, pp. 565–574, 2019, ISSN: 1540-9295, 1540-9309. DOI: 10.1002/fee.2125.

[3] B. Lascelles, G. Notarbartolo Di Sciara, T. Agardy, A. Cuttelod, S. Eckert, L. Glowka, E. Hoyt, F. Llewellyn, M. Louzao, V. Ridoux, and M. J. Tetley, "Migratory marine species: Their status, threats and conservation management needs," *Aquatic Conservation: Marine and Freshwater Ecosystems*, vol. 24, no. S2, pp. 111–127, 2014, ISSN: 10527613. DOI: 10.1002/aqc.2512.

[4] S. K. Hooker and L. R. Gerber, "Marine Reserves as a Tool for Ecosystem-Based Management: The Potential Importance of Megafauna," *BioScience*, vol. 54, pp. 27–39, 2004. DOI: 10.1641/0006-3568(2004)054[0027: MRAATF]2.0.CO;2.

[5] R. Furness and Camphuysen, "Seabirds as monitors of the marine environment," *ICES Journal of Marine Science*, vol. 54, no. 4, pp. 726–737, 1997, ISSN: 10543139. DOI: 10.1006/jmsc.1997.0243.

[6] E. Wakefield, R. Phillips, and J. Matthiopoulos, "Quantifying habitat use and preferences of pelagic seabirds using individual movement data: A review," *Marine Ecology Progress Series*, vol. 391, pp. 165–182, 2009, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps08203.

[7] E. A. Schreiber and J. Burger, Eds., *Biology of Marine Birds*. Hoboken, USA: CRC press, 2001.

[8] H. Weimerskirch, "Are seabirds foraging for unpredictable resources?" *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 54, no. 3-4, pp. 211–223, 2007, ISSN: 09670645. DOI: 10.1016/j.dsr2.2006.11.013.

[9] C. A. Bost, C. Cotté, P. Terray, C. Barbraud, C. Bon, K. Delord, O. Gimenez, Y. Handrich, Y. Naito, C. Guinet, and H. Weimerskirch, "Large-scale climatic anomalies affect marine predator foraging behaviour and demography," *Nature Communications*, vol. 6, no. 1, p. 8220, 2015, ISSN: 2041-1723. DOI: 10.1038/ncomms9220.

[10] R. E. Meier, R. B. Wynn, S. C. Votier, M. McMinn Grivé, A. Rodríguez, L. Maurice, E. E. van Loon, A. R. Jones, L. Suberg, J. M. Arcos, G. Morgan, S. A. Josey, and T. Guilford, "Consistent foraging areas and commuting corridors of the critically endangered Balearic shearwater Puffinus mauretanicus in the northwestern Mediterranean," *Biological Conservation*, vol. 190, pp. 87–97, 2015, ISSN: 00063207. DOI: 10/f7m4k2.

[11] S. Oppel, S. Weber, N. Weber, D. Fox, E. Leat, J. Sim, J. Sommerfeld, M. Bolton, A. Broderick, and B. Godley, "Seasonal shifts in foraging distribution due to individual flexibility in a tropical pelagic forager, the Ascension frigatebird," *Marine Ecology Progress Series*, vol. 585, pp. 199–212, 2017, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps12377.

[12] M. Kappes, H. Weimerskirch, D. Pinaud, and M. Le Corre, "Variability of resource partitioning in sympatric tropical boobies," *Marine Ecology Progress Series*, vol. 441, pp. 281–294, 2011, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps09376.

[13] M. Bolton, G. Conolly, M. Carroll, E. D. Wakefield, and R. Caldow, "A review of the occurrence of inter-colony segregation of seabird foraging areas and the implications for marine environmental impact assessment," *Ibis*, vol. 161, no. 2, pp. 241–259, 2019, ISSN: 0019-1019, 1474-919X. DOI: 10/gmw8sv.

[14] Y. Ropert-Coudert, M. Beaulieu, N. Hanuise, and A. Kato, "Diving into the world of biologging," *Endangered Species Research*, vol. 10, pp. 21–27, 2009, ISSN: 1863-5407, 1613-4796. DOI: 10.3354/esr00188.

