



OPEN Towards automated long-term acoustic monitoring of endangered river dolphins: a case study in the Brazilian Amazon floodplains

Florence Erbs¹, Marina Gaona^{1,2}, Mike van der Schaar¹, Serge Zaugg¹, Emiliano Ramalho², Dorian Houser³ & Michel André¹✉

Using passive acoustic monitoring (PAM) and convolutional neural networks (CNN), we monitored the movements of the two endangered Amazon River dolphin species, the boto (*Inia geoffrensis*) and the tucuxi (*Sotalia fluviatilis*) from main rivers to floodplain habitats (várzea) in the Mamirauá Reserve (Amazonas, Brazil). We detected dolphin presence in four main areas based on the classification of their echolocation clicks. Using the same method, we automatically detected boat passages to estimate a possible interaction between boat and dolphin presence. Performance of the CNN classifier was high with an average precision of 0.95 and 0.92 for echolocation clicks and boats, respectively. Peaks of acoustic activity were detected synchronously at the river entrance and channel, corresponding to dolphins seasonally entering the várzea. Additionally, the river dolphins were regularly detected inside the flooded forest, suggesting a wide dispersion of their populations inside this large area, traditionally understudied and particularly important for boto females and calves. Boats overlapped with dolphin presence 9% of the time. PAM and recent advances in classification methods bring a new insight of the river dolphins' use of várzea habitats, which will contribute to conservation strategies of these species.

In recent years, the International Union for the Conservation of Nature (IUCN) reassessed the status of the two river dolphin species of the Amazon, the pink river dolphin (*Inia geoffrensis*) and the tucuxi (*Sotalia fluviatilis*) from 'Data deficient' to 'Endangered'^{1,2}. With these new categorizations, all five remaining river dolphin species are now officially considered threatened with extinction. This alarming situation reflects the intricate combination of direct and indirect threats posed to river dolphins worldwide where conflict with commercial fisheries (i.e. competition for resources and damages to fish nets)^{3–6} are aggravated by the high level of anthropogenic pressure on tropical freshwater ecosystems⁷. In the Amazon basin, the main direct threats on river dolphin populations are being captured for bait in the commercial fishery of piracatinga *Calophysus macropterus* and entanglements in gillnets^{3,8–13}. Furthermore, disruption of hydrological connectivity through dam construction, mining, agriculture and cattle ranching is profoundly impacting river ecological functions and increasingly degrading freshwater ecosystems¹⁴. As a consequence, Amazon River dolphin populations are declining. Recent studies highlight an alarming population reduction of 50% every 10 years for boto and every 9 years for tucuxi⁶. Current models of population viability predict a 95% reduction of boto population within 50 years¹⁵. These two studies were conducted in a protected area, the Mamirauá Sustainable Development Reserve (Reserva de Desenvolvimento Sustentável Mamirauá—RDSM), where anthropogenic pressures on river dolphins are likely reduced compared to non-protected areas.

Amazonian River dolphins inhabit a unique environment characterised by radical seasonal changes in water regimes. Half of the year, large areas of the riverine forests are flooded¹⁶, extending the habitat of the aquatic animals from the main rivers to large areas of fringing floodplains locally called várzea and igapó. These provide access to a highly complex and resource-rich environment formed by submerged vegetation. The seasonal 'flood pulse' is the major factor driving the distribution and movements of many Amazonian aquatic species, including freshwater fish that undertake small-scale seasonal movements between the main rivers and floodplains for completion of their life cycle^{17,18}. Fish families known to constitute the major part of the river dolphin diets, such

¹Laboratori d'Aplicacions Bioacústiques, Universitat Politècnica de Catalunya - BarcelonaTech, Barcelona, Spain. ²Instituto de Desenvolvimento Sustentável Mamirauá, Tefé, Brazil. ³National Marine Mammal Foundation, San Diego, USA. ✉email: michel.andre@upc.edu

as the Characids (Characiforms) and Doradid catfish (Siluriforms)¹⁹, display such synchronised lateral migrations and Amazon dolphin seasonal changes in habitat density have been related to the migration of fish^{20–23}.

During the low-water season, river dolphins are concentrated in the main rivers^{21,24}. Preferred habitats are confluences where two or more water streams join together (e.g. small tributaries or larger channels connect to the main rivers), bays, lakes, and river margins^{23–26}. When water levels start to rise, river dolphins follow fish movements and enter the floodplains through river channels. The botos are highly adapted to the complex and cluttered environment of the floodplains; anatomical specialisations such as unfused cervical vertebrae provide extra neck flexibility and a unique shoulder joint allows for a broader rotation range of the flippers, greatly increasing manoeuvrability²⁷. The botos tend to disperse across the mosaic of newly inundated habitats, including floodplain channels, internal lakes (seasonally isolated from the rivers) and flooded forests, while the tucuxis usually remain in the deeper parts of the floodplain, i.e. occupying the river channels and small confluences^{21,28–30}. Once water levels start dropping, dolphins move back to the main rivers, probably to avoid entrapment^{5,21}. Botos seem to use preferentially different habitats according to their age and reproductive status. Male botos and females without calves appear to have similar habitat preferences and exhibit similar movement patterns from main rivers to the várzea²¹. On the other hand, females with calves and immatures animals spend more time in the várzea habitats (bays or small confluences, channels) than in the main river, sometimes not returning to the main rivers during the low water season, if the water level remains sufficiently high⁵.

The complexity of the floodplain habitats makes surveying the internal lakes and the flooded forest extremely challenging. Until now, the vast majority of information on river dolphin distribution has relied on boat-based visual surveys that are usually conducted along rivers. Unfortunately, inconspicuous surface behaviour and occurrence in remote areas make river dolphins difficult to monitor through visual techniques. Aerial visual monitoring methods have recently emerged as an alternative or complementary approach, with the use of unmanned aerial vehicles (drones) or non-rigid airship systems (blimp) focusing on surveying dolphins in the main rivers^{31,32}. These techniques, while promising in improving count estimates³¹, are restricted to open areas and cannot be applied in the flooded forest where the tree canopy prevents visual detection from above. Satellite tracking studies can inform on movements and habitat preferences of animals^{33,34} but come with major limitations regarding cost and operational complexity, risks associated with animal capture and tag attachment³⁵. A study linked mortality events with tag implants in belugas³⁶, whereas Martin et al.³⁷ did not find tagging affected survival rate of the tagged botos in the RDSM. While risk levels might be partially species specific, the recently published 'best practice guidelines for cetacean tagging' recommends that this technique should be limited to research questions that cannot be addressed by other methods³⁸.

The use of Passive Acoustic Monitoring (PAM) to conduct surveys of river dolphins takes advantage of the quasi-continuous vocal production of the river dolphins³⁹, resulting in a high acoustic detectability. River dolphins produce echolocation clicks to sense their environment, to orientate and forage, as well as other vocalisations including boto-specific low frequency pulsed vocalisations⁴⁰, and whistles⁴¹. Echolocation clicks are produced almost continuously and constitute a reliable, consistent, highly detectable acoustic means for investigating dolphin presence (for a review of methods and applications see⁴²). PAM has been successfully applied in Amazon River dolphin studies including population distribution and habitat use^{43–45}, and vocal behaviour^{39,44,46–51}. With technological advances, the acoustic presence of dolphins can be detected in real time⁵². However, the Amazonian environment produces numerous challenges to signal detection and classification. The complexity of the freshwater ecosystem soundscape includes various impulsive sound sources in addition to dolphin echolocation clicks, such as cavitation noise from ship engines, rain, and high frequency stridulating insects. Additionally, the complex acoustic propagation conditions in the constrained, shallow-water environments are complicated by suspended sediment and detritus that alters signal propagation via reflection, refraction and scattering.

Several methods exist to automatically detect impulsive sounds such as echolocation clicks. Support vector machine methodology and energy-based click detectors have been used for odontocete clicks (Amazon River dolphin⁴³, beaked whales⁵³), some of which have been coupled with neural networks (Indo-Pacific humpback dolphins⁵⁴). Other studies have used Gaussian mixture models (GMMs) with signals represented by cepstral features⁵⁵, entropy⁵⁶, or Gaussian-kernel-based networks⁵⁷ and feed forward neural networks, the first kind of artificial neural networks⁵⁸. The most recent advances in the field of automatic classification of acoustic signals use Deep Neural Networks (DNN). For cetacean echolocation clicks, this approach has been developed for sperm whale⁵⁹ and other odontocete clicks^{60,61}. So far, Convolutional Neural Networks (CNN), a class of DNN, have not been used for classifying river dolphin signals.

In this study we combined PAM techniques with state-of-the-art automatic classification algorithms based on CNN to monitor river dolphin presence in different floodplain habitats inside the Mamirauá Sustainable Development Reserve (RDSM) in the state of Amazonas, Brazil. Specifically, we focused on (1) developing a reliable classification model that can accurately discriminate between several types of impulsive sounds present in the floodplain soundscapes (echolocation clicks from dolphins, boat engine, and rain); (2) automatically detecting dolphin acoustic presence in large datasets from different floodplain habitats, including permanent and seasonally flooded sites; (3) identifying temporal overlap between dolphin and boat presence. The classifiers developed here will form part of the conservation strategy of the RDSM.

