University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

5-2018

Longitudinal tracking of physiological state with electromyographic signals.

Robert Warren Stallard
University of Louisville

LONGITUDINAL TRACKING OF PHYSIOLOGICAL STATE WITH

ELECTROMYOGRAPHIC SIGNALS

By

Robert Warren Stallard

M.S. Physics, 2013

A Dissertation

Submitted to the Faculty of the

J.B. Speed School of Engineering of the University of

Louisville

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in Electrical Engineering

Department of Electrical and Computer Engineering

University of Louisville

Louisville, Kentucky

May 2018

LONGITUDINAL TRACKING OF PHYSIOLOGICAL STATE WITH ELECTROMYOGRAPHIC SIGNALS

By

Robert Warren Stallard

A Dissertation approved on

April 11, 2018

by the following dissertation Committee:

_____

Dissertation Director
Dr. Karla Conn Welch

_____

Dr. Enrico Rejc

_____

Dr. John Naber

_____

Dr. Jacek Zurada

DEDICATION

I would like to dedicate this work to my Grandmother and Grandfather, for teaching me at an early age the rich rewards that wisdom and discipline earn; my Father for teaching me that the universe is limitless, yet in front of you, and perceivable, yet unadvertised; my Mother for teaching me patience in adverse times, and Amber, for bearing these burdens with me – and thriving. This work is also dedicated to Bella, whose forced rigors have kept me disciplined, wise, and stable.

Finally, this work is dedicated to the toiling masses who are responsible for the material foundations which make possible any and all scientific inquiry. May future scientific endeavours be in your service!

ACKNOWLEDGMENTS

It is a meandering path that carries one with formal training as a physicist to this sort of work, which has been most fruitful and intriguing. It is not without hurdles, naturally, but I am fortunate to have many people to help me through the process.

I would like to thank Dr. Jacek Zurada for introducing me to these concepts in an elegant and intriguing way that reinvigorated my imagination of the possible in realm of, especially, computational physics. I would like to thank Dr. John Naber for introducing me to the spinal cord rehabilitation efforts at the University of Louisville, and who also introduced me to Dr. Karla Conn Welch, who has guided me throughout my doctoral work in an outstandingly organized and patient fashion. I could not have asked for a better mentor. A physicist's journey into the realm of physiology can be terrifying, but I have Dr. Enrico Rejc to thank for guiding me through it and assisting with my first contact with the field.

ABSTRACT

LONGITUDINAL TRACKING OF PHYSIOLOGICAL STATE WITH
ELECTROMYOGRAPHIC SIGNALS

Robert W. Stallard

April 11, 2018

Electrophysiological measurements have been used in recent history to classify instantaneous physiological configurations, e.g., hand gestures. This work investigates the feasibility of working with changes in physiological configurations over time (i.e., longitudinally) using a variety of algorithms from the machine learning domain. We demonstrate a high degree of classification accuracy for a binary classification problem derived from electromyography measurements before and after a 35-day bedrest. The problem difficulty is increased with a more dynamic experiment testing for changes in astronaut sensorimotor performance by taking electromyography and force plate measurements before, during, and after a jump from a small platform. A LASSO regularization is performed to observe changes in relationship between electromyography features and force plate outcomes. SVM classifiers are employed to correctly identify the times at which these experiments are performed, which is important as these indicate a trajectory of adaptation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND MOTIVATION

Observing and diagnosing the changing conditions of various physiological states in humans with algorithmic approaches is an important research field, and one that could aid in advancements of diagnostic and clinical technologies for non-invasive observation of the neuromuscular system [1, 2]. For example, NASAs Strategic Knowledge Gap (SKG) reports from LEAG (Lunar Exploration Analysis Group) [3] and MEPAG (Mars Exploration Program Analysis Group) [4] indicate broadly that "how to maintain peak human health and performance in dusty, high-radiation, partial gravity environments" is a matter of concern. In particular, methods for gauging the relative state or physiological conditioning of individuals in spaceflight is both lacking and necessary, for example in sensorimotor perception [5]. In addition to the MEPAG and LEAG reports, Cermack (2006) argues for autonomous, on-site decision-making capabilities, such as in a medical scenario, where crew can operate independently of ground control [6]. Concern for human health and well-being naturally arises in many areas during manned spaceflight. Specifically, the effects of long-term spaceflight on sensorimotor perception [5], balance [7], visual target acquisition [8], functional mobility [9], and walking strategy [10] have been studied. On Earth, these methodologies have broad applications with numerous approaches studied, ranging from characterization of improvement from spinal cord injury, to characterization of the advance of degenerative neurological diseases such as Parkinsons disease. It is the principal

aim of this work to examine the application of algorithmic approaches to longitudinal physiological models.

Up until this point, it appears that much work has yet to be done exploring the idea of employing algorithms to characterize, predict, or discriminate changes in physical state over time. We suggest that, as it is possible to diagnose a condition, it should well be possible to observe how such a state changes over time (e.g., generally becoming worse or better, but described more appropriately depending on the situation). This has value in the case of recovery from spinal cord injury, just as it does in the case of quantifying a recovery trajectory in some deficient area after spaceflight.

Therefore, in this work, we explore a synthesis of algorithmic and nonstationary signal processing methods in an effort to elucidate situational physiological details from noisy surface electromyographic (sEMG) signals [11]. In particular, we aim to analyze sEMG data from surface electrodes to characterize some detail or state of human physiology. The main contribution of this work is the establishment of a method for characterizing, classifying, or predicting physiological states undergoing longitudinal study. In this specific case, we use electrophysiological data in concert with other variables such as ground reaction forces and force production as the impetus for developing this method [12, 13].

