A New Lens on Life: Cross-Contextual Sensing Technologies from Human Insights to Wildlife Conservation

by

Patrick C. Chwalek

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MEDIA ARTS AND SCIENCES

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2025

Authored by: Patrick C. Chwalek
Responsive Environments, MIT Media Lab
August 22, 2025

Certified by: Joseph A. Paradiso
Alexander W Dreyfoos Professor of Media Arts and Sciences
Thesis Supervisor

Accepted by: Joseph A. Paradiso
Academic Head
Program in Media Arts and Sciences

A New Lens on Life: Cross-Contextual Sensing Technologies from Human Insights to Wildlife Conservation

by

Patrick C. Chwalek

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on August 22, 2025 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MEDIA ARTS AND SCIENCES

ABSTRACT

The growing complexity of human-centric and ecological systems demands new sensing technologies capable of capturing holistic, contextual insights in real-world environments. However, a critical gap exists in the availability of integrated, intelligent platforms that can be adapted across these diverse domains. This dissertation addresses this challenge by designing, engineering, and validating a series of novel, multi-modal sensing platforms to provide a new and more insightful lens on a broad range of life.

The research spans two primary contexts. First, to explore the nuances of human well-being, the AirSpecs smart-eyeglass platform was developed and deployed. This system holistically measures an individual's proximate environment and physiological parameters, and was validated in a multi-site international study to investigate the dynamics of human comfort "in-the-wild". Second, to advance ecological monitoring, a progression of acoustic platforms was engineered. The SoundSHROOM system was created as a robust, multi-channel recorder for harsh environments and successfully deployed in the Arctic. Building on this, the BuzzCam system was developed for targeted pollinator monitoring, culminating in an end-to-end pipeline for on-device AI classification of endangered and invasive bee species in Patagonia. Finally, the CollarID platform was engineered and characterized as a versatile, low-power, multi-modal animal-borne sensor for wildlife tracking, integrating inertial, bioacoustic, and comprehensive environmental sensing to move beyond the limitations of location-only devices.

Key contributions of this work include the validated hardware platforms themselves; several unique, publicly available datasets from urban, Arctic, and Patagonian deployments; and a demonstrated methodology for implementing on-device AI to address the data-to-insight bottleneck in ecological monitoring. Collectively, this research provides the scientific community with a new suite of powerful research tools and demonstrates a cross-contextual design philosophy, leveraging engineering principles across disparate fields to enable a deeper understanding of organisms and their complex interactions with their environments.

Thesis Supervisor: Joseph A. Paradiso
Title: Alexander W Dreyfoos Professor of Media Arts and Sciences

THESIS COMMITTEE

Thesis Supervisor

Joseph A. Paradiso
Alexander W. Dreyfoos Professor of Media Arts and Sciences
Director, Responsive Environments Group, MIT Media Lab

Thesis Readers

Thad Starner
Professor, School of Interactive Computing
Georgia Institute of Technology

Rébecca Kleinberger
Assistant Professor, Khoury College of Computer Sciences
and College of Arts, Media and Design
Northeastern University

Acknowledgments

This dissertation is the culmination of years of work, and it would not have been possible without the guidance, support, and encouragement of a vast community of mentors, colleagues, family, and friends. I am deeply grateful to everyone who has been a part of this journey.

First and foremost, I extend my sincerest gratitude to my advisor, Professor Joseph A. Paradiso. Thank you, Joe, for assembling the incredible Responsive Environments family and for granting me the intellectual freedom to pursue the diverse projects that form the heart of this thesis. Your trust and guidance created an environment where curiosity and innovation could flourish.

I am also profoundly thankful to the members of my PhD committee, Professor Thad Starner and Professor Rébecca Kleinberger. Your insightful critiques and critical questions were instrumental in shaping the direction and depth of this work. Thank you for pushing me to think more deeply and for helping me refine my research into a coherent whole.

I have been fortunate to learn from several incredible mentors. I owe a special thanks to Dan, whose expertise and connections in the world of audio were invaluable. Thank you for teaching me everything I needed to complete my projects and for being so open about your decades of experience in tackling problems both inside and out of the lab. I am also grateful to Professor Christoph Reinhardt and Professor Holly Samuelson for their unwavering support during the general exam process. The advice you gave not only allowed me to complete the work on AirSpecs but also taught me invaluable lessons about the design process.

To my Responsive Environments family, both past, present, and adjacent, thank you. You made coming into the lab each day feel like coming home. Your constant support, camaraderie, and intellectual sparring were essential every step of the way. Specifically, I want to thank Caroline, Ariel, Ishwarya, Artem, Juliana, Spencer, Fangzheng, Perry, Nathan, Sam, Cedric, Elena, Lancelot, Vald, Samantha, Cathy, Rob, and Abhi. Within this family, I must give special thanks to two incredible friends and engineers, Brian and Mark. Brian, your advice over the years was foundational to my growth as an engineer, and I cannot thank you enough. Mark, your perfect blend of engineering know-how and healthy cynicism kept me grounded and my work focused; I wouldn't have been able to finish my projects to this caliber without you both.

This journey also brought one of the most important people into my life, David. I came into the program looking up to you, and as I leave, you have become one of the best friends I could have ever asked for. We started the glasses project together, and though our paths may diverge, I know we will always be there to support one another. I am also grateful to have started my time at MIT with Erik, Alex, and Joanne; we truly supported each other through all the ups and downs of this program.

This work was enriched by collaborations that spanned the globe. Thank you to all my international collaborators for being such amazing people to work with. To Sailin, thank you for all your work on AirSpecs and for helping to bridge the Swiss and MIT communities. To Marie and Marco, thank you for your technical work on BuzzCam, and to Marie especially for your incredible support during the challenging fieldwork. To my team in South America—Marina, Cristian, Vicky, and Jose—thank you for being so welcoming and making my visits so pleasant. And to Maggie, Zivvy, Ganit, and Eve, thank you for not only being great colleagues but for making the trip to the Arctic an adventure I will never forget.

I am also immensely grateful to the National Geographic Society. The Exploration Technology Lab, in particular, provided foundational support for all of my animal-related projects. They connected me to their incredible explorer network and gave me the encouragement and resources needed to pursue this work. A special thank you to Kasie, Alan, Kyler, Cody, and Brad for being so supportive. None of this research would have been possible without the support of the ExTech team.

I owe everything to the three most impactful women in my life: my grandmother, Helena; my mom, Elizabeth; and my sister, Angelika. As a first-generation American and a first-generation college student, this path was filled with challenges and mistakes. It was only because of your unconditional love and support that I was always able to fail up and keep moving forward.

I am also deeply grateful to Beata for a decade of unwavering support. I simply would not have made it this far without you, and this achievement is as much yours as it is mine.

Finally, I want to thank the friends outside of academia who kept me sane throughout this process. To my hometown friends who I grew up with—Greg, Leighanne, Meleta, Eva, Michael, and Marcelo—thank you for your lifelong support. To Andriana, thank you for our time together, which impressed upon me the importance of work-life balance. To Geri, Pepi, George, Stella, Ari, Alex, Kat, Shawn, Ted, and Korinna—many of whom I met in the last two years of my PhD—thank you. Your friendship was foundational in helping me find balance during the most intense times. In the final, stressful months of writing, I hosted over 25 dinners—a quirky tradition I called Buzz & Banter—where friends would join me to cook and share animal facts. This gave me the break I needed and the foundation to keep going. To everyone who participated, thank you for the food, the facts, and the friendship.

Biographical Sketch

Patrick Chwalek was born in Chicago, Illinois on July 16, 1993. He received his PhD in Media Arts and Sciences from the MIT Media Lab's Responsive Environments Group, where he developed and deployed novel sensor systems for real-world applications. He broadly identifies as a system developer and integrator with expertise spanning embedded hardware, firmware, and mechanical design. His academic journey is interdisciplinary, with a Bachelor of Science in Mechanical Engineering from the University of Illinois at Urbana-Champaign, a Master of Science in Computer Science from the Georgia Institute of Technology, and a Master of Science from the MIT Media Lab. The son of parents who immigrated from two small villages in Poland, he is the first in his family to attend graduate school and the first in his close and distant family to receive a PhD. The core of his doctoral research is dedicated to creating new technological lenses to better understand life, both animal and human. A significant portion of this work, often in collaboration with National Geographic Explorers, focuses on developing technologies for wildlife conservation. This has led to the design of systems to monitor endangered native and invasive bees in Patagonia (BuzzCam), track the environmental context of larger animals such as lions, hyenas, and dogs (CollarID), and listen to avian life in the Arctic Circle. The other major thrust of his research explores the intersection of technology and human well-being. To this end, he developed AirSpecs, a low-profile smart eyewear platform designed to collect rich data on a user's physiology and their immediate environmental conditions. This system was successfully deployed across three continents to create a diverse dataset for understanding how comfort perception and environmental factors vary in daily life.

Before and during his PhD, Patrick has applied his skills across a wide range of academic, government, and industry settings. He worked for over two years at MIT Lincoln Laboratory, where he designed wireless electroglottography systems for vocal monitoring and led the development of multi-spectrum imaging systems for off-grid surveillance. He has also completed impactful internships, including at the National Geographic Society's Exploration Technology Lab enhancing deep ocean camera systems, at the startup Gridware, Inc. architecting a low-cost wildfire detection system, and an early co-op at UTC Aerospace Systems investigating aircraft component failures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation: The Confluence of Sensing for Humans and Wildlife

The last few decades have witnessed a fundamental paradigm shift in computing. We have moved from an era of static, room-sized mainframes to one of dynamic, personal, and increasingly invisible technology, a progression often termed ubiquitous computing. Central to this evolution is the proliferation of sensing technologies that can perceive, interpret, and react to their context. This enhanced awareness has unlocked unprecedented opportunities to understand and augment our interaction with the world around us. This dissertation is situated at the confluence of two seemingly disparate, yet technologically related frontiers in this sensing revolution: capturing the nuanced experience of human life within built environments, and gaining a deeper, less intrusive understanding of the lives of wild animals in their natural habitats.

The first motivation for this work stems from the challenge of understanding the human experience. In developed nations, individuals spend the vast majority of their time—often over 90%—indoors. The quality of these built environments profoundly impacts our comfort, health, well-being, and productivity. Historically, building design has relied on standardized, static metrics for environmental quality, often measured by a single thermostat or sensor serving a large area. However, a growing body of research recognizes that human comfort is an inherently complex and subjective state, influenced not only by ambient conditions but by an individual's physiology, psychology, and personal preferences. This highlights a critical limitation of traditional sensing: it fails to capture the dynamic, personalized micro-environments that people actually experience as they move through their daily lives. To truly design human-centric spaces, we need to move beyond room-level measurements and develop wearable tools that can holistically capture an individual's proximate environmental conditions and relevant physiological responses simultaneously, "in-the-wild". This need for personalized, contextual sensing in the human domain is a primary driver of this thesis.

The second, parallel motivation is the pursuit of a new lens on the natural world. Ecosystems globally face unprecedented pressures from climate change, habitat loss, and invasive species, making effective biodiversity monitoring more critical than ever. Traditional ecological field methods, such as visual surveys or physical trapping, while foundational, are often labor-intensive, geographically limited, and can be invasive, potentially altering the behavior of the animals being studied. Technology offers a powerful alternative. The use of animal-attached electronic devices, known as biologging, has already revolutionized the study of animal movement. However, while modern collars increasingly include accelerometry, their focus on location and basic activity often leaves more nuanced questions about behavior, physiology, and environmental context unanswered. Similarly, Passive Acoustic Monitoring (PAM) has emerged as a potent, non-invasive tool for assessing biodiversity by "listening" to ecosystems. Yet, the utility of these advanced methods is frequently constrained by hardware not robust enough for harsh environments, by the immense volume of data they generate, and by the lack of on-device intelligence to process that data efficiently. This creates an urgent need for more robust, scalable, and intelligent ecological sensors that can provide a richer, more holistic picture of wildlife.

Central to this work is its interdisciplinary nature, which bridges Human-Computer Interaction (HCI), ecology, and engineering. HCI principles ensure that technologies like AirSpecs are user-friendly and contextually relevant. Ecological insights inform the design of non-invasive sensors like BuzzCam and CollarID. Engineering expertise underpins the robust, field-ready hardware across all projects. This cross-disciplinary approach not only enhances each project's effectiveness but also establishes a versatile methodology for future sensing challenges.

While sensing human comfort in an office and monitoring endangered species in the wild may appear to be unrelated challenges, they are connected by a shared set of underlying engineering and methodological hurdles. Both domains require the development of low-power, miniaturized, and robust hardware capable of surviving real-world deployment. Both demand the integration of multiple sensor modalities—from environmental to acoustic to inertial—to create a complete contextual picture. And both are increasingly reliant on advanced machine learning to transform vast streams of raw sensor data into actionable insights. This dissertation explores this cross-contextual opportunity, asserting that the principles and technologies developed to solve challenges in one domain can inform and accelerate progress in the other. By designing, engineering, and validating a series of novel sensing platforms across these contexts, this work aims to provide a new and more insightful lens on life.

1.2 Problem Statement

While sensing technologies have become widespread, there remains a critical gap in the availability of integrated, context-aware, and intelligent platforms that can be adapted to capture holistic insights across different domains. In the human dimension, we lack wearable tools to holistically measure the confluence of environmental, physiological, and subjective factors that define well-being. In the ecological domain, we lack robust, scalable, and intelligent sensors to non-invasively monitor the behavior, health, and environmental interactions of vulnerable species. This dissertation addresses the challenge of designing, engineering, and validating a series of novel sensing platforms to bridge these technological gaps.

1.3 Research Objectives and Questions

To address the technological gaps identified in the problem statement, this dissertation pursues a series of targeted research objectives. These objectives guided the design, development, and validation of the novel sensing platforms that form the core of this work. The overarching goal is to demonstrate that by engineering highly integrated and context-aware sensors, we can gain new insights into both human and animal life. The specific objectives are as follows:

1. **To design, deploy, and evaluate a novel, head-worn wearable platform (AirSpecs) capable of holistically sensing an individual's proximate environment and physiological state.** This objective seeks to move beyond static, room-level measurements of comfort by creating a tool for collecting rich, "in-the-wild" data. The central research question is: How can a multi-modal wearable sensor, combined with subjective feedback, deepen our understanding of human comfort and environmental awareness in real-world settings?

2. **To engineer and field-validate robust acoustic monitoring hardware for challenging ecological applications.** This objective is twofold. First, it involves developing a resilient, multi-channel acoustic recorder (SoundSHROOM) for harsh environments like the Arctic. Second, it aims to tackle the "data-to-insight" bottleneck by creating a specialized platform (BuzzCam) that uses on-device Artificial Intelligence to perform real-time classification of endangered pollinators, thereby enabling scalable, non-invasive monitoring.

3. **To engineer and characterize a versatile, low-power, multi-modal animal-borne platform (CollarID) for diverse wildlife monitoring.** This objective addresses the limitations of traditional location-only tracking devices. The goal is to develop and rigorously validate a single, integrated platform that combines inertial, bioacoustic, and comprehensive environmental sensing to enable a more holistic understanding of animal behavior, health, and their interactions with the immediate environment.

1.4 Scope and Limitations

To provide a clear framework for the contributions of this dissertation, this section defines the scope of the research undertaken and acknowledges its inherent limitations.

1.4.1 Scope of Research

The primary focus of this work is on the design, engineering, and validation of novel sensing platforms as research tools to enable new forms of data collection in both human-centric and ecological contexts. The scope includes:

- The complete design, prototyping, and characterization of three distinct sensing platforms: the AirSpecs smart eyeglasses, the SoundSHROOM and BuzzCam acoustic recorders, and the CollarID multi-modal wildlife collar.

- The empirical validation of these platforms through a series of targeted laboratory and field-based tests. This includes an international, multi-site "in-the-wild" user study with AirSpecs; a deployment of SoundSHROOM in the Arctic to test for robustness; a field study in Patagonia with BuzzCam to collect foundational data; and the comprehensive engineering characterization of the CollarID prototype, which involved mechanical stress simulations, environmental sealing tests, sensor performance co-location experiments, communication range trials, and an initial integrated system deployment on a farm animal, with upcoming deployments on lions, hyenas, and dogs in Central Africa.
- The development of an end-to-end pipeline for on-device artificial intelligence, demonstrating the process from data collection and annotation to model optimization and deployment on a resource-constrained microcontroller for real-time ecological classification with the BuzzCam system.
- The creation and public dissemination of multiple, high-value datasets resulting from these deployments to benefit the broader scientific community.

1.4.2 Limitations

It is equally important to define the boundaries of this research. The following limitations are acknowledged. While the field deployments provided crucial data for platform validation, they were not designed as long-term, multi-season ecological or sociological studies; their primary purpose was to demonstrate technological capabilities and collect foundational datasets, rather than to draw definitive scientific conclusions about the populations or environments studied. Furthermore, the machine learning models developed, particularly for the BuzzCam project, were trained on data from a single geographical region and season. Although this demonstrated a powerful proof-of-concept, their generalizability may require further site-specific data collection and fine-tuning. Moreover, the work on the CollarID platform centers on its rigorous engineering development and validation as a field-ready prototype; full-scale deployments on target wildlife species are a critical next step but are considered future work beyond the scope of this thesis.

Additionally, the research encountered specific challenges inherent to the diverse deployment contexts. In the human-centric AirSpecs study, for instance, the international deployment across Boston, Fribourg, and Singapore introduced cultural variability that could influence participants' comfort perceptions and technology interactions, a known complexity in field-based behavioral research [1]. These cultural differences may limit the generalizability of findings across populations. In the ecological projects, deployments in extreme environments posed significant challenges. The SoundSHROOM system, tested in the Arctic, faced high winds and limited accessibility, testing hardware reliability and data collection protocols. Similarly, the BuzzCam system in Patagonia contended with variable weather and the need for unobtrusive placement to avoid disrupting bee behavior. These environmental extremes influenced design choices and highlighted the importance of robust engineering for field deployments.

1.5 Thesis Contributions

This dissertation makes several key contributions to the fields of environmental sensing, human-computer interaction, and ecological technology. The contributions include novel hardware platforms, unique public datasets that enable new research, and demonstrated methodologies for applying on-device artificial intelligence to real-world challenges. The primary contributions are:

1. **A Novel Wearable Platform and Dataset for Human-Centric Sensing.** This work presents AirSpecs, a smart-eyeglass platform uniquely integrating a comprehensive suite of environmental and physiological sensors for "in-the-wild" human comfort research. The platform was successfully deployed in a multi-site international study, resulting in the creation of a rich, public dataset that captures synchronized environmental, physiological, and subjective comfort data. The value of this contribution has already been demonstrated through its use by independent researchers to develop and validate new, more advanced personalized comfort models.

2. **Advanced Platforms and On-Device AI for Ecological Acoustic Monitoring.** This dissertation contributes two novel acoustic platforms and their associated datasets:

    - The SoundSHROOM system, a robust, field-tested, multi-channel acoustic recorder designed for and validated in the harsh Arctic environment. This work yielded a unique public dataset of Arctic soundscapes suitable for spatial audio analysis .
    - The BuzzCam system, a specialized platform for pollinator monitoring, and an end-to-end methodology for deploying an AI model for real-time, on-device classification of endangered and invasive bee species on a low-power microcontroller. This contribution provides a tangible solution to the data-to-insight bottleneck in passive acoustic monitoring and includes a foundational, annotated dataset of bee bioacoustics from Patagonia.

3. **A Versatile Multi-Modal Platform for Wildlife Biologging.** This thesis details the design, engineering, and rigorous characterization of CollarID, a lightweight, low-power, and robust prototype for a multi-modal animal-borne monitoring platform . By integrating inertial, high-fidelity bioacoustic, and comprehensive environmental sensing capabilities into a single device, CollarID provides a more holistic sensing tool than many existing commercial alternatives, establishing a new hardware standard for contextual wildlife research.

4. **A Cross-Contextual Design Methodology.** Beyond the specific platforms, this dissertation demonstrates a replicable methodology for adapting and evolving sensing technologies across seemingly disparate research domains. It shows how engineering principles and learnings—from low-power design to robust field deployment and data management—can be translated from human-centric sensing to inform and accelerate the development of advanced tools for wildlife conservation.

Chapter 2

Background and Related Work

2.1 Introduction

The research presented in this dissertation is situated at the confluence of several dynamic and evolving fields, including Human-Computer Interaction (HCI), environmental sensing, ecological monitoring, wearable computing, and applied machine learning. This chapter provides a comprehensive review of the relevant background and related research. My aim is to establish the theoretical and technological underpinnings upon which the AirSpecs, SoundSHROOMs, BuzzCam, and CollarID projects are built, critically examine existing approaches to identify their limitations, and thereby highlight the specific opportunities and needs that my research seeks to address.

Given the interdisciplinary nature of my doctoral work—which spans from sensing human comfort in built environments to developing on-device artificial intelligence for monitoring endangered pollinators and engineering robust platforms for diverse wildlife tracking—the review that follows will necessarily traverse literature concerning human-centric environmental quality, wearable sensor technologies, passive acoustic monitoring techniques, advanced biologging for wildlife, and the application of machine learning and edge computing (TinyML) in these contexts. While an exhaustive survey of each of these vast fields is beyond the scope of a single chapter, I have attempted to synthesize the most pertinent advancements and critical discussions that directly inform and motivate the specific research questions and engineering objectives of this dissertation.

To provide a clear and logical progression through these diverse yet interconnected areas, this chapter is organized into several thematic sections. Section 2.2, "Understanding and Sensing Human Comfort in Built Environments," will delve into the complexities of human comfort and review technologies for its assessment, laying the groundwork for the AirSpecs platform. Section 2.3, "Ecological Monitoring: Traditional and Technological Approaches," will shift focus to the challenges and advancements in monitoring non-human animals, providing context for SoundSHROOM, BuzzCam, and CollarID projects. Subsequently, Section 2.4, "Machine Learning in Ecological Monitoring and Sensing," will explore the increasing role of artificial intelligence in extracting insights from complex sensor data, with particular relevance to the BuzzCam project. Finally, Section 2.5, "Synthesis and Identification of Research Gaps," will recapitulate the key limitations identified throughout the review and articulate how the collective body of work presented in this dissertation aims to address these gaps.

By systematically examining these areas, this chapter aims to provide a robust foundation for the detailed descriptions of the methodologies, systems, and findings presented in the subsequent chapters of this dissertation.

2.2 Understanding and Sensing Human Comfort in Built Environments

The quality of the built environments in which individuals spend the vast majority of their time—be it homes, offices, or public spaces—profoundly impacts their comfort, health, well-being, and productivity [2,3]. Historically, the design and operation of buildings have often prioritized energy efficiency and standardized comfort metrics. However, there is a growing recognition within both architectural research and Human-Computer Interaction (HCI) that a more nuanced, personalized, and human-centric approach is necessary to truly foster comfortable and supportive indoor environments [4–6]. This section reviews the literature on human comfort, explores established and emerging methods for sensing relevant environmental and physiological parameters, and discusses the role of technology in mediating occupant interaction and awareness.

2.2.1 Defining Human Comfort: A Multifaceted and Subjective Experience

Human comfort is an inherently complex and subjective state, defined by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) as "that condition of mind which expresses satisfaction with the thermal environment" [7]. While this definition focuses on thermal comfort, the broader concept encompasses satisfaction with multiple facets of the Indoor Environmental Quality (IEQ). Frontczak Wargocki [8] provide a comprehensive review highlighting these key IEQ dimensions:

- Thermal Comfort: Influenced by air temperature, radiant temperature, air velocity, humidity, metabolic rate, and clothing insulation. Seminal models like Fanger's Predicted Mean Vote (PMV) [9] and the Adaptive Comfort Model [10,11] have long guided building design, though they primarily predict average comfort for groups and may not capture individual variations or dynamic conditions effectively.

- Visual Comfort: Determined by factors such as illuminance levels, glare, light distribution, color temperature, and access to views. Inadequate visual conditions can lead to eye strain, fatigue, and reduced performance [12].

- Acoustic Comfort: Relates to the absence of unwanted sound (noise) and the presence of a suitable acoustic environment for the task at hand. Noise from external sources or internal activities can be a significant source of distraction and annoyance [13].

- Indoor Air Quality (IAQ): Influenced by concentrations of pollutants such as carbon dioxide (CO2), Volatile Organic Compounds (VOCs), particulate matter (PM), and biological contaminants. Poor IAQ is linked to various health issues, discomfort, and cognitive impairment [14,15].

Beyond these physical IEQ parameters, our understanding of comfort has evolved to embrace its subjective and contextual nature [5,16]. Individual factors such as personal preferences, physiological state, psychological factors (e.g., stress, mood), expectations, past experiences, cultural background, and perceived control over the environment significantly modulate comfort perception [17,18]. This inherent subjectivity makes it challenging to design "one-size-fits-all" environments and underscores the need for personalized approaches.

2.2.2 Technologies for Sensing Indoor Environmental Quality (IEQ)

A variety of technologies have been employed to measure IEQ parameters. Modern commercial buildings, for instance, are often equipped with extensive Building Management Systems (BMS) that monitor and control HVAC, lighting, and other systems. While BMS can provide rich data at a building or zone level [19], they typically lack the granularity to assess an individual's direct micro-environment or personal exposure, and occupants often have limited access to or understanding of this data [20]. To achieve more localized measurements, researchers and practitioners also utilize a wide array of standalone sensors and sensor networks. The advent of low-cost sensor platforms (e.g., Arduino, Raspberry Pi based) has spurred numerous "Do It Yourself" (DIY) environmental monitoring projects within HCI and ubiquitous computing research [21], though these can require manual setup and may lack calibration or long-term stability. A key challenge that pervades these fixed-sensor approaches is that of sensor placement. For instance, CO2 or VOC levels relevant to inhalation exposure are best measured near the breathing zone, which fixed room sensors cannot achieve [3,22], and personal thermal microclimates can vary significantly even within a single room.

2.2.3 Wearable Sensing for Personalized Comfort, Health, and Environmental Exposure

The proliferation of wearable technology offers a promising avenue for overcoming these limitations and for capturing personalized, continuous data streams relevant to comfort and health. Smartwatches, fitness trackers, and specialized research wearables now commonly measure physiological signals such as heart rate, skin temperature, and electrodermal activity, which can be indicative of stress, cognitive load, or thermal strain [23,24].

While physiological sensing is mature, the integration of comprehensive IEQ sensors into comfortable, everyday wearable form factors is a less developed, emerging area. Some research prototypes have explored thermal comfort wearables [25,26] or air quality monitors [27,28], but as noted previously, existing wearables often cover only one dimension of IEQ or focus solely on physiology [5]. Smart eyeglasses, like the AirSpecs platform, offer a unique advantage due to their proximity to multiple sensory organs and the user's breathing zone, allowing for more direct measurement of inhaled air quality and facial skin temperature changes that can correlate with cognitive or affective states [29,30].

This review thus highlights a gap for a holistic, head-worn wearable system like AirSpecs, capable of simultaneously sensing a broad suite of IEQ parameters and relevant physiological signals for in-the-wild comfort research.

2.2.4 Human-Building Interaction (HBI) and Occupant Engagement with Environmental Data

Beyond merely sensing the environment, a critical aspect of human-centric design involves how occupants interact with and understand their surroundings and the technologies that mediate them.

- Control and Agency: Research consistently shows that providing occupants with a greater degree of personal control over their environmental conditions (e.g., operable windows, personal thermostats) can lead to increased comfort tolerance and satisfaction, even if the physical conditions are not strictly within "standard" optimal ranges [17,31].

- Data Feedback and Awareness: Providing occupants with feedback on their environmental conditions or energy use can raise awareness and potentially influence behavior [32,33] . However, the design of such feedback (what data to show, how to show it, when to show it) is crucial for it to be effective and not overwhelming.

- Ecological Momentary Assessment (EMA): EMA, or experience sampling, is a valuable methodology for capturing subjective states like comfort, mood, and focus in real-time, within the participant's natural environment, thereby minimizing recall bias [34,35]. The Cozie app [36] is an example of leveraging mobile technology for EMA in comfort studies.

- Subtle Notifications and Peripheral Cues: To avoid disrupting occupants' primary tasks, researchers have explored subtle notification mechanisms. Peripheral light cues, for instance, have been investigated as a "naturalistic measure of focus" and a means to gently prompt interaction without demanding immediate attention [37,38].

This review of literature pertaining to human comfort and its sensing underscores the need for novel tools and methodologies to capture the rich, contextual, and personalized nature of human-environment interactions in the built environment. The AirSpecs study detailed in this dissertation directly addresses this gap, exploring how a wearable platform providing real-time data, coupled with EMA prompted by peripheral cues, influences environmental awareness and how users imagine future interactions with smart building systems [5].

2.3 Ecological Monitoring: Traditional and Technological Approaches

Shifting from the human-centric built environment to the broader ecological realm, the imperative for accurate and effective monitoring of biodiversity and ecosystem health is equally, if not more, pressing. Understanding species distributions, population dynamics, behavioral patterns, and responses to environmental change is fundamental to ecological science and essential for informing effective conservation strategies [39,40]. This section reviews traditional methods employed in ecological field research, then delves into the advancements and ongoing challenges associated with technological approaches, particularly Passive Acoustic Monitoring (PAM) and animal-borne biologging devices. This review will establish the context and highlight the technological gaps that my work on SoundSHROOM, BuzzCam, and CollarID aims to address.

2.3.1 Traditional Methods in Ecological Field Research: Foundations and Limitations

For centuries, ecological knowledge has been built upon a foundation of direct observation and manual data collection techniques. These traditional methods remain valuable but often come with inherent limitations, especially when aiming for broad spatial or temporal coverage. For example, direct visual observation by trained ecologists along predefined transects is a cornerstone for assessing diurnal species but is highly labor-intensive, costly, and can be prone to observer bias that underrepresents cryptic or nocturnal species [41–43]. For many other animal groups, trapping techniques and mark-recapture methods are employed to estimate population size and density [44,45]. While providing valuable demographic data, these methods can be invasive, potentially stressful to the animals, and ethically problematic for endangered species, as is the case with pan trapping for threatened bee species [46,47]. Additionally, the collection of specimens or surveys of indirect signs like tracks and scat can provide evidence of species presence but often offer limited insight into population size, behavior, or fine-scale habitat use [48].

While these traditional methods have been instrumental in advancing ecological understanding, their limitations in terms of labor, cost, invasiveness, and spatio-temporal resolution have spurred the development and adoption of technology-driven approaches.

2.3.2 Passive Acoustic Monitoring (PAM) in Ecology: Listening to Biodiversity

Passive Acoustic Monitoring (PAM)—the use of autonomous recorders to capture soundscapes—has emerged as a powerful, non-invasive tool for ecological research and biodiversity assessment [49,50]. PAM leverages the fact that many animal species produce characteristic sounds for communication, navigation, or movement. By deploying acoustic sensors, researchers can collect continuous data over long periods and across broad spatial scales, often in remote locations, with key advantages including the ability to detect nocturnal and cryptic species and create permanent acoustic records for re-analysis [51]. This technique has been widely applied to study a diverse range of taxa, from birds and bats to marine mammals and insects, for applications such as biodiversity inventories, behavioral studies, and soundscape health assessments [52–57].

Despite its advantages, PAM presents significant hardware challenges that often involve trade-offs between cost, performance, and robustness. At one end of the spectrum is the open-source AudioMoth, a low-cost mono recorder that, with a waterproof housing, can be deployed for under \$150. Its affordability has made it a widely used tool in research, greatly increasing the accessibility of PAM for large-scale studies [58]. However, the AudioMoth is not designed for high-end audio fidelity or with significant wind noise mitigation in mind. At the other end are commercial solutions from companies like Wildlife Acoustics. Their product line ranges from the Song Meter Micro, which is similar in size and cost to the AudioMoth, to the more advanced Song Meter SM4. The SM4, costing around \$700, offers higher-quality stereo recording and includes foam covers over the microphones to reduce wind noise. Yet, it is limited to approximately 440 hours of recording time on four large D-cell batteries, though it can be powered externally.

Crucially, neither of these commonly used recorders offers users the ability to run machine learning models on board. This lack of on-device intelligence is a significant limitation, as it prevents researchers from experimenting with edge AI models and creating autonomous detection systems. Such systems are vital for enabling large-scale, time-sensitive field experiments in remote areas where real-time data processing can provide immediate ecological insights and trigger timely conservation actions.

These existing systems highlight persistent challenges in the field. Deploying any acoustic recorder in harsh environments requires robust, weatherproof enclosures and efficient power management for extended unattended operation, a key motivation for the development of the SoundSHROOM system [59]. Wind is another pervasive issue, as it can mask biological sounds and degrade recording quality, making effective microphone windshielding a critical design consideration [60]. Furthermore, while most PAM studies use single or stereo microphones, multi-channel arrays that offer the potential for sound source localization and beamforming are less common due to engineering challenges in synchronization, data volume, and power—areas the SoundSHROOM project was designed to explore [61,62]. Beyond hardware, the primary challenge in PAM remains the analysis of the vast quantities of acoustic data generated, which necessitates automated approaches and leads to the discussion of machine learning in Section 2.4. The SoundSHROOM project directly addressed the need for robust, field-tested multi-channel PAM systems for extreme environments, while the BuzzCam project specialized this acoustic sensing approach for the challenging task of monitoring specific insect pollinators [56,59].

2.3.3 Biologging and Animal-Borne Tracking Technologies: Insights from the Animal's Perspective

Biologging—the use of animal-attached electronic devices—has provided unprecedented insights into the movement, behavior, physiology, and energetics of free-ranging animals [63,64]. Early wildlife tracking relied on VHF radio telemetry, but the advent of GPS and satellite telemetry revolutionized the field, enabling precise tracking over vast distances [65,66]. Modern biologgers increasingly move beyond location by integrating a suite of additional sensors. Inertial Measurement Units (IMUs) provide fine-grained data on movement and posture to classify behaviors like foraging or resting, while physiological and environmental sensors can offer insights into an animal's energetic expenditure, stress levels, and immediate ambient conditions [67–69].

However, current biologging devices still face significant challenges and limitations. A fundamental trade-off exists between power consumption and size, as smaller animals require lighter tags with smaller batteries, which limits sensor capabilities and deployment duration. Data retrieval can also be difficult, with remote download options often having bandwidth limitations and device recovery being uncertain. Furthermore, the high cost of advanced commercial collars can limit study sizes. A key limitation is the lack of platforms that cohesively integrate high-fidelity bioacoustic recording, comprehensive environmental characterization, and fine-grained movement sensing in a single, power-efficient package. This siloed approach makes it difficult to answer complex ecological questions about how an animal's immediate environmental context and its vocal communication are linked to its behavior and physiology. The CollarID platform detailed in Chapter 5 specifically aims to address this gap by engineering a lightweight, low-power, yet highly multi-modal animal-borne sensor package that integrates inertial sensing, bioacoustics, a comprehensive suite of environmental sensors, and long-range communication.

This review of ecological monitoring approaches highlights a clear trend towards leveraging technology for more detailed, less invasive, and broader-scale data collection. However, challenges related to hardware robustness, power management, data volume, and the integration of diverse sensing modalities persist, particularly for achieving truly comprehensive, contextualized insights. The following section will explore how machine learning is being applied to help overcome some of these data-related challenges.

2.4 Machine Learning in Ecological Monitoring and Sensing

The increasing deployment of advanced sensor technologies in ecological research, as discussed in Section 2.3, has led to an exponential growth in the volume and complexity of collected data. Manually analyzing these vast datasets—whether hours of acoustic recordings or continuous streams of accelerometer data—is often infeasible, creating a significant "data-to-insight" bottleneck [70]. Machine Learning (ML) has emerged as an indispensable tool for automating the processing and interpretation of this ecological sensor data, enabling researchers to extract meaningful patterns, classify events, and ultimately enhance their understanding of ecological systems. This section reviews key applications of ML in ecological monitoring, with a particular focus on bioacoustic analysis and the emerging potential of edge computing (TinyML) for in-situ intelligence, which directly informs the ML development we undertook for the BuzzCam project.

2.4.1 Automated Species Identification and Analysis from Acoustic Data

Passive Acoustic Monitoring (PAM) generates extensive audio archives, and Machine Learning (ML) has become central to efficiently processing these recordings for species detection, classification, and abundance estimation. Early attempts at automating bioacoustic analysis relied on traditional signal processing techniques to extract acoustic features for use in conventional classifiers like Support Vector Machines (SVMs) or Hidden Markov Models

(HMMs) [71,72]. While successful for specific tasks, these methods often required extensive manual feature engineering. The advent of deep learning, particularly Convolutional Neural Networks (CNNs) for analyzing spectrograms, has since revolutionized the field by enabling models to automatically learn relevant features from data [73,74]. Numerous studies have demonstrated the success of CNNs in classifying bird songs, frogs, and marine mammal vocalizations, with frameworks like ANIMAL-SPOT providing generalized architectures for such tasks [52,54,75–78]. Despite these successes, challenges remain, including the need for large, accurately labeled training datasets, model generalizability across different regions, and robustness to background noise, which the creation of the BuzzCam dataset was, in part, an effort to address [56,70,79].

2.4.2 Machine Learning for Behavioral Classification from Animal-Borne Sensor Data

Beyond acoustic data, ML is also extensively used to interpret data from animal-borne sensors (biologgers), particularly Inertial Measurement Units (IMUs), to automatically classify animal behaviors. Accelerometers, which capture fine-scale body movements, are often used to train ML algorithms ranging from decision trees to deep learning models to classify activities such as foraging, resting, or traveling [67,80]. To achieve more accurate and nuanced classifications, researchers are increasingly fusing data from multiple sensors—such as accelerometers, magnetometers, and GPS—using ML as a powerful framework for integration [64]. In situations where labeled data is scarce, unsupervised techniques like clustering are also being explored to discover novel behavioral patterns directly from raw sensor data [81]. The potential to apply these ML techniques to data from the multi-sensor CollarID platform is a key future direction, aiming to move beyond just tracking location to understanding detailed animal behavior in context.

2.4.3 Edge Computing and TinyML: Enabling In-Situ Intelligence for Ecological Sensors

While cloud-based or powerful offline ML processing has enabled many advances, there is a growing trend towards deploying ML directly on the sensor devices themselves—a paradigm known as edge computing or, for highly resource-constrained devices, TinyML [82,83]. This on-device approach offers several advantages for ecological applications, including reduced power consumption by minimizing data transmission, lower latency for real-time event detection, enhanced data minimization and privacy, and greater autonomy for scalable sensor networks [84].

However, deploying ML on microcontrollers presents significant challenges. Standard deep learning models are often too large, necessitating techniques like pruning and quantization to create efficient models that can operate within the limited memory and processing power of MCUs [85,86]. This process also requires specialized software frameworks and toolchains. While still a nascent field, examples of TinyML for ecological applications are beginning to emerge, such as on-device detection of animal presence or specific calls [87–91].

A key contribution of my PhD research is the development of a complete end-to-end pipeline for creating an on-device ML classifier for a specific and challenging ecological target—differentiating *Bombus* bee species by their flight buzzes—on a commercially available, low-power MCU. This work addresses the need for practical, field-deployable intelligent sensors for real-time ecological monitoring. The integration of ML, particularly through the advancements in TinyML, holds immense potential to transform ecological sensing from a data-collection-centric activity to an insight-driven one, enabling more responsive, scalable, and intelligent approaches to understanding and conserving our natural world.

2.5 Synthesis and Identification of Research Gaps

The preceding review has traversed several interconnected domains pertinent to this dissertation: the nuanced understanding and sensing of human comfort in built environments (Section 2.2), traditional and technological approaches to ecological monitoring (Section 2.3), and the transformative role of machine learning in processing complex sensor data, including the burgeoning field of on-device TinyML (Section 2.4). Across these areas, while significant advancements have been made, the literature also reveals persistent challenges and salient research gaps. This section synthesizes these identified limitations and articulates how the overarching aims and specific projects undertaken in my PhD research are designed to address these gaps, thereby establishing the novelty and contribution of this work.

2.5.1 Recapitulation of Key Limitations and Identified Research Gaps

Synthesizing the discussions from Sections 2.2, 2.3, and 2.4, several key research gaps emerge:

1. Holistic and In-Situ Human Comfort Sensing: Despite a sophisticated understanding of IEQ factors and the availability of various physiological sensors, there remains a scarcity of integrated, wearable platforms that can holistically capture an individual's proximate environmental conditions and relevant physiological responses simultaneously, in real-world, everyday contexts. Existing solutions are often limited in the scope of IEQ parameters measured by a single wearable, or they focus primarily on physiology without rich, co-located environmental data. Furthermore, research exploring novel interaction modalities to enhance occupant awareness and agency based on such personalized, in-situ data is still developing.

2. Robust and Specialized Acoustic Monitoring Hardware for Ecology: While Passive Acoustic Monitoring (PAM) is increasingly utilized, there is a continued need for robust, field-tested hardware specifically designed for challenging environmental conditions (e.g., Arctic deployments) and for specialized applications like multi-channel spatial audio analysis in ecological contexts. Moreover, dedicated, optimized PAM systems for specific, acoustically subtle taxa like insect pollinators (e.g., bees) are not widely available, hindering scalable monitoring efforts.

3. Comprehensive, Multi-Modal Animal-Borne Biologging Platforms: Current commercial and research-grade animal tracking collars, while providing invaluable location and

often activity data, frequently lack comprehensive, co-registered environmental sensing capabilities (temperature, humidity, air quality, light) and high-fidelity bioacoustic recording. There is a significant opportunity to develop more versatile, lightweight, low-power platforms that integrate a richer suite of sensors to provide a more holistic understanding of an animal's experience, behavior, and its interaction with its immediate environment, applicable to a diverse range of wildlife.

4. Practical Application of On-Device ML (TinyML) for Real-Time Ecological Insight: While machine learning has revolutionized the offline analysis of ecological sensor data, the practical, end-to-end development and field validation of on-device ML solutions (TinyML) for real-time classification of specific ecological targets (like endangered bee species) on resource-constrained, low-power microcontrollers remains an emerging and challenging area. Bridging the gap from powerful research models to efficient, deployable on-device intelligence is critical for enabling scalable and autonomous ecological monitoring.

5. Lack of Cross-Contextual Methodologies and Adaptable Technologies: Much research tends to be siloed within either human-centric sensing or specific ecological domains. There is a less explored area concerning the development of adaptable sensing methodologies and technological foundations that can be evolved and applied across these different contexts, leveraging learnings from one domain to inform another.

2.5.2 The Overarching Need for Integrated, Context-Aware, and Intelligent Sensing Solutions Across Domains

Collectively, these identified gaps point towards an overarching need for research that focuses on developing more integrated, context-aware, and intelligent sensing solutions. "Integrated" implies the fusion of multiple sensing modalities to create a richer, more complete picture of the animal-environment interaction. "Context-aware" suggests systems that can understand and adapt to the specific environment, the state of the animal being monitored (be it human or other living organism), and the particular research question at hand. "Intelligent" points towards the incorporation of on-device processing and artificial intelligence to transform raw sensor data into actionable insights efficiently and autonomously, directly at the point of collection. Furthermore, my review highlights the value in exploring how technological platforms and design principles can be adapted and translated across seemingly disparate domains—from understanding human well-being in buildings to conserving biodiversity in natural ecosystems. The engineering challenges of miniaturization, low-power design, robust field deployment, multi-modal data fusion, and practical ML implementation often share common principles, even if the specific applications differ. A research approach that embraces this cross-contextual learning and technological evolution can lead to more versatile, impactful, and innovative sensing solutions.

2.5.3 How This Dissertation Addresses the Identified Gaps

The body of work presented in this dissertation directly confronts these identified gaps and contributes to the overarching need for advanced sensing solutions through the development, deployment, and evaluation of three novel sensing platforms and associated methodologies:

- The AirSpecs platform (detailed in Chapter 3) addresses the gap in holistic and in-situ human comfort sensing. By uniquely integrating a comprehensive suite of IEQ and physiological sensors into a smart-eyeglass form factor, coupled with a mobile application for EMA and interaction, AirSpecs provides a novel tool for exploring comfort dynamics and environmental awareness in real-world settings, as demonstrated by the international study and the resulting public dataset [5,56].

- The SoundSHROOM and BuzzCam projects (detailed in Chapter 4) tackle critical needs in ecological acoustic monitoring.

  - SoundSHROOM addresses the demand for robust acoustic monitoring hardware for extreme environments through its design and successful Arctic deployment, yielding a unique multi-channel dataset [59].
  - BuzzCam specifically targets the challenge of scalable, non-invasive pollinator monitoring and the practical application of on-device ML. My development of the BuzzCam system, its associated Patagonian bee dataset [56], and critically, the on-device ML classifier for *Bombus* bees, represents a significant step towards real-time, intelligent ecological monitoring at the sensor edge.

- The CollarID platform (detailed in Chapter 5) is my response to the need for more comprehensive, multi-modal animal-borne biologging platforms. Its engineering, integrating inertial, bioacoustic, extensive environmental sensing, and long-range communication capabilities, aims to provide a versatile and robust tool for studying a diverse range of wildlife, thereby addressing limitations in current commercial and research-grade animal tracking devices.

Through these distinct yet interconnected projects, this dissertation embodies a cross-contextual exploration in environmental and ecological sensing. I demonstrate how iterative design processes, methodological adaptations, and technological innovations can be leveraged to develop tailored solutions for diverse sensing challenges, ultimately aiming to provide deeper insights into the complex interactions between organisms and their environments. The following chapters will detail the methodologies, specific contributions, results, and implications of each of these research endeavors.

A critical challenge in advancing sensing technologies across both human and ecological domains is the integration of multi-modal sensors into cohesive, functional systems. This integration presents several technical difficulties, including data synchronization, power management, and miniaturization. Data synchronization is essential for ensuring that diverse sensor streams (e.g., environmental, physiological, and inertial data) are temporally aligned, which is crucial for accurate context-aware analysis. Power management is another significant hurdle, particularly in wearable and animal-borne devices, where energy efficiency directly impacts deployment duration and data collection continuity. Additionally, miniaturization poses a challenge, as devices must be compact and lightweight to be unobtrusive, especially in wildlife monitoring where animal comfort and natural behavior are priorities. These challenges underscore the need for innovative engineering solutions, which this thesis addresses through the development of platforms like AirSpecs and CollarID.

