Data Mining on Cardiac Excitation-Contraction Coupling Data

myotwin GmbH

Syed Shameer Sarwar

Matriculation number: 20631673
syed.sarwar@stud.uni-goettingen.de

Supervisor
Prof. Dr. Constantin Pape (University)
Dr. rer. nat. Hendrik Windel (Industry)

M.Sc. Applied Data Science

Institute of Computer Science
Georg-August University of Göttingen
17. March 2025

I hereby declare that I have composed this thesis independently and without unauthorized aid, using only the sources and materials listed in the bibliography. All sources and quotations are clearly acknowledged and correctly cited according to academic standards.

Göttingen, 17. March 2025

Abstract

In recent years, three-dimensional (3D) cardiac organoids have emerged as a powerful tool in preclinical drug development, enabling more physiologically relevant disease modeling and compound screening. Unlike traditional high-throughput screening methods, which typically rely on single-parameter endpoints, myotwin leverages high-content screening (HCS) to capture comprehensive multimodal data streams. This thesis establishes a data-driven analytical framework designed to integrate these multimodal datasets, facilitating a detailed characterization of pharmacological effects on cardiac excitation-contraction coupling (ECC) in human-induced pluripotent stem cell-derived Myocardial tissues (hiPSC-MTs). Multimodal datasets—including electrophysiological (field potential), calcium transient, and contractile force signals—were acquired under interventions with E-4031, Nifedipine, and Calcium Titration. A preprocessing pipeline enabled noise reduction, baseline correction, and arrhythmia removal, followed by extraction of 18 biologically relevant features. Advanced analytics included t-SNE visualization, mixed-effects modeling, Hill equation dose-response fitting, and LASSO regression. Results demonstrated E-4031-induced prolongation of field potential duration, Nifedipine-mediated reduction in contractile force and calcium transient amplitudes, and calcium-dependent restoration of contraction strength. Correlation and regression analyses revealed conserved and treatment-specific feature associations. However, the study's scope is constrained by limited tissue replicates, which restricts statistical power and broad applicability. Additional limitations, such as sparse concentration sampling and parameter-sensitive denoising and peak detection heuristics, further underscore the need for more robust methods. Overall, this work advances preclinical cardiac research by integrating signal processing, statistical learning, and multimodal data analysis to characterize treatment-induced ECC modulation.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Cardiovascular diseases remain a major global health issue, driving the need for advanced in vitro models to study cardiac physiology and pharmacological responses accurately. Traditionally, animal models have dominated preclinical cardiac research; however, their translational utility is constrained by inherent species-specific differences in ion channel expression, calcium handling, and contractile mechanics [1]. To overcome these limitations, myocardial tissues (MTs) derived from human-induced pluripotent stem cells (hiPSC) have emerged as promising alternatives that provide a more human-relevant context for studying cardiac excitation-contraction coupling (ECC) dynamics [2]. Nevertheless, rigorous validation of their physiological fidelity, particularly regarding electrical, calcium handling, and mechanical properties, remains critical for establishing hiPSC-MTs as reliable platforms for disease modeling and drug screening.

1.1   Research Problem

Despite significant advancements, current pharmacological studies utilizing hiPSC-MTs often examine excitation-contraction coupling (ECC) parameters in isolation, neglecting critical interdependencies among electrophysiological, calcium, and mechanical signals. Additionally, there is a lack of standardized analytical pipelines capable of systematically preprocessing noisy multimodal data and performing robust statistical analyses. Existing research also inadequately addresses variability arising from differences in cell lines, differentiation methods, and culture conditions. This thesis aims to bridge these methodological gaps by developing a concise, data-driven analytical framework that integrates advanced signal processing, comprehensive feature extraction, and rigorous statistical modeling to enhance the reliability and interpretability of hiPSC-MT-based cardiac drug screening.

1.2   Research Objectives

The central goal of this research is to establish and validate a reproducible analytical pipeline to characterize treatment-induced modulation of ECC dynamics in hiPSC-MTs. Specifically, this study aims to:

- **Extract Biologically Relevant Features:** Implement preprocessing workflows to denoise multimodal signals (field potential, calcium, and force signals), detect and remove arrhythmic contractions, and reliably extract interpretable ECC biomarkers (e.g., field potential duration, contraction amplitudes).

- **Modeling & Validating Treatment Responses**: Quantify and validate pharmacological effects available in existing literature through robust statistical modeling and dose-response analysis.

- **Generate Actionable Insights**: Identify conserved and treatment-specific feature associations to inform mechanistic understandings and guide future MTs based drug screening studies and experimental designs.

1.3 Scope of the Study

This thesis examines the response of hiPSC-derived MTs subjected to three pharmacological stimuli: E-4031, Nifedipine, and Calcium Titration. Key elements of the scope include:

- **Preprocessing Pipeline**: Custom workflow for signal denoising, baseline correction, arrhythmia detection and removal, and the extraction of 18 biologically meaningful ECC biomarkers (see Chapter 4).

- **Multidimensional Analysis**: t-SNE visualization, mixed-effects statistical modeling, correlation trend analysis, principal component analysis, and regression modeling to identify pharmacological impacts. (Chapter 5)

- **Biological Validation**: Comparative analyses with known physiological drug mechanisms to evaluate hiPSC-MT fidelity (Sections 5.2.1-5.2.3).

The study deliberately excludes molecular-level mechanistic modeling, instead focusing exclusively on empirical, tissue-level functional responses to pharmacological treatments.

1.4 Structure of the Thesis

The thesis is structured to progressively address the research objectives, ensuring methodological rigor and interpretability:

- **Chapter 2 (Literature Review)**: Critically evaluates advancements in hiPSC-MT characterization and identifies gaps in multimodal data integration. Establishes the necessity of a unified preprocessing and analysis framework.

- **Chapter 3 (Experimental Setup)**: Details data acquisition process (Section 3.2) and pharmacological interventions (Section 3.3). Provides the foundational dataset for downstream analysis.

- **Chapter 4 (Data Preprocessing and Feature Extraction)**: Develops signal enhancement techniques (Sections 4.1) and arrhythmia classification and removal (Section 4.3). These steps are critical for ensuring data quality before feature extraction (Section 4.4), which directly supports Objective 1.

- **Chapter 5 (Data Mining and Analysis [Objective 2 and 3])**:

- *Global Outlook* (Section 5.1): t-SNE visualization provides a unified view of feature space.
- *Statistical Modeling* (Section 5.2): Validates hiPSC-MT physiological relevance through hypothesis-driven mixed-effects modeling.
- *Drug Response Modeling, Relative Features Comparison and PCIA* (Section 5.3): Quantifies dose-response relationships, establishes relative feature changes and identifies similarly varying feature clusters via Principal Components Influence Analysis (PCIA).
- *Correlation Analysis* (Section 5.4) and *Feature Associations* (Section 5.5): Uncover conserved and treatment-specific interactions through exploration of correlation trends and application of mixed modeling based LASSO regression.

- **Chapter 6 (Integrative Discussion)**: Synthesizes findings, discusses methodological limitations (e.g., empirical denoising thresholds), and proposes future directions for improving hiPSC-MT-based drug screening.

This structure ensures a logical progression from raw data acquisition to actionable insights, with each chapter building on the prior to address the research objectives comprehensively. By integrating signal processing, statistical learning, and pharmacological validation, this work advances preclinical cardiology toward more precise cardiac drug development through the application of MT-based assays.

Chapter 2

Literature Review

Human-induced pluripotent stem cell-derived cardiomyocytes (hiPSC-MTs) have significantly transformed preclinical cardiac drug screening methodologies, enabling the testing of compounds in human-relevant, patient-specific contexts. Unlike traditional animal models, hiPSC-MTs allow researchers to examine pharmacological effects in vitro, significantly reducing reliance on clinical trials for initial efficacy and safety evaluation [3, 4]. Despite these advantages, hiPSC-MTs still face critical limitations, primarily related to incomplete maturation and differentiation variability, posing challenges in reliably interpreting drug responses [5].

Advances in hiPSC-CM-based studies have increasingly integrated multimodal physiological measurements—electrical activity (field potentials), calcium transients, and contractile force—to better understand excitation-contraction coupling (ECC). Multimodal data acquisition platforms such as Multi-Electrode Arrays (MEAs) coupled with impedance and optical voltage imaging have improved the detection and characterization of drug-induced arrhythmias, surpassing conventional single-modal methods [6]. Recent studies have developed sophisticated software tools like CardioMDA, which enhances the analysis of field potentials from hiPSC-MTs through semi-automated processes that utilize correlation analyses and ensemble averaging techniques [7]. This software significantly improves accuracy, reliability, and throughput, thereby addressing limitations in manual analysis methods previously encountered in electro-physiological screening assays. Similarly, innovative studies have successfully utilized flexible electronics integrated with engineered cardiac cell sheet tissues to simultaneously measure electrophysiological parameters and contractile force. Such simultaneous measurements of electrical and mechanical activity provide more accurate insights into excitation-contraction coupling dynamics and potential drug-induced cardiac dysfunction [8].

Several specific hypotheses concerning drug effects have been validated using hiPSC-CM models. For example, studies have demonstrated that the potassium channel blocker E-4031 reliably induces prolonged field potential duration (FPD), a hallmark of arrhythmogenic risk, validating hiPSC-CM platforms for screening pro-arrhythmic compounds. This validation has been clearly established using engineered 3D cardiac tissue sheets, reinforcing their potential as predictive in vitro models for Torsades de Pointes (TdP), a severe and clinically relevant arrhythmia [9]. Additionally, research involving the calcium channel blocker Nifedipine has validated the hypothesis that increasing drug concentrations result in reductions in contractile and relaxation times, as well as decreased calcium transient and contraction force amplitudes[10]. These findings confirm the predictive accuracy of hiPSC-CM models for assessing cardiac contractility effects in response to

pharmacological intervention.

However, systematic analytical approaches that comprehensively integrate these multimodal signals remain relatively underdeveloped. Additionally, most existing ECC-focused pharmacological studies have limited analyses concerning inter-feature relationships among extracted functional parameters. ECC biomarkers, including field potential duration, calcium transient amplitude, and contractile force, exhibit complex interdependencies that may influence drug response interpretations. Therefore, comprehensive analyses that clearly characterize these interrelationships and their variability across treatments are essential but currently sparse [11].

Moreover, engineered heart tissues inherently exhibit variability arising from genetic differences among donors, differentiation protocols, and culture conditions. This variability significantly influences drug responses [12], yet existing studies rarely incorporate robust statistical methods to control for such heterogeneity. Recent literature advocating for the use of mixed-effects modeling highlights its utility in accounting for experimental site-to-site variability, demonstrating enhanced reliability of pharmacological conclusions in multi-site drug testing setups [11].

This thesis specifically addresses these methodological gaps by establishing a robust, data-driven analytical framework. It incorporates advanced preprocessing techniques for signal denoising, arrhythmia detection, and biologically meaningful feature extraction. The developed pipeline systematically integrates multimodal ECC data, facilitating comprehensive statistical analyses that include correlation studies, principal component analysis, and mixed-effects regression modeling. By employing mixed-effects modeling, this study explicitly accounts for the variability inherent in hiPSC-CM datasets, providing statistically validated insights into drug-induced modulation of ECC biomarkers. Consequently, this thesis not only fills existing gaps but also enhances the interpretative power and scalability of hiPSC-CM-based drug screening platforms, supporting more precise pharmacological evaluations.

Chapter 3

Experimental Setup

3.1 Cardiac Excitation-Contraction Coupling (ECC) Cycle

The excitation-contraction coupling (ECC) cycle in cardiomyocytes is a fundamental process that drives the heart's rhythmic contractions. It is initiated by an electrical impulse, known as an action potential, which propagates across the cardiomyocyte membrane, triggering a series of intracellular events. This depolarization activates voltage-gated L-type calcium channels, allowing extracellular calcium influx. This influx subsequently stimulates the release of a larger quantity of calcium ions from the sarcoplasmic reticulum (SR), a process known as calcium-induced calcium release (CICR). After calcium-induced calcium release (CICR), the transient rise in intracellular calcium triggers muscle contraction by enabling actin-myosin interactions. Calcium is then reabsorbed into the sarcoplasmic reticulum (SR) and expelled from the cell to restore baseline levels, allowing muscle relaxation.[13]. Understanding these cellular mechanisms is critical for designing experimental setups that accurately capture the interplay of electrical, calcium, and mechanical activity in cardiomyocytes.

3.2 Data Acquisition

The myotwin's data acquisition system used in this study was designed to capture multimodal physiological measurements during cardiomyocyte contractions. It consisted of three key components:

- **Multi-Electrode Arrays (MEAs)**: MEAs were employed to record extracellular field potentials, providing insights into the electrical activity of cardiomyocyte networks. A total of 32 microelectrodes were positioned at different spatial locations within the MT to capture the collective action potential propagation. The sampling frequency for MEA recordings was set at $2\,\mathrm{kHz}$, ensuring high temporal resolution for detecting rapid electrophysiological events.

- **Fluorescence Microscopy**: This imaging technique, combined with a calcium-sensitive fluorescent dye, was used to visualize and quantify intracellular calcium transients in cardiomyocytes. Image sequences were captured at $100\,\mathrm{Hz}$ using a fluorescence microscope, and regions of interest (ROIs) were manually selected for analysis. The fluorescence intensity from these ROIs was extracted and normalized using the F/F_0 method to generate $1D$ calcium intensity time-series data, representing calcium handling in the tissue.

- **Isometric Force Sensors**: Isometric force transducers were used to measure contractile force generation by the cardiomyocytes. Force measurements were recorded at a sampling rate of $100\,\mathrm{Hz}$, allowing for the assessment of contractile strength and mechanical properties of the MT.

By integrating these three modalities—electrical (field potential), optical (calcium transient), and mechanical (contractile force)—the experimental setup provided a comprehensive characterization of cardiomyocyte function, facilitating a deeper understanding of ECC dynamics in MTs. With these multimodal recording techniques in place, the next step involves perturbing the system using pharmacological stimulus to observe changes in ECC dynamics.

3.3   Pharmacological Experiments and Dataset Generation

The three physiological parameters—**Field Potential (mN)**, **Calcium Intensity (a.u.)**, and **Contractile Force (mN)** were measured simultaneously in 60-second[1] trials, capturing multiple contraction cycles. These measurements were conducted under three different pharmacological conditions, each applied at varying concentration levels:

1. **E-4031**: A drug that blocks a specific potassium channel responsible for repolarization of the action potential. By prolonging repolarization, E-4031 disrupts the normal timing of ECC, leading to abnormal calcium handling and potential arrhythmias.

2. **Nifedipine**: A calcium channel blocker that inhibits voltage-gated L-type calcium channels, reducing the influx of extracellular calcium. This directly limits calcium-induced calcium release (CICR) from the sarcoplasmic reticulum, leading to a weaker contraction and impaired actin-myosin cross-bridge formation.

3. **Calcium Titration**: An experiment designed to assess the role of extracellular calcium in the ECC cycle. Initially, calcium is removed from the surrounding medium, preventing calcium entry through L-type channels and thereby suppressing contraction. It is then gradually reintroduced, allowing for controlled observation of how calcium influx influences CICR, myofilament activation, and contractile force. This experiment validated the MTs' functional integrity by assessing their contractile responsiveness to varying extracellular calcium levels, which is crucial for verifying the fidelity of excitation-contraction coupling.

Baseline measurements were recorded before treatment application to serve as a reference for evaluating pharmacological effects. Each experimental condition was tested on two different tissue samples. However, data was unavailable for both tissues due to experimental constraints for some concentration levels. Figure 3.1a highlights a single contraction event, while Figure 3.2 presents the averaged contraction trends, offering insight into how different treatment concentrations applied in this study modulate tissue response.

---

[1]All subsequent plots in this thesis display time axes relative to the start of each experiment.
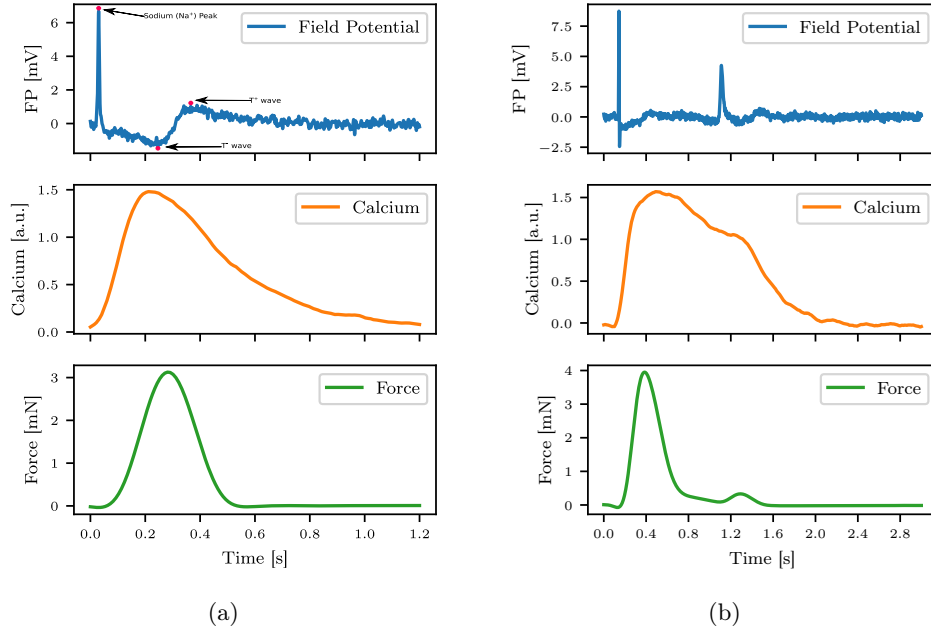
Figure 3.1: **Contraction Types (a) Non-arrhythmic contraction**: Synchronized view of a single contraction event across modalities: (i) Field potential (Na^+ peak, T^-, T^+ waves), (ii) Calcium transient, and (iii) Force signal. **(b) Arrhythmic Contraction**: Irregular contractions, excluded from analysis, characterized by multiple force peaks, abnormal calcium decay, or disrupted field potential morphology.

With multimodal data collected under pharmacological perturbations, the next step involves preprocessing to enhance signal fidelity. This includes filtering noise, identifying contraction events, and removing arrhythmias (Figure 3.1b). The following chapter details the preprocessing workflow and feature extraction techniques used to derive biologically meaningful insights from the recorded signals.
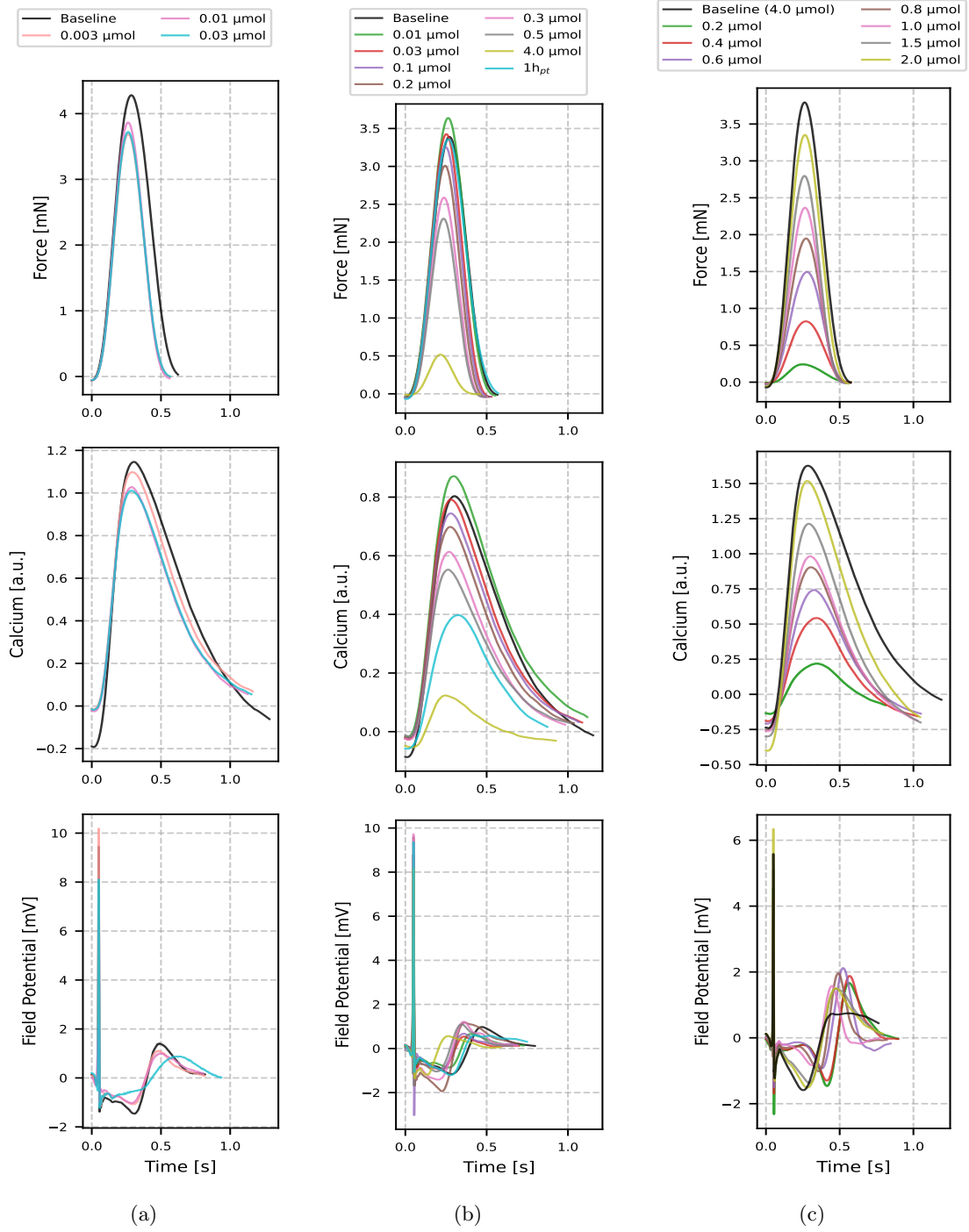
Figure 3.2: **Average Tissue Response in Pharmacological Experiments. (a) E-4031**: Averaged contraction signals recorded under baseline conditions and after E-4031 treatment, presented for three modalities (Force, Calcium and Field Potential) along a shared time axis. **(b) Nifedipine**: Tissue contraction response under baseline and Nifedipine treatment. **(c) Calcium Titration**: Contraction dynamics observed across different calcium concentrations.

Chapter 4

Data Preprocessing and Feature Extraction

After acquiring raw multimodal data, a structured preprocessing pipeline was applied to improve signal fidelity and remove systematic noise. Once preprocessed, contraction events were identified using peak detection algorithms, followed by arrhythmia classification to exclude irregular contractions. Finally, biologically relevant features were extracted from each contraction event to facilitate downstream analysis. The peak detection methodology, originally developed in an arrhythmia classification study [14], was refined in this work, particularly for Field Potential (FP) signal. The following sections detail these preprocessing and feature extraction steps.

4.1 Preprocessing and Signal Enhancement

Each signal in functional MT data i.e Force, Calicum and FP underwent a series of preprocessing steps to improve data quality and eliminate systematic noise.

4.1.1 Force Signal Preprocessing

1. **Systematic Noise Removal**
   - High-frequency noise, primarily originating from external interference, was identified through frequency-domain analysis.
   - A 5th-order Butterworth low-pass filter was applied, with the cutoff frequency set to eliminate components exceeding the heart's physiological frequency limit ($3.6\,\text{Hz}$) while preserving critical contraction-related components.

2. **Baseline Offset Correction**
   - The force signal often contains an inherent baseline shift due to sensor drift or passive tissue tension.
   - A histogram-based approach is used, where a Gaussian curve is fitted to the histogram peak for approximating the noise density.

- The mean of the fitted Gaussian is taken as the offset value and subtracted from the entire signal, ensuring accurate force measurements (Figure 4.1a).

3. **Smoothing and Noise Reduction**

- A Savitzky-Golay (SavGol) [15] filter is applied to the baseline-corrected signal for smoothing while preserving peaks and critical features associated with cardiomyocyte contractions.

The final preprocessed force signal is **filtered, baseline-corrected, and smoothed** (Figure 4.1b), providing a robust representation of contractile behavior. For a more detailed discussion of the methodology, refer to [14, Section 3.1].

(a) (b)
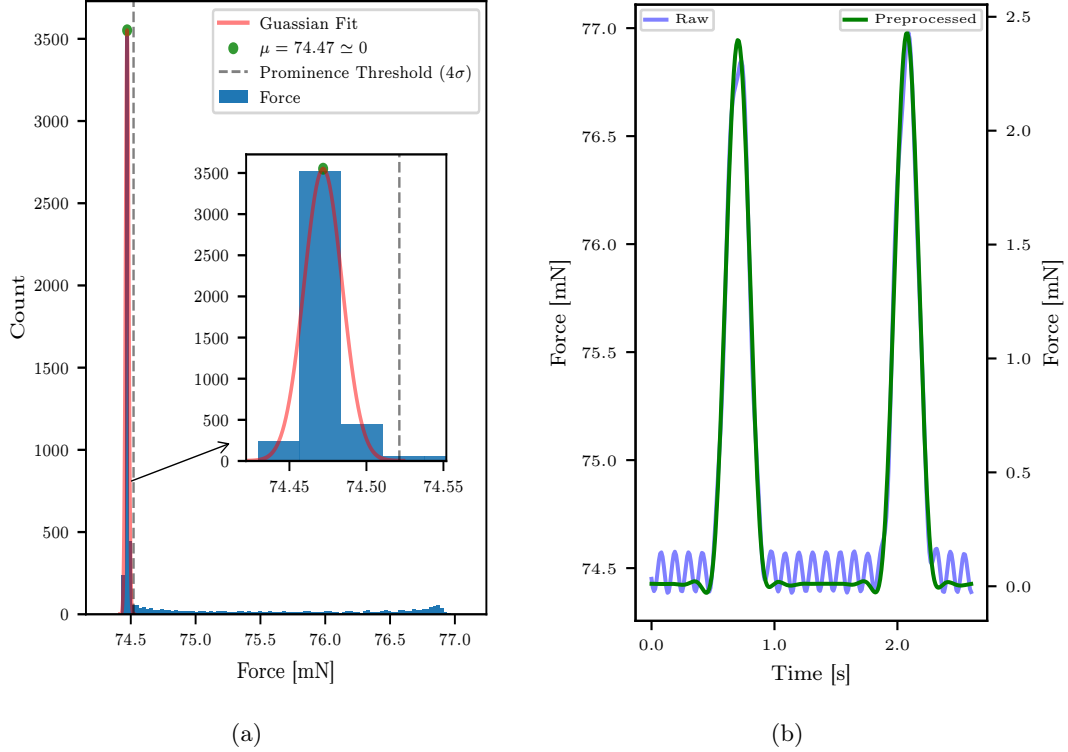
Figure 4.1: **Force Preprocessing. (a)** Histogram of the force signal with a fitted Gaussian curve centered at the histogram peak. The Gaussian mean, representing the baseline offset, is subtracted from the signal during baseline correction to achieve a zero-mean baseline. A 4σ threshold from this fit is then used as the prominence threshold for peak detection (see Section 4.4.1). **(b)** Comparison of the raw force signal (blue) and its preprocessed version (green), demonstrating the combined effects of noise removal, baseline correction, and smoothing on enhancing the signal quality.