This work attempts to improve the use of sEMG in, at a minimum, building statistical models for prediction and/or classification of physiological outcomes (e.g., days since bedrest, state of disuse, sensorimotor performance). One hurdle common to such physiological studies is the subject-to-subject variation in performance observed over time. We aim to negotiate this hurdle by investigating the issue algorithmically, by predicting or classifying these states better than chance, as described in more detail later in this dissertation. A more challenging goal is the discovery of underlying physiological connections or mechanisms responsible for the change in state and subsequent change in motor function, such as observed post-spaceflight locomotor

failures. The challenge lies in the fact that, in many cases, experiments are limited in size, complexity, and number.

Ultimately, sEMG measurement alone has limits in what can be explored by its measurement. Such measurements are merely produced by motor unit activity and do not directly indicate what is occurring in the entire sensorimotor system of humans. A broader systemic approach to studying, for example, the effects of spaceflight on human neurophysiology would require observations of changes in all parts of the sensorimotor system; including measurements in the central nervous system, sEMG measurements, the vestibular system and its associated processing areas in the brain, etc. This work focuses on sEMG, but adjacent studies involving mapping the brain via fMRI before and after spaceflight are underway [14] to investigate the adaptations of the central nervous system in long-term immersion in microgravity environments. These adjacent studies complement this work and support the importance of developing analytical techniques to assist in maintaining human health and well-being. See the review by Van Ombergen et al. (2017) on spaceflight-induced neuroplasticity for an overview [15] and the work by Cassady et al. (2016) for an example of the use of fMRI techniques in investigating changes in sensory network connections in the brain in simulated spaceflight environments [16].

Raw sEMG is difficult to employ for direct physiological interpretation. Such data simply indicates the presence of some motor unit activity in addition to a great deal of noise. Instead, one would prefer to extract potentially-informative features from the signal. Features, or statistics of interest, include: the root-mean-square (RMS) amplitude, zero crossings (ZC), median frequency (F_m), integrated EMG (iEMG), autoregression coefficients, moving averages, and pulse intervals [17–19]. This work utilizes many of these features in addition to the less-commonly employed wavelet coefficient-derived features.

As a brief orientation, in general, as sEMG activity and amplitude increases, so

does the force output. The particular relationship between force and motor unit activity depends on the muscle and movement in question and is not typically linear, but the logic is simple: as the number of motor units exhibiting electrical activity increases, so too does the voltage measured by the electrode and the net force output. Examples of adverse conditions that may have an effect on the signal include various forms of neuropathy (motor neuron decay) and myopathy (muscle fiber decay). Clinically, a professional may take a qualitative look at signals and make judgements based on a subjective reading of the data, other symptoms, and professional experience. A more algorithmic approach is exemplified in the work of Kamali et al. (2014), whose approach utilizes a variety of classification algorithms and wavelet-derived features which will be described in Chapter 2 [2]. It is appropriate here to reiterate the principle aim of this work to explore the development of methodology for algorithmically marking long-term changes in physiological state. The development of algorithms and analytical processes for longitudinal study of physiological state has numerous potential applications including quantifying human health and well-being under the duress of long-term space exploration and enhancing the long-term efficacy of human-machine prosthetic or robotic interfaces.

Other limitations exist in the development of useful features to inform algorithms. For example, a look at the frequency spectra of sEMG data can give information about the fatigability of the neuromuscular system in the form of shifts in median frequencies [20]. The power spectra of an sEMG signal may additionally give some idea as to the fiber type makeup within a muscle, as a higher median frequency indicates a stronger contribution to the signal by fast twitch fibers. As fast twitch motor units (MUs) participate less due to easy fatigability, the bulk of the action is taken up by slow twitch MUs and the power spectrum shifts left to a lower frequency range. Experimenters, for practical limitations of the Fourier Transform (FT), choose relatively smooth isometric motions to minimize the nonstationary elements (e.g.,

transient bursts) inherent in sEMG signals. When examining non-smooth motions, analysis with FT is limiting. The wavelet transform is employed in this work to work around the limitations concerning nonstationarities that FTs have.

1.1 Summary of Contributions

We shall briefly describe the principle contributions of this work.

1.2 Classify Differences of Physiological State from Longitudinal Measurements

A physiological state may vary depending on various factors such as disuse (e.g., bedrest), presence in a microgravity environment as in spaceflight, or the progression or recovery from disease. This work seeks to identify means and methods to discriminate between observations made at different points in time. One dataset examined in Chapter 3 involves discriminating between isometric contraction observations taken for bedrest subjects before and after a 35-day bedrest [11]. The second example used in this work is more complex in that sensorimotor performance observations of dynamic, nonisometric motions are made at different intervals before and after spaceflight or bedrest by measuring sEMG (muscle activation) during a jumpdown task (JDT) consisting of a jump from a platform, adjustment in air in preparation for landing, the landing itself, and reaction to the landing toward achieving balance [12, 13].

We demonstrate algorithmic classification of endpoint changes in physiological state well-above chance (i.e., greater than 50% cross-validated accuracy for a binary case) for both binary and multi-classification problems. The classification results are presented in Chapters 3 and 5.

1.3 Identify Physiologically-relevant Features Derived from the Data

Congruent with the two other aims described here is the notion of identifying physiologically-relevant features. In addition to prediction and classification capabilities, the discovery of physiologically interpretable features that explain variance in the data may enhance the capabilities of diagnostic and clinical technologies. Therefore, in addition to a "bag of features" approach in which field-standard quantities are calculated from each signal, new composite features will be constructed in an attempt to elucidate underlying physiological or kinematic processes. A tracjectory of change is also identified in addressing this approach.

The results addressing this contribution are presented in Section 5.2.

1.4 Overarching Theme and Area of Exploration

Section 2.2.1 describes the process of discriminating between observations involving physical conditions. In each case, experimental conditions are planned to generate the signals and some data wrangling is necessary to prepare the data as inputs for an algorithm. The particular methods of organizing the data may involve some domain knowledge, organization natural to the task, or observation at hand. In Nair et al. (2009), for example, the sEMG signals are normalized with respect to gait cycle that is, the raw signal is compacted to some fixed-length vector where the first vector element represents the initiation of gait and the final point represents the conclusion of the cycle [21]. Further examples are given in detail in the next two sections. Section 2.2.2 describes classification of signals corresponding to particular physical configurations such as hand-gestures. These tasks, as with the first set, have pre-processing steps specifically tailored for the experimental conditions involved and algorithms employed.

