Automated Detection of *Brachyphisis nattecantor* in Ebony Forest Soundscapes Using Machine Learning

Haylee Jackson

Thesis presented for the degree of
Bachelor of Advanced Science, Computing
(Honours)

Centre for Marine Science and Technology
Curtin University
Australia

October 2024

Automated Detection of *Brachyphisis nattecantor* in Ebony Forest Soundscapes Using Machine Learning

Haylee Jackson

Abstract

In Mauritius most of the forests are severely degraded with efforts currently being undertaken by Ebony Forrest to restore the endemic forest and the species within. Katydids are known for acoustically communicating and can be used as indicators of an ecosystems health. The combination of passive acoustic monitoring (PAM) and automated detectors can be used to monitor the presence of a species. *Brachyphisis nattecantor* is a species of katydid found in Ebony Forest with little to no research on its acoustic behaviour. This thesis aims to develop an accurate detector of *Brachyphisis nattecantor* signal in the Ebony Forest soundscape. Three detectors were developed and their performance investigated. The CNN performed the best with an F1-score of 83%, while both the Correlation and Envelope Detectors had an F1-score less than 70%. The CNN developed can be used to help researchers to monitor *Brachyphisis nattecantor* population in Ebony Forest as bioindicators of environmental health and contribute to biodiversity conservation efforts.

Acknowledgements

This thesis has been completed under the supervision of Dr Evgeny Sidenko, Dr Paul Nguyen Hong Duc, Professor Christine Erbe, and Dr Shyam Madhusudhana, whose guidance was invaluable.

I would like to thank Dr Shyam Madhusudhanna for his work in collecting the PAM data used in this study and for his help in identifying species of interest.

Contents

1 Introduction **6**
1.1 Background . . . 6
1.1.1 Mauritius . . . 6
1.1.2 Environmental Monitoring . . . 7
1.1.3 Passive Acoustic Monitoring (PAM) . . . 8
1.1.4 Soundscape . . . 9
1.1.5 Automated Tools . . . 9
1.2 Project Aims and Objectives . . . 10

2 Methodology **12**
2.1 Data Collection . . . 12
2.2 Manual Annotation of Data . . . 15
2.3 Signal Analysis . . . 15
2.4 Building Detectors . . . 16
2.4.1 Correlation Detector . . . 16
2.4.2 Envelope Detector . . . 17
2.5 Machine Learning . . . 18
2.5.1 Preparation of Model Inputs . . . 18
2.5.2 CNN Architecture . . . 19
2.5.3 Training . . . 21
2.5.4 Testing . . . 21
2.5.5 Detections for Other Deployments . . . 21
2.6 Performance Assessment . . . 22

3 Results **23**
3.1 Signal Analysis . . . 23
3.2 Machine Learning . . . 26
3.3 Comparison of Detectors . . . 28

4 Discussion **29**

5 Conclusion **33**

References **34**

List of Figures

1 Map of Passive Acoustic Monitoring Recording Sites Locations 13
2 Recording Schedule for Each Site 14
3 Template Signal Used For Correlation Detector 17
4 Architecture of CNN Developed 20
5 Spectrogram of *Brachyphisis nattecantor* on 16 March 2023 at 5:00 AM to 5:10 AM 23
6 *Brachyphisis nattecantor* Signal for 10 seconds on 16 March 2023 at 5:03 AM 24
7 *Brachyphisis nattecantor* Signal for 1 second on 16 March 2023 at 5:03 AM 25
8 Presence of *Brachyphisis nattecantor* in Site 6 Deployment 1 Recordings Over 24 Hour Period 26
9 Autocorrelation of *Brachyphisis nattecantor* Signal (top) and Representation of Peaks of Autocorrelation on Spectrogram (bottom) . . . 26
10 Precision-Recall Curve of CNN on Testing Dataset 27
11 F1-Score for Thresholds for CNN on Testing Dataset 27
12 Confusion Matrices for Detectors 28

List of Tables

1 Passive Acoustic Monitoring Recording Site Locations 12
2 Recording Schedule of PAM in All Deployments 14
3 Measurements Extracted when Labelling Acoustic Data 15
4 Audio and Spectrogram Pre-processing Parameters 19
5 CNN Model Hyperparameter Values Investigated and Chosen 20
6 Hyperparameters Investigated and Chosen to Train CNN 21
7 *Brachyphisis nattecantor* Signal Characteristics 26
8 Comparison of Detectors 29

1 Introduction

The biodiversity of rain forests in Mauritius is of significant ecological importance, yet the acoustic signatures of its inhabitants, particularly insect species, remain largely unexplored. This project aims to leverage machine learning techniques to identify insect species noises within recordings of Ebony Forest soundscapes, providing valuable insights into the biodiversity and ecological health of these ecosystems.

Due to the large and complex nature of soundscape datasets, manual annotation is too resource-expensive. Machine learning has been proposed as a more computationally efficient method than traditional methods. Previous studies have demonstrated the effectiveness of machine learning algorithms in identifying animal sounds within complex acoustic environments (Aodha et al., 2017; Faiß & Stowell, 2023; Kohlberg et al., 2024; Noda et al., 2019). However, few studies have focused specifically on insect species in Mauritian rain forests, presenting an opportunity to fill this knowledge gap.

1.1 Background

1.1.1 Mauritius

Mauritius is part of the Mascarene Islands, located in the West Indian Ocean to the east of Madagascar. Mauritius has a tropical maritime climate with two seasons: a warm, humid monsoon season extending from November to May and a relatively cool, dry winter from June to October (Mauremootoo et al., 2003). The climate and the isolation of the island have resulted in a unique ecosystem of plants and animals that are endemic to the Mascarene Islands (Republic of Mauritius, 2017).

In Mauritius, 25% of the land is covered by forest. However, most of the forests are severely degraded, with only 2% of the land having forests considered to be 'good quality' (Forestry Service, 2018). Ebony Forest is a private forest with several innovative restoration projects which started in 2006. They aim to restore and save the forest and species that inhabit it and reintroduce the locally extinct species. The development of nature-based tourism implemented by Ebony Forest is used to educate and inspire visitors and is also helping to fund research and restoration (Ebony Forest, n.d.).

Young (2000) suggests that it is not enough to protect the habitat and species of the area if the ecosystem is already degraded. Alongside protecting the environment, effort needs to be put into the restoration of the ecosystem. This makes Ebony Forest and the work being done there an important example of restoring a degraded ecosystem. Styring et al. (2022) suggests that in order to determine if an area has high conservation value, a thorough assessment of the biodiversity of the area needs to be conducted.

The Orthoptera insect order contains insect species known for acoustically communicating. These species, such as crickets and katydids, can be used as indicators of an ecosystem's health (Lehmann et al., 2014). The Mascarene Islands has over 100 of these species, yet they have only been the focus of a few studies (Hugel, 2010; Hugel & Desutter-Grandcolas, 2021; Hugel et al., 2021). None of these studies used acoustic monitoring. In order to assess the progress of reforestation efforts in Mauritius, the environment needs to be monitored over a period of time, at different locations, to ascertain if these species' presence in the environment is stable.

Brachyphisis nattecantor is a species of katydid found mostly on the endemic Sapotacae trees in Mauritius. There is only limited research on this species as it was newly discovered in 2010 by Hugel (2010). The males of this species are found on the higher branches, whilst the females can also be found on lower branches when laying eggs.

Although Hugel (2010) conducted his research on very limited data (8 specimens), he was able to establish that the *Brachyphisis nattecantor* call contains fundamental peaks between 15 to 20 kHz with long sequences of disyllabic chirps rarely interrupted by short breaks. He found that the species sang during night hours. Due to the limited research on this species, there has been a need to extrapolate from research on other species of katydids for this thesis.

1.1.2 Environmental Monitoring

Environmental monitoring allows researchers to understand baseline population levels. The continued monitoring then allows researchers to identify changes in the environment over time. This data can then be used to help monitor vulnerable species and develop effective conservation plans (Dröge et al., 2021; Styring et al., 2022).

Conducting environmental monitoring using traditional methods is time-consuming and expensive, especially in remote areas like Mauritius (Dröge et al., 2021). For invertebrates, there are various traditional techniques for monitoring their population. These include sweep netting and deploying traps such as sticky, pan, pitfall and pheromone traps (Kohlberg et al., 2024). They provide the ability to gain detailed taxonomy, biomass and abundance data. Due to their extensive use, this traditional form of monitoring is efficient and well-understood. These techniques require taxonomic expertise to identify the species that are trapped. Often, these techniques are difficult to standardise and are subject to both observer and sampling bias. The monitoring usually occurs over a short duration and during fine weather, leading to temporal bias. Trapping can also lead to negative effects on the species due to its potential lethality (Kohlberg et al., 2024). Reducing the lethality of environmental monitoring is important due to the vulnerability of species in areas already affected by climate change and human disruption. Rare or isolated species, which researchers are actively trying to monitor, could be significantly impacted by these techniques (Miller et al., 2022).

Due to the dense vegetation and difficult terrain in tropical rainforests, visual observation techniques are often not viable and acoustic methods are considered to be more accurate and cost-effective (Winiarska et al., 2024). Acoustic methods do not have adverse effects on the species due to their noninvasive nature (Browning et al., 2017; Eldridge et al., 2016). One such acoustic monitoring method is Passive Acoustic Monitoring (PAM), which uses acoustic sensors to record and monitor ecosystems (Browning et al., 2017).

1.1.3 Passive Acoustic Monitoring (PAM)

Passive Acoustic Monitoring (PAM) "involves surveying and monitoring wildlife and environments using sound recorders (acoustic sensors)" (Browning et al., 2017, p. 3). The acoustic sensors are comprised of a sound recorder and a microphone. This differs from active acoustic monitoring, which involves the use of sound-emitting devices such as sonar to detect and track species (Stein, 2011).

Microphones, to record species' acoustic presence, can be set up in a variety of ways depending on the information needed and the planned use of the data. Single microphones can be deployed to record the diversity and amount of a species in a particular area. It has been shown particularly effective in detecting the presence of endangered species in remote areas that are difficult to monitor (Blumstein et al., 2011). Stereo microphones and quadraphonic microphone arrays, with the use of localisation algorithms, can determine the geographical position of a sound source and can show species behaviour, territory dynamics, and population density (Blumstein et al., 2011; Browning et al., 2017).

