Markov Chain Wave Generative Adversarial Network for Bee Bioacoustic Signal Synthesis

Kumudu Harshani Samarappuli , Iman Ardekani * , Mahsa Mohaghegh , Abdolhossein Sarrafzadeh

Posted Date: 3 December 2025

doi: 10.20944/preprints202512.0214.v1

Keywords: bee bioacoustic; synthetic data; generative adversarial networks; Markov Chain; smart beekeeping

 doi:10.20944/preprints202512.0214.v1

Article

Markov Chain Wave Generative Adversarial Network for Bee Bioacoustic Signal Synthesis

Kumudu Samarappuli [1], Iman Ardekani [1,2,*], Mahsa Mohaghegh [3], Abdolhossein Sarrafzadeh [4]

[1] Unitec Institute of Technology, Auckland, New Zealand
[2] The University of Notre Dame Australia, Sydney, NSW, Australia
[3] Auckland University of Technology, Auckland, New Zealand
[4] Old Dominion University, VA, USA
* Correspondence: iman.ardekani@nd.edu.au

Abstract

This paper presents a framework for synthesizing bee bioacoustic signals associated with hive events. While existing approaches like WaveGAN have shown promise in audio generation, they often fail to preserve the subtle temporal and spectral features of bioacoustic signals critical for event-specific classification. The proposed method, MCWaveGAN, extends WaveGAN with a Markov Chain refinement stage, producing synthetic signals that more closely match the distribution of real bioacoustic data. Experimental results show that this method captures signal characteristics more effectively than WaveGAN alone. Furthermore, when integrated into a classifier, synthesized signals improved hive status prediction accuracy. These results highlight the potential of the proposed method to alleviate data scarcity in bioacoustics and support intelligent monitoring in smart beekeeping, with broader applicability to other ecological and agricultural domains.

Keywords: bee bioacoustic; synthetic data; generative adversarial networks; Markov Chain; smart beekeeping

1. Introduction

Bioacoustic signals convey critical information about behavior, health, and interactions of living organisms with the environment [1]. With recent advancements in AI, these signals have become increasingly important in applications such as ecological monitoring, medicine, and agriculture [2]. A key challenge is the scarcity of event-specific sounds, where the acoustic signatures are essential for identifying and distinguishing particular biological events. Bioacoustic signals are comprised of unique characteristics, including temporal, spectral, and structural properties, which help to extract meaningful insights for real-world applications. There are various categories of bioacoustics; however, this study focuses explicitly on ***Bee Bioacoustics*** due to its vital significance in sustainable pollination, ecosystem health, and agricultural productivity.

Bee bioacoustics is particularly interesting because it helps to improve sustainable pollination, and that is fundamental for ensuring global food security [2]. Recent studies have shown that bee bioacoustics can support efficient hive and behavioural management by detecting changes in sound patterns associated with stress, swarming, or queenlessness. Distinct acoustic signatures have been observed for key colony events such as queen presence, queen absence, and swarming, each carrying important implications for colony stability and productivity. For example, queen-less hives represent characteristic shifts in acoustic patterns that can be detected well before visual inspection, while swarming events generate specific pre-swarm signals that serve as early indicators for imminent colony reproduction and potential hive division [3]. This represents bee bioacoustics as an active area of research, attracting attention across agriculture, AI, and environmental monitoring domains. This research focuses specifically on bee bioacoustics as a case study, due to its critical ecological and

 doi:10.20944/preprints202512.0214.v1

agricultural importance. However, analyzing bee Bioacoustic signals comes with unique technical challenges.

With recent advancements in AI and ML technologies, Bioacoustic signals are now being effectively analyzed and applied across a wide range of domains such as apiculture, wildlife conservation, medical diagnosis, and more [2]. Developing robust machine learning models for bioacoustics analysis requires large volumes of clean, representative data, which is extremely challenging to obtain. Collecting sufficient high-quality bioacoustics signals is difficult due to contamination from environmental noise such as wind or water sounds, the labor-intensive nature of data collection, and the high associated costs [4–6]. This scarcity of quality data presents a significant bottleneck for both research and practical deployment of bioacoustics systems. A key challenge is the scarcity of event-specific sounds, where the acoustic signature is essential for identifying and distinguishing particular biological events. These challenges are equally evident in *Bee Bioacoustics*, where data quality and sufficient data availability are critical factors for advancing research and applications in this field [6].

A practical solution is to generate *synthetic* bioacoustics datasets [6] that are reliable, balanced, and capture *event-specific* characteristics. This is especially valuable for rare or hard-to-capture events that are essential for intelligent monitoring systems. While Generative Adversarial Networks (GANs) have shown potential in audio synthesis, they often fail to fully capture the long-term temporal dependencies inherent in bee bioacoustics. This limitation becomes particularly critical when generating event-specific sounds, where acoustic signals carry the essential signatures needed to accurately identify and distinguish events. In this paper, we investigate the use of Markov Chains (MC) as a refinement strategy for synthesizing event-specific Bioacoustic signals generated by a specialized GAN variant, WaveGAN. MC is a well-established mathematical framework widely used in predictive modeling [7]. It has also been successfully integrated with GANs to better handle high-dimensional data and produce more realistic and reliable outputs [8,9]. However, to the best of our knowledge, the integration of MC with the WaveGAN (MCWaveGAN - Markov Chain Wave Generative Adversarial Network) approach remains underexplored in the context of ***Bee Bioacoustic*** signal generation. In this work, we aim to bridge this gap by addressing the challenges posed by limited high-quality datasets in real-world bee bioacoustics intelligence applications.

Considering the motivation outlined above, this study focuses on addressing one of the main challenges in ***Bee Bioacoustics*** research; the lack of noise-free and sufficient real-world data by generating synthetic signals. The primary objective is to develop and evaluate a generative model with the capability to synthesize realistic bee signals that accurately capture event-specific acoustic patterns while preserving temporal dynamics. To achieve this, first, this study reviews existing synthetic data generation methods for bioacoustic data and identifies their limitations. Then, design and experiment with machine learning models capable of capturing unique characteristics associated with specific bee hive events, such as the presence or absence of a queen, and develop a robust model to generate synthetic bee bioacoustic signals. Additionally, produce a synthetic bee bioacoustic signal dataset and validate its realism using quantitative measures such as classification performance, and qualitative analyses, including spectral and statistical similarity to real bee sounds.

2. Materials and Methods

In ***Bee Bioacoustics*** research, scarcity and noise of real-world datasets are a critical challenge when developing and training robust machine learning models. To overcome it, this study proposes a framework that can generate synthetic event-specific bee bioacoustic signals that are close to real bee signals. Specifically, this framework aims to produce synthetic recordings representing queen-present and queen-absent hive conditions.

The proposed model consists of two main components: a WaveGAN module, which generates initial synthetic signals, and a Markov Chain module, which refines them to preserve temporal dependencies and acoustic fidelity. The proposed hybrid approach first uses the WaveGAN module to generate bee signals, which are then refined by Markov Chain module as depicted in Figure 1. In

 doi:10.20944/preprints202512.0214.v1

addition, we include a lightweight classification stage, making the framework a complete pipeline for event-specific bioacoustic analysis.

