SEISMICA

An Open Source Hydroacoustic Benchmarking Framework for Geophonic Signal Detection: supplementary materials

Pierre-Yves Raumer * [1,2], **Sara Bazin** [1], **Dorian Cazau** [2], **Vaibhav Vijay Ingale** [1,3], **Aude Lavayssière** [1], **Jean-Yves Royer** [1]

[1]Geo-Ocean, Univ Brest, CNRS, Ifremer, UMR6538, F-29280 Plouzane, France, [2]Lab-STICC, ENSTA-Bretagne, UMR6285, F-29200, Brest, France, [3]Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, CA, USA

This document is intended to provide additional materials regarding the main paper, focusing on dataset and model analysis.

Contents

*Corresponding author: pierre-yves.raumer@univ-brest.fr

1 Supplementary information S1: Aplose

APLOSE, a tool developed by the OSmOSE team from ENSTA Bretagne, enables to annotate a set of spectrograms. In this work, it was used on a set of spectrograms covering the whole periods of OHASISBIO-2020 and HYDROMOMAR. Two screenshots from the software are available at FIGURE S1, and show a spectrogram before and after time-frequency boxes annotations.

2 Supplementary information S2: annotation analysis

The main materials provide, regarding the analysis of the annotations obtained from the campaign, a figure showing the absolute number of annotations of each type for each annotator. Given the discrepancies of activity between the different annotators, a relative label importance analysis may be relevant. The FIGURE S2 presents these values for the campaign regarding the OHASISBIO-2020 and HYDROMOMAR-2013 dataset.

Given the geographical variability among the stations used causes a variability in the observed soundscapes, a figure showing the absolute number of annotations per label and per station resulting from the campaigns is available at FIGURE S3.

Some annotations present in the catalog used in this study have been seen by several annotators, other have only been seen by one. FIGURE S7 shows some examples of spectrograms of events seen by one annotator, while FIGURE S8 shows events seen by three annotators.

3 Supplementary information S3: model evaluation

The distributions of errors for segmentation models in the case of "conservative" datasets are given by FIGURE S6.