PAM addresses the bias inherent in traditional monitoring methods, such as disturbance due to human interaction and limited standardisation. The acoustic sensors can be deployed in remote locations and left to record data over months (Pérez-Granados & Traba, 2021). While PAM is used as an alternative to traditional methods, sensors are often expensive and the electronics can be vulnerable to damage from weather and animals (Blumstein et al., 2011; Browning et al., 2017; Pérez-Granados & Traba, 2021).

A study by Sugai et al. (2018) found that of 460 articles investigating terrestrial PAM between 1992 and 2018, only 5% of articles focused on invertebrates. Studies have also found a bias towards research in North America and Europe (Kohlberg et al., 2024; Sugai et al., 2018). Over half of the studies investigated by Sugai et al. (2018) manually analysed PAM recordings with only 19% using fully automated analysis. This was mainly in studies focused on bats, as for other species, the fully automated methods are still being developed.

Once the data is collected from the acoustic sensors, it is in a raw waveform signal. This waveform signal needs to be analysed to extract relevant information. One of the most common methods is using Fourier analysis to extract the signal's frequency information. This allows the sounds to be viewed visually in a spectrogram, which is a representation of sound in the time-frequency domain, allowing manual identification and labelling (Browning et al., 2017). The data can be labelled manually by suitably trained experts or using automated methods such as machine learning. Although humans can detect subtle differences in the spectrograms of species vocalisations, the amount of data makes human annotation unrealistic (Blumstein et al., 2011; Browning et al., 2017; Dröge et al., 2021).

1.1.4 Soundscape

A soundscape is composed of a variety of sounds, including biophonies, geophonies, and anthropophonies (Browning et al., 2017). Biophonies are the sounds produced by living organisms, such as animal vocalisation. Geophonies are the sounds produced by non-living natural environments, such as rain, thunder, and wind. Anthropophonies are the sounds produced by humans, such as people talking, car engines, and planes. Studying soundscapes allows researchers to gain insight into an ecosystem's biodiversity and habitat features (Erbe et al., 2015; Nieto-Mora et al., 2024).

Soundscape recordings can be used to monitor many species and environmental factors simultaneously and can be used to measure ecosystem health and help develop conservation plans (Blumstein et al., 2011). Soundscapes can be used to analyse species diversity and population dynamics. They capture the acoustic landscape and show the impact of ecological changes such as habitat loss and the effects of anthropogenic noises (Blumstein et al., 2011; Nieto-Mora et al., 2024).

Copious amounts of data are contained within soundscape recordings, making it difficult to analyse without leveraging automated tools (Blumstein et al., 2011; Nieto-Mora et al., 2024).

1.1.5 Automated Tools

Machine Learning has been investigated as a method of analysing the large amount of data generated from PAM. It has the ability to automatically detect and classify species by the sound they produce. Several studies have investigated different feature extraction methods and different machine learning algorithms across different species.

The main machine learning algorithms used are neural networks (NN), support vector machines (SVM), and random forests (Kohlberg et al., 2024). These models are usually trained on datasets of the sounds of interest and the categories to classify these sounds. The process of using one of these models involves, first, extracting features, such as spectral and temporal characteristics. Then, the model detects sounds of interest and classifies them as one of the categories defined in its training. Usually, the model will produce a probability of how likely it is to be classified in the category the model has classified it as (Browning et al., 2017; Usman et al., 2020).

Deep learning has been found to have high accuracy when classifying complex acoustic signals with minimal pre-processing of the input data (Faiß & Stowell, 2023; Kohlberg et al., 2024). Deep learning requires large training datasets. However, the sound datasets available for training are limited even in the areas they are biased towards, temperate climate and vertebrates (Browning et al., 2017). Deep learning has been shown to have a higher accuracy than other methods, especially with noisy data (Aodha et al., 2017; Browning et al., 2017). Kohlberg et al. (2024) found that over the 176 studies they investigated, deep learning had an average accuracy of 89.1% compared to machine learning, which had an average accuracy of 83.1%. More studies use machine learning than deep learning. However, the accuracy of the machine learning models is often overstated as they were trained specifically for their data set and do not generalise well. Alternatively, deep learning is known for being able to generalise well, which is why, on average, it has a higher accuracy than machine learning models (Aodha et al., 2017; Kohlberg et al., 2024).

There are very few classifiers available for tropical ecosystems with a high biodiversity (Browning et al., 2017). Kohlberg et al. (2024) also found that limited studies have focused on classifying species from tropical climates, with only 9.6% of the 176 studies they investigated focused on data from the tropics. Mauritius's climate is classified as tropical (Mauremootoo et al., 2003), and despite the high degree of biodiversity, there is very limited research on detecting species using machine learning.

1.2 Project Aims and Objectives

Only a limited number of studies have focused on invertebrates or species in tropical climates. The soundscape recordings for this project were collected in 2023 and were yet to be analysed. They contain a large amount of acoustic data of the soundscape of a topical rainforest with a rich biodiversity. This project aims to investigate the *Brachyphisis nattecantor* species found in this ecosystem, which has rarely been studied before in a non-invasive, acoustic manner.

The soundscape recordings from Ebony Forest have a lot of background noise. Deep learning has proven effective at detecting and classifying sounds from noisy data with little pre-processing but requires large training datasets (Aodha et al., 2017; Faiß & Stowell, 2023; Kohlberg et al., 2024). There is a limited number of datasets on the acoustics produced by invertebrates and none on Mauritian invertebrates. This makes identifying species difficult as there is no frame of reference. The manual annotation of data is needed so that there is a sufficient amount of data to train a model in detecting and classifying these species.

This project aims to address the current gap in the literature regarding using machine learning to detect and classify insects in tropical climates such as Mauritius. It aims to build a machine learning model to detect and classify the signal of *Brachyphisis nattecantor* in Ebony Forrest. This detector will be able to be used in the future to analyse the Ebony Forrest soundscape and investigate the occurrence of *Brachyphisis nattecantor* in Ebony Forrest as a bio-indicator of the health and stability of the ecosystem.

2 Methodology

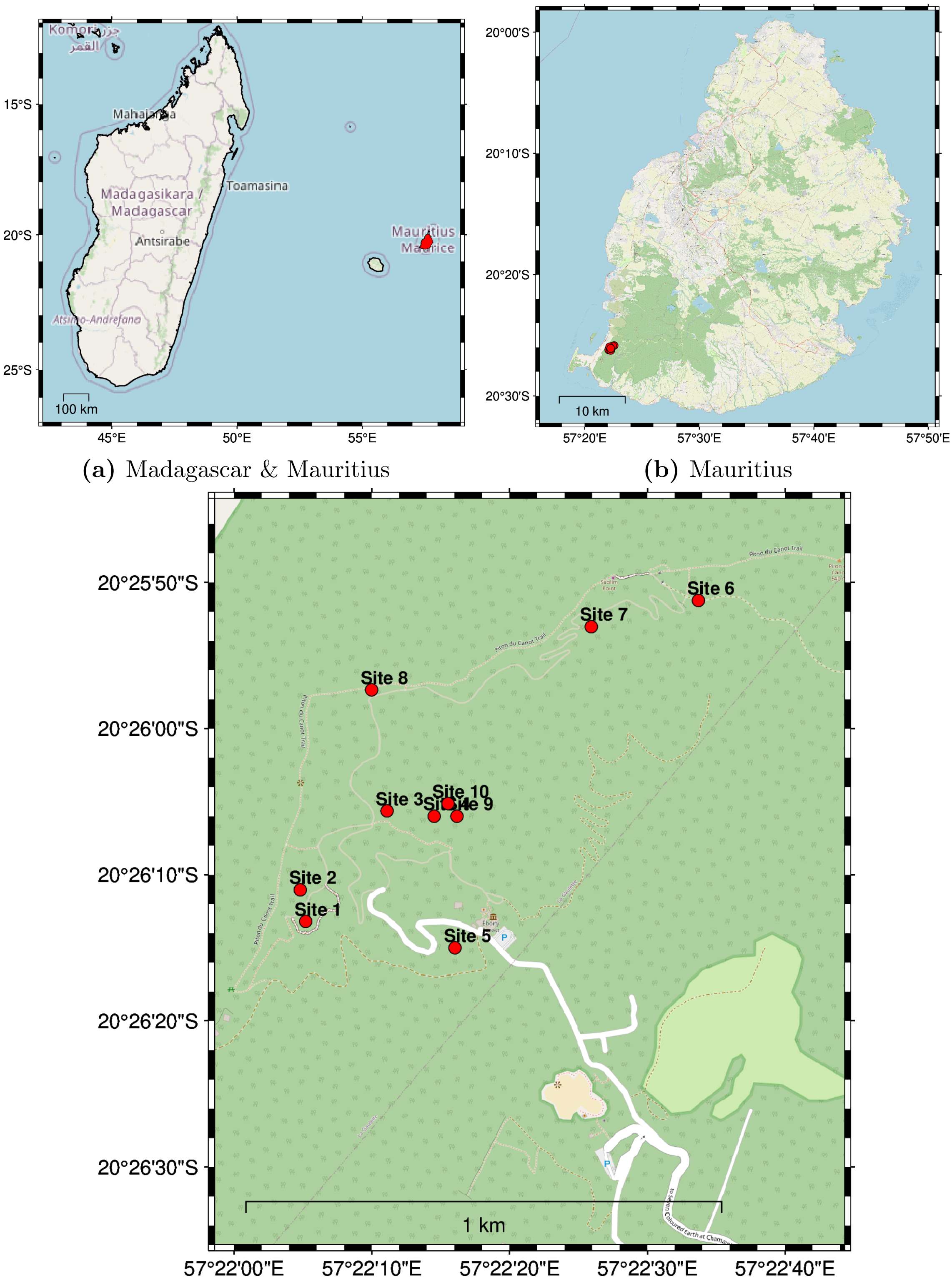
2.1 Data Collection

The Passive Acoustic Monitoring (PAM) data used in this study was collected from ten sites within Ebony Forest, Mauritius. The sites are in a forest that is degraded and in various stages of being rehabilitated. The PAM data contains the acoustic signals of a variety of species, some of which are endangered. Investigating this data allows researchers to better understand the Ebony Forest ecosystem and provide insight on the environmental restoration of the ecosystem. Each of the ten sites had up to eleven deployments from 16 March 2023 to 2 March 2024.

