



ORIGINAL RESEARCH

Seismic savanna: machine learning for classifying wildlife and behaviours using ground-based vibration field recordings

Alexandre Szenicer¹ , Michael Reinwald², Ben Moseley³, Tarje Nissen-Meyer¹ , Zachary Mutinda Muteti⁴, Sandy Oduor⁴, Alex McDermott-Roberts², Atilim G. Baydin^{3,5} & Beth Mortimer²

¹Department of Earth Sciences, University of Oxford, Oxford, UK

²Department of Zoology, University of Oxford, Oxford, UK

³Department of Computer Science, University of Oxford, Oxford, UK

⁴Mpala Research Centre, Nanyuki, Kenya

⁵Department of Engineering Science, University of Oxford, Oxford, UK

Keywords

African elephant, machine learning, seismic waves, wildlife monitoring

Correspondence

Tarje Nissen-Meyer, Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK. Tel/Fax: +44 1865 282149; E-mail: tarje.nissen-meyer@earth.ox.ac.uk

Editor: Vincent Lecours

Associate Editor: Francesco Rovero

Received: 1 June 2021; Revised: 9 September 2021; Accepted: 13 September 2021

doi: 10.1002/rse2.242

Introduction

The ecosystems on our planet are facing an existential crisis, which authors increasingly call the sixth extinction (Barnosky et al., 2004; Ceballos et al., 2020). In mere decades, hundreds of species have become either endangered or extinct because of human activity, either directly through hunting, or indirectly through side effects such as ocean acidification and spread of invasive species (Ceballos et al., 2020). For instance, about 90 amphibian species have disappeared in the last few decades, with a total of about 500 species on the decline (Scheele et al., 2019), or worldwide declines of coral reefs that support some of the most diverse ecosystems in otherwise barren tropical waters (Bellwood et al., 2004). Mammals, too, are adversely impacted and increasingly jeopardised, and many hold the unenviable

Abstract

We develop a machine learning approach to detect and discriminate elephants from other species, and to recognise important behaviours such as running and rumbling, based only on seismic data generated by the animals. We demonstrate our approach using data acquired in the Kenyan savanna, consisting of 8000 h seismic recordings and 250 k camera trap pictures. Our classifiers, different convolutional neural networks trained on seismograms and spectrograms, achieved 80%–90% balanced accuracy in detecting elephants up to 100 m away, and over 90% balanced accuracy in recognising running and rumbling behaviours from the seismic data. We release the dataset used in this study: SeisSavanna represents a unique collection of seismic signals with the associated wildlife species and behaviour. Our results suggest that seismic data offer substantial benefits for monitoring wildlife, and we propose to further develop our methods using dense arrays that could result in a seismic shift for wildlife monitoring.

position of being among the most famous endangered species, such as giant pandas, tigers or blue whales.

Of particular interest to us are African elephants *Loxodonta africana*, who are poached for their ivory that is falsely believed to have healing powers and reaches astronomical prices on the black market (Wasser et al., 2008), but are also killed because of increasing human–elephant conflicts. Populations of African elephants have generally been in decline during the 1970s and 1980s (Douglas-Hamilton, 1987). After a general recovery until the mid-2000s thanks to new laws and conservation efforts (Chase et al., 2016), the numbers are now once again on the decline and the species is now classified as endangered following its split with *Loxodonta cyclotis* (Gobush et al., 2021), with about 415 000 individuals estimated in 2016 (Thouless et al., 2016).

A crucial aspect of elephant conservation is real-time monitoring of the animals and their behaviour, for instance to understand if a herd is getting dangerously close to a human settlement, or if the animals start running, which could signify an attack by poachers (Sukumar, 2003). There are several approaches currently employed to monitor elephants that each have strengths and weaknesses. One possibility is to fit individual GPS collars to animals (Blake et al., 2001; Galanti et al., 2000; Ngene et al., 2010), which allows tracking in real-time, but unfortunately this is very time- and labour-intensive as it requires deploying a specialised team to tranquilise the elephant and fit the collar. A cheaper and commonly used option is to deploy camera traps, which trigger through an infrared motion sensor (Kays et al., 2009; Smit et al., 2019). A drawback of that method is the limited detection range of the infrared sensor, typically a field-of-view of roughly 60° and a viewing range of about 30 m (Randler & Kalb, 2018) in open terrain. The sensor is further influenced by environmental conditions such as temperature, and limited visibility due to vegetation and topography. One could perhaps deploy continuously recording cameras that do not have triggering mechanisms, but then again one runs into difficulties with obstacles to the field of view, massive data volumes and expensive computations, which are all drawbacks for real-time monitoring. There are also promising new airborne methods emerging for real-time monitoring by using drones (Mangewa et al., 2019), but they also have limitations such as field of view obstruction and environmental disruption due to noise.

In this study, we investigated the possibility of using seismometers and machine learning to distinguish elephants from other species, as well as classify their behaviour, from the ground motion resulting from their footsteps. Traditionally, seismic recordings belong to the realm of geophysics and seismology, where they are used to determine earthquake properties and the structure of the planet (Dziewonski et al., 1977; Hosseini et al., 2020; Szenicer et al., 2020). Additionally, seismic recordings have also been employed for footfall detection in humans (Clemente et al., 2019) and they seem a natural candidate for detecting much larger animals such as elephants. Mortimer et al. (2018) performed modelling with limited recorded data that suggested that, depending on behaviour and geology, seismic signals from elephants can be detected at distances up to hundreds of metres or even kilometres, and thus potentially have much larger ranges than camera traps do. In addition, seismometers are not limited by obstacles in the line of sight, nor do they have an azimuthal range limit. Finally, because of their low data volume and power requirements, seismic stations are routinely deployed for months at a time without the need

for maintenance or inspection, and therefore can be good candidates for monitoring in remote locations. All of the above suggest that using seismic data can be a promising approach to real-time monitoring, with potentially significant benefits over existing methods. Some work has been done in this direction, for example Wood et al. (2005) employed geophones and seismic data to recognise elephants, but on a limited scale, with only a few dozen datapoints, and handcrafted features extracted from the seismic data. Crucially, they only had one seismic station, and therefore could not investigate the generality of their method. Sugumar and Jayaparvathy (2013) also investigated using seismic data for elephant classification, but only tested their method on a few dozen simulated datapoints. Lamb et al. (2021) recently deployed Raspberry Shake sensors in South Africa to assess their viability for monitoring seismic vocalizations and locomotion of elephants, but not designed algorithms to detect and classify the animals. All these results are encouraging, and in this study, we take this idea to the next step, with large amounts of field data consisting of seismic and camera-trap recordings, and a fully data-driven automated deep learning approach.

Deep learning has pervaded many aspects of our lives in recent years, and has achieved great successes in fields as varied as natural language processing (Brown et al., 2020) and medical diagnostics (Shamout et al., 2019). It is also increasingly used in the sciences, with results such as solving protein folding (Senior et al., 2020), virtually fixing satellite instruments (Szenicer et al., 2019), or learning to solve the seismic wave equation (Moseley et al., 2020). Deep learning is a natural candidate to tackle complex unstructured data such as seismograms (i.e. ground displacement over time), and has already been employed on seismic data to detect earthquakes and estimate their magnitude (Meier et al., 2019; Mousavi & Beroza, 2020; Mousavi et al., 2020), image the subsurface (Bianco et al., 2019), or detect volcanic seismicity (Falcin et al., 2021), amongst many other applications. It is steadily being adopted more in to conservation, for example for counting elephant populations from high resolution satellite imagery (Duporge et al., 2020), or detecting whales from their vocalisations (Shiu et al., 2020), and we believe it is vital to further incorporate this technology into conservation efforts.

We conducted field work at the Mpala Research Centre in Kenya, where we deployed 22 seismometers, 30 camera traps and 8 microphones, running continuously for 3 weeks in February–March 2019. The cameras were deployed in the vicinity of some of the seismometers, and constitute the basis for defining labels for the seismic signals—the associated species and behaviour. The microphones are not the focus of the present study, apart from

using them to label seismic signals containing elephant rumbles. We acquired a rich dataset of more than 8000 h of seismic recordings and over 250 000 images, which were later annotated with labels, so that for each image containing an animal we extracted the corresponding seismogram (see Fig. 1). In Figure 1, the seismograms exemplify the diversity and complexity of the signals present in our data. We can see a clear difference between the giraffe and the elephant seismograms, the hooved footsteps producing a clearer, more impulsive signal. And yet the relatively small and lightfooted leopard, simply by its fortuitous close proximity to a seismometer, could easily be mistaken for a five ton elephant. One must therefore bear in mind the central roles played by several variables such as distance, soil and geology, number of animals, behaviour.