Materials and methods

Study area. The study area covers about 800 km² in the Mamirauá Sustainable Development Reserve (RDSM), in the state of Amazonas, Brazil (Fig. 1). The RSDM comprises approximately 11,000 km² at the confluence of the Solimões river (upper Amazon River) and the Japurá River and is the largest Brazilian protected area dedicated to the conservation of flooded rainforests. The RDSM is inhabited by local populations along riv-

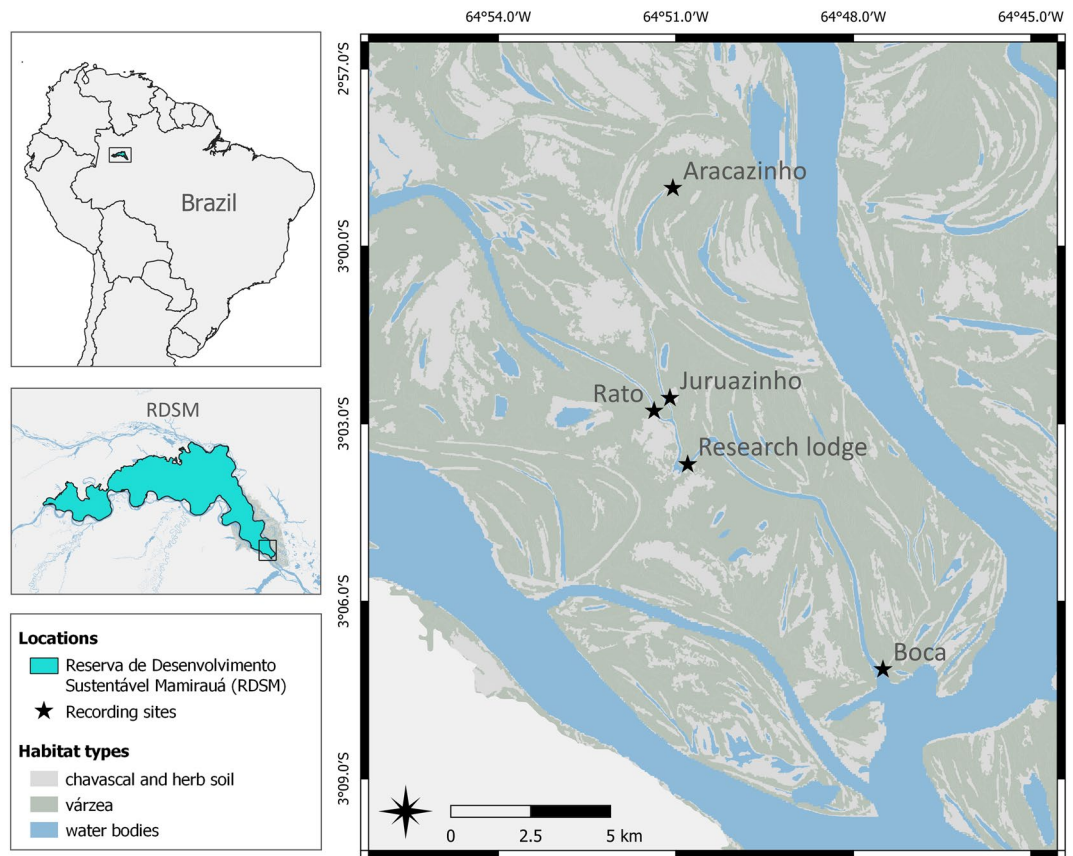


Figure 1. Map of the study site, the Mamirauá Sustainable Development Reserve (RDSM), Amazonas state, Brazil, and recording locations inside the RDSM. The habitat types are modified from Ferreira-Ferreira et al.⁶⁴.

ers and lakes, that are involved in managing and monitoring biodiversity through sustainable development. This protected status ensures that the areas under protection contain predominantly unmodified natural systems.

The region is formed by *várzea* (white-water river floodplain) habitat, a lowland forest seasonally flooded by white-waters from the Amazon, with an average annual variation in water levels of 10–12 m⁶². The region also contains patches of dense vegetation dominated by shrubs (*chavascal*) and herbaceous vegetation. During the dry season (September to March), the forest is intersected by numerous lakes and channels. During the wet season (April to August), floodwaters progressively inundate the forest, submerging most of the dry land. Two river dolphin species are present in the RDSM, the boto or pink river dolphin and the tucuxi. Each species has specific habitat preferences inside and outside the *várzea* (i.e. the main rivers bordered by the floodplains)²⁴.

This study encompasses four different *várzea* habitat types comprising permanently and temporarily flooded habitats. (1) ‘*Ressacas*’ are defined as shallow bays adjacent to the river channel, with low velocity current and often fringed with floating vegetation²¹. (2) River channels (*paraná*s) are minor aquatic systems that connect the rivers to the floodplain lakes. (3) Internal lakes are permanently flooded systems inside the floodplain. Depending on their geomorphological type and their location, lakes display various degrees of connectivity with the main rivers and can be covered with free floating aquatic macrophytes. (4) Low *várzea* forests, like the flooded forest in this study, are inundated for more than 3 months of the year (as opposed to high *várzea* forests) and covered with trees and shrubs⁶³.

Data acquisition. The acoustic data were acquired through different recording systems. An overview of the data collection is shown in Table 1. The Providence node, funded by the National Marine Mammal Foundation (NMMF), is composed of an icListen digital hydrophone (24-bit Smart hydrophone SB2-ETH model, Ocean Sonics, Canada, sensitivity –170 dB re 1 V/uPa) connected to a SONS-DCL real-time processing system (Sonsetc, Spain). This node was deployed from the Mamirauá floating research lodge, with direct access to the Mamirauá channel, at a depth of 5 m. The system sampled at 128 kHz with 24-bit resolution and without additional gain. Raw data from this system was transferred to Network Attached Storage (NAS) at the Uakari floating lodge (close to the research lodge) whenever the network was available. Data used from this system was recorded between July 2019 and June 2020.

Additionally, four Wildlife Acoustics SM4 recorders (Wildlife acoustics, USA) equipped with HTI-96-Min hydrophones (High Tech Inc., USA, sensitivity –165 dB re 1 V/uPa) were deployed in different *várzea* habitat types. One system was located at *Boca*, a *ressaca* (or bay) at the entrance of the Mamirauá Channel (situated at 2 km from the confluence with the Japurá River), deployed from a floating house at a depth of 3–5 m, and

Recording site	Habitat type	Recorder	Latitude / longitude	Start / end date	Sampling Frequency(kHz)	Duty cycle	Recording duration (h)
Research lodge	River channel	Providence node	3°3'41.81"S	2019-06-23	128	Continuous	2121
			64°50'48.59"W	2020-06-13			
Boca	Bay	SM4	3° 7'9.60"S	2019-11-05	96	1 on/9 off	393
			64°47'30.88"W	2020-04-29			
Araçazinho	Internal lake	SM4	2°59'1.34"S	2019-07-09	192	1 on/9 off	139
			64°51'3.28"W	2019-09-05			
Juruazinho	Flooded forest	SM4	3°2'34.69"S	2019-07-09	192	1 on/9 off	139
			64°51'6.07"W	2019-09-05			
Rato	Flooded forest	SM4	3°2'47.18"S	2019-02-06	96	10 on/110 off	139
			64°51'22.13"W	2019-05-29			
Multiple boat-based surveys	River channel, bay, internal lakes	Sound Trap	Multiple sites within MSDR	2019-06-01 2020-05-01	512	Continuous during survey	83

Table 1. Overview of the data collection at the RDSM.

recorded from November 2019 to April 2020. Two other systems were deployed inside the flooded forest, fixed to trees: *Rato* (between the Mamirauá channel and a depression lake) from February 2019 to May 2019; *Juruazinho* (between a *parana* and a depression lake) from July 2019 to September 2019. The last system was deployed at *Araçazinho* (a scroll lake with intermediate connectivity with the Japurá River) from July 2019 to September 2019. The latter three were deployed in areas that are not connected to the main rivers during the dry period of the year. All SM4 systems were sampled at 96 kHz with 16-bit resolution and without additional gain.

Separately from the autonomous recorders, manual recordings were collected during monthly boat-based surveys in the Mamirauá reserve, from July 2019 to May 2020. The recordings were made in close vicinity of dolphin groups (< 50 m) using a SoundTrap 300 HF (Ocean instruments, New Zealand, sensitivity – 176 dB re 1 V/uPa) deployed from the boat at 3 m depth. This data provided ground truth data for the training of the classifier as the signals came with visual identification.

CNN classification procedure. *Data annotation.* Data labelling was performed using a Python-based custom graphical interface (labelling tool) that displayed segments of spectrograms of a given duration (here, 5 s) and allowed to annotate signal extension in time and frequency with bounding-boxes and assign soundtypes (classes). Sound types were assigned only one label and multiple labels could be present independently at the segment level. The annotations were incrementally stored in a dedicated database (the Controlled Acoustic Repository database, or CAR DB). Frequency and time boundaries of each signal were then easily extracted from the bounding-boxes.

A subset (2.63 h) of data were selected from the boat-based recordings (9% manually selected to contain river dolphin signals and 91% randomly selected to include a representative sample of soundscape variability). A subset (6.57 h) of data was also randomly selected for the Providence node over a 2-day period. Both subsets of data were initially labelled with two target classes: echolocation clicks and boat engine noise, hereafter referred to as ‘click’ and ‘boat’ classes. Additional background sounds (e.g. aquatic insects, dolphin whistles, fish calls, ...) were labelled to ensure proper representation in the training set. In addition to labelled segments, segments containing only background sounds were included. Echolocation clicks were not separated by species as there is currently no click-based method developed to discriminate between the two river dolphin species (boto and tucuxi) present in the study area. After a first training of the classifier, this initial labelling was completed by an active learning labelling (see paragraph below). An overview of the data annotation is given Table 2.

