DEEP LEARNING AND TRAJECTORY REPRESENTATION FOR THE PREDICTION OF SEABIRD DIVING BEHAVIOUR

Amédée Roy
Institut de Recherche pour le Développement,
UMR248 MARBEC (IRD/CNRS/IFREMER/UM)
Avenue Jean Monnet, 34200, Sète, France
amedee.roy@ird.fr

Sophie Bertrand
Institut de Recherche pour le Développement,
UMR248 MARBEC (IRD/CNRS/IFREMER/UM)
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

February 25, 2021

ABSTRACT

1. Seabirds are often considered as suitable indicators for the study of marine ecosystems, since their foraging strategies give us a real-time response to the complex dynamics of the ecosystem. By deploying sufficiently light GPS sensors on seabirds, it is possible to obtain their trajectories. Behaviours at sea and foraging areas are then usually inferred by statistical models (e.g. Hidden Markov models). Deep learning has recently shown promising results for the classification of animal behaviour, yet there is still lot of investigation needed in terms of network architectures, data representation but also to demonstrate the relevance and generalization properties of such approaches.

2. From a database of about 250 foraging trajectories derived from GPS data deployed simultaneously with pressure sensors for the identification of dives, this work consisted in training deep networks in a supervised manner for the prediction of seabird dives from GPS data only. 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).

3. In this work we confirm that deep learning tools can predict better dive locations than usual methods for two seabirds species. We show that the representation of trajectory data as distance matrix greatly increases the ability of deep learning architectures to infer behaviours. We also point out the

18 potential impact of such methods on the estimation of dives' distribution maps with estimation error

19 divided approximately by 2 when using deep networks rather than usual Hidden Markov Models.

20 4. More generally, this study demonstrates the importance of data representation when using deep

21 learning for trajectory data. Longitude and latitude time series might not be the best representation

22 of the underlying geospatial information for learning schemes. Important benefits might be achieved

23 by using different trajectory representations such as matrix distance.

24 **Keywords** machine learning · neural network · matrix distance · Peruvian booby · Guanay cormorant · diving

25 behaviour

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) has received significant attention in marine ecology over the last decades (Authier *et al.*, 2017). These species offer a unique perspective into ocean processes and dynamics, given that they can amplify information on the structure of the seascape across multiple spatiotemporal scales due to their relatively high mobility and longevity (Hazen *et al.*, 2019). 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 (Lascelles *et al.*, 2014; Hooker & Gerber, 2004).

Among marine megafauna species, seabirds are considered as suitable indicators because they are both sensitive to variations in food supply and relatively easy to observe (Furness & Camphuysen, 1997; Wakefield *et al.*, 2009). During breeding season, seabirds 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 (Schreiber & Burger, 2001). 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 (Weimerskirch, 2007).

Over the last decades, the study of animal movements has been revolutionized by great technical advances in the miniaturization and autonomy of devices (Ropert-Coudert *et al.*, 2009). GPS loggers have been at the forefront of this breakthrough, and can now provide precise and accurate data on the movements of many free-ranging species, such as seabirds (Wakefield *et al.*, 2009; Yoda, 2019). The derived data is particularly useful in studying a variety of behavioral aspects including habitat selection, migration or dispersion patterns, and foraging strategies (Nathan *et al.*, 2008). Detailed information on the foraging behaviour has been gained through the combined use of GPS and 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 (Cox *et al.*, 2016; Shoji *et al.*, 2015).

Yet, for practical, 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 individual-based studies aim to infer behavioral state directly by applying thresholds from various ecological metrics of movement data, such as speed, direction and tortuosity (Dean *et al.*, 2015; Seidel *et al.*, 2018). A common example is the so-called First-Passage Time (FPT) method, which is defined as the time taken for an individual to cross a virtual circle of given radius (Carter *et al.*, 2016). Here foraging behaviour is assumed to occur when birds fly at very low speeds (Weimerskirch *et al.*, 2008). Modelling methods have also been used to predict diving behaviour taking the sequence of trajectory data into account through hidden Markov models (HMM) typically with 2 or 3 distinct behavioural modes (Boyd, 2014; McClintock & Michelot, 2018; Oppel *et al.*, 2015), and more occasionally through gaussian mixtures

models (Guilford *et al.*, 2008; Mendez *et al.*, 2017), or supervised machine learning approaches such as artificial neural networks, support vector machine, and random forests (Guilford *et al.*, 2009; Wang, 2019). We may refer the reader to (Joo *et al.*, 2020) 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 (Valletta *et al.*, 2017). Deep learning refers to a neural network with multiple layers of processing units (LeCun *et al.*, 2015). 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 scales (Nathan *et al.*, 2008), deep learning might be able to extract relevant representations of trajectories. 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 (Weinstein, 2018; Christin *et al.*, 2019). Fewer studies have explored deep learning for animal trajectory data. Recurrent neural networks (RNNs) have been used for movement prediction (Ardakani & Hashimoto, 2017; Rew *et al.*, 2019), and for the identification of representative movement patterns (Peng *et al.*, 2019). Very recently, an attention network has also been proposed for comparative analysis of animal trajectories (Maekawa, 2020). Related to our study, a fully-connected network (FCN) has been used to predict seabirds' diving in European shags, common guillemots and razorbills (Browning *et al.*, 2018). Using a fully-connected network 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 (Browning *et al.*, 2018), this work addresses the inference of seabird diving behaviour from GPS data using Deep Learning methods. From a database of about 250 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-Passage Time).

2 Materials and Methods

2.1 Dataset

GPS and TDR devices were fitted to Peruvian boobies (*Sula Variegata*) and Guanay Cormorant (*Leucocarbo Bougainvilli*) breeding off the coast of Peru at Isla Pescadores (11.775°S, 77.265°W) every year in December from 2008 to 2013. Similar devices were also deployed on Peruvian boobies from Isla Guanuape (8.566°S, 78.966°W) in December 2007. In total, GPS devices (Gipsy GPS, 2530 g, Technosmart, Rome, Italy) and time-depth recorders (TDR, 3 g; 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 tape on the tail feathers for boobies and on the back feathers for cormorants for 1 to 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. This dataset consists therefore in a total of 234 foraging trips of seabirds with doubled-deployment GPS and TDR (see Table 1).

Missing fixes in GPS data were linearly interpolated. The coverage ratio was computed as in (Browning *et al.*, 2018). 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

As baseline architecture, we considered fully-connected network proposed in (Browning *et al.*, 2018). 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. We describe below these two architectures and the associated supervised training procedure. We refer the reader to (Christin *et al.*, 2019) for an introduction to deep neural networks dedicated to ecologists.

2.2.1 Fully-Connected Network (FCN)

The first architecture implemented was similar to the fully-connected network presented by (Browning *et al.*, 2018). As input vector, we used the concatenation of longitude, latitude 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. Overall, this architecture involves 5.10^5 parameters.