The means by which we inform our classification algorithms are equally important

6

as the algorithms themselves. Indeed, much of the effort involved in this work involves extracting relevant portions of a signal, or organizing signals, then processing the signals in a way that extracts meaningful features for algorithms. Therefore, in Section 2 we briefly discuss, by way of exhibition, sEMG features in the time domain, the frequency domain, and the time-frequency domain.

The examples described in Sections 2.2.1 and 2.2.2 address cases of instantaneous differences between populations or trials (e.g., gestures). The specific aims of this work are geared toward addressing the issue of differences in populations over time, or longitudinal classifications, and Chapters 3, 4, and 5 constitute the purpose of this dissertation which is to demonstrate these classifications. The approaches outlined here both inform and illustrate the approach that is taken to analyze the preliminary work of the next section in which we are utilizing similar sEMG pre-processing approaches as described in this section. The primary differences are situational (the JDT data includes force features, for example) and the fact that these features are being examined over time.

CHAPTER 2

BACKGROUND AND STATE OF THE ART

For purposes of orientation, this section introduces, in a concise manner, three topics that set the stage for the work that follows. Section 2.1 introduces the otolith structure and its relationship with the vestibular system overall and the sensorimotor system in particular. Finally, a showcase of recent work taking a machine learning approach to physiological conditions and gesture recognition is presented. The sEMG signal itself is described on a case-by-case basis in relation to its particular investigation.

Not included in this section is a treatment of the interesting work done involving electroencephalogram (EEG) and machine learning techniques in a variety of capacities. These will be speculated upon later in this dissertation as an outgrowth of possible future investigations, but will not be expounded upon otherwise.

2.1 The Vestibular System and Sensorimotor Deficiency

Chapters 4 and 5 relate to the JDT performed by NASA researchers and astronauts as one means of examining the changes in sensorimotor performance over time. This section describes elements of the physiology involved in perceiving motion and their microgravity or bedrest induced changes. Sensation of motion and postural control rely on an integration of many sensory components in the body. Some of these include

visual stimuli, somatosensory stimuli (e.g., touch), and vestibular stimuli, discussed here in more detail.
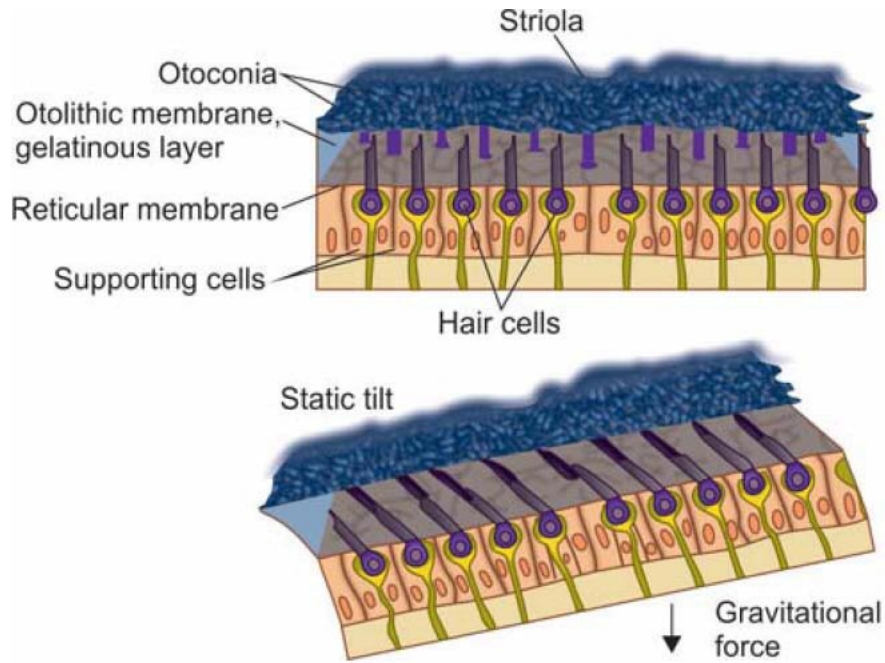
Figure 1. Illustration of the otolith function under gravity. Adapted from [22].

Much of the research revolves around the otolith function pictured in Figure 1. This system consists of calcium carbonate crystals (otoconia) resting on hairs supported by a gelatinous membrane. These crystals constitute an inertial mass element that resists changes in motion, thus introducing a deflection angle in the hairs supporting them. This introduces a potential which is then processed by the brain, informing sensations of linear acceleration and tilt.

In a microgravity environment, the tilt component is missing as the gravitational reference plane is no longer present in a significant sense. With this signal input missing, but still present in, for example, visual stimuli, adaptations occur that confound the signal's interpretation. A tilt may be perceived as a linear acceleration [23] and the mismatch with other sensory inputs can induce postural instability, dizziness and vomitting [24]. This effect may also persist for some time upon return to earth, and may pose risks regarding vehicle control during missions [5]. This *tilt-translation*

illusion is particularly common in the days of readaptation, so the concern for performance impact is primarily around crucial stages of a mission such as landing on a planetary surface or readjusting to orbital flight [25].

The subject population exhibits considerable subject-to-subject variability in adaptation and recovery. First-principle causes of this variability are difficult to study due to the complexities of investigating human neurophysiology and not well represented in the literature. However, there are some intriguing investigations in mass differences of otoconia demonstrating a correlation (and likely at least partially causal) with abormalities in, e.g., swimming behavior of salmon [26] with varying otoconia mass differentials, suggesting that in addition to adaptation to a microgravity environment itself being a source of disorientation, this phenomena may be in part explained by the organ responsible for tilt perception itself. While subject-to-subject variability is high, it nonetheless remains true that incidence of discomfort (e.g., headache, vomitting, and clusminess) is high even for shorter duration flights (space shuttle) and the effect tends to be greater for longer duration missions, such as those on the ISS [27].

