Manual

Automated acoustic classifiers for northeast Atlantic delphinids

Contents

Manual prepared by T. Kleyn

tristankleyn@gmail.com – github.com/tristankleyn/which.dolphin

Contributors: T. Kleyn, V. Janik, J. Macaulay, and J. Oswald

OVERVIEW

This manual provides guidance specific to using acoustic classifier models compatible with the open-source software PAMGuard [1], along with the eventClassifier graphical user interface, to automatically identify northeast Atlantic delphinid species in passive acoustic recordings. The method outlined consists of two stages of acoustic classification: base classification, where individual vocalizations or segments containing multiple vocalizations detected in PAMGuard are classified to species, and event classification, where a vector containing the outputs of the base classifiers is used to classify events containing multiple vocalizations or multiple segments of vocalizations to species. This workflow was motivated by the demonstrated success of the *BANTER* classification method [2], which has been used to classify whistles and clicks of Pacific delphinids [2] and clicks of arctic odontocetes [3] and beaked whale species [4]. In developing delphinID, we supplement this two-stage classification workflow with deep learning, which allows a higher level of model complexity and learning of latent features in the acoustic characteristics of delphinid vocalizations. ROCCA (Real-time Odontocete Call Classification Algorithm), developed by Oswald & Oswald [5], offers an alternative classification method based on selected acoustic characteristics and Random Forest analysis. The methods outlined in the following sections concern acoustic classification using whistles and echolocation clicks, two distinct types of vocalizations commonly produced by delphinids. Event classification, implemented intuitively through the eventClassifier interface, integrates information from both whistles and clicks to inform species prediction at the event level – cross-validated testing showed this level of classification to classify the seven Atlantic species listed below with up to 86% accuracy:

- Atlantic white-sided dolphin *(Lagenorhynchus acutus)*
- Common bottlenose dolphin *(Tursiops truncatus)*
- Killer whale *(Orcinus orca)*
- Long-finned pilot whale *(Globicephala melas)*
- Risso's dolphin *(Grampus griseus)*
- Short-beaked common dolphin *(Delphinus delphis)*
- White-beaked dolphin *(Lagenorhynchus albirostris)*

TERMINOLOGY

The following terminology is used throughout this manual.

Term	Explanation
Accuracy	The proportion of correct (out of total) predictions made by a classifier.
Bandwidth	A frequency range
Base classification	Classification of individual vocalizations or short segments containing multiple vocalizations
Broadband	Containing a wide range of frequencies
Click	Broadband transient signal used by dolphins for echolocation
Detection frame	The unit for base classification with delphinID: a 1D representation of the relative concentration of acoustic energy in 100 Hz bands ranging from 2-20 kHz for whistles or 10-40 kHz for clicks.
Event/encounter	Used interchangeably, events or encounters refer to periods containing multiple detections of whistles and/or clicks
Event classification	Classification of events, or the average classification result across all base classifications with an event
Narrowband	Containing a narrow range of frequencies
Nyquist frequency	Half the sampling rate – the highest frequency that can be captured
Sampling rate	The number of samples taken from an analog acoustic signal per second to convert it into a discrete digital signal.
Spectrogram	A 2D visual representation of the frequencies within a signal as they vary over time.
Whistle	Narrowband tonal vocalization produced by dolphins for social signalling.

BASE CLASSIFICATION IN PAMGUARD

To use our classifier models to identify northeast Atlantic delphinid species, it is first necessary to detect whistle and click vocalizations in your passive acoustic data. This can be done using the open-source acoustic analysis software PAMGuard [1], which can be freely downloaded at PAMGuard.org. For the best possible user experience, please download the latest version of PAMGuard. An installer for PAMGuard 2.02.15.e (April 2025 release) is provided in the master folder.

With delphinID

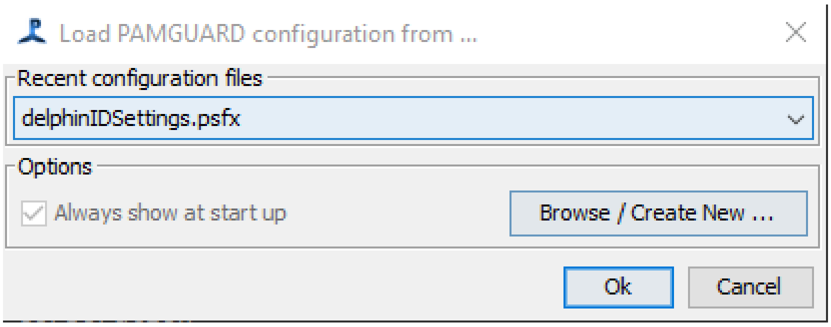
delphinID classifier models are convolutional neural networks trained on 1D acoustic representations of thousands of visually-verified examples of whistle and click vocalizations to classify seven northeast Atlantic delphinid species:

- Atlantic white-sided dolphin *(Lagenorhynchus acutus)*
- Common bottlenose dolphin *(Tursiops truncatus)*
- Killer whale *(Orcinus orca)*
- Long-finned pilot whale *(Globicephala melas)*
- Risso's dolphin *(Grampus griseus)*
- Short-beaked common dolphin *(Delphinus delphis)*
- White-beaked dolphin *(Lagenorhynchus albirostris)*

These 1D representations, herein referred to as detection frames, are vectors calculated from detections made in PAMGuard representing the relative concentration of acoustic energy in 100 Hz frequency bands ranging from 2-20 kHz for whistles and 10-40 kHz for clicks. For clicks, this vector is a normalised frequency power spectrum calculated click detections fed from the PAMGuard Click Detector module, while for whistles it is calculated as a normalised histogram of frequency values contained in detections of whistle peak frequency contours traced by PAMGuard's Whistle & Moan Detector. The following steps are required for classifying detection frames of whistles and clicks using the delphinID classifiers for northeast Atlantic delphinids:

Tutorial 1 | base classification with delphinID

1. Download PAMGuard (version 2.02.15 or later).
2. Load the delphinIDsettings.psfx configuration in PAMGuard.

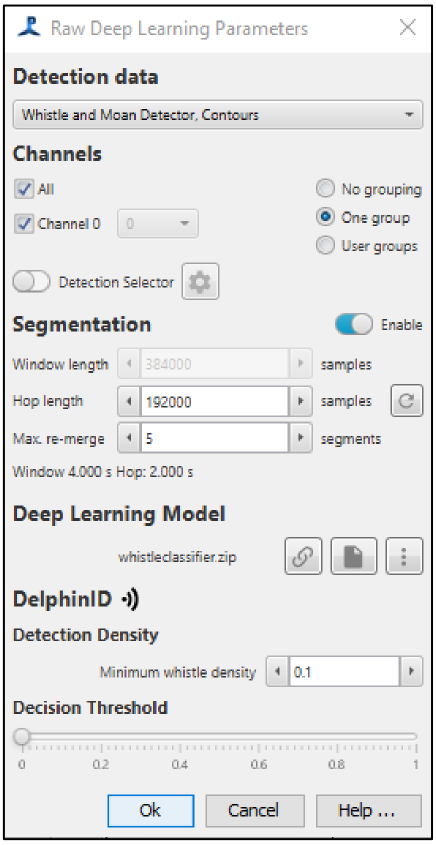


3. Go to **Settings → Sound Acquisition** and select sound file(s) for analysis. **Click Ok.**

4. Go to **Settings → Decimator.** Confirm that the Output sample rate is 96000 Hz. If your Source sample rate is not an integer multiple of 96000, change the **Interpolation** to **Quadratic. Click Ok.**

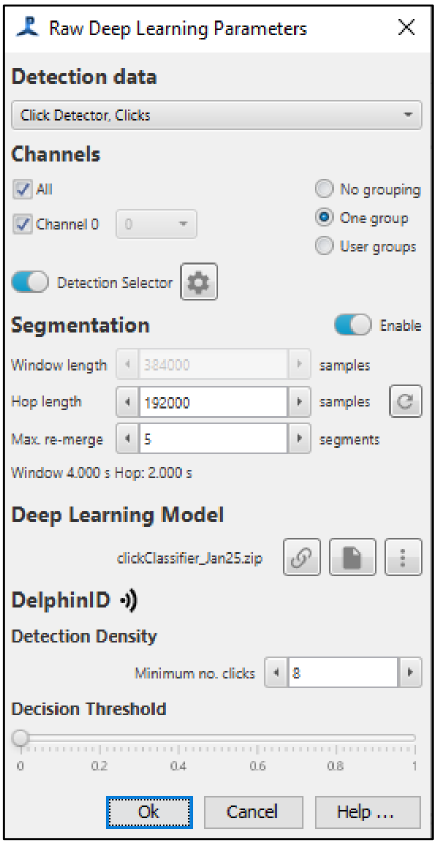
5. In the **Settings** for **Sound Output, Decimator, FFT (Spectrogram) Engine settings, Click Detector,** and **Whistle & Moan Detector,** make sure that the acoustic channel targeted for analysis is ticked in all cases.