4.1.2 Calcium Signal Preprocessing

1. **Baseline Offset Correction**
   - The calcium signal exhibited a sloped baseline or pedestal offset, requiring correction to ensure accurate intracellular calcium measurements.
   - A *two-step correction process* is applied:
     - First, a method similar to force signal correction is used, where a Gaussian curve is fitted around the peak corresponding to baseline noise in the histogram. The mean of the fitted Gaussian serves as the initial estimated baseline.
     - To refine this estimation, a Random Sample Consensus (RANSAC) algorithm is applied. RANSAC fits a regression line to data points around the initial estimated baseline by utilizing the standard deviation of the fitted guassian, ensuring a more robust correction (Figure 4.2a).
   - The refined baseline estimate is then subtracted from the entire calcium signal to achieve accurate baseline correction.

2. **Smoothing**
   - A Savitzky-Golay (SavGol) filter is applied to the baseline-corrected signal for further smoothing of the signal which enhances the signal-to-noise ratio.

3. **Time-Axis Correction**
   - A time-axis adjustment is performed to correct for a hardware-induced delay affecting calcium signal recordings.
   - This step ensures accurate synchronization with other signals, allowing for more precise multimodal analysis.

The final preprocessed calcium signal is **baseline-corrected, smoothed, and time-adjusted** (Figure 4.2b). A more detailed explanation of these preprocessing methods can be found in [14, Section 3.3].

4.1.3 Field Potential Signal Preprocessing

The preprocessing of Field Potential (FP) signal was performed in two stages. The initial preprocessing, conducted as part of prior work [14], focused on identifying high-quality signal channels and applying basic denoising techniques. This involved classifying channels based on signal quality, averaging selected channels to enhance stability, and applying a low-pass filter for noise reduction. Although initial preprocessing enhanced signal clarity, residual high-frequency noise and channel variability still affected the data. To address this, advanced denoising techniques, including Discrete Wavelet Transform (DWT) decomposition, were applied in this thesis to further refine signal quality and improve robustness in downstream analysis.

4.1.3.1 Channel Quality Classification and Averaging

- **Channel Quality Classification**
  - As stated in Section 3.2, FP recordings consisted of signals from 32 MEA channels, each exhibiting varying signal quality.
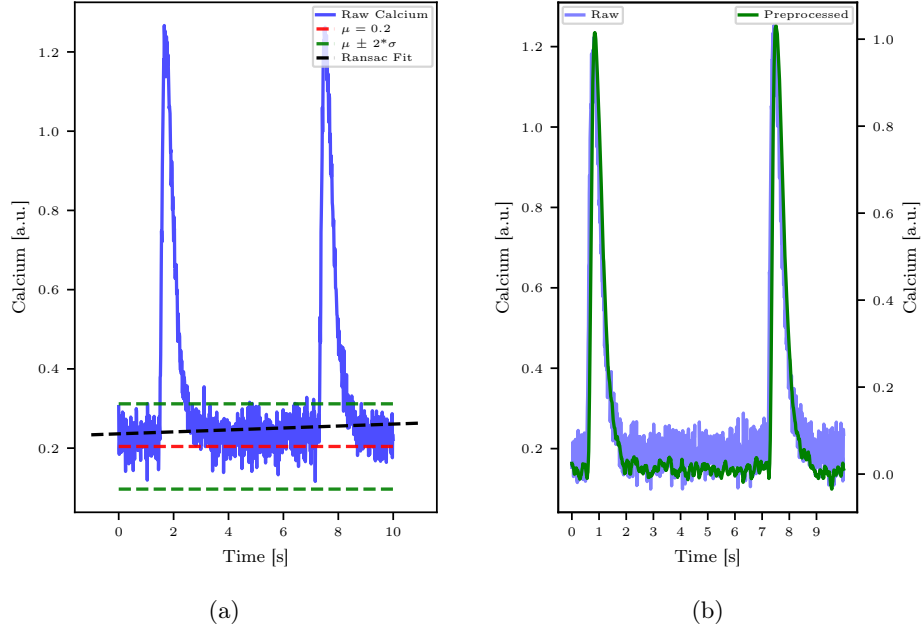
Figure 4.2: **Calcium Preprocessing (a)** A 10-second segment of the calcium signal with vertical lines indicating the mean (red) and ± 2 standard deviations (σ, green) of the fitted gaussian distribution around signal's histogram maximum. A RANSAC linear fit (black) is applied to the data points within a region of $\mu \pm 2\sigma$ to estimate baseline slope and use it as subtraction offset for baseline correction. **(b)** Comparison of the raw calcium signal with its preprocessed version.

- The channels from a random subset of FP recordings were selected and manually labeled into three categories: **bad, medium** and **good**, based on their signal characteristics (Figure 4.3a).
- Statistical features such as *mean, median, mode, standard deviation*, and various other time and frequency domain descriptors were extracted from each channel using the **Time Series Feature Extraction Library (TSFEL)** [16].
- A Random Forest classifier with 'balanced' class weights (due to high class imbalance) was trained on these extracted features to learn the mapping between statistical features and channel quality.
- The trained model achieved a high classification accuracy, effectively distinguishing high-quality channels for further analysis.

- **Selection and Averaging of Good Channels**
  - Only channels classified as "good" were retained for further analysis.
  - The signals from all good channels were averaged to obtain a single representative FP signal.

- **Initial Noise Filtering**
  - A low-pass filter was applied to the averaged signal, preserving key features of the FP

signal while attenuating high-frequency noise.

For a more detailed explanation of these preprocessing steps, refer to *Section 3.4* of previous work [14]. An averaged FP signal is shown in Figure 4.3b

(a)

(b)
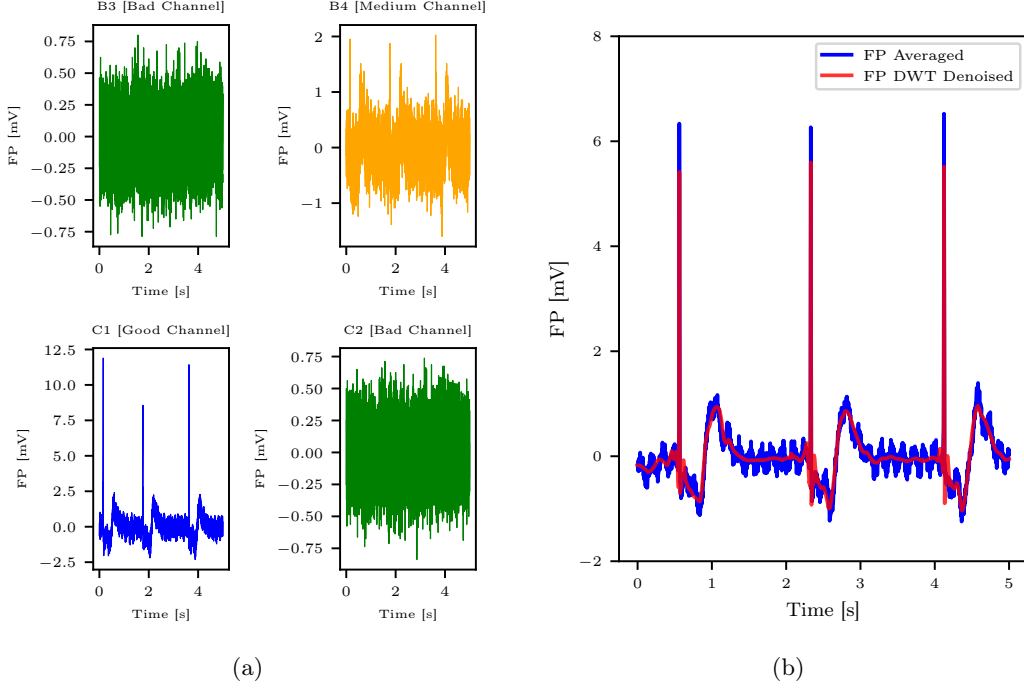
Figure 4.3: **Field Potential (FP) Signal Processing. (a) MEA Channel Quality Classification**: Channels are categorized as (i) Good (blue), capturing all aspects of the FP signal; (ii) Medium (orange), capturing peaks despite higher noise levels; and (iii) Bad (green), dominated by noise. **(b) Processed FP**: Comparison of the averaged and low-pass filtered FP signal (blue) from high-quality channels with its DWT-denoised version (red), illustrating the improvement in signal clarity.

4.1.3.2 DWT-Based Denoising

In this thesis, further FP signal denoising was performed to enhance signal quality. Despite the initial noise reduction, high-frequency noise was still present, as seen in Figure 4.3b. To address this, Discrete Wavelet Transform (DWT) [17] was applied to decompose the signal into different frequency components and selectively remove noise.

The frequency composition of the FP was first analyzed using the Fourier Transform (Figure 4.4b), which revealed the dominant signal energy components within the **0-200 Hz range**, as components outside this range were attenuated through the low-pass filter (Section 4.1.3.1). This information guided the denoising approach by selectively modifying the wavelet coefficients corresponding to high-to-medium frequency components.

DWT decomposition was performed using the **Daubechies-4 ('db4') wavelet** with whole-sample symmetric padding (implemented as 'reflect' in PyWavelets [18]). This decomposition method separates the signal into **approximation coefficients (A_n)**, which capture low-frequency trends, and **detail coefficients (D_n)**, which represent high-frequency variations. Each level of decomposition halves the frequency range of the previous level, with the first level spanning 500-1000 Hz, the second level covering 250-500 Hz, and so on. At each step, only the detail coefficients (D_n) are further decomposed, while the approximation coefficients remain unchanged. The decomposition continued up to **14 levels**, starting from the **Nyquist frequency ($F_s/2 = 1000$ Hz)** where F_s is the sampling rate of FP signal.

Table 4.1: Decomposition and Transformation of Averaged FP Signal using DWT

Level	Frequency Range (Hz)	Transformation Applied
1	500 - 1000	Detail Coefficients Removed ($D_1 = 0$)
2	250 - 500	Detail Coefficients Removed ($D_2 = 0$)
3	125 - 250	Soft Thresholding ($\lambda = 0.5$)
4	62.5 - 125	Soft Thresholding ($\lambda = 0.5$)
5	31.25 - 62.5	Soft Thresholding ($\lambda = 0.5$)
6	15.62 - 31.25	Soft Thresholding ($\lambda = 0.5$)
7	7.81 - 15.62	Soft Thresholding ($\lambda = 2.5$)
8	3.90 - 7.81	Soft Thresholding ($\lambda = 2.5$)
9 - 14	< 3.9	No Transformation Applied

Denoising was performed by modifying the detail coefficients while preserving the essential signal structure. The high-frequency detail coefficients at levels 1 and 2 were completely removed ($D_1, D_2 = 0$), eliminating components above 250 Hz. Soft thresholding was applied to the detail coefficients of levels 3 to 8. A threshold of $\lambda = 0.5$ for levels 3 to 6 and a higher threshold of $\lambda = 2.5$ for levels 7 and 8 to suppress low-frequency fluctuations. These threshold values were selected based on empirical evaluation, testing different values and choosing those that visibly demonstrated optimal noise suppression while preserving signal integrity. Levels 9 to 14, corresponding to frequencies below 3.9 Hz, were retained as they contribute to slow moving trends.

Soft thresholding was applied using the following transformation:

$$D_n'(x, \lambda) = \begin{cases} \text{sign}(x) \cdot (|x| - \lambda) & \text{if } |x| > \lambda \\ 0 & \text{otherwise} \end{cases}$$

where x represents the wavelet coefficients, and λ is the threshold value. This function reduces small fluctuations while preserving larger signal components, effectively suppressing noise without distorting the true signal. The frequency breakdown and transformations applied at each level are summarized in Table 4.1. Figure 4.4a shows exemplary levels to illustrate the effect of thresholding transformations on detail coefficients.

After modifying the wavelet coefficients, the Inverse Discrete Wavelet Transform (IDWT) was applied to reconstruct the denoised signal. The comparison of the original (averaged FP) and DWT denoised signal is presented in Figure 4.3b, demonstrating the removal of high-frequency noise and the preservation of the essential FP signal features.

Additionally, the Fourier Transform of the original and denoised signals (Figure 4.4b) confirms the suppression of noise components, validating the effectiveness of the denoising process. The preprocessed FP signal obtained after denoising was used for subsequent analysis.
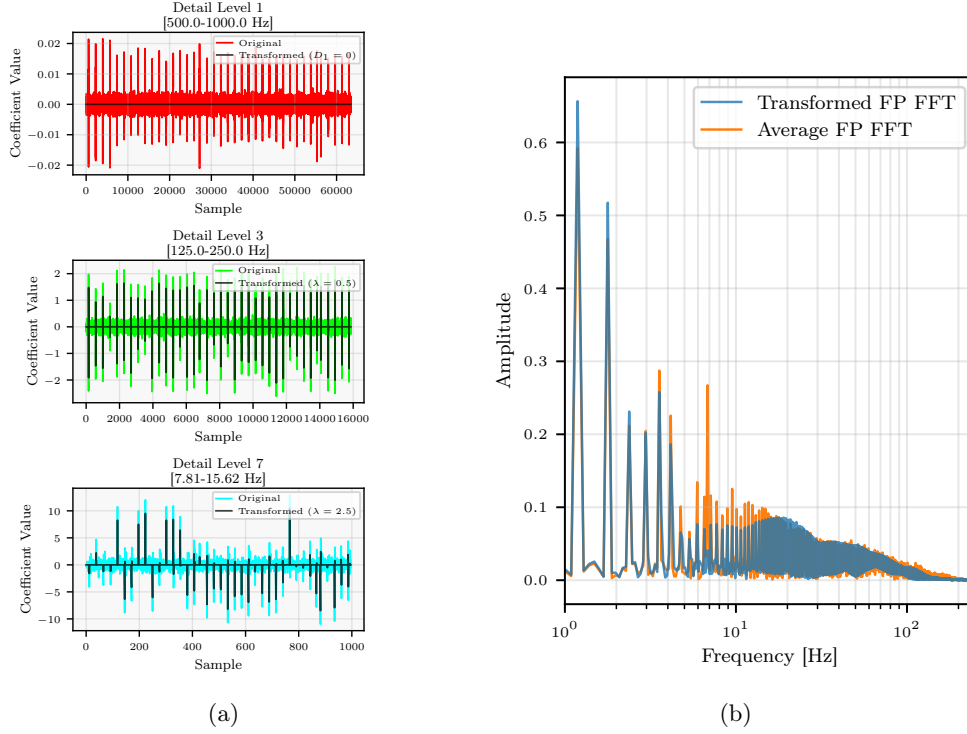
Figure 4.4: **(a) Effect of DWT-based Denoising on FP Signal**: Detail coefficients of the FP signal at selected decomposition levels: Level 1 (D_1) shows complete removal of coefficients for high-frequency noise reduction, Level 3 applies soft thresholding with $\lambda = 0.5$, and Level 7 applies soft thresholding with $\lambda = 2.5$ to suppress medium-to-low frequency noise. **(b) FFT of FP signals**: Comparison of the Fourier Transform (FFT) components for the original (averaged FP, orange) and DWT-denoised FP signal (blue) is shown, with the amplitude axis log-transformed to clearly illustrate the attenuation of noise components after denoising.

4.2   Identification of Contraction Events

After preprocessing, the next step was precise contraction event detection, which is critical for extracting biologically relevant features. This required accurate peak identification, as each detected peak corresponds to a distinct contraction. A prominence-thresholding-based peak detection logic was adapted from an arrhythmia classification study [14] to detect peaks in force and calcium signals. The high-level overview of the method is described in the following section. For the FP signal, a new method for deriving the prominence threshold was developed.

4.2.1   Force and Calcium Peak Detection

A peak's **prominence** quantifies its relative importance by measuring the vertical difference between the peak's apex and its lowest contour point. A higher prominence indicates a more distinct peak, making it a robust criterion for peak selection. To implement this, the `find_peaks`[19] method from SciPy signal processing module was utilized, with a dynamically determined prominence threshold tailored to each measurement.

To define an appropriate prominence threshold, the histogram of the preprocessed signal was analyzed. A Gaussian curve was fitted around the peak of the preprocessed signal. Baseline correction (Section 4.1.1 & 4.1.2) ensures that the dominant noise distribution is centered around zero, i.e $\mu = 0$. The standard deviation (σ) of this fitted Gaussian was utilized for determining prominence threshold, and peaks were retained if their prominence exceeded a threshold set at **+4 standard deviations**. This thresholding method is based on the statistical principle that approximately 99.9% of data points in a normal distribution lie within four standard deviations, ensuring that only the most relevant peaks are selected.

This prominence-based approach effectively identifies contraction peaks in both force and calcium signals, ensuring robust detection of contraction events (Figure 4.5). A more detailed explanation of this methodology can be found in [14, Section 4.1]
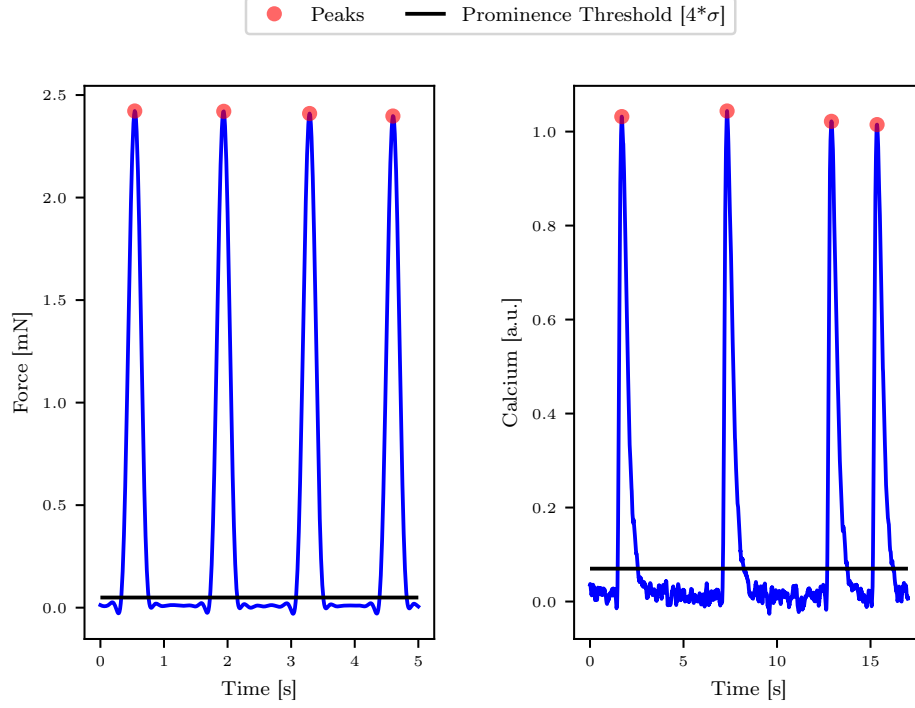
Figure 4.5: **Peak Detection via Prominence Threshold for Force and Calcium signals**: Using SciPy's `find_peaks`, force and calcium peaks are detected by applying a dynamic prominence threshold set at 4 standard deviations (σ) of the noise distribution, ensuring that only significant contraction events are identified.

4.2.2 Field Potential Peak Detection

To detect FP peaks, a Gaussian Mixture Model (GMM) [20] was fitted to the signal amplitude distribution. As shown in Figure 3.1a, a single contraction in the FP signal is characterized by a sharp **sodium** ($\mathbf{Na^+}$) peak, followed by two **T-waves**, where the negative component ($\mathbf{T^-}$) appears before the positive component ($\mathbf{T^+}$).

The histogram of FP signal (Figure 4.6a) exhibits three distinct regions:

- A sharp density peak around zero, indicative of baseline noise.

- A moderate-density negative peak, corresponding to the T^- wave.

- A lower-density positive peak, corresponding to the T^+ wave.

The remaining distribution represents the density of the high-amplitude sodium (Na^+) peak.

To model this distribution, a GMM with four components was fitted to the signal. This optimal number of components was selected using the Elbow criterion [21] applied to the Akaike Information Criterion (AIC) [22] score (Figure 4.6b). The fitted GMM parameters (mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$) were then used to derive prominence thresholds for peak detection.

Two different prominence thresholding methods were evaluated:

1. **Algorithm 1:** The Gaussian component with the highest mean ($\boldsymbol{\mu}_{\mathrm{hmean}}$) was selected, and the prominence threshold was set as:

$$\text{threshold} = \boldsymbol{\mu}_{\mathrm{hmean}} + 0.8 \cdot \boldsymbol{\sigma}_{\mathrm{hmean}}.$$

2. **Algorithm 2:** The sharpest Gaussian component (i.e., the one with the lowest standard deviation $\boldsymbol{\sigma}_{\mathrm{sharp}}$) was chosen, with the threshold defined as:

$$\text{threshold} = \boldsymbol{\mu}_{\mathrm{sharp}} + 7.5 \cdot \boldsymbol{\sigma}_{\mathrm{sharp}}.$$
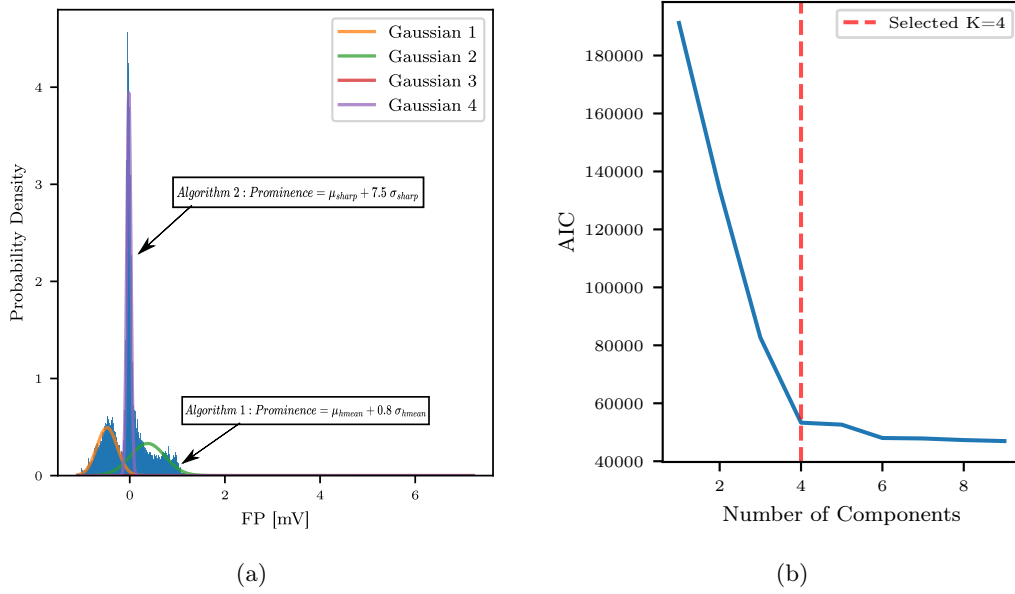
Figure 4.6: **Field Potential Peaks Detection Algorithms Based on GMM. (a) GMM-based Prominence Threshold Determination:** A Gaussian Mixture Model (K=4) is fitted to the denoised FP signal to model its amplitude distribution. Two thresholding methods are evaluated: Algorithm 1 sets the prominence threshold as $\mu_{\mathrm{hmean}} + 0.8\,\sigma_{\mathrm{hmean}}$ (green density), while Algorithm 2 uses the sharpest Gaussian component with a threshold of $\mu_{\mathrm{sharp}} + 7.5\,\sigma_{\mathrm{sharp}}$ (purple density). **(b) Optimal GMM Component Selection:** The optimal number of GMM components is determined using the AIC Knee criterion, which identifies K=4 as optimal for the FP signal.

Following peak detection, a post-processing step was applied to filter out multiple peaks occurring within a 0.05s window of each peak, retaining the absolute maximum signal value as the peak for that window. After filtering, each peak was categorized as Sodium (Na^+), T^-, or T^+ based on its half-width at half-maximum (HWHM) and amplitude.

Peak classification was performed based on the following criteria:

$$L(P) = \begin{cases} \mathrm{Na}^+, & \text{if } \mathrm{HWHM}(P) < 0.01\,\mathrm{s}, \\ \mathrm{T}^-, & \text{else if } \mathrm{Amplitude}(P) < 0, \\ \mathrm{T}^+, & \text{otherwise}, \end{cases} \tag{4.1}$$

where **HWHM** represents the peak width at half its maximum height. Sodium (Na$^+$) peaks are sharp with **HWHM < 0.01 s** (Figure 4.7), while T$^-$ and T$^+$ peaks are differentiated by their negative or positive amplitude. The complete peak detection and classification process is summarized in Procedure 1.
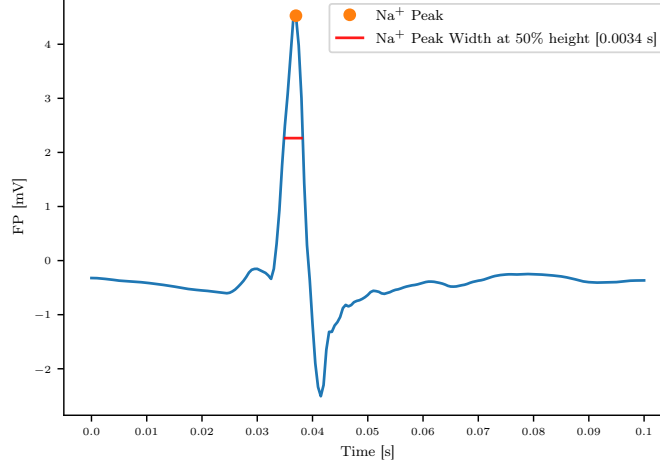
Figure 4.7: **Na$^+$ Peak Decision Factor.** The plot illustrates the criterion used for identifying Na$^+$ peaks in the FP signal. In this plot, Na$^+$ peak (shown in orange) is characterized by a half-width at half-maximum (HWHM) of less than 0.01 seconds (marked in red). This threshold serves as a decisive factor for distinguishing sharp Na$^+$ peaks.

