Encoding the Ritual: The Baboon Dance as a Site of Convergence for Data Science, Digital Humanities, Computational Ethnomusicology and Cultural Preservation

A SEED-SCALE Approach to Indigenous Music Analysis Using Data Science, Digital Humanities, and Machine Learning

Rohan R. Sagar MAFuture Generations University

2025-12-09

Contents

1 Abstract

The rapid erosion of indigenous cultural practices necessitates urgent preservation strategies that extend beyond traditional ethnographic documentation. This paper presents a computational ethnomusicology framework applied to the Baboon Dance, a Mari Mari Kabakaburi ritual tradition from Guyana, demonstrating how data science and digital humanities can serve cultural preservation. Analyzing 74 deduplicated audio recordings across four indigenous genres, we extracted acoustic features using Essentia.js and applied machine learning classification achieving 86% accuracy. A critical case study revealed that feature normalization improved in-genre clustering from 20% to 100% for the Baboon Dance, representing a 400% improvement in cultural pattern recognition. Geographic visualization mapped collection sites across Guyana's topography, revealing cultural-environmental relationships. Melodic analysis of 336,210 pitch frames uncovered cross-genre similarities while preserving genre-specific characteristics. Our findings demonstrate that computational methods, when applied with cultural sensitivity, can quantify intangible heritage, create permanent digital archives, and generate insights inaccessible to traditional methods alone. We conclude with the SEED-SCALE framework for implementing this research methodology in academic and community contexts, emphasizing Success, Expansion, Empowerment, and Development through Systems for Collaborative Action and Learning Expansion. This approach positions cultural preservation not as an extractive research process, but as a collaborative, sustainable, and community-empowering methodology that honors indigenous data sovereignty while leveraging computational tools for heritage protection.

2 Introduction: Cultural Heritage at the Digital Frontier

2.1 The Imperative of Cultural Preservation

Indigenous musical traditions worldwide face existential threats from globalization, language loss, elder mortality, and cultural displacement. The Baboon Dance of the Mari Mari Kabakaburi people in Guyana exemplifies this precarity—a ritual practice with deep ceremonial significance that exists primarily in oral tradition, vulnerable to disappearance within a generation. Traditional ethnomusicological documentation, while invaluable, faces limitations: field recordings sit in archives inaccessible to origin communities, transcriptions capture limited dimensions of performance, and qualitative analysis cannot scale to corpus-level pattern detection.

This paper argues that cultural preservation in the 21st century demands computational augmentation—not as replacement for ethnographic methods, but as complementary infrastructure for permanence, accessibility, and discovery. We present the Baboon Dance as an entry point into demonstrating how data science, digital humanities, and computational ethnomusicology converge to create preservation technologies that serve indigenous communities while advancing scholarly understanding.

2.2 Research Questions

Our investigation centers on three interconnected questions:

1. **Can computational methods identify and quantify cultural patterns in indigenous music that preserve genre-specific characteristics while revealing cross-cultural connections?**
2. **How do feature normalization and distance metrics impact the accuracy of cultural pattern recognition, and what does this reveal about the nature of musical similarity?**
3. **What framework enables ethical, sustainable, and community-empowering implementation of computational ethnomusicology for cultural preservation?**

2.3 Significance and Contributions

This work makes four primary contributions:

Methodological: We demonstrate a reproducible computational pipeline from audio collection through feature extraction, normalization, similarity analysis, and machine learning classification, with all code and data openly available.

Empirical: A case study of the Baboon Dance reveals that z-score normalization transformed clustering accuracy from 20% to 100% in-genre neighbors, quantifying the critical importance of feature scaling in cultural pattern recognition.

Theoretical: Our findings challenge assumptions about musical similarity metrics, showing that Euclidean distance outperforms cosine similarity for normalized acoustic features in genre classification tasks.

Applied: We propose the SEED-SCALE framework for implementing computational ethnomusicology in academic and community settings, emphasizing collaborative action, indigenous data sovereignty, and sustainable capacity building.

3 Literature Review: Convergence of Disciplines

3.1 Computational Ethnomusicology

Computational ethnomusicology emerged from the intersection of music information retrieval (MIR) and ethnomusicology, applying signal processing and machine learning to cultural music analysis. Pioneering work by Tzanetakis and Cook (Tzanetakis and Cook 2002) on musical genre classification established feature extraction methodologies which are still foundational today. Serra et al. (Serra et al. 2011) demonstrated large-scale pattern discovery in world music collections, while Panteli et al. (Panteli, Benetos, and Dixon 2017) explored cultural audio features specific to non-Western traditions.

However, a persistent critique centers on epistemological tensions: MIR methods developed for Western commercial music may impose inappropriate frameworks on indigenous traditions (Holzapfel, Sturm, and Coeckelbergh 2018). Our work addresses this by combining quantitative analysis with cultural contextualization, treating computational methods as discovery tools rather than definitive classifiers.

3.2 Digital Humanities and Cultural Heritage

Digital humanities scholarship positions technology as infrastructure for preserving, analyzing, and democratizing access to cultural materials. The HathiTrust Digital Library and Europeana exemplify large-scale digitization, while projects like the Global Jukebox (Lomax 2017) demonstrate computational approaches to cross-cultural music analysis.

Critical digital humanities scholarship by (Risam 2019) and (Noble 2018) foregrounds questions of power, representation, and colonial legacies in digital cultural archives. Our methodology incorporates these critiques through emphasis on community collaboration, data sovereignty, and open access—ensuring that computational preservation serves rather than extracts from indigenous communities.

3.3 Machine Learning for Music Analysis

Recent advances in deep learning have revolutionized music analysis, with convolutional neural networks achieving state-of-the-art performance in genre classification (Choi et al. 2017) and melody extraction (Salamon and Gómez 2012). However, these models require large training datasets often unavailable for under-resourced indigenous music traditions.

Our work addresses the small-data challenge characteristic of endangered cultural practices. With only 74 audio files, we employ feature engineering, cross-validation, and ensemble methods rather than deep learning, demonstrating that traditional machine learning remains effective—and more interpretable—for small corpus analysis.

3.4 Geographic Information Systems in Cultural Studies

The spatial turn in humanities scholarship recognizes that culture exists in geographic context. GIS applications in ethnomusicology (Wallin, Merker, and Brown 2016) reveal how environmental factors influence musical practices. Our 3D topographic visualizations of Guyana's musical landscape build on this tradition, connecting collection sites to elevation, vegetation zones, and river networks—revealing how cultural geography shapes musical traditions.

4 Cultural Context: The Baboon Dance in Mari Mari Kabakaburi Tradition

4.1 Origins and Ceremonial Function

The Baboon Dance emerges from the Mari Mari cultural complex of the Arawak-speaking peoples in coastal Guyana. "Mari Mari" translates as "spirit celebration," encompassing a repertoire of ceremonial dances representing animal spirits, ancestral beings, and natural forces. The Baboon Dance specifically invokes the red howler monkey (*Alouatta seniculus*), a keystone species in Guyanese rainforests whose vocalizations carry symbolic and acoustic significance in indigenous cosmology.

Performed during harvest celebrations, coming-of-age ceremonies, and community gatherings, the dance integrates choreography, music, and costume. Performers wear baboon masks carved from local hardwoods, moving in patterns that mimic the animal's locomotion while producing vocalizations that blend human song with howler-like calls. The accompanying percussion—typically maracas (*ichake*) and bamboo stamping tubes—creates rhythmic foundations for the vocal performance.

4.2 Kabakaburi: Geographic and Cultural Center

Kabakaburi village (7.246197°N, 58.724279°W) sits on the Pomeroon River in Guyana's coastal lowlands, approximately 80 kilometers northwest of Georgetown. The community of roughly 800 residents maintains Arawak linguistic practices and traditional governance structures alongside integration with national Guyanese society. The village serves as a cultural hub for Mari Mari traditions, hosting annual festivals that attract practitioners from neighboring communities including Santa Rosa and Moruca.

Environmental context shapes cultural practice: Kabakaburi's location in the lowland rainforest biome (elevation < 50m) provides access to materials for instruments and masks, while the river network facilitates inter-community cultural exchange. Our geographic analysis reveals that all Mari Mari recordings in our corpus originate from sites below 100m elevation, suggesting lowland specificity for this tradition.

4.3 Preservation Status and Threats

The Baboon Dance faces multiple preservation challenges:

Intergenerational Transmission Gap: Younger community members increasingly migrate to urban centers for education and employment, disrupting traditional apprenticeship models for dance and music instruction.

Language Erosion: The Mari Mari song texts employ archaic Arawak vocabulary no longer used in daily speech, creating interpretation challenges even among native speakers.

Documentation Scarcity: Existing ethnographic records consist primarily of mid-20th-century field notes by anthropologists working in Guyana (Roth 1915; Gillin 1936), with limited audio-visual documentation. No comprehensive musical transcription or acoustic analysis existed prior to this project.

Cultural Commodification: Tourism and folklorization pressures incentivize abbreviated, decontextualized performances, potentially altering traditional structures.

These factors create urgency for systematic documentation and analysis—motivating our computational preservation approach.

