

Análisis Espectral de los Géneros Colombianos: La Cumbia

Angie Daniela Millán Córdoba
Juan Manuel Perea Coronado

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2023

Análisis Espectral de los Géneros Colombianos: La Cumbia

Angie Daniela Millán Córdoba
Juan Manuel Perea Coronado

Trabajo de grado presentado como requisito parcial para optar al título de:
Estadístico(a)

Director(a):
Jeison Mesa Diez
Codirector(a):
Ph.D. Cesar Andres Ojeda Echeverri

Universidad del Valle
Facultad de Ingeniería, Escuela de Estadística
Santiago de Cali, Colombia
2023

Dedicatoria

Dedico este trabajo a mi familia.

A mis padres, quienes con su esfuerzo, paciencia y constante amor, me apoyaron durante todo este proceso, acompañándome en los buenos momentos y en los no tan buenos.

A mis hermanos que, mediante el ejemplo, han sido mis modelos a seguir, motivándome a continuar con el cumplimiento de mis objetivos.

A mis sobrinos, porque con su alegría y su incondicionalidad me han reconfortado cuando lo he necesitado.

Como dice la letra de aquella canción que se reproduce al culminar cierto juego: "El camino es tan esencial como llegar".

Juan Manuel Perea Coronado

Dedicatoria

A mi padre, por enseñarme el valor de la fortaleza y aún en su ausencia, brindarme todo su apoyo en momentos de dificultad.

A mi madre, por ser ejemplo de fuerza y disciplina durante toda mi vida.

A mis hermanas, quienes son el motor de motivación y felicidad.

Angie Daniela Millán Córdoba

Agradecimientos

Agradecemos a todas las personas que contribuyeron de manera significativa en la realización de este trabajo de grado. Sus contribuciones y apoyo han sido fundamentales para el éxito de esta investigación.

A todos los docentes que han aportado a la formación profesional y personal, en especial a los profesores: Jeison Mesa Diez, por dirigir nuestro trabajo de grado, por brindarnos su tiempo incondicionalmente y su paciencia durante este proceso. Cesar Andrés Ojeda Echeverri, por ser codirector en nuestro trabajo de grado, siendo una guía elemental en el desarrollo de este y a Luz Adriana Pereira Hoyos, quién asesoró y nos brindó su completa disposición, conocimiento y apoyo a lo largo de este trabajo.

A nuestros familiares y amigos, por confiar en nosotros y ser incondicionales en el transcurso de nuestro desarrollo como profesionales.

A nuestra Alma Mater, la Universidad del Valle, por su invaluable contribución a nuestra formación académica y por brindarnos la oportunidad de realizar este trabajo de grado, permitiéndonos culminar nuestro estudio de pregrado.

Resumen

El presente trabajo propone una metodología estadística de carácter exploratorio, con el fin de identificar patrones y tendencias en canciones del género cumbia colombiana, mediante la descomposición de sus características usando los paquetes del software de computación estadística y gráfica, R. La metodología se centra en la extracción de la frecuencia fundamental de los registros musicales de 30 canciones variadas del género, empleándolas como insumo principal para el desarrollo de análisis de componentes principales de series temporales, y así obtener componentes interpretables que resuman el comportamiento de las frecuencias de todas las canciones. En el análisis se encontró que, aunque individualmente los registros musicales poseen características tonales distintas entre sí, tienen en común el repetir cada medio segundo la misma secuencia de notas que; al escuchar detenidamente las canciones; están asociadas a los sonidos producidos por la percusión de fondo. Este comportamiento prevalece durante toda la reproducción en todas las canciones.

Palabras clave: Cumbia Colombiana, Análisis de Componentes Principales, Series de Tiempo, Patrones, Frecuencia Fundamental, Periodicidad

Abstract

This paper proposes a statistical methodology of an exploratory nature, in order to identify patterns and trends in songs of the Colombian cumbia genre, by detecting their characteristics using the packages from the statistical and graphical computing software, R. The methodology focuses on the extraction of the fundamental frequency from musical registers of 30 varied songs of the genre, using them as the main input for the development of time series principal component analysis, and thus obtain interpretable components that summarize the behavior of the frequencies of all the songs. In the analysis it was found that, although individually the musical registers have different pitch features, they have in common the repetition of the same notes sequence every half second; by listening carefully to the songs; they are associated with the sounds produced by background percussion. This behavior prevails during all playback on all songs.

Keywords: Colombian Cumbia, Principal Component Analysis, Time series, Patterns, Fundamental Frequency, Periodicity

Contenido

Resumen vi

Lista de Figuras xi

Lista de Tablas xii

1. Introducción 1
1.1. Planteamiento del Problema . . . 2
1.2. Justificación . . . 4

2. Objetivos 5
2.1. Objetivo General . . . 5
2.2. Objetivos Específicos . . . 5

3. Antecedentes 6
3.1. Antecedentes contextuales . . . 6
3.2. Antecedentes estadísticos . . . 7

4. Marco Teórico 10
4.1. Marco teórico contextual . . . 10
4.2. Marco teórico estadístico . . . 13

5. Metodología 22
5.1. Contrucción de fuentes de datos . . . 24
5.1.1. Revisión de calidad de los archivos . . . 24
5.1.2. Selección de la muestra . . . 25
5.1.3. Análisis preliminar de las señales de audio . . . 26
5.2. Extracción de la frecuencia fundamental . . . 27
5.2.1. Frecuencia fundamental de un sonido . . . 27
5.2.2. Método de Welch . . . 27
5.2.3. Extracción de la frecuencia fundamental en R . . . 29
5.2.4. Análisis preliminar de las series temporales . . . 29
5.3. Resumen de características mediante ACP para series temporales . . . 31
5.3.1. Análisis de Componentes Principales para series temporales . . . 31

5.3.2. Transformación de resultados en el dominio de las frecuencias al dominio del tiempo 32
5.4. Inspección y análisis de resultados 33
5.5. Software 33

6. Resultados **34**
6.1. Análisis preliminar de las señales de audio 34
6.2. Análisis preliminar de las series temporales 35
6.3. Análisis de la estimación del espectro de la primera componente principal . . 38
6.4. Interpretación de resultados en el dominio del tiempo 40
6.5. Inspección de resultados 42

7. Conclusiones y recomendaciones **44**
7.1. Conclusiones 44
7.2. Recomendaciones 45

A. Anexos: Gráficos, tablas y código en R **49**

Lista de Figuras

5-1. Esquema de metodología y actividades por fases 23
5-2. Representación de la segmentación de la señal realizada en el método de Welch. 28

6-1. Oscilogramas de las primeras tres canciones de la muestra, generados en R. El eje verticual corresponde a la amplitud de la onda, mientras que el eje horizontal corresponde al tiempo medido en segundos. 35
6-2. Representación gráfica de las series temporales asociadas a las primeras tres canciones. En el eje vertical se encuentran los valores de frecuencia fundamental, mientras que en el horizontal las unidades temporales. Cada dato está asociado a 1/21.53 de segundo. 36
6-3. Representación gráfica detallada de las periodicidades presentes en la serie temporal asociada a la tercera canción, con valores temporales entre 1750 y 2500. 37
6-4. Periodogramas de las series temporales asociadas a las primeras ocho canciones. En el eje vertical se encuentra la estimación del periodograma, mientras que en el eje horizontal están las frecuencias $w = \frac{j}{n}$, medida en ciclos por unidad de tiempo. Las líneas verticales rojas indican la frecuencia donde se genera el pico máximo de poder espectral, siendo para cada uno de los ocho periodogramaslos valores 0.0022, 0.0729, 0.0428, 0.0402, 0.0006, 0.0421, 0.0405 y 0.0363, respectivamente. 38
6-5. Estimación del espectro de la primera componente principal, donde se señalan las frecuencias asociadas a los primeros veinte máximos, agrupadas por cercanía usando colores: rojo, azul y amarillo. 39

A-1. Oscilogramas de las primeras 6 canciones 49
A-2. Oscilogramas de las canciones desde la 7 hasta la 12 50
A-3. Oscilogramas de las canciones desde la 13 hasta la 18 50
A-4. Oscilogramas de las canciones desde la 19 hasta la 24 51
A-5. Oscilogramas de las canciones desde la 24 hasta la 30 51
A-6. Representación gráfica de las primeras 6 series temporales asociadas a la frecuencia fundamental de las canciones 52
A-7. Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 7, hasta la 12 52

A-8. Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 13, hasta la 18 53
A-9. Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 19, hasta la 24 53
A-10. Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 25, hasta la 30 54
A-11. Periodogramas estimados de las primeras 8 series temporales asociadas a la frecuencia fundamental de las canciones, con sus valores máximos señalados en rojo. 55
A-12. Periodogramas estimados de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 9, hasta la 16 55
A-13. Periodogramas estimados de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 17, hasta la 24 56
A-14. Periodogramas estimados de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 25, hasta la 30 56

Lista de Tablas

6-1. Promedio de la proporción de poder espectral explicada por las primeras cinco componentes y su acumulado, por cada grupo de frecuencias. 40
6-2. Tabla de las descripciones de los grupos de frecuencias asociadas a los máximos encontrados en la estimación del espectro de la primera componente. 41

A-1. Tabla de metadatos asociados a la muestra de canciones. 57
A-2. Tabla de las proporciones de poder espectral explicada por las primeras cinco componentes en las frecuencias destacadas 58
A-3. Tabla de las frecuencias asociadas a los primeros veinte valores máximos dentro de la estimación del espectro de la primera componente principal, con su respectiva representación en el dominio del tiempo. 59
A-4. Tabla de caracterización de los atributos de la muestra de canciones. 59
A-5. Tabla de estadísticas descriptivas de las frecuencias fundamentales de cada canción. 60

Declaración

Nos permitimos afirmar que hemos realizado el presente Trabajo de Grado de manera autónoma y con la única ayuda de los medios permitidos y no diferentes a los mencionados en el propio trabajo. Todos los pasajes que se han tomado de manera textual o figurativa de textos publicados y no publicados, los hemos reconocido. Ninguna parte del presente trabajo se ha empleado en ningún otro tipo de Tesis o Trabajo de Grado.

Igualmente declaramos que los datos utilizados en este trabajo están protegidos por las correspondientes cláusulas de confidencialidad.

Santiago de Cali, 06.07.2023

Angie Daniela Millán Córdoba

Juan Manuel Perea Coronado

1. Introducción

Cada vez es aún mayor los avances tecnológicos que se implementan o interceden en la vida cotidiana, uno de ellos es el uso de archivos multimedia, como audio, videos, entre otros. En el caso de las señales de audio se han desarrollado diferentes tipos de aplicaciones, adquiriendo mejoras a través del procesamiento de audio y el aprendizaje automático, es aquí donde toma importancia la clasificación de audio, ya que nos permite entender y analizar el contenido de audio (Martín, 2017). Teniendo en cuenta que, el audio puede verse representando a través de señales dependientes del tiempo, describiendo así un fenómeno físico y catalogándose como una serie de tiempo (Zapata, 2021), es posible emplear análisis por medio de las diversas técnicas y métodos de estudio para estas.

El uso de metodologías enfocadas en análisis de series de tiempo ha logrado el reconocimiento de tendencias y patrones en los diversos estudios de información. Consecuente a lo anterior, el uso de técnicas estadísticas es aún más frecuente al tratarse de datos de series temporales, esto con el fin de explorar y entender las tendencias resultantes entre observaciones, o en su efecto, explicar la dependencia entre ellas (Cruz et al., 2019). Algunas herramientas por las cuales se han llevado a cabo análisis de series de tiempo son: Análisis de Fourier, donde a través de la sumatoria de funciones periódicas nos permite determinar las frecuencias principales; Análisis Espectral, el cual permite descomponer un conjunto de datos en términos de sus componentes cíclicos o repetitivos (González, 2001), logrando encontrar patrones de comportamiento entre series y por último, se tiene la combinación de las anteriores técnicas con metodologías de Aprendizaje automático, como Redes Neuronales, Modelos de Deep learning y Arboles de clasificación (Cruz et al., 2019).

El estudio de patrones musicales se ha convertido en un área de interés crecienciente en la comunidad científica, proporcionando información valiosa sobre la estructura subyacente de las canciones, en el caso de Molina (2020) hace uso del análisis espectral con el fin de modelar los armónicos de diferentes instrumentos, para así obtener los "valores objetivos" necesarios para sintetizarlos por medio de un modelo de síntesis aditiva lo cual permite fusionar diferentes ondas para obtener formas más complejas y robustas, por otro lado Osorio et al. (2012) recurre al Análisis de Fourier para descomponer el audio en términos de frecuencia, amplitud y fase.

Las técnicas antes mencionadas han demostrado grandes resultados para el análisis de

tendencias musicales, sin embargo, hay pocas investigaciones centradas en la caracterización de géneros musicales y particularmente en géneros musicales Colombianos, los cuales cuentan con expresiones multiculturales, aportando características de nuestra sociedad (Cruz et al.2019).

En este trabajo de investigación, se propone utilizar una combinación de técnicas estadísticas basadas en descomposición de la densidad espectral y análisis de componentes principales aplicado a series de tiempo. El objetivo principal es desarrollar un enfoque integral que permita identificar patrones musicales relevantes, tanto en términos de su contenido espectral como de su estructura temporal por medio del uso de paquetes del software estadístico R.

La metodología implementada se centró en la extracción de frecuencias fundamentales de registros musicales, empleando esto como insumo principal para el desarrollo de análisis de componentes principales de series temporales. Se llevaron a cabo 4 fases, donde la primera corresponde a la construcción de fuentes de datos, realizando así una revisión de calidad de los archivos, para seleccionar la muestra y por último realizar un análisis preliminar de las señales de audio; la segunda fase corresponde a la extracción de las frecuencias fundamentales haciendo uso del algoritmo de detección de pitch, generando así un análisis preliminar de las series temporales; en la tercer fase se aplica el análisis de componentes principales de las series temporales antes mencionado; finalmente, se obtienen componentes interpretables, resumiendo las frecuencias de todas las canciones, permitiendo conocer su registro tonal.

1.1. Planteamiento del Problema

La música hace parte de la historia étnica de un país, desarrollando funciones sociales y culturales dentro de él y en las distintas etapas de evolución colectiva de su población, de manera que; por cada región, departamento y/o ciudad se encuentra una expresión artística. La música tradicional colombiana enmarca características de la región en cuestión; fruto de la mezcla histórica de ritmos indígenas, africanos y europeos; contribuyendo a la construcción de la identidad de esta y abriendo paso a los múltiples géneros que radican en ella (Cruz et al., 2019)

La música tradicional colombiana puede ser clasificada a través de sus distintas regiones: Caribe, Pacífico, Orinoquía, Amazonía y Andina; cada una con sus propios atributos y particularidades que logran diferenciar los géneros musicales entre una región y otra; ya sea a través del tempo del ritmo, los instrumentos que se emplean o las composiciones en sí. Teniendo en cuenta lo anterior, el caracterizar o esquematizar un género musical de una

región colombiana en particular, es a su vez una forma de identificar a la región misma y al país al que pertenece. (Cruz et al., 2019)

Entre la gran gama de géneros y ritmos que se pueden encontrar en el folclor colombiano, se encuentra la Cumbia. La Cumbia es un género musical originario de la costa colombiana, en la región Caribe. Se caracteriza por la incorporación de elementos rítmicos e instrumentales de orígenes africanos, indígenas y en menor medida, españoles. Como género musical, suele ser acompañado por baile y movimientos enérgicos, galantes y seductores; en su mayoría, provenientes de la cultura africana, incluyendo vestimentas y trapos originarios de la cultura española (Arias, 2021).

Con los avances tecnológicos, la clasificación de los géneros musicales a través de plataformas digitales de música es más sencillo. Sin embargo, la presencia de géneros folclóricos o tradicionales como la cumbia es escasa en este tipo de espacios. "Los géneros musicales tradicionales de Colombia no son ampliamente reconocidos a nivel mundial y, en algunos casos, han perdido popularidad y permanecen exclusivamente en las regiones a las que pertenecen culturalmente" (Cruz et al., 2019).

Así mismo, se debe resaltar el interés del estudio del género musical en la comunidad científica; analizando y destacando las piezas musicales, considerando los sonidos como datos no estructurados, los cuales se perciben por dominios de tiempo y dominios armónicos (Reyes, 1998) resaltando que, los dominios de tiempo hacen referencia a los patrones de ritmo o repeticiones del sonido en función del tiempo, mientras que los dominios armónicos se limitan a la orquestación y manipulación tímbrica de los sonidos, siendo el sonido el conjunto de ondas mecánicas, percibidas a través del aire, de modo que, el audio es una representación digital del sonido (Olivares, 2019).

Con el objetivo de obtener las características más relevantes de las canciones, se lleva a cabo la clasificación de señal de audio, teniendo en cuenta los diversos tipos que incluye como lo son las características temporales (el centroide temporal, la autocorrelación, la tasa de cruce por cero), características espectrales (centroide espectral, ancho, asimetría, curtosis, planitud son momentos estadísticos obtenidos del espectro), MFCC- Mel Frequency Cepstral Coefficients , descriptores de energía (energía total, energía de sub-banda, energía armónica y energía del ruido) y características armónicas (frecuencia fundamental, ruido, falta de armonía) (Yu & Slotine, 2009).

Las representaciones de tiempo-frecuencia presentan patrones o tendencias en el dominio visual. Por lo que una herramienta fundamental de análisis son los espectrogramas los cuales representan de forma visual la señal de audio, proporcionando el beneficio de la clasificación. Así, al hablar de un grupo de canciones que hacen participe de un género, hablamos de

un conjunto de observaciones indexadas por el tiempo, las cuales pueden ser estudiadas y analizadas a través de métodos paramétricos y no paramétricos, proporcionando análisis interpretable de la serie temporal. De forma general, el audio es perteneciente a datos no estructurados, el cual a través de de herramientas estadísticas logran ser datos estructurados, ofreciendo así un análisis del audio, por medio de la densidad espectral. Teniendo en cuenta lo anterior, la pregunta de investigación que se desarrollará: ¿Cómo caracterizar un género musical tradicional colombiano como la Cumbia a través de componentes espectrales interpretables que permitan definir particularidades en su composición?

1.2. Justificación

Cómo se mencionó anteriormente, la clasificación de audio ha venido adquiriendo importancia en conjunto con la evolución digital y tecnológica del mundo, siendo relevante tanto como en su contenido teórico como en el reconocimiento de patrones en el campo práctico, aportando así a la inteligencia artificial y a la automatización de procesos (Rong, 2016).

Con el propósito de analizar e identificar patrones, elementos musicales y/o frecuenciales distintivos en las señales de audio, surge la implementación de herramientas y métodos para obtener todas las características y componentes esenciales, sin generar cálculos redundantes. De forma que, al hablar de una clasificación de audio, hablamos del análisis espectral, el cual nos permite obtener información contenida en la serie temporal, extrayendo los elementos y componentes característicos mediante técnicas no paramétricas o paramétricas (Yu & Slotine, 2009).

La posibilidad de caracterizar un género representativo culturalmente para Colombia, como lo es la Cumbia, a través de análisis espectral, sin escatimar información relevante, proporciona un avance cultural de la mano con la evolución tecnológica, destacando el comportamiento de señales en el dominio de la frecuencia, detectando variaciones o tendencias en el género.

El uso de técnicas conjuntas como el método de Welch y el Análisis de Componentes principales para series temporales, nos permite realizar caracterización espectral al género Cumbia, donde en el primero se realiza un promedio del periodograma, identificando las frecuencias en la serie de tiempo, reduciendo su varianza y protegiendo la información de pérdida; el segundo, se encarga de obtener un grupo de componentes representativas. Es necesario señalar que, estas técnicas empleadas para el análisis espectral consideran aplicaciones en campos diversos, teniendo en cuenta el objetivo principal, como lo es el comportamiento dinámico de un sistema determinado.

2. Objetivos

2.1. Objetivo General

Caracterizar las frecuencias que componen al género regional cumbia en Colombia.

2.2. Objetivos Específicos

- Estimar el espectro para los registros musicales, usando la transformada rápida de Fourier y el método de Welch.
- Analizar la descomposición del espectro para los registros musicales mediante análisis de componentes principales para series temporales.
- Caracterizar la descomposición del espectro resultante del análisis de componentes principales para series temporales para los registros musicales.

3. Antecedentes

A continuación, se presentan los antecedentes investigados tanto contextuales como estadísticos. El primer grupo de investigaciones son estudios sobre análisis de características en géneros musicales. A pesar de que emplean análisis estadístico, no recurren al método de estudio del trabajo realizado, sin embargo, son de gran aporte para reconocer el amplio análisis de las características musicales. El segundo grupo de investigaciones corresponde a estudios en los cuales se aplicó el análisis espectral singular o el análisis de componentes principales en diversas áreas de conocimiento; además del método de Welch como técnica de extracción de la frecuencia fundamental.

3.1. Antecedentes contextuales

Se realizó una revisión de artículos y trabajos de grado, relacionados con el análisis de características de géneros musicales, con el objetivo de indagar las diversas formas de estudio de dichas características.

Bernal et al. 2009. Realizan un estudio de los parámetros que interceden en el cálculo de la transformada de Fourier con el fin de exhibir esta herramienta como útil para el estudio de la fonética acústica, haciendo uso del Software Matlab. Realizaron la grabación de la voz en formato WAV, posterior a esto, muestrearon la señal de voz a analizar, una vez muestreada se limitan el número de puntos que se toman, y finalmente, se grafican diversos gráficos de espectros de frecuencias con su respectivo rango.

Sarria et al. 2016. Describen la construcción de un sistema computacional con el objetivo de la creación de una base de conocimiento para caracterizar el género de música salsa a través de información acústica e inteligencia artificial. Recopilarón 24.108 canciones de colecciones privadas, a las cuales se les aplicó análisis de señales de audio, contando con las siguientes características: espectro de potencia (power spectrum), magnitud de espectro (magnitude spectrum), variabilidad de espectro (spectral variability), centroide espectral (spectral centroid), compacidad (compactness), fluencia espectral (spectral flux), Coeficientes cepstrales de frecuencia de tipo mel (mel-frequency cepstral coefficients o MFCC), raíz cuadrática media (root mean square o RMS) y codificación lineal predictiva (linear predictive coding o LPC). Se evaluarón diversas implementaciones de algoritmos de inteligencia artificial y se opto por la mas adecuada, para después analizar e interpretar los resultados.