Broadly, then, the perception of motion and position begin with the sensing mechanisms themselves and are integrated and processed by the central nervous system, thus informing strategies for, as a rather convenient example, landing a jump from a short platform. First-principles investigations might be concerned with investigating disparities in the structure of sensing organs (as in the mass-difference otolith example) or with investigating the adaptations themselves by way of, say, fMRI studies of the brain pre- and post-flight [14]. In Chapters 4 and 5 we take steps to complement this by building an approach around sEMG measurements.

2.2 State-of-the-Art in Physiological Assessments

Already, much work has been done in classifying instantaneous physiological configurations – be it an arthritic condition or a hand gesture, for example. It is this

instantaneous classification that is counterposed against the impetus of this work, in which we are chiefly concerned with physiological changes in a population over time. Sections 2.2.1 and 2.2.2 outlines two broad categories of work in which machine learning algorithms have been employed toward instantaneous physiological classification.

2.2.1 Physiological Conditions

Instanenous identification of physiological conditions has been investigated in a number of studies including the work performed by Nair, French, Laroche, and Thomas (2010) [28]. Subjects with either rheumatoid arthritis, osteoarthritis, or no condition (control group) were studied, and their data were labeled as one of these three conditions. Bipolar surface electrodes were used to measure sEMG from six different muscles of the leg for the arthritic groups and seven different muscles for the control group. The same muscles from each group were used for purposes of comparison, however. These measurements were taken during gait.

Following sEMG collection, input vectors constructed by concatenating sampled signal measurements from each muscle informed the algorithms employed in the study. Several algorithmic approaches were employed including Least-Squares Kernel (LSK) (multiple variations), neural networks, and Linear Discriminant Analysis (LDA). Rheumatoid arthritis was correctly distinguished from the control group with 90.72% accuracy while osteoarthritis was correctly classified against the control group 98.23% of the time. Further, the study investigates how these algorithms might illuminate some therapeutic approach or condition monitoring guidelines by removing certain muscle groups from data prior to inputting the data into the algorithm. It was found that the muscles contributing most to the classification effort for rheumatoid arthritis vs. the control group were the soleus and biceps femoris, whereas for the osteoarthritis vs. control group case the gluteus medialis muscle was the primary driver in classification accuracy. This kind of effort exhibits the potential for using

feature selection methods to highlight areas of inquiry, similar to the second specific aim described in Section 1.3.

As previously described, needle electrodes can be employed to acquire EMG signals from single fibers. Kamali et al. (2014) , look at the shape of the MUAPs themselves in order to distinguish between healthy, neuropathic, and myopathic conditions [2]. After decomposing an EMG signal into a MUAP train (a sequence of MUAPs), they extract features describing the MUAP in both the time domain (rise time, duration, peak-to-peak amplitude) and time-frequency domain via discrete wavelet transform (sub-band coefficients mean absolute value, standard deviation). These features are then fed into a variety of classification schemes including a variety of Support Vector Machine (SVM) and k-Nearest Neighbor (k-NN) configurations, and fusions of these. The fusion methods, in which multiple algorithms are employed and the final class of the feature set is decided upon by majority voting, were found to be the most effective with near-perfect classification accuracies. These approaches warrant caution, however, as these more complex ensemble approaches often sacrifice interpretability for accuracy.

Subasi et al. (2006) consider the use of artificial neural networks (ANN) and wavelet neural networks (WNN) fed by autoregression coefficients (as features) to characterize MUAPs associated with neuromuscular disorders [29]. Their study measures intramuscular EMG data from 7 healthy subjects, 7 myopathic subjects, and 13 neuropathic subjects. Concentric needle electrodes are inserted into the biceps brachii (upper front arm) and subjects are instructed to perform an isometric contraction (i.e., perform a curl). At 30% of the maximal voluntary contraction (MVC), the EMG is recorded for each subject.

The signals are treated as an autoregressive model in which the signal is considered as a linearly-filtered Gaussian random process whose output is essentially the signal [30]. The autoregressive filter is described by its coefficients which are supposed to be

unique to the signal. These are then used as inputs into the classification algorithm. For purposes of novelty, only the WNN is described here. The WNN is a variant of the ANN in which the node activation function is described by a wavelet function (rather than, say, a threshold value) described further in [31]. The WNN method offered a best-case classification accuracy of 92% for myopathy, 86% for neuropathy, and 94% for healthy subjects.

These studies demonstrate the feasibility of an algorithmic approach to classifying otherwise noisy signals into their appropriate categories for the cases where the categories describe some physical condition. In the first case, Nair et al. (2010) employ surface electrodes in a particular experimental condition while in the last two cases needle electrodes are employed. In each case the end result is, essentially, feature extraction and algorithmic classification.

2.2.2 Gesture recognition

Another area of interest involves improvements over prosthetic devices. Advanced replacements in the case of loss of hands, for example, are sought so that more complex grasp and gesture motions may be performed. This would be a vast improvement over the more traditional hook. Already, firms like Lego(2015) [32] and Touch Bionics(2017) [33] have produced new prosthetic devices that utilize signals from residual muscle tissue. However, these are still immature technologies that may rely on external control mechanisms to allow a wider variety of grips and gestures. Touch Bionics device, for example, comes with a phone app that allows the user to select from some a set of grips, or even Grip Chips that are placed on surfaces where the user may have frequent need of interacting nearby. On the Touch Bionics website an example is provided whereby a user has placed a Grip Chip on the side of a pencil holder. When the bionic limb is moved nearby, the phone detects this proximate signal and the bionic hand moves into a pre-programmed configuration. These extra configurations

require external devices simply because either the device inputs (sEMG) are insuffi-ciently varied or their post-processing methods cannot identify the desired grip from the existing inputs. More sophisticated approaches are needed to improve these pros-thetic devices in order to reduce dependence on external control mechanisms. Some approaches include adding accelerometer data to the inputs [34], or even employing electroencephalogram (EEG) inputs [35].