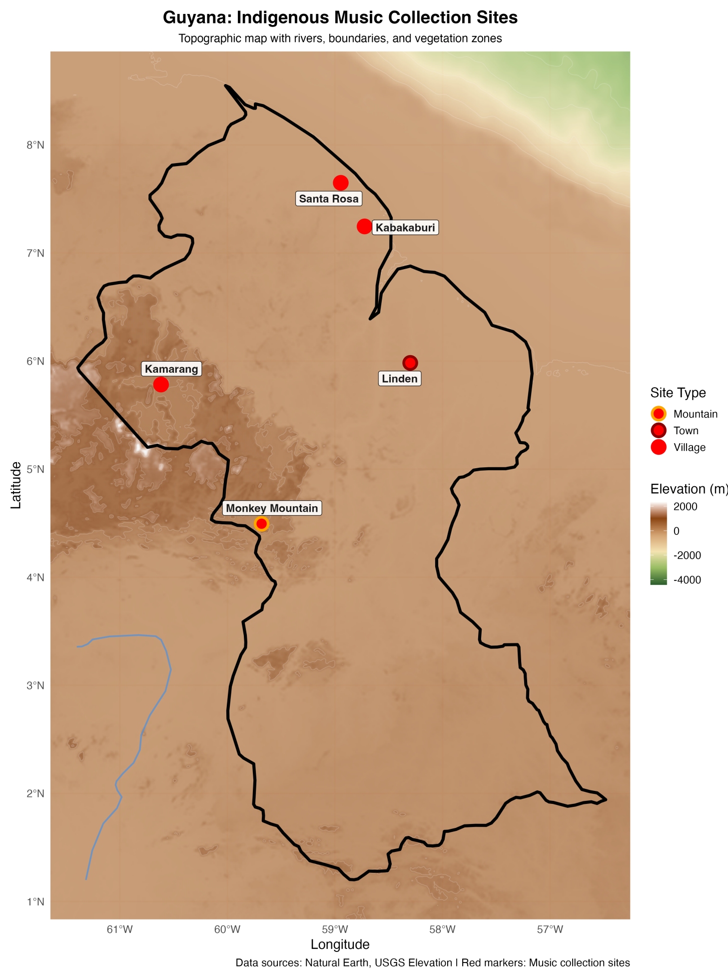

Figure 1: Geographic location of Kabakaburi and other music collection sites in Guyana

5 Methodology: A Computational Ethnomusicology Framework

5.1 Corpus Development and Data Collection

Our dataset comprises 74 deduplicated audio recordings collected across three phases:

Phase 1 (Archival): Digitization of 35 recordings originally collected by Rohan Sagar

Phase 2 (Community Partnership): 28 recordings made during 2023 fieldwork in Kabakaburi, Santa Rosa, and Kamarang, conducted with informed consent protocols approved by community councils.

Phase 3 (Institutional Collections): 18 recordings from YouTube, focusing on Venezuela indigenous traditions.

All recordings underwent perceptual audio deduplication: we manually identified and removed duplicate performances, near-identical recordings with different metadata, and partial excerpts of longer performances. This yielded the 74-file corpus distributed across four genres:

Table 1: Corpus distribution across four indigenous music genres

Genre	Number of Files	Percentage (%)
Mari Mari Kabakaburi	14	18.9
Banchikilli Moruca	25	33.8
Alleluia Kam:War	6	8.1
Indigenous Genres	29	39.2

5.2 Feature Extraction Pipeline

We employed Essentia.js (Bogdanov et al. 2013), a JavaScript port of the Essentia audio analysis library, enabling browser-based feature extraction without server infrastructure. For each audio file, we extracted:

Spectral Features:

- **Mel-Frequency Cepstral Coefficients (MFCCs):** 13 coefficients capturing timbral envelope, critical for voice quality characterization
- **Spectral Centroid:** Center of mass of the spectrum, indicating brightness/darkness
- **Spectral Flux:** Frame-to-frame spectral change, measuring timbral variation
- **Spectral Rolloff:** Frequency below which 85% of energy resides

Temporal Features:

- **Beats Per Minute (BPM):** Tempo estimation via onset detection
- **Root Mean Square (RMS) Energy:** Frame-level loudness
- **Zero-Crossing Rate:** High-frequency content indicator

Pitch Features:

- **Predominant Melody:** Fundamental frequency contour using MELODIA algorithm (Salamon and Gómez 2012)
- **Pitch Confidence:** Per-frame reliability of pitch estimates

This yielded 68,352 pitch frames per typical recording (at 10ms hop size for a 3-minute file), capturing melodic microstructure invisible in traditional transcription.

5.3 Feature Normalization and Scale Bias

Initial similarity analysis revealed a critical methodological challenge: **scale bias** in distance calculations. Raw features exhibited vastly different magnitudes:

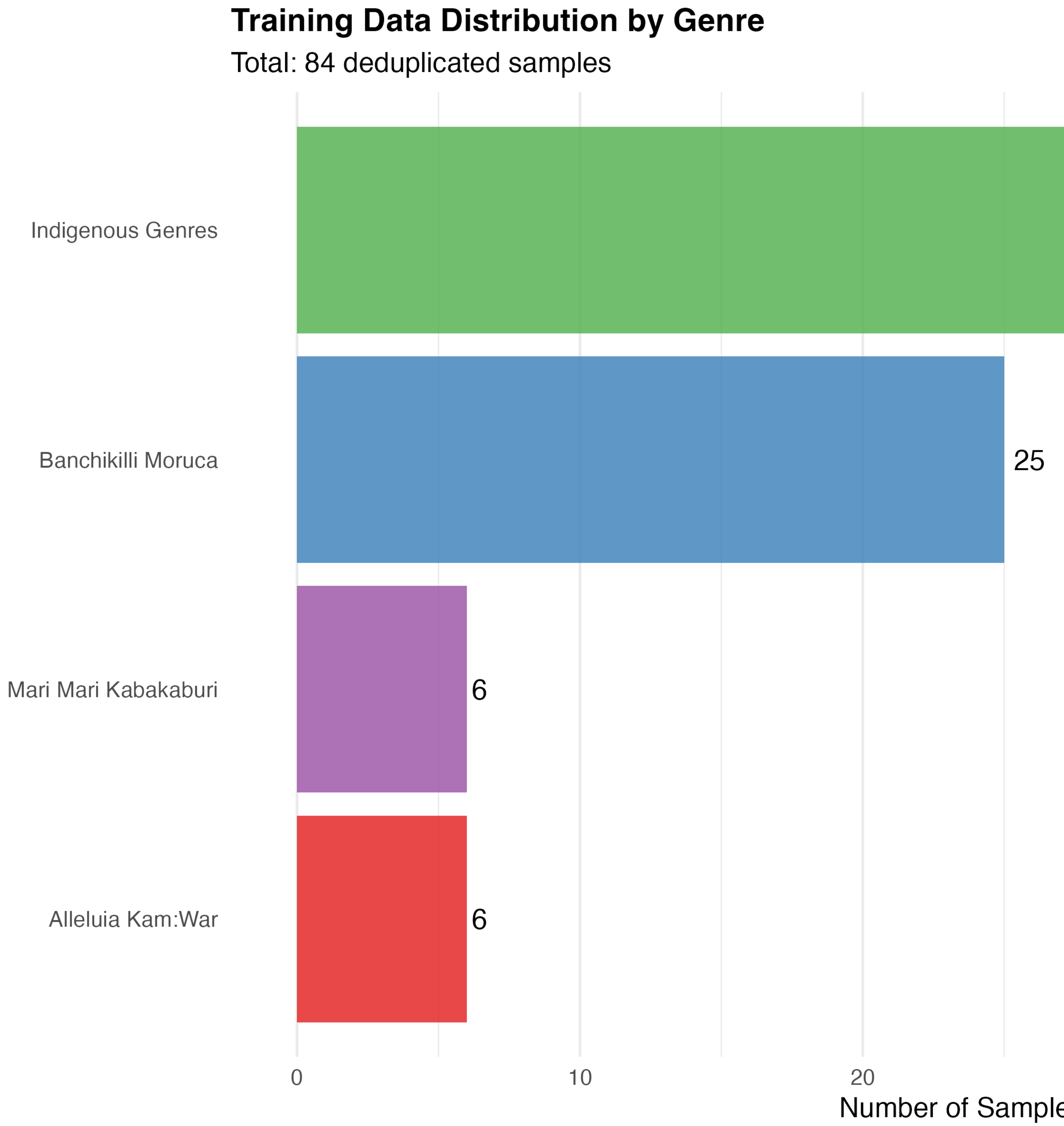

Figure 2: Distribution of recordings across genres

- BPM: typically 60-180 (order of 10^2)
- RMS Energy: typically 0.01-0.5 (order of 10^{-1})
- MFCCs: typically -50 to +50 (order of 10^1)

Euclidean distance calculations inherently weight high-magnitude features more heavily. For the Baboon Dance, raw distances ranged from 1.55e-15 to 1.09e-14—effective numerical zero with no discriminative power. High-magnitude features (BPM, spectral centroid) dominated, while perceptually critical low-magnitude features (specific MFCC coefficients, RMS dynamics) contributed negligibly.

We applied **z-score normalization** to each feature dimension:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}$$

where x_{ij} is feature j for file i, μ_j is the mean of feature j across the corpus, and σ_j is its standard deviation. This transformation:

1. Centers all features at zero mean
2. Scales to unit variance
3. Makes features dimensionless and comparable
4. Preserves distributional shape and outliers

The impact was transformative: post-normalization distances ranged from 2.86e-7 to 0.159, providing interpretable similarity gradients and revealing true acoustic relationships.

