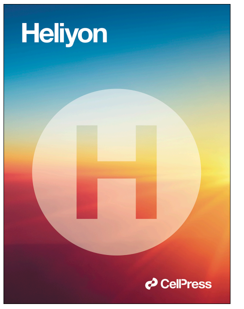
Journal Pre-proof



Ecology & computer audition: Applications of audio technology to monitor organisms and environment

Björn W. Schuller, Alican Akman, Yi Chang, Harry Coppock, Alexander Gebhard, Alexander Kathan, Esther Rituerto-González, Andreas Triantafyllopoulos, Florian B. Pokorny

PII: S2405-8440(23)10350-1

DOI: https://doi.org/10.1016/j.heliyon.2023.e23142

Reference: HLY 23142

To appear in: *HELIYON*

Received Date: 22 June 2023

Revised Date: 8 November 2023

Accepted Date: 27 November 2023

Please cite this article as: , Ecology & computer audition: Applications of audio technology to monitor organisms and environment, *HELIYON* (2024), doi: https://doi.org/10.1016/j.heliyon.2023.e23142.

Ecology & Computer Audition: Applications of Audio Technology to Monitor Organisms and Environment

Björn W. Schuller[1,2,3,*], Alican Akman[1], Yi Chang[1], Harry Coppock[1], Alexander Gebhard[2], Alexander Kathan[2], Esther Rituerto-González[2,4], Andreas Triantafyllopoulos[2], Florian B. Pokorny[2,5]

[1]GLAM – Group on Language, Audio, & Music, Imperial College London, UK
[2]EIHW – Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Germany
[3]audEERING GmbH, Gilching, Germany
[4]GPM – Group of Multimedia Processing, University Carlos III of Madrid, Spain
[5]Division of Phoniatrics, Medical University of Graz, Austria

[*]Correspondence: schuller@ieee.org

Abstract

Among the 17 Sustainable Development Goals (SDGs) proposed within the 2030 Agenda and adopted by all the United Nations member states, the 13th SDG is a call for action to combat climate change. Moreover, SDGs 14 and 15 claim the protection and conservation of life below water and life on land, respectively. In this work, we provide a literature-founded overview of application areas, in which computer audition – a powerful but in this context so far hardly considered technology, combining audio signal processing and machine intelligence – is employed to monitor our ecosystem with the potential to identify ecologically critical processes or states. We distinguish between applications related to organisms, such as species richness analysis and plant health monitoring, and applications related to the environment, such as melting ice monitoring or wildfire detection. This work positions computer audition in relation to alternative approaches by discussing methodological strengths and limitations, as well as ethical aspects. We conclude with an urgent call to action to the research community for a greater involvement of audio intelligence methodology in future ecosystem monitoring approaches.

Keywords Computer audition, audio technology, ecology, environment, organism, call to action

1 Introduction

Our climate is rapidly changing. According to the 2021 Assessment Report of the International Panel for Climate Change (IPCC) "human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years" [1] with CO_2 emissions having been the major driver. This global rise in temperature [2] has impacted the environment in various ways: increased precipitation, rise of sea levels, loss of glacier mass, desertification, heatwaves, and an increased frequency of extreme weather events [1]. A massive loss of natural habitats and biodiversity [3,4], dangers to human health [5] and food security [6], as well as increased armed conflicts [7] are among the consequences.

Accordingly, climate change has been described by many as 'the greatest challenge of our time', calling for an equally outstanding response by the international community. Efforts at governmental and institutional level to curb CO_2 emissions and limit environmental pollution have largely dominated public conversations. Future technological breakthroughs are expected to play a major role in curbing, and even reversing emissions, for example through the development of renewable energy sources [8] and carbon sequestration technology [9].

Artificial intelligence (AI) and its sub-paradigm machine learning (ML) are among those technological concepts with great potential for capturing changes in our ecosystem and assisting the fight against ecological catastrophe [10,11]. The major promise of ML and its latest mainstay deep learning [12] comes from their capacity to automatically analyse vast amounts of complex unknown data based on knowledge previously gained from a set of given training data. This data is acquired from various sensors, such as satellites or cameras, temperature and humidity sensors, or microphones that monitor the acoustic environment. With the advent of cheaper, high-fidelity audio sensors, acoustic data has shown increasing promise in environmental monitoring, sparking the development of a novel research sub-field at the intersection of audio and ecological research – ecoacoustics [13,14].

Ecoacoustic research relies on utilising a set of auditory "indices" (essentially handcrafted features extracted from audio signals), which are indicative of changes in the underlying environment, typically by tracking changes in vocalising species [15]. However, this coarse analysis of soundscape properties fails to capitalise on recent advances in data-driven audio research spearheaded by ML. These advances fall under the umbrella of computer audition (CA) research. CA is that particular sub-field of ML, which encompasses all facets of the auditory information stream, including advances in identifying vocalising species [16] or on disentangling a soundscape into the underlying constituents of biophony, anthropophony, and geophony [17].

Community interest in the auditory monitoring of the environment has been increasingly rising [14]. By end of 2022, there were 907 615 articles indexed in the Web Of Science Core Collection related to ecology (*search term: "ecolog*"*), with the number of articles per year having more than doubled over the last 10 years (33 419 in 2013 vs 69 024 in 2022). 5 237 of these articles deal with AI (*search term: "ecolog*" AND ("artificial intelligence" OR "machine learning" OR "deep learning")*) and were by the majority (≈ 55 %) published in the last two years (2021 or 2022). With a

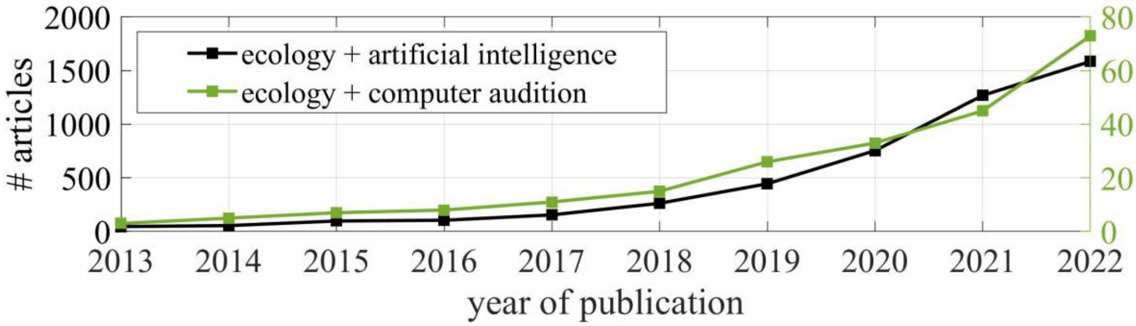


Figure 1: Number of (#) articles indexed in the Web of Science Core Collection over the last 10 full calendar years (as of 3 May 2023) related to ecology and artificial intelligence (black; left y-axis) vs ecology and computer audition (green; right y-axis).

total amount of just 236 Web of Science Core Collection indexed articles (*search term: "ecolog*" AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND (audio OR acoustic* OR sound)*), CA has played a minor role in ecology-related research so far. However, same as for AI in general, there has also been a significant increase of CA-related articles in the context of ecology in recent years (see Figure 1). Exactly half of the 236 articles were published in 2021 or 2022. These numbers indicate that CA represents a relatively underexplored, but at the same time emergent method for ecological applications in the future.

This article aims to give a comprehensive, but non-exhaustive overview of applications, in which CA has already been employed to retrieve information that potentially allows for the identification of ecologically critical processes and states. The catastrophic effects of climate change are spread over different domains of our planet's ecosystem. In our attempt to categorise the different capabilities of CA in the context of ecology, we distinguish between application areas related to organisms and applications related to the physical environment (see Figure 2). Each of the following two sections is dedicated to one of these categories; concrete applications of CA are given in italics at the beginning of paragraphs. Thereafter, we discuss the advantages and limitations of CA for ecology-related applications as compared to alternative methods, and address ethical issues. Finally, we conclude and call to action for further research on automatic audio-based ecosystem information retrieval.

2 Organism Monitoring

2.1 Animal Surveillance

Animal detection and classification. CA has already been used to detect and classify several animals in the past [18–20]. The availability of audio recordings over a wide area can not only identify which animals are present in the region, but also roughly how many of them and where they are moving. Animal monitoring therefore presents a great opportunity to use CA [18].

