Faculdade de Engenharia da Universidade do Porto

Computational Corpus Analysis of Expressive Timing in Galician and Irish Folk Traditions

Mário Pereira

Masters in Multimedia

Supervisor: Gilberto Bernardes

July 2, 2025

Computational Corpus Analysis of Expressive Timing in Galician and Irish Folk Traditions

Mário Pereira

Masters in Multimedia

July 2, 2025

Abstract

This research presents a computational corpus analysis of expressive timing in Galician and Irish folk music traditions, addressing a significant gap in Music Information Retrieval (MIR) and Music Performance Analysis (MPA) research where oral tradition music has received limited systematic computational study. Conducted as part of the broader EA-DIGIFOLK project, this investigation develops and applies methodological frameworks for analysing temporal characteristics in folk music performances to reveal how performers use timing variation as a means of musical expression.

The study analyses 253 performances across 49 songs from the two traditions: 154 Irish performances spanning 19 songs (including reels, jigs, hornpipes, and polkas) and 99 Galician performances across 30 songs (featuring muiñeiras, jotas, foliadas, and other traditional forms). Due to limitations in automatic beat tracking algorithms in this kind of corpus, the methodology employs manual beat tracking combined with computational analysis techniques at both performance and phrase levels. Statistical analysis and hierarchical clustering are applied to identify distinct temporal expression patterns within the corpus.

The research contributes methodologically to the field by establishing a systematic, replicable framework for computational analysis of expressive timing in oral tradition music that is expandable to other traditions and supports the digital preservation of cultural heritage.

This work demonstrates how computational methods can complement traditional ethnomusicological approaches while respecting the fluid and communal nature of folk music performance, providing new insights into the systematic analysis of cultural timing practices.

Acknowledgments

Over the past week, this document was brewed under more pressure than an espresso shot. And there were many of those. Not that they have any effect, I can have one in the evening and sleep just fine. I just like the taste. The materialisation of this work would have still been possible, but far harder, without the people surrounding me. The studio squad, Pedro "Maquinão" Magalhães, Nawaraj "NN" Khatri, João "Fonsi" Fonseca, Shayne "the tall Irish guy" Fogarty, constituted a core working dynamic throughout the semester, and we are going for a drink after this. Along with the rest of our colleagues and friends throughout these last two years.

Someone without whom this actually would not have been possible is my supervisor Gilberto Bernardes, so I would like to express my gratitude. I must also thank Hilda Romero Velo for providing essential materials and helping with their analysis. And also professor António Pinto for the valuable insights.

Outside of university, I would thank my mother , father and sister for continuous support and my *fellas* at *Ryo Tynto NiteClub*, who occasionally watched me fumble with the code at 2 a.m.

Even farther outside, I will thank Magnus Carlsen and Anna Muzychuk, winners of Norway Chess 2025, along with all the other participants. That was an exciting event to follow. And I will thank Tobias Forge and the entire team behind *Ghost* for the magical concert a couple of months ago.

I am gonna go have another coffee now. Cheers.

Mário Pereira

“Guys, I might have finished my thesis, and I might deliver it now.”

Pedro "Maquinão" Magalhães

Contents

List of Figures

List of Tables

Abbreviations and symbols

MIR	Music Information Retrieval
MPA	Music Performance Analysis
DTW	Dynamic Time Warping
ITMA	Irish Traditional Music Archive
CE	Computational Ethnomusicology
KDE	Kernel Density Estimation
IQR	Interquartile Range
BPM	Beats Per Minute

Chapter 1

Introduction

In Section 1.1, we will discuss the concept of folk music, followed by 1.2 and 1.3, where we go over the aspects of Irish and Galician musical traditions relevant to our research. In Section 1.4, we define the context and objectives of the dissertation. And finally, in Section 1.5 we describe the structure of the document.

1.1 Folk music

Folk music is traditionally understood as music that exists primarily through oral transmission within specific communities, often rural, where repertoire is learned informally rather than via notation. It is characterised by functional association with social activities, such as work, ritual or communal gatherings, and by an emphasis on participation rather than formal performance. Unlike art music or popular music, tied respectively to aesthetic contemplation or mass-mediated entertainment, folk song embodies "living tradition", evolving through continual communal recreation (Nettl, 2025).

Barry (1911) reinforces this view by distinguishing folk-song from art-song on the basis of dynamic communal ownership: a folk song's words and melody may shift with each singer's rendition, reflecting a collective creative process rather than adherence to a single authored version. As Barry (1914) emphasises, the folk singer is the primary medium of transmission, absorbing and reshaping material according to communal memory and individual interpretation. He also observes that folk singers, whether solitary labourers or participants in social societies, learn folk songs from diverse sources, be it other singers or circulated text, and integrate new and ancient material into a shared repertoire. This process produces variants, or versions, that coexist within and across communities, ensuring both continuity and innovation (Nettl, 2025).

A comprehensive performance analysis of folk song can combine ethnomusicological, Music Information Retrieval (MIR) and Music Performance Analysis (MPA) approaches. Ethnographic methods like participant observation, interviews, and contextual analysis elucidate the cultural meanings behind stylistic traits (Aldred, 2005). And quantitative audio analysis methods (e.g. feature extraction) provides objective metrics for comparing performances across recordings

(Guimarães, 2024). Integrating these methods allows scholars to trace how oral tradition patterns manifest as measurable musical features, while respecting the fluid and communal nature of folk song performance.

1.2 Irish Folk Music

Irish folk music represents a living cultural tradition that has undergone significant transformation whilst maintaining its essential character. The performance contexts have evolved substantially from historical practice. Originally, Irish traditional music served as functional accompaniment for social dancing and was predominantly performed within domestic settings. However, a fundamental shift occurred during the 1960s-1980s, when the tradition relocated "from the kitchen to the pub" (Falc'her-Poyroux, 2013).

The contemporary session culture, now present in Irish pubs globally, represents a new social ritual that has fundamentally altered the music's cultural function. The music has also transitioned from serving dancers to entertaining listeners. This functional shift has significant implications for musical interpretation and performance style, as performers no longer prioritise the rhythmic demands of dance accompaniment (Falc'her-Poyroux, 2013).

Irish traditional music encompasses several distinct performative styles, also known as tune types. Some of these being reels, hornpipes, jigs and its variants, polkas, waltzes, marches, mazurkas, just to name a few. Over this dissertation we will be lending a bit more focus towards jigs and reels.

There are four variants of jigs in the Irish music repertoire, double jig, single, slip and slide. Double jigs are the most common, and when anyone refers to a jig without specifying, this is likely what is implied (Guimarães, 2024). They are typically notated in a 6/8 time signature (Figure 1.1), with the typical pattern of two groups of three eighth notes per measure and are known for their generally lively tempo (Vallely, 2011). The rhythmic emphasis is usually on the first and fourth eighth notes, creating the characteristic *ONE-two-three, ONE-two-three* pattern.

The reel represents the most prevalent form in Irish traditional music repertoire. Reels are written in simple duple time, creating a more driving, continuous rhythmic pulse. They are typically notated in 4/4 time (Figure 1.2), even if felt in two (or cut time), with accents on the first and third beats of each bar (Hast and Scott, 2004; Vallely, 2011).

Both these types of tunes follow the same overall structure, consisting of two main sections commonly designated as A and B parts. Each section is usually eight bars in length and repeated, creating an AABB pattern (Hast and Scott, 2004; Vallely, 2011).

Within the context of phrase structure, we can also mention the concept of "hooks". These are memorable elements necessary for a tune's popularity and transmission within the tradition (Hillhouse, 2005). They often appear at phrase beginnings or cadential points, providing structural landmarks for both performers and listeners.

Irish traditional music employs a distinctive instrumental palette that has evolved considerably over time. Since the eighteenth century, major instruments include the violin (also referred to as

Banish Misfortune

Figure 1.1: Transcription of *Banish Misfortune*, a jig, retrieved from The Session.

The Mason's Apron

Figure 1.2: Transcription of *The Mason's Apron*, a reel, retrieved from The Session.

fiddle), flute, whistle, and two of the most emblematic, uilleann pipes (an Irish variant of bagpipes) and the harp. A bit later, the accordion and concertina were also introduced, and also became a core instruments of the tradition (Hast and Scott, 2004).

Irish traditional music preservation has undergone significant transformation through technological advancement. The shift from oral to media-transmitted channels represents a fundamental change in cultural transmission methods (Falc'her-Poyroux, 2013).

Several online repositories facilitate tune sharing and preservation. Websites such as The Session[1] and the Irish Traditional Music Tune Index[2] enable musicians to submit new compositions, comment on existing repertoire, and access notated materials. These platforms represent a democratisation of tune transmission, allowing global participation in repertoire development. The Irish Traditional Music Archive (ITMA) serves as a primary institutional repository, providing professional curation and scholarly access to traditional music materials. The EA-DIGIFOLK project, with which this work is associated, also encompasses music from the Irish tradition.

Digital preservation addresses both practical and theoretical concerns. It ensures continuous access to traditional materials regardless of geographical constraints, while also raising questions about authenticity, transmission methods, and the relationship between recorded and live performance traditions. Contemporary preservation efforts thus balance traditional concerns with technological capabilities, creating new opportunities for cultural continuity whilst potentially adding to the traditions through mediated transmission.

Computational methods have been used in the study of Irish traditional music for some years. Duggan et al. (2008) worked on tune identification by melodic similarity through feature extraction, which led to the development of the MATS algorithm (Vallely, 2011).

Recent developments in computational analysis have also provided new insights into Irish traditional music. Jiménez-Bravo et al. (2024) designed a system to classify different tune types based on audio analysis. Navarro-Cáceres et al. (2025) used machine learning techniques for diatonic mode detection. And in his dissertation, Guimarães (2024) employs both feature engineering and deep learning approaches to analyse Irish folk music similarity.

1.3 Galician Folk Music

Galician folk music serves as a fundamental expression of cultural identity and has played a crucial role in the promotion of national consciousness (Aldred, 2005). The music functions within multiple contexts that reflect its deep integration into Galician society, accompanying religious festivals, social gatherings, dances, and community celebrations.

The contemporary revival has expanded these traditional contexts to include concert halls, folk festivals, and international venues. Since the 1970s, the recovery of Galician musical heritage has

[1] https://thesession.org/ (last accessed June 2025)
[2] https://www.irishtune.info (last accessed June 2025)

Figure 1.3: Transcription of *Muiñeira de Albite*, retrieved from Folkoteca Galega.

occurred hand in hand, somewhat paradoxically, with innovation, transformation and hybridization. This transformation has allowed Galician folk music to maintain strong and direct links with its own past, while heading towards new horizons (Colmeiro, 2014).

Three genres have been generally accepted as the most characteristic of Galician traditional music: the alborada, the muiñeira and the alalá (Aldred, 2005). Additional important styles include the foliada, jota, marcha, pasadoble, to name a few. Each of these forms demonstrates the rich diversity within Galician folk music while maintaining distinctive characteristics.

Jotas and Muiñeiras will be more prevalent in our analysis further ahead. According to Schubarth (1984), and both have an archaic character, which lends to a lot of freedom in interpretation, both in terms of embellishment and phrase structure.

Muiñeiras are faster and are manifested in a binary meter - three groups of two quavers. Phrase structure present great variability, and arrangements can go from AAAA to ABCD (Schubarth, 1984). They are typically notated in 6/8 time (Figure 1.3), but there are variations (Aldred, 2005).

Jotas are slower, more melodic, and have a ternary meter - two groups of three quavers. Structurally, they have two recognisable parts, the *punto* and the *volta*, which can be freely combined, much like in the muiñeira (Schubarth, 1984). They are typically notated in 3/4 or 3/8 (Figure 1.4) (Mira et al., 2023).

From observation of ensemble performances, in both of these styles, the percussion instruments tend to accentuate the first beat in each measure, implying this may the typical rhythmic feel.

The bagpipes, locally known as *gaita*, stands as the most iconic instrument in Galician folk music and has well-established tradition of representation in Galician literature and the visual arts. The *pandeireta* (tambourine) and the *bombo* (bass drum) appear frequently in traditional ensembles, providing a rhythmic foundation. Many other instruments have been introduced in the past century, including keyboards, flutes and a series of imported instruments from other traditions such as the Irish bodhrán, tin whistle, harp and even bouzoukis (Colmeiro, 2014).

Figure 1.4: Transcription of *Churrusqueira*, a jota, retrieved from Folkoteca Galega.

In recent decades, the revival and reinvention of Galician folk music have been propelled by cultural associations, music schools, and festivals such as the Festival Internacional do Mundo Celta de Ortigueira, which attract both local and international audiences (Aldred, 2005). Other preservation efforts include traditional collections and documentation methods, like the Cancioneiro Popular Galego compiled by Schubarth (1984). In terms of digital formats, the website Folkoteca Galega [3] provides a wide collection of transcriptions.

In terms of computational approaches to the analysis of Galician folk music, not much can be found. (Carvalho et al., 2021) and (Orouji, 2025) have developed studies on Iberian folk music but not specifically Galician. In terms of purely musicological studies not much beyond Torner and Bal y Gay (1973)'s and Schubarth (1984)'s collections is available.

1.4 Context and Objectives

This research addresses a significant gap in computational music analysis by focusing on expressive timing in folk music traditions, specifically Irish and Galician repertoires. It is developed in the context of EA-DIGIFOLK, a broader project dedicated to promoting the study and dissemination of folk music from European and Ibero-American traditions. It aims to make folk music collections widely available to the public, researchers, and educators, in an effort to facilitate advanced musical and ethnomusicological analysis.

This work aims to develop systematic methods for analysing and comparing musical performances by extracting musically relevant information that captures the nuances and stylistic aspects important to ethnomusicologists.

The primary objective is to develop and apply computational methods for analysing expressive timing patterns across multiple performances, revealing how performers use temporal variation as a means of musical communication. This investigation seeks to understand expressive timing

[3] https://folkotecagalega.gal (last accessed June 2025)

not merely as deviation from a steady tempo, but as a deliberate and informed aspect of musical performance that carries meaning within traditional music contexts.

This investigation operates within the interdisciplinary domains of Music Information Retrieval (MIR) and Music Performance Analysis (MPA), addressing the relative scarcity of computational studies focused on folk music traditions.

1.5 Document Structure

This dissertation is organized into six chapters that progress from theoretical foundations through methodological development to empirical analysis and conclusions.

Chapter 1 establishes the research context by examining the nature of folk music, providing detailed background on Irish and Galician musical traditions, and defining the scope and objectives of the study. Chapter 2 reviews existing literature on tempo estimation techniques and music performance analysis, establishing the theoretical framework for computational analysis of expressive timing. Chapter 3 describes the preparation of the dataset, including data collection methods, curation processes, and organisational structures. Chapter 4 details the annotation procedures and temporal information extraction methods, comparing automatic versus manual beat tracking approaches and explaining the segmentation and tempo estimation processes. Chapter 5 presents the core analytical work, examining expressive timing through comparative analysis at multiple levels and implementing phrase-level analysis using hierarchical clustering techniques. Chapter 6 synthesizes the findings, discusses their implications, and outlines directions for future research.

Chapter 2

Tempo Information Extraction and Performance Analysis

This chapter outlines the most recent and most relevant techniques for the extraction of temporal information from music (Section 2.1, as well as MPA approaches that focus on rhythm and expressive timing (Section **??**.

2.1 Tempo Estimation

Tempo information extraction constitutes a fundamental task within MIR, encompassing both automatic computational approaches and manual annotation methods.

2.1.1 Automatic Approaches

The foundation of automatic tempo estimation lies in onset detection and periodicity analysis. Onset detection is the cornerstone of tempo estimation. It identifies the temporal locations of musical events that correlate with perceptual beats. Multiple approaches have been developed for this purpose, namely, the novelty functions described by Müller (2021) - energy-based, spectral-based, phase-based and complex-domain.

Periodicity analysis determines the underlying temporal regularity. Several techniques can been employed for this. Autocorrelation-based methods analyse the onset detection function for repeating patterns by computing correlation with time-shifted versions of itself (McKinney et al., 2007). Fourier-based analysis transforms the onset detection function into the frequency domain to identify dominant periodicities (Schreiber, 2020). Dynamic programming approaches optimize tempo paths through time by incorporating cost functions that penalize tempo changes, phase shifts, and deviations (Robertson, 2012). This last approached is also described by Ellis (2007), and it is the one employed in the Librosa module for Python, which will be mentioned later on.

This two stage approach is common but not exclusive. In Schreiber et al. (2020)'s comprehensive analysis, the authors examine some other current relevant techniques employed in this field and their applications. Böck et al. (2015) use neural networks in combination with comb filters to determine periodicity. And Schreiber stray completely from the traditional approach and use solely a convolutional neural network (CNN).

2.1.2 Manual Approaches

Manual annotation remains crucial for creating ground truth data and understanding expressive timing variations. Several methodologies have been developed for systematic beat annotation.