6. Go to **Settings → Deep Learning Classifier – Whistles.** You should see a display similar to this:



Here, you must import the delphinID whistle classifier (**NEAtlanticWhistleClassifier.zip**). Copy the settings shown in the screenshot above apart from the selected channel, window length, and hop length, which will be specific to your analysis. Click the refresh button to the right side of the **Segmentation** sub-section to automatically calculate the window length and hop length for your file. **Click Ok.**

7. Go to **Settings → Deep Learning Classifier – Clicks.** You should see a display similar to this:

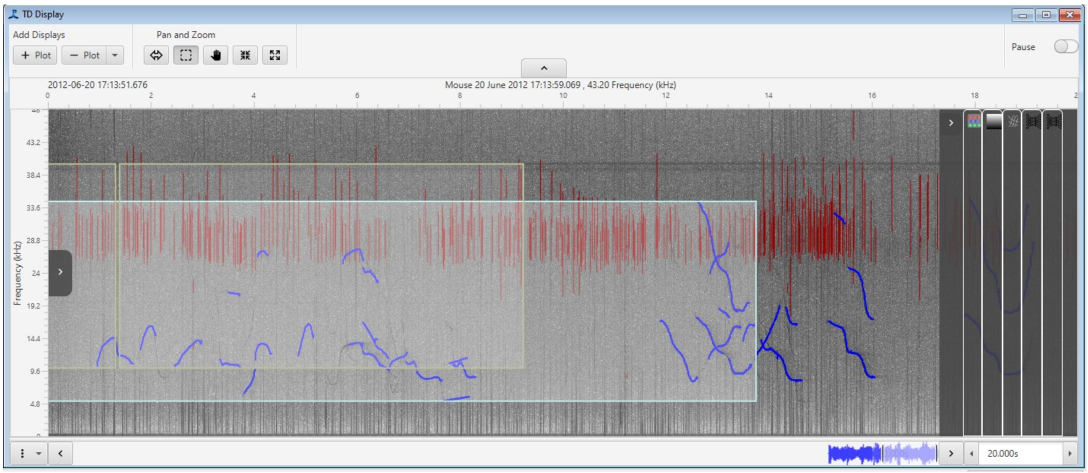


Same as for the whistle classifier. Under **Deep Learning Model**, import the delphinID click classifier (**NEAtlanticClickClassifier.zip**) and copy the above settings apart from channel, window length, and hop length. Refresh to update the window and hop length. **Click Ok.**

8. Before running the classifiers, we need to set up storage for saving results. Go to **File → Database → Database selection.** Select the file **trackDB.sqlite3** in the **eventClassifier** folder. This is necessary for event classification later on.

9. (Optional) To save binary files in addition to a database, go to **Add modules → Utilities → Binary Storage** and click **Ok.** Now go to **File → Binary storage** to select a folder where binaries will be stored.

10. We are now ready to run our delphinID classifiers! **Click the red button** near to the top of the PAMGuard display to run detection and classification. As the file is analysed, you should see a visual display similar to this:



The display can be interpreted as follows:

- Red lines show click detections
- Blue contours outline peak frequency contours of whistle detections
- Coloured boxes that appear indicate groups of detections classified by delphinID – the colour indicates the most likely species (species key available in display settings).

11. With the results from base classification in PAMGuard exported to your .sqlite3 database, you are now ready to use the eventClassifier interface to classify events containing whistles and/or clicks to species.

With ROCCA

ROCCA (Real-time Odontocete Call Classification Algorithm) classifier models are Random Forest models trained on select acoustic measurements from thousands of visually-verified examples of whistle and click vocalizations to classify seven northeast Atlantic delphinid species [6]:

- Atlantic white-sided dolphin *(Lagenorhynchus acutus)*
- Common bottlenose dolphin *(Tursiops truncatus)*
- Killer whale *(Orcinus orca)*
- Long-finned pilot whale *(Globicephala melas)*
- Risso’s dolphin *(Grampus griseus)*
- Short-beaked common dolphin *(Delphinus delphis)*
- White-beaked dolphin *(Lagenorhynchus albirostris)*

Fifty time-frequency measurements are used for whistles, describing characteristics pertaining to the frequency and frequency modulation contained in whistle contours, while thirteen click measurements capture the distribution of energy across frequency in clicks – these ROCCA parameters are extracted via the ROCCA module in PAMGuard. Full descriptions of these parameters are given in Appendix XXX, while more support can be found online here.

Tutorial 2 | base classification with ROCCA

1. Download PAMGuard (latest version).
2. Load the ROCCAsettings.psfx configuration in PAMGuard.
3. Go to **Settings → Sound Acquisition** and select sound file(s) for analysis. **Click Ok.**

4. Go to **Settings → Decimator.** Confirm that the Output sample rate is 96000 Hz. If your Source sample rate is not an integer multiple of 96000, change the **Interpolation** to **Quadratic. Click Ok.**

5. In the **Settings** for **Sound Output, Decimator, FFT (Spectrogram) Engine settings, Click Detector,** and **Whistle & Moan Detector,** make sure that the acoustic channel targeted for analysis is ticked in all cases.

