Fall 2022

Modelling Muscle Activation Using EMG Signal

Mercy U. Okonna

Fall 2022

Modelling Muscle Activation Using EMG Signal

Mercy U. Okonna

MODELLING MUSCLE ACTIVATION USING EMG SIGNAL

by

MERCY UZUNMA OKONNA

B.S., Federal University of Technology, Owerri, Nigeria, 2017

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

MODELLING MUSCLE ACTIVATION USING EMG SIGNAL

by

MERCY UZUNMA OKONNA

(Under the Direction of Stephen Carden)

ABSTRACT

Electromyography (EMG) is a method for measuring muscle activity by an electrical signal, and is useful in studying motor control, postural control, and in physical therapy. A current research topic is creating an algorithm that can use the EMG signal to reliably classify a muscle as active or inactive. This thesis presents a classification algorithm for leg muscles with a single activation spike while walking. Time is rescaled into steps, which are identified using data from cameras measuring joint angles while walking. The algorithm is based on moving averages and a convex combination of mean signal strength in active and inactive regions. The algorithm is tested on EMG signals from the gastrocnemius, vastus lateralis, and tibialis muscles, using angle measurements from hip, knee, and ankle joints.

INDEX WORDS: EMG, muscle, gait, joint

2009 Mathematics Subject Classification: 62H30, 68H10

MODELLING MUSCLE ACTIVATION USING EMG SIGNAL

by

MERCY UZUNMA OKONNA

Major Professor:	Stephen Carden
Committee:	Li Li
	Ionut Iacob

Electronic Version Approved:
December 2022

ACKNOWLEDGMENTS

I would love to express my sincere gratitude to my supervisor Dr. Stephen Carden for his unallayed support and guidance throughout the project phase. Also want to thank the defense committee, who generously provided knowledge and expertise.

I am also grateful to my friends and colleagues for their moral support in keeping up with the work load during my research.

Lastly, I would be remiss in not mentioning my family. Their belief in me has kept my spirits and motivation high during this process.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 The Basics of Electromyography

During neuromuscular activities, electromyography (EMG) is used to represent the biomedical signal that represents the electrical currents generated by the muscle cells during their contraction, which is primarily used in neuromuscular activities. This device is a useful tool when it comes to measuring the electrical output of the skeletal muscles during physical activities. It can provide access to a wide range of physiological information which explains how muscles generate force in order to produce movement through the generation of force. It is the nervous system that controls the activity of muscles, as well as their contractions and relaxations[7]. As a result, this implies that the EMG signal is indirectly controlled by the nervous system. Biomedical engineers are increasingly relying on EMG signals for a variety of purposes. This can help in distinguishing between myopathic (a condition that occurs when the muscle fibers do not function properly) and neurogenic (an illness caused by abnormal blood circulation in the body) muscle wasting or weakness. This system has the capability of detecting abnormalities such as chronic denervation, which is a condition where there is an interruption of the nerve connection to an organ, or fasciculations in muscles that seem to be clinically normal (fasciculations are involuntary movements of the muscles). The diagnosis of peripheral neuropathy can be achieved by determining the location of neurogenic abnormalities, by distinguishing focal nerve, plexus, or radicular pathology, in addition to providing supportive evidence of the pathophysiology and treatment of the disease. Among these conditions is demyelination, which occurs when myelin tissue is damaged or worn away. At any place on the surface of the skin, EMG detectors can collect signals from several motor units at the same time. In this way, different motor units can interact with each other at the same time.

EMG uses tiny devices called electrodes to translate these signals into graphs, sounds, or numerical values that can be understood by humans. There are several reasons why EMG signal analysis has gained popularity in recent years, including clinical diagnosis and biomedical applications. An effective application area for this technology is in the area of managing and rehabilitating people who have motor disabilities. In order to make a reliable diagnosis of neuromuscular disorders, it is necessary to know the shapes of the Motor Unit Action Potentials (MUAPs) in EMG signals as well as their firing rates. In addition to muscle disorders and nerve disorders, EMG can be used to diagnose problems dealing with the connection between muscle and nerve.

In addition to the use of EMG for medical purposes, it has many other advantages as well. EMG is employed clinically for the diagnosis of neurological and neuromuscular problems. By analyzing the activation timing for a muscle, we are able to determine when the excitation of the muscle begins and ends. We can also determine the duration of the activation. It is used to estimate the force produced by the muscle. Through analyzing the spectrum of the signal, it is possible to obtain an index of the rate at which muscle fatigue occurs based on how fast the signal is changing. In an EMG signal, the frequency ranges between 0 and 500Hz. In terms of the energy available to be used by EMG signals, 50-150Hz is thought to be the most useful range. It is used as a diagnostic tool by gait laboratories and clinicians who have been trained in the use of biofeedback and ergonomic assessment in order to improve their gait efficiency. Besides being used in the medical field, EMG has also been used in a number of other fields such as biomechanics, motor control, neuromuscular physiology, movement disorders, postural control, and physical therapy.

1.2 The History of EMG

Francesco Redi performed an experiment in 1666 which is regarded as the discovery of EMG. He discovered that the highly specialized muscle of the electric ray fish generates electricity[10]. By 1773, Walsh demonstrated that an eel fish's muscle tissue could generate a spark of electricity. In 1792, a publication entitled "De Viribus Electricitatis in Motu Musculari Commentarius" appeared, written by A. Galvani, where the author showed that electricity could initiate muscle contractions. Six decades later, in 1849, Dubios-Raymond discovered that it was also possible to record electrical activity during a voluntary muscle contraction. The first recording of this activity was made by Marey in 1890, who also introduced the term electromyography[8]. In 1922, Gasser and Erlanger used an oscilloscope to show the electrical signals from muscles. Because of the stochastic nature of the myoelectric signal, only rough information could be obtained from its observation. The capability of detecting electromyographic signals improved steadily from the 1930s through the 1950s and researchers began to use improved electrodes more widely for the study of muscles.

In 1960s, the clinical use of surface EMG for the treatment of more specific disorders began, Cram and Steger introduced a clinical method for scanning a variety of muscles using an EMG sensing device[3]. It is not until the middle of the 1980s that integration techniques in electrodes had sufficiently advanced to allow batch production of the required small and lightweights instrumentation and amplifiers. At present, a number of suitable amplifiers are commercially available. In the early 1980s, cables became available which produce artifacts in the desired microvolt range. During the past 15 years, research has resulted in a better understanding of the properties of surface EMG recording. In recent years, surface electromyography is increasingly used for recording superficial muscles in clinical protocols, where intramuscular electrodes are used for deep muscle only. Two types of electrodes have been used to acquire muscle signal, they are the invasive electrode

and non-invasive electrode.

The discoveries of Borelli, Galvani, Newton, Descartes, Marey, Carlet, the Weber brothers, Scherb, Duchenne, Muybridge and Braune, and Fischer provided a solid scientific foundation for the current understanding of human walking[9].

1.3 Gait Analysis

In the medical dictionary, gait analysis is described as an evaluation of the style or manner in which an individual walks, which is normally done by observing the individual walking in a straight line naturally[11]. A forward step normally consists of two phases: the stance phase, during which the foot and leg have to bear most or all of the body weight in order to move forward, and the swing phase, during which the foot is not touching the walking surface and the body weight must be borne by the other leg and foot. Two-step cycles last for about 25 percent of the time, which is the amount of time during which both feet are in contact with the floor at the same time. This part of the cycle, we refer to it as the double-support phase of the cycle.

An analysis of each component of the three phases of ambulation is an essential part of the diagnosis of various neurologic disorders and an important part of the assessment of a patient's progress during rehabilitation and recovery from the effects of neurologic disease, musculoskeletal injuries or disease processes, or amputation of a lower limb.

Around the world, lower limb amputations are one of the leading causes of disability and death. Controlling locomotion modes is dependent on the accurate classification of locomotion modes and the transitions between these modes. As a result, an efficient representation of the signal source is required in order to develop efficient pattern recognition (PR) systems[6]. Recent advances in signal processing and pattern recognition algorithms have led to EMG becoming a source of neural information that can be used for motion

classification and device control as a result of recent advances in signal processing. As a consequence, pattern recognition systems rely on the repeatability of the feature sets of EMG signals corresponding to different modes of locomotion to be analyzed. Due to this, it is possible to select a unique set of features that provide robust locomotion classification based on the available features. Despite their relative simplicity in computation and implementation, time-domain features such as auto-regressive coefficients (AR) have been used more often than frequency-domain features in the control of lower limb prosthetics due to their relative simplicity in computation and implementation in real-time. In order to achieve high classification accuracy, however, limited signal representation in the time domain requires a high number of features, leading to high dimensionality.