Using this dataset, we develop deep learning classifiers which use convolutional neural networks (CNN) to recognise elephants and their behaviour from the three-component (vertical, north and east axes of the seismometer) spectrograms of the seismic signals, which represent the temporal evolution of the frequency content of

the signal. In particular, we train a classifier to recognise elephants with good accuracy up to 150 m when trained and tested on the same set of seismometers. We also demonstrate that the network can recognise elephants on previously unseen seismic stations, albeit with lower accuracy when environmental conditions are very different from the training set, for example, on geographically distant stations with different terrains. Finally, we train a CNN to distinguish running elephants from walking ones, and a CNN to detect elephant rumbles which couple to the ground. Both signal types represent important behaviours to monitor from the perspective of conservation and biology.

Materials and Methods

Fieldwork

We conducted fieldwork in the Mpala research centre, located on the Laikipia Plateau in central Kenya, between February and March 2019. The research centre is geographically diverse, defined by a semi-arid savanna

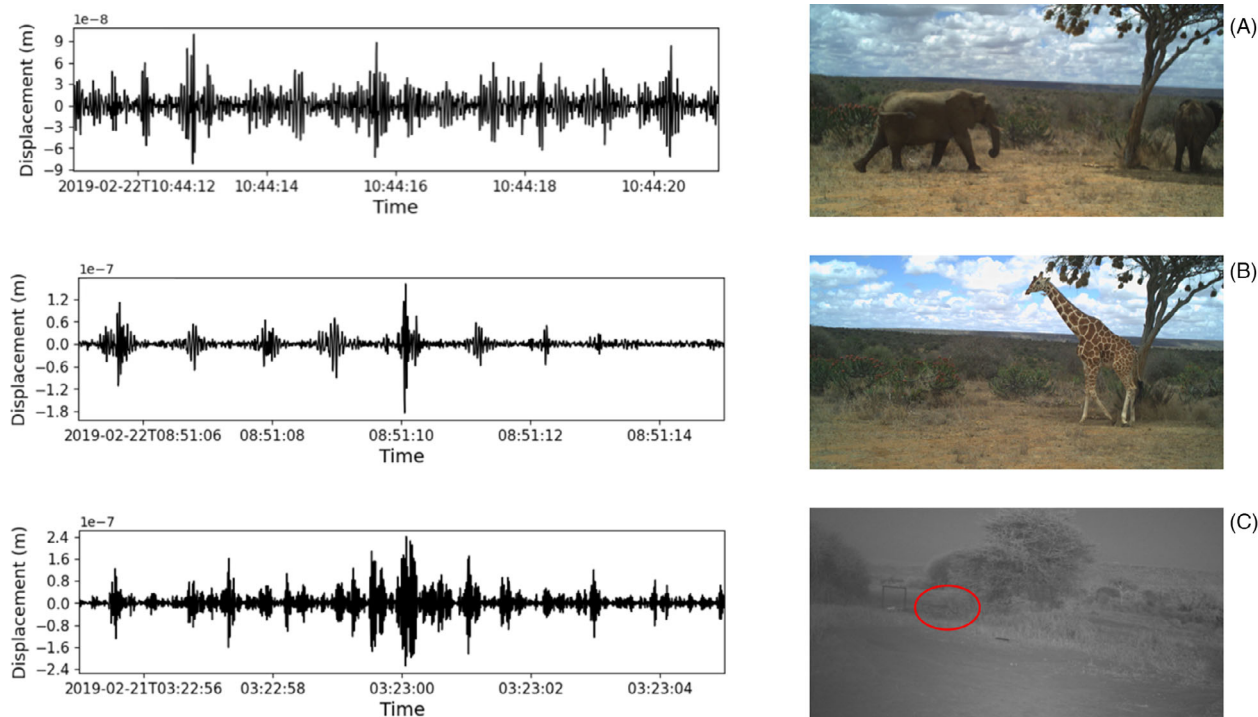


Figure 1. Camera trap pictures providing animal species and behaviour labels, and the corresponding vertical component seismograms extracted from nearby seismometers. While we only show the vertical component for clarity, our deep learning methods use all three components. The amplitude scales are purposefully not equalized between the plots, so that details of each seismogram can be distinguishable. In panel (A), we can see footsteps generated by two elephants on their way to a tree, whose shade provides relief from the heat, and its bark from itches. In (B) in the same location, a giraffe walking. In panel (C), a leopard sneaking past a seismometer at night, which can be seen in full size for more detail in Figure S1. All times on the seismograms are local times.

landscape and the catchments of two rivers, supporting a wide range of wildlife. In particular, Mpala is the temporary home of over 6000 migrating elephants, making it a great location for our experiment.

We deployed 22 6TD Güralp seismometers on a loan from SEIS-UK and 30 Reconyx Hyperfire 2 camera traps for a duration of 3 weeks. The seismometers recorded ground motion data in three orthogonal components at a 200 Hz sampling rate and were buried around 70 cm beneath the surface. The camera traps had an infrared motion detector that triggered the recording of 10 pictures at 1s intervals, and were mostly attached to trees and poles at human-eye height. The theoretical range of the sensor is up to 30 m in a 60° cone, although this is often lower in practice due to environmental conditions.

All the instruments were deployed in an area centred around a watering hole where many species congregated every day. As can be seen in Figure 2, most of the instruments were deployed within a few hundred metres of the hole in order to maximize the amount of data; but others were also deployed in more remote locations, so as to add diversity in terms of geology and terrain. For example, the terrain around the waterhole was wet, muddy and with trees and bushes on the outskirts, the NWP05

station was more arid with rocky and sandy terrains, while the remote NNL62 had a rocky geology as well as a river gorge that was often crossed by animals.

We recorded over 250 000 pictures on the camera traps and 8000 h of seismograms. One of the contributions of this paper is a unique dataset of seismic recordings with the corresponding label of the species generating the signal. To achieve this, the camera trap images were manually labelled for animal species, behaviour, number of animals and distance from the camera trap. The labels were then paired with seismic data from stations within a user-defined range from the animal sighting. More details on the processing and creation of the dataset can be found in the supplementary material.

Modelling

We investigated multiple machine learning models on this classification task, such as logistic regression, gradient boosted trees and multilayer perceptrons, on either seismograms or spectrograms. Our best-performing models were different architectures of convolutional neural networks (CNN), a complete introduction to which can be found in Goodfellow et al. (2016).

