Enhancing the Recording and Analysis of Antarctic Soundscapes

Emiliano Acevedo*, Maria Noel Espinosa*, Ilana Stolovas*, Martín Rocamora*†, Leonardo Steinfeld*
*Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay
†Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain
Email: {eacevedo, maria.espinosa, ilana.stolovas, leo}@fing.edu.uy, martin.rocamora@upf.edu

***Abstract*—Since 2020, the 'Antarctic Soundscapes' (AS) project has aimed to study anthropogenic impact in Antarctica through sound by deploying a network of AudioMoth recorders at key locations across the Fildes Peninsula (King George Island). These devices continuously capture bioacoustic data during summer campaigns, providing insights into the impacts of human activity on this ecosystem. However, the project faces two significant challenges: the limited autonomy of the recording devices, which typically require battery replacements every 15-20 days, and the extensive human effort needed to process and analyze months of continuous recordings. To address these limitations, we enhanced the AudioMoth-based recording system by integrating a solar panel, rechargeable battery, and charge controller, enabling continuous operation for up to 84 days without human intervention during the 2022-2023 campaign. Additionally, we implemented a machine learning-based audio tagging system that automates the identification of anthropogenic sounds, reducing the time and resources required for manual analysis. These contributions represent a significant step towards fully autonomous monitoring and analysis of Antarctic soundscapes, enabling more efficient and sustainable research.**

I. Introduction

Understanding the sound impact of human activity in Antarctica is crucial for protecting its environment and wildlife [1]. As a vulnerable and unique ecosystem, Antarctica requires protection against acoustic pollution, which poses risks to its fauna [2]–[4]. Studying these impacts not only helps establish appropriate noise limits to safeguard the ecosystem but also holds scientific significance by providing critical insights into human influence on protected areas.

The 'Antarctic Soundscapes' (AS) project, led by Dr. Lucía Ziegler of the Department of Ecology and Environmental Management at Universidad de la República (CURE, Centro Universitario Regional Este), focuses on studying anthropogenic impacts on the Antarctic sound environment [5]. Since 2020, the project has deployed a series of AudioMoth (AM) [6] acoustic recorders across nine points on the Fildes Peninsula (King George Island) (See Figure 1).

The methodology followed by AS is a structured three-stage process. Recording devices are prepared and deployed in key locations. These devices are setted on a predefined schedule, recording the first five minutes of each hour—a widely practice in bioacoustics [7]. Then, devices are serviced every 15 days to prevent battery depletion. Finally, a post-campaign analysis is done to identify anthropogenic sound sources.

The AS methodology encounters two primary challenges. First, the limited battery life of AM devices, which need replacements every 15-20 days, poses logistical difficulties for field operations in remote areas. Second, analyzing the extensive audio data collected during campaigns demands substantial human effort, making the process both time-consuming and resource-intensive.

The project implements three key contributions. First, we extend the autonomy of AM recording devices by integrating solar panels, rechargeable batteries, and optimized power management systems, ensuring uninterrupted operation. We also increased SD card capacity to accommodate extended recording periods. Field tests during the 2022-2023 Antarctic Summer Campaign confirmed the prototype's effectiveness, achieving continuous operation for up to 84 days. Second, we evaluated the feasibility of solar power in Antarctica by employing a custom-built data logger to monitor the power system and correlating the collected data with irradiance levels, providing insights into sustainable energy solutions tailored for remote and extreme environments. Finally, we enhanced post-campaign audio analysis by developing a machine learning-based audio tagging system that automates the detection of anthropogenic sounds, such as aerial, terrestrial, and aquatic vehicles, significantly reducing manual effort.

Collectively, our contributions represent significant advancements in data acquisition technology for Antarctic research, addressing critical challenges and opening new possibilities for conducting impactful scientific investigations in one of the most extreme environments on Earth.