Procedure 1 Field Potential Peak Detection

Input: FP signal, method $\in$ {'Algorithm1', 'Algorithm2'}
Output: Categorized peaks (Na$^+$, T^-, T^+)

1: Fit GMM with 4 components to the signal.
2: Extract Gaussian parameters ($\boldsymbol{\mu}, \boldsymbol{\sigma}$) from the fitted model.
3: **if** method == 'Algorithm1' **then**
4:     $hmean\_idx = \arg\max(\boldsymbol{\mu})$
5:     $\boldsymbol{\mu}_{\mathrm{hmean}}, \boldsymbol{\sigma}_{\mathrm{hmean}} = \boldsymbol{\mu}[hmean\_idx], \boldsymbol{\sigma}[hmean\_idx]$
6:     threshold $= \boldsymbol{\mu}_{\mathrm{hmean}} + 0.8 \cdot \boldsymbol{\sigma}_{\mathrm{hmean}}$
7: **else if** method == 'Algorithm2' **then**
8:     $sharp\_idx = \arg\min(\boldsymbol{\sigma})$
9:     $\boldsymbol{\mu}_{\mathrm{sharp}}, \boldsymbol{\sigma}_{\mathrm{sharp}} = \boldsymbol{\mu}[sharp\_idx], \boldsymbol{\sigma}[sharp\_idx]$
10:     threshold $= \boldsymbol{\mu}_{\mathrm{sharp}} + 7.5 \cdot \boldsymbol{\sigma}_{\mathrm{sharp}}$
11: Detect peaks using prominence threshold.
12: Apply post-processing: filter multiple peaks within 0.05s window.
13: Categorize peaks using $L(P)$ (see Equation 4.1).

To evaluate the two prominence thresholding methods, eight cases were randomly selected from the dataset. Each peak, along with its category, was manually annotated using the TRAINSET

annotation tool [23]. The detected peaks and their corresponding labels were compared against the manually annotated ground truth.

Matching was performed using linear cost assignment for bipartite matching, where the cost matrix was defined as:

$$\text{cost}[P_i, P_j] = \begin{cases} |t_i - t_j| + 1, & \text{if } L_i \neq L_j \text{ and } L_i = \text{Na}^+, \\ |t_i - t_j| + 0.5, & \text{if } L_i \neq L_j \text{ and } L_i \neq \text{Na}^+. \end{cases}$$

where:

- i belongs to the indices of ground truth peaks,
- j belongs to the indices of algorithm-detected peaks,
- t_i, t_j are the time indices of peaks P_i and P_j, and
- L_i, L_j are their respective labels.

The Hungarian algorithm [24] was then applied to determine the optimal assignment score of peaks and labels for each Algorithm. The total cost of the optimal assignment over labeled cases was used to evaluate the performance of the two thresholding methods. The results are summarized in Table 4.2.

According to the results, Algorithm 2, which uses the sharpest Gaussian component as the threshold, outperformed Algorithm 1 by detecting peaks at a lower overall cost. A single case comparison for each method is shown in Figure 4.8. Algorithm 1 in contrast captured a lot of noise peaks. As a result, Algorithm 2 was selected for FP peak detection across the entire dataset.

Table 4.2: Linear Cost Assignment Comparison of Field Potential Peak Detection Methods on Ground Truth.

Algorithm/Case	1	2	3	4	5	6	7	8	Total
GMM–μ_{hmean}	94.0	104.0	68.0	60.0	61.43	36.0	44.0	30.0	497.43
GMM–μ_{sharp}	94.0	104.0	68.0	60.0	42.10	36.09	45.25	30.0	479.44
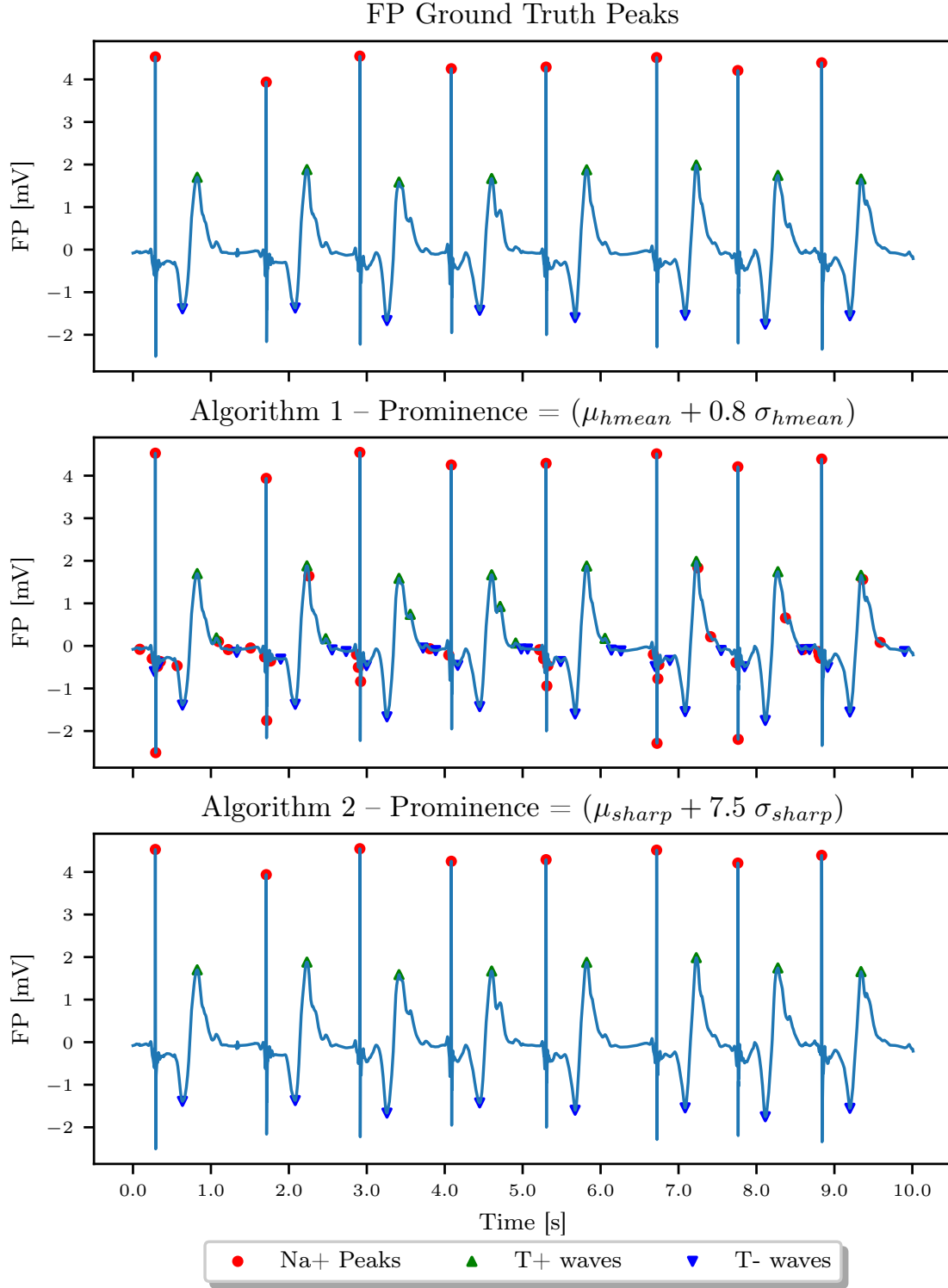
Figure 4.8: **Field Potential Peak Detection Comparison.** The figure presents a side-by-side comparison for a single case: Ground Truth (top), Algorithm 1 (middle), and Algorithm 2 (bottom). Algorithm 2, which uses the sharpest Gaussian component as its threshold, achieves a lower overall cost in peak assignment (see Table 4.2), whereas Algorithm 1 captures numerous noise peaks. This comparison underlines the superior performance of Algorithm 2 for FP peak detection.

4.3   Arrhythmia Detection and Removal

Arrhythmic contractions are irregular heartbeats that disrupt the normal ECC process. These anomalies require individual analysis and, for this study, they were systematically identified and removed using a machine learning-based arrhythmia classification approach developed during a prior study [14]. The models from that project were integrated into this work to eliminate arrhythmic contractions, ensuring that only non-arrhythmic valid contractions were analyzed. A representative arrhythmic contraction is shown in Figure 3.1b, characterized by abnormal signal decay and the presence of multiple peaks.

4.3.1   Arrhythmia Classification Framework

After signal preprocessing and identification of contraction events, the next step was labeling each contraction event. The force signal was selected as the primary reference for arrhythmia labeling due to its higher signal-to-noise ratio (SNR) and consistent peak morphology, making it more reliable for detecting abnormal contractions. A 3.0-second time window around each force peak was extracted. These windows were manually labeled as **arrhythmic** or **non-arrhythmic** based on signal morphology, and the labels were transferred to the corresponding calcium and FP signals in the same time window.

To build the classification model, the 3.0-second windows were further divided into smaller time windows ranging from 0.6 to 3.0 seconds. From each window, statistical features (such as mean, median, mode, percentiles, and 40+ others) were extracted using the TSFEL. These extracted features, along with their respective arrhythmic/non-arrhythmic labels, formed the final dataset for training the arrhythmia classification models (Figure 4.9).

Given the class imbalance (only 8% of contractions were arrhythmic), Random Forest Classifiers with 'balanced' weights were used for binary classification. Separate models were trained for each signal type and window size. The best test-set results were achieved with:

- A 2.4-second window model for force and calcium, achieving an **F1-score of 0.97 and 1.0** for arrhythmic and non-arrhythmic classes respectively.

- A 0.6-second window model for FP, achieving an **F1-score of 0.99 and 1.0** for arrhythmic and non-arrhythmic classes respectively.

For more details on modeling methodology and evaluation, refer to *Section 4.5* of Arrhythmia Classification work [14].

4.3.2   Integration to current work

The trained arrhythmia detection models demonstrated high accuracy and generalizability, allowing for their integration into this thesis project. Before conducting further analysis, **all arrhythmic contractions were identified and removed** using these models, ensuring that only non-arrhythmic contractions were included in the dataset. This preprocessing step helped maintain data integrity and improved the robustness of subsequent analyses.
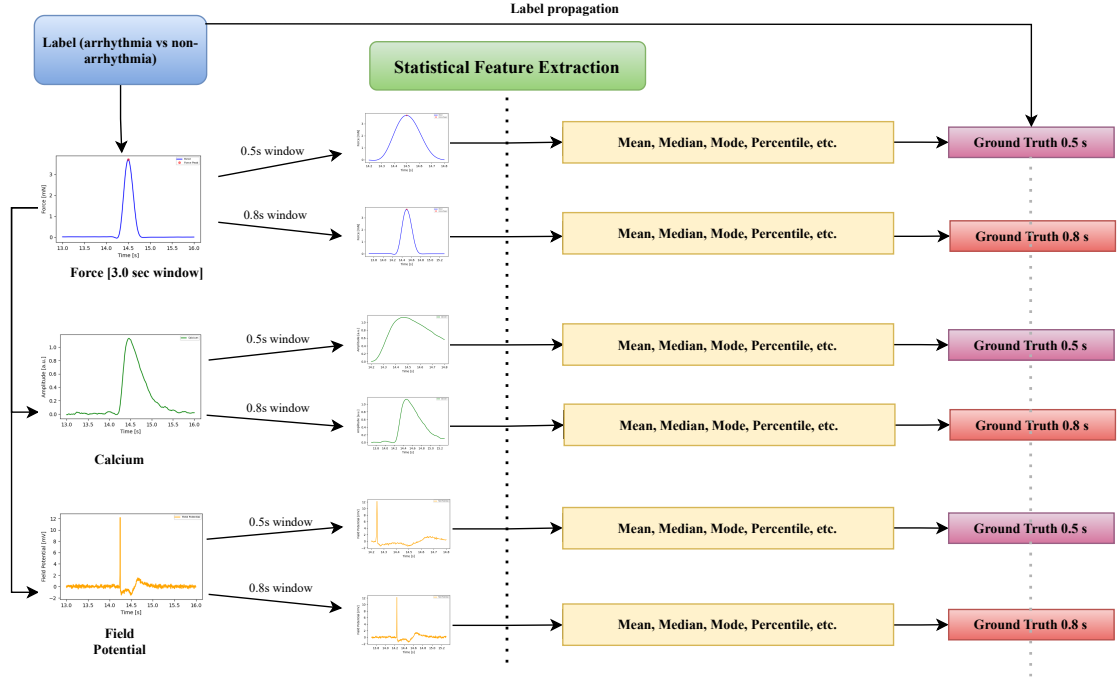
Figure 4.9: **Arrhythmia classification pipeline**

Machine learning workflow: (1) Force-based peak detection, (2) 3s window extraction, (3) Manual labeling, (4) Window sub-division and TSFEL feature extraction (44 features), (5) Label Assignment, (6) Random Forest Modeling (Balanced weights). Final models for Force and Calcium achieved F1=0.97 and F1=1.0 on arrhythmic and non-arrhythmic classes respectively, while for FP, best model achieved F1=0.99 and F1=1.0 for arrhythmic and non-arrhythmic events.

4.4 Feature Extraction

With preprocessed signals and accurately detected non-arrhythmic peaks, the next step involved extracting quantitative biomarkers that characterize tissue contraction strength, relaxation time, and electrical activity patterns. These features provide physiological insights into cardiac function and treatment-induced effects.

4.4.1 Biological Features from Force & Calcium

Features extracted from the force and calcium signals (Figure 4.10) were computed using SciPy's signal processing module, specifically the 'peak_widths' [25] function. This function, given the peak indices, signal values, and a relative height, determines the peak widths at that relative height with respect to peak amplitude. Additionally, it identifies the left and right intersection points of a horizontal line with the signal at a given relative height of the peak. These intersection points were further utilized to extract contraction and relaxation features. The extracted features are as follows:

- **Peak Amplitude**: The maximum signal value at the detected peak, measured in **mN** for force and **arbitrary units (a.u.)** for calcium, respectively.

- **Peak Widths (0.2, 0.5, 0.8) [s]**: The signal width at relative **20%**, **50%**, and **80%** of the peak height, measured in seconds (**s**).

- **Rise Time (RT 0.2-0.8) [s]**: The absolute time taken for the signal to rise from **20% to 80%** of the peak height, representing **contraction time**. Measured in seconds (**s**).

- **Rise Time (RT 0.8-Max) [s]**: The absolute time taken for the signal to rise from **80% to peak amplitude**, measured in seconds (**s**).

- **Decay Time (DT Max-0.8) [s]**: The absolute time taken for the signal to decay from **peak amplitude to 80%** of peak height, measured in seconds (**s**).

- **Decay Time (DT 0.2-0.8) [s]**: The absolute time taken for the signal to decay from **80% to 20%** of peak height, representing **relaxation time**. Measured in seconds (**s**).
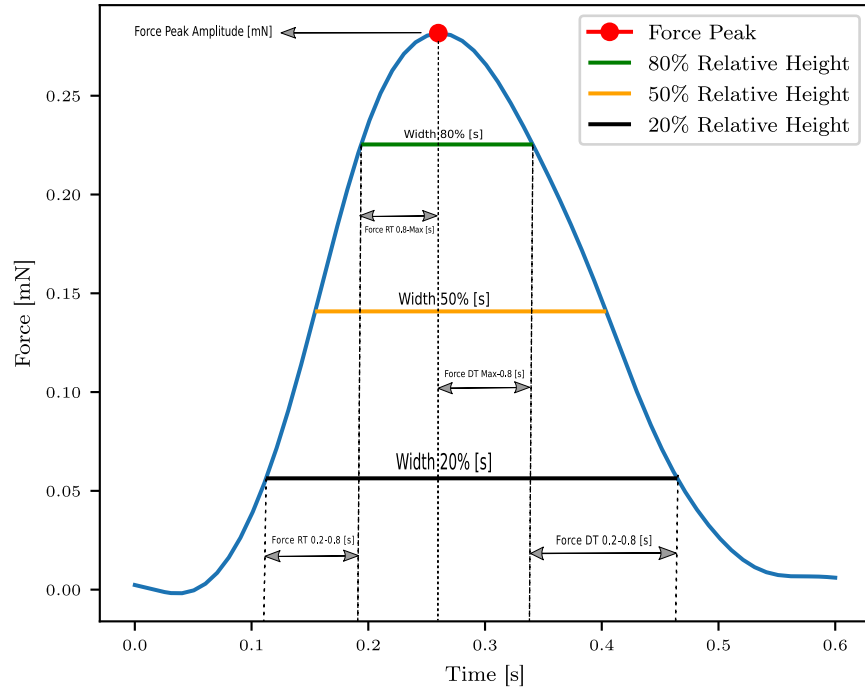
Figure 4.10: **Biologically Relevant Feature Extraction for Force.** This figure illustrates the key quantitative parameters extracted from the force signal, including peak amplitude (red), peak widths at 20%, 50%, and 80% of the peak height (black, orange, green), rise times ($RT_{20\text{-}80}$, $RT_{80\text{-}max}$), and decay times ($DT_{max\text{-}80}$, $DT_{20\text{-}80}$). These parameters characterize contraction strength, timing, and relaxation dynamics. The same extraction approach applies to the calcium signal.

4.4.2 Features from Field Potential

For the FP signal, two key features were extracted:

- **Field Potential Duration (FPD) [s]**: The absolute time difference (**s**) between the **Sodium (Na^+) peak** and the T^+ **wave**, representing the **total duration of the contraction event**.

- **Local Frequency [Hz]**[1]: Estimated using the time intervals between consecutive **Sodium (Na^+) peaks**, calculated as:

$$f_{\text{local}} = \frac{t_p + t_n}{2 \cdot (t_p \cdot t_n)}$$

where t_p and t_n represent the time difference to the **previous** and **next** Sodium peak, respectively (measured in Hertz [Hz]). The formula is basically the reciprocal of *harmonic mean* between two time periods, which provides a more appropriate averaging method for frequency-based measurements, as it better accounts for variations in peak intervals by reducing the influence of disproportionately large values. The **first and last detected peaks** were excluded from this calculation to avoid boundary effects (Figure 4.11).
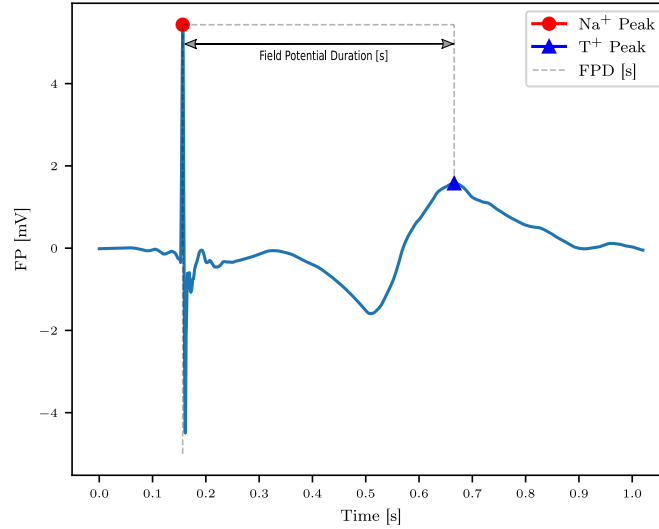
Figure 4.11: **Field Potential Duration (FPD) Measurement.** The field potential duration (FPD) is computed as the time difference between the Na^+ peak and the T^+ wave in the FP signal. This measurement represents the total duration of a contraction event and serves as a key biomarker for assessing electrical activity and contraction timing.

---

[1]The accuracy of this feature depends on the performance of the FP peak detection algorithm. If consecutive FP peaks are missed, the computed local frequency deviates from its intended definition, as it will measure time intervals between non-adjacent detected events. Improving the peak detection algorithm would enhance the reliability of this feature.

Table 4.3: Summary of Extracted Features

Feature	Description	Signal Type	Unit
Peak Amplitude	Maximum peak value	Force, Calcium	mN, a.u.
Peak Width (0.2, 0.5, 0.8)	Width at 20%, 50%, 80% peak height	Force, Calcium	s
Rise Time (RT 0.2-0.8)	Time for 20%-80% peak rise	Force, Calcium	s
Rise Time (RT 0.8-Max)	Time from 80% to peak rise	Force, Calcium	s
Decay Time (DT Max-0.8)	Time from peak to 80% decay	Force, Calcium	s
Decay Time (DT 0.2-0.8)	Time from 80%-20% decay	Force, Calcium	s
Field Potential Duration (FPD)	Time between Na^+ peak and T^+ wave	Field Potential	s
Local Frequency	Instantaneous beat frequency	Field Potential	Hz

The extracted features across all signal types are summarized in Table 4.3. To integrate extracted features across modalities, time alignment was performed based on detected peak indices. Specifically:

- Force and calcium peaks **co-occurring within the FPD of a detected field potential event** were linked to the corresponding field potential event.

- Any force or calcium peaks without an associated FP peak were excluded from the final dataset.

- Likewise, FP peaks without corresponding force and calcium peaks were discarded.

This filtering ensured that the final dataset contained only fully matched contraction events, allowing for a coherent multimodal analysis of contraction dynamics. The resulting dataset was subsequently used for analyzing treatment-induced changes in tissue response quantified through the above extracted features.

Chapter 5

Data Mining and Analysis

This chapter builds upon the rigorous preprocessing and feature extraction methodologies detailed in the preceding chapter, setting the stage for comprehensive data mining and analytical exploration of ECC responses. Having established robust datasets through careful preprocessing and feature extraction, this chapter leverages advanced analytical methods, including statistical hypothesis testing, dimensionality reduction (t-SNE), and regression-based modeling, to uncover meaningful biological insights from multimodal data. Specifically, the analysis systematically evaluates the effects of pharmacological stimulus—E-4031, Nifedipine, and Calcium titration experiments—on cardiac tissue functionality, elucidates dose-response relationships, assesses feature correlation dynamics, and identifies intricate treatment-induced feature associations. Collectively, these analyses aim to enhance understanding of treatment-induced modulation in ECC, reinforcing the utility of MTs as predictive human cardiac models.

5.1 Global Overview

Following the extraction of biologically relevant features, the analysis began with establishing a global perspective of the dataset. A total of 18 features were extracted (Section 4.4), and t-distributed Stochastic Neighbor Embedding (t-SNE) [26] was applied to project high-dimensional feature representations of contraction events onto a 2D space, preserving local structure while reducing dimensionality. This aids in visualizing the data space at both treatment and concentration levels.

5.1.1 Data Projection to 2D Space

As described in Section 3.3, the dataset consists of three different pharmacological treatments, each applied at varying concentrations[1]. The treatments and their respective concentrations are summarized in Table 5.1.

---

[1]These pharmacological interventions exhibit sufficiently rapid onset within hiPSC-derived MTs, thereby eliminating the need for a formal washout period between concentrations and minimizing the risk of carry-over effects.

Table 5.1: Treatments and corresponding concentration levels. * For Calcium Titration, the highest calcium concentration (4.0 µmol) was considered the baseline due to its distinct experimental nature. For **Nifedipine**, a one-hour post-treatment condition ($1h_{pt}$) was included as a washout experiment to evaluate tissue recovery after drug removal. Additional **E-4031** concentrations (0.1 µmol to 1.0 µmol) were tested but discarded as all contractions were arrhythmic at these levels.

Treatment	Concentration Levels (µmol)
E-4031	0.003, 0.01, 0.03, 0.1, 0.3, 0.6, 1.0
Nifedipine	0.01, 0.03, 0.1, 0.2, 0.3, 0.5, 1.0, 4.0, 10.0, $1h_{pt}$
Calcium Titration (Ca^{2+})	0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 4.0*

For **Nifedipine**, an additional condition referred to as **one-hour post-treatment ($1h_{pt}$)** was included in the dataset. This condition represents a **washout experiment**, where after the removal of the drug, measurements were taken one hour later. The purpose of this measurement is to assess whether the tissue recovers to its baseline state after drug exposure or if any residual effects persist. In the case of **E-4031**, an extended concentration range (0.1 µmol to 1.0 µmol) was tested to investigate its effects at higher doses. However, at these concentrations, all contractions exhibited **arrhythmic behavior**, leading to their exclusion from further analysis. The retained concentration levels (0.003 µmol to 0.03 µmol) were selected as they provided physiologically relevant responses without inducing complete arrhythmic activity. Unlike E-4031 and Nifedipine, which were investigated for their dose-dependent effects, Calcium Titration primarily served to assess the influence of varying extracellular Ca^{2+} concentrations on contraction dynamics. The highest calcium **4.0 µmol** concentration was considered the baseline due to its distinct experimental nature compared to the other two drugs.

Each drug was tested on two different tissues. To visualize the global feature space, features from all treatment concentrations were pooled before computing Scipy's [27] t-SNE using default parameters (**perplexity = 30**, **iterations = 1000**). The resulting low-dimensional representation was averaged over contraction events per tissue and concentration, such that each point in Figure 5.1 represents a unique combination of tissue and concentration. This visualization reveals structural relationships between contraction events across different experimental conditions.

5.1.2   Key Findings

A notable observation in Figure 5.1 is the close proximity of contraction events for tissue **'1.20b_B01'** under **E-4031 at 0.003 µmol** and **Ca^{2+} at 2.0 µmol** (from Calcium Titration) indicating a high degree of **feature similarity**. Further analysis of their feature distributions (Figure 5.2) confirms that their values are closely aligned. This may be attributed to the fact that the same tissue was used in both experiments, leading to baseline-like feature distributions at these specific concentrations. Since 0.003 µmol for E-4031 and 2.0 µmol for Ca^{2+} are close to their respective baseline conditions, their contraction dynamics may exhibit similar characteristics. These findings highlight potential overlapping physiological effects, which could provide insight into alternative drug actions or shared mechanisms of action across treatments.