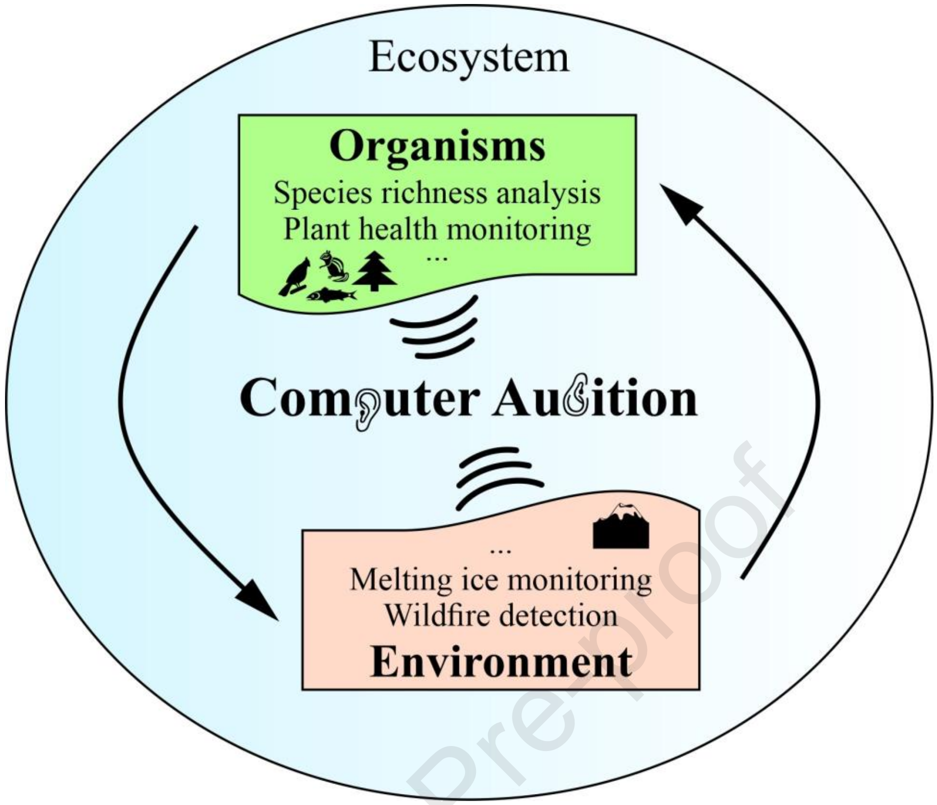


Figure 2: Categorisation of ecology-related application areas for computer audition as adopted in this work.

Animal movement and population analysis. Changes in the movement of animals could indicate possible dangers such as natural disasters [21]. However, animal movement profiles can also be used to analyse population dynamics at different locations. This would provide insight into the evolution of species populations and give the opportunity to save endangered species from extinction through early interventions.

2.1.1 Land Mammals

Animal trait analysis. Land mammals produce diverse sounds in nature, such as when defending territories, fighting with other animals, or during mating season [22]. With the help of these sounds, zoologists can gain a variety of valuable insights into the animals' way of life and use this information for further analyses. Moreover, the automatic characterisation of biological traits in animals, such as by sex and age groups, can help monitor the health of packs and herds in the wild [22].

*Species richness analysis.*A major problem of our time is the change in biodiversity and species extinction. It is a great advantage to be able to cover extensive areas with microphones to gain insights of animal populations. Audio recording and subsequent automatic audio analysis are particularly useful in places where it is difficult to make full-coverage camera recordings, such as in forests. The presence and composition of different animal sounds provides information about the composition of vocalising species and the orthopteran species richness, as well as about the structure of the landscape and the intensity of land use [23].

Detection of danger for animals. Another application area of CA is the protection of animals. Especially in agriculture, attacks by wild animals are a major hazard. Therefore, several solutions have already been developed in the field of computer vision to detect relevant situations [24–26]. Nevertheless, visual solutions have the problem that cameras usually cannot fully cover large pastures. CA offers the advantage that large fields for animals can be fully ranged, recording audio with microphones. In case of a foreign animal entering the pasture or, at the latest, when the herd is attacked by wild animals, the farmer can be informed immediately. Even an attack by humans on endangered animals – e. g., by poachers [27] – can be detected by means of CA at an early stage through gunshot recognition [28,29].

2.1.2 Maritime Life

Recording sounds with microphones under the water surface covers a larger space than, for example, diving in with cameras. In water, sound propagates with a higher velocity and over greater distances than in air. This is conversely to vision, which is dramatically hindered in water.

Coral reef integrity evaluation. Coral reefs are called the ‘rainforests of the sea’ [30] and represent ecosystems of high biodiversity. Underwater soundscapes are recently being used to monitor coral reefs [31]. The practice of dynamite fishing causes severe harm to the reefs [32,33]. Such endangering human activity could be automatically ‘heard’ by a CA system. In addition, due to rising sea temperatures and the acidification of the oceans, mass ‘bleaching’ of coral reefs is now a common phenomenon. A healthy coral reef can be acoustically distinguished from a damaged one, as its bright, loud, and diverse soundscapes guides the recruitment of reef organisms [34].

Whale monitoring. By sensing and analysing underwater animal sounds, such as sounds produced by whales for communication purposes or caused by their movement patterns, conclusions can be drawn with respect to their population, behaviour, and habitat [35–37]. Large bioacoustic archives like the Orchive [38,39] represent a useful data foundation for a CA system [40,41].

2.1.3 Birds

Birds represent organisms which are rapidly affected by changes in the environment. In the mining industry, caged canaries were carried down by miners into the mine tunnels. Whenever there was a leak of dangerous gases, such as carbon monoxide, the gases killed the canaries, which served as a warning for the miners to exit the tunnels immediately [42].

Birds are messengers that tell us about the health of the planet because they are widespread, they connect habitats, resources and biological processes. They also contribute to ecosystem services – as natural enemies of pests, pollinators of fruit, and seed transporters [43]. Birds also play a key role in cycling nutrients and helping to fertilise marine ecosystems [44]. Whether ecosystems are made for agricultural production, wildlife or water, success can be measured by the health of birds.

Ecosystem evaluation. Bird sound classification aids to determine their presence, tracking their migrations, and measuring their population [16,19,45]. A decline in bird numbers informs about a damaged environment [46], e. g., due to habitat fragmentation and destruction, pollution, and pesticides introduced species, etc. Furthermore, birds provide insect and rodent control, which results in tangible benefits to humans. Insect outbreaks can annually have a huge negative economic impact in agricultural and forest products, and some birds can be effective to substantially reduce insect pest populations without the health, environmental, and economic risks of harmful pesticides [47]. Microphone arrays can record bird data in a continuous manner, where other sensors such as cameras would struggle – e. g., for the tracking of nocturnally migrating birds [48], night vision cameras (requiring expensive complex manufacture with high-voltage power supplies to operate) would be required. The acoustic performance of bird communities reaches its maximum at dawn and dusk, when species are contemporarily singing and producing choruses [49]. Measuring the length, energy, and frequency components of choruses can, e. g., reveal the ambient temperature [50] and, thus, the subsequent changes caused in the physiology of organisms.

2.1.4 Insects

Insects are essential in food chains and cycles; they pollinate fruits, flowers, and vegetables, and are also very important as primary or secondary decomposers. Insects are under immense pressure from land use intensification and climate change effects, threatened with extinction or showing significant population declines [51]. Many insects are omnivorous, they eat a variety of foods including plants, fungi, dead animals, and decaying organic matter. Thereby, insects help breaking down and disposing wastes. Predatory or parasitic insects help keep pest populations, such as other insects or weeds, at a tolerable level. They are also the sole food source for many amphibians, reptiles, birds, and mammals. Further more, CA could give us relevant insights into insect control.

Hive integrity monitoring. Bees contribute to complex, interconnected ecosystems that allow a diverse number of different species to co-exist [52]. Acoustical non-intrusive sensors are being introduced along with temperature and moisture sensing for subsequent machine learning-based bee hive colony activity health and status monitoring [53].