II. Recording Device

A. Audiomoth (AM)

AudioMoth, developed by Open Acoustic Devices, is a low-cost, energy-efficient recording device designed to capture high-quality audio economically, supporting extended recording intervals due to its minimal power consumption. Its affordability allows the deployment of multiple devices, making it a valuable tool for bioacoustic research and environmental monitoring. Leveraging its advantages, the AS project has used AM extensively in recent Antarctic campaigns, establishing it as the primary choice for this project. A literature review

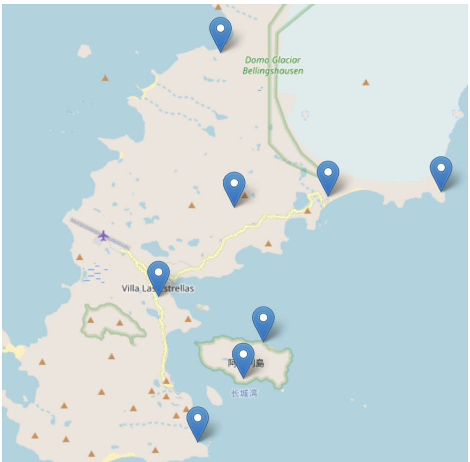

Fig. 1. Deployment locations of AM devices across the Fildes Peninsula, King George Island.

was conducted to evaluate alternatives such as the Song Meter Mini and BAR-LT, but AM stands out as significantly more cost-effective, making it especially suitable for projects that require deploying a large number of units.

The AM is based on the Silicon Labs Gecko EFM32 processor and includes an integrated MEMS (Micro-Electro-Mechanical System) microphone. The microcontroller contains internal SRAM memory and stores acquired recordings uncompressed (16-bit linear PCM format) in `.wav` format on an external SD card. The AM was designed to capture a wide range of frequencies, allowing for the selection of the sampling frequency between fixed values from 8k to 384k samples per second. A notable feature of the device is its small size (58 x 48 x 15 mm). The device features a wide input voltage range and typically is powered at 3.6V nominal, utilizing three AA NiMH rechargeable batteries. The AM also includes red and green LEDs to provide users with operational feedback, such as indicating recording statusor SD card issues.

B. Power profile

The AM operates in three modes: Continuous recording, USB/OFF (for PC connection or turning off), and Custom recording. The AS project uses the Custom mode, configured to record the first five minutes of each hour. In this mode, as shown in Figure 2, the AM alternates between two primary states: a low-power sleeping state to conserve energy and a recording state, where it captures the soundscape.

Next, we analyze current consumption in more detail based on the consumption profile shown in Figure 2. During the sleeping phase (T_s), the device enters an ultra-low power mode (EM4), drawing a minimal average current, I_s, which is approximately 0.092 µA. In this mode, periodic interrupts from a timer briefly wake the device to check the recording schedule (denoted as *Sleeping* in Figure 3).

Before transitioning to the recording phase, the device enters a preparation stage, lasting T_{pre}, during which it con-

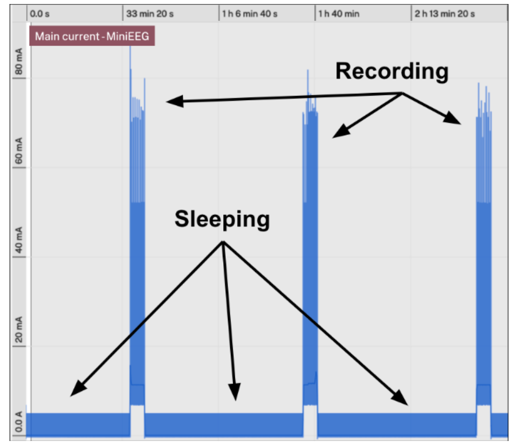

Fig. 2. AM's power consumption profile showing the alternation between Sleeping (low power mode) and Recording (audio capture with higher current consumption).