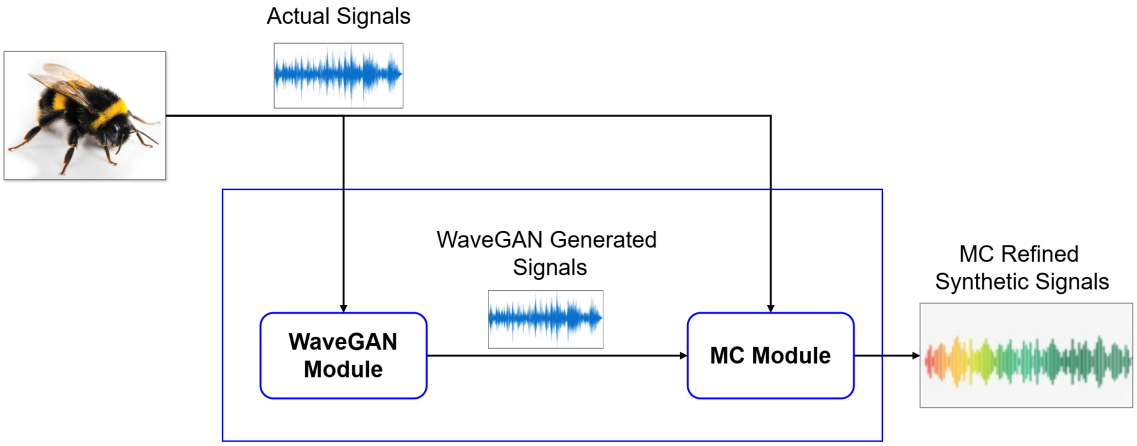


Figure 1. High-level architecture of the proposed MCWaveGAN Framework for synthetic bee bioacoustic signal generation.

2.1. WaveGAN Module

First proposed by Goodfellow et al. in 2014 [10], GANs are regarded as one of the most promising generative models for producing data. The ability of GANs to learn from real data has made them superior in synthetic data generation compared to traditional and other ML-based approaches. Today, GANs are most popular due to their widespread applicability and exceptional capability to model complex and high-dimensional data [11]. GAN-based models consist of a *generator (G)*, which produces synthetic samples, and a *discriminator (D)*, which distinguishes between real and generated data. The adversarial training process between *G* and *D* enables GANs to learn the underlying distribution of complex datasets.

In our proposed framework, the *WaveGAN Module* employs WaveGAN as the baseline model for generating synthetic bioacoustic signals [12]. It is recognized as the first GAN-based model designed for unsupervised raw audio waveform generation [12]. Its architecture was adopted from DCGAN, originally designed for image synthesis [12]. Technically, WaveGAN adapts the DCGAN architecture by reshaping convolutional layers to directly process audio waveforms [13]. The generator of WaveGAN produces floating-point audio signals, while the discriminator employs 1D convolutions to distinguish between real and synthetic waveforms. Training stability is enhanced through the use of Wasserstein GAN with Gradient Penalty (WGAN-GP) [12]. WaveGAN has demonstrated promising results across various audio synthesis domains. It has been applied to speech generation (eg, spoken digits) [14] from raw audio waveform and to generate drum beats and instrumental sound effects for the entertainment domain [15]. In the bioacoustics domain, WaveGAN has been used to generate bird calls and dolphin sounds, demonstrating its capability to model natural acoustic patterns [16]. However, the review of the literature indicates that the application or research of WaveGAN to bee bioacoustics has not yet been explored.

Although WaveGAN can capture general acoustic properties, it struggles to preserve essential event-specific sound characteristics, which are critical for identifying acoustic signatures. These characteristics are particularly critical for downstream classification tasks where the goal is to identify event-specific sounds (e.g., distinguishing hive events). To overcome these limitations, we introduce a *Markov Chain Module*, which refines the WaveGAN outputs to better capture event-related signal characteristics.

 doi:10.20944/preprints202512.0214.v1

2.2. Markov Chain Module

A **Markov Chain** is a mathematical model to represent a randomly changing system and its state transitions according to a set of probabilistic rules where the future states only depend on the current state and not on the sequence of past states (*Markov property*) [7]. In recent studies, integration of **Markov Chain process** with GANs has shown great advancements in various domains, particularly in handling sequential and time-dependent data. MCGAN represents a combination of the strengths of the *Markov Chain* process with the adversarial learning capability of *GANs*. One of the most prominent applications of MCGAN is to solve Bayesian inverse problems in physics and engineering [9], and it has shown a great performance compared to traditional methods. There, MCGAN generates high-quality outputs from complex, high-dimensional input data with higher accuracy and faster computational efficiency [9]. Incorporation of the MC process has helped to improve convergence and produce stable results.

Since WaveGAN generation does not guarantee to preserve temporal dynamics and other finer details of the original data, the MC module is used to address that limitation in this proposed method. The key idea behind MC is to design a process that moves step by step between samples, where each sample depends on the previous one. This sequential progression allows the target distribution to be accurately represented [17]. Among the available implementations, the Metropolis–Hastings (MH) algorithm [17] is the most widely adopted, and it is integrated into our MC module. Once the WaveGAN converges, candidate features $\mathbf{v}$ are drawn from the pool of the generated signals, and the MH algorithm decides whether to accept or reject them. During this process, a transition vector $\boldsymbol{\theta}_{\text{prop}}$ is defined as:

$$\boldsymbol{\theta}_{\text{prop}} = (1-\beta)(\mathbf{v}-\mathbf{s}_t) + \beta(\mathbf{v}-\mathbf{r}_t), \tag{1}$$

where $\mathbf{s}_t$ is the current synthetic signal at step t, $\mathbf{r}_t$ is the real audio signal at step t, and $\mathbf{v}$ is the candidate feature sampled from the WaveGAN-generated audio set. The balance parameter β controls the trade-off between similarity to the real features ($\mathbf{r}_t$) and smooth progression from the current synthetic step ($\mathbf{s}_t$). Here, the acceptance probability is given by:

$$\alpha = \min\left(1, \frac{p(\boldsymbol{\theta}_{\text{prop}})}{p(\boldsymbol{\theta}_t)}\right), \tag{2}$$

where $p(\cdot)$ represents the Gaussian prior [18] that models the transition probabilities of real audio features. A random number u from a uniform distribution is taken to determine the acceptance. If $u < \alpha$, the candidate is accepted. If $u \geq \alpha$, no action is taken. Accepted samples are further stabilized using Exponential Moving Average (EMA) smoothing [19] to reduce sudden jumps between features.

This proposed **MCWaveGAN** model integrates the strengths of **WaveGAN** and **Markov Chain** process to generate realistic synthetic bee bioacoustic signals. It combines WaveGAN's capability to learn raw audio waveforms with the temporal modeling power of a Markov Chain to produce signals that not only replicate the natural acoustic patterns but also preserve event-specific characteristics and essential temporal features. It is expected that this approach enhances the realism of generated signals to support advanced research and applications in bee bioacoustics and ecological monitoring.