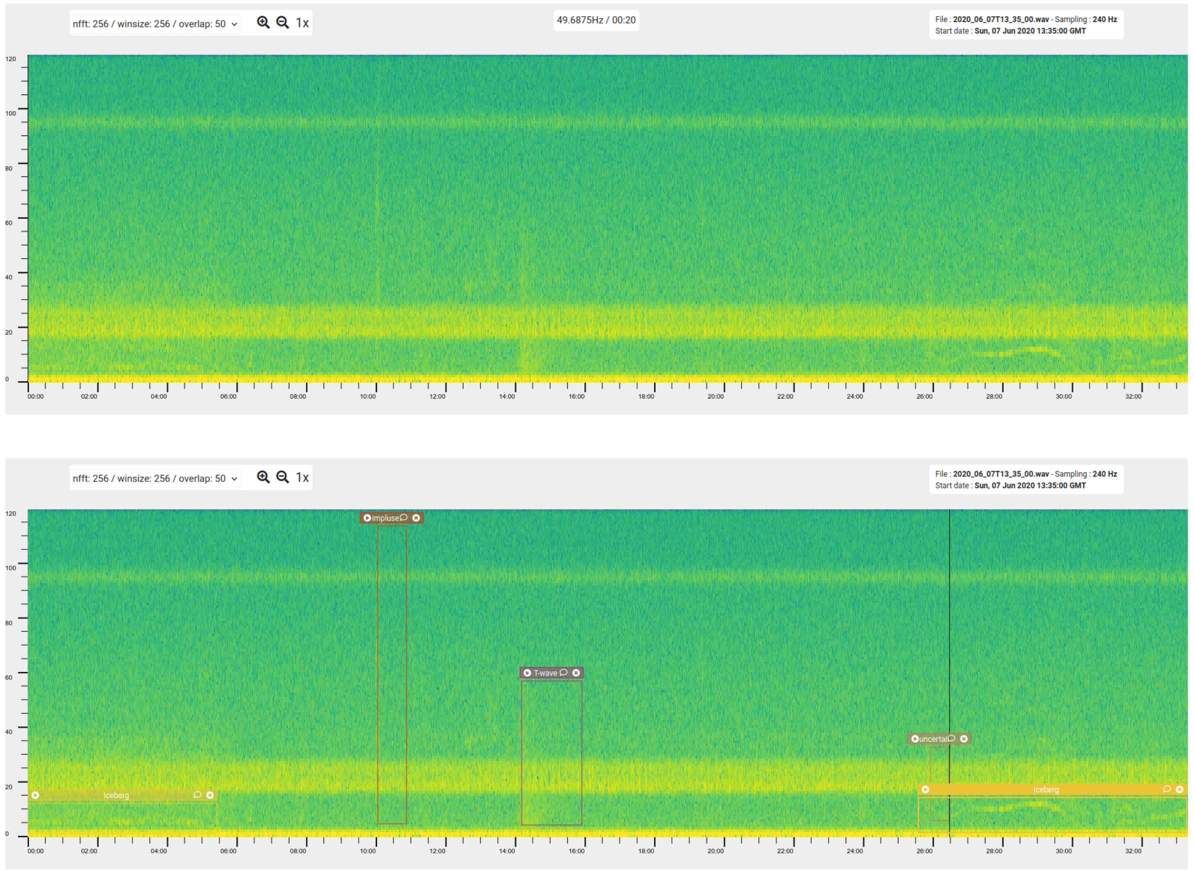

Figure S1 Aplose tool screenshots. Top : spectrogram of 2000 seconds showing a diversity of signals. Bottom : same spectrogram annotated with time-frequency boxes. The orange boxes account for cryogenic signals, the left brown box for a H-wave, the purple box for a T-wave and the right brown box for an uncertain signal.

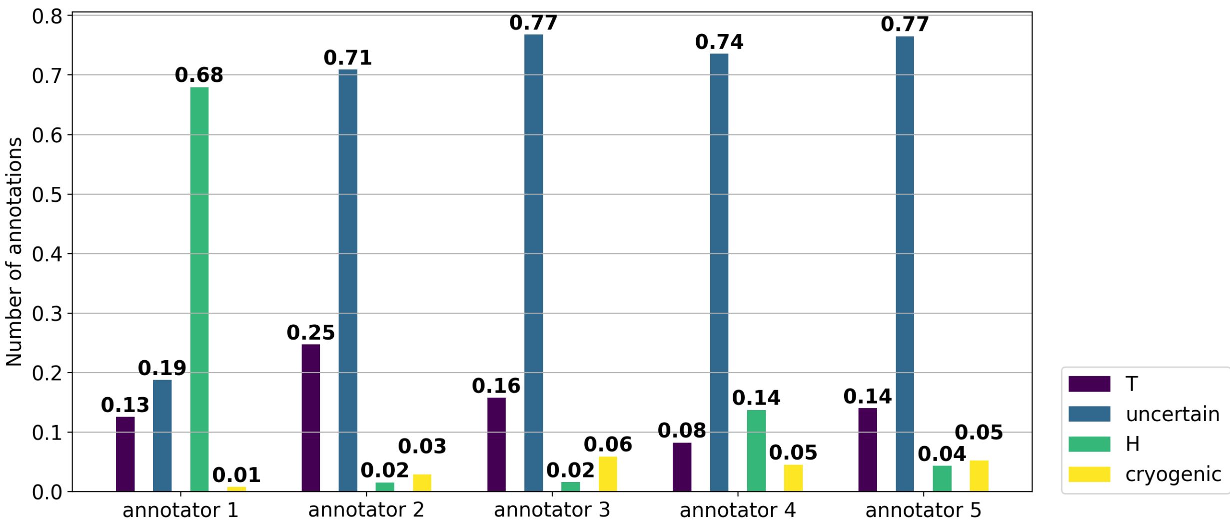

Figure S2 Proportion of each label in the annotation catalog produced by different annotators.

Pursuing the exploration of scores, examples of spectrograms of "positive" segments missed by all models are shown at FIGURE S7 and FIGURE S8.

