



Acoustic Pipeline

Classifier Technical Specifications: Nocturnal Breeding Birds Europe

1. Introduction

The BTO Acoustic Pipeline's **Nocturnal Breeding Birds Europe** classifier (version 3) detects the main calls of a suite of 22 largely nocturnal European breeding birds. This document provides technical notes on the derivation of the classifier, its use and how to interpret results. The rest of this document is arranged into five parts:

- Training data sample sizes
- Precision and Recall statistics on withheld data
- False positive rates on independent data
- Known issues
- Credits and citation

2. Training data sample sizes

This classifier is trained on strongly labelled audio recordings compiled by Simon Gillings (BTO) and collaborators, with additional recordings from xeno-canto. We are grateful to our collaborators and the sound recordists who share recordings via xeno-canto. The following table gives the number of audio samples used to train and evaluate the classifier. Sample sizes vary across species; for species with smaller sample sizes, classifier performance is typically poorer and detections should be checked carefully. The Background class encompasses ambient and anthropogenic sounds plus non-target wildlife (i.e. other bird species). Three additional classes (Insects [largely Orthoptera], Amphibians and Dogs) were included during model training but are not exported in results.

Class	Scientific name	Train sample size	Evaluation sample size
Dog	<i>Mammalia</i>	5000	151
Amphibian	<i>Amphibia</i>	5000	398
Insect	<i>Insecta</i>	5000	400
Background		125000	10000
Quail	<i>Coturnix coturnix</i>	5000	280
Red-necked Nightjar	<i>Caprimulgus ruficollis</i>	5000	400
Nightjar	<i>Caprimulgus europaeus</i>	5000	400
Corncrake	<i>Crex crex</i>	5000	400
Spotted Crane	<i>Porzana porzana</i>	5000	214
Little Crane	<i>Zapornia parva</i>	5000	151
Crane	<i>Grus grus</i>	5000	380
Woodcock	<i>Scolopax rusticola</i>	5000	400
Barn Owl	<i>Tyto alba</i>	5000	400

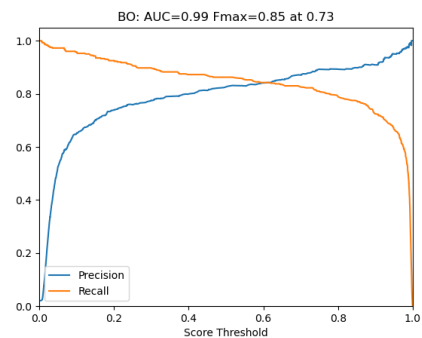
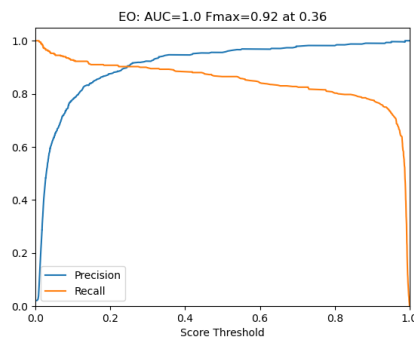
Class	Scientific name	Train sample size	Evaluation sample size
Tengmalm's Owl	<i>Aegolius funereus</i>	5000	400
Little Owl	<i>Athene noctua</i>	5000	400
Eurasian Pygmy Owl	<i>Glaucidium passerinum</i>	5000	400
Scops Owl	<i>Otus scops</i>	5000	400
Long-eared Owl	<i>Asio otus</i>	5000	400
Eurasian Eagle-Owl	<i>Bubo bubo</i>	5000	400
Tawny Owl	<i>Strix aluco</i>	5000	400
Ural Owl	<i>Strix uralensis</i>	5000	400
Great Grey Owl	<i>Strix nebulosa</i>	5000	400
River Warbler	<i>Locustella fluviatilis</i>	5000	400
Savi's Warbler	<i>Locustella luscinioides</i>	5000	400
Grasshopper Warbler	<i>Locustella naevia</i>	5000	400
Thrush Nightingale	<i>Luscinia luscinia</i>	5000	400