Table 1 shows the latitude and longitude of each of the sites. The sites are either "Weeded" and the invasive plant species have been removed or "Unweeded" and the invasive plant species still remain amongst the endemic species. The weeding is part of the reforestation process currently undergoing within Ebony Forest (Ebony Forest, n.d.). Figure 1 shows a map of the locations of the PAM recording sites. Figure 1 was made using the Python package PyGMT (Tian et al., 2024) and the OpenStreetMap tile (OpenStreetMap Foundation contributors, 2022).

Table 1: Passive Acoustic Monitoring Recording Site Locations

Site	**Latitude**	**Longitude**	**Management**
1	-20.437	57.36811	Weeded
2	-20.4364	57.368	Unweeded
3	-20.4349	57.36975	Weeded
4	-20.435	57.3707	Unweeded
5	-20.4375	57.37112	Weeded
6	-20.4309	57.37603	Unweeded
7	-20.4314	57.37387	Unweeded
8	-20.4326	57.36944	Weeded



(a) Madagascar & Mauritius

(b) Mauritius

(c) Ebony Forest

Figure 1: Map of Passive Acoustic Monitoring Recording Sites Locations

The same recording schedule is used for each Deployment, as noted in Table 2, recording a total of 2 hours and 40 minutes per day.

Song Meter SM4 Acoustic Recorders (Wildlife Acoustics, n.d.) were used to record

PAM data with a sample rate of 96 kHz. Each recording is 10 minutes long, with the duty cycle noted in Table 2.

Table 2: Recording Schedule of PAM in All Deployments

Start Time	End Time	Duty Cycle
12:00 am	12:10 am	ON
12:10 am	5:00 am	OFF
5:00 am	8:00 am	10 minutes ON, 20 minutes OFF
8:00 am	12:00 pm	OFF
12:00 pm	12:10 pm	ON
12:10 pm	4:30 pm	OFF
4:30 pm	8:10 pm	10 minutes ON, 20 minutes OFF
8:10 pm	12:00 am	OFF

Figure 2 shows the timeline for each site's deployments. Sites 4 and 7 were recorded from March to August in 2023 before they were deemed less informative locations and the PAM recorders were moved to Sites 9 and 10 in September 2023 to record until March 2024. Site 2 only recorded from March to late June 2023.

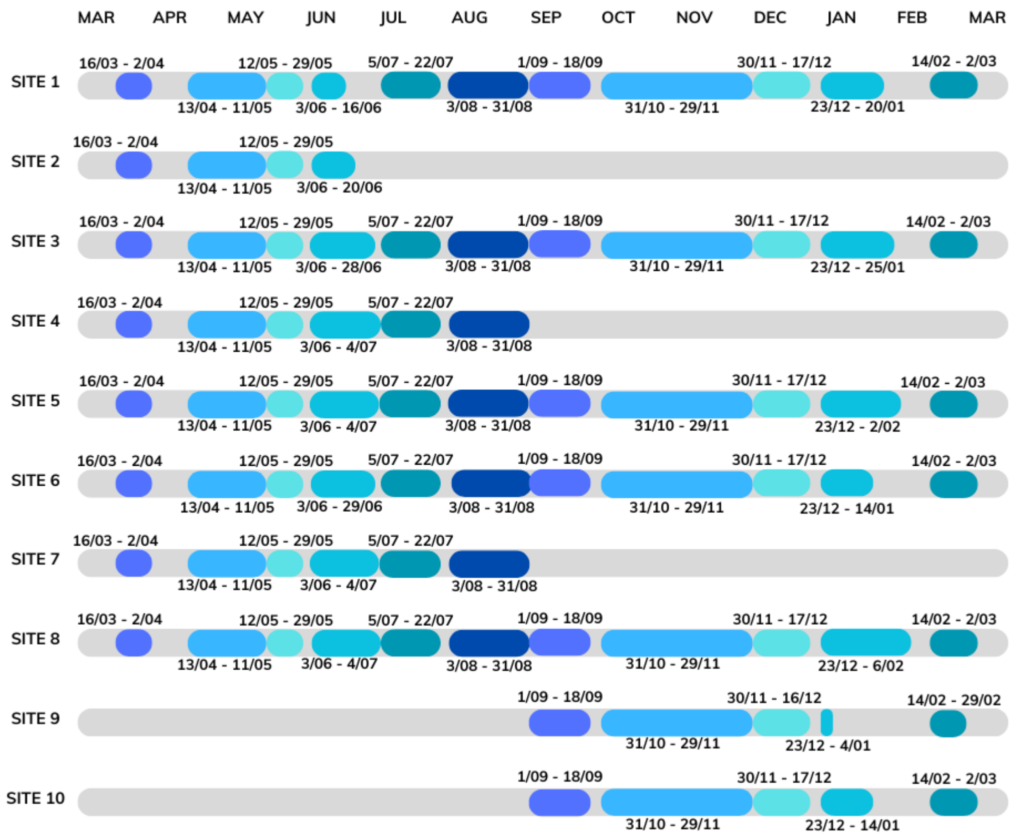


Figure 2: Recording Schedule for Each Site

This thesis focused on data from Site 6. Site 6 was chosen because all deployments were recorded, giving a complete view of the year. Some preliminary annotations identifying species of interest had already been done on Site 6, Deployment 1.

2.2 Manual Annotation of Data

Site 6, Deployment 1 was manually annotated using Raven Pro v.1.6.5 (K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology, 2024). Site 6, Deployment 1 contains 179 files with a total duration of 29 hours and 50 minutes. These manual annotations contained 152 samples of *Brachyphisis nattecantor*.

Brachyphisis nattecantor are the focus of this thesis as katydids are known to be a good indicator of ecosystem health (Lehmann et al., 2014). The *Brachyphisis nattecantor* are a food source for bats and bird species in Ebony Forest and very little research has been conducted on this species (Hugel, 2010). By focusing on this species, this thesis will provide more research into the acoustic activity and patterns of *Brachyphisis nattecantor*.

The measurements of each annotation of *Brachyphisis nattecantor* were then extracted into a table for further analysis. A description of the measurements of interest can be seen in Table 3.

Table 3: Measurements Extracted when Labelling Acoustic Data

Measurement	Description
Begin Time	The start time of the selection
End Time	The end time of the selection
Low Freq	The lower frequency bound of the selection
High Freq	The upper frequency bound of the selection
Begin Path	The file path
Begin File	Name of the sound file the selection is located
Beg File Samp	The starting point for the selection
File Offset	The amount of time between the beginning of the file and the start of the selection
Peak Freq	The frequency at which the maximum power in the selection occurs
Peak Freq Contour	Shows how the peak frequencies in the selection change over time
Peak Time	The first time in the selection at which a sample with the peak amplitude occurs
Bandwidth 90%	The frequency range of the bulk of the call energy
Tags	The annotation column

2.3 Signal Analysis

The data was manually analysed to get a broad understanding of the soundscape of Ebony Forrest. The manual annotations were analysed to find the key features of the *Brachyphisis nattecantor* signal, including the inter-pulse interval, the frequency range, and duration.

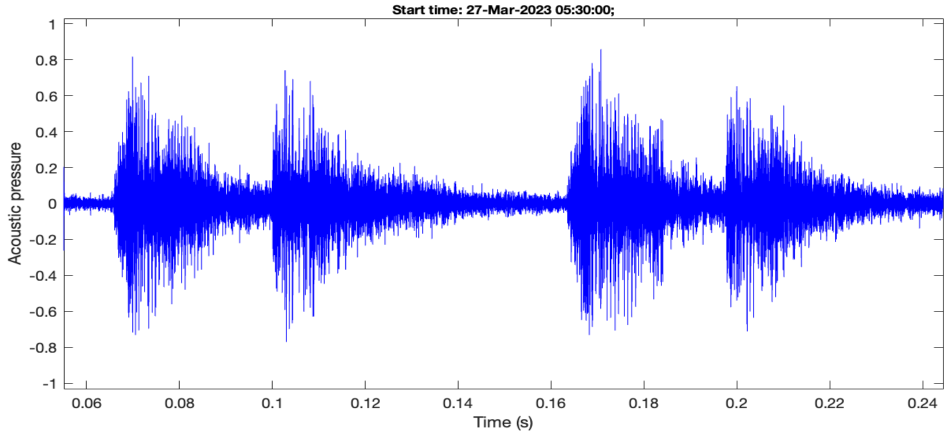
The inter-pulse interval was found through the use of autocorrelation. The signal peaks from this autocorrelation represent key values, including pulse duration and inter-pulse interval.

2.4 Building Detectors

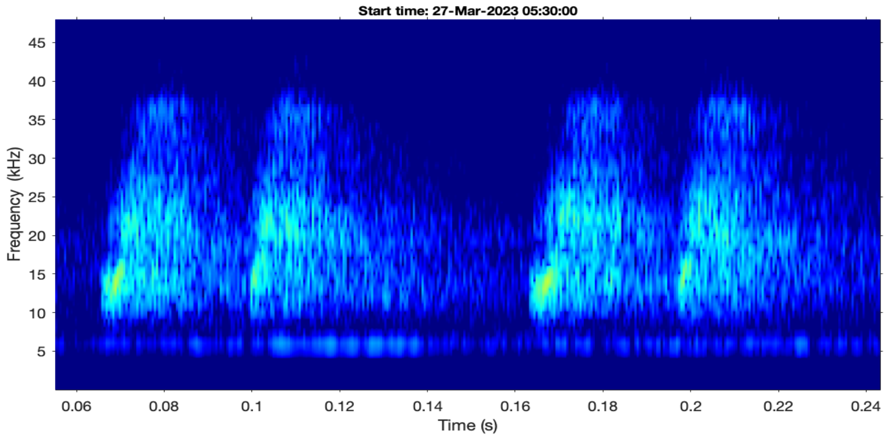
Two detectors were developed using manual annotations and signal analysis to inform parameters. These detectors were developed to speed up the manual labelling of data. They were then compared with the machine learning model developed. The parameters used in the detectors were calculated by investigating their effect on identifying signals in Site 6, Deployment 1. When comparing the detectors with the machine learning model, Site 6, Deployment 2 was used to assess performance.

2.4.1 Correlation Detector

The data was filtered using a 5th order Butterworth high-pass filter of 17 kHz. The template to correlate the signal with was two disyllabic pulses of the *Brachyphisis nattecantor* signal from the file for Site 6, Deployment 1, 27 March 2023 at 05:30 AM with a start time of 0.062 seconds and an end time of 0.225 seconds. This template captured two pulses of the *Brachyphisis nattecantor* signal. The waveform and spectrogram of the template can be seen in Figure 3.