Cruz et al. 2019. Estudiarón diversos métodos de aprendizaje automático para identificar el género de la música colombiana. Tomarón 180 piezas musicales de seis géneros musicales folclóricos colombianos como: Bambuco, Carranga, Cumbia, Joropo, Pasillo y Vallenato, a las cuales se les realiza un proceso de extracción de características de canciones, para posteriormente realizar el entrenamiento de modelos de aprendizaje automático. Se seleccionarón métodos de clasificación supervisados (Random Forest y Redes neuronales), y no supervisados (Clustering con ayuda de algoritmos basos en k-means).

Rodríguez 2021. Estudió un modelo de deep learning para el reconocimiento automático de géneros musicales a partir de una muestra de audio. Fueron seleccionados dos audios por cinco géneros musicales distintos como: Pop, Música clásica, Rock, Metal y Reggae, posterior a esto, se tomaron fragmentos de audio de 30 segundos, a los cuales se les aplicó una transformación que se realiza a partir de un frame work de Matlab denominado Audio Toolbox, obteniendo así espectrogramas. Una vez generados los espectogramas, se diseña, analiza y entrena una red neuronal convolucional, analizando resultados y predicciones.

3.2. Antecedentes estadísticos

En la revisión de literatura encontramos que el uso de análisis espectral es utilizado en diversas areas, como lo son: economía, geología, climatología, etc. Finalmente se realiza una descripción general de artículos y trabajos de grado.

Mesa & Carvajal 1994. Ilustran 2 metodologías para así realizar una predicción de lluvias y caudales, aplicadas a diversas cuencas del departamento de Antioquia, así como caudales en el territorio nacional. El Análisis Espectral Singular (AES) y las redes neuronales son las empleadas en este estudio, siendo la primera metodología útil en el filtrado de señales, de manera que se toma la componente no lineal dentro de las series temporales naturales para así estudiar el conjunto de valores dentro del sistema, teniendo en cuenta la utilización de métodos tradicionales de predicción en modelos autorregresivos AR(P), y con la segunda metodología, es posible estudiar fenómenos no lineales, considerando información de precipitación, caudales, manchas solares, temperatura superficial del mar (TSM), el Índice de Oscilación del Sur (IOS) y datos de la oscilación cuasi-bienal del viento (OQB).

García & Moreno 2006. Describen 3 metodologías para el análisis de información estratigráfica: Análisis de Fourier (AF), donde se asume que a través de la sumatoria de funciones periódicas, es posible determinar las frecuencias principales; Análisis Espectral Singular (AES), esta técnica se encuentra relacionada con la separación de la señal estratigráfica en componentes independientes y Redes Neuronales Artificiales (RNA), permiten extraer componentes principales no lineales de la señal estratigráfica, permitiendo la combinación con el clasificador de señal múltiples con el fin de determinar las frecuencias

predominantes en la señal estratigráfica. Cada uno de los métodos son aplicados a un caso sintético. Para la metodología AES se presenta el procedimiento tradicional (GHIL et al. 2002), y se propone una variación basada en redes neuronales.

Con el fin de aplicar las metodologías a datos reales, se emplearon herramientas como Matlab en el caso del AF, AES tradicional y su respectiva variación con redes neuronales, sin embargo, para las redes neuronales en combinación con el clasificador de señal múltiple se hizo uso de un lenguaje de programación Fortan 90, el cual es utilizado con facilidad en matemáticas y cálculos científicos.

Para su aplicación consideraron la definición de parámetros especificados por el usuario, detallando los más importantes de la siguiente manera:

- Tamaños de ventana o dimensión de segmentación
- Número de componentes principales
- Tasas de aprendizaje
- Valores óptimos de la activación
- Errores mínimos de entrenamiento

Se concluyó que del análisis espectral singular y de las redes neuronales se obtienen componentes asociados a frecuencias particulares, y que estas últimas en conjunto con el clasificador de señal múltiple detectan las frecuencias predominantes en una señal estratigráfica.

Echeverry et al. 2007. Realizan un análisis armónico de señales de MicroElectrodos de Registro (MER) aplicado a la enfermedad de Parkinson, mediante análisis del periodograma empleando el método de Welch, con el objetivo del reconocimiento de características de las diferentes zonas cerebrales. Tomaron pruebas sobre 66 señales de las diferentes zonas, etiquetando las señales por neurofisiólogos y neurocirujanos, donde 25 corresponden al núcleo subtalámico (STN) y 41 a la substancia nigra (SN), obteniendo frecuencias de muestreo de 24000Hz por medio de procedimientos quirúrgicos en pacientes de Parkinson, se calcula la densidad espectral de potencia mediante el método propuesto logrando suavizar el espectro de la señal y disminuir la varianza entre estimaciones.

Carvalho et al. 2011. Presentan un análisis económico, donde el objetivo es separar en diversas componentes el indicador económico PIB por medio del método AES (Análisis Espectral Singular). El método empleado incluye dos principales fases, la descomposición y la reconstrucción, donde la primera se subdivide en los pasos de incrustación y descomposición

de valores singulares, por otro lado, la fase de reconstrucción, incluye la agrupación y el promedio diagonal (Golyandina et al., 2001). Tuvieron en cuenta los periodos de contracción fechados por el Buró Nacional de Investigación Económica (NBER) y se aplicó un ejercicio en tiempo real, obteniendo un indicador compuesto por 7 componentes relevantes para el análisis llamado crr-filter y se compara con los indicadores producidos por los métodos de Christiano–Fitzgerald; Baxter–King; and Hodrick–Prescott, concluyendo que el indicador resultante caracteriza la dinamina cíclica de la actividad económica, arrojando estimaciones más bajas, y mayor rendimiento que otros.

Pascual & Traver 2014. Desarrollan y presentan la aplicabilidad del Análisis de Componentes Principales (PCA) en el área de operaciones de valor relativo sobre la curva de rendimientos. El PCA se llevó a cabo sobre las curvas de tipos de interés, reduciendo el análisis a tres principales factores, como el nivel, la pendiente y la curvatura, explicando la variabilidad de la Estructura Temporal de Tipos de Interés (ETTI). Por último, se analiza e implementan estrategias de inversión en curvatura no correlacionas con otros factores de riesgo.

4. Marco Teórico

En esta sección se expone el marco teórico de la investigación. En la primera parte se presenta el marco contextual, en el cual se incluyen algunas definiciones y conceptos importantes que ayudarán en la contextualización del análisis músical. En la segunda parte se presentan los temas o definiciones estadísticas, que aportan como teoría y herramientas con las cuales se buscaría resolver el problema planteado.

4.1. Marco teórico contextual

Sonido

El sonido puede definirse como la vibración del aire generada por la oscilación de alguna fuente sonora, ocasionando que las partículas de aire sean inducidas a compresiones y rarefacciones de forma cíclica. Cabe resaltar que su concepto no sólo se asocia al movimiento de las partículas y variación en la presión del aire, sino que, se le atribuyen características como, la fase, la duración del sonido, al amplitud de la variación en la presión del aire y la frecuencia de amplitud (Sueur 2018).

Música

La música, es una de las manifestaciones artísticas provenientes del arte. Al igual que las otras manifestaciones artísticas, la música cuenta con referencias o coordenadas espacio – temporales; es decir, cuenta con información relacionada al contexto en que fue desarrollada (Castro 2003).

Además, la música es un conjunto organizado de sonidos con propósitos estéticos y que además puede comunicar al creador de la música con sus oyentes. (Godt 2005). La música; estudiada como señales; puede verse como un conjunto de características acústicas y estructurales; como el tono o las armonías (Muller et al. 2011).

Género musical

El género musical es la manera más conocida de agrupar, clasificar o describir piezas musicales, actuando como una dimensión de la música misma, como el tempo, el timbre o el idioma de la canción (Pachet & Aucouturier 2003).

Los géneros musicales se encargan de caracterizar las similitudes existentes entre grupos de piezas musicales, cantantes y compositores; además de ser usados para organizar la música en catálogos y librerías (Scaringella & Zoia (2006) Scaringella et al. 2006).

Cumbia

La cumbia, como género musical, es un fenómeno latino - caribeño nacido en Colombia que se ha extendido por todo el continente americano, mutando continuamente de acuerdo a su contexto, conservando sus raíces mestizas, su expresión coreográfica y musical, producto de la confluencia histórica entre africanos, quienes contribuyeron con el ritmo y los tambores; indígenas, quienes aportaron con la caña de millo y la gaita, y el vestuario que es de origen español (Márquez 2016).

El significado del género reside en el testimonio de lo que el hombre negro realizó para conquistar a la mujer indígena, dando paso a una nueva generación. La cumbia; originalmente; es aquella en la que se emplean la tambora, el tambor alegre, el llamador, el guache o maracas, la flauta de millo y dos tipos diferentes de gaita; sin embargo, distintas localidades de Colombia; y en toda latinoamérica; adaptaron su propia versión de cumbia, agregando otros instrumentos o cambiando los originales, como el clarinete en reemplazo de las gaitas en la cumbia sabanera (Ochoa 2016).

Archivo de audio

Los archivos de audio, son una representación digital de señales de sonido usando números y que son almacenados en algún medio digital como CD's o computadoras (Fries & Fries 2005). Los archivos de audio cuentan con algunas características:

- **Frecuencia de muestreo:** Es un conjunto de valores que son asignados a cada muestra de la señal del sonido y que son tomados en un lapso temporal determinado. La frecuenca de muestreo se mide en Hertz (Hz) y a mayor cantidad de muestras, mejor es la calidad de audio. Normalmente, en los CD's de música, se usa una frecuencia de muestreo de 44.1 kHz (Stockdale (2002) Stockdale 2002).
- **Formato:** Es el esquema de compresión del audio en un archivo. Existen en una gran variedad, cada uno con distintas características dependiendo del uso que se le vaya a dar al archivo de audio y a la calidad que se necesita para este. Entre los más populares está el formato MPEG 1 Layer 3 (MP3), el formato Advanced Audio Coding (AAC) y el formato de onda Waveform Audio File Format (WAV) (Stockdale (2002) Stockdale 2002).
- **Profundidad de bits:** Es la amplitud de bits o el número de valores que se le puede

asignar a cada muestra de la señal del sonido. A mayor amplitud, mejor calidad de audio. Generalmente, los archivos de audio son codificados con valores entre 8 y 16 bits (Stockdale (2002) Stockdale 2002).

- **Canales de audio:** Estos indican en el archivo de audio la cantidad de señales que se están reproduciendo de forma simultánea, siendo frecuente el tener uno (mono) o dos canales (estereo) (Stockdale (2002) Stockdale 2002).

Waveform Audio File Format (WAV)

El formato de onda WAV, es un formato de audio de Windows muy popular, con el que se puede digitalizar directamente la información del audio en un archivo sin necesidad de comprimirla; es decir, mantiene toda la información original del audio en el archivo (Santoso & Hilman 2018).

En todos los casos, primero debemos encontrar una manera de pasar las muestras de datos de audio voluminosas, a una representación adecuada del contenido de audio. Lo anterior, se realiza a través de la extracción de características, lo cual mantiene una relevancia en la aplicación de aprendizaje automático y reconocimiento de patrones. De forma que, una de las diversas formas de análisis de características puede evidenciarse en el dominio visual donde las representaciones de tiempo-frecuencia suelen presentar tendencias (Yu & Slotine (2009) Yu, Guoshen. Slotine, Jean-Jacques 2009).

Representaciones gráficas

Existen formas de representar gráficamente algunas de las características de las señales de sonido, como son:

- **Oscilogramas:** Permiten visualizar la intensidad o la amplitud de la señal del sonido en lapsos cortos de tiempo, representándola con una forma de onda; es decir, es un gráfico de la amplitud de la señal del sonido contra el tiempo (Iriarte et al. 2010).
- **Periodogramas:** Muestran la variación de la densidad espectral; calculada a través de algún método como la transformada de Fourier; con respecto a su frecuencia (Bloomfield 2000).

Densidad espectral

La densidad espectral de potencia o PSD por sus siglas en inglés para "Power Spectral Density", es una medida que describe cómo se distribuye la potencia de una señal en componentes de frecuencia (Bisina & Azeez 2017).
Aritméticamente, la densidad espectral de potencia se puede describir como el cuadrado del espectro de frecuencia de la transformación de Fourier discreta o DFT; por "Discrete Fourier Transformation" (Sueur 2018).

Frecuencia fundamental de un sonido

La frecuencia fundamental es la frecuencia que determina la nota musical o el tono principal que se percibe en el sonido. Cuando se escucha un sonido complejo, como una nota musical o la voz humana, dicho sonido se compone de múltiples componentes de frecuencia o armonías que conforman su espectro; la frecuencia fundamental es la más baja de estas componentes y define la base tonal.

Para determinar la frecuencia fundamental de un sonido proveniente de un archivo de audio, se utilizan técnicas de análisis espectral y algoritmos de detección de pitch o tono en la señal del sonido, como la transformada rápida de Fourier y el método de Welch respectivamente.

4.2. Marco teórico estadístico

Series de tiempo

Una serie de tiempo $\{x_t, t \in T\}$ es una realización de un proceso estocástico, donde T es una indexación en el tiempo, para cualquier variable aleatoria, $x_t, t \in T$, están definidas en muestras espaciadas.

Definición 1 *si $x_t, t \in T$ es una serie de tiempo, entonces para t, t+k, para todo t y $k \in T$, se define:*

a. La función de autocovarianza, $\gamma(k)$ por:

$$\gamma(x_t, x_{t+k}) = E\{(x_t - \mu_t)(x_{t+k} - \mu_t + k)\}$$

b. La función de autocorrelación, $\rho(k)$ por:

$$\rho(x_t, x_{t+k}) = \frac{\gamma(x_t, x_{t+k})}{\sigma_t \sigma_{t+k}}$$

Transformada rápida de Fourier

Definición 2. Dada una serie de tiempo x_t con $t \in T$ que puede estar descrita en función del dominio periódico y teniendo en cuenta que dicho proceso también puede representarse en el dominio de la frecuencia x_w, es decir, dos representaciones de la misma función; la transformada de Fourier permite transformar señales tanto en el dominio del tiempo o de la frecuencia, según se requiera.

Para hacer uso de esta herramienta en la práctica es necesario conocer la transformada discreta de Fourier de una serie de tiempo $x(nT)$, la cual se encuentra definida como:

$$X\left(\frac{k}{NT}\right) = \sum_{n=0}^{N-1} x\left(nT\right) e^{\frac{-j2\pi nk}{N}} \quad ; \quad k = 0, ..., N-1$$

Métodos no paramétricos o libres de distribución

Los métodos estadísticos no paramétricos o de libre distribución, son aquellos que no necesitan el cumplimiento de asunciones o supuestos; generalmente relacionados a la distribución o al tamaño de los datos; para su correcto uso, desarrollo o validación (Pértega & Fernández 2016). A pesar de ser métodos no tan estrictos como los paramétricos; no están excentos de forma absoluta al cumplimiento de supuestos de otro tipo distinto a los relacionados con la distribución de los datos.

Análisis espectral singular (SSA)

El análisis espectral singular o SSA; por sus siglas del inglés "Singular Spectrum Analysis"; es un modelo no paramétrico aplicado a series temporales. Se desarrolla asumiendo que la serie temporal se puede representar en componentes que, a su vez, resumen la información contenida en dicha serie temporal (Zhigljavsky 2020).

El SSA se usa con fines de extraer componentes de tendencia, suavizar las series temporales mediante su reconstrucción, extraer componentes periódicas como los ciclos o la estacionalidad, interpretar la serie temporal a través de su separación en componentes más simples y ruido, entre otras.

La metodología del análisis espectral singular se lleva a cabo en cuatro etapas o pasos:

1. **Incrustración** es donde se define el tamaño de ventana (L) con el que se construye la llamada matriz de trayectorias perteneciente al espacio R^L

2. **Descomposición** descomposición en valores singulares de la matriz de trayectoria para obtener L valores y vectores propios

3. **Agrupación** de un número r de vectores propios en una matriz perteneciente al subespacio L_r del espacio R^L y su proyección en el mismo subespacio.

4. Consiste en el **promedio diagonal** de la matriz de proyecciones mencionada anteriormente, obteniéndose así una matriz de Hankel cuyos valores son una aproximación a la serie original o alguna de sus componentes.

Algunos de los conceptos a tener en cuenta dentro del análisis espectral singular, son los siguientes:

- **Relación de recurrencia lineal:** Consiste en considerar que los valores de una serie de tiempo, $X = \{ x_1, x_2, x_3, ..., x_N \}$ con N finito, se establecen como:

  $$X_n = \sum_{j=1}^{t} a_j x_{n-j}$$

  Es decir, los valores de la serie temporal son resultado de una combinación lineal de sus valores pasados. Esta característica es de vital importancia al momento de desarrollar un análisis espectral singular bajo un paradigma paramétrico.

- **Separabilidad de componentes:** Establece que una serie de tiempo $X = \{ x_1, x_2, x_3, ..., x_N \}$, puede representarse como la suma de; inicialmente; dos series temporales

  $$X1 = \{ x1_1, x1_2, x1_3, ..., x1_N \} \text{ y } X2 = \{ x2_1, x2_2, x2_3, ..., x2_N \}; \text{ con } N \text{ finito}$$

  Escrito de otra forma, los valores de la serie temporal X se pueden representar como

  $$x_i = x1_i + x2_i, \text{ con } i = 1, 2, ..., N; \text{ con } N \text{ finito}$$

  El análisis espectral singular, pretende descomponer o separar las series temporales originales en otras series temporales denominadas como componentes.

- **Tamaño de ventana:** El tamaño o longitud de ventana; denotado por L; es uno de los parámetros principales a tener en cuenta para el desarrollo del análisis singular espectral. Este parámetro representa el tamaño con el que se construye la matriz de trayectoria dentro del primer paso de ejecución; el de **incrustación**.

  Se sugiere que el valor de L se encuentre en el intervalo $2 \leq L \leq N - 1$, donde N es el número de datos que componen la serie de tiempo original. Es bueno que el tamaño de ventana sea lo suficientemente grande para que los vectores rezagados que componen la matriz de trayectoria, conserven buena información sobre el comportamiento de la serie temporal original, dando la posibilidad de pensar en cada vector como una serie temporal individual y analizar dinámicas particulares que son más tediosas de ver a través de la serie temporal original completa.

- **Tripletas propias (Eigentriple):** Las triplas o tripletas propias son un conjunto conformado por elementos resultantes del paso dos, de **descomposición**, en el análisis espectral.

  Sean λ_1, λ_2, ..., λ_L y V_1, V_2, ..., V_L los valores y vectores propios; respectivamente; obtenidos de la realización de la descomposición en valores singulares a la matriz de trayectoria.

  Sean también vectores V, tal que

  $$V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$$

  Por tanto las componentes de la matriz de trayectoria se pueden obtener como

  $$X_i = \sqrt{\lambda_i} U_i V_i^T$$

  Entonces una tripleta propia es aquella compuesta como ($\sqrt{\lambda_i}$, U_i, V_i); es decir, la compone un valor propio $\sqrt{\lambda_i}$, un vector propio o de carga U_i y un vector de trayectoria V_i. Estas tripletas propias son importantes para la selección de grupos en el paso de agrupación.

Análisis espectral singular multivariado (MSSA)

El análisis espectral singular multivariado, es una extensión del análisis espectral singular, con la capacidad de desarrollarse en múltiples series temporales de forma simultánea y no solo en una. Procedimentalmente, en términos generales, no hay diferencia con su contraparte simple, exceptuando algunas consideraciones al momento de construir la matriz de trayectoria en el paso de incrustación (Zhigljavsky 2020).

En la versión original, se construye una matriz de trayectoria con dimensiones L x K, donde $K = N - L + 1$; es decir, se compone por K columnas de vectores rezagados de longitud L. Por otra parte, en la versión multivariada, la construcción de la matriz de trayectoria se desarrolla con normalidad para las series temporales, obteniéndose una matriz por cada una, a la que se le denotará como X_i, con $i = 1, 2, ..., M$, donde M es el número de series temporales. Luego, dichas matrices pueden organizarse de dos formas: apilándolas de manera horizontal o de manera vertical; la elección de una u otra afecta superficialmente el procedimiento, donde se debe tener en cuenta que el apilamiento horizontal no exige un mismo tamaño para todas las series temporales, mientras que en el vertical es preferible que si se cumpla. Al final, se obtiene una matriz de trayectoria con dimensiones L x $(K)(M)$ para el caso de apilamiento horizontal, o $(L)(M)$ x K para el caso de apilamiento vertical (Zhigljavsky 2020).

Método de Welch

En resumen, el método de Welch es una técnica comúnmente utilizada para estimar el espectro de potencia de una señal. Se basa en dividir la señal en segmentos solapados, calcular la transformada rápida de Fourier de cada segmento y promediar los periodogramas resultantes para obtener una estimación del espectro de potencia. La ecuación para calcular el periodograma promediado es:

$$\hat{P}_w(f_k) = \frac{1}{m}\sum_{m=1}^{M}\hat{P}_m(f_k) \qquad 0 \leq f_k \leq f_s/2$$

Al dividir la señal de audio en segmentos solapados de longitud fija, estos se superponen para evitar discontinuidades en el análisis. El tamaño de los segmentos determina la resolución en frecuencia del resultado; es decir, entre más particiones haya en la señal, más valores se obtienen para la frecuencia fundamental.

Antes de calcular la transformada rápida de Fourier de cada segmento de la señal del sonido, se aplica una función de ventana para minimizar los efectos de las discontinuidades en los bordes del segmento. Luego, a cada partición, se le aplica una transformada rápida de Fourier (FFT) para obtener el espectro de frecuencia correspondiente, a partir del cual se calcula el espectro de potencia. El espectro de potencia se obtiene tomando el cuadrado del valor absoluto del espectro de frecuencia.