Traditional manual annotation involves human annotators tapping along with musical recordings or manually placing beat markers using software tools like Sonic Visualizer. This approach captures human perceptual responses to musical beats but requires careful consideration of annotator consistency and bias. In Robertson (2012)'s study, these kind of annotations are used to determine tempo, implementing an approach that takes into account smaller and often neglected variations that may convey important musical information.

McKinney et al. (2007) employed 40 different listeners to annotate beat positions in 140 musical excerpts, acknowledging the inherent subjectivity in beat perception. This approach provides statistical reliability and captures the variability in human beat perception across different listeners and musical styles.

2.1.3 Challenges

Current practices face several limitations. McKinney et al. (2007) found that algorithm performance varies significantly across musical genres.

Despite significant advances, several challenges remain. Deep learning systems exhibit high sensitivity to training data characteristics, performing poorly on musical content outside their training distribution (Pinto et al., 2021). The implicit bias toward straightforward musical material in annotated datasets contributes to reduced performance on challenging content featuring expressive tempo variation, non-percussive instruments, or changing meters.

Schreiber et al. (2020) emphasized the need for improved evaluation methodologies that align with actual user requirements rather than purely academic metrics. The research community must better understand real-world applications and their specific accuracy requirements to guide future development efforts effectively.

2.2 Music Performance Analysis

MPA represents a multidisciplinary field that examines how performers interpret and transform musical compositions into acoustic realizations. According to Lerch et al. (2020), musical performance constitutes a fundamental aspect of music as a performing art, requiring performers to render a musical blueprint into an acoustic realization that transforms compositional content through intentional parameter variations. The field encompasses the systematic study of performance parameters including timing, dynamics, pitch, and timbre, which collectively contribute to expressive interpretation.

Performers do not merely reproduce notated material but convey musical ideas to the listener by actively engaging in creative interpretation and modifying musical parameters while preserving compositional content. The audio recording becomes the primary analytical focus, as it represents the integral acoustic component present in all musical performances (Lerch et al., 2020). Traditional transcription methods often produce idealised representations that reflect the presumed intention of the performer rather than the actual performance, potentially losing valuable information contained in the audio recordings (Müller et al., 2010).

The fundamental performance parameters can be systematically categorized into four primary domains (Lerch et al., 2020):

- Tempo and timing encompass both overall tempo choices and expressive temporal variations, recognizing that rhythmic content and tempo indications in scores are often interpreted as suggestions rather than strict requirements.
- Dynamics involve loudness variations and accent placement, as score information regarding dynamics is frequently missing or imprecisely defined, requiring performers to make interpretive decisions about phrasing and tension.
- Pitch-based parameters include not only fundamental frequency accuracy but also expressive techniques such as vibrato and intonation choices.
- Timbre represents the least specifically encoded parameter category, often only implicitly indicated in scores through instrumentation designation.

Within the tempo and timing category, some studies have utilised this feature to investigate interpretive choices, tempo variations, and expressive timing in musical performances.

Dixon (2001)'s landmark study presents a computational framework for quantifying expressive tempo and beat structure directly from recorded music. This work revealed how performers shape tempo continuously and how it is perceived by listeners. Solely through low-level features, his algorithm managed to combine classic MIR tempo extraction techniques with human perception and achieve robust results.

In a study by Fabian and Schubert (2008), they examine how three key performance variables - dotting, articulation, and tempo - contribute to listeners' perceptions of musical character. Their results provide some insights into the key distinction between the performer's and the listener's perspective of timing. A discrepancy between performance and perception of dotting was found. Actual measured dotting ratios did not reliably predict listeners' judgments of rhythmic "sharpness" or "laziness," indicating that perceptual cues to dottedness involve more than literal duration ratios. To achieve this, a combination of manual annotations and computational methods were used.

Repp (1991)'s work on expressive timing in piano performance demonstrates how both global and local tempo variations serve to articulate musical structure and expression. He analysed multiple classical piano performances through manual beat tracking and among his many findings was a consistent slowing pattern at certain points in the score. These tempo modulations were remarkably uniform across artists, suggesting some shared expressive conventions and reflecting performers' employment of tempo modulation in systematic ways to highlight hierarchical structures, like phrases sections. Among other measurements and perceptual studies, these results provide a rigorous, quantitatively grounded account of how expressive tempo functions as a primary means of musical communication in Western art music.

Performance analysis traditionally relies on reference-based comparisons, typically utilizing either multiple performances of the same piece, variations by the same performer, or deviations from a mechanically quantized performance derived from symbolic notation (Lerch et al., 2020). This reference problem creates a bias toward classical music genres, where abstract score representations facilitate quantised performance models and provide abundant comparative material through numerous interpretations over time.

2.3 Computational Ethnomusicology

The term "computational ethnomusicology" (CE) was formally introduced to describe the usage of computer tools with the potential to assist in ethnomusicological research. This field emerged as researchers recognised the need to consider the large diversity of music cultures from around the world, expanding beyond the traditional MIR focus on popular and classical Western music (Tzanetakis, 2014).

The computational approach enables large-scale analysis that would be impossible through manual methods alone. While traditional forms of musicological analysis provide valuable expert knowledge, manual analysis can be very time-consuming, thus limiting the potential for larger-scale evaluations. The application of MIR techniques provides a plausible alternative for analysis of large collections (Panteli et al., 2018).

A fundamental methodological questions in computational folk music analysis concerns the relative effectiveness of global versus local features. Van Kranenburg et al. (2013) found that local features outperform global features in classification tasks for folk song melodies, indicating that local features carry more information about what characterises a melody. Other challenges also include complex rhythms, addressed in the study by Holzapfel and Stylianou (2009). Or intonation, since many recordings feature non-professional singers, significant deviations from expected pitches (Müller et al., 2010). These limitations require specialised approaches to handle the inherent variability and inconsistencies in the source material.

Despite significant advances, challenges remain in developing methods that are both scalable and generalizable across different folk music traditions. Panteli et al. (2018) note that while several research projects have focused on developing MIR tools for specific folk music cultures, tools for comparison of large corpora are yet to be explored.

Chapter 3

Dataset preparation

In this chapter we will go over all the information regarding the Irish and Galician performances and the dataset as a whole: how the recordings were collected; the treatment of the audio files; how they were organised and labelled with metadata; plus all the challenges that were presented in this stage.

3.1 Irish Performances

The Irish performances used for this analysis come from a collection organised in previous work done in the context of the EA-DIGIFOLK project. The original collection consists of 353 songs, each with multiple performances. From these, I ended up narrowing down to 19 songs, with varying numbers of performances per song (4 to 14), adding up to 154 performances. The styles represented were four in total, and include *reels*, *jigs*, *hornpipes*, and *polkas*, the first two being the most represented. Is it also composed both by ensemble and solo performances the latter with instruments including the bagpipes, flute and whistle, among others.

This big reduction was made, as we will see, because all the beat tracking and phrase segmentation had to be annotated manually, and due to the time consumed by these tasks, I had to adapt to the time I had.

For the original dataset, the performances were collected online, from publicly available video platforms, such as YouTube. A source like this can provide a large amount of data, with a great variety of instruments played in solo and in multiple ensemble configurations. One of the consequences, however, is that the collection displays a wide range of contexts and recording qualities.

For our analysis, the inconsistent audio quality won't be such a big issue. But some things had to be filtered out. The most troublesome inconsistency was the variable length of some performances compared to others. As is common in Irish traditional music, repetitions occur a lot. As such, very often, one performance would be three or four times longer than another. Predicting that further down the process it would be useful to align the performances and compare things side to side, it was important that at that stage we made sure to be looking at the same musical content

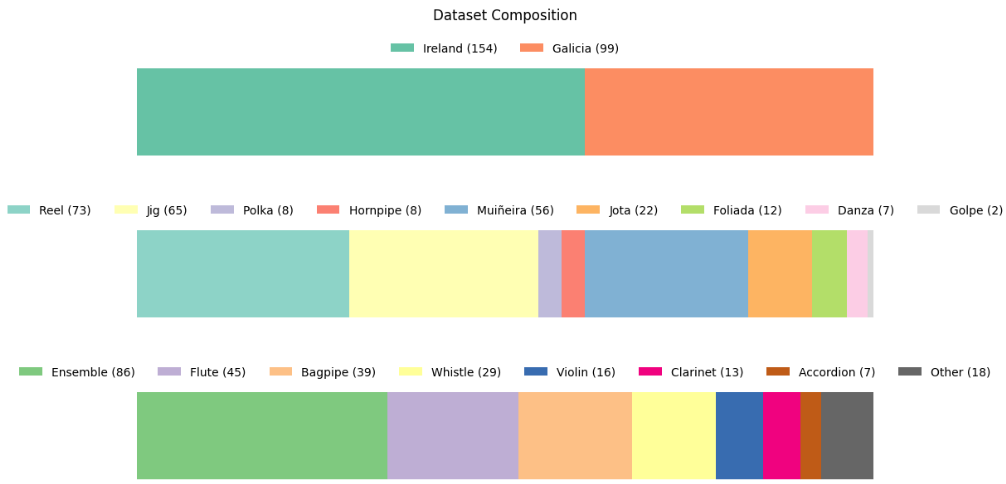


Figure 3.1: Dataset composition, colour coded by culture (top), style (middle) and instrumentation (bottom)

without constantly verifying the audio files. For this, each file was cropped accordingly, whenever possible. As a reference, the shortest performance of each song was used.

3.2 Galician Performances

The Galician performances used were provided by Hilda Romero Velo, a PhD candidate from the Universidad da Coruña, who had them recorded for her own research. This dataset included 47 songs, with multiple performances each, for a total of 145 performances. These were narrowed down to 30 songs, with 2 to 7 performances per song, for a total of 99 performances. In total, five styles were included and these were *danzas*, *foliadas*, *golpes*, *jotas*, and *muiñeiras*. *Muiñeira* is the most represented by far, but *jotas* are also present in a fair amount. All of them are peformed solo, with a variety of instruments, mainly bagpipes, flute, whistle and clarinet.

These required little to no curation. Since they were all recorded for the same purpose, all in the same context, with very consistent audio quality. As for the issue of the performances' length and alignment we mentioned above, a few odd cases were also present, but not nearly as many.

In Figure 3.1 we can see a visual representation of the complete dataset's compositions, based on the metadata attributes we will be exploring ahead.

Regarding the treatment of the audio files, for both portions of the dataset, everything was converted to MP3, 44.1 kHz, 64 kbps. It's a fairly heavy compression, but since the role of the audio will mostly serve as reference, priority was given to storage effectiveness.

3.3 Organisation and Metadata

As explained, our dataset comprises performances of multiple songs, in a number of different styles, with multiple instrumentation settings. This all had to be organised into an accessible

culture	title	id	style	file_name	instrumentation	instrument
ireland	Colonel Fraser	ir001	reel	ir001_v01.mp3	solo	bagpipe
galicia	Fogueteira	ga012	jota	ga012_v02.mp3	solo	whistle

Table 3.1: Example table to demonstrate how metadata is organised in the spreadsheet.

format, like a spreadsheet, displaying all the performances' relevant information. For each, we must know its culture of origin, song title, style, an assigned ID for easier organisation, file name to facilitate access, instrumentation - solo or ensemble - and instrument(s) used.

This spreadsheet was stored in a Google Drive folder, along with the audio files, annotations, and code. It is accessed through a Google Colab notebook, through which we can read and write information. This not only facilitates the workflow but also makes this dissertation's woks easily accessible and shareable.

The culture column simply identifies the culture of origin for each song, that is, Ireland or Galicia. This creates these two distinct groups that will be later used for analysis within themselves or for comparison between them.

The song ID column displays the ID assigned to each song. These consist of a prefix - *ir* for Irish and *ga* for Galician songs - and a three-digit number, starting at *001* (e.g., *ir001*, *ga002*). This makes it easier to name all the audio files in an organised and intuitive way, in order to facilitate access and indexing. The IDs were not assigned according to any particular criteria, but simply according to the order in which the files were added to the dataset.

The title column keeps the original names of the songs, so that we're not only dealing with isolated and arbitrary IDs. If we ever need to find further information on a particular song, we still need some original reference.

The style column identifies the style of the song. These serve as a reference for comparison of data later on, where we will analyse songs within a certain style or compare different styles.

The file name column indicates explicitly the audio file we need to access for this performance. They consist of the song's ID, the number of the version and the extension, formatted like so: *id_version.mp3* (e.g. *ir001_v01.mp3*).

The instrumentation column indicates the instrumentation setting of the performance in question. The only distinction made here is between *solo* and *ensemble*. For most performances, this was a straightforward distinction. However, there were a few specific cases where it was a bit more ambiguous. For example, a single musician playing along with a backing track, or two musicians, in which one of them is playing alone most of the time, and the second one does some accents here and there. In the end, it was settled that only a single musician playing alone, with no other musical cues, would be considered a *solo* performance. All the rest would be classified as *ensemble*. This classification also serves for grouping and comparison later on.

The instrument column indicates the instrument, or list of instruments, used in the performance. This is also a tool for comparison, especially for instruments played in a *solo* setting. To clarify a couple of particular cases, in which we will find the performances classified as *ensemble* but only with a single instrument listed, this means it is played on two or more of the same instrument.

Flawless accuracy cannot be guaranteed here, when trying to identify all the instrument in *ensemble* performances. Since the recordings were obtained from YouTube, it was sometimes possible to verify this through video footage. But for the videos that didn't show the musicians, it had to be done by ear.

Afterwards, more columns were added for the beat informations and segmentation, plus any other data that was calculated.

Chapter 4

Annotations and Extraction of Temporal Information

In this chapter we will go over the approaches taken in order to extract tempo information from the recordings. Section 4.1 goes over the first approach attempted, with automatic methods. Sections 4.2 and 4.3 explain the process of manual annotation of beats and phrase segmentation, along with its challenges. Lastly, Section 4.4 explains how the manual annotations were used to extract temporal data and how that data was prepared for analysis.

4.1 Automatic beat tracking

The first approach tried was automatically tracking the beats of each performance using the beat tracking function in the Librosa module for Python. Right away, the results weren't very promising. Unless the performance had some kind of clear percussive element strongly keeping the time, the algorithm often failed to detect or maintain a steady tempo throughout the performance. There were several attempts at tweaking the algorithm's parameters, and testing different configurations to see if the output could be improved, but that didn't lead anywhere particularly useful either.

Even in cases where the beat tracking seemed acceptable at first glance, a closer inspection revealed some issues. One of the main problems was that the algorithm quantises the output. In other words, each beat gets snapped to the nearest match in a fixed grid and not to the actual note onsets. As a result, even when the algorithm "worked well", what it produced was only an approximation, and not a realistic representation of the timing in the recording. To see this, we can take a look at a preliminary graph, drawn early in the process of experimentation, in Figure 4.1, where the tempo (in BPM) obtained through the automatic beat information is represented over time. Each red point represents a beat event, and we can see some regions where they form a straight horizontal line, meaning a 100% consistent tempo. That might look like a good result, but it's not a natural outcome — it reflects a mechanical, machine-like performance that's clearly not what's happening in the actual recording.

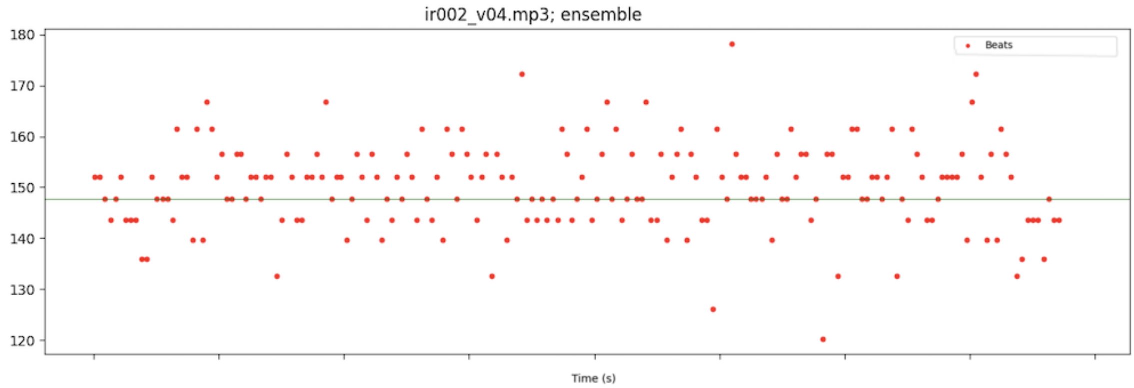


Figure 4.1: BPM over time calculated using automatic beat information.

With this, the plan to use automatic annotations was abandoned. Given these issues, it was evaluated that manually tracking the beats from scratch would probably be less time-consuming than trying to clean up the automatic output.

4.2 Manual Beat Tracking

The manual beat tracking was done using Sonic Visualizer (Figure 4.2). This tool allows us to import the audio, listen to it, and tap along to lay down the beats. Since the process is done by ear, the initial tapping is not always perfect, so we usually go over the same performance again and adjust the timing as needed.

Depending on the instrument that is being played and its attack envelopes, it can be easier or harder to nail down the exact timing of each beat. And while the waveform can be a reference sometimes, the final verdict should always judged by ear.