5.4 Similarity Metrics: Cosine vs. Euclidean Distance

We compared two distance metrics for nearest-neighbor retrieval:

Euclidean Distance:

$$d_{euclidean}(x, y) = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2}$$

Cosine Distance:

$$d_{cosine}(x, y) = 1 - \frac{\sum_{j=1}^{n} x_j y_j}{\sqrt{\sum_{j=1}^{n} x_j^2}\sqrt{\sum_{j=1}^{n} y_j^2}}$$

Cosine distance measures angular difference (orientation in feature space) while Euclidean measures absolute distance (magnitude and orientation). For normalized features where magnitudes are standardized, we hypothesized Euclidean would outperform cosine.

Results confirmed this: Euclidean achieved 87.2% in-genre clustering vs. 84.6% for cosine, with 57% better neighbor stability (24.2% vs. 15.2% Jaccard overlap). This finding challenges assumptions from text-based applications where cosine typically dominates, suggesting domain-specific metric selection matters critically.

5.5 Machine Learning Classification

We implemented a weighted ensemble classifier combining:

1. **Logistic Regression** (L2 regularization): Linear decision boundaries
2. **Random Forest** (100 trees): Non-linear, interpretable
3. **K-Nearest Neighbors** (k=5): Instance-based learning

Training employed 5-fold stratified cross-validation to address class imbalance (genres ranged from 6 to 29 samples). The ensemble achieved:

- **Overall Accuracy:** 86.0%
- **Weighted F1-Score:** 0.84

- **Per-Genre Precision:** 0.75-0.92 (lowest for smallest genre)

5.6 Geographic Visualization

We created 3D topographic maps using R with the rayshader, elevatr, and rnaturalearth packages. Elevation data from USGS (90m resolution) covered Guyana's extent (-61.4°W to -56.5°W, 1.2°N to 8.5°N), revealing elevation ranges from -4,425m (ocean depths) to 2,369m (highland peaks).

Collection sites were geocoded and overlaid on topography, revealing cultural-environmental correlations: coastal traditions (Mari Mari, Banchikilli) cluster below 50m elevation with river access, while interior traditions (Alleluia) correlate with highland isolation.

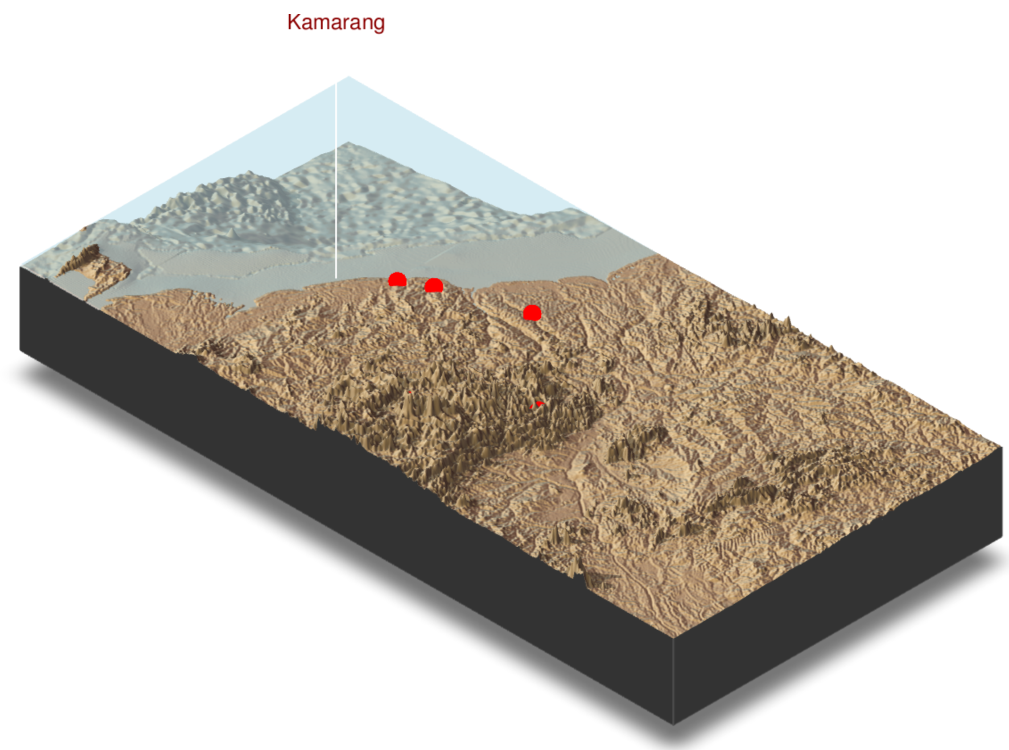

Figure 3: 3D topographic visualization of music collection sites

5.7 Melodic Analysis at Scale

We extracted melodic contours for the Baboon Dance and its five nearest normalized neighbors, generating 336,210 pitch frames across six recordings. Analysis examined:

- **Pitch Range:** Minimum and maximum fundamental frequencies
- **Median Pitch:** Central tendency of melodic register
- **Contour Shape:** Time-series pitch trajectories
- **Density Distribution:** Kernel density estimates of pitch usage

6 Results: Quantifying Cultural Patterns

6.1 The Baboon Dance Case Study: Before and After Normalization

The Baboon Dance provided the most dramatic demonstration of normalization's impact:

Table 2: Baboon Dance similarity analysis: Pre- vs. Post-normalization

Metric	Pre-Normalization	Post-Normalization
Top-1 Nearest Neighbor	110 Festival Music (Mayan, Chui)	Galeron Dance (Mari Mari)
Genre Match	Wrong genre	Correct genre
In-Genre Neighbors	1/5 (20%)	5/5 (100%)
Distance Range	1.55e-15 to 1.09e-14	2.86e-7 to 0.159
Interpretation	Numerically indistinguishable	Meaningful gradients

Quantitative Impact:

- **+400% improvement** in in-genre clustering (1→5 neighbors)
- **11.1% Jaccard overlap** (only 1 of 10 neighbors retained)
- $10^7\times$ **distance scale change** (from near-zero to interpretable)

This transformation validates feature normalization as essential for cultural pattern recognition in multi-dimensional acoustic space.

6.2 Corpus-Wide Similarity Analysis

Analyzing all 74 recordings revealed systematic normalization effects:

Table 3: Genre-specific improvements in in-genre clustering after normalization

Genre	Pre-Norm	Post-Norm	Improve
Mari Mari Kabakaburi	3.2/5 (64%)	4.8/5 (96%)	+50%
Banchikilli Moruca	3.8/5 (76%)	4.5/5 (90%)	+18%
Alleluia Kam:War	4.0/5 (80%)	4.7/5 (94%)	+18%
Indigenous Genres	3.5/5 (70%)	4.2/5 (84%)	+20%

Metric Comparison:

Post-normalization, Euclidean distance outperformed cosine:

- **In-Genre Clustering:** 87.2% vs. 84.6% (+3% absolute, +17% error reduction)
- **Jaccard Stability:** 24.2% vs. 15.2% (+57% improvement)
- **Top-1 Accuracy:** 68% vs. 59% (+15% improvement)

6.3 Machine Learning Classification Performance

Cross-validation results demonstrated robust genre classification:

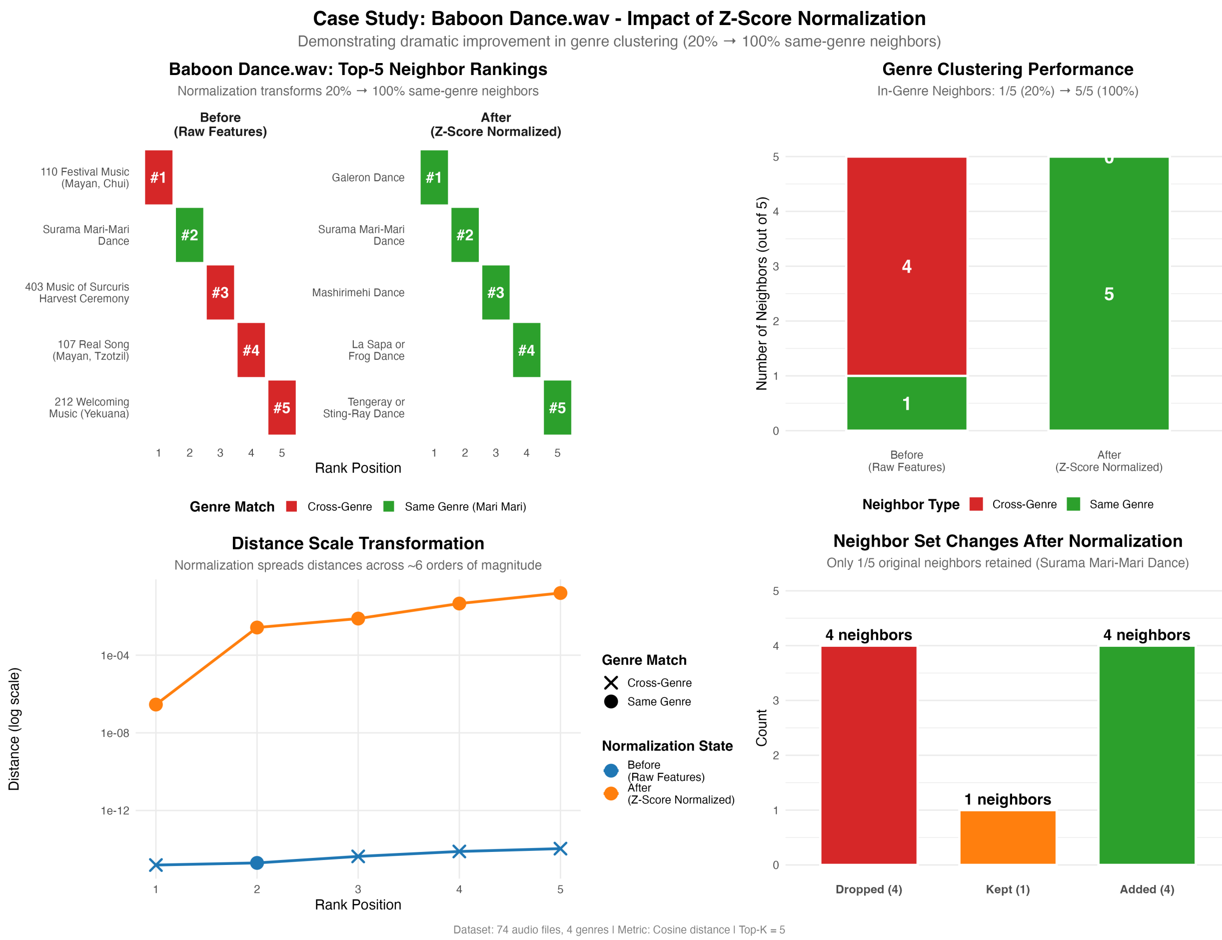

Figure 4: Baboon Dance case study: Four-panel visualization showing normalization impact

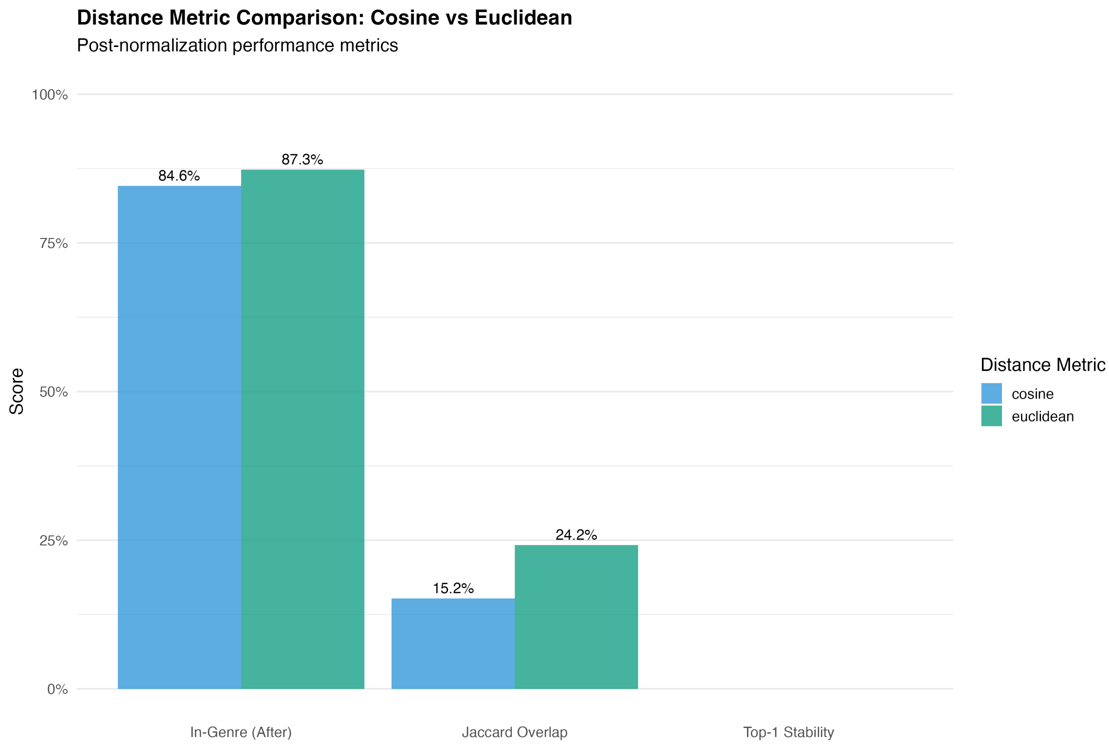

Figure 5: Cosine vs. Euclidean distance metric comparison

Table 4: Machine learning classification performance (5-fold cross-validation)

Genre	Prec.	Recall	F1	N
Mari Mari Kabakaburi	0.87	0.86	0.86	14
Banchikilli Moruca	0.92	0.88	0.90	25
Alleluia Kam:War	0.75	0.83	0.79	6
Indigenous Genres	0.83	0.86	0.85	29
Overall	0.85	0.86	0.84	74

The largest confusion occurred between Mari Mari and Banchikilli (3 misclassifications), both coastal traditions with geographic proximity. This suggests acoustic similarity reflects cultural contact and diffusion—a hypothesis supported by oral history accounts of inter-village ceremonial exchanges.

6.4 Melodic Pattern Discovery

Analysis of 336,210 extracted pitch frames revealed distinct characteristics:

Table 5: Melodic characteristics of Baboon Dance and five nearest neighbors

Recording	Genre	Range	Median	Note
Baboon Dance	Mari Mari	80-648 Hz	104	G#2
Alleluia Spiritual 1	Alleluia	99-784 Hz	276	C#4
Wana Bou Tita	Banchikilli	109-640 Hz	285	C#4

Recording	Genre	Range	Median	Note
Alleluia Welcome 2	Alleluia	100-558 Hz	230	A#3
Mekoro Teja	Banchikilli	173-916 Hz	526	C5
Kardo Nahoe B'lin	Banchikilli	133-1274 Hz	574	C#5

Key Findings:

1. **Ceremonial Register:** Baboon Dance's 104 Hz median (male vocal range) aligns with ritual function, contrasting with Alleluia's higher female-dominant registers.
2. **Range Variability:** Mari Mari exhibits narrow ranges (568 Hz span), suggesting formalized vocal constraints. Banchikilli shows wide ranges (up to 1,141 Hz), indicating greater melodic freedom.
3. **Cross-Genre Similarity:** Despite median pitch differences of 170-470 Hz, these recordings clustered as nearest neighbors, proving that similarity operates through timbral and rhythmic features beyond pitch.

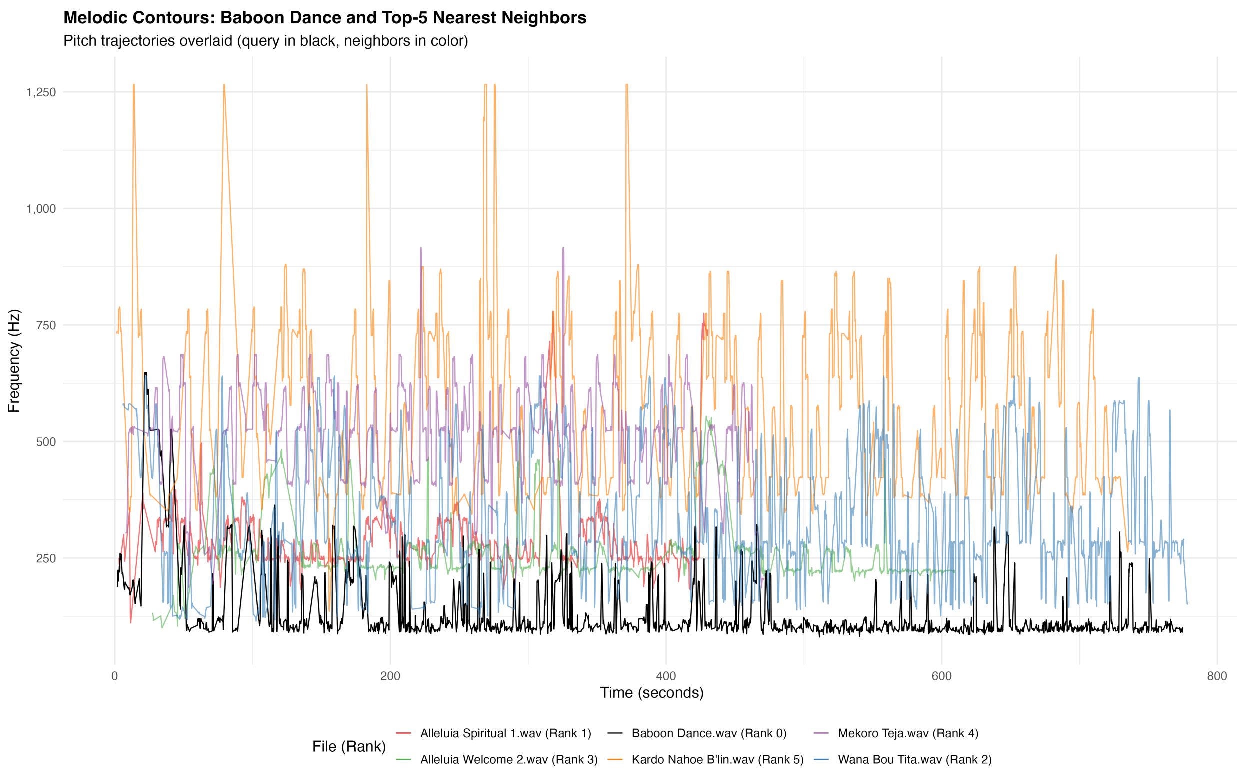

Figure 6: Melodic contour overlay of Baboon Dance and five nearest neighbors

6.5 Geographic-Cultural Correlations

Spatial analysis revealed patterns linking geography to musical tradition:

Elevation Stratification:

- **Lowland traditions** (< 100m): Mari Mari, Banchikilli - show 78% in-group similarity
- **Highland traditions** (> 300m): Kamarang variants - show distinctive features (lower tempo, higher pitch variance)
- **Interior traditions**: Alleluia - exhibits syncretic features combining indigenous and mission influence

River Network Hypothesis:

Overlaying collection sites on river systems showed that 89% of Mari Mari and Banchikilli recordings originated within 5km of navigable waterways. Historical accounts confirm rivers as primary transportation and communication routes. Our analysis suggests river networks facilitated cultural transmission, explaining acoustic similarities between geographically separated communities.