Una vez que se tiene el espectro de potencia de cada segmento, se promedian para obtener una estimación del espectro de potencia final. Esto se hace promediando los valores correspondientes de cada frecuencia. A partir del espectro de potencia promediado, se puede encontrar la frecuencia fundamental buscando el valor máximo en el espectro. Este valor máximo representa la frecuencia dominante en la señal de audio.

Las siguientes descripciones se presentan usando como base la literatura presentada en el libro “Time series analysis and its applications : with R examples”, de Robert Shumway y David Stoffer (Shumway & Stoffer 2017), donde exhiben cómo se desarrolla el análisis de componentes temporales teniendo en cuenta el tiempo; ideal para series temporales.

Análisis de Componentes Principales

El Análisis de Componentes Principales (ACP); también conocido como PCA por sus siglas en inglés para “Principal Component Analysis”; es una técnica de análisis multivariante utilizada para explorar y resumir la información contenida en un conjunto de datos. El objetivo principal del ACP es reducir la dimensionalidad de los datos mientras se mantiene la mayor parte de su variabilidad.

Para desarrollar un análisis de componentes principales primero se calcula la matriz de covarianzas o la matriz de correlación de los datos; dependiendo de si las variables están

medidas en diferentes escalas o no. Los datos deben estar debidamente estructurados en una tabla y con valores numéricos.

El ACP se desarrolla bajo el supuesto que para un conjunto de datos con variables $x = (x_1, x_2, x_3, ..., x_p)$, su varianza puede ser explicada a través de la combinación lineal de las componentes de x. Teniendo:

$$y = c'x = c_1x_1 + c_2x_2 + ... + c_px_p$$

Se busca encontrar un valor para c; dado que $c'c = 1$; tal que $var(y)$ sea lo más grande posible. La expresión que resume lo anterior es:

$$var(y) = c'\Sigma_{xx}c$$

Donde Σ_{xx} es la matriz p_xp de varianzas y covarianzas de x. La cuestión, es buscar $c \neq 0$ tal que:

$$max\frac{c'\Sigma_{xx}c}{c'c}$$

Para dar respuesta a lo anterior, primero se debe denotar la tuplas conformados por los valores y vectores propios de la matriz de covarianza Σ_{xx} como $(\lambda_1, e_1), (\lambda_2, e_2), ..., (\lambda_p, e_p)$; donde para cada valor propio $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ... \geq \lambda_p \geq 0$, y los vectores propios tienen longitud unitaria.

La solución para un valor de c que maximize la varianza de y, es asignar el valor e_1; es decir, $c = e_1$; obteniéndose la combinación lineal $y = e_1'x$, con varianza máxima $var(y) = \lambda_1$. La combinación lineal $y = e_1'x$ es la primera compononente principal y; siguiente la misma idea; la combinación lineal $y = e_2'x$ es la segunda componente principal, sujeta a que $cov(y_1, y_2) = 0$.

En general, el k-ésimo componente principal; para $k = 1, 2, ..., p$; es la combinación lineal, tal que $c'c = 1$ y $cov(y_k, y_j) = 0$; para cualquier valor de $j \neq k$. Entonces, la manera de maximizar la varianza de la componente, es tomar $c = e_k$, tal que $var(y_k) = \lambda_k$. De este resultado se tiene que la varianza total de x es:

$$var(x) = var(x_1) + var(x_2) + ... + var(x_p) = \sigma_{11} + \sigma_{22} + ... + \sigma_{pp}$$

Donde σ_{jj} es el j-ésimo elemento de la diagonal de la matriz Σ_{xx}. Por lo que $var(x)$ se puede escribir como $tr(\Sigma_{xx})$. Con lo anterior en cuenta, la proporción de varianza total atribuida a la k-ésima componente principal está dada por la expresión:

$$\frac{var(y_k)}{tr(\Sigma_{xx})} = \frac{\lambda_k}{\sum_{j=1}^{p}\lambda_j}$$

Aunque el análisis de componentes principales es una técnica poderosa y ampliamente utilizada, trabaja bajo algunos supuestos, como el de asumir que los datos son independientes entre sí, lo que significa que las observaciones no están correlacionadas temporalmente, debido a que es probable que los resultados no sean interpretables o útiles, ya que la estructura temporal de los datos puede llevar a que los componentes principales no representen adecuadamente la variabilidad de los datos o que las relaciones temporales importantes se pierdan.

Dado lo anterior, no se puede desarrollar un análisis de componentes principales clásico con el tipo de datos que se propone en esta investigación; sin embargo, existen variaciones que permiten desarrollar un análisis de componentes principales a series temporales multivariadas, cambiando la matriz de covarianza mencionada anteriormente, por la matriz de densidad espectral.

Matriz Hermitiana de densidad espectral

La matriz de densidad espectral, es una matriz Hermitiana o autoadjunta; es decir, es una matriz cuadrada cuyos elementos son equivalentes a los de la misma matriz, pero conjugada y transpuesta. Esta matriz proporciona información sobre la distribución de la energía de una señal en el dominio de la frecuencia.

La matriz de densidad espectral se obtiene mediante la implementación de la transformada de Fourier en la función de autocovarianza de la señal. Sea $x_t = (x_{t1}, x_{t2}, ..., x_{tp})'$ un proceso estacionario de dimensiones $p_x 1$, cuya matriz de autocovarianza está dada por:

$$\Gamma(h) = E[(x_{t+h} - \mu)(x_t - \mu)'] = \{\gamma_{jk}(h)\}$$

Se cumple que:

$$\sum_{h=-\infty}^{\infty} |\gamma_{jk}(h)| \leq \infty$$

Para todo $j, k = 1, 2, ..., p$. Entonces, la matriz de autocovarianza $\Gamma(h)$ se puede representar como:

$$\Gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i w h} f(w)\, dw$$

Con $h = 0, \pm 1, \pm 2, ...$ Sea la transformada inversa de densidad espectral $f(w) = \{\gamma_{jk}(w)\}$, para $j, k = 1, 2, ..., p$, se puede escribir $f(w)$ como:

$$f(w) = \sum_{h=-\infty}^{\infty} \Gamma(h) e^{2\pi i w h}$$

Con $-\frac{1}{2} \leq w \leq \frac{1}{2}$. Dando lugar a la matriz hermitiana de densidad espectral. Para el caso de series temporales, se asume que estas son estacionarias, con media 0 y que están integradas por valores complejos, pudiéndose representar de la forma $x_t = x_{1t} - ix_{2t}$; donde x_{1t} es la parte real y x_{2t} la imaginaria. La función de autocovarianza de x_t de dimensiones $p_x p$ se define a través de la expresión:

$$\Gamma_{xx}(h) = E(x_{t+h} x_t^*) - E(x_{t+h}) E(x_t^*)$$

Obedeciendo condiciones similares a las del caso anterior con valores reales. Teniendo:

$$\Gamma_{xx}(h) = \gamma_{ij}(h)$$

Para valores de $i, j = 1, 2, ..., p$. Cumpliéndose que:

- $\gamma_{ij}(0) \geq 0$ es real.
- $|\gamma_{ij}(h)|^2 \leq \gamma_{ii}(0)\gamma_{jj}(0)$ para todo h.
- $\Gamma_{xx}(h)$ es una función no negativa.

Si $\Sigma_h ||\Gamma_{xx}(h)|| < \infty$, la matriz de densidad espectral de las series x_t está dada por la expresión:

$$f_{xx}(w) = \sum_{h=-\infty}^{\infty} \Gamma(h) e^{2\pi i w h}$$

Obteniéndose una expresión análoga al caso anterior. Este procedimiento se aplica a cada una de las series temporales de frecuencias fundamentales extraídas de las muestras de las señales de las canciones; sin embargo, debido a su distribución a través del tiempo, estas deben ser centradas primero, realizando la diferencia entre cada valor de la serie con su promedio.

Con el paquete "astsa" del software de computación estadística y gráfica R, se pueden calcular las matrices de densidad espectral asociadas a las series temporales de las frecuencias fundamentales, haciendo uso de la función "mvspec()", que; a través de su extensión "$fxxrealiza los cálculos relacionados a la expresiones de matrices de autocovarianza y de densidad espectral presentadas anteriormente, obteniéndose un conjunto de m matrices de densidad expectral de dimensiones 30x30 (Stoffer & Poison, 2023).

Análisis de Componentes Principales para series temporales

Análogamente al caso clasico, el análisis de componentes principales para series temporales tiene como objetivo el descomponer cada serie en una combinación lineal de componentes principales, donde cada componente representa una fuente de variabilidad. Estos componentes se ordenan en función de su contribución a la variabilidad total de las series.

Sea x_t un proceso vectorial estacionario de media 0, con una matriz de densidad espectral de dimensiones $p_x p$ denotada f_{xx}, la cual es una matriz Hermitiana de valores complejos no negativa. Ahora, se supone que para un valor fijo de w se quiere encontrar un proceso univariado tal que:

$$y_t(w) = c(w)^* x_t$$

Donde $c(w)$ maximiza la densidad espectral de $y_t(w)$ en la frecuencia w; con $c(w)^* c(w) = 1$. Como en la frecuencia w, la densidad espectral de $y_t(w) = c(w)^* f_{xx} c(w)$, entonces se busca el valor complejo $c(w) \neq 0$, tal que:

$$max \frac{c(w)^* f_{xx}(w) c(w)}{c(w)^* c(w)}$$

Para dar solución a lo anterior, se denotan la tuplas conformadas por valores y vectores propios de la matriz de densidad espectral $f_{xx}(w)$ como $(\lambda_1(w), e_1(w)), \lambda_2(w), e_2(w)), ..., \lambda_p(w), e_p(w))$; donde para cada valor propio $\lambda_1(w) \geq \lambda_2(w) \geq ... \geq \lambda_p \geq 0$, y los vectores propios son de longitud unitaria.

Los valores propios de la matriz f_{xx} son reales, por lo que la solución es hacer $c(w) = e_1(w)$, haciendo:

$$max \frac{c(w)^* f_{xx}(w) c(w)}{c(w)^* c(w)} = \frac{e_1(w)^* f_{xx}(w) e_1(w)}{e_1(w)^* e_1(w)}$$

Por lo que la primera componente principal en la frecuencia w estaría dada por la combinación lineal:

$$y_t(w) = e_1(w)^* x_t$$

En general, el k-ésimo componente principal en la frecuencia w; para $k = 1, 2, 3, ..., p$ son las series temporales de valores complejos $y_{t1}(w) = e_k(w)^* x_t$, y la densidad espectral de $y_{tk}(w)$ en la frecuencia w es:

$$f_{yk}(w) = e_k(w)^* f_{xx} e_k(w) = \lambda_k(w)$$

5. Metodología

En esta sección se presenta la propuesta metodológica utilizada para llevar a cabo el análisis espectral de canciones del genero Cumbia Colombiana, enfocándose en la extracción de la frecuencia fundamental de los registros musicales. Esta frecuencia se emplea como insumo principal para el desarrollo de un análisis de componentes principales de series temporales con el que se obtienen componentes interpretables y con la capacidad de resumir las frecuencias de todas las pistas, caracterizando su registro tonal. Todo se realiza con la finalidad de dar solución a la pregunta de investigación y alcance a los objetivos propuestos.

El objetivo de este trabajo es analizar las características espectrales de las canciones de cumbia colombiana, con el fin de identificar patrones, elementos musicales y componentes frecuenciales distintivas presentes en este género. Para lograrlo, se emplearon técnicas de procesamiento de señales y análisis de datos, usando el software para computación estadística y gráfica R.

En una primera fase se realizaron consultas a diversas fuentes de datos con el fin de obtener y recopilar archivos de audio del género cumbia colombiana, de acuerdo a su disponibilidad. Posteriormente, se evaluó la calidad de los ficheros multimedia para seleccionar las canciones integrantes de la muestra a analizar, tratando que haya homogeneidad en sus características y considerando criterios como el formato, la tasa de muestreo y la tasa de compresión (Anders & Anders 2017). En esta fase también se realiza un análisis preliminar de las canciones de la muestra, a través de la revisión de sus oscilogramas examinando algunas de sus características acústicas (Iriarte et al. 2010).

La segunda fase de la metodología corresponde a la obtención de las series temporales de interés. De la muestra seleccionada se extraen series temporales con datos relacionados a la frecuencia fundamental de los registros musicales utilizando un algoritmo de detección de pitch para estimar la frecuencia fundamental de cada segmento de la canción, permitiendo así obtener una secuencia de valores a lo largo del tiempo (Potamitis & Rigakis 2016). Además, se generan gráficos de las series temporales y se calculan sus periodogramas para visualizar y analizar diferentes aspectos de su registro sonoro como la frecuencia, la periodicidad y la energía (Iriarte et al. 2010).

En la tercera fase, se desarrolla el análisis de componentes principales de las series temporales, descomponiéndolas en componentes independientes e interpretables que resumen

los elementos que las forman e identificando patrones y estructuras subyacentes en las canciones de cumbia (Shumway & Stoffer 2017).

Finalmente; en la cuarta fase; se realiza una interpretación de los resultados obtenidos en el análisis de componentes principales, relacionándolos con los aspectos musicales característicos de la cumbia, buscando comprender cómo se combinan e interactúan estos elementos sonoros en el género.

Esta propuesta metodológica pretende proporcionar una comprensión más profunda de las características y estructuras presentes en el género cumbia colombiana, ofreciendo una perspectiva novedosa sobre este estilo musical. Los resultados obtenidos pueden ser de gran relevancia tanto para la investigación musical como para la producción y el análisis de música relacionada a la cumbia colombiana. La **Figura** $5-1$ muestra el resumen de las actividades realizadas por fases.

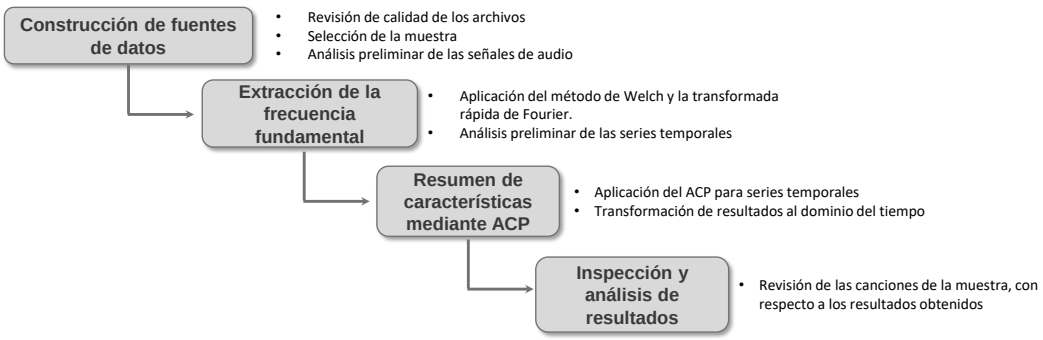

Figura 5-1.: Esquema de metodología y actividades por fases

5.1. Contrucción de fuentes de datos

La fuente de datos utilizada consiste en un conjunto no estructurado, compuesto por canciones pertenecientes al género cumbia colombiana. Estas canciones fueron seleccionadas de acuerdo a su disponibilidad; es decir, se usaron pistas pertenecientes a una recopilación de álbumes de cumbia colombiana, donde se encuentran varias canciones populares y representativas del género.

Cada canción en la muestra es representada como un archivo de audio digital, caracterizado por una serie de atributos que son tenidos en cuenta para el desarrollo de este trabajo; tales como su formato de compresión o su tasa de muestreo. Estos archivos son de calidad óptima y tienen una duración de alrededor de cuatro minutos; tratándose de las pistas completas de las canciones; suficiente para permitir un análisis detallado y preciso.

Además de los archivos de audio, se tiene información complementaria sobre cada canción en la muestra, incluyendo metadatos como el título de la canción, el nombre del artista o banda, el álbum al que pertenece y el año de lanzamiento. Estos metadatos pueden ayudar a organizar y contextualizar los resultados del análisis de componentes principales.

Se pretende que la población de interés para esta metodología sea la compuesta por las canciones de música del género cumbia colombiana integrantes de la muestra seleccionada. Por lo que la investigación se desarrolla bajo un enfoque exploratorio, cuyos resultados son solo aplicables a la muestra en sí, pero ofreciendo; mediante esta propuesta metodológica; la oportunidad de explorar y descubrir nuevas perspectivas, tendencias y características únicas de este género.

Una vez seleccionada la muestra de pistas musicales, se lleva a cabo la extracción de la frecuencia fundamental de cada canción, que es la frecuencia relacionada al registro tonal que hace parte de las pistas y que será usada para el desarrollo metodológico de esta investigación. Con la frecuencia fundamental de cada canción como insumo, se realiza una descomposición en componentes principales.

5.1.1. Revisión de calidad de los archivos

En total se cuenta con 128 pistas musicales de diferentes álbumes de intérpretes de cumbia colombiana; sin embargo, algunas canciones no pertenecen al género y solo hacen parte del proyecto musical; además que estás cuentan con distintos atributos; por lo que es necesario evaluar las cualidades de cada fichero para seleccionar una muestra de canciones homogénea y adecuada para la investigación.

Para seleccionar las pistas de la muestra, se evalúa la calidad de las canciones recopiladas,

considerando aspectos relacionados a la aptitud de los archivos de audio; tal y como es sugerido por Anders & Anders (2017). Inicialmente, se revisan las propiedades de cada uno de los ficheros, comprobando que posean características técnicas similares. Debido a la variedad de géneros existentes entre las canciones de los álbumes, se realiza una revisión de las pistas para filtrar solo aquellas pertenecientes al género cumbia colombiana. También es importante que el formato, la tasa de muestreo, los canales, la profundidad de bits y la tasa de compresión sean similares entre las canciones.

Es importante tener en cuenta estos aspectos debido a que, por ejemplo, de acuerdo al formato un archivo de audio podría contener más o menos información que otros. Algo similar pasa con las otras características a evaluar; tanto una tasa de compresión mayor, como una de muestreo, apunta a mayor información contenida en el archivo y; por consiguiente; mayor tamaño al momento de comprimir en un determinado formato. En cuanto a los canales; generalmente; cada canal corresponde a una fuente de sonido separada en una grabación o reproducción; por lo que, a mayor cantidad de canales, más información contiene el archivo. Por último, cuanto mayor sea la profundidad de bits, mayor será la precisión y la calidad de audio, traduciéndose en mayor información contenida. Como se puede notar, la cantidad de información almacenada en un fichero se refleja directamente en la calidad del audio y en el tamaño de este, por lo que es importante que las canciones cuenten con atributos similares para poder obtener resultados adecuados (Stockdale 2002).

Al revisar los ficheros, se tiene que el formato de todos los archivos de audio es MP3, con dos canales (estéreo), una profundidad de bits de 16 bits y tasas de muestreo de 44,1 kHz; que es lo mismo que decir que se toman 44100 muestras por segundo. Por otro lado, las tasas de compresión tenían un rango aproximado comprendido entre 110 kpps y 320 kpps. Al final, de 128 archivos de audio se agrupan 50 del género cumbia colombiana, entre los cuales hay 30 canciones con tasa de compresión entre los 110 kpps y 130 kpps, y el resto con 320 kpps.

5.1.2. Selección de la muestra

La selección del tamaño del conjunto de canciones que componen la muestra a analizar, se decide teniendo en cuenta el material multimedia disponible, esto debido a las restricciones existentes de derechos de autor y a la falta de bases de datos aprovechables, traduciéndose en limitaciones en los recursos para adquirir o recopilar las canciones. También se tiene en consideración que, para una investigación con un enfoque exploratorio como esta; tal y como hace mención Clarke 2000 sobre su estudio exploratorio; un tamaño muestral pequeño puede ser útil como un primer paso preliminar en el estudio de la cumbia colombiana, proporcionando una base inicial para identificar tendencias, patrones y características principales que pueden ser exploradas en estudios futuros con tamaños muestrales más grandes.

Teniendo en cuenta lo anterior, de las 128 canciones recopiladas de entre todos los álbumes, solo cerca de 50 son del género cumbia, pues; como se hizo mención anteriormente; los álbumes incluyen canciones de otros géneros dentro de su repertorio. En un segundo filtro, la muestra se recorta a 30 canciones para facilitar el procesamiento de la misma a través del software de computación estadística y gráfico R y tener una muestra más homogénea en cuanto a la información contenida dentro de los archivos de audio. Estas canciones corresponden a aquellas con tasa de compresión entre 110 kpps y 130 kpps, provenientes de álbumes recopilatorios de canciones populares del género, variando en intérpretes, subgéneros, álbumes y años de publicación. La tabla de metadatos correspondiente a la muestra de canciones se presenta en los anexos, en la **Tabla** $A - 1$.

5.1.3. Análisis preliminar de las señales de audio

Una forma común de realizar un análisis preliminar de las señales de audio es a través de oscilogramas, los cuales representan gráficamente el sonido en función del tiempo, permitiendo visualizar la amplitud de la señal en diferentes momentos y otras características como su duración y la presencia de eventos o patrones repetitivos. Para obtener el oscilograma de una canción, se toma la muestra de la señal de audio y se ubica en el eje vertical de un plano, mientras que en el eje horizontal se ubica el tiempo en segundos. La muestra de la señal contiene las amplitudes de la onda del sonido secuencialmente a través del tiempo y su cantidad está definida por la tasa de muestreo con la que se comprimió el archivo de audio. La gráfica resultante es el oscilograma, representando la amplitud de la onda en función del tiempo. Generalmente, el eje vertical se mantiene sin escala ni unidad, debido a que la amplitud que se obtiene de la señal del sonido, podría referirse a distintas cantidades relacionadas a la onda; como la presión o el voltaje; cantidades que suelen medirse en distintas unidades. Por otra parte, no se coloca la escala ya que se desconocen las condiciones de grabación y compresión del audio, traduciéndose en que no hay información sobre el sistema de calibración del mismo, y no se puede realizar una apropiada conversión de los valores de la amplitud (Sueur 2018).