[15] H. Chung, J. Lee, and W. Y. Lee, "A Review: Marine Bio-logging of Animal Behaviour and Ocean Environments," *Ocean Science Journal*, vol. 56, no. 2, pp. 117–131, 2021, ISSN: 1738-5261, 2005-7172. DOI: 10/gk76v5.

[16] K. Yoda, "Advances in bio-logging techniques and their application to study navigation in wild seabirds," *Advanced Robotics*, vol. 33, no. 3-4, pp. 108–117, 2019. DOI: 10.1080/01691864.2018.1553686.

[17] R. Nathan, W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse, "A movement ecology paradigm for unifying organismal movement research," *Proceedings of the National Academy of Sciences*, vol. 105, no. 49, pp. 19 052–19 059, 2008, ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0800375105.

[18] S. L. Cox, P. I. Miller, C. B. Embling, K. L. Scales, A. W. J. Bicknell, P. J. Hosegood, G. Morgan, S. N. Ingram, and S. C. Votier, "Seabird diving behaviour reveals the functional significance of shelf-sea fronts as foraging hotspots," *Royal Society Open Science*, vol. 3, no. 9, p. 160 317, 2016, ISSN: 2054-5703, 2054-5703. DOI: 10.1098/rsos.160317.

[19] S. Lewis, E. Schreiber, F. Daunt, G. Schenk, S. Wanless, and K. Hamer, "Flexible foraging patterns under different time constraints in tropical boobies," *Animal Behaviour*, vol. 68, no. 6, pp. 1331–1337, 2004, ISSN: 00033472. DOI: 10.1016/j.anbehav.2004.04.007.

[20] A. Shoji, K. Elliott, A. Fayet, D. Boyle, C. Perrins, and T. Guilford, "Foraging behaviour of sympatric razorbills and puffins," *Marine Ecology Progress Series*, vol. 520, pp. 257–267, 2015, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps11080.

[21] E. Browning, M. Bolton, E. Owen, A. Shoji, T. Guilford, and R. Freeman, "Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds," *Methods in Ecology and Evolution*, vol. 9, no. 3, J. McPherson, Ed., pp. 681–692, 2018, ISSN: 2041210X. DOI: 10/gc5vjk.

[22] B. Dean, H. Kirk, A. Fayet, A. Shoji, R. Freeman, K. Leonard, C. Perrins, and T. Guilford, "Simultaneous multi-colony tracking of a pelagic seabird reveals cross-colony utilization of a shared foraging area," *Marine Ecology Progress Series*, vol. 538, pp. 239–248, 2015, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps11443.

[23] D. P. Seidel, E. Dougherty, C. Carlson, and W. M. Getz, "Ecological metrics and methods for GPS movement data," *International Journal of Geographical Information Science*, vol. 32, no. 11, pp. 2272–2293, 2018, ISSN: 1365-8816, 1362-3087. DOI: 10.1080/13658816.2018.1498097.

[24] M. I. D. Carter, K. A. Bennett, C. B. Embling, P. J. Hosegood, and D. J. F. Russell, "Navigating uncertain waters: A critical review of inferring foraging behaviour from location and dive data in pinnipeds," *Movement Ecology*, vol. 4, no. 1, p. 25, 2016, ISSN: 2051-3933. DOI: 10/ggkjc2.

[25] D. Pinaud and H. Weimerskirch, "At-sea distribution and scale-dependent foraging behaviour of petrels and albatrosses: A comparative study," *Journal of Animal Ecology*, vol. 76, no. 1, pp. 9–19, 2007, ISSN: 0021-8790, 1365-2656. DOI: 10.1111/j.1365-2656.2006.01186.x.

[26] J. Sommerfeld, A. Kato, Y. Ropert-Coudert, S. Garthe, and M. A. Hindell, "Foraging Parameters Influencing the Detection and Interpretation of Area-Restricted Search Behaviour in Marine Predators: A Case Study with the Masked Booby," *PLoS ONE*, vol. 8, no. 5, G. Sorci, Ed., e63742, 2013, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0063742.