2.3. LDA-SVM Evaluation Framework

We consider a simple classification task to evaluate the usefulness of the synthetically generated data. The goal is to classify audio samples into three event categories. (e.g., *QueenPresent*, *QueenAbsent*, and *NoBee*). To achieve this, the audio signals are first converted into Mel-Frequency Cepstral Coefficients (MFCCs), where each audio segment is represented using 20 coefficients. To reduce dimensionality and highlight the discriminative structure of the data, we apply Linear Discriminant Analysis (LDA). Since the classification task involves three distinct classes ($C = 3$), LDA projects the MFCC features onto a 2-dimensional subspace (i.e., $C - 1$ components). This transformation simplifies the feature space and makes the classification task more tractable, enabling a clear visualization of

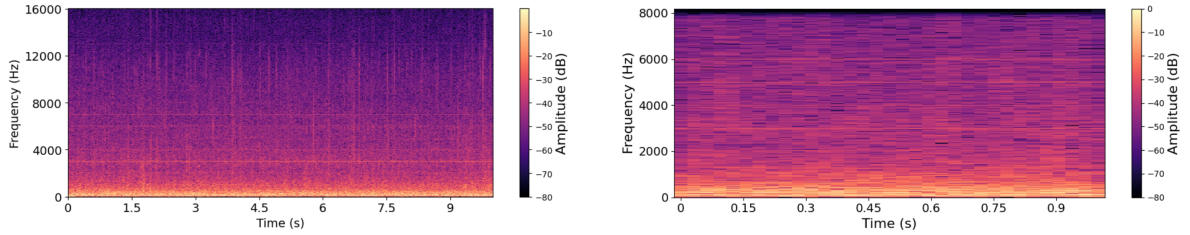
 doi:10.20944/preprints202512.0214.v1

class separation. For the classifier, we adopt a simple SVM model trained on the LDA-transformed features.

3. Experiments, Results and Discussion

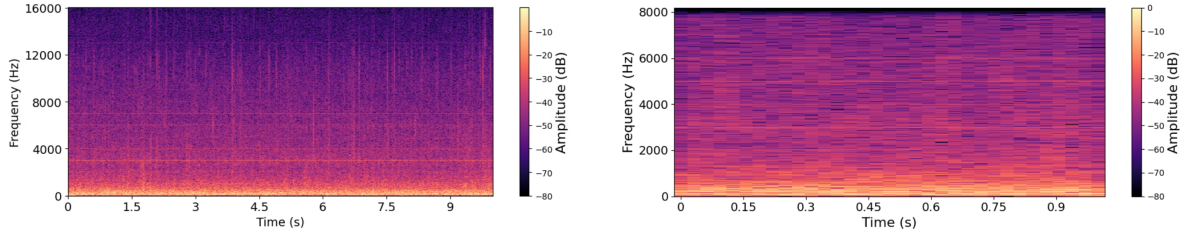
3.1. Dataset Description and Pre-Processing

In order to demonstrate the validity of the proposed methods, we apply them to a bee bioacoustics dataset [20,21] that includes three distinct events with their acoustic signatures: *QueenPresent*, *QueenAbsent*, and *NoBee*. Each sound recording in the dataset has a duration of 10 seconds, with most of the acoustic energy concentrated below 2 kHz. In total, the dataset provides 4,000 *QueenPresent* and 2,000 *QueenAbsent* recordings at an original sampling rate of 32 kHz. In addition to the bee categories, a *NoBee* dataset was included to strengthen validation. Several preprocessing steps were applied to standardize and enhance the dataset. First, the recordings were downsampled to 16,384 Hz and converted to mono. Silence intervals were removed to ensure that each clip starts and ends with an active bee sound. Band-pass filtering (20–2000 Hz) was applied to remove irrelevant frequency components. Finally, the recordings were segmented into 1-second clips, with padding or truncation applied to maintain fixed durations. After processing, the final dataset consisted of 39,838 *QueenPresent*, 19,897 *QueenAbsent*, and 9,964 *NoBee* samples. After preprocessing of raw bee audio signals, the spectrograms of *QueenPresent* and *QueenAbsent* samples showed significant improvements in spectral clarity and harmonic structures. Representative spectrograms of both bee-event categories are presented in Figures 2 and 3, which demonstrates the effectiveness of the processing pipeline.



(a) (b)

Figure 2. Spectrograms of *Queen Present* bee signals: (a) before processing and (b) after processing.



(a) (b)

Figure 3. Spectrograms of *Queen Absent* Bee signals: (a) before processing and (b) after processing.

Depending on the objective of the experiment, such as pretraining, synthetic sample generation, or classification performance evaluation, appropriate portions of the dataset were utilized. This approach ensured that each experiment setup was optimized for its intended usage while maintaining methodological consistency across the study.

3.2. MCWaveGAN Outperforms WaveGAN in Accuracy and Acoustic Realism

The first experimental claim evaluates the effectiveness of the proposed MCWaveGAN model compared to the WaveGAN alone in generating realistic bee audio signals. This analysis focused on both event-specific bee sound categories: *QueenPresent* bee and *QueenAbsent* bee.

 doi:10.20944/preprints202512.0214.v1

3.2.1. WaveGAN Generation

The *QueenPresent* subset of the processed dataset (15,942 samples) was first used to train the WaveGAN model for 120,000 iterations with a batch size of 64, in order to generate synthetic *QueenPresent* samples. The model producesC 1-second audio segments at a sampling rate of 16,384 Hz, generating 64 samples and saving both checkpoints and outputs every 1,000 iterations. After training, the generated samples from each checkpoint were evaluated using an LDA-SVM model (trained on real data) to determine the point of convergence. At iteration 48,000, the WaveGAN achieved its best performance, producing the highest proportion of correctly classified *QueenPresent* samples, with 45 out of 64 correctly identified. To further assess temporal dynamics, 20,000 synthetic samples were generated from this converged model and classified using the LDA-SVM classifier. With WaveGAN alone, only 59.8% of the generated samples were correctly classified as *QueenPresent*, while 40.1% were misclassified. The 59.8% accuracy doesn't necessarily mean the WaveGAN outputs are completely poor, as it could also reflect that the classifier (LDA-SVM trained only on real data) has not generalized well to the distribution of the synthetic data. However, as later demonstrated through LDA projections of the generated signals, these misclassifications are primarily due to poor data generation by WaveGAN.

The same procedure was then applied to the 15,941 *QueenAbsent* samples and 20,000 synthetic *QueenAbsent* samples were generated using the WaveGAN. Classification using the pre-trained LDA-SVM model revealed that out of a total of 20,000 samples, 85.8% were classified as *QueenAbsent*, while the remaining 14.2% were misclassified into the other two categories.