4 Supplementary information S4: model analysis

The SGBT models have several hyperparameters, two of which are the maximal number of trees and the learning rate. The tuning of these two hyperparameters is of tremendous importance for the training of the model

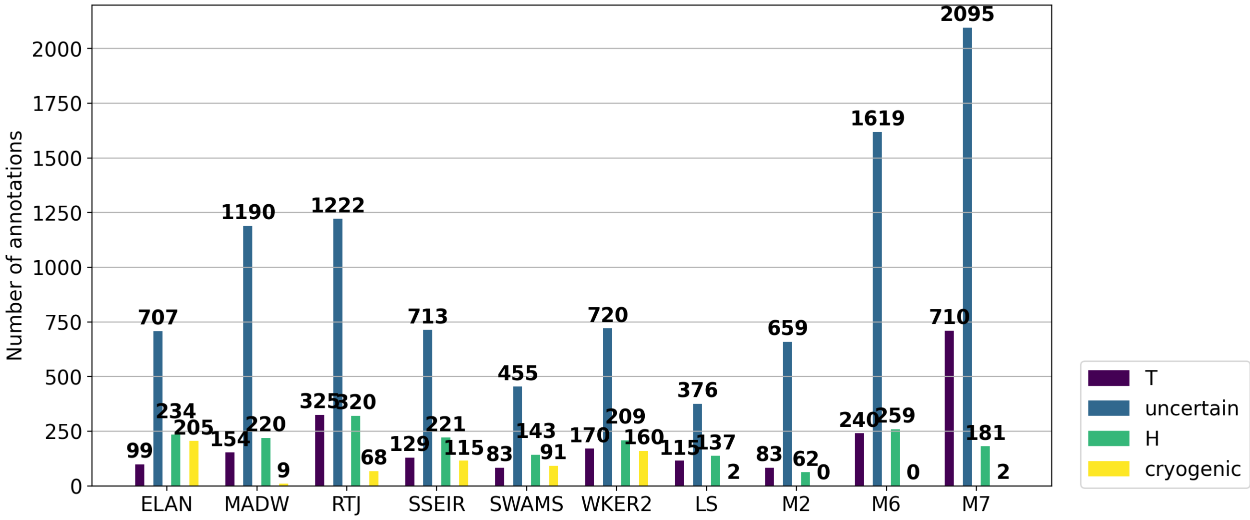

Figure S3 Number of annotations of each label in the annotation catalog produced by different stations.