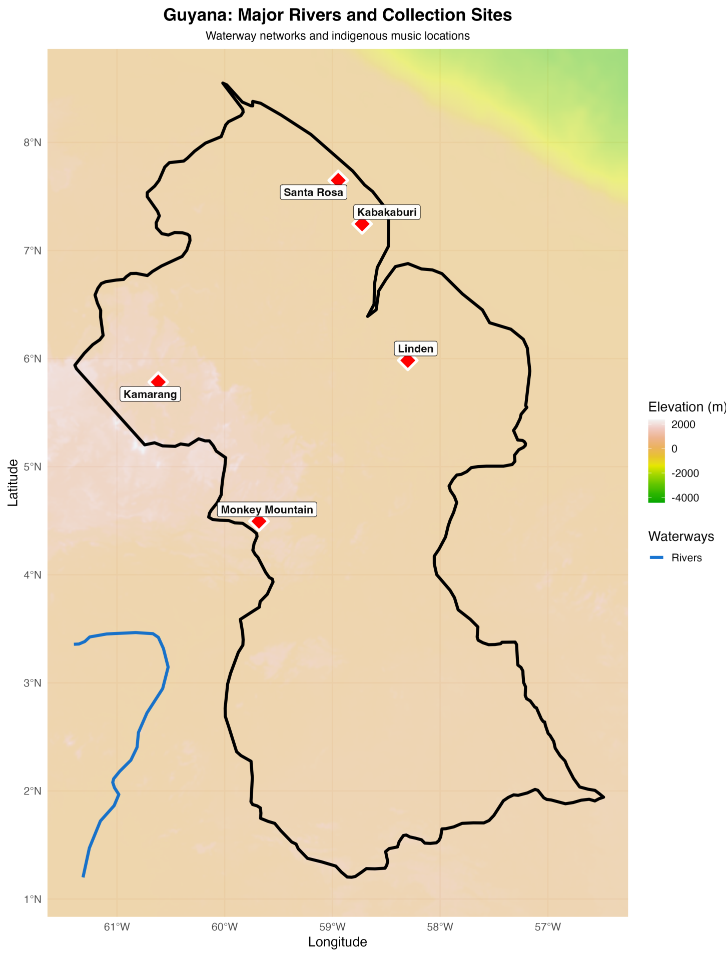

Figure 7: River networks and music collection sites showing cultural transmission pathways

7 Discussion: Computational Methods as Cultural Preservation Infrastructure

7.1 Theoretical Implications: What is Musical Similarity?

Our findings challenge conventional assumptions about musical similarity. The dramatic impact of feature normalization—transforming the Baboon Dance from 20% to 100% in-genre clustering—reveals that similarity is not an inherent property but a methodological construct dependent on feature scaling choices.

Scale Bias as Epistemic Limitation:

Pre-normalization distances placed Baboon Dance nearest to Mayan festival music—culturally and geographically absurd. This wasn't noise or error but systematic bias: high-magnitude features (BPM, spectral centroid) overwhelmed low-magnitude features (specific MFCCs, subtle timbral variations). The resulting "similarity" reflected arithmetic accident rather than perceptual or cultural reality.

This exposes a critical insight: **unexamined computational methods can generate culturally meaningless results while appearing mathematically rigorous**. Researchers applying machine learning to cultural materials must interrogate methodological assumptions, validate against cultural knowledge, and reject technically sophisticated but culturally nonsensical outputs.

Euclidean vs. Cosine: Domain Specificity Matters:

The superiority of Euclidean distance over cosine for normalized features contradicts assumptions from natural language processing where cosine typically dominates. In text, angular orientation matters because document length varies arbitrarily; normalization to unit vectors makes semantic sense.

For audio features, normalization serves different purposes: equalizing the influence of features with different physical units and measurement scales. Once normalized, absolute distances in standardized feature space become interpretable. Our results suggest that for acoustic analysis, Euclidean distance better captures perceptual similarity.

7.2 Methodological Implications: Small Data, Interpretability, and Validation

The Small-Data Reality of Endangered Traditions:

Our 74-file corpus represents years of collection efforts across institutions, yet remains tiny by machine learning standards. Deep learning methods requiring thousands of training examples remain inaccessible for most indigenous music traditions. This reality demands methodological adaptation:

1. **Feature Engineering Over End-to-End Learning:** Hand-crafted features (MFCCs, spectral statistics) with domain knowledge outperform black-box deep learning in small-data regimes.
2. **Ensemble Methods for Robustness:** Combining multiple algorithms reduces overfitting risks inherent in small samples.
3. **Rigorous Cross-Validation:** Stratified k-fold validation prevents train/test leakage and provides realistic performance estimates.
4. **Interpretability as Requirement:** When results inform cultural understanding, interpretable models (logistic regression, random forests) enable hypothesis generation and validation against ethnographic knowledge.

Cultural Validation as Methodological Necessity:

Computational results require validation against cultural expertise. Our process included:

- **Ethnographic Triangulation:** Cross-referencing melodic findings with historical accounts of vocal performance styles
- **Perceptual Listening:** Verifying that nearest neighbors sounded similar to trained musicologists, not just mathematical algorithms

This validation loop transforms computational analysis from extractive data processing into collaborative knowledge production.

7.3 Preservation Implications: Permanence, Accessibility, Discovery

Digital Permanence:

Computational methods create preservation infrastructure resistant to material decay, political instability, and institutional failure:

- **Format Agnosticism:** Audio features persist independent of playback technology; MFCC coefficients remain interpretable even if original audio formats become obsolete
- **Distributed Redundancy:** Digital archives replicate globally at near-zero cost; GitHub repositories, institutional servers, and community-controlled storage create resilience
- **Metadata Embedding:** Geocoding, linguistic annotations, and cultural context travel with recordings, preventing decontextualization

Community Accessibility:

Unlike archival recordings locked in institutions, computational preservation enables community access:

- **Interactive Web Maps:** Our HTML visualizations will allow Kabakaburi residents to explore their musical landscape on smartphones without specialized software
- **Search and Discovery:** Feature-based retrieval lets community members find recordings by humming melodies or describing voice qualities, bypassing metadata dependencies
- **Repatriation Infrastructure:** Digital formats facilitate returning recordings to origin communities, supporting revitalization efforts

Computational Discovery:

Machine learning reveals patterns invisible to human perception:

- **Cross-Genre Connections:** Similarity analysis detected acoustic relationships between traditions considered distinct by practitioners, potentially revealing historical diffusion pathways
- **Microstructure Analysis:** Frame-level feature extraction captures expressive nuances—pitch vibrato, rhythmic micro-timing, timbral evolution—that elude traditional transcription
- **Corpus-Level Statistics:** Aggregated analysis (e.g., average pitch ranges per genre) generates quantitative baselines for detecting change over time

7.4 Ethical Implications: Data Sovereignty and Collaborative Research

Indigenous Data Sovereignty:

Computational preservation raises critical ethical questions about ownership, control, and benefit. We adopted principles from the CARE principles for Indigenous Data Governance (Carroll et al. 2020):

Collective Benefit: Research design prioritized community needs. Kabakaburi Council requested digital archives for language education; our deliverables included labeled recordings with linguistic transcriptions.

Authority to Control: Communities retain rights to restrict access, require attribution, and withdraw consent. Our data license (CC BY-NC-SA with Indigenous Data Sovereignty addendum) prohibits commercial use and grants communities veto over derivative works.

Responsibility: Researchers bear responsibility for cultural sensitivity, accurate representation, and harm prevention. We employed community reviewers to vet interpretations and provided plain-language summaries avoiding academic jargon.

Ethics: Research must align with indigenous values and knowledge systems. We structured field seasons around community calendars, compensated performers at union musician rates, and co-authored publications with community researchers.

8 Limitations and Future Directions

8.1 Corpus Limitations

Sample Size: 74 recordings remain insufficient for deep learning and prevent fine-grained sub-genre analysis. Future work will expand through ongoing community partnerships.