Despite the fact that we are able to observe movements visually, we are not able to measure them by visual observation alone. Active movements are produced by muscles, which are the engines that power them. It is therefore reasonable to conclude that an understanding of the forces causing or contributing to movements requires the use of kinesiological electromyography (KEMG)[5]. We used KEMG to gain insights into normal and pathological gait in clinical gait analysis. We can define KEMG as a technique that can be used to determine the relationship between muscle activation signals (EMG) and joint movement as well as the gait cycle in a clinical setting. It was in the early years of KEMG that most of the commercially available electromyographs had too few channels to monitor multiple muscles simultaneously. Due to the fact that subjects were usually examined while seated or lying down, only minimal attention had been given to removing or reducing the extraneous noises caused by the cables.

When it comes to the analysis of clinical gaits, kinesiological EMG has proven to be beneficial. After the adoption of clinical gait analysis into the treatment decision-making process for cerebral palsy, the entire structure of treatment decisions has changed radically. In addition to the kinematics and kinetics of many of the changes that have taken place,

EMG has also been incorporated into all of the major developments that have taken place. A clearer understanding of the functional link between the knee and the ankle is one example of one of the benefits. As a result of the use of EMG in conjunction with movement measurements, it was possible to shed light on the mechanism of how the ankle plantar flexors are linked to the extension of the knee when a person is walking normally. The eccentric action of the plantar flexors at the beginning of the motion restrains the upward rotation of the tibia on the talus, which facilitates the passage of the body-weight force vector in front of the knee as a result. During terminal stance, the eccentric action is followed by a concentric action, which contributes to knee stability as well as energy conservation by keeping the body's center of mass from oscillating vertically.

The introduction of a new operation for patients with cerebral palsy has been one of the most significant contributions of EMG in the field. As a result of the gait documentation that accompanied and followed the introduction of this operation, it has been widely adopted and is in common use around the world today[1]. Among other things, the ability to plan better for surgical correction of the varus foot in cerebral palsy would be of significant assistance. For the purpose of transfer, EMG studies were conducted on the tibialis anterior and tibialis posterior muscles. Considering the first example, that of an experimental study of the action of the ankle plantar flexors, this study is of substantial value in broadening our understanding of the normal gait mechanism; the other two are prime examples of how EMG can be applied in clinical practice. As well as improving the management of patients with neuromuscular disabilities, EMG has contributed to the development of many other areas in which it has contributed to improvement in medical practice.

There is a common practice of validating computational activation methods in comparison to visually detecting surface EMG signals. Besides being time-consuming, visual detection has inherent variability that comes from both natural human error as well as the differences between researchers when it comes to making observations. By using simu-

lated or modeled surface EMG data, it is also possible to create a known EMG activation; however, we should keep in mind that models are not designed to be exactly the same as reality, but rather a tool used to understand reality[12].

The purpose of the present study is to use experimental surface EMG data in order to develop statistical algorithms for the systematic determination of the onset or activation of EMG activity. The development of an algorithm or set of algorithms that are consistently reliable for EMG activation determination should enable more objective EMG activation detection across studies, providing a standard analytical method for clinical research studies.

CHAPTER 2

DATA

2.1 Data Collection

Data for this study were collected at the kinesiology laboratory at Georgia Southern University on October 25th, 2021. A team of laboratory personnel and student assistants applied electrodes to kinesiology researcher Dr. Li Li, a faculty member of the school mentioned above. The laboratory was equipped with equipment and instruments such as a treadmill where locomotion takes place, cameras positioned strategically to capture joint movements, electrodes, a large screen to display the movement, and computers for displaying and converting electromyograms. The laboratory was set up before Dr. Stephen Carden and I were present during data collection. Three muscles and three joints were measured.

2.1.1 Surface Electromyography

Surface EMG was used in the laboratory for the purpose of this study. As measured by electrodes placed on the skin overlying the muscle, it represents the sum of the electrical contributions made by the active motor units (MUs) that are detected by electrodes placed on the skin. In this method, you do not have to be a medical doctor to conduct the test. It is a non-invasive method, which can be carried out by anyone other than a medical doctor. It is possible to extract the timing of the onset and offset of the activation of muscular activity using surface electromyography (EMG). In clinics, this approach has been widely adopted to quantitatively characterize the various neuromotor disorders, monitor patients' progress, and improve rehabilitation strategies as well.

Surface electrodes can be divided into two different types based on their uses. Dry electrodes are in direct contact with the skin and gelled electrodes are used as a chemical

interface between the skin and the metallic part of the electrode through the use of an electrolytic gel.

In most applications, dry electrodes are used when the geometry or size of electrodes makes it impossible for gel to be used. Dry electrodes include bar electrodes and array electrodes, which are examples of dry electrodes. When it comes to dry electrodes, it is common for circuitry to be attached to the electrodes in order to serve as a pre-amplifier[4]. As a result, the weight of dry electrodes (typically over 20g) is significantly higher than the weight of gelled electrodes (typically under 1g). As a result of this increased inertial mass, electrode fixation can be more difficult than it would be with gelled electrodes due to the increased difficulty in maintaining fixation. In this research, gelled electrodes were used as electrodes.

It is important to note that the amplitude, time, and frequency domain properties of a surface electromyography signal are influenced by a variety of factors, including the timing and intensity of muscle contraction, the distance of the electrode from the active muscle area, the properties of the surrounding tissues (e.g. the thickness of skin and the amount of adipose tissue surrounding it), the electrode and amplifier properties, as well as the quality of the contact between the electrode and the skin.

As part of the application of surface EMG electrodes, the skin surface was shaved so that no hair would distort the EMG signal or impede its collection. In order to obtain a reliable quality EMG signal, it was necessary to apply an alcohol rub to the surface to ensure that any wetness or sweat had been eliminated. This helped to achieve a clean surface. There is a direct current voltage potential generated at the electrode-skin interface mainly due to the large increase in impedance that occurs in the outermost layer of the skin, which consists of dead skin material and oil secretions[2].

EMG electrodes were placed in a number of different positions where muscle signals were to be recorded, according to the design. The recording was taken from the gastrocne-

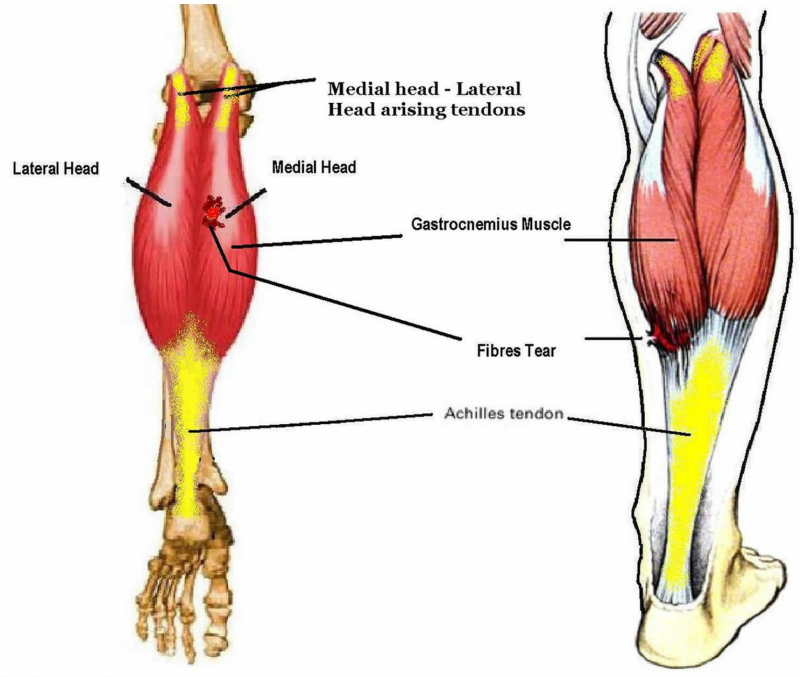

Figure 2.1: Location of the gastrocnemius muscle.

mius muscle, vastus lateralis muscle, and tibialis anterior muscle of the tibia.

The gastrocnemius muscle is one of the two major muscles in the lower leg that is located on the back part of the calf. This particular muscle is responsible for the normal walking and running actions of the lower leg as it is one of the most prominent components of the lower leg. Besides the gastrocnemius muscle, the soleus muscle is another major calf muscle, which is a flat muscle that lies beneath the gastrocnemius muscle. As both gastrocnemius and soleus run the length of the lower leg, they are connected behind the knee and at the heel and cover the entire length of the lower leg. A diagram showing where the gastrocnemius is located is shown in Figure 2.1.