2.5.4 Ethical Considerations in Cross-Contextual Sensing

The development and deployment of sensing technologies in both human-centric and ecological contexts raise important ethical considerations. In human studies, privacy is a paramount concern, particularly when collecting personal data such as physiological signals and environmental exposure. Ensuring informed consent, data anonymization, and secure storage are essential to protect participants' rights and maintain trust. In ecological monitoring, especially with animal-borne devices, ethical considerations focus on minimizing stress and ensuring the welfare of the animals. This includes designing lightweight, non-invasive sensors and adhering to strict deployment protocols to avoid disrupting natural behaviors. Addressing these ethical dimensions is crucial for the responsible advancement of sensing technologies, and this thesis integrates these considerations into the design and deployment of its platforms.

Chapter 3

AirSpecs - Sensing Human-Environment Interaction and Comfort In-the-Wild

3.1 Introduction: Understanding Human Comfort in Context

The environments in which individuals live, work, and transit profoundly influence their well-being, productivity, and overall quality of life. Within the domain of Human-Computer Interaction (HCI) and architecture, there is a growing acknowledgment that human comfort in built environments is a deeply subjective and multifaceted experience, extending beyond standardized metrics of temperature or illumination [5,16]. This chapter details my development of AirSpecs, a novel smart-eyeglass platform, and its application in an international in-the-wild study, designed to capture a more holistic and contextualized understanding of human comfort and the intricate dynamics of human-environment interaction.

3.1.1 The Complexity of Human Comfort: Beyond Standardized Metrics

For decades, research into human comfort, particularly thermal comfort, has been heavily influenced by laboratory-based studies aiming to define universal standards [9]. While foundational, these approaches often simplify the rich tapestry of factors that contribute to an individual's perceived comfort. As we highlighted in Zhong et al. [5], "recent studies on human experiences in smart built environments...widely acknowledge that comfort should be studied and designed first and foremost as a subjective experience." This subjective experience is shaped not only by physical Indoor Environmental Quality (IEQ) parameters—such as thermal conditions, air quality, lighting, and acoustics [8]—but also by individual physiological responses, psychological states (e.g., stress, focus), personal preferences, cultural backgrounds, and the degree of control individuals have over their surroundings [17,18]. The increasing automation in modern buildings, while aimed at energy optimization, can inadvertently lead to a "diminished perception of comfort and compromised long-term user awareness and satisfaction" if human agency and subjective experience are not prioritized [5].

3.1.2 The Need for In-Situ, Personalized, and Holistic Sensing

To truly understand and design for human comfort in its natural complexity, research methodologies must move beyond controlled laboratory settings and into the "wild" – the diverse, dynamic environments where people actually spend their time. As we noted in our urban comfort dataset study [3], "In developed countries, we typically spend over 90% of our time indoors, yet indoor environments exhibit significant variability both between and within buildings." Longitudinal studies that encompass periods spent outside buildings and during commuting are also crucial, as these transitional experiences can significantly impact overall comfort and exposure [3]. This necessitates sensing tools that are:

- *In-Situ and Context-Aware:* Capable of capturing data within the real-world contexts of daily life.
- *Personalized:* Able to link environmental measurements directly to an individual's physiological responses and subjective feedback.
- *Holistic:* Covering multiple dimensions of IEQ and relevant physiological indicators simultaneously.
- *Continuous or Semi-Continuous:* Providing data over extended periods to capture dynamic changes and long-term patterns.

While commercially available sensor-rich systems exist, they often abstract away system intricacies, and dedicated solutions for continuous environmental monitoring on the person are scarce [3]. Existing DIY or commercial wearables often focus on a single IEQ dimension or primarily on physiological sensing, "lacking holistic evaluation and feedback on physiological, mental, and physical dimensions that interactively contribute to the complexity of comfort" [5,16].

3.1.3 Introducing AirSpecs: Vision and Specific Objectives for This Research

It was to address these gaps that I conceived and undertook the development of the AirSpecs platform (Figure 3.1). The vision behind AirSpecs was to create a research tool in the form of smart eyeglasses, leveraging their unique position near key human sensory organs (eyes, ears, nose, mouth) to provide rich, contextualized data on an individual's immediate micro-environment and physiological state. As I described in earlier work [30], AirSpecs evolved from prior work on informed decision-making using sensor networks [92] and the open-source Captivates platform [93], but with a significantly redesigned sensor suite tailored for environmental monitoring and comfort research.

The specific objectives for the AirSpecs project within my PhD research were multifaceted. The first was to design and engineer a novel, extensible smart-eyeglass platform capable of holistically sensing key IEQ parameters and relevant physiological signals in close proximity to the user's face. This was followed by the development of an accompanying mobile application ecosystem for iOS and watchOS, designed to interface with AirSpecs for data streaming, provide users with intuitive data visualizations, and facilitate Ecological Momentary

Figure 3.1: Prototypes of the AirSpecs Smarteyeglass Platform [30]

Assessment (EMA) through novel interaction modalities like peripheral LED cues. The complete system was then deployed in a multi-site, international in-the-wild study across Boston (USA), Fribourg (Switzerland), and Singapore to collect a comprehensive dataset [3]. Finally, this deployment allowed us to investigate the interplay between environmental conditions, physiological responses, and subjective comfort, and to explore how access to such personalized data influences occupant awareness and interaction with their environment [59].

Ultimately, the AirSpecs project sought to contribute not only a novel sensing tool and a valuable dataset but also deeper insights into how technology can mediate a more nuanced and human-centric understanding and management of comfort in our increasingly intelligent built environments.

3.1.4 Chapter Overview

This chapter will detail the AirSpecs project in its entirety. Section 3.2 will describe the system design and architecture, covering both the AirSpecs hardware and the accompanying software ecosystem. Section 3.3 will outline the methodology of the international "Urban Comfort" study conducted using the AirSpecs platform. Section 3.4 will present the key findings and analyses from this study, particularly focusing on insights related to comfort awareness and human-building interaction, drawing significantly from our publication "Sensors and Sensibilities: Exploring Interactions for Habitat Comfort with An Environmental-Physiological Sensing Eyewear In the Wild" [5]. Section 3.5 will detail the "Dataset exploring urban comfort through novel wearables and environmental surveys" [3] as a significant output of this research.

Finally, Section 3.6 will discuss the broader implications and learnings from the AirSpecs project, and Section 3.7 will conclude the chapter, providing a transition to the subsequent ecological sensing work undertaken in this PhD.

3.2 AirSpecs: System Design and Architecture

To enable the in-situ investigation of human comfort and environmental interaction as outlined in the objectives (Section 3.1.3), I designed and developed the AirSpecs smart-eyeglass hardware platform and co-developed its accompanying mobile application. This section details the rationale behind the smart-eyeglass form factor, the specific hardware components and sensor suite integrated into AirSpecs, and the architecture of the mobile application and data infrastructure that support its operation.

3.2.1 Rationale for Smart Eyeglasses as a Sensing Platform: Proximity and Context

The decision to base the AirSpecs platform on a smart-eyeglass form factor was a deliberate one, driven by the unique advantages eyeglasses offer for environmental and physiological sensing in the context of human comfort research. As noted in Chwalek et al. [3] and Zhong et al. [5], smart glasses are "positioned near the eyes, ears, mouth, and nose—critical areas for visual and auditory perception, as well as inhalation and exhalation of contaminants." This proximity allows for the measurement of Indoor Environmental Quality (IEQ) parameters at locations highly relevant to an individual's direct exposure and perception. For instance, air quality sensors on the nose bridge can provide more representative data on inhaled air compared to room-level sensors [22]. Furthermore, the head-worn nature allows for continuous monitoring as the user moves through different micro-environments, both indoors and during transit. It is important to note, however, that during such transit, certain sensor channels, such as blink detection and skin temperature, can be sensitive to motion artifacts and rapidly changing environmental parameters, like quick variations in infrared radiation from sunlight. These periods of high motion, which can be identified using accelerometry data, introduce noise that must be considered when filtering and analyzing the data.

My development of AirSpecs built upon the foundation of the open-source Captivates smart-eyeglass platform [93], which prioritized user comfort and extensibility for psychophysiological monitoring. However, for AirSpecs, I significantly redesigned the sensor suite to have a strong "environmental monitoring focus" and developed a new "toolset for prototyping and software development" [30].

3.2.2 AirSpecs Hardware Design: A Holistic Environmental and Physiological Sensor Suite

The AirSpecs hardware integrates a comprehensive array of sensors into a standard eyeglass frame, aiming for a balance between sensing capability, wearer comfort, and unobtrusiveness. My engineering effort focused on miniaturizing and robustly packaging these sensors. The key hardware components are detailed below (and summarized in Table 3.1 and Figure 3.2):

Figure 3.2: Labeled overview of sensors in the AirSpecs Smarteyeglass Platform [30]

- Integrated IEQ Sensor Suite:
  - *Air Temperature and Humidity:* An SHT45 sensor (Sensirion) is located on the temple (sideboard) and the nose bridge to capture ambient thermal conditions. The dual placement allows for comparison and understanding of microclimates around the face.
  - *Volatile Organic Compounds (VOCs) and Nitrogen Oxides (NO_x):* An SGP41 sensor (Sensirion) is integrated into both the temple (sideboard) and the nose bridge to measure these key air pollutants. Placement on the nose bridge aims to capture data close to the user's breathing zone [22,30].
  - *Illuminance (lux) and Light Spectrum:* A TSL27721 ambient light sensor (AMS) and an AS7341 11-channel spectral sensor (AMS) are located on the nose bridge to measure light intensity and its spectral composition, crucial for assessing visual comfort.
  - *Indoor Air Quality Index (IAQ) and Equivalent CO_2 (eCO_2):* A BME688 sensor (Bosch Sensortec) on the nose bridge provides an IAQ index and an estimated eCO_2 level, offering further insights into air quality.
  - *Noise Level (dBA) and Audio Frequency Analysis:* An ICS-43434 MEMS microphone (TDK InvenSense) is positioned on the left temple. The firmware I developed processes its output to calculate ambient noise levels (dBA) and perform a basic audio frequency analysis. As stated in Zhong et al. [5], "raw acoustic information is not transmitted" to preserve privacy, only the derived metrics.
- Integrated Physiological Sensor Suite:
  - *Skin Temperature:* To explore potential correlates of cognitive load or stress [29], I integrated multiple TPIS 1S 1385 non-contact thermopile sensors (Excelitas). These are strategically placed to measure skin temperature on the temple (right,

inside), the back of the nose bridge, and the right nose pad. As noted in Chwalek et al. [30], "For the face temperature sensing, we improved the original design by also including two additional contact-less temperature sensors to increase the accuracy of the readings across different head sizes."

  - *Blink Detection:* A QRE1113 miniature reflective object sensor (onsemi) is placed on the left nose pad, directed towards the eye, to detect blinks, which can also be an indicator of visual comfort, fatigue, or cognitive state.

- Core Electronics and Physical Construction:

  - *Microcontroller Unit (MCU):* STM32WB5MMGH6TR module was chosen for its Bluetooth Low Energy (BLE) capabilities, multiple communication peripherals allowing us to have multiple data busses, physical size, and low-power operation.
  - *Power:* The system is powered by two 550 mWh lithium-polymer batteries, designed for a full day of operation.
  - *Connectivity:* BLE is used for wireless communication with the companion mobile application.
  - *Printed Circuit Boards (PCBs):* I designed custom flexible and rigid PCBs to integrate all sensors and electronics within the constraints of the eyeglass frame.
  - *Peripheral LED:* A key interactive element I incorporated is a peripheral RGB LED light, used for providing subtle cues to the user, such as survey availability [37].

Parameter	**Sensor**	**Sample rate**	**Accuracy**	**Location on glasses**
Air temperature Humidity	SHT45	Every 5 sec	±0.1°C ±1.0%	Temple (right, outside as sideboard), bridge (front)
VOC NO_x	SGP41	Every 5 sec	±15% ±50 ppb	Temple (right, outside as sideboard), bridge (front)
Illuminance (lux)	TSL27721	Every 1 sec	-	Bridge (front)
Spectrum	AS7341	Every 5 sec	-	Bridge (front)
IAQ (e)CO_2	BME688	Every 5 sec	±15% ±15%	Bridge (front)
Noise (dBA) Audio Frequency	ICS-43434	48000 Hz (activate 85 ms every min)	- -	Temple (left, outside)
Skin temperature	TPIS 1S 1385	Every 1 sec	±0.3°C	Temple (right, inside), bridge (back), nose pad (right)
Blink	QRE1113	1000 Hz	-	Nose pad (left)

Table 3.1: Summary of sensing parameters, their sampling settings, and corresponding locations on the AirSpecs device [5].

Figure 3.3: System overview of AirSpecs device and apps [30]. The sensor readings of the device in the top figure (except those in *italic*) are accessible by the user. Most readings are presented in raw form, except that skin temperature data was used to estimate cognitive load. Screenshots of the AirSpecs apps design in iOS and the corresponding watchOS are shown, including home, historical records, settings, and the survey. The watchOS app aims to provide spontaneous information, so we eliminated the historical records and settings screens in its design. Two sets of sensors that measure air temperature, humidity, VOC, and NO_x are located in the left temple (on a sideboard) and bridge, considering that measurements near the breathing area and the surrounding environment can differ [5]

3.2.3 AirSpecs Software Ecosystem: Mobile Applications and Data Infrastructure

To manage data flow, provide user interaction, and enable research data collection, I co-developed a software ecosystem centered around an iOS mobile application, with a companion watchOS app. The primary iOS application managed the BLE connection to the AirSpecs glasses, receiving and processing the continuous stream of sensor data. It offered users several key features, including a "Home" screen for real-time visualizations of key IEQ dimensions (Figure 3.3b), a "My Data" screen for viewing historical trends to facilitate self-reflection (Figure 3.3a), and customizable comfort range settings that influenced the visual feedback on the main display (Figure 3.3c). For research purposes, the app incorporated a micro-EMA survey module, inspired by the Cozie app [36], which presented brief, contextual questions about comfort and focus triggered by a peripheral LED cue from the glasses (Figure 3.3d-e).

A companion watchOS app provided a convenient alternative for quick EMA responses and viewing essential real-time data (Figure 3.3f-i). All sensor data from AirSpecs and EMA responses from the app were timestamped (UTC), tagged with a unique participant ID, and forwarded to an external server for monitoring and storage in an InfluxDB time-series database [3]. Data serialization across the AirSpecs device, iPhone, and server was managed using Protocol Buffers to ensure data integrity and minimize privacy concerns during transmission [5].

This integrated hardware and software system provided the technological foundation for the in-the-wild urban comfort study detailed in the subsequent sections, enabling the collection of rich, multi-modal, and contextualized data.

3.3 International "Urban Comfort" Study: Methodology

With the AirSpecs platform engineered and its software ecosystem developed (Section 3.2), we embarked on an in-the-wild study to investigate human comfort dynamics across diverse urban environments. This international study was designed to collect rich, contextualized data leveraging the unique capabilities of AirSpecs. This section details the objectives of this study, the participant recruitment strategy across three global cities, the experimental design and procedure followed, and the types of data collected. The methodology described herein formed the basis for the findings presented in Zhong et al. [5] and the dataset published in Chwalek et al. [3].

3.3.1 Study Objectives: Exploring Comfort In-the-Wild

The primary objectives of this international "Urban Comfort" study were multifaceted:

1. To collect a comprehensive, longitudinal dataset capturing synchronized indoor environmental quality (IEQ) parameters, personal physiological responses, and subjective perceptions of comfort and environmental conditions from individuals going about their daily routines in their natural work/live environments.

2. To explore the complex interplay between objective environmental conditions (measured by AirSpecs), physiological reactions (partially captured by AirSpecs and other

wearables), and subjective human experiences (assessed via Ecological Momentary Assessment - EMA).

3. To evaluate the usability and effectiveness of the AirSpecs platform (glasses and mobile applications) as a research tool for in-the-wild, multi-modal data collection over an extended period (five days).

4. To investigate the efficacy of using subtle peripheral LED cues, integrated into the AirSpecs glasses, as a method for prompting participant engagement with micro-EMA surveys with minimal disruption, and to explore this as an objective measure of environmental awareness (a key focus of Zhong et al. [5]).

5. To gather qualitative insights into users' experiences with wearable environmental sensing technology, their interpretation of the data provided, and their imagined future interactions with smart building systems.

3.3.2 Participant Recruitment and Cross-Cultural Demographics

To capture a diversity of cultural backgrounds, climatic conditions, and built environment typologies, we conducted the study across three distinct geographical regions: Boston, Massachusetts, USA (March/April 2023); Fribourg, Switzerland (May/June 2023); and Singapore (June/July 2023). At each location, we recruited ten participants, resulting in a total of 30 participants for the entire study. Recruitment was primarily conducted via lab-wide email advertisements and university bulletin boards, targeting graduate students, university staff, and researchers who were likely to have work tasks requiring concentration and spend significant time in indoor environments. Specific selection criteria included prioritizing individuals likely to have concentration-demanding tasks, no requirement to wear prescription glasses (or ability to wear contact lenses), and not being extremely satisfied with all their current work environments [3].

The final cohort of 30 participants comprised 14 women, 14 men, 1 non-binary/third gender individual, and 1 who preferred not to disclose their gender, with ages ranging from 21 to 52 (Table 3.2). This diverse group provided a range of perspectives and experiences. Participants received compensation in local currency vouchers for their time and involvement.

3.3.3 Experimental Design and In-the-Wild Procedure

The study used a longitudinal, in-the-wild design, with each participant collecting data continuously over five working days, excluding evenings and nighttime hours. We obtained Institutional Review Board (IRB) approval from each respective study site (Massachusetts Institute of Technology, University of Fribourg, and National University of Singapore) prior to commencing recruitment and data collection [3]. The overall study procedure is depicted in Figure 3.4.

The study began with a pre-screening and onboarding process. Interested individuals completed a survey covering their university status, time spent in work locations, and prerequisites such as vision status [3]. Selected participants were then invited to an in-person onboarding session where we explained the study's purpose, data collection procedures, and

PID	Age	Gender	Race/ethnicity	Occupation	Site	Work in built env
1	24	Female	White	Master student	1	No
2	25	Female	White	University staff	1	No
3	29	Non-binary	Hispanic/Latinx	Master student	1	No
4	24	Male	Hispanic/Latinx, White	PhD student	1	No
5	26	Female	East Asian	University staff	1	No
6	21	Male	White	Undergraduate student	1	No
7	22	Female	Asian-American	Undergraduate student	1	No
8	32	Male	East Asian	PhD student	1	Yes
9	21	Female	Hispanic/Latinx, Middle Eastern	Undergraduate student	1	No
10	24	Male	White	PhD student	2	No
11	46	Female	Hispanic/Latinx	Professor	1	Yes
12	31	Male	White	PhD student	2	No
13	27	Male	East Asian, White	PhD student	2	No
14	27	Male	Hispanic/Latinx, White	PhD student	2	No
15	30	Male	White	PhD student	2	No
16	45	Prefer not to say	White	Manager	2	No
17	33	Female	White	PhD student	2	No
18	52	Female	White	PhD student	2	No
19	27	Male	White	PhD student	2	Yes
20	25	Female	White	PhD student	2	No
21	23	Male	Southeast Asian	Undergraduate student	3	No
22	27	Female	East Asian	PhD student	3	Yes
23	23	Female	East Asian	Undergraduate student	3	No
24	24	Male	East Asian	PhD student	3	Yes
25	23	Female	East Asian	Master student	3	No
26	35	Male	South Asian	Master student	3	Yes
27	26	Female	East Asian	PhD student	3	No
28	24	Female	Southeast Asian	PhD student	3	Yes
29	23	Male	South Asian	Undergraduate student	3	No
30	29	Male	East Asian	Postdoc	3	Yes

Table 3.2: 30 participants were selected from 23, 14, and 45 registrations at sites 1–3 based on pre-screening responses prioritizing: 1) graduate students, staff, and researchers with concentration-intensive tasks, 2) those not requiring glasses (or can wear contacts), and 3) those not extremely satisfied with all work environments [3].

privacy measures. We presented a detailed datasheet of all sensor data to be collected and emphasized that no dialogue would be recorded. Participants were then equipped with the AirSpecs glasses, an Empatica E4 wristband, and an Apple Watch if needed, and completed an onboarding survey to collect demographic and baseline sensitivity data [3].

Following onboarding, participants began the five-day data collection period. They were instructed to wear the devices as continuously as possible during their working hours in their usual environments. The AirSpecs glasses continuously streamed environmental and physiological data to the connected iPhone. A core element of the methodology was the use of a peripheral LED on the glasses to prompt Ecological Momentary Assessment (EMA) surveys at random intervals. If a participant did not respond to the light cue, a vibration notification was sent to their Apple Watch. The micro-EMA survey queried their perceived focus, comfort state, and other contextual factors [3,5].

Upon completion of the data collection, participants took part in a post-study session. This began with an exit survey, followed by a semi-structured interview focusing on their experience with the system and their comfort assessment process [5]. Finally, to understand which sensor readings they found most relevant, we conducted a short sensor arrangement co-design session where participants were asked to ideate their preferred data configurations on the AirSpecs app interface (Figure 3.5).

3.3.4 Data Collected: A Multi-Modal Longitudinal Dataset

This comprehensive methodology allowed us to collect a rich, multi-modal dataset for each participant, synchronized via UTC timestamps and unique participant IDs [3]:

1. *Environmental Data from AirSpecs:* Continuous streams of temperature, humidity, VOC, NO_x, illuminance, light spectrum, IAQ, eCO_2, and loudness (dBA).
2. *Physiological Data from AirSpecs:* Continuous streams of skin temperature from multiple facial locations and blink events.
3. *Physiological Data from Empatica E4:* Electrodermal activity (EDA), blood volume pulse (BVP), skin temperature, and 3-axis accelerometry.
4. *Subjective EMA Data from AirSpecs App:* Responses to micro-EMA surveys regarding comfort, focus, flow, context, and discomfort sources, including reaction times to peripheral LED cues.
5. *Activity and Location Data from Cozie App Framework / Apple HealthKit:* Steps, heart rate, and generalized location context (e.g., home, work) via the iOS application.
6. *Survey Data:* Responses from pre-screening, onboarding, and exit surveys.
7. *Qualitative Data:* Transcripts from post-study interviews and outputs from the co-design session.

This multi-faceted data collection approach provided the necessary foundation for both the quantitative and qualitative analyses presented in subsequent sections and publications.

3.4 Key Findings and Analysis from the AirSpecs Study

The international in-the-wild study conducted using the AirSpecs platform yielded a wealth of quantitative and qualitative data, providing novel insights into human comfort dynamics, environmental awareness, user interaction with personalized sensing technologies, and the potential for new modeling approaches. This section synthesizes key findings primarily drawn from our analysis presented in "Sensors and Sensibilities: Exploring Interactions for Habitat Comfort with An Environmental-Physiological Sensing Eyewear In the Wild" [5], and highlights the utility of the collected dataset [3] as demonstrated by subsequent independent research such as Zhang et al. [94].

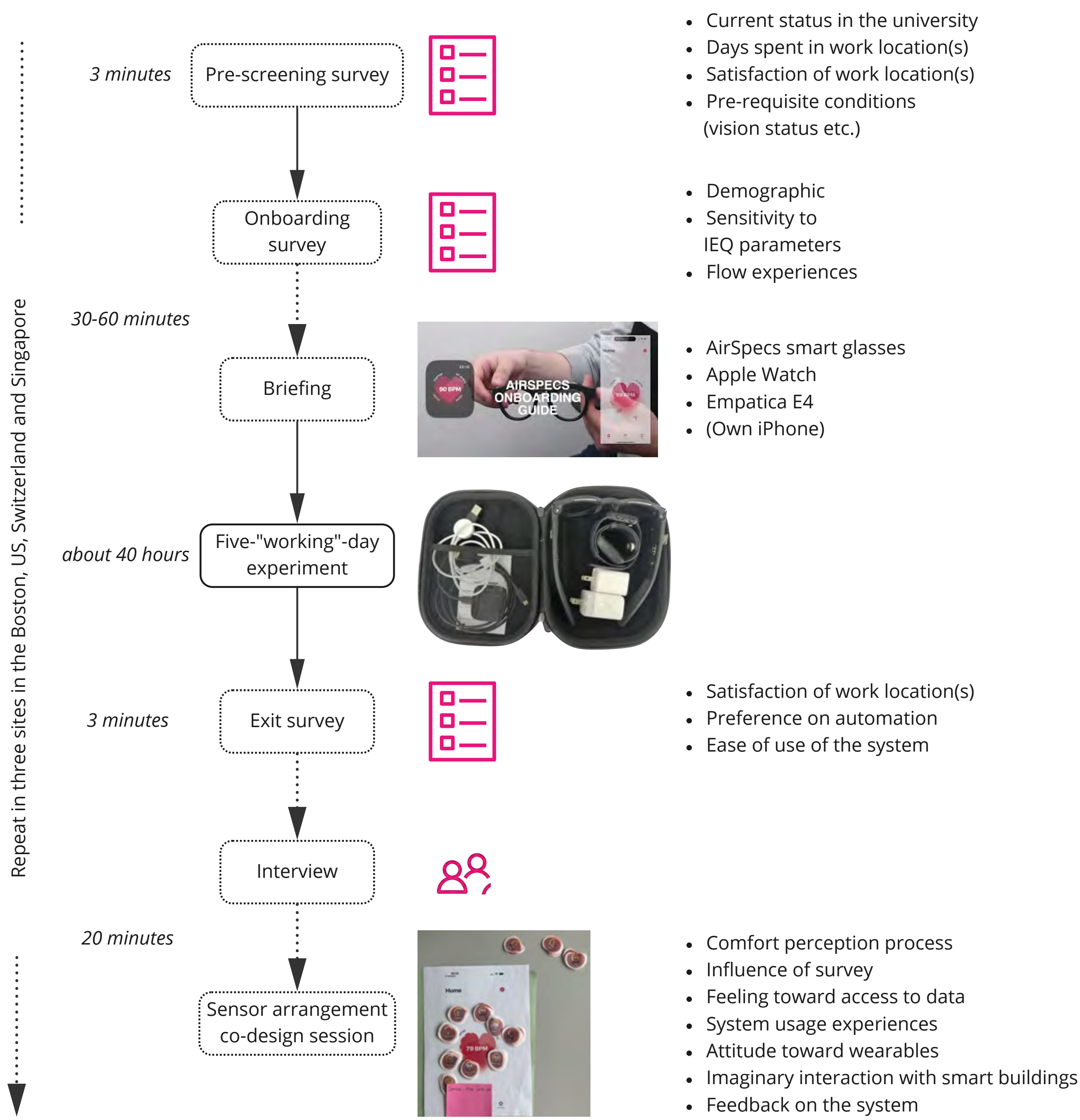


Figure 3.4: Overview of study design [3].

Figure 3.5: Sensor rearrangement co-design examples. The sample results are from P30 (left), P1 (middle), and P18 (right) [5].

3.4.1 Environmental Awareness, Comfort States, and the Efficacy of Peripheral Cues

A core investigation within the AirSpecs study was to explore the relationship between an occupant's comfort state, their focus level, and their awareness of their surrounding environment. As detailed in Zhong et al. [5], we utilized the reaction time to the peripheral LED cue on the AirSpecs glasses as an objective proxy for environmental awareness.

- *Preconscious Nature of Comfort Assessment:* The data on reaction times to the LED prompts indicated that participants did not constantly consciously attend to comfort assessment. Reactions were spread between immediate responses (within seconds) and significantly delayed responses (up to 1750 seconds, near the 15-minute vibration reminder), supporting our hypothesis that the assessment of comfort can indeed be preconscious, existing below the immediate threshold of focused attention before transitioning to conscious evaluation (Figure 3.7).

- *Comfort State Influences Attention to Periphery:* Our quantitative analysis revealed that participants in a "comfy" state were more likely to have prolonged reaction times, sometimes missing the LED cue entirely until the subsequent vibration reminder, compared to when they were in a "not comfy" state. This supported our hypothesis that when users are comfortable, their attention is harder to shift to peripheral environmental cues (Figure 3.6).

- *Interaction between Comfort, Focus, and Environmental Awareness:* Further analysis using Generalized Linear Mixed Models (GLMM) showed a significant interaction effect between comfort state and focus state on reaction time (environmental awareness). As we reported in Zhong et al. [5], the difference in reaction times between "comfy" and "not comfy" states was more pronounced when users reported being highly focused on their primary task (Figure 3.7). Conversely, when users were in a more distracted state, their reaction times (and thus, environmental awareness) were more similar regardless of their comfort level. This suggests that both internal state (comfort) and cognitive engagement (focus) modulate an individual's receptiveness to environmental information.

These findings demonstrate the potential of using subtle, integrated cues in wearable devices to objectively assess environmental awareness and highlight the dynamic interplay between comfort, focus, and perception.

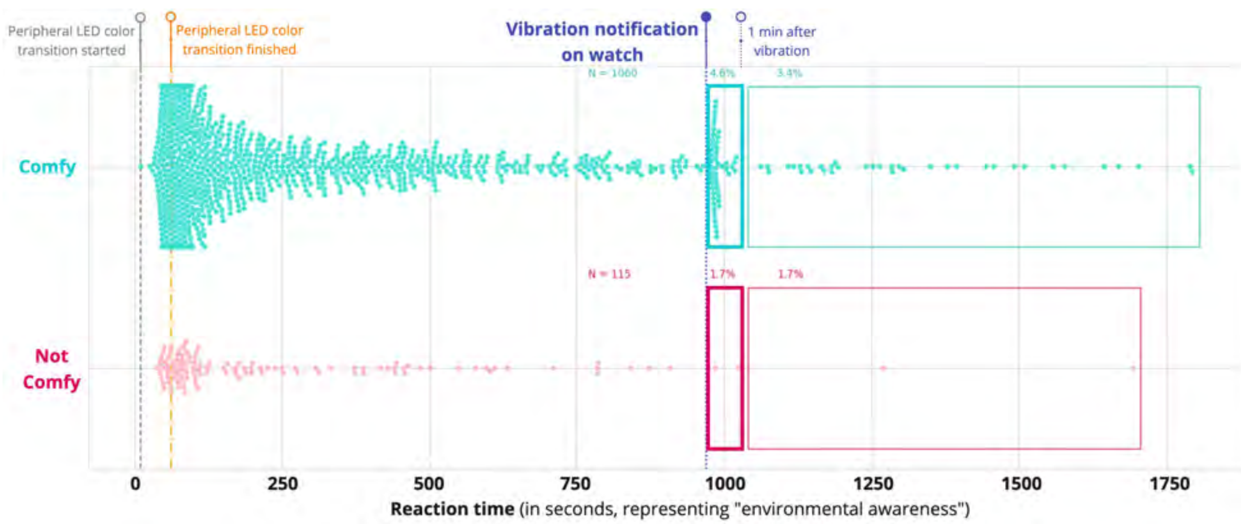


Figure 3.6: Swarm plot of reaction time with respect to comfort state. N is the number of EMA responses in each comfort state. The percentage value is the proportion of points that fall in the highlighted rectangular area with respect to N [5].

3.4.2 User Experiences: Self-Inquiry, Intervention, and Data Interpretation

The qualitative data gathered from post-study interviews and co-design sessions provided rich insights into how participants engaged with the real-time environmental and physiological data provided by AirSpecs [5].

- *Facilitating Self-Inquiry and Action:* Access to personalized data often prompted participants to investigate their environments and how their actions might impact sensor readings. For example, one participant, upon seeing elevated CO_2 levels (around 800 ppm) via the AirSpecs app, "immediately opened the window" [5]. Others used the device to confirm suspicions about environmental issues (e.g., high VOCs after

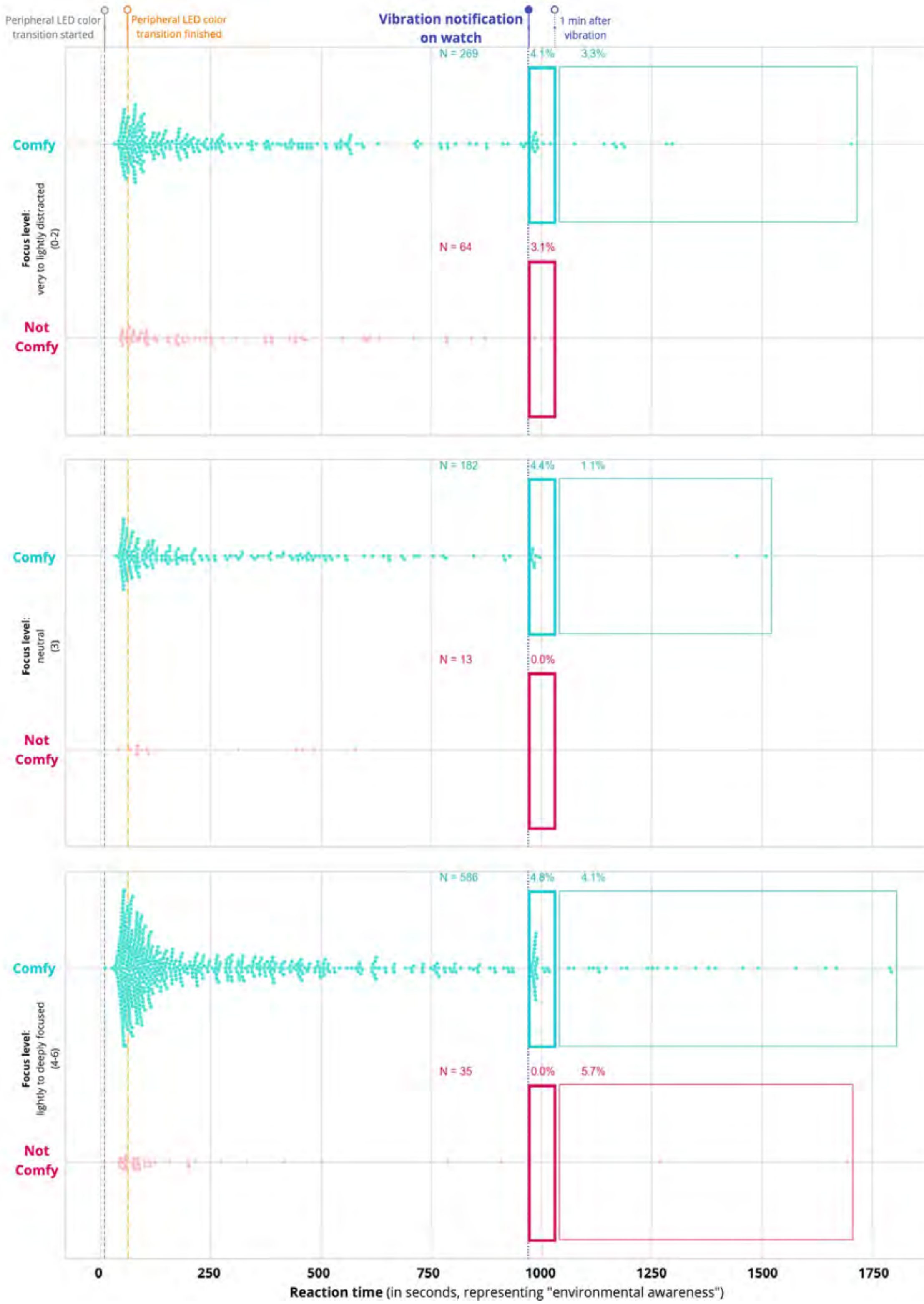


Figure 3.7: Swarm plot of reaction time with respect to comfort state in three focus levels. N is the number of EMA responses in each comfort state. The percentage value is the proportion of points that fall in the highlighted rectangular area with respect to N [5].

painting) or to compare conditions across different spaces (e.g., home vs. office, or during cooking).

- *Varying Data Needs and Preferences:* We observed that while some participants enjoyed exploring all available sensor parameters, others preferred simplified metrics or notifications only for hazardous conditions. A common theme was the desire for data to be presented in a more intuitively understandable way, especially for less familiar metrics like VOCs or NO_x. The proximity of sensors on the eyeglasses (e.g., "nose" vs. "sideboard" for air quality) was also a point of interest, with users finding measurements closer to their breathing zone more personally relevant [5].

- *Personalized Tracking Interests:* The sensor arrangement co-design session provided direct insight into user data priorities. While some participants valued the comprehensive IEQ data, many used the opportunity to ideate the inclusion of more personalized physiological or psychological states, such as stress, hydration, or mood. As shown in the sample results from P30 and P1 , users often grouped or prioritized sensors differently than the default layout, demonstrating a desire to customize the interface to support their own self-inquiry and correlate environmental conditions with their personal well-being [5].

3.4.3 Design Implications for Human-Centric Smart Buildings and Interactions

The findings from the AirSpecs study led us to propose several design implications for future human-building interaction (HBI) systems that aim to be more responsive, adaptable, and human-centric [5].

- *Fluidity of Interaction Modes:* We proposed a framework of three interaction modes—"focus mode" (high automation, minimal interruption), "ambient assistance" (gentle nudges, personalized information when users are preconsciously assessing comfort), and "reflective explanation" (detailed information when users are consciously engaged)—that building systems could dynamically transition between based on the occupant's comfort state, focus level, and context (Figure 3.8).

- *Personalization and Explainability:* The study underscored the need for greater personalization in how environmental data is presented and how automated systems respond. Users desired control over automation levels and explanations for system actions.

- *Designing for Diversity and Inclusiveness:* Our study hinted at the unique needs of neurodivergent individuals or those with heightened sensitivities, suggesting that future "smart agendas" for buildings must move beyond standardized approaches to accommodate cognitive variations and diverse personal strategies for comfort management [5].

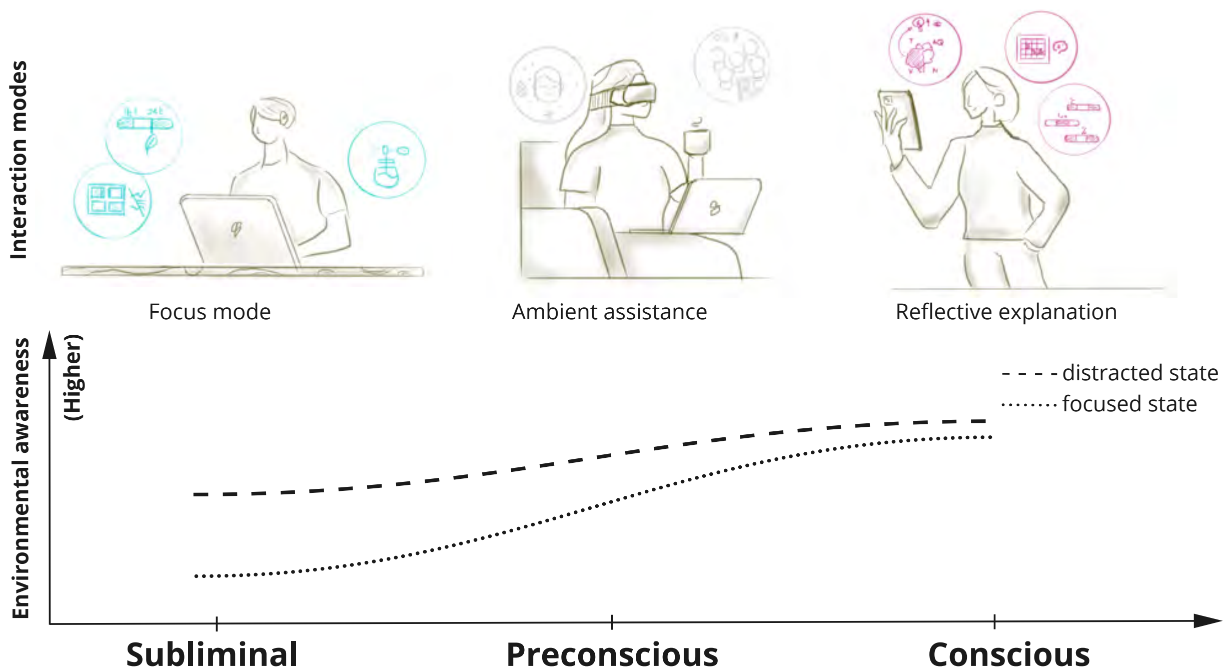


Figure 3.8: Illustrative sketch of the interaction modes with respect to three states of comfort awareness, its interaction with focus state [94].

3.4.4 Utility of the AirSpecs Dataset for Advanced Comfort Modeling: The "Mind the Hazard" Study

A key outcome of our international study was the creation of a comprehensive dataset [3], which includes the AirSpecs sensor data, EMA responses, and contextual information. The value and utility of this dataset for advancing personalized comfort modeling has been independently demonstrated by Zhang et al. [94] in our BuildSys paper, "Mind the Hazard: Modeling and Interpreting Comfort with Personalized Sensing." In Zhang et al., the dataset we collected was utilized to develop a novel modeling framework based on a Neural Ordinary Differential Equations (Neural ODEs)-based Continuous-Time Markov Chain (CTMC) to model the transitions in comfort states over time. The study reports that this model not only predicts comfort states more accurately and stably than conventional classification models but also provides a unique representation of how the hazards of state transitions are influenced by changing ambient and contextual conditions. This "Mind the Hazard" framework facilitates an analysis of how environmental and contextual factors influence the likelihood of transitioning between comfort states, with this analysis showing that factors like temperature, IAQ, and internal unease play critical roles. Ultimately, this suggests that this model's ability to predict comfort transition hazards can enable proactive building management interventions to avoid occupant discomfort, offering a more holistic approach to building control.

The successful application of the AirSpecs dataset by Zhang et al. [94] to develop and validate a sophisticated new comfort modeling technique underscores the richness and research value of the data we collected. It serves as a strong external validation of the utility of the

AirSpecs platform and the dataset itself for advancing the state-of-the-art in personalized comfort research.

3.5 The AirSpecs Urban Comfort Dataset Contribution: A Resource for In-the-Wild Comfort Research

Beyond the specific interaction studies and comfort modeling explorations enabled by the AirSpecs platform, a principal and lasting contribution of this phase of my PhD research is the comprehensive, multi-modal "Dataset Exploring Urban Comfort Through Novel Wearables and Environmental Surveys" [3]. This dataset, meticulously collected during the international in-the-wild study described in Section 3.3, has been curated, documented, and made publicly available to the broader scientific community, serving as a valuable resource for advancing research in human-building interaction, environmental health, and personalized comfort.

3.5.1 Description and Scope of the Published Dataset

The dataset, as detailed in our Scientific Data publication [3], encapsulates a rich array of information captured from 30 participants across three distinct global cities: Boston (USA), Fribourg (Switzerland), and Singapore. Each participant engaged in a five-day data collection period, during which the AirSpecs smart eyeglasses, companion mobile applications, and supplementary wearable sensors (Empatica E4, Apple Watch) were utilized.

- Key Data Streams Included:
  1. *High-Resolution Environmental Data from AirSpecs:* Continuous measurements of ambient air temperature, humidity, Volatile Organic Compounds (VOCs), Nitrogen Oxides (NO_x), illuminance (lux), light spectrum, an Indoor Air Quality (IAQ) index, equivalent CO_2 (eCO_2), and ambient noise levels (dBA), all captured in close proximity to the user's facial sensory organs (Figure 3.3).
  2. *Physiological Data from AirSpecs:* Continuous recordings of skin temperature from multiple facial locations (temple, nose bridge, nose pad) and blink events.
  3. *Subjective Ecological Momentary Assessment (EMA) Data:* Over 1,175 micro-EMA survey responses capturing participants' self-reported comfort states (e.g., "comfy" vs. "not comfy"), perceived focus levels, flow states, desired changes in environmental conditions, sources of discomfort, and their current context. This includes the reaction times to the peripheral LED survey prompts [3].
  4. *Supplementary Physiological and Activity Data:* Data streams from the Empatica E4 wristband (electrodermal activity, blood volume pulse, skin temperature, accelerometry) and Apple Watch/HealthKit (heart rate, step counts) were also collected and are part of the overall dataset structure, providing additional physiological and activity context [3].
  5. *Contextual and Demographic Information:* The dataset includes participants' experiment schedules, pre-screening survey responses (university status, time in

work locations, satisfaction with environments, IEQ sensitivities), demographic information, exit survey responses, and qualitative data from sensor rearrangement co-design sessions [3].

- Data Synchronization and Formatting: All quantitative data streams are synchronized using UTC timestamps and unique participant IDs. The raw sensor data, originally exported from InfluxDB, was consolidated into easily parsable data frames (e.g., pickle format per sensor type) for ease of use by other researchers. Detailed descriptions of all variables, column meanings, and data formats are provided in the supplementary information accompanying the dataset publication ([95], `Summary_of_derived_data.xlsx`).
- Accessibility: The complete dataset has been made openly accessible via the Figshare repository [95], along with detailed documentation and code for data processing where applicable.

3.5.2 Novelty and Significance of the Dataset

The AirSpecs Urban Comfort Dataset offers several novel aspects and holds significant value for the research community. It is one of the first publicly available datasets to utilize smart eyeglasses for holistic and proximate sensing of an individual's micro-environment and physiological responses in real-world settings over multiple days, providing data that is arguably more representative of direct exposure than traditional room-level or wrist-worn sensors. The rich, multi-modal nature of the dataset, which combines detailed environmental sensing, multiple physiological streams, and fine-grained subjective EMA reports, provides an unparalleled opportunity to explore the complex drivers of human comfort. Furthermore, its cross-cultural and cross-climatic scope, with data collected across Boston, Fribourg, and Singapore, allows for investigations into how comfort perceptions may vary across different contexts. The five-day continuous data collection period for each participant provides a longitudinal perspective, enabling the study of dynamic comfort transitions not easily captured in short-term lab studies. As demonstrated by the work of Zhang et al. [94], this dataset serves as a robust foundation for developing and validating new personalized comfort models and advancing methodologies in ubiquitous and wearable computing, with its open availability intended to spur further innovation.

3.5.3 Data Quality Considerations, Validation, and Known Limitations

In publishing this dataset [3], we took care to document aspects related to data quality and any known limitations to ensure transparency and guide its appropriate use by other researchers. We provided technical validation for the non-contact skin temperature sensors on AirSpecs by comparing them against reference iButton temperature loggers, demonstrating their reliability within specified tolerances [3]. As is common in longitudinal in-the-wild studies, some data discontinuities exist due to participants occasionally removing devices, connectivity issues, or data retention policy issues that led to the loss of some physiological data. These known discontinuities are documented to aid researchers in data cleaning and analysis. Finally, the participant pool was primarily composed of university students and staff, which, while diverse across the three sites, may not be fully representative of the broader population in each city ([3], "Usage Notes").

Despite these limitations, the AirSpecs Urban Comfort Dataset represents a substantial and carefully curated contribution. My efforts in leading the design of the AirSpecs platform and our collective efforts in executing the international study and processing the data have resulted in a resource that we believe will significantly benefit the research community by enabling new avenues of inquiry into the nuanced dynamics of human comfort in real-world urban environments.