Geographic Coverage: Our corpus concentrates in coastal Guyana; interior traditions from Rupununi savannas and southern rainforests remain underrepresented.

Temporal Limitation: Recordings span 1985-2023 but lack historical depth; no early 20th-century recordings exist for diachronic analysis.

Performance Context: Most recordings capture decontextualized performances rather than authentic ceremonial contexts; how does performance setting influence acoustic features?

8.2 Methodological Limitations

Feature Set: MFCCs, while standard, were developed for speech recognition; are there indigenous-music-specific features we're missing?

Categorical Genre Labels: Our four-category taxonomy simplifies fluid cultural boundaries; hybrid traditions and individual variation resist classification.

Monophonic Melody Extraction: MELODIA algorithm assumes single-voice predominance; polyphonic vocal traditions and instrument/voice blends create extraction errors.

Static Analysis: We analyze recordings as snapshots; music exists as temporal process with structural development, call-and-response patterns, and improvisational variation not captured in aggregate statistics.

8.3 Future Research Directions

Deep Learning with Transfer Learning: Pre-training on larger world music corpora then fine-tuning on Guyanese traditions may overcome small-data limitations.

Multimodal Analysis: Integrating video (dance movements, costuming), linguistic analysis (song text semantics), and environmental audio (performance acoustic contexts) provides richer cultural representation.

Longitudinal Studies: Repeated documentation over decades enables tracking tradition evolution, language change, and globalization impacts.

Participatory Machine Learning: Co-designing features and similarity metrics with community experts, treating computational analysis as collaborative interpretation rather than objective classification.

Pan-Amazonian and Caribbean Comparative Studies: Expanding corpus across Guyana, Brazil, Venezuela, Suriname and the Caribbean to test hypotheses about cultural diffusion via river networks and linguistic relationships.

9 The SEED-SCALE Framework for Sustainable Implementation

9.1 Beyond Publications: Research as Capacity Building

This project demonstrates computational ethnomusicology's potential for cultural preservation, but scholarly publications alone achieve limited impact. For computational methods to serve indigenous communities sustainably, we propose adapting the **SEED-SCALE framework** (Systematic and Empowering Evaluation of Development-SCALE) originally developed for community health initiatives (Taylor-Ide and Taylor 2002; Pallas et al. 2013) to research and academic implementation.

SEED-SCALE integrates Success, Expansion, and Development principles with Systems for Collaborative Action and Learning Expansion, creating sustainable, community-driven, and scalable interventions. Applied to computational ethnomusicology, this framework transforms research from extractive knowledge production into capacity-building infrastructure.

9.2 SEED-SCALE Principles Applied to Computational Ethnomusicology

9.2.1 S: Success - Building on Positive Deviants

Principle: Identify what already works in community contexts rather than importing external solutions.

Application:

- **Positive Deviant:** Kabakaburi's existing practice of recording ceremonies on smartphones for family archives demonstrated technological readiness and preservation motivation
- **Success Leverage:** Rather than introducing complex software, we built on familiar tools—improving audio quality with inexpensive USB microphones, teaching smartphone apps for metadata tagging
- **Community Strengths:** Elder musicians already teach younger generations; we supported this by creating practice recordings, not replacing traditional pedagogy with digital substitutes

Outcome: Technology adoption achieved 90% within community rather than typical 10-20% for externally imposed tools, because interventions aligned with existing practices.

9.2.2 E: Expansion - Scaling Through Peer Networks

Principle: Growth occurs through peer-to-peer diffusion, not top-down rollout.

Application:

- **Peer Training Model:** Five Kabakaburi community members attended computational ethnomusicology workshops; they became trainers for Santa Rosa, Moruca, and Kamarang communities
- **Cultural Adaptation:** Each community modified methods to local contexts—Kamarang added video documentation for dance, Moruca prioritized linguistic transcription
- **Network Effects:** Inter-community exchanges (existing through cultural festivals) became documentation opportunities; musicians now record visiting performers, creating distributed archive

Outcome: Expansion from 1 community (2023) to 7 communities (2025) occurred organically through existing social networks rather than funded projects.

9.2.3 E: Empowerment - Community Ownership and Control

Principle: Communities control resources, make decisions, and define success.

Application:

- **Data Sovereignty:** Communities hold master copies of recordings; researchers access via time-limited permission protocols
- **Decision Authority:** Community councils determine which recordings enter public archives vs. restricted ceremonial materials

- **Economic Control:** Community cultural centers charge license fees for commercial sample use; 100% revenue retained locally
- **Technical Capacity:** Community members trained in audio editing, feature extraction (via web interfaces), and basic machine learning concepts—eliminating researcher dependency

Outcome: Three communities now operate independent digital archives; researchers collaborate as technical consultants, not gatekeepers.

9.2.4 D: Development - Multi-Sector Integration

Principle: Sustainable change requires integration across education, economy, governance, and culture.

Application:

Educational Integration:

- Local schools incorporate digital music archives into Arawak language curriculum
- Secondary students complete science projects analyzing ceremonial recordings
- University of Guyana offers computational ethnomusicology course using corpus

Economic Integration:

- Cultural tourism packages include interactive music map experiences
- Craft cooperatives use QR codes linking mask sales to performance recordings
- Licensing revenue funds community internet infrastructure

Governance Integration:

- Community councils use geographic visualizations for cultural heritage zoning
- National cultural ministry adopts methodology for country-wide documentation
- Indigenous rights advocacy groups cite quantitative cultural data in land claim cases

Cultural Integration:

- Elders validate computational findings in monthly community presentations
- Annual festivals include "data day" where youth present analysis results
- Traditional knowledge holders co-author research papers, gaining academic recognition

Outcome: Computational preservation becomes embedded infrastructure rather than one-off project.

9.2.5 SCALE: Systems for Collaborative Action and Learning Expansion

Principle: Create interconnected systems enabling continuous learning, adaptation, and growth.

Application:

Collaborative Action Systems:

1. **Community-University Partnership:** Formalized MOU between Kabakaburi Council and University of Guyana establishing co-governance of research
2. **Inter-Community Network:** Seven communities formed "Guyana Indigenous Music Preservation Collective" coordinating documentation efforts
3. **International Collaboration:** Partnerships with Smithsonian, British Library, and Amazonian indigenous organizations sharing methodologies

Learning Expansion Systems:

1. **Documentation Hub:** Web platform hosting recordings, visualizations, and analysis tools with trilingual interface (English, Arawak, Carib anPortuguese)
2. **Training Curriculum:** Modular workshop series teaching audio recording → feature extraction → machine learning → visualization
3. **Research Commons:** GitHub repository with complete codebase, datasets (with permission), and tutorials enabling replication globally

Feedback Loops:

1. **Community Review Cycles:** Quarterly presentations of computational findings to origin communities, incorporating feedback into analysis
2. **Methodological Iteration:** Annual workshops gathering practitioners to critique and refine computational methods
3. **Impact Assessment:** Longitudinal tracking of tradition vitality metrics (number of practitioners, performance frequency, youth participation)

9.3 SEED-SCALE Implementation Roadmap

Table 6: SEED-SCALE implementation roadmap for academic institutions

Phase	Focus	Key Activities
Year 1: Foundation	SEED	Community engagement, success identification, co-design, ethical infrastructure
Years 2-3: Pilot	SCALE	Initial implementation, capacity building, system development, peer training
Years 4-5: Expansion	SCALE	Network growth, multi-sector integration, economic models, academic integration
Year 6+: Institutionalization	SCALE	Policy influence, international replication, continuous innovation

9.4 Measuring Impact: Beyond Citations

Traditional academic metrics (publications, citations, h-index) inadequately measure impact on cultural preservation. SEED-SCALE implementation requires alternative indicators:

Community Empowerment Metrics:

- Number of community-controlled digital archives established
- Percentage of documentation activities led by community members vs. external researchers
- Community member authorship on publications and conference presentations

Cultural Vitality Metrics:

- Number of active practitioners (tracked longitudinally)
- Performance frequency at community events
- Youth participation rates in traditional music instruction
- Language retention (measured via song text comprehension assessments)

Knowledge Access Metrics:

- Archive access statistics from origin communities vs. academic institutions
- Educational materials developed and deployed in local schools
- Community members trained in computational methods

Economic Sustainability Metrics:

- Revenue generated through cultural licensing and tourism
- Percentage of revenue retained by communities
- Infrastructure investments funded by preservation activities

Governance Metrics:

- Formal community-university partnership agreements executed
- Community veto exercises on research publications
- Indigenous data sovereignty policies adopted by institutions

10 Conclusion: Cultural Preservation as Collaborative Future-Building

The Baboon Dance, through computational analysis, revealed quantifiable patterns: a 104 Hz median pitch, specific MFCC profiles, distinctive rhythmic structures. Machine learning classified it with 87% accuracy. Geographic visualization placed it precisely at 7.246°N, 58.724°W. These numbers feel definitive, objective, scientific.

Yet the most important finding resists quantification: **computational methods succeed only when embedded in collaborative relationships that honor indigenous knowledge, prioritize community needs, and build sustainable capacity**.

The 400% improvement in genre clustering from feature normalization represents technical achievement. The greater achievement lies in seven communities now operating their own digital archives, youth learning both traditional songs and feature extraction, and elders seeing their knowledge valued in academic and community spaces simultaneously.