2.2.2 U-Network with Distance Matrix Encoder (DME-UNet)

As described in Figure 3, we also introduced a novel architecture based on a U-Network (UNet) (Ronneberger *et al.*, 2015). This architecture exploits a Distance Matrix Encoder (DME) to represent geometrical features along a trajectory. Similarly to the FCNet architecture, its input vector is the concatenation of longitude, latitude and coverage time series over a 20-second window and it outputs a vector of diving probability of the same length.

Distance Matrix Encoder   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 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. 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 & Benhamou, 2008), but with circle of different radius which values will be adjusted automatically by the network.

U-Network   We adapt a U-Net architecture, which is the state-of-the-art neural architecture for segmentation tasks (Ronneberger *et al.*, 2015), 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. The key feature of the U-Net architecture is to combine the information extracted by convolutional blocks applied at 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.3. 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 architecture, the last layer applies a sigmoïd activation to transform the output into a time series of diving probabilities. Overall, this architecture (DME + U-Net) involves 2.10^4 parameters.

2.2.3   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 and 30 for boobies.

The minimisation of the training loss exploits the Adam stochastic optimizer (Kingma & Ba, 2015). 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

6

probability of all predictions derived from all 20 positions windows. These models were implemented, trained and
tested with python using pytorch library (Paskze *et al.*, 2019).

2.3 Benchmarked methods

Two classical methods for dive prediction First-Passage Time (FPT), and Hidden Markov Models (HMM) were evalu-
ated for intercomparison purposes. FPT was computed following (Fauchald & Tveraa, 2003), 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. We implemented HMMs with 3 behavioural modes associated to traveling, searching and
diving behaviours using the momentuHMM R package (McClintock & Michelot, 2018). This approach represents
trajectories as a sequence of steps and angles. It models steps as random variables following a gamma marginal dis-
tribution 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.

2.4 Evaluation scheme

We describe below the evaluation scheme we implemented to assess the performance of the proposed neural network
approaches. We first focus on the benchmarking of the performance of the considered approaches in terms of dive pre-
diction accuracy. For the proposed neural network architectures, we further analyze the impact on the dive prediction
performance of different data types used as inputs as well as their generalization performance.

Dive prediction performance   As evaluation metrics for dive prediction, we evaluated the receiver operating charac-
teristics curve (ROC), the area under the curve (AUC) as well as the binary cross entropy (BCE) for the test datasets.
For binary classification, the ROC curve plots the true positive rate (i.e. true predicted dives) against the false positive
rate (i.e. false predicted dives). We obtained this curve by varying the probability threshold defining dive/no dive
behaviours. Regarding the AUC, it was estimated by integrating the ROC curve along the x axis. For neural network
approaches, we also analyzed the value of training loss for the training and test datasets.

Data inputs   Both studied species breed in the most productive upwelling system (Humboldt Current System) and
feed on Peruvian anchovies (Jahncke & Goya, 1998). However, they 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 30 m depth, and spend up to 40% of the time in water (Weimerskirch *et al.*, 2012).
We assessed the dive prediction performance of the benchmarked methods when considering trajectory data derived
from these two seabirds and with different time resolutions (trajectory data sampled at 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. Overall, this led to the quantitative comparison of the performance of
17 models and 6 datasets all listed in Table 2.

7

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

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

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

193 Peruvian boobies trajectories but from a different breeding colony (i.e. Guañape datset). We expect Peruvian boobies

194 behaviour to share common features across colonies, so that models fitted to the trajectory data of a given colony might

195 still be meaningful when applied to another colony. In this experiment, we compared the dive prediction performance

196 of FPT and HMM methods to the best FCNet and DME-UNet models. Beyond AUC and BCE performance metrics,

197 we also evaluated the relevance of the estimated map of dive distributions. The later were computed using a weighted

198 Kernel Density Estimator (KDE) using dive probabilities as weighing factor. As groundtruth, we considered the map

199 of dive distributions estimated from true dive locations defined by TDR data. From these maps, we evaluated an

200 Hellinger distance as an integrated performance metrics for the different approaches (Wilson, 2011).

3   Results

202 We detailed below the numerical experiments performed in this study to assess the relevance of the proposed neural

203 network approaches to predict dive behaviour of Peruvian boobies and cormorants from trajectory data.

204 **Overall performance**   For all datasets, the different methods obtained contrasted performance results, with AUC

205 going from 0.71 to 0.97 (see Table 2), which corresponds in the best cases to correct prediction rates of diving and

206 non-diving behaviour of approximately 90% and of 70% in the worst cases (see Figure 1). Overall, neural networks-

207 based approaches obtained better performance than the state-of-the art methods, with highest AUC (around 0.9) and

208 lowest binary cross entropy (around 0.5). First-Passage Time obtained quasi systematically the lowest AUC, and the

209 Hidden Markov Models had most of the time AUC around 0.85 but the highest BCE. State-of-the-art methods were

210 outperformed by neural networks specifically for the 5s resolution datasets, with AUC improvements up to 15% with

211 neural networks over HMM. In particular, the DME-UNet was the most consistent method, being able to get better

212 performance on most datasets, and to improve substantially the performance of the FCNet proposed by (Browning

213 *et al.*, 2018). As illustration, on the boobies dataset, BCE was approximately two times lower with the DME-UNet

214 than with the FCNet, and the DME-UNet obtained systematically higher AUC.

215 **Impact of the temporal resolution**   Interestingly, on the cormorant dataset the sampling resolution did not affect

216 much the performance of the neural network approaches (DME-UNet with AUC of 0.95) whereas state-of-the-art

217 methods FPT/HMM had increasing AUC with larger resolution going from respectively 0.73/0.78 to 0.79/0.86 (see

218 Table 2). At the opposite, on the peruvian datasets the FPT/HMM methods had constant performance disregarding the

219 sampling resolution (AUC around 0.76/0.88), while deep learning tools had decreasing AUC with larger resolution. In

220 particular, the dataset of Peruvian boobies with the larger resolution (i.e. 30s) is the only one where neural networks

221 did not do any better than 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 1).

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 2) . Yet, without coverage information, the accuracy of neural networks was poorer. UNet and FCNet had decreased their performance with AUC around 0.7, and the network using a Distance Matrix Encoder (DME-UNet) maintained 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 2).

Maps of diving probabilities   We further analysed the relevance of the dive predictions through the estimated dive distributions maps reported in Figure 4 and 5. For both species, we observed that neural networks and HMM approaches have contrasted outputs (either very high or low diving probabilities), while the FPT approach discriminated the least dive from non-diving behaviours. Moreover, these maps also pointed 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). It was particularly visible from FCNet and HMM output and specifically in the vicinity of the colony.

Application to Guañape dataset   Both FCNet and DME-UNet outperformed HMM and FPT in terms of AUC (e.g., with AUC of around 0.97 for neural networks vs. 0.88 for HMM) on trajectories from birds from another colony (i.e. Guañape). The Hellinger distance of the estimated dive distribution maps stressed the greater relevance of DME-UNet predictions with a distance value 1.6 times smaller than the one derived from FCNet estimations and 2 times smaller than the one derived from HMM estimation.