3.2.2. MC-Refined WaveGAN Generation (MCWaveGAN)

In our proposed MCWaveGAN model, the first phase consists of the WaveGAN module to generate synthetic bee signals, and the second phase involves the MC refinement module to address the limitations of the WaveGAN output. The WaveGAN-generated samples were fed into this MC module, together with the corresponding real samples, for MC refinement. This process was configured to produce 20,000 audio clips per class (*QueenPresent* and *QueenAbsent*), each 1 second in duration and sampled at 16,384 Hz. The MH algorithm was employed with $\beta = 0.01$ for *QueenPresent* and $\beta = 0.9$ for *QueenAbsent*, which controls the trade-off between retaining the characteristics of the real samples and maintaining smooth transitions in the synthetic data. The refined samples were subsequently validated using the LDA-SVM model (trained on real data) for classification accuracy. After refinement with the MC module, performance improved significantly.

Table 1 and Figure 4 illustrate the comparison of classification accuracy between synthetic bee signals generated after MC refinement and those produced by the WaveGAN model. Approximately 99.9% of the generated *QueenPresent* samples and 99.6% of the *QueenAbsent* samples were classified correctly, while less than 1% were incorrectly classified. The results clearly demonstrate that the combined MCWaveGAN model substantially outperforms WaveGAN alone in generating realistic bee sound signals. Further, this indicates that the MC refinement achieves a strong balance between maintaining the acoustic fidelity of real bee signals and ensuring smooth transitions in the synthetic data.

Table 1. LDA-SVM classification of WaveGAN-generated and MC-refined bee signals for optimal β values: $\beta = 0.01$ for *QueenPresent* and $\beta = 0.9$ for *QueenAbsent*.

	QueenPresent	**QueenAbsent**	**NoBee**
WaveGAN-generated *QueenPresent*	11,963 (59.8%)	3	8,034
MC Refined *QueenPresent* ($\beta = 0.01$)	**19,999(99.9%)**	**0**	**1**
WaveGAN-generated *QueenAbsent*	313	17,159 (85.8%)	2,528
MC Refined *QueenAbsent* ($\beta = 0.9$)	**0**	**19,911 (99.6%)**	**89**

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 December 2025 doi:10.20944/preprints202512.0214.v1

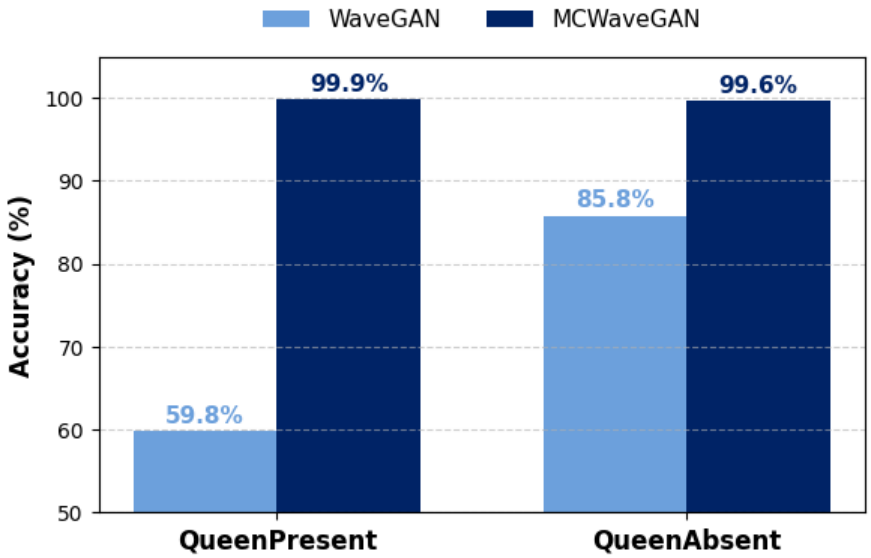


Figure 4. Comparison of classification accuracy for synthetic bee signals generated by proposed MCWaveGAN and WaveGAN alone for QueenPresent and QueenAbsent signals.

3.2.3. Statistical Analysis

Figures 5 and 6 compare frequency distribution, amplitude distribution, and spectral centroid for three randomly selected samples of the actual and generated data (for the *QueenPresent* and *QueenAbsent* classes). These characteristics are not features used in the AI models, but are included here solely for visualization and comparison. As shown, the MC-refined samples (green) align more closely with the distribution of real signals (blue) than those generated by WaveGAN alone (orange). This indicates that the MC refinement step reduces the divergence observed in WaveGAN outputs, resulting in synthetic signals whose observable acoustic behavior is more consistent with real bee sounds.

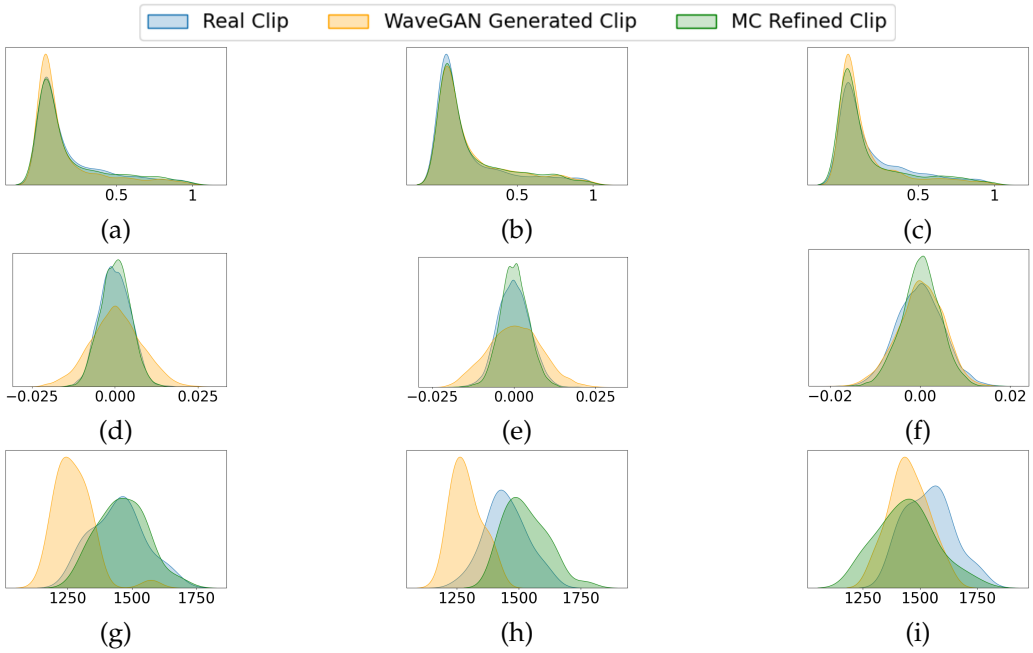


Figure 5. Comparative analysis of acoustic characteristics for actual, WaveGAN-generated, and MC-refined *Queen-Present* samples. Panels (a–c) show normalized frequency probability distributions, (d–f) amplitude probability distributions, and (g–i) spectral centroid probability distributions for three representative sample pairs.