A large, powerful muscle located on the outside of the thigh, the vastus lateralis, stretches from the pelvis all the way up to the kneecap. It is located on the lateral part, or outside, of the thigh. As one of the quadriceps muscles, this is one of the four muscles that are responsible for straightening a knee joint. A large and predominant part of it can be

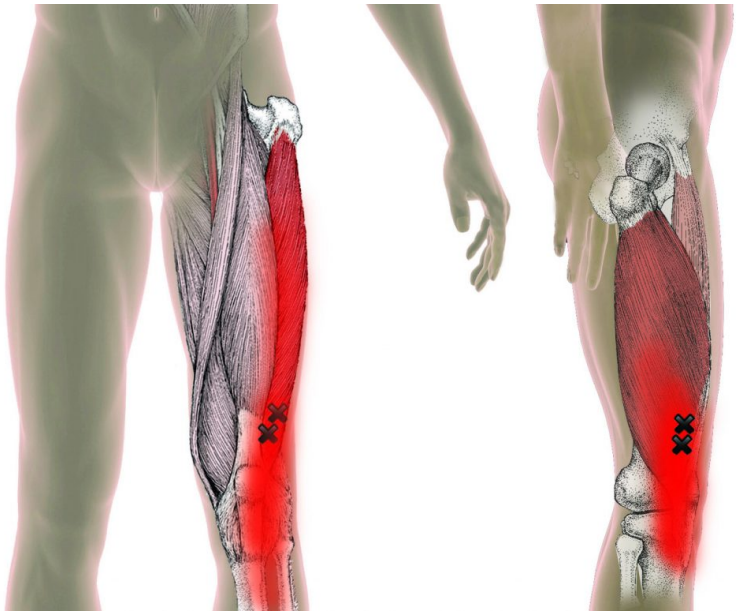

Figure 2.2: Location of the vastus lateralis muscle.

found in the middle of the femur. It is the largest muscle in the body. In conjunction with the other quad muscles, the vastus laterails help to extend the knee joint, as well as play a role in the extension of the hip joint. Additionally, it is also responsible for maintaining the position of the kneecap and the thigh while walking or running. Figure 2.2 illustrates the location where the vastus lateralis muscle is located.

the tibialis anterior muscle is located at the end of the front part of the shin bone of the lower leg. The muscle is found just below the knee, extends along the front of the shin, then attaches itself to the top of the foot in the middle of the foot. As when a foot is tapped, the anterior tibialis muscle in the lower leg serves to help you raise the foot and ankle off the ground by flexing them. The muscle in the foot also helps to pull the foot inwards, a motion that is known as the inversion motion. Since it is attached to the top of the foot, the anterior tibial muscle also plays a crucial role in raising the arch of the foot. It is shown in 2.3 where the tibialis muscle can be found.

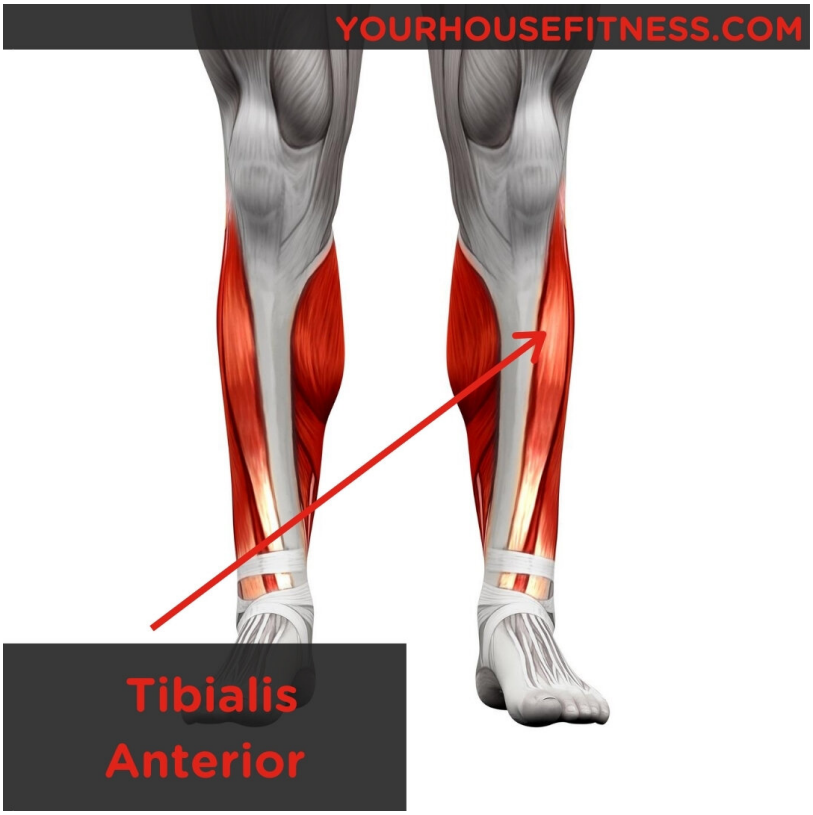

Figure 2.3: Location of the tibialis anterior muscle.

2.1.2 Cameras and Joints

The laboratory was furnished with the equipment that will help in collecting data for the study. This equipment consists of cameras to capture the movement, a treadmill to walk on, and electrodes to capture the EMG signals that occur during locomotion. As the locomotion was taking place, the cameras were capturing joint angles and projecting to a screen to see the pattern and the signals. The joints that were captured are the hip, ankle, and knee.

One of the most prominent joints in the human body is the hip joint. As a result, it allows us to move around freely, run, jump, and walk. The hip and leg bear the weight of our body along with the force exerted by the strong muscles in the hip and leg. Despite this, the hip joint is also one of the most flexible joints in the body, having a broader range of motion than any other joint, except for the shoulder, in the body. Hip joints are ball-and-socket synovial joints that form between the os coxa (hip bone) and the femur (thigh bone). As the socket for the hip joint is formed by the acetabulum, a round, cup-shaped structure on the os coxa, the acetabulum forms the socket for the joint. The hip joint must be able to support half of the body's weight along with any other forces acting upon the body at the same time. The force exerted on the hip joint as a result of running and jumping, for example, is multiplied many times by the force exerted by the body's weight. This is a result of the force applied to it by the body's movements. There is nothing more imperative to the hip joint than being able to accommodate these extreme forces repeatedly during intense physical activities.

The knee joint is the largest joint in the human body and one of the most complex joints in the entire body. Walking, running, sitting, and standing are all activities that require movement of the knee joint in order to function effectively. Movement is controlled by various muscles, ligaments are responsible for stability, special cartilage is responsible for absorbing pressure during movement, as well as various other structures that ensure

pain-free, smooth movement. As the knee bones work together to support and transfer forces between the hip and the foot, the leg is able to move easily and smoothly as the forces are transferred between the hip and foot. Knee joints serve a variety of functions, one of which is to support the body in an upright position without having to exert any effort on the part of the muscles involved. It also helps in lowering and raising the body in a variety of ways, such as sitting, climbing, and squatting. As a result, the leg is able to rotate or twist so that the foot is accurately placed and positioned. During running, the knee acts as a strong forward propeller for the body, working in conjunction with the ankle joint to propel the body forward. In addition to providing stability to the leg, it also provides proprioception to the leg. As a shock absorber, it plays a crucial role in the process.

The ankle joints, also called talocrural joints, are significant joints that play a critical role in ambulation. This is because they adapt to the surface your body is walking on in order to maintain stability. There is an upward and downward movement of the foot that is allowed by the ankle joints. It is a joint that connects the foot to the leg. A variety of different activities can be performed with it such as walking, running, jumping, and many others. In addition to allowing movement, it also contributes to the stability of the lower limbs. Dorsiflexion and plantar flexion are the two different types of movements that can be performed at this joint. There is a motion called dorsiflexion that occurs when the top of the foot moves towards the leg as part of its motion. In contrast to this, plantar flexion occurs when the top of the foot moves away from the leg in a downward motion.

2.2 Data Cleaning

In order to make EMG data suitable for use, the first thing that needs to be done after receiving it is to clean it up. This will enable us to make it suitable for use. It is widely known that data transformation refers to the process of changing the format, structure, or value of data. Data transformation is a process that involves adding, copying, or replicating

data, deleting fields or records, standardizing variable names, and renaming, moving, or combining columns in a database. It also involves joining one set of data with another and saving a file in a different format from what it was originally created in if necessary. After the spreadsheets were created, it was necessary to save them as comma-delimited values (CSV) files, along with many other processes to clean the data.