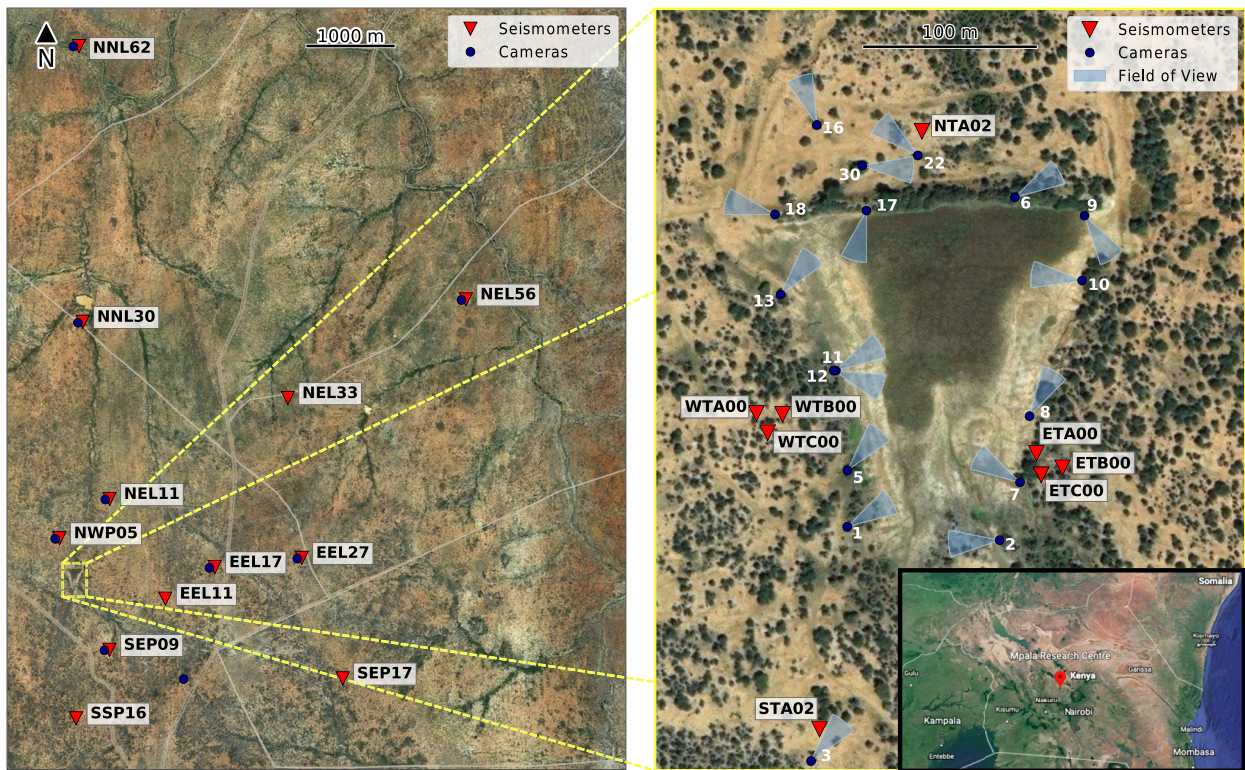


Figure 2. Site of fieldwork, with annotations for the positions of camera traps and seismometers. The name of seismometers is coded in reference to the watering hole, for example WTA00 is due west of the hole at a ‘distance’ of 0 m, STA02 is due south at distance of 150 m, NWP05 is north-west at a distance of 500 m.

In this study, we experimented with different types of CNNs: 1D CNNs trained from scratch on the seismograms, 2D CNNs trained from scratch on spectrograms and 2D CNNs pretrained on ImageNet (Deng et al., 2009) and then finetuned on spectrograms. The latter approach is very successful in environmental sound classification (Palanisamy et al.,

2020), and is therefore a natural candidate to apply to our dataset.

Metrics

As can be seen in Figure 3, we have a relatively imbalanced dataset with more sightings of elephants that non-

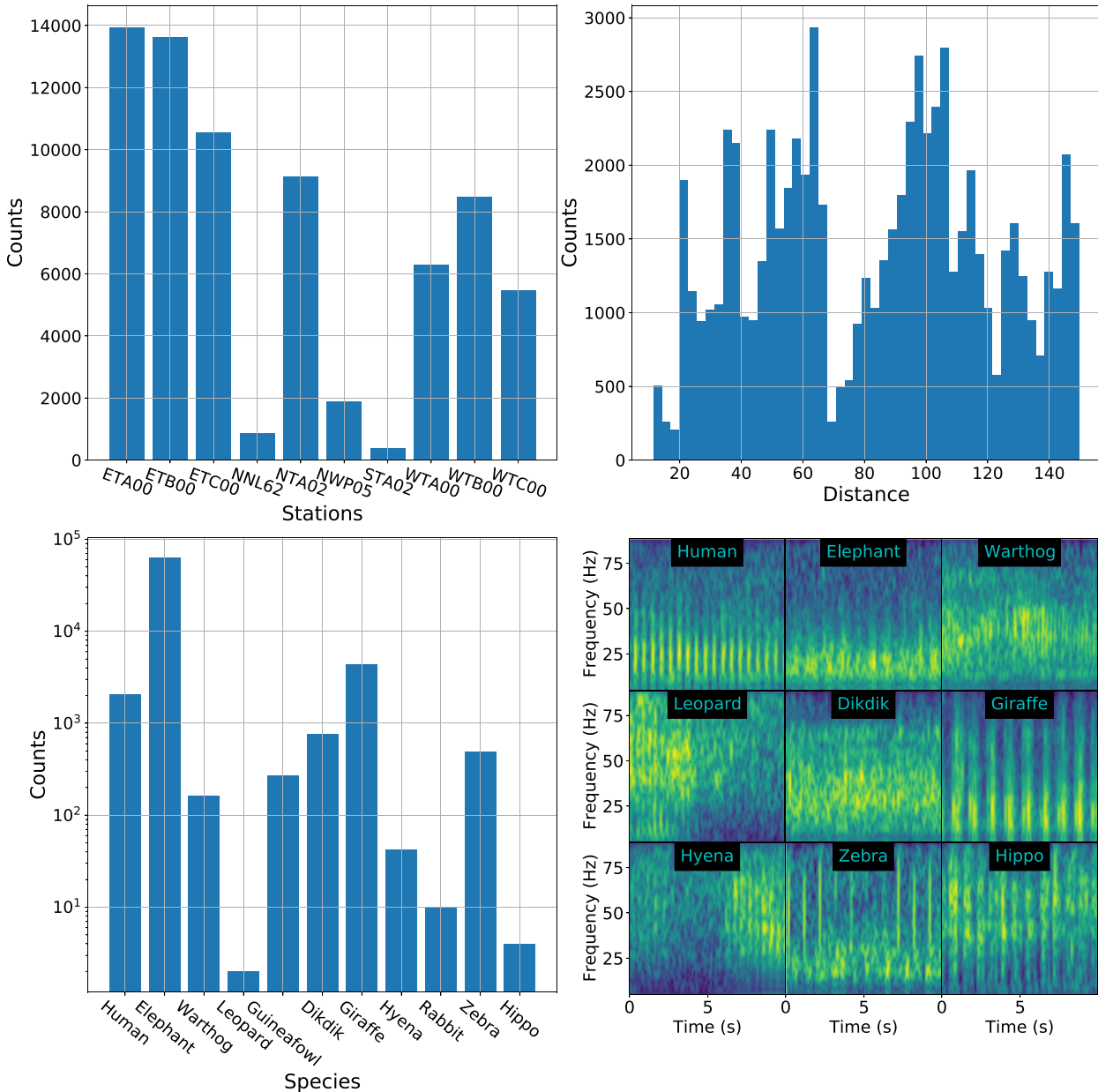


Figure 3. We provide a dataset of seismic chunks that contain labels of animals located up to 150 m distance from any given seismometer. For each chunk, we provide the class (species), the station it is taken from (see Fig. 2 for a map), the distance from the station and the image that generated the label. Panel (A) shows the data count for each station. We can see that the majority of the data comes from around the waterhole, particularly the ET and WT stations. Panel (B) shows the distribution of data with respect to distance. Panel (C) displays the class counts, where the y axis is on a log scale, given the overwhelming majority of elephant signals. Finally, panel (D) shows the spectrograms of several events from different classes in the dataset.

elephants, and class imbalances for the behavioural datasets as well, therefore we have to use appropriate metrics to account for that fact, as usual metrics (e.g. accuracy) are overly optimistic for imbalanced datasets. With TP denoting true positives, TN true negatives, FP false positives and FN false negative, we report true positive rate $TPR = TP/(TP + FN)$, true negative rate $TNR = TN/(TN + FP)$ and balanced accuracy $BA = (TPR + TNR)/2$. By averaging the true positive and negative rates, BA offers a balanced and interpretable measure of the quality of binary classification, even in the presence of class imbalance.

Data augmentation

Data augmentation is a standard machine learning method to improve the generalisation capabilities of a network. It involves modifying datapoints from the training set in realistic ways in order to expose the network to a more diverse dataset, which is more likely to be representative of the data in the test set.

A crucial requirement of data augmentation methods is that they must create realistic datapoints. While in traditional image classification of natural images such as cats and dogs, common augmentation techniques include image rotations and distortions (Shorten & Khoshgoftaar, 2019), it would be unsuitable for seismic data, since, for example rotating a spectrogram plot would not result in a realistic new datapoint.

There have been several methods proposed for earthquake seismic data such as channel dropout, source superposition, false positive noise (Zhu et al., 2020). We selected two methods that appear most suitable for our task, which are source superposition and addition of noise. The former corresponds to adding together different seismograms from the same class. Given the linearity of the wave equation, adding for instance several different signals of elephants walking should mimic the signal generated by several elephants walking together. Such augmentation clearly does not encompass the large variety of scenarios encountered in the wild, but nevertheless adds some tangible extensions to the dataset that are otherwise not recorded. Noise addition simply consists of adding seismic noise to the signal, to artificially create more diverse signal-to-noise ratios in the data. To create realistic noise, we randomly selected 10-second seismic chunks from time windows in the recorded data that did not correspond to any camera-trap detected activity, and inspected them visually to ensure they contained only noise.

Results

To provide context about the classification results, we begin by introducing the dataset used for this study,

which is specifically curated for machine learning applications and made open source. We then proceed to detail the results of our method on the classification of elephants as well as their behaviours, in particular running and rumbling.