Data augmentation. Data augmentation to artificially expand the training dataset was performed in two steps, focusing on providing additional spectrograms containing soundtypes that were underrepresented in the training data. The first step was data oversampling, where labels from small classes were duplicated to have a minimum number of labels per epoch for the training. In this case, the minimum number of labels was set to 300. Duplication was done by duplicating segments that contained the underrepresented classes. Since a segment may have contained multiple labels, this potentially also duplicated labels of other classes. Second, with the duplicated data set, “on-the-fly” data augmentation was performed^{65,66}. This was done by transforming the original data, including transformations in the frequency domain (small circular shift along the frequency axis), time shifts (small circular data shift along the time axis), contrast adjustments by modifying the spectrogram power, and time-warping (stretching the spectrogram along the time-axis and clipping it back to original length). For each epoch, the whole training set was run through the transformations, providing slightly different segments from the base data. A summary of the label dataset can be found in Table 2.

Automatic classification of acoustic signals using a Convolutional Neural Network (CNN). The approach to automatically classifying acoustic signals was based on a Convolutional Neural Network (CNN), a class of deep learning algorithms, that was trained to automatically perform image-based classification on the visual representation (spectrogram) of sounds⁶⁷. This approach required input spectrograms with annotations identifying

Dataset	Initial labelling			Active learning labelling		Total	Train Pre-DA # segments	Train DA # segments	Test #segments
	Boat-based survey	Boat-based survey	Research lodge	Research lodge	Várzea sites	All			
Selection	Manual	Automatic (random)	Automatic (10 s. per minute over 2 days)	Semi-automatic (Random above threshold)	Semi-automatic (Random above threshold)	All	NA	NA	NA
Duration of data subset (h)	0.23	2.40	6.57	0.41	1.56	11.17	NA	NA	NA
Use	Test	Train	Train/test (Equal split)	Train	Train/test (Equal split)	Train/test	NA	Train	Test
no. segm. with echolocation clicks	138	1230	96	25	76	1565	1337	4106	228
no. segm. with ship noise	0	308	104	69	178	659	530	1130	129
no. segm. with rain	0	0	0	0	378	378	188	382	190
no. segm. without any of the 3 classes (background)	30	403	4533	204	362	5532	3046	3204	2486

Table 2. Overview of annotated data used for the initial model (first 3 rows) and the final model (all 5 rows).

the sound classes to be classified. Network prediction consisted in the successive convolution of trained filters (or kernels) on the spectrogram image to extract relevant features for class prediction (classification). Over a series of epochs (iterations of the network over the training dataset), the CNN model described below was then trained using the binary cross-entropy loss function (with the loss summing over the labels and the batch) and the Adam optimizer⁶⁸ to perform predictions on the presence of the classes within a spectrogram. The dataset used to produce the CNN classifier was split between a training and a testing dataset, each set containing data from the boat-based survey (different recording sessions), the research lodge (different days), and the várzea (equal split on a random selection) (see Table 2). Based on the output prediction, the classifier performance was evaluated.

Data pre-processing and CNN architecture. We used a convolutional neural network (CNN) with the architecture as shown in Supplementary Materials Figure S1. To compute the input features, the wave forms that were not already sampled at 96 kHz were first down-sampled to 96 kHz by low pass filtering and decimation by 2 or 4 depending on the original sampling frequency. Then all wave form data was combined and segmented into 5-s non-overlapping segments. The classifier operated in a frequency band from DC to 48 kHz. Each segment was Fourier transformed using a 2048 sample Hamming window with 1112 samples of overlap to compute the power spectral density (PSD). This resulted in a time–frequency matrix of 1024 by 1024 PSD values. The PSD values were log-transformed element wise. Then the frequency dimension of the matrix was Mel scale-transformed such that 1024 linear frequencies were mapped to 128 log-spaced Mel bands. Finally, the matrix was equalised over the time dimension (subtraction of the median for each Mel row) and then presented as input to the classifier. The first part of the CNN classifier consisted of 5 blocks that were identically structured, with each block containing two 2D convolution layers that were using the same number of 3 × 3 filters with the rectified linear activation function, followed by a max-pooling layer (using 2 × 2 filter without overlap). The number of filters was changed each iteration of a block: 32, 64, 96, 128, 160. All convolution layers were preceded by a batch normalisation layer. After the last convolution block, the resulting feature maps were reshaped to obtain a two-dimensional matrix which was run through two one-dimensional convolution layers with batch-normalization, each convolution with 256 filters, kernel size 1 and rectified linear activations. Finally, the per-class output was obtained through a 1-dimensional convolution layer with sigmoid activation.

Active learning procedure. Considering the large dataset and relatively rare presence of some of the target signals, an active learning approach was followed. After first training using the boat-based and the Providence node dataset, the classifier model was evaluated on other data sets (Boca and várzea sites) to identify misclassifications, i.e. new sound types not present in the soundscape of the initial training data that conflicted with classified classes. Predictions from the initial model were manually checked for out-of-sample errors (generalisation errors). Custom Python scripts were created to automatically extract and display spectrograms of randomly selected positively classified sounds, above a threshold of the classification output selected from the model performance (see paragraph below). Misclassifications for the classes ‘click’ and ‘boat’ were identified through visual observation of spectrogram segments, and typically contained unrelated sounds with similar/overlapping acoustic characteristics to that of the sound types of interest (e.g. rain, see Fig. 2). The audio segments containing correctly classified and misclassified sounds were automatically extracted and combined with the initial training dataset for retraining of the model. An overview of sound labels by site used for initial classification and active learning can be found in Table 2.

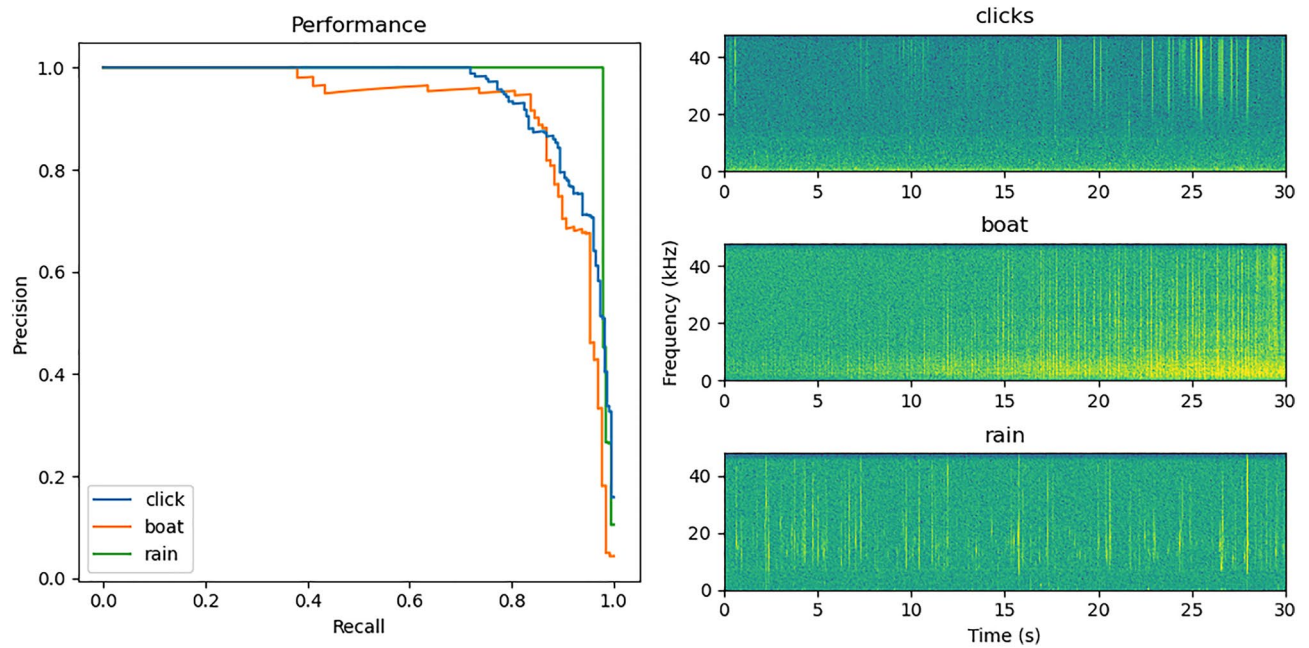


Figure 2. Precision-recall curves for class click, boat and rain (left) and spectrograms of the three corresponding impulsive sounds classified (right).

Classification performance evaluation. The final classifier was trained with a data set consisting of the initial and active learning data sets, and classified three impulsive sound types: echolocation clicks, shipping noise and rain. First, we produced a summarised classification output calculated over a 5-s segment. The dimension of the (scaled) spectrogram submitted to the classifier was a 512×128 matrix (time x frequency bins). After the iteration of convolution-pooling layers, this was reduced to a 16×3 matrix, and classification was performed on each column producing 16 output values per class (3) per segment. After CNN training, the 16 values per class were summarised by taking the mean over the values between the 75th and 100th percentile as the segment-based classification result, referred to as Q75. After evaluating several options we found that the 75th percentile had best performance for a signal that we expect to be repeated several times within the spectrogram. This approach should reduce the number of spurious high classification values under the assumption that our target classes have multiple high output values per 5-s segment.