Any approach dependent on myoelectric input is still subject to changing condi-tions over time, however. Electrodes may shift [36], the skin-electrode interface may become more (or less) moist over time [37], and the muscle fibers may experience fatigue thus altering the signal content [38]. Any of these time-dependent issues may invalidate any supervised model for mapping a set of signals to a grasp or gesture, thus effectively limiting the performance of the device. While this work is not focused on solving these problems specifically, it is focused on and indeed exploits long-term changes in physiology and an algorithmic approach to recognizing them.

Phinyomark et al. [39–43] have studied algorithmic approaches to gesture-recognition via sEMG input. For brevity and to avoid repetition as this work is expounded upon in greater detail in Section 3, we focus here on methods utilized in an earlier pub-lication by Phinyomark [39]. In this study, sEMG is recorded from the flexor carpi radialis and extensor carpi radialis longus muscles located, roughly, on the bottom and top of the widest part of the forearm, respectively. Six different hand-gestures including wrist flexion and extension, hand closed and opened, and forearm prona-tion and supination are performed by healthy subjects only. The objective here is to discriminate between each gesture based on the two input sEMG signals alone.

Signals from selected muscles were band-passed filtered from 10-500 Hz at 1000 Hz. These gesture trials were repeated 10 times per gesture and signals were post-processed to conform to a 256 ms window. A 4-level DWT decomposition was then applied to these signals using 7 different mother wavelets coming from the Daubechies,

Coiflet, Symlet, Biorthogonal Spline, and Reverse Biorthogonal Spline families. The inverse process was also performed such that each decomposition level was reconstructed. An illustration of this process will be given in Section 3, as it was employed in that work. From these signals and their derivatives, the average and standard deviations of the Mean Absolute Value (MAV) were calculated to examine how well the features clustered for each trial and gesture. The study demonstrated strong clustering trends but did not examine classification yields.

Work along these lines continues in [41] where similar experiments were performed using the same 6 gestures previously described plus two new gestures (wrist radial and ulnar deviations) for a total of 8 gestures. From the time-domain, ZC, iEMG, MAV, and RMS and more are calculated. In the frequency-domain, mean, median, and peak frequencies, and mean and total power are among calculated quantities. From the time-frequency domain, features such as the DWT coefficients, Wavelet Packet Transform Energy (WPTE) and Relative Wavelet Packet Transform Energy (rWPTE) are employed. There are 7 feature sets used to inform the algorithms that are constructed in advance. A number of classifiers are used to demonstrate 10-fold cross-validated algorithm performances. For brevity we highlight the simple LDA case in which two different feature sets yield 91.67% and 92.19% accuracy. This is a multi-classification approach in which all gestures are considered at once, as opposed to a series of binary classifiers (against some control perhaps). The first feature set obtaining 91.67% accuracy is the commonly employed time-domain feature set described by Hudgins et al. [44]. The second feature set obtaining 92.19% accuracy consists of the RMS values and first four autoregression coefficients.

The example cases of the above-described work performed by Touch Bionics and Phinyomark et al. illustrate, respectively, a use for instantaneous classification of physiological state and approaches of performing that work using sEMG. The next examples illustrate a commercial approach at integrating sEMG inputs with accelerom-

eter and gyroscope inputs.

In the commercial domain, Thalmic Labs has engineered an armband called Myo. The device fits over the forearm and utilizes both spatial data (via accelerometer and gyroscope) and sEMG data as inputs for gesture detection. The device natively supports 5 hand gestures. A Software Developers Kit (SDK) [45] is available in which one could add more detectable gestures, but the classification problem becomes more difficult as the number of classes increases. The introduction of spatial data allows more degrees of freedom in human-computer interaction in addition to gesture-detection from the forearm, users can move their arms around in particular patterns (forming "shapes") as device inputs. Currently, the Myo has rather limited use but its commercial emphasis is as another way of controlling, for example, computers, gaming systems, and TVs, with gestures.

A similar, custom forearm armband device was employed by Riillo et al. [46] to examine the use of gesture-recognition on healthy subjects and a transradial amputee. The particular subject in this case was amputated 11 cm below the elbow. The study measured sEMG using six sensors spaced equidistant on an armband placed around the forearm. The armbands were oriented such that the first sensor recorded sEMG from the flexor carpi-radialis and the last sensor recorded sEMG from the brachio-radialis. The amputee subject had sufficient residual muscle so that this could be replicated in their case.

Following a similar bag-of-features scheme, a number of quantities are calculated from the sEMG signals including, but not limited to: mean sEMG amplitude, RMS, variance, slope sign change, and simple square integral. Three classifiers are tested with the feature data including LDA, non-linear Gaussian Radial Basis SVM, and Artificial Neural Networks (ANN) with a hidden layer of 10 neurons and an output of 5 neurons trained with a back-propagation algorithm. The success of these classifiers was studied before and after PCA was applied to reduce the dimensionality of any

16

selected feature sets.

Features were tested individually and then the input feature vector was added to based on those outcomes constituting a sort of manual forward feature selection process. In the end, the healthy subjects group gestures were accurately classified as high as 87% of the time. Curiously, the lone amputee subject had a higher classification accuracy of 93%. The 25:1 sample size difference must be noted here with appropriate caution, of course.

This section emphasizes methods of acquiring sEMG signals, extracting features, and performing classification tasks rather than the specifics of forearm measurements. A similar approach will be seen in Chapter 3 as applied to a bedrest study and the NASA JDT data in Chapters 4 and 5, where these methods are applied to sEMG obtained from the lower limbs.

CHAPTER 3

A BINARY CLASSIFICATION OF PRE/POST BEDREST
SUBJECTS

This section describes the first original constituent work of this dissertation extending
the concept of instantaneous classification of physiological state described in Section
2.2.1 to the temporal domain i.e., classifying populations as they are subjected
to some change over time. The first work [11] presented here examines a binary
classification case in which otherwise healthy subjects undergo a bedrest. The task
in this case is to separate the pre- and post-bedrest cases.