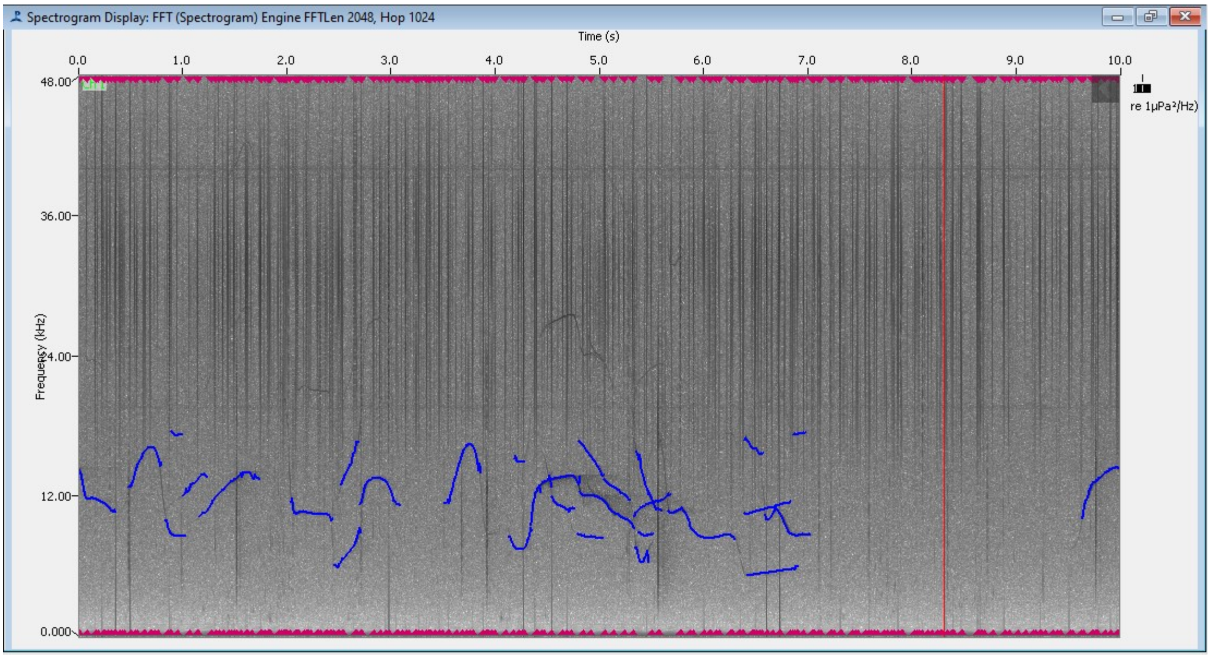
6. Go to **Settings → Rocca parameters → Contours/Classifier.** Import the ROCCAwhistleclassifier.model and ROCCAclickclassifier.model files for the **Whistle Classifier** and **Click Classifier,** respectively.

7. Go to **Settings → Rocca parameters → Output.** Select a folder to export ROCCA .csv data (This will not be used directly for event classification but can be useful for inspecting the results). **Click Ok.**

8. Before running the classifiers, we need to set up storage for saving results. Go to **File → Database → Database selection.** Select the file **trackDB.sqlite3** in the **eventClassifier** folder. This is necessary for event classification later on.

9. (Optional) To save binary files in addition to a database, go to **Add modules → Utilities → Binary Storage** and click **Ok.** Now go to **File → Binary storage** to select a folder where binaries will be stored.

10. We are now ready to run our ROCCA classifiers! **Click the red button** near to the top of the PAMGuard display to run detection and classification. As the file is analysed, you should see a visual display similar to this:



The display can be interpreted as follows:

- Red lines show click detections
- Blue contours outline peak frequency contours of whistle detections
- Counts of base classifications for individual detections are counted for each species in the lefthand sidebar display under **Rocca**.

12. With the results from base classification in PAMGuard exported to your .sqlite3 database, you are now ready to use the eventClassifier interface to classify events containing whistles and/or clicks to species.

Optional setting adjustments for delphinID or ROCCA

Several parts of the workflow for base classification with delphinID can be tweaked to achieve optimal results on your data:

- Click detection settings including **trigger threshold, click length, digital pre-filter** and **trigger filter** can be adjusted to control features of the click waveform targeted for detection.
- Clicks can be categorised into spectral profiles through **click classification.** delphinID.psfx includes a default target spectral profile for delphinids.
- Whistle & moan detection settings including **minimum length, min and max frequency, and signal-to-noise threshold (trigger)** can be adjusted to control the time-frequency features of whistles targeted for detection.