[27] H. Weimerskirch, M. Le Corre, and C. Bost, "Foraging strategy of masked boobies from the largest colony in the world: Relationship to environmental conditions and fisheries," *Marine Ecology Progress Series*, vol. 362, pp. 291–302, 2008, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps07424.

[28] L. Mendez, P. Borsa, S. Cruz, S. de Grissac, J. Hennicke, J. Lallemand, A. Prudor, and H. Weimerskirch, "Geographical variation in the foraging behaviour of the pantropical red-footed booby," *Marine Ecology Progress Series*, vol. 568, pp. 217–230, 2017, ISSN: 0171-8630, 1616-1599. DOI: 10.3354/meps12052.

[29] J. Garriga, J. R. B. Palmer, A. Oltra, and F. Bartumeus, "Expectation-Maximization Binary Clustering for Behavioural Annotation," *PLOS ONE*, vol. 11, no. 3, A. Gursoy, Ed., e0151984, 2016, ISSN: 1932-6203. DOI: 10/f8wh5g.

[30] C. Boyd, "Movement models provide insights into variation in the foraging effort of central place foragers," *Ecological Modelling*, p. 13, 2014. DOI: 10.1016/j.ecolmodel.2014.03.015.

[31] B. J. Dunphy, S. I. Vickers, J. Zhang, R. L. Sagar, T. J. Landers, S. J. Bury, A. J. R. Hickey, and M. J. Rayner, "Seabirds as environmental indicators: Foraging behaviour and ecophysiology of common diving petrels (Pelecanoides urinatrix) reflect local-scale differences in prey availability," *Marine Biology*, vol. 167, no. 4, p. 53, 2020, ISSN: 0025-3162, 1432-1793. DOI: 10.1007/s00227-020-3672-4.

[32] B. T. McClintock and T. Michelot, "momentuHMM: R package for generalized hidden Markov models of animal movement," *Methods in Ecology and Evolution*, vol. 9, no. 6, S. Goslee, Ed., pp. 1518–1530, 2018, ISSN: 2041-210X, 2041-210X. DOI: 10.1111/2041-210X.12995.

[33] S. Oppel, A. Beard, D. Fox, E. Mackley, E. Leat, L. Henry, E. Clingham, N. Fowler, J. Sim, J. Sommerfeld, N. Weber, S. Weber, and M. Bolton, "Foraging distribution of a tropical seabird supports Ashmole's hypothesis of population regulation," *Behavioral Ecology and Sociobiology*, vol. 69, no. 6, pp. 915–926, 2015, ISSN: 0340-5443, 1432-0762. DOI: 10.1007/s00265-015-1903-3.

[34] T. Guilford, J. Meade, J. Willis, R. Phillips, D. Boyle, S. Roberts, M. Collett, R. Freeman, and C. Perrins, "Migration and stopover in a small pelagic seabird, the Manx shearwater *Puffinus puffinus* : Insights from machine learning," *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, no. 1660, pp. 1215–1223, 2009. DOI: `10.1098/rspb.2008.1577`.

[35] G. Wang, "Machine learning for inferring animal behavior from location and movement data," *Ecological Informatics*, vol. 49, pp. 69–76, 2019. DOI: `10.1016/j.ecoinf.2018.12.002`.

[36] R. Joo, M. E. Boone, T. A. Clay, S. C. Patrick, S. Clusella-Trullas, and M. Basille, "Navigating through the R packages for movement," *Journal of Animal Ecology*, vol. 89, no. 1, L. Börger, Ed., pp. 248–267, 2020, ISSN: 0021-8790, 1365-2656. DOI: `10.1111/1365-2656.13116`.

[37] J. J. Valletta, C. Torney, M. Kings, A. Thornton, and J. Madden, "Applications of machine learning in animal behaviour studies," *Animal Behaviour*, vol. 124, pp. 203–220, 2017. DOI: `10.1016/j.anbehav.2016.12.005`.

[38] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, ISSN: 0028-0836, 1476-4687. DOI: `10.1038/nature14539`.