4   Discussion

This study aimed at predicting seabirds dives from GPS data only by training a deep network in a supervized manner based on TDR data to define the true dives. In line with (Browning *et al.*, 2018), it further supports the relevance of deep learning approach over classical methods for dive predictions. Moreover, we introduced 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 reported even better results with higher stability to the different data inputs.

Seabird Dives Distribution Map Estimation   The proposed network allows to better predict dive behaviour but also results in better seabird dive distribution maps. 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 (Delord *et al.*, 2020; Weimerskirch *et al.*, 2020). Here, we show that the error in the estimation of dive distributions maps can be 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. Sulids and cormorants spend time bathing

253   near their breeding territories involving vigorous splashing and beating the water with the wings, (Nelson, 2005).
254   Such behaviours associated to low speed might be erroneously classified as diving behaviour by HMMs which could
255   explain the observed bias. For the same reason, we might explain that HMM are better on the boobies dataset than
256   with cormorants because these birds spend more time resting at the surface (see Table 1). The high BCE obtained by
257   HMM also suggests that in these situations HMM predicts erroneously dive with high confidence. In opposition, deep
258   learning seem better for discriminating such resting/bathing behaviours from dives. It would be nice to see further
259   why and how deep networks succeed in outperforming HMMs, but the interpretation is one of the inherent difficulty
260   of deep learning, and would require further analysis.

261   **Deep Learning for high-resolution GPS data analysis**   Better performance reported with DME-UNet closely re-
262   lated to the temporal resolution of the sampled dataset (Figure 1). In particular, from Peruvian boobies' trajectories
263   resampled at 15s and 30s, similar results were obtained through DME-UNet and HMM. However, we observed that
264   with the 5s resolution dataset, deep networks had outperformed HMM. This suggests that HMMs are great approaches
265   to trajectory segmentation with relatively large temporal resolution, and that they are unable to interpret movement
266   process at very fine spatio-temporal scales. Knowing that Peruvian boobies dives last about 2 seconds (see Table 1),
267   this suggests that the deep networks would be able to capture diving behavioural movements from a GPS track sampled
268   with a resolution in the same order of magnitude than dive durations. For the Guanay cormorants dataset, DME-UNet
269   had better results than HMM for all tested sampling resolution. However, duration in dives for cormorants is about
270   20s (Weimerskirch *et al.*, 2012), which could explain why even with the 30s sampled dataset deep networks had better
271   results than HMMs. With technological advances in sensor technology, ecologists are able to collect larger amount
272   of data than ever before. We might expect GPS with higher resolution in the future. Such an expected trend would
273   make more critical the exploitation of development of the proposed deep learning approaches to make the most of the
274   collected high-resolution animal trajectories (Malde *et al.*, 2020; Yoda, 2019)

275   **Importance of coverage data**   The GPS tracks were characterized by short gaps in the regularly sampled sequence of
276   locations, since these devices do not receive a satellite signal while submerged (Boyd, 2014; Wilson & Vandenabeele,
277   2012). These gaps are therefore sometimes directly considered indicative of diving behaviour (Weimerskirch *et al.*,
278   2012). Surprisingly, HMM did not take benefits from such variable for inferring dives. Nevertheless, the use of this
279   coverage information as input helped substantially deep networks to predict dives. Indeed, without coverage data,
280   the FCNet proposed by (Browning *et al.*, 2018) predicted dives with relatively low accuracy (see Figure 2). One
281   could therefore argue that deep networks are not relevant with longitude and latitude data only. However, the DME-
282   UNet proposed in this study demonstrated the ability of deep networks to classify animal behaviours in such situation.
283   Indeed, without the coverage ratio variable it obtained still an AUC about 5 % more than HMM.

284   **Trajectory data Representation in deep learning approaches**   A key breakthrough of our approach was thus to
285   use the so-called Distance Matrix Encoder (DME), which is a simple neural network block which aims to extract
286   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 (Bengio *et al.*, 2014). 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 method, and used instead one-hot representations initially introduced for text data (Nguyen *et al.*, 2021) as well as trajectory image (Endo *et al.*, 2016). Here, we explored distance matrix as this representation has been used to capture relevant spatial information for protein structure prediction (Senior, 2020). 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 (Ardakani & Hashimoto, 2017; Goodfellow, 2016; Rew *et al.*, 2019).

Application to other datasets and species   When considering learning-based schemes, the assessment of the generalization performance is of key importance. Beyond the evaluation of dive prediction performance on a trajectory 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 (Kawaguchi *et al.*, 2020; Zhang *et al.*, 2017) 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. Peruvian boobies from Guañape Island did have different foraging strategies from their counterparts from Pescadores island, with trips two times longer and dives slightly longer (see Table 1). It turned out that the two deep networks were very stable on this testing dataset predicting dives with high accuracy. The improvement of neural networks over the HMM fitted to the dataset was about 10%. These results 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. To make easier such applications, we share online the different models we pre-trained on the considered dataset [1]. Beyond such a direct application, trained models may also be 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. We then expect the models trained in this work to be potentially of interest for predicting dives in seabirds from other colonies and species of same genus.

Acknowledgements

---

[1]https://github.com/AmedeeRoy/BirdDL/results

Authors' Contribution

A.R., R.F. and S.B. developed the idea. A.R. conducted the analysis. All authors contributed to the redaction of the manuscript.

Data Accessibility

Trajectory data, pytorch code and fitted models are available on github repository, https://github.com/AmedeeRoy/BirdDL/

References

Ardakani, I.S. & Hashimoto, K. (2017) Encoding Bird's Trajectory using Recurrent Neural Networks. *Proceedings of 2017 IEEE International Conference on Mechatronics and Automation*, p. 6. Japan.

Authier, M., Spitz, J., Blanck, A. & Ridoux, V. (2017) Conservation science for marine megafauna in Europe: Historical perspectives and future directions. *Deep Sea Research Part II: Topical Studies in Oceanography*, **141**, 1–7.

Barraquand, F.D.R. & Benhamou, S. (2008) Animal movements in heterogeneous landscapes: Identifying profitable places and homogeneous movement bouts. *Ecology*, **89**, 3336–3348.

Bengio, Y., Courville, A. & Vincent, P. (2014) Representation Learning: A Review and New Perspectives. *arXiv:12065538 [cs]*.

Boyd, C. (2014) Movement models provide insights into variation in the foraging effort of central place foragers. *Ecological Modelling*, p. 13.

Browning, E., Bolton, M., Owen, E., Shoji, A., Guilford, T. & Freeman, R. (2018) Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds. *Methods in Ecology and Evolution*, **9**, 681–692.

Carter, M.I.D., Bennett, K.A., Embling, C.B., Hosegood, P.J. & Russell, D.J.F. (2016) Navigating uncertain waters: A critical review of inferring foraging behaviour from location and dive data in pinnipeds. *Movement Ecology*, **4**, 25.