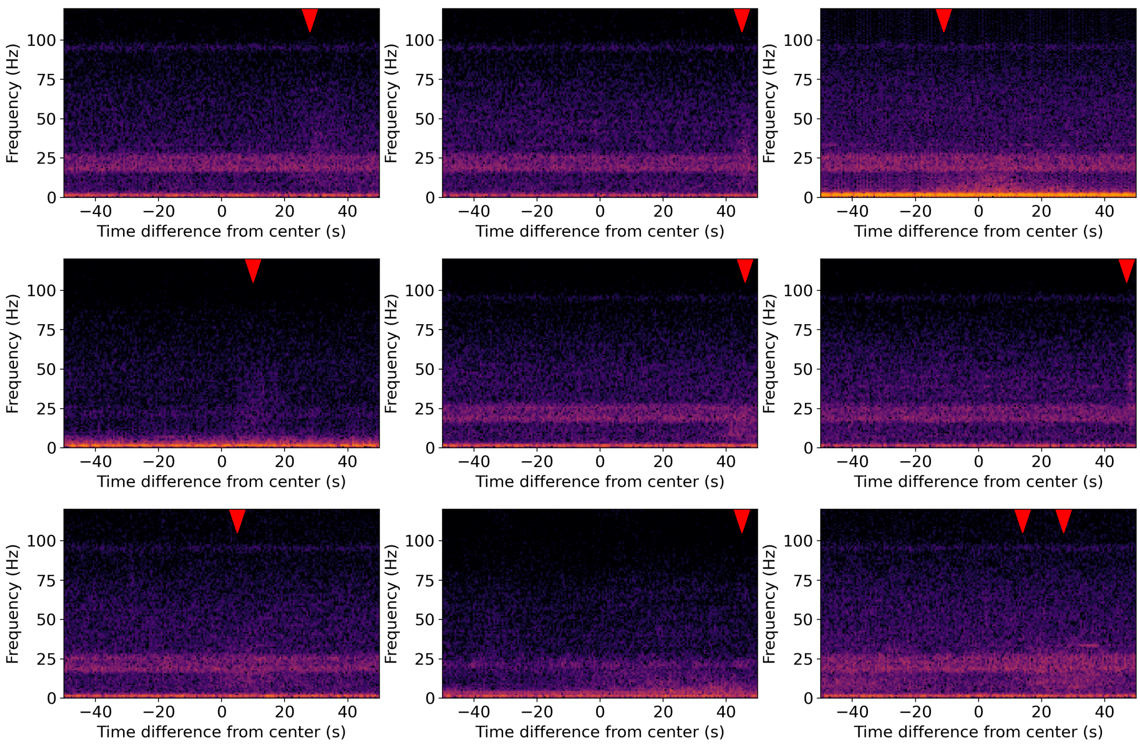

Figure S4 Spectrograms of examples of events seen by one annotator only. Red triangles show the Time of Arrival (ToA) of these events. Spectrograms are normalized between 60 and 140 dB.

and is highly visible on its final score. We performed 10,000 trainings of the model using different values for these two hyperparameters, using a bayesian optimization algorithm to guide the search in a more clever way than a grid-search. The Tree Parzen Estimator algorithm was used with the library hyperopt, using some uniform a-priori laws for the variables, linearly from 0 to 10,000 for the number of trees and logarithmically from 0.01 to 1 for the learning rate. The search space explored by the algorithm is plotted on FIGURE S9, showing the tried values together with their scores, and the position of the best couple tried.

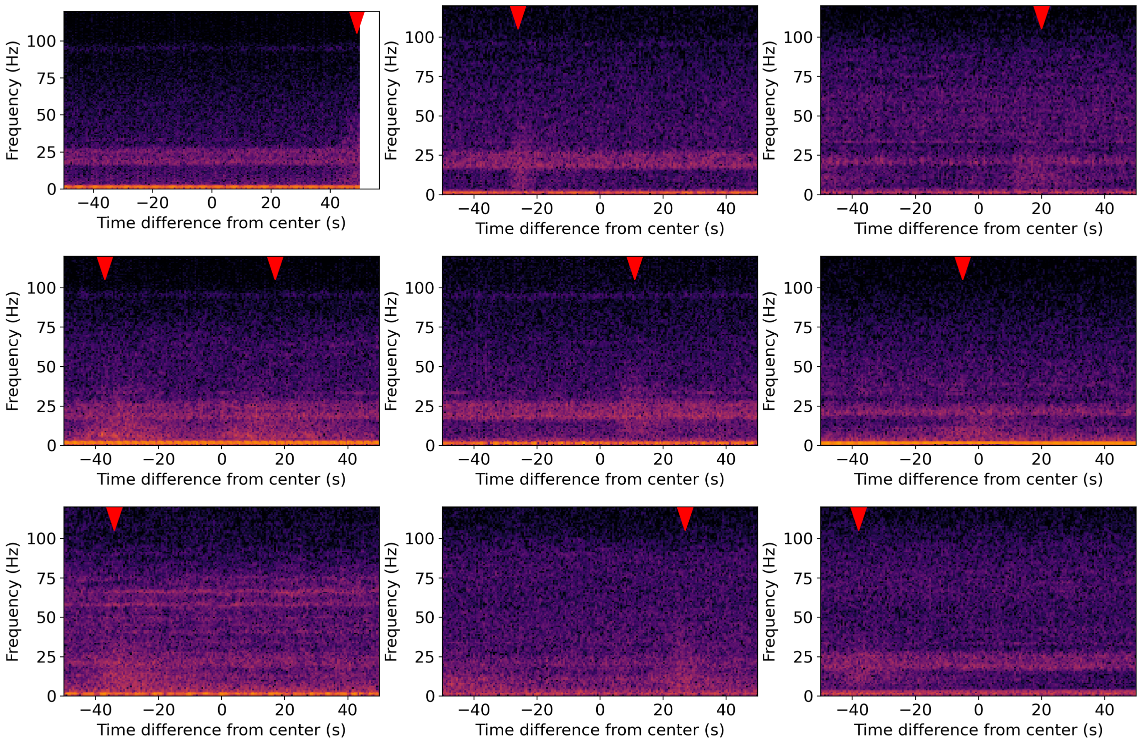

Figure S5 Spectrograms of examples of events seen by exactly three annotators. Red triangles show the ToA of these events. Spectrograms are normalized between 60 and 140 dB.