(a) Waveform of Template Signal



(b) Spectrogram of Template Signal

Figure 3: Template Signal Used For Correlation Detector

This template was cross-correlated with the signal and normalised using a normal factor calculated from the maximum value of the autocorrelation of the template. To calculate the cross-correlation SciPy's signal correlation method was used (Virtanen et al., 2020). If this normalised correlation was above the threshold of 0.0125, the *Brachyphisis nattecantor* signal was determined to be present. The threshold 0.0125 was chosen as it had the highest F1-score on the PAM data from Site 6, Deployment 1 of all the thresholds investigated.

2.4.2 Envelope Detector

The data was filtered using a 5th order Butterworth high-pass filter of 17 kHz. The analytical signal was computed using the Hilbert transform (Virtanen et al., 2020). The envelope is the amplitude of the analytical signal. SciPy's 1-D maximum filter (Harter, 2009) followed by SciPy's 1-D uniform filter was used on the result to smooth the curve (Virtanen et al., 2020). This produced an envelope of the curve. If the envelope was above the threshold of 0.003, it was defined as a *Brachyphisis* *nattecantor* signal. The threshold 0.003 was chosen as it had the highest F1-score on the PAM data from Site 6, Deployment 1 of all the thresholds investigated.

2.5 Machine Learning

To develop a machine learning model to detect the *Brachyphisis nattecantor* signal in the soundscape, the Python package Koogu v.0.7.2 (Madhusudhana, 2023) was used. Koogu uses TensorFlow v.2.10.0 (Abadi et al., 2015) as an underlying framework to develop machine learning solutions for bioacoustic applications. The models were developed on an Intel Xeon CPU with eight virtual CPUs and 51GB of RAM and an NVIDIA Tesla T4 GPU with 15GB of RAM. The machine was running Python v.3.10.12.

Various machine learning architectures and hyperparameters were tested to find the optimal parameters for detecting *Brachyphisis nattecantor* signal.

2.5.1 Preparation of Model Inputs

The audio data was preprocessed into clips with a length of 1 second, with 0.4 seconds overlap in clips. The 1-second clip length was chosen as this is shorter than the *Brachyphisis nattecantor* call sequence but long enough to contain multiple disyllabic pulses (8 to 10 per second). The 0.4 overlap was chosen to increase the dataset by a factor of 2.5 and provide a sliding window of the signal. The audio had a sampling rate of 96,000 Hz and was filtered with 5th order Butterworth high-pass filter of 17,000 Hz. The audio clips were also normalised, meaning that the waveform in each clip was scaled to be in the range of [-1.0, 1.0]. Power spectral density spectrograms were computed from the audio clips. The spectrograms were computed with a 6 ms Hann window with 50% overlap between frames. The spectrograms were then clipped to be between 17,000 Hz and 40,000 Hz. This resulted in a time and frequency resolution of 3 ms and 166.67 Hz, respectively. The spectrograms had the dimensions 139 x 332 and were the input for the model. A summary of the settings used to preprocess the audio and produce the spectrograms can be seen in Table 6.

Table 4: Audio and Spectrogram Pre-processing Parameters

Type	Parameters	Value
Audio	Clip Length	1
	Clip Advance	0.4
	Sample Rate (Hz)	96,000
	Filter Audio	5th order, 17 kHz highpass
	Normalise Clips	True
Spectrogram	Window Length (s)	0.006
	Percentage of Window Overlap	50%
	Bandwidth Clip	[17000, 40000]

2.5.2 CNN Architecture

The model trained is a Convolutional Neural Network (CNN) with three layers with 16, 32, and 64 filters per layer. Each convolutional layer applies a 3 x 3 filter with a stride of 1 and padding with a 1 pixel border. ReLu is used as the activation function. The model uses average pooling with a pool size of 2 x 2 and a stride of 2. The output of the final convolutional layer is flattened and inputted into a fully connected layer that uses sigmoid activation.

Several different architectures were investigated including CNNs and DenseNets, as described in Huang et al. (2018). The number of filters and layers, the pooling type, pool size, and pooling stride was investigated. The CNNs had a higher validation accuracy during training then the DenseNets. The final hyperparameters chosen had the highest validation accuracy. A summary of this architecture and the parameters investigated can be seen in Table 5.

Table 5: CNN Model Hyperparameter Values Investigated and Chosen

Hyperparameter	**Values Investigated**	**Value Chosen**
Number of Layers	[2, 8]	3
Number of Filters Per Layer	[2, 64]	16, 32, 64
Filter Size	N\A	3×3
Filter Stride	N\A	1
Filter Padding	N\A	1
Activation Function	N\A	ReLu
Pooling Type	[Max, Average]	Average
Pool Size	[2, 4]	2×2
Pooling Stride	[2, 4]	2
Pooling Padding	N\A	None
Final Layer Activation Function	N\A	Sigmoid

Figure 4 depicts the architecture of the final CNN model. The light orange blocks represent the convolutional layers with the ReLu activation function. The dark orange blocks represent the pooling layers. The light purple block is the fully connected layer with the Sigmoid activation function attached as a dark purple block.The dimensions of each layer can be seen in the figure.

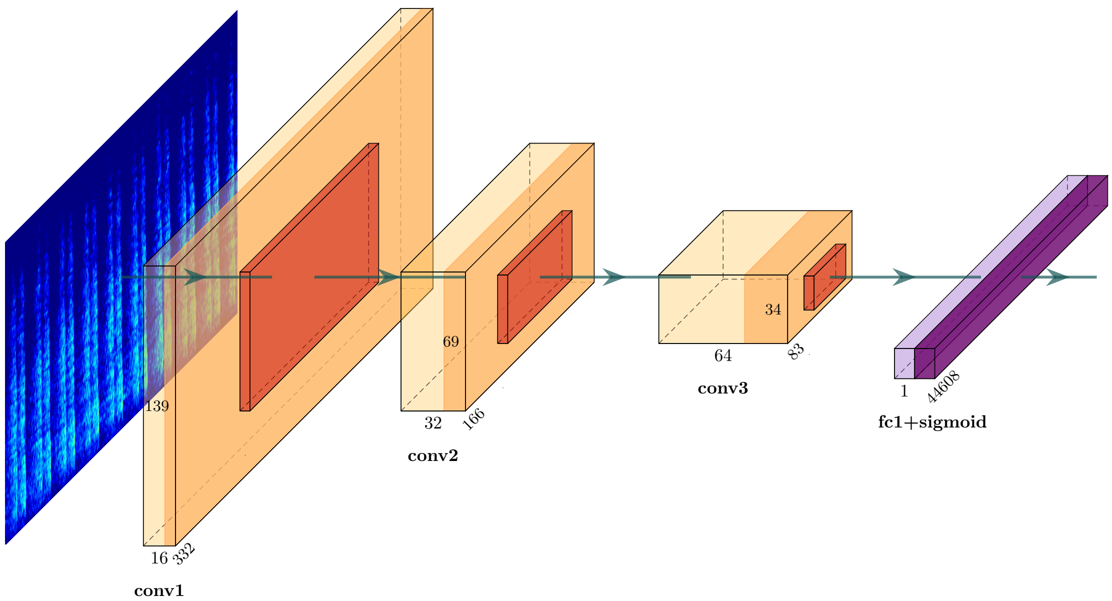


Figure 4: Architecture of CNN Developed

2.5.3 Training

The model was trained on 80% of the annotated data from Site 6, Deployment 1. This is equivalent to 24 hours of data. The files in the training set were randomly chosen from all the files from Site 6, Deployment 1. 15% of the training set was used as a validation set by Koogu during training.

The model was trained for 30 epochs with a batch size of 64 using the Adam optimiser. A dropout rate of 0.05 and an L2 regularisation with a weight decay of 1e-4 were used to improve model generalisation. It had an initial learning rate of 0.01, which decreased to 0.001 at 15 epochs.

Various training hyperparameters were investigated. This included weight decay, dropout, learning rate, and the optimiser. The final hyperparameters used in training the model were the ones that produced the highest validation accuracy during training. A summary of the hyperparameters investigated and values chosen in training the CNN can be seen in Table 6.

Table 6: Hyperparameters Investigated and Chosen to Train CNN

Hyperparameter	**Values Investigated**	**Value Chosen**
Weight Decay	[1e-6, 1e-3]	1e-4
Dropout	[0.0, 0.5]	0.05
Learning Rate	[1e-5, 0.1]	0.01, decrease to 0.001 after 15 epochs
Optimiser	[SGD, Adam]	Adam
Batch Size	[32, 128]	64
Loss Function	N/A	Binary Cross-Entropy Loss

The remaining 20% of the data from Site 6, Deployment 1 was used to test the model and determine a threshold using a Precision-Recall curve. Thresholds from 0.1 to 1 with a step of 0.01 were investigated.

2.5.4 Testing

To test the CNN, 51 files from Site 6, Deployment 2 were annotated. These files spanned from midnight on 14 April 2023 to 6:00 AM on 16 April 2023, containing 8.5 hours of data. The performance was assessed using the metrics described in Section 2.6.

2.5.5 Detections for Other Deployments

The CNN was used to detect the *Brachyphisis nattecantor* signal in Site 6, Deployments 5 and 6. The detections were manually inspected to ascertain how accurately it identified the signal. Manual inspection was used to evaluate the model on other data with different environmental parameters as it is known that katydids' acoustic activity and signal can be effected by temperature and moonlight.

2.6 Performance Assessment

A variety of performance metrics were used to assess the detectors. Each recording was segmented into 1-second long clips. These clips were then analysed so that if they contained a manual annotation and the detector's annotation, they were marked as true positive. If there was only a manual annotation, the segment was marked as a false negative. If there was only a detector annotation, the segment was marked as false positive. If no manual or detector annotation was in the segment, it was marked as a true negative. The count of true negatives, true positives, false positives, and false negatives was used to calculate precision, recall, F1-Score, and Balanced Accuracy.