SeisSavanna: an out-of-the-box dataset

One contribution of this paper is to open source the dataset used in this work, which we name SeisSavanna. The data are in netcdf format (Rew & Davis, 1990), and are separated into three main files. The first file is focused on the task of species classification. It is composed of seismograms, the corresponding picture that generated the sighting, the species (i.e. the label), and other information such as seismometer name and distance of the animals from the seismometer. We provide a 'master' dataset that includes animal sightings up to 150 m from any given seismometer. We also provide scripts that easily reduce this dataset to any desired maximal distance, such as 40 or 60 m. Likewise, we include sightings from all available seismometers, as well as tools to select desired subsets of stations. In this way, we want to provide maximal flexibility to the user and enable creative applications. The second netcdf file is specific to elephant behaviour, and only contains elephant sightings up to 150 m away, with the label being either 'running' or 'walking'. Finally, the last netcdf file is focused on the task of detecting elephant rumbles recorded on seismometers, and as such the class labels are 'rumble' or 'not rumble'. The rumble signals do not have distance information, while the non-rumble signals come from sightings of animals up to 150 m away. To the best of our knowledge, SeisSavanna is the first dataset of its kind, which contains large amounts of seismic data and images from wild animals, made open-source and curated for ML applications.

In Figure 3, we provide some statistics about the species classification file in SeisSavanna. In particular, we plot the data counts with respect to station, distance and species. We can see that most of the events come from the stations within a 200 m radius from the waterhole, since this location was both the most instrumented with camera traps, but also saw the heaviest animal traffic being one of only two locations within the area bearing water. The two more remote stations, NWP05 and>NNL62, have between one and two thousand samples each. There is also a class imbalance, where the majority of datapoints come from elephants (about 60 000), with other species having from a few dozen to a few thousand samples. We also plot several examples of spectrograms coming from different species in panel (D).

It is interesting to observe the differences in signal shape, and relate them to biological features. For example,

we can see a clear difference in the frequency of footsteps between a relatively small animal (e.g. human, warthog and hyena) and a giraffe or a hippo. Similarly, there are noteworthy differences in the frequency content of the signal. For example, some signals demonstrate more impulsive sources with a relatively white spectrum (e.g. giraffe or human), whereas others exhibit a softer original impact (e.g. elephants), with a more concentrated spectrum. Similar patterns in the spectra of footfalls have been observed by Wood et al. (2005). However, when making this observation, it is important to bear in mind the impact of signal propagation (and thus distance) on seismic data, which progressively removes high frequencies. A detailed description of our data collection process is given in the supplementary material.

Classifying elephants versus other seismic sources

Generalising in time

First, we focused on the task of discriminating elephants from other species, while keeping environmental conditions mostly constant, by training and testing on the same set of seismic stations. We selected a subset of stations (ETA00, WTA00, STA02, NTA02, NWP05 and>NNL62) for which the data were sorted in time. We then selected the first 15% of the data for the test set, the following 15% for validation and the remaining data for training. This split in time allowed us to have independent training/validation/test sets. Indeed, we cannot simply randomise our samples before splitting, as is common in many machine learning applications, since the samples were not independent and identically distributed. We report the results for several maximal distances of animals, namely 40, 60, 80, 100 and 150 m, which allow us to probe the limits of the detection range in SeisSavanna.

We trained different models on this task for the 60 m dataset: several baselines including a logistic regression, gradient boosted trees and a multi-layer perceptron, all on flattened seismograms/spectrograms, where flattened means that the three seismogram components are concatenated that is corresponding to a 1×6144 dimensional input rather than 3×2048 dimensional input. Then we trained an all-convolutional 1D-CNN inspired by Springenberg et al. (2014) on the seismogram data, an all-convolutional 2D-CNN on the spectrogram data, and finally, a Squeezenet (Iandola et al., 2016) either trained from scratch or pretrained on ImageNet and finetuned on spectrograms. The last approach was motivated by research in environmental sound classification, in which ImageNet-pretrained

models deliver state-of-the-art results (Palanisamy et al., 2020). All the models use the three components of the seismograms/spectrograms. While we have also tried using only the vertical component, it generally led to lower accuracy. Details of the models' architectures and hyperparameters used during training can be found in the supplementary materials.

Here, we only report results from the best-performing model on the task, but results from additional models are provided in the supplementary material. For the elephant/non-elephant classification with generalisation in time on the 60-m dataset, the best-performing model was the finetuned Squeezenet trained on spectrogram data, which we therefore used on the datasets for other maximal distances. In Figure 4A, we plot the balanced accuracy achieved by the network on the test set of datasets with increasing maximal distances. We can see that the model achieves close to 90% balanced accuracy up to 80 m distance, which then decreases to 82% at 100 m, and decreases further to 73% for 150 m. The drop in accuracy is further confirmed by inspecting classification accuracy as a function of distance, as can be seen in Figure 5. The plot suggests that beyond around 100 m, the accuracy drops off, likely because increasingly more seismograms have lower signal-to-noise ratios. This drop in accuracy is due to the labels becoming corrupted, that is a signal labelled 'elephant' while it is only noise on the seismometer, as can be seen in Figure 6. The extent of signal propagation depends on environmental conditions such as force of the impact, number of animals, local geology. As such, label corruption can also happen at closer ranges, but on average increases with distance from the seismometer.

We believe that these results are very promising, and they showcase the great potential of using seismic data to accurately detect elephants up to relatively large distances compared to other methods such as camera traps: around 100 m with our current dataset.

Generalising to new stations

In this section, we still classify signals into elephant/non-elephant, but in a more ambitious training scenario. Indeed, we trained on a subset of stations and tested on a different subset. Achieving across-station generalisation would in theory allow us to only conduct data-gathering for training once, and then add new stations to our existing network, or even perform the analysis with this same pretrained network in new field deployments and locations. Each split used a new seismometer as a test set, in order to investigate generalisation ability to different environments. It is noteworthy to appreciate that the terrain across our 22 stations varied quite significantly, including

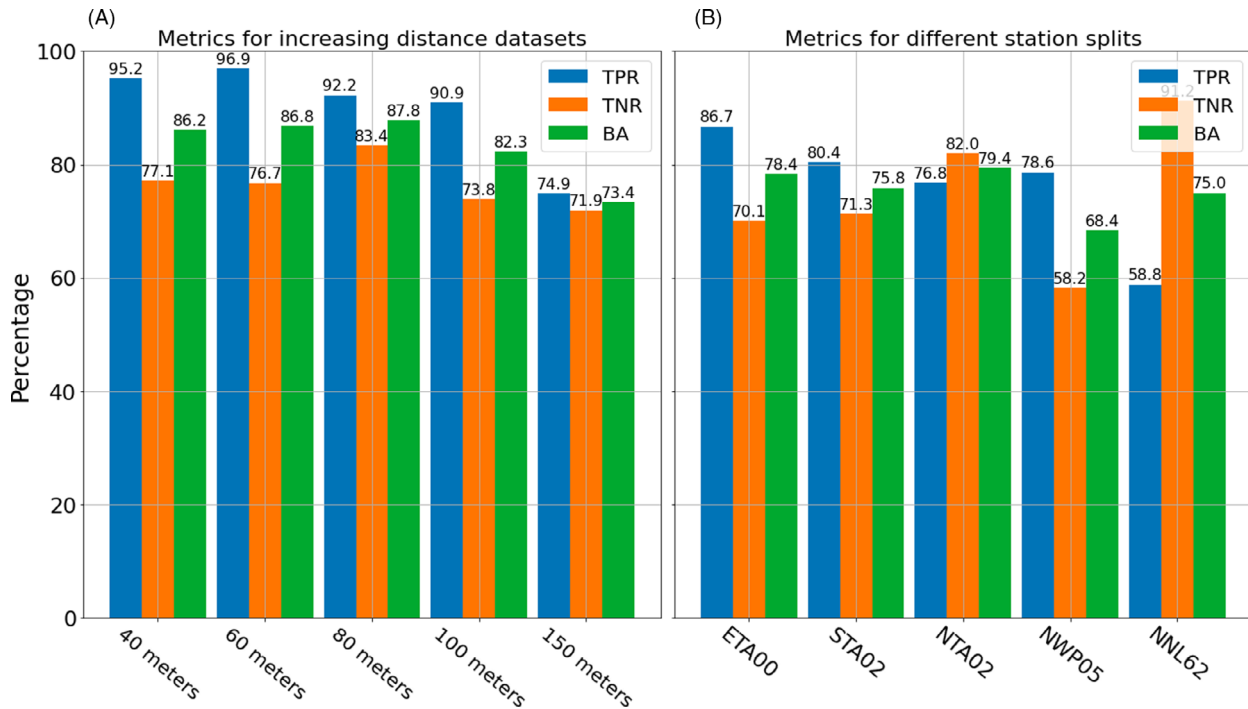


Figure 4. (A) True positive rate (TPR, elephant accuracy), true negative rate (TNR, non-elephant accuracy) and balanced accuracy (BA, average of elephant and non-elephant accuracies) achieved by the finetuned SqueezeNet on the test set for datasets with increasing maximal distance of animals. (B) True positive rate (TPR, elephant accuracy), true negative rate (TNR, non-elephant accuracy) and balanced accuracy (BA, average of elephant and non-elephant accuracies) achieved by the finetuned SqueezeNet on the test set for different test stations, which probed the ability of the network to generalise to new environmental conditions. We annotate the histograms with the value of the metric for easier interpretation.

topographic and vegetation changes, soil conditions from wet mud to unconsolidated sand and rocky terrain. All the splits were for a maximal distance of 60 m, because we wanted to decouple the effect of the environment from effects such as label degradation with distance mentioned in the previous section.