Christin, S., Hervet, É. & Lecomte, N. (2019) Applications for deep learning in ecology. *Methods in Ecology and Evolution*, **10**, 1632–1644.

Cox, S.L., Miller, P.I., Embling, C.B., Scales, K.L., Bicknell, A.W.J., Hosegood, P.J., Morgan, G., Ingram, S.N. & Votier, S.C. (2016) Seabird diving behaviour reveals the functional significance of shelf-sea fronts as foraging hotspots. *Royal Society Open Science*, **3**, 160317.

Dean, B., Kirk, H., Fayet, A., Shoji, A., Freeman, R., Leonard, K., Perrins, C. & Guilford, T. (2015) Simultaneous multi-colony tracking of a pelagic seabird reveals cross-colony utilization of a shared foraging area. *Marine Ecology Progress Series*, **538**, 239–248.

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

Endo, Y., Toda, H., Nishida, K. & Ikedo, J. (2016) Classifying spatial trajectories using representation learning. *International Journal of Data Science and Analytics*, **2**, 107–117.

Fauchald, P. & Tveraa, T. (2003) Using First-Passage Time in the analysis of area-restricted search and habitat selection. *Ecology*, **84**, 282–288.

Furness, R. & Camphuysen (1997) Seabirds as monitors of the marine environment. *ICES Journal of Marine Science*, **54**, 726–737.

Goodfellow, I. (2016) Generative Adversarial Networks. *NIPS 2016*.

Guilford, T., Meade, J., Willis, J., Phillips, R., Boyle, D., Roberts, S., Collett, M., Freeman, R. & Perrins, C. (2009) 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*, **276**, 1215–1223.

Guilford, T.C., Meade, J., Freeman, R., Biro, D., Evans, T., Bonadonna, F., Boyle, D., Roberts, S. & Perrins, C.M. (2008) GPS tracking of the foraging movements of Manx Shearwaters Puffinus puffinus breeding on Skomer Island, Wales: GPS tracking of breeding Manx Shearwaters. *Ibis*, **150**, 462–473.

Hazen, E.L., Abrahms, B., Brodie, S., Carroll, G., Jacox, M.G., Savoca, M.S., Scales, K.L., Sydeman, W.J. & Bograd, S.J. (2019) Marine top predators as climate and ecosystem sentinels. *Frontiers in Ecology and the Environment*, **17**, 565–574.

Hooker, S.K. & Gerber, L.R. (2004) Marine Reserves as a Tool for Ecosystem-Based Management: The Potential Importance of Megafauna. *BioScience*, **54**, 27–39.

Jahncke, J. & Goya, E. (1998) Diets of the Guanay cormorant and Peruvian booby as indicators of the abundance and distribution of anchovy. *Boletin Instituto del Mar del Peru*, **17**, 15–33.

Joo, R., Boone, M.E., Clay, T.A., Patrick, S.C., Clusella-Trullas, S. & Basille, M. (2020) Navigating through the R packages for movement. *Journal of Animal Ecology*, **89**, 248–267.

Kawaguchi, K., Kaelbling, L.P. & Bengio, Y. (2020) Generalization in Deep Learning. *arXiv:171005468 [cs, stat]*.

Kingma, D.P. & Ba, J. (2015) Adam: A Method for Stochastic Optimization. *ICLR 2015*.

Lascelles, B., Notarbartolo Di Sciara, G., Agardy, T., Cuttelod, A., Eckert, S., Glowka, L., Hoyt, E., Llewellyn, F., Louzao, M., Ridoux, V. & Tetley, M.J. (2014) Migratory marine species: Their status, threats and conservation

management needs: MIGRATORY MARINE SPECIES: CONSERVATION MANAGEMENT NEEDS. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **24**, 111–127.

LeCun, Y., Bengio, Y. & Hinton, G. (2015) Deep learning. *Nature*, **521**, 436–444.

Maekawa, T. (2020) Deep learning-assisted comparative analysis of animal trajectories with DeepHL. *Nature Communications*, **11**, 5316.

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

McClintock, B.T. & Michelot, T. (2018) momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, **9**, 1518–1530.

Mendez, L., Borsa, P., Cruz, S., de Grissac, S., Hennicke, J., Lallemand, J., Prudor, A. & Weimerskirch, H. (2017) Geographical variation in the foraging behaviour of the pantropical red-footed booby. *Marine Ecology Progress Series*, **568**, 217–230.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P.E. (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, **105**, 19052–19059.

Nelson, B. (2005) *Pelicans, Cormorants, and Their Relatives*. Oxford University Press.

Nguyen, D., Vadaine, R., Hajduch, G., Garello, R. & Fablet, R. (2021) GeoTrackNet-A Maritime Anomaly Detector using Probabilistic Neural Network Representation of AIS Tracks and A Contrario Detection. *arXiv:191200682 [cs, stat]*.

Oppel, S., Beard, A., Fox, D., Mackley, E., Leat, E., Henry, L., Clingham, E., Fowler, N., Sim, J., Sommerfeld, J., Weber, N., Weber, S. & Bolton, M. (2015) Foraging distribution of a tropical seabird supports Ashmole's hypothesis of population regulation. *Behavioral Ecology and Sociobiology*, **69**, 915–926.

Paskze, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G. & Killen, T. (2019) PyTorch: An Imperative Style, High-Performance Deep Learning Library. pytorch.

Peng, C., Duarte, C.M., Costa, D.P., Guinet, C., Harcourt, R.G., Hindell, M.A., McMahon, C.R., Muelbert, M., Thums, M., Wong, K.C. & Zhang, X. (2019) Deep Learning Resolves Representative Movement Patterns in a Marine Predator Species. *Applied Sciences*, **9**, 2935.

Rew, J., Park, S., Cho, Y., Jung, S. & Hwang, E. (2019) Animal Movement Prediction Based on Predictive Recurrent Neural Network. *Sensors*, **19**, 4411.

Ronneberger, O., Fischer, P. & Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham.

Ropert-Coudert, Y., Beaulieu, M., Hanuise, N. & Kato, A. (2009) Diving into the world of biologging. *Endangered Species Research*, **10**, 21–27.

Schreiber, E.A. & Burger, J., eds. (2001) *Biology of Marine Birds*. CRC press, Hoboken, USA.

Seidel, D.P., Dougherty, E., Carlson, C. & Getz, W.M. (2018) Ecological metrics and methods for GPS movement data. *International Journal of Geographical Information Science*, **32**, 2272–2293.

Senior, A.W. (2020) Improved protein structure prediction using potentials from deep learning. *Nature*, **577**, 706–710.

Shoji, A., Elliott, K., Fayet, A., Boyle, D., Perrins, C. & Guilford, T. (2015) Foraging behaviour of sympatric razorbills and puffins. *Marine Ecology Progress Series*, **520**, 257–267.