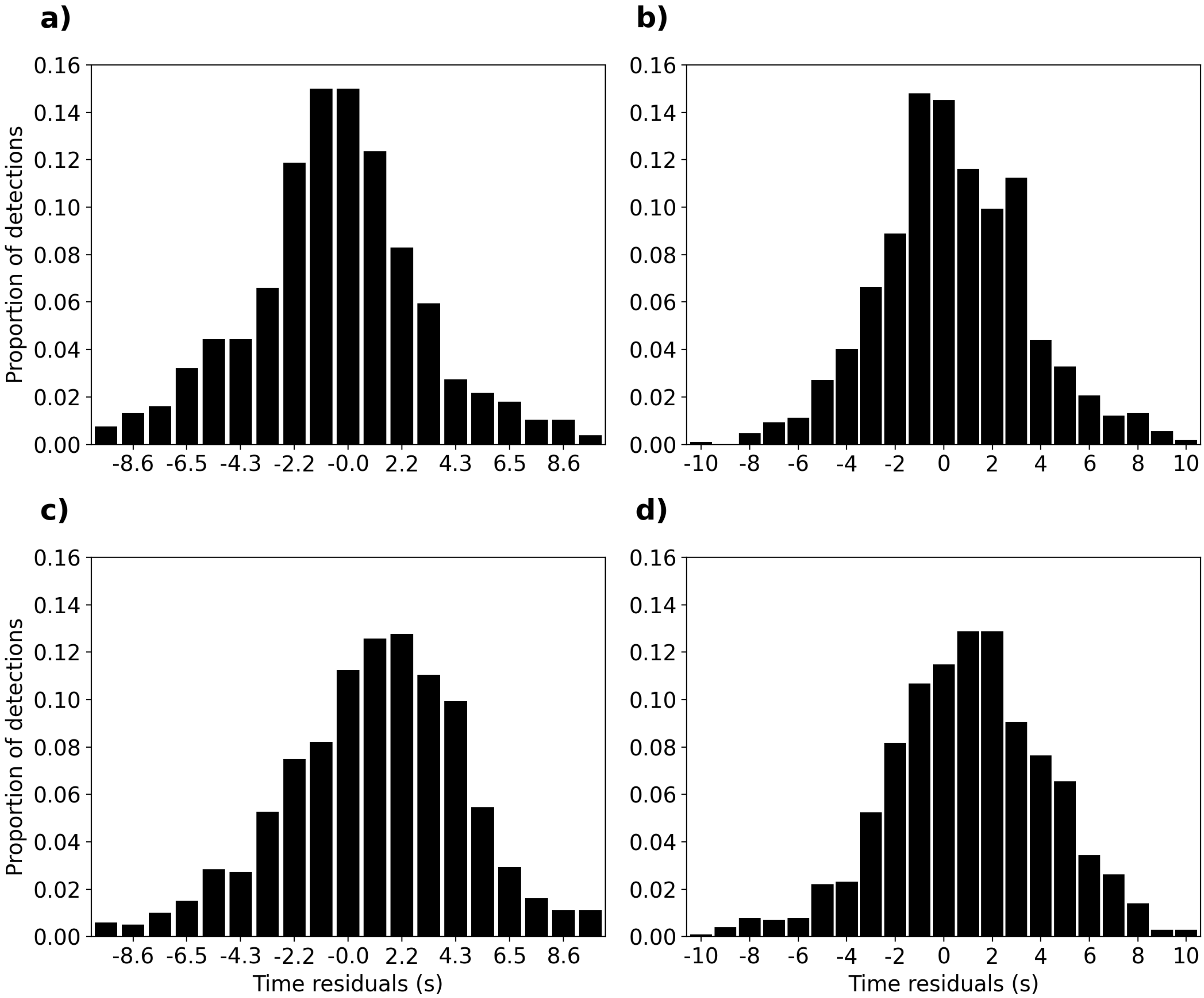

Figure S6 Distribution of pick errors of TiSSNet and AcousticPhaseNet on the two evaluation datasets in ”conservative” mode: (a) TiSSNet on OHASISBIO-2020 dataset, (b) AcousticPhaseNet on OHASISBIO-2020, (c) TiSSNet on HYDROMOMAR-2013 and (d) AcousticPhaseNet on HYDROMOMAR-2013. For (a) and (c), the x-axis increments are multiples of the spectrogram time resolution, which is close to 0.55s.