 doi:10.20944/preprints202512.0214.v1

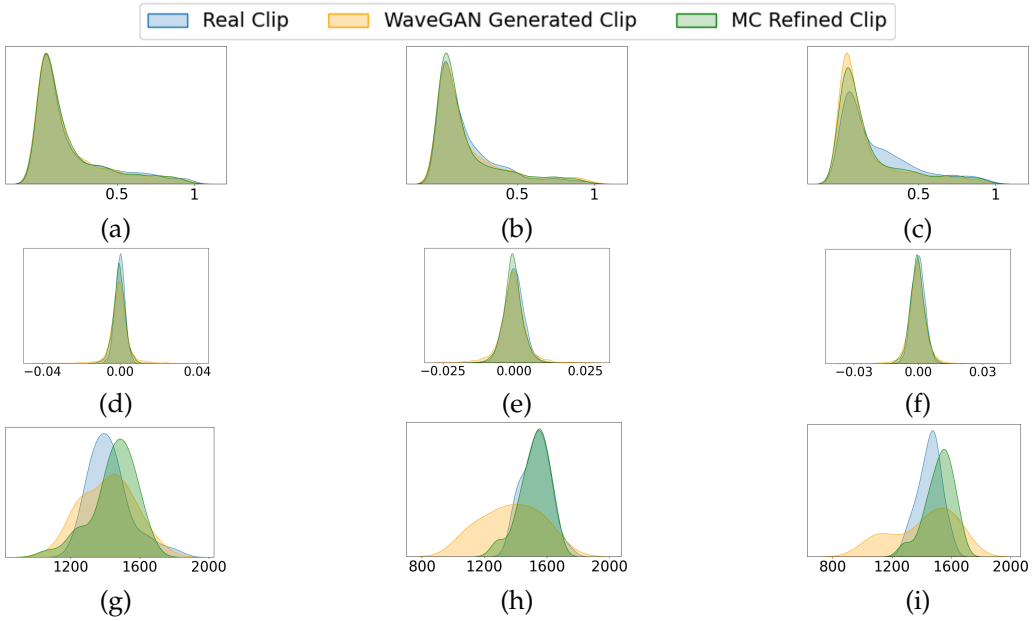
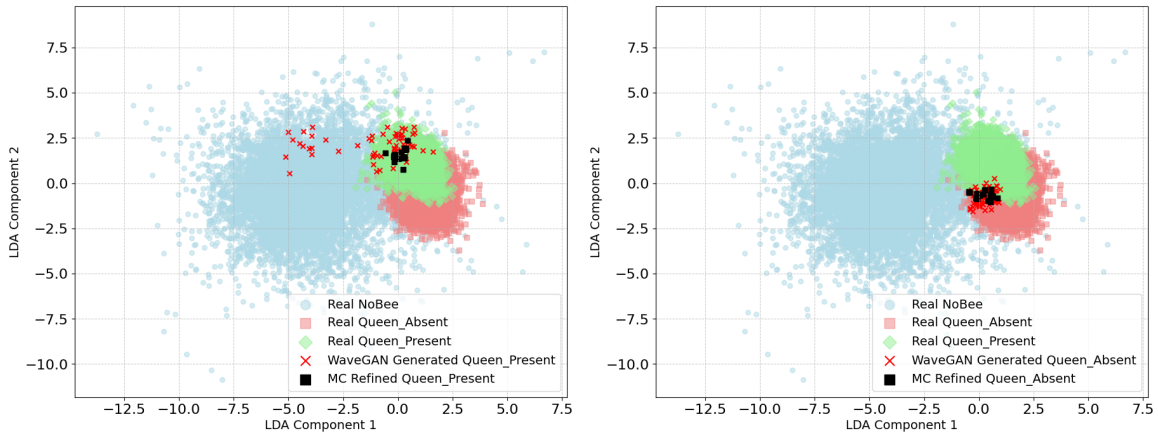


Figure 6. Comparative analysis of acoustic charactristics for Real, WaveGAN-generated, and MC-refined *Queen-Absent* samples. Panels (a–c) show normalized frequency probability distributions, (d–f) amplitude probability distributions, and (g–i) spectral centroid probability distributions for three representative sample pairs.

3.2.4. LDA Analysis

LDA projections were used to visually examine the alignment of the synthetic data distributions with the real data. The plot depicted in Figure 7 clearly shows that WaveGAN-generated samples are dispersed while MC-refined samples are tightly clustered around the real *QueenPresent* and *QueenAbsent* regions, respectively. This shift in data distribution illustrates that the MC refinement effectively reduces class overlap and pushes synthetic data close to real data. Hence, it improves realism and enforces structural similarity within the feature space.



(**a**) *Queen Present* (**b**) *Queen Absent*

Figure 7. LDA projections of actual bee bioacoustic signals compared with WaveGAN-generated and MC-refined synthetic samples. The plots illustrate how MC refinement shifts synthetic data closer to the real class distribution for both (**a**) *Queen Present* and (**b**) *Queen Absent* categories.

3.2.5. Real-Data Evaluation with Augmented Training

We evaluated the impact of synthetic bee signal augmentation by comparing classification performance for both *QueenPresent* and *QueenAbsent* samples generated by WaveGAN alone vs those refined through the MC process (MCWaveGAN). We define two models using previously generated bee audio samples: **Model 1** is trained on real data plus WaveGAN-generated bee signals, and **Model**

 doi:10.20944/preprints202512.0214.v1

2 is trained on real data plus MCWaveGAN-generated bee signals (i.e., WaveGAN-generated, and then refined by the MC process). This experiment was done for both *QueenPresent* and *QueenAbsent* datasets separately.

For *QueenPresent* samples, the classification results shown in Table 2 indicate that MCWaveGAN augmentation outperforms WaveGAN-generated augmentation. While recall remained stable at 98% for both models, the precision increased from 89% to 95%. That indicates a significant reduction in false positives. The F1-score also improved from 93% to 96%, and that reflects the realism improvement of the synthetic *QueenPresent* audio. Furthermore, the overall accuracy increased from 94% in **Model 1** to 96% in **Model 2**, demonstrating the effectiveness of the MC refinement stage in improving classification performance.

Table 2. Comparison of classification performance for *QueenPresent* and *QueenAbsent* datasets across Model 1 and Model 2.

Category	Metric	Model 1	Model 2
QueenPresent	Precision	89%	**95%**
	Recall	98%	**98%**
	F1-score	93%	**96%**
	Accuracy	94%	**96%**
QueenAbsent	Precision	94%	**97%**
	Recall	96%	**96%**
	F1-score	96%	**97%**
	Accuracy	96%	**97%**

A similar experiment was conducted for *QueenAbsent* using WaveGAN and MCWaveGAN generated samples. There, **Model 1** was trained on real data combined with WaveGAN-generated *QueenAbsent* samples, while **Model 2** was trained on real data augmented with MCWaveGAN-generated *QueenAbsent* samples. The Classification results comparison shown in Table 2 demonstrates that the **Model 2** achieves consistent improvement for the *QueenAbsent* as well. There, Model 1 achieved a precision of 94%, recall of 96%, and F1-score of 96%, with overall accuracy of 96%. In contrast, Model 2 improved precision to 97%, accuracy to 97%, F1-score to 97%, while maintaining recall at 96%. These results indicate that the MCWaveGAN-generated synthetic bee bioacoustic signals enhance the classifier's ability to more accurately distinguish *QueenPresent* and *QueenAbsent* acoustic patterns with higher accuracy.