Usando el software de computación estadística y gráfica de R y el paquete "tuneR" (Ligges et al., 2023), se convierte el formato MP3 de las canciones seleccionadas a WAV; o de tipo onda; a través de la función "extractWave()". Esto se logra debido a que la función lee el archivo de audio, extrayendo su información, para luego codificarla en formato WAV. Este cambio se hace dado que muchas de las funciones relacionadas a la manipulación de archivos de sonido y ondas, son aplicadas a través de cargar elementos o ficheros en formato WAV, debido a la calidad y la cantidad de información que se conserva en los audios con este formato, además de permitir manipular dicha información con mucha más versatilidad (Sueur 2022).

A continuación, se usa la función "mono()" para mezclar los canales estéreo de las canciones en uno solo, promediando los valores de las señales de audio del canal izquierdo y del derecho. Esto se hace debido a que, para el análisis, es más práctico tener una sola serie temporal por canción, ya que muchas funciones no aceptan múltiples canales dentro de sus parámetros de ejecución (Sueur 2022).

Con el paquete "seewave" (Sueur, 2022), se usa la función "oscillo()" para obtener los gráficos de los oscilogramas de las señales de audio. Esta función recibe los objetos tipo onda creados con el paquete "tuneR".

5.2. Extracción de la frecuencia fundamental

La propuesta presentada en esta subsección para la extracción de la frecuencia fundamental, está basada en la realizada por Potamitis & Rigakis 2016, en su estudio sobre mosquitos, donde exhiben una manera de analizar el sonido de su aleteo a través de la extracción de la frecuencia fundamental asociada.

5.2.1. Frecuencia fundamental de un sonido

La frecuencia fundamental se refiere a la frecuencia que determina la nota musical o el tono principal que se percibe en el sonido. Las canciones son sonidos complejos; al igual que una nota musical o la voz humana; que se componen de múltiples componentes de frecuencia o armonías que conforman su espectro; la frecuencia fundamental es la más baja de estas componentes y define la base tonal. Generalmente se denota por f_0 (Sueur 2018).

Para determinar la frecuencia fundamental de un sonido proveniente de un archivo de audio se utilizan técnicas de análisis espectral y algoritmos de detección de pitch o tono en la señal del sonido (Doval & Rodet 1991). Como propuesta metodológica basada en el estudio de Potamitis & Rigakis 2016, se implementa el método del Welch; donde interviene la transformada rápida de Fourier; como técnica para extraer la frecuencia fundamental, f_0, asociadas a las canciones de la muestra.

5.2.2. Método de Welch

El método de Welch es una técnica donde se estima el espectro de potencia de una señal, dividiéndola en segmentos a los cuales se les aplica la transformada de rápida de Fourier,

promediando lo espectros resultantes para obtener una estimación más suave (Potamitis & Rigakis 2016)

De acuerdo con Ahmad et al. 2010 en su estudio donde compararon distintos métodos de estimación de la frecuencia dominante, una forma óptima de aplicar el método de Welch, es usando un solapamiento de 2048 puntos por segmento, ejecutando la transformada rápida de fourier en 4096 puntos con una superposición del 50 %, y con una función de ventana de Hamming. De esta manera se puede capturar una cantidad significativa de información que permita un mejor y más detallado análisis espectral.

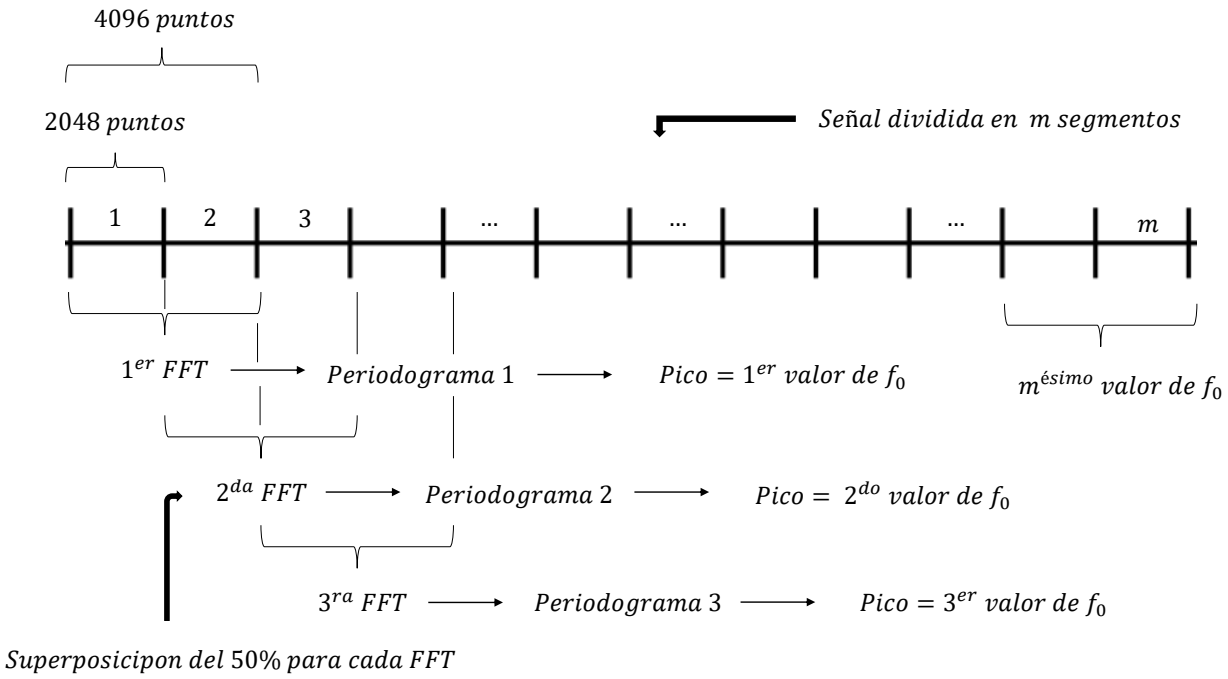

Figura 5-2.: Representación de la segmentación de la señal realizada en el método de Welch.

La señal de una de las canciones de la muestra se denota como $a_t = a_1, a_2, a_3, ..., a_p$; donde p es el tamaño o longitud de la señal. Teniendo en cuenta los parámetros definidos anteriormente, la señal es separada en m segmentos de tamaño 2048 cada uno. A cada segmento se le aplica la función de ventana, para luego ejecutar la transformada rápida de Fourier por pares consecutivos; es decir, se ejecuta el algoritmo cada 2048 datos sobre 4096 datos, superponiendo la mitad de ellos cada vez; véase **Figura** $5-2$. Con lo anterior se obtiene una estimación del periodograma para cada segmento.

Como se hizo mención antes, este método es usado para suavizar el periodograma de una señal; sin embargo, para propósitos de la investigación, no es necesario realizar el promedio de los periodogramas para obtener la frecuencia fundamental. Para esto, basta con identificar la frecuencia con mayor intensidad para cada una de las estimaciones de los resultados de aplicar la transformada rápida de Fourier; es decir, la frecuencia fundamental será aquella donde haya un valor máximo o pico (Potamitis & Rigakis 2016). Se puede notar que se obtiene un valor de frecuencia fundamental por cada segmento solapado de la señal; en otras palabras, se obtienen m valores de frecuencia fundamental por señal.

5.2.3. Extracción de la frecuencia fundamental en R

Usando el paquete "tuneR" (Ligges et al., 2023), del software de computación estadística y gráfica R, se obtienen los objetos tipo onda que se usaron como insumo para la generación de oscilogramas en el **Apartado 5.1.3**.

Como primer paso para obtener la frecuencia fundamental asociada a las señales, se usa la función "periodogram()" con los elementos de tipo onda de las canciones. Esta función desarrolla el método de Welch tal y como se ha explicado anteriormente, usando los parámetros mencionados.

Finalmente, se estima la frecuencia fundamental en Hertz de las señales usando; del mismo paquete; la función "FF()" sobre los resultados obtenidos de la función "periodogram()", hallando la frecuencia dominante para cada segmento de la señal y teniendo como base de escala tonal el diapasón A440 o de 440 Hertz, del sistema de notación musical occidental (Ligges et al., 2023).

Estas frecuencias son agrupadas en una tabla de datos usando las funciones del paquete base de R "data.frame()" y "cbind()", de tal manera que haya una columna por cada canción con su respectiva frecuencia fundamental. Como se extrae un número distinto de valores asociados a la frecuencia fundamental para cada canción, las dimensiones finales de la tabla están dadas por la longitud del vector asociado a la frecuencia fundamental con más valores; es decir, la canción más larga. Con la función de base "ts()", se declara la tabla de datos elaborada como serie temporal multivariada.

5.2.4. Análisis preliminar de las series temporales

Como primer acercamiento a las series temporales obtenidas anteriormente, se revisa su comportamiento de forma visual. Los gráficos permiten visibilizar patrones y tendencias en los datos a lo largo del tiempo, posibilitando identificar comportamientos recurrentes

o cambios en su trayectoria; además de facilitar el reconocer patrones estacionales, ciclos, tendencias, y otros fenómenos relevantes. Para elaborar un gráfico que refleje el comportamiento de una serie temporal, se toman sus valores y se ubican en el eje vertical de un plano, mientras que las unidades temporales están en el eje horizontal (Shumway & Stoffer 2017).

La estacionariedad con media cero en las series es un requisito para el desarrollo de técnicas como el análisis de componentes principales para series temporales, por lo que es importante procurar que se cumplan estas características (Shumway & Stoffer 2017). Se verifica que las series asociadas a la frecuencia fundamental de las canciones son estacionarias a través del test de raíz unitaria de Dickey - Fuller, y al no tener media cero; debido a que sus valores no oscilan a su alrededor; deben ser centradas, calculándole a cada una su valor promedio y restándoselo a cada frecuencia fundamental de su respectiva serie.

Una vez corregido la estacionariedad de las series, graficar sus periodogramas, es otra forma de obtener información preliminar de las mismas. Para estimar el periodograma de una serie temporal se usa la transformada de Fourier, asociando cada resultado a un valor de frecuencia $w_j = \frac{j}{n}$, definida en ciclos por unidad de tiempo. El gráfico se elabora ubicando los valores del periodograma en el eje vertical del plano, mientras que las frecuencias se ubican en el eje horizontal. El periodograma permite contemplar la distribución de energía en diferentes frecuencias dentro de la serie temporal, facilitando visualizar aquellas que son dominantes o la presencia de componentes frecuenciales, ayudando a detectar patrones periódicos o estacionales en los datos (Shumway & Stoffer 2017).

Shumway & Stoffer (2017), sugieren la estimación del periodograma mediante el uso de la expresión:

$$I(w_j) = \frac{|d(w_j)|^2}{n}$$

Donde $d(w_j)$ es el resultado de la transformada rápida de Fourier y n es el número de datos de la serie temporal.

Usando el paquete base del software de computación estadística y gráfica de R, se usa la función plot() para graficar las series temporales de las canciones, tomando como insumo las series en sí. Por otro lado, a través del paquete "stats", se usa la función "mvfft()" con la serie multivariada generada anteriormente; conformada por las series temporales asociados a la frecuencia fundamental de cada canción. Esta función ejecuta la transformada rápida de Fourier en cada columna o serie temporal. Luego se calcula el periodograma a través del cuadrado del valor absoluto de los resultados obtenidos, para finalmente dividirlos por el número total de datos, n. Con la función "plot()" del paquete base de R, se grafica el periodograma de cada serie, usando el anterior resultado calculado y una secuencia discreta de datos $j = 1, 2, ..., n$; correspondiente a los ciclos en $w_j = \frac{j}{n}$. Para elaborar el gráfico en

términos de las frecuencias w_j, es suficiente con dividir la secuencia discreta mencionada anteriormente entre el número de datos, n, y ejecutar de la misma forma la función "plot()".

5.3. Resumen de características mediante ACP para series temporales

El desarrollo de esta subsección está basado en la literatura expuesta por Shumway & Stoffer 2017, donde explican cómo se desarrolla el análisis de componentes principales para series temporales, ofreciendo teorías, definiciones y conceptos que se usaron como insumos para el planteamiento de esta metodología.

5.3.1. Análisis de Componentes Principales para series temporales

El análisis de componentes principales es utilizado para explorar y resumir la información contenida en un conjunto de datos, reduciendo su dimensionalidad mientras se mantiene la mayor parte de la variabilidad. Análogamente, el análisis de componentes principales para series temporales tiene como objetivo el resumir o caracterizar un conjunto de series temporales, descomponiendo cada una en una combinación lineal de componentes principales (Shumway & Stoffer 2017).

Teniendo la serie temporal multivariante $x_t = x_1, x_2, ..., x_30$; correspondiente a los n valores asociados a la frecuencia fundamental extraída de cada canción; se le calcula una matriz de densidad espectral f_{xx} de dimensiones 30_x30 por cada valor de frecuencia $w = \frac{j}{n}$; con $j = 1, 2, 3, ..., \frac{n}{2}$.

Shumway & Stoffer 2017, sugieren la estimación de la matriz de densidad espectral mediante el uso de la expresión:

$$\hat{f}_{xx}(w_j) = \sum_{m}^{\ell=-m} h_\ell I_n(w_j + \frac{\ell}{n})$$

Donde $I_n(w_j)$ es la matriz de periodograma suavizado. De estas matrices se obtienen 30 valores $(\lambda_k(w))$ y vectores propios $(e_k(w))$ los cuales se utilizan para maximizar la densidad espectral de las combinaciones lineales $y_{tk}(w) = c(w)^* x_t$; correspondientes a las componentes principales en un determinado w.

Se calcula la primera componente principal y_{t1} en la frecuencia w, haciendo $c(w) = e_1(w)$;

es decir, la combinación lineal queda como $y_{t1}(w) = e_1(w)^* x_t$; donde $e_1(w)$ es el primer vector propio. El primer valor propio $\lambda_1(w)$, representa la densidad espectral de la primera componente principal y_{t1}.

Luego, se obtiene una estimación del espectro de la primera componente y_{t1}, calculando los primeros valores propios $\lambda_1(w)$ de la matriz de densidad espectral $f_{xx}(w)$; donde $w = \frac{j}{n}$, para cada $j = 1, 2, 3, ..., \frac{n}{2}$. Esta estimación se grafica ubicando en el eje vertical del plano los valores propios $\lambda_1(w)$ y en eje horizontal los ciclos j o las frecuencias w. Con el gráfico, se analiza la frecuencia que está asociada a un pico; que se traduce en la frecuencia asociada al valor máximo entre los valores propios. Esta frecuencia da información sobre la periodicidad dominante entre las series temporales analizadas.

Con el paquete .ªstsa"del software de computación estadística y gráfica R, se puede calcular la matriz de densidad espectral asociada a la serie temporal multivariada; con los valores de frecuencia fundamental de las canciones; haciendo uso de la función "mvspec()", que; a través de su extensión "$fxx"; realiza los cálculos relacionados a las estimaciones de la matriz de densidad espectral para cada una de las frecuencias w, obteniéndose un arreglo tridimensional de dimensiones $30_x 30_x \frac{n}{2}$; es decir, $\frac{n}{2}$ matrices de dimensiones $30_x 30$ (Stoffer & Poison, 2023).

Con la función "eigen()" del paquete base de R, se calculan los valores y vectores propios asociados a las estimaciones de la matriz de densidad espectral para cada frecuencia w, obtenidas anteriormente. Con la función "plot()" se grafica el espectro de la primera componente principal, extrayendo el primer valor propio de cada matriz de densidad espectral de las diferentes frecuencias y usando una secuencia discreta de números $j = 1, 2, ..., \frac{j}{n}$. También se grafica en el dominio de las frecuencias, dividiendo la secuencia discreta generada previamente entre el número de datos, n, y ejecutar de la misma forma la función "plot()".

5.3.2. Transformación de resultados en el dominio de las frecuencias al dominio del tiempo

Los resultados que se han obtenido en el anterior apartado, están en función de la frecuencia w. Sin embargo, estos mismos resultados pueden transformarse para ser presentados en el dominio del tiempo (Shumway & Stoffer). Por la naturaleza de la investigación, los resultados son más versátiles si se desarrollan en función del tiempo, debido a la facilidad de inspección que ofrece.

Teniendo en cuenta que para cada canción se tiene una señal de 44100 datos por segundo que, posteriormente, es segmentada en particiones de 2048 datos para obtener un solo valor de la frecuencia fundamental por partición; entonces existe la relación de que hay 21,53 datos

de frecuencia fundamental por segundo, que es lo mismo a decir que se tiene una tasa de muestreo de $\frac{1}{21{,}53}$. Este resultado se denota como Δ.

Se transforman las frecuencias $w = \frac{j}{n}$; medida en ciclos por segundo; interpretándola como una fracción del ciclo por unidad de tiempo; por lo tanto, se necesitan $\frac{1}{w}$ fracciones para completar un solo ciclo de la serie temporal. Por lo que para obtener el tiempo en segundos al que corresponde el ciclo, se busca la cantidad de Δ necesarios para obtener un ciclo; es decir:

$$\frac{1}{w} * \Delta$$

El resultado de la anterior expresión indica cada cuántas unidades de tiempo; en este caso, segundos; hay un patrón cíclico o periódico en las series temporales.

5.4. Inspección y análisis de resultados

Una vez obtenida la frecuencia asociada al valor propio máximo representada en la estimación del espectro de la primera componente principal y que esta haya sido transformada al dominio del tiempo, se inspeccionan las canciones para verificar la existencia de patrones cíclicos en el tiempo encontrado. Usando los metadatos y la capacidad auditiva para identificar cambios en el sonido; como en los instrumentos, acordes o escalas tonales; se pretende justificar el resultado encontrado sobre el patrón cíclico que caracteriza al conjunto de canciones.

5.5. Software

Como se puede notar de anteriores secciones, se utilizó el software de computación estadística y gráfica R para la mayoría de actividades mencionadas en esta propuesta metodológica, entre las que se destaca la generación de gráficos, el procesamiento de las señales, la adecuación de los datos, la extracción de la frecuencia fundamental y caracterización a través de análisis de componentes principales. Se usaron varios paquetes de funciones, principalmente la librería "tuneR" (Ligges et al., 2023) para la manipulación de los archivos de audio, la librería "seewave" (Sueur, 2022) para la generación de gráficos, la librería "astsa" (Stoffer & Poison, 2023) para el desarrollo del análisis de componentes principales de series temporales y en gran medida la librería base y "stats" para desarrollar actividades más generales; como elaborar gráficos y calcular valores. La versión del software es la 4.2.1.

6. Resultados

En esta sección, se presentan los resultados obtenidos de realizar las actividades propuestas en la metodología, dando una respuesta a la pregunta de investigación y cumpliendo con los objetivos propuestos.

En un primer acercamiento, se presentan los análisis gráficos preliminares de las señales de las canciones y de las series temporales que se generaron a partir de ellas; correspondientes a los valores de frecuencia fundamental de cada canción. Luego se muestran los resultados obtenidos de ejecutar un análisis de componentes principales sobre las series temporales y su respectiva explicación para posteriormente interpretar sus valores en el dominio del tiempo. Por último, se exponen los hallazgos de realizar una inspección rigurosa de los resultados obtenidos del análisis de componentes principales sobre las canciones de la muestra.

6.1. Análisis preliminar de las señales de audio

Una vez desarrollado el proceso de recolección de canciones y su posterior selección, se obtiene un conjunto de 30 canciones al que se le realiza un análisis visual a través de graficar la señal de cada canción, con respecto al tiempo; en segundos.

En la **Figura** $6-1$, se muestran los oscilogramas asociados a las señales de las primeras tres canciones de la muestra. Esto es debido a la dificultad de apreciación de detalles en los gráficos al agrupar varios en una sola ventana y a que los comentarios realizados aplican a la muestra completa, con sus 30 canciones. Las demás representaciones gráficas se encuentran en la sección de anexos, **Figura** $A-1$, **Figura** $A-2$, **Figura** $A-3$, **Figura** $A-4$ y **Figura** $A-5$.

Como se puede observar en la **Figura** $6-1$, los oscilogramas de las 30 canciones cuentan con una duración media de 200 segundos aproximadamente. Visualmente son notables las variaciones positivas y negativas de las ondas musicales de forma aleatoria en las que; de acuerdo con Juan Byron (2012); se podría intuir la presencia de sonidos aperiódicos. Sin embargo, hay ondas caracterizadas por líneas verticales representadas de extremo a extremo, las cuales son interpretables como indicios de sonidos periódicos. Un patrón distintivo en

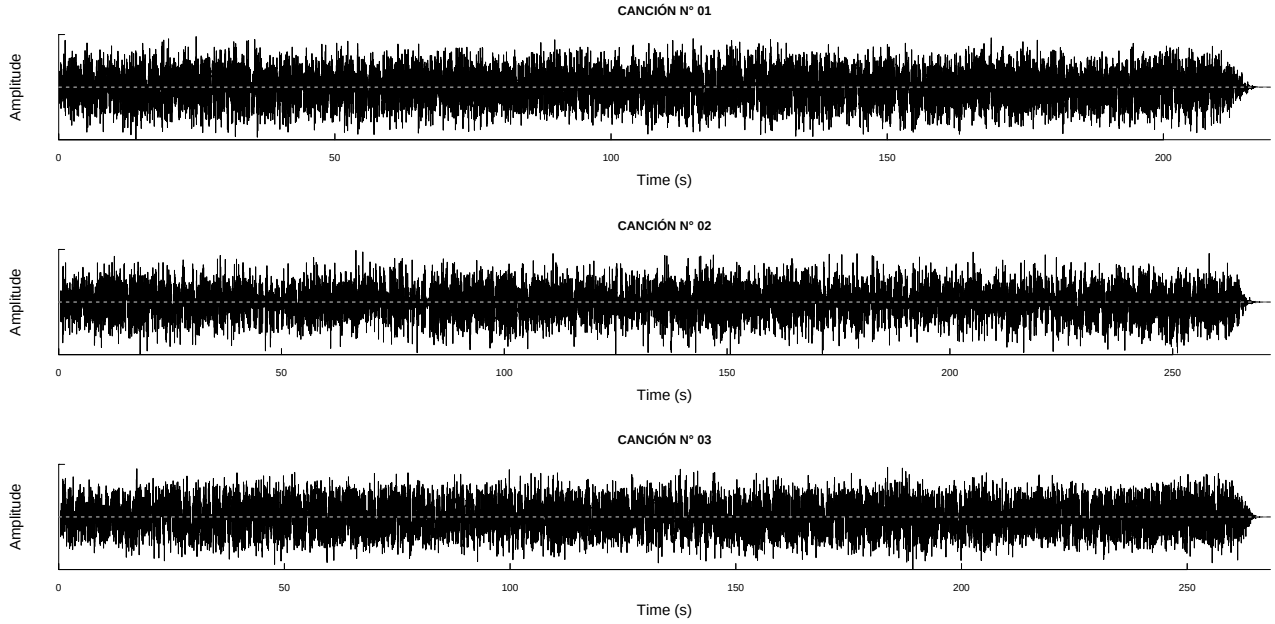

Figura 6-1.: Oscilogramas de las primeras tres canciones de la muestra, generados en R. El eje verticual corresponde a la amplitud de la onda, mientras que el eje horizontal corresponde al tiempo medido en segundos.

algunos oscilogramas son los lapsus de ausencia de energía durante un tiempo determinado, de forma que es interpretable que la presión del aire no intervenga con intensidad.