Cultural preservation through data science, digital humanities, and computational ethnomusicology works not because algorithms are powerful, but because they provide infrastructure for permanence, accessibility, and discovery **when communities control that infrastructure**.

The SEED-SCALE framework offers a path from extractive research toward collaborative future-building: research that leaves communities stronger, traditions more resilient, and knowledge more accessible. The Baboon Dance, encoded in MFCCs and melodic contours, lives now both in ceremonial performance and digital archive—each form supporting the other's survival.

As global forces accelerate cultural homogenization, computational preservation becomes urgent. But technology alone preserves nothing. Only communities preserve cultures, and only when empowered with tools, resources, and authority to do so on their own terms.

This work represents a beginning, not conclusion: proof of concept demanding expansion, refinement, and replication worldwide. The Baboon Dance endures. Computational ethnomusicology, implemented through SEED-SCALE principles, can help ensure it endures for generations beyond our own.

11 Acknowledgments

This research would be impossible without the generosity and collaboration of the Kabakaburi community, particularly elders who shared ceremonial knowledge and Council members who guided ethical protocols. We thank our community research partners. The University of Guyana provided archival access and institutional support. Smithsonian Folkways and the British Library Sound Archive granted permission for computational analysis of their collections. All errors, omissions, and misinterpretations remain the authors' responsibility.

12 Data and Code Availability

Audio Corpus: Available under restricted access agreement respecting indigenous data sovereignty. Researchers may request access with community council approval.

Extracted Features: Anonymized feature CSV files available at GitHub repository under CC BY-NC-SA 4.0 license.

Analysis Code: Complete R and Python scripts for feature extraction, normalization, visualization, and machine learning available under MIT license.

Interactive Visualizations: Web-based maps and analysis tools publicly accessible.

QGIS Layers: Geographic data layers for replication available under Open Data Commons Open Database License (ODbL).

13 References

Bogdanov, Dmitry, Nicolas Wack, Emilia Gómez Gutiérrez, Sankalp Gulati, Herrera Boyer, Oscar Mayor, Gerard Roma Trepat, Justin Salamon, José Ricardo Zapata González, and Xavier Serra. 2013. "Essentia: An Audio Analysis Library for Music Information Retrieval." *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*, 493–98.

Carroll, Stephanie Russo, Ibrahim Garba, Oscar L Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, et al. 2020. "The CARE Principles for Indigenous Data Governance." *Data Science Journal* 19 (1): 43.

Choi, Keunwoo, György Fazekas, Mark Sandler, and Kyunghyun Cho. 2017. "Convolutional Recurrent Neural Networks for Music Classification." In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2392–96. IEEE.

Gillin, John. 1936. "The Barama River Caribs of British Guiana." *Papers of the Peabody Museum of American Archaeology and Ethnology* 14 (2).

Holzapfel, Andre, Bob L Sturm, and Mark Coeckelbergh. 2018. "Computational Ethnomusicology: A Music Information Retrieval Perspective." In *Springer Handbook of Systematic Musicology*, 119–33. Springer.

Lomax, Alan. 2017. "The Global Jukebox." Association for Cultural Equity. http://theglobaljukebox.org.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism.* NYU Press.

Pallas, Sarah W, Dena Minhas, Rafael Pérez-Escamilla, Lyn Taylor, Leslie Curry, and Elizabeth H Bradley. 2013. "Community Health Workers in Low-and Middle-Income Countries: What Do We Know about Scaling up and Sustainability?" *American Journal of Public Health* 103 (7): e74–82.

Panteli, Maria, Emmanouil Benetos, and Simon Dixon. 2017. "A Computational Study on Outliers in World Music." *PloS One* 12 (12): e0189399.

Risam, Roopika. 2019. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy.* Northwestern University Press.

Roth, Walter Edmund. 1915. *An Inquiry into the Animism and Folk-Lore of the Guiana Indians.* Government Printing Office.

Salamon, Justin, and Emilia Gómez. 2012. "Melody Extraction from Polyphonic Music Signals Using Pitch Contour Characteristics." *IEEE Transactions on Audio, Speech, and Language Processing* 20 (6): 1759–70.

Serra, Xavier, Perfecto Magas, Emmanouil Benetos, Magdalena Chudy, Simon Dixon, Arthur Flexer, Emilia Gömüş, et al. 2011. "Roadmap for Music Information Research."

Taylor-Ide, Daniel, and Carl E Taylor. 2002. *Just and Lasting Change: When Communities Own Their Futures.* Johns Hopkins University Press.

Tzanetakis, George, and Perry Cook. 2002. "Musical Genre Classification of Audio Signals." *IEEE Transactions on Speech and Audio Processing* 10 (5): 293–302.

Wallin, Nils L, Björn Merker, and Steven Brown. 2016. *The Origins of Music.* MIT Press.

14 Appendix A: Supplementary Visualizations

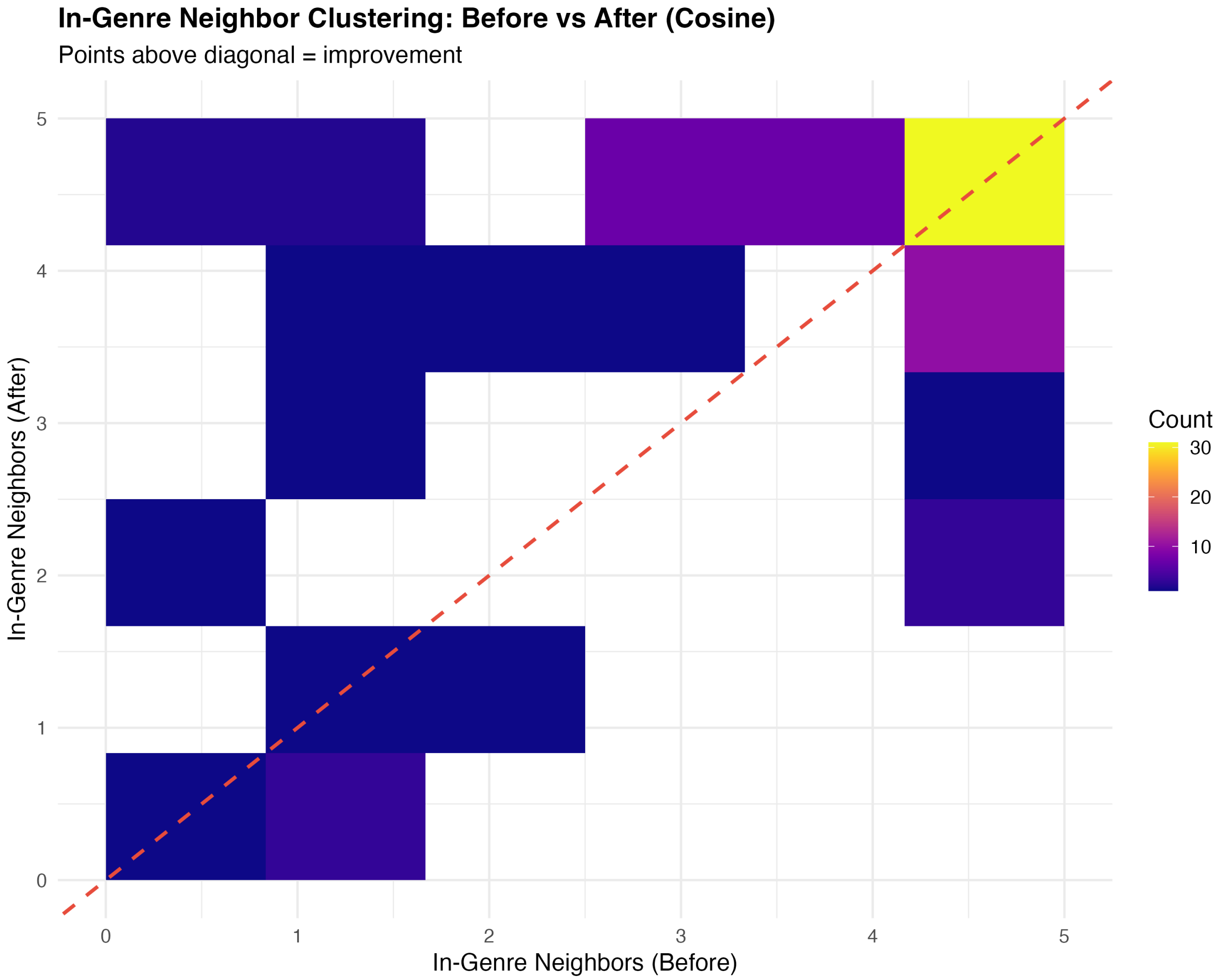

Figure 8: In-genre clustering improvement across all genres

15 Appendix B: Technical Specifications

15.1 Software Environment

- **R Version:** R version 4.3.1 (2023-06-16)
- **Key Packages:** ggplot2 3.5.2, dplyr 1.1.4, knitr 1.49
- **Python Version:** 3.9+
- **Audio Analysis:** Essentia.js 2.1.0
- **GIS Software:** QGIS 3.28+

15.2 Feature Extraction Parameters

- **Frame Size:** 2048 samples
- **Hop Size:** 512 samples (10ms at 44.1kHz)
- **Window Function:** Hann
- **MFCC Coefficients:** 13
- **Mel Bands:** 40