EVENT CLASSIFICATION IN eventClassifier

The output from base acoustic classification using delphinID or ROCCA in PAMGuard is not intended to be interpreted as a final classification result – thorough testing of our northeast Atlantic classifiers demonstrates that both the highest overall accuracy and most consistent performance across species is achieved through integration of predictions from whistles and clicks at the event, or encounter, classification level. We have developed an intuitive interface, using *R Shiny,* for this purpose that we call the eventClassifier, where users can import an .sqlite3 PAMGuard database containing the output from delphinID or ROCCA classifiers to classify acoustic events, which are separated by recording file, to species.

Tutorial 3 | event classification with eventClassifier

1. Download the latest version of R from https://cran.r-project.org/.

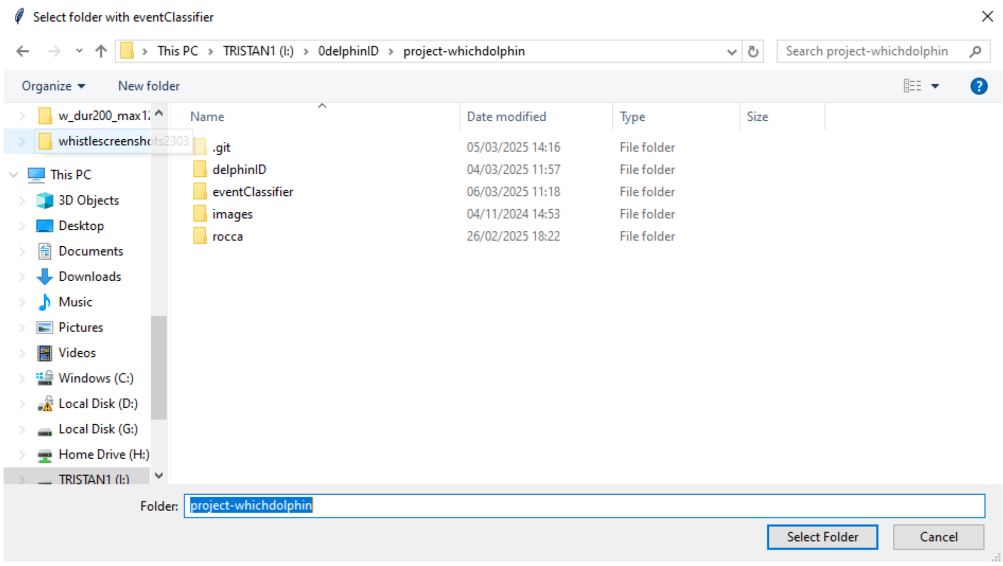
2. Download the which.dolphin repository from https://github.com/tristankleyn/which.dolphin, either by downloading the repository directly from the webpage or by cloning the repository to your device:

```
git clone https://github.com/tristankleyn/which.dolphin.git
```

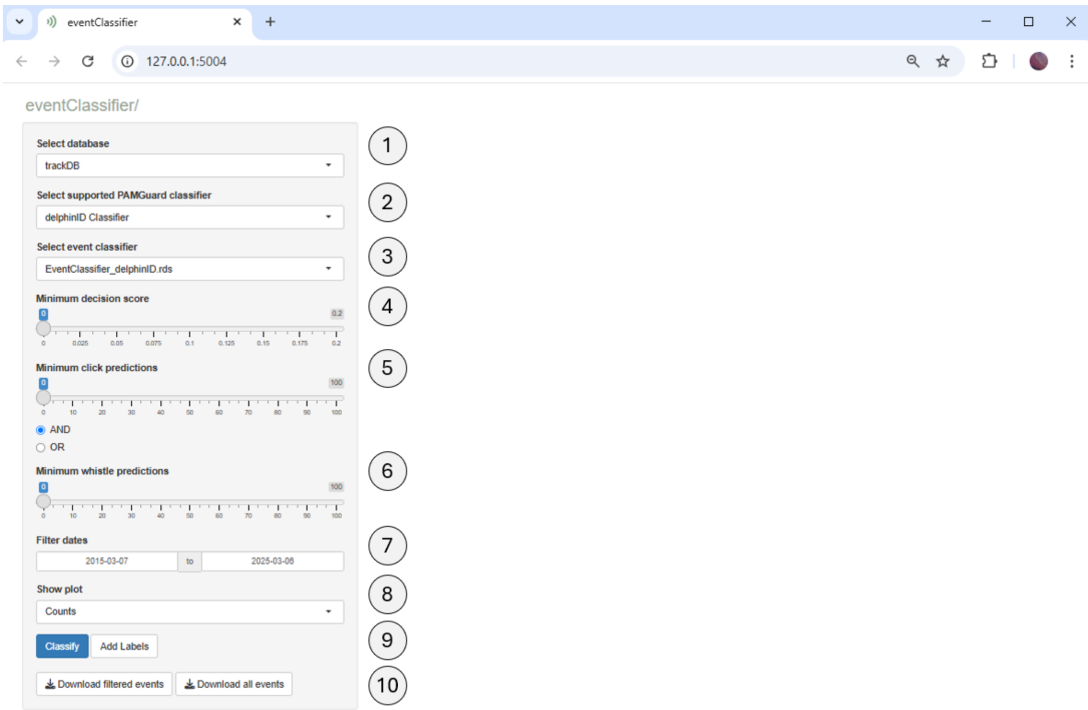
3. Run which.dolphin/eventClassifier/runApp.R. This can be done either by sourcing the filepath or by dragging the runApp.R file into your R console.

```
source('---INPUT PATH---/which.dolphin-main/eventClassifier/runApp.R')
```