Some performance metrics, for example accuracy, can be deceptive when considering the actual performance of a classifier, since the data are unbalanced (i.e., some classes are more prevalent than others in the dataset) and when models give the probability score⁶⁹. Here, we preferred reporting on precision and recall, metrics that can provide a better insight when dealing with unbalanced class representation^{69,70}. The precision (also called Positive Predicted Value, PPV) was defined as the fraction of observations predicted to be positive that were in fact positive. The recall (or True Positive Rate, TPR) was the fraction of observations classified as positive out of all positive observations (i.e. a probability of detection). A precision-recall curve was plotted wherein these two performance metrics were respectively plotted on the x and y axis for a sequence of decision thresholds. Then, the Average Precision (area under the precision-recall curve, AP) was computed for each class to evaluate the individual class performance. The global performance of the classifier over the three classes was computed using the mean average precision (mAP) through micro averaging, where all classification results from the different classes were combined to compute a single precision-recall curve, and then the AP from this curve was computed to give the mAP; this approach accounted for class unbalance.

$$\text{Precision (PPV)} = \text{True Positive} / (\text{True Positive} + \text{False Positive})$$

$$\text{Recall (TPR)} = \text{True Positive} / (\text{True Positive} + \text{False Negative})$$

$$\text{True Negative Rate (TNR)} = \text{True Negative} / (\text{True Negative} + \text{False Positive})$$

$$\text{Average Precision (AP)} = \sum_{t=0}^{t=T-1} [\text{TPR}(t) - \text{TPR}(t+1)] * \text{PPV}(t)$$

with T the number of thresholds and $\text{TPR}(T) = 0$, $\text{PPV}(T) = 1$.

Amazon River dolphin presence and boat passage. *Choice of optimal classification threshold.* The entire dataset except the boat-based surveys (2931 h of recordings) was evaluated with the final model. For each

5 s segment, this gives 3 predicted scores for the 3 target classes (*click*, *boat*, *rain*). Validation outputs (Average Precision curves and scatter plots) of the final model were used to identify the optimal classification threshold for each sound type. The optimal threshold (or decision threshold) was determined as the value point of the Average Precision curve where both Precision and Recall are equal. Decision thresholds were then tested and evaluated per location on randomly selected classified segments from the entire dataset. The TPR and the True Negative Rate (TNR) were assessed and the thresholds were adjusted and re-evaluated if necessary. The decision threshold for each class and location was then compared to the prediction score of each segment classified, and the segment was assigned as positive for dolphin (*click* class), ship (*boat* class) or rain (*rain* class) presence if the prediction score was above the class threshold.

Metric of daily acoustic presence. Five second segments with CNN scores above the threshold were counted as acoustic occurrences of river dolphins. Daily acoustic presence is a proportion that was obtained as the daily duration with acoustic occurrences divided by the total daily recording duration. Reporting a proportion compensates for the differences in the recording duty cycles between the Providence node and the SM4 recordings. For the várzea sites, where dolphin passages were assumed to be infrequent and further away from the recording equipment with complex propagation conditions (flooded forest), the positive detections were manually checked.

Additional filters. For the várzea sites, dolphin detections typically had very low signal-to-noise ratios (SNR) compared to the river channel detections. This led to an increase in click misclassifications with the rain sound-type. As the final classifier has a higher performance for rain than clicks (see Results), the rain class was used as a posterior filter for misclassifications from all recording sites (e.g. any 5-s segment that contained a value above threshold both for click and rain sound-types was attributed to rain).

Co-occurrence of dolphins and boat. To investigate the spatio-temporal overlap of dolphins and boats, a co-occurrence count was calculated at Boca, the location with the highest acoustic occurrence of dolphins. This count corresponds to the number of minutes that contain both positive click classification and positive boat classification.

Results

CNN Classifier performance. The initial and the final classifier performance were assessed based on the precision recall curve for the 3 impulsive sound types: Echolocation clicks from river dolphins, engine noise from passing boats and rain.

The initial classifier had an Q75 Average Precision of 0.90. After the active learning procedure, tests with reduced segment length, and addition of the *rain* class, the final classifier had an Average Precision of 0.95. This corresponds to an overall increase of 5.5% on the Average Precision after the active learning procedure. Highest Average Precision was achieved from the *rain* class with 0.98. *Click* and *boat* classes also showed high performance with respective values of 0.95 and 0.92 (Fig. 2). Positive classification threshold for echolocation clicks was set to values shown in Table 3, selected from the semi-automated performance evaluation performed for each recording location. Table 3 also shows the corresponding TPR and TNR for a given threshold at a given location. *Click* TPR and TNR were between 0.88 and 1, while these values ranged from 0.94 to 1 for the *boat* class.

Acoustic presence of river dolphins. Acoustic detection of dolphins was frequent at the two sites located in permanent bodies of water (Fig. 3). At Boca (the *ressaca* location), dolphin presence based on echolocation clicks was detected over the full deployment period, from November 2019 to May 2020. Dolphin presence increased from November to the beginning of January 2020, when detections peaked briefly in mid-January (dolphin presence detected approximately 70% of the daily time). Detections peaked similarly at the end of February and at the beginnings of April and May (Fig. 3, middle). At the Research lodge (the *river channel* location), there was a very low level of detections in the beginning of July and from September to November Detections

Class	Recording location	Class threshold	Rain filter threshold	TPR	TNR
Click	Rato	0.85	0.90	0.96	0.91
	Aracazinho	0.85	0.99	1	0.88
	Juruazinho	0.85	0.99	1	0.88
	Boca	0.85	0.99	1	1
	Research lodge	0.85	0.99	0.96	0.94
Boat	Rato	0.98	NA	1	1
	Aracazinho	0.98	NA	1	1
	Juruazinho	0.98	NA	1	0.97
	Boca	0.97	NA	0.94	0.97
	Research lodge	0.98	NA	1	0.97

Table 3. Classification threshold selected for echolocation click and boat at each recording location, and associated TPR and TNR assessed on 60 randomly selected classified segments (see Methods).

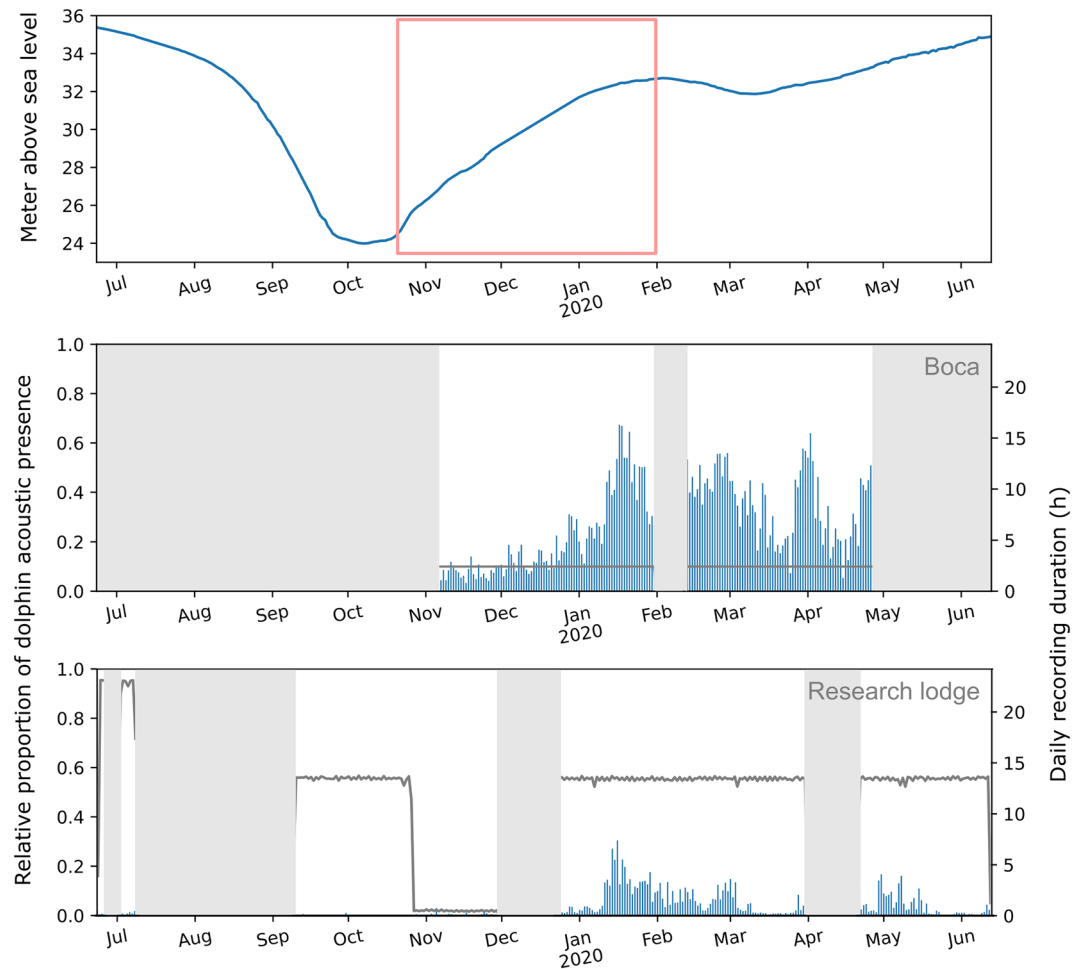


Figure 3. River dolphin acoustic presence at the *ressaca* and in the river channel (blue bars), based on CNN click classification. Water levels (top); Presence at the entrance of the river channel (middle); Presence at the Research lodge (bottom). The pink box indicates the rising water period. The grey line is the daily recording duration. Grey areas indicate an absence of recordings.

then started to increase until peak presence in Mid-January when dolphins were detected approximately 30% of the time. Between March and April and at the end of May, detections decreased to minimum values (Fig. 3, bottom).