Valletta, J.J., Torney, C., Kings, M., Thornton, A. & Madden, J. (2017) Applications of machine learning in animal behaviour studies. *Animal Behaviour*, **124**, 203–220.

Wakefield, E., Phillips, R. & Matthiopoulos, J. (2009) Quantifying habitat use and preferences of pelagic seabirds using individual movement data: A review. *Marine Ecology Progress Series*, **391**, 165–182.

Wang, G. (2019) Machine learning for inferring animal behavior from location and movement data. *Ecological Informatics*, **49**, 69–76.

Weimerskirch, H., Bertrand, S., Silva, J., Bost, C. & Peraltilla, S. (2012) Foraging in Guanay cormorant and Peruvian booby, the major guano-producing seabirds in the Humboldt Current System. *Marine Ecology Progress Series*, **458**, 231–245.

Weimerskirch, H., de Grissac, S., Ravache, A., Prudor, A., Corbeau, A., Congdon, B., McDuie, F., Bourgeois, K., Dromzée, S., Butscher, J., Menkes, C., Allain, V., Vidal, E., Jaeger, A. & Borsa, P. (2020) At-sea movements of wedge-tailed shearwaters during and outside the breeding season from four colonies in New Caledonia. *Marine Ecology Progress Series*, **633**, 225–238.

Weimerskirch, H., Le Corre, M. & Bost, C. (2008) Foraging strategy of masked boobies from the largest colony in the world: Relationship to environmental conditions and fisheries. *Marine Ecology Progress Series*, **362**, 291–302.

Weimerskirch, H. (2007) Are seabirds foraging for unpredictable resources? *Deep Sea Research Part II: Topical Studies in Oceanography*, **54**, 211–223.

Weinstein, B.G. (2018) A computer vision for animal ecology. *Journal of Animal Ecology*, **87**, 533–545.

Wilson, P.D. (2011) Distance-based methods for the analysis of maps produced by species distribution models. *Methods in Ecology and Evolution*, p. 11.

Wilson, R.P. & Vandenabeele, S.P. (2012) Technological innovation in archival tags used in seabird research. *Marine Ecology Progress Series*, p. 18.

Yoda, K. (2019) Advances in bio-logging techniques and their application to study navigation in wild seabirds. *Advanced Robotics*, **33**, 108–117.

Zhang, C., Bengio, S., Hardt, M., Recht, B. & Vinyals, O. (2017) Understanding deep learning requires rethinking generalization. *arXiv:161103530 [cs]*.
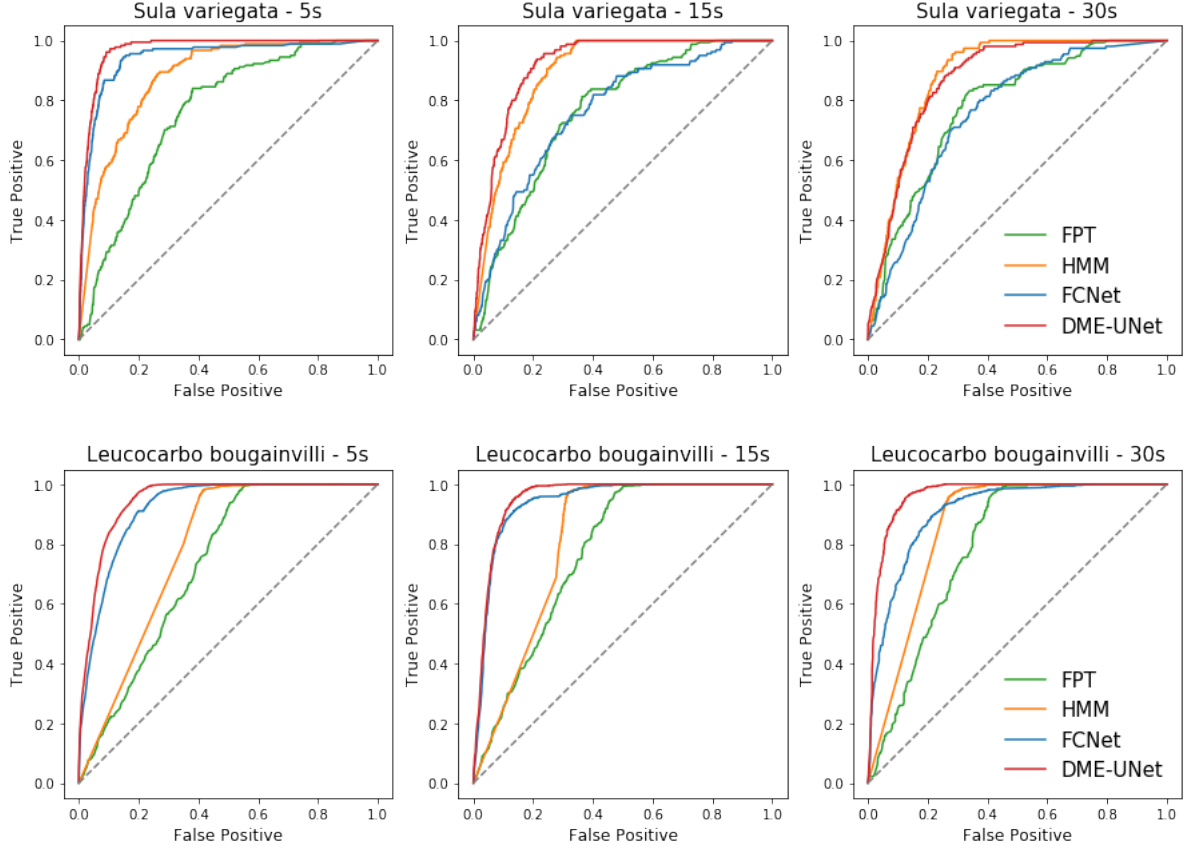
Figure 1: **Comparison of algorithms** - ROC curves obtained from the prediction of 4 algorithms, First-Time Passage (FPT), 3-states Hidden Markov Models (HMM), Fully-Connected Network (FCN) and U-Network with a Distance Matrix Encoder (DME-UNet) on 4 distinct test datasets derived from two seabirds species breeding in Pescadores Island from 2008 to 2013

Table 1: **Datasets Overview** - (m ± s) is for respectively mean and standard deviation. Resting has been defined as the proportion of time with speeds inferior to 1 m.s-1 associated to non-diving behaviour.