Some performances were very straightforward, but others, particularly certain Galician songs, presented more of a challenge. In many of these cases, the initial feeling for tapping along would seem correct but would suddenly switch to the off-beat at the turn of a phrase. To understand

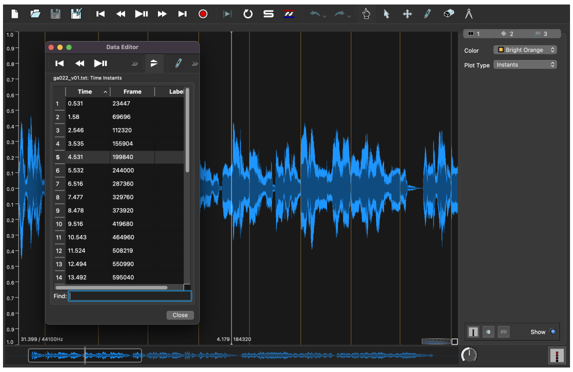


Figure 4.2: The Sonic Visualizer interface. The annotated beats correspond to the time values in the Data Editor window and represented over the waveform by the vertical lines.

Figure 4.3: How the beat initially felt (red
versus how it actually is accentuated (green) in the Galician song *Foliada de San Fins do Castro*

what was happening, the transcriptions of the songs available at Folkoteca Galega were consulted. These helped identify that, in some cases, that initial feeling of the beat was differently from how it was transcribed, and probably from how the musicians were feeling it. Figure 4.3

This is an important point to address. It concerns the issue of the performer's and the listener's perspective. While there's no "wrong" way to feel a rhythm, it was decided that prioritising how the musicians themselves were likely perceiving the tempo was more appropriate for the analysis, since it would probably better reflect their expressive timing patterns.

A similar issue arose when in some transcriptions the indicated tempo was *prestissimo*, which typically corresponds to around 200 BPM. However, in performance, the song clearly did not feel that fast. To clarify this, other interpretations of the same song were examined online, revealing ensemble versions in which a percussive element (e.g. a bass drum) would accentuate only the first beat of each measure in the transcription, suggesting that, even at theoretically high BPMs, the musicians were feeling the tempo at a slower pace (shown in Figure 4.4). This was also confirmed by Hilda Romero Velo, who provided the recordings, so the annotations were done accordingly.

The Irish songs, on the other hand, were generally easier to work with. Reels, in particular, because they follow a regular 4/4 time signature, make the beat tracking and phrase segmentation more predictable. Jigs posed slightly more of a challenge because of the 6/8 timing, but no particular difficulties worth noting. Whenever needed, transcriptions were also used as reference, in the platform The Session.

4.3 Segmentation

For the segmentation into musical phrases, it is important to note that the concept of musical phrase is not rigidly defined, more so in folk music.

However, in Irish tunes, there is a recurring structure that can be used as a general guideline. Using this as a rule of thumb, phrase segmentation in Irish performances was relatively straightforward.

Figure 4.4: How 200 BPM would be felt (blue) versus how it actually is accentuated in the Galician song *Foliada de San Fins do Castro*

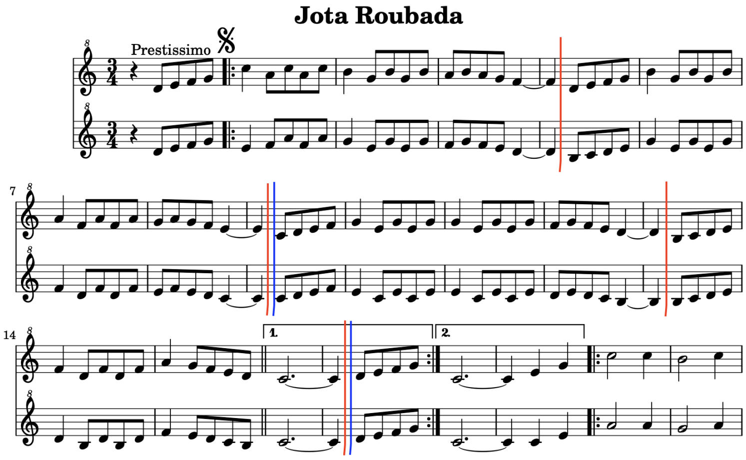


Figure 4.5: Two possible ways to divide the first section of *Jota Roubada*, in two (blue) or four (red) phrases.

The Galician songs, again, proved to be more challenging, often for the same reasons that made beat tracking harder. While the transcriptions show repeated sections and a general form, phrase boundaries were not always obvious. The lengths of the phrases also varied more than in the Irish repertoire.

In Figure 4.5 we can see an example of a song in which two different segmentations are acceptable.

Once the beat tracking and phrase segmentation were complete in Sonic Visualizer, the annotations were exported as plain text files, each line containing a single timestamp, with a label for the phrase end-points. These can then be loaded into the code as lists and updated in the spreadsheet.

Each phrase was identified by its end-point - the moment the next phrase begins, essentially. We are assuming the start of the first phrase is at the time stamp 0, which is a safe assumption, given that the audio files were previously cropped with that in mind. To avoid redundancy, this time stamp was not included in the annotations. This is later handled in the code.

As shown in Figure 4.6, each phrase is defined as the set of beats that occur between two segmentation points. For example, the first phrase includes all beats from time 0 up to, but not including, the first segmentation time stamp (pink area in the figure). The second phrase includes all beats from the first segmentation time stamp up to the second (green area), and so on.

We are not assuming the last segmentation point corresponds to the end of the audio file. For two reasons: first, some performances end on the first beat of a hypothetical next phrase, meaning there may be extra beats at the end that should not be included in the final phrase. Second, by explicitly marking the end of the last phrase with a segmentation point, we avoid having to load the audio file just to check its duration. This keeps the code cleaner and more efficient.

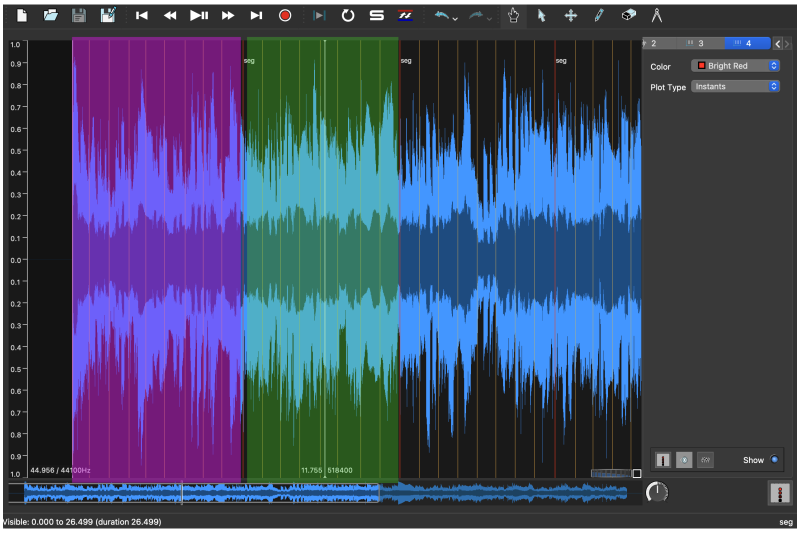


Figure 4.6: Phrase segmentation in Sonic Visualizer. The pink area encompasses the first outlined phrase, and the green area the second.

4.4 Tempo Estimation and Variation

4.4.1 Inter-beat intervals and BPM over time

Once the beat tracking was done, the next step was to calculate the inter-beat intervals. This just means subtracting each beat from the next, to get the time elapsed between them. Naturally, this gives us one fewer value than the total number of beats, since the first beat does not have a previous one to be compared to. To calculate the BPM we can divide 60 (seconds in a minute) by each inter-beat interval (δt).

$$BPM = \frac{60}{\delta t} \tag{4.1}$$

This results in a list of local BPMs, which we can then plot to visualize how the tempo evolves over time, as seen in the graphic in Figure 4.7. The red dots in the graph represent beat events at which we calculated the local BPM, and by connecting them we get a rough tempo curve for the whole performance. This is also known as a piecewise linear function. Additionally, we can also use the segmentation information to visualise the phrase boundaries. This is the simplest way for us to start looking at expressive tempo, both at the performance and phrase levels. It is not the most practical when looking at a large number number of performances, but perhaps the most visually intuitive. We will look into some examples in the next chapter.

4.4.2 Estimating Global BPM

One way to further work with these is to explore deviation, to try to understand how much a performance strays from some intended or reference tempo. But before we can calculate deviation, we must establish this reference BPM as the term of comparison.

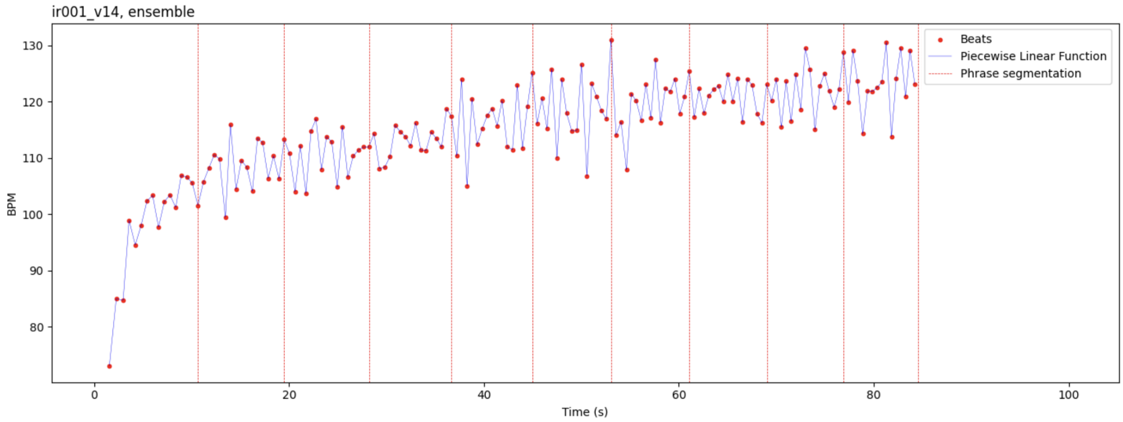


Figure 4.7: BPM over time for for the performance *ir001_v14* with phrase segmentation

There are a few ways to calculate a global BPM for a performance. A super simple one is taking the mean of the local BPM list. The issue with this is that it is very sensitive to outliers, like very accentuated peaks or dips, so it might not reflect the intended tempo very well if there are lots of fluctuations Another option is the median of the local BPM list, which is useful because it ignores these extreme values and gives a good sense of the "centre" if the performance is relatively stable.

There is also a more mechanical method, through the following formula, found in Lerch (2023).

$$globalBPM = \frac{N * 60}{l} \tag{4.2}$$

Where N is the total number of beats and l is the total length, in seconds, from the first beat to the last. This gives you a kind of "overall average BPM" across the whole piece. Conceptually it feels similar to the mean, but I'm still not entirely sure how it behaves in the presence of outliers

A slightly more advanced method is through a kernel density estimation (KDE) on the local BPM values, followed by finding the mode of that distribution.

This KDE mode estimation feels a bit more musically meaningful, because instead of just averaging, it tries to find the most "frequent" tempo.

All of these approaches look at the entire performance, but perhaps, from a musical point of view, that might not make the most sense. From a performer's perspective it's usually the first count-in that acts as a reference, even if the rest of the performance moves expressively around or away from it. With that in mind, we can estimate the tempo again, but this time using only the first few beats. How many beats exactly is more or less arbitrary. 10 seemed like an appropriate range. So, the same KDE mode estimation method was used, only on these first 10 BPM values, and this would be the "global" BPM value to be used as reference.

	Low Mean Deviation	High Mean Deviation
Low Standard Deviation	Steady tempo	Consistent shift/offset
High Standard Deviation	Symmetrical variation	Lots of variation

Table 4.1: How to cross reference mean deviation values with standard deviation values, to get a sense of how BPM varies throughout a performance

4.4.3 Deviation calculation: mean and standard deviation

We can now calculate each performance's deviation from its reference BPM. This is relatively straight forward, we simply subtract the reference BPM from the each value in the list of local BPMs, effectively turning it into a list of deviation values.

$$[deviation] = [localBPMs] - referenceBPM \tag{4.3}$$

With this list of deviation values, we can take their average, to get the mean deviation per performance. This value, even if very rough, can already tell us something about a performance's tempo. We should be careful though. A value close to 0 may lead us to think the performance has a rather stable tempo. However, it may be the case that it actually has a lot of symmetrical variation.

To get a more solid idea of how the tempo varies, we can combine the mean deviation value with the standard deviation for a more grounded interpretation, as shown in 4.1

We do not need to be looking at pairs of values one by one though. We can draw box plots that helps us visualise this in a more intuitive way. In figure 4.8 we can see three box plots. Each one draw with the list of tempo deviation values of a performance of the song *ga011*. The mean deviation is represented by the green dashed line and the median by the orange one. The boxes confine the middle 50% of the data, also know as the Interquartile Range (IQR). The vertical lines coming out of them, also called the whiskers, stretch to the minimum and maximum values, excluding outliers, which are represented by the outside points. The width of the box plus the whiskers, while not explicitly showing a the value, can give us an idea of the standard deviation.

In *ga006_v01*, we can see a relatively high mean deviation, along with a fairly wide box, mostly containing positive values.

We can infer then, that this performance has a tendency to speed up. The second performance, *ga006_v02*, has a mean value close to zero and a very narrow set of values, leading us to believe it has a fairly steady tempo.

Finally, *ga006_v03*, also has a mean value close to zero, but shows a wider distributions, along with a lot of outliers. Plenty of variation, but no particular tendency.

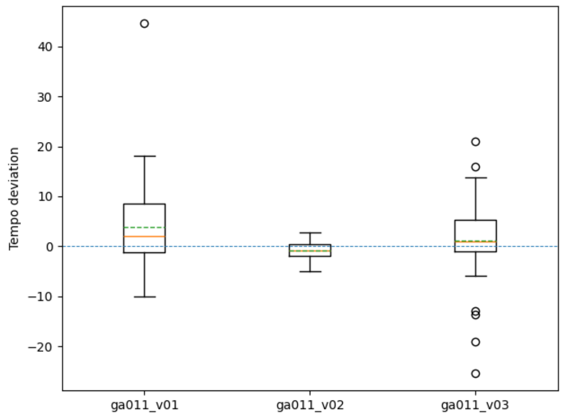


Figure 4.8: Box plots drawn from the tempo deviation in performances of the song *ga011*

Chapter 5

Analysis of Expressive Timing

In this chapter we will take the data obtained and evaluate it through different representations. We begin, in Section 5.1, by looking at BPM over time graphs of a few performances and doing a brief visual interpretation. Then we move one to comparative analysis. Starting by establishing meaningful groups of comparison, in Section 5.2. Then, using these groups, we will look at performance-level BPM deviation (Section 5.3 and phrase-level tempo curves, though hierarchical clustering 5.4.

5.1 Visualising BPM over Time

Let us briefly go over a couple of graphs that display how the BPM changes over the duration of the performance. Analysing these visually is not a scalable method. It would be very time consuming to look at graphs of a large number of performances and quite hard to draw conclusions. Yet, it can still be an interesting and intuitive way to look into expressive tempo. Figure 5.1 shows four performances of the song *ir002*, which is a jig. This is the same type of graph from Figure 4.7, with a few additions. We now have a horizontal line indicating the reference BPM we calculated for each performance, and a teal-coloured line showing an exponential moving average (EMA). The latter helps smooth the piecewise linear function, which is often quite jagged, and gives us a more interpretable sense of how the tempo is evolving. And to enhance the phrase structure information, connection lines were draw between the graphs, delineating the same phrases throughout the graphs.

I chose these four performances because they are a prime example of how different performers can stretch time. *ir002_v08* and *ir002_v12* performed at a slower tempo, thus the graphs dragging out a bit more. But more interestingly perhaps, is the way *ir002_v08* starts very slow and eases into a tempo to then stabilise. This is a possible tendency we can look out for, and it actually comes up later in the chapter when we look BPM curves at the phrase-level.

Then, in Figure 5.2 we have four performances of a foliada, *ga008*. In these, there is an interesting pattern. In all the performances, the BPM seems to dip at a lot phrase transitions. This could indicate a tendency to prolong the last note or pause of each phrase a little longer. However, without more detailed information, it is hard to formulate an hypothesis. It could be a particularity of the song itself, or a quirk of the performers - we do know the Galician recordings were done by a limited number of musicians, so this is possible. But we cannot pinpoint it with any certainty. This pattern does not emerge later in the phrase-level. Nevertheless, it is an interesting detail worth exploring, and we would not have seen had we not looked into this form of representation. These graphs are available for all the performances in C

5.2 Groups of Comparison

Before we can interpret any results, we need to first define how we're going to look at the data. Using the performances' metadata, we can establish musically meaningful groups of comparison. While we could randomly pick two performances and compare them, such a comparison would lack purpose and provide little insight into the corpus as a whole. By organizing our analysis around structured comparisons, we can also begin to set expectations — whether based on prior research, related studies, or just musical intuition.