6.2. Análisis preliminar de las series temporales

Tras la extracción de los valores de frecuencia fundamental asociados a las pistas de las canciones y su declaración como serie temporal multivariada, se realiza una revisión gráfica del comportamiento de cada una de las series temporales que la conforman.

En la **Figura** $6-2$, se muestran las representaciones gráficas de las primeras tres series temporales con los valores de frecuencia fundamental de cada una. Al igual que con los oscilogramas, en la sección de anexos, **Figura** $A-6$, **Figura** $A-7$, **Figura** $A-8$, **Figura** $A-9$ y **Figura** $A-10$, se encuentran los gráficos asociados a las demás series temporales.

A grandes rasgos, en las representaciones gráficas de las series temporales no parece que existan patrones o características destacables a lo largo del tiempo. Sus comportamientos son irregulares, sin tendencias pronunciadas o ciclos que se repitan frecuentemente de forma clara

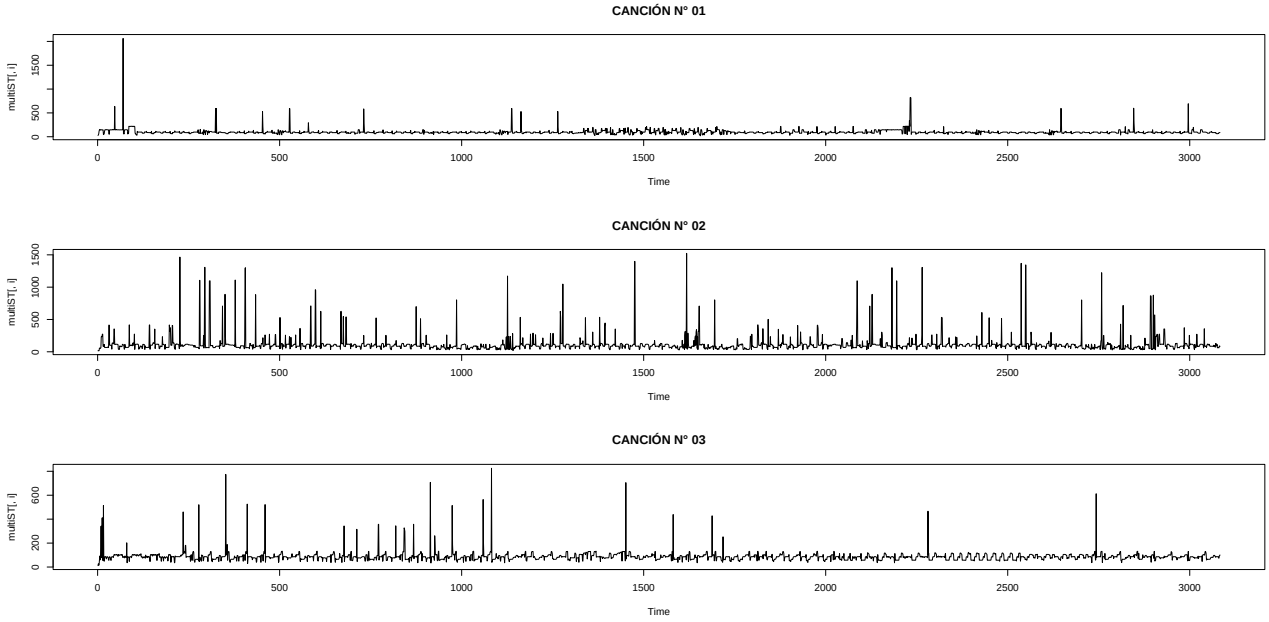

Figura 6-2.: Representación gráfica de las series temporales asociadas a las primeras tres canciones. En el eje vertical se encuentran los valores de frecuencia fundamental, mientras que en el horizontal las unidades temporales. Cada dato está asociado a 1/21.53 de segundo.

(Wei, 2006). Se puede notar la existencia de valores extremos a lo largo sus trayectorias, pero estos no parecen ser periódicos; de acuerdo con Ligges (2023), es posible que sean producto del exceso de ruido en algunos instantes de las canciones.

Por otro lado, al revisar de forma más minuciosa las series, se encuentran pequeños patrones ocasionales a lo largo del tiempo. Por ejemplo, la **Figura** $6-3$ muestra en detalle la serie temporal de la canción tres entre los valores 1750 y 2500 de unidades de tiempo, donde se puede observar un ligero comportamiento estacional cada cierto intervalo de valores. Esto puede deberse a solos instrumentales o vocales, donde el ruido es mínimo y los tonos o notas musicales son claras. Al igual que con la canción tres, este fenómeno ocurre en todas las demás series temporales en diferentes medidas.

Como parte del análisis preliminar, se estima el periodograma para cada serie temporal centrada y se grafica; tal y como se muestra en la **Figura** $6-4$, donde se exhiben los periodogramas de las ocho primeras series temporales. Los demás gráficos relacionados al resto de series temporales se encuentran en anexos, **Figura** $A-11$, **Figura** $A-12$, **Figura** $A-13$ y **Figura** $A-14$. A través de los periodogramas se puede apoyar los

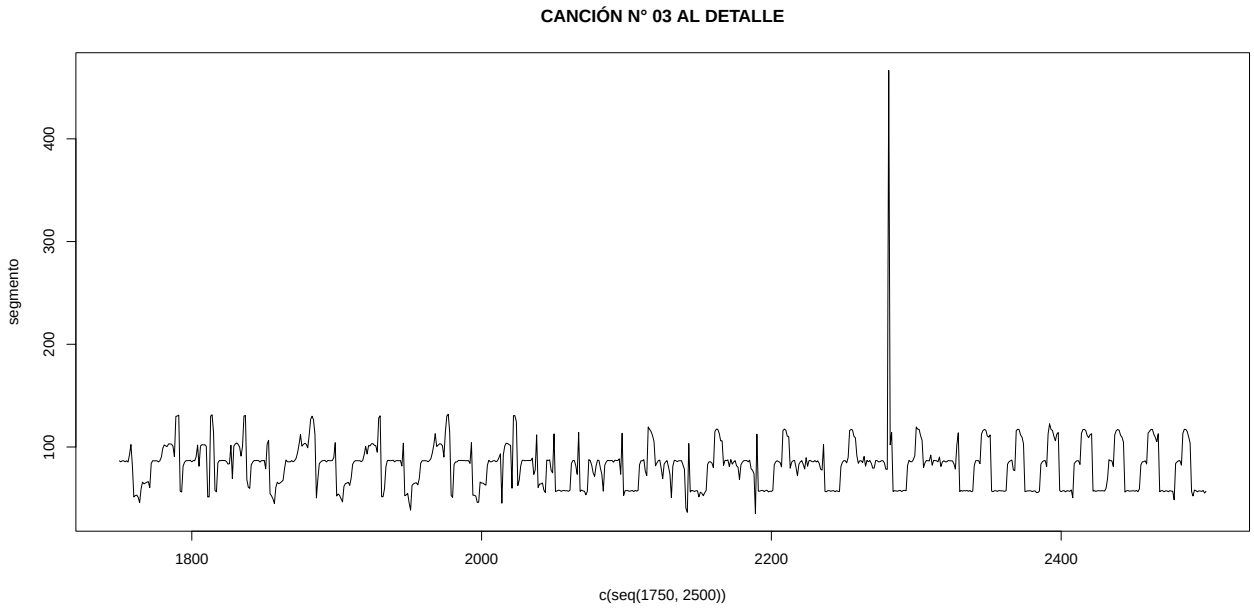

Figura 6-3.: Representación gráfica detallada de las periodicidades presentes en la serie temporal asociada a la tercera canción, con valores temporales entre 1750 y 2500.

hallazgos encontrados anteriormente sobre la existencia de periodicidades y ciclos en las series temporales (Shumway & Stoffer, 2017).

Como se observa, en todos los periodogramas la distribución de la energía es distinta, con picos diferentes, variando entre las frecuencias 0.0006 y 0.1047. Estos picos reflejan las frecuencias dominantes y sirven de indicio para reconocer periodicidades en las que hay mayor intensidad de energía o poder espectral (Shumway & Stoffer, 2017). Tomando de ejemplo el periodograma asociado a la primera canción, su pico se encuentra en el valor de frecuencia w=0.0022 con un poder espectral mucho menor al pico del periodograma de la segunda canción, el cual está en w=0.0729. Teniendo en cuenta lo anterior, no se visualiza un pico común entre todas las canciones, aunque hay canciones que comparten niveles similares de intensidad o poder espectral; como en el caso de la canción tres y la siete.

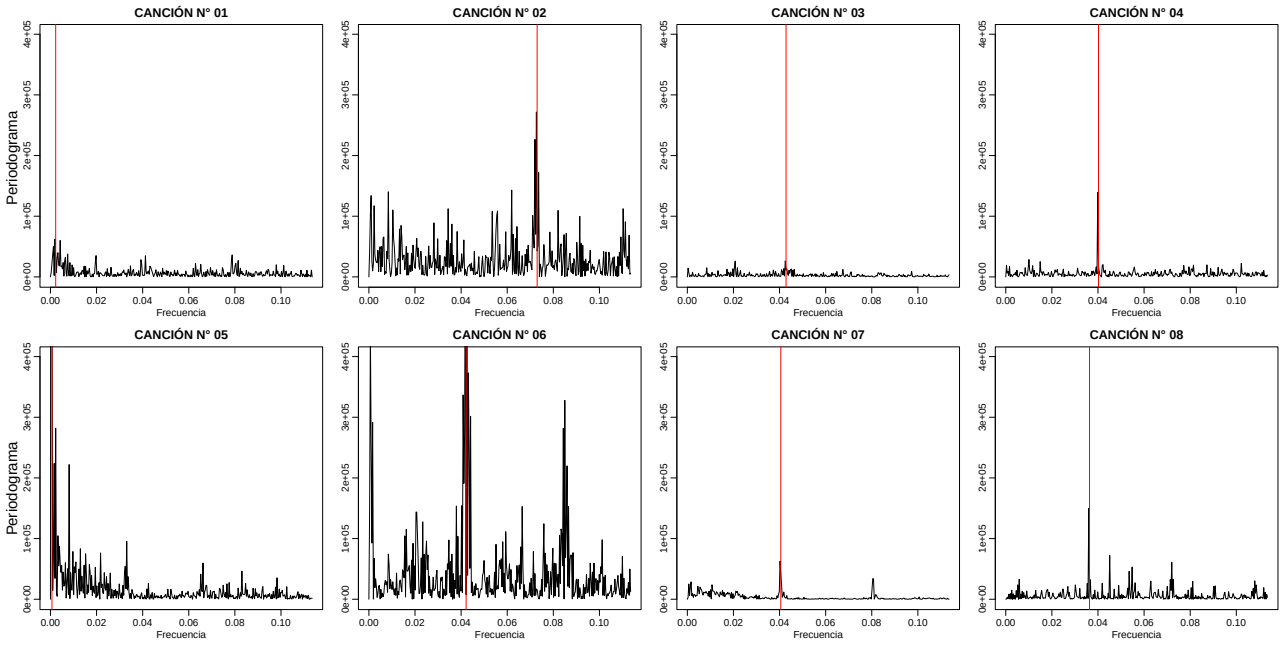

Figura 6-4.: Periodogramas de las series temporales asociadas a las primeras ocho canciones. En el eje vertical se encuentra la estimación del periodograma, mientras que en el eje horizontal están las frecuencias $w = \frac{j}{n}$, medida en ciclos por unidad de tiempo. Las líneas verticales rojas indican la frecuencia donde se genera el pico máximo de poder espectral, siendo para cada uno de los ocho periodogramaslos valores 0.0022, 0.0729, 0.0428, 0.0402, 0.0006, 0.0421, 0.0405 y 0.0363, respectivamente.

6.3. Análisis de la estimación del espectro de la primera componente principal

Con las matrices de densidad espectral calculadas por cada valor de frecuencia w y los valores $(\lambda_1(w))$ y vectores propios $(e_1(w))$ asociados, se construye el gráfico de la estimación del espectro de la primera componente principal y_{t1}, con la que se pretende resumir y caracterizar las series temporales relacionadas a la frecuencia fundamental de las canciones.

La **Figura** $6-5$ muestra el gráfico de la estimación del espectro de la primera componente principal, el cual se encuentra limitado horizontalmente hasta un valor de frecuencia de $w =$ 0,1135, debido a que valores de frecuencia superiores en el dominio del tiempo representan cantindades muy pequeñas en milisegundos, dificultando el análisis a través del oído humano; esto se examina con mayor detenimiento en subsecciones posteriores.

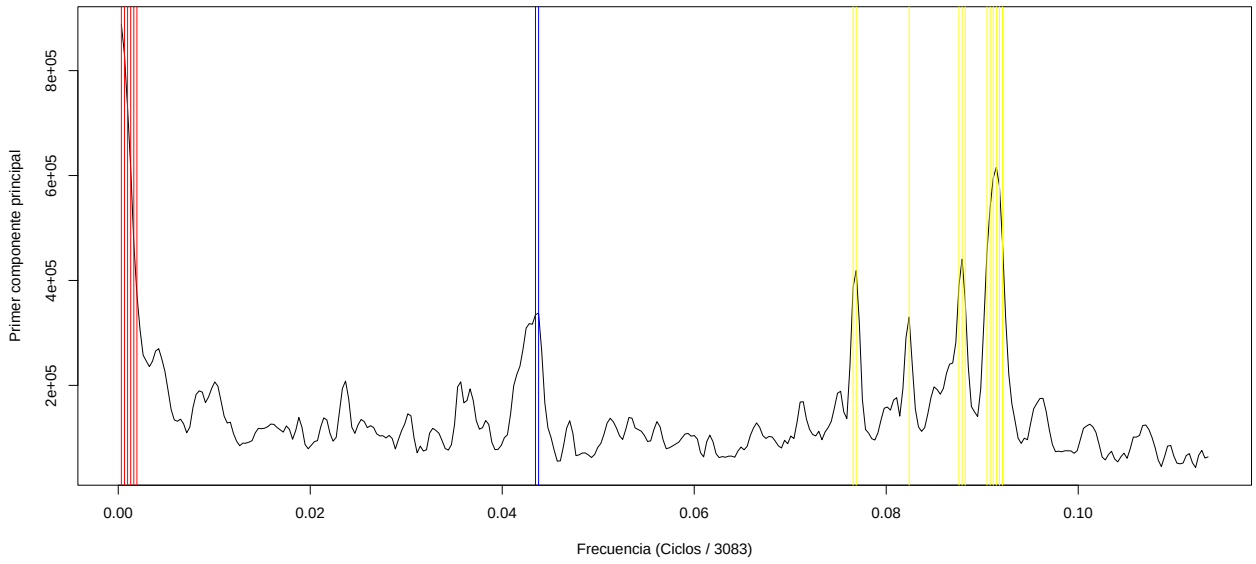

Figura 6-5.: Estimación del espectro de la primera componente principal, donde se señalan las frecuencias asociadas a los primeros veinte máximos, agrupadas por cercanía usando colores: rojo, azul y amarillo.

En la **Figura** $6-5$ se observa que existen secciones del gráfico destacadas por tener picos muy juntos en determinados valores de frecuencia, indicando que en ellos se acumula una buena cantidad de poder espectral. Al hallar las frecuencias asociadas a los primeros veinte valores máximos de la estimación del espectro, estas son agrupadas; de acuerdo a su proximidad; en tres grupos distintos, distinguidos por los colores rojo, azul y amarillo en la **Figura** $6-5$; los cuales en adelante denotaremos por $G1$ (Rojo), $G2$ (Azul) y $G3$ (Amarillo).

El color rojo representa las frecuencias más cercanas a cero. La existencia de valores máximos en estas frecuencias podría deberse a la presencia de valores atípicos en las series temporales que se usaron para la estimación, por lo que hay mucha incertidumbre alrededor de estos valores. El color azul y amarillo, represetan las frecuencias que se encuentran aledaños a 0,04 y 0,08, respectivamente.

Cada valor de frecuencia donde hay un pico, indica que hay un aspecto cíclico en común en los registros tonales de todas las canciones analizadas asociado a dicha frecuencia. La proporción de poder espectral atribuida a la primera componente de acuerdo a cada una de estas frecuencias, está dada por la expresión:

$$\frac{\hat{\lambda}_1(w)}{tr(\hat{f}_{xx}(w))}$$

Donde w es la frecuencia en particular. Por ejemplo, para el valor de frecuencia correspondiente a $w = 0{,}0003$, se hace:

$$\frac{\hat{\lambda}_1(w)}{tr(\hat{f}_{xx}(w))} = \frac{\hat{\lambda}_1(0{,}0003)}{tr(\hat{f}_{xx}(0{,}0003))} = \frac{888000{,}98}{1285888} = 0{,}6905$$

Indicando que la primera componente en la frecuencia $w = 0{,}0003$, explica un 69,05 % del poder espectral total; llegando casi al 70 %. El resultado del cálculo para las primeras cinco componentes en las veinte frecuencias destacadas, puede visualizarse en la sección de Anexos, **Tabla** $A - 2$. No se tienen en cuenta el resto de componentes, debido a que la proporción del poder espectral que explican es muy baja o muy cercana a cero, por lo que es suficiente con revisar solo las primeras cinco componentes.

Como los resultados entre las frecuencias de cada grupo son muy similares entre sí, estos son resumidos en la **Tabla** $6 - 1$, donde se ofrece información sobre la proporción de poder espectral promedio explicada por las primeras cinco componentes en cada grupo; $G1$, $G2$ y $G3$.

Grupo	Acumulado	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
G1	0.9996	0.7043	0.1953	0.0743	0.0220	0.0037
G2	0.9996	0.8132	0.1251	0.0457	0.0133	0.0023
G3	0.9997	0.8627	0.0837	0.0409	0.0107	0.0017

Tabla 6-1.: Promedio de la proporción de poder espectral explicada por las primeras cinco componentes y su acumulado, por cada grupo de frecuencias.

Como se puede observar en la tabla **Tabla** $6 - 1$, la primera componente es la que explica el mayor porcentaje del poder espectral total en cada caso.

6.4. Interpretación de resultados en el dominio del tiempo

Del análisis del espectro de la primera componente principal, se tiene que las series comparten varios fenómenos cíclicos asociados distintas frecuencias. Estos valores se llevan al dominio del tiempo considerando una cantidad d cuyo producto con cada frecuencia w sea 1; es decir,

$d * w = 1$. Cabe mencionar que w corresponde a una fracción de un solo ciclo, por lo que la cantidad d; en teoría; corresponde a las fracciones necesarias de w para obtener un ciclo completo. El valor que cumple con lo anterior es el recíproco de w, $d = \frac{1}{w}$.

Teniendo en cuenta que las series temporales tienen una tasa de muestreo de 21,53 datos por segundo; que es lo mismo a tener un dato cada $\Delta = \frac{1}{21,53}$ de segundo; se busca cuántos Δ; en segundos; se necesitan para tener un ciclo, lo cual se consigue haciendo:

$$\frac{1}{w} * \Delta$$

Tomando el ejemplo particular de la frecuencia $w = 0{,}0003$, se obtiene:

$$\frac{1}{w} * \Delta = \frac{1}{0{,}0003} * \frac{1}{21{,}53} = 154{,}82$$

El anterior resultado indica que existe un patrón cíclico característico de las series temporales analizadas, que se presenta cada 154.82 segundos aproximadamente. La conversión de las veinte frecuencias destacadas al dominio del tiempo, puede visualizarse en la sección de Anexos, **Tabla** $A - 3$.

Nuevamente, como las conversiones de las frecuencias al dominio del tiempo de cada grupo son muy similares entre sí, estas son resumidas en la **Tabla** $6 - 2$, donde se presentan estadísticos descriptivos por cada grupo, relacionados a la frecuencia y su conversión al dominio del tiempo en segundos.

Característica	G1	G2	G3
Rango Frecuencia	0.0003 - 0.0019	0.0434 - 0.0437	0.0765 - 0.0921
Promedio Frecuencia	0.0011	0.0436	0.0872
Desviación Estándar Frecuencia	0.0006	0.0002	0.0056
Rango Tiempo (S)	24.44 - 154.82	1.06 - 1.07	0.5043 - 0.6071
Promedio Tiempo (S)	62.66	1.065	0.5346
Desviación Estándar Tiempo (S)	48.98	0.0070	0.0371

Tabla 6-2.: Tabla de las descripciones de los grupos de frecuencias asociadas a los máximos encontrados en la estimación del espectro de la primera componente.

De la **Tabla** $6 - 2$, se puede notar que en el primer grupo ($G1$) hay una gran varianza en cuanto a las cantidades en segundos. Como se mencionó anteriormente, esto podría deberse a los valores atípicos que contienen las series temporales, por lo que los resultados asociados a este primer grupo no son seguros para el análisis. Un buen ejemplo de lo anterior, es

la frecuencia $w = 0{,}0003$, cuya conversión al dominio del tiempo es de 154,82 segundos por ciclo. Este resultado genera duda, debido a que hay canciones dentro de la muestra que tienen una duración alrededor de este valor, sugiriendo que son consumidas por el ciclo entero y costándole sentido al análisis.