Pest infestation recognition. In recent decades, acoustic approaches have provided non-destructive, remote, automated detection and monitoring of insect and pest infestations for pest managers, regulators, and researchers [54]. Microphones are useful sensors for airborne signals, specially ultrasonic sensors, which are particularly effective for detecting wood-boring pests like termites at frequencies of more than 20 kHz.

Pest control assistance. Insect pests can also pose a serious threat to agricultural and forest ecosystems, but they are difficult to control. Novel research on methods for acoustic data analysis based on active sound production by larvae (i. e., stridulations) can give insight into larval ecology produced by pests [55] and opens up new pest control avenues.

2.2 Plant Bioacoustics

The field of plant bioacoustics is a rising exploration field focussing on measurement and interpretation of sounds and vibrations created by or within plants, with special interest to signals produced by plant-dwelling insects/animals.

Plant health monitoring. Bioacoustic tools have been applied to measure mechanical properties of plant structures, optimise mechanical harvesting, and detect the distribution of root systems, as well as to monitor plant health [56], photosynthesis, and ecology [57]. Recent experimental studies open the possibility of assessing the stress of plants by using ML algorithms on acoustic signals, e. g., analysis of ultrasound emitted by plants to determine their health [56,58]. In addition, acoustic and vibration sensors are used by entomologists to detect hidden infestations of invasive insect species, and to monitor insect movement, feeding, and mating activities on host plants [59], and as a reliable assessment of hidden pest infestations, including invasive insect species of importance for plant biosecurity [56].

Drought detection. The acoustic performance of some organisms shows the magnification of the effects of climate change. Drought [60], for example, produces stress on trees and leads to an increased vulnerability to insect attacks [61] as insects are drawn to stressed trees using chemical signals, but also are attracted by the sounds emitted by tree cells [62]. These sounds, which are produced by forest trees when being under drought stress, are known as cavitation, which is the result of cells collapsing by gradual dehydration. The majority of these sounds emitted are within a frequency range of 20 kHz to 200 kHz [61] and carry information for insects, that can perceive such signals.

Defoliation monitoring. Defoliating insects have a large impact on ecosystems and are influenced by climate change as well [63]. Therefore, changes of their behaviour can be used as an indicator of, e. g., an increasing amount of CO_2 in the atmosphere. Early detection of these changes is of great importance. Thus, this acoustic feedback from insects can have a positive effect, as it can be detected by CA and thus indicate defoliation, forest decline, and CO_2 increase.

3 Environment Monitoring

Identification of short- and long-term environmental changes: detection of natural disasters – deforestation recognition. Environmental changes are becoming increasingly rapid and have an impact on many areas of life. Therefore, it is important that these changes are identified at an early stage so that appropriate countermeasures can be taken. Such environmental changes can be natural changes in the form of natural disasters, such as earthquakes or volcanic eruptions [64], but also man-induced changes [65], like deforestation [66]. Earthquakes induce sounds in various natural objects, which can be detected [67]; and the progressing deforestation of the rain forest can be ‘heard’ by recognising and analysing sounds of chainsaws or large machines.

3.1 Water

Water serves as nutritional source [68], as a habitat for various animals [69,70], or can simply be used to maintain hygiene and therefore avoid several diseases [71,72]. Every living being on our planet needs water to survive. Therefore, we have to ensure the quality and quantity of this vital resource.

Water can be encountered in three physical states: solid, liquid, and gaseous. Thus, we have various ways to make use of this element as an input for a CA system.

3.1.1 Melting Ice

Melting ice monitoring. In times of melting polar ice caps [73] it would be helpful to monitor the process of the melting ice. Cracking sounds in icebergs, ice floes, or glaciers can serve as acoustic indicators [74–76] detectable and interpretable by means of CA. On the one hand, the monitoring of melting ice can be important for study purposes, e. g., to gain knowledge about how long the melting process lasts before a chunk of ice breaks down from a glacier [77]. On the other hand, it can be utilised for the prediction of avalanches or floods and, therefore, enable precautions or appropriate countermeasures. Melting glaciers can also cause enormous landslides and tsunamis [78,79]. According to researchers, there is a ticking time bomb at the moment in Alaska within the Barry Arm area, the Barry glacier, which has the potential of causing a mega-tsunami [80].

3.1.2 Floods

Flood prediction. Not only the sound of cracking ice might be helpful in predicting floods or flood waves, but also the sound of flowing water (e. g., in rivers), since this sound reflects the water flow velocity [81,82]. The sound is generally generated by particle collisions through streamed sediment movements as well as the flow of water over submerged obstructions [82]. These factors might be very helpful for roughly inferring the amount of flowing water and predicting overflowing rivers and lakes in order to appropriately prepare surrounding areas for such crises. While flood prediction models (even ML-based ones) already exist [83,84], they might not have the ideal temporal or spatial granularity for dealing with flash floods, especially in densely populated areas. Complementing existing systems with area-wide auditory sensing could facilitate a more comprehensive real-time monitoring of how flooding spreads throughout a particular city or neighborhood and help authorities adapt their evacuation protocol accordingly.

3.1.3 Water Supply

Water scarcity prediction. In addition to the early prediction of floods due to the sound of running water, CA could also predict the opposite. That is, if there is minimal audible sound it might be a clue for drying up rivers, lakes or natural fountains and, thus, be a sign of imminent water shortage in a certain region. Via early prediction of water scarcity in specific areas, artificial irrigation facilities could be constructed in advance or the people living in such areas could be relocated.

3.2 Meteorological Phenomena

The mass of air that we denominate as the atmosphere represents only about 5% of the total volume of our planet, but is crucial for all forms of life on it. The discipline of meteorology describes the study of atmospheric processes including weather phenomena. Acoustic measurements can give relevant meteorological information, such as air pressure-related measures and wind characteristics [85].

Wind analysis. The movement of winds have huge implications for storm systems and precipitation patterns. Specifically, winds transport dust from desert regions to faraway locations, making changes in the environment [86]. Recording aeroacoustic noise generated by wind flowing past a microphone, and decomposing the acoustic spectra into low-frequency components, can provide wind speed and wind orientation. A recent study analysed the frequencies composing the wind-induced acoustic signal measured by microphones [85]. The acoustic spectra recorded under a wind flow can be decomposed into low-frequency components, mainly reflecting the wind velocity, and higher frequency components, regarded to depend on the wind direction relative to the microphone. Therefore, CA as approach to monitor the wind has a huge potential to show climate disruptions and provide potential help, adaptively controlling, for instance, energy-generating wind mill farms and at the same time recognising potential disruptions in their routine.

Hurricane and tornado detection. Numerous geophysical and anthropogenic events, such as hurricanes and tornadoes, emit infrasound, i. e., acoustic waves below the human hearing range of about 20 Hz that can be captured by low frequency microphones. The rate of severe storm environments becomes greater in the northern hemisphere due to temperatures rises [87]. Tornado-producing storm systems emit infrasound up to 2 hours before tornado genesis, which can be detected from large distances (in excess of 150 km) due to weak atmospheric attenuation at these frequencies. Thus, infrasound could be used for intelligent, long-range, passive monitoring and detection of tornado genesis, as well as for the delineation of tornado properties [88].

3.3 Fire

The world's flora and fauna are under threat from the increased frequency and strength of wildfires, that also more and more affect residential areas causing immense damage to communities.

3.3.1 Wildfires

Wildfire detection and classification. Wildfire seasons are getting longer and more intense [89], and fire brigades need to figure out how to best spread their limited resources. In that respect, crown fires, which burn through the upper layers of trees, are more intense and have a higher velocity than surface or ground fires. Early detection as well as distinguishing between different types of fires are crucial for combating large wildfires, as it determines the type of response needed. Image information can be utilised for fire detection and categorisation. However, in practice, categorisation is hampered by limited visibility due to smoke. In contrast, temperature, humidity, and smoke sensors are tailored to detect the presence of fire, but are insensitive to its type. To that end, CA can be used to 'hear' fire from a large distance, while acoustic fire properties have been also previously shown to vary across different fire types, thus, enabling their audio-based classification [90,91]. Furthermore, audio information captured via microphone arrays can be used to determine the location of lightning thunder sources, since lightning strikes are one of the major causes of forest fires.