3.6 Discussion and Learnings from the AirSpecs Project

The AirSpecs project, encompassing the development of a novel smart-eyeglass sensing platform and its deployment in a multi-site international study, provided a rich learning experience and yielded several insights that extend beyond the specific findings presented earlier. This section discusses my reflections on the challenges and successes of conducting in-the-wild wearable sensing studies, the broader implications of this work for smart building design and Human-Computer Interaction (HCI), and acknowledges the limitations of the AirSpecs project itself, which in turn inform future research directions.

3.6.1 Reflections on Conducting In-the-Wild Wearable Sensing Studies

Undertaking a longitudinal, multi-modal, in-the-wild study of this nature presented numerous practical and methodological challenges, from which valuable lessons were learned. The experience highlighted the ongoing challenge of balancing comprehensive sensing capabilities with long-term wearer comfort and the social acceptability of head-worn devices like AirSpecs. Ensuring the robustness of custom-built prototypes against the rigors of daily use also required meticulous engineering and contingency planning ([5], Section 9.4). Maintaining participant adherence and engagement over five days in their natural environments proved to be a significant task, requiring careful study design, clear communication, and robust technical support. Even with attempts to minimize disruption, such as using peripheral LED cues, user engagement still varied based on their current task and focus [5].

Furthermore, collecting and synchronizing diverse data streams from multiple devices across different time zones presented significant data management complexities, underscoring the need for robust backend infrastructure and meticulous data cleaning protocols. The data loss incident for the Fribourg site served as a stark reminder of the importance of redundant backup strategies in longitudinal field studies [3]. Finally, conducting the study across Boston, Fribourg, and Singapore highlighted subtle cross-cultural differences in how participants interacted with the technology and perceived environmental factors, suggesting that future HBI systems need to be culturally sensitive and adaptable ([5], Section 7.4).

3.6.2 Implications for Smart Building Design and Human-Computer Interaction (HCI)

The AirSpecs project and its findings challenge a core assumption in the design of many smart environments: that occupants are always rational actors, ready to receive and act upon environmental information. Our results suggest a more complex reality. The discovery that occupants in a state of comfort and deep focus are *less* receptive to peripheral environmental cues reveals a fundamental paradox for HBI design: the very moment a building is successfully supporting its occupant's primary task is also the moment the occupant is least likely to engage with the building's systems or feedback. This insight implies that 'smart' systems designed to constantly provide data may be fundamentally at odds with human cognitive states. Therefore, the true challenge is not just to provide data, but to design systems with the intelligence to know *when* and *how* to interact. The work from this project offers several implications for the future design of HCI systems aimed at enhancing occupant well-being:

- *Moving Towards Human-Centric Automation:* The study reinforces the idea that purely techno-centric automation in buildings can be suboptimal. Instead, systems that understand and adapt to the occupant's state (comfort, focus), provide personalized feedback, and offer appropriate levels of control are more likely to be effective and well-received. The interaction modes we proposed (focus, ambient assistance, reflective explanation; [5]) offer a conceptual framework for such systems.

- *The Potential of Wearables as Mediators:* AirSpecs demonstrated that personal wearable devices can serve as powerful mediators between occupants and building systems. They can provide personalized environmental data directly to the user, gather subjective feedback, and potentially even actuate personal comfort systems or communicate preferences to building management systems. This shifts the locus of sensing and control closer to the individual.

- *Designing for Awareness and Agency:* By making occupants more aware of their micro-environment and its potential impact on them, systems like AirSpecs can empower them to take proactive steps to improve their comfort and well-being. However, this data must be presented in an understandable and actionable manner ([5], Section 7.2).

- *Rich Data for Personalized Models:* The type of multi-modal, longitudinal data collected by AirSpecs is crucial for developing the next generation of personalized comfort models, as exemplified by the work of [94]. These models can move beyond population averages to predict and cater to individual needs and preferences with much greater accuracy.

3.6.3 Limitations of the AirSpecs Project

While the AirSpecs project achieved its primary objectives, I recognize several limitations. As a research prototype, AirSpecs may not have achieved the level of polish and long-term wear comfort of a commercial product, and some design aspects like pressure points could be further refined ([5], Section 9.4). There were also inherent challenges in measuring IEQ parameters accurately with sensors placed so close to the human body, where factors like breath can affect readings. While participants were advised to interpret these as relative, micro-environmental readings, further research into calibration algorithms or sensor fusion could improve their absolute accuracy, and a more thorough analysis of accelerometry data is needed to better understand the impact of motion artifacts [3,5]. Furthermore, the scope of the study focused primarily on IEQ dimensions and basic physiological correlates, whereas a truly holistic understanding of well-being would require integrating more diverse data streams. Finally, the study's duration of five days, while providing valuable longitudinal data, was not sufficient to understand seasonal variations or long-term adaptation, and the participant sample, being drawn from university communities, may not be fully representative of broader urban populations.

These limitations, however, also point towards exciting avenues for future research, building upon the foundations laid by the AirSpecs project. The development of this platform and the insights gained from its deployment were instrumental in shaping my approach to tackling even more complex sensing challenges in the ecological domain, as will be detailed in subsequent chapters.

While the analysis in this thesis focused on interaction modalities and reaction time as a proxy for focus, the rich physiological data collected by AirSpecs enables deeper, more mechanistic inquiries into human well-being. The platform's unique ability to co-locate environmental and physiological sensors provides a powerful tool for testing specific hypotheses within the existing dataset. For example, future analysis of this dataset could directly test whether acute "in-the-wild" exposure to elevated indoor VOC or eCO_2 levels correlates with measurable changes in blink rate and facial skin temperature. Confirming such a link would provide a more objective physiological grounding for subjective reports of discomfort or "stuffiness," moving beyond correlation to hint at the causal pathways affecting occupant well-being and cognitive performance. This represents a significant opportunity to leverage the existing data to its fullest potential.

3.6.4 Broader Applications and Future Directions

The validated AirSpecs platform and the insights from our initial study open up several avenues for specific, hypothesis-driven future research:

- **Testing the Link Between Air Quality and Cognitive Fatigue:** The platform's ability to co-locate environmental and physiological sensors enables targeted studies into occupational health. Future work could use AirSpecs to directly test the hypothesis that acute, "in-the-wild" exposure to elevated indoor VOC or eCO_2 levels correlates with measurable changes in blink rate and subjectively reported focus, providing a more nuanced understanding of Sick Building Syndrome symptoms [96,97].

- **Developing Proactive, Personalized Thermal Comfort Interventions:** The AirSpecs system provides the necessary data streams to close the loop in human-building interaction. A future study could directly use the real-time skin temperature data and subjective comfort reports from AirSpecs to control a personal comfort device (e.g., a smart desk fan or heater). This would allow for a real-world test of the proposed interaction framework (focus, ambient, and reflective modes) and an evaluation of its impact on both comfort satisfaction and energy consumption.

- **Investigating Environmental Triggers for Clinical Conditions:** For clinical applications, AirSpecs could be deployed to investigate environmental triggers for conditions sensitive to IEQ. For example, a study could equip individuals with asthma to track how personal exposure to particulate matter or specific VOCs, as measured by the glasses, correlates with the timing and severity of their symptoms, potentially leading to more effective personalized avoidance strategies.

- **Validating Advanced Comfort Models:** The public dataset from this work has already been used to validate a novel comfort model. A direct future step is to use the AirSpecs platform for longitudinal data collection specifically designed to train and compare competing personalized comfort models, assessing not only their predictive accuracy but also their stability and computational efficiency for real-world implementation in smart building systems.

3.7 Chapter Conclusion: Laying the Groundwork for Cross-Contextual Sensing

This chapter has detailed the conception, design, development, and empirical application of the AirSpecs smart-eyeglass platform. My primary goal with this project was to create a novel research tool capable of providing rich, contextualized insights into the complex dynamics of human comfort and environmental interaction within real-world built environments. Through a multi-site international study, we successfully demonstrated the AirSpecs system's utility in collecting holistic, in-situ data encompassing environmental parameters, physiological responses, and subjective user feedback. The key achievements presented in this chapter include the engineering of the AirSpecs hardware and its accompanying software ecosystem [3,5]; the generation of key findings on the interplay between comfort, focus, and environmental awareness [5]; and the creation and public dissemination of a unique and valuable multi-modal dataset that has already begun to facilitate new research in personalized comfort modeling [3,94].

The AirSpecs project directly contributed to my overarching PhD objective of developing and applying novel sensing technologies across diverse contexts. It provided me with invaluable experience in designing user-centric wearable sensor systems, managing complex multi-modal data from challenging in-the-wild deployments, considering the nuances of human-computer interaction with environmental data, and navigating the intricacies of international research collaborations. The methodological insights gained, particularly regarding participant engagement, data integrity, and the iterative design of field-deployable prototypes, were instrumental.

This foundational work in understanding human-environment dynamics through personalized, wearable sensing, and the practical engineering skills honed during the AirSpecs project, directly spurred my interest and equipped me to apply similar principles of custom sensor development and robust field deployment to address pressing challenges in a vastly different domain: ecological monitoring. The insights into sensor integration, data management, and the importance of contextual information gathered with AirSpecs provided a strong springboard for tackling the complexities of monitoring elusive or sensitive species in their natural habitats. The subsequent chapter will detail how I leveraged this foundation to develop novel acoustic sensing platforms, SoundSHROOM and BuzzCam, aimed at advancing our understanding and conservation of biodiversity.

Chapter 4

SoundSHROOM & BuzzCam: Advancing Ecological Acoustic Monitoring with On-Device AI for Bees

4.1 Introduction to Ecological Acoustic Monitoring

The natural world is replete with sounds, from the intricate songs of birds to the subtle rustling of leaves and the distinct flight buzzes of insects. These acoustic signals form a rich tapestry of information that, if systematically captured and analyzed, can offer profound insights into the state and dynamics of ecosystems. This chapter details the development and application of novel acoustic sensing technologies aimed at harnessing this acoustic information for ecological monitoring, with a particular focus on advancing tools for pollinator conservation through on-device machine learning. We begin by exploring the value of sound in ecological studies and then introduce the specific projects undertaken: SoundSHROOM, a multi-channel acoustic recorder designed for harsh environments, and BuzzCam, a specialized system culminating in an on-device artificial intelligence (AI) classifier for endangered bees.

4.1.1 The Value of Sound in Ecological Studies

Bioacoustics, the study of sound production, dispersion, and reception in animals (including humans), has emerged as a powerful and increasingly accessible discipline within ecology and conservation biology [57,98]. Sound serves as a primary mode of communication for a vast array of species, playing critical roles in mate attraction, territory defense, parent-offspring interactions, predator-prey dynamics, and social cohesion [99]. Consequently, the acoustic environment, or soundscape, of a habitat can provide a wealth of information about its biological composition, the behavior of its inhabitants, and its overall ecological health [100].

The use of sound as a modality for monitoring biodiversity offers several distinct advantages over traditional survey methods. Firstly, passive acoustic monitoring (PAM) is inherently non-invasive, allowing researchers to collect data without disturbing or physically capturing animals, which is particularly crucial when studying sensitive, elusive, or endangered species [49,50]. Secondly, acoustic sensors can operate continuously (24/7) over extended periods, capturing temporal patterns in animal activity, including nocturnal or crepuscular behaviors that are often missed by visual surveys limited to daylight hours [51]. This continuous data stream enables the study of diel, seasonal, and long-term trends in species presence and vocal activity.

Furthermore, acoustic monitoring can detect species often missed by visual surveys, especially in complex habitats like dense forests or aquatic environments where visibility is limited, or for species that are small, camouflaged, or cryptic [55]. For instance, many insect species, including important pollinators like bees, produce characteristic sounds (e.g., flight buzzes, stridulation) that can be used for their detection and even identification, offering a pathway to monitor populations that are otherwise difficult to assess visually at scale [59,101]. Beyond simple presence/absence, acoustic data can also provide behavioral insights. Variations in call rates, song complexity, or the timing of vocalizations can indicate breeding activity, stress levels, or responses to environmental changes [102].

Despite its considerable potential, ecological acoustic monitoring is not without its challenges. The sheer volume of data generated by continuous, long-term deployments can be immense, creating significant hurdles for storage, transmission, and processing [75]. Extracting meaningful biological signals from recordings often requires sifting through substantial environmental noise (e.g., wind, rain, anthropogenic sounds), which can mask or degrade target sounds; a high microphone signal-to-noise ratio is therefore crucial for enhancing the acoustic sampling of wildlife [103]. Moreover, the audio quality of many existing solutions can be a significant limitation. The fidelity of the recording hardware varies widely, and while high-end recorders are ideal for capturing the nuanced acoustic details needed for robust analysis, they are not always accessible [104]. Perhaps the most significant challenge lies in accurate and efficient species identification from acoustic data. Manual analysis of recordings by expert listeners is time-consuming and not scalable for large datasets. This has spurred the development of automated methods, increasingly leveraging machine learning, to classify sounds and identify species [52,78]. However, developing robust and generalizable automated classifiers often requires large, high-quality annotated datasets and sophisticated algorithms capable of handling the variability inherent in biological sounds and complex acoustic environments [105].

This chapter presents work aimed at addressing some of these challenges, particularly in the context of developing field-ready hardware and leveraging on-device AI to move from raw acoustic data to actionable ecological insights more efficiently.

4.1.2 Chapter Overview

Building on the recognized value and inherent challenges of ecological acoustic monitoring, this chapter details a focused research effort to develop and apply novel acoustic sensing technologies for enhanced environmental and ecological insight. The work presented herein progresses through two key interconnected projects: SoundSHROOM and BuzzCam. These projects represent a systematic exploration, beginning with the development of robust, general-purpose acoustic data collection hardware suitable for demanding field conditions, and culminating in a highly specialized application of on-device machine learning for the conservation monitoring of endangered pollinators.

The first part of this chapter, Part A (Sections 4.2 - 4.6), introduces the SoundSHROOM project. This initiative was driven by the need for advanced acoustic recording systems capable of withstanding harsh environmental conditions, such as those found in the Arctic, while also providing multi-channel capabilities for potential spatial audio analysis. I will detail the motivation, design, and architecture of the SoundSHROOM units, providing details of their application in Svalbard, Norway. Key findings from this deployment, including an evaluation of microphone windshield performance and the creation of a unique multi-channel Arctic acoustic dataset [59], will be presented. Crucially, this section will also highlight the practical lessons learned regarding hardware robustness, power management in remote settings, and multi-channel data handling, which provided invaluable foundational experience for subsequent work.

The second part of this chapter, Part B (Sections 4.7 - 4.11), transitions to the BuzzCam project, a more targeted application focused on addressing the critical issue of pollinator decline. This work leverages the expertise gained from SoundSHROOM to develop a specialized acoustic and environmental sensing system specifically designed for monitoring bee populations, with a primary focus on the endangered Patagonian bumblebee, *Bombus dahlbomii*. I will describe the BuzzCam system design, its deployment in Patagonia, Argentina, and the meticulous process of creating a high-resolution, annotated dataset of native and invasive bee flight buzzes [56]. The central contribution detailed in this part is the development, rigorous optimization, and successful deployment of an efficient machine learning model for the real-time, on-device classification of bee presence and species. This involved adapting existing neural network architectures, employing quantization-aware training, and porting the model to a resource-constrained microcontroller (Analog Devices MAX78000). The performance of this on-device AI system, in terms of classification accuracy, inference latency, power consumption, and model size, will be thoroughly evaluated and discussed.

Finally, Section 4.12 will conclude the chapter, summarizing the key achievements of both the SoundSHROOM and BuzzCam projects and reiterating their collective contribution to advancing the tools and methodologies available for ecological acoustic monitoring. The progression from general-purpose robust hardware development to a specialized, AI-driven conservation tool illustrates a key theme of this dissertation: the iterative development and cross-contextual application of sensing technologies to address pressing environmental and ecological challenges.

Part A: SoundSHROOM - Multi-Channel Acoustic Data Collection in Arctic Environments

4.2 SoundSHROOM: Motivation and Objectives

The impetus for the SoundSHROOM (Sensor-equipped Habitat Recorders for Outdoor Omnidirectional Monitoring) project stemmed from a recognized need within the ecological research community for more advanced and resilient tools capable of operating in scientifically critical yet environmentally demanding locations. My development of this system was driven by several key motivations: the pressing requirement for robust sensing technologies that could reliably function in extreme environments like the Arctic; an interest in exploring the rich ecological information that could be gleaned from multi-channel audio data; and a desire to contribute a unique, high-quality dataset to the broader scientific community. This section elaborates on these motivating factors and outlines the specific objectives I set for the SoundSHROOM initiative, much of which formed the basis of our subsequent publication [59].

4.2.1 Need for Robust Sensing in Extreme Environments: Challenges of Arctic Deployments

The Earth's extreme environments, particularly polar regions such as the Arctic, serve as crucial indicators of global climate change and are undergoing rapid ecological transformations [106,107]. These regions, including areas like Svalbard, Norway, are home to unique ecosystems and species that are often highly specialized and acutely vulnerable to environmental shifts, such as rising temperatures, dimin ishing sea ice, and altered weather patterns [59,108]. Effective monitoring of biodiversity and ecological processes within these sensitive locales is therefore essential for understanding the multifaceted impacts of climate change, for developing informed conservation strategies, and for assessing the adaptive capacity of Arctic flora and fauna.

However, the logistical and technological challenges associated with conducting field research in such environments are substantial. As noted by Callaghan et al. [109] and Laidre et al. [110], factors such as extreme weather (including persistent low temperatures, strong winds, and various forms of precipitation), difficult and often remote terrain, limited accessibility, and constrained field seasons, impose severe limitations on traditional ecological survey methods and the equipment deployed. While Passive Acoustic Monitoring (PAM) offers a promising, non-invasive approach to ecological data collection, standard commercially available recording devices are often not engineered to withstand the prolonged, unattended operation required in harsh polar conditions. This specific challenge directly motivated my endeavor to engineer a sensor system with enhanced durability and operational reliability. As we highlighted in Chwalek et al. [59], "Bioacoustic monitoring of biodiversity in harsh and remote locations addresses several critical challenges in ecological research and conservation," underscoring my aim to develop a tool that could effectively collect valuable data despite the significant environmental adversities inherent in Arctic deployments.

4.2.2 Exploring Multi-Channel Audio: Potential for Spatial Audio Applications

Beyond the basic detection of species presence through their vocalizations, the spatial characteristics of sound represent a rich, yet often underutilized, dimension of information in ecological studies. The majority of conventional PAM systems, from low-cost open-source recorders like the AudioMoth to more expensive commercial units like the Song Meter line from Wildlife Acoustics, employ either single (monophonic) or dual (stereophonic) microphones, which provide limited or no capability for determining the location or directionality of sound sources [111,112]. Recognizing this limitation, I was motivated to explore the potential of multi-channel microphone arrays. Such arrays, as established in the field of acoustics [113], form the basis for advanced spatial audio processing techniques, including beamforming and sound source localization [61,62]. However, there are no commercial microphone arrays designed for long-term PAM use, especially in outdoor environments, which often leads researchers to create custom solutions for short-term deployments [114].

The application of these spatial audio techniques in ecological contexts holds significant promise. For instance, sound source localization could facilitate the differentiation of individual animals within a complex chorus, thereby improving population density estimates. It could also be used to track the movement of animals through their habitat or to study fine-scale spatial patterns in habitat utilization based on vocal activity [49]. Beamforming, by contrast, could significantly enhance the signal-to-noise ratio for faint or distant vocalizations by electronically "steering" the array's sensitivity towards the target sound and attenuating off-axis noise, effectively increasing the detection radius of the monitoring system. While the full implementation and field validation of sophisticated spatial audio algorithms were considered secondary objectives for the initial SoundSHROOM deployment, a core motivation for me was to engineer a hardware platform and collect a high-quality dataset that would specifically enable and encourage such future work. As stated in our publication [59], "These devices were designed to synchronously capture audio from multiple microphones for beamforming and localization applications and to serve as a testbed..."

4.2.3 Specific Objectives for the SoundSHROOM Project

The motivations for the SoundSHROOM project—the need for robust sensing in extreme environments and the potential of multi-channel audio—led to several specific objectives, many of which were subsequently reported in Chwalek et al. [59]. The first was to design and build a robust, multi-channel acoustic recording system capable of extended, unattended deployment in harsh conditions like the Arctic. A second objective was to deploy these units in Svalbard to collect a unique, spatially rich acoustic dataset and meticulously document the system's real-world performance. This deployment also enabled the systematic evaluation of various microphone windshield configurations to yield practical data on effective wind noise mitigation. Finally, the project aimed to comprehensively assess the overall deployment process, system durability, and operational reliability of the custom-developed hardware in an Arctic context.

The successful achievement of these objectives, would not only result in a valuable and novel dataset for the broader scientific community but also furnish critical experiential knowledge and technical insights that would directly inform my subsequent design and deployment of more specialized environmental and ecological sensing systems, most notably the BuzzCam project detailed later in this chapter.

4.3 SoundSHROOM: System Design and Architecture

The SoundSHROOM system represents a custom-engineered solution, which I designed and developed to address the specific challenges and opportunities identified for multi-channel acoustic data collection in demanding outdoor environments. My design philosophy centered on achieving synchronous audio capture from a versatile microphone array, ensuring physical robustness against harsh environmental elements, and integrating sufficient onboard processing and storage capabilities for sustained, unattended field deployments. The core system architecture and initial deployment outcomes were presented in our collaborative publication [59]. This section provides a detailed description of the key hardware components and the software architecture I implemented to manage data acquisition, elaborating on the engineering decisions and their underlying rationale.

4.3.1 Hardware Design

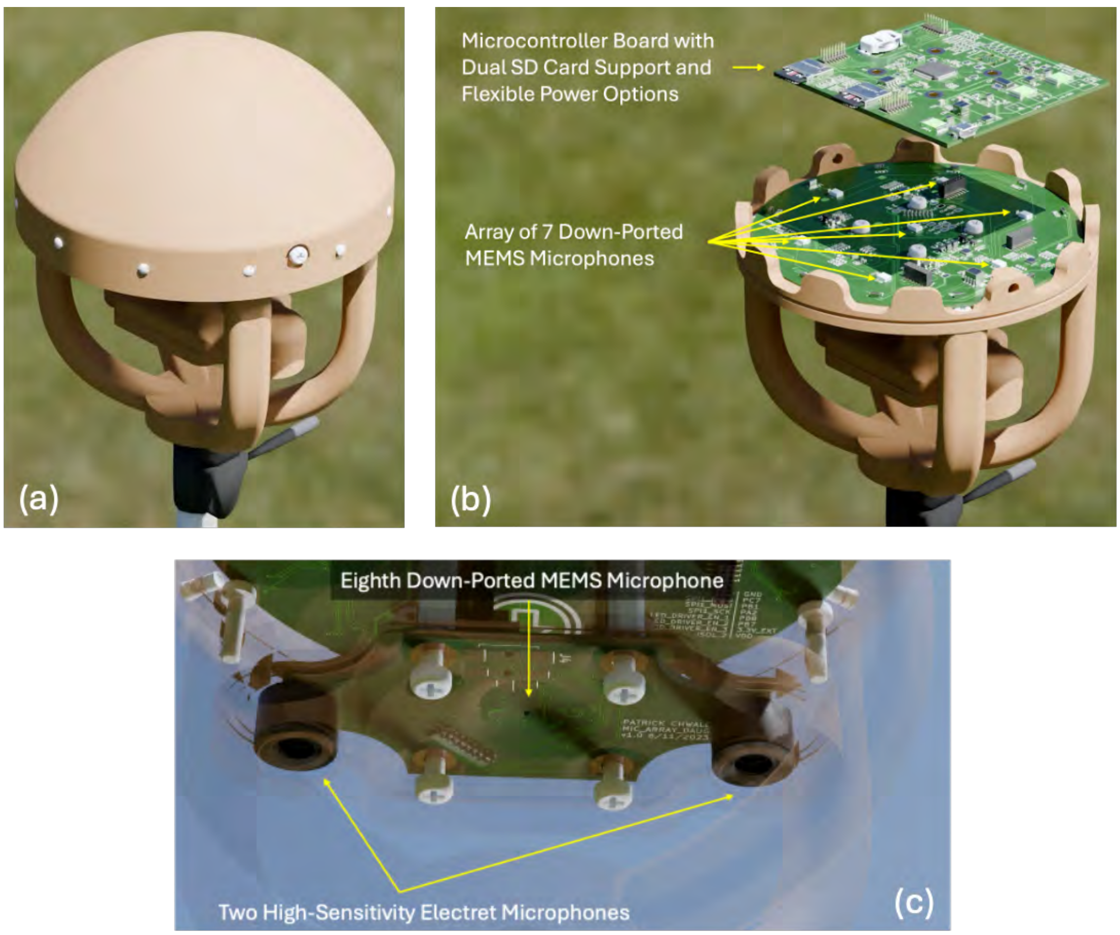


Figure 4.1: SoundSHROOM device. (a) Full device view. (b) Top view showing seven MEMS microphone locations and the main processing board. (c) Bottom view showing two electret microphones and one MEMS microphone.

The hardware implementation of each SoundSHROOM unit involved the careful selection and integration of a specialized microphone array, a powerful yet energy-efficient microcontroller, reliable data storage, a robust power delivery system, and a custom-designed protective enclosure (Figure 4.1).

Table 4.1: Microphone WAV Channels and Relative Positions on Unit[59]

WAV File Channel	Type	Name	Relative Location (mm)		
			x	y	z
1	MEMS	MK2	18.8	-32.5	0
2		MK1	37.5	0	0
3		MK4	-37.5	0	0
4		MK3	-18.8	-32.5	0
5		MK6	18.8	32.5	0
6		MK5	-18.8	32.5	0
7		MK8	0	0	-41.6
8		MK7	0	0	0
9	Electret	ELEC1	30	0	-41.6
10		ELEC2	-30	0	-41.6

Microphone Array: A Hybrid Approach for Versatility and Acoustic Quality

A cornerstone of the SoundSHROOM system is its ten-channel hybrid microphone configuration. While ambisonic microphone arrangements are often used for spatial audio as they are good for steering sound, they are not typically designed for sound localization. For this reason, I chose to use a custom omnidirectional array as it offered greater flexibility in post-processing and allowed for a more weather-resistant design. Specifically, this approach enabled the use of downward-facing omnidirectional microphones, which are inherently more protected from rain and surface moisture—a critical consideration for long-term outdoor deployments. The array was designed to offer both the spatial sampling necessary for advanced applications like beamforming and the high-fidelity stereo capture needed for general bioacoustic analysis. The spacing of the microphones (Table 4.1) allows for a theoretical frequency range of 2.3 kHz to 5.7 kHz. The array is composed of eight Vesper VM3011 MEMS microphones, chosen for their robustness against dust, shock, and weather [59]. These microphones were strategically placed to allow for comprehensive 3D spatial sampling. To complement the MEMS array and provide a high-quality audio reference, I integrated two PUI Audio AOM-5024P-HD-MB-R analog electret condenser microphones, selected for their excellent acoustic properties, including a high signal-to-noise ratio and a flat frequency response suitable for capturing nuanced avian vocalizations [59]. To demonstrate the array's localization capabilities, we used the Acoular Python package [115] to simulate three white noise sound sources. As shown in Figure 4.2, the simulation, using our device's microphone arrangement, was able to identify distinct beams in the direction of each sound source, as well as point clusters indicating their probable locations. These preliminary results, further discussed in [59], illustrate the inherent potential of the system for spatial audio analysis.

Signal Conditioning and Digital Conversion Architecture

The conversion of analog and PDM microphone signals into a synchronized digital data stream was a critical aspect of the electronic design. The PDM output signals from the eight MEMS microphones are directly interfaced with a Knowles ADAU7118 8-channel PDM-to-I²S/TDM

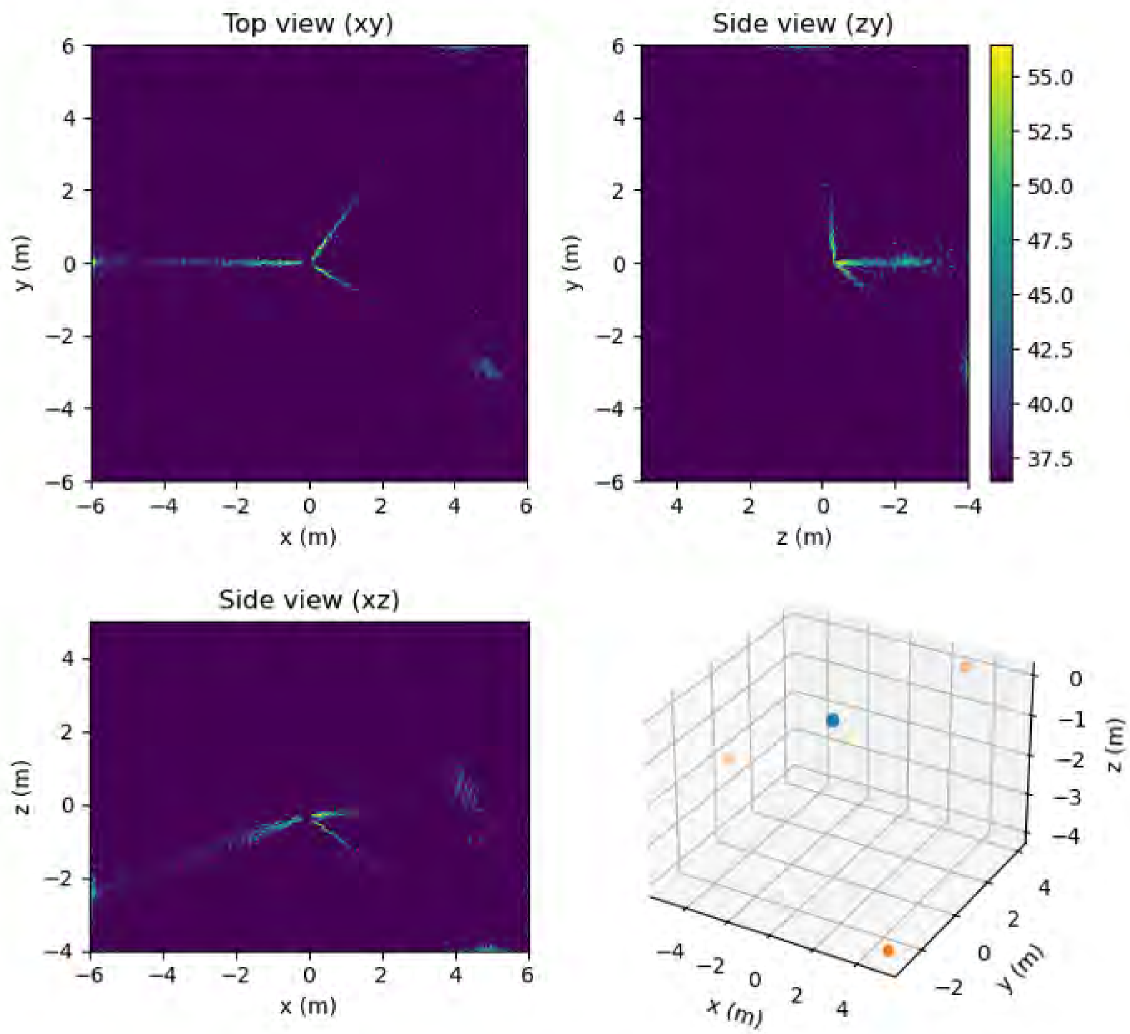


Figure 4.2: A simulated beamforming example using the Acoular [115] Python package with CLEAN-SC algorithm. Three sound sources were defined at locations (5, -3, 4), (3, 4, 0), and (-5, 0, 2). The SoundSHROOM was positioned at the origin with the 8 MEMS microphones simulated at their real locations [59].

converter with a built-in antialiasing filter, which efficiently aggregates the signals into a standard time-division multiplexed (TDM) digital audio stream [59]. For the two analog electret microphones, I designed a front-end that feeds into an Analog Devices ADAU1979 24-bit ADC. A crucial element of the architecture was configuring this ADC to output its digitized audio data onto the same TDM bus used by the MEMS converter. This shared bus topology ensures that all ten audio channels are precisely synchronized at the hardware level, which is essential for any subsequent inter-channel phase-dependent processing like beamforming [59].

Processing Unit, Data Storage, and Power Management

The core operational capabilities of SoundSHROOM are managed by a central processing unit, the STMicroelectronics STM22U575RIT6, an ARM Cortex-M33 based microcontroller. This MCU was selected for its comprehensive feature set, including multiple Serial Audio Interface (SAI) peripherals necessary for handling the 10-channel TDM audio stream, and its extensive low-power modes critical for maximizing battery life [59]. All acoustic data is recorded onto standard microSD cards, managed by firmware that includes an exFAT file system implementation. The units were engineered for autonomous operation, powered by a high-capacity rechargeable 10,000 mAh LiPo battery managed by a Microchip MCP73831T-2ACI/OT controller. For the Svalbard deployment, the system was powered by external, solar-assisted battery banks.

Enclosure and Environmental Protection

Given the intended deployment in the harsh Arctic environment, the physical enclosure was a critical design consideration. I designed a custom, robust enclosure fabricated using FDM 3D printing with PETG filament, selected for its mechanical strength and UV resistance (Figure 4.3). The two-part clamshell design incorporates an O-ring and silicone gasket to provide a weather-resistant seal. The microphone apertures were engineered to securely hold the microphone elements while also allowing for the easy attachment of various experimental windshields, which were a key part of the experimental setup to evaluate their effectiveness in mitigating wind noise (Figure 4.5).

4.3.2 Software and Data Acquisition Firmware

The custom embedded firmware I developed for the STM22U575RIT6 microcontroller is responsible for the entire data acquisition pipeline, including initializing all peripherals, managing the synchronous sampling of the ten audio channels, processing the data stream, and reliably writing the audio data to the microSD card.

A fundamental requirement addressed in the firmware architecture was the precise, synchronous sampling of all ten microphone channels. Leveraging the shared TDM bus, the firmware configures the MCU's Serial Audio Interface (SAI) peripheral to receive the interleaved audio data. A Direct Memory Access (DMA) scheme transfers blocks of incoming audio to RAM buffers with minimal CPU intervention, ensuring no data loss. For the Svalbard deployment, all ten channels were configured to be sampled at a rate of 32 kHz with 16-bit resolution, a deliberate choice to balance file size while still capturing the entire audible spectrum up to 16 kHz [59].

The firmware processes the buffered audio data and writes it sequentially to the microSD card using the industry-standard, uncompressed multi-channel WAV format. To enhance file system manageability, the recording is segmented into individual WAV files limited to a maximum size of 2 GB, which corresponds to approximately 52 minutes of audio per

Figure 4.3: SoundSHROOM pictured deployed in Svalbard [59]

segment. The firmware ensures these segments are named sequentially for straightforward concatenation during post-processing. To provide temporal context, an onboard Real-Time Clock (RTC) was manually set at the beginning of each deployment, and the firmware logs timestamps to allow correlation of acoustic events with world time.

4.4 SoundSHROOM: Arctic Deployment and Data Collection

Following the successful design, engineering, and bench-testing of the SoundSHROOM units, the next critical phase involved deploying these custom-developed systems in their intended target environment: the Arctic. This section details the fieldwork I planned and co-led in Svalbard, Norway, during the summer of 2023. It covers the selection of the study site, the deployment strategy for the SoundSHROOM units, and an overview of the multi-channel acoustic data that we successfully collected. The methodology and outcomes of this deployment formed a significant part of our publication [59].

4.4.1 Study Site: Longyearbyen, Svalbard – A Unique Arctic Setting

The archipelago of Svalbard, Norway, situated deep within the Arctic Circle, was chosen as the primary study site for the initial SoundSHROOM deployment. Specifically, our fieldwork, as detailed in Chwalek et al. [59], was centered around Longyearbyen, the main settlement, and its immediate surroundings. This location offers a unique combination of factors making it highly suitable for this research. Ecologically, Svalbard is a critical breeding ground for numerous Arctic bird species and supports a unique ecosystem that is highly sensitive to climatic variations, making it a vital location for monitoring biodiversity [108]. At the same time, the challenging environmental conditions of the Arctic summer, including the "midnight sun," strong winds, and diverse terrain, provided an ideal natural testbed for rigorously evaluating the robustness of the SoundSHROOM units. While remote, Longyearbyen offers a degree of logistical support that makes such a deployment feasible. Finally, the areas around the settlement encompass a range of habitat types, from tundra valleys and coastal estuaries to areas near human activity and glacial forelands (Figure 4.4, Table 4.2), which allowed us to capture a breadth of environmental sounds and test system performance under different conditions.

Table 4.2: Deployment Locations and Times in Central European Time (CET). Reference Figure 4.4 for more location details and Figure 4.5 for windshield details [59].

Deployment	Description	Lat	Lon	Start	End	Windshield Type
1	near stream, tundra	78.1914	15.8691	4th July, 18:00	5th July, 11:24	a, b
2	dog yard, tundra	78.2180	15.7116	5th July, 23:20	6th July, 20:50	b, c
3	dog yard, tundra	78.2184	15.7081	7th July, 16:30	9th July, 13:42	b, c
4	tundra valley flat	78.1736	16.0180	9th July, 20:00	11th July, 10:12	b, c
5	estuary edge, tundra	78.2131	15.7522	11th July, 13:15	12th July, 17:03	b, c, d
6	glacial foreland	78.1925	15.5496	12th July, 11:25	12th July, 15:13	d

Figure 4.4: Deployment locations in Svalbard as seen in Google Earth [59,116].

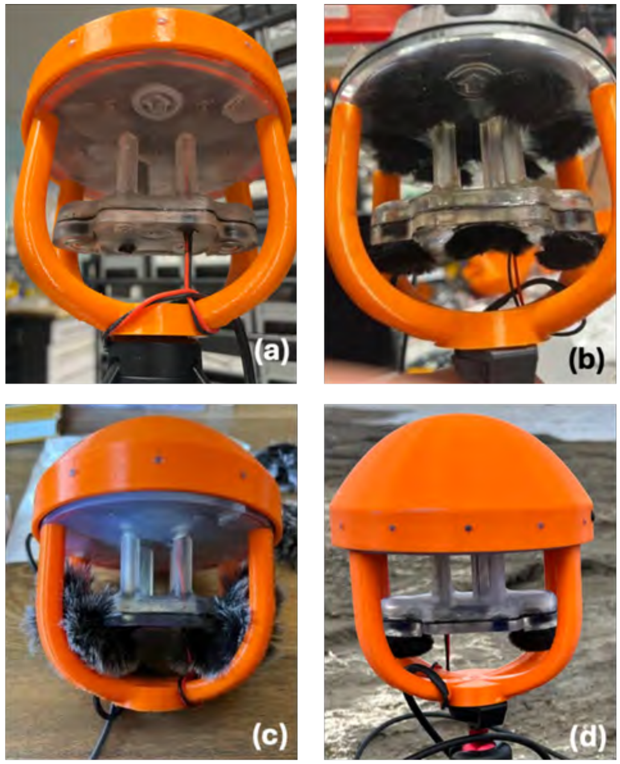


Figure 4.5: SoundSHROOM was tested in four windshield configurations: (a) no windshield, (b) small windshield on all microphones, (c) two large windshields on just the two electret microphones, and (d) two small windshield on just the two electret microphones [59].

4.4.2 Deployment Strategy and Execution

The deployment strategy aimed to achieve several goals simultaneously: test the SoundSHROOM hardware under different environmental stresses, evaluate the performance of various microphone windshield configurations, and collect a diverse multi-channel acoustic dataset. As reported in Chwalek et al. [59], we deployed three SoundSHROOM units across nine distinct locations over a period of six days in July 2023. Deployment locations were carefully chosen based on observed biodiversity, proximity to different environmental features (streams, tundra flats, estuary edges, glacial areas), and varying levels of proximity to anthropogenic sound sources to ensure a diverse range of acoustic inputs. For example, some deployments were near streams or active dog yards, while others were in more remote tundra or glacial foreland areas.

A key experimental aspect of the deployment was the testing of different microphone windshield configurations to assess their impact on wind noise reduction. The SoundSHROOM units were configured in four distinct setups across the deployments: (a) no windshield, (b) small individual windshields on all ten microphones, (c) two large windshields covering only the two electret microphones, and (d) two small windshields on only the two electret microphones (Figure 4.5). This systematic variation allowed for a comparative analysis of windshield efficacy.

Logistically, each SoundSHROOM unit was typically mounted on a staked tripod to elevate it slightly and provide stability. I ensured that each unit was initialized with a fully charged external solar-assisted battery (25Ah, 3.7V) and a formatted 512 GB microSD card. The internal Real-Time Clock (RTC) of each unit was synchronized to local time (Central European Time, CET) at the start of each deployment session to enable temporal alignment of the collected data. The units were then left to record continuously until battery depletion or manual retrieval by our team, with some sessions involving a reboot to digitally mark a visit or to replace the battery and SD card.

4.4.3 Data Collected: A Multi-Channel Arctic Soundscape Archive

The fieldwork in Svalbard resulted in a substantial and unique acoustic dataset. As reported in Chwalek et al. [59], we successfully collected approximately 400 GB of uncompressed 10-channel audio over the 6-day collection. The recordings captured a rich tapestry of the Arctic summer soundscape, including prominent avian vocalizations from species such as Barnacle Geese, Dunlins, and Arctic Terns (Figures 4.6 and 4.7), environmental geophony like wind and glacial melt, and limited anthrophony from nearby human activity.

Crucially, all recordings consist of ten synchronized audio channels, sampled at 32 kHz with 16-bit resolution. This multi-channel nature is a distinctive feature of the dataset, providing the raw material for future explorations into spatial audio analysis of Arctic soundscapes. For each deployment, associated metadata such as location, start and end times, unit ID, and windshield configuration were meticulously logged, providing vital context for the interpretation and analysis of the acoustic data (Table 4.2 and start_end_times.xlsx in [59]).

The successful execution of this deployment phase and the acquisition of this rich dataset demonstrated the field-readiness of the SoundSHROOM system and provided the empirical data necessary for the subsequent analyses of system performance and dataset characterization presented in Chwalek et al. [59] and discussed in the following section of this chapter.

4.5 SoundSHROOM: Key Findings and Dataset Contribution

The deployment of the SoundSHROOM system in the Arctic yielded significant findings regarding the system's performance in extreme conditions, the efficacy of different wind noise mitigation strategies, and culminated in the creation of a valuable, publicly available multi-channel acoustic dataset. These outcomes, which we presented and analyzed in detail in Chwalek et al. [59], are summarized in this section. This work also provided critical lessons that directly informed my subsequent development of more specialized ecological sensors.

4.5.1 System Performance and Robustness in Arctic Conditions

A primary objective of the SoundSHROOM project was to assess the operational viability and robustness of the custom hardware I designed when subjected to the rigors of an Arctic field deployment. Overall, the SoundSHROOM units demonstrated satisfactory performance and durability throughout the six-day deployment period in Svalbard. As reported in Chwalek

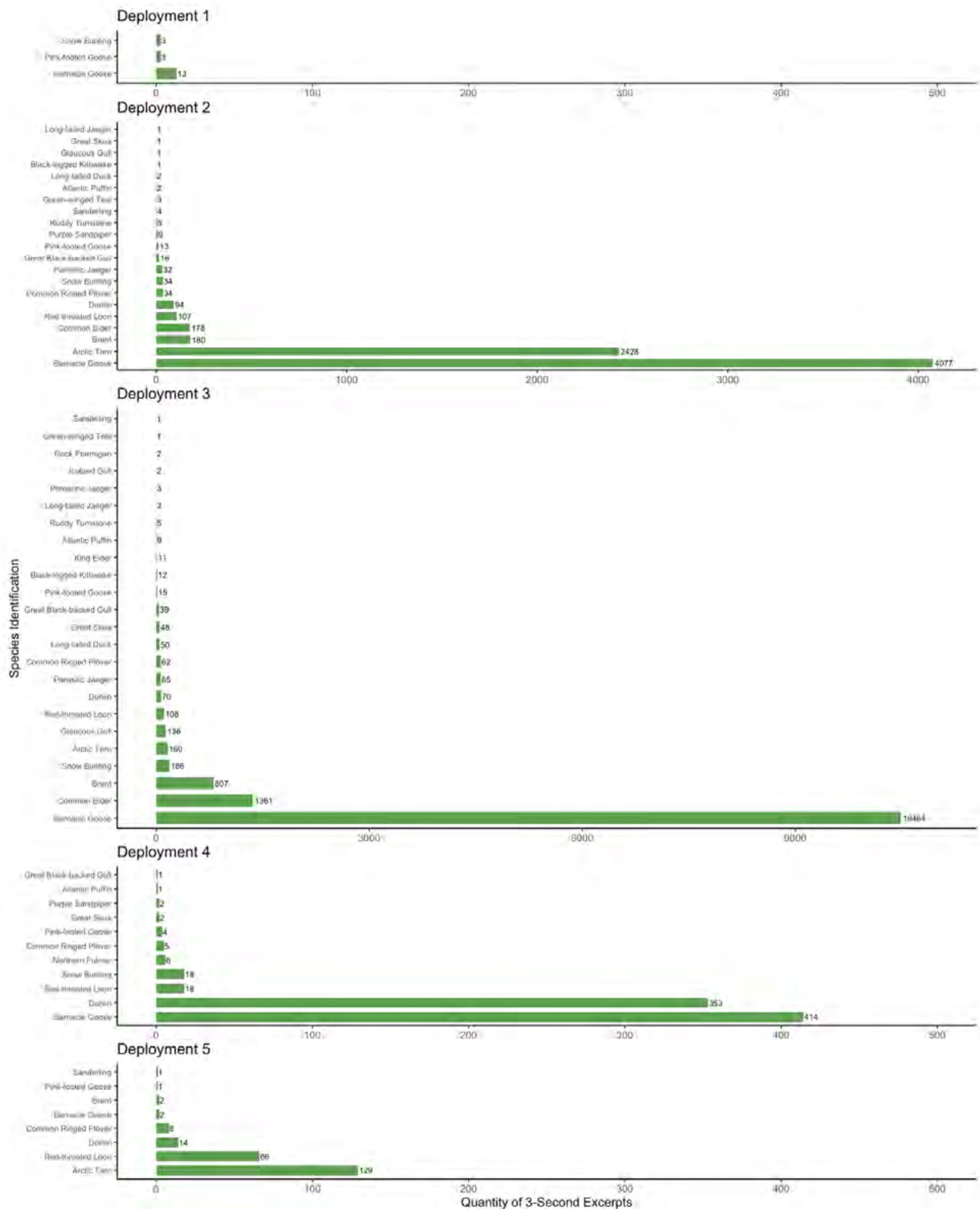


Figure 4.6: Frequency of bird calls identified by species in each deployment (n=5). Each data point represents one 3-second excerpt taken with a 1.5-second moving window. Deployment 6 is excluded due to a limited sample set [59].