Considerando lo anterior, el análisis se llevará a cabo teniendo en cuenta solo los otros dos grupos; $G2$ y $G3$.

6.5. Inspección de resultados

Tomando algunos de los resultados obtenidos del análisis preliminar de las series de tiempo, se tiene que para la primera canción cuya frecuencia máxima era $w = 0{,}002$, indica que cada 23,22 segundos se repiten las notas; mientras que, para la tercera canción con frecuencia máxima $w = 0{,}04$, el ciclo sucede cada 1,16 segundos.

Revisando las canciones se encontró algunas coherencias con lo expuesto anteriormente, estos son algunas de las observaciones obtenidas:

Primera canción:

- **Primer ciclo:** Inicia la canción con sonidos de tambora, para luego iniciar una secuencia tonal o conjunto de notas interpretadas con piano y trompetas.
- **Segundo ciclo:** Inicia la primera estrofa o verso de la canción, cantada con una secuencia tonal similar a la anterior.
- **Tercer ciclo:** Continúa con un nuevo ciclo de trompetas y piano con la misma secuencia tonal con la que inició la canción.
- **Cuarto ciclo:** Cambia la secuencia tonal, debido probablemente a que se presenta el coro de la canción.
- **Quinto ciclo:** Se reproduce una nueva secuencia tonal diferente a las dos anteriores, interpretada en trompetas.
- **Sexto ciclo:** A partir de este ciclo, se retoma la sucesión de ciclos desde el primero, como si la canción tuviese dos periodos de cinco ciclos cada uno.

Tercera canción:

- **Primer ciclo:** Inicia con una secuencia tonal ascendente interpretada en acordeón.
- **Segundo ciclo:** Continúa con una secuencia tonal similar, pero descendente. También está interpretada en acordeón.
- **Tercer ciclo:** Se repite la secuencia tonal del primer ciclo, aún con acordeón.
- **Cuarto ciclo:** Finaliza con una nueva secuencia de notas descendente en acordeón.
- **Quinto ciclo – Octavo ciclo:** Se repite el primer y segundo ciclo de forma consecutiva hasta llegar al octavo.

Las siguientes secuencias de notas, parecen ser más largas que 1,16 segundos y no son frecuentes, pero se considera que; en realidad; lo que detecta como ciclo son las notas que se interpretan constantemente con instrumentos de percusión en el fondo y que prevalecen hasta el final de la canción. Este caso es similar al de la segunda canción, que tiene un ciclo cada 0,66 segundos, pero el tiempo es tan corto, que es difícil identificarlo con el oído y sin práctica.

Como se puede notar, ocurren ciclos cada cierta cantidad de segundos en las canciones, pero estos son distintos para cada una de ellas; más entre unas, que entre otras.

El resultado del espectro de la primera componente principal del ACP realizado, resume la periodicidad de las series temporales, sugiriendo que ocurren fenómenos cíclicos cada 1,065 segundos (Segundo grupo - $G2$) y cada 0,5346 segundos (Tercer grupo - $G3$), aproximadamente. Al inspeccionar qué ocurre durante esos intervalos de tiempo en las canciones de la muestra, se encuentra que ambos tiempos están asociados al mismo fenómeno musical en todas las pistas, tratándose de las notas de fondo interpretadas por instrumentos de percusión, como la tambora y la guacharaca, y que se mantienen durante toda la canción o en gran parte de ella.

7. Conclusiones y recomendaciones

7.1. Conclusiones

Inicialmente, las gráficas de las señales de los archivos de audio (Oscilogramas) no muestran patrones cíclicos o estacionales de forma notoria, ni características de interés para la investigación; esto es debido a la complejidad de la señal al estar compuesta por varios elementos sonoros, como a su vez es el resultado de la combinación lineal de distintas componentes frecuenciales con formas sinusoidales.

El extraer la frecuencia fundamental de la señal de audio de la canciones como series, permite el análisis de los registros tonales de cada pista. Las gráficas de las series temporales presentan comportamientos periódicos en lapsos cortos de tiempo, pero que aparecen con frecuencia en toda la serie. El análisis de los periodogramas individuales de cada canción, corrobora la existencia de estos fenómenos cíclicos en todas las canciones y permite notar diferencias en cuanto a intensidad y duración se refiere, siendo mas similares entre unas, que entre otras; como es el caso de las canciones 3, 4, 7, 8 y 9; cuyas periodicidades están asociadas a frecuencias cercanas al valor 0,04, indicando que hay secuencias de notas que se repiten cada 1,16 segundos aproximadamente.

Del análisis de las series temporales, también se contempla la existencia de puntos atípicos que ocasionan picos muy cercanos al valor cero en la estimación del periodograma; pero por su naturaleza y la incertidumbre que acarrean estos valores, no son tenidos en cuenta en la investigación, más que para ejemplificar algunos procedimientos.

Se consigue a través del análisis de componentes principales para series temporales, una caracterización del conjunto de series a través de una sola primera componentes en diferentes frecuencias; componente que en promedio es capaz de explicar hasta un 80 % del poder espectral total; siendo adecuada para resumir el conjunto de datos. De esta caracterización se generan tres grupos principales de frecuencias donde se concentra una buena cantidad de poder espectral; descartando el primero al ser el que tiene los valores de frecuencia cercanos a cero. Los otros dos grupos obtienen resultados similares, en cuanto a la detección de periodicidades comunes entre las canciones, encontrándose que existen secuencias de tonos y notas que se interpretan o reproducen cada un segundo y cada medio segundo.

La inspección de las canciones demuestra que es real la existencia de fenómenos cíclicos cada medio segundo y cada segundo en todas las canciones, asociándose al sonido de fondo para el compás, que se interpreta con intrumentos de percusión como la tambora y la guacharaca.

7.2. Recomendaciones

Como recomendación para futuras investigaciones, se sugiere un enfoque distinto al propuesto en este trabajo; es decir, distinto al exploratorio; tratando de recopilar un conjunto de datos más robusto y con mayor rigurosidad en la selección de las canciones, de tal forma que haya mayor representatividad en la muestra seleccionada, con participación de la aleatoriedad. También se podrían considerar más subgéneros actuales, como la tecnocumbia o la cumbia urbana, para representar mejor las épocas.

También se recomienda realizar análisis similares usando como insumo otras características de las canciones, como las demás frecuencias armónicas que las componen o la misma señal de la onda. Además, se podría sugerir el implementar otra clase de métodos para encontrar características entre las canciones, como el análisis espectral singular o el agrupamiento por clustering.

Bibliografía

Ahmad, A., Schlindwein, F. & Ng, G. A. (2010), 'Comparison of computation time for estimation of dominant frequency of atrial electrograms: Fast fourier transform, blackman tukey, autoregressive and multiple signal classification'.

Anders, E. & Anders, F. (2017), 'Long-term average spectrum in popular music and its relation to the level of the percussion', *Audio Engineering Society, Convention Paper.* .

Arias, A. (2021), 'La cumbia colombiana', *Universidad INCCA de Colombia* .

Bernal, J., Gómez, P. & Bobadilla, J. (2009), 'Una visión práctica en el uso de la transformada de fourier como herramienta para el análisis espectral de la voz', *Estudios de fonética experimental* .

Bisina, K. & Azeez, M. (2017), 'Optimized estimation of power spectral density'.

Bloomfield, P. (2000), *Fourier Analysis of Time Series. Wiley-Interscience.*

Carvalho, M., Rodrigues, P. & Rua, A. (2011), 'Economic letters', *Elsevier* .

Castro, M. (2003), *Música para todos: Una introducción al estudio de la música*, Universidad de Costa Rica.

Clarke, J. (2000), 'Tourism brands: An exploratory study of the brands box model', *Journal of Vacation Marketing* .

Cruz, D. A., Lopez, S. S. & Camargo, J. E. (2019), 'Automatic identification of traditional colombian music genres based on audio content analysis and machine learning technique'.

Doval & Rodet (1991), 'Estimation of fundamental frequency of musical sound signals'.

Echeverry, J., Gutiérrez, A. & Guarnizo, C. (2007), 'Análisis de la densidad espectral de potencia en registros mer'.

Fries, B. & Fries, M. (2005), *Digital Audio Essentials*, O'Reilly Media, Inc.

García, O. & Moreno, G. (2006), 'Sobre la utilización del análisis de fourier, análisis espectral singular y redes neuronales artificiales en estratigrafía. parte 1: Teoría y caso sintético', *Universidad Nacional de Colombia* .

Godt, I. (2005), 'Music: A practical definition'.

González, D. (2001), 'Análisis espectral: Consideraciones teóricas y aplicabilidad.', *Economía y sociedad* .

Iriarte, J., Fernández, S., Fernandez, N., Urrestarazu, E., Pagola, I., Alegre, M. & Artieda, J. (2010), 'Sound analysis of catathrenia: a vocal expiratory sound'.

Ligges, U., Krey, S., Mersmann, O., Schnackenberg, S., Guenard, G., Preusser, A., Thieler, A. & Mielke, J. (2023), 'Package "tuner"'.

Martín, F. A. (2017), DDesarrollo y análisis de clasificadores de señales de audio, PhD thesis.

Mesa, O. & Carvajal, L. (1994), 'Análisis espectral singular y redes neuronales aplicadas en hidrología. avances en recursos hidráulicos', *Revista Unal* .

Molina, Y. (2020), 'Empleo de técnicas de análisis tiempo-frecuencia para el modelado de instrumentos musicales y su aplicación a la síntesis musical aditiva'.

Muller, M., Ellis, D., Klapuri, A. & Gaël, R. (2011), 'Signal processing for music analysis'.

Márquez, I. (2016), 'Cumbia digital: Tradición y postmodernidad'.

Ochoa, J. (2016), 'La cumbia en colombia: invención de una tradición'.

Olivares, O. (2019), 'Análisis de audio para extracción de características, segmentación, clasificación y predicción', *Centro de Investigación en Matemáticas, A.C.* .

Osorio, J., Knott, A. & Chaves, J. (2012), 'Apróximación de la síntesis de la música a través del análisis de fourier'.

Pachet, F. & Aucouturier, J.-J. (2003), 'Representing musical genre: A state of the art'.

Pascual, A. & Traver, M. (2014), 'Aplicaciones prácticas del análisis de componentes principales en gestión de carteras de renta fija (iii): Estrategias de valor relativo mediante el análisis de componentes principales: Pca butterflies'.

Potamitis & Rigakis (2016), 'Measuring the fundamental frequency and the harmonic properties of the wingbeat of a large number of mosquitoes in flight using 2d optoacoustic sensors. applied acoustics'.

Pértega, S. & Fernández, S. (2016), 'Métodos no paramétricos para la comparación de dos muestras'.

Reyes, J. (1998), 'Modelos de expresión musical basados en análisis y tratamiento de señal de audio'.

Rodríguez, M. L. (2021), Detección Automática de Géneros Músicales, PhD thesis.

Rong, F. (2016), 'Método de clasificación de audio basado en aprendizaje automático'.

Santoso, H. & Hilman, M. (2018), 'Perancangan aplikasi keamanan file audio format wav (waveform) menggunakan algoritma rsa'.

Sarria, G., Mora, M. & Lopera, C. (2016), 'Salsa dataset: primera base de conocimiento de música salsa', *Ricercare* .

Scaringella, N. & Zoia, M. (2006), 'Automatic genre classification of music content: a survey'.

Shumway, R. & Stoffer, D. (2017), *Time Series Analysis and Its Applications*, Springer.

Stockdale, A. (2002), 'Tools for digital audio recording in qualitative research'.

Stoffer, D. & Poison, N. (2023), 'Package "astsa"'.

Sueur, J. (2018), *Sound Analysis and Synthesis with R.*

Sueur, U. (2022), 'Package "seewave"'.

Yu, G. & Slotine, J.-J. (2009), 'Audio classification from time-frequency texture', *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* .

Zapata, J. R. (2021), 'Minería de audio'.