Precision measures the number of correctly identified positives out of all positives detected. A low precision score means a high number of negative segments have been incorrectly labelled as containing the signal of interest.

$$Precision = \frac{True\ Positives}{True\ Positives\ +\ False\ Positives} \tag{1}$$

Recall measures the number of correctly identified positives out of all real positives. A low recall score means a high number of segments containing the signal of interest were incorrectly identified as a negative detection.

$$Recall = \frac{True\ Positives}{True\ Positives\ +\ False\ Negatives} \tag{2}$$

F1-Score assesses the precision-recall trade-off. It is a useful metric for imbalanced data as it assesses the performance on each class, instead of overall performance.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision\ +\ Recall} \tag{3}$$

Balanced accuracy was used as the dataset was imbalanced, with less than 5% of segments containing the *Brachyphisis nattecantor* signal. Whilst accuracy measures the number of correctly predicted segments out of all predictions, balanced accuracy takes into account the imbalanced dataset. It looks at both the correctly identified positives out of all real positives (sensitivity) and the correctly identified negatives out of all real negatives (specificity) and gives them an equal weighting towards the balanced accuracy.

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{4}$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \tag{5}$$

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2} \tag{6}$$

3 Results

3.1 Signal Analysis

Initial manual signal analysis showed that birds were the main source of noise in the soundscape around dawn and dusk. The *Brachyphisis nattecantor* signal was only detected during the night hours when heavy rain was not present.

Figures 5, 6, and 7 were all created using CHORUS (Gavrilov & Parsons, 2014). Figure 5 depicts the power density spectrogram of the Ebony Forest soundscape on 16 March 2023 from 5:00 AM to 5:10 AM. The figure shows five instances of the *Brachyphisis nattecantor* signal during the ten minutes. The signal is one of the only signals above 15 kHz and in this spectrogram is displayed as columns of sound.

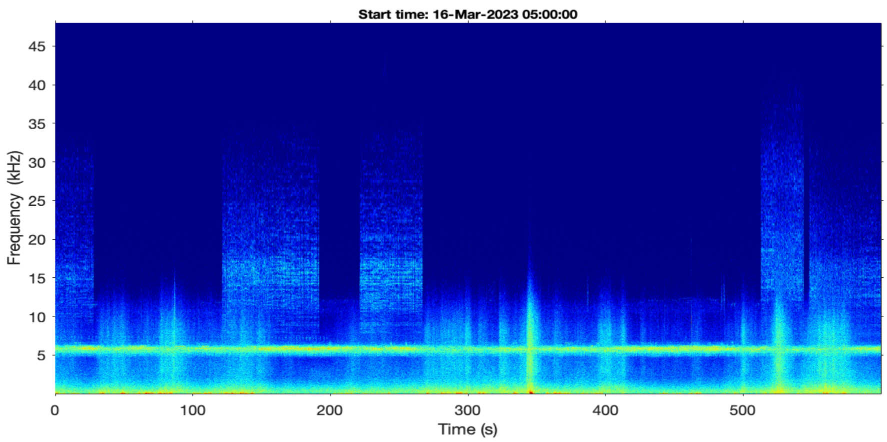
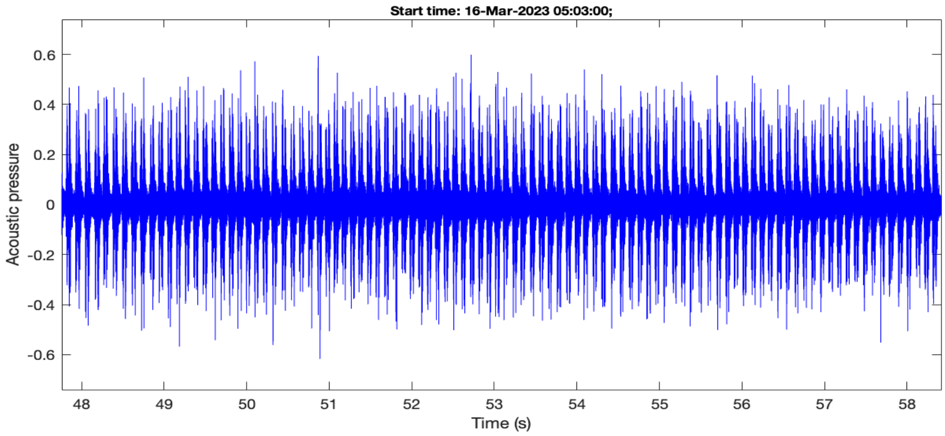
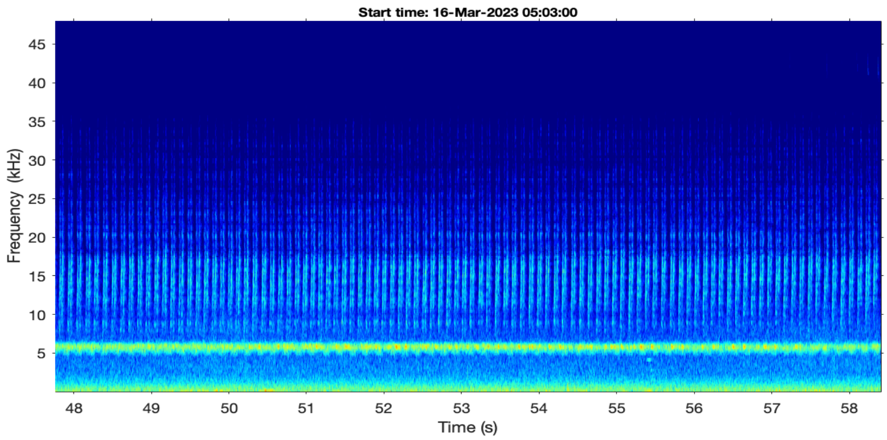


Figure 5: Spectrogram of *Brachyphisis nattecantor* on 16 March 2023 at 5:00 AM to 5:10 AM

Figure 6 depicts the power density spectrogram and waveform of the Ebony Forest soundscape on 16 March 2023 at 5:03 AM for 10 seconds. The figure shows a series of uninterrupted *Brachyphisis nattecantor* calls. Figure 7 depicts the power density spectrogram and waveform of the Ebony Forest soundscape on 16 March 2023 at 5:03 AM for 1 second. The disyllabic nature of the *Brachyphisis nattecantor* pulses can be clearly seen. Figures 6a and 7a have a highpass filter at 12 kHz to filter out other signals and show the *Brachyphisis nattecantor* signal more clearly.

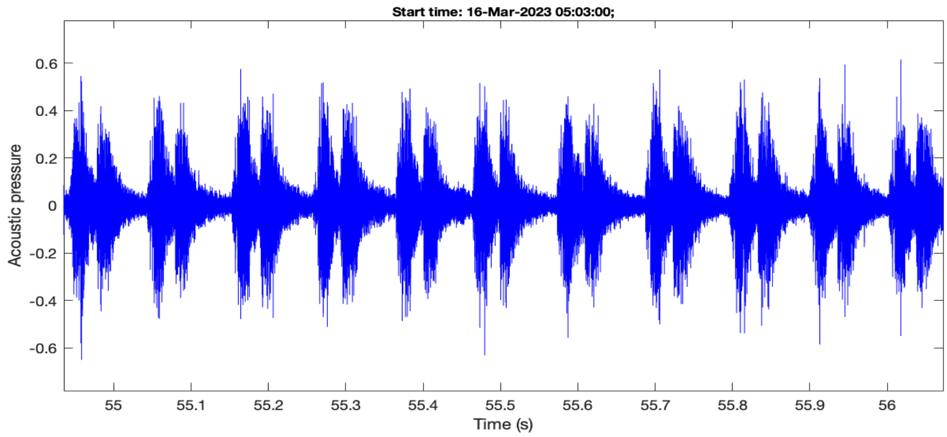


(a) 10 Second Waveform of *Brachyphisis nattecantor*

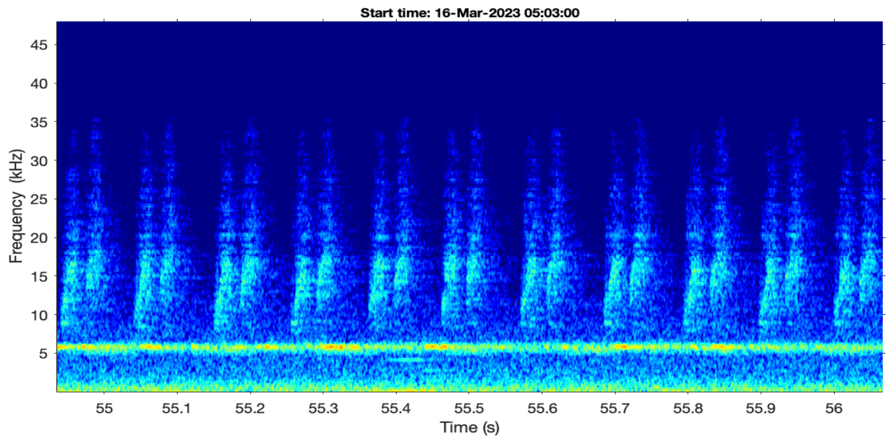


(b) 10 Second Spectrogram of *Brachyphisis nattecantor*

Figure 6: *Brachyphisis nattecantor* Signal for 10 seconds on 16 March 2023 at 5:03 AM



(a) 1 Second Waveform of *Brachyphisis nattecantor*



(b) 1 Second Spectrogram of *Brachyphisis nattecantor*

Figure 7: *Brachyphisis nattecantor* Signal for 1 second on 16 March 2023 at 5:03 AM

The *Brachyphisis nattecantor* signal has an average peak frequency of 13.8 kHz. Table 7 shows a summary of the characteristics, as defined in Table 3, of the *Brachyphisis nattecantor* signal. These statistics were derived from 241 samples of the *Brachyphisis nattecantor* signal.

The *Brachyphisis nattecantor* signal has an average duration of 56.91 seconds but can go for over 10 minutes. The *Brachyphisis nattecantor* was most active at night, with 83.4% of annotations between 7:00 PM and 5:40 AM. All annotations of the *Brachyphisis nattecantor* signal were between 4:30 PM and 5:40 AM. 53.11% of annotations of the *Brachyphisis nattecantor* signal were in the early morning between 12:00 AM and 5:40 AM. Figure 8 depicts a visual representation of the duty cycle of the recordings and when the *Brachyphisis nattecantor* signal is present in the PAM data. As the recordings only covered 11% of each 24 hour period, the presence of *Brachyphisis nattecantor* during the times recordings did not occur cannot be known.