Recognition of reignition. A further challenging factor is that fires naturally disrupt monitoring systems put in place too close, as the fire itself, or the water used to extinguish it, damages the sensors. Thus, immediately after a fire, affected areas are left without proper sensory coverage. This constitutes a crucial risk as the reignition of fires in already burnt-out areas is a major problem for firefighters. A poignant example is the August 2021 fire of Varympompi (near Athens, Greece)[1], which was initially controlled by the fire brigade, only to be reignited a few hours later due to insufficient supervision, with the majority of its destruction coming with the second wave. This example illustrates that it is imperative to rapidly (re-)deploy sensing equipment in allegedly cleared areas. Those areas, however, might be heavily affected by smoke (especially if the fire is still ongoing in nearby land), making it hard for image or smoke sensors to detect potential sources of rekindling. Therefore, CA might be a good option for this scenario as well. However, domain adaptation is going to be an issue as the sound of fire might differ between forests.

3.3.2 Fire Damage Evaluation

Communities around the world struggle to reconvene their lives in the aftermath of a catastrophic fire, especially if the fire affected residential areas. One of the usual reaction to such fires is the promise to rebuild all destroyed or affected buildings. Unfortunately, building is a major source of CO_2 emissions, necessitating an environmentally-friendly rebuilding paradigm. This includes the effort to salvage as much as possible from the remnants of an urban fire.

Structural integrity evaluation. A major consideration after a building fire is its effect on structural integrity, which is the ability of a structure to withstand the required load without collapsing. Determining the extent of damage, however, is not an easy feat, especially as any investigation should be conducted by means of non-invasive techniques in order not to further compromise the structure. Generally, a building's integrity is dependent on the strength of the materials the building is made from, which is in turn reflected by the way sound propagates through the materials. Thus, CA presents a novel avenue of investigating changes in material strength, which can provide useful information on the damage a building has sustained. As a recent example of such work, Schabowicz and colleagues [92] studied the condition of materials subjected to fire, and utilised acoustics for identifying the degree of degradation of fibre-cement boards.

4 Discussion

The identified applications of CA in the context of ecology demonstrate the potential of this so far largely disregarded methodology to allow for a more comprehensive picture of our planet's ecosystems and to facilitate a prevention of nature and living beings from harm. Even though the number of studies on ecological information retrieval by means

[1]https://go.ifrc.org/reports/14615 (as of 15 July 2022)

of CA is comparatively low – some work presented here dealt with acoustics, but did not employ machine learning methodology, our overview shows that we are steering the right course by being open-minded for novel approaches at a time where our planet undergoes significant changes and we humans need to find solutions for capturing these changes to countersteer or timely protect ourselves and other organisms from their consequences. However, what are the advantages of CA over alternative approaches? What are its limitations? Are there considerations from an ethics perspective? Is CA already set to reasonably contribute? In the remainder of this work, we aim to give answers to these questions.

4.1 Computer Audition vs Alternatives

There are several ways to capture ecological processes on our planet with audio recording being just one example suited for intelligent/ML-based analysis. Visual sensing is another example of established data collection for intelligent computer-based analysis, i. e., computer vision. In addition, various other sensing alternatives, e. g., other physical or chemical sensors, exist that can be used to gather input data for intelligent organism or environment monitoring as well as associated harm detection systems. Finally, sensors can be compared with the non-technology-based abilities of the human individual. In Table 1 we compare acoustic sensing with other eligible sensing modalities for subsequent ML analysis on the basis of specific criteria. In doing so, we indicate to the best of our knowledge, whether the different modality-specific entries can be regarded as an advantage, a disadvantage, or as neutral, respectively.

4.2 Limitations

In Table 1 we benchmarked audio against other data modalities as the input for intelligent ecology monitoring systems. According to this comparison, audio has certain advantages over other modalities, such as low sensor costs in combination with a wide range of recording coverage, richness of information within one recording, and a number of pre-trained ML models available. However, we have to point out that our comparison is based on the theoretical assumption of an optimal recording setting and an ideal data storage, transmission, pre-processing, and analysis workflow. Unfortunately, this assumption is usually not fulfilled for ‘in-the-wild’ scenarios. For example, one challenge with regard to audio might be the installation of microphones in ‘free nature’ and the long-term power supply of the recording device as well as a potentially necessary data transmission module operating across a wide temperature range and different (possibly extreme) weather conditions. Another challenge might be the automatic separation of application-relevant audio information from (background) noise. Nevertheless, the question if audio represents a suitable modality for an (intelligent) organism and/or environment monitoring scenario generally depends on the specific task of interest. For many tasks, a combination of different sensing modalities and, thus, an extension of some existing approaches for audio might be ideal.

In the following, we disclose limitations that are not specific to audio, but inherent to intelligent data analysis in general.

Table 1: Comparison of acoustic sensing for intelligent organism and environment monitoring vs other sensing modalities with regard to seven key criteria. Classification by the authors: +/green shading = advantageous, +-/orange shading = neutral, -/red shading = disadvantageous; m = meter(s); ML = machine learning

Criteria \ Modalities	Acoustic Sensing of airborne sound including infrasound, sound in human audible frequency range, and ultrasound using microphones	Visual Optical sensing by means of 2D, 3D, high-speed, thermal, aerial, and microscope cameras	Other physical Sensing of physical parameters other than airborne sound and visual information, such as temperature, current, humidity, fluid level, acceleration, pressure, and solid-borne sound	Chemical Sensing of chemical information, such as analyt composition, presence of specific elements, element concentration, and chemical activity	Human Non-technology-based sensing by means of human sight, hearing, smell, taste, and touch
Data throughput Maximum amount of data that can be transmitted/processed per time	+ Extent of audio data is modest most of the time	+- 2D images tend to be petite, while 3D images or videos tend to be comparatively large	+ Data stream is manageable in size	+ Data stream is manageable in size	+- Sensing continuous-ly; however, cannot really be transmitted from one human to another or to a computer
Covered area Sensor range regarding spatial coverage	+ Several hundred to several thousand m^2 (e. g., thunder); 360° recording angle possible (omnidirectional microphone)	+ Variable dependent on specific camera type; especially high range for aerial/satellite cameras; limited angle of view	+ From focused on one location but poten-tially representative for a bigger area, e. g., temperature, to nearly global cover-age in satellite use	+- Very focused on one location; sometimes representative for a bigger area, e. g., soil composition, atmos-pheric chemistry (via satellites) [93]	+- Depends on the modality (vision, touch, etc.); can be large but also quite limited
Privacy Extent of personal data collected by the sensor; importance/possibility of data anonymisation [94]	+- Critical in case human voice is recorded [95]; not applicable for most scenarios in the context of ecology	+- Critical in case human faces are recorded [96]; not applicable for most scenarios in the context of ecology	+ Not critical	+ Critical blood or saliva analysis, or genetic sequencing not applicable in context of nature	+ Unproblematic as no human data are recorded at all
Info richness Amount of information extractable from recorded data	+ Many sound sources possible within one audio recording, e. g., animal sounds, rain, cars, etc.	+ Many objects, classes, locations can be captured in one photo or video	+- Mostly built to sense specific information, i. e., only the desired information is recorded	+- Mostly built to sense specific information, i. e., only the desired information is recorded	+ Due to synchroneous multimodal sensing of the human body, lots of information is captured and processed in the brain
ML models Availability of (pre-trained) machine learning models	+ Many pre-trained models available, which use raw audio or acoustic features as input	+ Many pre-trained models available, which use raw video or visual features as input	+ Several available models [97–99]	- Very few to none existent models with respect to climate	- No models available
Costs Costs incurred by sensor	+ Microphones are very cheap; even specialised microphones are not too expensive	+- Some cameras can be relatively cheap, others are very expensive; satellites or microcameras are extremely expensive	+ Mostly relatively cheap	+ Sensor materials/resources can be expensive, i. e., sensor costs vary a lot	+ The human body needs no further sensors; auxiliary means such as glasses are not too expensive
Pollution Pollution caused by sensor	+ Each sensor pollutes the environment to some degree; can be reused very often	+- Shooting satellites into orbit emits lots of burnt gas and precipitates debris [100,101]	+ Each sensor pollutes the environment to some degree; can be reused quite often	+- Each sensor pollutes the environment to some degree; can be reused only sometimes	+ Presumably the most environment-friendly option

4.2.1 Generalisability

One of the major critiques of ML systems – including CA systems – is their (lack of) ability to generalise. This is especially relevant for the applicability of models in real world settings. Moreover, the deployment environment of an intelligent system should be identical to the training environment – a constraint that is hard to satisfy without vastly increasing the amount of bias-free training data. Alternatively, domain adaptation algorithms [102] can be used to explicitly minimise the discrepancy between source and target domains. For intelligent audio applications in particular, the notion of generalisation is closely linked to that of robustness to different perturbations. Traditionally, robustness has been studied under the auspices of speech enhancement [103], where (human) speech constitutes the signal of interest and (environmental) noise the unwanted interference that needs to be removed. However, in our case the opposite is required – human voices would need to be removed in order to get more robust measurements of the environmental conditions [104].