[39] B. G. Weinstein, "A computer vision for animal ecology," *Journal of Animal Ecology*, vol. 87, no. 3, L. Prugh, Ed., pp. 533–545, 2018. DOI: `10.1111/1365-2656.12780`.

[40] S. Christin, É. Hervet, and N. Lecomte, "Applications for deep learning in ecology," *Methods in Ecology and Evolution*, vol. 10, no. 10, H. Ye, Ed., pp. 1632–1644, 2019. DOI: `10.1111/2041-210X.13256`.

[41] I. S. Ardakani and K. Hashimoto, "Encoding Bird's Trajectory using Recurrent Neural Networks," in *Proceedings of 2017 IEEE International Conference on Mechatronics and Automation*, Japan, 2017, p. 6. DOI: `10.1109/ICMA.2017.8016063`.

[42] J. Rew, S. Park, Y. Cho, S. Jung, and E. Hwang, "Animal Movement Prediction Based on Predictive Recurrent Neural Network," *Sensors*, vol. 19, no. 20, p. 4411, 2019. DOI: `10.3390/s19204411`.

[43] C. Peng, C. M. Duarte, D. P. Costa, C. Guinet, R. G. Harcourt, M. A. Hindell, C. R. McMahon, M. Muelbert, M. Thums, K.-C. Wong, and X. Zhang, "Deep Learning Resolves Representative Movement Patterns in a Marine Predator Species," *Applied Sciences*, vol. 9, no. 14, p. 2935, 2019. DOI: `10.3390/app9142935`.

[44] T. Maekawa, "Deep learning-assisted comparative analysis of animal trajectories with DeepHL," *Nature Communications*, vol. 11, p. 5316, 2020. DOI: `10.1038/s41467-020-19105-0`.

[45] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597 [cs]*, 2015. arXiv: `1505.04597 [cs]`.

[46] A. Paskze, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, and T. Killen, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," *arXiv preprint arXiv:1912.01703*, 2019. arXiv: `1912.01703`.

[47] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. Van Essen, A. A. S. Awwal, and V. K. Asari, "A State-of-the-Art Survey on Deep Learning Theory and Architectures," *Electronics*, vol. 8, no. 3, p. 292, 2019, ISSN: 2079-9292. DOI: `10/gfw52f`.

[48] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," *arXiv:1206.5538 [cs]*, 2014. arXiv: `1206.5538 [cs]`.

[49] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint arXiv:1412.6980*, 2014. arXiv: `1412.6980`.

[50] P. Fauchald and T. Tveraa, "Using First-Passage Time in the analysis of area-restricted search and habitat selection," *Ecology*, vol. 84, no. 2, pp. 282–288, 2003, ISSN: 0012-9658. DOI: `10.1890/0012-9658(2003)084[0282:UFPTIT]2.0.CO;2`.

[51] J. Jahncke and E. Goya, "Diets of the Guanay cormorant and Peruvian booby as indicators of the abundance and distribution of anchovy," *Boletin Instituto del Mar del Peru*, vol. 17, pp. 15–33, 1998.

[52] P. D. Wilson, "Distance-based methods for the analysis of maps produced by species distribution models," *Methods in Ecology and Evolution*, p. 11, 2011. DOI: `10/b3bnhq`.

[53] H. Weimerskirch, S. Bertrand, J. Silva, C. Bost, and S. Peraltilla, "Foraging in Guanay cormorant and Peruvian booby, the major guano-producing seabirds in the Humboldt Current System," *Marine Ecology Progress Series*, vol. 458, pp. 231–245, 2012, ISSN: 0171-8630, 1616-1599. DOI: `10/f34jq3`.

[54] R. P. de Santana Campelo, F. K. Bonou, M. de Melo Júnior, X. F. G. Diaz, L. E. A. Bezerra, and S. Neumann-Leitão, "Zooplankton biomass around marine protected islands in the tropical Atlantic Ocean," *Journal of Sea Research*, vol. 154, p. 101 810, 2019, ISSN: 13851101. DOI: `10.1016/j.seares.2019.101810`.

[55] B. Nelson, *Pelicans, Cormorants, and Their Relatives*. Oxford University Press, 2005.