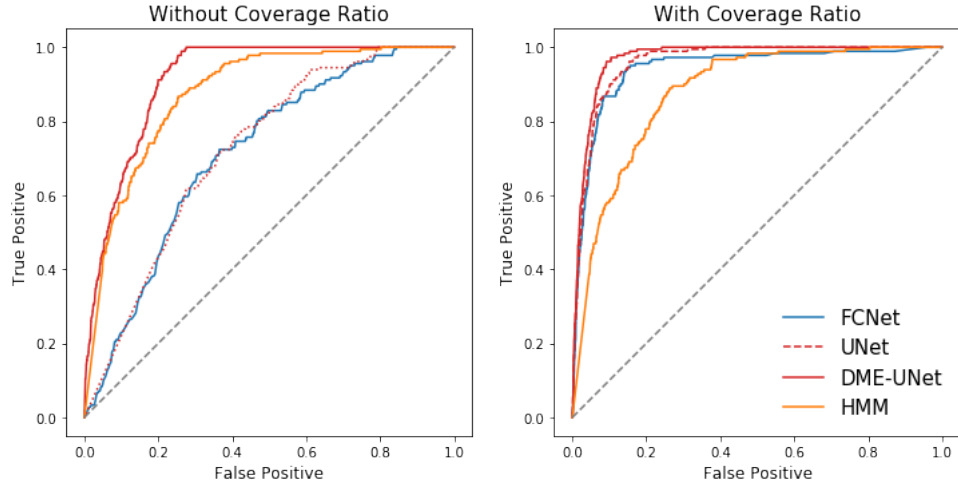| Species	Colony Location	Nb of trips	Trip Duration (min)	Dives (%)	Dives Duration (s)	Gaps (%)	Resting (%)
Sula variegata	Guañape Island	25	178 ± 88	0.5 %	3.6 ± 2.6	8.9 %	6.4 %
Sula variegata	Pescadores Island	133	65 ± 43	0.8 %	2.2 ± 0.9	13.4 %	6.7 %
Leucocarbo bougainvilli	Pescadores Island	76	123 ± 45	22.9 %	17.6 ± 12.6	40.9 %	21.8 %

Figure 2: **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
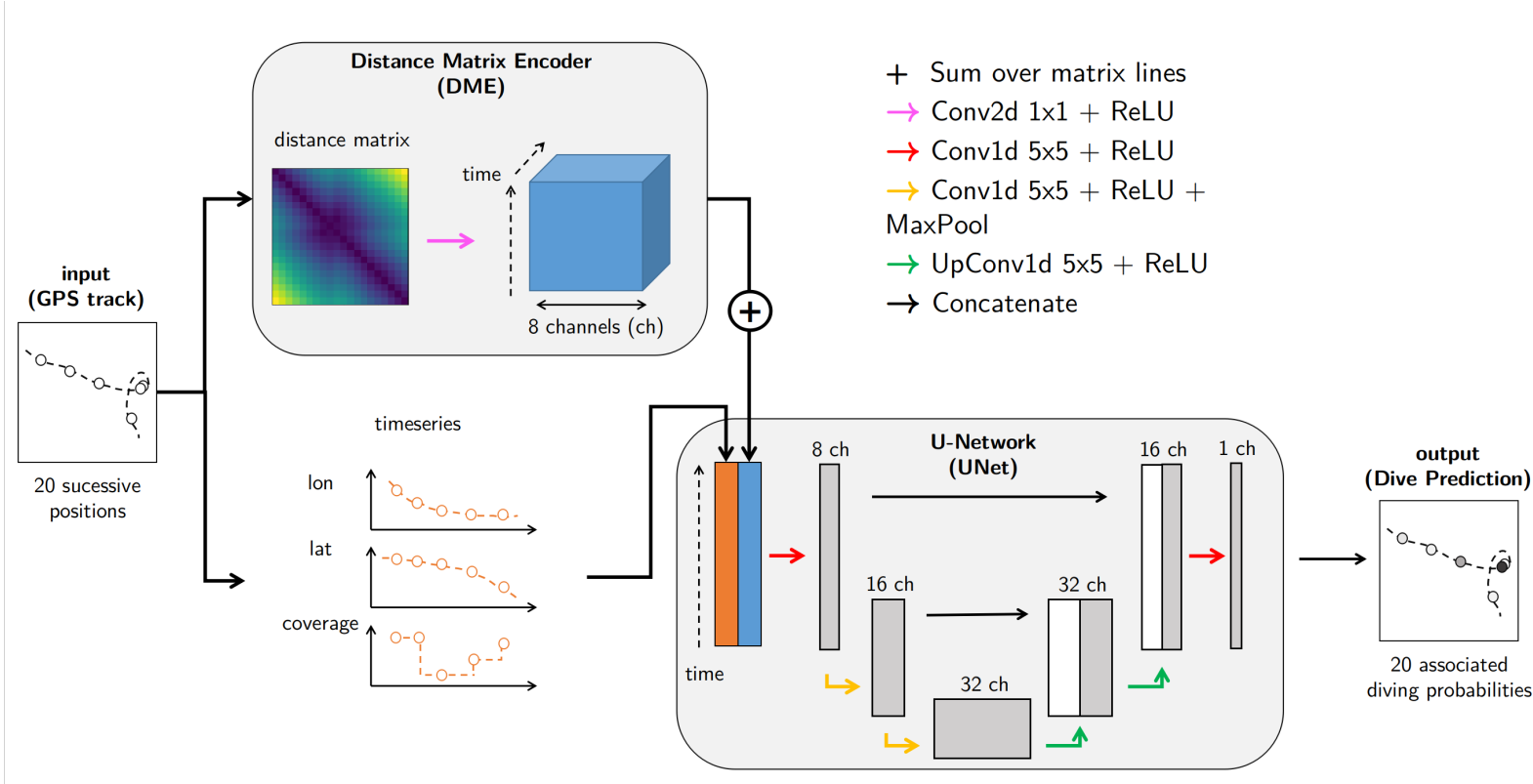