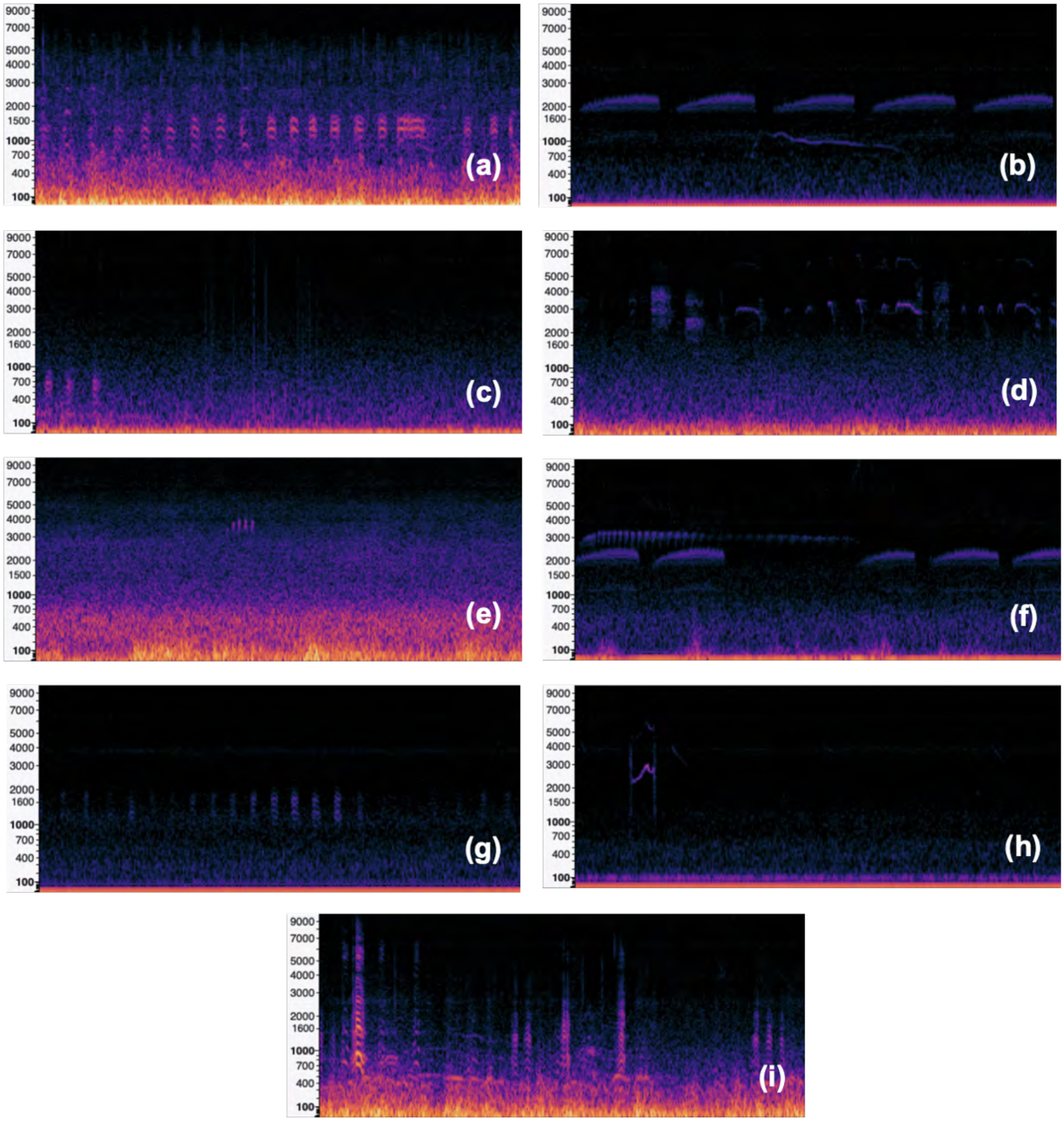


Figure 4.7: Spectrograms corresponding to BirdNET Analyzer classifications and deployment. (a) Red-throated Loon, dep. 5, (b) Barnacle Goose, dep. 4, (c) Common Eider, dep. 3, (d) Arctic Tern, dep. 2, (e) Snow Bunting, dep. 2, (f) Dunlin, dep. 4, (g) Northern Fulmar, dep. 4, (h) Common Ringed Plover, dep. 2, (i) Brant, dep. 3. [59].

et al. [59], the units successfully recorded continuous multi-channel audio for extended periods, typically limited by battery capacity rather than hardware failure. The systems were also exposed to occasional rainfall and continued to operate without failure, affirming the soundness of the mechanical enclosure and electronic component choices.

The power system I integrated generally performed as anticipated, enabling multi-day unattended recordings. The batteries rarely fully depleted, thanks to frequent sun exposure during the Arctic summer and near-daily check-ins. The data storage and file management system in the firmware also proved reliable, with consistent recording of 10-channel WAV files to the microSD cards. The segmentation of recordings into 2GB files facilitated easier data handling and backup in the field.

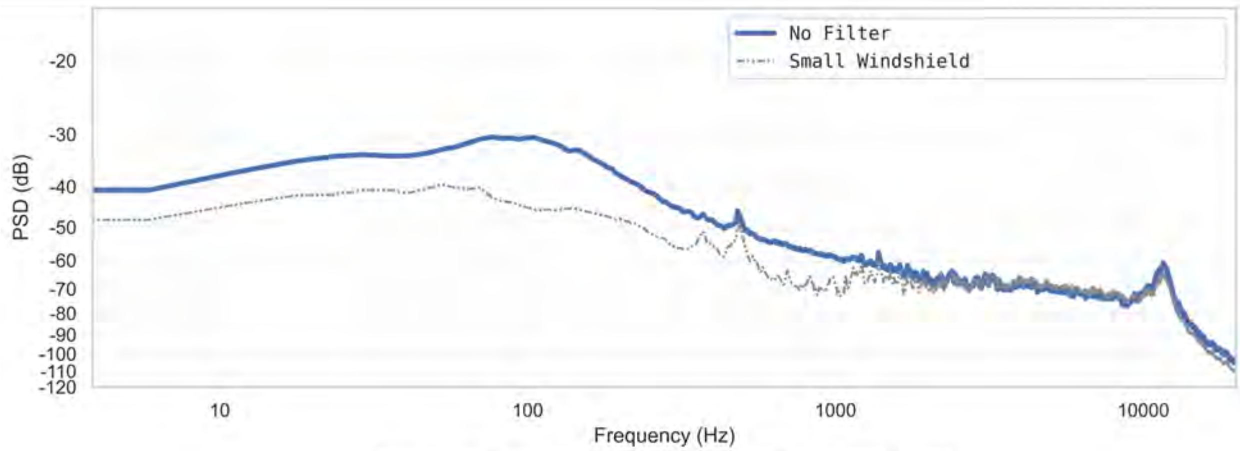
4.5.2 Evaluation of Microphone Windshield Performance

A significant experimental component of the Svalbard deployment was the systematic evaluation of different microphone windshield configurations. Wind-induced noise is a major contaminant in outdoor acoustic recordings, and understanding effective mitigation strategies is crucial for obtaining high-quality data. As described in Chwalek et al. [59], our methodology involved deploying units with four different windshield setups: no windshield, small individual windshields on all microphones, large windshields covering only the electret stereo pair, and small windshields on only the electrets (Figure 4.5).

Our analysis involved both quantitative and qualitative approaches. For the quantitative analysis, we calculated Power Spectral Density (PSD) plots from recordings made in a controlled wind tunnel environment (Figures 4.8 and 4.9). These plots clearly demonstrated that the small windshields provided significant noise reduction in the low-frequency range most affected by wind. For the MEMS microphones, the reduction was as much as -14.3 dB at 100 Hz, while for the electret microphones, the reduction was up to -13.4 dB at 300 Hz. In both cases, the most significant attenuation occurred below 1 kHz, with negligible differences at higher frequencies [59]. For the qualitative analysis, we utilized an Audio Spectrogram Transformer (AST) model to classify field recordings for the probability of "wind noise (microphone)". This analysis confirmed that any microphone with a windshield had less perceivable wind noise than those without, that larger windshields outperformed smaller ones, and that they were more effective on the electret channels [59]. These findings provide empirical evidence supporting the necessity of appropriate windshielding for outdoor acoustic monitoring and offer insights into the relative performance of different windshield types. This directly informed my subsequent design choices for the BuzzCam system.

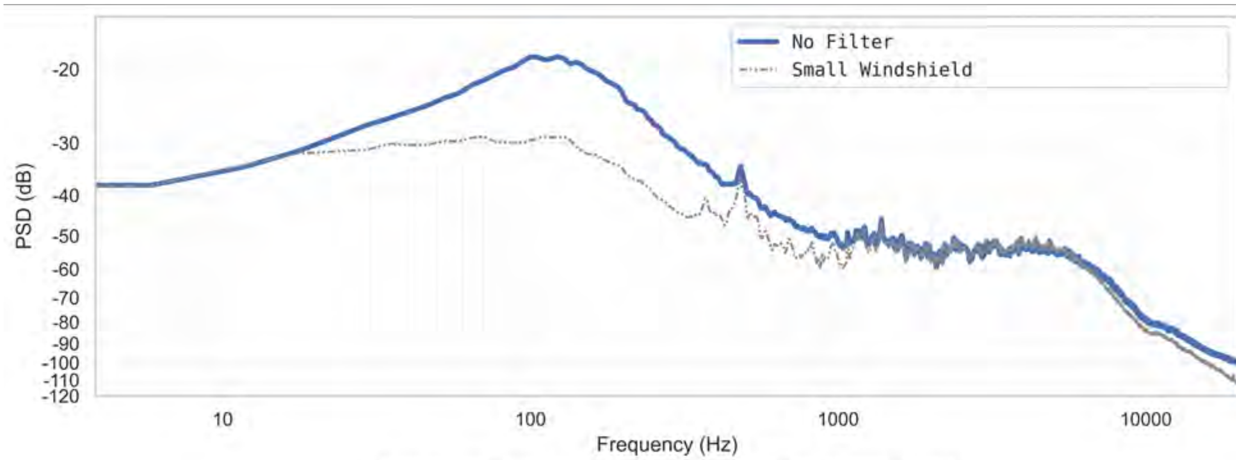
4.5.3 Data Collected: A Multi-Channel Arctic Soundscape Archive

A primary tangible output of this project is the "Acoustic data collection in arctic environments during the midnight sun using multi-channel SoundSHROOMs" dataset, which we curated and made publicly available via Figshare, as documented in Chwalek et al. [59]. The dataset comprises approximately 400 GB of 10-channel, 32kHz/16-bit synchronized audio recordings from nine distinct locations around Longyearbyen, Svalbard. It includes detailed metadata regarding deployment locations, times, unit configurations (including windshield type), and approximate start/end times for each recording file (Table 4.2 and start_end_times.xlsx in [59]). The data also includes the results of an analysis identifying various avian species present in the recordings (Figure 4.6). To generate these species identifications, we utilized the BirdNET Python library (version 2.4-V2), which can identify and classify around 6,000 bird sounds [52]. With our raw audio as input, the system segmented the recordings



	100	300	500	750	1000	2500	5000	7500	10000	20000
Small Windshield	-14.3	-11.0	-3.4	-12.6	-9.8	-1.3	0.7	0.5	-0.5	-5.3

Figure 4.8: The Power Spectral Density (PSD) plot for a 10 m/s airflow across MEMS microphones (VM3011) equipped with small windshields is shown. The table displays the difference in decibels between using no filter and using a windshield [59].



	100	300	500	750	1000	2500	5000	7500	10000	20000
Small Windshield	-10.9	-13.4	-2.9	-7.9	-5.1	0.1	1.2	-3.3	-3.9	-11.8

Figure 4.9: The Power Spectral Density (PSD) plot for a 10 m/s airflow across electret microphones (AOM-5024P-HD-MB-R) equipped with small windshields is shown. The table displays the difference in decibels between using no filter and using a windshield [59].

into overlapping 3-second excerpts with a 1.5-second stride, ensuring no significant audio information was missed. We provided the microphone locations, week of the year, and a minimum confidence level as parameters to the analyzer, and each segment was labeled with the species found to be present with a confidence level over 50%.

This dataset is significant for several reasons. It provides a unique acoustic snapshot of the Arctic summer soundscape during the midnight sun period and serves as a testbed

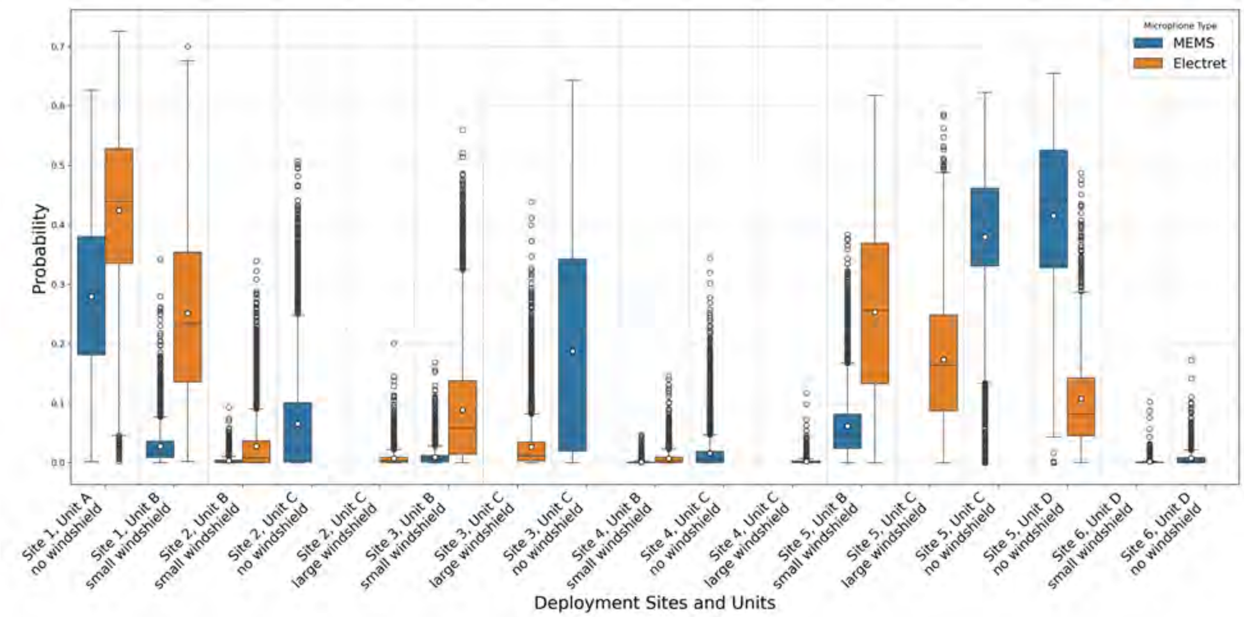


Figure 4.10: Probability results for "wind noise (microphone)" class when the Audio Spectrogram Transformer is applied on 10-second excerpts from one of the MEMS microphones (MK8, channel 7) and electret microphones (ELEC1, channel 9) on SoundSHROOM. Higher probabilities mean that wind noise is more likely to be observed in the acoustic data for that site and unit configuration [59].

for developing and evaluating species detection algorithms for Arctic avian fauna. The multi-channel synchronized nature of the data offers a rare resource for researchers interested in spatial audio analysis, such as sound source localization or beamforming of wildlife vocalizations. Furthermore, the associated metadata and documented deployment conditions make it valuable for studies on the impact of environmental factors on soundscapes and sensor performance. By making this dataset openly accessible, we aim to foster further research in Arctic bioacoustics, ecoacoustics, and sensor technology development.

4.5.4 Key Lessons Learned for Future Acoustic Deployments and Sensor Design

Beyond the specific findings and the dataset itself, the SoundSHROOM project yielded invaluable experiential knowledge that significantly informed my subsequent research and development efforts, particularly for the BuzzCam system:

- **Hardware Robustness is Paramount:** The Arctic deployment reinforced the critical need for meticulous attention to enclosure design, component selection, and weatherproofing for any sensor system intended for long-term, unattended outdoor use.
- **Wind Noise Mitigation is Essential:** The quantitative and qualitative data on windshield performance underscored that effective wind noise reduction is not an

afterthought but a primary design consideration for acquiring high-quality acoustic data in exposed environments.

- **Data Management for Large Datasets:** The experience of handling and organizing a large multi-channel audio dataset highlighted the need for robust metadata logging and clear file organization protocols from the outset of any field campaign.
- **Value of Iterative Design and Field Testing:** The SoundSHROOM deployment served as a crucial real-world test of the system, revealing both its strengths and areas for potential minor improvements, embodying an iterative design philosophy.

These lessons provided a strong foundation of practical engineering knowledge and field experience that I carried directly into the design and execution of the BuzzCam project, which aimed to apply acoustic sensing to a very different ecological challenge.

4.6 SoundSHROOM: Summary of Outcomes and Foundational Learnings

The SoundSHROOM project successfully validated a novel, robust multi-channel acoustic recorder in the harsh Arctic environment of Svalbard. Key outcomes included a unique, publicly available Arctic soundscape dataset and empirical data on mitigating wind noise in outdoor recordings. The practical experience gained in low-power hardware design, multi-channel data acquisition, and sensor protection proved foundational. These lessons were directly applied to the subsequent development of the BuzzCam system, a specialized sensor aimed at addressing the pressing ecological challenge of pollinator monitoring using on-device machine learning.

Part B: BuzzCam - On-Device Machine Learning for Endangered Bee Monitoring

4.7 BuzzCam: Motivation and Specific Objectives

The development of the BuzzCam system and its associated on-device machine learning capabilities was driven by an urgent ecological crisis: the well-documented global decline of pollinator populations, and the specific conservation crisis facing the endemic Patagonian bumblebee, *Bombus dahlbomii* (Figure 4.11). This section outlines the motivations stemming from this crisis, the potential of passive acoustic monitoring (PAM) as a novel solution, the critical need to overcome data processing bottlenecks through edge artificial intelligence (AI), and the specific objectives for the BuzzCam project.

Figure 4.11: *Bombus terrestris* (left) and *Bombus dahlbomii* (right)

4.7.1 The Pollinator Crisis and *Bombus dahlbomii*: An Urgent Call for Innovative Monitoring

Pollinators, particularly bumblebees (*Bombus* spp.), are indispensable for maintaining terrestrial ecosystem health and ensuring global food security [117,118]. However, as highlighted in both my prior work [56] and the broader literature [119,120], many bumblebee populations worldwide face unprecedented declines. These declines are attributed to a confluence of anthropogenic stressors, including habitat degradation, climate change, pesticide use, and the pervasive threat of invasive species and associated pathogen spillover [121].

The endemic Patagonian bumblebee, *Bombus dahlbomii*, serves as a stark example of this crisis. As one of the world's largest bumblebees and a crucial pollinator for a diverse array of native Patagonian flora [56,122], its populations have suffered a dramatic decline. This decline is largely attributed to the introduction and rapid spread of invasive congeners like Bombus terrestris [123]. The International Union for Conservation of Nature (IUCN) has classified *B. dahlbomii* as endangered, and its diminishing numbers have cascading negative impacts on the native plant communities it serves and the broader ecosystem stability of Patagonia [56,124].

Effective conservation of *B. dahlbomii* and mitigation of the impacts from invasive species necessitate robust, long-term monitoring programs. However, traditional methods for monitoring Bombus species present significant limitations. Visual transect surveys or lethal trapping (e.g., pan traps) are often labor-intensive, limited in spatio-temporal coverage, potentially disruptive to bee behavior, or ethically problematic for endangered species [125,126]. For instance, transect walks are laborious and depend heavily on observer expertise and visibility, which can be poor in dense vegetation [56,126]. Netting and detailed identification require significant expertise, and pan traps, while low-effort, result in bee mortality – an unacceptable outcome for conservation-focused studies of endangered species [47,56]. These limitations underscored the urgent need for innovative, non-invasive, and scalable monitoring solutions.

4.7.2 The Potential of Passive Acoustic Monitoring (PAM) for Bees: Harnessing Buzz Signatures

Passive Acoustic Monitoring (PAM) has emerged as a promising non-invasive alternative for ecological monitoring [127]. For bees, PAM offers the potential to leverage their characteristic flight buzzes for detection and even species differentiation [128,129]. PAM offers the potential for continuous, scalable data collection on bee activity and diversity [56,130,131]. The unique acoustic signatures produced by the wingbeats of different bee species, influenced by factors like body size and wing morphology, can serve as a non-invasive proxy for their presence and activity [129,132]. Our hypothesis was that these distinct buzz signatures, if accurately captured and analyzed, could form the basis of an effective monitoring tool for Bombus species, including the differentiation of *B. dahlbomii* from its invasive congener, *B. terrestris*. This acoustic approach would overcome many of the ethical and logistical drawbacks of traditional methods.

4.7.3 Addressing the "Data-to-Insight" Bottleneck with On-Device Machine Learning: Towards Real-Time, Scalable Monitoring

While PAM offers a powerful data collection method, the sheer volume of data generated by PAM necessitates efficient and automated analysis techniques. Manually listening to and analyzing countless hours of audio recordings is impractical for large-scale or long-term monitoring. Machine learning (ML), particularly deep learning models like Convolutional Neural Networks (CNNs), has shown considerable success in bioacoustic classification tasks [52,78]. Deploying these computationally intensive models in remote field settings, however, presents significant challenges with power scarcity, unreliable connectivity, and the need for real-time feedback.

My research confronts this bottleneck through edge computing—specifically TinyML—which leverages low-power microcontrollers (MCUs) with AI accelerators [82]. On-device data processing and inference minimizes data transmission needs, reduces latency, and enhances the autonomy and sustainability of long-term monitoring deployments. The core idea was to shift the analytical burden from centralized servers or powerful computers to the sensor itself, enabling near real-time ecological insights directly in the field. This would transform PAM from a purely data-collection paradigm into an active ecological insights engine.

4.7.4 Specific Objectives for the BuzzCam Project

Based on these motivations, I defined the following specific objectives for the BuzzCam project:

1. To design and deploy a specialized acoustic and environmental sensing system (BuzzCam) tailored for monitoring *Bombus* bee populations in their natural habitat in Patagonia. This involved engineering a device capable of capturing high-quality bee

buzz acoustics alongside relevant environmental parameters (e.g., temperature, humidity) that might influence bee activity, building upon the hardware experience gained from the SoundSHROOM project.

2. To collect and annotate a comprehensive, high-resolution acoustic and environmental dataset from Puerto Blest, Argentina—a key habitat for *B. dahlbomii*. This dataset (now published as Chwalek et al. [56]) was specifically designed to "capture the bioacoustics of native and invasive Bombus species amidst ongoing habitat pressures" and serve as foundational data for training and validating machine learning models.

3. To adapt, systematically prune, and optimize an established bioacoustic CNN architecture (ANIMAL-SPOT; Bergler et al. [78]) for deployment on a resource-constrained, low-power microcontroller (Analog Devices MAX78000). This objective focused on meeting the stringent memory (e.g., 442 KB for weights) and computational constraints of the target MCU while preserving classification accuracy.

4. To investigate and implement optimal data preprocessing techniques and quantization-aware training (QAT) strategies to maximize classification accuracy for differentiating between *B. dahlbomii*, *B. terrestris*, and background noise, while ensuring model compatibility with the 8-bit precision of the hardware accelerator on the MAX78000.

5. To evaluate the performance of the deployed on-device model in terms of classification accuracy, inference latency, and power consumption on the integrated BuzzCam-MAX78000 platform, thereby demonstrating the feasibility of this edge AI approach for advancing autonomous, large-scale pollinator monitoring.

The successful achievement of these objectives would pave the way for cost-effective, scalable, and autonomous sensor networks capable of providing near real-time ecological insights, significantly enhancing our ability to understand pollinator dynamics and inform timely conservation strategies for endangered species like *B. dahlbomii*.

4.7.5 Hardware Design: Tailored for Bee Bioacoustics and Environmental Context

The BuzzCam hardware (Figure 4.12) comprises several key subsystems that I engineered for this specific application. The acoustic sensing subsystem was optimized for bee buzzes, featuring two high-SNR electret microphones (AOM-5024L-HD-R) spaced 144 millimeters apart. My choice of these microphones was based on their appropriate frequency response for capturing the characteristic low-frequency fundamental tones and higher harmonics of Bombus flight buzzes (typically below 200 Hz fundamental, with harmonics up to 5 kHz) [56].

A significant challenge in outdoor bee acoustic monitoring is wind noise, which can easily mask subtle buzz sounds. Based on the findings from the SoundSHROOM deployment (Section 4.5.2), a systematic wind tunnel test was conducted to refine the windshield design specifically for bee bioacoustics. We tested a variety of fabrics and covers over the electret diaphragm in a 10m/s wind—a moderately strong natural gust—with pink noise playing through the tunnel.

As shown in Figure 4.13, while many materials performed similarly at higher frequencies where they generally converged, the best-performing architecture for mitigating low-frequency wind noise was a cylindrical design covered with SAATI B260HY fabric. This systematic approach to windshield selection was directly informed by the quantitative and qualitative wind noise evaluations from the SoundSHROOM Arctic deployment, which underscored the critical importance of mitigating low-frequency noise for acquiring high-quality bioacoustic data. This design was compared against the small "deadcat" windshield from the SoundSHROOM project (Figure 4.14). The results showed that the smooth SAATI fabric on the cylinder provided comparable wind noise reduction to the deadcat design, with the added practical advantages of being less prone to fouling from debris and drying much quicker after wet weather conditions.

For the original dataset collection, each microphone signal path included a 20 dB preamplifier and was synchronously sampled with a 24-bit ADC, with the signal then digitally amplified by 15 dB before recording the 16 most significant bits (MSB) at a 48 kHz sampling rate. The updated BuzzCam platform features a redesigned, lower-power ADC (TLV320ADC3101) that provides a direct I^2S stream accessible by both the application processor and the on-device AI microcontroller.

To provide contextual data, I integrated a Bosch BME688 environmental sensor, configured using Bosch's Sensortec Environmental Cluster (BSEC) API, to log raw temperature, pressure, humidity, and equivalent CO_2, allowing for the study of correlations between bee acoustic activity and local weather conditions [56]. The processing and power systems were designed for both initial data collection and long-term on-device AI deployment. The original system used an STM32WB5MMGH6 microcontroller for data acquisition and storage to a microSD card. The updated version integrates an Analog Devices MAX78000 microcontroller, chosen for its energy-efficient CNN hardware accelerator, making it ideal for running the pruned bee classification model directly on the device. Both versions utilize microSD cards for local data archival.

The BuzzCam units are designed for long-term, battery-powered field deployment, typically using a single 3.7V, 3500mAh 18650 lithium-ion battery, which allows for over 86 hours of continuous recording. For extended deployments, a custom solar panel assembly was developed, consisting of five SM141K09L panels (277 mW each). Each panel is protected by a Schottky diode and wired in parallel to maximize harvesting efficiency, even in partial shading conditions. The diodes prevent a shaded cell from being backfed by the current from the other cells, which can cause heating or minor damage. This configuration can deliver up to 1.385 W at peak sunlight. With this design, continuous 24/7 operation is feasible in locations like Puerto Blest, where typical solar conditions can sustain the system's ≈ 150 mW load. In many ecological monitoring applications where 24/7 operation is unnecessary, recording can be scheduled for periods of peak species activity, substantially reducing the daily power budget. To facilitate accurate data labeling during the dataset collection phase, a Bluetooth Low Energy (BLE) radio was included, allowing our field team to connect to each device with a custom iOS application to annotate bee sightings in real-time, a significant improvement over traditional asynchronous methods.

Figure 4.12: BuzzCam with sensor locations shown. The system can be mounted to a standard tripod so the height is adjustable [56].

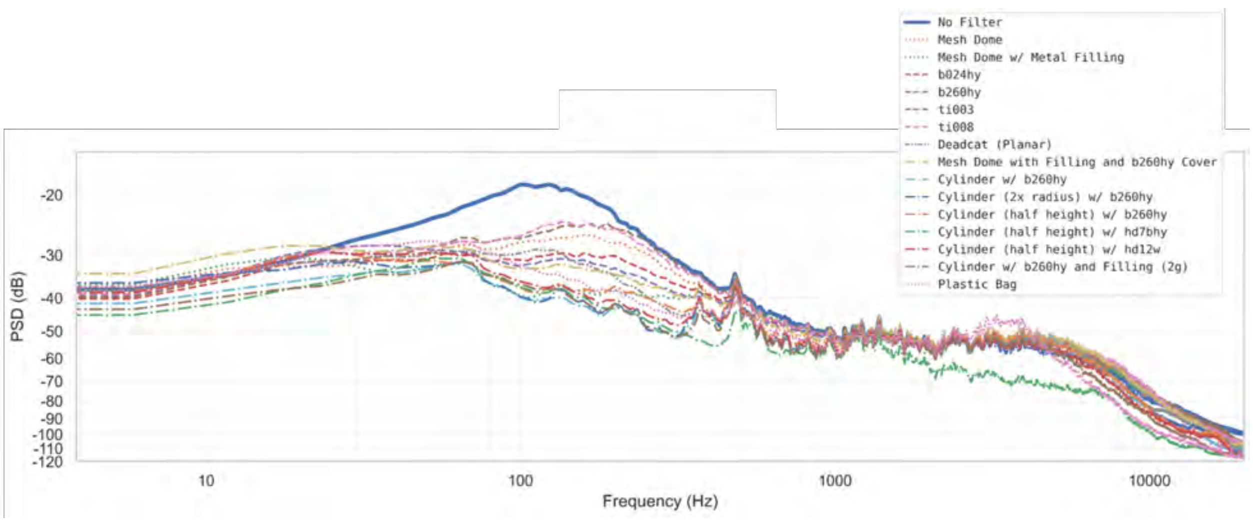
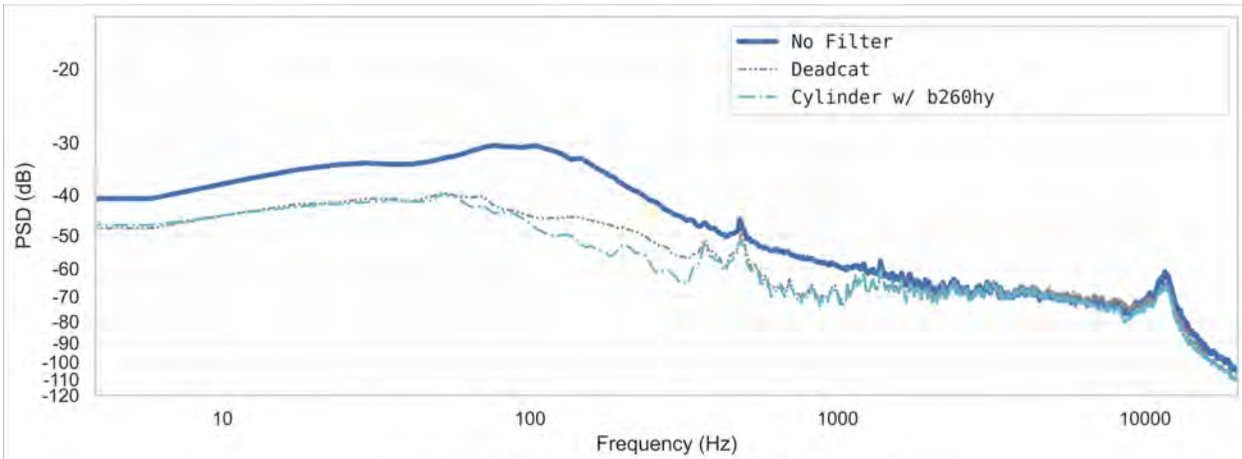


Figure 4.13: The Power Spectral Density (PSD) plot for a 10 m/s airflow across an electret microphone equipped with a variety of windshields is shown. The tested configurations include an aluminum mesh dome with and without metal fillings, different types of SAATI acoustic-grade fabrics, a deadcat similar to the one used in the SoundSHROOM project, and various cylindrical designs with varying dimensions to measure the impact on wind noise mitigation. A plastic bag was also tested, as this method is anecdotally used to weatherize inexpensive microphones in the field.

4.7.6 Comparing BuzzCam Acoustic Recording Performance with AudioMoth

Figure 4.15 shows spectrograms of an audio segment recorded simultaneously in Puerto Blest, where a BuzzCam and an AudioMoth were positioned next to each other. For this comparison, the recordings were scaled so that their noise floors (during periods with no animal, natural, or anthropogenic sounds) have an equal RMS. The segment includes a bee



	100	300	500	750	1000	2500	5000	7500	10000	20000
Deadcat	-14.3	-11.0	-3.4	-12.6	-9.8	-1.3	0.7	0.5	-0.5	-5.3
Cylinder w/ b260hy	-16.2	-18.3	-4.5	-14.1	-10.4	-1.1	0.8	-2.2	-3.1	-3.9

Figure 4.14: The Power Spectral Density (PSD) plot for a 10 m/s airflow across an electret microphone equipped with the deadcat-type windshield used in SoundSHROOM and with the custom cylindrical design used in BuzzCam. The accompanying table quantifies the wind noise reduction in decibels at various frequencies, referenced from the baseline condition of having no filter.

buzz in the first half and a bird sound in the second. A qualitative visual inspection reveals significant differences in recording quality between the BuzzCam and the AudioMoth, a low-cost system commonly used for bioacoustic recordings due to its affordability. The BuzzCam recording exhibits a much lower noise floor, visible as a darker background, particularly in the higher frequencies. In contrast, the Audiomoth spectrogram shows persistent, broadband high-frequency electrical noise. While the RMS levels for the bird call are similar between the two microphones, the BuzzCam recording maintains energy concentrated in the expected harmonic bands, resulting in a cleaner, crisper capture with more defined harmonics.

To quantify these differences, a further analysis was performed on the bee buzz portion of the recordings. A bandpass filter (200 Hz to 4 kHz with a 12 dB rolloff) was applied to both clips to isolate the buzz and its prominent harmonics. Within this band, the RMS level of the background noise was -40.3 dB for the Audiomoth and -37.8 dB for the BuzzCam. For the bee buzz segment itself, the RMS was -29.9 dB for the Audiomoth and -24.0 dB for the BuzzCam. These measurements yield a Signal-to-Noise Ratio (SNR) of 10.4 dB for the AudioMoth and 13.8 dB for the BuzzCam. The 3.4 dB higher SNR of the BuzzCam demonstrates its superior ability to capture a stronger, more prominent signal relative to background noise. Overall, while RMS alone may be comparable for some signals, the combination of spectral characteristics and a higher SNR reveals that the BuzzCam provides superior signal fidelity, which is critical for robust bioacoustic analysis and machine learning applications.

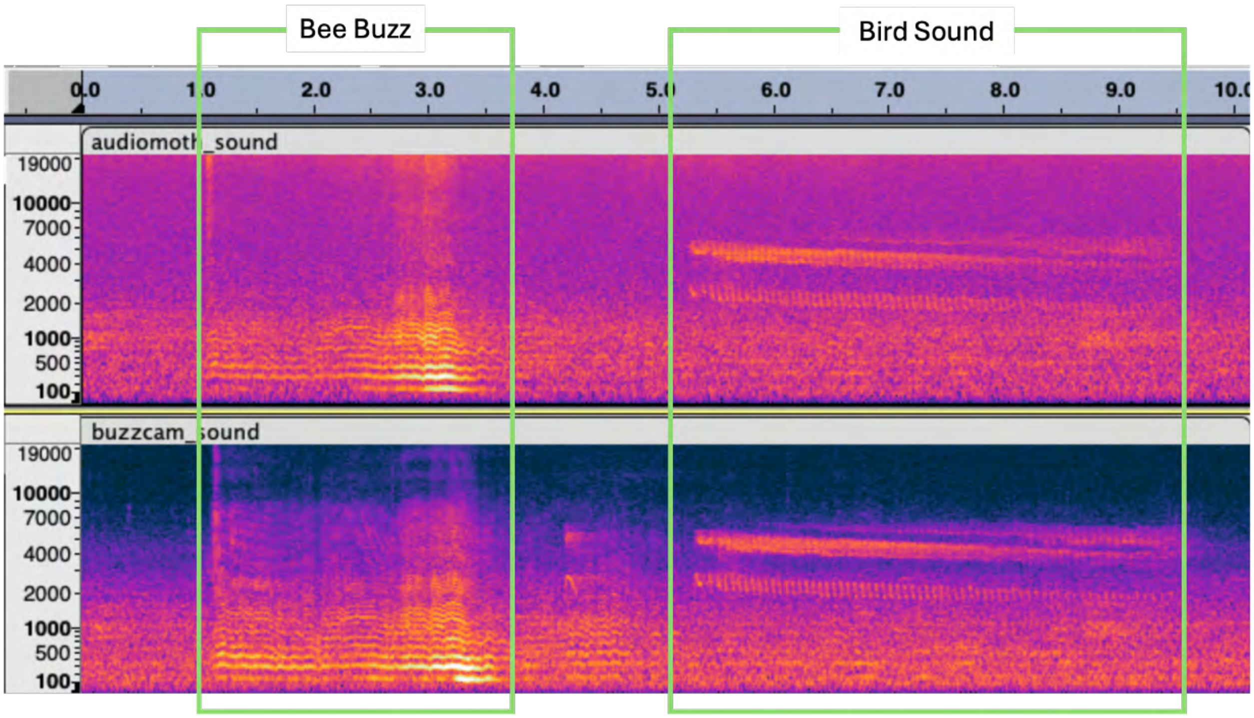


Figure 4.15: A side-by-side spectrogram comparison of an audio segment recorded simultaneously by an AudioMoth (top) and a BuzzCam (bottom) in Puerto Blest. The recordings were scaled to have an equal RMS noise floor for direct comparison. The segment contains a bee buzz (approx. 1-4 seconds) and a bird call (approx. 5-10 seconds). Note the lower background noise and clearer harmonic structure in the BuzzCam recording, which corresponds to a 3.4 dB higher Signal-to-Noise Ratio for the bee buzz.

4.7.7 Custom iOS Application for In-Field Annotation

A critical component for the creation of the high-quality labeled dataset [56] was the development of a custom iOS application (Figure 4.16), which I co-developed with Isamar Zhu, an undergraduate researcher at MIT, to interface with the BuzzCam units via BLE. The application provided several key functionalities for field researchers, allowing them to connect to individual BuzzCam units, observe real-time sensing metrics, and reconfigure device settings such as sample rates and bit resolution. Most importantly, it enabled them to trigger real-time annotations corresponding to observed bee activity. Using dedicated buttons, researchers could immediately label a sighting as either *B. dahlbomii* ("Native") or *B. terrestris* ("Invasive"), transmitting a precise timestamp to the connected BuzzCam unit, which recorded it alongside the continuous audio and environmental data streams.

This in-field, real-time annotation capability, as highlighted in Chwalek et al. [56], "is a significant improvement from traditional asynchronous manual methods, where paper observations are correlated to recorded acoustic data." It allowed for precise temporal alignment between observed bee activity and the captured sensor data, which was crucial for extracting accurate training examples for the subsequent machine learning model development. The architecture of the updated BuzzCam, integrating the MAX78000, allows this system not only to collect data as described but also to perform the on-device inference detailed in later sections, transforming it from a passive logger into an active, intelligent ecological monitor.

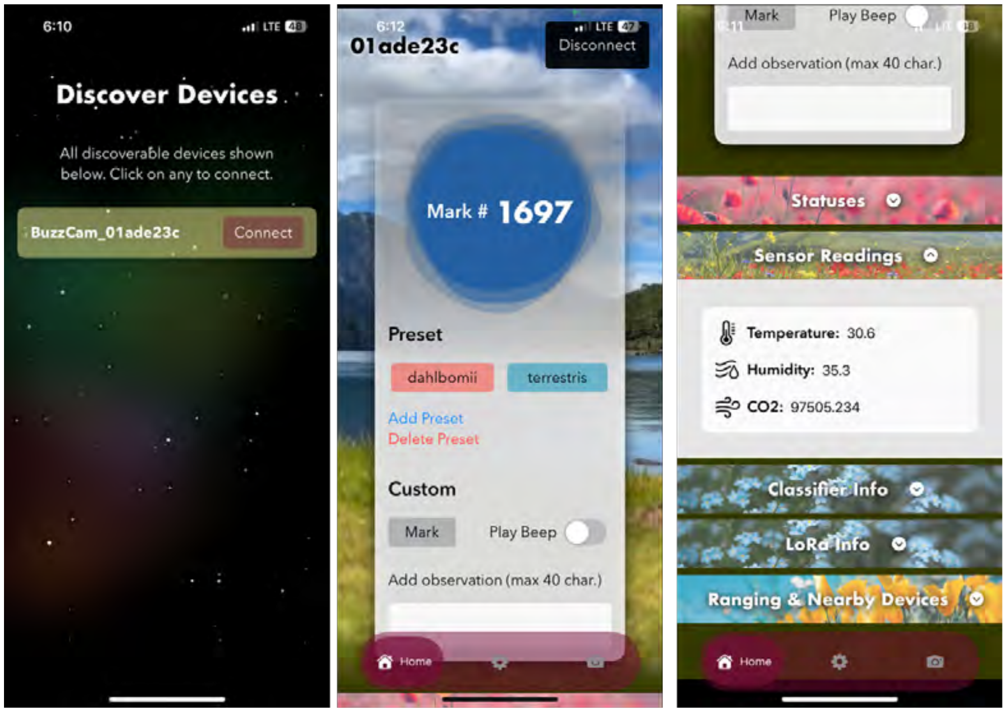


Figure 4.16: BuzzCam iOS App: device discovery (left), annotation window (middle), and real-time environmental sensor readings (right).

4.8 BuzzCam: Field Deployment and Dataset Creation in Patagonia

Following the design and development of the BuzzCam system and its accompanying iOS annotation application, the next phase involved deploying these tools in the field to collect the acoustic and environmental data necessary for monitoring *Bombus dahlbomii* and training machine learning models. This section details the fieldwork we conducted in Puerto Blest, Argentina, a key habitat for the endangered Patagonian bumblebee, and describes the comprehensive methodology we employed for data collection, annotation, and subsequent processing to create the "High-Res Acoustic and Environmental Data to Monitor *Bombus dahlbomii* Amid Invasive Species, Habitat Loss" dataset [56].

4.8.1 Study Site: Puerto Blest, Argentina – A Focal Point for *Bombus dahlbomii* Conservation

Puerto Blest, located within the Nahuel Huapi National Park in Patagonia, Argentina, was selected as the primary study site for this research. As highlighted in Chwalek et al. [56], this region is recognized for its ecological significance as a habitat for *B. dahlbomii*. The area presents a complex mosaic of temperate rainforest, characterized by diverse native flora, including *Fuchsia magellanica*, a key forage plant for Bombus species in the region. Importantly, Puerto Blest is also an area where the invasive *Bombus terrestris* has established, creating a zone of sympatry and potential competition with the native B. dahlbomii. This co-occurrence made it an ideal location to collect comparative acoustic data from both species under natural conditions.

4.8.2 Deployment Strategy and In-Field Data Collection Protocol

The fieldwork in Puerto Blest took place between March 10-15, 2024, under the National Park Administration (APN) permit #1839 [56]. Our team deployed a total of nine custom-built BuzzCam stereo acoustic recorders, each identified by a unique 4-character hexadecimal identifier.

As described in Chwalek et al. [56], we first identified several locations within Puerto Blest (Figure 4.20) where both native and invasive bees were observed actively foraging, particularly around blooming Fuchsia shrubs. BuzzCam units were then placed at these locations, all at a consistent height of approximately 0.3 meters above ground and mounted on tripods to standardize recording conditions. Some larger locations with high bee activity were designated more than one recorder. The initial locations were nearer to areas with some tourist or motor vehicle activity, which occasionally introduced anthropogenic noise into recordings (Figure 4.17). Consequently, for later deployments, we focused on several Fuchsia shrubs along the shoreline of Nahuel Huapi Lake, which exhibited higher *B. dahlbomii* density and lower ambient noise (Figure 4.18). It was noted that "only worker and male Bombus activity was observed, with no queens present due to it being late in the season" [56].

A core component of our data collection methodology was the use of the custom iOS application for real-time, in-field annotation of bee sightings. As detailed in Chwalek et al. [56], when bees were most active, a team member would position themselves near a target BuzzCam device. Upon observing either *B. dahlbomii* or *B. terrestris* (identified visually by trained field ecologists), the annotator pressed a corresponding button on the iOS application. This action transmitted a matching timestamp to the connected recorder, allowing for precise temporal linkage between the visual observation and the acoustic/environmental data record. Due to the initially low observed activity of *B. dahlbomii* at some static recorder positions, we also adopted a hybrid annotation approach. This involved annotating sounds captured by both the stationary BuzzCam units ("static" recordings) and, as shown in Figure 4.19, using a mobile setup where microphones were held closer to any observed *B. dahlbomii* activity ("dynamic" recordings) to capture clearer recordings of the target species [56].

Throughout the deployment, the BuzzCam units recorded continuous stereo audio at 48 kHz/16-bit, along with environmental data from the BME688 sensor at 0.2 Hz. In total, we amassed roughly 250 hours of uncompressed stereo audio data over the six-day collection period [56].

Post-Collection Data Processing and Annotation Refinement

Following the fieldwork, we undertook a meticulous multi-stage process to refine the raw data and annotations to create a high-quality dataset suitable for machine learning. The

Figure 4.17: Road nearby sensor locations 1 through 3 (reference Figure 4.20) [56].

Figure 4.18: Example of sensor deployment. BuzzCam sensor pictured near Fuchsia bush [56].

first step was to ensure privacy and data cleanliness by removing human speech. To do this, "the team manually scanned the spectrograms for each recorded audio file, selected segments where the power within the spectrogram correlated with human speech frequencies, listened to confirm it was human speech, and silenced the segment if human speech was detected... By silencing, rather than deleting sections of audio, we preserved the true length of the recordings for reference back to world time" [56]. Next, using the timestamps from the in-field iOS annotations, we developed a script to extract 10-second stereo audio snippets centered on each "Native" (*B. dahlbomii*) or "Invasive" (*B. terrestris*) annotation.

Figure 4.19: Example of dynamically-positioned microphone [56].