The best-performing model for the across-station splits was the finetuned SqueezeNet. In Figure 4B, we plot the balanced accuracy, true positive rate (elephant accuracy) and true negative rate (non-elephant accuracy), for the test set of each station split. We still achieved good balanced accuracy, albeit lower than when generalising in time. Remarkably, the network achieved lower balanced accuracy when the test set was NWP05 or NNL62, compared to the other splits. This is noteworthy because these two stations were the furthest away from the waterhole where most of the data came from, and had therefore the most different geological and environmental conditions, for instance hard rocky ground at NNL62, as opposed to wet muddy conditions at the waterhole. This suggests that, as expected, we need to acquire data from a varied collection of environmental conditions to improve our generalisation capabilities.

Classifying behaviour

Locomotion versus vocalisation

Besides recognising and tracking elephants, a central objective of a monitoring system is to be able to alert authorities to potentially perilous situations for the animals. A behaviour of interest is elephant vocalisation, in particular in the form of rumbling, because they couple to the ground and can be observed in our seismic signals. Elephants use these vocalisations as communication, for example for greetings, warnings about imminent threats, or communicating movement (Poole, 2011). While there is research looking at automatically detecting and classifying elephant vocalisations (Leonid & Jayaparvathy, 2020), it is performed with small sample sizes, and crucially on acoustic data from microphones. To the best of our knowledge, this work is the first time this is done using seismic signals with large amounts of data, and using deep learning.

Because rumbles cannot be identified from images, we manually went through a portion of the data to look for vocalisations. The rumbles have a characteristic shape that

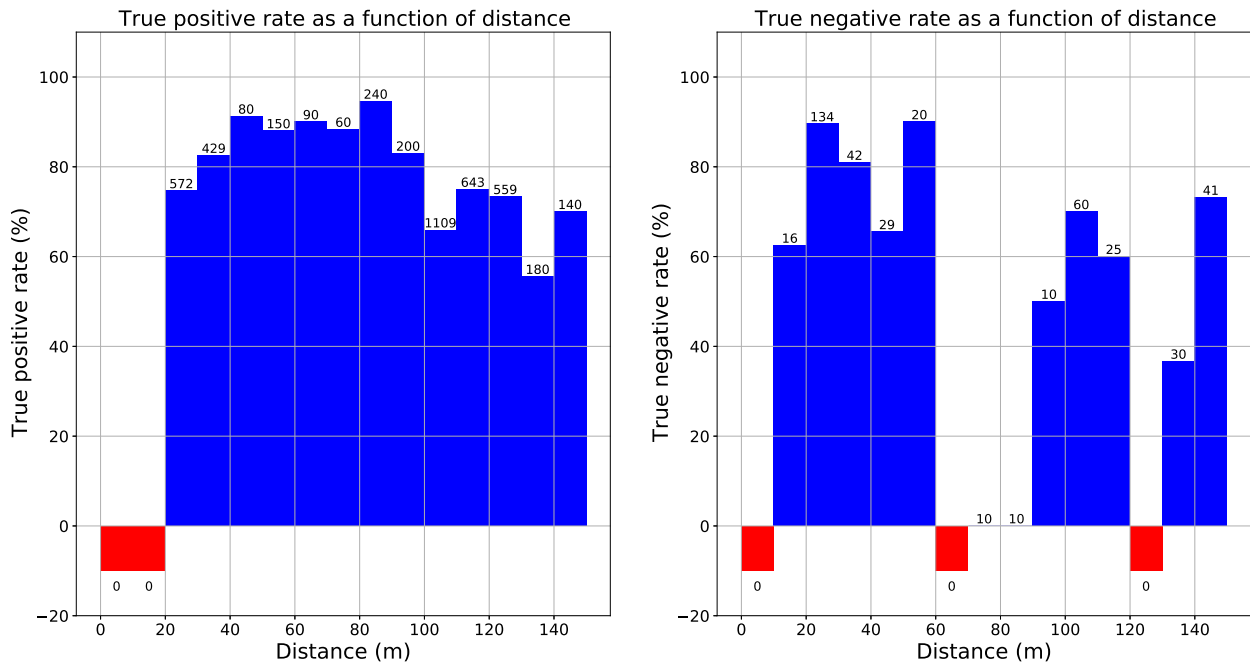


Figure 5. (A) True positive rate (elephant accuracy) and (B) true negative rate (non-elephant accuracy) as a function of distance on the test set of the 150-m dataset, for a finetuned squeezeNet. We annotate the histogram with numbers of samples to clarify the absence of data for certain distances.

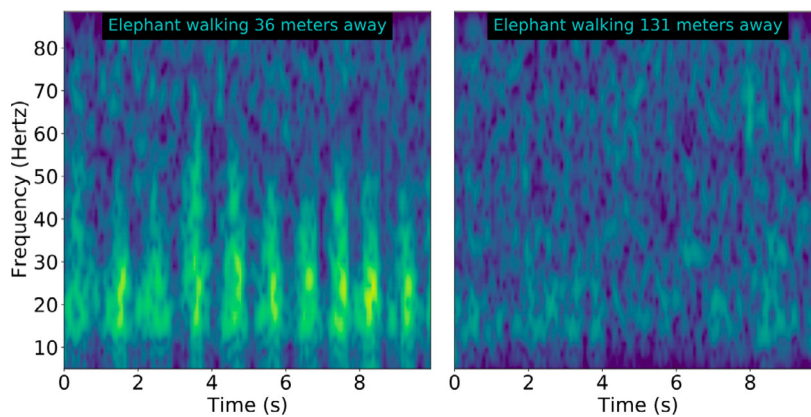


Figure 6. Spectrograms of two elephant samples, recorded on the station NWP05 at different distances. We can see that for the elephant walking 131 m away, the camera trap provides an elephant label, but the seismic signal becomes so faint after propagation that the datapoint looks like noise.

is easily recognisable, and in stark contrast to an impulsive locomotion signal, as we can see in Figure 7. Note that seismic rumble signals do not contain as many harmonics as seen in acoustic signals, presumably due to their coupling characteristics to the ground Reinwald et al. (2021); this will need to be investigated in a future study. However, it is important to note that the lowest seismic frequency of the rumble is the most interesting and useful one, as that signal will propagate furthest. We

therefore manually labelled 1500 rumbles on two seismic stations (ETA00 and EEL11), and combined these into a dataset with locomotion signals from all other species together (elephants, giraffes, etc.). The rumbles detected in the seismic data were also validated by their presence in the microphone data. The spectrograms were processed differently than in the previous section (see supplementary material for details), and following Reinwald et al. (2021), we applied the structure tensor (Harris &

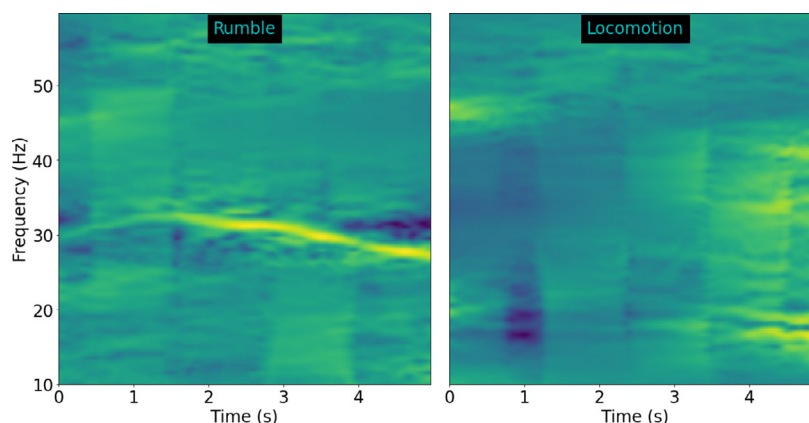


Figure 7. Seismic spectrograms of an elephant rumble (left) versus elephant locomotion (right). The rumble has a very distinctive spectral signature, although it is different from the shape of the spectrogram one obtains when recording rumbles on microphones, because the ground coupling has a strong effect on the characteristics of the signal. We process the spectrograms to enhance the rumble signal and attenuate the locomotion signals, as can be seen in Figures S2 and S3.