Zhigljavsky, G. N. . A. (2020), *Singular Spectrum Analysis for Time Series.*

A. Anexos: Gráficos, tablas y código en R

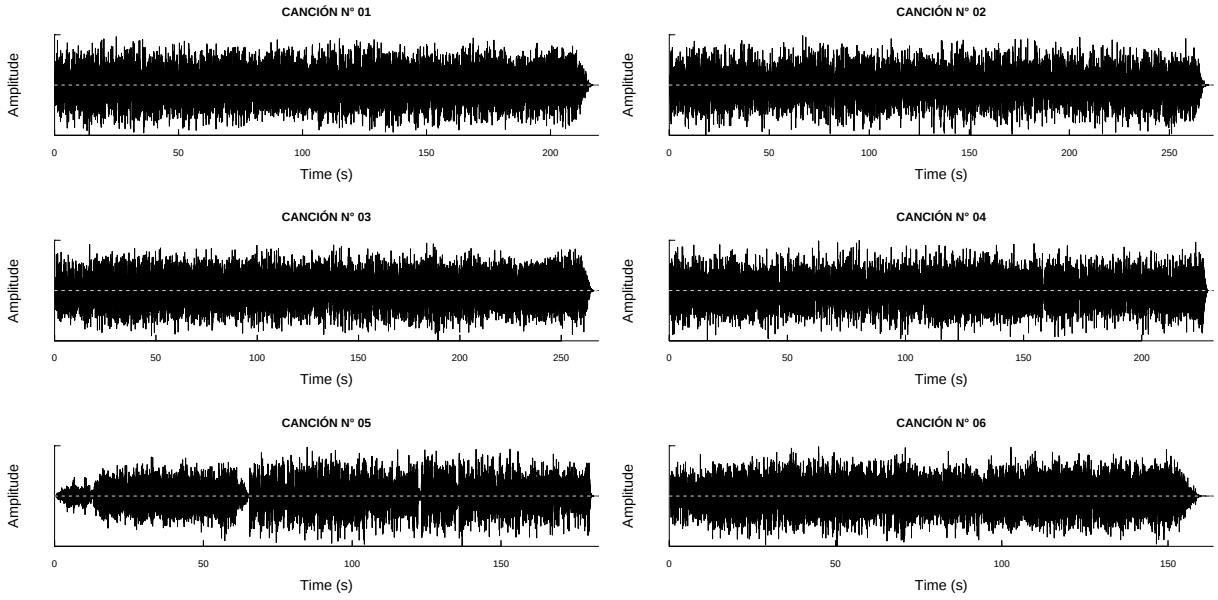

Figura A-1.: Oscilogramas de las primeras 6 canciones

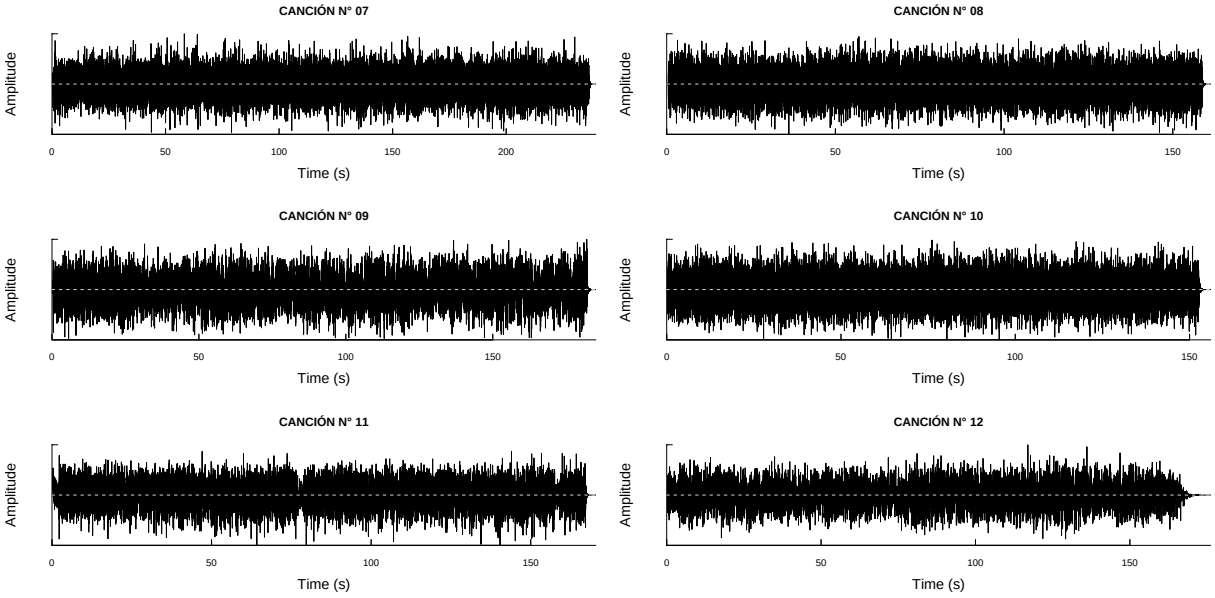

Figura A-2.: Oscilogramas de las canciones desde la 7 hasta la 12

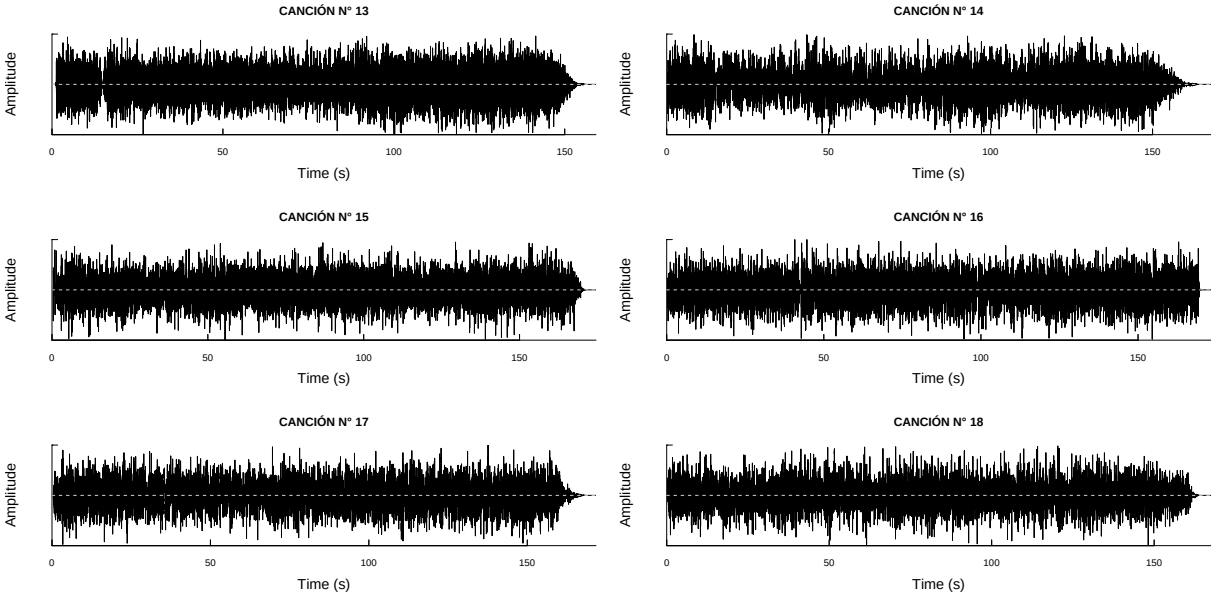

Figura A-3.: Oscilogramas de las canciones desde la 13 hasta 18

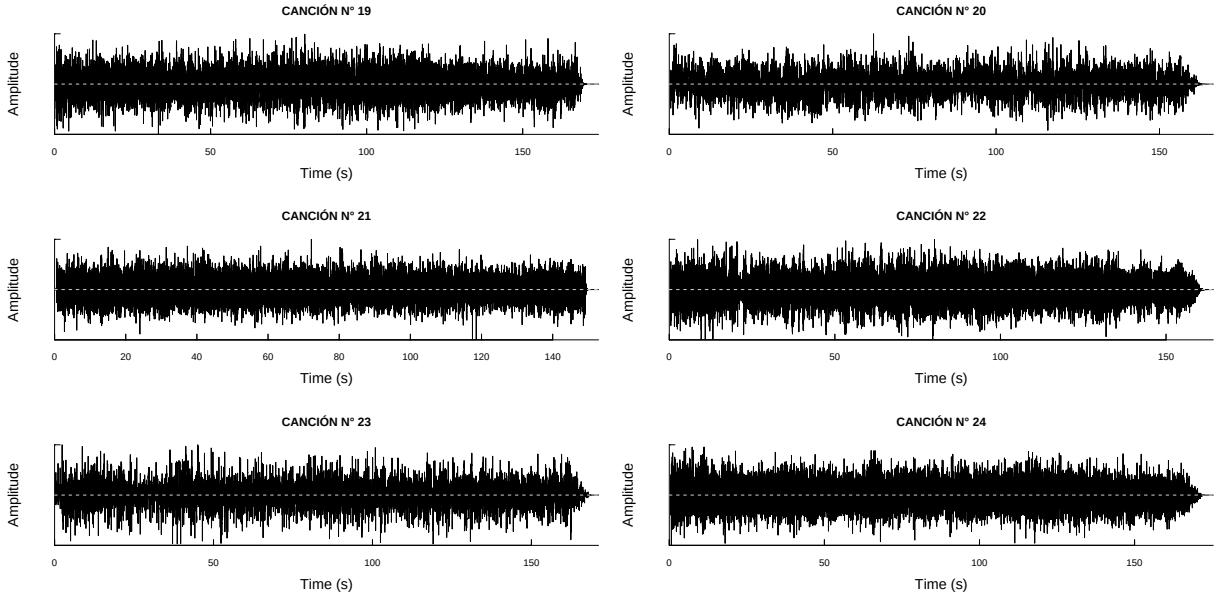

Figura A-4.: Oscilogramas de las canciones desde la 19 hasta la 24

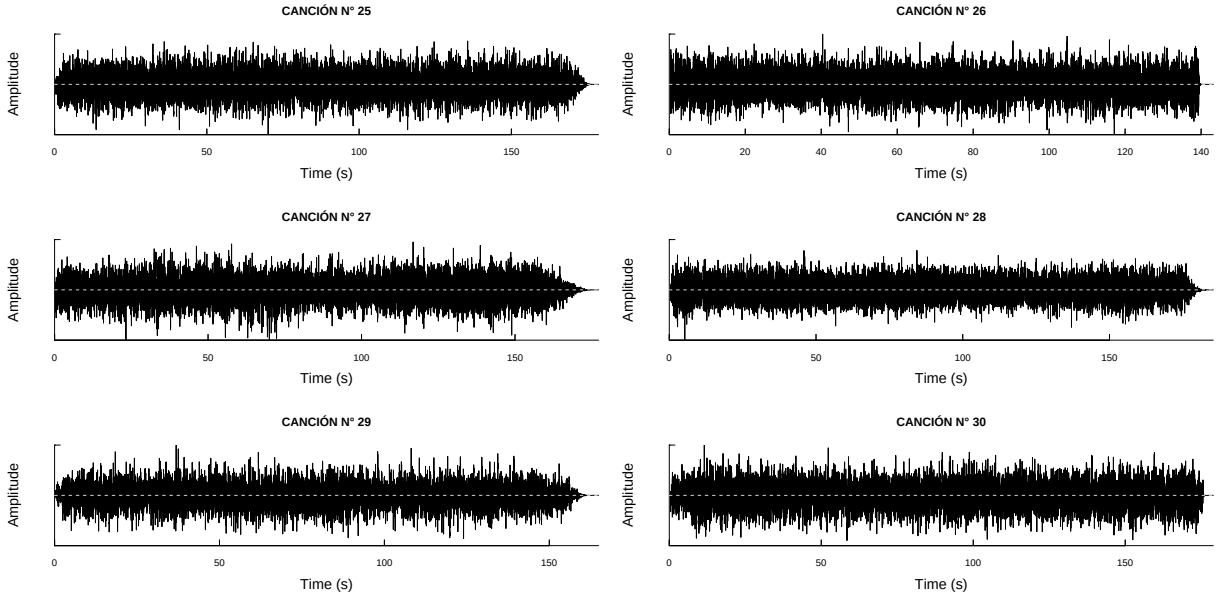

Figura A-5.: Oscilogramas de las canciones desde la 24 hasta la 30

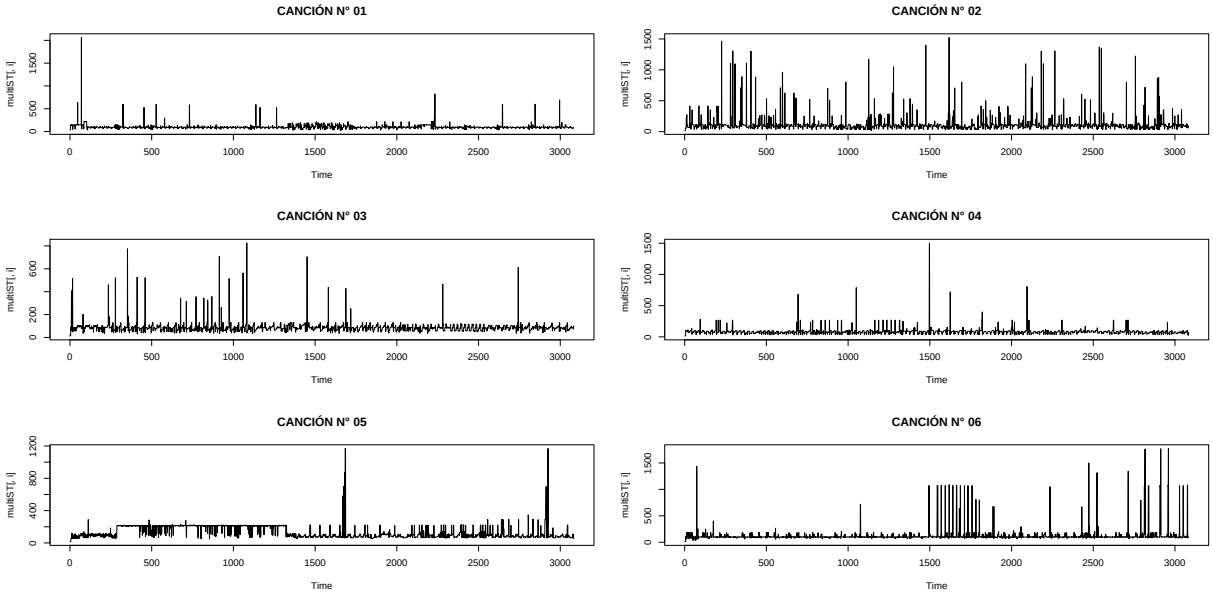

Figura A-6.: Representación gráfica de las primeras 6 series temporales asociadas a la frecuencia fundamental de las canciones

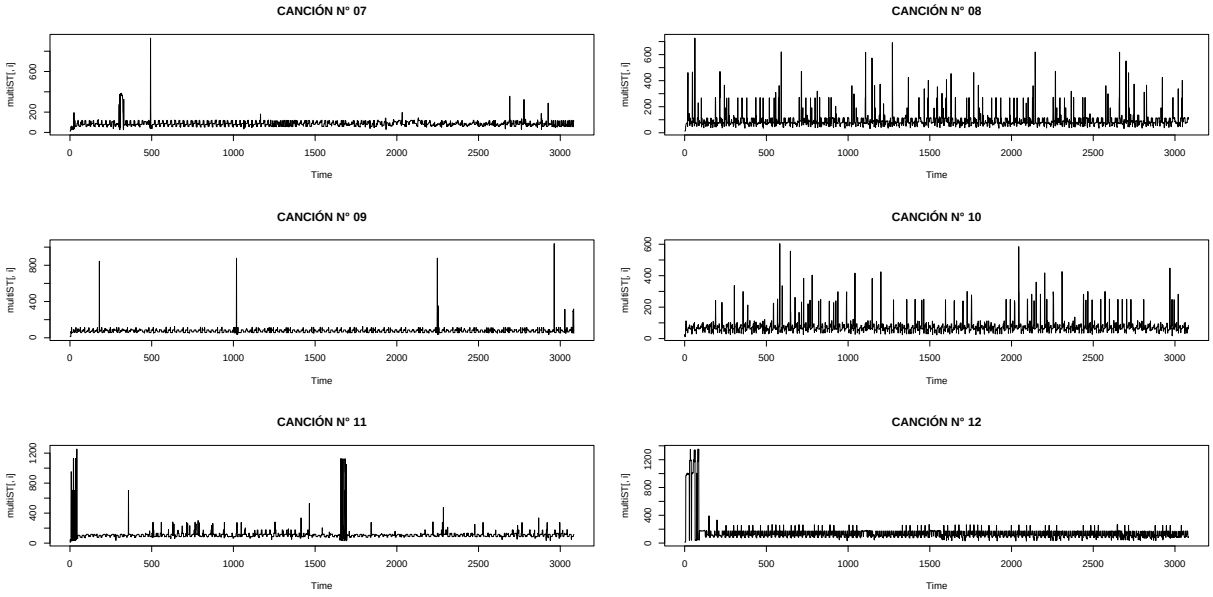

Figura A-7.: Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 7, hasta la 12

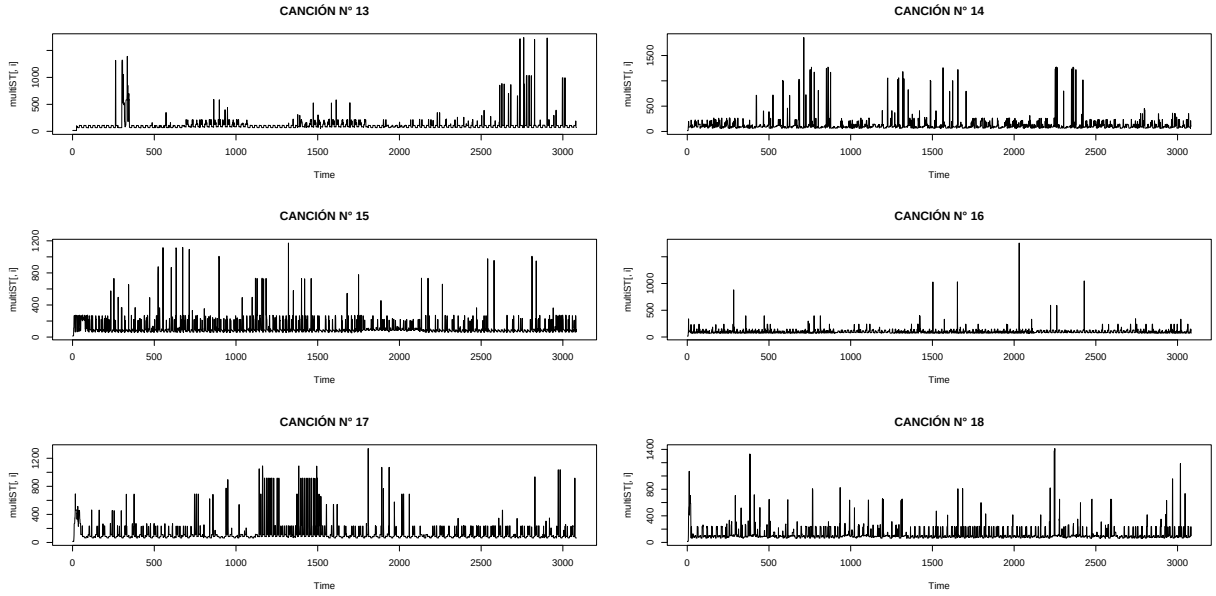

Figura A-8.: Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 13, hasta la 18

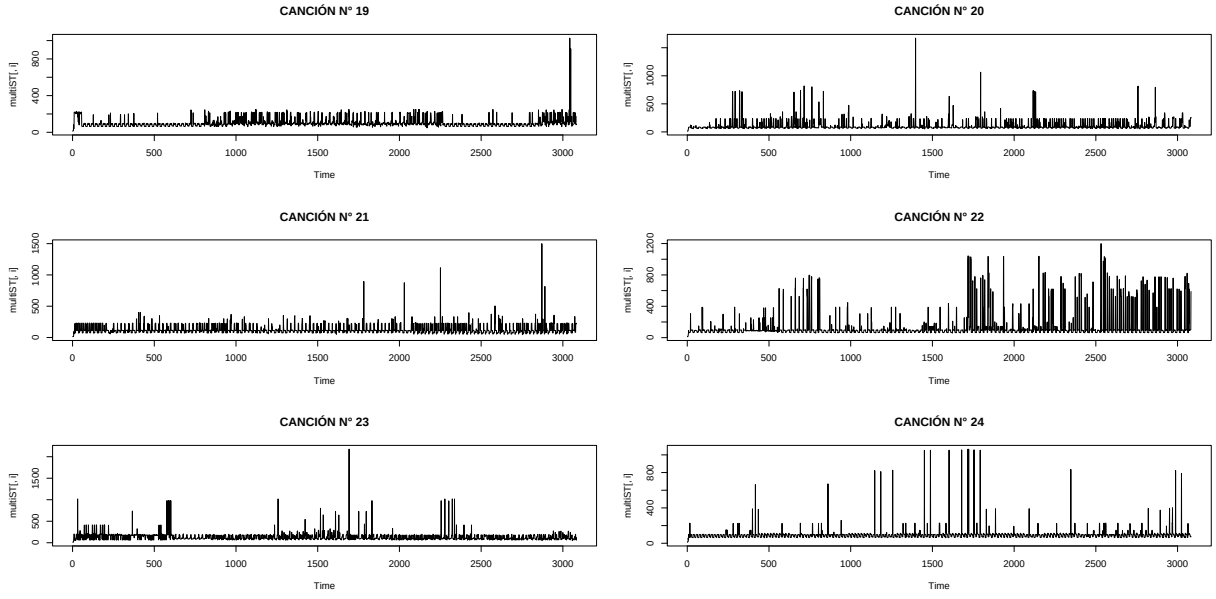

Figura A-9.: Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 19, hasta la 24

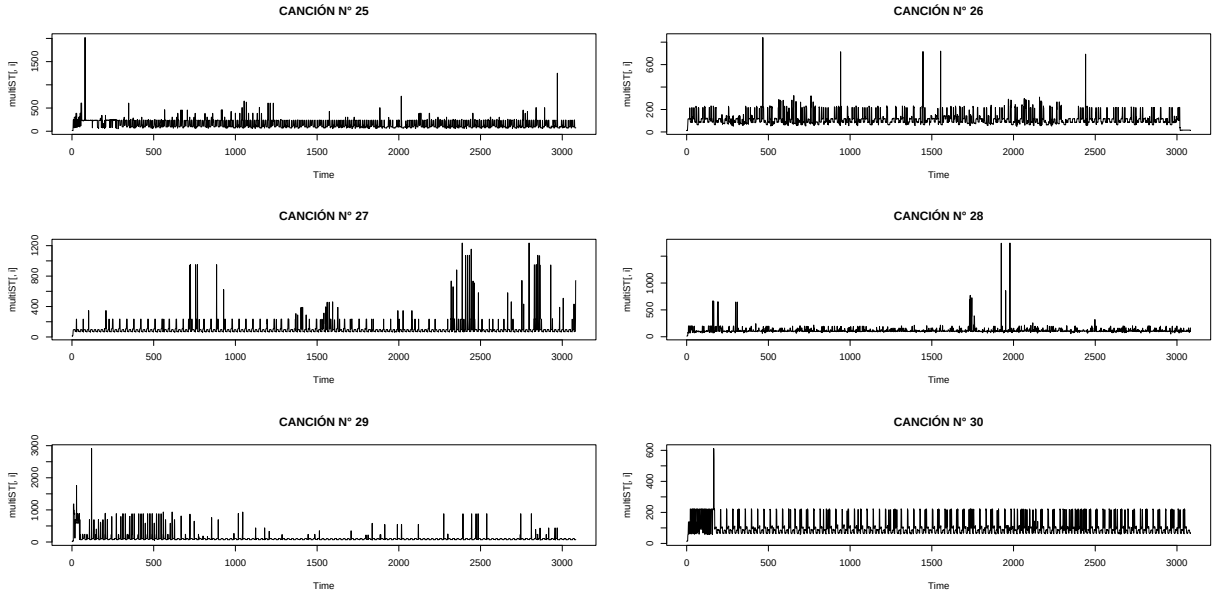

Figura A-10.: . Representación gráfica de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 25, hasta la 30

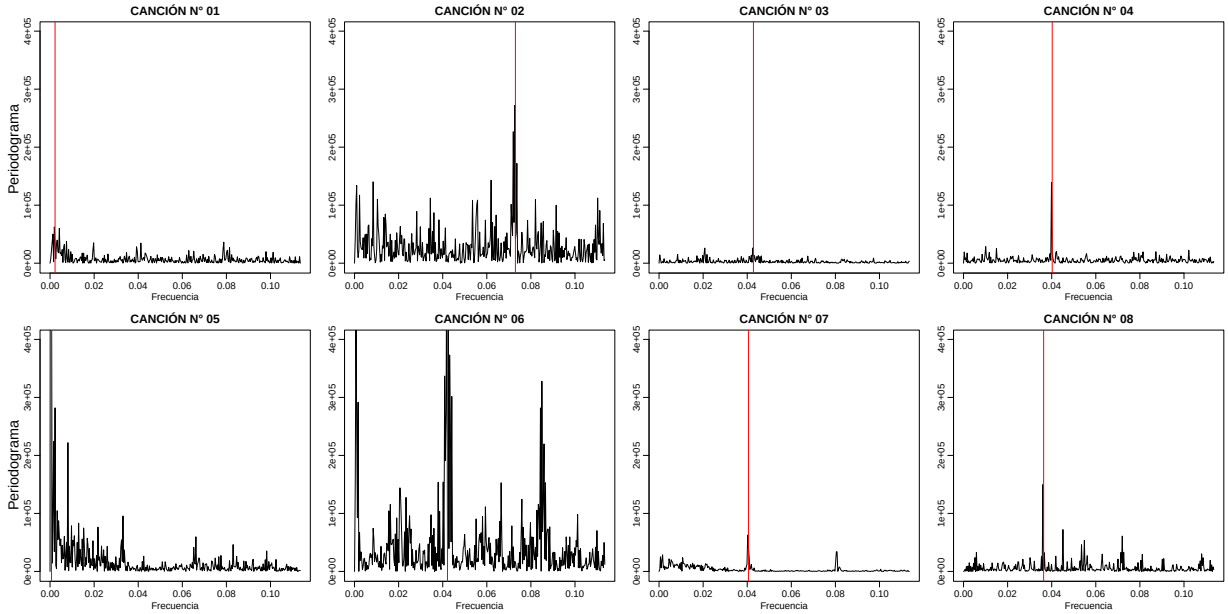

Figura A-11.: . Periodogramas estimados de las primeras 8 series temporales asociadas a la frecuencia fundamental de las canciones, con sus valores máximos señalados en rojo.

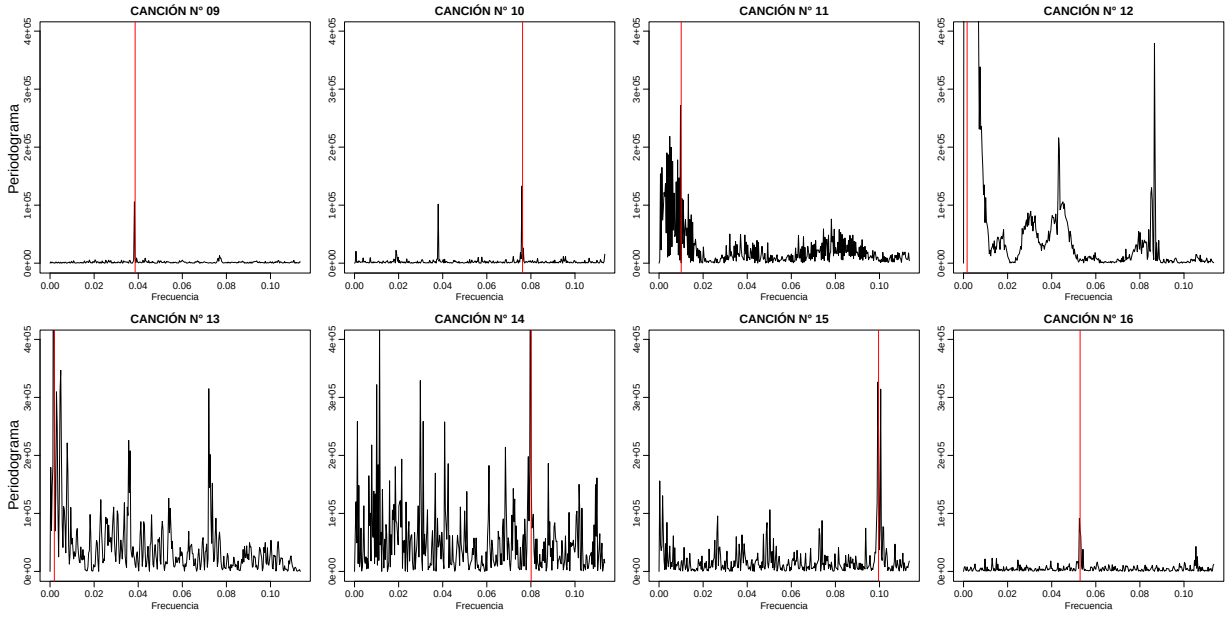

Figura A-12.: . Periodogramas estimados de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 9, hasta la 16

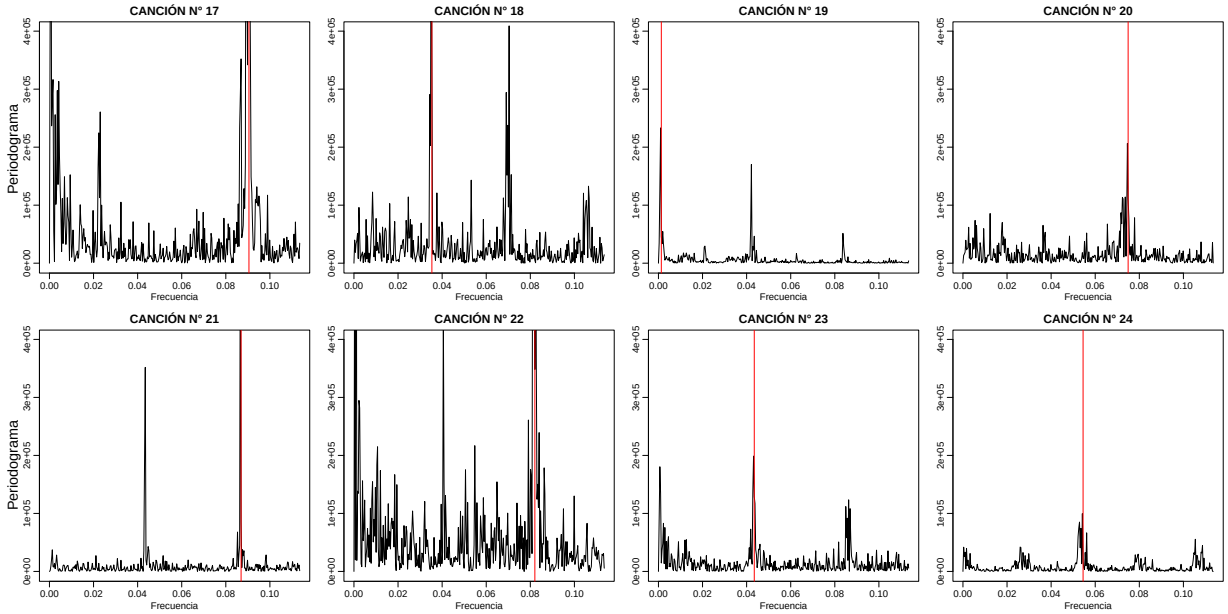

Figura A-13.: . Periodogramas estimados de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 17, hasta la 24