Another key observation from this representation is that **different tissues exhibit distinct mean responses to drug exposure**, as shown by their spatial distribution in t-SNE space. This suggests potential tissue-specific variations in drug response, which may be relevant for identifying

phenotypic differences across tissues and refining treatment strategies.
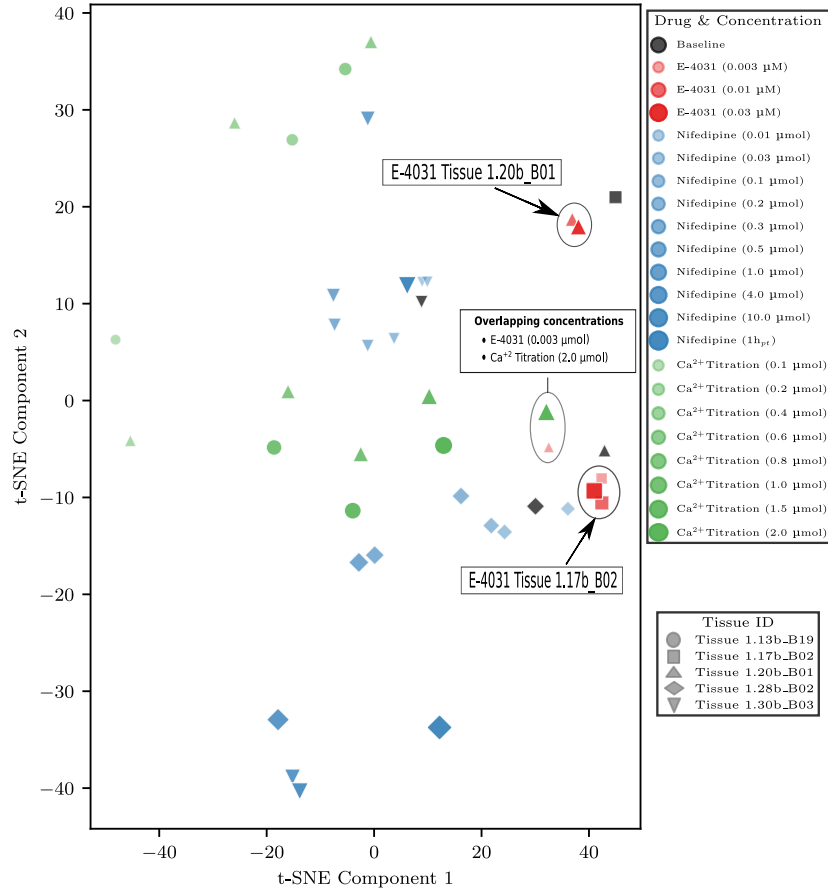
Figure 5.1: **Global t-SNE Projection of All Treatments and Concentrations (Perplexity = 30)**: This t-SNE visualization provides an overview of the feature space across all treatments and concentrations. Each point represents the averaged low-dimensional embedding for a unique tissue and treatment concentration combination. Proximity between points indicates similarity in feature distributions, highlighting potential physiological relationships. Notably, the overlap between contraction events under E-4031 (0.003 µmol) and Ca^{2+} from Calcium Titration (2.0 µmol) for tissue '1.20b_B01' suggests a high degree of feature similarity, possibly due to shared baseline-like characteristics. This finding may indicate similar mechanisms of action or alternative pharmacological effects, warranting further investigation. Additionally, distinct spatial separation between tissues highlights variations in their mean drug response (for instance, E-4031 '1.20b_B01' and '1.17b_B02' tissues, shown in red triangle and squares, respectively), suggesting potential tissue-specific effects that could be relevant for phenotypic characterization and treatment optimization.
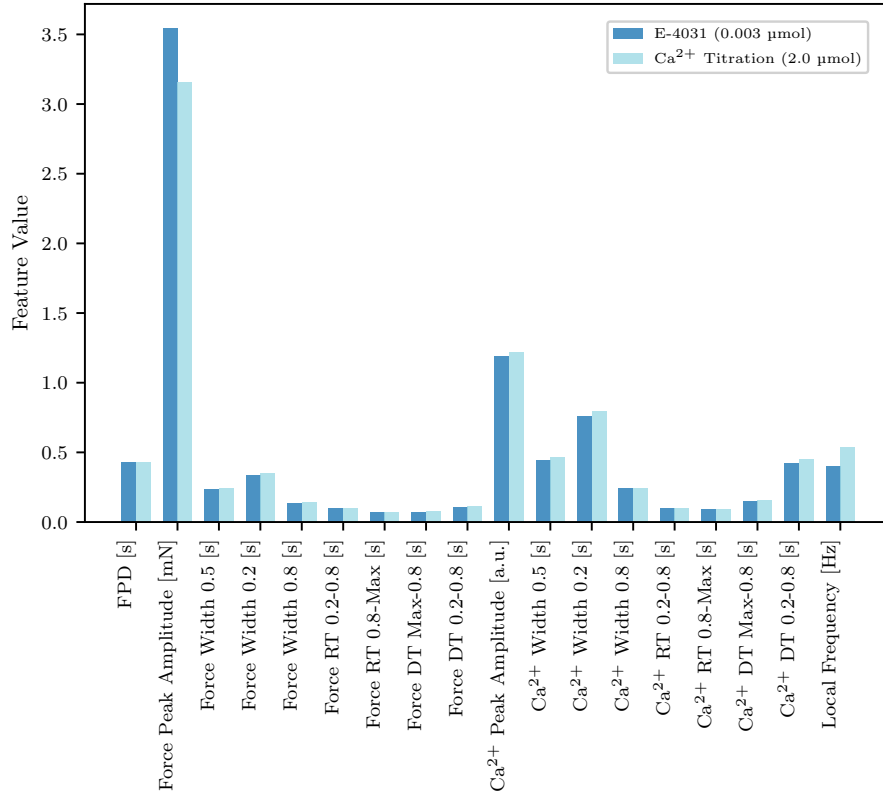
Figure 5.2: **Feature Distribution for Overlapping E-4031 (0.003 µmol) and Calcium Titration (2.0 µmol) Concentrations**: This plot displays the averaged feature values of contraction events recorded from tissue '1.20b_B01' under two distinct treatment conditions–E-4031 (0.003 µmol) and Calcium (Ca^{2+}) Titration (2.0 µmol). Averaged representation from these conditions exhibited significant overlap in the t-SNE representation (Figure 5.1) and are also comparable in this plot. The similarity in feature values suggests that these conditions induce comparable physiological effects, potentially due to their proximity to baseline conditions. Such overlaps in feature space can help identify alternative treatment strategies or reveal shared pharmacodynamic mechanisms between drug treatments.

5.1.3   Section Summary

In summary, the t-SNE visualization provided a global perspective on treatment-induced contraction dynamics, revealing potential relationships between treatments and tissues. Notably, treatment conditions E-4031 (0.003 µmol) and Ca^{2+} (2.0 µmol) from Calcium Titration experiment for tissue '1.20b_B01' exhibited feature similarity, suggesting potential shared physiological effects. Additionally, distinct clustering across tissues highlighted potential tissue-specific variations in drug response, relevant for phenotypic characterization.

5.2 Statistical Analysis of Treatment Effects

Having established a global overview of the data space, the next crucial step was to determine the statistically significant effects of pharmacological treatments. This analysis was conducted at both the overall treatment level and the concentration level using mixed modeling and post-hoc tests, respectively. Specific hypotheses regarding known treatment effects, previously presented in the literature (see Literature Review, Chapter 2; [9, 10]), were reformulated and systematically evaluated.

5.2.1 E-4031

5.2.1.1 Methodology

E-4031, known to induce arrhythmias at higher concentrations, was tested. Measurements from higher concentrations were discarded in the arrhythmia classification study (Section 4.3), leaving only data from three concentrations and baseline. A summary of available data is shown in Table 5.2. Since measurements were available for both tissues at each concentration, a crossed design[2] with repeated measures[3] was employed. To account for tissue variability, a mixed-effects model was used, treating tissue as a random effect while keeping concentration as a fixed effect[4]. This approach improves control of the Type I error rate and increases statistical power, leading to more accurate inference. The model decomposes error variation into two components: one due to tissue variability (σ_α) and the other due to measurement residuals (σ_ϵ).

The model was fitted using the `lmerTest` [30] package in R and is expressed as:

$$Y = \beta_0 + \beta_1 \cdot [\text{Concentration}] + Z_\alpha \cdot [\text{Tissue ID}] + \epsilon,$$

where:

- Y is the target feature.

- β_0 is the intercept, capturing the overall mean response.

- β_1 is the fixed effect of concentration and is the primary effect of interest in this analysis, capturing the direct influence of concentration on the feature response (i.e. Y). [Concentration] is the corresponding factor variable[5] with four levels (baseline and 3 concentrations).

- Z_α represents the random effect (a.k.a random intercept) of tissue, modeled as a normally distributed variable: i.e $Z_\alpha \sim \mathcal{N}(0, \sigma_\alpha^2)$. The variable [Tissue ID] is the 2-level factor distinguishing different tissues. Since Z_α is a random effect, the model estimates its variance σ_α^2 rather than individual Z_α values.

- ϵ is the error term, with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

---

[2]A **crossed design** in statistical modeling refers to an experimental setup where each level of one factor appears in combination with all levels of another factor [28].

[3]As each tissue is repeatedly exposed to multiple concentrations, its a repeated measure design.

[4]A **fixed effect** represents a factor whose levels are of direct interest and are assumed to be consistent across different samples, while a **random effect** is a factor whose levels are drawn from a larger population and introduce variability that is accounted for in the model. [29]

[5]A **factor variable** is a categorical variable with a fixed set of levels, used in statistical models to represent group differences where one level is typically chosen as the **reference category (baseline)** for comparison.

For post-hoc analyses, the Dunnett's test [31] was used to compare concentration levels with the baseline. This test adjusts the significance level for multiple comparisons and was implemented using the `multcomp` [32] package in R. The significance level was set to 0.05.

The results for a subset of features are shown in Figure 5.3. For each feature, the global p-value[6] (with Welch's correction) is displayed at the top of the plot, along with the decomposition of error variance into σ_α and σ_ϵ. This global p-value indicates whether the treatment has a statistically significant effect on feature response compared to the baseline, with a significant value suggesting that the treatment induces measurable changes beyond natural variability. The fit's singularity (or non-singularity) is also indicated. A singular fit occurs when the estimated value of σ_α is zero. Welch's correction [33] was applied to adjust for heteroscedasticity[7] (unequal variances between tissues), which was observed in residual diagnostics (Figure 5.5). This ensures more reliable significance testing despite violating the homoscedasticity (equal variance) assumption. For example, the $Ca^{2+} Peak\ Amplitude$ feature exhibits heteroscedasticity, with residuals increasing as fitted values increase, following a funnel-shaped pattern.

The p-values from the Dunnett's test are shown as horizontal significance lines above each concentration on the x-axis, where "ns" indicates non-significance and asterisks (*) indicate significance compared to the baseline. The significance levels are: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. A significant p-value for a concentration indicates a statistically significant deviation in feature response at that concentration of treatment compared to the baseline. If the global p-value is not significant, no post-hoc concentration-level tests are performed (as for $Ca^{2+} Peak\ Amplitude$).

The Quantile-Quantile (QQ) plots of residuals for each feature (Figure 5.4) were used to assess the residuals normality assumption. While some residuals deviate slightly from normality, linear mixed models are known to be robust to such deviations [34], and the residuals are not excessively non-normal. Results for additional features are provided in the supplementary figures (Figures A.1–A.3).

Table 5.2: Contraction event counts across E-4031 concentrations (Crossed Design).

Tissue ID	E-4031 Concentration (μmol)				Total
	0.0	0.003	0.01	0.03	
1.17b_B02	19	11	11	9	50
1.20b_B01	12	26	35	32	105
Total	31	37	46	41	**155**

---

[6]Statistical significance of β_1.

[7]**Heteroscedasticity** refers to the condition where the variance of the errors in a regression model is not constant across observations, potentially violating model assumptions and affecting inference.
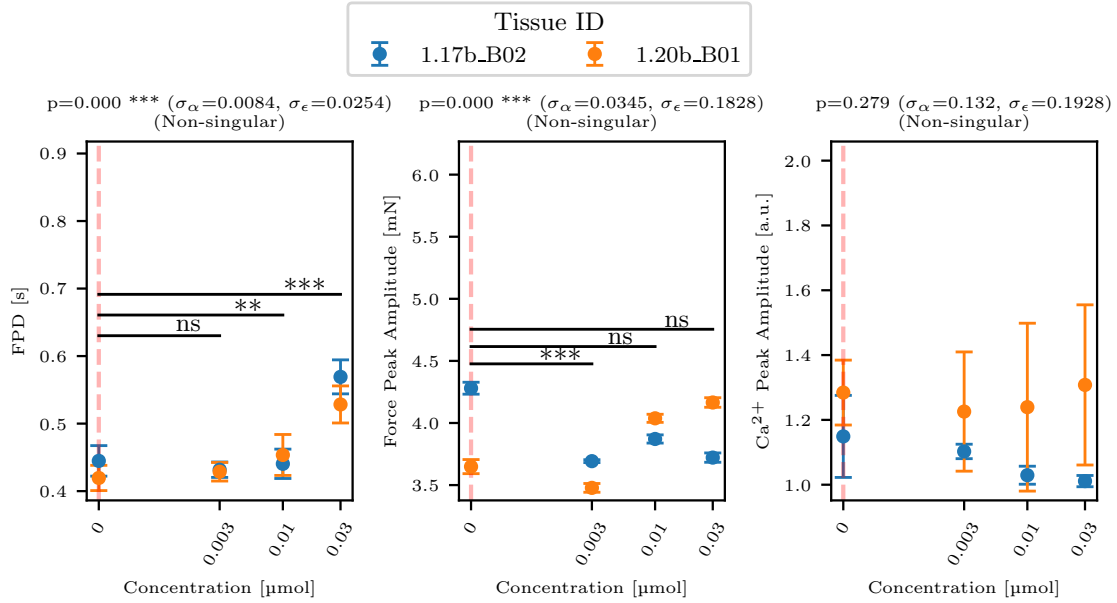
Figure 5.3: **Statistical significance of E-4031 treatment on subset of Biological features**:
This figure presents the results of a mixed-effects model assessing the impact of E-4031 on a subset
of biological features, controlling for tissue-specific variability through inclusion of a random-effect.
Global p-values (with Welch's correction for heteroscedasticity) are shown atop each subplot, while
horizontal significance lines (from post-hoc Dunnett's tests) compare each treatment concentration
with baseline (orange, dotted). Significance is indicated as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$,
and "ns" denotes non-significance. Additionally, each subplot displays the estimated decomposition
of error variance into tissue variability (σ_α) and residual noise (σ_ϵ). The hypothesis that increasing
E-4031 concentration leads to a significant increase in field potential duration (FPD) is strongly
supported by these results (*FPD* ** $p < 0.01$ and *** $p < 0.001$ for 0.01 µmol and 0.03 µmol
respectively.)
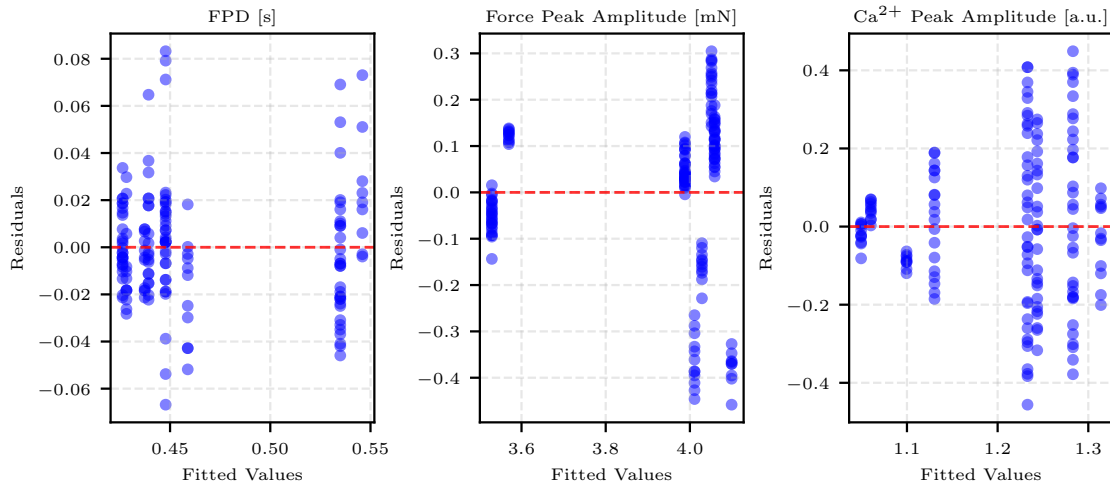
Figure 5.5: **Residuals vs. Fitted Values for Mixed Model Fit of E-4031 Treatment**: This
diagnostic plot shows the relationship between residuals and fitted values. A funnel-shaped pattern
in Ca^{2+} *Peak Amplitude* indicates heteroscedasticity, which was addressed via Welch's correction
during significance testing.
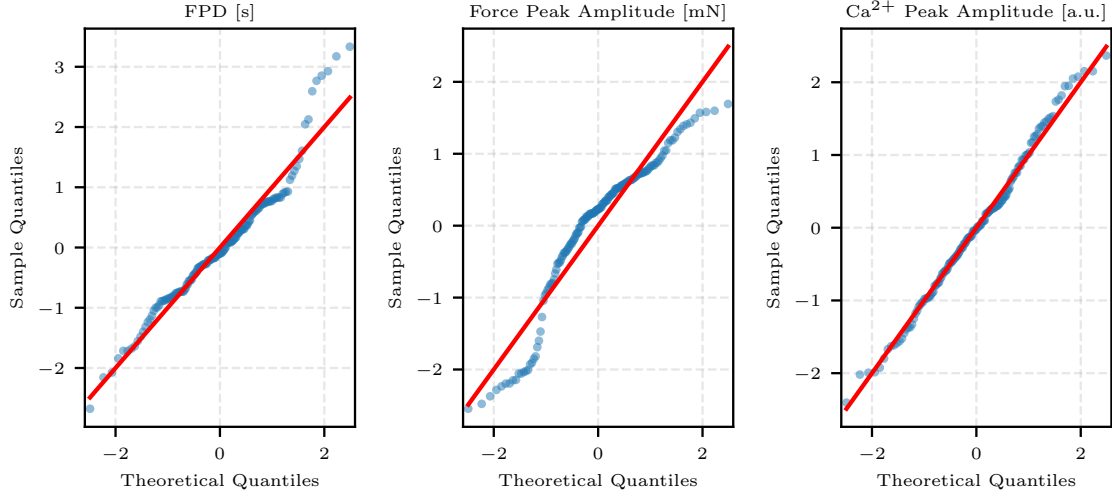
Figure 5.4: **QQ Plot of Residuals for Mixed Model Fit of E-4031 Treatment.**: This QQ plot evaluates the normality of residuals from the mixed-effects model. Although some deviation from the normal line is apparent, linear mixed models are robust to such deviations [34].

5.2.1.2 Hypothesis Evaluation

The hypothesis for E-4031 is that it increases field potential duration (FPD) at lower concentrations, ultimately inducing arrhythmia's at higher concentrations. Thus, the hypothesis statement is:

The field potential duration (FPD) should increase significantly with increasing E-4031 concentration.

As shown in Figure 5.3, the overall drug effect on field potential duration (FPD) is highly significant ($p < 0.001$). Post-hoc analysis reveals that concentrations of 0.01 µmol and 0.03 µmol significantly increase FPD compared to baseline ($p < 0.01$ and $p < 0.001$, respectively). This confirms the hypothesis, demonstrating that E-4031 significantly impacts FPD as E-4031 concentration increases.

5.2.2 Nifedipine

5.2.2.1 Methodology

For Nifedipine, measurements were not available for all concentrations in both tissues (Table 5.3), resulting in a nested design[8]. To account for this, a nested random effect[9] for concentration within tissue was added to the model:

$$Y = \beta_0 + \beta_1 \cdot [\text{Concentration}] + Z_\alpha \cdot [\text{Tissue ID}] + Z_\beta \cdot [\text{Tissue ID:Concentration}] + \epsilon,$$

where:

---

[8]A **nested design** is an experimental setup where the levels of one factor are unique to each level of another factor, meaning they do not repeat across different levels of the higher-order factor [28].

[9]The model design was motivated from this StackExchange thread.

- Z_β is the random effect of concentration nested within tissue, with $Z_\beta \sim \mathcal{N}(0, \sigma_\beta^2)$. Since nested effects are modeled as random, the model estimates only the variance of the nested effect σ_β^2, rather than individual values Z_β.

- [Tissue ID:Concentration] is a factor variable that uniquely identifies each combination of tissue and its corresponding concentrations. Based on the Nifedipine data structure (Table 5.3), the number of levels for this nested effect is 20, considering that 9 concentrations are applied to both tissues, while 2 concentrations (1 µmol and 10 µmol) are exclusive to tissue 1.30b_B03.

The methodology otherwise remains consistent with the E-4031 analysis. Significance results for a subset of features are shown in Figure 5.6. The results for remaining features along with their diagnostic plots are provided in the supplementary figures (Figures A.4–A.6).

Table 5.3: Contraction event counts across Nifedipine concentrations (Nested Design).

Tissue ID	Nifedipine Concentration (µmol)											Total
	0.0	0.01	0.03	0.1	0.2	0.3	0.5	1.0	4.0	10.0	$\mathbf{1h}_{pt}$	
1.28b_B02	24	18	15	15	19	16	19	0	32	0	29	187
1.30b_B03	16	16	16	13	11	10	15	19	24	9	1	150
Total	40	34	31	28	30	26	34	19	56	9	30	**337**
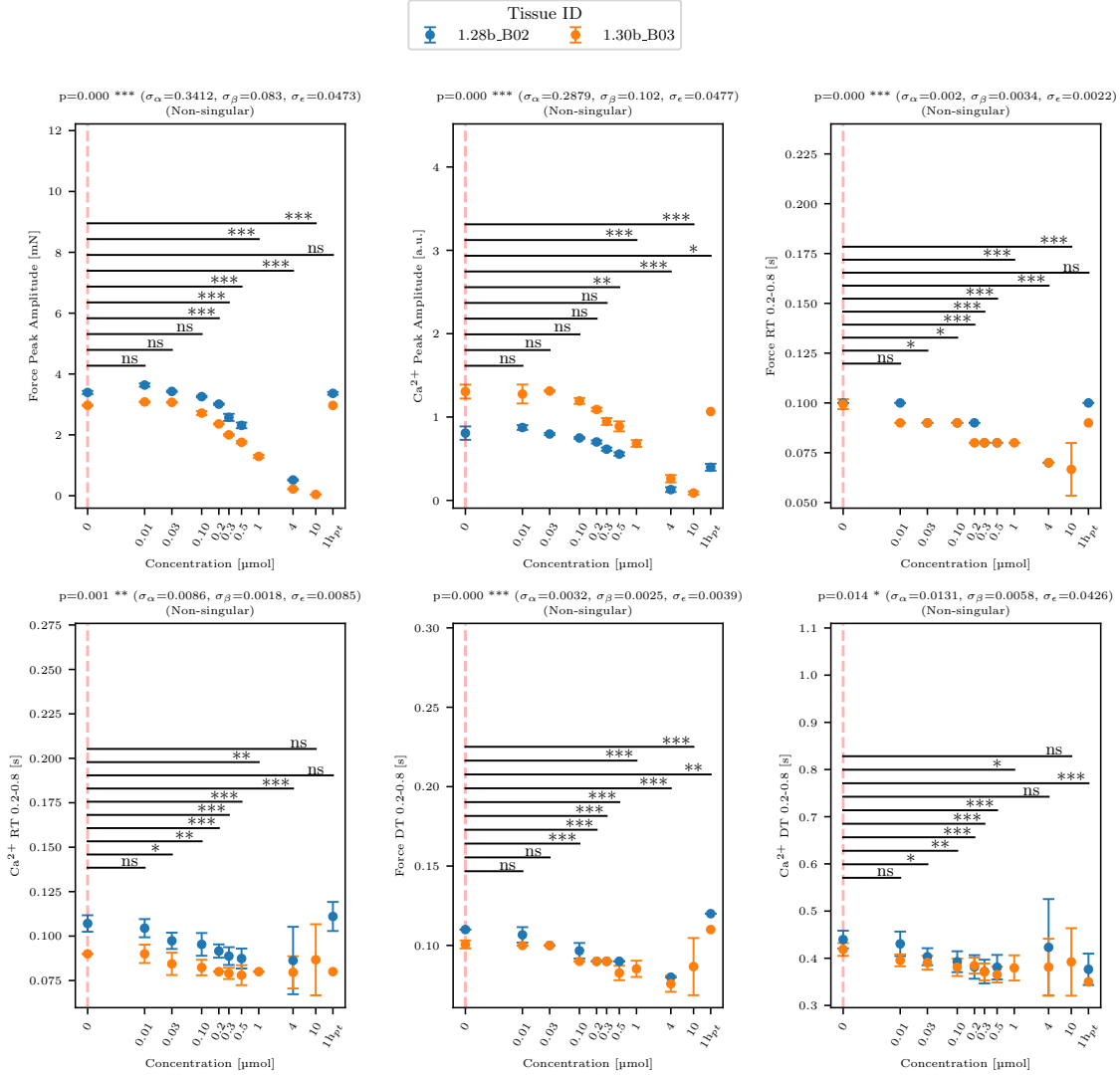
Figure 5.6: **Statistical significance of Nifedipine treatment on subset of Biological features.**: This figure presents the results of an extended mixed-effects model, incorporating a nested design to account for concentration effects within tissues. Global p-values (with Welch's correction) are displayed on-top of each feature, with post-hoc Dunnett's test comparisons (denoted by * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and "ns" for non-significance) marking significant differences from baseline. Both force and calcium peak amplitudes, as well as contraction and relaxation times (e.g., *Force RT 0.2–0.8, Ca^{2+} RT 0.2–0.8, Force DT 0.2–0.8*, and *Ca^{2+} DT 0.2–0.8*), are shown in the plot. The estimated variance for the nested concentration effect (σ_β) is also provided on top. These results confirm that increasing Nifedipine concentrations significantly reduce peak amplitudes and shorten contraction and relaxation durations, supporting the hypothesis that Nifedipine weakens contractions by disrupting calcium-dependent ECC (*** $p < 0.001$ for *Force Peak Amplitude* & *Calcium Peak Amplitude* at 0.2–10 µmol and 1–10 µmol respectively. Contraction and Relaxation features also show significant decline at increasing concentrations of Nifedipine.)