[56] P. L. Mancini, K. A. Hobson, and L. Bugoni, "Role of body size in shaping the trophic structure of tropical seabird communities," *Marine Ecology Progress Series*, vol. 497, pp. 243–257, 2014. DOI: `10.3354/meps10589`.

[57] K. Delord, C. Barbraud, D. Pinaud, B. Letournel, D. Jaugeon, H. Goraguer, P. Lazure, and H. Lormée, "Movements of three alcid species breeding sympatrically in Saint Pierre and Miquelon, northwestern Atlantic Ocean," *Journal of Ornithology*, p. 30, 2020.

[58] H. Weimerskirch, S. de Grissac, A. Ravache, A. Prudor, A. Corbeau, B. Congdon, F. McDuie, K. Bourgeois, S. Dromzée, J. Butscher, C. Menkes, V. Allain, E. Vidal, A. Jaeger, and P. Borsa, "At-sea movements of wedge-tailed shearwaters during and outside the breeding season from four colonies in New Caledonia," *Marine Ecology Progress Series*, vol. 633, pp. 225–238, 2020, ISSN: 0171-8630, 1616-1599. DOI: `10.3354/meps13171`.

[59] J. Zhang, M. Rayner, S. Vickers, T. Landers, R. Sagar, J. Stewart, and B. Dunphy, "GPS telemetry for small seabirds: Using hidden Markov models to infer foraging behaviour of Common Diving Petrels (*Pelecanoides urinatrix urinatrix*)," *Emu - Austral Ornithology*, vol. 119, no. 2, pp. 126–137, 2019, ISSN: 0158-4197, 1448-5540. DOI: `10/ghtzsr`.

[60] R. P. Wilson and S. P. Vandenabeele, "Technological innovation in archival tags used in seabird research," *Marine Ecology Progress Series*, p. 18, 2012.

[61] C. Beyan and H. Browman, "Setting the stage for the machine intelligence era in marine science," *ICES Journal of Marine Science*, vol. 77(4), pp. 1267–1273, 2020. DOI: `doi:10.1093/icesjms/fsaa084`.

[62] K. Malde, N. O. Handegard, L. Eikvil, and A.-B. Salberg, "Machine intelligence and the data-driven future of marine science," *ICES Journal of Marine Science*, p. 12, 2020.

[63] D. Nguyen, R. Vadaine, G. Hajduch, R. Garello, and R. Fablet, "GeoTrackNet-A Maritime Anomaly Detector using Probabilistic Neural Network Representation of AIS Tracks and A Contrario Detection," *arXiv:1912.00682 [cs, stat]*, 2021. arXiv: 1912.00682 [cs, stat].

[64] Y. Endo, H. Toda, K. Nishida, and J. Ikedo, "Classifying spatial trajectories using representation learning," *International Journal of Data Science and Analytics*, vol. 2, no. 3-4, pp. 107–117, 2016, ISSN: 2364-415X, 2364-4168. DOI: 10.1007/s41060-016-0014-1.

[65] A. W. Senior, "Improved protein structure prediction using potentials from deep learning," *Nature*, vol. 577, pp. 706–710, 2020. DOI: 10.1038/s41586-019-1923-7.

[66] J. Xu and S. Wang, "Analysis of distance-based protein structure prediction by deep learning in CASP13," *Proteins: Structure, Function, and Bioinformatics*, vol. 87, no. 12, pp. 1069–1081, 2019, ISSN: 0887-3585, 1097-0134. DOI: 10/gh3mv4.

[67] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," *arXiv:1406.2661*, 2014. arXiv: 1406.2661.

[68] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *arXiv:1611.07004 [cs]*, 2018. arXiv: 1611.07004 [cs].

[69] K. Kawaguchi, L. P. Kaelbling, and Y. Bengio, "Generalization in Deep Learning," *arXiv:1710.05468 [cs, stat]*, 2020. arXiv: 1710.05468 [cs, stat].

[70] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," *arXiv:1611.03530 [cs]*, 2017. arXiv: 1611.03530 [cs].