Stephens, 1988) of the spectrogram as a filter. This approach enhances the rumbles which contain sharp contours along the frequency axis, and diminishes the broadband locomotion signals, as can be seen in Supplementary Figures S2 and S3.

The finetuned squeezeNet achieved the best performance on the rumble classification task, with 96.1% balanced accuracy, of which the true positive rate (rumble accuracy) was 97.3%, and true negative (non-rumble accuracy) was 94.8%. This provides us with a highly accurate automated seismic rumble detector, which is yet another reason to push further with using seismic data for elephant conservation.

Walking versus running

Another common behaviour exhibited in times of danger is running (Sukumar, 2003), either as an attack or

defence mechanism from poaching for instance. Therefore, once we identified an elephant signal, we additionally attempted to train a classifier to determine whether the elephants were running or not.

This task is particularly difficult for elephants, because their gait transition is not as pronounced and clear as other animals such as in humans or zebras (Ren & Hutchinson, 2008). Therefore, the main differences we can expect in the signal that comes either from the frequency of the footfalls (as can be seen in Figure 8) or their intensity. However, these two effects can also be confounded by the number of animals or the distance of the animals, for example several elephants walking close to a seismometer could look the same as elephants running further away. Nevertheless, we included all sightings with up to 100 animals in the image, in order to be as realistic as possible in the type of signals we cover, and also increase our sample size.

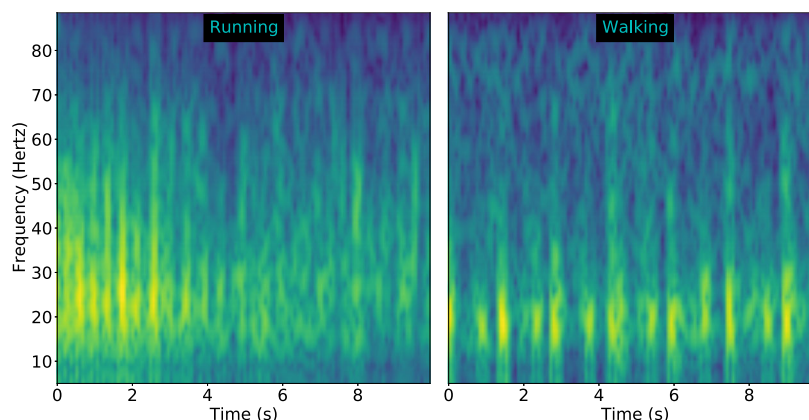


Figure 8. Spectrograms of an elephant running (left) versus elephant walking (right). In this instance we can see a clear difference in footstep frequency between the two signals.

The behaviour dataset was generated similarly to the species dataset, except that we only selected sightings of elephants, and used behavioural notes for class attribution to ‘running’ or ‘walking’. The 60-m dataset contained 309 samples of the running class, and 18 726 of non-running. We took particular care to only keep running data for which the signal is very clear, since having erroneous labels with such low sample size drastically degraded performance. To this end, we visually inspected the running data and refined the labels by discarding samples for which it is clear that the signal has not been recorded at the seismometer. While we only refined labels for the samples from the ‘running’ class (the large amount of samples in the ‘walking’ class made visual inspection too time consuming), and it is inevitable that the ‘walking’ class will contain some unnoticed instances of running, the majority of the labels should be correct and allowed the network to successfully discriminate between the signals.

The best-performing model on this task was the 2D-CNN trained from scratch, which on the test set achieved a balanced accuracy of 93.9%, with a perfect true positive rate (running accuracy) of 100%, and a true negative rate (walking accuracy) of 87.8%. Interestingly, this was the only task for which the 2D-CNN trained from scratch outperformed the pretrained Squeezenet, which achieved 75% balanced accuracy.

Once again, these results show great promise for the use of seismic data to not only detect elephants, but also understand whether they display behaviour that potentially indicates a perilous situation.

Discussion

In this work, we introduce a novel approach to monitoring wildlife, including African elephants, which utilises deep learning to classify seismic signals recorded on seismometers.

Crucially, we open-source SeisSavanna, the seismic wildlife dataset used in this work, with the hope that it will foster further collaboration between the fields of machine learning, geophysics, and biology. The species file in SeisSavanna contains 70 k sightings from 11 different species, for distances up to 150 m. We also provide two elephant behaviour files, one containing sightings of elephants running and walking, with respectively 1.5 k and 61 k samples up to 150 m; and a rumbles file containing 1.5 k rumbles and 25 k locomotion signals. While this dataset is intended to benefit the investigation of new methods to use seismic data for conservation, it is worth noting that it is also very useful for computer vision tasks, since for each seismogram in the species and

running files, we provide the corresponding picture that generated the sighting.

By applying deep learning methods to these datasets, we were able to distinguish elephants from other species with 80%–90% balanced accuracy for distances of up to 100 m, recognise elephant vocalisations in the form of rumbles with 96% balanced accuracy, and distinguish walking from running in elephants with 94% balanced accuracy for distances up to 60 m. Our best-performing approaches were CNNs applied on spectrograms, in particular finetuning a Squeezenet pretrained on ImageNet, or training an all convolutional 2D-CNN from scratch.

The aforementioned results are very encouraging, as deep learning methods performed very well in the tasks we set them, with good accuracy despite the challenging natural context in which data were collected. In particular, the variable field terrain (e.g. wet and muddy material, rocky ridges, hard stone and sand, gorges) and the noisy conditions of the sites (e.g. many different animal species, large variation in number of animals and their concurrent behaviour, wind, cars) are usually seen as compromising factors for generalisation of such an approach, but our results indicate that these can be handled well. Our results also highlight some of the limitations and key variables of the dataset which we discuss in more detail below: (1) uncertainty in the labels due to partial camera trap coverage; (2) decrease of signal-to-noise ratios (SNRs) with distance, which causes corruption in the labels; and (3) effect of local environmental conditions and their impact on the ability of the network to generalise to new locations.

We now address each of these points in more detail, before outlining how these limitations can be mitigated.

To begin with, one crucial consideration about the data, which is relevant to all the tasks we are tackling, is the relatively high level of uncertainty present in the signal. Indeed, the seismometers record data with a 360° azimuthal range, whereas the animal species labels are provided by camera traps, which have a range of about 60° and about 30 m. This means that there is a sizeable blind spot in our labelling procedure, and to give an extreme example, there could be an elephant walking a few metres from a seismometer, while the label provided by the camera is that of a zebra 50 m away. Naturally, the fieldwork was designed to try and minimize these scenarios, notably by placing camera traps covering all directions in areas most propitious to animal traffic, as recommended by our local field guide. Therefore, the corruption of labels should not be too prevalent, but must nevertheless be kept in mind. Considering this uncertainty, the accuracies achieved in this study are all the more encouraging.

Another important limitation of the data to bear in mind, as highlighted by the decrease in accuracy in classifying elephants versus non-elephants for the 150 m dataset, is the increasing corruption of the labels in the dataset and decreasing SNRs. Indeed, as we include sightings of animals at increasingly large distances, seismic signals are less and less likely to be recorded by a seismometer. Therefore, if we have a label saying ‘elephant at 500 m’, the extracted chunk will most likely contain only noise. This causes problems both during training (showing very similar noise signals both for the ‘elephant’ and ‘non-elephant’ classes) and during testing (labels are corrupted and therefore metrics are biased). Notably, this does not preclude signals to propagate over such distances or further in principle; only the complexity and noise environment of our specific dataset prevents us from training the algorithms with such distant propagation.