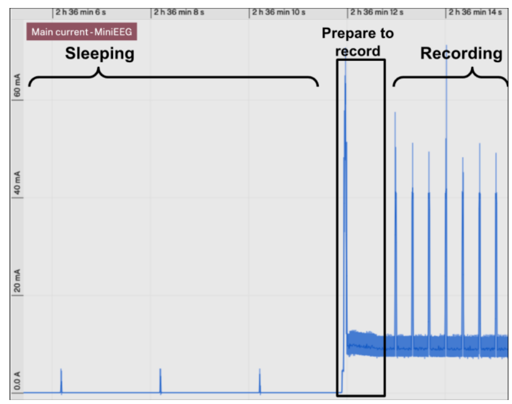

Fig. 3. AM's power consumption profile in Custom mode: Sleeping (low power), Preparation to Record (system initialization), and Recording (audio capture with data transfer spikes).

sumes an average current, I_{pre} (denoted as *Prepare to record* in Figure 3). This stage involves initializing components such as the microphone, memory buffers, and performing power checks. Power consumption during this phase is higher, with an average current of around 12.7 mA over a duration of 750 ms for a SD card of 32 GB.

During the recording phase, the device captures audio data through the microphone, stores it in SRAM, and periodically transfers it to an SD card via SPI. This phase consumes I_g during the recording interval T_g (denoted as *Recording* in Figure 3), with average current of about 11 mA.

Since the recording time is fixed, any preparation time is subtracted from the total available sleep time, ensuring that the overall cycle remains within the designated time frame.

Based on the measured current consumption and time spent in each phase, we developed a simple yet effective energy model for the AM, which can predict energy consumption under various configurations, presented in Eq. (1):

$$Q_{hour} = T_g I_g + (T_s - T_{pre}) I_s + T_{pre} I_{pre} \quad (1)$$

TABLE I
MEASURED AND MODELED CURRENT CONSUMPTION VALUES FOR CONFIGURATIONS.

Sleeping		Preparing		Recording		I_{mod}	I_{meas}	error
T_s	I_s	T_{pre}	I_{pre}	T_g	I_g			
3300	0.092	3	12.7	300	11.4	1.045	1.04	0.47%
55	0.029	3	23.3	5	20	2.857	2.86	-0.11%

To validate the model, we measured the current consumption and the corresponding times for two different configurations using a 32 GB SD card, as shown in Table I (time is expresed in s and current in mA.). The first experiment corresponds to the typical schedule (five minutes of recording per hour) and the second one the schedule was modified to five seconds of recording per minute.

Based on the measurement values presented in Table I, the total charge for one cycle was calculated. The average current consumption per cycle (I_{mod}) was then determined by dividing the total charge by the cycle period. This calculated value was compared with direct measurements taken throughout the period (I_{meas}), yielding a relative error below 1%. This result confirms the accuracy of the proposed model. Interestingly, the results show that for more frequent recording, while maintaining the recording-to-sleep ratio, the pre-recording charge consumption becomes increasingly significant, thereby raising the overall energy consumption.

The power consumption of 32 GB, 64 GB, and 128 GB Kingston SD cards was measured under the typical AM configuration. The analysis showed minimal variation in energy consumption across different SD card sizes in this setup.

The analysis provides a daily energy consumption of 24.28 mAh for the AM, which is crucial for accurately sizing its power supply and understanding its energy requirements.

C. Enhanced Device: Extending Autonomy with Solar Power

To extend the autonomy of AM devices, a solar-powered system was developed, consisting of an Adafruit bq24074 charger, a 6V 1W solar panel, and a 3.7V 2500mAh Li-Po battery. Considering Antarctica's extreme conditions, a conservative approach was taken by over-sizing the solar panel and battery. Despite this, the solution proved to be both cost-effective and reliable. Experiments conducted in Antarctica validated the design and provided data for future improvements, aiming to reduce the safety margin in redesigns.

III. EXPERIMENTS

To evaluate the solar-powered design's performance and its feasibility for autonomous audio recording in extreme conditions, field tests were conducted during the 2022-2023 Antarctic Summer Campaign in two main experiments.

A. Short-Term Experiment

A custom data logger was built using an Arduino Uno with an RTC/SD card shield to sample signals with a timestamp. Powered by an independent source, it collects time-based data without affecting the AM's operation. The logger samples the PGOOD and CHG signals every eight seconds. PGOOD indicates sufficient solar irradiance for battery charging, while CHG shows active charging. When PGOOD is high and CHG is low, it confirms the battery is fully charged, demonstrating effective energy storage.