The most obvious first step is to compare the two cultures present in the dataset. From there, we can look within each culture and analyse distinctions across styles, instruments, and multiple renditions of the same tune. In the Irish dataset specifically, we can also compare solo versus ensemble performances.

When comparing cultures, we might observe broad timing tendencies or stylistic differences. However, such a high-level comparison should be treated with caution due to inherent differences in how the datasets were compiled and recorded. These structural differences may influence the results independently of musical practice.

Focusing on the Irish dataset, one of the key internal comparisons is between solo and ensemble performances. As detailed earlier, a solo performance is defined as one without any form of accompaniment — whether other musicians or a backing track. Based on common musical reasoning, we might expect ensemble performances to exhibit more consistent tempo, as the presence of multiple musicians usually provides rhythmic reinforcement. Solo performers, on the other hand, may exhibit more tempo fluctuation, since they rely solely on their own internal sense of time.

We can also compare stylistic differences — particularly between Jigs and Reels, which are the two most represented categories in the dataset. As discussed in the beat tracking section, these styles differ in both time signature and metrical feel. A comparison may reveal how these structural differences manifest in expressive timing and tempo variation. Instrument-based comparison is also relevant, but here we'll focus only on solo performances to avoid confounding variables. While ensemble performances often have a leading instrument, identifying it consistently across the dataset is not always straightforward. Restricting our analysis to isolated instruments allows us to explore how factors such as playing technique, physical engagement with the instrument, and timbre characteristics might influence timing expression. Finally, we can compare multiple performances of the same song. This allows us to see whether certain pieces tend to elicit particular

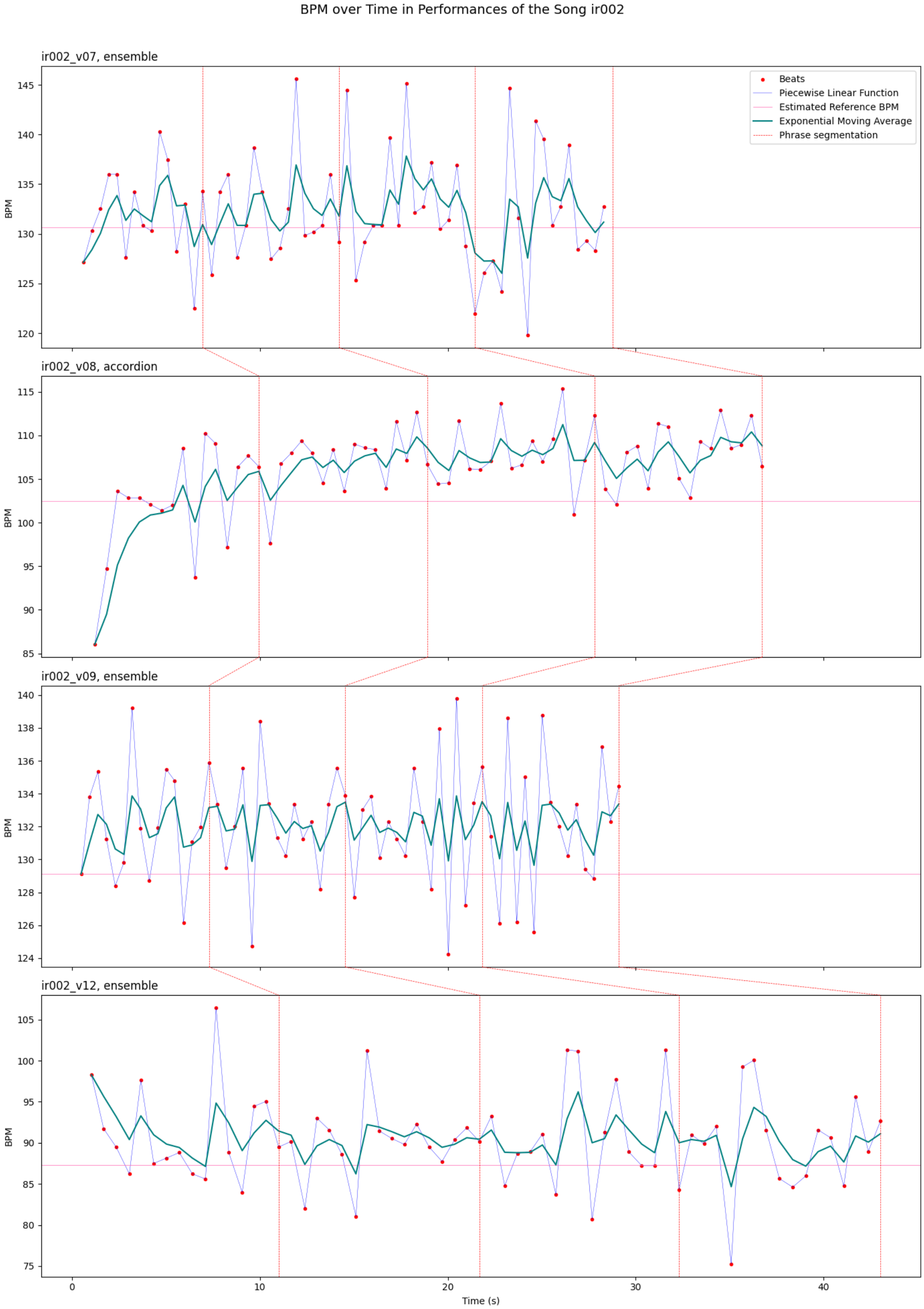


Figure 5.1: BPM over time in four performances of the song *ir002*

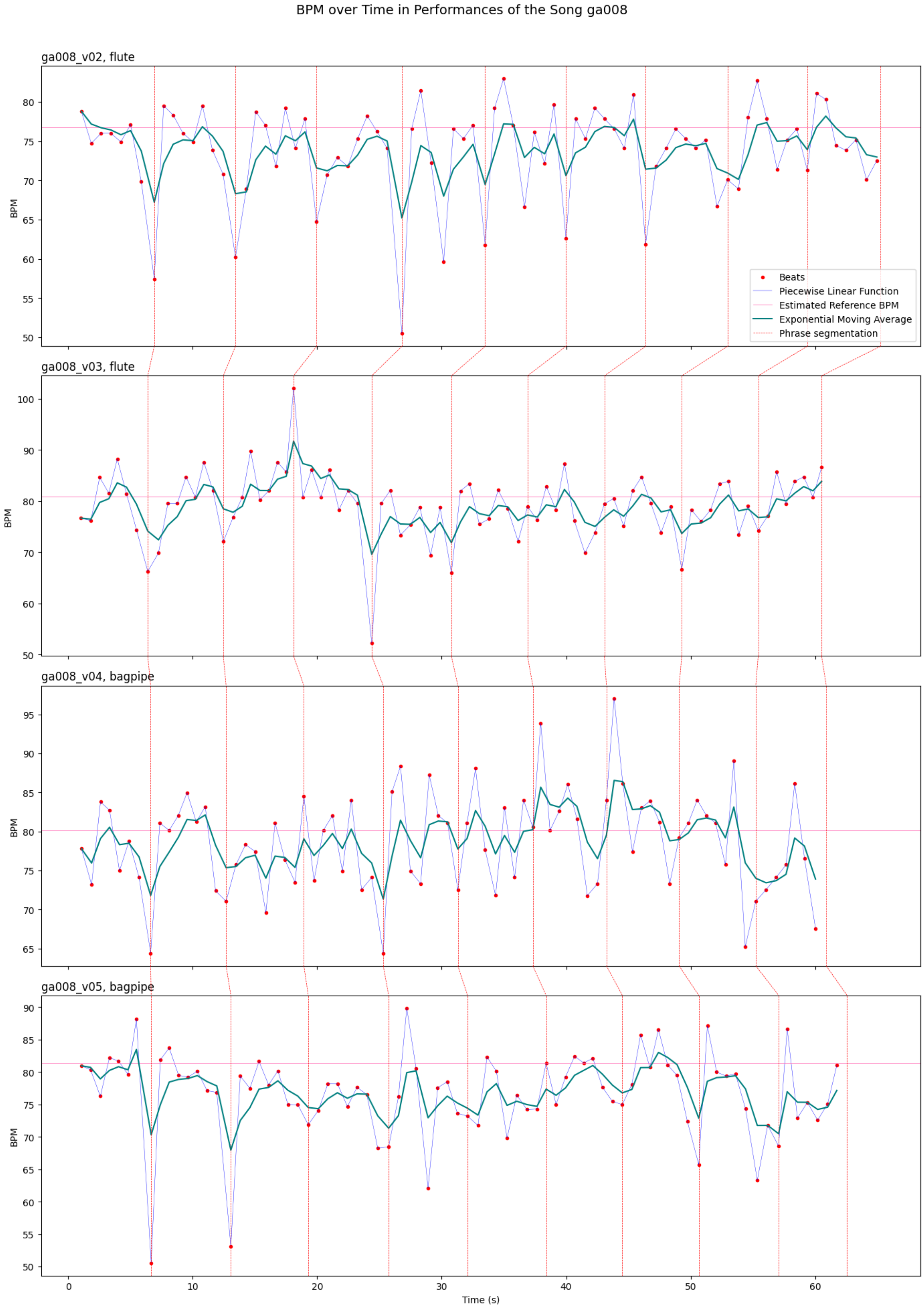


Figure 5.2: BPM over time in four performances of the song *ga008*

timing behaviours, or whether interpretations vary widely. In either case, it offers insight into how expressive timing may be shaped by the musical content itself.

As for the Galician dataset, we can conduct many of the same comparisons — with the exception of solo versus ensemble, since all performances are solo. Within styles, the most represented are Jotas and Muiñeiras, which differ in both time signature and rhythmic feel. While both tend to emphasize the start of each measure, their metric structures create distinct temporal sensations. Thanks to the fact that all performances are solo, we also have more consistent conditions for instrument-based comparison. This gives us a stronger foundation to examine how different instruments might shape expressive timing. Lastly, while most songs only have one recorded version, there are a few that appear multiple times, which opens up the possibility for intra-song performance comparisons in the Galician tradition as well.

Throughout this chapter, we will explore a range of analytical strategies and these comparison groups will serve as a consistent framework for interpreting the results.

5.3 BPM Deviation Analysis at Performance-level

Based on the calculated mean tempo deviation values for each performance, box plots were drawn, with the established groups of comparison in mind. Looking at these, we can try to find any global tendencies on the performance level.

Starting with both cultures, side by side, in Figure 5.3. We can see a clear difference between them. Irish songs seem to have a lot more tendency to speed up while the Galician ones tend to be more contained. This is hardly enough to draw hard conclusions though. Do keep in mind, we are looking at mean deviation values per performance, which is a global and rather low resolution metric. This may also simply be a consequence of our limited analysed corpus. Interesting nonetheless. To avoid repetition, let us note right away that these points are valid for any results we observe.

Next up, we can take a look within the Irish part of the dataset. Starting with the solo and ensemble performances, in Figure 5.4. The difference between these is not as accentuated as expected. Ensemble performances would be good candidates to have considerably less tempo variation, but they seem to have a tendency to accelerate almost as much as solo performances.

Looking at, different styles now, focusing on jigs and reels. In Figure 5.5, we can see a few more outliers in the reels, but not much difference besides that. The other styles show very little variation, but there are barely represented, so we cannot make any assumptions.

Lastly we have instruments in Irish performances in Figure 5.6. The flute, whistle and violin were the most represented, despite the absolute numbers not being very significant. We can see a little more tendency for acceleration in the flute, but with these, we really draw anything concrete.

Moving on to the Galician part of the dataset. All performances are solo, so we do not have that distinction here. We can look at different styles though, in Figure 5.7. They all have a fairly low amount of variation. Makes sense, from what we observed when comparing both culture above.

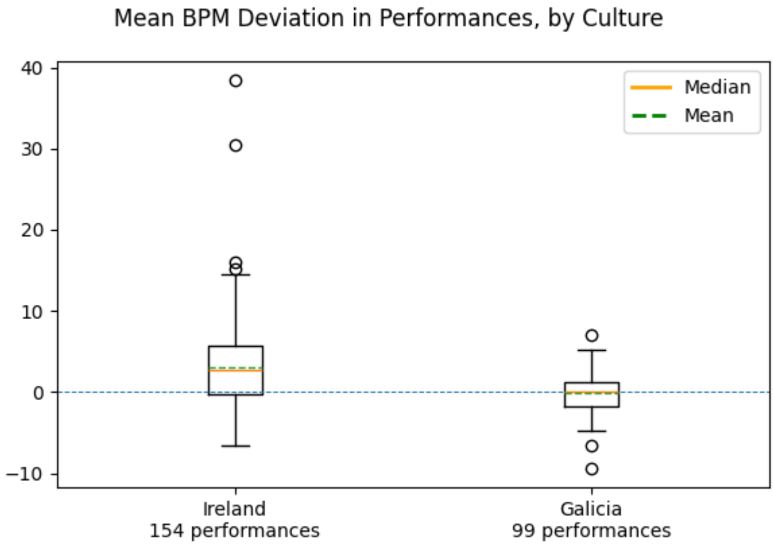


Figure 5.3: Mean tempo deviation per performance in both cultures

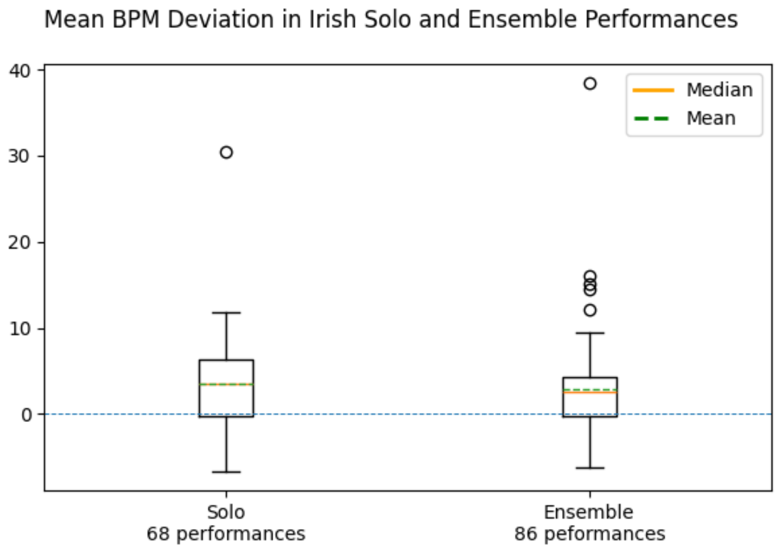


Figure 5.4: Mean tempo deviation in Irish solo and ensemble performances.

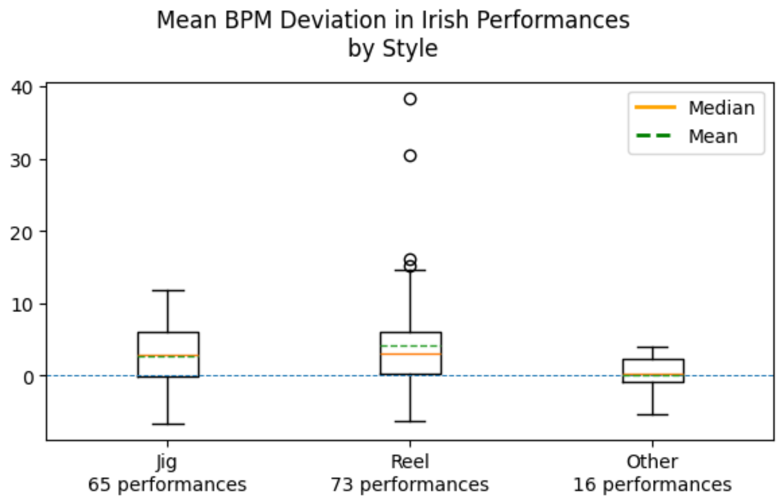


Figure 5.5: Mean tempo deviation in Irish performances of different styles.

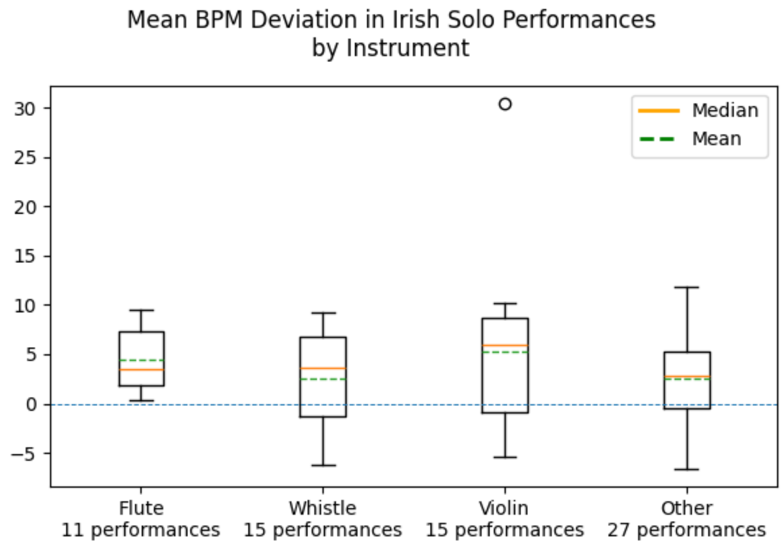


Figure 5.6: Mean tempo deviation in Irish performances with different instruments.