The annual cycle of flooding between June 2019 and June 2020 (Fig. 3, top) showed the typical pattern of fluctuations between high and low water levels. From June to October, water levels decreased until reaching a minimum. This corresponded to a null or very low number of clicks detected at the Research lodge in the river channel. Later, at both sites (*ressaca* and *river channel*), dolphin presence increased over a 3-month period from November to January, with a clear synchronised peak in mid-January at the two locations. This time of the year corresponds to the rising water period of the annual cycle of flooding. In 2019, water levels rose 8 m from October to reach their highest level in mid-January (Fig. 3, top). Finally, during the following month (February to end of May), water levels remained high, and dolphins were detected regularly at both sites with similar variations in the rate of detection (matching peaks in March, April and May).

River dolphin acoustic presence was lower in the periodically flooded sites than at the permanent bodies of water. At *Rato*, one of the flooded forest sites, dolphins were mainly present in February–March and in May, during 64% of the sampling days, with a maximum value of 1.7% of the daily time. The recording period corresponds to the end of the rising water period (Fig. 4, top). At *Juruazinho*, the other flooded forest site, during the 2 and half months of recording available, detection rates were very low and dolphin presence was detected essentially during 8 days (12% of the sampling days) between the end of July and the beginning of August. The detections were of the same order of magnitude at *Aracazinho*, the internal lake location (less than 0.6% of the time, 9 days of detections). In terms of water level, this period corresponds to the receding water, where water levels in early July began to drop to reach approximately 30 m above sea level in early September.

Spatio-temporal co-occurrence of river dolphins and boats. The dolphin activity varied strongly by month, but when they were present there did not seem to be a strong difference between day or night activity (Fig. 5, left). A Wilcoxon rank-sum test was used for each month to compare the distributions of the number of

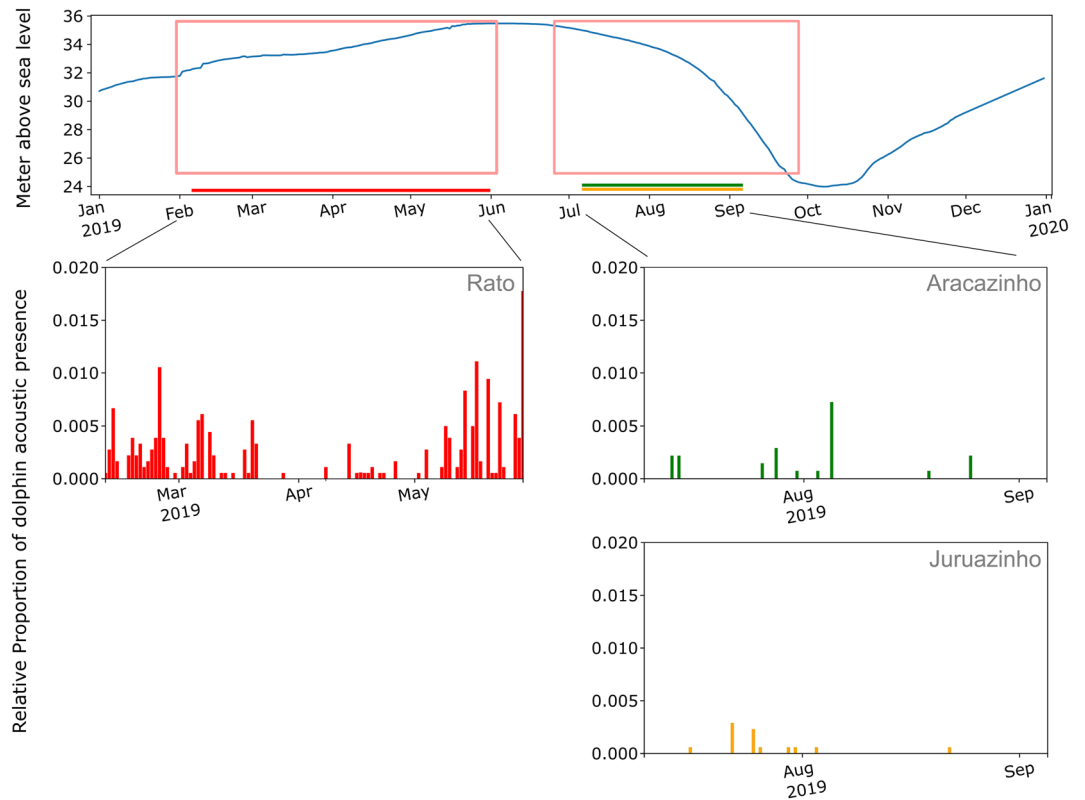


Figure 4. River dolphin acoustic presence at three periodically flooded sites, based on CNN click classification. Water levels (top); Dolphin presence at Rato (middle left); Presence at Aracazinho (middle right). Presence at Juruazinho (bottom right). The pink boxes indicate the high water period (left box) and the receding water period (right box).

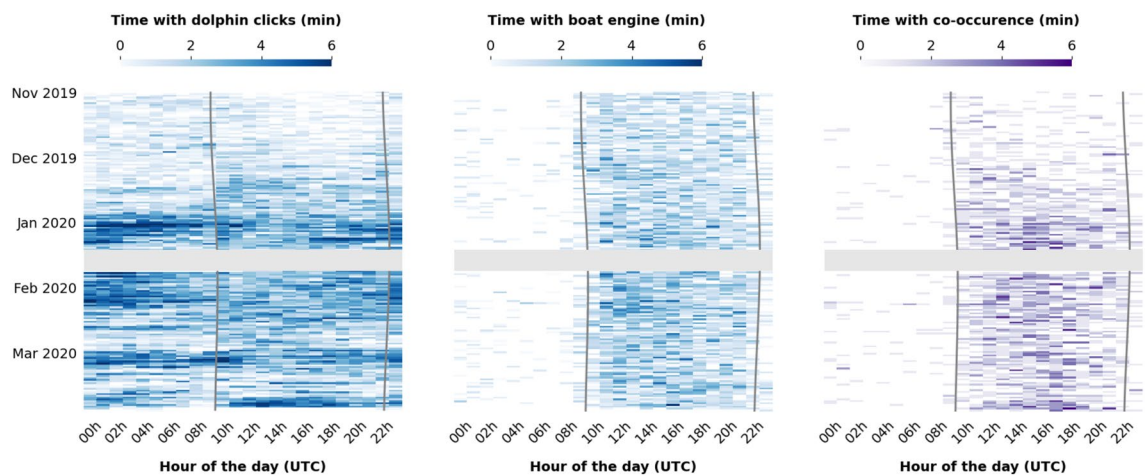


Figure 5. Acoustic presence of dolphin, boat, and co-occurrence grouped by hour of the day at the Boca site. The dark grey lines represent dusk and dawn. The light grey areas represent a period without data collection (battery replacement). Please note the values of 6 min per hour is the maximum value that can be obtained based on the data collection duty cycle (1 min on, 9 min off, see the Methods section).

positively classified segments between day (dawn–dusk) and night (dusk–dawn), both scaled by the time period being measured in minutes as the duration of day and night changed over time (H0: the day/night activity is the same; H1: one time of the day has higher activity than the other). A non-parametric test was selected because the distribution shapes between day/night appeared to be different and not normal. Only February showed a significant difference between day/night ($p = 0.00$ with $N = 18$ as there was missing data at the start of the month).

But for the other months such a difference was not found. Boat passages were almost exclusively detected during daylight hours without noticeable difference between the hours of the day or between months (Fig. 5, middle).

The temporal overlap between boats and dolphins was estimated through the co-occurrence of segments with positive clicks and boat classification within a 1-min time period. Figure 5 shows relatively low values in terms of time with co-occurrence. Over the full 5-month dataset, the number of minutes with co-occurrence is low (mean 0.54, SD = 0.94), representing an average of 9% of the recorded time. These values remain low when computed over the 3-month period corresponding to the higher dolphin presence in January–March (mean = 0.67, SD = 1.05) with 11% of the recorded time.

Discussion

The results of the CNN automatic classification of river dolphin echolocation clicks revealed patterns of presence in relation to the period of the annual water flood. In the bay and in the river channel, dolphin acoustic presence clearly increased during the period of rising waters, from November to January. This pattern was especially conspicuous in the bay (entrance of the river channel), where the daily acoustic presence rose from approximately 10% to 70% during this period. Interestingly, the main peak of acoustic presence was detected at both sites simultaneously. The synchronised detection peak at both locations suggests local population scale movements of dolphins entering the várzea from the main river through the river channel. These results are in agreement with published data on dolphin movements in the várzea in relation to the flooding cycle. Martin and da Silva²¹ reported the movements of the dolphins inside the RDSM through the river channel during rising waters based on visual surveys and radio-tracking data, with a rapidly increasing presence of dolphins peaking “at about a level of 10 m”. Our results show a similar pattern of detections, with a peak at rising water, when water levels reached 8 m above the lowest level. At high water levels (from February to June), dolphins remained present in the bay, where they were acoustically detected between 20 and 60% of the time. Dolphins were also regularly detected in the river channel, although not continuously, and with a considerably lower acoustic presence (0–15% of the time).

These findings further support the idea that bays formed at confluence areas are an important habitat for river dolphins. Mintzer et al.⁵ studied the seasonal movements of botos inside the RDSM estimating a transition probability between habitats, and characterised the entrance of the Mamirauá lake system as a core area for botos (i.e. where animals spend a maximum amount of time). Especially mother/calf pairs and immatures seemed to spend more time in the bays before moving back into the Mamirauá lake system at low/rising water. This preference was also demonstrated through a PAM study from the same location at the end of rising waters⁴³. Tucuxis also seems to favour confluences^{29,71}. The importance of this habitat appears to be shared by Amazonian River dolphin populations and subpopulations across their distribution^{33,34}. A recent study covering several locations in both the Orinoco and the Amazon basin highlighted that the highest dolphin densities for both Amazonian River dolphin species were found in the confluence areas, with densities averaging 23 and 16 ind./km² for botos and tucuxis respectively, and reaching 61 and 64 ind./km² in the confluences of the Mamirauá Reserve⁷².