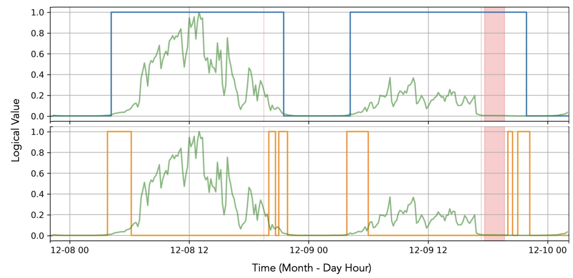

Fig. 4. Time series plot showing the digital pins PGOOD (blue) and CHG (orange), alongside the normalized irradiance time series (green). Red bands indicate missing data in the irradiance series.

The analysis, as shown in Figure 4, demonstrated that the device discharged during periods of no irradiance (night) and efficiently recovered lost energy when sunlight returned. Table II presents the irradiance values for the analyzed days, comparing them to the mean and worst days of the campaign. Notably, the system achieved a fully charged battery on both a good day (08/12, the best day of the campaign) and a relatively bad day (09/12), demonstrating its robustness under varying irradiance conditions. On average, the PGOOD pin was active for 16.2 hours per day, while the CHG pin was active for 4.1 hours daily, highlighting the over-dimensioning of the solar-powered solution to ensure reliable operation even in challenging Antarctic conditions.

TABLE II
STATISTICS OF DAILY IRRADIANCE. ALL VALUES ARE IN W/m^2

Date	Maximum	Mean	Std. Deviation
Worst Day	122.5	39.3	41.2
Mean Day	525.5	143.6	160.2
08/12	1026.0	303.9	314.5
09/12	364.5	89.6	101.5

B. Long-Term Experiment

One key experiment involved deploying a solar-powered AM alongside an original device used in the AS project. Both devices were tested at Ardley Island from December 3, 2022, to March 4, 2023, and housed in weather-resistant enclosures. The prototype achieved 84 days of continuous operation, far exceeding the estimated 16-day autonomy of the original device. The test demonstrated that the prototype's limitations were due to storage capacity rather than power, as the SD card became full, highlighting the potential for even longer autonomy with higher-capacity storage.

IV. Post-Campaign Audio Analysis

Following the summer campaign, the AS project faces another significant challenge: analyzing two months' worth of audio recordings. This task heavily relies on human personnel for manual analysis, making it time-consuming and resource-intensive. To address this, we developed a machine learning-based audio tagging system capable of processing Antarctic soundscape recordings and automatically identifying anthropogenic sources, specifically aircraft and motorized vehicles, including boats.

Due to the limited availability of labeled anthropogenic sound data for training a model, we created a synthetic dataset using Scaper [8], combining foreground (anthropogenic sounds) and background (ambient Antarctic sounds) audios. Backgrounds were easily sourced from AS project recordings, as most lacked human-made sounds, while foregrounds were selected from public datasets like AudioSet [9] and Freesound [10]. Inspired by a DCASE competition [11], we synthesized 6,500 ten-second audio files, simulating realistic Antarctic soundscapes to provide essential training data for training audio tagging model.

TABLE III
Performance comparison of the baseline and different models for audio tagging

Method	Accuracy	PR_AUC	EER Motors	EER Air
Baseline	0.308	0.368	0.432	0.402
MobileNetV2	0.913	0.970	0.045	0.079
VGG13	0.783	0.867	0.088	0.215
AST	0.487	0.838	0.231	0.152
PaSST	0.515	0.903	0.187	0.104

To address this audio tagging problem, we used log-mel spectrograms as input, optimizing the parameters to $n\_mels = 128$, $window\ length = 2048$, and $hop\ length = 511$, based on a baseline model using MobileNetV1 from the DCASE 2019 framework. We then evaluated various models for audio tagging, including MobileNetV2, VGG13, PaSST and AST; using Accuracy, Precision-Recall AUC and Equal Error Rate (EER) as evaluation metrics. In order to mitigate overfitting, we applied data augmentation techniques such as Mixup, Random Erasing, Scaling, Shifting, and Rotating. The performance results are summarized in Table III.