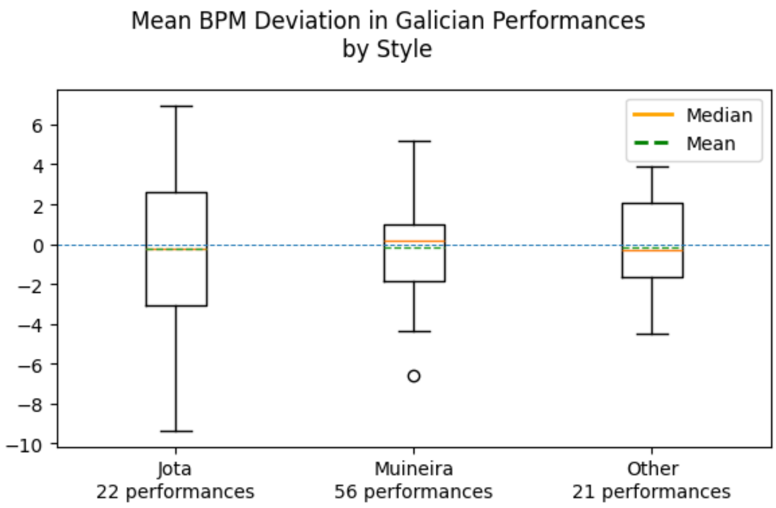


Figure 5.7: Mean tempo deviation in Galician performances of different styles.

Finally, different instruments in Galician performances, in Figure 5.8. The flute and the bagpipes (or gaita) had the largest number of performances, a number we can already start considering as significant. And interestingly they show different tendencies. The flute seems relatively steady, but the bagpipes seem to have a tendency for retardation, which we do not see in any other instrument, Irish or Galician

5.4 Phrase-level Analysis

5.4.1 Phrase Preparation

Before we could apply the similarity measures, the BPM data from each phrase had to go through a few preprocessing steps to make them comparable. First, we subtracted the first BPM value of each phrase from all subsequent values. This aligns all phrases to start at zero, letting us focus on relative tempo changes within the phrase rather than absolute BPM values. Next, we smoothed the curves using an exponential moving average. This technique gives more weight to recent points, which helps reduce small fluctuations and noise while preserving the overall shape and direction of the phrase. Finally, the BPM curves were resample using linear interpolation so that all phrases would have the same number of beats — matching the longest phrase in the dataset. While interpolation introduces minor distortion, it preserves the expressive shape of the curve, which is our main focus.

5.4.2 Hierarchical Clustering

We now turn to hierarchical clustering to further investigate the relationships between tempo curves at the phrase-level. This approach allows us to not only compare phrases, but also group them into a tree-like structure that reveals how closely they relate to each other.

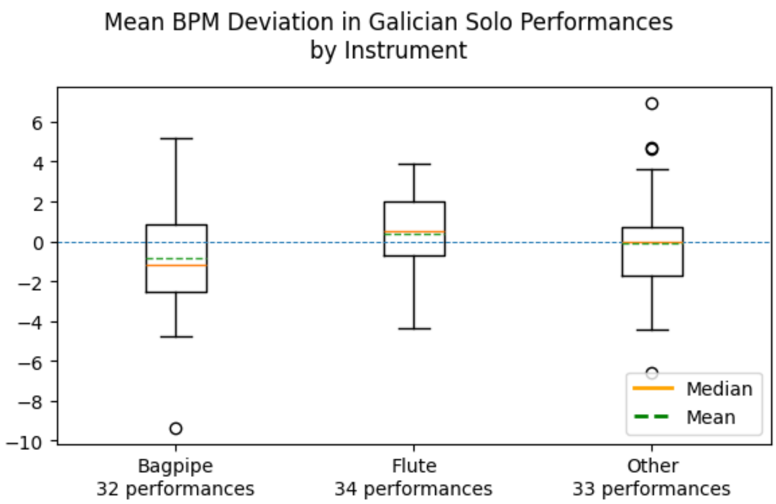


Figure 5.8: Mean tempo deviation in Galician performances with different instruments.

To do this, we build a linkage matrix by directly feeding our list of resampled and normalised phrases into the clustering function. The linkage matrix encodes the order and distance at which phrases or clusters of phrases are merged. Even though we did not use a precomputed distance matrix beforehand, the function internally measures dissimilarity based on the shape of the time series, using a specified linkage strategy - Ward, in our case. This way, phrases with similar expressive timing patterns are grouped earlier, and those that are more dissimilar are only merged later, higher up the hierarchy.

The idea is simple: phrases that have a more similar curve will be grouped earlier (with lower distances), while phrases that are more different will only be merged later, further up the tree. Once the linkage matrix is built, we can generate a dendrogram (Figure 5.9), a visual representation of the hierarchical structure. The dendrogram shows each phrase as a leaf, and branches connect phrases or groups of phrases according to their distance, giving us a clear picture of how phrases cluster together based on the shape of their BPM curves. The threshold distance, shown by the red dashed line can be adjusted until we find the groups that make the most sense to what we are analysing. Here, an optimal distance seemed to be the one that created the four distinct curve shapes we can see in Figure 5.10. In it, each of the graphs show all the curves of that group, superimposed, in grey, and the mean curve of the group, in red.

This method adds yet another layer of insight. Dendrograms give us a hierarchical view of similarity, where we can observe and potentially identify meaningful clusters. For example, we might see whether all phrases from a particular instrument, song, or style tend to fall under the same branch, or whether expressive timing patterns cut across these categories in unexpected ways. Another advantage is that this method does not require us to define the number of clusters in advance. We can "cut" the dendrogram at different distances by adjusting its threshold and explore how the structure changes.

Altogether, this technique gives us a more structured and interpretable overview of the data

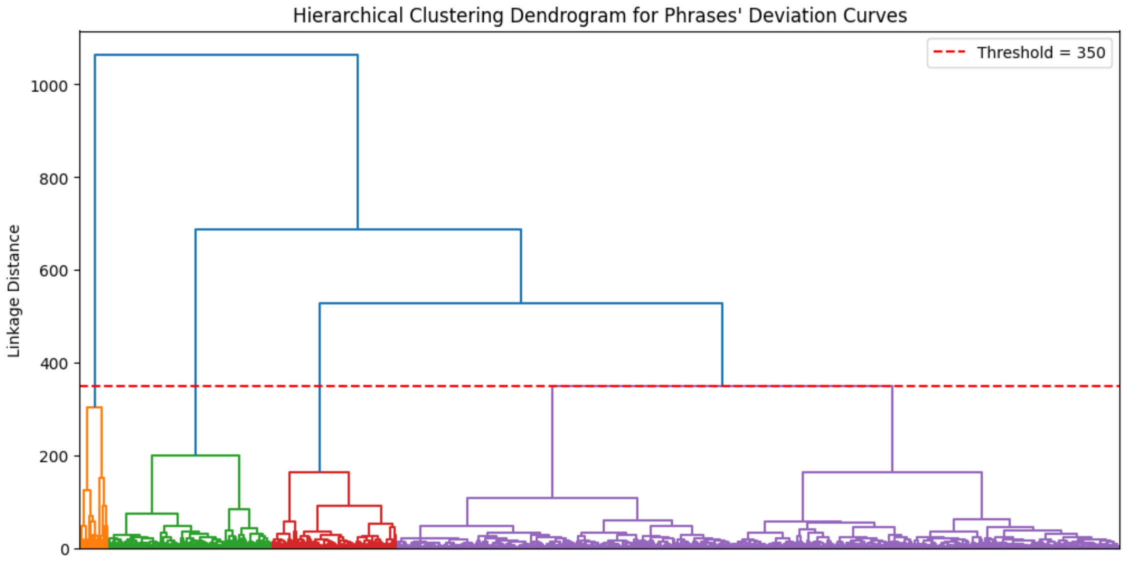


Figure 5.9: Dendrogram generated from hierarchical clustering of all phrases' BPM deviation curves.

and it complements the other methods by showing how phrases relate not just side by side, but as part of a nested, hierarchical system.

5.4.3 Tempo Curve Analysis at Phrase-level

A good way to visualise how these four curve shapes are distributed throughout the dataset is through stacked plot, or stacked bar charts. By assigning one bar for each cluster, we can colour code whatever metadata attribute we wish to track and it will show us how much of each curve type is present in that attribute. We can then go through our groups of comparison once more and see if there are any notable tendencies in any of them.

Starting once again with both cultures, in Figure 5.11. The accentuated acceleration curves seem to be exclusive to the Irish performances, which corroborates what we observed before. The remaining curves shapes seem to be proportionally distributed.

Now for different styles, in Figure 5.12. Most notable here is probably the Muiñeira, with most of its phrase curves landing on cluster 4, which corresponds to a steady tempo.

In Figure 5.13 we are looking at different instruments. Nothing seems to particularly stick out here, we can see an even distribution in all clusters.

Finally, now that we are dealing with phrases and not entire performances, we can add yet another layer, by identifying where in the performance the phrase appears. This distinction was made simply by classifying each phrase by whether it was the first phrase of the performance, the last, or anywhere else in between. Here we can see that the accentuated accelerations are happening only at the start of the performances. Thinking back to the manual beat tracking phase, while listening the audios, a few Irish performances did start at a very slow pace and quickly picked up the tempo, so that is probably the kind of tempo behaviour we are looking at here.

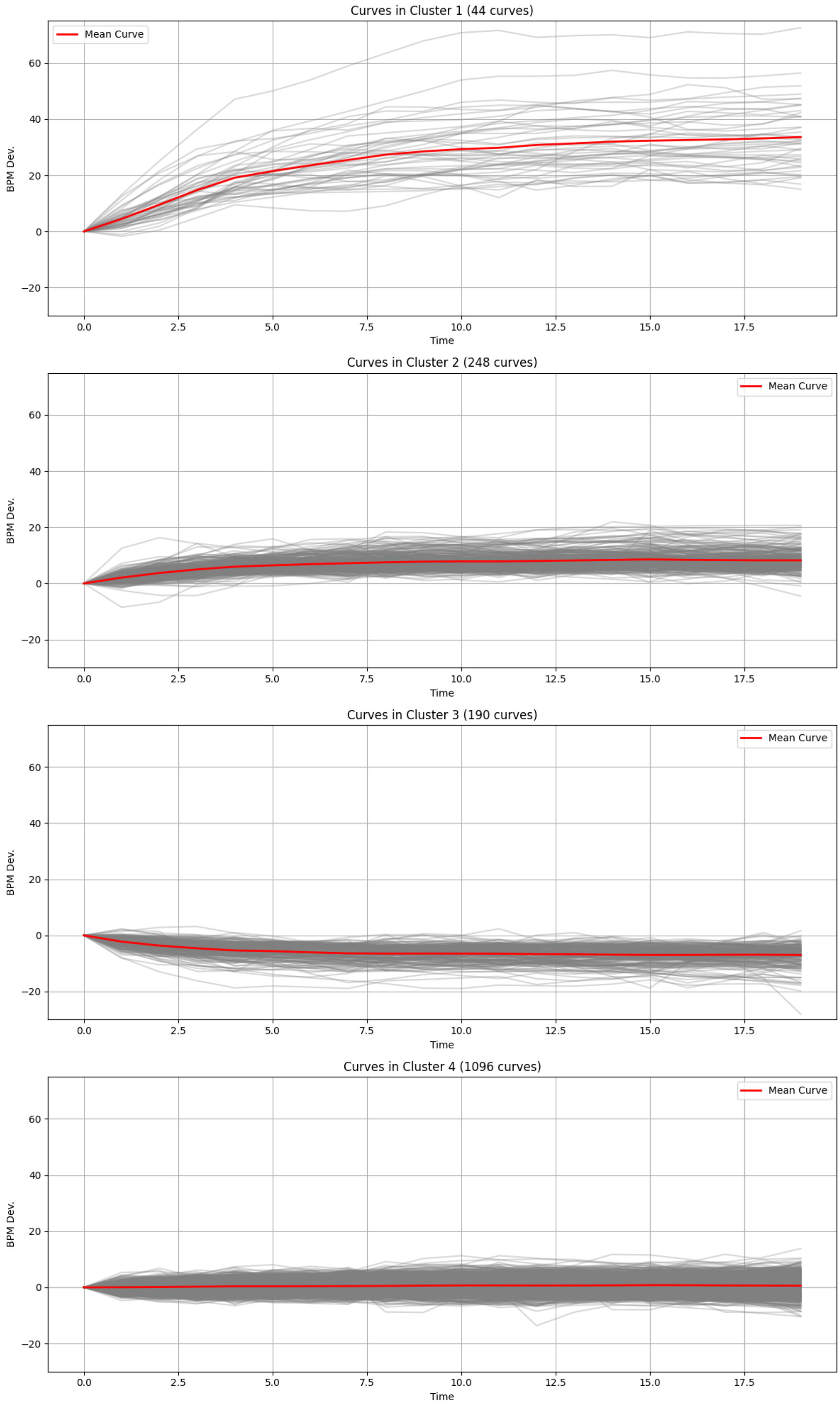


Figure 5.10: Four distinct BPM deviation curve shapes found through hierarchical clustering. Accentuated acceleration (Cluster 1), slight acceleration followed by stabilisation (Cluster 2), slight deceleration followed by stabilisation (Cluster 3), stable tempo (Cluster 4).

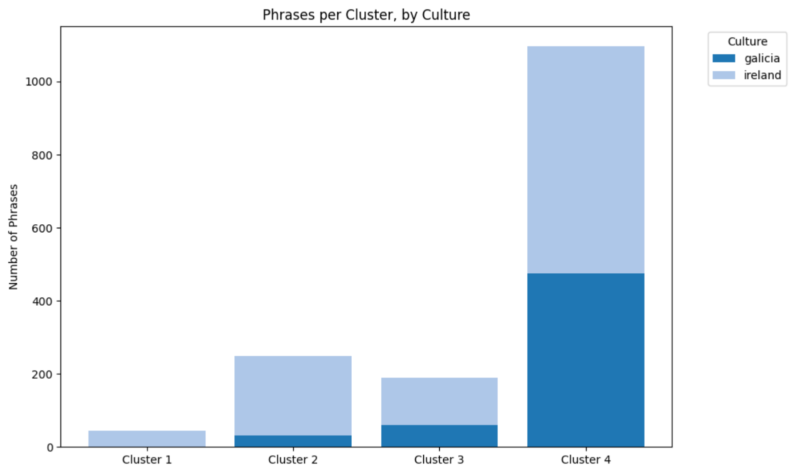


Figure 5.11: Stacked bar charts colour coded for both cultures.

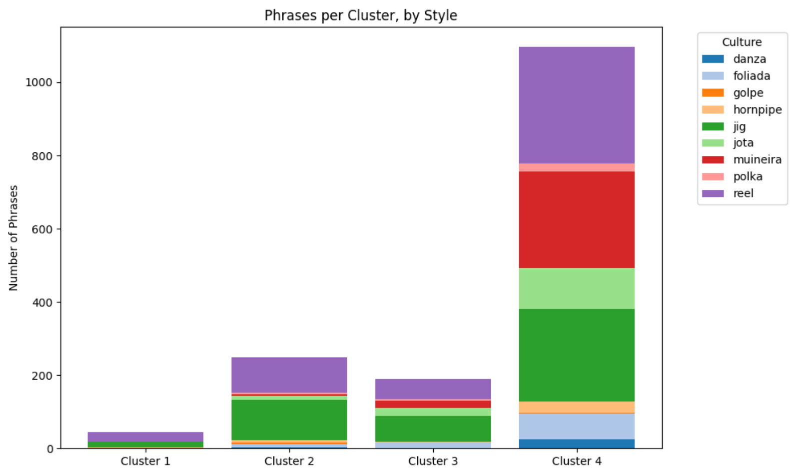


Figure 5.12: Stacked bar charts colour coded for different styles.

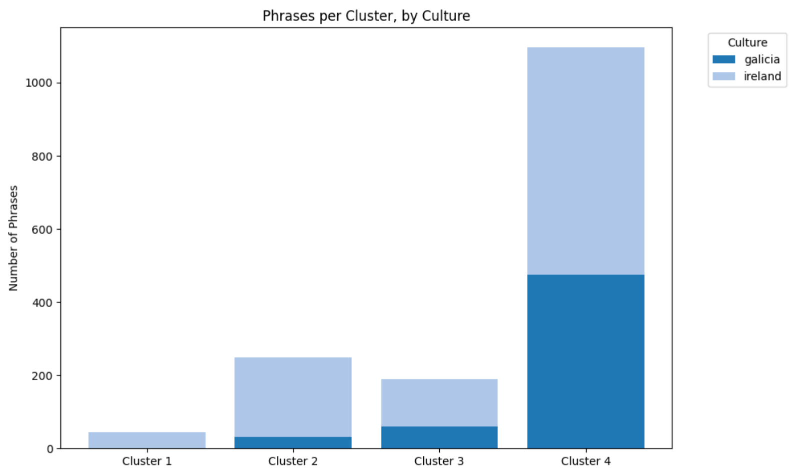


Figure 5.13: Stacked bar charts colour coded for different instruments.