Additionally, river channels appeared to be used by dolphins, especially botos, as a gateway to access remote parts of the várzea. Results from past PAM study in the RDSM investigating dolphin click trains and trajectories showed that the animals mainly used the Mamirauá channel as a passage to other locations of the várzea⁴³. From this channel, botos could access either permanent lakes connected (e.g. Mamirauá lake) or disconnected from the riverine system at low water (e.g. Rato lake) and seasonally flooded lakes (e.g. Juruazinho and Aracazinho lakes). Tucuxis are also known to be present in the channel, although limiting their use of the channel to the lower part, closer to the main river²⁹.

Thus, the difference between detection values at the two sites (situated 10 km apart on the same river channel) could be explained not only by a difference in the number of dolphins present but also by a difference in their habitat use. The entrance of the channel is a bay (*ressaca*), close to a confluence of two major waterways, the Solimões (upper part of the Amazon) and the Japurá rivers. The local environmental conditions create favourable low-current prey-rich habitats for the dolphins. Higher acoustic activity could reflect either an increase in time that individual dolphins spent in the area, an increase in the number of dolphins using this area, but also an increase in the acoustic activity due to the higher click production used for foraging compared to travelling behaviours^{47,73,74}.

Dolphin detections at high water in the flooded forest (Rato) were very low in terms of duration (less than 2% of the daily time) but regular in terms of presence (i.e. number of days). The Rato flooded forest site is an access to the Rato lake and dolphin detections likely reflect the regular passage of botos to the remote parts of the floodplains. Várzea lakes, especially the ones with floating vegetation that provide refuge for a great variety and abundance of fish, are also a favourite habitat of river dolphins²¹. Nevertheless, due to the difficulties to penetrate the intricate flooded forest ecosystem with boats, there is very little information on the distribution of dolphins in the mosaic of floodplain habitats. Even using alternative monitoring techniques, such as tracking animals through tags, data on dolphin distribution once they leave the river channel is excessively difficult to collect. In a study using VHF transmitters on 24 botos, Martin and da Silva²¹ reported that during high waters the tagged botos were out of range up to 100% of the time, preventing their localization inside the area. Our results indicate here the regular use of flooded forest passages connecting várzea channels to internal lakes by the dolphins during high waters.

Females botos with calves and immatures animals spend more time in the várzea habitats than males²¹. One of the reasons is that várzea habitats provide access to rich prey resources. Floodplain systems, which combine high levels of habitat complexity with nutrient-rich waters, host a great diversity of fish associated with high biomass^{17,75,76}. Another hypothesis is that várzea habitats seem to grant shelter against males' aggressive behaviour, especially towards calves^{77,78}. Finally, habitats such as internal lakes, flooded forests and small channels

provide resting areas with lower currents^{5,21} that are usually favoured by river dolphins. This unique and beneficial combination of environmental conditions make várzea habitats of major importance for females with calves and immatures, and therefore for boto populations survival.

From July to September, at the end of high water and at falling waters, our study collected data on two sites: one inside the flooded forest (Juruazinho), and one in an internal lake (Aracazinho). The detection levels were very low and of the same order of magnitude at both sites, with 12% of the days with dolphin detections, mostly occurring in July and early August. Unfortunately, no data could be collected during the same time period at the other study sites, and it was not possible to draw strong conclusions about the relationship between the limited presence of dolphins and the decrease of water levels. Nevertheless, our results are aligned with known dolphin movements outside the várzea. Both river dolphin species are known to move back into the main rivers during this time period, following fish movements outside the floodplains, and anticipating the upcoming risk of entrapment^{5,21,29,79}.

Quantifying the extent of spatiotemporal co-occurrence between dolphins and boats in freshwater environments is of critical importance, especially in core areas where the animals spend a great amount of their time budget in vital activities such as foraging. Passage of boats in these spatially restricted areas introduces engine cavitation noise in the underwater environment. Chronic stress, masking effects, behavioural and acoustic responses were reported for marine populations in the open ocean⁸⁰ and it can be assumed that such responses are even stronger in freshwater systems of (e.g. rivers, channels, bays) that are spatially restricted and provide less opportunities for animals to evade disturbance. Our results show that at Boca, where continuous and significant dolphin presence was detected during the 5-month recording period, the level of co-occurrence was approximately 10% of the recorded time. Nevertheless, the data was collected on a 10% duty cycle, and can only provide an estimation of how much boat traffic overlaps with dolphin presence. It is likely though that at this location the level of disturbance is reduced due to the low level of boat traffic (essentially fishermen from local communities and tourism boats that pass by a few times per week). However, effects of underwater noise on river dolphin populations are remarkably understudied⁸⁰ and so far only a handful of publications have addressed the effects of shipping traffic on freshwater cetacean populations^{81–83}. Therefore, the effects of noise exposure representing 10% of the time spent by river dolphins in core areas can not be evaluated. Continuous recordings are needed to accurately assess the overlap in boat and dolphin presence and evaluate the potential disturbance caused by this source of underwater noise. Further studies will be conducted at other confluence habitats regularly used by the river dolphins, where boat traffic is higher (e.g. Lake Tefé).

The CNN method developed here classified three sound types (echolocation clicks from river dolphins, boat noise from engine cavitation, and rain), with a mean Average Precision of 0.95. The sounds are represented as spectrogram images and this demonstrates the validity of using an image-based approach for classifying and discriminating underwater acoustic events of impulsive nature. Once labels were created (a time-intensive task), our CNN-based workflow required only a few pre-processing steps. Furthermore, with the integration of on-the-fly data augmentation, this workflow allowed to train initial models for undersampled classes. These models were used to automatically retrieve additional true positive (sound type of interest) and false positive (misclassifications) examples in an active learning loop in order to swiftly strengthen the model performance.

There is an increasing interest in using convolutional neural networks to automatically detect and classify odontocete echolocation clicks. Some studies focus on classifying single clicks in a single class classifier and achieve high performance^{60,84}. However, compared to monospecific recordings (i.e. containing one sound class), classification tasks using soundscape recordings (i.e. containing multiple sound classes) represent an important challenge for classification algorithms. Recent classification contests demonstrated that the performance scores achieved on soundscape recordings were 4 times smaller than on monospecific recordings⁸⁵. We chose to base our workflow on 5-s soundscape recordings with the objective of furthermore developing a general classification model for different types of sounds (impulsive or tonal) from biological (Amazon aquatic species), anthropogenic (e.g. boat) and natural (e.g. rain) sources.

Conclusion

This study demonstrates the suitability of using CNN-based classification to automatically detect river dolphin echolocation clicks in the complex soundscape of freshwater habitats. The efficiency and speed of the CNN method allow to analyse the totality of the data collected without having to subsample as usually done for manual analysis, making it possible to detect major movements of dolphins in the study area, and rare passages in specific habitats or seasons. The use of Passive Acoustic Monitoring coupled with automatic analytical methods such as CNN-based classification of dolphin signals can efficiently increase our knowledge on endangered dolphin populations across a range of flooded habitats, especially in remote and understudied habitats of flooded forest and lakes, and allows to precisely time the movements of river dolphins between várzea habitats in relation to the flooding pulse. The classifier in this study was extended to include automatic detection of boat passages in dolphin core areas to assess the extent of underwater noise disturbance on river dolphins.

Our study calls for a generalisation of the use of PAM inside the mosaic of floodplain habitats to understand habitat preferences and requirements of river dolphins, especially the boto females and calves. Practical applications in forecasting the dolphins' response to habitat loss and degradation (e.g. deforestation for pastures, plantations, selective logging, ...) will contribute to the management strategies of the aquatic-terrestrial transition zone (ATTZ), critical for the maintenance of habitat connectivity^{16,86}. Another area of applications is towards developing and implementing standardised protocols to monitor distribution shifts in relation to the recent amplification of drought and flood events in the Amazon basin⁸⁷. As sentinel species of the aquatic systems they inhabit, river dolphins can constitute an early detection system of ecosystem unbalance²⁶.

Data availability

The datasets generated during the current study are available from the corresponding author on reasonable request.