3.1   Motivation and Background

The algorithmic approach to previously collected bedrest data here, demonstrates
the use of wavelet-derived features from surface electromyography (sEMG) signals
to classify subjects undergoing a 35-day bed rest using the k-Nearest Neighbors (k-
NN) and Support Vector Machine (SVM) classification schemes. The greater role
that this study serves is toward the development of measurement and classification
techniques that may shorten an otherwise long feedback loop for rehabilitation or
physiological diagnostics. Establishing non-invasive, high resolution, and less time-
consuming methods of categorizing or describing an individual's physiological state is
also valuable for the development of autonomous diagnostic or rehabilitation systems.
Toward these goals, it is necessary to compare measurements made on the same

individuals using the same methods but with a larger span of time between the two to mark physiological changes, rather than characterization of sEMG output to mark immediate physiological conditions (e.g., gestures or presence of arthritis). In the present discussion, we investigate sEMG data taken from the lower limb of subjects performing three different prescribed motions before and after bedrest. Together, these approaches and this investigation constitute progress toward the development of tools for the description of physiological processes (i.e., diagnoses, monitoring), and prediction of outcomes.

Well-controlled trials with careful observations of physiological changes ease the development of methodology for understanding and quantifying the stages of rehabilitation in other areas including exposure to microgravity, recovery from traumatic injury, and rehabilitation of individuals suffering from spinal cord injury. The intent here is to develop analytical tools and capabilities to further identify clinical variables marking physiological changes over time. This particular study's data is sourced from a simulation of microgravity via bed rest, which is known to induce a decline in muscle function due to muscle atrophy, reduction of neural drive and muscle specific force, particularly in the postural muscles of the lower limb [47]. The ability to observe, predict, and mitigate the loss of muscle mass and function is of particular interest to agencies interested in human-based space exploration [3, 4] and organizations interested in quantifying the effects of various neurological or physical disorders, such as muscular dystrophy [48].

In the recent decade, a substantial body of work has been published concerning a variety of methods for the use of sEMG signals as a means to control prosthetic devices [49–52] and diagnose physiological disorders [21, 53]. Phinyomark et al. (2011-2012) have made use of features derived from the Discrete Wavelet Transform (DWT) for the purpose of classifying hand gestures [39, 42]. Additionally, statistical comparisons were performed for a several wavelets, features, and algorithms [40] and have

demonstrated classification accuracy of greater than 98% with six different hand gestures. Karlik (2014) has reviewed studies utilizing both wavelet methods and time-series modeling to extract features for the purpose of prosthetic-related classification with upper classification accuracies ranging from 90% to 98% for time-series methods and 92% to 98% percent with wavelet methods [54]. While their performances are comparable, the DWT is preferred here due to its unique properties including the ability to process the signal in both frequency and time domains while providing a "zoomed-in" view across a range of frequency bins without distortion. These reviewed studies report on a number of features to be examined when analyzing sEMG signals with a focus on classification of hand-based gestures by utilizing an array of features measured from lower limb muscles, across time before and after bed rest, in an effort to identify changes due to physical deconditioning, rather than instantaneous movements.

Work has also been done in utilizing features derived from sEMG signals to determine whether or not an individual has a disorder such as arthritis. Nair et al. (2009), employ time-based features and test eight algorithms to classify between healthy subjects, subjects with rheumatoid arthritis, and subjects with osteoarthritis [21]. Nair et al. emphasize that the purpose of these classification schemes is to provide insight into muscle behavior in order to improve rehabilitation. sEMG recordings of six muscle groups within each leg of the patient are performed while the subject is in gait (walking). Utilizing both neural nets and kernel methods, they demonstrate up to 97% accuracy for the three-class problem.

One method of evaluating overall neuromuscular changes due to prolonged disuse includes a statistical evaluation of changes in lower limbs' power and force output. A complementary study analyzed the effects of bed rest in comparison to power output recorded during explosive extensions of the lower limbs [55]. The study supports its conclusions that muscle atrophy was mitigated by a moderate caloric diet restric-

tion that significantly reduced the body fat gained throughout 35 days of bed rest; however, this positive outcome was not sufficient for reducing the loss of maximal explosive power of the lower limbs.

By examining the bedrest data, the aim is to demonstrate the use of direct electro-physiological measurements – the surface electromyogram (sEMG) – to discriminate between two physiological states, pre- and post-bed rest that might otherwise be done by comparing morphological characteristics or force output measurements. This effort aligns with similar efforts at algorithmic classification of instantaneous states as previously described in Section 2.2.1, by exploring the similar problem of classification along a time-dimension. Additionally, we demonstrate the improvement in classification yield obtained when employing wavelet-derived features.

3.2   Experimental Methodology

The data analyzed in this study were collected from 10 healthy male subjects who underwent 35 days of bed rest under a moderate caloric diet restriction (1.2 times their resting energy requirement). Physical characteristics of the subjects are reported by Rejc et al. [55]. Before and after bed rest, the subjects performed non-fatiguing, isometric, maximal voluntary contractions of knee extensors, knee flexors, and plantar flexors with the right (dominant) lower limb. Force output and sEMG from the gastrocnemius medialis (GM), rectus femoris (RF), vastus lateralis (VL), and bicep femoris (BF) were sampled at 1 kHz using an EMG100C BIOPAC Systems, Inc. amplifier connected to an MP100 BIOPAC data acquisition system. The experimental setup is described in detail by Rejc et al. [56]. Briefly, the isometric contractions were performed by the subjects seated on either a special chair (knee flexion and extension) or sledge ergometer (plantar flexion). Two maximal voluntary contractions of 4–5 s under each isometric effort were performed. To prevent fatigue, after each contraction the subject rested for 2 minutes.