Finally, attempting to classify elephants from non-elephants and generalise to new stations highlights the importance of the local environment of the station, and the need for more diverse datasets to improve generalisation capabilities to new environments. Indeed, it is noteworthy that the balanced accuracy is higher when the test set station is similar to the stations in the training set, and lower when the test station is remote or with a different environment, which is physically coherent, because factors such as terrain and geology have important impacts on signal propagation (Mortimer et al., 2018). It is also in line with general results in deep learning, whereby training and test data have to be generated from the same distribution for the network to produce good results, and is the reason why the machine learning community strives to produce datasets containing enormous amounts of labelled data (e.g. ImageNet (Deng et al., 2009) for image classification with 14 197 122 pictures in 1000 classes, or MS COCO (Lin et al., 2014) for object detection with 328 k pictures annotated with bounding boxes, segmentation masks, natural language captions).

How might we address or mitigate the limitations outlined above? An obvious solution is to acquire more data, from more diverse environments. This will allow us to improve accuracy, generalise better to new stations, but also test refined classification tasks, such as multispecies classification. A straightforward way to get more labelled data is to label the remaining camera trap pictures. Another possibility worth exploring is to deploy dense arrays of devices such as geophones or accelerometers. These devices are inexpensive and easy to deploy, and can therefore be used to sample many different locations, which is a boon both for the monitoring range since we can cover more area, but also for the performance of the method, as we will sample more varied environmental conditions and generalise better. Dense arrays of

instruments can also be beneficially exploited, for example to enhance signal to noise ratios with stacking, but also with more advanced array methods (Rost & Thomas, 2002). However, more data are not the only solution to overcome the outlined limitations of our approach.

A very successful way to improve generalisation ability in machine learning is the use of data augmentation methods, whereby one creates fake training examples by modifying existing data in realistic ways (e.g. rotating or distorting pictures in image classification (Shorten & Khoshgoftaar, 2019)). We attempted to use data augmentation methods designed for seismic signals (Zhu et al., 2020), but have not seen a consistent improvement in accuracy. This likely means that for our task, these methods do not produce realistic or useful datapoints, and in future work we will address the design of specialised data augmentation techniques that will help us generalise to new environments and therefore across stations. For instance, we are planning to investigate synthetic signal based augmentation, by exploiting generative adversarial networks (GAN), which can be used to produce realistic new training examples (Frid-Adar et al., 2018; Li et al., 2018). In particular, we would like to be able to create data that appears to come from new environments, which is key to across-station generalisation.

To tackle the uncertainty in the labels, future fieldwork should surround each seismometer with several camera traps in order to provide full azimuthal coverage and remove label corruption altogether. To address the issue of label corruption due to increasing distances, one can improve the existing dataset by sifting through all the datapoints and manually refining the labels by removing samples that contain no signal. It is a slow and time consuming process, which unfortunately is hard to automate by simply removing low SNR samples, because doing so also removes many valuable low amplitude signals that have comparable SNRs.

Overall, our results show that information-bearing, discriminatory signals propagate over large distances compared to other methods (over 100 m), and that seismic data are a very promising avenue for monitoring wildlife and their behaviour with automated techniques. Together with the release of seismic wildlife datasets, this opens up the door for many applications. First and foremost, this approach paves the way for the development of elephant seismic monitoring systems for conservation, with the potential for near real-time capability. We have shown that seismic data offers great benefits for detecting elephants, in terms of sensitive detection up to tens of metres range and accurate classification of both the species and behaviours. The detection of elephant running is of particular interest since this potentially indicates a situation of welfare concern (Sukumar, 2003). In the future,

in combination with the deployment of dense arrays of geophones or accelerometers in optimised array geometries to improve signal-to-noise ratios and sample more varied environments, we envisage autonomous sensor systems that can record, collate and analyse seismic data streams in near real-time. Moreover, a variety of different sensors can be investigated, including ones that are cheaper and more practical to deploy. This has broad applications for elephant monitoring, whether for the study of their behaviour and communication in the wild, or as information for rangers to respond to behaviours of concern, such as elephants running. With further research, this approach could have applications beyond elephants to detect, classify and monitor a range of animals within their remote habitats.

Acknowledgments

We would like to express our gratitude to all the staff at the Mpala Research Centre for assisting with fieldwork operations and for creating a warm and welcoming environment, in particular Dino Martins and Cosmas Nzomo. We thank Frank Pope and staff at Save the Elephants for help getting research permits and Kenyan Wildlife Service affiliations, and Paula Koelemeijer for help with sensor deployment. A. Szenicer thanks the anonymous donor of his PhD grant. We thank National Geographic (NGS-50019R-18), Royal Society (URF R1 191033), John Fell Oxford University Press Research Fund, Royal Commission for the Exhibition of 1851 for funding. This research has been supported by the Centre for Doctoral Training in Autonomous Intelligent Machines and Systems at the University of Oxford, Oxford, UK, and the UK Engineering and Physical Sciences Research Council. A. G. Baydin is supported by EPSRC/MURI grant EP/N019474/1 and by Lawrence Berkeley National Lab.

Data Availability Statement

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

The SeisSavanna dataset, as well as codes to replicate the main results from the paper, can be found at: <https://github.com/AlexSzen/SeisSavanna><https://github.com/AlexSzen/SeisSavanna>.

References

Barnosky, A.D., Koch, P.L., Feranec, R.S., Wing, S.L., & Shabel, A.B. (2004) Assessing the causes of late pleistocene extinctions on the continents. *Science*, **306**(5693), 70–75.

- Bellwood, D.R., Hughes, T.P., Folke, C., & Nyström, M. (2004) Confronting the coral reef crisis. *Nature*, **429**(6994), 827–833.
- Bianco, M.J., Gerstoft, P., Olsen, K.B., & Lin, F.-C. (2019) High-resolution seismic tomography of long beach, ca using machine learning. *Scientific Reports*, **9**(1), 1–11.
- Blake, S., Douglas-Hamilton, I., & Karesh, W. (2001) Gps telemetry of forest elephants in Central Africa: results of a preliminary study. *African Journal of Ecology*, **39**(2), 178–186.
- Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P. et al. (2020) Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.
- Ceballos, G., Ehrlich, P.R., & Raven, P.H. (2020) Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *Proceedings of the National Academy of Sciences*, **117**(24), 13596–13602.
- Chase, M.J., Schlossberg, S., Griffin, C.R., Bouché, P.J., Djene, S.W., Elkan, P.W. et al. (2016) Continent-wide survey reveals massive decline in African savannah elephants. *PeerJ*, **4**, e2354.
- Clemente, J., Song, W., Valero, M., Li, F., & Liy, X. (2019) Indoor person identification and fall detection through non-intrusive floor seismic sensing. In: *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 417–424.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009) Imagenet: a large-scale hierarchical image database. In: *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
- Douglas-Hamilton, I. (1987) African elephants: population trends and their causes. *Oryx*, **21**(1), 11–24.
- Duporge, I., Isupova, O., Reece, S., Macdonald, D.W., & Wang, T. (2020) Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. *Remote Sensing in Ecology and Conservation*. <https://doi.org/10.1002/rse2.195>
- Dziewonski, A.M., Hager, B.H., & O'Connell, R.J. (1977) Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research*, **82**(2), 239–255.
- Falcin, A., Métaixian, J.-P., Mars, J., Stutzmann, É., Komorowski, J.-C., Moretti, R. et al. (2021) A machine-learning approach for automatic classification of volcanic seismicity at La Soufrière Volcano, Guadeloupe. *Journal of Volcanology and Geothermal Research*, **411**, 107151.
- Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018) Gan-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, **321**, 321–331.
- Galanti, V., Tosi, G., Rossi, R., & Foley, C. (2000) The use of GPS radio-collars to track elephants (*Loxodonta africana*) in the Tarangire National Park (Tanzania). *Hystrix-the Italian*