4. In the window prompt that appears, select the folder containing eventClassifier.



5. After a few seconds, the eventClassifier interface should launch in an offline browser window, showing the below display.



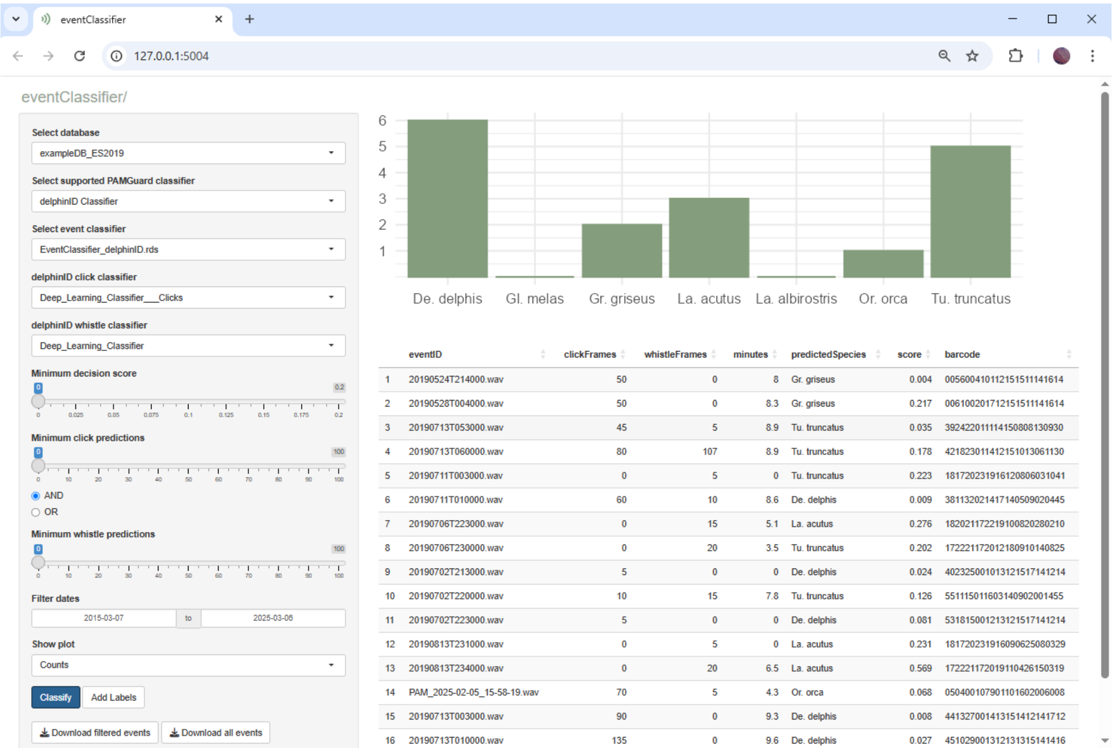
The controls on the sidebar are:

1. **Select database** | choose a PAMGuard database containing base classification results.
2. **Select supported PAMGuard classifier** | choose delphinID or ROCCA
3. **Select event classifier** | choose an event classifier model
4. **Minimum decision score** | a filter for event classifications based on classification confidence. Decision score = M x P , where M is the maximum species classification probability (i.e. the classification probability 0-1 of a given event for the most likely

species) and P is the prominence of this probability, that is the difference between the maximum species probability and the second highest species probability.