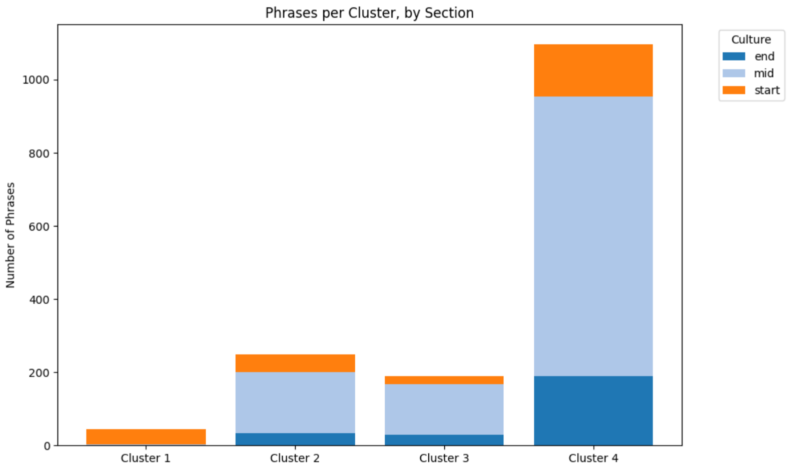


Figure 5.14: Stacked bar charts colour coded for different phrase placement in the performance.

Chapter 6

Conclusions and Future Work

Realistically, the results obtained should be considered carefully, for a number of reasons we have outlined throughout the document. First, the sample size is not as significant as it could be. Especially due to the constraints of having to manually track the beats for every individual performance analysed. This was a very time-consuming time and eventually had to be cut short when faced with the deadlines. Second, the very different nature of the datasets. The performances from the Irish portion of the dataset display a wide range of instrumentation settings, recording contexts and performers, while the Galician performances all have the same context and a limited amount of musicians. So there is a good chance any differences found were due to this contrast. Regardless, the analysis framework that was developed is replicable and expandable, to more performances and potentially more cultures.

With these considerations in mind, some of the next steps for the research can be easily outlined. Finishing the beat and segmentation annotations is a good place to start. This would certainly grant a more solid ground for our findings or even help uncover more insights. Beyond that, an expansion of the dataset. New recordings can be included and indexed seamlessly, and through the same methodology we can analyse more styles, more cultures, or even do a more focused study on a specific instrument or performer.

From a broader perspective, the work, despite current limitations, constitutes an overall positive contribution to the digital preservation and analysis of cultural heritage by demonstrating how computational methods can complement traditional ethnomusicological approaches.

References

Sarah Aldred. *Cultural Identity and Galician Folk Music*. PhD thesis, University of Strathclyde, 2005.

Phillips Barry. Irish Folk Song. 24(93), 1911.

Phillips Barry. The Transmission of Folk-Song. *The Journal of American Folklore*, 27(103):67, January 1914. ISSN 00218715. doi: 10.2307/534796. URL https://www.jstor.org/stable/534796?origin=crossref.

Sebastian Böck, Florian Krebs, and Gerhard Widmer. Accurate Tempo Estimation based on Recurrent Neural Networks and Resonating Comb Filters. 2015.

Nádia Carvalho, Sara Gonzalez-Gutierrez, Javier Merchan Sanchez-Jara, Gilberto Bernardes, and Maria Navarro-Cáceres. Encoding, Analysing and Modeling I-Folk: A New Database of Iberian Folk Music. In *8th International Conference on Digital Libraries for Musicology*, pages 75–83, Virtual Conference GA USA, July 2021. ACM. ISBN 978-1-4503-8429-2. doi: 10.1145/3469013.3469023. URL https://dl.acm.org/doi/10.1145/3469013.3469023.

José Colmeiro. Bagpipes, Bouzoukis and Bodhráns: The Reinvention of Galician Folk Music. In *A Companion to Galician Culture*, number 344 in Colección Támesis. Serie A: Monografías. Tamesis, Woodbridge, Suffolk, UK, 2014. ISBN 978-1-85566-277-3. OCLC: ocn876432107.

Simon Dixon. Automatic Extraction of Tempo and Beat From Expressive Performances. *Journal of New Music Research*, 30(1):39–58, March 2001. ISSN 0929-8215. doi: 10.1076/jnmr.30.1.39.7119. URL http://www.tandfonline.com/doi/abs/10.1076/jnmr.30.1.39.7119.

Bryan Duggan, Brendan O'Shea, Mikel Gainza, and Padraig Cunningham. Machine Annotation of Sets of Traditional Irish Dance Tunes. pages 401–406, January 2008.

Daniel P. W. Ellis. Beat Tracking by Dynamic Programming. *Journal of New Music Research*, 36(1):51–60, March 2007. ISSN 0929-8215, 1744-5027. doi: 10.1080/09298210701653344. URL http://www.tandfonline.com/doi/abs/10.1080/09298210701653344.

Dorottya Fabian and Emery Schubert. Musical Character and the Performance and Perception of Dotting, Articulation and Tempo in 34 Recordings of Variation 7 from J.S. Bach's *Goldberg Variations* (BWV 988). *Musicae Scientiae*, 12(2):177–206, July 2008. ISSN 1029-8649, 2045-4147. doi: 10.1177/102986490801200201. URL https://journals.sagepub.com/doi/10.1177/102986490801200201.

Erick Falc'her-Poyroux. Irish Music Redefined, An Outsider's Viewpoint. In Thierry Dubost and Alexandra Slaby, editors, *Music and the Irish Imagination*, pages 29–41. Presses universitaires de Caen, 2013. ISBN 978-2-84133-429-2 978-2-84133-793-4. doi: 10.4000/books.puc.7078. URL http://books.openedition.org/puc/7078.

Luís M R Guimarães. *Irish Folk Music Similarity: Comparing Engineered Features with Deep Learning Embeddings*. PhD thesis, University of Porto, 2024.

Dorothea E. Hast and Stanley Scott. *Music in Ireland: Experiencing Music, Expressing Culture*. Global Music Series. Oxford University Press, Oxford, New York, July 2004. ISBN 978-0-19-514555-7.

Andrew Neil Hillhouse. *Tradition and innovation in Irish instrumental folk music*. PhD thesis, University of British Columbia, 2005.

Andre Holzapfel and Yannis Stylianou. RHYTHMIC SIMILARITY IN TRADITIONAL TURKISH MUSIC. *Poster Session*, 2009.

Diego M. Jiménez-Bravo, Álvaro Lozano Murciego, Juan José Navarro-Cáceres, María Navarro-Cáceres, and Treasa Harkin. Identifying Irish Traditional Music Genres Using Latent Audio Representations. *IEEE Access*, 12:92536–92548, 2024. ISSN 2169-3536. doi: 10.1109/ACCESS.2024.3421639. URL https://ieeexplore.ieee.org/document/10579819/.

Alexander Lerch. *An introduction to audio content analysis: music information retrieval tasks & applications*. John Wiley & Sons, Inc, Hoboken, New Jersey, second edition edition, 2023. ISBN 978-1-119-89096-6 978-1-119-89097-3 978-1-119-89098-0.

Alexander Lerch, Claire Arthur, Ashis Pati, and Siddharth Gururani. An Interdisciplinary Review of Music Performance Analysis. *Transactions of the International Society for Music Information Retrieval*, 3(1):221–245, November 2020. ISSN 2514-3298. doi: 10.5334/tismir.53. URL http://transactions.ismir.net/articles/10.5334/tismir.53/.

M. F. McKinney, D. Moelants, M. E. P. Davies, and A. Klapuri. Evaluation of Audio Beat Tracking and Music Tempo Extraction Algorithms. *Journal of New Music Research*, 36(1):1–16, March 2007. ISSN 0929-8215, 1744-5027. doi: 10.1080/09298210701653252. URL http://www.tandfonline.com/doi/abs/10.1080/09298210701653252.

Rodrigo Mira, Eduardo Coutinho, Emilia Parada-Cabaleiro, and Björn W. Schuller. Automated composition of Galician Xota—tuning RNN-based composers for specific musical styles using deep Q-learning. *PeerJ Computer Science*, 9:e1356, May 2023. ISSN 2376-5992. doi: 10.7717/peerj-cs.1356. URL https://peerj.com/articles/cs-1356.

Meinard Müller. *Fundamentals of Music Processing: Using Python and Jupyter Notebooks*. Springer International Publishing, Cham, 2021. ISBN 978-3-030-69807-2 978-3-030-69808-9. doi: 10.1007/978-3-030-69808-9. URL https://link.springer.com/10.1007/978-3-030-69808-9.

Meinard Müller, Peter Grosche, and Frans Wiering. Automated analysis of performance variations in folk song recordings. In *Proceedings of the international conference on Multimedia information retrieval*, pages 247–256, Philadelphia Pennsylvania USA, March 2010. ACM. ISBN 978-1-60558-815-5. doi: 10.1145/1743384.1743429. URL https://dl.acm.org/doi/10.1145/1743384.1743429.

Juan José Navarro-Cáceres, Diego M. Jiménez-Bravo, and María Navarro-Cáceres. Evaluating Preprocessing Techniques for Unsupervised Mode Detection in Irish Traditional Music. *Applied Sciences*, 15(6):3162, March 2025. ISSN 2076-3417. doi: 10.3390/app15063162. URL https://www.mdpi.com/2076-3417/15/6/3162.

Bruno Nettl. Folk music | Definition, History, Artists, Songs, Bands, Instruments, & Facts | Britannica, June 2025. URL https://www.britannica.com/art/folk-music.

Amir Abbas Orouji. *Music Segmentation Analysis in Iberian Folk Music*. PhD thesis, University of Porto, 2025.

Maria Panteli, Emmanouil Benetos, and Simon Dixon. A review of manual and computational approaches for the study of world music corpora. *Journal of New Music Research*, 47(2):176–189, March 2018. ISSN 0929-8215, 1744-5027. doi: 10.1080/09298215.2017.1418896. URL https://www.tandfonline.com/doi/full/10.1080/09298215.2017.1418896.

António Pinto, Sebastian Böck, Jaime Cardoso, and Matthew Davies. User-Driven Fine-Tuning for Beat Tracking. *Electronics*, 10(13):1518, June 2021. ISSN 2079-9292. doi: 10.3390/electronics10131518. URL https://www.mdpi.com/2079-9292/10/13/1518.

Bruno H. Repp. Patterns of Expressive Timing in Performances of a Beethoven Minuet by Nineteen Famous Pianists. 1991.

Andrew Robertson. DECODING TEMPO AND TIMING VARIATIONS IN MUSIC RECORDINGS FROM BEAT ANNOTATIONS. 2012.

Hendrik Schreiber. A SINGLE-STEP APPROACH TO MUSICAL TEMPO ESTIMATION USING A CONVOLUTIONAL NEURAL NETWORK.

Hendrik Schreiber. *Data-Driven Approaches for Tempo and Key Estimation of Music Recordings*. PhD thesis, January 2020.

Hendrik Schreiber, Julián Urbano, and Meinard Müller. Music Tempo Estimation: Are We Done Yet? *Transactions of the International Society for Music Information Retrieval*, 3(1):111, August 2020. ISSN 2514-3298. doi: 10.5334/tismir.43. URL https://transactions.ismir.net/article/10.5334/tismir.43/.

Dorothé Schubarth, editor. *Cancioneiro popular galego*. Fundación "Pedro Barrié de la Maza, Conde de Fenosa", La Coruña, 1984. ISBN 978-84-85728-42-8.

Eduardo M. Torner and Jesús Bal y Gay. *Cancionero Gallego*. 1973.

George Tzanetakis. Computational ethnomusicology: a music information retrieval perspective. 2014.

Fintan Vallely, editor. *The companion to Irish traditional music*. Cork University Press, Cork, 2nd ed., substantially revised & expanded edition, 2011. ISBN 978-1-85918-450-9. OCLC: ocn772584731.

Peter Van Kranenburg, Anja Volk, and Frans Wiering. A Comparison between Global and Local Features for Computational Classification of Folk Song Melodies. *Journal of New Music Research*, 42(1):1–18, March 2013. ISSN 0929-8215, 1744-5027. doi: 10.1080/09298215.2012.718790. URL http://www.tandfonline.com/doi/abs/10.1080/09298215.2012.718790.

Appendix A

Mean Deviations Box Plots with Histograms

Here are presented the box plots described in Chapter 5, Section 5.3, along with histograms showing the distribution of the mean deviation in each groups. At the bottom of the graphs we also have the P-value scores from different statistical significance test. The usual convention is that a P-value under 0.05 tells us there is a notable difference between the two groups of data.

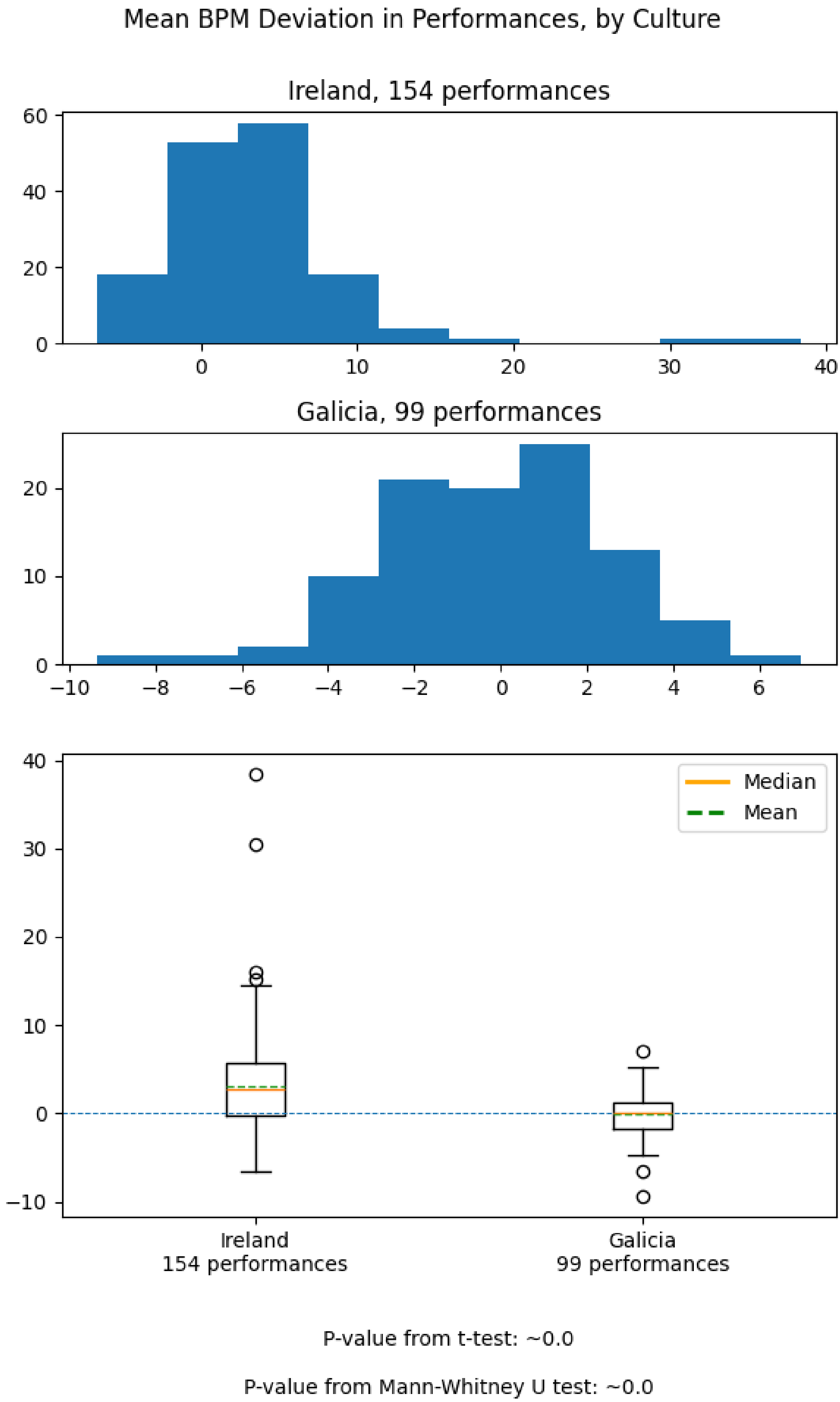


Figure A.1: Mean deviation in both cultures.