Received: 3 December 2022; Accepted: 5 June 2023

Published online: 27 July 2023

References

- da Silva, V. *et al.* *Inia geoffrensis*, Amazon River dolphin. in *IUCN Red List Threat. Species 2018 ET10831A50358152* <https://doi.org/10.2305/IUCN.UK.2018-2.RLTS.T10831A50358152.en>. (2018).
- da Silva, V., Martin, A. R., Fettuccia, D. de C., Bivaqua, L. & Trujillo, F. *Sotalia fluviatilis*. in *IUCN Red List Threat. Species 2020 ET190871A50386457* <https://doi.org/10.2305/IUCN.UK.2020-3.RLTS.T190871A50386457.en>. (2020).
- Trujillo, F., Crespo, E., Van Damme, P. A. & Usma, J. S. Status and conservation of river dolphins *Inia geoffrensis* and *Sotalia fluviatilis* in the Amazon and Orinoco basins in Colombia. in *The action plan for South American River Dolphins 2010–2020. WWF, Fundación Omacha, WDS, WDCS, Solamac*. 29–57 (2010).
- Sinha, R. K. & Kannan, K. Ganges river dolphin: An overview of biology, ecology, and conservation Status in India. *Ambio* **43**, 1029–1046 (2014).
- Mintzer, V. J., Lorenzen, K., Frazer, T. K., da Silva, V. M. F. & Martin, A. R. Seasonal movements of river dolphins (*Inia geoffrensis*) in a protected Amazonian floodplain. *Mar. Mammal Sci.* **32**, 664–681 (2016).
- da Silva, V. M. F., Freitas, C. E. C., Dias, R. L. & Martin, A. R. Both cetaceans in the Brazilian Amazon show sustained, profound population declines over two decades. *PLoS One* **13**, e0191304 (2018).
- Dudgeon, D. *et al.* Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biol. Rev.* **81**, 163–182 (2006).
- Loch, C., Marmontel, M. & Simoes-Lopes, P. C. Conflicts with fisheries and intentional killing of freshwater dolphins (Cetacea: Odontoceti) in the Western Brazilian Amazon. *Biodivers. Conserv.* **18**, 3979–3988 (2009).
- da Silva, V. M. F., Martin, A. R. & Do Carmo, N. A. Boto bait: Amazonian fisheries pose threat to elusive dolphin species. *Species* **53**, 10–11 (2011).
- Alves, L. C. P. D. S., Zappes, C. A. & Andriolo, A. Conflicts between river dolphins (Cetacea: Odontoceti) and fisheries in the Central Amazon: a path toward tragedy?. *Zool. Curitiba* **29**, 420–429 (2012).
- Iriarte, V. River Dolphin (*Inia geoffrensis*, *Sotalia fluviatilis*) mortality events attributed to artisanal fisheries in the western Brazilian Amazon. *Aquat. Mamm.* **39**, 116–124 (2013).
- Brum, S. M., da Silva, V. M. F., Rossoni, F. & Castello, L. Use of dolphins and caimans as bait for *Calophrys macropterus* (Lichtenstein, 1819) (Siluriforme: Pimelodidae) in the Amazon. *J. Appl. Ichthyol.* **31**, 675–680 (2015).
- Mosquera-Guerra, F. *et al.* Strategy to identify areas of use of Amazon River dolphins. *Front. Mar. Sci.* **9**, 838988 (2022).
- Castello, L. & Macedo, M. N. Large-scale degradation of Amazonian freshwater ecosystems. *Glob. Change Biol.* **22**, 990–1007 (2016).
- Martin, A. R. & da Silva, V. M. F. Amazon river dolphins *Inia geoffrensis* are on the path to extinction in the heart of their range. *Oryx* **56**(4), 587–591. <https://doi.org/10.1017/S0030605320001350> (2021).
- Junk, W. J., Bayley, P. B. & Sparks, R. E. The flood pulse concept in river-floodplain systems. *Can. Spec. Publ. Fish. Aquat. Sci.* **106**, 110–127 (1989).
- Goulding, M. *The Fishes and the Forest: Explorations in Amazonian Natural History* (University of California Press, 1980). <https://doi.org/10.1525/9780520316133>.
- Duponchelle, F. *et al.* Conservation of migratory fishes in the Amazon basin. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **31**, 1087–1105 (2021).
- da Silva, V. M. F. in *Ecologia Alimentar dos Golfinhos da Amazônia* (FUA/INPA, 1983).
- Vidal, O. *et al.* Distribution and abundance of the Amazon river dolphin (*Inia geoffrensis*) and the tucuxi (*Sotalia fluviatilis*) in the upper Amazon river. *Mar. Mammal Sci.* **13**, 427–445 (1997).
- Martin, A. R. & da Silva, V. M. F. River dolphins and flooded forest: seasonal habitat use and sexual segregation of botos (*Inia geoffrensis*) in an extreme cetacean environment. *J. Zool.* **263**, 295–305 (2004).
- Gomez-Salazar, C., Trujillo, F., Portocarrero-Aya, M. & Whitehead, H. Population, density estimates, and conservation of river dolphins (*Inia* and *Sotalia*) in the Amazon and Orinoco river basins. *Mar. Mammal Sci.* **28**, 124–153 (2012).
- Reis, R. E. in *Check List of the Freshwater Fishes of South and Central America* (Edipucrs, 2003).
- Martin, A. R., Silva, V. M. F. & Salmon, D. L. Riverine habitat preferences of botos (*Inia geoffrensis*) and tucuxis (*Sotalia fluviatilis*) in the central Amazon. *Mar. Mammal Sci.* **20**, 189–200 (2004).
- McGuire, T. L. & Winemiller, K. O. Occurrence patterns, habitat associations, and potential prey of the river dolphin, *Inia geoffrensis*, in the Cinaruco river Venezuela. *Biotropica* **30**, 625–638 (1998).
- Gomez-Salazar, C., Coll, M. & Whitehead, H. River dolphins as indicators of ecosystem degradation in large tropical rivers. *Ecol. Indic.* **23**, 19–26 (2012).
- Gutstein, C. S., Cozzuol, M. A. & Pyenson, N. D. The Antiquity of riverine adaptations in Iniidae (Cetacea, Odontoceti) documented by a humerus from the late Miocene of the Ituzaingó formation, Argentina: Antiquity of riverine adaptations in Iniidae. *Anat. Rec.* **297**, 1096–1102 (2014).
- Best, R. C. & da Silva, V. M. Amazon river dolphin, boto *Inia geoffrensis* (de Blainville, 1817). *Handb. Mar. Mamm.* **4**, 1–23 (1989).
- Faustino, C. & Da Silva, V. M. F. Seasonal use of Amazon floodplains by the tucuxi *Sotalia fluviatilis* (Gervais 1853), in the central Amazon, Brazil. *Lat. Am. J. Aquat. Mamm.* <https://doi.org/10.5597/lajam00100> (2006).
- Trujillo, F., Diazgranados, M. C., Galindo, A. L. F. & Fuentes, L. Delfin gris *Sotalia fluviatilis*. in *2006 Libro Rojo Los Mamíferos Colomb. Ser. Libr. Rojos Especies Amenazadas Colomb. Conserv. Int. Minist. Ambiente Vivienda Desarro. Territ. Bogota Colomb.* 273–278 (2006).
- Fürstenau Oliveira, J. S., Georgiadis, G., Campello, S., Brandão, R. A. & Ciuti, S. Improving river dolphin monitoring using aerial surveys. *Ecosphere* **8**, e01912 (2017).
- Oliveira-da-Costa, M. *et al.* Effectiveness of unmanned aerial vehicles to detect Amazon dolphins. *Oryx* **54**(5), 696–698. <https://doi.org/10.1017/S0030605319000279> (2019).
- Mosquera-Guerra, F. *et al.* Home range and movements of Amazon river dolphins *Inia geoffrensis* in the Amazon and Orinoco river basins. *Endanger. Species Res.* **45**, 269–282 (2021).
- Mosquera-Guerra, F. *et al.* Identifying habitat preferences and core areas of Amazon River dolphin activity using spatial ecology analysis. *Landsc. Ecol.* **37**, 2099–2119 (2022).
- Norman, S. A., Hobbs, R. C., Foster, J., Schroeder, J. P. & Townsend, F. I. A review of animal and human health concerns during capture-release, handling and tagging of odontocetes. *J. Cetacean Res. Manag.* **6**, 53–62 (2004).
- Burek-Huntington, K. A. *et al.* Postmortem pathology investigation of the wounds from invasive tagging in belugas (*Delphinapterus leucas*) from Cook Inlet and Bristol Bay, Alaska. *Mar. Mammal Sci.* **39**(2), 492–514. <https://doi.org/10.1111/mms.12981> (2022).