5. **Minimum click predictions** | a filter for event classifications based on the number of base classifications from the click classifier within an event. For ROCCA, this number is equal to the number of individual clicks classified while for delphinID it is equal to the number of click detection frames.
6. **Minimum whistle predictions** | a filter for event classifications based on the number of base classifications from the whistle classifier within an event. For ROCCA, this number is equal to the number of individual whistles classified while for delphinID it is equal to the number of whistle detection frames.
7. **Filter dates** | a filter for event classifications based on the starting UTC timestamps of events.
8. **Show plot** | choose between **counts**, a barplot display of the number of classified events for each species, and **map**, a scatterplot display showing clustering of events in acoustic space.
9. **Classify/Add labels** | Click **Classify** to calculate and display results. Click **Add labels** to add an optional label column to the results table.
10. **Download filtered or all events** | Export events to .csv (saves to eventClassifier folder)

6. Click **Classify** to classify events and display results. After some time, depending on the size of the database, plot and table representations of the results should appear in the main panel of the interface.



The columns in the results table are:

- **eventID** | identifier for each event (recording file) extracted from the database.
- **clicks** or **clickFrames** | the number of click predictions within each event
- **whistles** or **whistleFrames** | the number of whistle predictions within each event
- **predictedSpecies** | Latin names for the classified species for each event
- **score** | decision score (described in step 5.4) for each event classification
- **barcode** | a 24-digit embedding representing the acoustic profile of each event

7. Click **Download filtered events** to export filtered classification results or **Download all events** to export all classification results to a .csv file, which is saved in the eventClassifier folder.

PERFORMANCE

The below tables show various performance metrics for event classification of northeast Atlantic delphinid species using the delphinID and ROCCA base classifiers. Metrics were estimated through cross-validated testing across more than 250 recording files representing 105 events, or encounters, which were separated from one another by a distance of at least 6 km. Performance shown below reflects that using decision thresholds to filter out less confidence predictions (see table captions).

Table 1 | Confusion matrix of event classification using **delphinID whistle and click classifiers** using a base classification threshold of 30% and a minimum event decision score of 0.05. Overall average accuracy across 30 random trials was estimated at 86.3% (SD=1.8%), while accuracies varied amongst species with lowest accuracy for *Delphinus delphis* at 80.0% and highest accuracy for *Lagenorhynchus albirostris* at 92.0%. Standard deviations in accuracies are shown in parentheses and the average number of events tested in each trial is given in the right-hand column. Event predictions below the decision score threshold of 0.05 and thus discarded made up 29.3% of all events.

	Classified as							
True species	***De. delphis***	***Gr. griseus***	***Gl. melas***	***La. acutus***	***La. albirostris***	***Oo. orca***	***Tu. truncatus***	**N**
De. delphis	**0.80 (0.03)**	0.01 (<0.01)	0.0 (<0.01)	0.10 (0.01)	0.06 (0.01)	0.0 (<0.01)	0.06 (0.02)	**20**
Gr. griseus	0.02 (0.03)	**0.90 (0.02)**	0.0 (<0.01)	0.01 (<0.01)	0.09 (0.02)	0.0 (<0.01)	0.02 (0.01)	**10**
Gl. melas	0.01 (0.01)	0.0 (<0.01)	**0.86 (0.01)**	0.0 (<0.01)	0.0 (<0.01)	0.14 (0.03)	0.01 (<0.01)	**14**
La. acutus	0.12 (0.03)	0.0 (<0.01)	0.0 (<0.01)	**0.88 (0.05)**	0.0 (<0.01)	0.0 (<0.01)	0.0 (0.01)	**8**
La. albirostris	0.01 (0.01)	0.08 (0.02)	0.0 (<0.01)	0.01 (0.01)	**0.92 (0.02)**	0.0 (<0.01)	0.0 (<0.01)	**12**
Oo. orca	0.0 (<0.01)	0.02 (0.01)	0.16 (0.02)	0.0 (<0.01)	0.0 (<0.01)	**0.83 (0.06)**	0.01 (0.01)	**6**
Tu. truncatus	0.01 (0.03)	0.10 (0.03)	0.02 (0.01)	0.06 (0.01)	0.0 (<0.01)	0.0 (<0.01)	**0.90 (0.04)**	**10**

Table 2 | Confusion matrix of event classification using **ROCCA whistle and click classifiers** using a vocalization threshold of 30% and a minimum event decision score of 0.05. Overall average accuracy across 30 random trials was estimated at 78.0% (SD=1.3%), while accuracies varied amongst species with lowest accuracy for *Grampus griseus* at 60.0% and highest accuracy for *Orcinus orca* at 92.0%. Standard deviations in accuracies are shown in parentheses and the average number of events tested in each trial is given in the right-hand column. Event predictions below the decision score threshold of 0.05 and thus discarded made up 26.9% of all events.