5.2.2.2 Hypothesis Evaluation

Nifedipine, a calcium channel blocker, weakens contractions by disrupting the calcium-induced calcium release (CICR) cycle. The hypothesis is:

> The amplitude of calcium and force, as well as tissue contraction and relaxation times, should decrease with increasing Nifedipine concentration.

As shown in Figure 5.6, the force peak and calcium peak amplitudes decrease significantly with increasing Nifedipine concentration ($p < 0.001$). Post-hoc analysis indicates that concentrations of 0.2–10 µmol and 1–10 µmol significantly reduce force and calcium peak amplitudes compared to baseline. Similarly, features representing tissue contraction and relaxation times (e.g., *Force RT 0.2–0.8, Ca^{+2} RT 0.2–0.8, Force DT 0.2-0.8* and *Ca^{+2} DT 0.2-0.8)* also show significant decreases. These results confirm the hypothesis, demonstrating that Nifedipine weakens contractions and reduces tissue contraction and relaxation times.

5.2.3 Calcium Titration Experiment

5.2.3.1 Methodology

The methodology for the calcium titration experiment follows the same design as the Nifedipine experiment, using a nested design approach (Table 5.4). In this context, *Force Peak Amplitude* is the only relevant feature for hypothesis evaluation due to its sensitivity to changes in calcium concentrations, making it a key indicator of the physiological response in this experiment. The result for *Force Peak Amplitude* is shown in Figure 5.7.

5.2.3.2 Hypothesis Evaluation

The hypothesis for the calcium titration experiment is straightforward:

> Increasing calcium concentration enhances contractile force.

As shown in Figure 5.7, the force peak amplitude increases significantly with calcium concentration ($p < 0.001$), confirming the hypothesis.

Table 5.4: Contraction event counts across Calcium Titration experiment concentrations (Nested Design).

Tissue ID	Calcium Concentration (µmol)									Total
	4.0	0.1	0.2	0.4	0.6	0.8	1.0	1.5	2.0	
1.13b_B19	0	40	0	35	9	0	34	31	25	174
1.20b_B01	22	0	98	47	37	39	36	33	32	344
Total	22	40	98	82	46	39	70	64	57	**518**
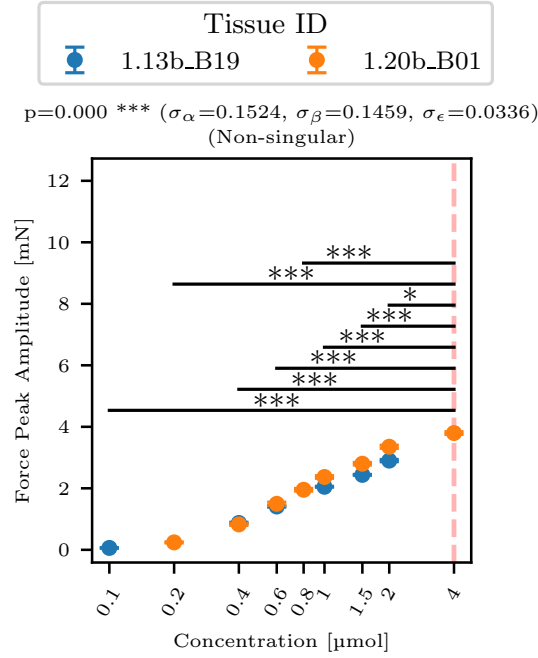
Figure 5.7: **Statistical significance of Calcium Titration on Force Peak Amplitude**: This figure displays the results of a mixed-effects model (nested design, see 5.2.2.1) applied to the Calcium Titration experiment. Only *Force Peak Amplitude* is shown here since it is the primary feature for hypothesis evaluation, testing the hypothesis that increasing calcium concentration during Calcium Titration enhances contractile force. The results reveal that *Force Peak Amplitude* increases significantly with calcium concentration (*** $p < 0.001$), confirming the expected effect.

5.2.4   Section Summary

The statistical analysis confirms that the observed treatment effects align with expected physiological mechanisms, supporting the validity of the experimental approach. The results suggest that MTs exhibit pharmacological responses consistent with known treatment actions, making them a promising in vitro model candidates for studying cardiac treatment effects. However, further research is required to fully establish their translational relevance to human heart physiology, including comparing responses to human cardiac tissues and validating findings in additional experimental conditions.

5.3 Drug Response Modeling and Feature Analysis

After assessing treatment effects, the next step was to identify biologically informative drug concentrations and analyze feature variations across treatments. This was achieved through the following steps:

- **Hill equation fitting** to model drug response curves and estimate the half-maximal effective concentration (EC_{50}) or inhibitory concentration (IC_{50}).

- **Relative comparison of treatment effects** by assessing changes in biological features at the EC_{50}/IC_{50} concentration in relation to baseline values.

- **Principal Component Influence Analysis (PCIA)** to systematically identify features that exhibit similar variation across different treatment concentrations, enabling a more data-driven understanding of treatment-induced effects.

The following sections provide a detailed explanation of each step.

5.3.1 Hill Fitting

To model the dose-response relationship of the treatments, the Hill equation was fitted to the extracted features. The Hill equation is defined as:

$$\frac{E}{E_{\max}} = \frac{[A]^n}{EC_{50} + [A]^n} \tag{5.1}$$

where:

- E represents the measured feature response.

- $E_{\max}$ is the maximum observed feature response.

- $[A]$ is the treatment concentration.

- EC_{50} is the treatment concentration at which the response reaches half of $E_{\max}$.

- n is the Hill coefficient, describing the steepness of the dose-response curve.

For this analysis, four key features were considered: *Field Potential Duration (FPD)*, *Force Peak Amplitude*, *Calcium Peak Amplitude*, and *Local Frequency*. The Hill equation was fitted to the data for Calcium Titration and Nifedipine, while E-4031 was excluded due to an insufficient number of concentration levels.

Since the primary focus of this analysis was on overall treatment effects rather than inter-tissue variability, response values were averaged across tissues to ensure a generalized pharmacological interpretation.

Hill fitting was performed using the `Hillfit` package [35] in Python (Figure 5.8), and the estimated EC_{50}/IC_{50} values of above selected features were averaged and used for further analysis. The closest available experimental concentration available to this value was selected:

- $IC_{50} \approx 1.22$ µmol for Calcium Titration, mapped to the available concentration of 1.0 µmol.

- $EC_{50} \approx 0.91$ µmol for Nifedipine, mapped to the available concentration of 1.0 µmol.

These values serve as a reference point for assessing the relative impact of pharmacological treatment in the next section.
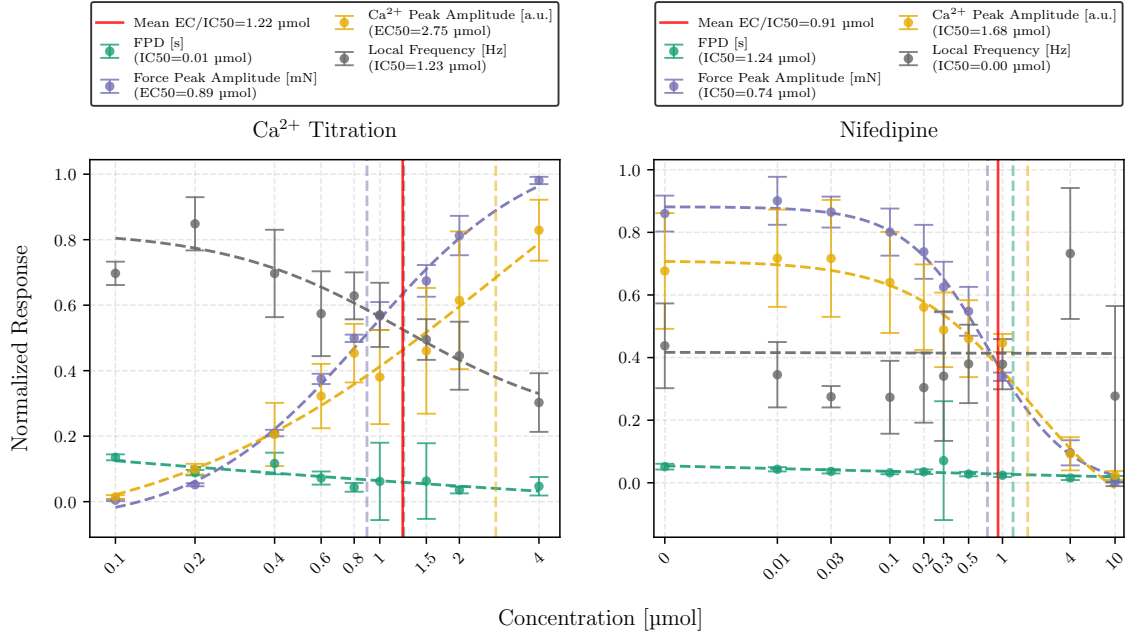
Figure 5.8: **Hill Equation Fitting for Calcium Titration Experiment and Nifedipine**: This figure presents the normalized dose-response curves for Calcium Titration (left) and Nifedipine (right), fitted using the Hill equation. Four key biological features were analyzed: *Field Potential Duration (FPD) [s], Force Peak Amplitude [mN], Ca^{+2} Peak Amplitude [a.u], Local Frequency [Hz]*. The Hill equation was used to estimate the half-maximal effective concentration (EC_{50}) and the inhibitory concentration (IC_{50}) for Calcium Titration experiment and Nifedipine, respectively. The fitted curves help quantify the non-linear drug response and allow comparison between experimental data and expected pharmacological behavior. The experimental concentration closest to the mean of estimated EC_{50}/IC_{50} values (1.0 µmol) was selected as a reference for subsequent analyses, ensuring that treatment-induced effects are evaluated at a biologically relevant dosage.

5.3.2 Relative Comparison of Treatment Effect

To evaluate how pharmacological treatments modify biological features, responses at EC_{50}/IC_{50} were compared to baseline measurements. The baseline was defined by aggregating feature values across all conditions before treatment.

For each feature, the mean baseline value was set to 100%, and the percent change at the selected EC_{50}/IC_{50} concentration was computed as:

$$\text{Relative Change (\%)} = \frac{\text{Feature Value at } EC_{50} - \text{Feature Value at Baseline}}{\text{Feature Value at Baseline}} \times 100 \qquad (5.2)$$

By normalizing feature changes relative to baseline, this approach enables a direct and standardized comparison of drug effects across treatments, improving interpretability. The results are

presented in Figure 5.9. For E-4031, where the mean EC_{50} could not be determined, the largest available concentration (0.003 µmol) was used as a substitute.
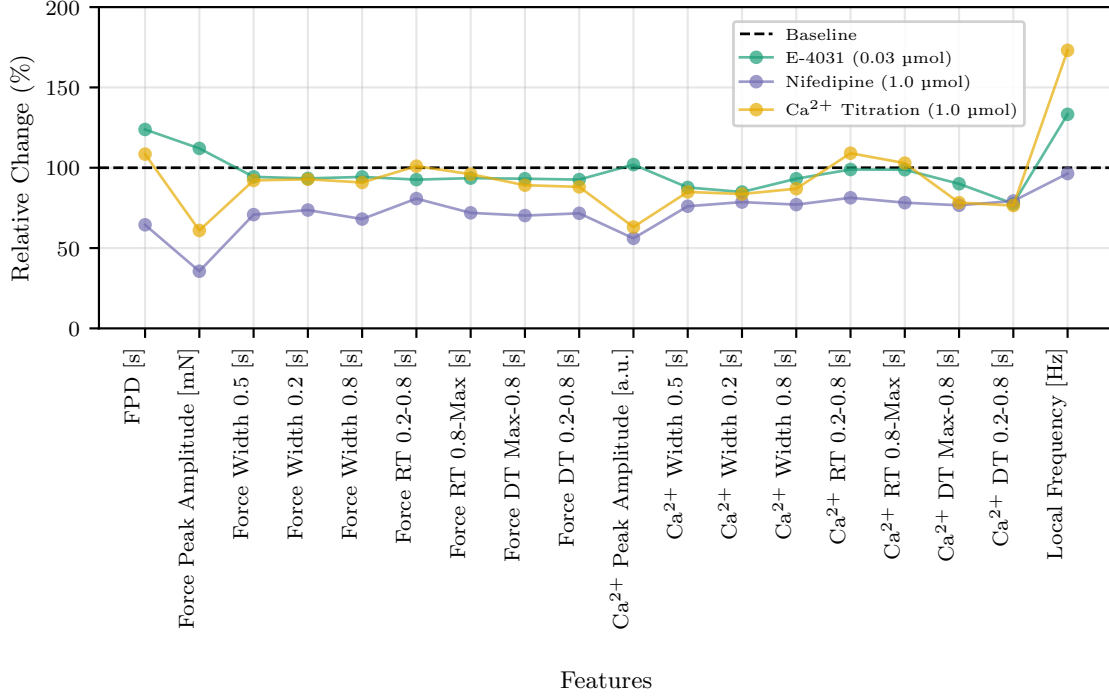
Figure 5.9: **Mean Relative Change of Features Across Treatments at EC_{50}/IC_{50}**: This figure compares the percentage change in biological features relative to baseline at treatment concentrations approximating EC_{50}/IC_{50} for Nifedipine and Calcium Titration (1.0 µmol). The baseline is defined as the mean feature value before treatment, set at 100%. The observed changes reflect the extent of treatment-induced modifications in cellular function. For E-4031, where EC_{50}/IC_{50} could not be determined, the highest available concentration (0.003 µmol) was used as a reference. These relative comparisons facilitate cross-drug analysis by normalizing responses, enabling a clearer interpretation of each treatment's impact on tissue function.

5.3.3 Principal Components Influence Analysis (PCIA)

Beyond individual feature analysis, a systematic approach was developed to identify groups of features that respond similarly across different treatment concentrations. This helps in recognizing shared biological mechanisms and uncovering functionally related feature groups.

The approach is based on **Principal Component Analysis (PCA)** to reduce data dimensionality while preserving variance and **DBSCAN clustering** of principal component (PC) loadings to group features with similar contribution patterns. More briefly, PCA was applied to each treatment data, and principal components explaining 95% of the variance were retained (Figure 5.10a). Feature contributions (loadings) to these components were then extracted and clustered

using DBSCAN. To determine the optimal ε parameter for DBSCAN, the k-nearest neighbors (k-NN) heuristic was employed with $k = 2$, ensuring that even a cluster with only two features having similar concentration-dependent variations could be identified. The knee point in the decreasingly sorted k-NN distance curve (d_i) was identified and used as ε value (Figure 5.10b). The resulting clustered feature groups captured functionally related responses, while features categorized as noise points were identified as independently varying.

To determine the optimal ε for DBSCAN, the Kneedle algorithm [36] was applied using the `Kneed` [37] package in Python. Since the algorithm is designed for concavely increasing functions, it first inverts the k-nearest neighbors (k-NN) distance curve (d_i^*) to adapt it for knee point detection. Before identifying the knee point, the algorithm normalizes both the sample indices (x-axis) and the corresponding k-distances (y-axis), ensuring a scale-independent detection of the knee region. Once the knee point is detected, the algorithm returns the corresponding unnormalized k-distance value, which is then used as the final ε parameter for DBSCAN.

(a)

(b)
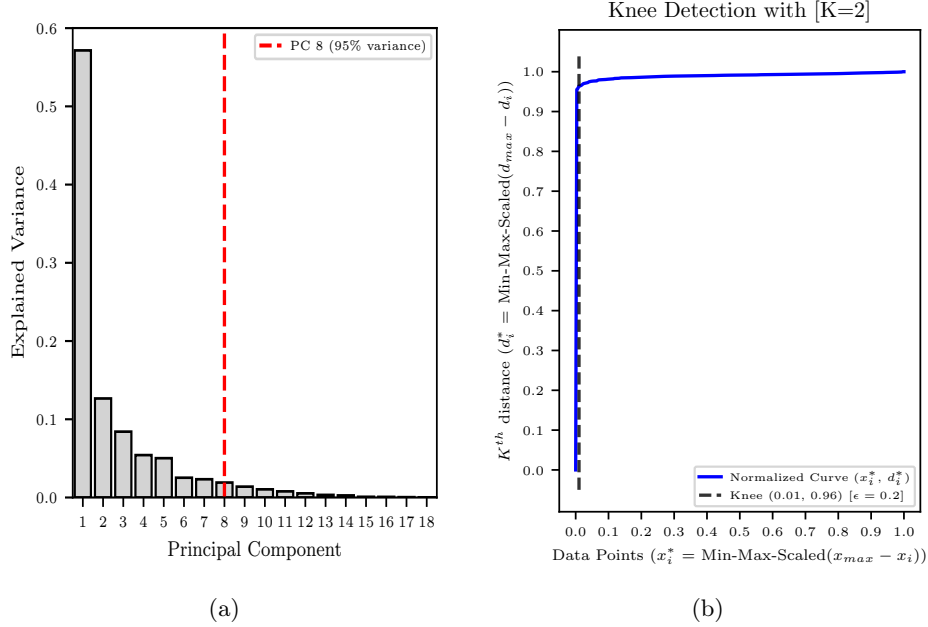
Figure 5.10: **(a) Variance Explained by Principal Component Analysis (PCA) for Nifedipine**: Cumulative variance explained by principal components (PCs) derived from Nifedipine-treated samples. The red vertical line marks the retention of 8 components (out of 18 total) required to capture 95% of the variance. These 8 PCs were used in subsequent clustering of feature loadings to group concentration-dependent responses. **(b) Kneedle Algorithm Fit for Determining DBSCAN Clustering Parameter in Nifedipine Analysis**: K-nearest neighbors (k-NN, k=2) distance curve (d_i) sorted in decreasing order is used to determine the optimal ε for DBSCAN clustering. The Kneedle algorithm inverts the distance curve for concavity-based detection. It then normalizes sample indices (x-axis) and distances (y-axis) to identify the knee point (black dotted line). The unnormalized distance (i.e. d_i) at this knee (0.2) was selected as ε.

The full procedure is summarized in Procedure 2.

Procedure 2 Principal Components Influence Analysis (PCIA)

Input: Treatment response data with features across concentrations
Output: Feature clusters exhibiting similar variation and Independently varying features (Noise points)

1: Perform PCA on the treatment data and retain components explaining 95% variance.
2: Compute k-nearest neighbors distances for features with $k = 2$.
3: Sort distances in decreasing order.
4: Apply Kneedle algorithm to determine the knee point; take the unnormalized distance (d_i) at knee as ε.
5: Perform DBSCAN clustering using ε and min_samples $= 2$ (minimum samples required to form a cluster).
6: Return noise points and feature clusters.

Applying PCIA to *Nifedipine* resulted in one major feature cluster, consisting of six features with similar PC loading contributions (Figure 5.11a). The similarity in their behavior was further confirmed by visualizing their relative change across different drug concentrations (Figure 5.11b).

- Most features exhibited a consistent trend across concentrations.

- One feature, *Force Peak Amplitude*, was clustered with the other features but showed greater variation at higher concentrations in the relative change plot. This suggests that its response pattern deviates from the cluster assignment made by DBSCAN, potentially due to the heuristic-dependent parameters of the algorithm. Further refinement of the clustering approach, such as adjusting distance metrics, could improve feature grouping accuracy.

Overall, this methodology systematically identifies biologically meaningful feature clusters by grouping features with similarly varying responses across treatment concentrations. It reveals functional relationships between biological signals and provides a basis for generating hypotheses for further experimental validation, aiding in the understanding of treatment effects and their physiological implications.
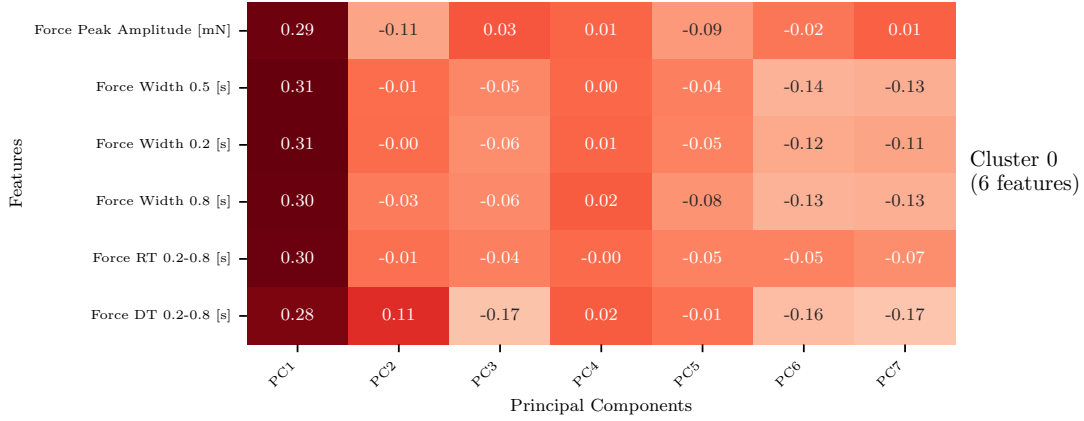
(a)
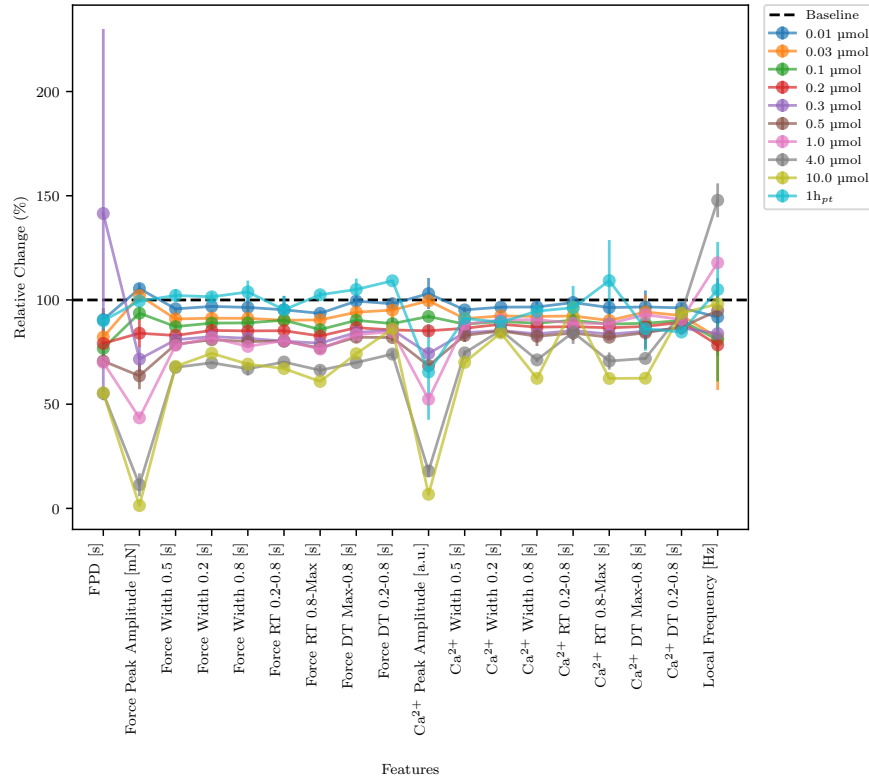
(b)

Figure 5.11: (a) **Feature Clusters Derived from DBSCAN on Nifedipine Principal Component Loadings**: This figure presents the DBSCAN clustering results applied to the principal component loadings of Nifedipine-treated samples. The heatmap visualizes the identified feature cluster (comprising six features) that exhibit similar loading contributions, highlighting their shared variation patterns across principal components. (b) **Relative Change in Nifedipine-Induced Feature Modifications Across Concentrations**: This figure displays the relative changes in biological features under Nifedipine treatment, averaged across tissues. The trends illustrate how feature values progressively shift with increasing drug concentrations. Most features identified in cluster from (a), namely *Force Peak Amplitude, Force Width 0.2, Force Width 0.5, Force Width 0.8, Force RT 0.2-0.8, and Force DT 0.2-0.8* follow a consistent varying pattern across concentrations, confirming their functional grouping as identified by DBSCAN. However, the deviation observed in *Force Peak Amplitude* at higher concentrations suggests that further refinement of clustering methods, such as alternative distance metrics, may improve clustering accuracy.

5.3.4 Section Summary

This chapter presented a systematic approach to modeling drug responses and identifying biologically relevant feature variations at treatment-effective concentrations. Hill equation fitting successfully quantified drug potency, with estimated EC_{50}/IC_{50} values. The relative comparison framework enabled cross-treatment analysis by normalizing responses, while PCIA revealed feature clusters with shared variation patterns across treatment concentrations, providing insight into underlying biological mechanisms. These findings demonstrate the utility of quantitative modeling in drug effect analysis, supporting further experimental validation and refinement of in-vitro MT based cardiac models.

5.4   Correlation Trend Analysis

After establishing the dose-response relationships and assessing relative changes in features under pharmacological treatments in the previous chapter, this analysis focuses on investigating correlation patterns between feature pairs across different pharmacological conditions. The goal is to establish robust estimates of inter-feature relationships and observe how they align or deviate from baseline conditions. This can help identify biologically relevant feature associations that persist or shift under treatment influence.

5.4.1   Methodology

To analyze correlations that are consistent irrespective of tissues and concentrations of pharmacological experiment, bootstrapping was used to estimate the distribution of correlation coefficients between feature pairs under different pharmacological conditions and the baseline. Since the interest lies in identifying global correlation trends persisting under treatment exposure rather than tissue and concentration specific effects, the data from all baseline experiments and all pharmacological conditions were combined, giving rise to a distinct dataset for baseline and each pharmacological condition.