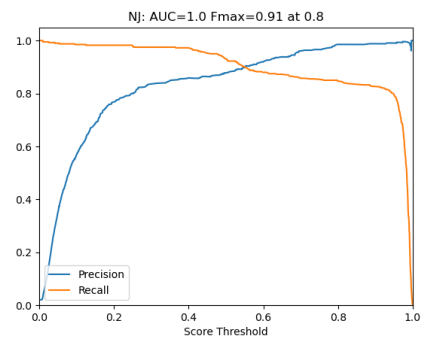
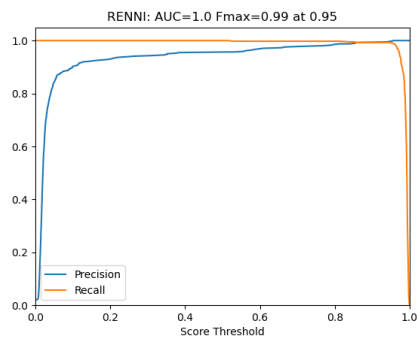
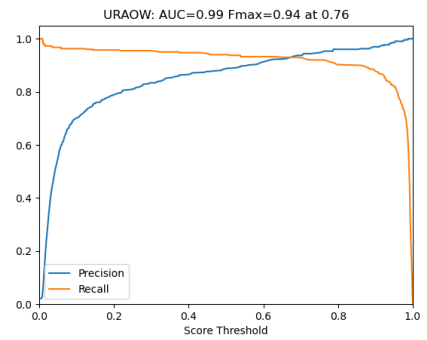
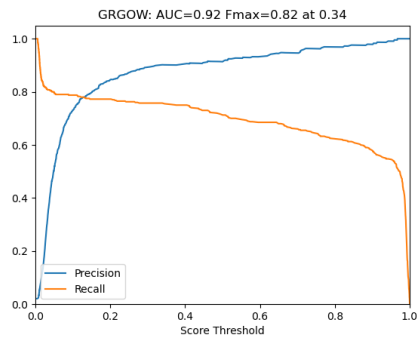
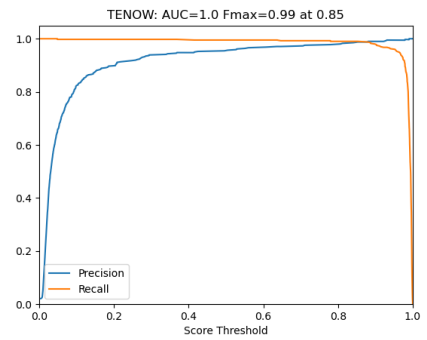
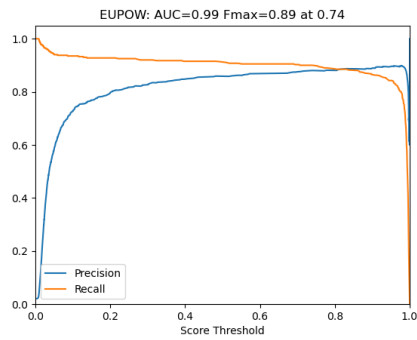
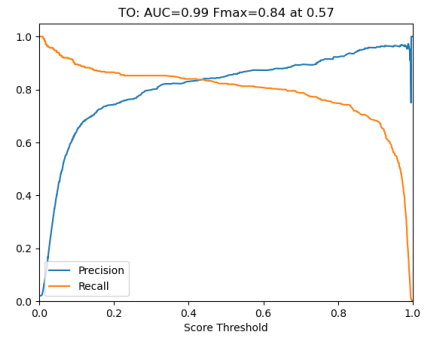
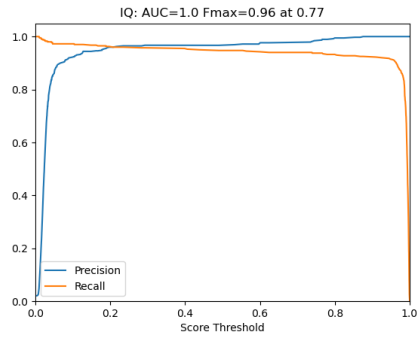
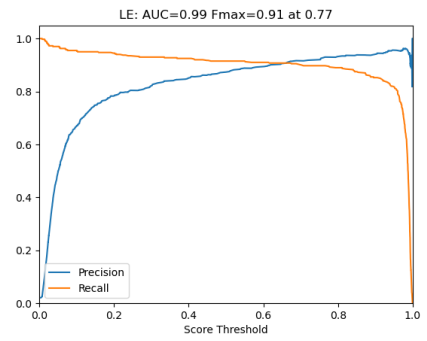
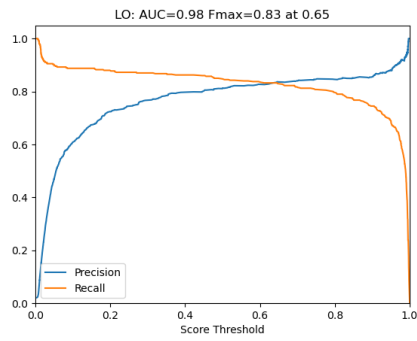
3. Precision and Recall