4.2.2 Efficiency

Data is the fuel that drives contemporary AI applications. Bigger and better data usually leads to bigger and better models, which require (vastly) more computational power to be trained [105]. It is thus imperative to seek more resource efficient approaches for AI in general and CA in particular, including the ability to learn from sparse data. Two learning paradigms that may be of use here are those of transfer learning [106] and zero-shot learning [107,108]. While transfer learning describes the approach of reusing learned knowledge/a pre-trained model on a new task, the method of zero-shot learning allows a model to classify samples for whose class no example data were available during the training phase. In general, the field of sustainable AI [109] is about developing AI that is compatible with sustaining environmental resources for current and future generations, while keeping in mind that there are environmental costs to AI itself.

4.2.3 Bias

All real-world AI technologies are plagued by bias. Audition in particular is affected, as sources of signal are difficult to determine and can come from a wide range of generative factors. The typical AI setup involves sampling from the real world, creating a dataset, and training a model to perform a task in this sampled space. Through identifying spurious correlations and bias from study design sampling methods, AI models are often able to perform well in these sampled datasets. This wrongly suggests a true underlying signal and so opportunity for AI to help in a field. This has been demonstrated in the recent COVID-19 pandemic, where early research suggested COVID-19 was uniquely identifiable from infected individuals' respiratory sounds [110]. Moving forward, CA for the environment should be directed down biologically plausible avenues and significant work should be done on study design and bias mitigation.

4.3 Ethical Considerations

4.3.1 Privacy

Data protection is a critical issue in the era of big data [111]. It regards the possibility of mass surveillance, especially in applications such as large-scale environmental monitoring. A widespread deployment of sensors, in our case microphones, in combination with an ML system potentially violates the privacy of citizens around the world by collecting, storing, sharing, and analysing data without their knowledge or permission. A straightforward solution for this problem is to automatically remove all speech information immediately after the recording and before any subsequent processing is done [104]. However, this strategy might lead to suboptimal results as any source separation algorithm invariably introduces unwanted artefacts and information loss.

4.3.2 Fairness

When intending to deploy an ML system for a global endeavour such as tackling ecological challenges, it is essential to provide adequate performance guarantees for all parts of our planet's ecosystem, irrespective of where in the world those may be [112]. This becomes particularly pressing as the areas of the world that are most in danger, tend to be more strained for resources. Unfortunately, it is easy to imagine a scenario where the richest countries collect most of the data within their borders, leading to an underrepresentation of the world's more vulnerable countries in the training, and, consequently, to the underperformance of the algorithms when deployed on their premises.

4.3.3 Trustworthiness

While AI is often touted as a technology capable of operating completely autonomously, AI applications are usually embedded in an ecosystem involving other entities, such as humans. These entities have to trust the AI system under consideration of known limitations. For humans, trustworthiness is largely related to explainability [113], i. e., the understanding of how the machine came to a certain decision.

4.3.4 Beneficiary

Some applications outlined in this work seem to be beneficial for human individuals at first glance, such as the early detection of a tornado. However, in this work, we regard the basic need of mankind to further populate the Earth as its habitat as given and, under these circumstances, the prevention of human beings from physical harm or material damage might prevent costly and pollutive medical interventions or the resources-consuming replacement of destroyed goods. Thereby, we followed the dominant paradigm, putting humanity and its interests in the centre of discussion, but we acknowledge the need for a holistic consideration of our natural ecosystem.

5 Conclusion

In this article, we first pointed out that there is a lack of research on intelligent acoustic solutions for monitoring ecological phenomena. In order to motivate more work in this field, we gave an overview of applications, in which CA has already been employed in the context of ecology or in which CA could be employed in future as acoustic sensing (without subsequent ML stage) has already shown promise. We depicted that CA has the potential to complement alternative methods due to properties like a cost-effective coverage of large areas and the availability of pre-trained ML models. Moreover, in the audio domain certain ecological changes become apparent before they are noticeable in other modalities. This allows CA for an early detection of potentially harmful, ecologically critical processes and states. However, more work needs to be done to exploit the full potential of audio intelligence in this area. In particular, it is important to overcome some methodological limitations, while always meeting legal requirements and respecting ethical values. The combined processing of audio data and data from other domains in intelligent multimodal systems would be worth striving for in the future as well.

With this work, we would like to convince researchers of the potential of audio intelligence for ecology-related monitoring tasks, in times in which our planet undergoes a severe transformation and we need to tackle the situation in order to prevent nature and living beings from harm.

Acknowledgements

This work was funded by the German Research Foundation (DFG; Reinhart Koselleck project, No. 442218748: AUDI0NOMOUS). The authors want to thank everyone who takes responsibility for our beautiful planet and makes an effort to prevent nature from harm. Zero GPU hours were consumed to provide the contents discussed herein.

Conflict of Interest Statement

The author Björn W. Schuller was employed by the company audEERING GmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This manuscript has been released as a preprint in an extended version at arXiv.org [114].

Data Availability Statement

No data was used for the research described in the article.

References

[1] V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, *et al.*, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth*

Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, 2021.

[2] P. Stassen, "Global warming then and now," *Nature Geoscience*, vol. 9, no. 4, pp. 268–269, 2016.

[3] O. Hoegh-Guldberg and J. F. Bruno, "The impact of climate change on the world's marine ecosystems," *Science*, vol. 328, no. 5985, pp. 1523–1528, 2010.

[4] G. Ceballos, P. R. Ehrlich, and R. Dirzo, "Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines," *Proceedings of the National Academy of Sciences*, vol. 114, no. 30, pp. 6089–6096, 2017.

[5] N. Watts, W. N. Adger, P. Agnolucci, J. Blackstock, P. Byass, W. Cai, S. Chaytor, T. Colbourn, M. Collins, A. Cooper, *et al.*, "Health and Climate Change: Policy responses to protect Public Health," *The Lancet*, vol. 386, no. 10006, pp. 1861–1914, 2015.

[6] T. Wheeler and J. Von Braun, "Climate Change impacts on global food security," *Science*, vol. 341, no. 6145, pp. 508–513, 2013.

[7] C. Raleigh and H. Urdal, "Climate Change, Environmental Degradation and Armed Conflict," *Political Geography*, vol. 26, no. 6, pp. 674–694, 2007.

[8] N. Panwar, S. Kaushik, and S. Kothari, "Role of renewable energy sources in environmental protection: A review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 3, pp. 1513–1524, 2011.

[9] J. D. Figueroa, T. Fout, S. Plasynski, H. McIlvried, and R. D. Srivastava, "Advances in CO2 capture technology—the US Department of Energy's Carbon Sequestration Program," *International Journal of Greenhouse Gas Control*, vol. 2, no. 1, pp. 9–20, 2008.

[10] J. Cowls, A. Tsamados, M. Taddeo, and L. Floridi, "The AI Gambit—Leveraging artificial intelligence to combat climate change: Opportunities, challenges, and recommendations," *AI & Society*, pp. 1–25, 2021.

[11] D. Rolnick, P. L. Donti, L. H. Kaack, K. Kochanski, A. Lacoste, K. Sankaran, A. S. Ross, N. Milojevic-Dupont, N. Jaques, A. Waldman-Brown, A. S. Luccioni, T. Maharaj, E. D. Sherwin, S. K. Mukkavilli, K. P. Kording, C. P. Gomes, A. Y. Ng, D. Hassabis, J. C. Platt, F. Creutzig, J. Chayes, and Y. Bengio, "Tackling climate change with machine earning," *ACM Computing Surveys*, vol. 55, no. 2, pp. 1–96, 2022.

[12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT press, 2016.