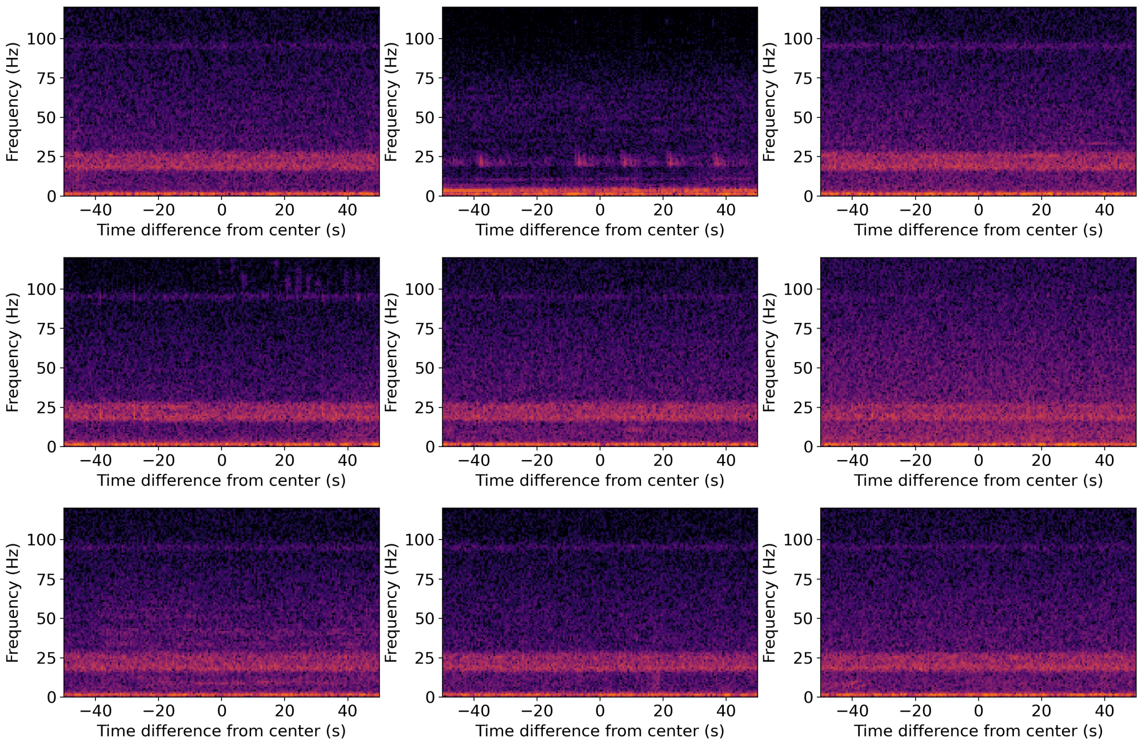

Figure S7 Spectrograms of examples of events seen by exactly 1 annotator, but missed by all models. Events seen by 1 annotator only represent 259 of the 400 common false positive segments. Spectrograms are normalized between 60 and 140 dB.