From each dataset, $N = 1000$ bootstrap samples were generated, each containing half the size of the original dataset, with replacement. This approach distorts tissue-level and concentration-level effects, allowing for a more generalized view of correlation patterns under pharmacological influence. For each bootstrap sample, the **Spearman correlation coefficient** was computed for all feature pairs, and its distribution was estimated. The same procedure was repeated for each pharmacological condition as well as for the baseline condition.

Spearman correlation was chosen due to its ability to capture non-linear relationships, which are often biologically relevant. The distribution of bootstrapped Spearman correlations for a subset feature pairs under baseline conditions, along with their kernel density estimations, is shown in Figure 5.12.
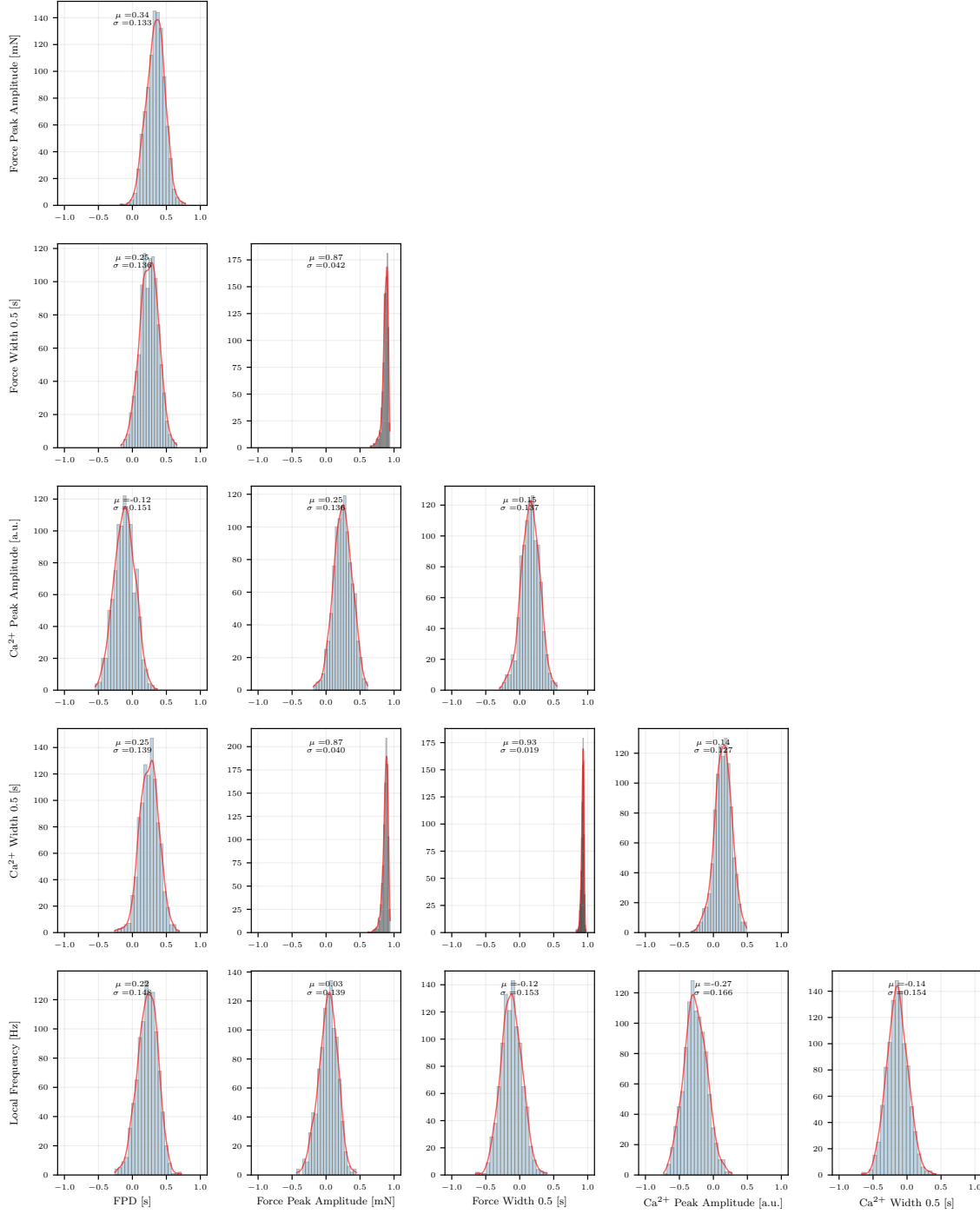
Figure 5.12: **Bootstrapped Distribution of Spearman Correlations for subset of Feature Pairs Under Baseline Conditions**: This figure presents the distribution of bootstrapped Spearman correlation coefficients for a subset of feature pairs in the baseline condition (N=1000). The kernel density estimation (KDE) curves overlaying the histograms provide a smooth representation of correlation distributions. The estimated mean(μ) and sigma(σ) of bootstrapped correlation distribution for each feature pair are also shown.

5.4.2 Results and Interpretation

To determine whether a correlation was significant, a 95% confidence interval (CI) criterion was applied. The nature of correlation was classified based on the following conditions:

- **Positive correlation**: The 95% confidence interval does not include zero, and all values within the interval are positive.

- **Negative correlation**: The 95% confidence interval does not include zero, and all values within the interval are negative.

- **No significant correlation**: The 95% confidence interval includes zero, indicating that the relationship is not statistically significant.

To compare correlation patterns across treatment conditions, a Venn diagram was generated to visualize the overlap of correlation classifications (positive, negative, or no correlation) across pharmacological conditions (Figure 5.13). To specifically analyze treatment-induced effects, the Venn diagram focuses on how correlations persist or shift under different treatments, excluding baseline conditions. This helps differentiate between treatment-specific and shared biological interactions:

- **37** feature pairs exhibited the same correlation nature (either positive or negative) across all three conditions, suggesting that these relationships persist regardless of treatment.

- **E-4031** shared correlation patterns with **Nifedipine** for **4** feature pairs and with the **Calcium Titration experiment** for **9** feature pairs, respectively.

- **Nifedipine** and the **Calcium Titration experiment** exhibited **67** shared feature pair correlations, indicating that their effects on feature relationships are more similar to each other than to **E-4031**.

- Unique correlation patterns were observed for **1, 14**, and **16** feature pairs under **E-4031, Nifedipine**, and the **Calcium Titration experiment**, respectively, suggesting that these pharmacological stimuli induce distinct interactions between certain features.
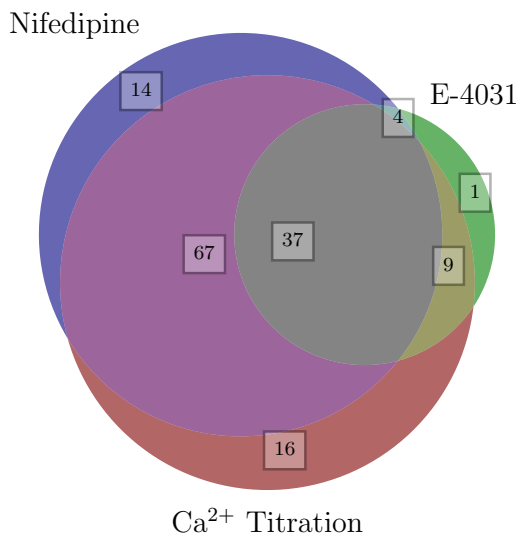
Figure 5.13: **Venn Diagram of Shared Correlation Patterns Across Treatments**: This figure illustrates the overlap in correlation classifications (positive, negative, or non-significant) among the three pharmacological treatments: E-4031, Nifedipine, and Calcium (Ca^{2+}) Titration. A total of 37 feature pairs exhibit consistent correlations across all three treatments, suggesting stable biological associations. Nifedipine and Calcium Titration share 67 feature pair correlations, indicating a stronger similarity in their effects. In contrast, E-4031 shares only 4 correlations with Nifedipine and 9 with the Calcium Titration experiment. Unique correlation patterns are observed for 1, 14, and 16 feature pairs under E-4031, Nifedipine, and Calcium Titration, respectively, highlighting treatment-specific effects on feature relationships.

Since the baseline condition was not included in the Venn diagram, a separate visualization was developed to compare treatment-induced correlation trends to baseline. More briefly, a trend comparison was visualized in a line plot (Figure 5.14). In this plot:

- The x-axis represents feature pairs (mapped to a unique id).

- The y-axis shows the bootstrapped mean of Spearman correlation coefficients for each feature pair, along with the shaded 95% confidence interval.

Feature pairs were sorted based on their baseline correlation, with strongly positive correlations on the left side of the plot transitioning toward negative correlations on the right. Due to the large number of feature pairs and for visualization clarity, each feature pair on the x-axis is mapped to a unique identifier instead of displaying full feature names. The dictionary defining the mapping between feature pairs and their corresponding IDs is provided in the supplementary section (Section A.4).

This visualization provides a high-level, bird's-eye view of how different treatments alter biological interactions. It captures general trends in feature associations, offering insights into potential drug and treatment mechanisms and side effects. However, due to the large number of feature pairs, detailed interpretation requires zooming in on specific associations.

A key consideration in interpreting the results is the degree of overlap in confidence intervals. Many confidence intervals, particularly those around zero (seen in the right portion of the plot for baseline), suggest that the correlation is not statistically significant. Therefore, only feature pairs where the confidence interval does not include zero should be considered for further investigation.

Example Interpretation: For instance, the correlation between *Duration* and *Force Peak Amplitude* shows a significant positive correlation under baseline conditions, which is retained under Nifedipine and E-4031 but not in the Calcium Titration experiment, where the mean correlation shifts toward negative values. This suggests that the relationship between *Duration* and *Force Peak Amplitude* remains stable under Nifedipine and E-4031, but is altered by Calcium Titration, possibly due to calcium availability significance in ECC (Figure 5.14).
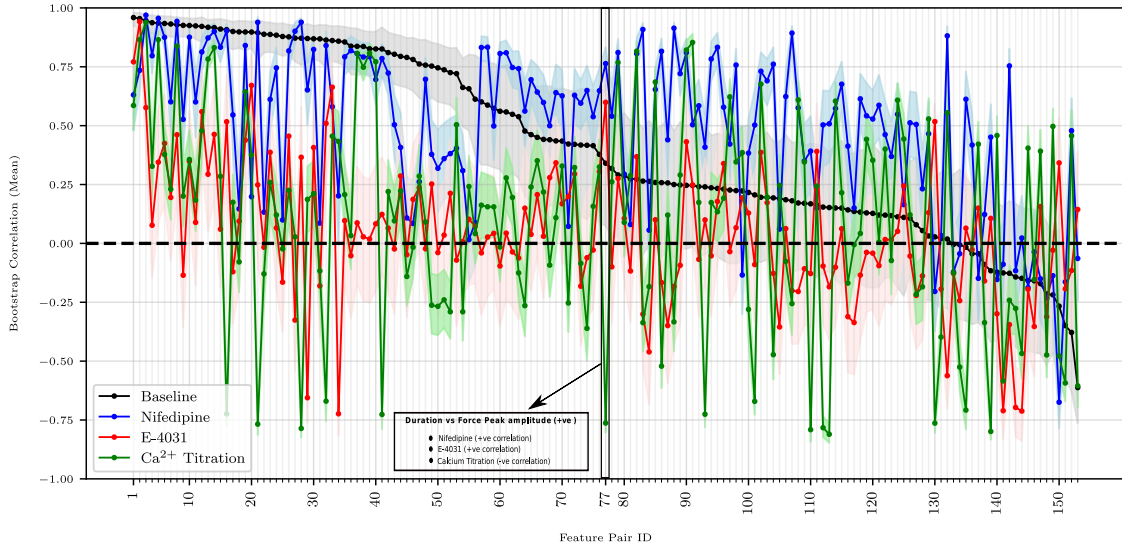
Figure 5.14: **Treatment-Induced Changes in Feature Pair Correlation Trends Compared to Baseline**: This figure visualizes the mean of bootstrapped distribution of Spearman correlation coefficients for **153 feature pairs** across baseline and treatment conditions. The x-axis represents feature pairs, each mapped to a unique identifier for visualization clarity instead of displaying full feature names (Section A.4). Feature pairs are sorted from **strong positive correlations** (left) to **strong negative correlations** (right) based on baseline correlation. The y-axis shows the mean correlation values, with shaded regions indicating **95% confidence intervals**. Feature pairs with confidence intervals overlapping zero indicate weak or non-significant correlations. **Example Interpretation (Middle Box):** The correlation between *Duration* and *Force Peak Amplitude* is significantly positive under baseline conditions and remains stable under Nifedipine and E-4031. However, in Calcium Titration, the mean correlation shifts toward negative values, suggesting a disruption in their relationship.

Such a correlation trend can be used to identify feature pair associations that differ or persist across drug and treatment conditions. Additionally, some correlations may strengthen or weaken under treatment conditions compared to baseline, providing insights into treatment-induced interactions.

5.4.3   Biological Implications and Summary

This analysis provides a high-level, bird's-eye view of how pharmacological treatments influence biological feature associations by examining global correlation patterns across treatments. Bootstrapped Spearman correlations were used to identify both stable biological relationships and treatment-specific shifts, offering initial insights into potential modifications in biological pathways. Feature pairs that maintain consistent correlations across all conditions may represent fundamental biological interactions, whereas treatment-specific changes could indicate distinct pharmacological effects or potential side effects due to cross-treatment interactions.

It is important to emphasize that these findings represent broad statistical associations rather than mechanistic causal relationships among feature pairs. As an initial screening tool, this analysis helps characterize treatment effects at a systems level, guiding future studies toward more targeted experimental validation or predictive modeling to explore specific biological mechanisms in greater depth.

5.5 Modeling Treatment-Induced Feature Associations

Following the exploration of global correlation trends among feature pairs, the final phase of the analysis focuses on identifying statistical associations between individual features under both baseline and pharmacological/experimental conditions through multiple regression. While this analysis does not provide a mechanistic explanation of tissue dynamics, it establishes an essential statistical framework for identifying potential regulatory interactions. This framework serves as a starting point for further modeling and experimental validation of treatment effects on MTs.

To quantify these associations, a mixed modeling approach with **LASSO regression** was applied, where each feature was considered as the target variable and regressed against all other features. This was performed across both baseline and three experimental conditions of this study, allowing for the identification of statistical relationships that persist or change in response to pharmacological stimuli.

5.5.1 Model Specification

As in previous analyses, data across different treatment concentrations and tissues were merged to study the overall effect. However, as discussed in significance analysis (Section 5.2), some tissues lacked data points for certain concentrations (Tables 5.3 and 5.4), creating an unbalanced repeated-measures design. This was addressed using an appropriately nested mixed-effects model. However, in the current analysis, where each feature is regressed over all remaining features, the high-dimensional nature of the dataset necessitated the use of **LASSO regularization** to shrink coefficients of non-contributory predictors.

Due to current limitations in applying LASSO to mixed models, only the `glmmLASSO` [38] package was available, which does not support a fully nested structure. To accommodate this, concentrations with missing data for one or more tissues were removed, converting the dataset into a crossed design. The final mixed model incorporated **tissue** and **concentration** as **random effects** with an intercept, while all other features were included as **fixed effects**. The general model structure is given by:

$$x_j = \beta_0 + \sum_{i \neq j} \beta_i x_i + Z_\alpha \cdot [\text{Tissue ID}] + Z_\beta \cdot [\text{Concentration}] + \epsilon$$

where:

- x_j is the target feature, analogous to Y in the statistical significance model 5.2.1.1.

- β_0 is the intercept, capturing the overall mean response.

- x_i represents all other features ($i \neq j$) included as fixed effect covariates with coefficients β_i, extending the previous model by incorporating additional predictors.

- Z_α is the random intercept for **tissue**, modeled as $Z_\alpha \sim \mathcal{N}(0, \sigma_\alpha^2)$, consistent with the statistical significance model.

- Z_β is the random intercept for **concentration**, modeled as $Z_\beta \sim \mathcal{N}(0, \sigma_\beta^2)$, thereby transitioning concentration from a fixed effect to a random effect in this model.

- $\epsilon \sim \mathcal{N}(0, \sigma^2)$ represents the residual error.

This formulation assumes that the random effects of tissue and concentration are uncorrelated. However, in a more generalized setting, a full covariance structure[10] could be used to model potential correlations between these random effects.

For each regression, the optimal LASSO shrinkage parameter (λ) was determined through **5-fold cross-validation**, selecting the value that minimized residual deviance. Baseline conditions from all experiments were combined to serve as a reference.

5.5.2 Results and Interpretation

The results of the regression analysis are visualized in Figure 5.15, where instead of displaying numerical coefficient values, a color-coded heatmap representation is used to facilitate comparison between baseline and treatment-induced conditions. Each regression's optimal five-fold cross-validated λ value is also provided.

As highlighted earlier, this analysis is not intended to serve as a definitive mechanistic model but rather as a **high-level exploratory framework** for identifying statistical associations among multiple features that may guide further investigations. Some consistent coefficient patterns are observable across both treatment and baseline conditions, particularly in the bottom right of Figure 5.15, where strong statistical associations persist regardless of pharmacological intervention. The observed differences in regression coefficients across treatment conditions may provide valuable insights into how specific features interact under varying pharmacological influences.

5.5.3 Biological Implications and Summary

This chapter applied LASSO-regularized based mixed modeling regression to identify statistical associations between biological features under baseline and pharmacological/experimental conditions. By regressing each feature against all others while accounting for tissue and concentration variability, the analysis revealed both persistent feature interactions and condition-specific shifts in association patterns. While these results do not imply causality, they offer a high-level statistical framework to guide future investigations into regulatory changes induced by pharmacological stimulus.

---

[10]Mixed models allow for the specification of covariance structures between random effects, such as an **unstructured** or **compound symmetry** covariance matrix. These structures model dependencies between random effects rather than assuming independence. For further details, see [39]
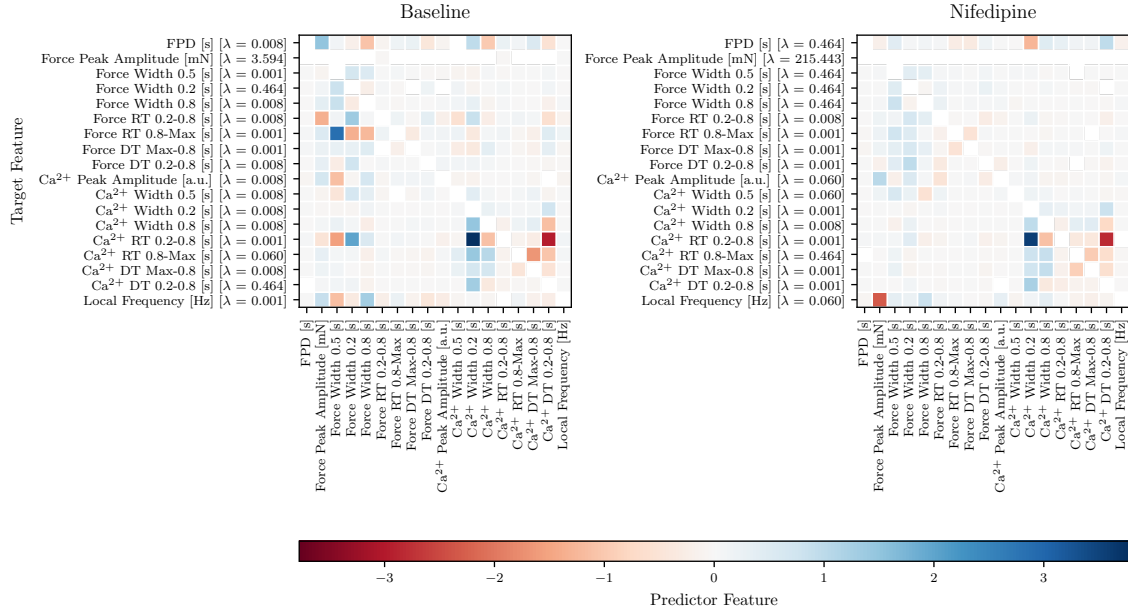
Figure 5.15: **Feature Associations Identified via Generalized Linear Mixed Model with LASSO Regularization (Baseline & Nifedipine)**: The heatmaps visualizes the regression coefficients obtained from the LASSO-regularized generalized linear mixed model (GLMM-LASSO) applied to the dataset under baseline and Nifedipine drug. The analysis identifies statistical associations between features while accounting for inter-tissue and concentration-level variability. Each cell represents the magnitude and direction of the estimated coefficient for a given feature pair, with a color gradient indicating positive or negative associations. The LASSO shrinkage parameter (λ) was selected using 5-fold cross-validation, ensuring optimal coefficient regularization while minimizing residual deviance. Persistent feature relationships across both baseline and drug condition (observed in the lower right region of the heatmap) suggest stable biological interactions, while changes in coefficient patterns under drug treatments indicate pharmacologically induced modifications in feature associations. This model serves as an exploratory framework for identifying key regulatory interactions that may guide further experimental validation and mechanistic modeling of treatment-induced effects in MTs. The heatmaps for E-4031 and Calcium Titration are provided in supplementary figure (Figure A.7)

Chapter 6

Integrative Discussion and Conclusion

6.1 Integrative Summary of Thesis Work

This thesis develops a comprehensive, data-driven pipeline to characterize how different pharmacological interventions modulate ECC in iPSC-derived MTs. The motivation stems from limitations in existing approaches, which often consider ECC parameters independently without accounting for their interrelationships or variability due to experimental conditions. To address these gaps, the thesis integrates multimodal data—contractile force, calcium transients, and field potentials—under interventions with E-4031, Nifedipine, and Calcium Titration.

The analytical workflow begins with meticulous preprocessing designed to improve signal fidelity. Signal noise was reduced using Butterworth and Savitzky-Golay filters for force and calcium signals, and discrete wavelet transform-based denoising for field potentials. Baseline offsets due to sensor drift were corrected using Gaussian-based histogram fitting methods, ensuring accurate signal representations. Subsequently, peak-detection algorithms identified contraction events, which were further filtered to exclude arrhythmic contractions using machine learning classifiers.

Following preprocessing, biomarkers quantifying the ECC cycle—such as peak amplitudes, contraction and relaxation times, and field potential durations—were systematically extracted and fed into a suite of analytical methods. Initially, a global overview was obtained using t-SNE visualization to detect inherent patterns and potential clustering across treatments and concentration levels, providing immediate insights into data structure and treatment similarity.

Statistical significance was assessed through mixed-effects modeling, accounting explicitly for tissue-to-tissue and tissue-to-concentration variability, confirming known pharmacological effects. Dose–response (Hill) modeling further characterized the relative impact of treatments, establishing the potency and effect magnitude for each treatment compared to baseline conditions. Principal Components Influence Analysis (PCIA) was developed to identify groups of features responding similarly across treatment concentrations, revealing clusters indicative of jointly modulated biological processes.

Subsequent correlation analysis established global feature correlation trends, identifying both conserved interactions suggestive of shared pathways and novel, treatment-specific correlations in-

dicative of treatment-specific effects. Finally, LASSO-regularized regression modeling evaluated how each individual feature was influenced by other features, providing foundational information necessary for future mechanistic modeling and deeper insights into feature interplay under pharmacological influence. A concise summary of the applied methods and their outcomes is provided in Table 6.1.

The combined outcomes validated known mechanisms, such as E-4031 prolonging repolarization and Nifedipine reducing contraction force, while also uncovering novel feature interactions that inform future research into ECC modulation. Collectively, this thesis provides a framework for future researchers aiming to understand and characterize pharmacological effects of different treatments on ECC modulation of MTs.

6.2 Discussion on Methodology

6.2.1 Preprocessing and Signal Enhancement

The preprocessing approach encompassed noise reduction through low-pass filtering for force and calcium signals, and Discrete Wavelet Transform (DWT) based denoising for field potential (FP) signals (Section 4.1). While effective, the empirical determination of wavelet threshold values and baseline correction parameters can introduce biases—particularly if denoising inadvertently delays signal peaks or attenuates amplitudes, which may significantly affect subsequent analyses. Future studies could explore automated, data-driven methods for selecting optimal denoising thresholds to enhance robustness and reliability.

Peak detection for contraction-event identification relied on prominence-based algorithms, demonstrating effectiveness under most conditions (Section 4.2). However, the thresholds chosen here may not generalize well to datasets with different noise characteristics, and inaccurate thresholds can bias results by either missing real peaks or including noise spikes. Developing more adaptive or data-driven thresholding methods remains a critical avenue for improvement, especially for low signal-to-noise ratio scenarios.

6.2.2 Feature Visualization and Statistical Significance

Feature space visualization using t-SNE provided valuable insights into global data structure, highlighting overlapping pharmacological effects and distinct tissue-specific responses (Section 5.1). Nevertheless, the inherent limitation of t-SNE—its inability to accurately preserve original data distances—restricts quantitative interpretation, suggests linear dimensionality reduction techniques might provide complementary insights.

Statistical significance testing via mixed-effects modeling effectively handled repeated measures and nested treatment designs, confirming drug-induced physiological effects consistent with literature (Sections 5.2.1–5.2.3). Given only two tissue replicates, simpler statistical methods like ANOVA could have sufficed, but the chosen mixed modeling approach is justified due to its scalability for future larger studies involving multiple tissues.

6.2.3 Drug Response Modeling and Limitations

The application of Hill equation modeling facilitated quantitative drug response characterization, providing an estimate of EC_{50}/IC_{50} values for the studied pharmacological stimulus (Section 5.3.1).

Table 6.1: Summary of Methodologies and Outcomes

Method Step	Techniques Applied	Outcome	Section
Signal Preprocessing	Butterworth filtering, Savitzky-Golay smoothing, DWT denoising	Enhanced signal quality across modalities	4.1
Peak Detection	Prominence thresholding, Gaussian Mixture Models	Robust contraction event detection	4.2
Arrhythmia Detection	Machine learning classification (Random Forest)	High-fidelity arrhythmia exclusion	4.3
Feature Extraction	Peak characterization, Duration measurements	Comprehensive biomarker extraction	4.4
Global Visualization	t-SNE dimensionality reduction	Global overview of feature space and potential identification of clusters	5.1
Statistical Testing	Mixed-effects modeling	Establishment of significant treatment effects on biomarkers	5.2
Dose-response & Relative Feature Analysis	Hill equation modeling	Quantitative treatment effect estimation against baseline	5.3.1–5.3.2
PCIA	PCA with DBSCAN clustering	Identification of co-varying biomarkers	5.3.3
Correlation Analysis	Bootstrapped correlation analysis	Discovery of treatment-resistant and treatment-specific associations	5.4
Regression Modeling	Regularized (LASSO) mixed-effects regression	Establishment of foundational information for mechanistic modeling	5.5

However, limited concentration points restricted model accuracy. The substantial gaps between tested concentrations impeded precise curve-fitting, highlighting a clear need for increased data points at intermediate concentrations for improved curve fitting and reliable pharmacological characterization in future work.