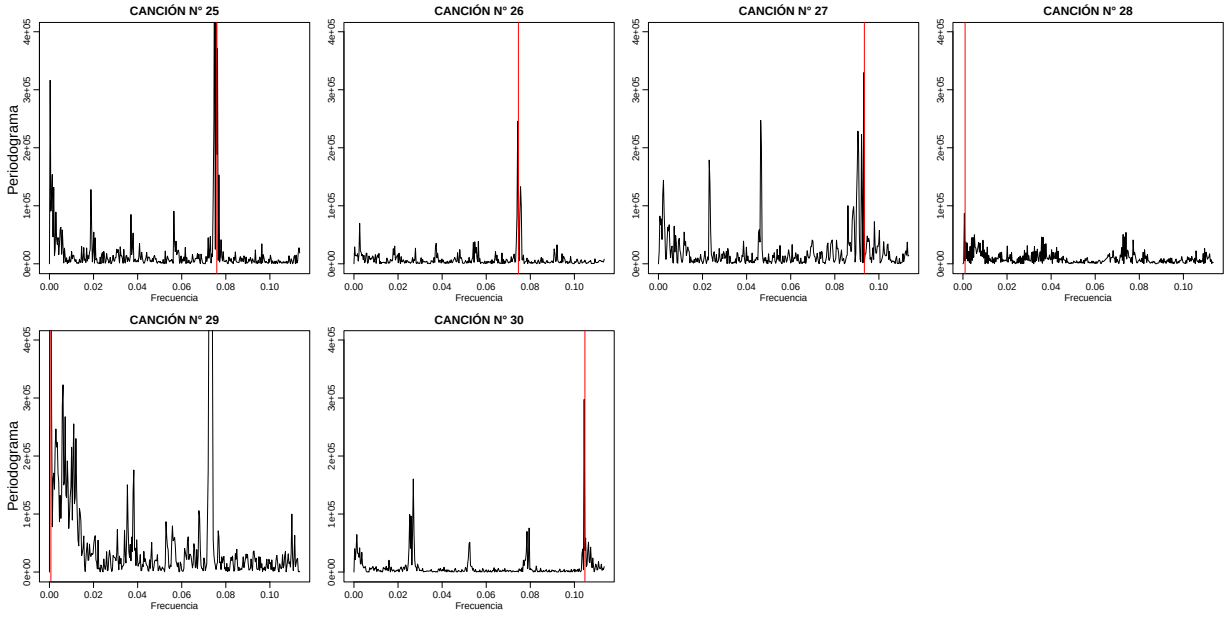

Figura A-14.: . Periodogramas estimados de las series temporales asociadas a la frecuencia fundamental de las canciones, desde la 25, hasta la 30

Tablas

Código	Título	Álbum	Intérprete	Año	Duración
01	Canción N°01	Álbum N°01	Intérprete N°01	1982	219
02	Canción N°02	Álbum N°02	Intérprete N°02	1980	271
03	Canción N°03	Álbum N°03	Intérprete N°03	1984	268
04	Canción N°04	Álbum N°04	Intérprete N°04	1976	230
05	Canción N°05	Álbum N°05	Intérprete N°05	1969	182
06	Canción N°06	Álbum N°06	Intérprete N°06	1953	163
07	Canción N°07	Álbum N°07	Intérprete N°01	1999	239
08	Canción N°08	Álbum N°08	Intérprete N°07	1982	161
09	Canción N°09	Álbum N°09	Intérprete N°08	1990	185
10	Canción N°10	Álbum N°10	Intérprete N°09	1978	156
11	Canción N°11	Álbum N°11	Intérprete N°10	1982	170
12	Canción N°12	Álbum N°06	Intérprete N°06	1953	176
13	Canción N°13	Álbum N°12	Intérprete N°05	1961	159
14	Canción N°14	Álbum N°13	Intérprete N°11	1963	167
15	Canción N°15	Álbum N°14	Intérprete N°12	1989	174
16	Canción N°16	Álbum N°15	Intérprete N°13	1993	173
17	Canción N°17	Álbum N°16	Intérprete N°14	1967	171
18	Canción N°18	Álbum N°17	Intérprete N°15	1960	167
19	Canción N°19	Álbum N°18	Intérprete N°16	1975	174
20	Canción N°20	Álbum N°19	Intérprete N°09	1965	166
21	Canción N°21	Álbum N°20	Intérprete N°17	1968	152
22	Canción N°22	Álbum N°21	Intérprete N°18	1972	164
23	Canción N°23	Álbum N°22	Intérprete N°19	1966	171
24	Canción N°24	Álbum N°13	Intérprete N°11	1963	175
25	Canción N°25	Álbum N°23	Intérprete N°14	1962	178
26	Canción N°26	Álbum N°24	Intérprete N°20	1966	143
27	Canción N°27	Álbum N°25	Intérprete N°21	1994	177
28	Canción N°28	Álbum N°26	Intérprete N°22	1965	185
29	Canción N°29	Álbum N°27	Intérprete N°23	1975	164
30	Canción N°30	Álbum N°28	Intérprete N°24	1962	178

Tabla A-1.: Tabla de metadatos asociados a la muestra de canciones.

Frecuencia	Componente 1	Componente 2	Componente 3	Componente 4	Componente 5
0,0003	0.6905	0.1975	0.1013	0.0105	0.0000
0,0006	0.7053	0.1997	0.0653	0.0247	0.0048
0,0009	0.7140	0.1967	0.0601	0.0260	0.0030
0,0914	0.9164	0.0537	0.0224	0.0063	0.0009
0,0012	0.7281	0.1533	0.0907	0.0223	0.0053
0,0911	0.9187	0.0460	0.0259	0.0081	0.0011
0,0917	0.9125	0.0498	0.0286	0.0078	0.0010
0,0908	0.9252	0.0437	0.0224	0.0072	0.0013
0,0016	0.6965	0.2210	0.0563	0.0209	0.0050
0,0921	0.8992	0.0577	0.0328	0.0084	0.0015
0,0904	0.9107	0.0449	0.0341	0.0081	0.0019
0,0879	0.8177	0.1075	0.0585	0.0145	0.0017
0,0768	0.8674	0.0658	0.0479	0.0162	0.0025
0,0875	0.7608	0.1671	0.0553	0.0146	0.0020
0,0765	0.8235	0.1185	0.0447	0.0102	0.0028
0,0019	0.6913	0.2037	0.0724	0.0279	0.0045
0,0882	0.7912	0.1321	0.0545	0.0197	0.0023
0,0437	0.8297	0.0999	0.0567	0.0117	0.0018
0,0434	0.7967	0.1504	0.0348	0.0149	0.0029
0,0823	0.8088	0.1176	0.0641	0.0077	0.0016

Tabla A-2.: Tabla de las proporciones de poder espectral explicada por las primeras cinco componentes en las frecuencias destacadas

Posición	Ciclo	Frecuencia	Tiempo (S)	Grupo
1	1	0.0003	154.82	G1
2	2	0.0006	77.41	G1
3	3	0.0009	51.60	G1
4	282	0.0914	0.5081	G3
5	4	0.0012	38.70	G1
6	281	0.0911	0.5098	G3
7	283	0.0917	0.5065	G3
8	280	0.0908	0.5115	G3
9	5	0.0016	29.02	G1
10	284	0.0921	0.5043	G3
11	279	0.0904	0.5137	G3
12	271	0.0879	0.5284	G3
13	237	0.0768	0.6047	G3
14	270	0.0875	0.5308	G3
15	236	0.0765	0.6071	G3
16	6	0.0019	24.44	G1
17	272	0.0882	0.5266	G3
18	135	0.0437	1.06	G2
19	134	0.0434	1.07	G2
20	254	0.0823	0.5643	G3

Tabla A-3.: Tabla de las frecuencias asociadas a los primeros veinte valores máximos dentro de la estimación del espectro de la primera componente principal, con su respectiva representación en el dominio del tiempo.

Atributo	Cuenta
N° Canciones	30
N° Álbumes	28
N° Intérpretes	24
N° Años	30

Tabla A-4.: Tabla de caracterización de los atributos de la muestra de canciones.

Canción	Rango	Media	Desviación Estándar
01	2050.76	98.32	56.48
02	2036.38	110.98	120.22
03	840.23	88.95	51.57
04	1492.37	82.74	57.97
05	1157.89	125.68	79.29
06	1761.87	85.72	148.65
07	917.63	95.96	32.02
08	715.26	75.84	64.60
09	1574.80	73.88	40.72
10	592.83	115.79	44.17
11	1241.83	143.24	99.90
12	1341.20	114.93	137.58
13	1727.37	124.90	130.24
14	1844.57	114.99	145.73
15	1161.18	98.68	107.44
16	1744.46	126.76	57.75
17	1327.39	121.12	141.21
18	1402.38	104.33	109.10
19	1351.69	103.06	56.58
20	1660.26	121.15	85.46
21	1489.02	141.00	73.33
22	1187.70	141.00	180.20
23	2165.06	136.06	105.78
24	1053.53	95.31	68.81
25	2007.29	152.67	96.13
26	829.78	111.55	56.44
27	1223.93	103.86	98.10
28	1733.33	115.21	63.50
29	2900.77	119.94	145.02
30	2530.21	108.13	118.92

Tabla A-5.: Tabla de estadísticas descriptivas de las frecuencias fundamentales de cada canción.

Código en R

```
# DESARROLLO DE CODIGO PARA LA IMPLEMENTACION DEL METODO:

#### PREPARATIVOS: ####

### INTALANDO PAQUETES:
install.packages("tuneR")        ## EXTRACCION DE AUDIO.
install.packages("stringr")      ## MANIPULACION DE CADENAS DE TEXTO.
install.packages("astsa")        ## DESARROLLO ACP.
install.packages("seewave")      ## GRAFICAR OSCILOGRAMAS.
install.packages("ggplot2")      ## GRAFICAR FRECUENCIAS.
install.packages("dplyr")        ## ORDENAR DATA FRAMES.
install.packages("ggridges")     ## GRAFICAR FRECUENCIAS.
install.packages("plotly")       ## GRAFICAR FRECUENCIAS.
install.packages("reshape2")     ## MODIFICAR TABLAS.
install.packages("gridExtra")    ## ACOMODAR GRAFICOS.
install.packages("tidyr")        ## MANIPULAR TABLAS.
install.packages("tseries")      ## TESTEAR ESTACIONARIEDAD.
install.packages("astsa")        ## DESARROLLAR ACP.

### LLAMANDO LIBRERIAS:
library(tuneR)
library(stringr)
library(astsa)
library(seewave)
library(ggplot2)
library(dplyr)
library(ggridges)
library(plotly)
library(reshape2)
library(gridExtra)
library(tidyr)
library(tseries)
library(astsa)

### ESTABLECIENDO VARIABLES DE DIRECTORIO DE ENTRADA DE
### INSUMOS Y SALIDA DE RESULTADOS:
entradas <- paste0(getwd(), "/entradas")
salidas <- paste0(getwd(), "/salidas")

#### LEYENDO ARCHIVOS DE MUSICA EN FORMATO MP3: ####

### ESTABLECIENDO DIRECTORIO DE ENTRADA DE INSUMOS:
setwd(entradas)

### ENLISTANDO CANCIONES EN FORMATO MP3:
ficheros <- list.files(pattern = "\\.mp3")
```

```
### QUITANDO FORMATOS DE LOS NOMBRES:
nombres <- str_sub(ficheros, 1, nchar(ficheros) - 4)

### ALMACENANDO CANCIONES:
for (i in 1:length(nombres)) {
  assign(nombres[i], readMP3(ficheros[i]))
}

#### EXTRAYENDO INFORMACION DE LAS CANCIONES: ####

### EXTRAYENDO INFORMACION WAV:
for(i in 1: length(nombres)) {
  assign(paste0(nombres[i], "_WAV"), extractWave(get(paste0(nombres[i])),
  interact = FALSE))
}

### COMBINANDO CANALES:
for(i in 1: length(nombres)) {
  assign(paste0(nombres[i], "_WAV"), mono(get(paste0(nombres[i], "_WAV")),
  "both"))
}

### GRAFICANDO OSCILOGRAMAS:
windows()
par(mfrow = c(3, 2))
for (i in 7:12) {
  oscillo(get(paste0(nombres[i], "_WAV")),
  fastdisp = TRUE, title = paste0("CANCION_N_", nombres[i]))
}

### ASIGNANDO FRECUENCIAS:
for(i in 1:length(nombres)) {
  assign(paste0(nombres[i], "_Freq"), FF(periodogram(get(paste0(nombres[i],
  "_WAV")), fastdisp = TRUE, title = nombres[i], width = 4096,
  overlap = 2048), peakheight = 0.1))
}

### ALMACENANDO EN UN DATA FRAME:
cancionesDT <- data.frame(cancion = paste0("CANCION_N_", nombres[1]),
frecuencia_fundamental = get(paste0(nombres[1], "_Freq")))

for(i in 2: length(nombres)) {
  cancionesDT <- rbind(data.frame(cancion = paste0("CANCION_N_", nombres[i]),
  frecuencia_fundamental = get(paste0(nombres[i], "_Freq"))), cancionesDT)
}

### PREPARANDO DATA FRAME:
```

```
## OMITIENDO VALORES VACIOS.
cancionesDT <- na.omit(cancionesDT)
## CAMBIANDO TIPO DE VARIABLE PARA DELIMITAR.
cancionesDT$frecuencia <- as.numeric(cancionesDT$frecuencia_fundamental)
## DELIMITANDO FRECUENCIA PARA GRAFICAR.
cancionesDT <- subset(cancionesDT, cancionesDT$frecuencia_fundamental < 120)
## DELIMITANDO FRECUENCIA PARA GRAFICAR
cancionesDT <- subset(cancionesDT, cancionesDT$frecuencia_fundamental > 50)
## ORDENANDO POR CANCIONES.
cancionesDT <- arrange(cancionesDT, cancion)

### GRAFICANDO FRECUENCIAS:
windows()
ggplot(cancionesDT, aes(x = frecuencia_fundamental, y = cancion)) +
geom_density_ridges() +
ggtitle("Densidad marginal de la frecuencia fundamental de las canciones")

### ELIMINANDO ELEMENTOS INNECESARIOS:
for (i in 1:length(nombres)) {
  rm()
}

#### ORGANIZANDO TABLA DE DATOS: ####

### HALLANDO TAMANO MAXIMO ENTRE LAS FRECUENCIAS DE LAS CANCIONES:
tamanos <- c()

for (i in 1:length(nombres)) {
  tamanos[i] <- length(get(paste0(nombres[i], "_Freq")))
}

tamMax <- min(tamanos)

### ASIGNANDO UN MISMO TAMANO PARA TODAS LAS CANCIONES Y UNIENDOLAS EN UN
### SOLO DATA FRAME:
temp <- get(paste0(nombres[1], "_Freq"))
length(temp) <- tamMax
frecuencias <- temp

for (i in 2:length(nombres)) {
  temp <- get(paste0(nombres[i], "_Freq"))
  length(temp) <- tamMax
  frecuencias <- cbind(frecuencias, temp)
}

colnames(frecuencias) <- nombres
```

```
### CAMBIANDO LOS VALORES NA'S EN CEROS:
frecuenciasDF <- frecuencias

### CONVIRTIENDO DATOS A SERIE TEMPORAL MULTIVARIADA:
multiST <- ts(frecuenciasDF)

### GRAFICANDO SERIES TEMPORALES:
windows()
par(mfrow = c(3, 1))
for (i in 1:3) {
  plot(multiST[ , i], main = paste0("CANCION N ", nombres[i]))
}

### GRAFICANDO UN SEGMENTO DETALLADO:
segmento <- `03 Freq`[1750:2500]
windows()
plot(c(seq(1750, 2500)), segmento, main = paste0("CANCION N ", nombres[3],
" AL DETALLE"), type = "l")

### CENTRANDO SERIES TEMPORALES:
medias <- apply(multiST, MARGIN =  2, FUN = mean)

for (i in 1:length(nombres)) {
  multiST[ , i] <- multiST[ , i] - medias[i]
}

### GRAFICANDO SERIES TEMPORALES CENTRADAS:
windows()
par(mfrow = c(3, 1))
for (i in 1:3) {
  plot(multiST[ , i], main = paste0("CANCION N ", nombres[i]))
}

### EVALUANDO ESTACIONARIEDAD MEDIANTE DICKEY FULLER:
r.test <- c()
for (i in 1: length(nombres)) {
  r.test[i] <- adf.test(multiST[ , i], k = 0)$p.value
}

#### DESARROLLANDO ACP PARA SERIES TEMPORALES: ####

### DECLARANDO EL NUMERO DE OBSERVACIONES DE LAS SERIES:
n <- length(multiST[ , 1])

### ESTIMANDO PERIODOGRAMAS PARA CADA SERIE:
Per <- abs(mvfft(multiST))^2 / n

### ENCONTRAR PICOS MAXIMOS EN CADA CANCION:
```

```
picos <- c()

for (i in 1:length(nombres)) {
  picos[i] <- which.max(Per[ , i])
}

picoMed <- mean(picos)
picosFreq <- picos / 3083

### GRAFICANDO PERIODOGRAMAS ESTIMADOS EN EL DOMINIO DE LOS CICLOS:
windows()
par(mfrow = c(2,4), mar = c(3,2,2,1), mgp = c(1.6,.6,0), oma = c(0,1,0,0))
for (i in 1:8){
  plot(0:350, Per[1:351, i], type = "l", ylim = c(-1000, 400000),
  main = paste0("CANCION N ", nombres[i]), xlab = "Ciclos", ylab = "",
  xaxp = c(0, 350, 7))
  abline(v = picos[i], col = "red")
}
mtext("Periodograma", side = 2, line = -.3, outer = TRUE, adj = c(.2, .8))

### GRAFICANDO PERIODOGRAMAS ESTIMADOS EN EL DOMINIO DE LAS FRECUENCIAS:
windows()
par(mfrow = c(2,4), mar = c(3,2,2,1), mgp = c(1.6,.6,0), oma = c(0,1,0,0))
for (i in 1:8){
  plot((c(0:350) / 3083), Per[1:351, i], type = "l", ylim = c(-1000, 400000),
  main = paste0("CANCION N ", nombres[i]), xlab = "Frecuencia", ylab = "")
  abline(v = picosFreq[i], col = "red")
}
mtext("Periodograma", side = 2, line = -.3, outer = TRUE, adj = c(.2, .8))

### GRAFICANDO PERIODOGRAMAS CON MVSPEC:
windows()
mvspec(multiST, kernel("daniell", c(1, 1)), taper = .5, plot = TRUE,
main = "PERIODOGRAMAS DE LAS CANCIONES")

### OBTENCION DE MATRIZ DE DE DENSIDAD ESPECTRAL HERMITIANA:
fxx <- mvspec(multiST, kernel("daniell", c(1, 1)), taper = .5,
plot = FALSE)$fxx

### OBTENCION DE VALORES PROPIOS:
l.val <- rep(NA, 350)

for (k in 1:350) {
  u <- eigen(fxx[ , , k], symmetric = TRUE, only.values = TRUE)
  l.val[k] <- u$values[1]
}

### LOCALIZANDO VALORES MAXIMOS:
```

```
valMax <- which.max(l.val)
maxFreq <- valMax / 3083

valores <- l.val
valoresMax <- c()
valoresMax[1] <- which.max(valores)
valores[valoresMax[1]] <- 0

for (i in 2:20){
  valoresMax[i] <- which.max(valores)
  valores[valoresMax[i]] <- 0
}

frecuenciasMax <- valoresMax / 3083

### CREANDO GRUPOS:
grupo1 <- valoresMax[c(1, 2, 3, 5, 9, 16)]
grupo2 <- valoresMax[c(18, 19)]
grupo3 <- valoresMax[c(4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 17, 20)]
grupo1f <- grupo1 / 3083
grupo2f <- grupo2 / 3083
grupo3f <- grupo3 / 3083

### GRAFICANDO PRIMER COMPONENTE EN EL DOMINIO DE LOS CICLOS:
windows()
plot(l.val, type = "n", xaxt = "n", xlab = "Ciclos (Frequencia x 3083)",
ylab = "Primer componente principal")
axis(1, seq(0, 350, by = 50))
grid(lty = 2, nx = NA, ny = NULL)
abline(v = seq(0, 350, by = 50), col = 'lightgray', lty = 2)
lines(l.val)
abline(v = grupo1, col = "red")
abline(v = grupo2, col = "blue")
abline(v = grupo3, col = "yellow")

### GRAFICANDO PRIMER COMPONENTE EN EL DOMINIO DE LA FRECUENCIA:
windows()
plot((c(1:350) / 3083), l.val, type = "l",
xlab = "Frecuencia (Ciclos / 3083)", ylab = "Primer componente principal")
abline(v = grupo1f, col = "red")
abline(v = grupo2f, col = "blue")
abline(v = grupo3f, col = "yellow")

### EXPLORANDO LA FRECUENCIAS DESTACADAS Y HALLANDO PROPORCION DE PODER
### EXPLICADA POR LAS COMPONENTES EN CADA UNA:
proporciones = matrix(0, length(valoresMax), 6)

u <- eigen(fxx[ , , valoresMax[1]], symmetric = TRUE)
```

```
lam <- u$values
evec <- u$vectors

p1 <- lam[1] / sum(lam)
p2 <- lam[2] / sum(lam)
p3 <- lam[3] / sum(lam)
p4 <- lam[4] / sum(lam)
p5 <- lam[5] / sum(lam)

proporciones[1, ] <- c(valoresMax[1], p1, p2, p3, p4, p5)

for (i in 2:length(valoresMax)) {
  u <- eigen(fxx[ , , valoresMax[i]], symmetric = TRUE)
  lam <- u$values
  evec <- u$vectors

  p1 <- lam[1] / sum(lam)
  p2 <- lam[2] / sum(lam)
  p3 <- lam[3] / sum(lam)
  p4 <- lam[4] / sum(lam)
  p5 <- lam[5] / sum(lam)

  proporciones[i, ] <- c(valoresMax[i], p1, p2, p3, p4, p5)
}

proporciones[ , 1] <- proporciones[ , 1] / 3083
colnames(proporciones) <- c("Frecuencia", "Componente 1", "Componente 2",
"Componente 3", "Componente 4", "Componente 5")

}
```