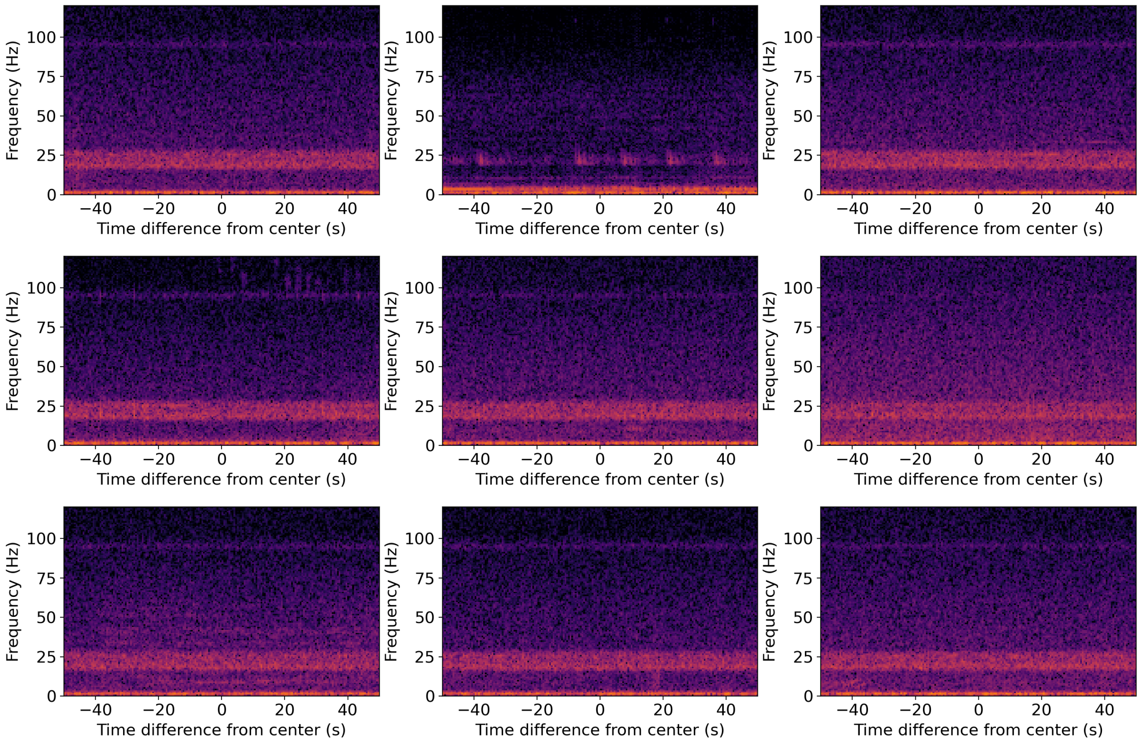

Figure S8 Spectrograms of examples of events seen by exactly 3 annotators, but missed by all models. Events seen by 3 annotators represent 55 of the 400 common false positive segments. Spectrograms are normalized between 60 and 140 dB.

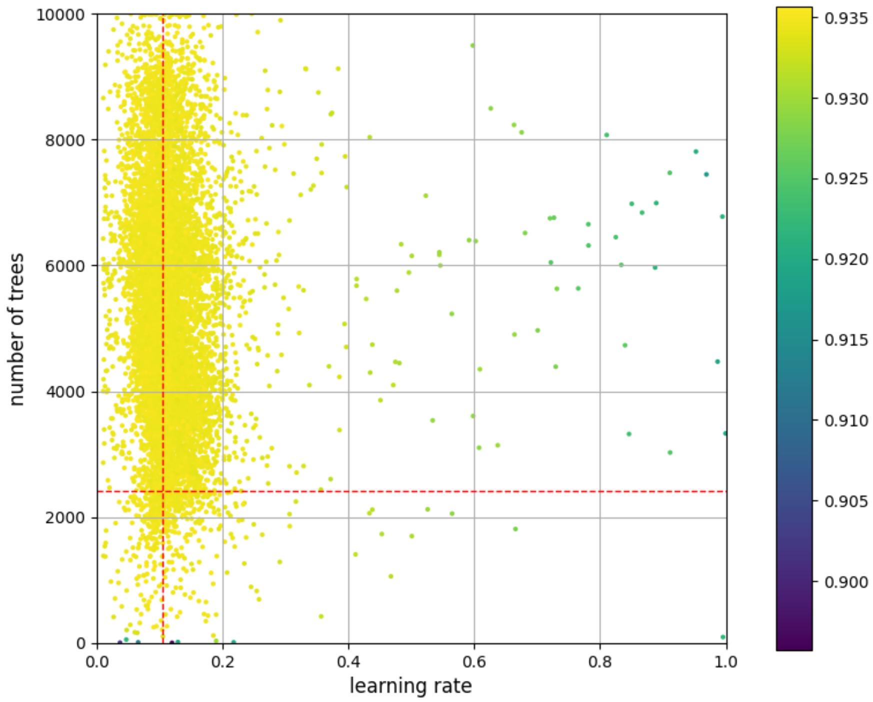

Figure S9 Search space of the learning rate and maximum number of tree for the SGBT model, with segments of 20 seconds. Points represent the values tried by the Tree Parsen Estimator meta-optimization algorithm, colored by their objective function value. The red dot lines represent the chosen values of the parameters that led to the maximal objective function value.