A multi-step annotation validation and filtering process followed to ensure the quality of the training data. Recognizing that some field annotations might be imprecise, we first used Amazon Mechanical Turk (ATurk) to validate if the initially extracted 10-second excerpts contained noticeable bee buzzes, providing an initial filter. To address potential misinterpretations by human annotators on ATurk, we then employed an additional filtering step using an Audio Spectrogram Transformer (AST) model [133] pre-trained on the AudioSet dataset [134]. We segmented the 10-second clips into 1-second snippets with 0.5-second overlap and used the AST model to identify snippets with a high likelihood of containing bee-related sounds. This multi-stage validation process allowed us to create robust sets of 1-second and 10-second audio clips with high confidence of either containing target bee buzzes ("Positive" sets) or containing only background sounds ("Negative" sets), with the final breakdown presented in Table 4.3.

4.8.3 The BuzzCam Patagonian Bee Acoustic Dataset

The culmination of this fieldwork and rigorous data processing is the publicly available dataset described in Chwalek et al. ([56]. The dataset is comprehensive, including raw, continuous stereo acoustic recordings (with human speech silenced); raw environmental sensor data synchronized with the audio; the initial 10-second excerpts based on field annotations; and the refined and validated 1-second and 10-second positive (Native/Invasive buzzes present)

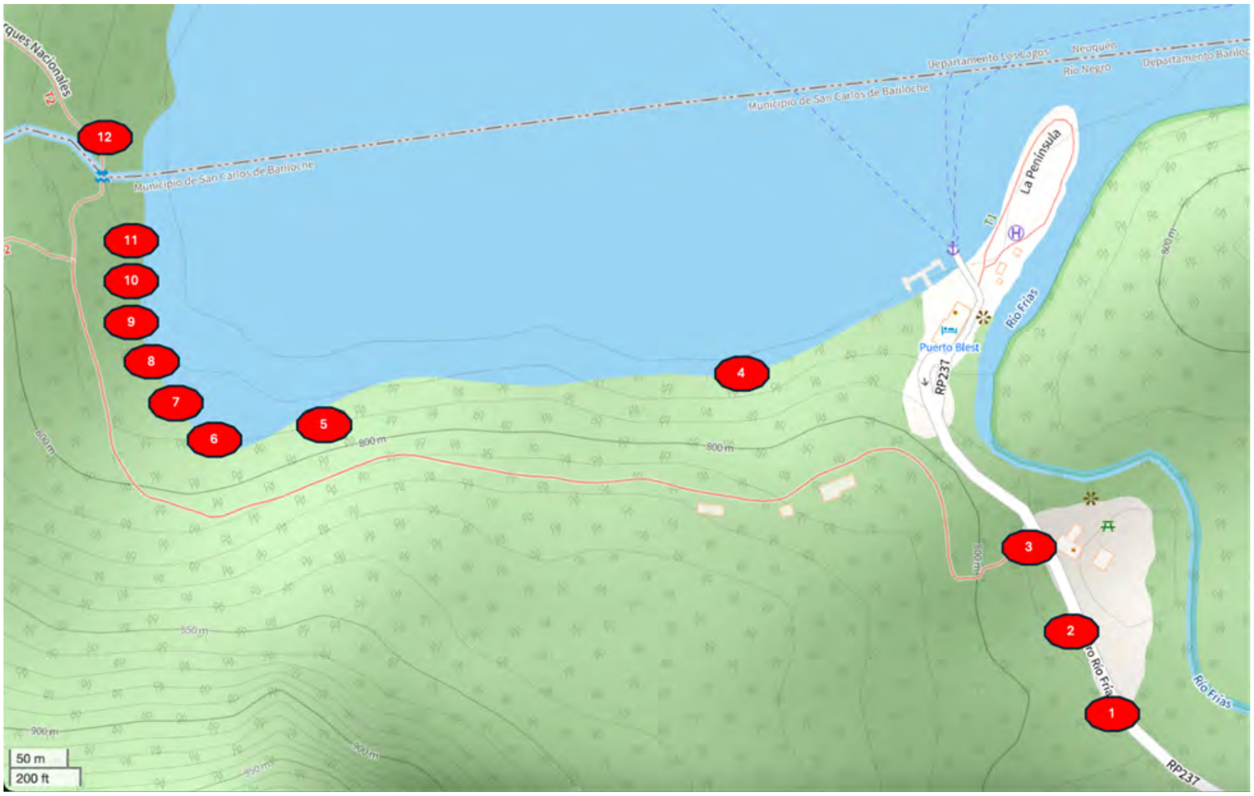


Figure 4.20: Locations of test sites in the Puerto Blest Area. Numbers represent numbered field sites [56].

and negative (background only) audio clips. It also includes detailed metadata linking all data components, such as timestamps, original field annotations, microphone type, and validation results, as well as supporting weather station data from the Puerto Blest area.

As stated in Chwalek et al. [56], "This dataset facilitates the development of machine learning models for monitoring Bombus populations, crucial for conservation efforts. Additionally, our robust data annotation techniques enhance the dataset's reliability for future modeling work." It provides a unique resource for researchers working on bee bioacoustics, machine learning for ecological applications, and understanding the interactions between native and invasive species. The availability of synchronized environmental data further enables the development of multi-modal ML models. This comprehensive dataset, created through careful fieldwork and a multi-layered validation process we designed and executed, forms the critical foundation upon which the on-device machine learning classifier, detailed in the subsequent sections, was developed and evaluated.

4.9 BuzzCam: On-Device Machine Learning for Bee Classification

Having established a high-quality, annotated dataset of Bombus bee buzzes from Patagonia [56], the subsequent and central goal of the BuzzCam project was to develop an efficient and

Table 4.3: Results from Acoustic Data Validation [56].

		Samples	
		1-second	**10-second**
Negative	**Terrestris**	4,812	278
	Dahlbomii	6,907	404
Positive	**Terrestris**	5,455	1,023
	Dahlbomii	4,572	1,020

accurate machine learning (ML) model capable of classifying these buzzes in real-time, directly on a resource-constrained microcontroller. This section details the entire pipeline I developed for this on-device AI system, from audio data preprocessing and model architecture selection to quantization-aware training (QAT) and deployment on the Analog Devices MAX78000 MCU. The methodologies and findings are further elaborated in my ongoing work.

4.9.1 Data Preprocessing

The audio data underwent a multi-stage preprocessing pipeline adapted from the ANIMAL-SPOT framework [78] and optimized for the target MAX78000 MCU. The preprocessing pipeline operates in two distinct stages optimized for the MAX78000's memory architecture (Figure 4.21). The first stage performs low-pass filtering and resampling on contiguous DMA-transferred audio data, while the second stage executes spectral analysis and feature extraction once sufficient samples accumulate for STFT computation. An ablation study was conducted using systematic hyperparameter optimization to optimize preprocessing parameters, including low-pass filter cutoff frequency, mel spectrogram configurations (resampling rate, window/hop size, pooling), the use of pre-emphasis, and spectral representation (mel vs. linear). The final stages and their parameters listed below represent the best-performing configuration identified through these experiments, optimized for the MAX78000 input requirements. The steps, applied sequentially, were as follows:

1. Low-Pass Filtering: A Finite Impulse Response (FIR) low-pass filter with a cutoff frequency of 6,114 Hz and 10 taps was applied to the raw 48 kHz audio. This cutoff was chosen to preserve the crucial harmonic content of Bombus buzzes, which extends to approximately 5 kHz , while effectively filtering out irrelevant higher-frequency environmental noise and preventing aliasing during the subsequent downsampling step [56,129,135].

2. Resampling: The filtered audio was resampled from 48 kHz to 16,384 Hz. This step reduces the computational load for the microcontroller and aligns the data rate with subsequent processing stages without losing the essential frequency information retained by the low-pass filter.

3. Pre-emphasis: A first-order pre-emphasis filter (coefficient of 0.97) was applied. This process amplifies higher-frequency components, which helps to balance the spectrum

and enhance the subtle structural details of the bee wingbeat harmonics, making the unique acoustic signatures of each species more prominent.

4. Spectrogram Generation: A Short-Time Fourier Transform (STFT) was computed using a 1024-sample Hann window and a 256-sample hop size. This configuration provides an effective balance between temporal resolution and frequency resolution (16.0 Hz per bin), which is critical for analyzing the continuous, tonal nature of flight buzzes.

5. A 512-channel Mel filter bank was applied to the linear spectrogram, spanning from 0 Hz to the 6,114 Hz cutoff. This step transforms the frequency axis to the Mel scale, which emulates human auditory perception by grouping frequencies in a perceptually relevant way. This focuses the model on the most diagnostically important spectral features for differentiating bee species.

6. Amplitude-to-dB Conversion: The Mel spectrogram energies were converted to a logarithmic decibel (dB) scale. This compresses the dynamic range of the audio, making the model more robust to variations in the loudness of the bee buzz, which can change based on the bee's distance from the microphone.

7. Normalization: The log-mel spectrograms were linearly scaled by a factor of 1/50 and clamped, then scaled to the integer range [-128, 127] for conversion to 8-bit signed integers.

8. Max Pooling: Finally, a 2D max pooling operation was applied with a window size of 2×8 and a stride of 2×8. This operation reduced the feature map dimensions from 61×512 (time x frequency) to 30×64 to meet the input dimensionality constraints of the MAX78000's hardware accelerator.

4.9.2 Model Architecture for On-Device Implementation

The core model architecture was adapted from the ResNet-18-inspired CNN encoder of the ANIMAL-SPOT framework [78], significantly pruned to meet the resource constraints of the MAX78000 MCU (Figure 4.22). The MAX78000 provides 442 KB of SRAM for weights and 512 KB for data, supporting 1 to 8-bit quantized weights [136].

Encoder

The pruned encoder, designed to operate on the 30×64 preprocessed log-mel spectrograms, consists of the following:

- **Input Layer:** A 3×3 convolutional layer with 64 output channels, a stride of 1, and padding of 1.

- **Residual Stage 1:** One residual block containing two 3×3 convolutional layers (64 channels each) and an identity shortcut. This stage preserves dimensions, outputting a 64 × 30 × 64 (Channels × Time × Frequency) feature map.

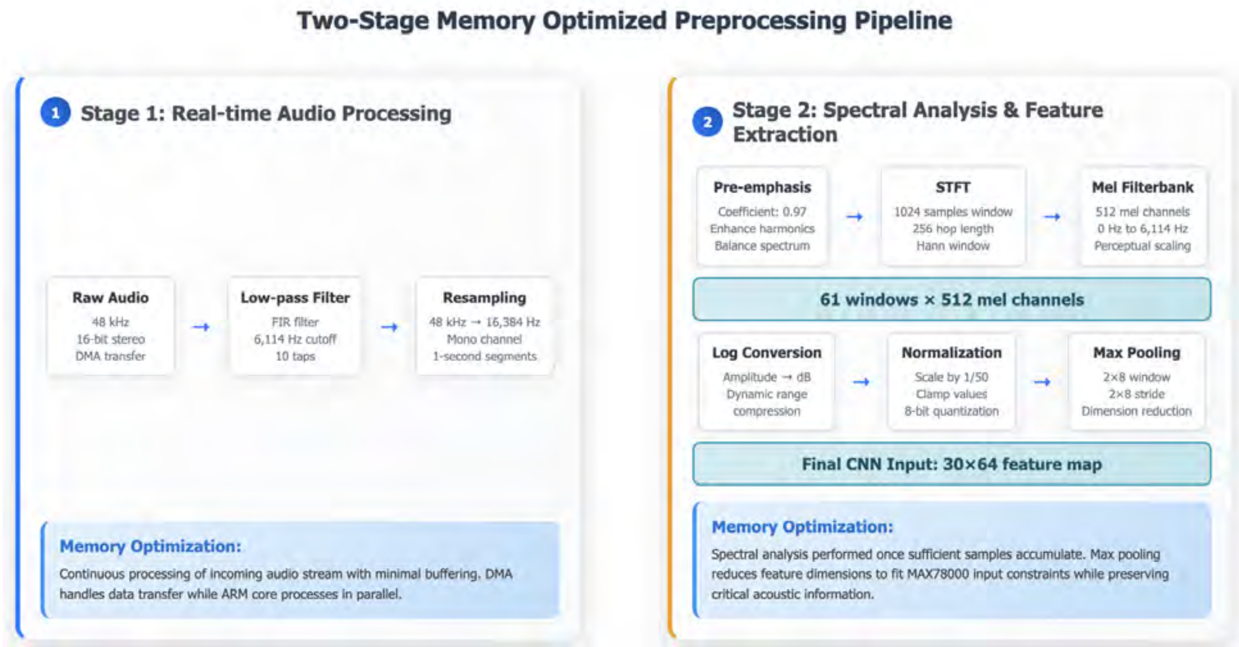


Figure 4.21: Two-stage Memory Optimized Preprocessing Pipeline

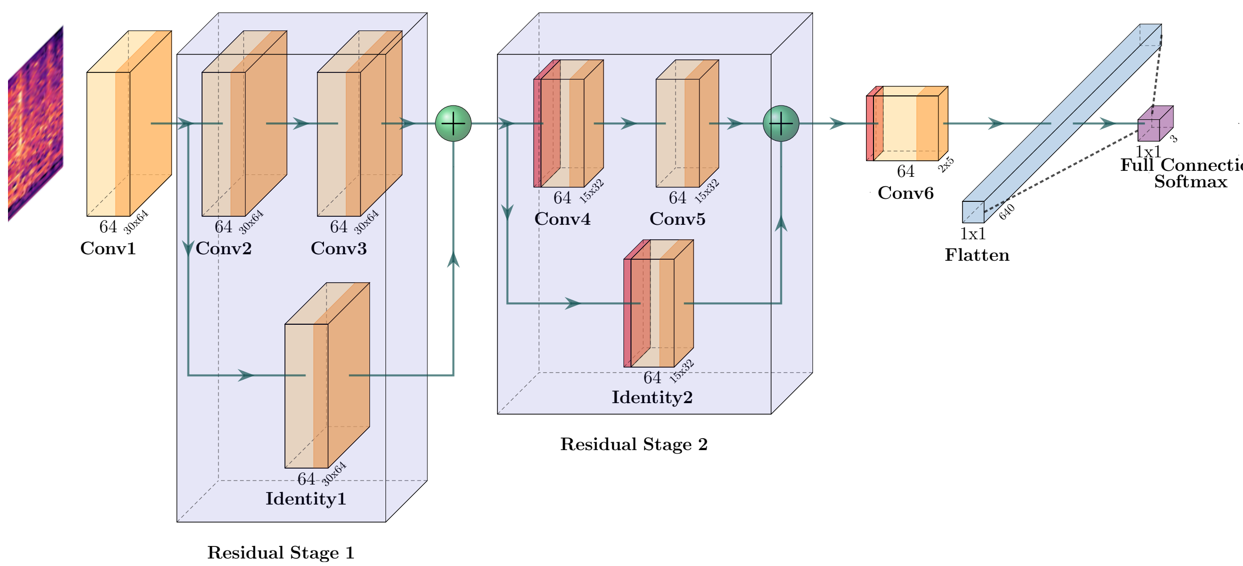


Figure 4.22: ResNet based Model Architecture on MAX78000. Image generated using PlotNeuralNet open-source tool [137].

- **Residual Stage 2:** One residual block performs downsampling using 2x2 average pooling layer with a stride of 2, which halves the spatial dimensions, followed by a 3x3 convolutional layer with a stride of 1. This is followed by a second 3x3 convolutional layer that preserves the resolution. An identity shortcut path also undergoes downsampling using a 1×1 convolution and average pooling to match the dimensions of the main path. The output feature map has dimensions $64 \times 15 \times 32$.

- **Projection Layer:** A 1×1 convolution with 64 output channels), followed by 6x6 average pooling. This results in a final feature map of $64 \times 2 \times 5$.

This architecture was selected through a series of ablation studies that varied architectural hyperparameters to optimize performance within the MAX78000's strict hardware constraints (442,000 8-bit parameters) and activation memory limits. We employed systematic hyperparameter optimization using Bayesian optimization method to optimize the top-1 validation accuracy. Only network configurations that fit within the MAX78000's memory limits were considered for training, and the best-performing model on the evaluation set was advanced to subsequent ablation studies.

Classifier

The flattened output of the encoder's projection layer was fed into a single fully connected layer with 3 output units (corresponding to the three classes), followed by a softmax activation. During inference, the final classification result was obtained by applying an argmax operation to the softmax output. This classifier was quantized to 8-bit integers. A single-layer int8 classifier was chosen over a two-layer configuration due to superior accuracy observed in ablation studies.

4.9.3 Model Training and Quantization

Models were trained using the PyTorch framework [138]. Training was performed for 200 epochs using the Adam optimizer. The learning rate was swept logarithmically in the range $[10^{-4}, 10^{-2}]$, and the batch size was set to either 64 or 96, depending on the specific configuration being tested. In the configuration that achieved the highest accuracy, training was performed for 200 epochs with a batch size of 96, yielding a total of approximately 57,000 training steps. The loss function employed was CrossEntropyLoss. To ensure the model was optimized for the target hardware, we implemented Quantization-Aware Training (QAT) using the `ai8x-training framework` from Analog Devices [139]. This process simulates the effects of quantization during training, allowing the model to adapt its weights to the reduced 8-bit precision used for all weights, biases, and activations in the deployed model. During QAT, weights and activations get quantized during the forward pass, while the gradient and parameter update is still performed in floating point. QAT helps the model to adapt its weights and activations to the reduced precision, making the network more robust to quantization noise and yielding lower accuracy degradation when converting weights and activation from floating point to integer.

4.9.4 Deployment on MAX78000

The QAT-trained PyTorch model was converted into synthesizable C code for the MAX78000's CNN accelerator using the official Analog Devices development pipeline. This process involves model parsing, quantization parameter application, weight extraction, and generation of hardware-specific instructions. The final MAX78000 firmware, including the preprocessing pipeline and the CNN model, was flashed onto the updated BuzzCam hardware for on-device evaluation.

4.9.5 Model Evaluation

Off-Device Performance Metrics The performance of the trained models was evaluated on the held-out test set using standard classification metrics: overall accuracy, precision, recall, F1-score (macro-averaged), and Area Under the Precision-Recall Curve (AUPRC) for each class.

On-Device Performance Metrics Once deployed on the updated BuzzCam [56] platform featuring the MAX78000, the model's real-world performance was assessed by measuring:

1. **Inference Latency:** The time taken for the MAX78000 to process one 1-second audio input segment from the live audio stream, reporting both preprocessing on the ARM core and CNN inference on the accelerator.

2. **Inference Efficiency:** The inference efficiency of the top-performing model was also evaluated using the MACs-per-cycle (MAC/Cycle) metric, which quantifies the number of multiply-accumulate operations performed per clock cycle by the CNN accelerator.

3. **Accuracy on Device:** Verification of classification accuracy using a representative subset of the test data fed through the live ADC frontend of the updated BuzzCam to the MAX78000, to confirm consistency between simulated quantized performance and actual on-hardware performance.

4. **Power Consumption:** Power consumption was measured on the MAX78000-based BuzzCam platform using the windowed accumulation mode in the Power Monitor firmware provided by Analog Devices. To establish a baseline, idle power was recorded over a 1-second period of inactivity. Active power was then measured across the three distinct CNN phases: kernel loading, input loading, and inference. To ensure a stable average, each phase was repeated 100 times. The final energy consumption per inference was calculated by first subtracting the idle power from the active power measurement to find the net power draw, and then multiplying that result by the active measurement duration.

5. **Parameter Count:** The total number of trainable parameters in the deployed model, including all convolutional layer weights, biases, and fully connected layer parameters. This metric indicates the model's complexity and memory requirements for weight storage on the target hardware.

4.10 BuzzCam: Results and Evaluation

4.10.1 Model Performance Comparison

Table 4.4: Model Performance - General Metrics

Model	#Params	Test Acc. (%)	Precision (%)	Recall (%)	F1 (%)
ANIMAL-SPOT	11,176,003	85.4	83.5	81.7	82.4
MAX78000	158,144	86.3	84.1	83.1	83.5

The performance evaluation of our optimized CNN architecture against the baseline ANIMAL-SPOT model demonstrates that substantial model compression can actually improve classification accuracy for edge deployment. Tables 4.4 and 4.5 present a comprehensive comparison of both models across multiple performance metrics.

Our compressed MAX78000-optimized model achieved an overall test accuracy of 86.3%, actually surpassing the full ANIMAL-SPOT baseline (85.4%) by 0.9 percentage points (Table 4.4). This performance improvement is particularly remarkable given the substantial reduction in model complexity: our optimized architecture contains only 158,144 parameters compared to ANIMAL-SPOT's 11,176,003 parameters—a 70-fold reduction in model size while actually improving classification performance.

The precision and recall metrics further validate the effectiveness of our optimization approach. The MAX78000 model achieved 84.1% precision and 83.1% recall, compared to ANIMAL-SPOT's 83.5% precision and 81.65% recall. Notably, our model showed improved recall performance (+1.45 percentage points), indicating better sensitivity in detecting positive instances of *Bombus* species, which is critical for conservation monitoring applications where missing rare endangered species detections could have significant ecological implications.

The macro-averaged F1-score of 83.5% for the MAX78000 model (compared to 82.42% for ANIMAL-SPOT) demonstrates well-balanced performance across all three classes, suggesting that the architectural reduction and quantization-aware training successfully preserved the model's discriminative capabilities across the *B. dahlbomii*, *B. terrestris*, and negative classes.

4.10.2 Per-Class Performance Analysis

Table 4.5: Model Performance - Per-class AUPRC

Model	Negative	*Bombus dahlbomii*	*Bombus terrestris*
ANIMAL-SPOT	0.973	0.725	0.828
MAX78000	0.983	0.768	0.865

The per-class Area Under the Precision-Recall Curve (AUPRC) values provide critical insights into the model's species-specific classification capabilities (Table 4.5). These results are contextualized by the dataset's class distribution: 20,370 negative samples (52.3%), 10,070 *Bombus terrestris* samples (25.8%), and 8,532 *B. dahlbomii* samples (21.9%).

The MAX78000 model excels at identifying negative samples (environmental background noise), achieving an AUPRC of 0.983, slightly surpassing ANIMAL-SPOT (0.973). This high performance, while partly influenced by the majority class size, is essential for minimizing false positives in autonomous monitoring, ensuring reliable operation in noisy field conditions such as wind or ambient sounds.

For the invasive *B. terrestris*, the MAX78000 model achieves an AUPRC of 0.865, improving upon ANIMAL-SPOT's 0.828 (+0.037). This enhanced detection is vital for conservation, enabling accurate monitoring of *B. terrestris*' spread and impact in Patagonia.

The endangered *B. dahlbomii* poses the greatest classification challenge, with the MAX78000 model achieving an AUPRC of 0.768 compared to ANIMAL-SPOT's 0.725 (+0.043). The smaller sample size (8,532 samples) likely contributes to this lower performance, as deep learning models benefit from larger datasets to capture robust features. Despite this, the AUPRC of 0.768 significantly exceeds the random baseline of 0.219 (class proportion), indicating meaningful discriminative capability.

Several factors affirm the model's robustness despite class imbalances: (1) Consistent AUPRC improvements over ANIMAL-SPOT across all classes suggest genuine enhancements rather than overfitting to the majority class; (2) A macro-averaged F1-score of 83.5% indicates balanced performance across classes, mitigating bias toward the negative class; (3) The improvement over ANIMAL-SPOT, a bioacoustic-specific baseline, validates the model's ability to capture subtle acoustic signatures; (4) The preprocessing pipeline and quantization-aware training enhance feature extraction for *B. dahlbomii*, as evidenced by the performance gain.

To evaluate robustness to field noise, the model's strong performance on the negative class (0.983 AUPRC) demonstrates its ability to distinguish bee buzzes from environmental noise, including wind and ambient sounds recorded in Patagonia. However, untested noise variations (e.g., overlapping insect vocalizations or extreme weather in other ecosystems) could challenge generalization. Future work should validate performance across diverse ecological soundscapes to ensure scalability for global conservation applications.

The AUPRC of 0.768 for *B. dahlbomii*, though the lowest, supports reliable detection for conservation monitoring, where identifying rare endangered species is critical. The model's overall robustness, driven by optimized preprocessing and on-device efficiency, positions it as a practical tool for real-time, scalable ecological monitoring.

4.10.3 Model Efficiency and Resource Utilization

The dramatic reduction in model parameters from over 11 million to 158,144 (99.99% reduction) while achieving improved accuracy demonstrates the effectiveness of our systematic architecture compression and quantization approach. This compression enables deployment on resource-constrained microcontrollers where the full ANIMAL-SPOT model would be computationally infeasible.

This improved performance is likely attributable to several complementary factors. The architectural reduction strategy used to adapt the ResNet-18-based ANIMAL-SPOT architecture for the MAX78000 preserved the initial convolutional layers, which are critical for feature extraction in environmental sound classification, while removing deeper, more complex stages that may have been prone to overfitting on the training data. Additionally, the quantization-aware training process may have acted as an implicit regularization mechanism, forcing the model to learn more robust feature representations that are less sensitive to numerical precision. The refined preprocessing pipeline, optimized specifically for *Bombus* acoustic characteristics through systematic ablation studies, likely enhanced the quality of input features compared to the more general-purpose ANIMAL-SPOT preprocessing. Finally, the architectural constraints imposed by the MAX78000's memory limitations may have prevented the model from learning overly complex decision boundaries, resulting in better generalization to unseen test data.

The successful deployment of the 8-bit quantized model on the MAX78000 platform validates our quantization-aware training methodology. The minimal accuracy degradation—indeed, the accuracy improvement—observed during the transition from floating-point to integer arithmetic indicates that the model weights and activations adapted effectively to the reduced precision constraints imposed by the target hardware.

These results collectively demonstrate that sophisticated bioacoustic classification capabilities can be successfully compressed and deployed on edge devices with improved performance, challenging the conventional assumption that model compression necessarily involves accuracy trade-offs. The superior performance of our optimized model suggests that the systematic architecture compression, refined preprocessing pipeline, and quantization-aware training not only enabled edge deployment but also enhanced the model's discriminative capabilities for *Bombus* species classification.

4.10.4 On-Device Performance and Energy Efficiency

The deployment of our optimized model on the MAX78000 microcontroller demonstrated exceptional real-time performance characteristics suitable for autonomous field monitoring applications. The complete processing pipeline for a 1-second audio segment achieved an inference latency of 10.4 ms, enabling real-time classification with substantial computational headroom for additional system tasks.

The energy consumption analysis revealed highly efficient operation across the system components. The CNN accelerator consumed only 0.794 mJ per inference. The MAX78000 ARM core's average energy consumption, which includes both the active preprocessing of acoustic data and the deep-sleep intervals between processing, was 15 mJ. For context, the broader BuzzCam system components including the STM32WB processor, environmental sensors, ADC, microphones, and SD card logging consumed 150 mJ.

To contextualize this efficiency for field deployment, the system can operate continuously for 2.8 days on a single 18650 battery (11.1 Wh) while running inference in real-time. More practically, if the system runs only 12 hours per day when bees are present, the battery extends operational duration to approximately 5.6 days. The integration of a 1.4 W solar panel provides sufficient energy that, with just an average of 1.7 hours of direct sunlight daily, the system can operate indefinitely.

These performance metrics validate the practical feasibility of deploying sophisticated bioacoustic classification in resource-constrained edge environments, enabling scalable monitoring networks that can operate autonomously for extended periods while providing near real-time ecological insights.

4.11 BuzzCam: Discussion

This study successfully demonstrated the deployment of an on-device machine learning classifier for distinguishing endangered *B. dahlbomii* from invasive *B. terrestris* and background noise on the resource-constrained MAX78000 microcontroller. The achievement of 86.3% accuracy with 70-fold parameter reduction while maintaining 10.4 ms inference latency and consuming only 794 µJ per classification represents a significant advancement in autonomous pollinator monitoring technology.

4.11.1 Key Technical Achievements

The superior performance of our compressed model over the full ANIMAL-SPOT baseline is not merely an engineering achievement; it is a significant insight into the nature of applying deep learning to specialized bioacoustic tasks. This counter-intuitive result suggests that for a well-defined problem like *Bombus* classification, larger, more generalized architectures may be prone to overfitting, learning spurious correlations from the background acoustic environment rather than honing in on the most salient features of the target signal. Our systematic process of architectural pruning, combined with a preprocessing pipeline optimized specifically for the harmonic structure of bee buzzes, likely acted as a form of architectural regularization. It forced the model to learn a more robust and efficient feature representation, effectively filtering out irrelevant complexity. This finding has profound implications for conservation technology, demonstrating that the pursuit of on-device AI is not necessarily a compromise on accuracy. In fact, the constraints of edge computing can drive the development of more focused, specialized, and ultimately better-performing models, making advanced AI more accessible and effective for real-world conservation.

The improved recall performance (+1.45 percentage points) is particularly significant for conservation applications where missing endangered species detections could have critical ecological implications. The AUPRC of 0.768 for *B. dahlbomii*, while the lowest per-class performance, substantially exceeds random baseline expectations (0.219) and demonstrates reliable detection capabilities suitable for practical monitoring applications.

4.11.2 Conservation and Deployment Implications

This technology addresses fundamental limitations of traditional pollinator surveys by enabling continuous, autonomous surveillance with unprecedented temporal resolution. The system's energy efficiency—operating 5.6 days on battery alone or indefinitely with minimal solar charging—makes widespread deployment feasible in remote locations. Near real-time transmission of summarized data (e.g., hourly species counts) enables proactive conservation responses, such as identifying invasive species hotspots or tracking habitat restoration success.

The democratization of sophisticated monitoring tools is particularly valuable, enabling smaller research groups and local conservation agencies to implement evidence-based programs. Continuous acoustic data can reveal temporal niche partitioning, peak foraging periods, and correlations with environmental variables, providing insights crucial for predictive modeling and targeted conservation strategies.

Furthermore, the on-device intelligence of BuzzCam fundamentally changes the workflow for ecological experts. Instead of being faced with the daunting task of manually listening to hundreds of hours of raw audio, researchers can leverage the system's output to dramatically improve efficiency. The model not only provides quantitative estimates of buzz activity over time but can also generate a log of high-confidence temporal events. This allows experts to hone in on specific, relevant sections of the audio for verification or more detailed analysis, transforming the data review process from an exhaustive search into a targeted investigation.

4.11.3 Limitations and Considerations

Several limitations warrant acknowledgment. The model was trained primarily on data from Puerto Blest, Argentina, potentially limiting generalizability to different acoustic environments without site-specific adaptation. The dataset predominantly captured worker and male activity due to seasonal timing, and performance on different castes or behavioral contexts remains untested. While preprocessing aims to capture distinctive *Bombus* features, potential misclassification with other acoustically similar insects in complex environments requires further validation. The current acoustic-only approach could benefit from multi-modal sensor integration for enhanced accuracy and richer contextual data. While not explored in the machine learning pipeline of this work, the co-registered environmental data from BuzzCam's sensors offers significant system-level potential. For instance, local weather data could be used as a contextual prior to inform the classifier's confidence. During a heavy rain event, which can be identified by both the acoustic channels and a rapid increase in humidity, the system could be programmed to anticipate that bee activity is highly unlikely. This context would allow it to down-weight or flag potential acoustic false positives from the sound of rain, leading to a more robust and reliable monitoring system.

The current acoustic-only approach could benefit from multi-modal sensor integration for enhanced accuracy and richer contextual data. Additionally, long-term system durability under variable environmental conditions has not been extensively demonstrated, though the core ML performance and hardware integration are well-validated.

4.11.4 Potential Applications and Broader Impact

Despite these limitations, the BuzzCam system and the on-device AI methodology we have developed hold considerable promise:

- Enhanced Pollinator Monitoring Programs: BuzzCam offers a tool for more scalable, cost-effective, and less invasive monitoring of *Bombus* populations compared to traditional methods.

- Research into Bee Ecology and Behavior: Continuous acoustic data, potentially augmented by on-device classification, can reveal fine-scale temporal niche partitioning, identify periods of peak foraging activity, and correlate these with environmental variables also captured by BuzzCam [56].

- A Blueprint for Other On-Device Bioacoustic Systems: The methodologies we employed for model pruning, QAT, and deployment on the MAX78000 can serve as a valuable blueprint for developing similar intelligent sensor solutions for a diverse array of other acoustically active species and varied ecological contexts.

The successful demonstration of a pathway to on-device AI for *Bombus* classification paves the way for transforming ecological monitoring from a reactive, often data-limited endeavor into a proactive, data-rich paradigm capable of providing timely insights for evidence-based conservation.

4.11.5 Future Directions

Future work will focus on two primary areas: enhancing model capabilities and completing full system deployment. Model improvements include exploring advanced architectures through Neural Architecture Search, expanding training datasets across diverse geographical and seasonal conditions, and developing finer classification capabilities for caste differentiation and behavioral analysis [140,141].

System integration priorities include completing LoRa communication integration, developing centralized data visualization dashboards, and conducting extensive multi-season field trials. A significant extension of the system's capabilities would be to network multiple BuzzCam units to function as a dynamic, distributed microphone array. This would enable larger-scale spatial analysis of pollinator activity. By leveraging GPS for coarse time synchronization and acoustic chirps from integrated piezoelectric buzzers for fine-grained calibration, the network could achieve the tight temporal alignment necessary for sound source localization. Onboard IMUs would provide device orientation, and the system could be further enhanced with ultrawideband (UWB) radios to achieve even higher-resolution spatial and temporal tracking. Once validated, this methodology can be adapted for monitoring other important pollinators and acoustically active fauna, extending utility across diverse ecological applications and contributing to broader biodiversity assessment efforts.

4.12 Chapter Conclusion: Advancements in Acoustic Sensing for Ecological Understanding

This chapter detailed the dedicated efforts to advance ecological acoustic monitoring through the development of two distinct yet interconnected platforms: SoundSHROOM and BuzzCam. The progression from designing a general-purpose, robust recorder to an AI-enabled, specialized sensor illustrates a key facet of this doctoral research—iterative technological development tailored to specific ecological challenges.

The SoundSHROOM project served as a practical validation for robust hardware design in extreme environments. Its deployment in the Arctic confirmed the system's durability and yielded two key contributions: a unique, public dataset of multi-channel Arctic soundscapes suitable for spatial audio analysis, and empirical data on effective wind noise mitigation techniques [59].

These foundational learnings in hardware engineering were then applied to the BuzzCam project, which addressed the "data-to-insight" bottleneck in passive acoustic monitoring. This work encompassed the design of a specialized sensor, the meticulous creation of a high-resolution annotated dataset of *Bombus* buzzes from Patagonia, and the development of a complete end-to-end pipeline for on-device machine learning [56]. By successfully adapting and optimizing a convolutional neural network for the resource-constrained Analog Devices MAX78000 microcontroller, this research demonstrated a practical method for real-time, low-power acoustic classification of endangered and invasive bee species in the field.

Collectively, the work in this chapter contributes new tools, datasets, and a repeatable methodology for on-device bioacoustics to ecological technology and conservation science. The experience gained from these acoustic-centric systems, from hardware engineering for extreme conditions to the complexities of on-device AI, provided the necessary foundation to address the broader challenge of developing a comprehensive, multi-modal sensing platform for larger wildlife, as detailed in the following chapter on the CollarID project.

Chapter 5

CollarID: Engineering a Robust, Multi-Modal Platform for Diverse Wildlife Monitoring

5.1 Introduction: The Need for Advanced Wildlife Monitoring Technologies

The effective conservation and management of wildlife populations in an increasingly human-dominated and rapidly changing world depend critically on our ability to understand their behavior, spatial ecology, health, and interactions with their environment. Traditional methods for studying wild animals, while foundational, often face significant limitations in terms of the scope, resolution, and type of data they can provide. As ecological questions become more complex—spanning from individual physiological responses to landscape-level population dynamics and the nuanced impacts of anthropogenic disturbances—there is a growing and urgent need for more advanced, integrated, and versatile monitoring technologies. This chapter details my development of CollarID, a novel multi-modal sensing platform engineered to address these evolving research needs and to provide richer, more contextualized insights into the lives of diverse wildlife species.

5.1.1 Challenges in Monitoring Diverse Wildlife: Moving Beyond Location-Only Data

For decades, biologging devices, particularly GPS collars, have revolutionized wildlife ecology by allowing researchers to track animal movements over vast distances and gain unprecedented insights into home ranges, migration patterns, and habitat selection [64,65]. However, while location data is invaluable, it often tells only part of the story.

Many existing commercial tracking collars face several challenges. A primary issue is their limited behavioral and physiological context; standard GPS collars primarily provide location data, offering little direct information on what an animal is doing (e.g., foraging, resting, socializing, exhibiting stress behaviors) or its physiological state, and inferring behavior from movement patterns alone can be ambiguous and lacks fine-grained detail [67]. Furthermore, there is a lack of co-registered environmental data, as most collars do not have comprehensive sensing suites to measure ambient conditions (e.g., temperature, humidity, air quality), making it difficult to directly link animal behavior or physiology to specific microclimatic variables or environmental stressors. Power constraints and deployment longevity are also major limiting factors, as continuous high-frequency sensing (e.g., for accelerometry or audio) or operating power-intensive subsystems such as GPS at aggressive fix rates can rapidly deplete power reserves, especially for smaller animals where battery size is constrained. Moreover, the size, weight, and cost of collars can impact animal welfare and behavior, particularly for smaller species, and high costs can limit the scale of studies or their accessibility to researchers with constrained budgets. Finally, data retrieval poses challenges, often requiring animal recapture for non-telemetered devices—which can be difficult, stressful for the animal, and may lead to data loss—or facing bandwidth limitations that restrict the type and volume of data that can be transmitted remotely.

These limitations highlight a critical gap: the need for wildlife monitoring tools that can provide a more holistic picture by integrating location data with fine-grained behavioral metrics, physiological indicators, and co-registered, high-resolution environmental data, all within a power-efficient and robust package.

5.1.2 Introducing CollarID: Vision and Specific Objectives

It was with these challenges and opportunities in mind that I conceived and embarked upon the development of the CollarID platform. The overarching vision for CollarID is to create a lightweight, low-power, robust, and highly configurable multi-modal sensing system that is not limited to a single taxonomic group but is versatile enough to be adapted for monitoring a diverse range of terrestrial mammals, including species of significant conservation concern such as lions, hyenas, and wild dogs, as well as animals of agricultural or ecological interest like camels and cattle. This vision for a holistic sensing package builds directly on the insights from the AirSpecs project, which demonstrated the necessity of fusing environmental and physiological data to understand human comfort, a principle this work extends to the ecological domain.

Within the scope of my PhD research, the specific objectives for the development and engineering validation of the CollarID platform were:

1. To design and engineer a novel hardware prototype (CollarID) that integrates a comprehensive suite of sensors. This includes inertial measurement units (IMUs) for detailed movement and activity analysis, microphones for bioacoustic recording (capturing vocalizations and environmental sounds), and a suite of environmental sensors to measure ambient temperature, humidity, barometric pressure, and potentially air quality indicators like gases and particulate matter.

2. To prioritize low-power operation and miniaturization in the hardware and firmware design. This is crucial for maximizing deployment longevity on free-ranging animals and minimizing the physical burden of the device.

3. To engineer a robust physical enclosure and system design capable of withstanding the harsh conditions and physical stresses associated with deployment on active wild

animals across diverse environments.

4. To integrate long-range, low-power communication capabilities into the platform to enable remote status updates, device configuration, and potentially low-bandwidth data telemetry for critical event notification or summary data transmission.

5. To conduct rigorous engineering validation and comprehensive characterization of the CollarID prototype's key subsystems. This includes detailed power consumption analysis across different operational modes, mechanical robustness testing, communication range testing in realistic field conditions, and, where applicable, bench-level validation of sensor performance, as well as a comparative assessment against existing commercial sensor solutions.

The successful achievement of these engineering objectives would result in a validated, field-ready prototype platform, representing a significant technological advancement for wildlife research and forming a key contribution of this dissertation. While full-scale deployments on target wildlife species are planned as immediate future work beyond the primary scope of this thesis, the rigorous development and characterization of the CollarID platform itself is the central focus of this chapter.

5.1.3 Chapter Overview

This chapter will describe the development and validation of the CollarID platform (Figure 5.1). The subsequent sections are organized by subsystem to provide a cohesive narrative for each component's design and performance. Section 5.2 details the mechanical housing design and its rigorous structural and environmental validation. Section 5.3 covers the core electronics, power management, and the performance validation of the integrated environmental sensors. Section 5.4 describes the long-range communication subsystem and the results of its field-based range testing. Section 5.5 presents results from an integrated system field trial on a farm. Finally, Section 5.6 discusses the broader capabilities and novelty of the CollarID platform and its limitations, and Section 5.7 concludes the chapter, summarizing its contributions and transitioning to the final chapter of the dissertation.

5.2 Mechanical Design and Validation

5.2.1 Mechanical Housing Design

The mechanical housing is engineered for extreme durability and environmental resistance. It is composed of a custom machined transparent polycarbonate top cover and a machined 6061 aluminum base, which has been anodized for corrosion resistance (Figure 5.2). The polycarbonate cover is vapor polished for optical clarity and coated with a transparent hardened layer for scratch and UV resistance. The two halves of the enclosure are secured with six M3 16mm flat head screws, and a custom-cut rubber gasket seated between them provides a watertight seal. The design philosophy for this robust environmental sealing draws from the practical lessons of the SoundSHROOM project, where ensuring the integrity of the

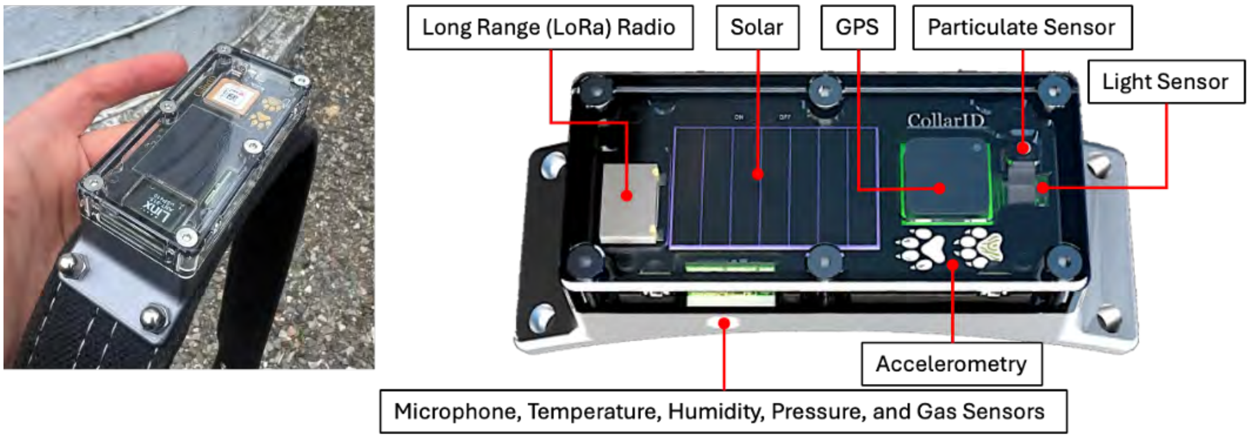


Figure 5.1: The CollarID prototype mounted on a cotton collar to show its physical form factor and scale (left), and an annotated top view detailing the integrated multi-modal sensor suite (right).

housing against the harsh, wet conditions of the Arctic was paramount for successful data collection. The complete housing, including the internal electronics and battery, weighs 155 grams.

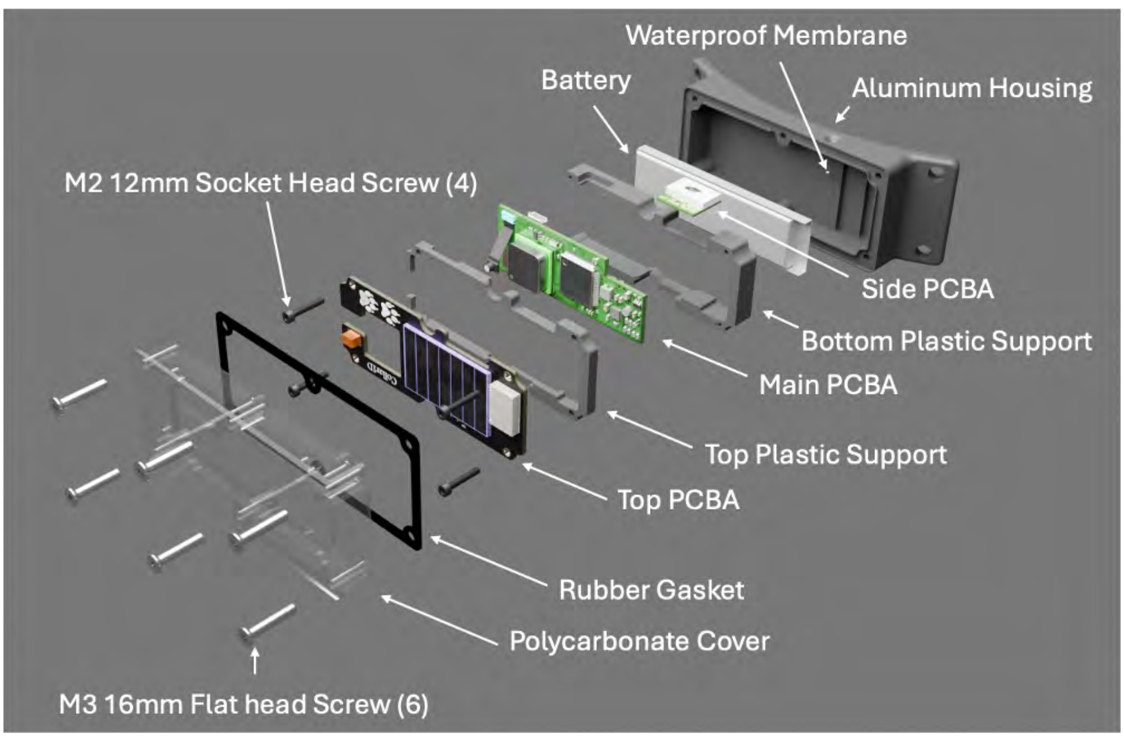


Figure 5.2: CollarID Exploded View

The aluminum base features a 0.5mm port for the microphone and environmental sensor. This port is sealed with a specialized membrane that allows sound to pass through while preventing water or dust ingress. The side board containing these sensors interfaces to the port with an additional adhesive gasket, providing a secondary seal to protect the main chamber if the external membrane is punctured. The top board also serves as a heat sink for the particulate sensor, interfacing with it via a custom copper heat sink and thermal pad. For deployment, the ergonomically curved aluminum base attaches to a collar with four M5 screws.