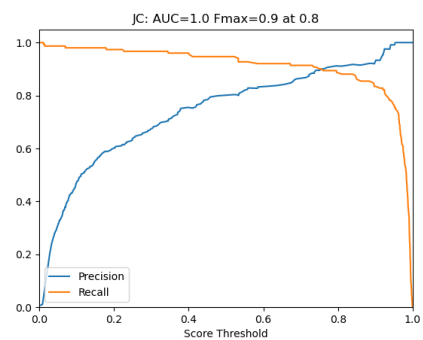
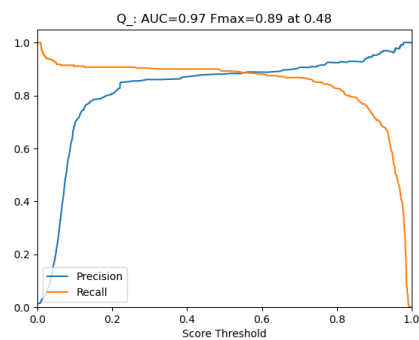
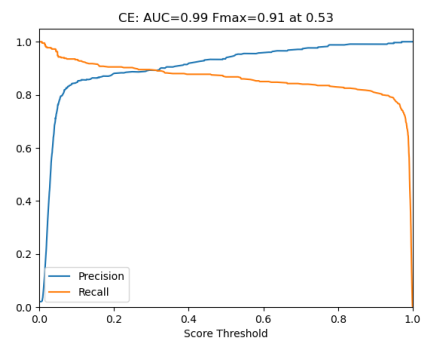
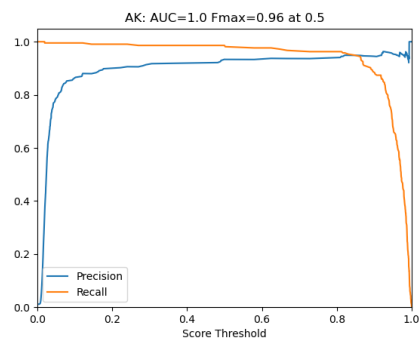
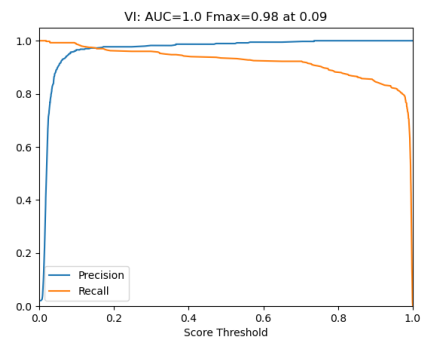
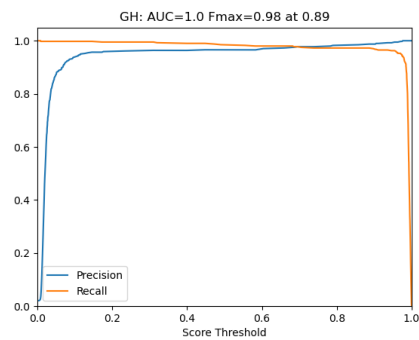
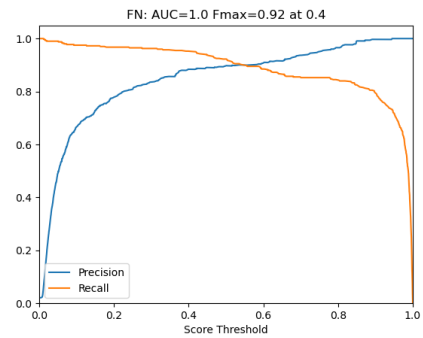
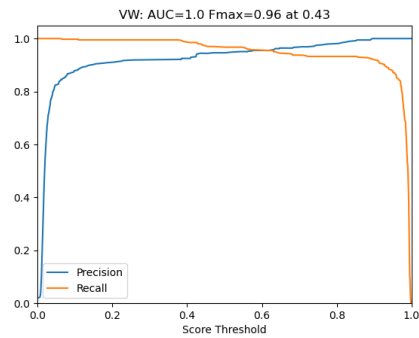
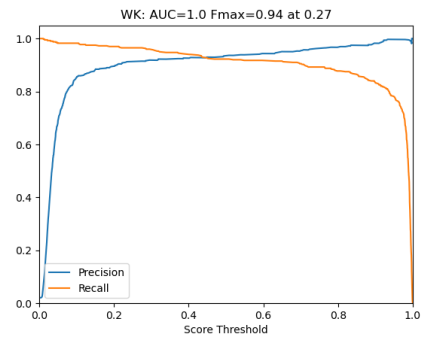
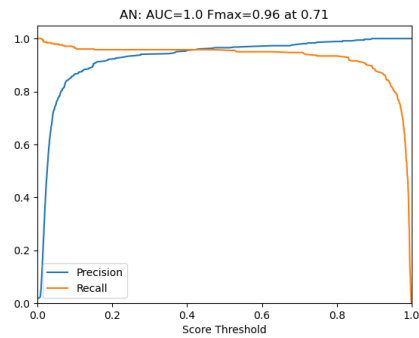
Classifier performance is typically evaluated using the metrics *Precision* and *Recall*:

- **Precision** is the % of detections returned by a classifier that truly are of that species. If the classifier thinks 100 clips contain Redwing calls, Precision is the percentage of these that actually do contain Redwing calls.
- **Recall** is the percentage of true instances that are detected. For example, if 100 clips contain Redwing calls, Recall is the percentage of these that are found by the classifier.

These metrics are *threshold dependent*. If we say that all classifier scores greater than 0.5 constitute a detection we will get different Precision and Recall values than if we use a more stringent score threshold of 0.9. Figures below shows how Precision and Recall vary with score threshold for each species treated by this classifier. These figures are based on application of the classifier to withheld training data.







The subsequent table provides Precision and Recall values for commonly used score thresholds. Additionally, *Best T* gives the threshold at which the F-score is maximised, which is a way of optimising both Precision and Recall. The Precision and Recall statistics for that threshold (*P|R Best*) are also shown.

Detections exported from the Acoustic Pipeline by default use a stringent threshold of 0.9 to reduce false positives.

Species	Scientific	P R 0.50	P R 0.90	P R 0.95	P R 0.99	Best T	P R Best
Quail	<i>Coturnix coturnix</i>	88.3 89.3	95.2 71.4	96.3 55.7	100.0 4.3	0.48	88.0 90.0
Red-necked Nightjar	<i>Caprimulgus ruficollis</i>	95.7 100.0	99.5 99.2	99.7 99.0	100.0 66.5	0.95	100.0 99.0
Nightjar	<i>Caprimulgus europaeus</i>	87.9 93.0	98.8 82.5	99.1 79.2	99.2 29.8	0.80	99.0 85.0
Corncrake	<i>Crex crex</i>	94.0 86.8	99.1 80.8	99.7 78.2	100.0 57.2	0.53	95.0 87.0
Spotted Crake	<i>Porzana porzana</i>	93.3 98.1	94.5 88.3	95.4 67.3	93.8 21.0	0.50	93.0 99.0
Little Crake	<i>Zapornia parva</i>	80.3 94.7	93.3 83.4	99.1 76.2	100.0 33.8	0.80	91.0 89.0
Crane	<i>Grus grus</i>	96.5 95.5	100.0 88.4	100.0 80.8	100.0 45.5	0.71	98.0 95.0
Woodcock	<i>Scolopax rusticola</i>	93.7 92.2	98.2 83.2	99.7 78.0	99.5 47.0	0.27	91.0 96.0
Barn Owl	<i>Tyto alba</i>	82.7 86.2	90.9 72.2	95.0 66.5	98.4 44.8	0.73	88.0 82.0
Tengmalm's Owl	<i>Aegolius funereus</i>	95.7 99.5	99.0 97.8	99.5 96.0	100.0 68.5	0.85	99.0 99.0
Little Owl	<i>Athene noctua</i>	81.3 84.8	85.6 74.5	88.1 68.8	93.2 48.0	0.65	84.0 83.0
Eurasian Pygmy Owl	<i>Glaucidium passerinum</i>	85.8 91.0	88.9 86.2	89.6 84.2	88.1 66.5	0.74	88.0 90.0
Scops Owl	<i>Otus scops</i>	96.9 94.8	100.0 92.2	100.0 91.2	100.0 70.2	0.77	99.0 94.0
Long-eared Owl	<i>Asio otus</i>	87.4 91.5	94.2 85.2	95.6 81.8	94.4 38.2	0.77	93.0 90.0
Eurasian Eagle-Owl	<i>Bubo bubo</i>	95.8 86.5	99.0 77.5	99.3 72.8	100.0 37.5	0.36	95.0 89.0
Tawny Owl	<i>Strix aluco</i>	85.2 82.2	95.8 68.2	96.5 55.8	95.9 11.8	0.57	87.0 81.0
Ural Owl	<i>Strix uralensis</i>	88.9 94.0	97.0 87.5	98.5 82.0	100.0 51.7	0.76	95.0 92.0
Great Grey Owl	<i>Strix nebulosa</i>	91.3 71.2	97.9 57.8	99.1 54.5	100.0 30.0	0.34	90.0 76.0
River Warbler	<i>Locustella fluviatilis</i>	94.6 96.8	100.0 91.8	100.0 87.0	100.0 50.0	0.43	94.0 98.0
Savi's Warbler	<i>Locustella luscinioides</i>	98.9 93.2	100.0 84.5	100.0 82.0	100.0 68.2	0.09	96.0 99.0
Grasshopper Warbler	<i>Locustella naevia</i>	96.6 98.2	98.7 96.8	99.5 96.2	100.0 72.2	0.89	99.0 97.0
Thrush Nightingale	<i>Luscinia luscinia</i>	89.8 92.2	99.7 79.2	100.0 71.8	100.0 44.8	0.40	88.0 95.0