A .xlsx file with two sets of data, corresponding to joint angles and EMG signals, was transferred to us by the kinesiology researcher after the data collection event. The file contains both sets in one .xlsx file. Due to different time scales, the data was not in the standard rectangular format. In examining the data, the researcher found that the data consisted of eight columns, each of which contains an array of data that is arranged in a certain way. The first four columns describe joint variables with a low frequency and 1500 observations, while the remaining four describe muscle variables with a high frequency and about 65000 observations. In order to make sense of the data, it was split into two separate data frames. Due to the fact that these datasets had different observations, causing them to have different lengths, the separation of these datasets was capable of taking care of the missing values in both datasets. After the datasets were separated, there were no missing values in any of the datasets. Two files were created. The name of them are `kin_data` and `EMG_data`.

The `EMG_data` data frame has a row for each point in time and a column for each muscle. The `kin_data` data frame has a row for each point in time and a column for each joint. The columns have been renamed as follows: `Time-kinematics`, `Hip`, `Knee`, `Ankle`, `Time-EMG`, `VastusLateralis`, `Gastrocnemius`, `Tibialis`.

In order to clean the EMG data, the following steps are followed. It was necessary to import some packages into R in order to help clean the data. These packages or libraries are:

tidyverse: This library has many functions, which are: Group by based on function, generate all combinations using the crossing, coord-flip to display counts more accurately, fct-

reorder to sort for charts creation, extract rows from the first table which are not matched in the second table, extract rows from the first table which are matched in the second table, filter groups without making a new column, str-replace-all to find and replace multiple options at once, case-when to create when conditions are met, select columns with starts-with and ends-with, number parsing, date column creation, randomly shuffle the data, create a new column basis group by, create a new column basis count option.

The magrittr library can be used for a number of things such as basic piping, argument placeholders, reusing placeholders for attributes, building (unary) functions, and piping variable expressions.

There is a library called readxl that is used to read Excel documents.

From a time series plot of the EMG signals, it was observed that some data was darker and denser at less than ten seconds. This can be seen by zooming in on Figure 2.4. It is caused by rounding time to a fixed number of significant digits that results in lower precision when the leading decimal place increases.

From Figure 2.5, the EMG signals at time 1.00, 1.01, 1.02, etc show that multiple observations are condensed into the same instant in time. This is an artifact due to rounding the time value to a lower precision than the frequency at which the measurements were made. Therefore the researcher calculated the hertz and overwrite the time variable in the data.

$$EMG\_hertz = \frac{1}{numberofrow} \tag{2.1}$$

The kinematics data was checked but it was not affected by this problem, therefore the time variable was not overwritten as it was in the EMG data.

Figure 2.6 Shows the graph of the EMG data which consists of the three different muscles, gastrocnemius, tibialis, and vastus lateralis. The graph is noisy and does not have a clean cyclic pattern. Figure 2.7 shows the graph of the kinematics data that consist of the

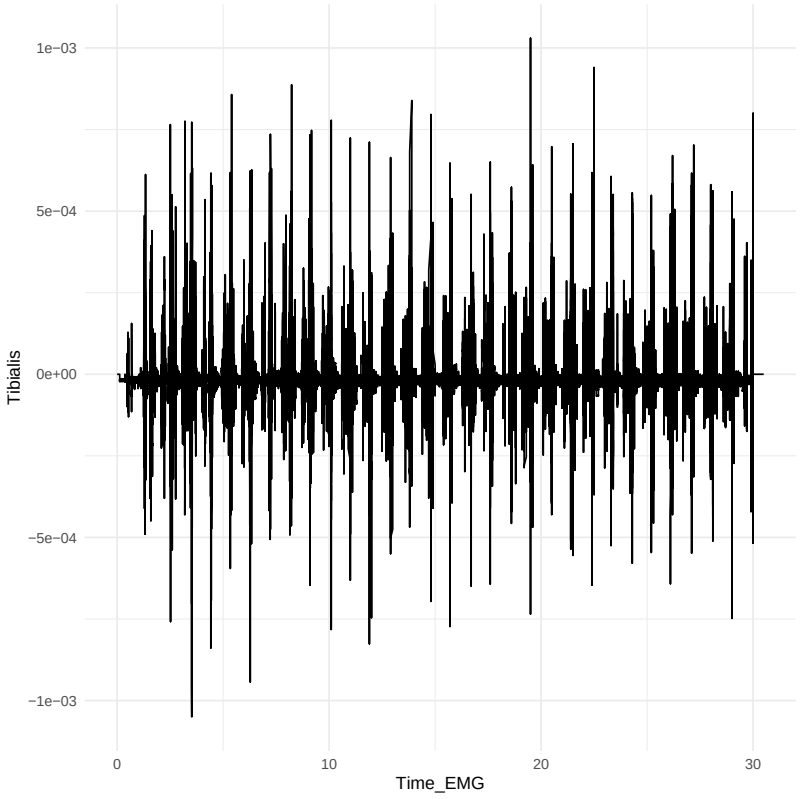

Figure 2.4: This is the EMG time variable

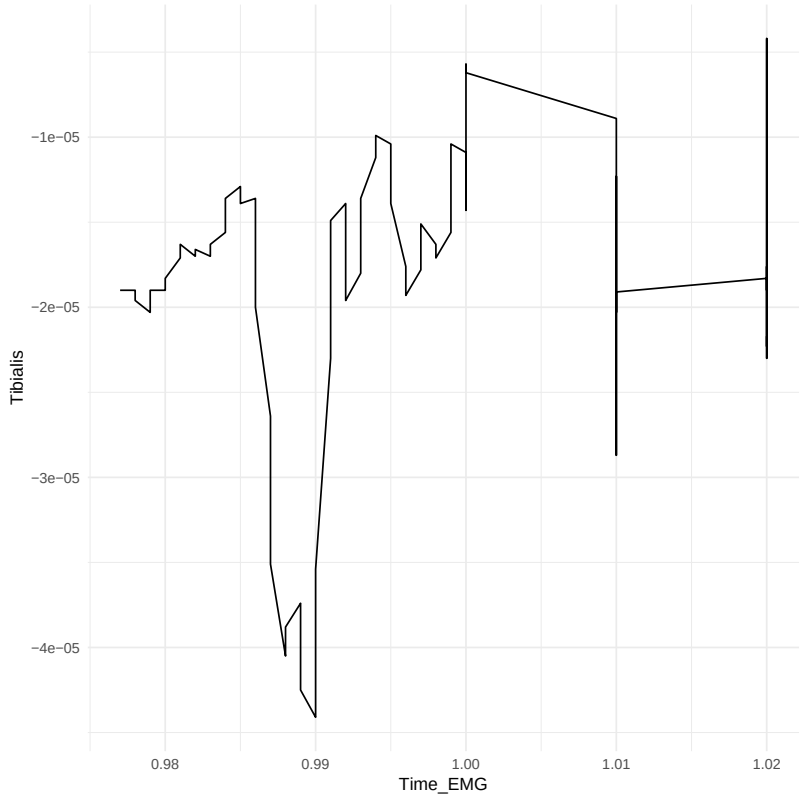

Figure 2.5: This is the zoomed-in time variable. Notice the repetition of signal values at the values after 1.00