Figure 3: **DME-UNet Architecture** - This network is composed of two blocks entitled Distance Matrix Encoder (DME) and U-Network (UNet). It takes as input a GPS of 20 successive positions and outputs a diving probability to each of these positions. A channel refer to deep learning terminology and describes 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 (Paskze *et al.*, 2019)
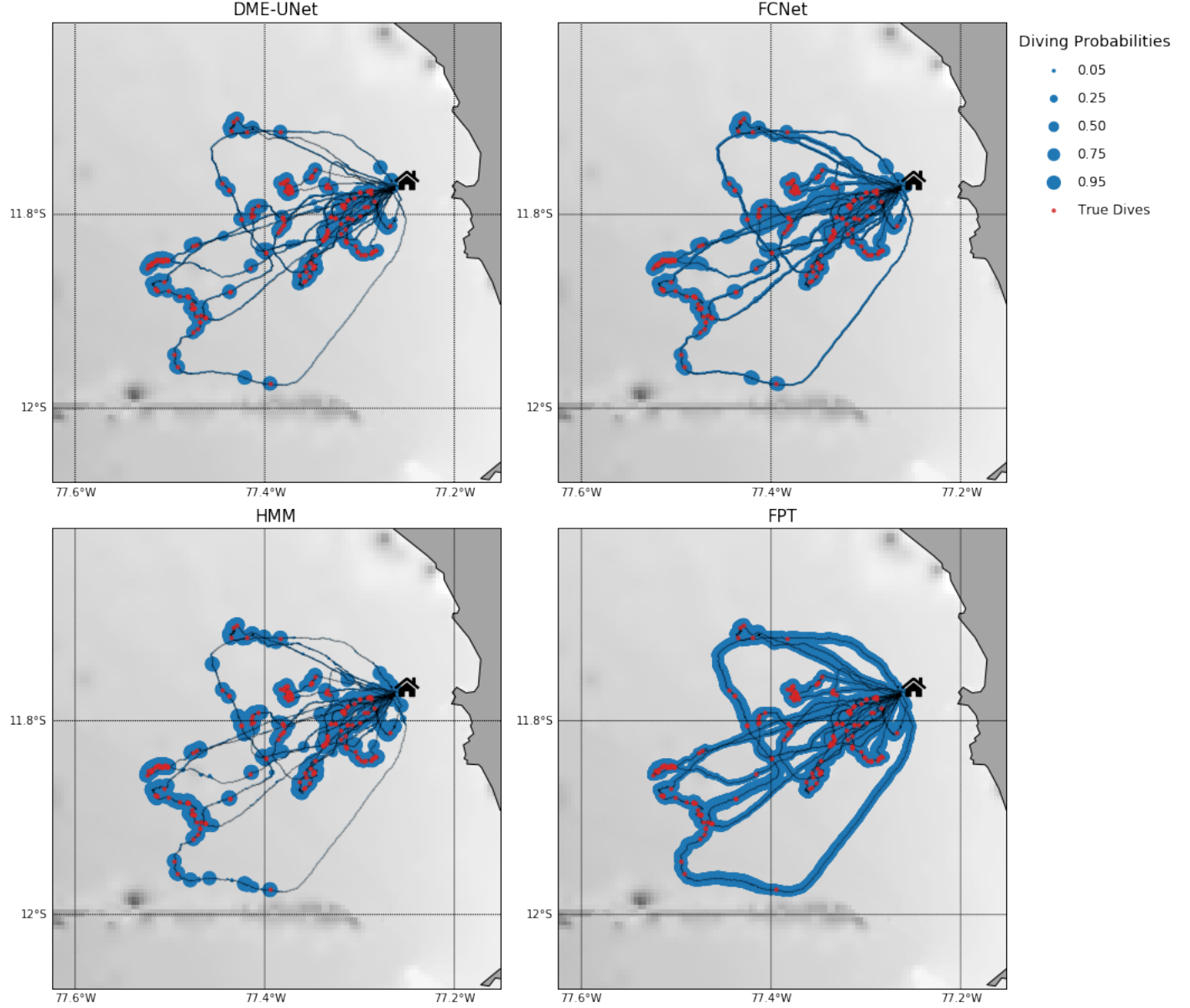
17

Figure 4: **Maps of Peruvian boobies predicted dives (test dataset)** - 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: our proposed network DME-UNet (top left), a fully connected network FCNet (top right), a Hidden Markov Model HMM (bottom left), and a First-Passage Time approach FPT (bottom right)
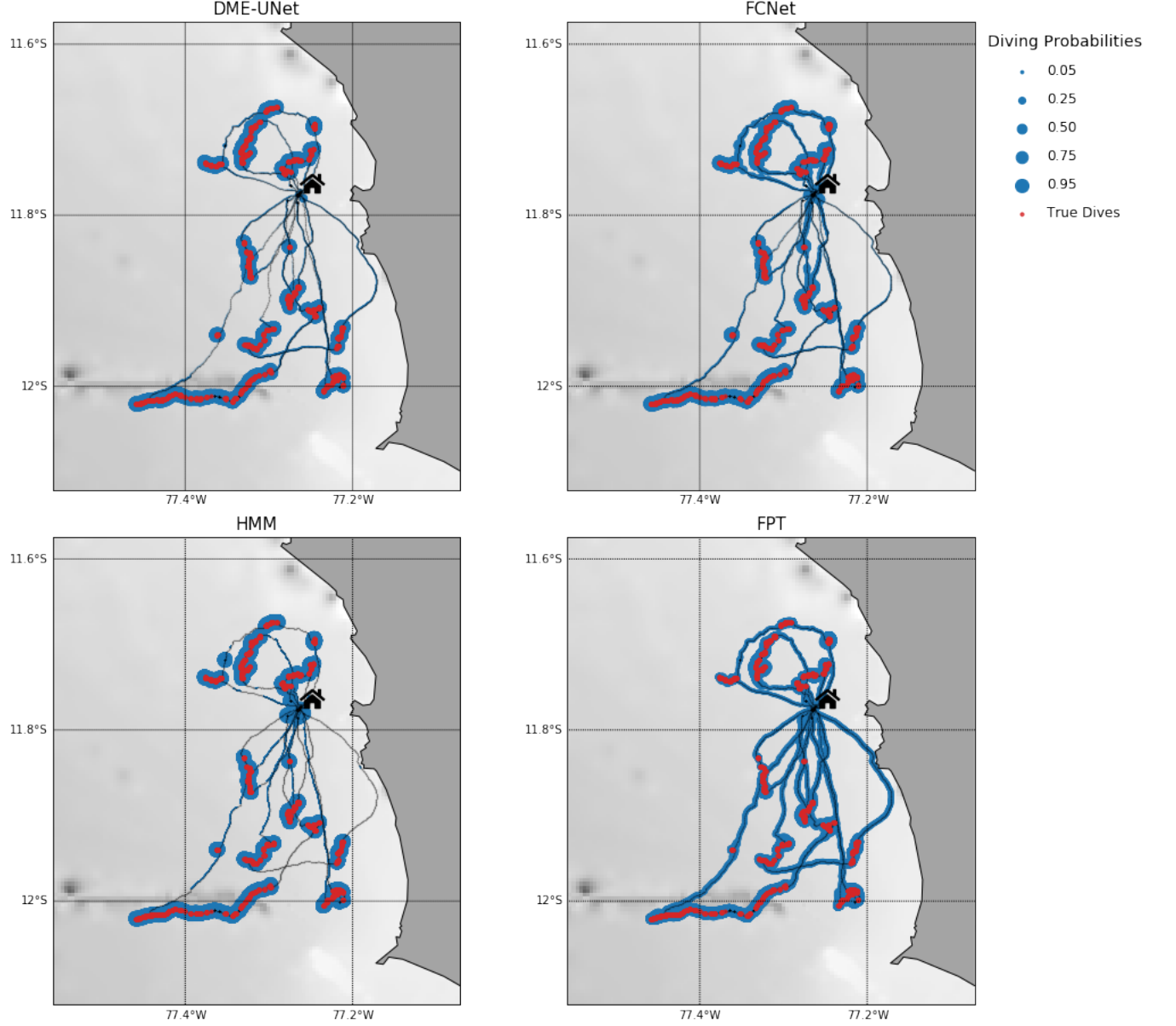
Figure 5: **Maps of Guanay cormorants predicted dives (test dataset)** - Red points represent true dive derived from TDR data. Blue points represents diving probabilities of each location with radius increasing for higher probabilities. These probabilities are the results of four methods: our proposed network DME-UNet (top left), a fully connected network FCNet (top right), a Hidden Markov Model HMM (bottom left), and a First-Passage Time approach FPT (bottom right)