[13] A. Farina and S. H. Gage, *Ecoacoustics: The ecological role of sounds*. John Wiley & Sons, 2017.

[14] J. Xie, K. Hu, M. Zhu, and Y. Guo, "Data-driven analysis of global research trends in bioacoustics and ecoacoustics from 1991 to 2018," *Ecological Informatics*, vol. 57, p. 101068, 2020.

[15] M. Towsey, J. Wimmer, I. Williamson, and P. Roe, "The use of acoustic indices to determine avian species richness in audio-recordings of the environment," *Ecological Informatics*, vol. 21, pp. 110–119, 2014.

[16] J. Xie, Y. Zhong, J. Zhang, S. Liu, C. Ding, and A. Triantafyllopoulos, "A review of automatic recognition technology for bird vocalizations in the deep learning era," *Ecological Informatics*, p. 101927, 2022.

[17] M. L. Clark, L. Salas, S. Baligar, C. A. Quinn, R. L. Snyder, D. Leland, W. Schackwitz, S. J. Goetz, and S. Newsam, "The effect of soundscape composition on bird vocalization classification in a citizen science biodiversity monitoring project," *Ecological Informatics*, vol. 75, p. 102065, 2023.

[18] R. Gibb, E. Browning, P. Glover-Kapfer, and K. E. Jones, "Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring," *Methods in Ecology and Evolution*, vol. 10, no. 2, pp. 169–185, 2019.

[19] D. Stowell, M. D. Wood, H. Pamuła, Y. Stylianou, and H. Glotin, "Automatic acoustic detection of birds through deep learning: The first bird audio detection challenge," *Methods in Ecology and Evolution*, vol. 10, no. 3, pp. 368–380, 2019.

[20] D. Nieto-Mora, S. Rodríguez-Buritica, P. Rodríguez-Marín, J. Martínez-Vargaz, and C. Isaza-Narváez, "Systematic review of machine learning methods applied to ecoacoustics and soundscape monitoring," *Heliyon*, 2023.

[21] M. Garstang and M. C. Kelley, "Understanding animal detection of precursor earthquake sounds," *Animals*, vol. 7, no. 9, p. 66, 2017.

[22] D. T. Blumstein, D. J. Mennill, P. Clemins, L. Girod, K. Yao, G. Patricelli, J. L. Deppe, A. H. Krakauer, C. Clark, K. A. Cortopassi, *et al.*, "Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus," *Journal of Applied Ecology*, vol. 48, no. 3, pp. 758–767, 2011.

[23] S. Müller, M. M. Gossner, C. Penone, K. Jung, S. C. Renner, A. Farina, L. Anhäuser, M. Ayasse, S. Boch, F. Haensel, *et al.*, "Land-use intensity and landscape structure drive the acoustic composition of grasslands," *Agriculture, Ecosystems & Environment*, vol. 328, p. 107845, 2022.

[24] N. Andavarapu and V. K. Vatsavayi, "Wild-animal recognition in agriculture farms using W-COHOG for agro-security," *International Journal of Computational Intelligence Research*, vol. 13, no. 9, pp. 2247–2257, 2017.

[25] T. Petso, R. S. Jamisola, and D. Mpoeleng, "Review on methods used for wildlife species and individual identification," *European Journal of Wildlife Research*, vol. 68, pp. 1–18, 2022.

[26] V. K. Vatsavayi and N. Andavarapu, "Identification and classification of wild animals from video sequences using hybrid deep residual convolutional neural network," *Multimedia Tools and Applications*, vol. 81, no. 23, pp. 33335–33360, 2022.

[27] J. Kamminga, E. Ayele, N. Meratnia, and P. Havinga, "Poaching Detection Technologies -— A Survey," *Sensors*, vol. 18, no. 5, p. 1474, 2018.

[28] J. Turian, J. Shier, H. R. Khan, B. Raj, B. W. Schuller, C. J. Steinmetz, C. Malloy, G. Tzanetakis, G. Velarde, K. McNally, M. Henry, N. Pinto, C. Noufi, C. Clough, D. Herremans, E. Fonseca, J. Engel, J. Salamon, P. Esling,

P. Manocha, S. Watanabe, Z. Jin, and Y. Bisk, "Hear 2021: Holistic evaluation of audio representations," *arXiv preprint arXiv:*, 2022.

[29] L. K. Katsis, A. P. Hill, E. Pina-Covarrubias, P. Prince, A. Rogers, C. P. Doncaster, and J. L. Snaddon, "Automated detection of gunshots in tropical forests using convolutional neural networks," *Ecological Indicators*, vol. 141, p. 109128, 2022.

[30] N. Knowlton, "The future of coral reefs," *Proceedings of the National Academy of Sciences*, vol. 98, no. 10, pp. 5419–5425, 2001.

[31] T.-H. Lin, T. Akamatsu, F. Sinniger, and S. Harii, "Exploring coral reef biodiversity via underwater soundscapes," *Biological Conservation*, vol. 253, p. 108901, 2021.

[32] L. M. Slade and B. Kalangahe, "Dynamite fishing in tanzania," *Marine Pollution Bulletin*, vol. 101, no. 2, pp. 491–496, 2015.

[33] S. Wells, "Dynamite fishing in northern Tanzania – pervasive, problematic and yet preventable," *Marine Pollution Bulletin*, vol. 58, no. 1, pp. 20–23, 2009.

[34] T. Lamont, B. Williams, L. Chapuis, M. Prasetya, M. Seraphim, H. Harding, E. May, N. Janetski, J. Jompa, D. Smith, A. Radford, and S. Simpson, "The sound of recovery: Coral reef restoration success is detectable in the soundscape," *Journal of Applied Ecology*, 12 2021.

[35] W. C. Cummings and D. V. Holliday, "Sounds and source levels from bowhead whales off pt. barrow, alaska," *The Journal of the Acoustical Society of America*, vol. 82, no. 3, pp. 814–821, 1987.

[36] H. Klinck, S. L. Nieukirk, D. K. Mellinger, K. Klinck, H. Matsumoto, and R. P. Dziak, "Seasonal presence of cetaceans and ambient noise levels in polar waters of the north atlantic," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. EL176–EL181, 2012.

[37] B. W. Schuller, A. Batliner, C. Bergler, F. Pokorny, J. Krajewski, M. Cychosz, R. Vollmann, S.-D. Roelen, S. Schnieder, E. Bergelson, A. Cristià, A. Seidl, L. Yankowitz, E. Nöth, S. Amiriparian, S. Hantke, and M. Schmitt, "The INTERSPEECH 2019 Computational Paralinguistics Challenge: Styrian Dialects, Continuous Sleepiness, Baby Sounds & Orca Activity," in *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH)*, (Graz, Austria), pp. 2378–2382, ISCA, 2019.

[38] S. Ness and G. Tzanetakis, "Human and machine annotation in the Orchive, a large scale bioacoustic archive," in *Proceedings of the Global Conference on Signal and Information Processing (GlobalSIP)*, (Atlanta, USA), pp. 1136–1140, IEEE, 2014.

[39] S. Ness, H. Symonds, P. Spong, and G. Tzanetakis, "The Orchive: Data mining a massive bioacoustic archive," *arXiv preprint arXiv:1307.0589*, 2013.

[40] C. Bergler, H. Schröter, R. X. Cheng, V. Barth, M. Weber, E. Nöth, H. Hofer, and A. Maier, "ORCA-SPOT: An automatic killer whale sound detection toolkit using deep learning," *Scientific Reports*, vol. 9, no. 1, pp. 1–17, 2019.

[41] C. Hauer, E. Nöth, A. Barnhill, A. Maier, J. Guthunz, H. Hofer, R. X. Cheng, V. Barth, and C. Bergler, "Orca-spy enables killer whale sound source simulation, detection, classification and localization using an integrated deep learning-based segmentation," *Scientific Reports*, vol. 13, no. 1, p. 11106, 2023.

[42] National Research Council (U.S.), *Animals as sentinels of environmental health hazards*. National Academy Press Washington, D.C, 1991.

[43] C. Whelan, D. Wenny, and R. Marquis, "Ecosystem services provided by birds," *Annals of the New York Academy of Sciences*, vol. 1134, pp. 25–60, 07 2008.