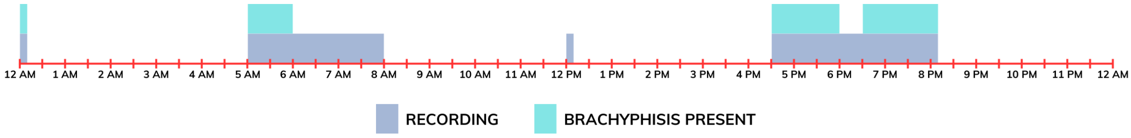


Figure 8: Presence of *Brachyphisis nattecantor* in Site 6 Deployment 1 Recordings Over 24 Hour Period

Table 7: *Brachyphisis nattecantor* Signal Characteristics

	Low Frequency (Hz)	High Frequency (Hz)	Peak Frequency (Hz)	Bandwidth 90% (Hz)	Peak Time (s)	Duration (s)
min	7384.62	28918.34	9187.5	9187.5	0.02	1.83
mean	11970.62	35739.19	13834.54	13496.89	22.48	56.91
max	18263.42	44700.49	24000	21000	547.44	600
std	2349.96	3383.15	2416.4	1523.16	49.54	99.17
25%	10369.89	32986.13	12187.5	12562.5	1.55	10.91
50%	11705.4	35901.8	13687.5	13500	6.62	28.37
75%	13164.87	38258.59	14812.5	14250	23.76	58.98

The *Brachyphisis nattecantor* signal was autocorrelated, and the peaks were extracted to find the disyllabic pulse duration, the interval between disyllabic pulses and the interval between pulses in disyllabic pulses. It was found that at 22 °C, the duration of the disyllabic pulse was 0.107 seconds, the interval between disyllabic pulses was 0.75 seconds, and the interval between the syllables was 0.032 seconds. Figure 9 shows a visual representation of these results.

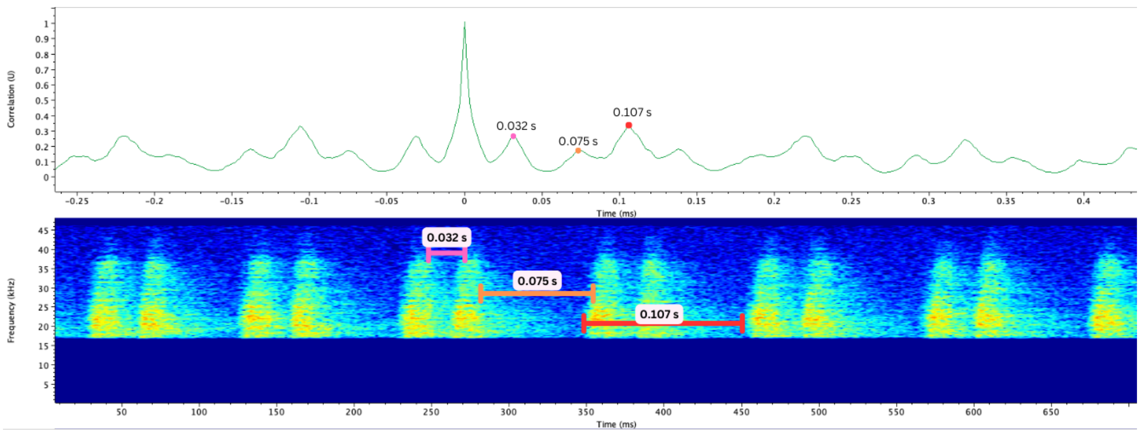


Figure 9: Autocorrelation of *Brachyphisis nattecantor* Signal (top) and Representation of Peaks of Autocorrelation on Spectrogram (bottom)

3.2 Machine Learning

Various thresholds were investigated to find the best threshold for detecting *Brachyphisis nattecantor*. After analysing the Precision-Recall curve (Figure 10) and F1-score (Figure 11) for varying thresholds, a threshold of 0.9 was chosen.

There was a high recall, greater than 0.89, for all thresholds investigated. The highest precision was 0.98 at a threshold of 0.99. The 0.99 threshold had a recall of 0.8949 and an F1-score of 0.939. This was the lowest recall of all thresholds.

The greatest F1-score was 0.9753 at a threshold of 0.88.

A threshold of 0.9 had an F1-score of 0.9748, a precision of 0.975 and recall of 0.9746.

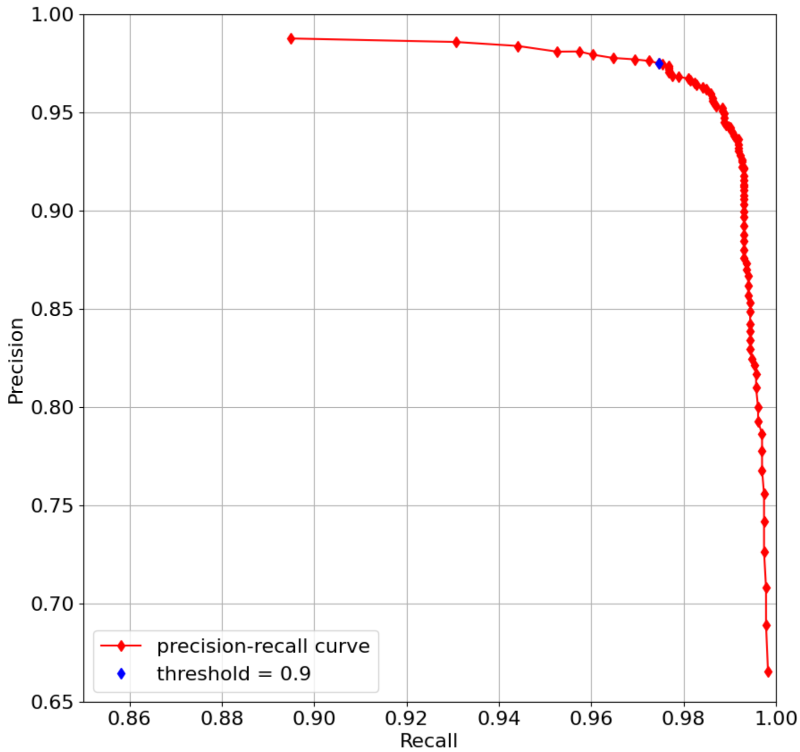


Figure 10: Precision-Recall Curve of CNN on Testing Dataset

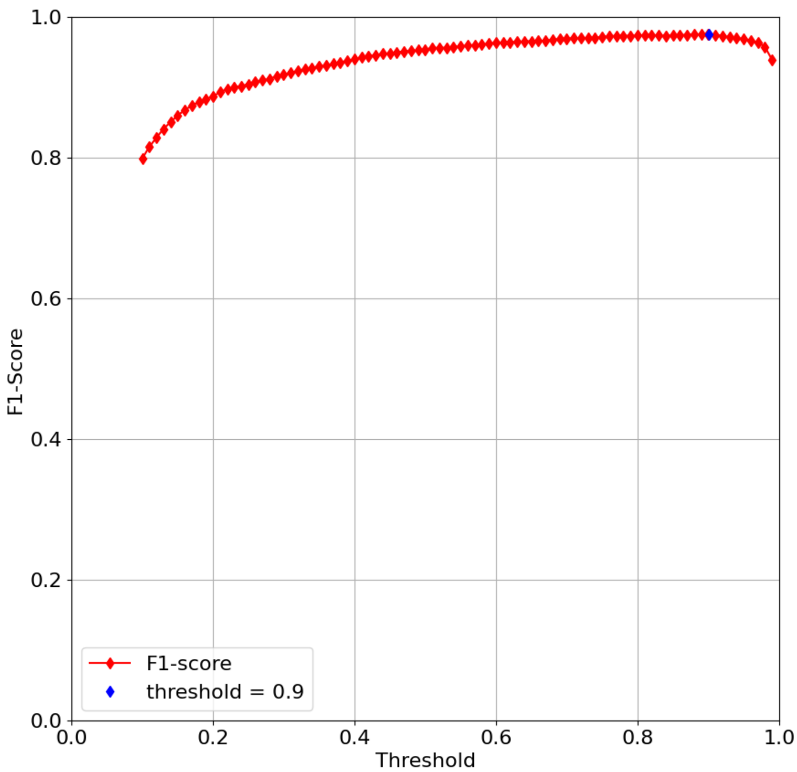
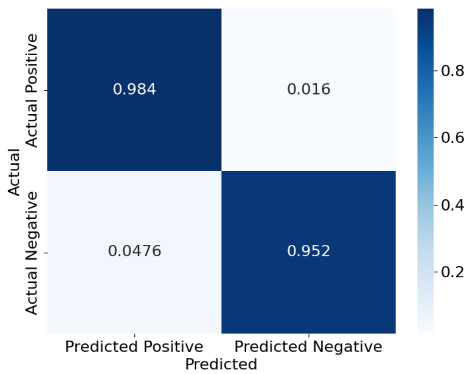


Figure 11: F1-Score for Thresholds for CNN on Testing Dataset

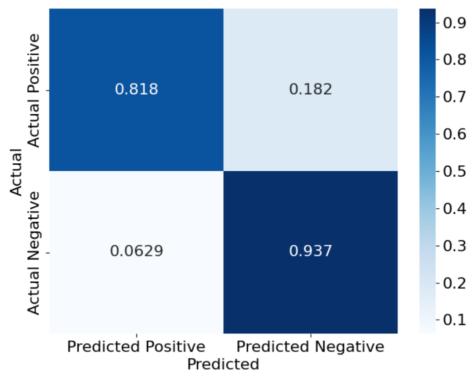
When tested on the 51 files from Site 6, Deployment 2, the CNN had a precision of 0.7108, a recall of 0.9933, and an F1-score of 0.8286. It had a balanced accuracy of 0.9863. The CNN was run on the data from Deployments 5 and 6 and it found that there were 1,085 and 3,829 detections respectively. Manual inspection of these detections showed that no *Brachyphisis nattecantor* signal was detected in Deployment 5, meaning all 1,085 detections were false positives. In Deployment 6, less than 50 were true positives, however there were no false negatives that could be found.

3.3 Comparison of Detectors

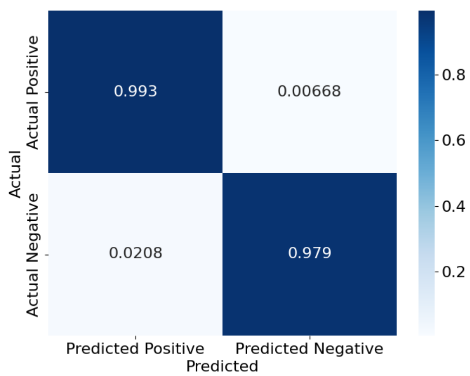
The detectors were tested on 51 files from Site 6, Deployment 2. Figure 12 shows the confusion matrices for each detector. Table 8 shows the results for precision, recall, F1-score, and balanced accuracy. The Envelope Detector had the lowest scores across all performance metrics. The CNN had the highest scores across all performance metrics.