5.2.2 Mechanical and Environmental Robustness Validation

To ensure the structural integrity and field survivability of the CollarID housing, the design was subjected to Finite Element Analysis (FEA) based on forces it may encounter from the study animals or their primary competitors. Of particular concern are the powerful jaws of the spotted hyena (*Crocuta crocuta*), a species renowned for its bone-cracking (durophagous) feeding ecology. Wroe, McHenry, and Thomason [142] estimates the bite force of a spotted hyena at the canine as 773 N. While lions (*Panthera leo*) are significantly larger and possess a greater absolute bite force, estimated by the same study to be 1768 N at the canine, their bite is primarily adapted for gripping and suffocating prey. Given that the collar may be subjected to the powerful, focused bite of a hyena or the strong grip of a lion, both of these scientifically documented bite forces were taken into consideration as critical inputs for the FEA to validate the enclosure's robustness.

FEA Simulation of Dorsal Bite Force

To establish a realistic load parameter for the FEA, canine tooth dimensions for *Panthera leo* and *Crocuta crocuta* were sourced from foundational biomechanics literature. The basal anteroposterior and mediolateral diameter measurements were obtained from Table 1 of Van Valkenburgh and Ruff [143], where they are presented in base-10 logarithmic format. These logarithmic values were converted to their corresponding linear measurements in millimeters, yielding basal tooth widths greater than 10 mm for both species. Based on these large dimensions, a conservative contact area representing a focused, tapered canine tip was estimated as a 2.5 mm diameter circle for the FEA simulation. It is important to note, however, that it is behaviorally and biomechanically unlikely the device would be subjected to the animal's absolute maximum physiological bite force. Vertebrates are known to actively modulate their bite force based on the properties of the object being bitten to prevent catastrophic self-injury, such as tooth fracture [144]. Therefore, by designing for the theoretical maximum force applied over a concentrated area, the CollarID housing is robustly validated against a worst-case scenario that likely exceeds the more probable, lower-force investigatory or agonistic bites it may encounter in the field.

To translate the documented bite forces into load parameters for the Finite Element Analysis (FEA), pressure values were calculated based on the predicted contact area. Assuming a concentrated load from a canine tooth tip, a circular contact area with a diameter of 2.5 mm was defined, yielding a surface area of approximately 4.91 mm^2. For the spotted hyena (*Crocuta crocuta*), the documented 773 N bite force results in a pressure of approximately 157.4 N/mm^2 (157.4 MPa). The substantially larger bite force of the lion (*Panthera leo*), at 1768 N, results in a pressure of approximately 360.1 N/mm^2 (360.1 MPa) over the same area. Given that the lion's bite exerts more than double the pressure of the hyena's in this puncture scenario, the lion's corresponding pressure of 360.1 MPa was selected as the primary baseline for the subsequent structural analysis, as it represents the more severe test of the CollarID enclosure's integrity.

The dorsal bite scenario was simulated using SolidWorks Finite Element Analysis (FEA) to evaluate the structural response of the polycarbonate enclosure. A pressure of 360.1 MPa, corresponding to the maximum force of a lion's bite, was applied uniformly over a 2.5 mm diameter circular area. This load was centered on the largest unsupported span of the enclosure, a region anticipated to experience the largest deflection from external forces (Figure 5.3). This method represents the localized stress concentration from a direct canine puncture. In this scenario, the maximum observed deflection was 1.5 mm, with an average von Mises stress of 98 MPa and a maximum of 197 MPa at the contact surface (Figure 5.4). These simulated stresses are well beyond the polycarbonate's yield strength (63 MPa), and the surface contact stress exceeds the ultimate tensile strength (70 MPa). Under this maximum bite force, the polycarbonate is predicted to experience significant permanent deformation (i.e., bite marks and dents). While complete fracture is uncertain due to the material's ductility and the presence of internal supporting features, repeated loading at these levels would be a significant fatigue concern.

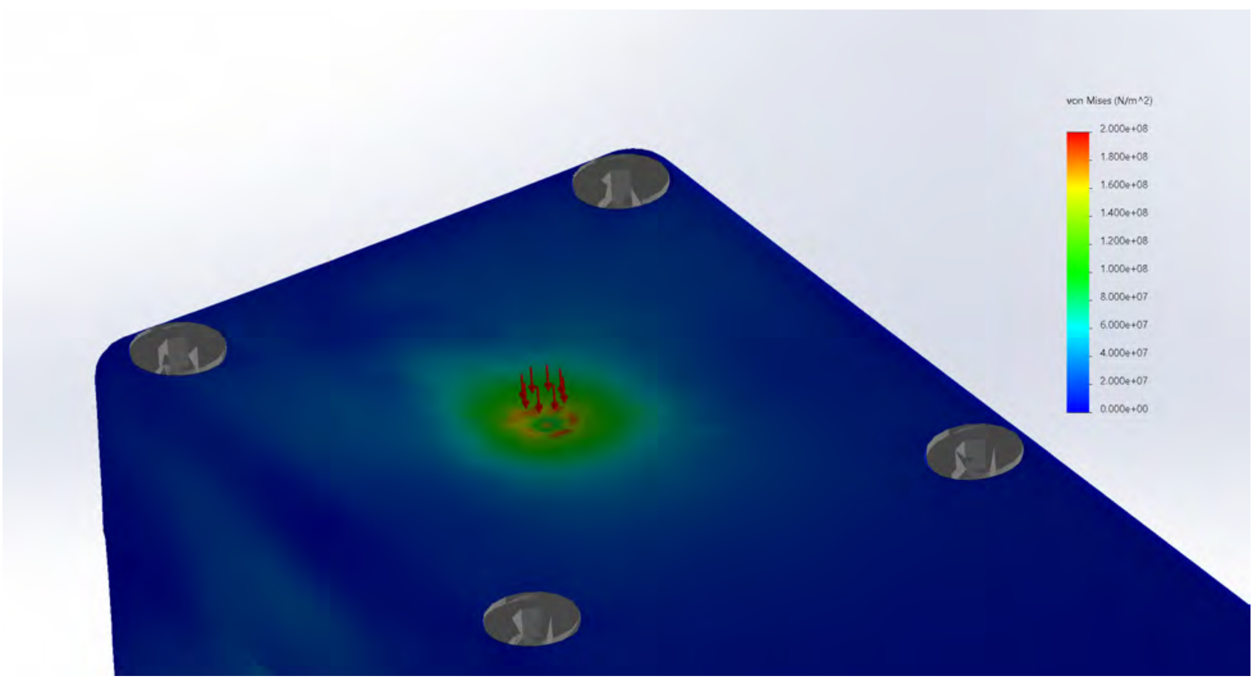


Figure 5.3: Dorsal loading simulation showing displacement from a 360.1 MPa pressure applied over a 2.5 mm diameter circle on the top polycarbonate surface.

To evaluate the system's response to a more probable, sub-maximal interaction, a second operational load case was defined using 50% of the maximum theoretical force. This value was chosen as a conservative engineering estimate for a significant but non-lethal investigatory or agonistic bite. This reduced simulation, corresponding to a pressure of approximately 180.1 MPa, resulted in an average von Mises stress of 55 MPa and a maximum surface stress of 96 MPa (Figure 5.5). While the average stress remains below the material's yield point,

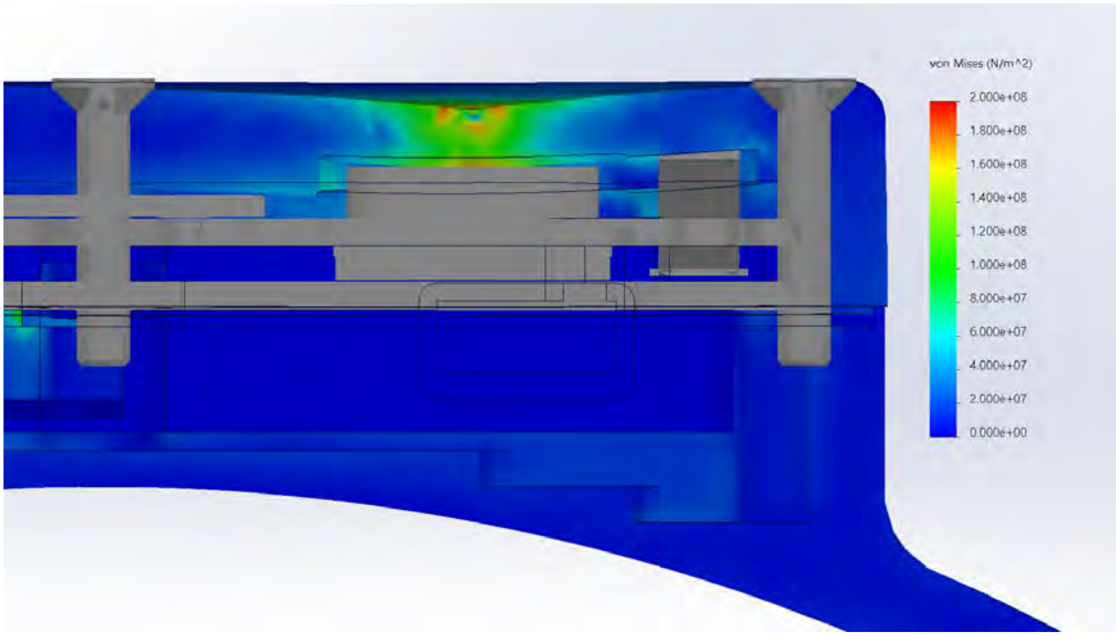


Figure 5.4: Section view of von Mises stress from the dorsal loading simulation (360.1 MPa).

the maximum localized stress of 96 MPa still exceeds the yield strength, indicating that even a sub-maximal bite is likely to cause permanent, localized surface deformities under these conditions. Further longitudinal testing is required to gain insight into the validity of these simulation conditions and to better understand the natural fatigue dynamics.

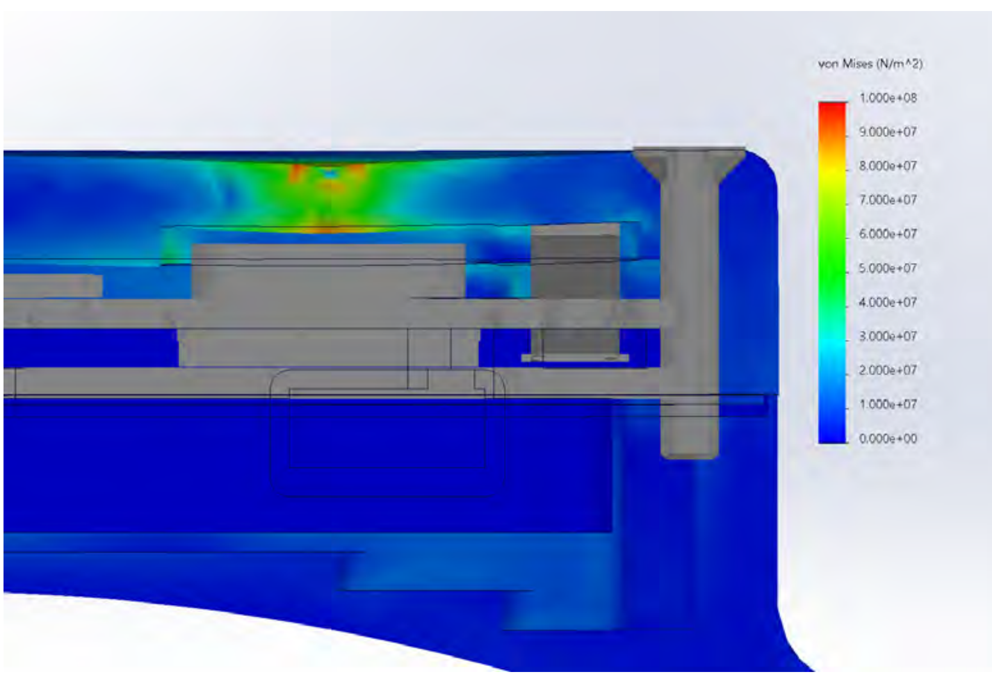


Figure 5.5: Section view of von Mises stress from the reduced dorsal loading simulation (180.1 MPa).

It should be noted that the presented FEA represents a conservative case, as the model did not include the pre-compressive supporting forces from the device's assembly hardware, assumes a fine contact point of 2.5 mm in diameter, and is performed on the structurally weakest points of the design. In the final assembly, the polycarbonate enclosure is secured to an aluminum base via six M3 screws, which uniformly compresses a rubber gasket seated between the two materials. This clamping pressure and the added rigidity from the full assembly would create boundary conditions that provide significant support against deformation. It is therefore likely that this added support would prevent as high of a bending moment along the center of the polycarbonate surface as was shown in the analysis, suggesting the observed deflection values are a conservative overestimation of what would occur in the real-world application.

FEA Simulation of Lateral Bite Force

In addition to the dorsal impact, the enclosure's resilience to a lateral bite was analyzed. A pressure of 360.1 MPa was applied over a 2.5 mm diameter circular area along the center of the longest lateral edge of the polycarbonate housing, between two supporting screws and centered 1 mm above the mating edge of the polycarbonate and aluminum (Figure 5.6). This simulation modeled a direct, worst-case puncture to the side of the device. The analysis showed a maximum deflection of 0.6 mm and a maximum von Mises stress of 198 MPa within the material (Figure 5.7). This observed stress is significantly above both the yield strength (approx. 63 MPa) and ultimate tensile strength (approx. 70 MPa) of polycarbonate. Therefore, local material failure is predicted, and the integrity of the environmental seal provided by the gasket could potentially be compromised under such an extreme load.

To simulate a more realistic contact scenario as in the previous section, 50% of the maximum theoretical force was selected and reapplied at the same lateral location. This refined analysis resulted in a maximum observed deflection of 0.4 mm and a maximum von Mises stress of 83 MPa, localized on the polycarbonate surface near the point of contact (Figure 5.8). While the stresses through the core of the material were significantly lower, this maximum surface stress exceeds the polycarbonate's yield strength (63 MPa). Therefore, permanent deformation (denting) of the enclosure is predicted even under this sub-maximal load. Furthermore, as this stress also surpasses the material's ultimate tensile strength (approx. 70 MPa), localized surface failure is a risk. Although immediate bulk fracture is unlikely, periodic loading at these levels will lead to fatigue and stress cracking over time, warranting further longitudinal testing.

These simulations represent a conservative estimate of performance, as they do not account for the structural support provided by the full device assembly. The pre-compressive force exerted by the six M3 screws pressing the polycarbonate into the rubber gasket and aluminum base would introduce significant boundary constraints. This assembly would add considerable rigidity and resist the lateral bowing shown in the simulation, indicating that the actual deflection and stresses experienced in the field would likely be lower than these calculated worst-case values.

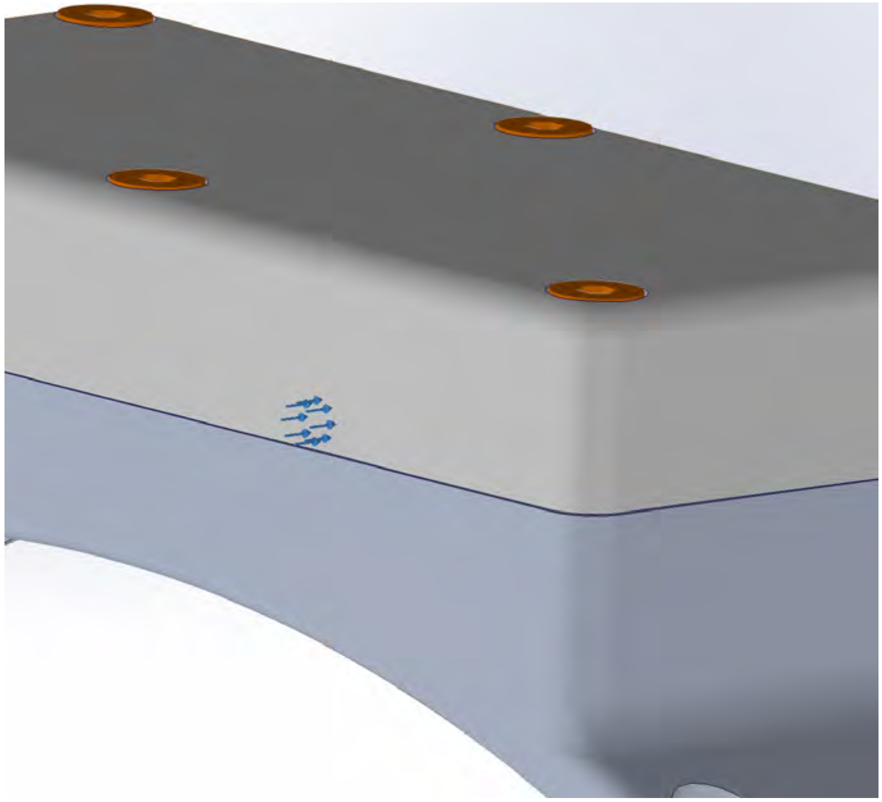


Figure 5.6: Location of the simulated lateral load application on the polycarbonate housing.

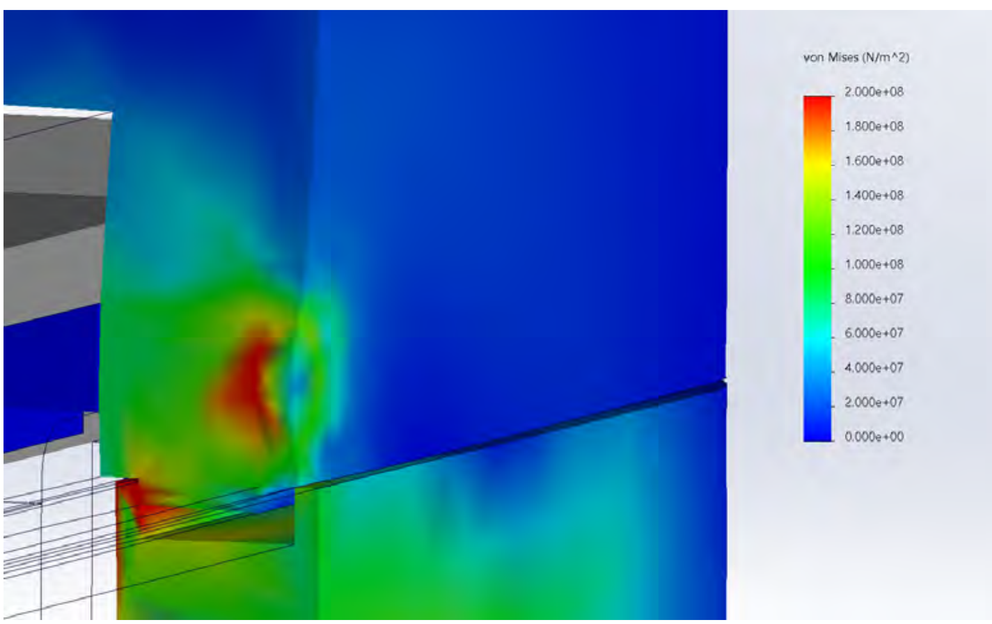


Figure 5.7: FEA results showing von Mises stress from the worst-case lateral loading (360.1 MPa).

Summary of Structural Simulations

The Finite Element Analyses of both dorsal and lateral bite scenarios provide a consistent and robust validation of the CollarID enclosure's structural integrity against the significant

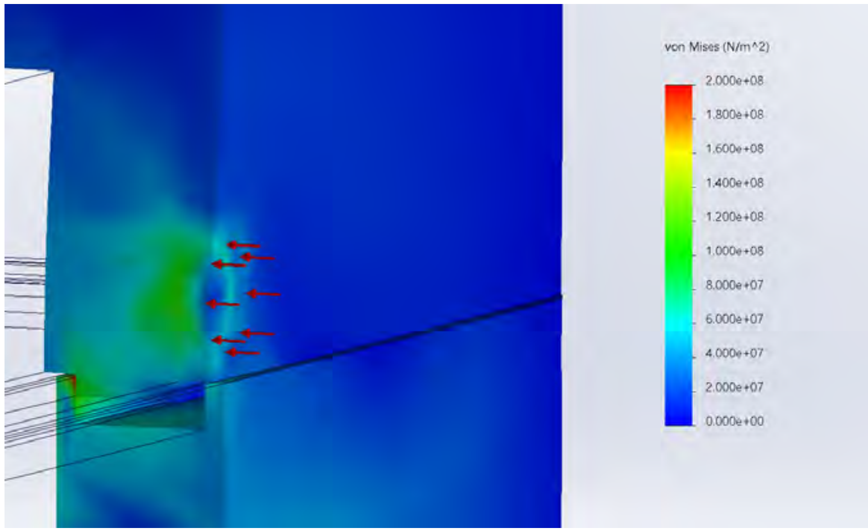


Figure 5.8: FEA results showing von Mises stress from the reduced lateral loading (180.1 MPa).

mechanical stresses it may encounter. In both loading orientations, the simulations predict that the polycarbonate housing will experience localized, permanent deformation (i.e., denting and bite marks) under both worst-case theoretical maximum forces and more probable sub-maximal forces, as the concentrated stresses at the point of contact exceed the material's yield strength.

However, the results also consistently indicate that catastrophic bulk fracture of the enclosure is unlikely. This resilience is attributed to the inherent ductility of polycarbonate and the conservative nature of the simulations, which modeled isolated components and did not account for the significant additional structural support and pre-compressive forces provided by the fully assembled device. While the housing is predicted to survive significant, isolated bite events, the primary long-term concern identified across both analyses is material fatigue. The potential for stress cracking and compromised environmental seals from repeated, lower-force impacts over time highlights the necessity of future empirical testing to fully characterize the enclosure's fatigue life and validate its long-term durability in the field.

Dishwasher Experiment

To validate the environmental sealing and operational resilience of the CollarID enclosure, the fully assembled device was subjected to an accelerated stress test. The unit was placed on the top rack of a standard dishwasher (Figure 5.9) and subjected to a complete cycle using detergent. This procedure exposed the enclosure to harsh conditions, including high temperatures, sustained high humidity, direct water jet spray, and corrosive chemical agents.

Throughout the cycle, internal sensors logged temperature and humidity, while the onboard microphone recorded the acoustic environment. The resulting data, presented in Figure 5.10, showed that internal conditions remained stable. A peak internal temperature of 162°F and humidity of 95% demonstrated the effectiveness of the enclosure's thermal insulation and watertight seal. Furthermore, as shown in Figure 5.11, a spectrogram of the recorded audio clearly revealed distinct acoustic signatures, such as the periodic frequency shifts corresponding to the dishwasher's rotating water jets, confirming the system and its microphone remained fully operational throughout the test.

A post-cycle inspection revealed no water ingress or damage to the internal components, validating the robustness of the enclosure's gasket seal. However, the polycarbonate coating was stripped by the dishwashing detergent. When the experiment was repeated without detergent, the coating remained intact, indicating the failure was due to the chemical agent rather than the heat or moisture.

Figure 5.9: CollarID Placement in Dishwasher

5.3 Core Electronics, Power, and Sensor Subsystems

5.3.1 System Electronics and Power Management

The CollarID system is centered around an ARM Cortex-M33 32-bit STM32U5 Microcontroller (STM32U595) with 4 MB of flash memory. This chip was chosen for its extensive peripherals (e.g., I2C, SPI, USB, SDIO) and its integrated Analog Front End, which allows for the efficient sampling of a PDM microphone (SPH0641LU4H-1). While the chosen microphone is capable of sensing ultrasound, the current focus is on sampling the audible spectrum.

The system's electronics are distributed across three interconnected circuit boards (Figure 5.2). The main board contains the core processing and sensing components, including the microcontroller, SD Card slot, a 3-axis accelerometer (BMA400), a light sensor (APDS-9306-065), a GPS module (SAM-M10Q-00B), and a particulate sensor (BMV080). The APDS-9306-065 is a digital ambient light sensor that measures illuminance in lux across a wide dynamic range, with a spectral response designed to approximate that of the human eye, making it sensitive primarily to visible light. A side board, connected to the main board

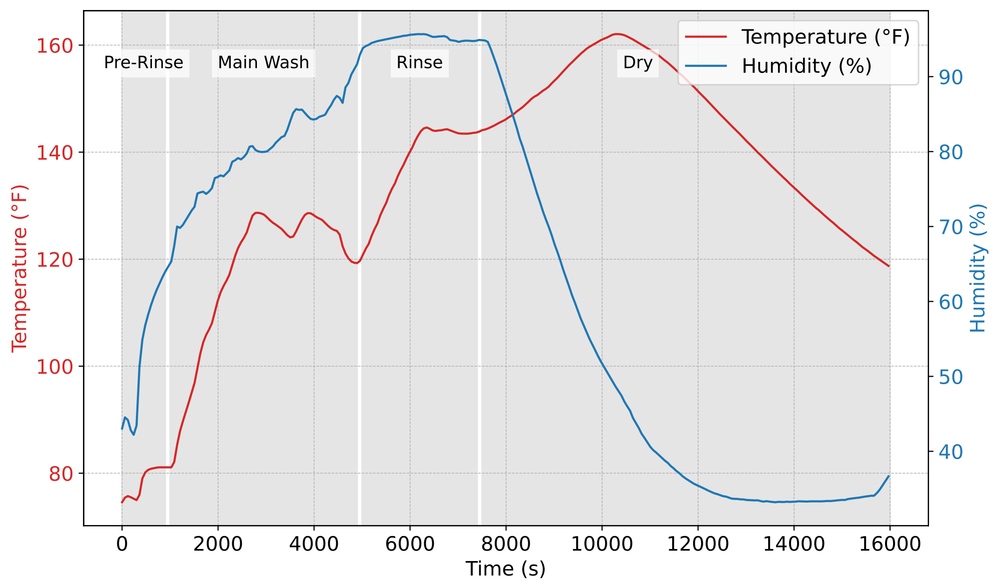


Figure 5.10: CollarID Temperature and Humidity in Dishwasher

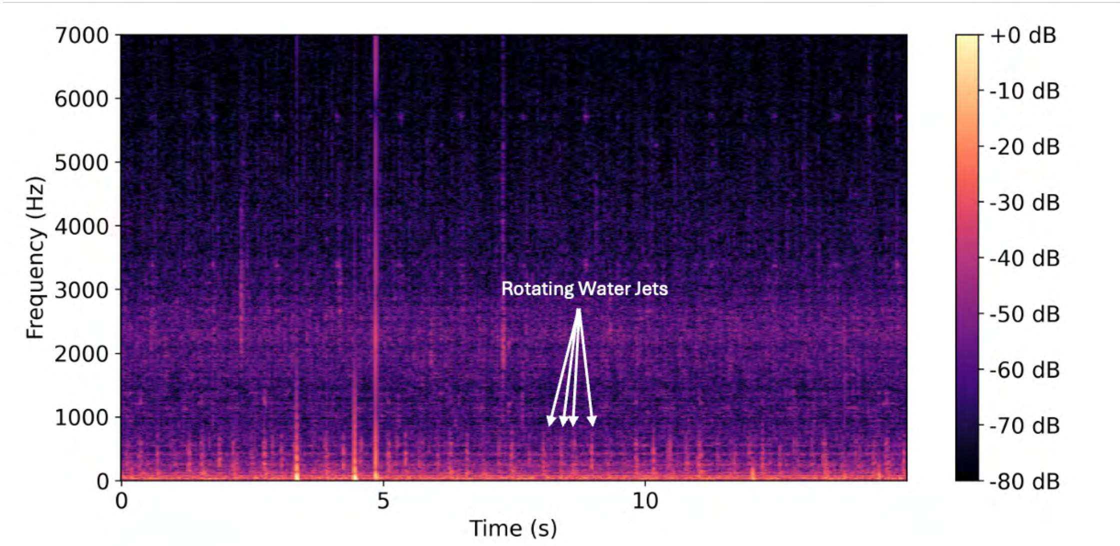


Figure 5.11: CollarID Microphone FFT Excerpt from Dishwasher Acoustic Collection

via a ribbon cable, houses the microphone subsystem and a BME688 environmental sensor capable of measuring temperature, humidity, air pressure, and gas. A top board provides mounting for external components and includes an optional status LED. The top and main boards are secured together with four M2 12 mm socket head screws and two FDM-printed plastic supports.

The system is designed for long-term, autonomous deployment through careful power management and solar energy harvesting. A 154 mW solar panel (SM101K07TF) is integrated into the top board, managed by a BQ25570 solar charger on the main board. The circuitry has been optimized for low power consumption; individual subcomponents are run on the lowest possible bias voltage, and each subsystem is powered via separate regulators. This allows for subsystems to be duty-cycled to save power or if a failure is detected. To minimize power draw during data logging, the microcontroller interfaces with the SD card via SDIO, which parallelizes data transfer and decreases the time the SD card must be powered on. In the event the battery (2.96 Wh, lithium-polymer) drains to critical levels, the system enters a hibernation state, shutting down all non-essential components and consuming only 118 μW until the battery has recharged beyond 30%. This fault-tolerant, autonomous power management strategy is a direct evolution of the need for long-term, unattended operation first addressed with the solar-assisted BuzzCam deployments, maturing the concept into a fully recoverable system for even more demanding, multi-year wildlife studies. With the given load during hibernation, the system can remain in that state for over 2.5 years, assuming a battery self-discharge rate of 3%.

5.3.2 Power Consumption Characterization

To estimate the field longevity of the CollarID, it is essential to create a detailed power budget. The total average power consumption of the system is a sum of the power consumed by its individual components, taking into account that many sensors are duty-cycled (i.e., periodically turned on and off) to conserve energy. The power draw of each key subsystem was measured empirically and is presented in Table 5.1. It is important to note that all power readings in Table 5.1 already include the additional power consumed when writing sensor data to the SD card, providing a realistic measure of in-field operational power.

The total average power consumption can be calculated using Equation 5.1, which models the contribution of each subsystem based on its specific operational parameters.

Table 5.1: Power Consumption Estimates with Variable References

Component	**Variable Name**	**Power (mW)**
Quiescent System (3 MHz)	P_{base}	2.8
Accelerometer (50 Hz) and Lux (5 m interval)	$P_{\text{accel}} + P_{\text{lux}}$	1.4
Environmental Sensor (5 m interval)	P_{env}	0.3
Microphone (16 kHz)	P_{mic}	1.9
GPS Acquisition	$P_{\text{GPS,acq}}$	53.6
GPS Standby Mode	$P_{\text{GPS,standby}}$	0.4
LoRaWAN (16 dBm TX power)	P_{LoRa}	310.0
Particulate Sensor during Integration	$P_{\text{part,on}}$	87.0
Particulate Standby Mode	$P_{\text{part,standby}}$	0.1

Table 5.2: LoRaWAN Transmission Duration Estimates (EU868, 16 dBm TX Power)

Data Rate	**Packet Syze (byte)**	**Approx. Transmission Time (s)**
DR4	220	0.30
DR3	110	0.68
DR2	50	0.65
DR1	50	1.52

$$\begin{aligned}\text{Total Power Consumption} = {} & P_{\text{base}} \\ & + P_{\text{accel}} \\ & + P_{\text{lux}} \\ & + P_{\text{env}} \\ & + P_{\text{GPS,acq}} \cdot \frac{T_{\text{GPS,acq}}}{T_{\text{GPS,period}}} \\ & + P_{\text{GPS,standby}} \cdot \frac{T_{\text{GPS,period}} - T_{\text{GPS,acq}}}{T_{\text{GPS,period}}} \\ & + P_{\text{LoRa}} \cdot \frac{T_{\text{LoRa}}}{T_{\text{LoRa,period}}} \\ & + P_{\text{part,on}} \cdot \frac{T_{\text{int}}}{T_{\text{part,period}}} \\ & + P_{\text{part,standby}} \cdot \frac{T_{\text{part,period}} - T_{\text{int}}}{T_{\text{part,period}}} \\ & + P_{\text{mic}} \cdot \frac{T_{\text{mic}}}{T_{\text{mic,period}}}\end{aligned} \tag{5.1}$$

Equation 5.1 calculates the average power by summing the system's base power (P_{base}) with the power consumed by its various subsystems, which are grouped into two categories for this analysis (see Table 5.1).

The first group consists of sensors that operate at fixed, predefined settings, and their power consumption is treated as a constant average. For instance, the accelerometer's sample rate is set to 50 Hz, a standard rate frequently used in on-animal systems. The low-power lux sensor samples infrequently (every 5 minutes) simply to determine broad light conditions (e.g., direct daylight vs. shade); its consumption is therefore coupled with the accelerometer's in the power budget. Similarly, the environmental sensor is sampled only once every five minutes, as ambient conditions like temperature and humidity are typically slow to change.

The second group includes components with variable, scenario-dependent duty cycles, such as the GPS, LoRaWAN radio, particulate sensor, and microphone. For these components, the power consumption is weighted by the fraction of time each is active, allowing the total power draw to be modeled for different deployment configurations.

Example Calculation

Let's calculate the total average power consumption for a specific, hypothetical deployment scenario.

Scenario Parameters:

- **GPS:** The average GPS acquisition time, $T_{\text{GPS,acq}}$, was experimentally determined to be 5 seconds in an urban collection environment when starting the GPS from a warm start. We will set the GPS sampling period, $T_{\text{GPS,period}}$, to be 5 minutes, meaning the GPS attempts a new fix every 5 minutes.
- **LoRaWAN:** We select Data Rate 1 (DR1) for its long-range performance. Let's assume a transmission period, $T_{\text{LoRa,period}}$, of 60 minutes and a time-on-air, T_{LoRa}, of 1.52 seconds for a 50-byte packet (Table 5.2).
- **Particulate Sensor:** The integration time, T_{int}, is set to the recommended 10 seconds based on the Bosch API. However, we empirically found that when the sensor is configured with this integration time, it consumes power for only $\sim$5 ms before returning to its standby state. With this in mind, we further define the sampling period, $T_{\text{part,period}}$, to be 30 minutes.
- **Microphone:** The microphone is active for $T_{\text{mic}} = 2$ minutes, with a period of $T_{\text{mic,period}} = 4$ minutes.

Calculation:

1. **Base Power:** $P_{\text{base}} = 2.8$ mW
2. **Accelerometer, Lux, and Environmental:** $P_{\text{accel}} + P_{\text{lux}} + P_{\text{env}} = 1.7$ mW
3. **GPS Power:** $P_{\text{GPS}} = (53.6 \text{ mW} \cdot \frac{5 \text{ s}}{300 \text{ s}}) + (0.4 \text{ mW} \cdot \frac{300 \text{ s} - 5 \text{ s}}{300 \text{ s}}) \approx 0.89 \text{ mW} + 0.39 \text{ mW} = 1.28 \text{ mW}$
4. **LoRa Power:** $P_{\text{LoRa}} = 310.0 \text{ mW} \cdot \frac{1.52 \text{ s}}{60 \times 60 \text{ s}} \approx 0.13 \text{ mW}$
5. **Particulate Sensor Power:** $P_{\text{particulate}} = (87.0 \text{ mW} \cdot \frac{0.005 \text{ s}}{30 \times 60 \text{ s}}) + (0.1 \text{ mW} \cdot \frac{30 \times 60 \text{ s} - 0.005 \text{ s}}{30 \times 60 \text{ s}}) \approx 0.10 \text{ mW}$
6. **Microphone Power:** $P_{\text{mic}} = 1.9 \text{ mW} \cdot \frac{2 \text{ min}}{4 \text{ min}} = 0.95 \text{ mW}$
7. **Total Average Power:** $P_{\text{total}} = 2.8 + 1.7 + 1.28 + 0.13 + 0.10 + 0.95 = \mathbf{6.96\ mW}$

In this example configuration, the total average power consumption is approximately 7 mW. The base system power is now the dominant power draw but it can be further reduced by leveraging lower power states of the ARM processor. Given the 2.96 Wh battery, this configuration would yield approximately 425 hours, or over 17 days, of operational life, excluding any input from solar. This demonstrates the effectiveness of a warm-start GPS strategy. By sampling the GPS at a period of 5 minutes, well within the 4-hour window required to maintain valid telemetry data, the receiver can acquire a fix much more quickly than the approximate 30 seconds required for a cold start. This ability to rapidly reacquire a fix, where possible, dramatically reduces the GPS duty cycle and extends deployment longevity.

5.3.3 Environmental Sensor Performance Validation

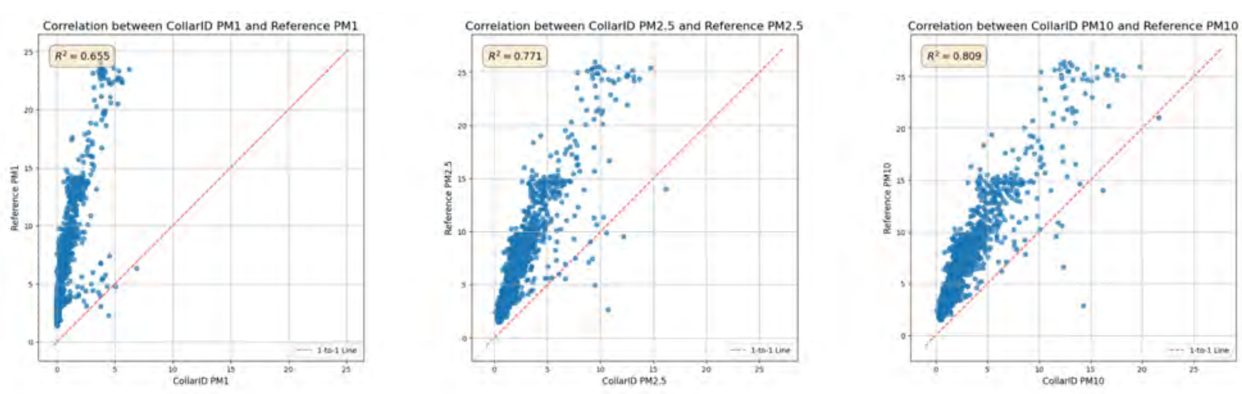
To validate the performance of the environmental sensors integrated into the CollarID system, a co-location experiment was conducted.

Experimental Setup

The experiment was performed over a five-day period, from June 9, 2025, to June 13, 2025, within a residential apartment kitchen in Somerville, MA. The CollarID system, which utilizes a BMV080 particulate sensor and a BME688 environmental sensor, was co-located with reference instruments. A reference sensor package featuring a Sensirion SPS30 was used for particulate matter, temperature, and humidity comparison, while a Bosch BME280 was used for pressure. The reference instruments sampled once per second, while the CollarID recorded data once every 15 seconds. The co-location ensures that all instruments were exposed to the same ambient atmospheric conditions, allowing for a direct comparison of their measurements.

Particulate Sensor Correlation Analysis

To quantify the agreement between the particulate sensors, the collected data was synchronized by resampling each time series into five-minute averages. This averaging window is sufficiently long to ensure that data points from both sensors are captured within each interval, while still being short enough to preserve significant variations in air quality. The correlation between the two instruments was then evaluated for PM1, PM2.5, and PM10 particle sizes. The results are presented as scatter plots in Figure 5.12.



(a) PM1 Correlation (b) PM2.5 Correlation (c) PM10 Correlation

Figure 5.12: Correlation plots for (a) PM1, (b) PM2.5, and (c) PM10 concentrations based on 5-minute averaged data. The red dashed line indicates a perfect 1-to-1 agreement.

As illustrated, there is a positive linear correlation between the CollarID and the reference sensor across all three particle size fractions. The strength of this correlation, indicated by the coefficient of determination (R^2), improves with increasing particle size: PM10 ($R^2 = 0.809$), PM2.5 ($R^2 = 0.771$), and PM1 ($R^2 = 0.655$). This variance indicates the CollarID sensor most reliably tracks relative changes in PM10 and PM2.5.

Across all three plots, the data points consistently fall above the 1-to-1 line, revealing a systematic negative bias where the CollarID sensor consistently reports lower concentration values than the reference instrument. This consistent bias suggests that a linear correction factor could be developed to calibrate the CollarID's output if absolute accuracy is required. For this study, the high degree of correlation validates the CollarID sensor as a reliable tool for tracking relative changes in air quality.

Case Study: Tracking Dynamic Indoor Air Quality Events

To assess the real-world performance of the optical particulate sensor when housed within its protective enclosure, the CollarID was tested with its custom polycarbonate cover. This cover was specifically designed according to the BMV080 particulate sensor's official design guidelines to ensure minimal interference with its optical measurements. A detailed analysis was conducted on data from June 10, 2025. Figure 5.13 presents the time-series data for PM2.5 concentrations from this day. The plot clearly shows two major pollution events: a sharp peak at 12:00 corresponding to a high-temperature cooking event, and a wider peak beginning at 14:00 corresponding to the cessation of an air conditioning unit and an air purifier. In both events, the CollarID sensor successfully tracked the changes in air quality in close concert with the reference instrument, demonstrating its efficacy for monitoring dynamic environmental changes even when operating behind its protective window.

Temperature, Humidity, and Pressure Correlation

The CollarID's capability to accurately measure other key environmental parameters was validated against the co-located reference sensors. The correlation plots for these three parameters are shown in Figure 5.14. The analysis reveals a high degree of correlation across all three measurements:

- **Temperature:** Exceptionally strong correlation ($R^2 = 0.960$), with a consistent offset where CollarID reports slightly higher values.
- **Humidity:** Strong correlation ($R^2 = 0.918$) with less systematic bias.
- **Pressure:** Outstanding correlation ($R^2 = 1.000$) with no meaningful bias or error.

These results confirm that the CollarID provides a robust and reliable suite of environmental sensors. It is particularly noteworthy that this high level of performance is achieved even with its sensors positioned behind a protective membrane, demonstrating that the enclosure successfully balances environmental sealing with high-fidelity sensing.

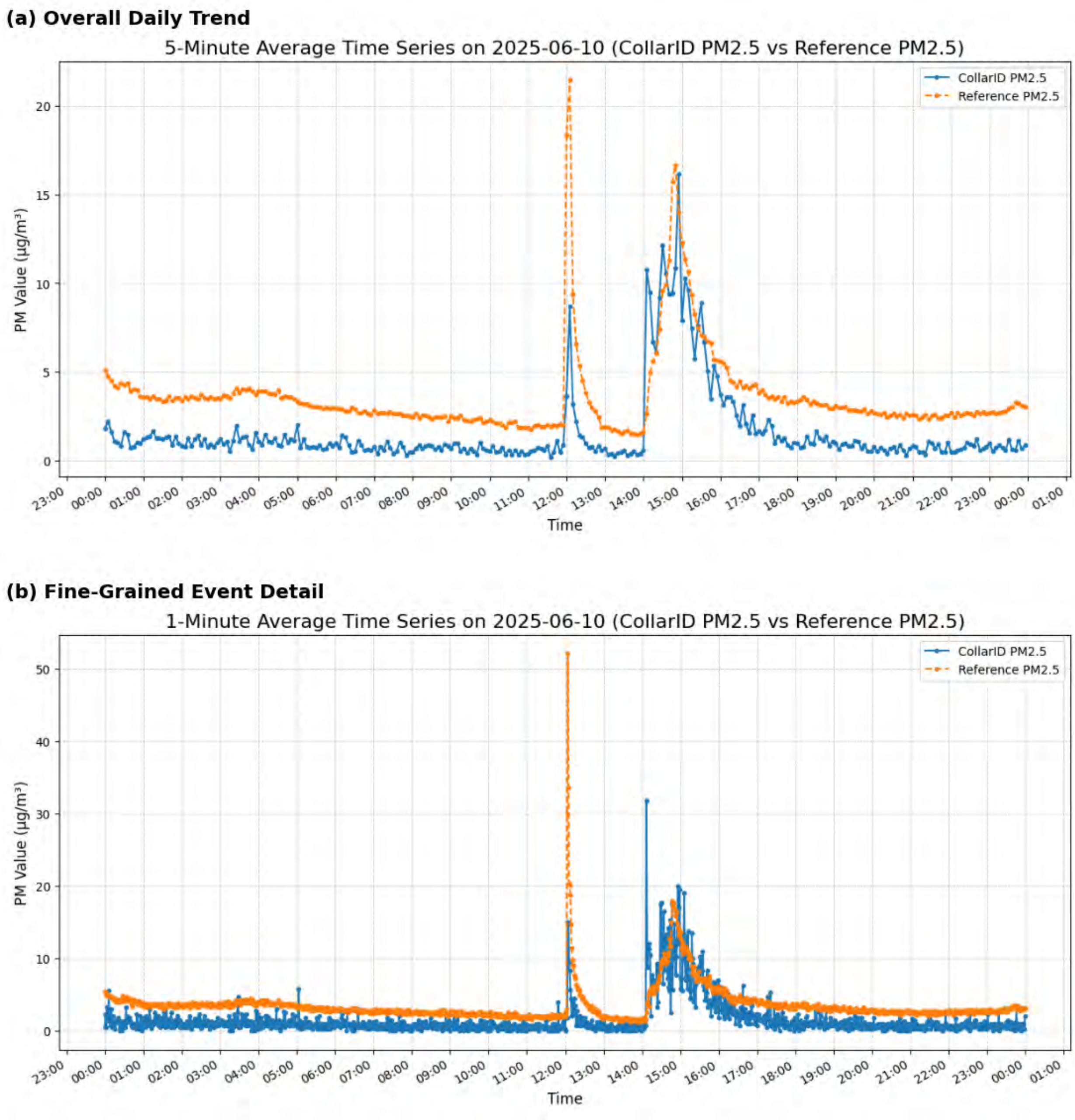


Figure 5.13: Time-series comparison of CollarID and reference sensor PM2.5 readings on June 10, 2025. Plot (a) shows data smoothed with a 5-minute average, highlighting the overall daily trend. Plot (b) uses a 1-minute average, revealing finer details and more rapid fluctuations.