The comparative evaluation between baseline WaveGAN (Model 1) and the proposed MCWaveGAN (Model 2) clearly demonstrates that integrating the Markov Chain refinement significantly enhances synthetic bee sound generation and classification performance. Both *QueenPresent* and *QueenAbsent* categories in Model 2 outperformed Model 1 across all evaluated metrics. These consistent performance gains highlight the advantage of incorporating temporal dependencies through Markov Chain transitions. The enhanced temporal coherence and smoother spectral continuity produced by MCWaveGAN significantly contributed to generating more realistic bee bioacoustic signals.

3.3. MCWaveGAN-Augmented Training Improves Classification Performance

We investigated the effectiveness of synthetic bee bioacoustic signals generated using the proposed MCWaveGAN model in enhancing the learning capability of machine learning models trained on limited and imbalanced real-world data. The occurrence of bee sounds in natural bee-hive environments varies based on the presence or absence of the queen, and that creates naturally unbalanced acoustic conditions. Many recent studies have highlighted the potential of acoustic monitoring for assessing the colony state, particularly for detecting queen presence/absence and swarming preparation, both of which are critically important for colony health and productivity [3,22].

To replicate natural bee hive conditions, an unbalanced dataset was intentionally constructed for both event types. The experiment evaluated whether the proposed MCWaveGAN generated synthetic

 doi:10.20944/preprints202512.0214.v1

bee bioacoustic signals of these two event types could augment the real training dataset to improve classification performance. The goal of the experiment was to enhance the model's sensitivity to rare bee events while maintaining robust performance for automated hive monitoring applications. It is expected to use the proposed MCWaveGAN model to generate the bee bioacoustic signals and overcome the data scarcity challenge for rare hive events while improving the reliability of acoustic classification models. It supports beekeepers in the timely detection of colony conditions. *QueenPresent* related experiment started with only 20% of *QueenPresent* samples and gradually increased its proportion with MCWaveGAN generated samples until matched with *QueenAbsent* samples, eventually the model balance is reached. Conversely, the second experiment began with 20% *QueenAbsent* samples and progressively increased its proportion with synthetically generated samples from MCWaveGAN to equal the *QueenPresent* samples. That helped us to identify the model's adaptability to changing class distributions and how synthetic bee bioacoustic signals impact the model performance.

3.3.1. QueenPresent Augmentation

For this experiment, we define **Model 1** as the baseline model, and it is trained using only real bee audio data, where *QueenPresent* is the minority class. When the **Model 1** was trained and classified with the LDA-SVM model, it achieved an overall accuracy of 94.22% along with a precision of 85%, a recall of 89%, and an F1-score of 87% as shown in Table 3. Synthetically generated *QueenPresent* samples from MCWaveGAN (WaveGAN-generated and MC refined) were used to augment **Model 1**'s training set to evaluate its impact on performance with real test data. The classifier's performance showed gradual improvement as the dataset approached balance through **Model 2 to 4** as depicted in Table 3. The augmented models showed increased accuracy and stability for the minority class, *QueenPresent*. And that reflects that MCWaveGAN-generated samples were both realistic and beneficial for model learning.

Table 3. Classification performance on real test data for models trained with varying levels of synthetic *QueenPresent* samples generated by the MCWaveGAN model.

Model	Queen Present	Queen Absent	NoBee	Precision	Recall	F1-score	Accuracy
Model 1	**4K[20%]**	8K [40%]	8K [40%]	85%	89%	87%	94.22%
Model 2	**5.3K[26%]**	8K [40%]	8K [40%]	86%	88%	87%	95%
Model 3	**6.8K[30%]**	8K [36%]	8K [34%]	88%	85%	87%	94%
Model 4	**8K[33.3%]**	8K [33.3%]	8K [33.3%]	89%	84%	87%	95%

3.3.2. QueenAbsent Augmentation

To evaluate the scarcity of *QueenAbsent* data in natural hive environments, the dataset was prepared to reflect the real-world class imbalance. The dataset was divided into *QueenPresent*, *QueenAbsent*, and *NoBee* categories with a 40:20:40 ratio, respectively, in the training set. In this setup, we define the baseline classifier model (**Model 1**) trained using the real bee data only while keeping the *QueenAbsent* as the minority class. Using the LDA-SVM model, the baseline model achieved an overall accuracy of 94.82% with class-specific precision, recall, and F1-score as detailed in the Table 4.

Synthetically generated *QueenAbsent* samples from the MCWaveGAN model were then gradually added to the training set of the **Model 1** to evaluate their effect on classification performance. Table 4 presents the classification results. The findings show that progressively augmenting the training set with MCWaveGAN-generated *QueenAbsent* samples led to improved classifier performance, indicating that the synthetic samples were both realistic and beneficial for enhancing model learning.

As depicted in the Figure 8, the results of this experiment demonstrated that augmenting a real bee audio dataset with synthetic bee signals generated by the proposed MCWaveGAN model can significantly improve machine learning model performance, specifically under conditions of natural data imbalance. It showed noticeable improvements in both overall accuracy and class-level performance. It demonstrated that the synthetic data generated by the MCWaveGAN model can effectively enhance model training when real-world data are limited. This finding provides substantial

 doi:10.20944/preprints202512.0214.v1

importance for beekeeping, ecological monitoring, and bee-hive management systems. In real-world conditions, collecting *QueenAbsent* recording is both rare and logistically challenging. Because the occurrence of the queenless state is unpredictable. The ability to synthesize acoustically realistic *QueenAbsent* signals helps researchers to train more robust classification models without requiring extensive real-data collection.

Table 4. Classification performance on real test data for models trained with varying levels of synthetic *QueenAbsent* samples generated by the MCWaveGAN model.

Model	Queen Present	Queen Absent	NoBee	Precision	Recall	F1-Score	Accuracy
Model 1	8K[40%]	**4K[20%]**	8K[40%]	87%	95%	91%	94.6%
Model 2	8K[37.5%]	**5.3K[25%]**	8K[37.5%]	88%	93%	91%	95.0%
Model 3	8K[35%]	**6.8K[30%]**	8K[35%]	89%	92%	91%	95.0%
Model 4	8K[33.3%]	**8K[33.3%]**	8K[33.3%]	90%	91%	91%	95.0%