(a) Correlation Detector



(b) Envelope Detector



(c) CNN

Figure 12: Confusion Matrices for Detectors

Table 8: Comparison of Detectors

Scores	Correlation Detector	Envelope Detector	CNN Model
Precision	0.5154	0.4007	0.7108
Recall	0.9840	0.8176	0.9933
F1-Score	0.6765	0.5378	0.8286
Balanced Accuracy	0.9682	0.8774	0.9863

4 Discussion

This thesis developed three detectors for the *Brachyphisis nattecantor* signal in the Ebony Forest soundscape. Each detector used a different method: signal cross-correlation, envelope, and machine learning. The performance of the detectors were investigated to find which detector would be most suited for detecting *Brachyphisis nattecantor*. The *Brachyphisis nattecantor* signal was analysed to help build on the limited existing research of this species' bioacoustics.

Hugel (2010) discovered the *Brachyphisis nattecantor* in 2010 and conducted research on eight specimens, with only three specimens used for the bioacoustic results. The recordings of the specimens were conducted in a lab at above 20 °C and, to date, is the only research conducted on this species' bioacoustics. Hugel (2010) found that the duration of a disyllabic pulse was 48 - 85 ms with an average of 66.3 ms, and the interval between syllables in a disyllabic pulse was between 29 - 63 ms, with an average of 46.1 ms. Hugel (2010) also found that the fundamental peaks were between 15 - 20 kHz. In this research Hugel (2010) does not measure the interval between disyllabic pulses and instead measures how many pulses in a second.

The signal analysis conducted for this thesis found that the *Brachyphisis nattecantor* signal was only detected between 4:30 PM and 5:40 AM, which aligns with what Hugel (2010) found with *Brachyphisis nattecantor* "singing by night hours". The results of the autocorrelation of the *Brachyphisis nattecantor* signal found that the interval between syllables was 32 ms at 22 °C., which also aligns with Hugel (2010). The duration of a disyllabic pulse was 107 ms, and the interval between pulses was 75 ms (See Figure 9). As shown in Table 7, the signal analysis found that the peak frequency was between 9.2 and 24 kHz, with a mean of 13.8 kHz.

The interval between syllables in the disyllabic pulses aligns with what Hugel (2010) found in his research. The difference between the duration of the disyllabic pulse and Hugel (2010) could be due to the way measurement was conducted, with this thesis possibly measuring both duration and interval. Autocorrelation is not usually used to calculate the duration and was mainly used in this thesis to find the intervals. However, a third peak in the autocorrelation was found and upon investigation was matched to the duration of the signal.

The signal analysis in this thesis, found the minimum of the peak frequency of the *Brachyphisis nattecantor* signal is 9.2 kHz, which is lower than expected. However, this is mainly due to the signal overlapping with another louder signal at 9 kHz. The duration of the signal ranged from 1.83 seconds to 10 minutes. The lower duration signals annotated were usually part of a longer signal; however, they were broken up into separate annotations due to a break in the signal for longer than 1 second. This was due to the annotations being used as the ground truth for the performance metrics, which used 1-second long clips.

In this study, all the detectors and the model had higher recall than precision. This means they were good at identifying positive signals but sometimes incorrectly marked a negative signal as positive. They were more likely to mark a negative signal as positive than a positive as negative.

The detectors' parameters were calculated by investigating the effect they had on identifying signals in Site 6, Deployment 1. Their performance was assessed on Site 6, Deployment 2 to compare with the machine learning model. Testing on another deployment ensured the results would show if the parameters were overfitted for Deployment 1. The correlation detector had a higher precision and F1-score on Deployment 1 but a lower recall and balanced accuracy. The envelope detector performed better on Deployment 1 for all 4 metrics but had more false positives. This shows that the envelope detector's parameters produced reasonable results on the deployment they were configured for but not on other deployments. The correlation detector was expected to have a higher precision based on its results on Deployment 1. This could be due to the template chosen having a variable that is featured more in Deployment 1 than Deployment 2.

The envelope detector performed the worst out of the three detectors, with the lowest score across all performance metrics (See Table 8). Both the envelope and correlation detectors had a precision lower than 0.52. There was a high proportion of negative signals to positive signals in the data. The low precision means that a large number of negative signals were identified as positive. The correlation and envelope detectors identified 4.76% and 6.29% of negatives as positive, respectively. Overall, the correlation detector had 1,385 false positives, and the envelope detector had 1,831 false positives. The data had 1,497 actual positive signals, which means the envelope detector detected more false positives than there were actual positives in the data. This shows that the envelope detector was not accurate.

In the soundscape, several signals were similar to the *Brachyphisis nattecantor* signal. The main differences were that they were only single pulses and not disyllabic. They also had a larger interval between pulses. These signals were centred around a lower frequency with an upper frequency bound of approximately 19 kHz compared to the *Brachyphisis nattecantor* signal, which had an average upper frequency bound of 35.74 kHz. To limit the detectors from identifying these other signals as *Brachyphisis nattecantor*, a high-pass filter of 17 kHz was used to filter out as much of the other signals as possible. This filter could not be at a higher frequency as while the *Brachyphisis nattecantor* signal is present well above 17 kHz, the bulk of the energy call is between 12 kHz and 25 kHz.

The envelope detector had trouble differentiating between these other signals and the *Brachyphisis nattecantor* signal, as when the other signals were present over 17 kHz, they had a high amplitude. The envelope detector also incorrectly classified rain as *Brachyphisis nattecantor*, showing that any signal over 17 kHz with a high enough amplitude was incorrectly classified.

The correlation detector had fewer false positives than the envelope detector, making it slightly better at identifying *Brachyphisis nattecantor* signals. This is due to correlation being a measure of similarity between two signals and takes into account more than just the amplitude of the signals, unlike the envelope detector which only measures the amplitude of the signal. The correlation detector used a template of the *Brachyphisis nattecantor* signal, which means instances of the *Brachyphisis nattecantor* signal will have a higher correlation than other signals. The low precision score is most likely due to the template only containing two pulses of the signal, leading to the other similar signals producing a high correlation with the template. The correlation detector correctly identifies rain as negative. It has a higher recall than the envelope detector, meaning it is better at detecting the *Brachyphisis nattecantor* signal when it is present.

The machine learning model developed had a balanced accuracy of 98.63% and an F1-score of 0.8286, making it the most accurate of the detectors developed for this thesis. The precision was 0.7108, meaning that while it performed better than the other detectors, it still had a large number of false positives, with 605 false positives.

A large number of annotations are needed to train a machine learning model. There were 21,402 *Brachyphisis nattecantor* one-second clips and 193,594 one-second clips of other signals after preprocessing. The more data to train the model, the more accurate the results will be. However, this can cause overfitting, which is when the model has a high accuracy on the training set but not on the testing set.

Annotating the PAM data collected for this study had several difficulties. Unlike birds and bats, there is a very small number of datasets of acoustic data of katydids and crickets. These datasets are all focused on insect species from America, Europe, and Australia. There is no labelled acoustic data of *Brachyphisis nattecantor* or any of the other katydids endemic to the Mascarene Islands. This made it difficult to find examples of *Brachyphisis nattecantor* signal to aid as a reference in the process of annotating the data collected for this study. The data collected for this study is currently the only labelled acoustic data of *Brachyphisis nattecantor*.

The model was used to detect signals in Site 6, Deployments 5 and 6, to ascertain its performance on PAM data from different times of the year. These detections were then manually checked. The detections for Deployments 5 and 6 were investigated as the temperature in these deployments was colder than Deployments 1 and 2. Deployments 5 and 6 had a low of 16.5 °C and a high of 27 °C, whilst Deployments 1 and 2 had a low of 19.5 °C and a high of 33 °C. The model detected 1,085 *Brachyphisis nattecantor* signal in Deployment 5. However, on manual inspection, it was found that all of these were false positives, and the *Brachyphisis nattecantor* signal was not found in Deployment 5. In Deployment 6, the model detected 3,829 detections. Upon manual inspection of these detections, roughly less than 50 were actual *Brachyphisis nattecantor*. This shows that although the *Brachyphisis nattecantor* is present in colder temperatures, they are not as prevalent in the soundscape. The false positives in Deployments 5 and 6 were a mixture of the similar signals previously described, bats echolocation, and chirps from birds that covered several frequency bands.

There may be many other variables effecting the acoustic activity of the *Brachyphisis nattecantor* and therefore the ability of the machine learning model to detect the signal. Research has found that the katydids' signals are effected by moonlight (Gomez-Morales & Acevedo-Charry, 2022; Symes et al., 2024). It has also been found that temperature, both nightly minimum and maximum had an effect on the acoustic activity of katydids, with an increase in signalling during warmer nights (Gomez-Morales & Acevedo-Charry, 2022). Temperature can also effect the signals frequency (Edes, 1899). To ensure that the machine learning model is accurate in the colder temperatures, it will need to be trained on annotations from the colder deployments.

To improve the model's accuracy, several steps can be taken. To increase the accuracy on PAM data from colder deployments, such Deployments as 5 and 6, the model could be trained on Deployment 5 with all signals identified as negative, as it has already been ascertained the *Brachyphisis nattecantor* signal is not present in Deployment 5's data. A way to improve the model's precision score would be to train it on other instances of incorrectly labelled signals, specifically the other similar signals to *Brachyphisis nattecantor* labelled as negative.

Passive Acoustic Monitoring (PAM) allow researchers to non-invasively monitor an environment and record acoustic data. The acoustic data can then be used by automated detectors, like machine learning models, to accurately detect a species' acoustic presence in the soundscape. The combination of PAM and machine learning is more accurate than traditional methods like sweep netting and trapping. However, it should be noted that machine learning can also be computationally intensive. This means the hardware the model is running on needs to have enough RAM, CPUs and GPUs to store the data and make the computations required for machine learning.

5 Conclusion

For this thesis, the CNN developed was the most accurate of the detectors, with a balanced accuracy of 98.63% and an F1-score of 0.8286. The results of the signal analysis conducted on the *Brachyphisis nattecantor* signal supports Hugel (2010) findings. Future work needs to be done to improve the precision of the model by training it on more instances of negative signals, and the soundscape in colder temperatures.