Grouping of features using Principal Component Influence Analysis (PCIA) through DBSCAN clustering effectively identified feature clusters which followed similar variation patterns across different concentrations of a treatment, demonstrating consistent pharmacological modulation (Section 5.3.3). Yet, heuristic parameter selection for DBSCAN is a clear limitation. Future research should implement more robust, parameter-agnostic clustering algorithms to enhance reproducibility and accuracy.

6.2.4 Feature Associations and Correlation Analysis

Correlation trend analysis, involving bootstrapped correlation distributions across treatments, effectively captured global feature relationships, highlighting persistent and unique feature pair associations (Section 5.4). However, these correlation analyses only suggest statistical relationships, not causal interactions, and should be regarded as exploratory rather than definitive. Additional experiments or mechanistic studies would be needed to ascertain causal pathways.

The regression-based association study using mixed-model LASSO regression provided a useful preliminary tool for identifying inter-feature relationships while controlling for experimental variability (Section 5.5). A key limitation was the inability of this approach to explicitly model nested structures (concentration within tissue). More sophisticated approaches, such as Bayesian hierarchical modeling for incorporating biological feature priors or non-linear interpretable methods like decision trees, could be explored in future studies to better capture biological complexity.

A further overarching limitation of this research is its reliance on only two tissue replicates for pharmacological testing. While sufficient for a proof of concept, larger sample sizes and additional tissues would provide stronger statistical power and broader generalizability of findings. Integrating more tissue replicates, increased concentration sampling, and parameter-agnostic algorithms in future work will enhance the robustness, reproducibility, and practical impact of these methodological advances.

6.3 Biological Implications

This thesis bridges a key gap established in the literature (Chapter 2) by integrating multimodal data (electrical, calcium, and mechanical) with robust statistical modeling to explore both individual biomarker changes and cross-biomarker relationships in ECC. Whereas many prior studies focused on isolated endpoints, the approach here systematically correlates multiple ECC parameters, revealing both conserved interactions (indicating fundamental physiological processes) and treatment-specific alterations (highlighting treatment-related mechanisms). By explicitly accounting for tissue-level variability through mixed-effects modeling, the framework advances beyond standard methods that often overlook inter-tissue differences. Overall, combining signal processing, correlation analyses, dose–response estimation, and regression in a unified pipeline demonstrates how high-content, multimodal data can yield deeper biological insights than conventional single-parameter approaches in ECC pharmacodynamics.

6.4   Future Research Directions

Although this pipeline effectively characterizes and interprets treatment-induced changes in ECC, several avenues remain for further enhancement. First, while arrhythmic events were excluded in this study, future work could incorporate arrhythmic events, potentially enhancing predictions of pro-arrhythmic risk. Second, applying the same analysis framework to other muscle tissue types—with appropriate parameter tuning for denoising and peak detection—would broaden its relevance beyond cardiac contexts. Third, increasing the number of tissue replicates and sampling finer concentration gradients would strengthen statistical power, refine dose–response modeling, and better capture biological variability. Finally, exploring more advanced or non-linear modeling strategies, such as hierarchical Bayesian and Tree-based approaches, could uncover additional biomarker interdependencies and help unify empirical data with mechanistic models of cardiac physiology.

6.5   Conclusion

This thesis developed a robust, data-driven analytical framework to systematically characterize how pharmacological interventions modulate ECC in hiPSC-derived MTs. Despite certain methodological constraints—such as empirical threshold settings, limited tissue replicates, and inadequate concentration sampling—the proposed approach effectively captured physiologically relevant treatment responses. By combining advanced signal preprocessing, peak detection, arrhythmia classification, and feature extraction with rigorous statistical learning–based modeling, the pipeline reliably characterized treatment-induced physiological responses.

While heuristic-dependent parameters in signal preprocessing and feature extraction present clear avenues for refinement, the integrative methodology nonetheless revealed treatment-specific and conserved patterns indicative of pharmacological modulation. These insights reinforce the growing utility of multimodal MTs in human cardiac pharmacology, underscoring the potential for improved preclinical screening.

Moving forward, future work should address current limitations by expanding the dataset, refining signal processing and arrhythmia detection, and exploring more sophisticated modeling strategies (e.g., decision trees). Such endeavors would further enhance the scalability and translational value of MT-based assays, laying the groundwork for more precise preclinical predictions in cardiac drug development.

References

[1] Dandan Zhao, Wei Lei, and Shijun Hu. "Cardiac organoid — a promising perspective of preclinical model". In: *Stem Cell Research & Therapy* 12.1 (May 2021), p. 272. ISSN: 1757-6512. DOI: 10.1186/s13287-021-02340-7. URL: https://doi.org/10.1186/s13287-021-02340-7.

[2] Gary Gintant et al. "Use of Human Induced Pluripotent Stem Cell–Derived Cardiomyocytes in Preclinical Cancer Drug Cardiotoxicity Testing: A Scientific Statement From the American Heart Association". In: *Circulation Research* 125.10 (2019), e75–e92. DOI: 10.1161/RES.0000000000000291. eprint: https://www.ahajournals.org/doi/pdf/10.1161/RES.0000000000000291. URL: https://www.ahajournals.org/doi/abs/10.1161/RES.0000000000000291.

[3] Anna P Hnatiuk et al. "Human iPSC modeling of heart disease for drug development". en. In: *Cell Chem. Biol.* 28.3 (Mar. 2021), pp. 271–282.

[4] TYT Lee, JG Coles, and JT Maynes. "iPSC-cardiomyocytes in the preclinical prediction of candidate pharmaceutical toxicity". In: *Frontiers in Pharmacology* 15 (2024), p. 1308217. DOI: 10.3389/fphar.2024.1308217.

[5] Peng Wu et al. "Maturation strategies and limitations of induced pluripotent stem cell-derived cardiomyocytes". en. In: *Biosci. Rep.* 41.6 (June 2021).

[6] Enrique G Navarrete et al. "Screening drug-induced arrhythmia [corrected] using human induced pluripotent stem cell-derived cardiomyocytes and low-impedance microelectrode arrays". en. In: *Circulation* 128.11 Suppl 1 (Sept. 2013), S3–13.

[7] Paruthi Pradhapan et al. "Cardiomyocyte MEA Data Analysis (CardioMDA) – A Novel Field Potential Data Analysis Software for Pluripotent Stem Cell Derived Cardiomyocytes". In: *PLoS ONE* 8 (2013). URL: https://api.semanticscholar.org/CorpusID:16267073.

[8] Takashi Ohya et al. "Simultaneous measurement of contractile force and field potential of dynamically beating human iPS cell-derived cardiac cell sheet-tissue with flexible electronics". In: *Lab Chip* 21 (20 2021), pp. 3899–3909. DOI: 10.1039/D1LC00411E. URL: http://dx.doi.org/10.1039/D1LC00411E.

[9] Masahide Kawatou et al. "Modelling Torsade de Pointes arrhythmias in vitro in 3D human iPS cell-engineered heart tissue". In: *Nature Communications* 8.1 (Oct. 2017), p. 1078.

[10] Sungjin Min et al. "Versatile human cardiac tissues engineered with perfusable heart extracellular microenvironment for biomedical applications". In: *Nature Communications* 15.1 (Mar. 2024), p. 2564.

[11] Ksenia Blinova et al. "International multisite study of human-induced pluripotent stem cell-derived cardiomyocytes for drug proarrhythmic potential assessment". en. In: *Cell Rep.* 24.13 (Sept. 2018), pp. 3582–3592.

[12] Gary Gintant et al. "Use of human induced pluripotent stem cell-derived cardiomyocytes in preclinical cancer drug cardiotoxicity testing: A scientific statement from the American heart association". en. In: *Circ. Res.* 125.10 (Oct. 2019), e75–e92.

[13] Donald M Bers. "Cardiac excitation-contraction coupling". In: *Nature* 415.6868 (2002), pp. 198–205. DOI: `https://doi.org/10.1038/415198a`.

[14] Shameer Sarwar. *Arrhythmia Classification Using Machine Learning on Cardiac Excitation-Contraction Coupling Data.* Tech. rep. Accessed: 2025-03-08. myotwin GmbH, Mar. 2024. URL: `https://github.com/myotwin/arrhythmia-classification/blob/main/ProjectReport.pdf`.

[15] Wikipedia contributors. *Savitzky–Golay filter — Wikipedia, The Free Encyclopedia.* Accessed: 2025-03-08. Mar. 2025. URL: `https://en.wikipedia.org/wiki/Savitzky%E2%80%93Golay_filter`.

[16] Marília Barandas et al. "TSFEL: Time Series Feature Extraction Library". In: *SoftwareX* 11 (2020), p. 100456.

[17] Wikipedia contributors. *Discrete wavelet transform — Wikipedia, The Free Encyclopedia.* Accessed: 2025-03-08. Mar. 2025. URL: `https://en.wikipedia.org/wiki/Discrete_wavelet_transform`.

[18] Gregory R. Lee et al. "PyWavelets: A Python package for wavelet analysis". In: *Journal of Open Source Software* 4.36 (2019), p. 1237. DOI: `10.21105/joss.01237`. URL: `https://doi.org/10.21105/joss.01237`.

[19] SciPy Developers. *scipy.signal.find_peaks.* Accessed: 2025-03-08. 2025. URL: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html`.

[20] Scikit-learn contributors. *sklearn.mixture.GaussianMixture.* Accessed: 2025-03-13. 2025. URL: `https://scikit-learn.org/stable/modules/generated/sklearn.mixture.GaussianMixture.html`.

[21] R. L. Thorndike. "Who belongs in the family?" In: *Psychometrika* 18.4 (1953), pp. 267–276. DOI: `10.1007/BF02289263`.

[22] Wikipedia contributors. *Akaike information criterion.* Accessed: 2025-03-13. 2025. URL: `https://en.wikipedia.org/wiki/Akaike_information_criterion`.

[23] Geocene Inc. and Contributors. *TRAINSET: A Graphical Tool for Labeling Time Series Data.* Maintained by Geocene Inc. with contributions from Rush Kapoor, Ajay Pillarisetti, Jeremy Coyle, Skot Croshere, Marc Paré, Hamza Benkhay, and Danny Wilson. Accessed: 14-Feb-2025. 2025. URL: `https://trainset.geocene.com/`.

[24] Wikipedia contributors. *Hungarian algorithm — Wikipedia, The Free Encyclopedia.* Accessed: 2025-03-08. Mar. 2025. URL: `https://en.wikipedia.org/wiki/Hungarian_algorithm`.

[25] SciPy Community. *scipy.signal.peak_widths.* Accessed: 2025-03-13. 2025. URL: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.peak_widths.html`.

[26] Laurens van der Maaten and Geoffrey Hinton. "Visualizing Data using t-SNE". In: *Journal of Machine Learning Research* 9.86 (2008), pp. 2579–2605. URL: http://jmlr.org/papers/v9/vandermaaten08a.html.

[27] Fabian Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[28] Holger Schielzeth and Shinichi Nakagawa. "Nested by design: model fitting and interpretation in a mixed model era". In: *Methods in Ecology and Evolution* 4.1 (2013), pp. 14–24. DOI: https://doi.org/10.1111/j.2041-210x.2012.00251.x. eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2041-210x.2012.00251.x. URL: https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-210x.2012.00251.x.

[29] Wikipedia contributors. *Mixed model — Wikipedia, The Free Encyclopedia*. Accessed: 9 March 2025. 2024. URL: https://en.wikipedia.org/wiki/Mixed_model.

[30] Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. "lmerTest Package: Tests in Linear Mixed Effects Models". In: *Journal of Statistical Software* 82.13 (2017), pp. 1–26. DOI: 10.18637/jss.v082.i13.

[31] Wikipedia contributors. *Dunnett's test — Wikipedia, The Free Encyclopedia*. Accessed: 9 March 2025. 2024. URL: https://en.wikipedia.org/wiki/Dunnett%27s_test.

[32] Torsten Hothorn, Frank Bretz, and Peter Westfall. "Simultaneous Inference in General Parametric Models". In: *Biometrical Journal* 50.3 (2008), pp. 346–363.

[33] Wikipedia contributors. *Welch–Satterthwaite equation — Wikipedia, The Free Encyclopedia*. Accessed: 9 March 2025. 2024. URL: https://en.wikipedia.org/wiki/Welch%E2%80%93Satterthwaite_equation.

[34] Hélène Jacqmin-Gadda et al. "Robustness of the linear mixed model to misspecied error distribution". In: *Computational Statistics and Data Analysis* 51.10 (June 2007), pp. 5142–5154. DOI: 10.1016/j.csda.2006.05.021. URL: https://inserm.hal.science/inserm-00084214.

[35] Himoto. *Fitting the Hill Equation to Modeled or Experimental Data*. https://github.com/himoto/hillfit. Accessed: 9 March 2025. 2024.

[36] Ville Satopaa et al. "Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior". In: *2011 31st International Conference on Distributed Computing Systems Workshops*. 2011, pp. 166–171. DOI: 10.1109/ICDCSW.2011.20.

[37] Kevin Arvai. *kneed: Knee Point Detection in Python*. https://github.com/arvkevi/kneed. Accessed: 9 March 2025. 2024.

[38] Lukas Meier Jürg Schelldorfer and Peter Bühlmann. "GLMMLasso: An Algorithm for High-Dimensional Generalized Linear Mixed Models Using 1-Penalization". In: *Journal of Computational and Graphical Statistics* 23.2 (2014), pp. 460–477. DOI: 10.1080/10618600.2013.773239. eprint: https://doi.org/10.1080/10618600.2013.773239. URL: https://doi.org/10.1080/10618600.2013.773239.

[39] José C. Pinheiro and Douglas M. Bates. *Mixed-Effects Models in S and S-PLUS*. 1st ed. Statistics and Computing. Springer New York, NY, 2000, pp. XVI, 528. ISBN: 978-0-387-98957-0. DOI: 10.1007/b98882.

Appendix A

Supplementary Material

A.1 Significance Analysis

The significance plots and diagnostic plots for E-4031 (Figures A.1, A.2, A.3) and Nifedipine (Figures A.4, A.5 and A.6) are presented below.

Figure A.1: **Statistical significance of E-4031 treatment on remaining Biological features.**.

Figure A.2: **QQ Plot of Residuals for Mixed Model Fit of E-4031 Treatment (Remaining Features)**