Figure 6: **Maps of dive distributions of Peruvian Boobies from Guañape Island (test dataset)** - These density maps were obtained through Kernel Density Estimation with automatic bandwith procedure. The left map has been computed from true dives derived from TDR data. The other maps are estimated using dives probabilities derived from four methods: our proposed network DME-UNet (top left), a fully connected network FCNet (top right), a Hidden Markov Model HMM (bottom left), and a First-Passage Time approach FPT (bottom right)
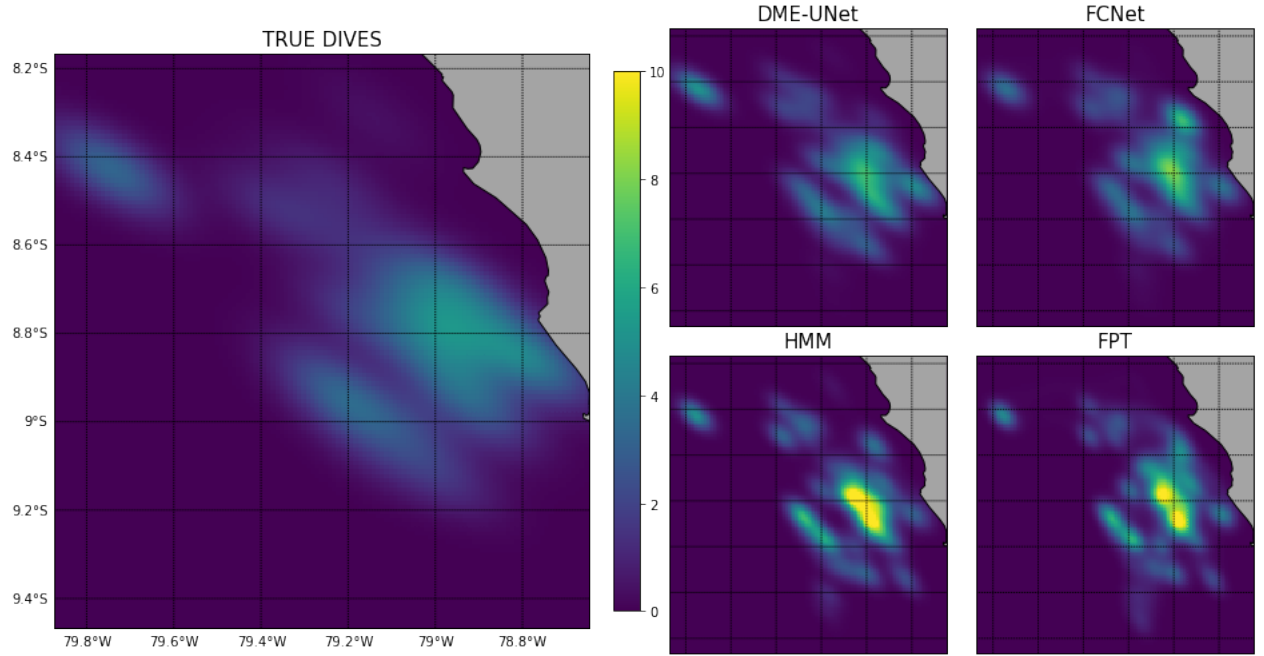
Table 2: **Overview of all Trained Deep Networks** - AUC means the Area Under the ROC curve, BCE is for binary cross entropy computed on the testing trajectories. Train and Validation Loss correspond to the loss estimation after model training on respectively training and validation trajectories. SV is for Peruvian boobies (*Sula variegata*), LB is for Guany cormorants (*Leucocarbo bougainvilli*)

Dataset	Resolution	Model	Variables	AUC	BCE	Train Loss	Validation Loss
SV (Pescadores)	5s	FPT	lon, lat	0.76	1.09	-	-
		HMM	step length, direction	0.87	2.98	-	-
		HMM	step length, direction, cov	0.88	2.83	-	-
		FCNet	lon, lat	0.71	0.52	1.05	1.12
		FCNet	lon, lat, cov	0.94	0.22	0.42	0.74
		UNet	lon, lat	0.72	0.48	0.97	1.12
		UNet	lon, lat, cov	0.96	0.21	0.47	0.61
		DME-UNet	lon, lat	0.92	0.27	0.57	0.66
		DME-UNet	**lon, lat, cov**	**0.97**	**0.16**	**0.36**	**0.56**
	15s	FPT	lon, lat	0.76	1.49	-	-
		HMM	step length, direction, cov	0.89	3.36	-	-
		FCNet	lon, lat, cov	0.76	0.94	1.65	1.80
		DME-UNet	**lon, lat, cov**	**0.92**	**0.45**	**0.86**	**1.07**
	30s	FPT	lon, lat	0.78	1.93	-	-
		HMM	**step length, direction, cov**	**0.88**	**1.91**	**-**	**-**
		FCNet	lon, lat, cov	0.76	1.30	2.13	2.40
		DME-UNet	lon, lat, cov	0.87	0.82	1.32	1.52
LB (Pescadores)	5s	FPT	lon, lat	0.73	5.2	-	-
		HMM	step length, direction, cov	0.78	7.10	-	-
		FCNet	lon, lat, cov	0.92	0.44	0.44	0.47
		DME-UNet	**lon, lat, cov**	**0.95**	**0.36**	**0.38**	**0.39**
	15s	FPT	lon, lat	0.77	5.7	-	-
		HMM	step length, direction, cov	0.81	4.76	-	-
		FCNet	lon, lat, cov	0.94	0.41	0.53	0.51
		DME-UNet	**lon, lat, cov**	**0.95**	**0.35**	**0.40**	**0.43**
	30s	FPT	lon, lat	0.79	6.0	-	-
		HMM	step length, direction, cov	0.86	4.26	-	-
		FCNet	lon, lat, cov	0.91	0.65	0.84	0.79
		DME-UNet	**lon, lat, cov**	**0.96**	**0.57**	**0.63**	**0.62**

Table 3: **Guañape projection** - The deep network fitted on the 5s sampled SV dataset in Pescadores have been used for dive prediction in Guañape. AUC is for area under the roc curve. BCE is the binary cross entropy. Hellinger Distance corresponds to the distance of the diving distribution maps estimated with kernel density estimations and plotted in Figure 6 to the correct diving distribution.

Dataset	Resolution	Model	Variables	AUC	BCE	Hellinger Distance
SV (Guañape)	5s	FPT	lon, lat	0.78	0.93	31.7
		HMM	step length, direction, cov	0.88	3.70	29.6
		FCNet	lon, lat, cov	0.97	0.10	25.0
		DME-UNet	lon, lat, cov	0.96	0.07	15.1