Initial investigation into the presence of *Brachyphisis nattecantor* in the soundscape shows that there is a significant decrease in its presence in the colder, drier months featured in the data. The acoustic activity of *Brachyphisis nattecantor* in the recordings of the Ebony Forest soundscape should be further investigated to find what causes this variation. The CNN developed can be used by researchers to help speed up the annotating and detections of *Brachyphisis nattecantor* in the data. The CNN can be used across the different sites and deployments to detect *Brachyphisis nattecantor* and investigate the correlation between their presence, the presence of endemic plant species and reforestation efforts.

This project will help researchers monitor *Brachyphisis nattecantor* population in Ebony Forest as bioindicators of environmental health, contribute to biodiversity conservation efforts, and advance the field of bioacoustics and machine learning.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., ... Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems [Computer software]. https://www.tensorflow.org/

Aodha, O. M., Gibb, R., Barlow, K. E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G. R., Newson, S. E., Pandourski, I., Parsons, S., Russ, J., Szodoray-Paradi, A., Szodoray-Paradi, F., Tilova, E., Girolami, M., Brostow, G., & Jones, K. E. (2017). Bat detective - deep learning tools for bat acoustic signal detection. *bioRxiv*. https://doi.org/10.1101/156869

Blumstein, D., Mennill, D., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J., Krakauer, A., Clark, C., Cortopassi, K., Hanser, S., McCowan, B., Ali, A., & Kirschel, A. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus. *Journal of Applied Ecology*, *48*, 758–767. https://doi.org/10.1111/j.1365-2664.2011.01993.x

Browning, E., Gibb, R., Glover-Kapfer, P., & Jones, K. (2017). Passive acoustic monitoring in ecology and conservation. https://doi.org/10.13140/RG.2.2.18158.46409

Dröge, S., Martin, D. A., Andriafanomezantsoa, R., Burivalova, Z., Fulgence, T. R., Osen, K., Rakotomalala, E., Schwab, D., Wurz, A., Richter, T., & Kreft, H. (2021). Listening to a changing landscape: Acoustic indices reflect bird species richness and plot-scale vegetation structure across different land-use types in north-eastern madagascar. *Ecological Indicators*, *120*, 106929. https://doi.org/10.1016/j.ecolind.2020.106929

Ebony Forest. (n.d.). *Conservation*. https://www.ebonyforest.com/about-us/conservation/

Edes, R. T. (1899). Relation of the chirping of the tree cricket (Oecanthus niveus) to temperature. *American Naturalist*, *33*(396), 935–938.

Eldridge, A., Casey, M., Moscoso, P., & Peck, M. (2016). A new method for ecoacoustics? toward the extraction and evaluation of ecologically-meaningful soundscape components using sparse coding methods. *PeerJ*, *4*, e2108. https://doi.org/10.7717/peerj.2108

Erbe, C., Verma, A., McCauley, R., Gavrilov, A., & Parnum, I. (2015). The marine soundscape of the perth canyon. *Progress in Oceanography*, *137*, 38–51. https://doi.org/10.1016/j.pocean.2015.05.015

Faiß, M., & Stowell, D. (2023). Adaptive representations of sound for automatic insect recognition. *PLOS Computational Biology*, *19*(10), 1–22. https://doi.org/10.1371/journal.pcbi.1011541

Forestry Service. (2018). *Annual report of the forestry service*. Mauritius. https://forestry.govmu.org/Documents/Annual%20Report/(1)%20ANNUAL%20%20REPORT%20%20OF%20%20THE%20%20FORESTRY%20%20SERVICE%202018.pdf

Gavrilov, A., & Parsons, M. (2014). A matlab tool for the characterisation of recorded underwater sound (chorus). *Acoustics Australia / Australian Acoustical Society*, *42*(3), 190–196.

Gomez-Morales, D. A., & Acevedo-Charry, O. (2022). Satellite remote sensing of environmental variables can predict acoustic activity of an orthopteran assemblage. *PeerJ*, *10*, e13969. https://doi.org/10.7717/peerj.13969

Harter, R. (2009). The minimum on a sliding window algorthim. http://www.richardhartersworld.com/cri/2001/slidingmin.html

Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2018). Densely connected convolutional networks. https://arxiv.org/abs/1608.06993

Hugel, S. (2010). New and little known predatory katydids from mascarene islands (ensifera: Meconematinae and hexacentrinae). *Zootaxa*, *2543*, 1–30. https://doi.org/10.11646/zootaxa.2543.1.1

Hugel, S., & Desutter-Grandcolas, L. (2021). New intertidal crickets from Comoros and Mascarene islands (Orthoptera: Trigonidiidae: Nemobiinae: Burcini). *Zootaxa*, *4995*(1), 1–26. https://doi.org/10.11646/zootaxa.4995.1.1

Hugel, S., Warren, B. H., & Desutter-Grandcolas, L. (2021). The Phalangopsidae crickets (Orthoptera, Grylloidea) of the Seychelles Archipelago: Taxonomy of an ecological radiation. *Zootaxa*, *5047*(3), 201–246. https://doi.org/10.11646/zootaxa.5047.3.1

K. Lisa Yang Center for Conservation Bioacoustics at the Cornell Lab of Ornithology. (2024). *Raven Pro: Interactive Sound Analysis Software* (Version 1.6.5) [Computer software]. Ithaca, NY: The Cornell Lab of Ornithology. https://www.ravensoundsoftware.com/

Kohlberg, A. B., Myers, C. R., & Figueroa, L. L. (2024). From buzzes to bytes: A systematic review of automated bioacoustics models used to detect, classify and monitor insects. *Journal of Applied Ecology*, *00*, 1–13. https://doi.org/10.1111/1365-2664.14630

Lehmann, G., Frommolt, K.-H., Lehmann, A., & Riede, K. (2014). Baseline data for automated acoustic monitoring of orthoptera in a mediterranean landscape, the hymettos, greece. *Journal of Insect Conservation*, *18*. https://doi.org/10.1007/s10841-014-9700-2

Madhusudhana, S. (2023). *Koogu* (Version 0.7.2). Zenodo. https://doi.org/10.5281/zenodo.8254287

Mauremootoo, J. R., Watt, I., & Florens, F. B. V. (2003). *Mauritius biodiversity* [Unpublished Book Chapter produced for Conservation International's State of the Hotspots publication: Madagascar & Indian Ocean Islands].

Miller, Z. J., Lynn, A., Oster, C., Piotter, E., Wallace, M., Sullivan, L. L., & Galen, C. (2022). Unintended consequences? lethal specimen collection accelerates with conservation concern. *American Entomologist*, *68*(3), 48–55. https://doi.org/10.1093/ae/tmac057

Nieto-Mora, D. A., Ferreira de Oliveira, M. C., Sanchez-Giraldo, C., Duque-Muñoz, L., Isaza-Narváez, C., & Martínez-Vargas, J. D. (2024). Soundscape characterization using autoencoders and unsupervised learning. *Sensors*, *24*(8). https://doi.org/10.3390/s24082597

Noda, J. J., Travieso-González, C. M., Sánchez-Rodríguez, D., & Alonso-Hernández, J. B. (2019). Acoustic classification of singing insects based on mfcc/lfcc fusion. *Applied Sciences*, *9*(19). https://doi.org/10.3390/app9194097

OpenStreetMap Foundation contributors. (2022). Openstreetmap.mapnik tile. https://www.openstreetmap.org/copyright

Pérez-Granados, C., & Traba, J. (2021). Estimating bird density using passive acoustic monitoring: A review of methods and suggestions for further research. *Ibis*, *163*(3), 765–783. https://doi.org/10.1111/ibi.12944

Republic of Mauritius. (2017). *National biodiversity strategy and action plan 2017 – 2025*. Mauritius. https://www.cbd.int/doc/world/mu/mu-nbsap-v2-en.pdf

Stein, P. J. (2011). Active acoustic monitoring systems for detecting, localizing, tracking, and classifying marine mammals and fish. *The Journal of the Acoustical Society of America*, *129*, 2369. https://doi.org/10.1121/1.3587669

Styring, A., Unggang, J., Ragai, R., Kueffner, K., Froehlich, D., Megom, N., Joseph, L., Jukie, A., Tarang, M., Nazrin, M., Sulok, K., Sekina, K., Setia, L., Giannone, L., Aron, B., Swartz, N., Hyde, P., & Tyler, B. (2022). Determining high conservation values in production landscapes: Biodiversity and assessment approaches. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.783794

Sugai, L. S. M., Silva, T. S. F., Ribeiro, J. W., & Llusia, D. (2018). Terrestrial passive acoustic monitoring: Review and perspectives. *BioScience*, *69*(1), 15–25. https://doi.org/10.1093/biosci/biy147

Symes, L. B., Madhusudhana, S., Martinson, S. J., Geipel, I., & ter Hofstede, H. M. (2024). Multi-year soundscape recordings and automated call detection reveals varied impact of moonlight on calling activity of neotropical forest katydids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1904), 20230110. https://doi.org/10.1098/rstb.2023.0110

Tian, D., Uieda, L., Leong, W. J., Fröhlich, Y., Schlitzer, W., Grund, M., Jones, M., Toney, L., Yao, J., Magen, Y., Tong, J.-H., Materna, K., Belem, A., Newton, T., Anant, A., Ziebarth, M., Quinn, J., & Wessel, P. (2024). *PyGMT: A Python interface for the Generic Mapping Tools* (Version 0.13.0). Zenodo. https://doi.org/10.5281/zenodo.13679420

Usman, A. M., Ogundile, O. O., & Versfeld, D. J. J. (2020). Review of automatic detection and classification techniques for cetacean vocalization. *IEEE Access*, *8*, 105181–105206. https://doi.org/10.1109/ACCESS.2020.3000477

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wildlife Acoustics. (n.d.). Song meter sm4 acoustic recorder. https://www.wildlifeacoustics.com/products/song-meter-sm4

Winiarska, D., Pérez-Granados, C., Budka, M., & Osiejuk, T. S. (2024). Year-round vocal activity of two african barbet species. *Emu - Austral Ornithology*, *124*(3-4), 221–231. https://doi.org/10.1080/01584197.2024.2352380

Young, T. P. (2000). Restoration ecology and conservation biology. *Biological Conservation*, *92*(1), 73–83. https://doi.org/10.1016/S0006-3207(99)00057-9