[44] D. Plazas-Jiménez and M. V. Cianciaruso, "Valuing ecosystem services can help to save seabirds," *Trends in Ecology & Evolution*, vol. 35, no. 9, pp. 757–762, 2020.

[45] R. Pillay, R. J. Fletcher Jr, K. E. Sieving, B. J. Udell, and H. Bernard, "Bioacoustic monitoring reveals shifts in breeding songbird populations and singing behaviour with selective logging in tropical forests," *Journal of Applied Ecology*, vol. 56, no. 11, pp. 2482–2492, 2019.

[46] Ç. H. S¸ ekerciog˘lu, G. C. Daily, and P. R. Ehrlich, "Ecosystem consequences of bird declines," *Proceedings of the National Academy of Sciences*, vol. 101, no. 52, pp. 18042–18047, 2004.

[47] N. A. B. C. Initiative, *The State of the Birds United States of America 2009*. U.S. Department of Interior, 2009.

[48] R. Bardeli, D. Wolff, F. Kurth, M. Koch, K. H. Tauchert, and K.-H. Frommolt, "Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring," *Pattern Recognition Letters*, vol. 31, no. 12, pp. 1524–1534, 2010.

[49] A. Farina, M. Ceraulo, C. Bobryk, N. Pieretti, E. Quinci, and E. Lattanzi, "Spatial and temporal variation of bird dawn chorus and successive acoustic morning activity in a Mediterranean landscape," *Bioacoustics*, vol. 24, 09 2015.

[50] N. Hasan and M. Badri, "Effect of ambient temperature on dawn chorus of house sparrows," *Environment & Ecology Research*, vol. 4, pp. 161–168, 05 2016.

[51] Ç. H. S¸ekercioğlu, R. B. Primack, and J. Wormworth, "The effects of Climate Change on tropical birds," *Biological Conservation*, vol. 148, no. 1, pp. 1–18, 2012.

[52] V. Patel, N. Pauli, E. Biggs, L. Barbour, and B. Boruff, "Why bees are critical for achieving sustainable development," *Ambio*, vol. 50, 04 2020.

[53] A. König, "Extending bee hive health state monitoring by integrated acoustical sensing and machine learning," in *Proceedings of the International Conference on Sensors and Electronic Instrumentation Advances (SEIA)*, (Corfu, Greece), pp. 168–173, 2019.

[54] R. W. Mankin, D. W. Hagstrum, M. T. Smith, A. Roda, and M. T. K. Kairo, "Perspective and promise: a century of insect acoustic detection and monitoring," *American Entomologist*, vol. 57, pp. 30–44, 2011.

[55] C.-M. Görres and D. Chesmore, "Active sound production of scarab beetle larvae opens up new possibilities for species-specific pest monitoring in soils," *Scientific Reports*, vol. 9, p. 10115, 07 2019.

[56] R. W. Mankin, D. Stanaland, M. Haseeb, B. Rohde, O. Menocal, and D. Carrillo, "Assessment of plant structural characteristics, health, and ecology using bioacoustic tools," *Proceedings of Meetings on Acoustics*, vol. 33, no. 1, p. 010003, 2018.

[57] I. Khait, R. Sharon, R. Perelman, A. Boonman, Y. Yovel, and L. Hadany, "Plants emit remotely detectable ultrasounds that can reveal plant stress," *bioRxiv*, 2019.

[58] I. Khait, O. Lewin-Epstein, R. Sharon, K. Saban, R. Goldstein, Y. Anikster, Y. Zeron, C. Agassy, S. Nizan, G. Sharabi, *et al.*, "Sounds emitted by plants under stress are airborne and informative," *Cell*, vol. 186, no. 7, pp. 1328–1336, 2023.

[59] R. B. Cocroft and R. L. Rodríguez, "The behavioral ecology of insect vibrational communication," *BioScience*, vol. 55, pp. 323–334, 04 2005.

[60] R. Birdsey and Y. Pan, "Drought and dead trees," *Nature Climate Change*, vol. 1, no. 9, pp. 444–445, 2011.

[61] M. Gagliano, S. Mancuso, and D. Robert, "Towards understanding plant bioacoustics," *Trends in Plant Science*, vol. 17, pp. 323–325, 03 2012.

[62] B. Krause and A. Farina, "Using ecoacoustic methods to survey the impacts of climate change on biodiversity," *Biological Conservation*, vol. 195, pp. 245–254, 2016.

[63] D. S. Pureswaran, A. Roques, and A. Battisti, "Forest insects and climate change," *Current Forestry Reports*, vol. 4, no. 2, pp. 35–50, 2018.

[64] S. Hallegatte, *Natural Disasters and Climate Change.* Springer, 2016.

[65] J. E. Hansen and M. Sato, "Paleoclimate implications for human-made Climate Change," in *Climate Change*, pp. 21–47, Springer, 2012.

[66] R. S. DeFries, T. Rudel, M. Uriarte, and M. Hansen, "Deforestation driven by urban population growth and agricultural trade in the twenty-first century," *Nature Geoscience*, vol. 3, no. 3, pp. 178–181, 2010.

[67] T. M. Müller, E. Caspari, Q. Qi, J. G. Rubino, D. Velis, S. Lopes, M. Lebedev, and B. Gurevich, "Acoustics of partially saturated rocks: Theory and experiments," in *Seismic Exploration of Hydrocarbons in Heterogeneous Reservoirs: New Theories, Methods and Applications*, pp. 45–75, Elsevier, 2015.

[68] B. M. Popkin, K. E. D'Anci, and I. H. Rosenberg, "Water, hydration, and health," *Nutrition Reviews*, vol. 68, no. 8, pp. 439–458, 2010.

[69] J. J. Alava, P. S. Ross, C. Lachmuth, J. K. Ford, B. E. Hickie, and F. A. Gobas, "Habitat-based PCB environmental quality criteria for the protection of endangered killer whales (orcinus orca)," *Environmental Science & Technology*, vol. 46, no. 22, pp. 12655–12663, 2012.

[70] Y. Samuel, S. Morreale, C. Clark, C. Greene, and M. Richmond, "Underwater, low-frequency noise in a coastal sea turtle habitat," *The Journal of the Acoustical Society of America*, vol. 117, no. 3, pp. 1465–1472, 2005.

[71] D. D. Mara, "Water, sanitation and hygiene for the health of developing nations," *Public Health*, vol. 117, no. 6, pp. 452–456, 2003.

[72] N. J. Ashbolt, "Microbial contamination of drinking water and disease outcomes in developing regions," *Toxicology*, vol. 198, no. 1-3, pp. 229–238, 2004.

[73] F. Paul, "Melting glaciers and ice caps," *Nature Geoscience*, vol. 4, no. 2, pp. 71–72, 2011.

[74] R. J. Urick, "The noise of melting icebergs," *The Journal of the Acoustical Society of America*, vol. 50, no. 1B, pp. 337–341, 1971.

[75] M. Ashokan, G. Latha, A. Thirunavukkarasu, G. Raguraman, and R. Venkatesan, "Iceberg cracking events as identified from underwater ambient noise measurements in the shallow waters of Ny-Alesund, Arctic," *Polar Science*, vol. 10, no. 2, pp. 140–146, 2016.

[76] K. M. Lee, P. S. Wilson, and E. C. Pettit, "Underwater sound radiated by bubbles released by melting glacier ice," *Proceedings of Meetings on Acoustics*, vol. 20, no. 1, p. 070004, 2013.

[77] G. B. Deane, O. Glowacki, M. D. Stokes, and E. C. Pettit, "The underwater sounds of glaciers," *Acoustics Today*, vol. 15, no. 4, pp. 12–19, 2019.

[78] A. V. Marchenko, E. G. Morozov, and S. V. Muzylev, "A tsunami wave recorded near a glacier front," *Natural Hazards and Earth System Sciences*, vol. 12, no. 2, pp. 415–419, 2012.

[79] B. Higman, D. H. Shugar, C. P. Stark, G. Ekström, M. N. Koppes, P. Lynett, A. Dufresne, P. J. Haeussler, M. Geertsema, S. Gulick, *et al.*, "The 2015 landslide and tsunami in Taan Fiord, Alaska," *Scientific Reports*, vol. 8, no. 1, pp. 1–12, 2018.