- Journal of Mammalogy*, **11**(2). <https://doi.org/10.4404/hystrix-11.2-4145>
- Gobush, K., Edwards, C., Balfour, D., Wittemyer, G., Maisels, F., & Taylor, R. (2021) *Loxodonta africana*. The IUCN red list of threatened species 2021: e. t181008073a181022663.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016) *Deep learning*, Vol. 1. Cambridge: MIT Press.
- Harris, C.G., Stephens, M. (1988) A combined corner and edge detector. In: Taylor, C.J. (Eds.) *Alvey vision conference*, Vol. 15. Plessey Research Roke Manor, UK: Alvey Vision Club, pp. 10–5244. <https://doi.org/10.5244/C.2.23>
- Hosseini, K., Sigloch, K., Tsekhmistrenko, M., Zaheri, A., Nissen-Meyer, T., & Igel, H. (2020) Global mantle structure from multifrequency tomography using p, pp and p-diffracted waves. *Geophysical Journal International*, **220**(1), 96–141.
- Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., & Keutzer, K. (2016) Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5 mb model size. *arXiv preprint arXiv:1602.07360*.
- Kays, R., Kranstauber, B., Jansen, P., Carbone, C., Rowcliffe, M., Fountain, T. et al. (2009) Camera traps as sensor networks for monitoring animal communities. In: *2009 IEEE 34th Conference on Local Computer Networks*, pp. 811–818. <https://doi.org/10.1109/LCN.2009.5355046>
- Lamb, O.D., Shore, M.J., Lees, J.M., Lee, S.J., & Hensman, S.M. (2021) Assessing raspberry shake and boom sensors for recording African elephant acoustic vocalizations. *Frontiers in Conservation Science*, **1**, 10.
- Leonid, T.T., & Jayaparvathy, R. (2020) Statistical-model based voice activity identification for human-elephant conflict mitigation. *Journal of Ambient Intelligence and Humanized Computing*, **12**(5), 5269–5275.
- Li, Z., Meier, M.-A., Hauksson, E., Zhan, Z., & Andrews, J. (2018) Machine learning seismic wave discrimination: application to earthquake early warning. *Geophysical Research Letters*, **45**(10), 4773–4779.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D. et al. (2014) Microsoft COCO: common objects in context. In: Fleet, D., Pajdla, T., Schiele, B. & Tuytelaars, T. (Eds.) *European conference on computer vision*. Cham, Switzerland: Springer, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48
- Mangewa, L.J., Ndakidemi, P.A., & Munishi, L.K. (2019) Integrating UAV technology in an ecological monitoring system for community wildlife management areas in Tanzania. *Sustainability*, **11**(21), 6116.
- Meier, M.-A., Ross, Z.E., Ramachandran, A., Balakrishna, A., Nair, S., Kundzicz, P. et al. (2019) Reliable real-time seismic signal/noise discrimination with machine learning. *Journal of Geophysical Research: Solid Earth*, **124**(1), 788–800.
- Mortimer, B., Rees, W.L., Koelemeijer, P., & Nissen-Meyer, T. (2018) Classifying elephant behaviour through seismic vibrations. *Current Biology*, **28**(9), R547–R548.
- Moseley, B., Nissen-Meyer, T., & Markham, A. (2020) Deep learning for fast simulation of seismic waves in complex media. *Solid Earth*, **11**(4), 1527–1549.
- Mousavi, S.M., & Beroza, G.C. (2020) A machine-learning approach for earthquake magnitude estimation. *Geophysical Research Letters*, **47**(1), e2019GL085976.
- Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., & Beroza, G.C. (2020) Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, **11**(1), 1–12.
- Ngene, S.M., Van Gils, H., Van Wieren, S.E., Rasmussen, H., Skidmore, A.K., Prins, H.H. et al. (2010) The ranging patterns of elephants in Marsabit protected area, Kenya: the use of satellite-linked GPS collars. *African Journal of Ecology*, **48**(2), 386–400.
- Palanisamy, K., Singhania, D. & Yao, A. (2020) Rethinking CNN models for audio classification. *arXiv preprint arXiv:2007.11154*.
- Poole, J.H. (2011) Behavioral contexts of elephant acoustic communication. In: *The Amboseli elephants: a long-term perspective on a long-lived mammal*. Chicago: The University of Chicago, pp. 125–161. <https://doi.org/10.7208/chicago/9780226542263.003.0009>
- Randler, C., & Kalb, N. (2018) Distance and size matters: a comparison of six wildlife camera traps and their usefulness for wild birds. *Ecology and Evolution*, **8**(14), 7151–7163.
- Reinwald, M., Moseley, B., Szenicer, A., Nissen-Meyer, T., Oduor, S., Vollrath, F. et al. (2021) Seismic localization of elephant rumbles as a monitoring approach. *Journal of the Royal Society Interface*, **18**(180), 20210264.
- Ren, L., & Hutchinson, J.R. (2008) The three-dimensional locomotor dynamics of African (*Loxodonta africana*) and Asian (*Elephas maximus*) elephants reveal a smooth gait transition at moderate speed. *Journal of the Royal Society Interface*, **5**(19), 195–211.
- Rew, R., & Davis, G. (1990) Netcdf: an interface for scientific data access. *IEEE Computer Graphics and Applications*, **10**(4), 76–82.
- Rost, S., & Thomas, C. (2002) Array seismology: methods and applications. *Reviews of Geophysics*, **40**(3), 2–1.
- Salamon, J., & Bello, J.P. (2017) Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Processing Letters*, **24**(3), 279–283.
- Scheele, B.C., Pasmans, F., Skerratt, L.F., Berger, L., Martel, A., Beukema, W. et al. (2019) Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. *Science*, **363**(6434), 1459–1463.
- Senior, A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T. et al. (2020) Improved protein structure prediction using potentials from deep learning. *Nature*, **577**(7792), 706–710.
- Shamout, F.E., Zhu, T., Sharma, P., Watkinson, P.J., & Clifton, D.A. (2019) Deep interpretable early warning system for the detection of clinical deterioration. *IEEE*

- Journal of Biomedical and Health Informatics*, **24**(2), 437–446.
- Shiu, Y., Palmer, K., Roch, M.A., Fleishman, E., Liu, X., Nosal, E.-M. et al. (2020) Deep neural networks for automated detection of marine mammal species. *Scientific Reports*, **10** (1), 1–12.
- Shorten, C., & Khoshgoftaar, T.M. (2019) A survey on image data augmentation for deep learning. *Journal of Big Data*, **6** (1), 1–48.
- Smit, J., Pozo, R.A., Cusack, J.J., Nowak, K., & Jones, T. (2019) Using camera traps to study the age-sex structure and behaviour of crop-using elephants *Loxodonta africana* in Udzungwa Mountains National Park, Tanzania. *Oryx*, **53**(2), 368–376.
- Springenberg, J.T., Dosovitskiy, A., Brox, T., & Riedmiller, M. (2014) Striving for simplicity: the all convolutional net. *arXiv preprint arXiv:1412.6806*.
- Sugumar, S., & Jayaparvathy, R. (2013) An early warning system for elephant intrusion along the forest border areas. *Current Science*, **104**(11), 1515–1526.
- Sukumar, R. (2003) *The living elephants: evolutionary ecology, behaviour, and conservation*. New York: Oxford University Press.
- Szenicer, A., Fouhey, D.F., Munoz-Jaramillo, A., Wright, P.J., Thomas, R., Galvez, R. et al. (2019) A deep learning virtual instrument for monitoring extreme UV solar spectral irradiance. *Science Advances*, **5**(10), eaaw6548.
- Szenicer, A., Leng, K., & Nissen-Meyer, T. (2020) A complexity-driven framework for waveform tomography with discrete adjoints. *Geophysical Journal International*, **223** (2), 1247–1264.
- Thouless, C., Dublin, H., Blanc, J., Skinner, D.P., Daniel, T.E., Taylor, R.D. et al. (2016) African elephant status report 2016: an update from the African Elephant Database. Occasional Paper Series of the IUCN Species Survival Commission, No. 60 IUCN / SSC Africa Elephant Specialist Group. IUCN, Gland, Switzerland, vi +309 pp.
- Wasser, S.K., Joseph Clark, W., Drori, O., Stephen Kisamo, E., Mailand, C., Mutayoba, B. et al. (2008) Combating the illegal trade in African elephant ivory with DNA forensics. *Conservation Biology*, **22**(4), 1065–1071.
- Wood, J.D., O’Connell-Rodwell, C.E., & Klemperer, S.L. (2005) Using seismic sensors to detect elephants and other large mammals: a potential census technique. *Journal of Applied Ecology*, **42**(3), 587–594.
- Zhu, W., Mousavi, S.M., & Berosa, G. (2020) Seismic signal augmentation to improve generalization of deep neural networks. *Advances in Geophysics*, **61**, 151–177.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Full size version of the leopard image seen in Figure 1.

Figure S2. Spectrogram of a rumble before and after applying the structure tensor. We can see that the locomotion signal is attenuated when applying the structure tensor, thus making the rumble signal more prominent.

Figure S3. Spectrogram of an elephant walking before and after applying the structure tensor. We can see that the locomotion signal is attenuated when applying the structure tensor.

Data S1. The SeisSavanna dataset.