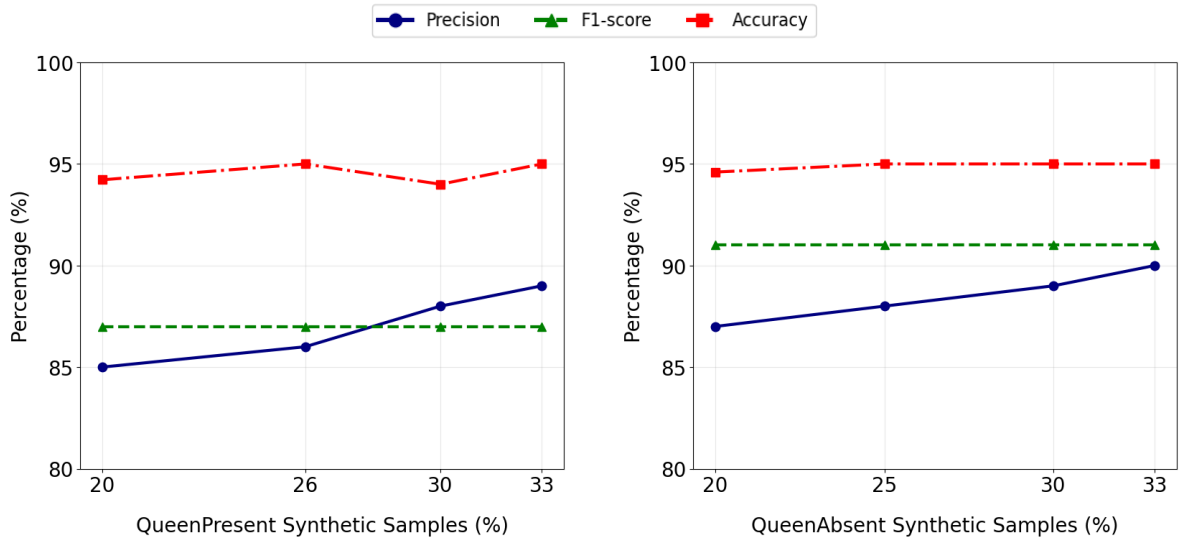


Figure 8. Impact of increasing synthetic samples generated by MCWaveGAN on classifier performance metrics (Precision, F1-score, Accuracy) for (a)*QueenPresent* and (b)*QueenAbsent* classes, evaluated on real test data.

4. Discussion

In summary, the results of the study demonstrated that the proposed MCWaveGAN model consistently outperforms the standard WaveGAN in generating realistic bee bioacoustic signals. Statistical analysis illustrated that MC-refined signals are more closely replicate the frequency, amplitude, and spectral centroid distributions of real bee signals compared to WaveGAN outputs. LDA projections further demonstrated that MCWaveGAN-generated samples are tightly clustered around real data, indicating more structural similarity, whereas WaveGAN-generated data are more dispersed. Additionally, real-data performance evaluation with augmented training confirmed that MCWaveGAN augmentation improves classification metrics for *QueenPresent* and *QueenAbsent* synthetic samples with higher precision, recall, F1-score, and overall accuracy than WaveGAN. Overall, these results confirm that the MCWaveGAN model is more effective than the standard GAN (WaveGAN) for realistic synthetic bee signal generation.

Further, augmenting the real bee audio dataset with synthetic bee signals generated by the proposed MCWaveGAN model significantly improves the machine learning model performance, specifically under conditions of natural data imbalance. It showed noticeable improvements in both overall accuracy and class-level performance. It demonstrated that the synthetic data generated by the MCWaveGAN model can effectively enhance model training when real-world data are limited. This finding provides substantial importance for beekeeping, ecological monitoring, and bee-hive

 doi:10.20944/preprints202512.0214.v1

management systems. In real-world conditions, collecting *QueenAbsent* recording is both rare and logistically challenging. Because the occurrence of the queenless state is unpredictable. The ability to synthesize acoustically realistic *QueenAbsent* signals helps researchers to train more robust classification models without requiring extensive real-data collection. For real-world hive monitoring applications, improved classification accuracy means that early detection of queenless or hive stress can be achieved with great reliability. This has direct implications for colony health management as it helps beekeepers to take preventive measures before colony collapse occurs. Further, generating balanced and realistic bee bioacoustic signals through the MCWaveGAN model reduces the dependency on manual data collection. This study did not include real-world validation within beehive environments, but the findings suggest that synthetic bee bioacoustic signals generated using the proposed MCWaveGAN model hold strong potential for validation by future researchers. Subsequent research studies could leverage this model to generate synthetic bee bioacoustic signals for practical validations in real beehives. Such that it will give a beneficial contribution to sustainable bee-hive monitoring and preserving bee populations, which is an essential aspect of global agricultural productivity and biodiversity.

5. Conclusions

We have explored and presented a novel approach to generate synthetic bee bioacoustic signals. It is a hybrid model that combines a WaveGAN and a Markov Chain process. The primary aim of this study was to address the scarcity of high-quality and balanced bee bioacoustic datasets that are required for the AI-driven bee hive monitoring applications and sustainable beekeeping activities. Nowadays, collecting real-world bee bioacoustic data is difficult and expensive. Even when such data is obtained, it is often contaminated by noise from surrounding environmental sounds. That creates a critical challenge in developing robust machine learning models. The integration of probabilistic modeling capabilities of Markov chains with the generative power of GANs has helped us to refine GAN-generated data to produce synthetic bee bioacoustic signals that closely resemble the original bee signals. Hence, the proposed MCWaveGAN model improved the realism, temporal coherence, and event-specific characteristics of the synthetically generated data. This study explored several research questions related to the generation of synthetic bee bioacoustic data, as outlined in Chapter 1. First, how can limited and noisy bee bioacoustic data be effectively augmented using generative models? Secondly, we explored how a Markov Chain process can enhance the quality of GAN-generated bee audio signals. The optimal parameters and configurations required to generate the most realistic bee bioacoustic signals are also one of the questions outlined. Lastly, how realistic are the MCWaveGAN-generated bee signals compared to real bee signals when evaluated quantitatively and qualitatively?

This research systematically addressed the challenge of generating realistic synthetic bee bioacoustic signals by first conducting a comprehensive review of existing synthetic data generation techniques, identifying key limitations in their ability to preserve the temporal and event-specific characteristics essential for bee communication. The study introduced a novel hybrid architecture, the Markov Chain Wave Generative Adversarial Network (MCWaveGAN), which integrates a WaveGAN-based generative module with a Markov Chain process implemented through the Metropolis–Hastings algorithm to enhance the acoustic realism of generated signals. Experimental evaluations demonstrated that MCWaveGAN consistently outperformed the baseline WaveGAN in both qualitative and quantitative aspects, producing synthetic signals that more closely matched real bee acoustics in terms of frequency, amplitude, and spectral centroid distributions. LDA projections further confirmed that MCWaveGAN-generated samples aligned more closely with real data, indicating stronger structural similarity. Moreover, augmenting real training datasets with MCWaveGAN-generated samples improved classification metrics—including precision, recall, F1-score, and overall accuracy—across both *QueenPresent* and *QueenAbsent* categories, effectively addressing data imbalance caused by rare hive conditions. These findings validate the MCWaveGAN model's capability to generate high-fidelity

 doi:10.20944/preprints202512.0214.v1

synthetic bee bioacoustic data that can serve as a reliable substitute for real recordings. Overall, the study contributes a novel and practical approach to computational bioacoustics, demonstrating how integrating probabilistic temporal modeling with adversarial learning can enhance data realism while supporting sustainable, AI-driven beekeeping and hive monitoring applications by reducing the need for extensive real-world data collection.