Figure A.3: **Residuals vs. Fitted Values for Mixed Model Fit of E-4031 Treatment (Remaining Features).**
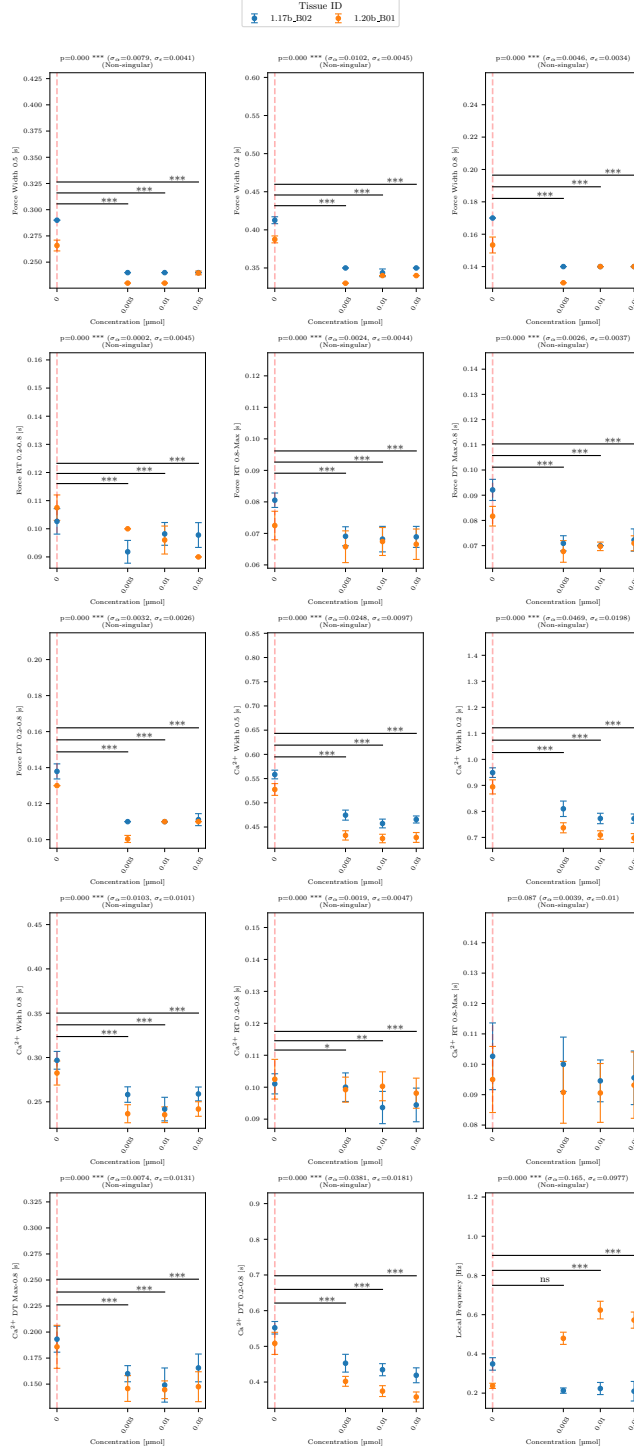
Figure A.4: **Statistical significance of Nifedipine treatment on remaining Biological features.**.
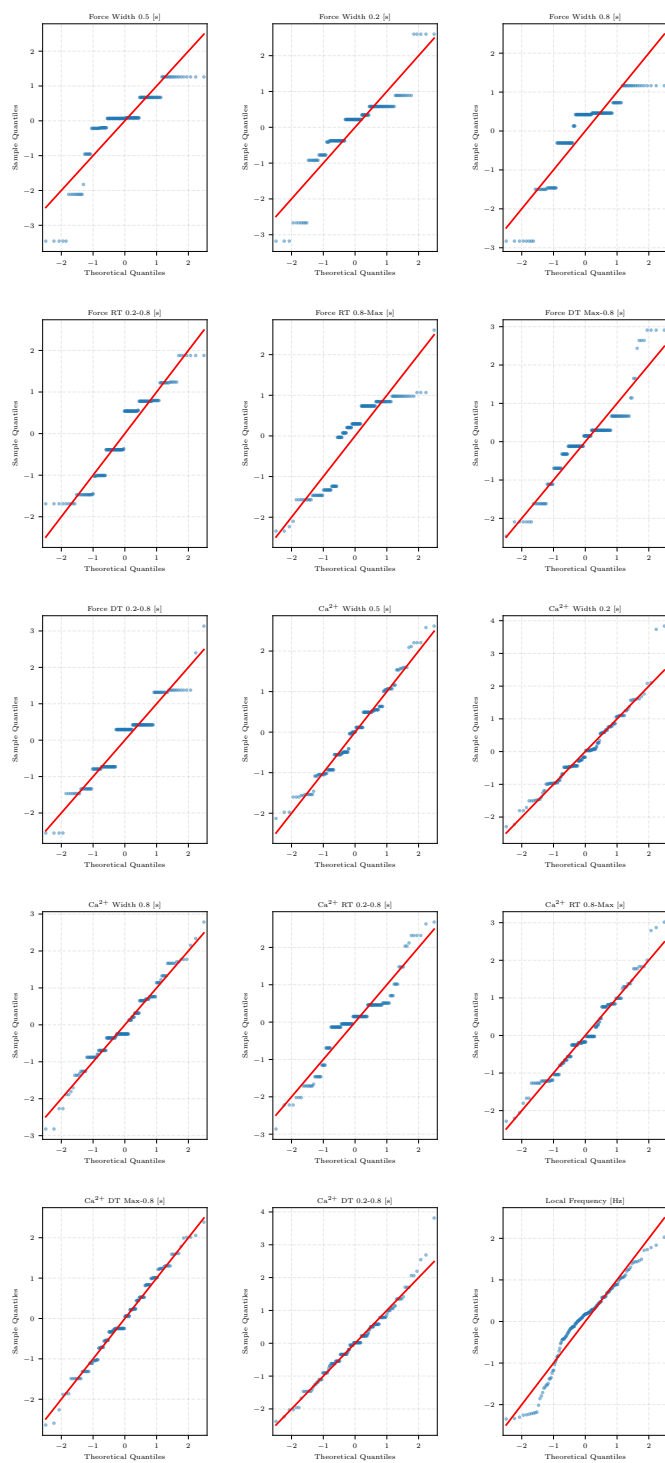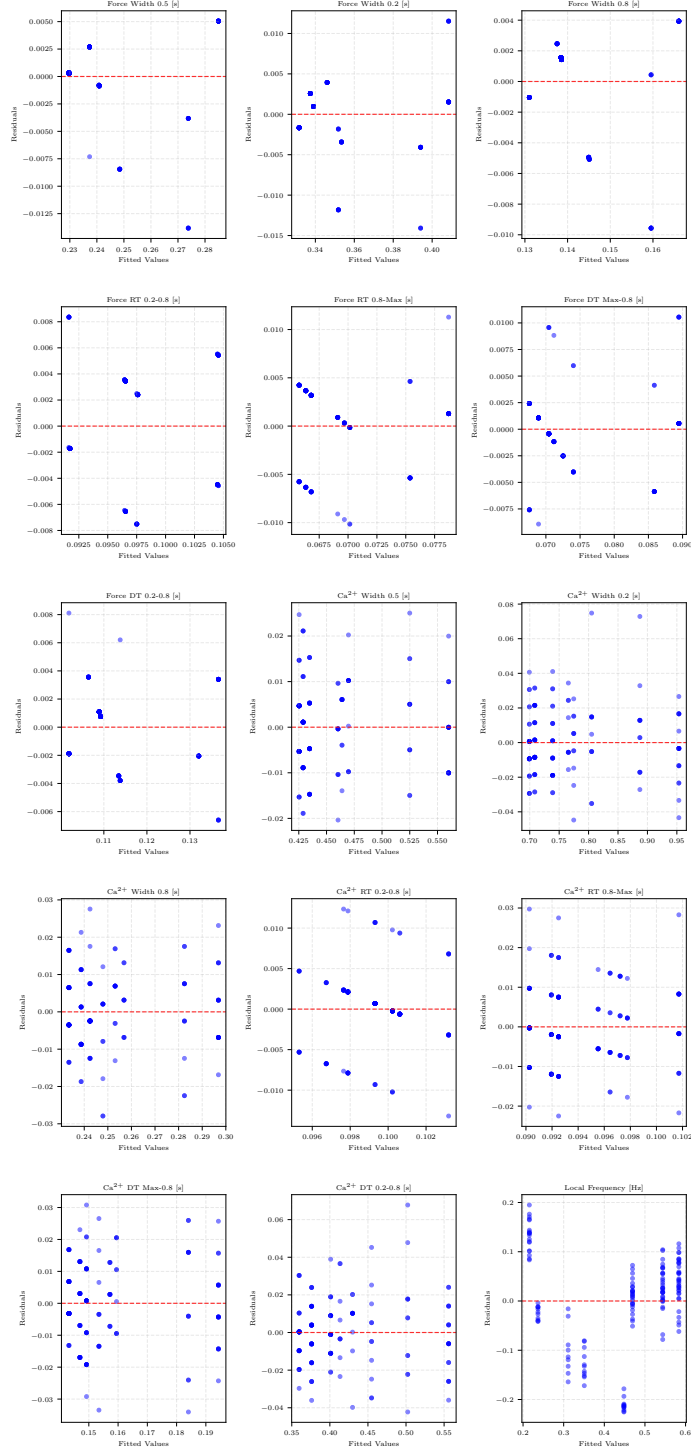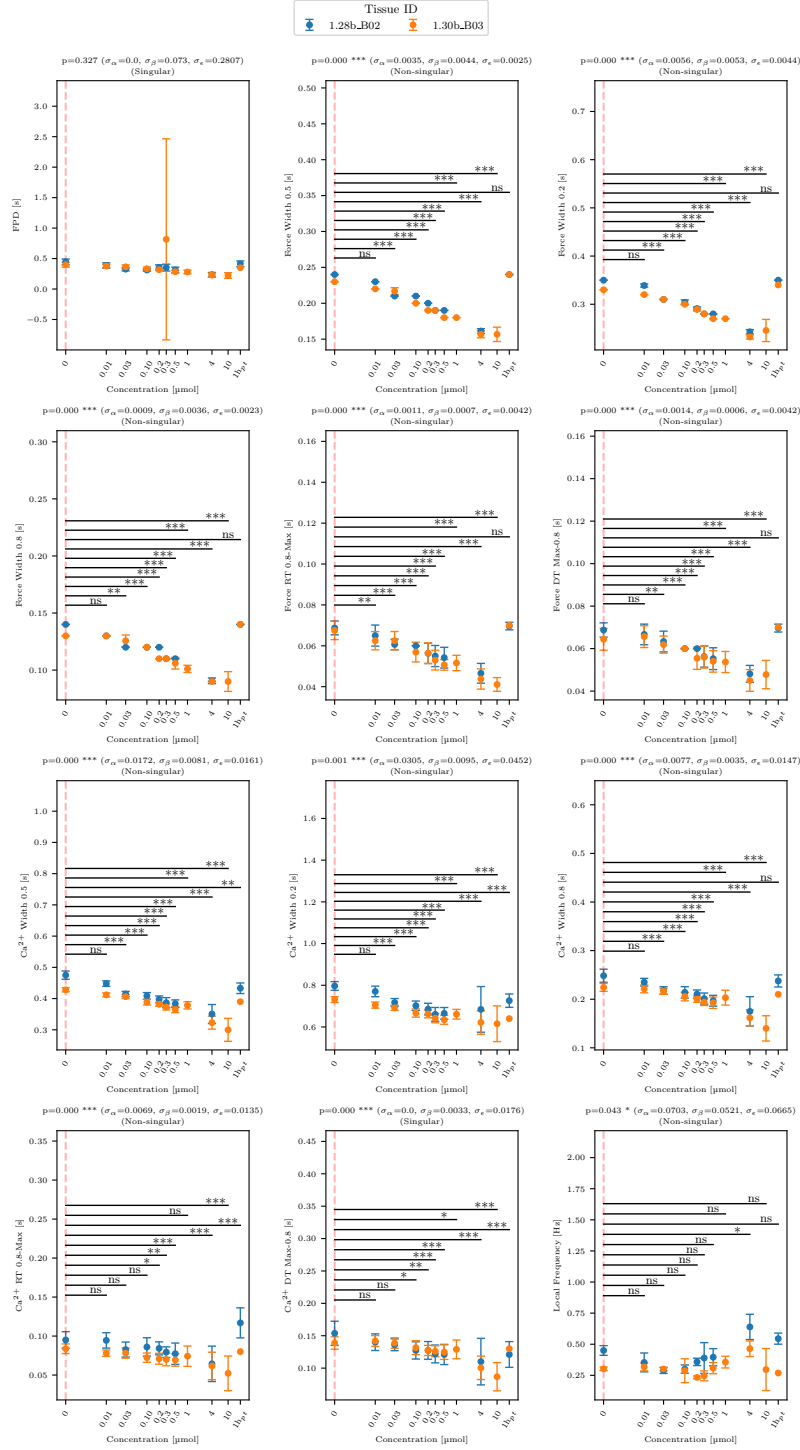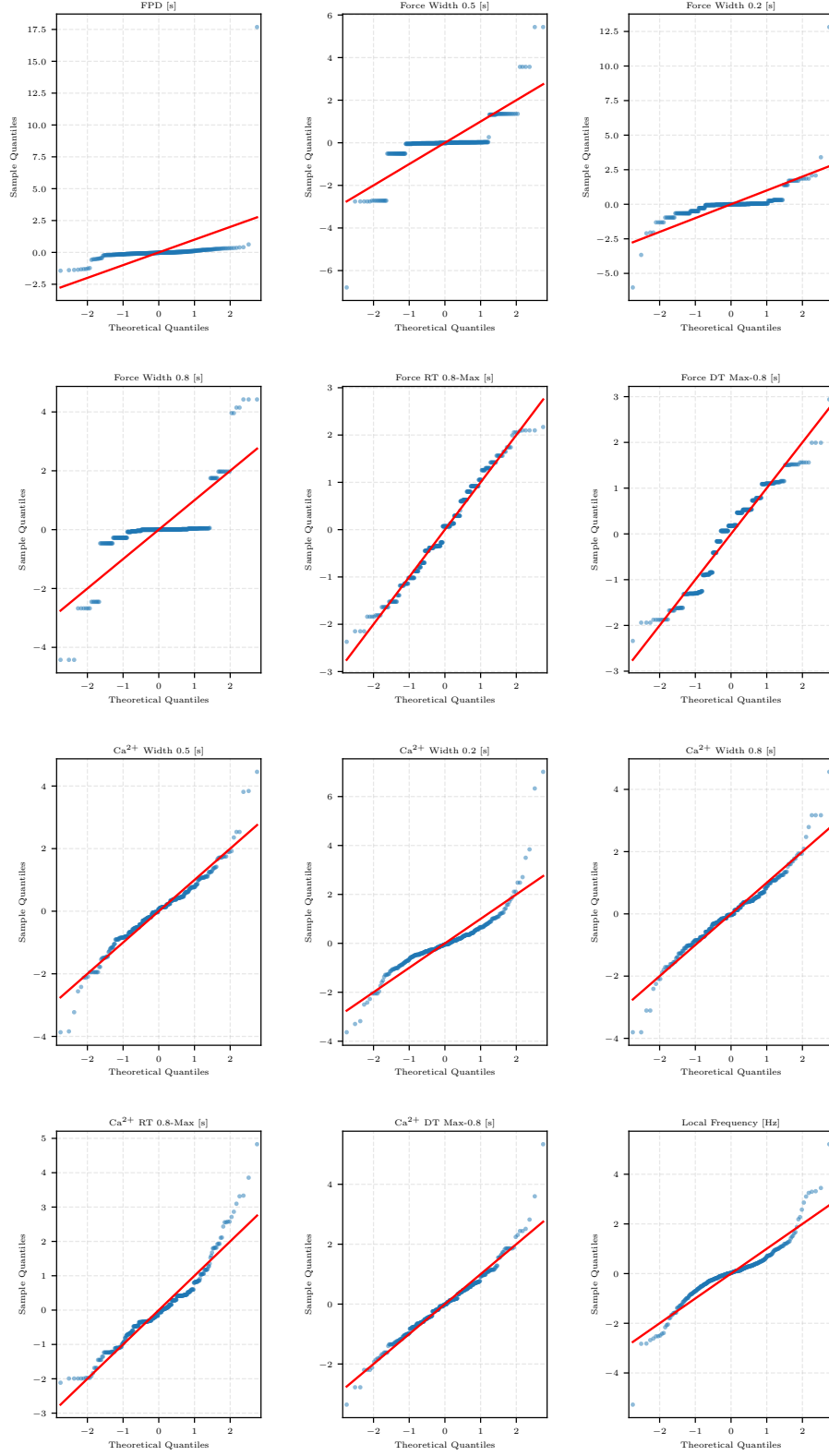
Figure A.5: **QQ Plot of Residuals for Mixed Model Fit of Nifedipine Treatment (Remaining Features)**
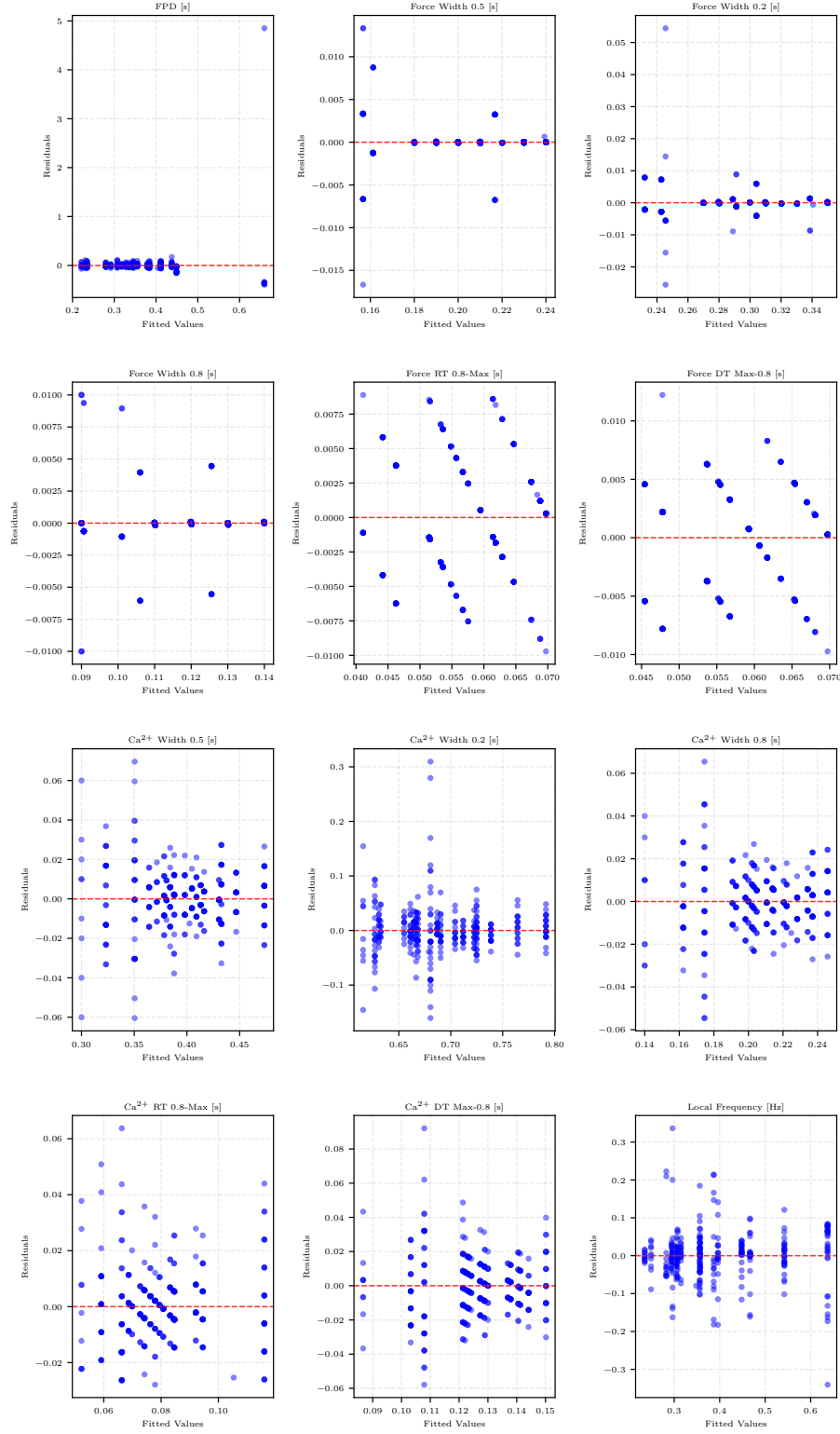
Figure A.6: **Residuals vs. Fitted Values for Mixed Model Fit of Nifedipine Treatment (Remaining Features).**
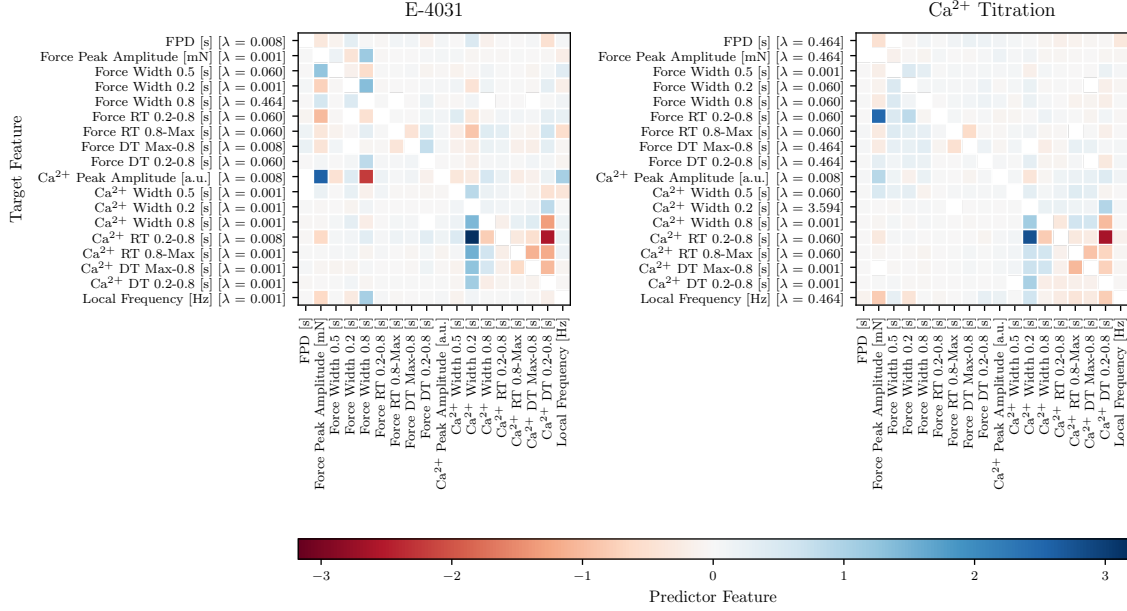
Figure A.7: **Feature Associations Identified via GLMM Regression modeling with LASSO Regularization (Section 5.3) for E-4031 and Calcium Titration Experiment.**

A.2 Feature Associations via LASSO based Mixed Modeling

This section contains the LASSO based Mixed modeling regressions plots (Section 5.5) for E-4031 and Calcium Titration experiment. (Figure A.7)

A.3 Code and Implementation Notes

The complete codebase for the analysis is available at `https://github.com/myotwin/MT-data-mining`. The analysis was primarily conducted using Python, while statistical analysis and post-hoc comparisons were performed in R.

A brief overview of the tools and libraries used:

- **Signal Processing and Feature Extraction:** Scipy Signal Processing Module and Py-Wavelets were used for FFT, DWT, and peak detection.

- **Feature Space Outlook:** t-SNE plots were generated using Scikit-learn.

- **Statistical Analysis:** Mixed modeling and LASSO regression were conducted in R using lmerTest, multcomp, and glmmLasso.

- **Data Handling:** Pandas was used for data transformation and preprocessing.

- **Visualization:** Matplotlib and Seaborn were employed for plotting, with annotations refined in Inkscape.

A.4   Feature Mapping Dictionary for Correlation Trend (5.14)

ID	Feature Pair
1	Ca^{2+} Width 0.2 [s] vs Ca^{2+} Width 0.5 [s]
2	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} Width 0.2 [s]
3	Force Width 0.2 [s] vs Force Width 0.5 [s]
4	Ca^{2+} Width 0.5 [s] vs Force DT 0.2-0.8 [s]
5	Force Width 0.5 [s] vs Force Width 0.8 [s]
6	Ca^{2+} Width 0.5 [s] vs Force Width 0.2 [s]
7	Ca^{2+} Width 0.2 [s] vs Force Width 0.2 [s]
8	Force Width 0.2 [s] vs Force Width 0.8 [s]
9	Ca^{2+} Width 0.2 [s] vs Force DT 0.2-0.8 [s]
10	Ca^{2+} Width 0.5 [s] vs Force Width 0.5 [s]
11	Ca^{2+} Width 0.2 [s] vs Force Width 0.5 [s]
12	Ca^{2+} Width 0.5 [s] vs Ca^{2+} Width 0.8 [s]
13	Force DT 0.2-0.8 [s] vs Force Width 0.5 [s]
14	Force DT 0.2-0.8 [s] vs Force Width 0.2 [s]
15	Ca^{2+} Width 0.5 [s] vs Force Width 0.8 [s]
16	Force Peak Amplitude [mN] vs Force Width 0.8 [s]
17	Ca^{2+} Width 0.2 [s] vs Force Width 0.8 [s]
18	Ca^{2+} DT 0.2-0.8 [s] vs Force Width 0.2 [s]
19	Force DT 0.2-0.8 [s] vs Force Width 0.8 [s]
20	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} Width 0.5 [s]
21	Force Peak Amplitude [mN] vs Force Width 0.2 [s]
22	Ca^{2+} DT 0.2-0.8 [s] vs Force Width 0.5 [s]
23	Ca^{2+} Width 0.2 [s] vs Ca^{2+} Width 0.8 [s]
24	Ca^{2+} Width 0.8 [s] vs Force DT 0.2-0.8 [s]
25	Ca^{2+} DT 0.2-0.8 [s] vs Force DT 0.2-0.8 [s]
26	Ca^{2+} Width 0.8 [s] vs Force Width 0.2 [s]
27	Ca^{2+} Width 0.5 [s] vs Force Peak Amplitude [mN]
28	Force Peak Amplitude [mN] vs Force Width 0.5 [s]
29	Ca^{2+} Width 0.2 [s] vs Force Peak Amplitude [mN]
30	Ca^{2+} Width 0.8 [s] vs Force Width 0.5 [s]
31	Ca^{2+} DT 0.2-0.8 [s] vs Force Width 0.8 [s]
32	Force DT 0.2-0.8 [s] vs Force Peak Amplitude [mN]
33	Ca^{2+} DT Max-0.8 [s] vs Ca^{2+} Width 0.8 [s]
34	Ca^{2+} DT 0.2-0.8 [s] vs Force Peak Amplitude [mN]
35	Ca^{2+} Width 0.8 [s] vs Force Width 0.8 [s]
36	Ca^{2+} Width 0.8 [s] vs Force Peak Amplitude [mN]
37	Force DT Max-0.8 [s] vs Force Width 0.5 [s]
38	Force DT Max-0.8 [s] vs Force Width 0.8 [s]
39	Force DT Max-0.8 [s] vs Force Width 0.2 [s]
40	Force DT 0.2-0.8 [s] vs Force DT Max-0.8 [s]
41	Force DT Max-0.8 [s] vs Force Peak Amplitude [mN]
42	Ca^{2+} Width 0.5 [s] vs Force DT Max-0.8 [s]
43	Ca^{2+} Width 0.2 [s] vs Force DT Max-0.8 [s]
44	Ca^{2+} DT Max-0.8 [s] vs Ca^{2+} Width 0.5 [s]
45	Ca^{2+} DT 0.2-0.8 [s] vs Force DT Max-0.8 [s]
46	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} Width 0.8 [s]

ID	Feature Pair
47	Ca^{2+} DT Max-0.8 [s] vs Ca^{2+} Width 0.2 [s]
48	Ca^{2+} Width 0.8 [s] vs Force DT Max-0.8 [s]
49	Ca^{2+} DT Max-0.8 [s] vs Force Width 0.2 [s]
50	Ca^{2+} DT Max-0.8 [s] vs Force DT 0.2-0.8 [s]
51	Ca^{2+} DT Max-0.8 [s] vs Force Width 0.8 [s]
52	Ca^{2+} DT Max-0.8 [s] vs Force Width 0.5 [s]
53	Ca^{2+} DT Max-0.8 [s] vs Force Peak Amplitude [mN]
54	Ca^{2+} DT Max-0.8 [s] vs Force DT Max-0.8 [s]
55	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} DT Max-0.8 [s]
56	Ca^{2+} DT 0.2-0.8 [s] vs Force RT 0.8-Max [s]
57	Force RT 0.8-Max [s] vs Force Width 0.5 [s]
58	Force RT 0.8-Max [s] vs Force Width 0.2 [s]
59	Ca^{2+} Width 0.2 [s] vs Force RT 0.8-Max [s]
60	Force Peak Amplitude [mN] vs Force RT 0.8-Max [s]
61	Force RT 0.8-Max [s] vs Force Width 0.8 [s]
62	Ca^{2+} Width 0.5 [s] vs Force RT 0.8-Max [s]
63	Force DT 0.2-0.8 [s] vs Force RT 0.8-Max [s]
64	Ca^{2+} RT 0.2-0.8 [s] vs Force RT 0.2-0.8 [s]
65	Ca^{2+} Width 0.8 [s] vs Force RT 0.8-Max [s]
66	Ca^{2+} RT 0.8-Max [s] vs Force Width 0.5 [s]
67	Ca^{2+} RT 0.8-Max [s] vs Force DT 0.2-0.8 [s]
68	Ca^{2+} RT 0.8-Max [s] vs Ca^{2+} Width 0.2 [s]
69	Ca^{2+} RT 0.8-Max [s] vs Ca^{2+} Width 0.5 [s]
70	Ca^{2+} RT 0.8-Max [s] vs Force Width 0.8 [s]
71	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} RT 0.8-Max [s]
72	Ca^{2+} RT 0.8-Max [s] vs Force Width 0.2 [s]
73	Force DT Max-0.8 [s] vs Force RT 0.8-Max [s]
74	Ca^{2+} RT 0.8-Max [s] vs Force Peak Amplitude [mN]
75	Ca^{2+} RT 0.8-Max [s] vs Force RT 0.8-Max [s]
76	Ca^{2+} RT 0.8-Max [s] vs Ca^{2+} Width 0.8 [s]
77	FPD [s] vs Force Peak Amplitude [mN]
78	Ca^{2+} RT 0.8-Max [s] vs Force DT Max-0.8 [s]
79	FPD [s] vs Force Width 0.8 [s]
80	Ca^{2+} DT Max-0.8 [s] vs Force RT 0.8-Max [s]
81	Ca^{2+} RT 0.8-Max [s] vs Local Frequency [Hz]
82	FPD [s] vs Force Width 0.2 [s]
83	Force RT 0.2-0.8 [s] vs Force Width 0.5 [s]
84	Ca^{2+} DT 0.2-0.8 [s] vs FPD [s]
85	FPD [s] vs Force DT Max-0.8 [s]
86	Force DT 0.2-0.8 [s] vs Force RT 0.2-0.8 [s]
87	Ca^{2+} Width 0.2 [s] vs FPD [s]
88	Force RT 0.2-0.8 [s] vs Force Width 0.2 [s]
89	Ca^{2+} Width 0.5 [s] vs FPD [s]
90	FPD [s] vs Force Width 0.5 [s]
91	Ca^{2+} Peak Amplitude [a.u.] vs Force Peak Amplitude [mN]
92	Ca^{2+} RT 0.2-0.8 [s] vs Ca^{2+} RT 0.8-Max [s]
93	Ca^{2+} Peak Amplitude [a.u.] vs Force DT Max-0.8 [s]
94	Ca^{2+} Width 0.8 [s] vs Force RT 0.2-0.8 [s]
95	Ca^{2+} Width 0.5 [s] vs Force RT 0.2-0.8 [s]
96	Ca^{2+} Width 0.2 [s] vs Force RT 0.2-0.8 [s]
97	Ca^{2+} RT 0.2-0.8 [s] vs FPD [s]
98	Force RT 0.2-0.8 [s] vs Force RT 0.8-Max [s]
99	FPD [s] vs Local Frequency [Hz]
100	Ca^{2+} DT Max-0.8 [s] vs FPD [s]
101	Ca^{2+} Peak Amplitude [a.u.] vs Force Width 0.8 [s]
102	FPD [s] vs Force DT 0.2-0.8 [s]
103	Ca^{2+} Width 0.8 [s] vs FPD [s]

ID	Feature Pair
104	Force DT Max-0.8 [s] vs Force RT 0.2-0.8 [s]
105	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} Peak Amplitude [a.u.]
106	Ca^{2+} RT 0.8-Max [s] vs Force RT 0.2-0.8 [s]
107	Force RT 0.2-0.8 [s] vs Force Width 0.8 [s]
108	Ca^{2+} RT 0.2-0.8 [s] vs Force Width 0.5 [s]
109	Ca^{2+} DT Max-0.8 [s] vs Force RT 0.2-0.8 [s]
110	Ca^{2+} Peak Amplitude [a.u.] vs Force DT 0.2-0.8 [s]
111	Ca^{2+} DT 0.2-0.8 [s] vs Force RT 0.2-0.8 [s]
112	Ca^{2+} Peak Amplitude [a.u.] vs Force Width 0.5 [s]
113	Ca^{2+} Peak Amplitude [a.u.] vs Force Width 0.2 [s]
114	Ca^{2+} RT 0.2-0.8 [s] vs Force Width 0.2 [s]
115	FPD [s] vs Force RT 0.8-Max [s]
116	Ca^{2+} Peak Amplitude [a.u.] vs Ca^{2+} Width 0.5 [s]
117	Ca^{2+} Peak Amplitude [a.u.] vs Ca^{2+} Width 0.2 [s]
118	Ca^{2+} RT 0.2-0.8 [s] vs Ca^{2+} Width 0.2 [s]
119	Ca^{2+} RT 0.2-0.8 [s] vs Force DT 0.2-0.8 [s]
120	Ca^{2+} RT 0.8-Max [s] vs FPD [s]
121	Ca^{2+} RT 0.2-0.8 [s] vs Ca^{2+} Width 0.5 [s]
122	Ca^{2+} DT Max-0.8 [s] vs Ca^{2+} Peak Amplitude [a.u.]
123	Ca^{2+} Peak Amplitude [a.u.] vs Ca^{2+} Width 0.8 [s]
124	Ca^{2+} RT 0.2-0.8 [s] vs Force Width 0.8 [s]
125	Ca^{2+} RT 0.2-0.8 [s] vs Local Frequency [Hz]
126	Ca^{2+} RT 0.2-0.8 [s] vs Force RT 0.8-Max [s]
127	Ca^{2+} RT 0.2-0.8 [s] vs Ca^{2+} Width 0.8 [s]
128	Ca^{2+} DT 0.2-0.8 [s] vs Ca^{2+} RT 0.2-0.8 [s]
129	Ca^{2+} RT 0.2-0.8 [s] vs Force DT Max-0.8 [s]
130	Force Peak Amplitude [mN] vs Local Frequency [Hz]
131	Ca^{2+} DT Max-0.8 [s] vs Ca^{2+} RT 0.2-0.8 [s]
132	Force Peak Amplitude [mN] vs Force RT 0.2-0.8 [s]
133	Force RT 0.8-Max [s] vs Local Frequency [Hz]
134	Ca^{2+} DT Max-0.8 [s] vs Ca^{2+} RT 0.8-Max [s]
135	Ca^{2+} RT 0.2-0.8 [s] vs Force Peak Amplitude [mN]
136	Ca^{2+} Peak Amplitude [a.u.] vs Force RT 0.8-Max [s]
137	Force Width 0.8 [s] vs Local Frequency [Hz]
138	Ca^{2+} Peak Amplitude [a.u.] vs Ca^{2+} RT 0.8-Max [s]
139	Ca^{2+} Peak Amplitude [a.u.] vs FPD [s]
140	Force Width 0.5 [s] vs Local Frequency [Hz]
141	Ca^{2+} DT 0.2-0.8 [s] vs Local Frequency [Hz]
142	FPD [s] vs Force RT 0.2-0.8 [s]
143	Ca^{2+} Width 0.5 [s] vs Local Frequency [Hz]
144	Ca^{2+} Width 0.2 [s] vs Local Frequency [Hz]
145	Force Width 0.2 [s] vs Local Frequency [Hz]
146	Ca^{2+} Width 0.8 [s] vs Local Frequency [Hz]
147	Force DT 0.2-0.8 [s] vs Local Frequency [Hz]
148	Ca^{2+} DT Max-0.8 [s] vs Local Frequency [Hz]
149	Force DT Max-0.8 [s] vs Local Frequency [Hz]
150	Ca^{2+} Peak Amplitude [a.u.] vs Local Frequency [Hz]
151	Force RT 0.2-0.8 [s] vs Local Frequency [Hz]
152	Ca^{2+} Peak Amplitude [a.u.] vs Force RT 0.2-0.8 [s]
153	Ca^{2+} Peak Amplitude [a.u.] vs Ca^{2+} RT 0.2-0.8 [s]