4. False positive rates

The false positive rate indicates how often the classifier suggests a species is present when it is not. For this test we

use a benchmark dataset of c.21,000 ambient sound clips that have been manually checked to confirm they contain no bird records. We run the classifier against this dataset and summarise the percentage of clips that are falsely assigned to a species. As for performance metrics, this measure is threshold dependent, with typically fewer errors when a more stringent score threshold is applied. False positive rates for each species are shown below. Note that this is a simple test: in reality false positive rates may be higher in natural soundscapes, for example, where a distant call of one species is mistaken for another species.

Species	Scientific	0.50	0.90	0.95	0.99	Best
Quail	<i>Coturnix coturnix</i>	0.071	0.019	0.014	0.000	0.071
Red-necked Nightjar	<i>Caprimulgus ruficollis</i>	0.009	0.000	0.000	0.000	0.000
Nightjar	<i>Caprimulgus europaeus</i>	0.085	0.005	0.005	0.000	0.019
Comcrake	<i>Crex crex</i>	0.056	0.000	0.000	0.000	0.047
Spotted Crake	<i>Porzana porzana</i>	0.000	0.000	0.000	0.000	0.000
Little Crake	<i>Zapornia parva</i>	0.009	0.000	0.000	0.000	0.000
Crane	<i>Grus grus</i>	0.169	0.028	0.005	0.000	0.085
Woodcock	<i>Scolopax rusticola</i>	0.071	0.019	0.009	0.000	0.104
Barn Owl	<i>Tyto alba</i>	0.089	0.009	0.005	0.000	0.019
Tengmalm's Owl	<i>Aegolius funereus</i>	0.056	0.009	0.009	0.000	0.019
Little Owl	<i>Athene noctua</i>	0.028	0.014	0.014	0.000	0.024
Eurasian Pygmy Owl	<i>Glaucidium passerinum</i>	0.005	0.000	0.000	0.000	0.000
Scops Owl	<i>Otus scops</i>	0.009	0.000	0.000	0.000	0.000
Long-eared Owl	<i>Asio otus</i>	0.292	0.094	0.047	0.000	0.188
Eurasian Eagle-Owl	<i>Bubo bubo</i>	0.042	0.014	0.009	0.000	0.071
Tawny Owl	<i>Strix aluco</i>	0.334	0.071	0.052	0.005	0.249
Ural Owl	<i>Strix uralensis</i>	0.202	0.071	0.042	0.009	0.160
Great Grey Owl	<i>Strix nebulosa</i>	0.099	0.042	0.019	0.000	0.136
River Warbler	<i>Locustella fluviatilis</i>	0.000	0.000	0.000	0.000	0.000
Savi's Warbler	<i>Locustella luscinioides</i>	0.019	0.000	0.000	0.000	0.155
Grasshopper Warbler	<i>Locustella naevia</i>	0.000	0.000	0.000	0.000	0.000
Thrush Nightingale	<i>Luscinia luscinia</i>	0.000	0.000	0.000	0.000	0.000

5. Known issues

Performance of this classifier is moderate to high depending on species. Low frequency owl hoots present a general problem for classifiers owing to overlap with poorly defined low frequency ambient sounds and can lead to locally poor performance with false positive detections. Distant barking dogs are a known issue for some owls and is an active area of work to collate more examples of such false positive recordings.

As with all classifiers we recommend careful manual verification of detections.

Warnings are currently applied on the basis of European breeding range. Note that for some scarce nocturnal species the range maps are incomplete (e.g. we have true records of Eagle Owl outside the known range).

6. Credits and citation

We are grateful to many sound recordists who shared individual recordings for training the classifier. In particular the following recordists provided large number of clips: Chris Hails, Enrico Caprio.

Dr Chris Scott helped in the development of the classifier.

Use of the classifier should be cited as **BTO Acoustic Pipeline: Nocturnal Breeding Birds Europe (version 3)**