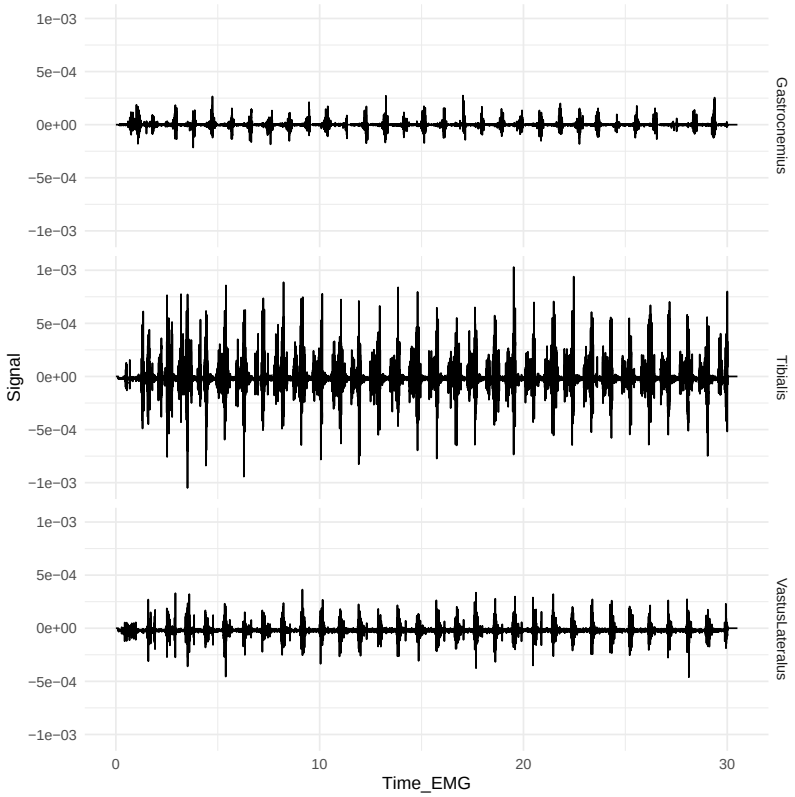

Figure 2.6: Graph showing the EMG data which consist of the muscles signals

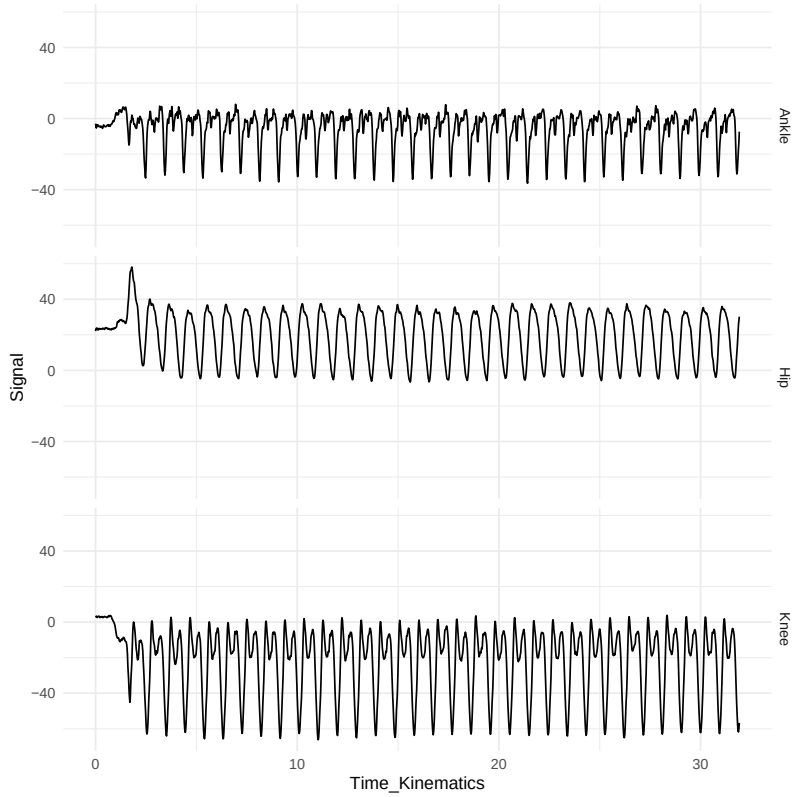

Figure 2.7: Graph of the kinematics data consisting of the joints signals

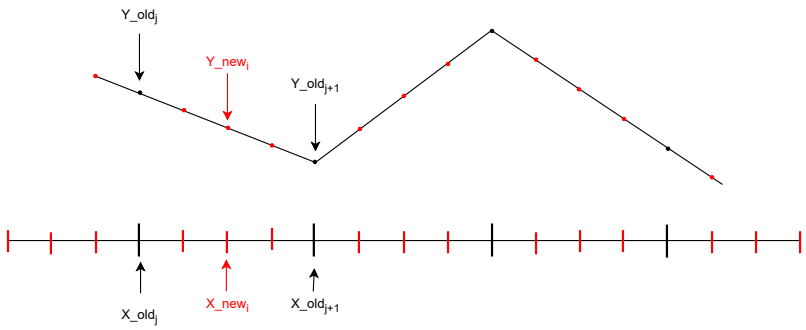

Figure 2.8: Figure of interpolation done on the kinematic data

ankle, hip, and knee joint angles. The graph has a clear cyclic pattern which is good and clean. The researcher interpolated the kinematics data to the same time scale as the EMG data and then combined the two datasets together with a common time variable. Figure 2.8 shows the interpolating figure.

It was observed in Figure 2.9 that before the cyclic pattern begins, there are irregularities as the individual starts to walk. This happened as the gait cycle has not started. The gait cycle describes the cyclic pattern of movement that occurs while walking. A single gait starts when the heel of one foot strikes the ground and ends when that same heel touches the ground again. This appeared to be within two seconds and it was manually removed from the data frame.

Interpolation is a method of deriving a simple function from the given discrete data set such that the function passes through the provided data points. This helps to insert data points in between the given observations. This method is used to compute the value of a function for an intermediate value of the independent variable. In short, interpolation is a process of approximating the unknown values that lie in between the known data points. The type of interpolation that was used is linear interpolation, which has the equation:

$$Y\_new_i = Y\_old_j + (Y\_old_{j+1} - Y\_old_j)\frac{X\_new_i - X\_old_j}{X\_old_{j+1} - X\_old_j}. \tag{2.2}$$

To implement this, we wrote a function called `interpolate()`.

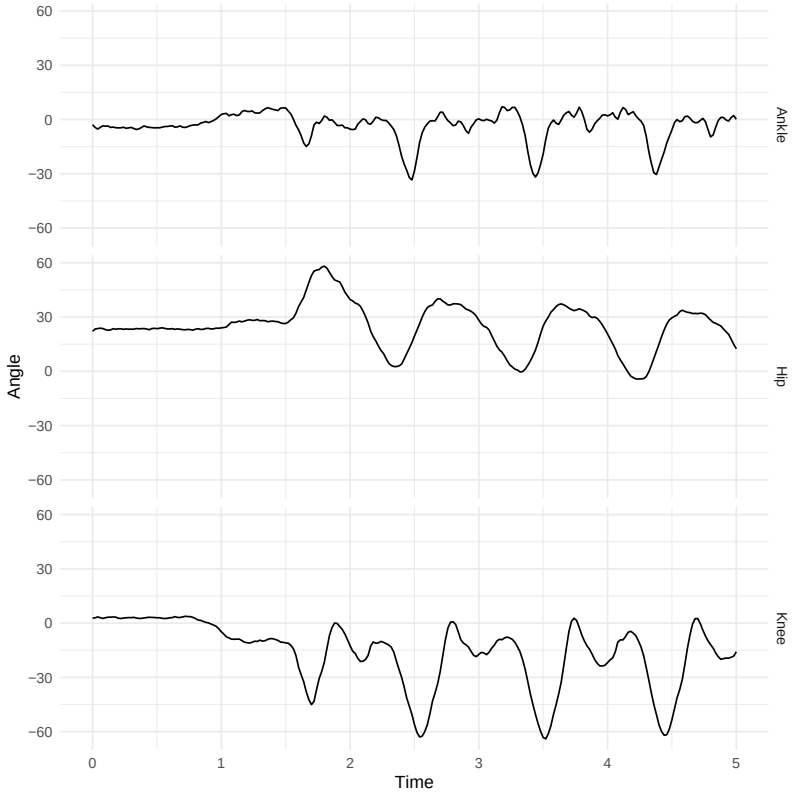

Figure 2.9: Graph showing the irregularities of the kinematics data within the first 2 seconds due to the subject just beginning to walk and not yet exhibiting cyclic behavior.

This is the tabular representation of the data after interpolation:

	Time	VastusLateralis	Gastrocnemius	Tibialis	Hip	Knee	Ankle
1	2.0000	-1.59e-05	-1.46e-05	-1.80e-05	39.7	-12.5	-5.33
2	2.0005	-1.80e-05	-1.96e-05	-1.70e-05	39.7	-12.6	-5.33
3	2.0009	-1.70e-05	-1.51e-05	-1.76e-05	39.7	-12.6	-5.34
4	2.0014	-1.56e-05	-8.56e-06	-1.76e-05	39.6	-12.7	-5.34
5	2.0019	-1.44e-05	-4.36e-06	-1.49e-05	39.6	-12.8	-5.35
6	2.0023	-1.63e-05	1.68e-07	-1.16e-05	39.6	-12.9	-5.35

Table 2.1: Table of the clean data consisting of the EMG signals from the muscles and joints at a given time

2.3 Functions to Transform Data Into Step Format

Some functions were built to transform the data from being indexed by time to being indexed by the individual's steps. These functions are `moving average()`, `get_joint_minimums()`, `interpolate()`, `transform_to_steps()`, `get_average_step()`. These functions are used to analyze the EMG data for the different muscles and joints considered in the limb gait analysis. For this analysis, the name of the function depicts what the function does. Functions are explained earlier.