37. Martin, A. R., Da Silva, V. M. F. & Rothery, P. R. Does radio tagging affect the survival or reproduction of small cetaceans?. *A test. Mar. Mammal Sci.* **22**, 17–24 (2006).
38. Andrews, R. D. *et al.* Best practice guidelines for cetacean tagging. *J. Cetacean Res. Manag.* **20**, 27–66 (2019).
39. Ladegaard, M., Jensen, F. H., de Freitas, M., Ferreira da Silva, V. M. & Madsen, P. T. Amazon river dolphins (*Inia geoffrensis*) use a high-frequency short-range biosonar. *J. Exp. Biol.* **218**, 3091–3101 (2015).
40. Podos, J., Da Silva, V. M. & Rossi-Santos, M. R. Vocalizations of Amazon river dolphins, *Inia geoffrensis*: insights into the Evolutionary origins of delphinid whistles. *Ethology* **108**, 601–612 (2002).
41. Wang, D., Würsig, B. & Leatherwood, S. Whistles of boto, *Inia geoffrensis*, and tucuxi *Sotalia fluviatilis*. *J. Acoust. Soc. Am.* **109**, 407–411 (2001).
42. Zimmer, W. M. X. *Passive Acoustic Monitoring of Cetaceans* (Cambridge University Press, 2011). <https://doi.org/10.1017/CBO9780511977107>.
43. Yamamoto, Y., Akamatsu, T., da Silva, V. M. F. & Kohshima, S. Local habitat use by botos (Amazon river dolphins, *Inia geoffrensis*) using passive acoustic methods. *Mar. Mammal Sci.* **32**, 220–240 (2016).
44. Campbell, E. C., Alfaro Shigueto, J., Godley, B. & Mangel, J. Abundance estimate of the Amazon River dolphin (*Inia geoffrensis*) and the tucuxi (*Sotalia fluviatilis*) in southern Ucayali Peru. *Lat. Am. J. Aquat. Res.* **45**, 957–969 (2017).
45. Muirhead, C. A. Passive acoustic monitoring of river dolphin (*Inia geoffrensis* and *Sotalia fluviatilis*) presence: A comparison between waters near the city of Iquitos and within the Pacaya-Samiria National Reserve. *Lat. Am. J. Aquat. Mamm.* **16**, 3–11 (2021).
46. Melo, J. F., Amorim, T. & Andriolo, A. Pulsed sounds produced by amazon river dolphin (*Inia geoffrensis*) in the Brazilian Amazon: Comparison between two water turbidity conditions. *J. Acoust. Soc. Am.* **138**, 1792–1793 (2015).
47. Yamamoto, Y., Akamatsu, T., da Silva, V. M. F., Yoshida, Y. & Kohshima, S. Acoustic characteristics of biosonar sounds of free-ranging botos (*Inia geoffrensis*) and tucuxis (*Sotalia fluviatilis*) in the Negro River, Amazon, Brazil. (2015).
48. Ladegaard, M., Jensen, F. H., Beedholm, K., da Silva, V. M. F. & Madsen, P. T. Amazon river dolphins (*Inia geoffrensis*) modify biosonar output level and directivity during prey interception in the wild. *J. Exp. Biol.* **220**, 2654–2665 (2017).
49. Melo-Santos, G. *et al.* The newly described Araguaian river dolphins, *Inia araguaiaensis* (Cetartiodactyla, Iniidae), produce a diverse repertoire of acoustic signals. *PeerJ* **7**, e6670 (2019).
50. Melo-Santos, G., Walmsley, S. F., Marmontel, M., Oliveira-da-Costa, M. & Janik, V. M. Repeated downsweep vocalizations of the Araguaian river dolphin *Inia araguaiaensis*. *J. Acoust. Soc. Am.* **147**, 748–756 (2020).
51. Melo, J. F. The biosonar of the boto: Evidence of differences among species of river dolphins (*Inia* spp.) from the Amazon. *PeerJ* **9**, e11105 (2021).
52. André, M. *et al.* Listening to the deep: Live monitoring of ocean noise and cetacean acoustic signals. *Mar. Pollut. Bull.* **63**, 18–26 (2011).
53. Jarvis, S., Dimarzio, N., Morrissey, R. & Moretti, D. A novel multi-class support vector machine classifier for automated classification of beaked whales and other small odontocetes. *Can. Acoust.* **36**(1), 34–40 (2008).
54. Liu, J., Yang, X., Wang, C. & Tao, Y. A convolution neural network for dolphin species identification using echolocation clicks signal. in *2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)* 1–4 (IEEE, 2018). <https://doi.org/10.1109/ICSPCC.2018.8567796>.
55. Roch, M. A., Soldevilla, M. S., Burtenshaw, J. C., Henderson, E. E. & Hildebrand, J. A. Gaussian mixture model classification of odontocetes in the Southern California Bight and the Gulf of California. *J. Acoust. Soc. Am.* **121**, 1737–1748 (2007).
56. Siddagangaiah, S. *et al.* Automatic detection of dolphin whistles and clicks based on entropy approach. *Ecol. Indic.* **117**, 106559 (2020).
57. Van der Schaar, M., Delory, E. & André, M. Classification of sperm whale clicks (*Physeter macrocephalus*) with Gaussian-kernel-based networks. *Algorithms* **2**, 1232–1247 (2009).
58. Zaugg, S., van der Schaar, M., Houégnigan, L., Gervaise, C. & André, M. Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories. *Appl. Acoust.* **71**, 1011–1019 (2010).
59. Bermant, P. C., Bronstein, M. M., Wood, R. J., Gero, S. & Gruber, D. F. Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. *Sci. Rep.* **9**, 12588 (2019).
60. Luo, W., Yang, W. & Zhang, Y. Convolutional neural network for detecting odontocete echolocation clicks. *J. Acoust. Soc. Am.* **145**(1), EL7–EL12. <https://doi.org/10.1121/1.5085647> (2019).
61. Ziegenhorn, M. A. *et al.* Discriminating and classifying odontocete echolocation clicks in the Hawaiian Islands using machine learning methods. *PLoS One* **17**, e0266424 (2022).
62. Ramalho, E. E. *et al.* Ciclo hidrológico nos ambientes de várzea da Reserva de Desenvolvimento Sustentável Mamirauá - Médio rio Solimões, Período de 1990 a 2008. **25** (2009).
63. Junk, W. J., Piedade, M. T. F., Schöngart, J. & Wittmann, F. A classification of major natural habitats of Amazonian white-water river floodplains (várzeas). *Wetl. Ecol. Manag.* **20**, 461–475 (2012).
64. Ferreira-Ferreira, J. *et al.* Combining ALOS/PALSAR derived vegetation structure and inundation patterns to characterize major vegetation types in the Mamirauá Sustainable development reserve, Central Amazon floodplain Brazil. *Wetl. Ecol. Manag.* **23**, 41–59 (2015).
65. Takahashi, N., Gygli, M., Pfister, B. & Van Gool, L. Deep Convolutional Neural Networks and Data Augmentation for Acoustic Event Detection. Preprint at <http://arxiv.org/abs/1604.07160> (2016).
66. Salamon, J. & Bello, J. P. Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Process. Lett.* **24**, 279–283 (2017).
67. Hershey, S. *et al.* CNN architectures for large-scale audio classification. in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 131–135 (2017).
68. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. in (arXiv, 2017).
69. Branco, P., Torgo, L. & Ribeiro, R. A survey of predictive modelling under imbalanced distributions. *ACM Comput. Surv. CSUR* **49**, 1–50 (2015).
70. Saito, T. & Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* **10**, e0118432 (2015).
71. Magnusson, W. E., Best, R. C. & Da Silva, V. M. F. Numbers and behaviour of Amazonian dolphins, *Inia geoffrensis* and *Sotalia fluviatilis fluviatilis*, in the rio Solimões. *Brasil. Aquat. Mamm.* **8**, 27–32 (1980).
72. Paschoalini, M. *et al.* Density and abundance estimation of Amazonian river dolphins: Understanding population size variability. *J. Mar. Sci. Eng.* **9**, 1184 (2021).
73. Jones, G. J. & Sayigh, L. S. Geographic variation in rates of vocal production of free-ranging bottlenose dolphins. *Mar. Mammal Sci.* **18**, 374–393 (2002).
74. Nowacek, D. P. Acoustic ecology of foraging bottlenose dolphins (*Tursiops truncatus*), habitat-specific use of three sound types. *Mar. Mammal Sci.* **21**, 587–602 (2005).
75. Junk, W. J. (ed.) *The Central Amazon Floodplain* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1997).
76. Crampton, W. G. R., Castello, L. & Viana, J. P. 6. Fisheries in the Amazon Várzea: Historical trends, current status, and factors affecting sustainability. In *People in Nature: Wildlife Conservation in South and Central America* (eds Silvius, K. M. *et al.*) 76–98 (Columbia University Press, 2004). <https://doi.org/10.7312/silv12782-006>.

77. Martin, A. R. & Da Silva, V. M. F. Sexual dimorphism and body scarring in the Boto (Amazon river dolphin) *Inia geoffrensis*. *Mar. Mammal Sci.* **22**, 25–33 (2006).
78. da Silva, V. M. F. *et al.* Aggression towards neonates and possible infanticide in the boto, or Amazon river dolphin (*Inia geoffrensis*). *Behaviour* **158**, 971–984 (2021).
79. Trebbau, P. & Van Bree, P. J. H. Notes concerning the freshwater dolphin *Inia geoffrensis* (de Blainville, 1817) in Venezuela. *Z Säugetierkd.* **39**, 50–57 (1974).
80. Erbe, C. *et al.* The effects of ship noise on marine mammals - A review. *Front. Mar. Sci.* **6**, 606 (2019).
81. Kreb, D. & Rahadi, K. D. Living under an aquatic freeway: Effects of boats on Irrawaddy dolphins (*Orcaella brevirostris*) in a coastal and riverine environment in Indonesia. *Aquat. Mamm.* **30**, 363–375 (2004).
82. Nabi, G., Hao, Y., McLaughlin, R. W. & Wang, D. The possible effects of high vessel traffic on the physiological parameters of the critically endangered Yangtze finless porpoise (*Neophocaena asiaorientalis* ssp. *asiaorientalis*). *Front. Physiol.* **9**, 11 (2018).
83. Dey, M., Krishnaswamy, J., Morisaka, T. & Kelkar, N. Interacting effects of vessel noise and shallow river depth elevate metabolic stress in Ganges river dolphins. *Sci. Rep.* **9**, 15426 (2019).
84. Buchanan, C. *et al.* Deep convolutional neural networks for detecting dolphin echolocation clicks. in *2021 36th International Conference on Image and Vision Computing New Zealand (IVCNZ)* 1–6 (IEEE, 2021). <https://doi.org/10.1109/IVCNZ54163.2021.9653250>.
85. Goeau, H. *et al.* Overview of BirdCLEF 2018: Monospecies versus soundscape bird identification. *Work. Notes CLEF 2018-Conf. Labs Eval. Forum* **2125**, 13 (2018).
86. Hurd, L. E. *et al.* Amazon floodplain fish communities: Habitat connectivity and conservation in a rapidly deteriorating environment. *Biol. Conserv.* **195**, 118–127 (2016).
87. Bodmer, R. *et al.* Major shifts in Amazon wildlife populations from recent intensification of floods and drought. *Conserv. Biol.* **32**, 333–344 (2018).

Acknowledgements

The authors are grateful to the IDSM field team and to the field guides Izael Miranda and Alcibdes 'Bide' Martins for their help during deployment and maintenance of the equipment. This work was funded in part by a generous grant from the National Marine Mammal Foundation.

Author contributions

F.E.: data collection, study design, data analysis, manuscript writing. M.G.: data collection. M.V.D.S.: study design, data analysis. S.Z.: classification framework. E.R.: data collection. D.H.: data collection, funding. M.A.: funding, study design, data collection, manuscript writing. All the authors contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-36518-1>.

Correspondence and requests for materials should be addressed to M.A.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023