sEMG signal data was digitally pre-processed by applying a Hilbert transform to the sEMG signal envelope (obtained via zero-phase, 5th order butterworth filter with a low-pass corner frequency of 10 Hz) and determining the starting and ending moments of the signal, ultimately selecting the active portion of the sEMG signal. This selection was then extracted from the original signal, rectified and high-pass filtered with corner frequency of 10 Hz to remove any motion artifacts [57]. The active signal time is obtained by examining the signal from the primary muscle driving the movement (e.g., VL drives the knee extension). See Figure 2.
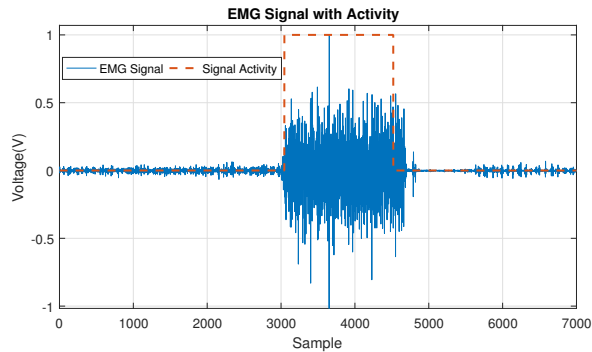
Figure 2. Illustration of active portion signal extraction.

The present study analyzes wavelet-derived features for classifying physiological changes indicative of prolonged disuse. The DWT is used to extract compact approximations of the original signal by decomposing it into coefficients corresponding to high and low frequencies (or detailed and approximate coefficients, respectively) [58], commonly referred to as a wavelet decimation. Chosen features are moving average (MAV), root-mean-square (RMS), the standard deviation (STD), median frequency (MF), waveform length (WL), skewness, kurtosis, and zero-crossings (ZC). These are derived from these wavelet coefficients, reconstructions from said coefficients, and the original filtered signal, followed by a forward feature selection algorithm to optimize classification results based on k-NN performance criteria. Additionally, a fourth-order autoregression model is estimated, using the Burg method, for non-wavelet transformed signals. Thus, there are two sets of features being compared.

The nonwavelet-derived features include all of the aforementioned feature types and the autoregression coefficients. The wavelet-derived features include all of the aforementioned features *except* the autoregression coefficients – these are calculated only on the wavelet-transformed signals. The comparison is therefore motivated to assess performance and utility of the wavelet-derived features as compared against more traditional sEMG measures in the context of longitudinal classification.

For the nonwavelet-transformed signals there are then 12 primary features per muscle for a total of 48 features per observation. There are 8 wavelet-derived signals per muscle for which there are 8 primary features for a total of 256 features per observation. The total count is then 304 features per observation. The primary feature types described in this section were chosen based on their commonality of use for analyzing sEMG signals. While there are no precise "equations of motion" for how these signals change over time, it is expected that changes in physiological behavior after bed rest should manifest themselves in some way in these primary features. The task then, is to determine if an algorithmic approach informed by these features can distinguish between these cases.

The final collected dataset included 15 pre-bed rest and 18 post-bed rest observations for the knee extension, 18 pre-bed rest and 16 post-bed rest observations for the plantar flexion, and 19 pre-bed rest and 18 post-bed rest observations for the knee flexion. The MAV, RMS, STD, skewness, kurtosis, and WL features are described as:

$$\text{MAV} = \frac{1}{N} \sum_{n=1}^{N} |x_n| \tag{1}$$

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} x_n^2} \tag{2}$$

$$\text{STD} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2} \tag{3}$$

$$\text{Skewness} = \frac{E[(X - \mu)^3]}{(E[X - \mu)^2])^{3/2}} \tag{4}$$

$$\text{Kurtosis} = \frac{E[(X - \mu)^4]}{(E[X - \mu)^2])^{4/2}} \tag{5}$$

$$\text{WL} = \sum_{n=1}^{N-1} |x_{n+1} - x_n| \tag{6}$$

where N is the total number of points, μ is the mean, and x_n represents the nth data point in the series. X indicates the signal time-series. E is the expectation value operator.

On the other hand MF is determined by finding the frequency bin, f_b, such that it divides the spectra into two halves with equal power amplitude:

$$\sum_{f_b=1}^{MF} P_{f_b} = \sum_{f_b=MF}^{TF} P_{f_b} \tag{7}$$

where P_{f_b} represents the power distributed within the frequency bin f_b and TF indicates the final frequency bin. ZC was obtained via MATLAB's `ZeroCrossingDetector` object located within the Digital Signal Processing toolbox. Similarly, MATLAB's implementation of the Burg method for estimating autoregressive coefficients was used. The selection of the AR(4) model was motivated by work showing these were robust for sEMG prosthethics applications [59, 60].

The wavelet decomposition was implemented using MATLAB's `wavedec` function for a total of 3 levels using the 7-tap Daubechies wavelet. The feature matrix columns

were normalized to prevent skewed feature selection due to the large magnitudes of, for example, ZC and MF as compared to MAV and RMS. Following this, forward feature selection was used with k-NN (k = 1-5) classification as the criterion for populating the feature vector. SVM classification with Radial Basis Function (RBF) and linear kernels was also performed in an otherwise identical fashion.

The forward feature selection is performed by iterating over available features and attempting the classification. A feature is added to the feature vector based on some criteria – in this case, minimizing the classification error rate. Table 1 illustrates an example case in which a feature vector consisting of four features are employed. The first of these features is the RMS value of the first-level details coefficients from the wavelet decomposition of the VL muscle with a 10-fold cross validated misclassification rate of approximately 20%. The last of these features is the RMS value of the third-level reconstruction of the details coefficients of the BF muscle. The final classification rate for this run was about 94%.