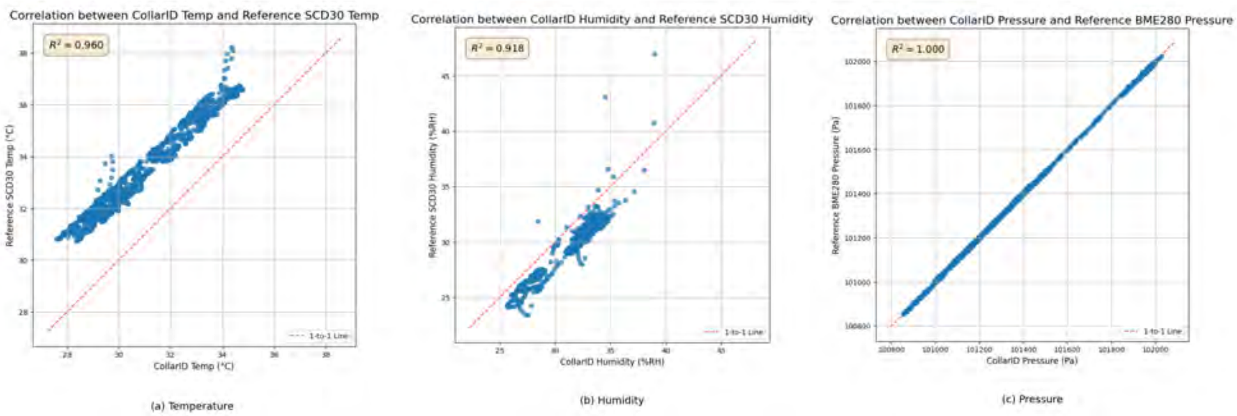


Figure 5.14: Correlation plots for (a) temperature, (b) humidity, and (c) pressure concentrations based on 5-minute averaged data. The red dashed line indicates a perfect 1-to-1 agreement.

5.4 Communications Subsystem Design and Validation

5.4.1 Long-Range Communication Design

For remote data transmission and status updates, the CollarID incorporates a low-power, long-range communication subsystem. The main board features an SX1262 radio module capable of operating in the 868 MHz and 915 MHz bands. The radio is supported by a custom RF matching network on the main board, which connects to either an ANT-868-USP410 or ANT-915-USP410 antenna mounted on the top board. The choice of antenna is dependent on the geographical region of deployment and local radio regulations. An additional RF matching network is included on the top board to ensure optimal antenna performance. CollarID supports LoRaWAN, interfaces with any LoRaWAN gateway, and has been configured to support gateways from SmartParks, a nonprofit which has been installing gateways in Botswana to aid in ecological research.

5.4.2 Field Performance Validation

The CollarID radio system was tested at the Mass Audubon Tidmarsh Wildlife Sanctuary, a 481-acre nature preserve in Plymouth, MA (Figure 5.15). For the test, an 8 dBi, 915 MHz omnidirectional antenna was mounted 10 m high on a barn on the property (Figure 5.16). This antenna was interfaced to a Seeed Studio SenseCAP M2 (US915) gateway, which connected to a custom ChirpStack LoRaWAN network server.

Two CollarID devices were used for range testing, one configured for a lower power transmission of 14 dBm and the other for a higher power of 23 dBm. Each device was configured to acquire a valid GPS fix and then transmit a data packet containing the coordinates every 10 seconds. The gateway appended Received Signal Strength Indicator (RSSI) and Signal-to-Noise Ratio (SNR) metadata to each received packet. To collect range data, each device was carried along several walking trails throughout the sanctuary (Figure 5.17), representing both open-field and densely forested environments. The devices cycled through LoRaWAN Data Rates (DR) 1 through 4 (Table 5.3).

Figure 5.15: Mass Audubon Tidmarsh Wildlife Sanctuary

Figure 5.16: Antenna mounted at the Mass Audubon Tidmarsh Wildlife Sanctuary.

Table 5.3: US915 LoRaWAN Data Rates with Spreading Factor, Bandwidth, Minimum SNR, and Receiver Sensitivity

DR	Spreading Factor (SF)	Bandwidth	Min SNR (dB)	Receiver Sensitivity (dBm)
DR0	SF10	125 kHz	-20.0	-121
DR1	SF9	125 kHz	-17.5	-118.5
DR2	SF8	125 kHz	-15.0	-116
DR3	SF7	125 kHz	-12.5	-113.5
DR4	SF8	500 kHz	-5.0	-106

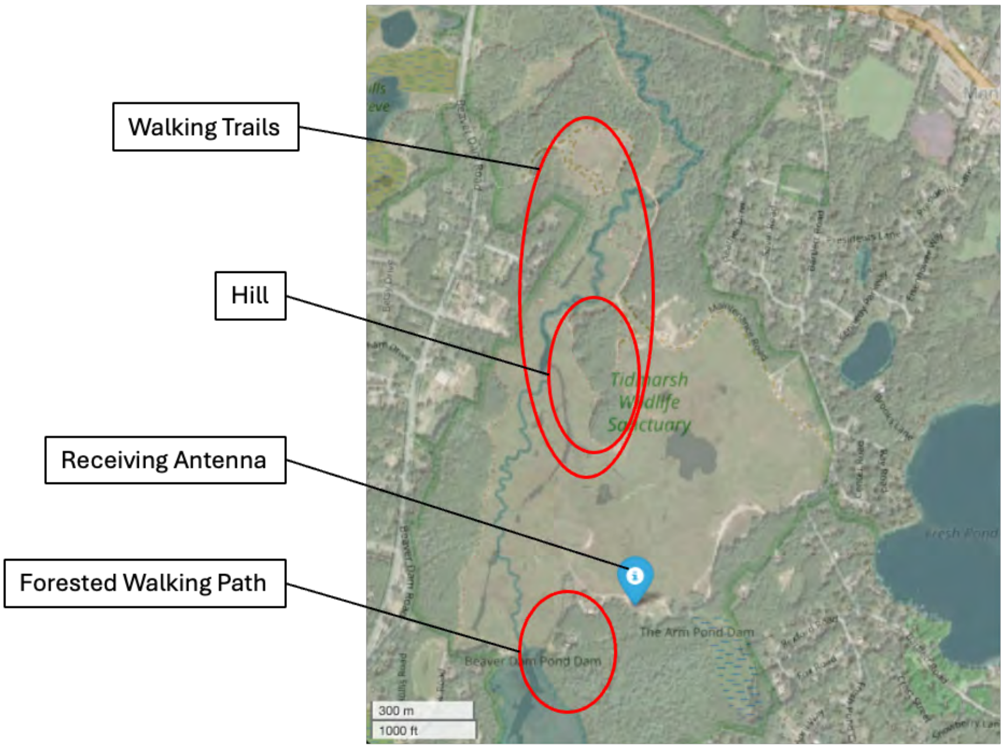


Figure 5.17: Tidmarsh Test Site.

Impact of Transmit Power and Data Rate on Signal Quality (SNR)

As anticipated, both transmit power and data rate had a significant impact on the communication link quality (Figures 5.18 and 5.4.2). At 14dBm, the robust DR1 maintained a reliable link, while the high-speed DR4 was marginal. Increasing the transmit power to 23dBm dramatically improved the link quality for all data rates, making the DR4 link highly viable. This demonstrates that higher transmit power can effectively compensate for the reduced robustness of faster data rates.

Impact of Transmit Power on Signal Strength (RSSI)

The RSSI data (Figure 5.20) showed a clear improvement with increased transmit power. In Figure 5.20, data rates (DR1 through DR4) at each power level were combined to show all received packets at that respective transmit power. At both power levels, we were able to receive packets from as far as 5000 ft (1524 m) under non-line-of-sight conditions. We did

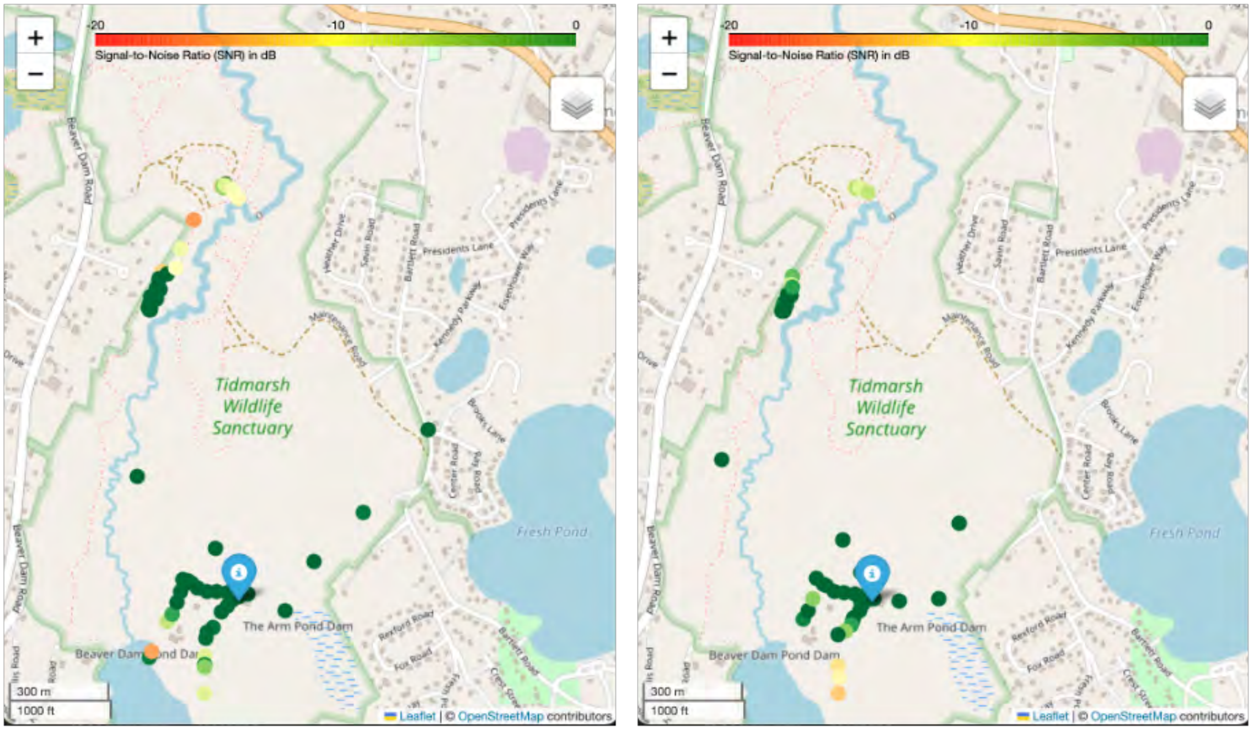


Figure 5.18: Spatial plot of Signal-to-Noise Ratio (SNR) for a low-power (14 dBm) transmission. The color of each point represents the quality of the received signal, with green indicating a strong, clear signal and yellow/orange indicating a weaker one. The left panel shows the robust, long-range Data Rate 1 (DR1), which maintains a viable signal quality across the test path. The right panel shows the faster, shorter-range Data Rate 4 (DR4), which struggles to maintain a consistently strong link at this lower power level, as indicated by the more frequent orange points.

not attempt to exceed this range during testing, so further work is warranted to estimate real-world range—particularly under line-of-sight conditions, which were not available at this site.

Discussion of Propagation Anomaly in Forested Environment

An interesting and counter-intuitive observation was made on the southern forested path (Figure 5.21). In this specific area, the lower-power 14dBm transmitter demonstrated higher transmission reliability (more packets successfully received) than the higher-power 23dBm transmitter. This phenomenon, where a stronger signal results in a worse link quality, is a classic indicator of complex multipath fading effects, which are common in forested environments.

Rather than the receiver being overloaded by the stronger signal, it is more likely that the dense foliage and terrain created a complex web of signal reflections. These reflected signals travel different distances to the receiver, arriving out of phase with the direct signal. At certain locations, the higher power of the 23dBm signal may have resulted in stronger reflected signals that caused significant destructive interference, creating deep nulls in the

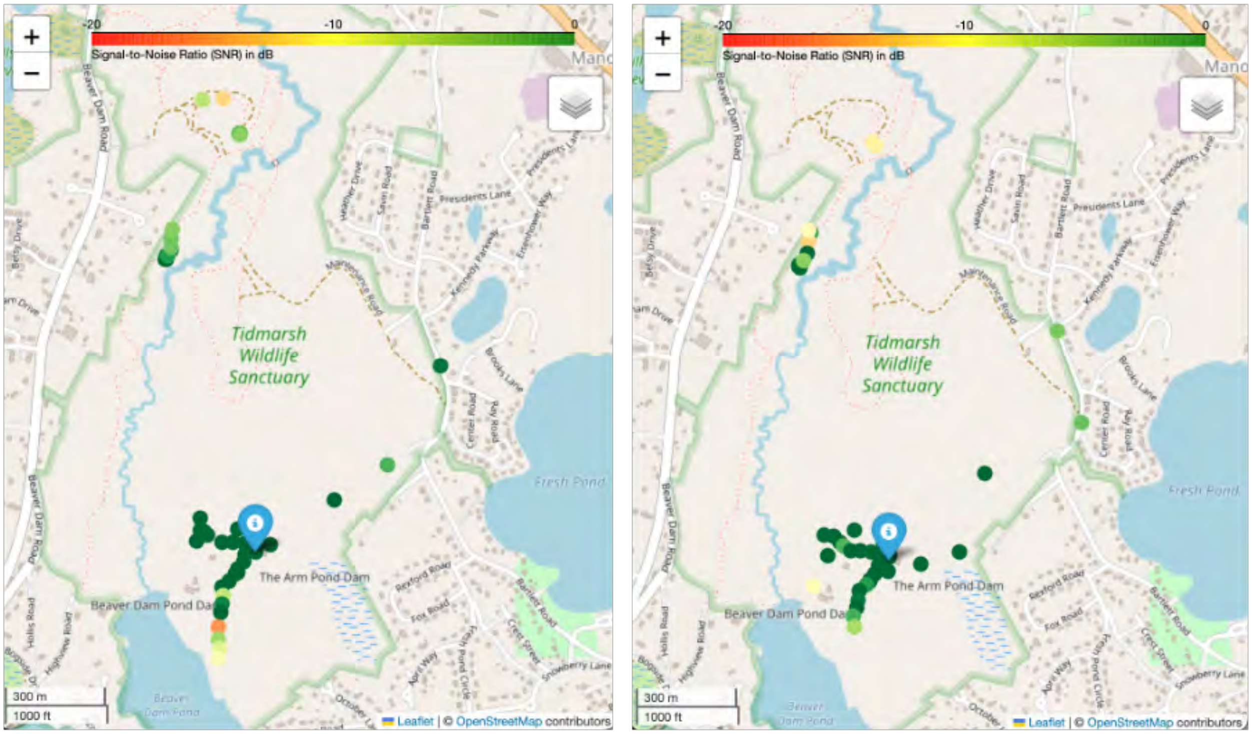


Figure 5.19: Spatial plot of Signal-to-Noise Ratio (SNR) for a high-power (23 dBm) transmission. The color of each point represents the quality of the received signal, with green indicating a strong, clear signal and yellow/orange indicating a weaker one. Compared to the low-power test (Figure 5.18), increasing the transmit power improves the link quality for both the robust DR1 (left) and the faster DR4 (right).

received signal strength and corrupting the data packets. The weaker 14dBm signal, while also subject to multipath, would have had less powerful reflections, potentially avoiding the most severe destructive interference and thus maintaining a more reliable, albeit weaker, link. This highlights a critical real-world consideration in RF network planning in complex environments: maximizing transmit power is not always the optimal strategy, and link quality can be highly dependent on the specific propagation environment and the resulting multipath phenomena. Overall, these field trials confirmed the LoRaWAN subsystem's capability for reliable data transmission up to 1.5 km under challenging non-line-of-sight conditions, validating its suitability for remote wildlife monitoring applications.

5.5 Integrated System Field Trials

To validate the CollarID system's performance and durability in a real-world agricultural setting, a field trial was established in partnership with Northaven Pastures, a family-owned farm in Red Hook, NY. The farm focuses on regenerative agriculture and rotational grazing, where cattle are continuously exposed to outdoor environmental conditions. This provided an ideal test case for the system's resilience. Following a similar setup to previous deployments, a LoRaWAN antenna and a SenseCAP M2 gateway were installed on the farm's main barn

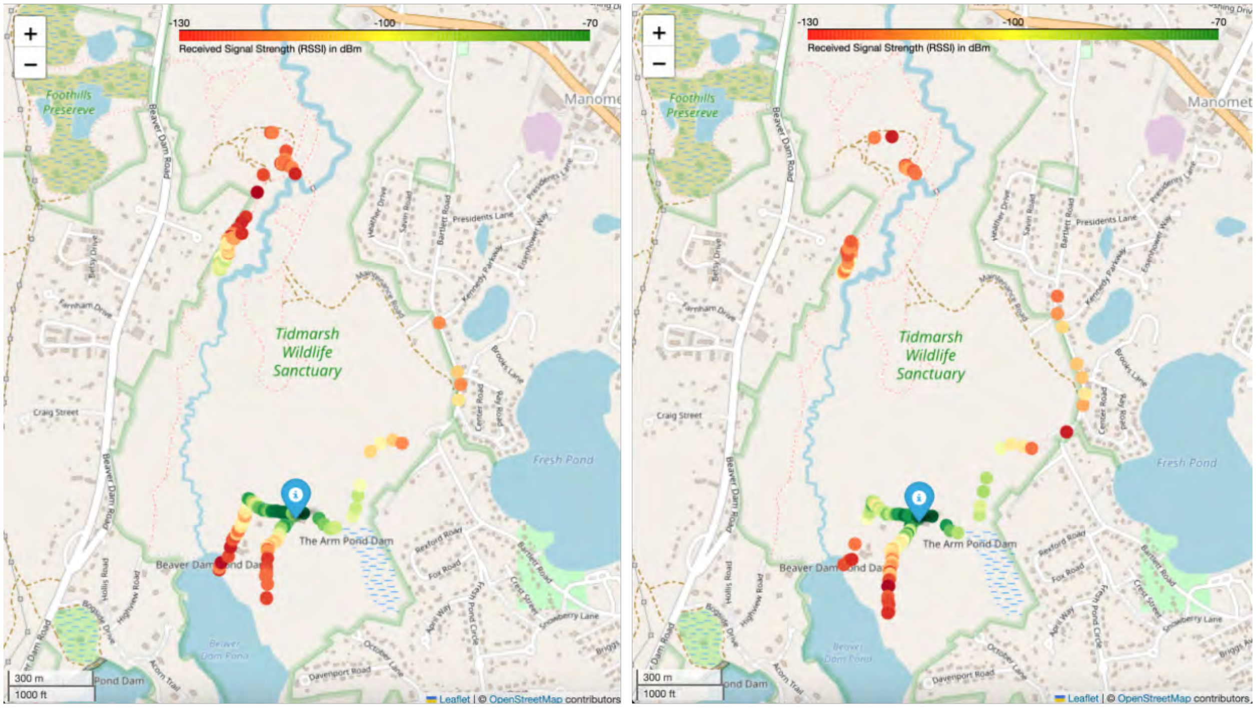


Figure 5.20: Spatial plot of Received Signal Strength Indicator (RSSI), a measure of the raw signal power received at the gateway. This figure combines all received packets across all tested data rates (DR1-DR4) to show the overall signal strength. The color scale indicates signal power, with red being the strongest. The left panel shows the results for a low-power (14 dBm) transmission, while the right panel shows the high-power (23 dBm) transmission.

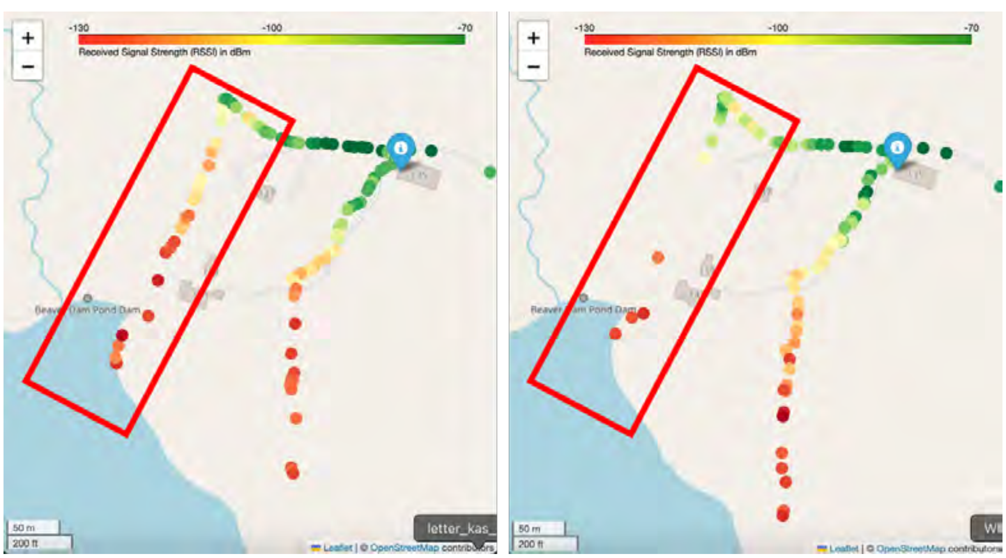


Figure 5.21: RSSI for forested path for 14 dBm (left) and 23 dBm (right) transmit power.

to ensure reliable data transmission from the collars.

A CollarID unit was fitted onto a cow named Mayra on June 6th, 2025, as shown in Figure 5.22. The system is configured to acquire and transmit a GPS fix and a suite of environmental readings every five minutes. Concurrently, the onboard microphone samples and stores two minutes of audio data every four minutes to a 2 TB Kioxia SD card for local retrieval and analysis. This initial deployment does not include a particulate sensor, as the API for the BMV080 is still under evaluation for robustness; however, future versions deployed at the farm will incorporate this sensor.

A preliminary analysis of the high-fidelity audio data reveals a rich acoustic tapestry of the farm environment (Figure 5.23). The onboard microphone successfully captured not only the cow's own vocalizations, such as bawling, but also subtle behavioral sounds like the rhythm of its walking. Furthermore, the system recorded key environmental and anthropogenic sounds, including the chirping of nearby birds, the sound of a passing airplane, and the engine noise of an all-terrain vehicle. This initial analysis demonstrates the platform's capability to concurrently gather bioacoustic data directly from the subject animal and ecoacoustic data from its surrounding habitat, providing a strong foundation for future automated behavioral classification and environmental context analysis.

Figure 5.22: Mayra wearing a CollarID at Northaven Pastures.

The collar has streamed data consistently since its deployment, demonstrating the system's overall reliability. A notable exception occurred between 22:00 on June 12th and 14:00 on June 13th, where the device was offline for a 16-hour period. This event triggered the collar's autonomous power-saving protocol. The system is designed to enter an 8-hour hibernation period upon battery depletion. If the battery has not adequately recharged via its solar panels after this period, it enters another 8-hour cycle. In this instance, the system

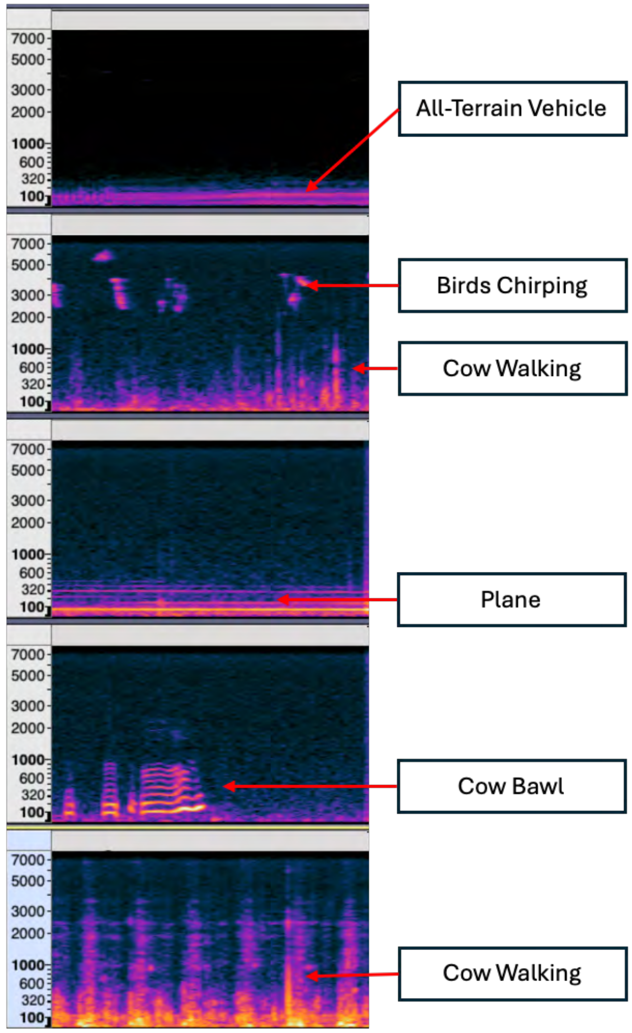


Figure 5.23: Example spectrograms from the CollarID audio data collected at Northaven Pastures, showcasing a variety of identified acoustic events including animal vocalizations, movement, and ambient environmental sounds.

successfully recovered and resumed normal operation after two hibernation cycles, validating the effectiveness of the fault-tolerant power management system.

To facilitate remote monitoring for both the research team and our farm partners, a custom web dashboard was developed (Figure 5.24). This dashboard provides a real-time view of all incoming collar data, including the animal's current location on a map, battery status, and time-series plots for various environmental sensors. Occasional sharp drops to zero in the sensor data, as seen in the figure, are visualization artifacts resulting from lost LoRa packets during transmission and do not represent actual environmental measurements. The data handling for the dashboard has since been updated to better manage these missing data points in future collections. Users can configure the time windows for data display, and the dashboard features an automatic refresh function to ensure the latest data is always visible.

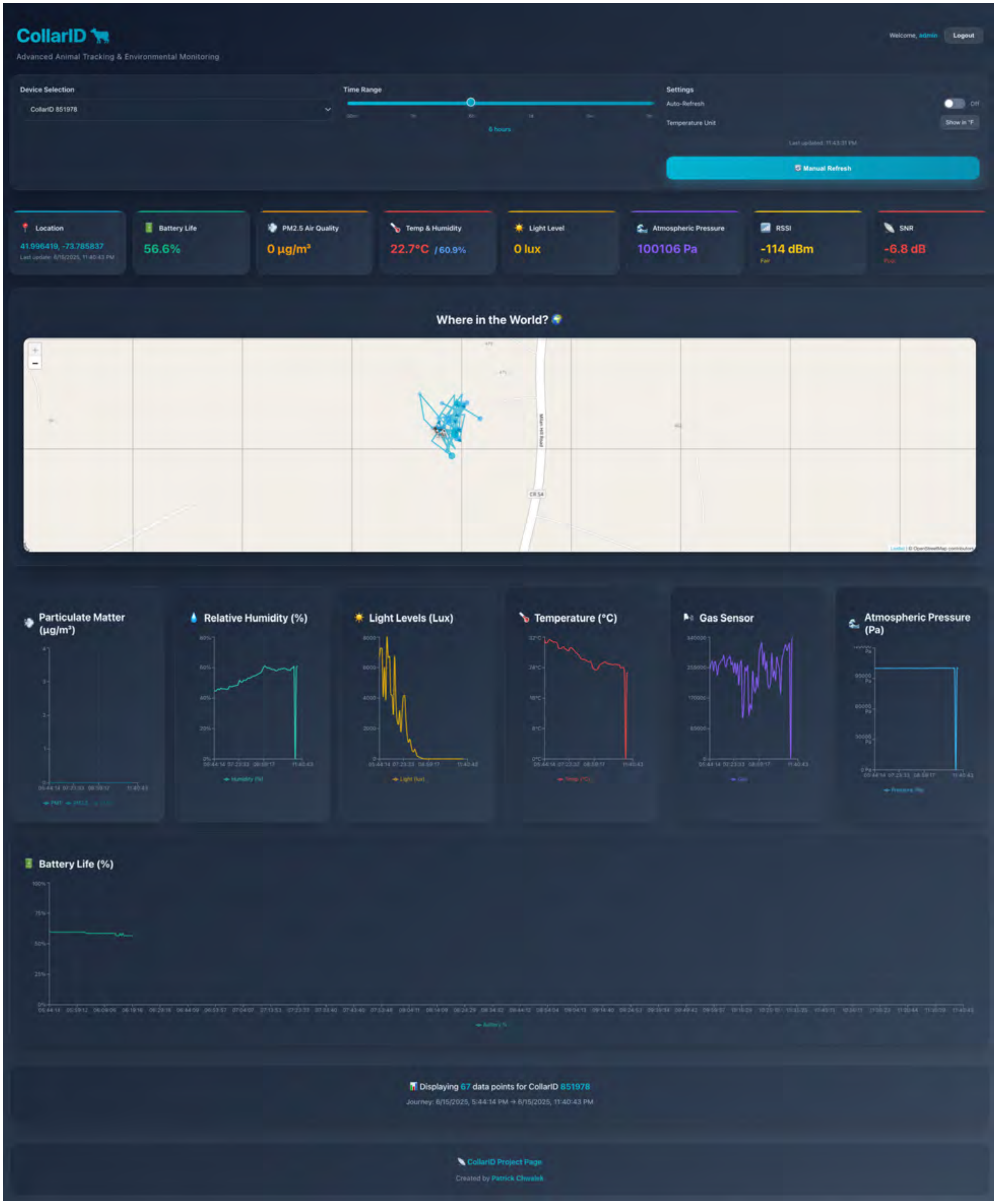


Figure 5.24: CollarID Web Dashboard

5.6 Discussion

The engineering and validation process detailed in this chapter demonstrates that the CollarID prototype is a robust, capable, and field-ready platform for advanced wildlife monitoring. The series of validation tests confirm the successful integration of its core subsystems and its resilience to environmental and mechanical stressors. The FEA simulations, grounded in biomechanical literature, established that the mechanical housing can withstand extreme, worst-case bite forces, while the dishwasher experiment proved its environmental sealing and operational durability in harsh, wet conditions. Furthermore, co-location experiments validated the accuracy of the environmental sensor suite, and extensive field testing confirmed the reliability of the long-range communication system over variable terrain. The initial successful deployment on a cow at Northaven Pastures serves as a capstone validation, proving the integrated system's reliability, autonomous power management, and data streaming capabilities in a real-world setting.

The primary novelty of the CollarID platform lies in its holistic, multi-modal sensing capabilities, which stand in contrast to many existing commercial and research platforms. By integrating high-fidelity bioacoustic recording, detailed inertial measurement, GPS tracking, and a comprehensive environmental suite, CollarID is designed to provide a much richer, more contextualized view of an animal's life. The inclusion of a particulate matter sensor, in particular, is not just an incremental addition but a feature that opens up entirely new and critical avenues of ecological inquiry. For decades, wildlife ecology has focused on visible landscape changes, but the threat of atmospheric pollutants has been largely invisible. In an era of increasing wildfire frequency and anthropogenic haze, the ability to directly measure an animal's real-time exposure to airborne particulates is transformative. It allows researchers, for the first time, to move from broad correlational studies to asking direct, mechanistic questions: How does acute smoke exposure during a wildfire affect the foraging efficiency or respiratory health of a lion? Do hyenas alter their denning site selection to avoid areas with consistently poor air quality from nearby human settlements? By co-locating this novel environmental data stream with high-resolution behavioral and physiological metrics, CollarID provides the tool needed to begin answering these urgent questions, bridging the gap between atmospheric science and large mammal ecology. This integrated approach also enables researchers to correlate vocalization patterns with specific activities and fine-scale microclimatic conditions and creates the potential for the platform to function as an early alert mechanism for animals in distress.

However, as a research prototype, it is important to acknowledge the platform's current limitations. The sensor validation, while rigorous, was conducted in controlled residential and rural farm settings; further characterization in dusty, highly dynamic outdoor environments is a necessary next step. Finally, and most importantly, this chapter has focused on the engineering validation of the platform. While this establishes its readiness, full-scale, long-term deployments on the target wildlife species (e.g., lions, hyenas) are the critical next phase of research needed to generate the ecological datasets the platform was designed to collect, and these are currently planned to start in August 2025.

5.7 Conclusion

This chapter has detailed the rigorous engineering and comprehensive validation of the CollarID platform, a novel, multi-modal sensor package designed for diverse wildlife monitoring. Through mechanical stress testing, environmental sensor co-location, long-range communication field trials, and an initial agricultural deployment, this work has successfully demonstrated the prototype's readiness for in-situ ecological research.

The development of CollarID represents the synthesis of principles and skills honed throughout this dissertation. Building on the experience of creating user-centric wearable sensors for human-environment interaction with AirSpecs, and tackling specialized ecological acoustic monitoring with BuzzCam, CollarID integrates these disparate sensing modalities—environmental, bioacoustic, and inertial—into a single, robust platform. It embodies the core thesis objective of creating a more holistic 'lens on life' by moving beyond single-purpose sensors to a truly integrated, context-aware system.

The integration of a particulate matter sensor, in particular, moves the platform beyond traditional biologging and into the critical domain of environmental health, enabling the formulation of direct, testable hypotheses about the sublethal impacts of air quality on wildlife. Building on the initial questions posed by this capability, future deployments are positioned to formally test hypotheses such as:

- **H1: Exposure and Foraging Efficiency.** Lions or hyenas exposed to PM2.5 levels exceeding a critical threshold during wildfire events will exhibit a statistically significant reduction in daily travel distance (measured by GPS) and an increase in resting behaviors (classified from IMU data) in the subsequent 48 hours.

- **H2: Habitat Selection and Pollution Avoidance.** Animals will actively alter their habitat use patterns to avoid areas with chronically elevated anthropogenic particulate matter, a preference that can be modeled by correlating fine-scale location data with on-collar PM10 measurements.

- **H3: Respiratory Health and Bioacoustics.** A higher cumulative exposure to airborne particulates will correlate with a greater incidence of sounds indicative of respiratory distress (e.g., coughing, wheezing) detectable through continuous on-collar bioacoustic analysis.

By empowering researchers to move from broad landscape-level correlations to direct, animal-borne exposure-response studies, CollarID provides a tool to address a crucial and previously invisible ecological stressor.

The successful development and validation of the CollarID platform lay the groundwork for its deployment in full-scale wildlife studies. Future work will focus on collaborating with field ecologists to deploy the platform on target species, such as lions, hyenas, and wild dogs, in their natural habitats. These deployments will enable long-term, multi-modal data collection, allowing researchers to investigate complex ecological questions—such as the impact of environmental changes on animal behavior and health. Additionally, the platform's "AI-ready" architecture will be leveraged to develop on-device machine learning models for real-time behavioral classification, further enhancing its utility as an autonomous monitoring tool. Specifically, the end-to-end methodology for model creation and deployment developed for BuzzCam provides a direct blueprint for this work. The same principles of data collection, model pruning, and quantization-aware training can be applied to CollarID's rich acoustic and inertial data streams to create on-device classifiers for specific events of high conservation interest, such as poaching activities (e.g., gunshots, vehicles) or animal distress calls, creating a powerful early-alert mechanism for park rangers.

Having now presented the design and validation of all three novel sensing platforms, the final chapter will synthesize the cross-contextual contributions of this research. It will discuss the broader implications of these new tools for human-computer interaction and conservation technology, reflect on the common principles of engineering 'in the wild,' and outline key directions for future work that build upon these foundations.

Chapter 6

Discussion, Future Work, and Conclusion

The central ambition of this dissertation has been to design, engineer, validate, and apply a new generation of sensing technologies capable of providing a more insightful, holistic, and contextualized understanding of life. Situated at the confluence of human-computer interaction, ecological monitoring, and embedded systems engineering, this work has traversed diverse environments—from the built interiors of our cities to the harsh wilderness of the Arctic and the vital ecosystems of Patagonia. The research journey was propelled by a foundational belief: that by creating better tools to see and interpret the world, we can foster a deeper understanding of both our own species and the myriad others with which we share the planet. This dissertation presented a suite of novel platforms, each acting as a new technological "lens," designed to move beyond simplistic, isolated measurements and instead capture the complex interplay between organisms and their environments.

The investigation began with an inward focus, exploring the nuanced and deeply subjective nature of the human experience within built environments. The AirSpecs smart-eyeglass platform was developed to challenge the paradigm of static, room-level environmental monitoring, providing instead a personalized, "in-the-wild" perspective on human comfort and well-being. By holistically sensing an individual's immediate micro-environment and physiological state, AirSpecs demonstrated a new methodology for capturing the data needed to design truly human-centric spaces.

The engineering principles and methodological insights honed during the development of AirSpecs—the challenges of robust field deployment, low-power design, and multi-modal data fusion—provided a direct bridge to a parallel and equally pressing challenge: the non-invasive monitoring of wildlife. The research then pivoted, applying this cross-contextual design philosophy to the ecological domain. This led to a progression of acoustic platforms designed to listen to the natural world with greater fidelity and intelligence. The SoundSHROOM system was engineered as a direct response to the need for robust, multi-channel recorders capable of surviving harsh environments like the Arctic, enabling new possibilities for spatial bioacoustic analysis. Building on this, the BuzzCam project tackled the critical "data-to-insight" bottleneck in conservation technology by developing a specialized platform and an end-to-end pipeline for on-device AI classification of endangered and invasive bee species. Finally, the research culminated in the CollarID platform, a synthesis of the dissertation's principles. This versatile, animal-borne sensor was engineered to provide a truly holistic view of an animal's life, moving beyond simple location tracking by integrating a rich suite of inertial, bioacoustic, and environmental sensors.

6.1 Summary of Research and Contributions

This dissertation pursued three primary research objectives aimed at developing novel sensing technologies to provide a more insightful lens on life across both human and ecological contexts. This section revisits each of those objectives, summarizing how the research projects undertaken—AirSpecs, SoundSHROOM, BuzzCam, and CollarID—successfully addressed the initial research questions and resulted in a series of tangible scientific and engineering contributions.

6.1.1 Objective 1 Revisited: Human-Centric Sensing with AirSpecs

The first objective was to design, deploy, and evaluate a novel, head-worn wearable platform (AirSpecs) capable of holistically sensing an individual's proximate environment and physiological state to deepen our understanding of human comfort in real-world settings. This work was motivated by the recognized limitation of traditional building sensing, which relies on static, room-level metrics that fail to capture the dynamic and subjective nature of human comfort.

This objective was successfully met through several key contributions. First, I developed a novel wearable sensing platform, AirSpecs, a smart-eyeglass that uniquely integrates a comprehensive suite of environmental sensors (temperature, humidity, VOCs, NO_x, light, CO_2, noise) and physiological sensors (skin temperature, blink rate). This platform was supported by a custom mobile application for data visualization and Ecological Momentary Assessment (EMA) and was engineered specifically for "in-the-wild" human comfort research. The system then underwent successful "in-the-wild" validation through a multi-site international study involving 30 participants in Boston, Fribourg, and Singapore, which demonstrated its viability as a research tool for collecting rich, longitudinal data. Finally, a key outcome of this work is a unique public dataset, the "Dataset Exploring Urban Comfort Through Novel Wearables and Environmental Surveys," which has been made available to the scientific community [3]. The value of this rich dataset, capturing synchronized environmental, physiological, and subjective data, has already been confirmed through its use by independent researchers to develop and validate new, more advanced personalized comfort models.

6.1.2 Objective 2 Revisited: Ecological Acoustics and AI with SoundSHROOM and BuzzCam

The second objective was twofold: to engineer and field-validate robust acoustic monitoring hardware for challenging ecological applications and to tackle the "data-to-insight" bottleneck by using on-device Artificial Intelligence for scalable monitoring.

This objective was addressed through two sequential projects. The first, focused on robust multi-channel acoustic hardware (SoundSHROOM), addressed the need for resilient hardware in extreme environments. For this, I developed the SoundSHROOM system, a robust, multi-channel acoustic recorder. Its successful deployment in the harsh Arctic environment of

Svalbard validated its design, produced a unique public dataset of Arctic soundscapes suitable for advanced spatial audio analysis, and provided crucial insights into designing hardware for robust field deployment. Building on this experience, the second project, focused on on-device AI for pollinator monitoring (BuzzCam), was developed to provide a tangible solution to the data bottleneck in passive acoustic monitoring. The key contributions of BuzzCam include a specialized platform tailored for pollinator monitoring and an associated high-resolution, annotated dataset of endangered and invasive bee bioacoustics from Patagonia. Critically, the project also delivered a complete end-to-end methodology that provides a replicable blueprint for creating intelligent, scalable ecological sensors, demonstrating the process of model adaptation, Quantization-Aware Training (QAT), and deployment on a low-power microcontroller (the MAX78000) for real-time bee classification.

6.1.3 Objective 3 Revisited: Multi-Modal Biologging with CollarID

The third objective was to engineer and characterize a versatile, low-power, multi-modal animal-borne platform (CollarID) to overcome the limitations of traditional location-only tracking devices and enable a more holistic understanding of wildlife. This work was driven by the need to fuse location data with fine-grained behavioral, physiological, and environmental context.

This objective was met through the rigorous development and validation of the CollarID platform. I first designed, engineered, and characterized CollarID as a holistic, multi-modal wildlife sensor, an integrated prototype that combines inertial measurement (IMU), high-fidelity bioacoustic recording, and a comprehensive environmental sensing suite (temperature, humidity, pressure, and particulate matter—a novelty for animal loggers) with GPS and long-range communications. This prototype then underwent rigorous engineering validation through a series of comprehensive tests to establish its field-readiness, including Finite Element Analysis (FEA) of its mechanical robustness against simulated predator bite forces, environmental sealing tests, co-location validation of its environmental sensors, and extensive field testing of its communication capabilities. The outcome is a validated field-ready platform; the successful integration of all subsystems was confirmed in an initial field trial on a farm, where the collar demonstrated its reliability, autonomous solar-powered operation, and data streaming capabilities. This work contributes a validated hardware platform that serves as

6.2 Synthesis: Cross-Contextual Insights from Engineering in the Wild

While the application contexts of this dissertation are diverse—spanning human comfort, Arctic soundscapes, pollinator monitoring, and large mammal tracking—the research journey has revealed a set of unifying engineering principles and methodological insights. The process of designing, building, and deploying sensor systems in uncontrolled, real-world environments surfaced common challenges and solutions. This section synthesizes these cross-contextual learnings, which represent a core contribution of this thesis, grouped into three key themes: the power of holistic sensing, the principles of robust field deployment, and the trajectory towards on-device intelligence.

6.2.1 The Power of Holistic, Multi-Modal Sensing

A consistent theme across all projects is that a deeper, more contextualized understanding of an organism arises from integrating multiple, diverse data streams. This research has repeatedly demonstrated the limitations of single-metric sensing and the power of a holistic, multi-modal approach. With AirSpecs, understanding human comfort required moving beyond a simple thermostat reading; by uniquely integrating a full suite of Indoor Environmental Quality (IEQ) sensors with physiological measurements and subjective feedback, the platform captured a far richer and more personalized picture of well-being. This principle was extended into the ecological domain with BuzzCam, which was engineered not only to capture high-fidelity bee buzz acoustics but also to co-register key environmental parameters like temperature and humidity, enabling the correlation of acoustic activity with the specific microclimatic conditions that influence pollinator behavior. The CollarID platform represents the culmination of this multi-modal philosophy, intentionally designed to break down data silos by fusing inertial sensors for behavior, bioacoustic microphones for vocalizations, a comprehensive environmental suite for context, and GPS for location into one cohesive system. This integration is what facilitates a truly holistic view, making it possible to simultaneously ask *what* an animal is doing, *where* it is, *what* it is experiencing, and *how* it is reacting.

6.2.2 Common Principles for Robust Field Deployment

Engineering "in the wild" imposes a set of harsh, practical constraints that are shared across all field-based sensing applications. The successful deployments of AirSpecs, SoundSHROOM, and CollarID were contingent on adhering to several common principles for robust design. Environmental hardening was a primary concern, driving the design of the SoundSHROOM system to withstand Arctic weather and culminating in the CollarID, whose housing was validated not only for environmental sealing but also for structural integrity against the worst-case bite forces of large predators. Power management and longevity were also critical; this ranged from designing AirSpecs for a full day of operation, to leveraging external solar-assisted batteries for SoundSHROOM, to integrating solar charging and an autonomous, fault-tolerant hibernation protocol directly into the CollarID platform for long-term survival. Finally, for acoustic platforms, wind noise mitigation was essential. This was systematically investigated with the SoundSHROOM project in the Arctic, and the practical lessons learned were then directly applied to the engineering of the BuzzCam microphone enclosures to ensure the highest possible data quality.

6.2.3 The Trajectory Towards On-Device Intelligence

The research in this dissertation follows a deliberate trajectory away from simple data logging and towards on-device intelligence. This evolution is driven by the practical need to address the "data-to-insight" bottleneck, where the sheer volume of sensor data collected in the field becomes a barrier to its timely analysis and use. The journey began with AirSpecs, where primary data processing was offloaded to a connected smartphone, but initial steps towards on-device processing were taken for metrics like real-time noise level (dBA) calculation. The BuzzCam project represented a focused leap into on-device AI, conceived as a direct answer to the data bottleneck problem in passive acoustic monitoring. The core contribution here was the development of a complete end-to-end pipeline to successfully deploy a sophisticated Convolutional Neural Network (CNN) onto a highly resource-constrained, low-power microcontroller (the MAX78000). This achievement transforms the sensor from a passive logger into an active, intelligent edge device that can provide real-time classifications. The CollarID platform is the final step in this trajectory. While the work in this thesis focused on its rigorous engineering validation, the platform was explicitly designed to be "AI-ready," with its powerful microcontroller and rich, time-synchronized, multi-modal data streams providing an ideal foundation for future work in on-device machine learning, such as the automated classification of animal behaviors directly on the collar.

This progression towards on-device intelligence is a critical theme, highlighting a path to creating more scalable, power-efficient, and responsive sensors capable of delivering actionable insights directly from the field.

6.2.4 Towards a Unified Framework for Organism-in-Environment Sensing

Beyond the shared engineering principles of robust design and low-power operation, this body of work points toward a more unified conceptual framework for sensing life. The research began by exploring human "comfort," a subjective state of well-being assessed through a fusion of environmental, physiological, and behavioral data. The subsequent ecological platforms, while not measuring subjective states, pursued a parallel goal: assessing animal "well-being" through proxies like behavior, health, and environmental interaction. This parallel suggests that the core challenge is fundamentally the same: to understand an organism's internal state and its dynamic relationship with its immediate surroundings.

The methodology developed and replicated across these contexts—fusing data from proximate environmental sensors (IEQ for AirSpecs, microclimate for CollarID), behavioral sensors (reaction time for AirSpecs, IMU for CollarID), and physiological-proxy sensors (skin temperature for AirSpecs, bioacoustics for CollarID)—represents a tangible step toward a common, species-agnostic approach. It posits that a holistic understanding of any organism, whether human or animal, requires moving beyond single-metric observation to a multi-modal assessment of the organism-environment interface. This dissertation, therefore, contributes not just a set of tools, but a validated methodology for creating a more complete, contextualized, and empathetic understanding of an organism's state.

6.3 Broader Implications

The contributions of this dissertation—the novel platforms, unique datasets, and demonstrated methodologies—extend beyond their immediate research contexts. The work carries broader implications for the future of human-computer interaction and smart buildings, for the practice of conservation technology and ecology, and for the foundational discipline of sensing systems engineering.