While this study and its results have made a successful contribution to generating synthetic bee bioacoustic signals, there are several future directions open for future explorations to advance further and extend this work. One such is validating the synthetic bee signal in real beehive environments. This research did not include the testing of synthetically generated *QueenPresent* or *QueenAbsent* sounds in a real beehive to explore their practical usage, as it was beyond the scope of this research. For instance, when a queen leaves a hive, beekeepers often face the risk of colony collapse. In such cases, playing synthetic *QueenPresent* sounds inside the hive could help maintain colony stability until a new queen is introduced and protect the hive. Similarly, synthetic *QueenAbsent* sounds can be used with experimental studies to observe how bees respond when the queen is absent. That will help the researchers to better understand beehive dynamics and communication patterns that are important for efficient beekeeping.

Generating sound signals of other bioacoustics domains is also a potential future direction. My study focused on bee bioacoustics as it is one of the most important domains due to its relevance to apiculture and environmental sustainability; the same generative approach can be extended to other bioacoustics domains. Generating realistic synthetic bioacoustic signals for species such as bats, birds, and marine animals would be highly valuable for advancing ecological research and conservation efforts.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: he source code and generated data samples used in this paper are publicly available at: https://github.com/KumuduS/Bioacoustic_MCWaveGAN

Acknowledgments: During the preparation of this manuscript/study, the author(s) used artificial intelligence assisted technologies for the purposes of refining the language and enhancing the readability of our text. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ML	Machine Learning
GAN	Generative Adversarial Network
MC	Markov Chains
MCGAN	Markov Chain Generative Adversarial Networks
MCWaveGAN	Markov Chain Wave Generative Adversarial Networks
SVM	Support Vector Machines
WGAN-GP	Wasserstein Generative Adversarial Network with Gradient Penalty
MH	Metropolis–Hastings
MFCCs	Mel-Frequency Cepstral Coefficients
LDA	Linear Discriminant Analysis
FIR	Finite Impulse Response

 doi:10.20944/preprints202512.0214.v1

References

1. Janovec, M.; Pap'an, J.; Tatarka, S. Approaches and Strategies for Identifying Bioacoustic Signals from Wildlife in IoT Forest Environments. In Proceedings of the 2024 International Conference on Emerging eLearning Technologies and Applications (ICETA). IEEE, 2024, pp. 226–231.
2. Penar, W.; Magiera, A.; Klocek, C. Applications of bioacoustics in animal ecology. *Ecological complexity* **2020**, *43*, 100847.
3. Uthoff, C.; Homsi, M.N.; Von Bergen, M. Acoustic and vibration monitoring of honeybee colonies for beekeeping-relevant aspects of presence of queen bee and swarming. *Computers and electronics in agriculture* **2023**, *205*, 107589.
4. Nieto-Mora, D.; Rodriguez-Buritica, S.; Rodriguez-Marin, P.; Martínez-Vargaz, J.; Isaza-Narvaez, C. Systematic review of machine learning methods applied to ecoacoustics and soundscape monitoring. *Heliyon* **2023**, *9*.
5. McEwen, B.; Soltero, K.; Gutschmidt, S.; Bainbridge-Smith, A.; Atlas, J.; Green, R. Automatic noise reduction of extremely sparse vocalisations for bioacoustic monitoring. *Ecological Informatics* **2023**, *77*, 102280.
6. Lu, Y.; Shen, M.; Wang, H.; Wang, X.; van Rechem, C.; Fu, T.; Wei, W. Machine Learning for Synthetic Data Generation: A Review. *arXiv preprint arXiv:2302.04062* **2023**.
7. Awiszus, M.; Rosenhahn, B. Markov Chain Neural Networks. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018, pp. 2180–2187.
8. Lin, C.; Killeen, P.; Li, F.; Yeap, T.; Kiringa, I. Preserving Temporal Dynamics in Synthetic Multivariate Time Series Using Generative Neural Networks and Monte Carlo Markov Chain.
9. M"ucke, N.T.; Sanderse, B.; Boht'e, S.M.; Oosterlee, C.W. Markov chain generative adversarial neural networks for solving Bayesian inverse problems in physics applications. *Computers & Mathematics with Applications* **2023**, *147*, 278–299.
10. Yuan, Y.; Guo, Y. A Review on Generative Adversarial Networks. In Proceedings of the 2020 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT). IEEE, 2020, pp. 392–401.
11. Nord, S. Multivariate Time Series Data Generation using Generative Adversarial Networks: Generating Realistic Sensor Time Series Data of Vehicles with an Abnormal Behaviour using TimeGAN, 2021.
12. Donahue, C.; McAuley, J.; Puckette, M. ADVERSARIAL AUDIO SYNTHESIS. *arXiv preprint arXiv:1802.04208* **2018**.
13. Radford, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv preprint arXiv:1511.06434* **2015**.
14. Hsu, P.c.; Wang, C.h.; Liu, A.T.; Lee, H.y. Towards Robust Neural Vocoding for Speech Generation: A Survey. *arXiv preprint arXiv:1912.02461* **2019**.
15. Lavault, A. Generative Adversarial Networks for Synthesis and Control of Drum Sounds. PhD thesis, Sorbonne Université, 2023.
16. Zhang, L.; Huang, H.N.; Yin, L.; Li, B.Q.; Wu, D.; Liu, H.R.; Li, X.F.; Xie, Y.L. Dolphin vocal sound generation via deep WaveGAN. *Journal of Electronic Science and Technology* **2022**, *20*, 100171.
17. Gilks, W.R.; Best, N.G.; Tan, K.K. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Journal of the Royal Statistical Society Series C: Applied Statistics* **1995**, *44*, 455–472.
18. Cao, H.; Tan, V.Y.; Pang, J.Z. A Parsimonious Mixture of Gaussian Trees Model for Oversampling in Imbalanced and Multimodal Time-Series Classification. *IEEE Transactions on Neural Networks and Learning Systems* **2014**, *25*, 2226–2239.
19. Bergmeir, C.; Hyndman, R.J.; Ben'ıtez, J.M. Bagging exponential AUDIO-BASED IDENTIFICATION methods using STL decomposition and Box–Cox transformation. *International Journal of Forecasting* **2016**, *32*, 303–312.
20. Cecchi, S.; Terenzi, A.; Orcioni, S.; Riolo, P.; Ruschioni, S.; Isidoro, N.; et al. A preliminary study of sounds emitted by honey bees in a beehive. In Proceedings of the 144th Audio Engineering Society Convention. AES, 2018.
21. Nolasco, I.; Terenzi, A.; Cecchi, S.; Orcioni, S.; Bear, H.L.; Benetos, E. AUDIO-BASED IDENTIFICATION OF BEEHIVE STATES. In Proceedings of the ICASSP 2019 - IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2019, pp. 8256–8260.
22. Ruvinga, S.; Hunter, G.; Duran, O.; Nebel, J.C. Identifying Queenlessness in Honeybee Hives from Audio Signals Using Machine Learning. *Electronics* **2023**, *12*, 1627.