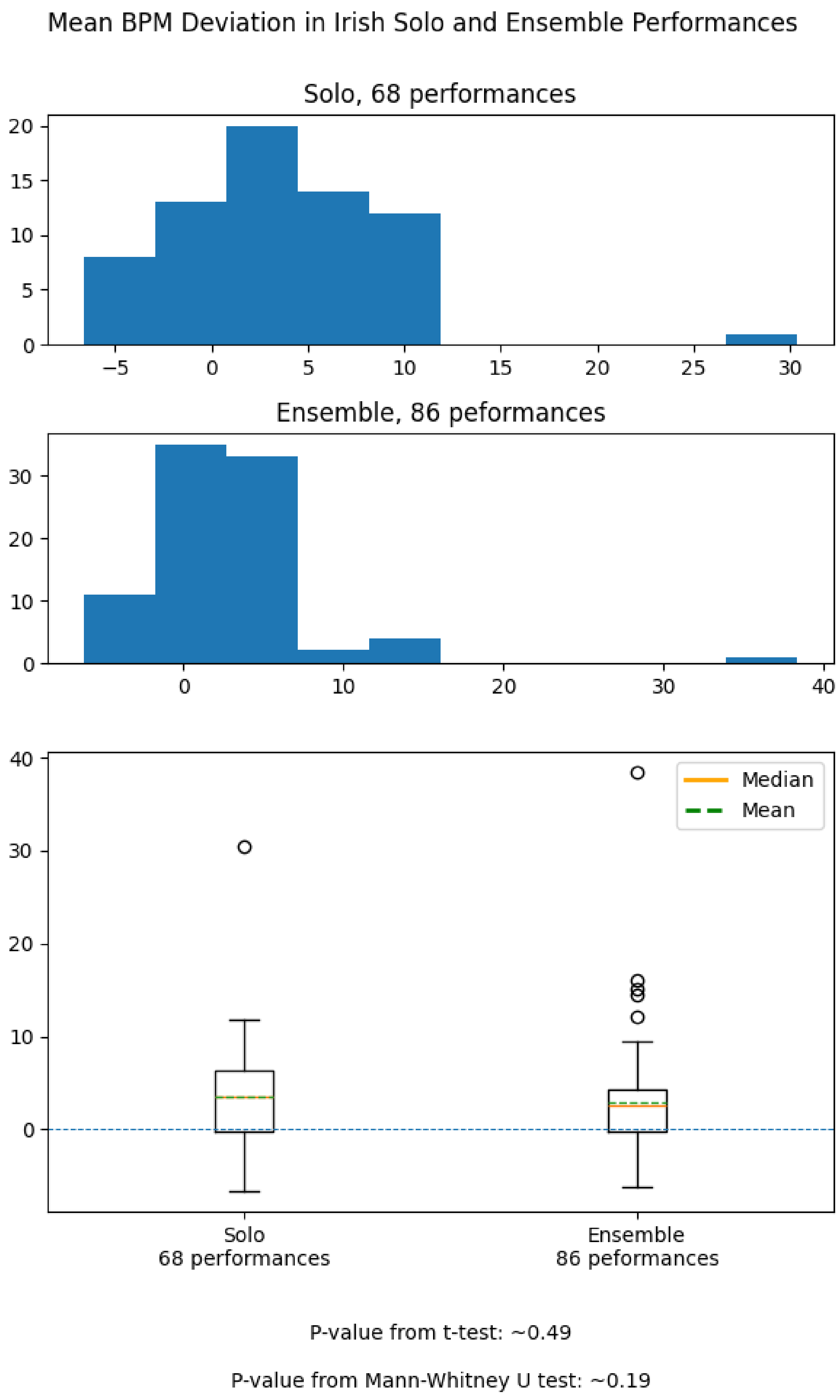
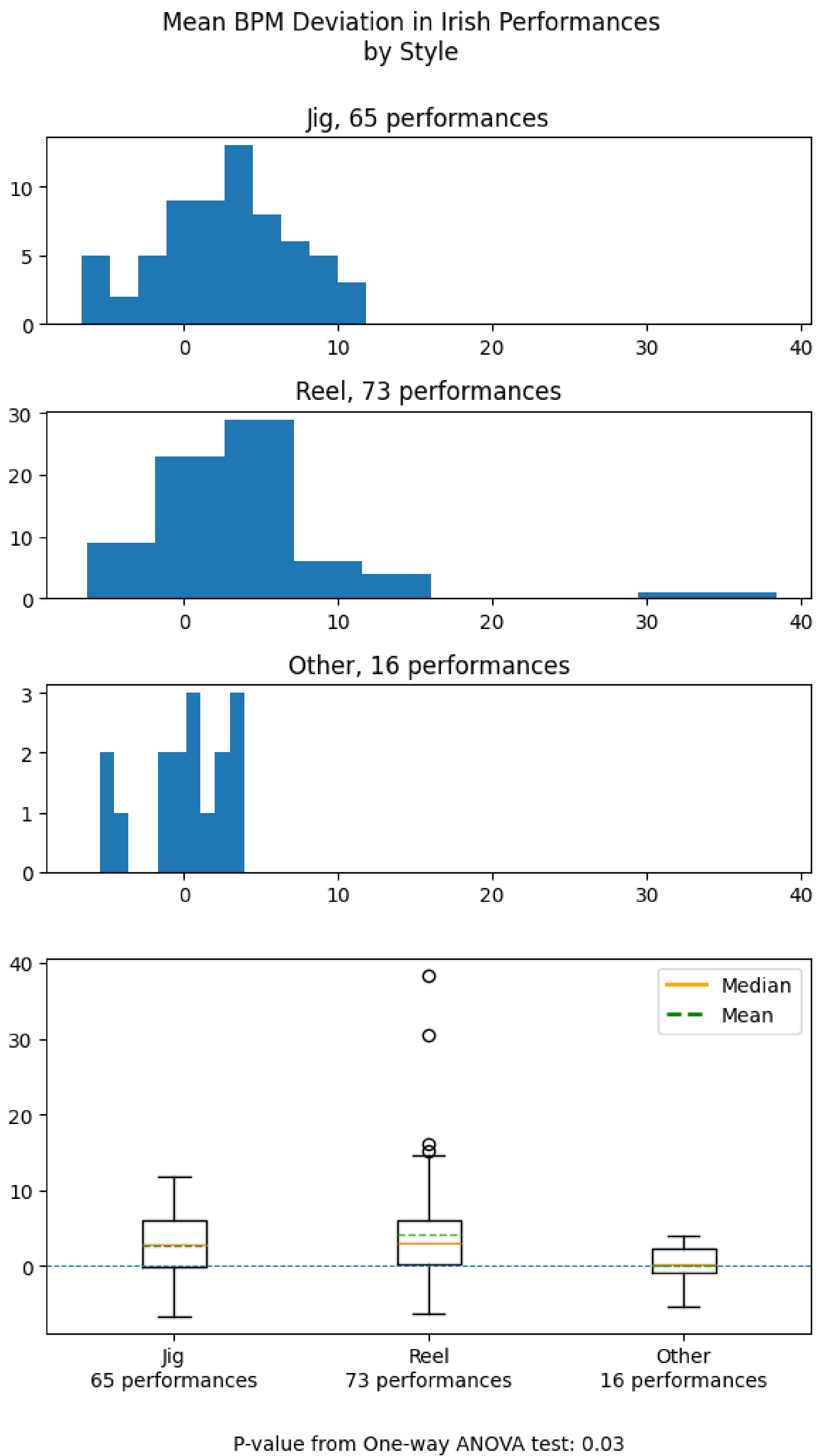
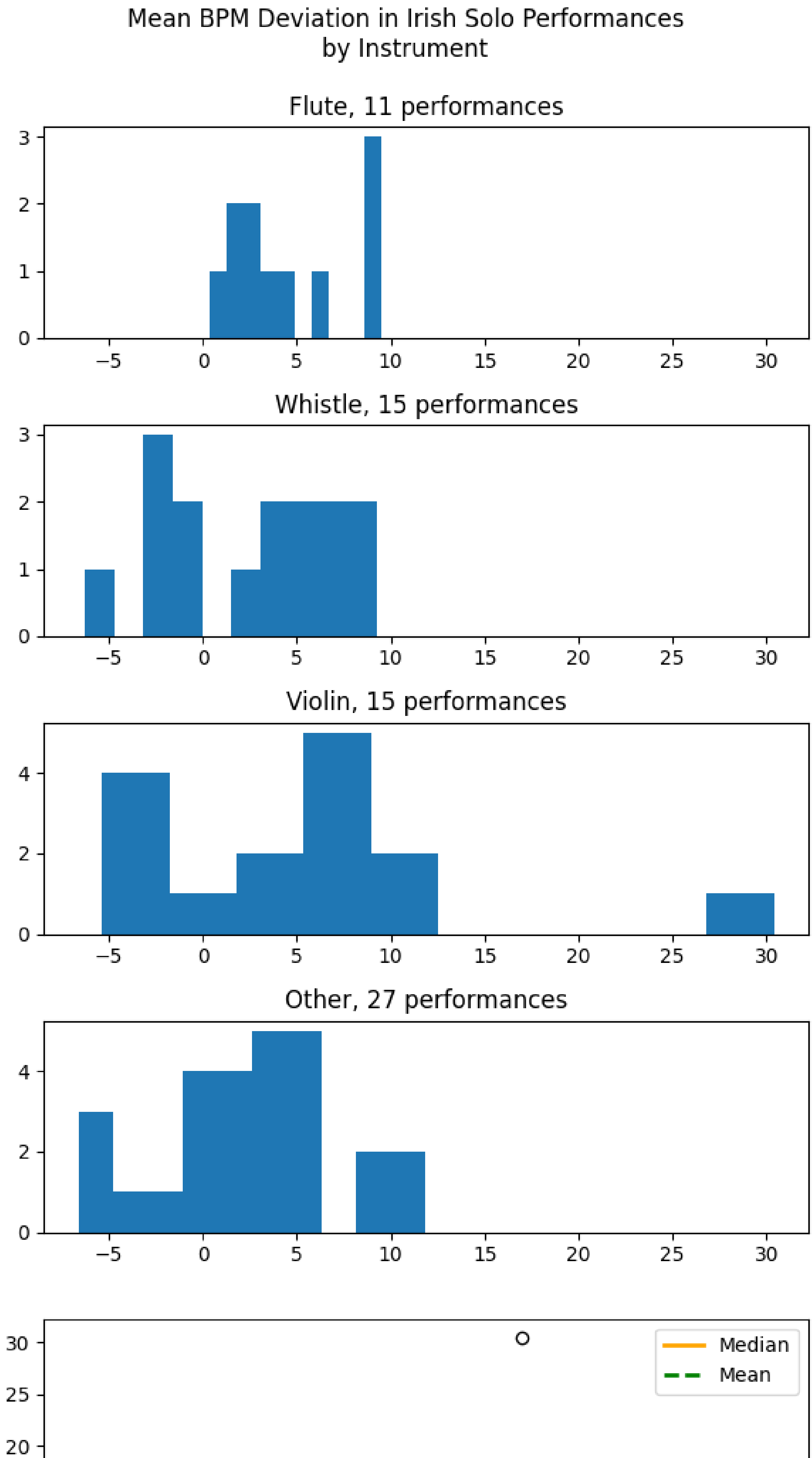
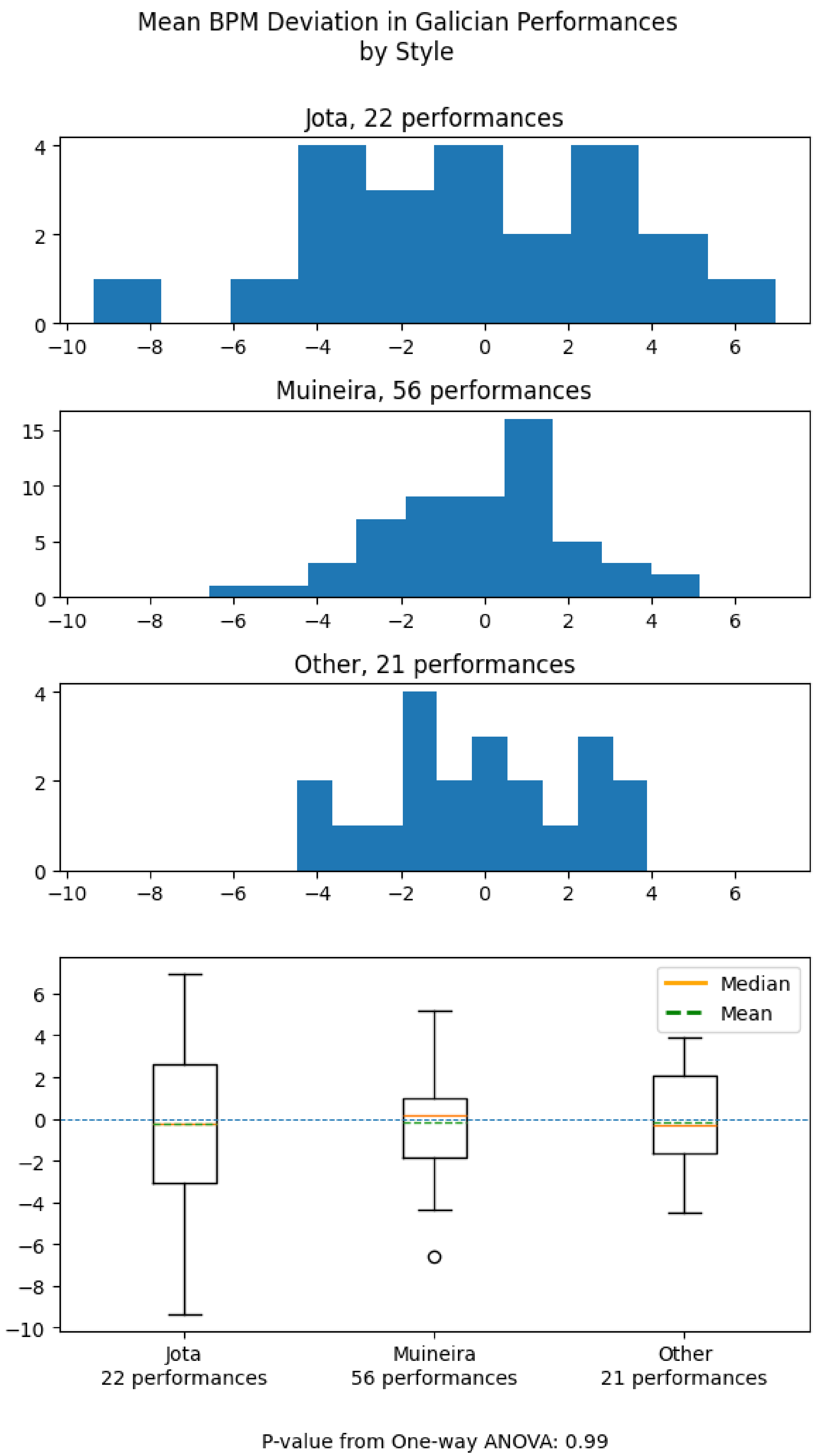
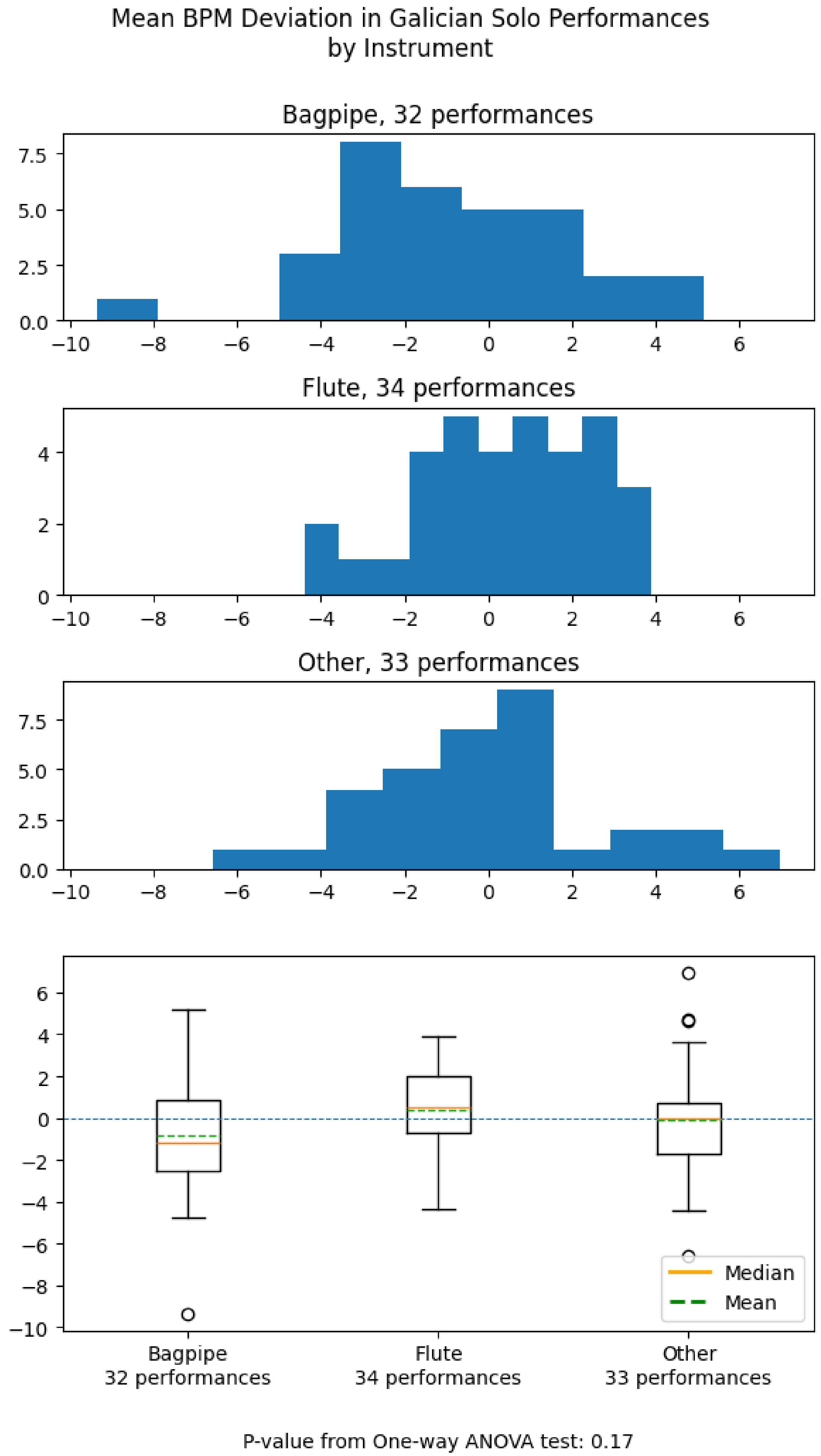


Figure A.2: Mean deviation in Irish solo and ensemble performances.