6.3.1 Implications for Human-Computer Interaction and Smart Buildings

The AirSpecs project, in particular, offers several key implications for how we can design more effective, humane, and responsive intelligent environments. This research reinforces the shift from standardized to personalized comfort by providing a practical methodology using wearables like AirSpecs to collect the essential data streams—co-located environmental data, physiological signals, and subjective feedback—needed to build and validate next-generation, adaptive comfort models. The findings also challenge the paradigm of the "invisible" smart building by promoting a design for awareness and agency; showing that access to personal data can empower occupants to take direct action, this suggests a future for Human-Building Interaction (HBI) that is collaborative rather than purely controlling. Finally, the work proposes new context-aware interaction modalities, with a framework of "focus mode," "ambient assistance," and "reflective explanation" that provides a blueprint for systems that adapt to an occupant's cognitive state to be helpful without being disruptive.

6.3.2 Implications for Conservation Technology and Ecology

The ecological sensing platforms developed in this thesis provide tangible tools and new methodologies that can help address some of the most pressing challenges in biodiversity monitoring and conservation. The SoundSHROOM project, for instance, enables monitoring in extreme and inaccessible environments; by demonstrating that robust systems can be deployed in hostile regions like the Arctic, it allows researchers to gather crucial baseline data on the impacts of climate change. The BuzzCam system offers a solution for scalable, non-invasive monitoring for cryptic species. Its on-device AI pipeline transforms passive acoustic monitoring into an autonomous solution, providing a powerful, non-lethal alternative to traditional methods and creating a viable pathway for large-scale, long-term monitoring networks. Finally, the CollarID platform is designed to provide a more holistic view of animal ecology, moving research beyond location-only data. By holistically integrating behavioral, bioacoustic, environmental, and spatial data—including the novel use of particulate matter sensing—it empowers ecologists to investigate more complex questions about animal-environment interactions, such as the sublethal effects of air quality on wildlife.

Beyond its academic contributions, this research has the potential to influence conservation policies and industry practices. For instance, the scalable monitoring enabled by BuzzCam's on-device AI could inform policies aimed at protecting endangered pollinators, while CollarID's holistic data collection could guide habitat management strategies. In the built environment, AirSpecs' personalized comfort insights could shape standards for smart building design, promoting occupant well-being and energy efficiency.

6.3.3 Implications for Sensing Systems Engineering

Finally, this body of work, viewed as a whole, offers several insights for the practice and philosophy of engineering sensor systems for real-world applications. This thesis serves as a case study for the value of a cross-contextual design philosophy, demonstrating how transferable engineering principles from a human-centric wearable like AirSpecs provided a direct foundation for tackling similar challenges in ecological platforms. This highlights the importance of end-to-end system development, where a successful solution is an integrated system, as exemplified by the BuzzCam project from its custom hardware and in-field annotation app to the final co-designed machine learning model. This approach is realized through prototyping as a research method; the iterative act of engineering "in the wild" is not merely a validation step but a generative process that stress-tests assumptions, reveals unforeseen challenges, and uncovers new research questions, championing the value of "research through design and deployment."

6.4 Limitations and Future Work

This dissertation has successfully developed and validated a suite of novel sensing platforms, yet it is essential to acknowledge the boundaries of the research and to outline the exciting avenues for future work that these new tools have opened.

6.4.1 Overall Research Limitations

While the projects achieved their primary objectives, the following limitations are acknowledged to provide a clear context for the contributions of this thesis. The scope of field deployments, while crucial for technological validation, was not designed for long-term, multi-season sociological or ecological studies; their primary purpose was to demonstrate hardware capabilities and collect foundational datasets, as exemplified by the limited duration and university-based participant pool of the AirSpecs study. Similarly, the generalizability of machine learning models is a consideration; the BuzzCam models, trained on data from a specific region and season, demonstrated a powerful proof-of-concept but may require site-specific fine-tuning for high-confidence deployment elsewhere. Finally, the work on the CollarID platform focused on its engineering validation vs. full-scale deployment. While this thesis presents its rigorous characterization as a field-ready prototype, full-scale, long-term deployments on target wildlife species are a critical next step considered future work.

Synthesized Limitations and the Sensing Frontier

Reviewing the limitations across the AirSpecs, BuzzCam, and CollarID projects reveals a set of common, fundamental challenges inherent to "in-the-wild" sensing. These include the trade-offs between deployment duration and data richness, the geographical specificity that can limit the generalizability of machine learning models trained in one environment, and the engineering reality of balancing cutting-edge capability with prototype robustness.

More profoundly, this body of work illuminates the "null space" of what these advanced platforms still do not capture. While they provide an unprecedented lens on an organism's external context and behavioral response, they lack direct insight into its internal biochemical state. Crucial drivers of well-being such as nutritional status, hormonal stress profiles (e.g., cortisol levels), immune response, and pathogen load remain largely invisible to this sensing paradigm.

The next frontier for truly holistic monitoring will involve bridging this gap between external context and internal physiology. Future work should therefore explore the integration of emerging non-invasive or minimally-invasive biochemical sensors. Advances in wearable patches for sweat analysis, sensors for detecting volatile organic compounds in breath, or microneedle arrays for interstitial fluid sampling could one day be integrated into platforms like AirSpecs and CollarID. Such a fusion of environmental, behavioral, and biochemical sensing would represent the next evolutionary step in creating a truly comprehensive "lens on life."

6.4.2 Future Research Directions

The platforms and methodologies developed in this thesis lay the groundwork for numerous avenues of future research. A key direction involves closing the loop in Human-Building Interaction; future work could use AirSpecs data to directly modulate personal comfort systems or provide feedback to a Building Management System, allowing for a real-world test of the proposed interaction framework. Another major area is expanding on-device AI for ecological intelligence. This includes enhancing the generalizability of the BuzzCam bee classification model with more diverse datasets and, for CollarID, leveraging its "AI-ready" architecture to develop on-device models for real-time behavioral classification. This process would directly leverage the on-device AI pipeline developed for BuzzCam, adapting the methodology to train classifiers on CollarID's acoustic data to detect specific, high-stakes events like gunshots or vehicle engines. Success in this area would transform the collar into a real-time early alert mechanism for poaching events, providing park rangers with timely and actionable intelligence. The most critical future work for the CollarID platform is its full-scale, multi-modal wildlife studies on target populations, which would allow ecologists to investigate complex questions, such as how microclimatic variations or air quality events impact animal behavior and health. Finally, the multi-channel SoundSHROOM platform enables future research in leveraging spatial audio for ecological insights, where the implementation of sound source localization and beamforming algorithms could enhance biodiversity assessments by tracking vocalizing animals or improving the detection of faint calls.

6.4.3 Concluding Remarks

This dissertation was founded on the pursuit of a new technological lens on life—one capable of capturing the world with greater context, holistically, and with increasing intelligence. The research journey moved from the microscopic scale of personal human environments to the vast and challenging landscapes of ecological conservation. It began with the development of AirSpecs, a platform to understand the nuances of our own well-being within the built world, and progressed to the engineering of SoundSHROOM, BuzzCam, and CollarID to better perceive, interpret, and protect the lives of the diverse species with which we share our planet.

Through the design, engineering, and real-world validation of these novel platforms, this thesis has contributed not just a set of new research tools and valuable public datasets, but also a demonstrated cross-contextual methodology. It has shown how the principles of robust field deployment, multi-modal data fusion, and on-device intelligence can be translated across disparate domains to solve fundamental sensing challenges.

The platforms created here represent a tangible step towards a future where technology can bridge the gap between raw data and actionable insight. They are instruments designed to listen more closely to the subtle buzz of a bee, to see the invisible environmental stressors affecting an animal, and to understand the deeply personal nature of human comfort. As these sensing technologies continue to evolve—becoming smaller, more power-efficient, and more intelligent—they hold the promise of empowering scientists, conservationists, designers, and individuals with a more profound and empathetic understanding of the intricate connections that define life on Earth. The work presented herein is a foundational step on that continuing journey. Ultimately, this thesis's demonstrated cross-contextual methodology—applying engineering principles from human-centric sensing to ecological conservation—offers a replicable blueprint for future interdisciplinary research, fostering innovation across diverse fields.

References

[1] M. Rampioni, P. Nicolini, and L. Mandolini. "Analysis of human behavior in everyday life's contexts, for the development of new technologies, in support of the improvement of life quality and well-being." In: 2017. URL: https://api.semanticscholar.org/CorpusID:158818457.

[2] C. Schweizer, R. D. Edwards, L. Bayer-Oglesby, W. J. Gauderman, V. Ilacqua, M. J. Jantunen, H. K. Lai, M. J. Nieuwenhuijsen, and N. Künzli. "Indoor time–micro-environment–activity patterns in seven regions of Europe." *Journal of Exposure Science and Environmental Epidemiology*, **17**(), 2007, pp. 170–181. URL: https://api.semanticscholar.org/CorpusID:205677619.

[3] P. Chwalek, S. Zhong, N. Perry, T. Liu, C. Miller, H. S. Alavi, D. Lalanne, and J. A. Paradiso. "Exploring Urban Comfort through Novel Wearables and Environmental Surveys." *ArXiv*, **abs/2408.08323**(), 2024. URL: https://api.semanticscholar.org/CorpusID:271891888.

[4] H. S. Alavi, E. F. Churchill, M. Wiberg, D. Lalanne, P. Dalsgaard, A. F. gen. Schieck, and Y. Rogers. "Introduction to Human-Building Interaction (HBI)." *ACM Transactions on Computer-Human Interaction (TOCHI)*, **26**(), 2019, pp. 1–10. URL: https://api.semanticscholar.org/CorpusID:83458858.

[5] S. Zhong, P. Chwalek, N. Perry, D. Ramsay, C. Miller, D. Lalanne, H. S. Alavi, and J. A. Paradiso. "Sensors and Sensibilities: Exploring Interactions for Habitat Comfort with An Environmental-Physiological Sensing Eyewear In the Wild." *International Journal of Human-Computer Studies*, **200**(), 2025, p. 103510. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2025.103510. URL: https://www.sciencedirect.com/science/article/pii/S1071581925000679.

[6] M. Feldmeier and J. A. Paradiso. "Personalized HVAC control system." *2010 Internet of Things (IOT)*, (), 2010, pp. 1–8. URL: https://api.semanticscholar.org/CorpusID:423027.

[7] ASHRAE. *ANSI/ASHRAE Standard 55-2020: Thermal Environmental Conditions for Human Occupancy*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2021. URL: https://www.ashrae.org/technical-resources/standards-and-guidelines/read-only-versions-of-ashrae-standards.

[8] M. J. Frontczak and P. Wargocki. "Literature survey on how different factors influence human comfort in indoor environments." *Fuel and Energy Abstracts*, (), 2011. URL: https://api.semanticscholar.org/CorpusID:94032215.

[9] P. O. Fanger. "Thermal comfort: analysis and applications in environmental engineering," in: 1972. URL: https://api.semanticscholar.org/CorpusID:110028033.

[10] R. de Dear and G. Brager. "Developing an adaptive model of thermal comfort and preference." *Center for the Built Environment*, **104**(), 1998, pp. 145–167. URL: https://api.semanticscholar.org/CorpusID:4666243.

[11] J. F. Nicol and M. A. Humphreys. "ADAPTIVE THERMAL COMFORT AND SUSTAINABLE THERMAL STANDARDS FOR BUILDINGS." *Energy and Buildings*, **34**(), 2002, pp. 563–572. URL: https://api.semanticscholar.org/CorpusID:17571584.

[12] M. Prendergast. "Lighting at Work." In: 2017. URL: https://api.semanticscholar.org/CorpusID:117488016.

[13] B. Berglund. "FROM THE WHO GUIDELINES FOR COMMUNITY NOISE TO HEALTHY SOUNDSCAPES." *ENVIRONMENTAL NOISE: WHO GUIDELINES*, (), 2023. URL: https://api.semanticscholar.org/CorpusID:108126924.

[14] P. Wargocki, D. P. Wyon, J. Sundell, G. Clausen, and P. O. Fanger. "The effects of outdoor air supply rate in an office on perceived air quality, sick building syndrome (SBS) symptoms and productivity." *Indoor air*, **10 4**(), 2000, pp. 222–36. URL: https://api.semanticscholar.org/CorpusID:28560826.

[15] O. Seppänen, W. J. Fisk, and M. J. Mendell. "Association of ventilation rates and CO2 concentrations with health and other responses in commercial and institutional buildings." *Indoor air*, **9 4**(), 1999, pp. 226–52. URL: https://api.semanticscholar.org/CorpusID:14264134.

[16] R. J. Cole, J. Robinson, Z. Brown, and M. O'shea. "Re-contextualizing the notion of comfort." *Building Research & Information*, **36**(), 2008, pp. 323–336. URL: https://api.semanticscholar.org/CorpusID:110328866.

[17] G. Brager, G. Paliaga, and R. de Dear. "Operable windows, personal control and occupant comfort." *Ashrae Transactions*, (), 2004, pp. 17–35. URL: https://api.semanticscholar.org/CorpusID:15884553.

[18] M. Schweiker et al. "Review of multi-domain approaches to indoor environmental perception and behaviour." *Building and Environment*, (), 2020. URL: https://api.semanticscholar.org/CorpusID:216326991.

[19] A. I. Dounis and C. Caraiscos. "Advanced control systems engineering for energy and comfort management in a building environment–A review." *Renewable & Sustainable Energy Reviews*, **13**(), 2009, pp. 1246–1261. URL: https://api.semanticscholar.org/CorpusID:17035059.

[20] E. Margariti, V. Vlachokyriakos, and D. S. Kirk. "Understanding occupants' experiences in quantified buildings: results from a series of exploratory studies." *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (), 2023. URL: https://api.semanticscholar.org/CorpusID:258217324.

[21] G. Coulby, A. K. Clear, O. Jones, and A. Godfrey. "A Scoping Review of Technological Approaches to Environmental Monitoring." *International Journal of Environmental Research and Public Health*, **17**(), 2020. URL: https://api.semanticscholar.org/CorpusID:219550723.

[22] S. Yun, S. Zhong, H. S. Alavi, A. Alahi, and D. Licina. "Proxy methods for detection of inhalation exposure in simulated office environments." *Journal of Exposure Science & Environmental Epidemiology*, **33**(3), 2023, pp. 396–406. DOI: https://doi.org/10.1038/s41370-022-00495-w.

[23] K. Nkurikiyeyezu, Y. Suzuki, and G. Lopez. "Heart rate variability as a predictive biomarker of thermal comfort." *Journal of Ambient Intelligence and Humanized Computing*, **9**(), 2018, pp. 1465–1477. URL: https://api.semanticscholar.org/CorpusID:52300086.

[24] J. Healey and R. W. Picard. "Detecting stress during real-world driving tasks using physiological sensors." *IEEE Transactions on Intelligent Transportation Systems*, **6**(), 2005, pp. 156–166. URL: https://api.semanticscholar.org/CorpusID:1409560.

[25] K. Knecht. "Micro-Atmospheres: Exploring Portable and Wearable Solutions to Support Individual Thermal Comfort." *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, (), 2015. URL: https://api.semanticscholar.org/CorpusID:20930128.

[26] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos. "Personal thermal comfort models with wearable sensors." *Building and Environment*, (), 2019. URL: https://api.semanticscholar.org/CorpusID:199083732.

[27] S. Zhong, H. S. Alavi, and D. Lalanne. "Hilo-wear: Exploring Wearable Interaction with Indoor Air Quality Forecast." *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, (), 2020. URL: https://api.semanticscholar.org/CorpusID:218483359.

[28] V. Nguyen. "myAQM: Interface Portable Air Quality Sensors with Apple Watch." *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, (), 2022. URL: https://api.semanticscholar.org/CorpusID:258301649.

[29] Y. Abdelrahman, E. Velloso, T. Dingler, A. Schmidt, and F. Vetere. "Cognitive Heat." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **1**(), 2017, pp. 1–20. URL: https://api.semanticscholar.org/CorpusID:23761664.

[30] P. Chwalek, S. Zhong, D. Ramsay, N. Perry, and J. A. Paradiso. "AirSpec: A Smart Glasses Platform, Tailored for Research in the Built Environment." *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing*, (), 2023. URL: https://api.semanticscholar.org/CorpusID:263742895.

[31] A. Brambilla and T. Jusselme. "Preventing overheating in offices through thermal inertial properties of compressed earth bricks: A study on a real scale prototype." *Energy and Buildings*, **156**(), 2017, pp. 281–292. URL: https://api.semanticscholar.org/CorpusID:115581179.

[32] D. Pisharoty, R. Yang, M. W. Newman, and K. Whitehouse. "ThermoCoach: Reducing Home Energy Consumption with Personalized Thermostat Recommendations." *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, (), 2015. URL: https://api.semanticscholar.org/CorpusID:7534015.

[33] J. E. Froehlich, L. Findlater, and J. A. Landay. "The design of eco-feedback technology." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (), 2010. URL: https://api.semanticscholar.org/CorpusID:262318070.

[34] C. A. Hoppmann and A. Ho. "Ecological Momentary Assessment." In: 2015. URL: https://api.semanticscholar.org/CorpusID:134420442.

[35] S. S. Intille, C. Haynes, D. Maniar, A. Ponnada, and J. Manjourides. "μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch." *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, (), 2016. URL: https://api.semanticscholar.org/CorpusID:2576931.

[36] F. Tartarini, S. Schiavon, M. Quintana, and C. Miller. "Personal comfort models based on a 6-month experiment using environmental parameters and data from wearables." *Indoor Air*, **32**(), 2022. URL: https://api.semanticscholar.org/CorpusID:254034412.

[37] D. B. Ramsay and J. A. Paradiso. "Peripheral Light Cues as a Naturalistic Measure of Focus." *Proceedings of the 2022 ACM International Conference on Interactive Media Experiences*, (), 2022. URL: https://api.semanticscholar.org/CorpusID:249928097.

[38] E. Costanza, S. A. Inverso, E. Pavlov, R. Allen, and P. Maes. "eye-q: eyeglass peripheral display for subtle intimate notifications." In: *Mobile HCI*. 2006. URL: https://api.semanticscholar.org/CorpusID:11385304.

[39] C. Kremen and A. M. Merenlender. "Landscapes that work for biodiversity and people." *Science*, **362**(), 2018. URL: https://api.semanticscholar.org/CorpusID:53012788.

[40] H. M. Pereira et al. "Essential Biodiversity Variables." *Science*, **339**(6117), 2013, pp. 277–278. DOI: 10.1126/science.1229931. eprint: https://www.science.org/doi/pdf/10.1126/science.1229931. URL: https://www.science.org/doi/abs/10.1126/science.1229931.

[41] C. J. Bibby, N. D. Burgess, and D. A. Hill. "Bird Census Techniques." In: 1992. URL: https://api.semanticscholar.org/CorpusID:83098529.

[42] S. T. Buckland. "Introduction to distance sampling : estimating abundance of biological populations." In: 2001. URL: https://api.semanticscholar.org/CorpusID:140535037.

[43] K. F. Kellner and R. K. Swihart. "Accounting for Imperfect Detection in Ecology: A Quantitative Review." *PLoS ONE*, **9**(), 2014. URL: https://api.semanticscholar.org/CorpusID:14513358.

[44] D. L. Otis, K. P. Burnham, G. C. White, and D. R. Anderson. "Statistical inference from capture data on closed animal populations." *Wildlife Monographs*, (), 1980, pp. 3–135. URL: https://api.semanticscholar.org/CorpusID:87934076.

[45] A. B. Lawson, K. H. Pollock, J. D. Nichols, C. Brownie, and J. E. Hines. "Statistical inference for capture-recapture experiments." *Biometrics*, **107**(), 1992, p. 656. URL: https://api.semanticscholar.org/CorpusID:123163371.

[46] R. A. Powell and G. Proulx. "Trapping and marking terrestrial mammals for research: integrating ethics, performance criteria, techniques, and common sense." *ILAR journal*, **44 4**(), 2003, pp. 259–76. URL: https://api.semanticscholar.org/CorpusID:14370302.

[47] T. H. Roulston, S. A. Smith, and A. L. Brewster. "A Comparison of Pan Trap and Intensive Net Sampling Techniques for Documenting a Bee (Hymenoptera: Apiformes) Fauna." In: 2007. URL: https://api.semanticscholar.org/CorpusID:86219562.

[48] G. J. Wilson and R. J. Delahay. "A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation." *Wildlife Research*, **28**(), 2001, pp. 151–164. URL: https://api.semanticscholar.org/CorpusID:86260323.

[49] D. T. Blumstein et al. "Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus." *Journal of Applied Ecology*, **48**(), 2011, pp. 758–767. URL: https://api.semanticscholar.org/CorpusID:33485268.

[50] L. S. M. Sugai, T. S. F. Silva, J. W. Ribeiro, and D. Llusia. "Terrestrial Passive Acoustic Monitoring: Review and Perspectives." *BioScience*, **69**(), 2018, pp. 15–25. URL: https://api.semanticscholar.org/CorpusID:92602132.

[51] 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*, **10**(), 2018, pp. 169–185. URL: https://api.semanticscholar.org/CorpusID:92554621.

[52] S. Kahl, C. M. Wood, M. Eibl, and H. Klinck. "BirdNET: A deep learning solution for avian diversity monitoring." *Ecol. Informatics*, **61**(), 2021, p. 101236. URL: https://api.semanticscholar.org/CorpusID:232359415.

[53] C. L. Walters et al. "A continental-scale tool for acoustic identification of European bats." *Science & Engineering Faculty*, (), 2012. URL: https://api.semanticscholar.org/CorpusID:18353091.

[54] J. H. Waddle, T. F. Thigpen, and B. M. Glorioso. "Efficacy of automatic vocalization recognition software for anuran monitoring." *Herpetological Conservation and Biology*, (), Dec. 2009. USGS Publication.

[55] T. A. Marques, L. Thomas, S. Martin, D. K. Mellinger, J. A. Ward, D. J. Moretti, D. V. Harris, and P. L. Tyack. "Estimating animal population density using passive acoustics." *Biological Reviews of the Cambridge Philosophical Society*, **88**(), 2012, pp. 287–309. URL: https://api.semanticscholar.org/CorpusID:16013311.

[56] P. Chwalek et al. "High-Res Acoustic and Environmental Data to Monitor Bombus dahlbomii Amid Invasive Species, Habitat Loss." *Scientific Data*, **12**(), 2025. URL: https://api.semanticscholar.org/CorpusID:277495481.

[57] B. C. Pijanowski, L. J. Villanueva-Rivera, S. L. Dumyahn, A. Farina, B. L. Krause, B. M. Napoletano, S. H. Gage, and N. Pieretti. "Soundscape Ecology: The Science of Sound in the Landscape." *BioScience*, **61**(3), Mar. 2011, pp. 203–216. ISSN: 0006-3568. DOI: 10.1525/bio.2011.61.3.6. eprint: https://academic.oup.com/bioscience/article-pdf/61/3/203/19404645/61-3-203.pdf. URL: https://doi.org/10.1525/bio.2011.61.3.6.

[58] A. P. Hill, P. Prince, E. P. Covarrubias, C. P. Doncaster, J. L. Snaddon, and A. Rogers. "AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment." *Methods in Ecology and Evolution*, **9**(), 2018, pp. 1199–1211. URL: https://api.semanticscholar.org/CorpusID:54932414.

[59] P. Chwalek, M. Coblentz, S. Montague, M. Kuronaga, I. Zhu, and J. A. Paradiso. "Acoustic data collection in arctic environments during the midnight sun using multi-channel SoundSHROOMs." *Scientific Data*, **12**(), 2025. URL: https://api.semanticscholar.org/CorpusID:276557168.

[60] S. S. Sethi, R. M. Ewers, N. S. Jones, C. D. L. Orme, and L. Picinali. "Robust, real-time and autonomous monitoring of ecosystems with an open, low-cost, networked device." *bioRxiv*, (), 2017. URL: https://api.semanticscholar.org/CorpusID:196634409.

[61] M. Allen, L. Girod, R. Newton, S. Madden, D. T. Blumstein, and D. Estrin. "VoxNet: An Interactive, Rapidly-Deployable Acoustic Monitoring Platform." *2008 International Conference on Information Processing in Sensor Networks (ipsn 2008)*, (), 2008, pp. 371–382. URL: https://api.semanticscholar.org/CorpusID:7575824.

[62] T. A. Rhinehart, L. M. Chronister, T. Devlin, and J. Kitzes. "Acoustic localization of terrestrial wildlife: Current practices and future opportunities." *Ecology and Evolution*, **10**(), 2020, pp. 6794–6818. URL: https://api.semanticscholar.org/CorpusID:220850742.

[63] Y. Ropert-Coudert and R. P. Wilson. "Trends and perspectives in animal-attached remote sensing." *Frontiers in Ecology and the Environment*, **3**(), 2005, pp. 437–444. URL: https://api.semanticscholar.org/CorpusID:86733941.

[64] C. C. Wilmers, B. A. Nickel, C. M. Bryce, J. A. Smith, R. E. Wheat, and V. Yovovich. "The golden age of bio-logging: how animal-borne sensors are advancing the frontiers of ecology." *Ecology*, **96 7**(), 2015, pp. 1741–53. URL: https://api.semanticscholar.org/CorpusID:26607170.

[65] R. W. Kays, M. C. Crofoot, W. Jetz, and M. Wikelski. "Terrestrial animal tracking as an eye on life and planet." *Science*, **348**(), 2015. URL: https://api.semanticscholar.org/CorpusID:206633498.

[66] S. M. Tomkiewicz, M. R. Fuller, J. G. Kie, and K. K. Bates. "Global positioning system and associated technologies in animal behaviour and ecological research." *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**(), 2010, pp. 2163–2176. URL: https://api.semanticscholar.org/CorpusID:2609381.

[67] D. D. Brown, R. W. Kays, M. Wikelski, R. P. Wilson, and A. P. Klimley. "Observing the unwatchable through acceleration logging of animal behavior." *Animal Biotelemetry*, **1**(), 2013, pp. 1–16. URL: https://api.semanticscholar.org/CorpusID:14578416.

[68] E. L. C. Shepard. "Identification of animal movement patterns using tri-axial accelerometry." *Endangered Species Research*, **10**(), 2008, pp. 47–60. URL: https://api.semanticscholar.org/CorpusID:32342781.

[69] S. J. Cooke, S. G. Hinch, M. Wikelski, R. D. Andrews, L. J. Kuchel, T. G. Wolcott, and P. J. Butler. "Biotelemetry: a mechanistic approach to ecology." *Trends in ecology & evolution*, **19 6**(), 2004, pp. 334–43. URL: https://api.semanticscholar.org/CorpusID:17085487.

[70] D. Tuia et al. "Perspectives in machine learning for wildlife conservation." *Nature Communications*, **13**(), 2021. URL: https://api.semanticscholar.org/CorpusID:239768714.

[71] T. S. Brandes. "Automated sound recording and analysis techniques for bird surveys and conservation." *Bird Conservation International*, **18**(), 2008, S163–S173. URL: https://api.semanticscholar.org/CorpusID:49583815.

[72] T. M. Aide, C. J. Corrada-Bravo, M. Campos-Cerqueira, C. Milan, G. Vega, and R. Alvarez. "Real-time bioacoustics monitoring and automated species identification." *PeerJ*, **1**(), 2013. URL: https://api.semanticscholar.org/CorpusID:1534630.

[73] Y. LeCun, Y. Bengio, and G. E. Hinton. "Deep Learning." In: 2015. URL: https://api.semanticscholar.org/CorpusID:3074096.

[74] D. Stowell. "Computational bioacoustics with deep learning: a review and roadmap." *PeerJ*, **10**(), 2021. URL: https://api.semanticscholar.org/CorpusID:245123578.

[75] D. Stowell. "Computational Bioacoustic Scene Analysis." In: 2018. URL: https://api.semanticscholar.org/CorpusID:64473642.

[76] C. Duhart, G. Dublon, B. D. Mayton, G. Davenport, and J. A. Paradiso. "Deep Learning for Wildlife Conservation and Restoration Efforts." In: 2019. URL: https://api.semanticscholar.org/CorpusID:197463451.

[77] G. Gubnitsky, Y. Mevorach, S. Gero, D. F. Gruber, and R. Diamant. "Automatic detection and annotation of eastern Caribbean sperm whale codas." *Scientific Reports*, **15**(), 2025. URL: https://api.semanticscholar.org/CorpusID:277782875.

[78] C. Bergler et al. "ANIMAL-SPOT enables animal-independent signal detection and classification using deep learning." *Scientific Reports*, **12**(), 2022. URL: https://api.semanticscholar.org/CorpusID:254904190.

[79] D. Stowell and M. D. Plumbley. "Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning." *PeerJ*, **2**(), 2014. URL: https://api.semanticscholar.org/CorpusID:3835905.

[80] R. Nathan, O. Spiegel, S. Fortmann-Roe, R. Harel, M. Wikelski, and W. M. Getz. "Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures." *Journal of Experimental Biology*, **215**(), 2012, pp. 986–996. URL: https://api.semanticscholar.org/CorpusID:10251682.

[81] G. Fehlmann, M. J. O'Riain, P. Hopkins, J. O'Sullivan, M. D. Holton, E. L. C. Shepard, and A. J. King. "Identification of behaviours from accelerometer data in a wild social primate." *Animal Biotelemetry*, **5**(), 2017, pp. 1–11. URL: https://api.semanticscholar.org/CorpusID:1887462.

[82] P. Warden and D. Situnayake. "TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers." In: 2019. URL: https://api.semanticscholar.org/CorpusID:226814321.

[83] M. Verhelst and B. Moons. "Embedded Deep Neural Network Processing: Algorithmic and Processor Techniques Bring Deep Learning to IoT and Edge Devices." *IEEE Solid-State Circuits Magazine*, **9**(), 2017, pp. 55–65. URL: https://api.semanticscholar.org/CorpusID:12602463.

[84] N. D. Lane, S. Bhattacharya, A. Mathur, P. Georgiev, C. Forlivesi, and F. Kawsar. "Squeezing Deep Learning into Mobile and Embedded Devices." *IEEE Pervasive Computing*, **16**(), 2017, pp. 82–88. URL: https://api.semanticscholar.org/CorpusID:27704761.

[85] S. Han, J. Pool, J. Tran, and W. J. Dally. "Learning both Weights and Connections for Efficient Neural Network." In: *Neural Information Processing Systems*. 2015. URL: https://api.semanticscholar.org/CorpusID:2238772.

[86] B. Jacob, S. Kligys, B. Chen, M. Zhu, M. Tang, A. G. Howard, H. Adam, and D. Kalenichenko. "Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference." *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (), 2017, pp. 2704–2713. URL: https://api.semanticscholar.org/CorpusID:39867659.

[87] L. Schulthess, S. Marty, M. Dirodi, M. D. Rocha, L. Rüttimann, R. H. R. Hahnloser, and M. Magno. "TinyBird-ML: An ultra-low Power Smart Sensor Node for Bird Vocalization Analysis and Syllable Classification." *2023 IEEE International Symposium on Circuits and Systems (ISCAS)*, (), 2023, pp. 1–5. URL: https://api.semanticscholar.org/CorpusID:260004271.

[88] R. Verma and S. Kumar. "AviEar: An IoT-Based Low-Power Solution for Acoustic Monitoring of Avian Species." *IEEE Sensors Journal*, **24**(), 2024, pp. 42088–42102. URL: https://api.semanticscholar.org/CorpusID:273842152.

[89] S. Gallacher, D. Wilson, A. J. Fairbrass, D. Turmukhambetov, M. Firman, S. Kreitmayer, O. M. Aodha, G. J. Brostow, and K. E. Jones. "Shazam for bats: Internet of Things for continuous real-time biodiversity monitoring." *IET Smart Cities*, (), 2021. URL: https://api.semanticscholar.org/CorpusID:244213736.

[90] K. F. A. Darras et al. "Eyes on nature: Embedded vision cameras for multidisciplinary biodiversity monitoring." *bioRxiv*, (), 2023. URL: https://api.semanticscholar.org/CorpusID:260354795.

[91] "Role of IoT and Edge Computing in addressing biodiversity and environmental monitoring." In: URL: https://api.semanticscholar.org/CorpusID:266364582.

[92] B. D. Mayton, G. Dublon, S. Palacios, and J. A. Paradiso. "TRUSS: Tracking Risk with Ubiquitous Smart Sensing." *2012 IEEE Sensors*, (), 2012, pp. 1–4. URL: https://api.semanticscholar.org/CorpusID:6882893.

[93] P. Chwalek, D. B. Ramsay, and J. A. Paradiso. "Captivates." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **5**(), 2021, pp. 1–32. URL: https://api.semanticscholar.org/CorpusID:237507919.

[94] Y. Zhang, M. Favero, P. Chwalek, S. Zhong, D. Lalanne, J. A. Paradiso, C. Miller, and A. Sonta. "Mind the Hazard: Modeling and Interpreting Comfort with Personalized Sensing." *Proceedings of the 11th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, (), 2024. URL: https://api.semanticscholar.org/CorpusID:273730844.

[95] P. Chwalek, S. Zhong, N. Perry, T. Liu, C. Miller, H. S. Alavi, D. Lalanne, and J. A. Paradiso. *A dataset exploring urban comfort through novel wearables and environmental surveys. figshare* https://doi.org/10.6084/m9.figshare.27625509. 2024.

[96] J. A. Stern, D. J. Boyer, and D. Schroeder. "Blink Rate: A Possible Measure of Fatigue." *Human Factors: The Journal of Human Factors and Ergonomics Society*, **36**(), 1994, pp. 285–297. URL: https://api.semanticscholar.org/CorpusID:23769702.

[97] K. C. Parsons. "Human Thermal Environments: The Effects of Hot, Moderate, and Cold Environments on Human Health, Comfort and Performance." In: 1999. URL: https://api.semanticscholar.org/CorpusID:109232791.

[98] J. W. Bradbury and S. L. Vehrencamp. *Principles of Animal Communication*. 2nd ed. Sinauer Associates, 2011. URL: https://api.semanticscholar.org/CorpusID:210462082.

[99] C. K. Catchpole and P. J. B. Slater. "Bird Song: Biological Themes and Variations." In: 1995. URL: https://api.semanticscholar.org/CorpusID:83055124.

[100] J. Sueur and A. Farina. "Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound." *Biosemiotics*, **8**(), 2015, pp. 493–502. URL: https://api.semanticscholar.org/CorpusID:16834742.

[101] W. H. Kirchner. "Acoustical communication in honeybees." *Apidologie*, **24**(), 1993, pp. 297–307. URL: https://api.semanticscholar.org/CorpusID:84974943.

[102] P. Laiolo. "The emerging significance of bioacoustics in animal species conservation." *Biological Conservation*, **143**(), 2010, pp. 1635–1645. URL: https://api.semanticscholar.org/CorpusID:84123556.

[103] K. F. A. Darras, F. Deppe, Y. Fabian, A. Kartono, A. Angulo, B. Kolbrek, Y. Mulyani, and D. M. Prawiradilaga. "High microphone signal-to-noise ratio enhances acoustic sampling of wildlife." *PeerJ*, **8**(), 2020. URL: https://api.semanticscholar.org/CorpusID:226238248.

[104] K. F. A. Darras, P. Batáry, B. J. Furnas, A. Celis-Murillo, S. L. V. Wilgenburg, Y. Mulyani, and T. Tscharntke. "Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis." *Journal of Applied Ecology*, (), 2018. URL: https://api.semanticscholar.org/CorpusID:90338571.

[105] E. C. Knight, K. C. Hannah, G. Foley, C. D. Scott, R. Brigham, and E. M. Bayne. "Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs." *Avian Conservation and Ecology*, **12**(), 2017. URL: https://api.semanticscholar.org/CorpusID:56168332.

[106] M. C. Serreze and R. G. Barry. "Processes and impacts of Arctic amplification: A research synthesis." *Global and Planetary Change*, **77**(), 2011, pp. 85–96. URL: https://api.semanticscholar.org/CorpusID:19437410.

[107] Arctic Monitoring and Assessment Programme (AMAP). *Arctic Climate Change Update 2021: Key Trends and Impacts. Summary for Policy-makers*. Tech. rep. Published May 2021. Tromsø, Norway: Arctic Council Secretariat, 2021, p. 16. URL: https://www.amap.no/documents/doc/arctic-climate-change-update-2021-key-trends-and-impacts.-summary-for-policy-makers/3508.

[108] S. Descamps, J. Aars, E. Fuglei, K. M. Kovacs, C. Lydersen, O. K. Pavlova, Å. Ø. Pedersen, V. T. Ravolainen, and H. Strøm. "Climate change impacts on wildlife in a High Arctic archipelago – Svalbard, Norway." *Global Change Biology*, **23**(), 2017. URL: https://api.semanticscholar.org/CorpusID:34897286.

[109] T. V. Callaghan et al. "Multi-Decadal Changes in Tundra Environments and Ecosystems: Synthesis of the International Polar Year-Back to the Future Project (IPY-BTF)." *AMBIO*, **40**(), 2011, pp. 705–716. URL: https://api.semanticscholar.org/CorpusID:4673429.

[110] K. L. Laidre et al. "Arctic marine mammal population status, sea ice habitat loss, and conservation recommendations for the 21st century." *Conservation Biology*, **29**(), 2015, pp. 724–737. URL: https://api.semanticscholar.org/CorpusID:17614752.

[111] D. J. Mennill. "Field tests of small autonomous recording units: an evaluation of in-person versus automated point counts and a comparison of recording quality." *Bioacoustics*, **33**(), 2024, pp. 157–177. URL: https://api.semanticscholar.org/CorpusID:269351383.

[112] C. A. Starbuck, L. M. DeSchepper, M. L. Hoggatt, and J. M. O'Keefe. "Tradeoffs in sound quality and cost for passive acoustic devices." *Bioacoustics*, **33**(), 2023, pp. 58–73. URL: https://api.semanticscholar.org/CorpusID:266277815.

[113] M. S. Brandstein and D. B. Ward. "Microphone Arrays Signal Processing Techniques and Applications." In: 2001. URL: https://api.semanticscholar.org/CorpusID:54115066.

[114] E. Verreycken, R. Simon, B. Quirk-Royal, W. Daems, J. R. Barber, and J. Steckel. "Bioacoustic tracking and localization using heterogeneous, scalable microphone arrays." *Communications Biology*, **4**(), 2021. URL: https://api.semanticscholar.org/CorpusID:243988937.

[115] G. Herold and E. Sarradj. "Open-source software for the application of microphone array methods." *Noise & Vibration Worldwide*, **48**(), 2017, pp. 44–51. URL: https://api.semanticscholar.org/CorpusID:113562054.

[116] Google Earth. *Google Earth*. Accessed: 2023-07-04. 2023. URL: https://www.google.com/earth/.

[117] A.-M. Klein, B. E. Vaissière, J. H. Cane, I. Steffan-Dewenter, S. A. Cunningham, C. Kremen, and T. Tscharntke. "Importance of pollinators in changing landscapes for world crops." *Proceedings of the Royal Society B: Biological Sciences*, **274**(), 2007, pp. 303–313. URL: https://api.semanticscholar.org/CorpusID:6720190.

[118] S. G. Potts, J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger, and W. E. Kunin. "Global pollinator declines: trends, impacts and drivers." *Trends in ecology & evolution*, **25 6**(), 2010, pp. 345–53. URL: https://api.semanticscholar.org/CorpusID:33073953.

[119] D. Goulson, E. Nicholls, C. Botías, and E. L. Rotheray. "Bee declines driven by combined stress from parasites, pesticides, and lack of flowers." *Science*, **347**(), 2015. URL: https://api.semanticscholar.org/CorpusID:206558985.

[120] S. A. Cameron, J. D. Lozier, J. P. Strange, J. B. Koch, N. Cordes, L. F. Solter, and T. L. Griswold. "Patterns of widespread decline in North American bumble bees." *Proceedings of the National Academy of Sciences*, **108**(), 2011, pp. 662–667. URL: https://api.semanticscholar.org/CorpusID:2776789.

[121] D. Goulson. "Effects of introduced bees on native ecosystems." *Annual Review of Ecology, Evolution, and Systematics*, **34**(), 2003, pp. 1–26. URL: https://api.semanticscholar.org/CorpusID:3892981.

[122] A. H. Abrahamovich, N. B. Díaz, and M. Lucía. "Identificación de las "abejas sociales" del género Bombus (Hymenoptera, Apidae) presentes en la Argentina: clave pictórica, diagnosis, distribución geográfica y asociaciones florales." *Revista De La Facultad De Agronomia De La Universidad Del Zulia*, **106**(), 2007, pp. 165–176. URL: https://api.semanticscholar.org/CorpusID:83815785.

[123] C. L. Morales, M. P. Arbetman, S. A. Cameron, and M. A. Aizen. "Rapid ecological replacement of a native bumble bee by invasive species." *Frontiers in Ecology and the Environment*, **11**(), 2013, pp. 529–534. URL: https://api.semanticscholar.org/CorpusID:86469248.

[124] IUCN. *Bombus dahlbomii*. https://www.iucnredlist.org/species/21215142/100240441. Accessed: 2025-06-03. 2024.

[125] C. Carvell, W. R. Meek, R. F. Pywell, D. Goulson, and M. Nowakowski. "Comparing the efficacy of agri-environment schemes to enhance bumble bee abundance and diversity on arable field margins." *Journal of Applied Ecology*, **44**(), 2006, pp. 29–40. URL: https://api.semanticscholar.org/CorpusID:16488687.

[126] T. J. Wood, J. M. Holland, and D. Goulson. "Pollinator-friendly management does not increase the diversity of farmland bees and wasps." *Biological Conservation*, **187**(), 2015, pp. 120–126. URL: https://api.semanticscholar.org/CorpusID:3948517.

[127] L. S. M. Sugai, C. Desjonquères, T. S. F. Silva, and D. Llusia. "A roadmap for survey designs in terrestrial acoustic monitoring." *Remote Sensing in Ecology and Conservation*, **6**(), 2019. URL: https://api.semanticscholar.org/CorpusID:210626486.

[128] G. B. Celli and B. Maccagnani. "Honey bees as bioindicators of environmental pollution." In: 2003. URL: https://api.semanticscholar.org/CorpusID:130663525.

[129] A. Gradiek, G. Slapniar, J. Sorn, M. Lutrek, M. Gams, and J. Grad. "Predicting species identity of bumblebees through analysis of flight buzzing sounds." *Bioacoustics*, **26**(), 2017, pp. 63–76. URL: https://api.semanticscholar.org/CorpusID:44256946.

[130] A. I. S. Ferreira, N. F. F. da Silva, F. N. Mesquita, T. C. Rosa, V. H. Monzón, and J. N. Mesquita-Neto. "Automatic acoustic recognition of pollinating bee species can be highly improved by Deep Learning models accompanied by pre-training and strong data augmentation." *Frontiers in Plant Science*, **14**(), 2023. URL: https://api.semanticscholar.org/CorpusID:258113088.

[131] Z. J. Miller. "What's the Buzz About? Progress and Potential of Acoustic Monitoring Technologies for Investigating Bumble Bees." *IEEE Instrumentation & Measurement Magazine*, **24**(), 2021, pp. 21–29. URL: https://api.semanticscholar.org/CorpusID:238221990.

[132] N. E. Miller-Struttmann, J. C. Geib, J. D. Franklin, P. G. Kevan, R. M. Holdo, D. Ebert-May, A. Lynn, J. A. Kettenbach, E. Hedrick, and C. Galen. "Functional mismatch in a bumble bee pollination mutualism under climate change." *Science*, **349**(), 2015, pp. 1541–1544. URL: https://api.semanticscholar.org/CorpusID:46616411.

[133] Y. Gong, Y.-A. Chung, and J. R. Glass. "AST: Audio Spectrogram Transformer." *ArXiv*, **abs/2104.01778**(), 2021. URL: https://api.semanticscholar.org/CorpusID:233024831.

[134] J. F. Gemmeke, D. P. W. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter. "Audio Set: An ontology and human-labeled dataset for audio events." *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (), 2017, pp. 776–780. URL: https://api.semanticscholar.org/CorpusID:21519176.

[135] N. E. Miller-Struttmann, D. Heise, J. Schul, J. C. Geib, and C. Galen. "Flight of the bumble bee: Buzzes predict pollination services." *PLoS ONE*, **12**(), 2017. URL: https://api.semanticscholar.org/CorpusID:5601945.

[136] *MAX78000 Artificial Intelligence Microcontroller with Ultra-Low-Power Convolutional Neural Network Accelerator*. Tech. rep. 19-100868; Rev 1. Document Number: 19-100868, Revision: 1. 2021.

[137] H. Iqbal. *PlotNeuralNet*. https://github.com/HarisIqbal88/PlotNeuralNet. Accessed: 2025-07-23. 2018.

[138] A. Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library." *ArXiv*, **abs/1912.01703**(), 2019. URL: https://api.semanticscholar.org/CorpusID:202786778.

[139] Analog Devices Inc. *ai8x-training: Model Training for ADI's MAX78000 and MAX78002 Edge AI Devices*. https://github.com/analogdevicesinc/ai8x-training. Last updated November 7, 2024. 2024.

[140] L. Zhao, T. Torchet, M. Payvand, L. Kriener, and F. Moro. "Quantizing Small-Scale State-Space Models for Edge AI." *ArXiv*, **abs/2506.12480**(), 2025. URL: https://api.semanticscholar.org/CorpusID:279402041.

[141] Y. Liang, Z. Wang, X. Xu, Y. Tang, J. Zhou, and J. Lu. "MCUFormer: Deploying Vision Tranformers on Microcontrollers with Limited Memory." *ArXiv*, **abs/2310.16898**(), 2023. URL: https://api.semanticscholar.org/CorpusID:264490968.

[142] S. Wroe, C. McHenry, and J. Thomason. "Bite club: comparative bite force in big biting mammals and the prediction of predatory behaviour in fossil taxa." *Proceedings of the Royal Society B: Biological Sciences*, **272**(), 2005, pp. 619–625. URL: https://api.semanticscholar.org/CorpusID:5803993.

[143] B. van Valkenburgh and C. B. Ruff. "Canine tooth strength and killing behaviour in large carnivores." *Journal of Zoology*, **212**(), 1987, pp. 379–397. URL: https://api.semanticscholar.org/CorpusID:86701588.

[144] M. B. Meers. "Maximum Bite Force and Prey Size of Tyrannosaurus rex and Their Relationships to the Inference of Feeding Behavior." *Historical Biology*, **16**(), 2002, pp. 1–12. URL: https://api.semanticscholar.org/CorpusID:86782853.