[80] C. Dai, B. Higman, P. J. Lynett, M. Jacquemart, I. M. Howat, A. K. Liljedahl, A. Dufresne, J. T. Freymueller, M. Geertsema, M. Ward Jones, *et al.*, "Detection and assessment of a large and potentially tsunamigenic periglacial landslide in barry arm, alaska," *Geophysical Research Letters*, vol. 47, no. 22, 2020.

[81] D. Tonolla, M. S. Lorang, K. Heutschi, and K. Tockner, "A flume experiment to examine underwater sound generation by flowing water," *Aquatic Sciences*, vol. 71, no. 4, pp. 449–462, 2009.

[82] A. E. Lumsdon, I. Artamonov, M. C. Bruno, M. Righetti, K. Tockner, D. Tonolla, and C. Zarfl, "Soundpeaking–hydropeaking induced changes in river soundscapes," *River Research and Applications*, vol. 34, no. 1, pp. 3–12, 2018.

[83] A. Mosavi, P. Ozturk, and K.-w. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, p. 1536, 2018.

[84] B. Basnyat, N. Roy, A. Gangopadhyay, and A. Raglin, "Environmental sound classification for flood event detection," in *2022 18th International Conference on Intelligent Environments (IE)*, pp. 1–8, IEEE, 2022.

[85] B. Chide, N. Murdoch, Y. Bury, S. Maurice, X. Jacob, J. P. Merrison, J. J. Iversen, P.-Y. Meslin, M. Bassas-Portús, A. Cadu, A. Sournac, B. Dubois, R. D. Lorenz, D. Mimoun, and R. C. Wiens, "Experimental wind characterization with the supercam microphone under a simulated martian atmosphere," *Icarus*, vol. 354, pp. 114060–114072, 2021.

[86] J. T. Abell, G. Winckler, R. F. Anderson, and T. D. Herbert, "Poleward and weakened westerlies during Pliocene warmth," *Nature*, vol. 589, pp. 70–75, 2021.

[87] C. Lepore, R. Abernathey, N. Henderson, J. T. Allen, and M. K. Tippett, "Future global convective environments in CMIP6 models," *Earth's Future*, vol. 9, no. 12, pp. 1–21, 2021.

[88] B. R. Elbing and C. Petrin, "Monitoring infrasound from a Tornado in Oklahoma," *The Journal of the Acoustical Society of America*, vol. 143, no. 3, pp. 1808–1808, 2018.

[89] M. Moreno, C. Bertolín, D. Arlanzón, P. Ortiz, and R. Ortiz, "Climate change, large fires, and cultural landscapes in the Mediterranean basin: An analysis in southern Spain," *Heliyon*, 2023.

[90] A. A. Khamukhin and S. Bertoldo, "Spectral analysis of forest fire noise for early detection using wireless sensor networks," in *Proceedings of the international Siberian conference on control and communications (SIBCON)*, (Kazan, Russia), pp. 1–4, IEEE, 2016.

[91] A. A. Khamukhin, A. Y. Demin, D. M. Sonkin, S. Bertoldo, G. Perona, and V. Kretova, "An algorithm of the wildfire classification by its acoustic emission spectrum using wireless sensor networks," *Journal Of Physics: Conference Series : Information Technologies in Business and Industry (ITBI2016)*, vol. 803, no. 1, p. 012067, 2017.

[92] K. Schabowicz, T. Gorzelan´czyk, and M. Szymków, "Identification of the degree of degradation of fibre-cement boards exposed to fire by means of the acoustic emission method and artificial neural networks," *Materials*, vol. 12, no. 4, pp. 656–674, 2019.

[93] S. Pryor, P. Crippa, and R. Sullivan, "Atmospheric chemistry," in *Reference Module in Earth Systems and Environmental Sciences*, Elsevier, 2015.

[94] F. Mireshghallah, M. Taram, P. Vepakomma, A. Singh, R. Raskar, and H. Esmaeilzadeh, "Privacy in deep learning: A survey," *arXiv preprint arXiv:2004.12254*, pp. 1–24, 2020.

[95] J. L. Kröger, O. H.-M. Lutz, and P. Raschke, "Privacy implications of voice and speech analysis–information disclosure by inference," in *IFIP International Summer School on Privacy and Identity Management*, (Windisch, Switzerland), pp. 242–258, Springer, 2019.

[96] E. Jose, M. Greeshma, M. T. Haridas, and M. Supriya, "Face recognition based surveillance system using FaceNet and MTCNN on Jetson TX2," in *Proceedings of the International Conference on Advanced Computing & Communication Systems (ICACCS)*, (Coimbatore, India), pp. 608–613, IEEE, 2019.

[97] M. S. Hanoon, A. N. Ahmed, N. Zaini, A. Razzaq, P. Kumar, M. Sherif, A. Sefelnasr, and A. El-Shafie, "Developing machine learning algorithms for meteorological temperature and humidity forecasting at terengganu state in malaysia," *Scientific Reports*, vol. 11, no. 1, pp. 1–19, 2021.

[98] R. Wang, K. Kashinath, M. Mustafa, A. Albert, and R. Yu, "Towards physics-informed deep learning for turbulent flow prediction," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, (online), pp. 1457–1466, ACM, 2020.

[99] A. Vamseekrishna, R. Nishitha, T. A. Kumar, K. Hanuman, and C. G. Supriya, "Prediction of temperature and humidity using IoT and machine learning algorithm," in *Proceedings of the International Conference on Intelligent and Smart Computing in Data Analytics (ISCDA)*, (Andhra Pradesh, India), pp. 271–279, Springer, 2020.

[100] M. Ross and J. A. Vedda, "The policy and science of rocket emissions," *Center for Space Policy and Strategy, The Aerospace Corporation*, pp. 2–10, 2018.

[101] N. Adilov, P. J. Alexander, and B. M. Cunningham, "An economic analysis of earth orbit pollution," *Environmental and Resource Economics*, vol. 60, no. 1, pp. 81–98, 2015.

[102] S. Ben-David, J. Blitzer, K. Crammer, F. Pereira, *et al.*, "Analysis of representations for domain adaptation," *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*, vol. 19, pp. 137–144, 2006.

[103] D. Wang and J. Chen, "Supervised speech separation based on deep learning: An overview," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 10, pp. 1702–1726, 2018.

[104] S. Liu, A. Triantafyllopoulos, Z. Ren, and B. W. Schuller, "Towards speech robustness for acoustic scene classification," in *Proceedings of the Annual Conference of the International Speech Communication Association (INTERSPEECH)*, (Shanghai, China), ISCA, 2020.

[105] R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green AI," *Communications of the ACM*, vol. 63, no. 12, pp. 54–63, 2020.

[106] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[107] W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A survey of zero-shot learning: Settings, methods, and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–37, 2019.

[108] H. Xie and T. Virtanen, "Zero-shot audio classification via semantic embeddings," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1233–1242, 2021.

[109] A. Wynsberghe, "Sustainable AI: AI for sustainability and the sustainability of AI," *AI and Ethics*, vol. 1, pp. 1–6, 02 2021.

[110] H. Coppock, L. Jones, I. Kiskin, and B. W. Schuller, "Covid-19 detection from audio: Seven grains of salt," *The Lancet Digital Health*, vol. 3, no. 9, pp. 537–538, 2021.

[111] S. Yu, "Big privacy: Challenges and opportunities of privacy study in the age of big data," *IEEE Access*, vol. 4, pp. 2751–2763, 2016.

[112] A. Triantafyllopoulos, M. Milling, K. Drossos, and B. W. Schuller, "Fairness and underspecification in acoustic scene classification: The case for disaggregated evaluations," in *Proceedings of the Detection and Classification of Acoustic Scenes and Events Workshop (DCASE)*, (online), 2021.

[113] A. Ignatiev, "Towards Trustable Explainable AI.," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, (Yokohama, Japan), pp. 5154–5158, 2020.

[114] B. W. Schuller, A. Akman, Y. Chang, H. Coppock, A. Gebhard, A. Kathan, E. Rituerto-González, A. Triantafyllopoulos, and F. B. Pokorny, "Climate change & computer audition: A call to action and overview on audio intelligence to help save the planet," *arXiv preprint arXiv:2203.06064*, 2022.

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Bjoern Schuller reports financial support was provided by German Research Foundation. Bjoern Schuller reports a relationship with audEERING GmbH that includes: employment.