MobileNetV2 outperformed other models in audio tagging tasks, showing strong performance on the synthetic dataset. However, real-world testing highlighted challenges from domain shifts, particularly noisy and windy conditions, leading to increased false positives. These issues, along with the limitation of recording only the first five minutes of each hour, make detecting sparse anthropogenic sounds difficult. Examples of detections and a model testing tool are available on the project site[1].

Adjustments to both the recording methodology and dataset composition are anticipated to enhance the model's performance in real-world applications.

[1]https://pfcserena.github.io/

V. Conclusions

The advancements in the AS project significantly enhance the viability of long-term bioacoustic monitoring in the harsh Antarctic environment. By integrating a solar-powered system into the AM-based devices, we extended the autonomy of the recorders to 84 days without the need for battery replacements, addressing a critical limitation of previous setups. The proposed solar power solution proved both reliable and cost-effective, as confirmed by an independent data logger monitoring the charging process, ensuring optimal energy system performance throughout the campaign. In addition the integration of machine learning for automatic audio tagging represented a key development in reducing the time and resources required for manual analysis of the collected recordings. This approach not only streamlined the processing of continuous recordings but also enhanced the efficiency of identifying anthropogenic impacts on the Antarctic ecosystem.

These innovations collectively contribute to a more sustainable, autonomous monitoring solution, empowering the AS project to track the effects of human activity on Antarctic soundscapes more effectively. With these improvements, the project is well-positioned to provide valuable insights that will inform future research and conservation efforts in this ecologically sensitive region.

References

[1] T. Tin and et al., "Impacts of local human activities on the antarctic environment," *Antarctic Science*, vol. 21, no. 1, p. 3–33, 2009.

[2] A. C. M. Kok and et al., "How chronic anthropogenic noise can affect wildlife communities," *Frontiers in Ecology and Evolution*, vol. 11, 2023.

[3] J. G. Morley Erica L. and R. A. N., "The importance of invertebrates when considering the impacts of anthropogenic noise proc. r. soc. b.2812013268320132683," 2014.

[4] G. Shannon, M. F. McKenna, L. M. Angeloni, K. R. Crooks, K. M. Fristrup, E. Brown, K. A. Warner, M. D. Nelson, C. White, J. Briggs, S. McFarland, and G. Wittemyer, "A synthesis of two decades of research documenting the effects of noise on wildlife," *Biol. Rev. Camb. Philos. Soc.*, vol. 91, no. 4, pp. 982–1005, Nov. 2016.

[5] L. Ziegler and A. Soutullo, "Anthropogenic noise in terrestrial antarctica: challenges and opportunities," *Polar Research*, vol. 43, Apr 2024.

[6] A. P. Hill, P. Prince, J. L. Snaddon, C. P. Doncaster, and A. Rogers, "Audiomoth: A low-cost acoustic device for monitoring biodiversity and the environment," *HardwareX*, vol. 6, p. e00073, 2019.

[7] L. Sugai, C. Desjonquères, T. Silva, and D. Llusia, "A roadmap for survey designs in terrestrial acoustic monitoring," *Remote Sensing in Ecology and Conservation*, vol. 6, 11 2019.

[8] J. Salamon, D. MacConnell, M. Cartwright, P. Li, and J. P. Bello, "Scaper: A library for soundscape synthesis and augmentation," in *2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2017, pp. 344–348.

[9] 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," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 776–780.

[10] F. Font, G. Roma, and X. Serra, "Freesound technical demo," in *Proceedings of the 21st ACM international conference on Multimedia*, 2013, pp. 411–412.

[11] A. Mesaros, T. Heittola, E. Benetos, P. Foster, M. Lagrange, T. Virtanen, and M. D. Plumbley, "Detection and classification of acoustic scenes and events: Outcome of the DCASE 2016 challenge," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 2, pp. 379–393, Feb 2018.