	Classified as							
True species	***De. delphis***	***Gr. griseus***	***Gl. melas***	***La. acutus***	***La. albirostris***	***Oo. orca***	***Tu. truncatus***	**N**
De. delphis	**0.76 (0.02)**	0.0 (<0.01)	0.03 (<0.01)	**0.21 (0.02)**	0.0 (<0.01)	0.0 (<0.01)	0.01 (<0.01)	**21**
Gr. griseus	0.08 (0.01)	**0.60 (0.05)**	0.09 (<0.01)	0.10 (<0.01)	0.04 (<0.01)	0.01 (<0.01)	0.10 (0.01)	**11**
Gl. melas	0.0 (<0.01)	0.0 (<0.01)	**0.86 (<0.01)**	0.0 (<0.01)	0.0 (<0.01)	0.16 (0.02)	0.0 (<0.01)	**15**
La. acutus	0.15 (0.04)	0.0 (<0.01)	0.0 (<0.01)	**0.76 (0.02)**	0.03 (<0.01)	0.0 (<0.01)	0.06 (0.01)	**9**
La. albirostris	0.02 (0.01)	0.17 (0.04)	0.0 (<0.01)	0.0 (<0.01)	**0.82 (0.01)**	0.0 (<0.01)	0.0 (<0.01)	**10**
Oo. orca	0.0 (<0.01)	0.0 (<0.01)	0.04 (0.01)	0.05 (<0.01)	0.0 (<0.01)	**0.92 (0.11)**	0.0 (<0.01)	**7**
Tu. truncatus	**0.21 (0.03)**	0.0 (<0.01)	0.0 (<0.01)	0.06 (0.01)	0.0 (<0.01)	0.0 (<0.01)	**0.83 (0.04)**	**14**

IMPORTANT CONSIDERATIONS

Our classifiers are only as good as the data you feed it.

While our classifiers available for base and event classifications were trained on thousands of examples of delphinids vocalizations of visually-confirmed recordings of Atlantic delphinid species, it is important to consider that the quality of your detections made in PAMGuard can influence the reliability of classification results. The performance estimated of our classifiers in Tables 2-3 shows variable accuracy and misclassification across species. As the use of false detections of whistles and clicks (caused by other noise sources in recordings, such as anthropogenic or environmental noise) in testing was found to be minimal for our study, we did not empirically investigate how false detections may influence classifier performance and misclassification rates. It is crucial, however, to consider how systematic false detections of whistles or clicks could alter the expected performance of classifiers and thus it is crucial to monitor and validate a portion of your PAMGuard detections when using our classifiers. Lastly, we recommend targeting high signal-to-noise (SNR) vocalizations for classification and to avoid using recordings with abnormally loud (so much as to confound vocalizations) background noises such as electrical interference or vessel noise.

No classifier is perfect!

Even if your detections are true detections and of high-quality, there is a limit to our classifiers' abilities to predict species accurately simply due to the fact that acoustically identifying dolphin species is no easy task! Anyone using these models must account for some inherent error, which can be estimated from the confusion matrices in Tables 2-3, when interpreting their results. In applying acoustic classifiers to large datasets, we have found that visualising classification results over time and propagating expected classification errors can provide additional insights into understanding likely patterns of true and false classifications.

REFERENCES

[1] Gillespie, D., Mellinger, D.K., Gordon, J.O.N.A.T.H.A.N., Mclaren, D., Redmond, P.A.U.L., McHugh, R., Trinder, P.W., Deng, X.Y. and Thode, A., 2008. PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans. *Journal of the Acoustical Society of America*, *30*(5), pp.54-62.

[2] Rankin, S., Archer, F., Keating, J.L., Oswald, J.N., Oswald, M., Curtis, A. and Barlow, J., 2017. Acoustic classification of dolphins in the California Current using whistles, echolocation clicks, and burst pulses. *Marine Mammal Science*, *33*(2), pp.520-540.

[3] Zahn, M.J., Rankin, S., McCullough, J.L., Koblitz, J.C., Archer, F., Rasmussen, M.H. and Laidre, K.L., 2021. Acoustic differentiation and classification of wild belugas and narwhals using echolocation clicks. *Scientific Reports*, *11*(1), p.22141.

[4] Rankin, S., Sakai, T., Archer, F.I., Barlow, J., Cholewiak, D., DeAngelis, A.I., McCullough, J.L., Oleson, E.M., Simonis, A.E., Soldevilla, M.S. and Trickey, J.S., 2024. Open-source machine learning BANTER acoustic classification of beaked whale echolocation pulses. *Ecological Informatics*, *80*, p.102511.

[5] Oswald, J.N., and M. Oswald. 2013. ROCCA (Real-time Odontocete Call Classification Algorithm) User's Manual. Prepared for Naval Facilities Engineering Command Atlantic, Norfolk, Virginia under HDR Environmental, Operations and Construction, Inc Contract No. CON005-4394-009, Subproject 164744, Task Order 03, Agreement # 105067. Prepared by Bio-Waves, Inc., Encinitas, California.