The `get_joint_minimums()` function takes a threshold and joint as arguments to get the minimum of the `kinematic_data` and returns the time at which the minimum joint angle is obtained for each step. Comparing Figures 2.6 and 2.7, it is clear the joint angle data has a clearer signal than the muscle EMG data. By separating the data into intervals in which the joint angle is below a threshold and finding the global minimum inside each interval, we obtain breakpoints that define the beginning and end of step cycles. In Figure 2.10, to get the minimum, the red dotted line is the threshold for the signal and

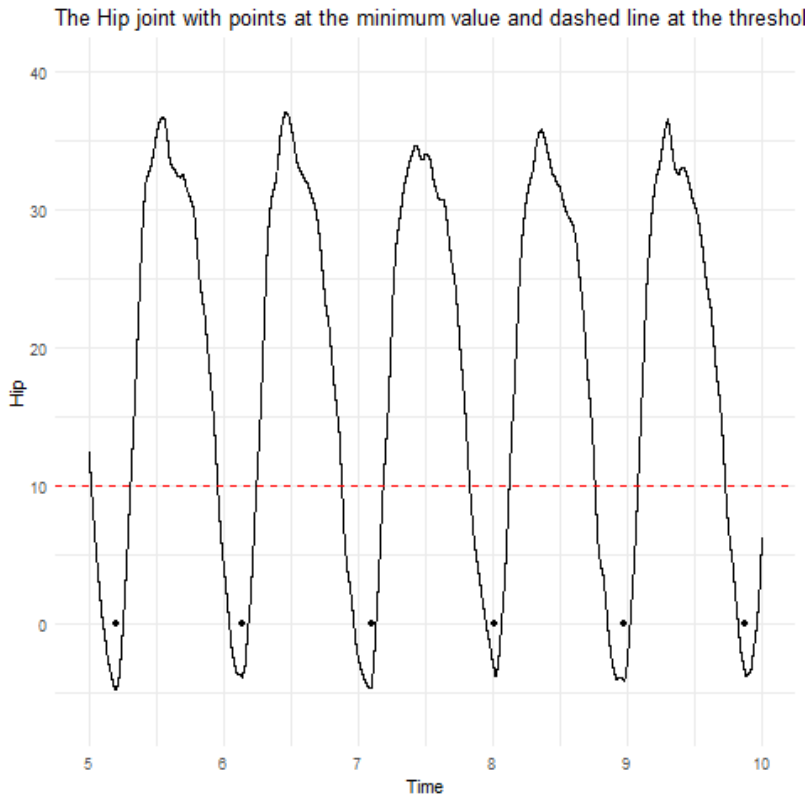

Figure 2.10: An illustration of how the `get_joint_minimums()` function works

the black solid points are the global minimum, therefore the distance between these points is considered a step of the `EMG_data`.

Each of the previous processes are called by the `transform_to_steps()` function which outputs the transfomred data in the desired format. Figure 3.2 shows the graph of the EMG signals transformed into steps.

CHAPTER 3

MODEL

3.1 The Model

Trying to analyze the activation of the muscle using the EMG signal directly will be difficult to work with as it has a symmetry that is centered at zero. For this reason, the researchers model the smoothed absolute EMG signals. The function that was used to smooth and find the average data is `moving_average()`. This function helps to smooth the EMG signal over a specified period by creating a constantly updated local average. The simple moving average was used by the researcher as they calculate the arithmetic mean of the absolute value of a given set of EMG signals over a specific time.

Given the variables $Y_1, Y_2, Y_3, ..., Y_n$ to be the EMG signal observations, the moving average of any given time t within a bandwidth $b > 0$ can be calculated as

$$\bar{Y}_{t,b} = \frac{\sum_{k=t-b}^{t+b} |Y_k|}{(2b+1)}.$$

Figure 3.1 shows an averaged graph of the gastrocnemius signal. The absolute value of the EMG signal is averaged to get a sense of the amplitude of the wave-like feature of the signal. Figure 3.2 shows the graph of the smoothed absolute EMG signal of gastrocnemius muscle with time broken into steps according to the minimum point of the hip joint. The researcher standardized time so that a single step takes place inside the unit interval. For instance, 0.3 is interpreted as being 30% of the way through a step cycle.

From Figure 2.6, notice the gastrocnemius signal has a single significant spike in each cycle. We will focus on muscles that have a single active region in each cycle. The active region has the starting point denoted α and the endpoint denoted β. The graph of the model is shown in Figure 3.3 where the lower blue line shows the inactive region and the upper red line is the active region. The muscle activates at the point α and remains active to the point β within a step. Any particular step will deviate from the ideal step represented

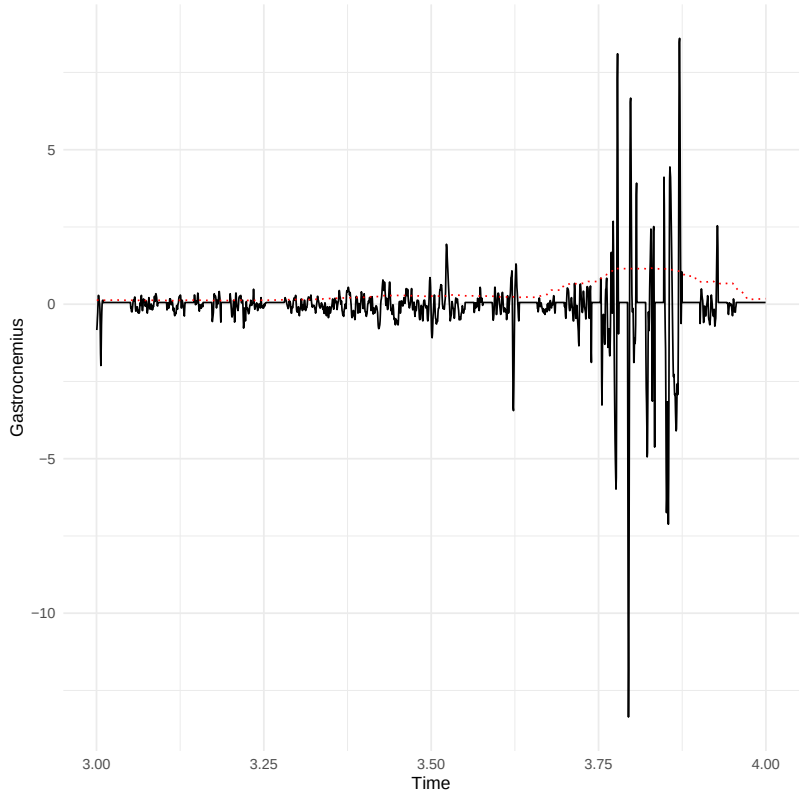

Figure 3.1: The black solid line is the standardized EMG signal for gastrocnemius muscle. The red dotted line is the smoothed absolute value of the standardized EMG signal for the gastrocnemius muscle.

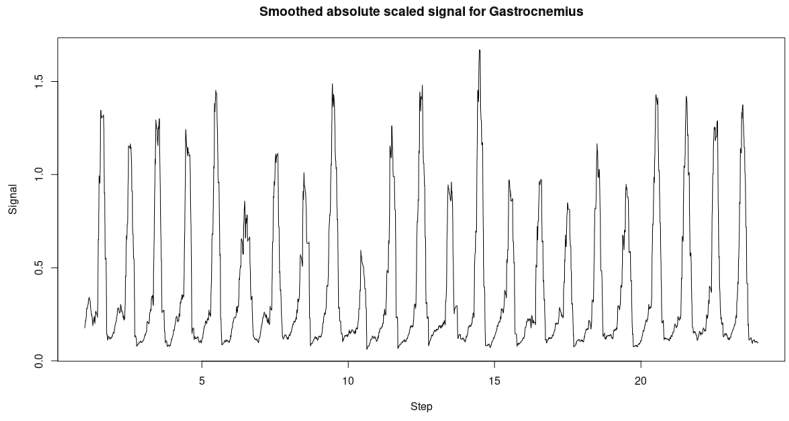

Figure 3.2: This is the smoothed absolute EMG signal of gastrocnemius muscle. Time was broken into steps according to the minimum point of the hip joint.