Mean BPM Deviation in Irish Performances
by Style
Jig, 65 performances
Reel, 73 performances
Other, 16 performances
Median
Mean
Jig
65 performances
Reel
73 performances
Other
16 performances
P-value from One-way ANOVA test: 0.03


Mean BPM Deviation in Irish Solo Performances
by Instrument
Flute, 11 performances
Whistle, 15 performances
Violin, 15 performances
Other, 27 performances
Median
Mean


Mean BPM Deviation in Galician Performances
by Style
Jota, 22 performances
Muineira, 56 performances
Other, 21 performances
Median
Mean
Jota
22 performances
Muineira
56 performances
Other
21 performances
P-value from One-way ANOVA: 0.99


Mean BPM Deviation in Galician Solo Performances
by Instrument
Bagpipe, 32 performances
Flute, 34 performances
Other, 33 performances
Median
Mean
Bagpipe
32 performances
Flute
34 performances
Other
33 performances
P-value from One-way ANOVA test: 0.17

Appendix B

Mean Deviations Box Plots with Histograms

The figures in this appendix illustrate the same process described in Section 5.4, but done at the performance-level. That, is applying the same hierarchical clustering and trying to group BPM curves for entire performances, rather than just phrases.

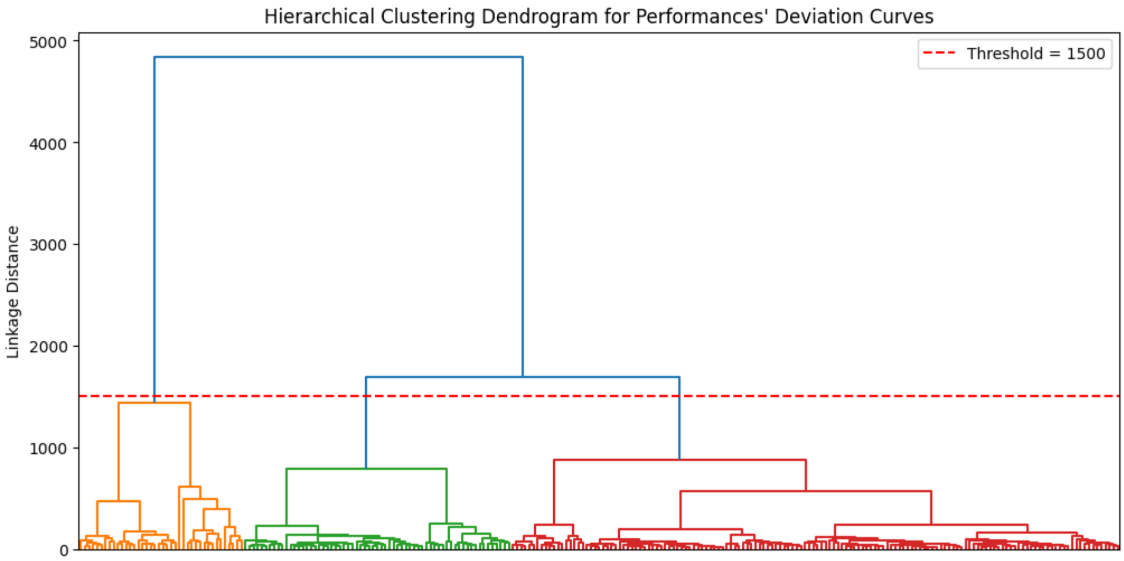


Figure B.1: Dendrogram generated from hierarchical clustering of all performances' BPM deviation curves.

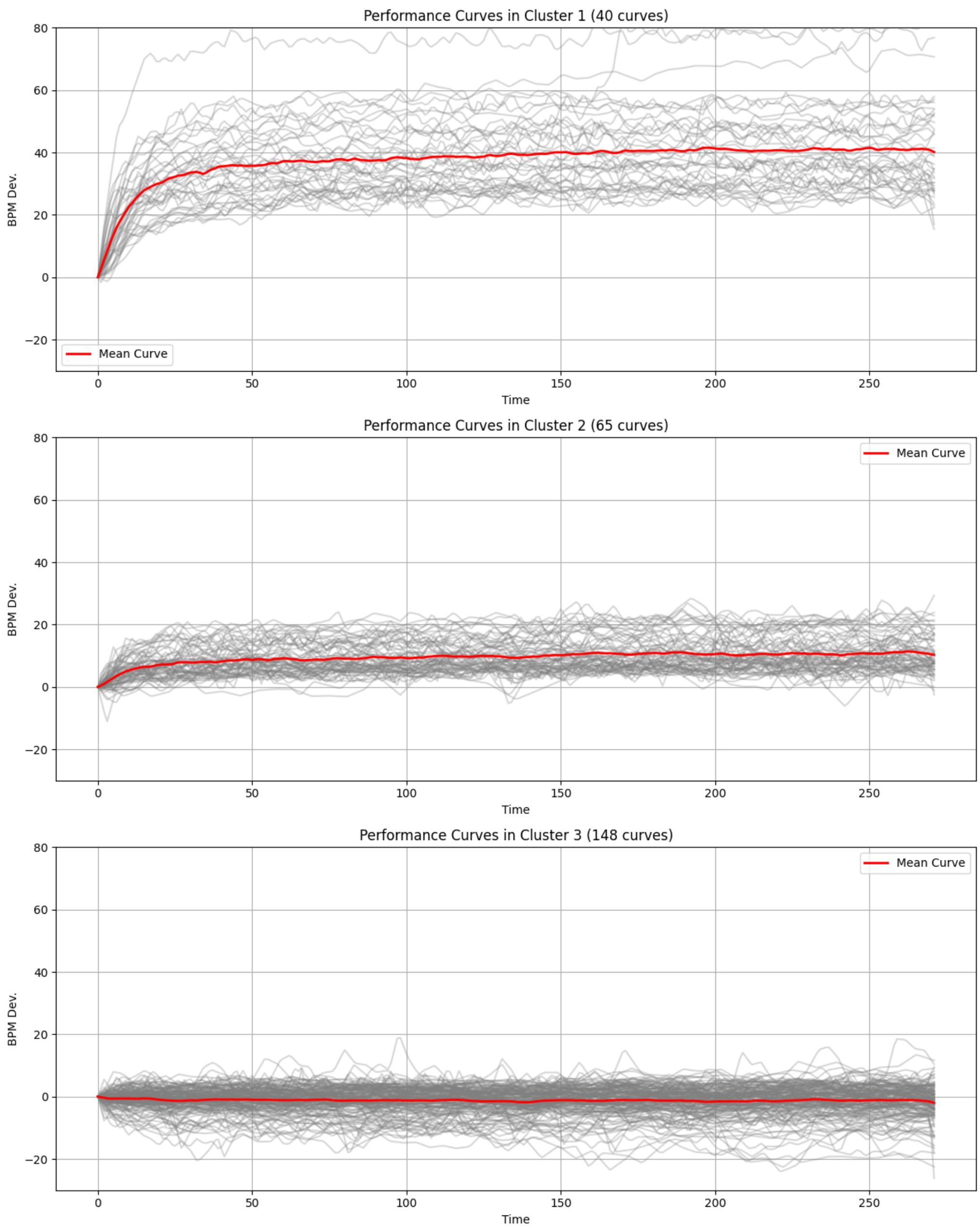


Figure B.2: Three distinct BPM deviation curve shapes found through hierarchical clustering. Accentuated acceleration (Cluster 1), slight acceleration followed by stabilisation (Cluster 2), stable tempo (Cluster 3).

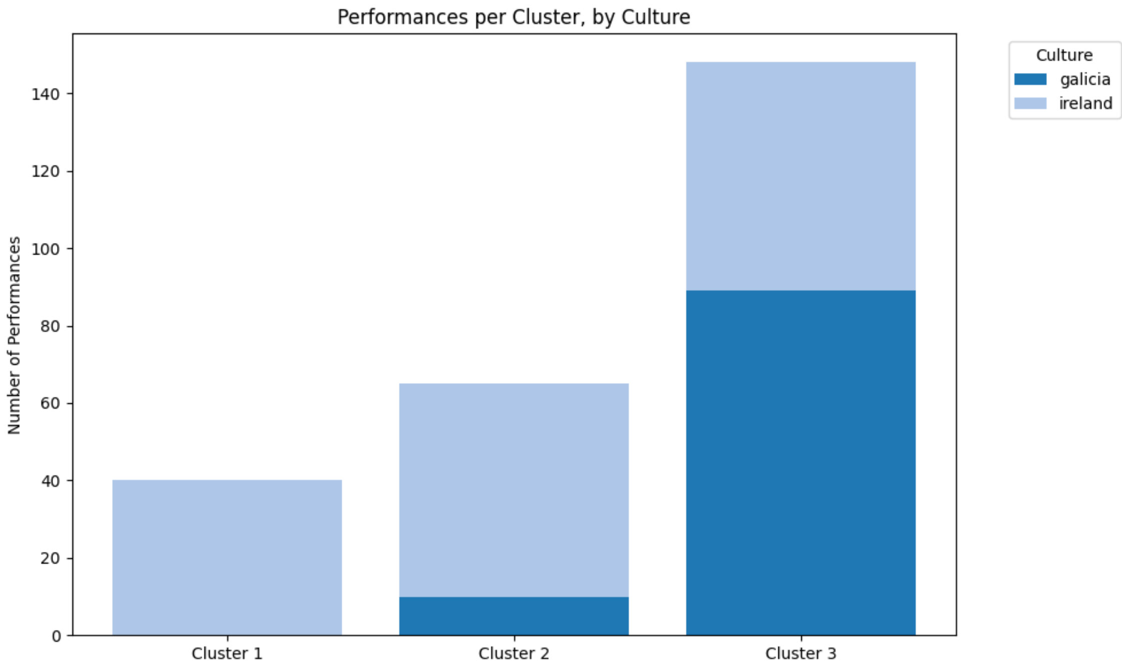


Figure B.3: Stacked bar charts colour coded for both cultures.

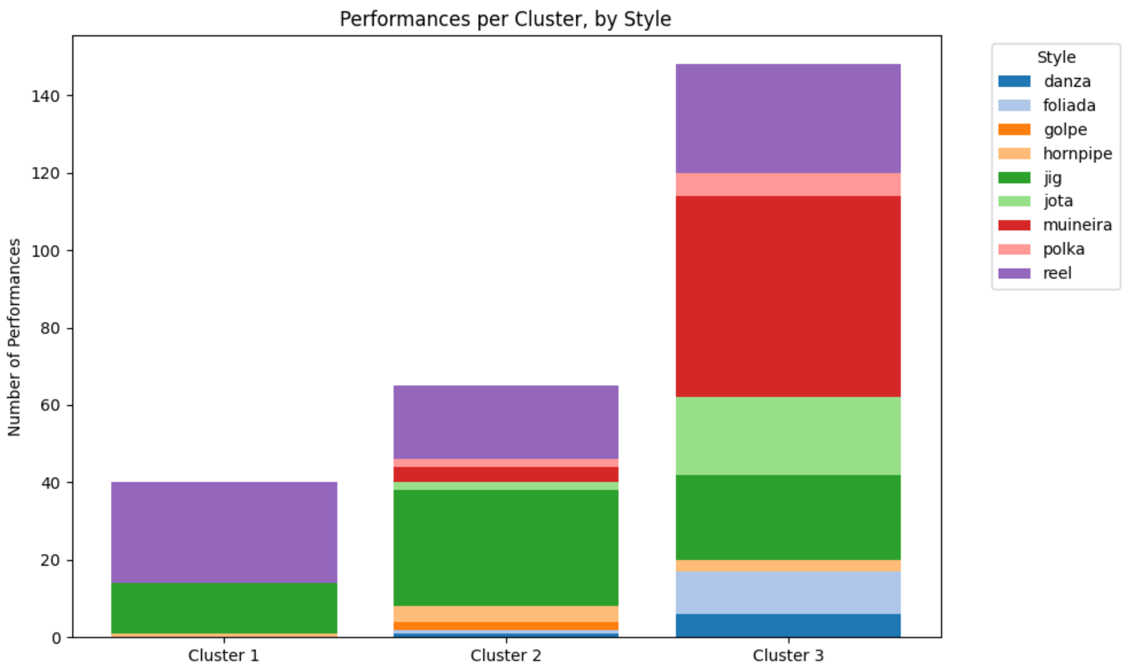


Figure B.4: Stacked bar charts colour coded for different styles.

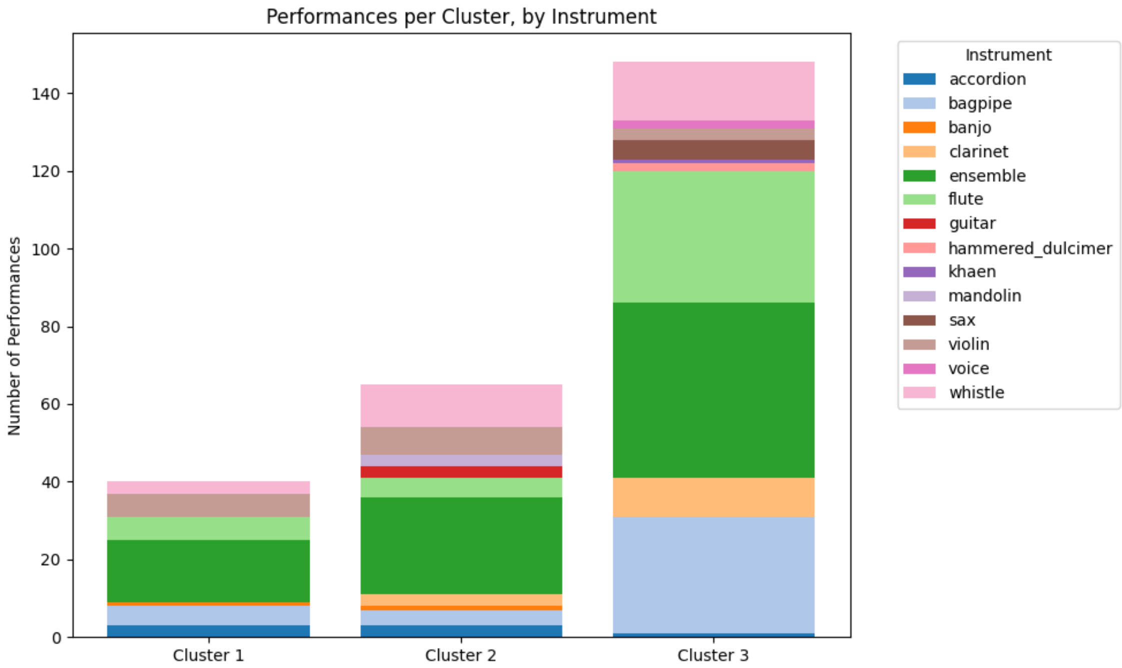


Figure B.5: Stacked bar charts colour coded for different instruments.

Appendix C

Complementary Materials

This final appendix serves just to provide a link to all the materials associated with the dissertation that were too bulky to include to be entirely included in the document. This includes the audio files, the txt files with the beats and segmentation, the code and the spreadsheet - with all the metadata, beat and segmentation information and the data calculated from it. All the generated graphs can also be found there, including the ones representing tempo over time, described in Section 5.1, for all the performances analysed, organised by song.

https://github.com/merio177/mm_dissertation