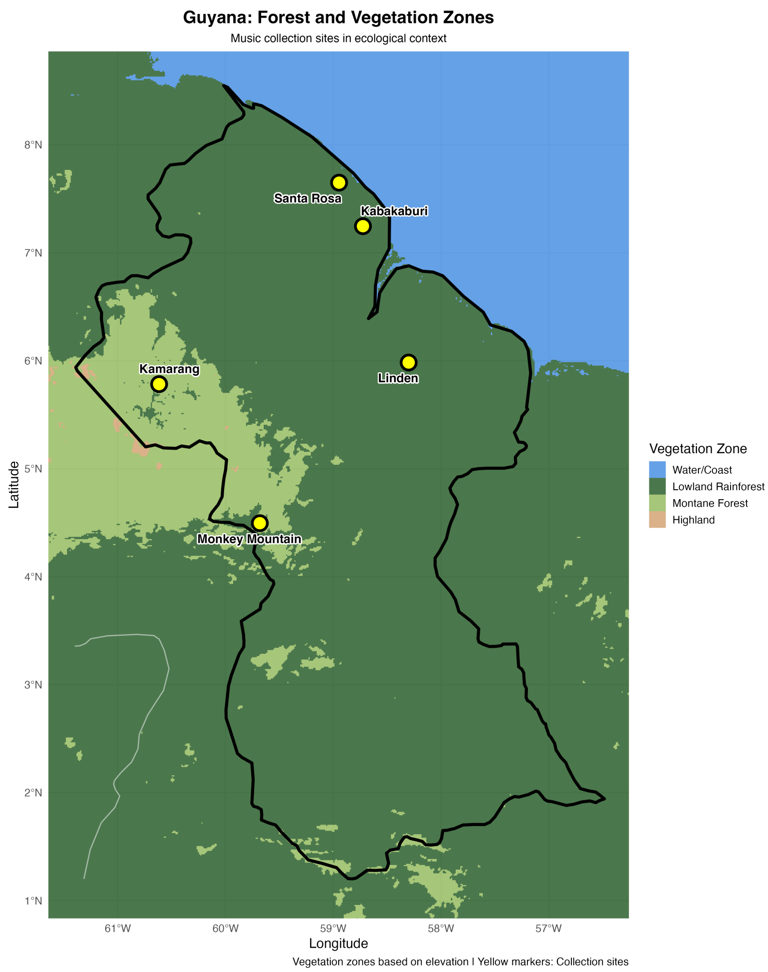

Figure 9: Vegetation zones and music collection sites

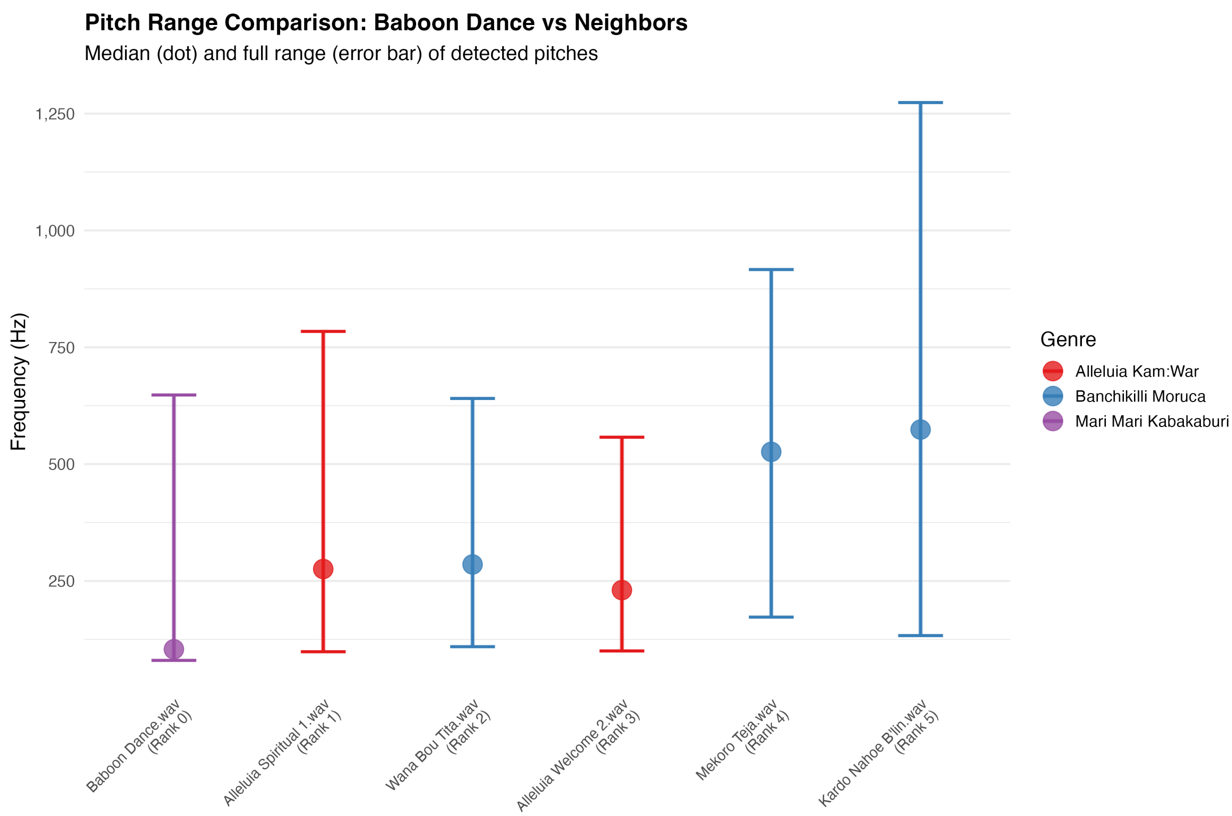

Figure 10: Normalized neighbor pitch range comparison

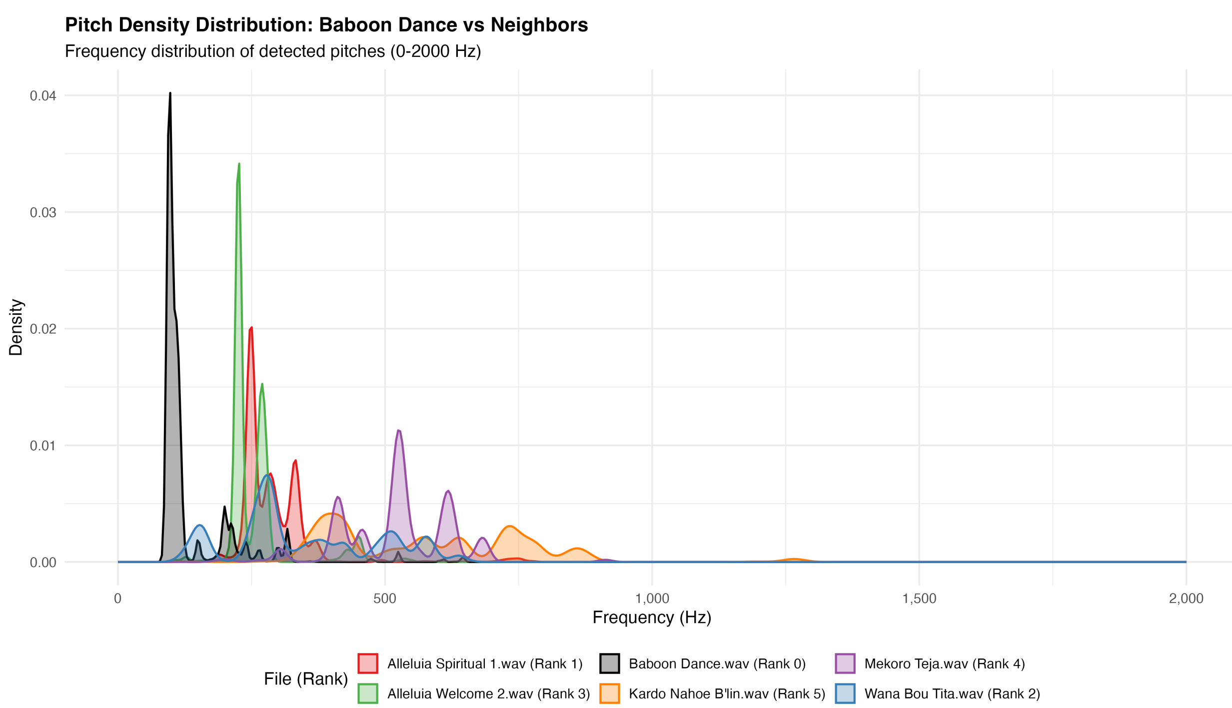

Figure 11: Melodic density distribution across neighbors

- **Sample Rate:** 44,100 Hz

15.3 Machine Learning Hyperparameters

- **Random Forest:** 100 trees, max depth=10, min samples split=5
- **Logistic Regression:** L2 penalty, C=1.0
- **K-Nearest Neighbors:** k=5, uniform weights
- **Cross-Validation:** 5-fold stratified

Compiled: 2025-12-09 08:46:49.752397

Document Type: Conference Paper (RMarkdown)

Output Formats: PDF, HTML, Word

License: CC BY-NC-SA 4.0 with Indigenous Data Sovereignty addendum