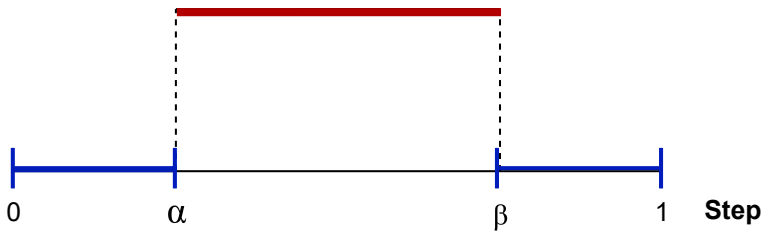

Figure 3.3: This is the graph of the ideal model for each step.

in the model. The muscle's activation period may begin and end before or after the α, β terms, and the strength of the signal between may not be uniform.

Estimating the $\alpha's$ and $\beta's$ to get the beginning and the end of the activation period is not trivial. The activation period during step i can be modeled as having endpoints

$$\alpha_i = \alpha + \epsilon_i, \tag{3.1}$$

$$\beta_i = \beta + \delta_i, \tag{3.2}$$

where ϵ_i, δ_i for $i \in 1, 2, ..., m$ (m is the number of steps) are random variable error terms. The α_i, β_i quantities are not directly observable, but must be estimated from the EMG signal data.

The question remains how to estimate the $\alpha's$ and $\beta's$. We present the following approach based on the heuristic that the signal in the active region should be as different as possible as the signal in the inactive region.

Consider the unit interval [0,1] to be the universal set, A =[α,β] $\subseteq$ [0,1] to be the subset on which the muscle is active, and I = A^c to be the part of the unit interval outside [α, β], that is, where the muscle is inactive. Define $\bar{Y}_A$ to be the average of the set of values $\{Y_i; i \in A\}$. That is, $\bar{Y}_A$ is the average absolute smoothed EMG signal on the active region. Similarly, $\bar{Y}_I$ is the average of $\{Y_i; i \in I\}$, and is the average absolute smoothed EMG signal on the inactive region. Then consider maximizing the quantity

$$f(\alpha, \beta) = \gamma \bar{Y}_A + (1 - \gamma)(-\bar{Y}_I) \tag{3.3}$$

for the range of values of $\gamma \in [0, 1]$.

- If γ is approximately 0, therefore $\bar{Y}_A$ is ignored and it cares most about minimizing $\bar{Y}_I$.
- If γ is approximately 1, therefore $\bar{Y}_I$ is ignored and cares most about maximizing $\bar{Y}_A$.

It is a familiar convex combination, but the value of γ is relatively subjective. The elements available to us are the observation from the data which is gotten from different steps.

The `mean_diff()` function calculates the difference between the active region and the inactive region of the `EMG_data` for a given step and returns the values of equation 3.3. The effect of γ on the resulting active or inactive classification will be visualized later. The `find_better_interval()` function accepts an initial interval and grid size as arguments, and searches within the given grid size (with restriction as shown in Figure 3.4) applying the `mean_diff()` function, then return equation 3.3. The big solid black circle is the initial interval, the red circles are the infeasible region and the black circles are the feasible region because of the criteria $0 <$ `lower_bound` $<$ `upper_bound` < 1.

Figures 3.5, 3.6, 3.7 visualize the effect of γ on the classification of the active and inactive regions. In Figure 3.5 the value of γ is 0.01. It lessens the weight on $\bar{Y}_A$, essentially ignoring it, while $(1 - \gamma)$ will be a large weight on $\bar{Y}_I$ therefore the equation priorities minimizing the inactive region. In Figure 3.6 the value of γ is 0.99. It gives a relatively large weight to $\bar{Y}_A$ therefore the equation prioritizes maximizing the active region of the step, while $(1 - \gamma)$ will be very small essentially ignoring $\bar{Y}_I$. The researcher iterated over the values of $\gamma's$ and based on visual observation, the value of 0.25 was considered to be a reasonable value and was used in the analysis.

The total number of steps that were obtained from the thirty seconds of data collection was twenty-four steps. Showing the graph and classification for all twenty-four steps is not reasonable in this thesis, but here is a link to a document containing all of them. In partic-

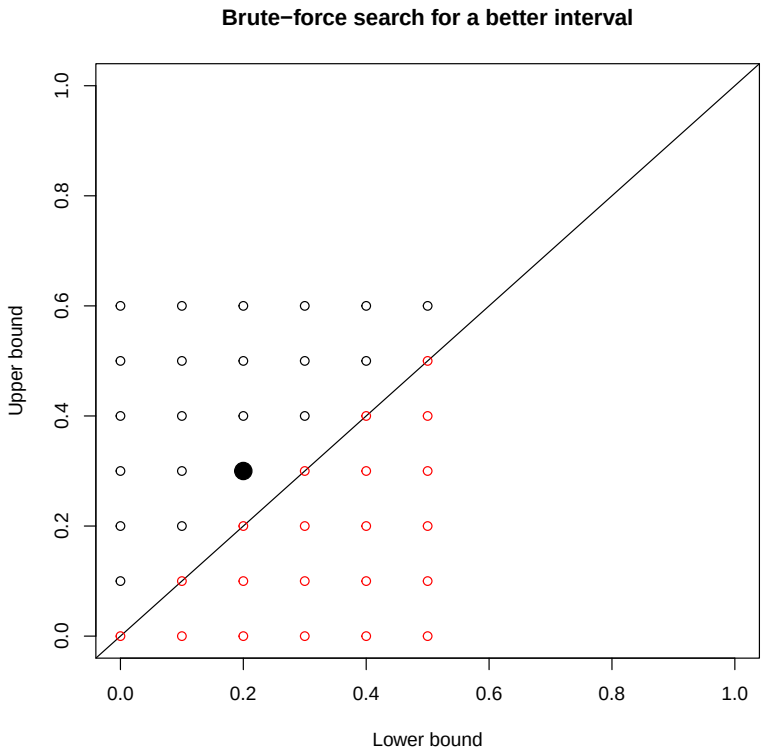

Figure 3.4: The black circles are the feasible regions and red circles are infeasible regions for the `find_better_interval`.

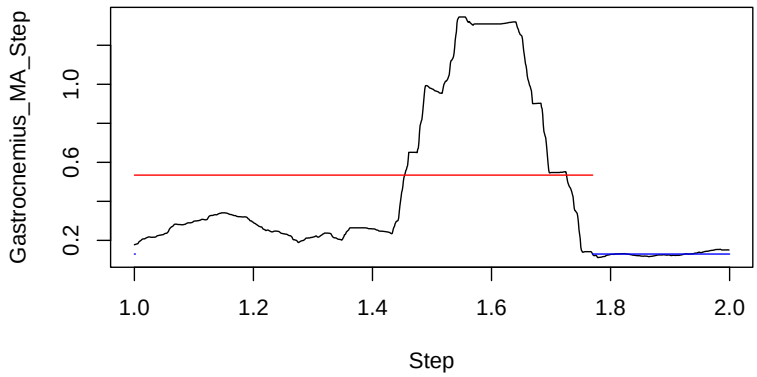

Figure 3.5: This is the graph where γ is 0.01, therefore the active region is made very low.

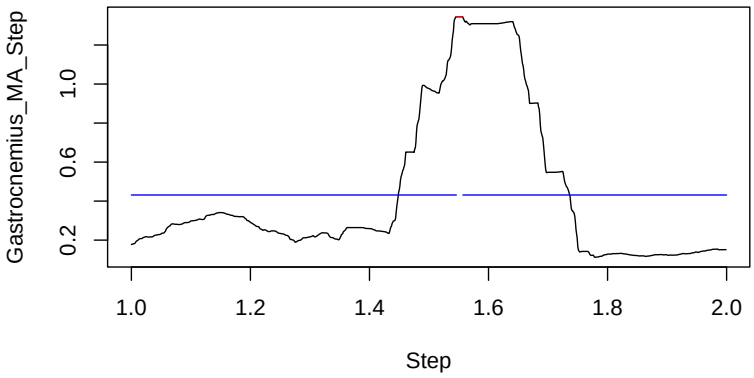

Figure 3.6: This is the graph where γ is 0.99, therefore the active region is very high.

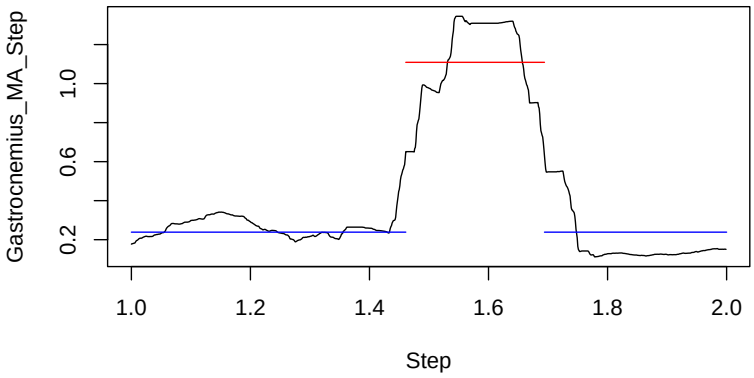

Figure 3.7: This is the graph of the reasonable γ value which is 0.2.

ular, see pages 13 to 38. `https://sites.google.com/a/georgiasouthern.edu/stephen-carden/research/emg-analysis`

The `find_best_interval()` function applies `find_better_interval()` function repeatedly until no better interval can be found. Calculating the average of the activation region of the steps, we have the activation region interval to be in the interval [0.3714783 0.6456087]. Therefore, we can say 37% into the step is where the muscle activates on average, and 65% into the step is where it deactivates on average.

3.2 Error Analysis

We model the ideal step to be active on the interval $[\alpha, \beta]$, and step i will be realized to be active on the interval $[\alpha_i, \beta_i]$ where

$$\alpha_i = \alpha + \epsilon_i,$$

$$\beta_i = \beta + \delta_i,$$

where ϵ_i and δ_i are the residual terms. For an overall estimate of α and β, we average the estimated $\hat{\alpha}$ and $\hat{\beta}$ across steps.

$$\hat{\alpha} = \frac{\sum_{i=1}^{m} \hat{\alpha_i}}{m},$$
$$\hat{\beta} = \frac{\sum_{i=1}^{m} \hat{\beta_i}}{m},$$

where m is the number of steps. Therefore, we have the residuals to be

$$\hat{\epsilon_i} = \hat{\alpha_i} - \hat{\alpha},$$

$$\hat{\delta_i} = \hat{\beta_i} - \hat{\beta},$$

A visualization of the residual in a scatter plot is in Figure 3.8. Observing an outlier in the scatter plot, the outlier was removed and we have the scatter plot of the residuals without the outlier in Figure 3.9.

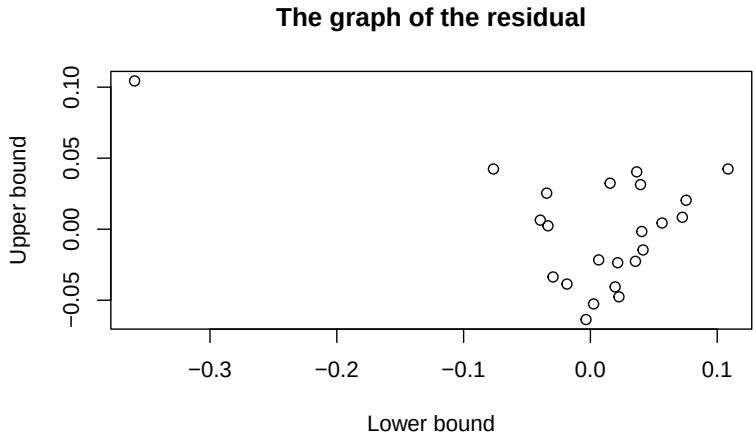

Figure 3.8: This is the scatter plot of the residual pairs ($\hat{\epsilon_i}$, $\hat{\delta_i}$). The correlation is -0.437.

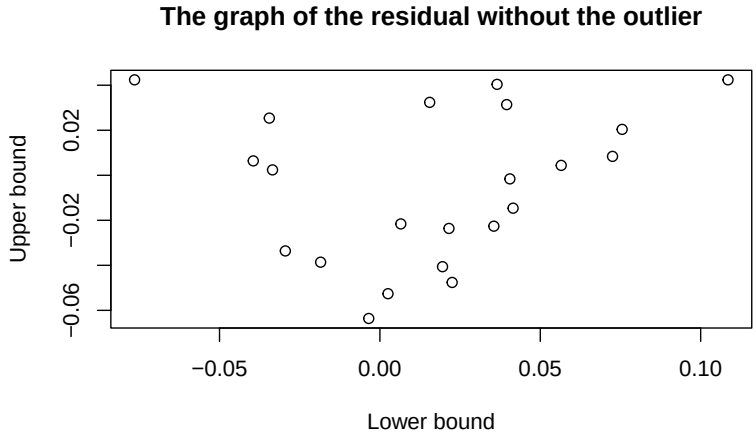

Figure 3.9: This is the scatter plot of the residual pairs without the outlier. The correlation is 0.1733.

Investigating the model, we wish to test the hypothesis

H_0 : There is no relationship between a deviation from where a muscle activates in a step and a deviation from where the muscle deactivates in the step.

H_1 : There is a relationship between a deviation from where a muscle activates in a step and a deviation from where the muscle deactivates in the step.

The correlation of the residual is 0.1733 and performing the test for significance of the correlation, it was observed that a 95$\%$ confidence interval contains zero, and the p-value is 0.4405, therefore it is not statistically significant. Therefore, we fail to reject the null hypothesis and cannot conclude that there is a relationship between the beginning of the activation of a muscle in a step and its ending. A confidence interval for the correlation is [-0.2678498 0.5544246].

CHAPTER 4

CONCLUSION

Working on this project, we were able to achieve these during the study: we were able to clean and transform the data, also smooth and average the EMG steps using `moving_average()`, we were able to transform the EMG data into a signal, and also were able to get the active interval region of the steps using the `find_better_interval()`.

Although all these mentioned above were done, there are some limitations. The limitation of the project are:

- The code is not efficient as it takes a considerable amount of time to execute.
- Our code only works for a muscle with only one spike and does not work for one with multiple spikes.
- A normality test for the residual distribution was not conducted.
- The data was collected from one person.
- The data was obtained from a healthy individual. Our procedure has not been tested on a person with neuromuscular issues.

Computerized detection methods have the advantage of increased objectivity of analysis, reduced time requirements, and fewer skill requirements for researchers.
With continued work, we hope our classification method will be competitive with the state of the art.

REFERENCES

[1] Evan D Crotty, Laura-Anne M. Furlong, Kevin Hayes, and Andrew J. Harrison, *Onset detection in surface electromyographic signals across isometric explosive and ramped contractions: a comparison of computer-based methods*, Physiological Measurement **42** (2021).

[2] Scott Day, *Important factors in surface emg measurement*, Bortec Biomedical Ltd Publishers (2002), 1–17.

[3] Raez M B I et al, *Techniques of emg signal analysis: detection, processing, classification and applications*, Biological procedures online vol. 8 (2006): 11-35. doi:10.1251/bpo115.

[4] Richard P. Di Fabio, *Reliability of computerized surface electromyography for determining the onset of muscle activity.*, Physical Therapy **67** (1987), 43–48.

[5] Paul W. Hodges and Bang H. Bui, *A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography*, Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control **101** (1996), no. 6, 511–519.

[6] Francesco Di Nardo, Teresa Basili, Sara Meletani, and David Scaradozzi, *Wavelet-based assessment of the muscle-activation frequency range by emg analysis*, IEEE Access **10** (2022), 9793–9805.

[7] Mamun Bin Ibne Reaz, M Sazzad Hussain, and Faisal Mohd-Yasin, *Techniques of emg signal analysis: detection, processing, classification and applications*, Biological procedures online **8** (2006), no. 1, 11–35.

[8] NM Sobahi, *Denoising of emg signals based on wavelet transform*, Asian Transactions on Engineering **1** (2011), no. 5, 17–23.

[9] David H Sutherland, *The evolution of clinical gait analysis part l: kinesiological emg*, Gait Posture **14** (2001), no. 1, 61–70.

[10] D.H Sutherland, *The evolution of clinical gait analysis: Part ii kinematics*, Gait Posture **16** (2002), no. 2, 159–179.

[11] D.H. Sutherland, *The evolution of clinical gait analysis part iii – kinetics and energy assessment*, Gait Posture **21** (2005), no. 4, 447–461.

[12] Jian Zhang, Rahul Soangra, and Thurmon E. Lockhart, *A comparison of denoising methods in onset determination in medial gastrocnemius muscle activations during stance*, Sci (2020).