Table 1. Example forward feature selection result

Feature	RMS_CD1_VL	RMS_CD1_RF	RMS_CD1_BF	RMS_D3_BF
Criterion	0.2059	0.1176	0.0882	0.0588

Table 1 shows the results of one iteration of a forward feature selection dimensionality reduction scheme. As the reader moves to the right, the feature vector increases in length and the net error rate decreases. This result is one of many, as performance depends on the random assignment of folds in the cross-validation scheme and any random initialization in the classification algorithm. In this particular result, all four features selected are RMS features. The first selected feature is the RMS value of the first level detail coefficients (from the DWT) of the VL muscle. The second and third are the same, but of the RF and BF muscles, respectively. The last is the RMS value of the reconstructed signal obtained by inverse DWT of the third level detail coefficients of the BF muscle. A brief tutorial on this point is given in

Appendix B. Note that in this labeling scheme, the first label (e.g., RMS) indicates the feature, the second (e.g., CD1) indicates the signal from which that feature is derived, and the third indicates the muscle from which that signal is derived. That is, "FEATURE_SIGNAL_MUSCLE."

k-NN and SVM classification models are developed using the nonwavelet-derived feature set and the wavelet-derived feature set separately for sake of comparison. A range of nearest-neighbor values, k, numbering from 1 to 5, are chosen to investigate the range of behavior in classification performance, from granular to more broad classification neighborhoods. With a total observation number for each experiment type (knee flexion, knee extension, and plantar flexion) between 34 and 38 observations, larger nearest-neighbor numbers become unwieldy and overbroad for classification. In the case of the SVM classifier, the two different kernel functions were chosen to similarly explore the range in classification behavior.

The classification models all employ a 10-fold cross validation scheme in the feature selection process. The cross-validation procedure is used to characterize the validity and generality of the models. This procedure is useful in cases such as these in which the total number of observations is small. The exact train-to-test ratio depends on the total number of observations and random assignments of particular observations into folds, but is approximately 90% training data to 10% test data.

The concomitant goal of these classification attempts is to explore the utility of the wavelet-derived features, establish feasibility of longitudinal classification, and point to further research in classification model development for similar purposes.

3.3   Results

The resultant feature-complete matrix consisted of 34, 35, and 38 rows (observations) for plantar flexion, knee extension, and knee flexion, respectively, with each observation consisting of 4 muscles for which there were either: 8 signals including 4 wavelet
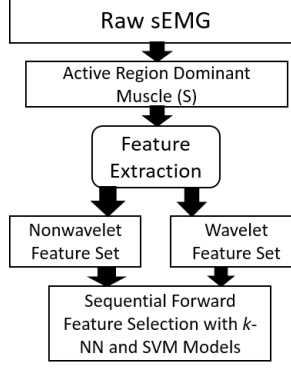
Figure 3. Features, including additional autoregressive coefficients, are calculated on the signals over the time period for which the muscles are active. This signal is transformed as described using the DWT and features, excluding the autoregressive coefficients, are calculated on these subsequent transformed signals. These two distinct feature sets are then used to inform cross-validated k-NN and SVM classification models as the selection criteria for the final feature vectors. That is, sequential feature selection is performed for each classifier configuration.

coefficients, and one reconstruction from each wavelet coefficient) obtained with 8 features each, resulting in 256 columns (features), or for the nonwavelet feature set, 12 features calculated on the signals from the 4 muscles. The issue of model instability (in which a variety of feature selection outcomes occur) under k-NN and SVM is additionally studied by examining the distribution of average classification accuracies for number of features selected for each nearest-neighbor number. These results are compared against models developed without the use of wavelet-derived features.

As outlined in [55], the subjects' physical characteristics changed in an expected manner. In particular, changes in total body composition and lean volume of the lower limb were limited yet statistically significant. Bed rest resulted in a decrease in total body fat mass and fat free mass (-4.8% and -2.4%, respectively) as well as in the lean lower limb volume (-3.6%).

3.3.1 k-NN Classification Results and Model Instability

A k-NN classification scheme was employed with the standard Euclidean distance metric using the optimal feature vector as deduced by the forward selection algo-

rithm. A k-fold cross-validation was performed to estimate the model's classification accuracy. The experiment's sample size is small, which yields an unfortunate variance in outcome based on how the samples are divided into the k-folds. These assignments into the folds are random, and, given that there are a number of outliers and ambiguous cluster boundaries, this randomness has an affect the outcome. As such, each experiment exhibits different classification performance. In assessing classifier performance, care must be taken to examine this performance variability.

This model instability is investigated by examining the average classification accuracy for 100 runs for $k = 1, 2, ..., 5$ for varying lengths of feature vectors. In this manner, we observe how model performance changes based on the number of features selected by the forward feature selection algorithm and based on nearest neighbor number.

Classification accuracy occupies a range of results for reasons previously described. The trend is promising, however, with typical classification averages as high as 91% and relatively tight-distribution of standard deviation values. Figure 5 demonstrates the range of results for varying nearest-neighbor number and experiment type.
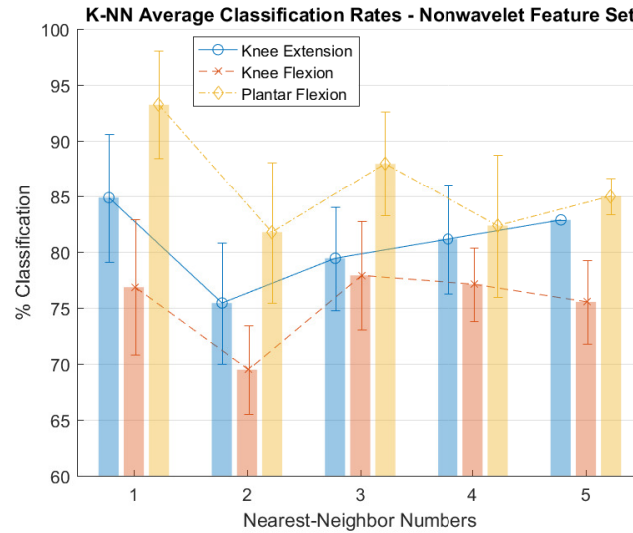
Figure 4. Average k-NN classification accuracy by experiment and nearest-neighbor value. Error bars indicate standard deviation of correct classifications. This figure does not include wavelet-derived features.
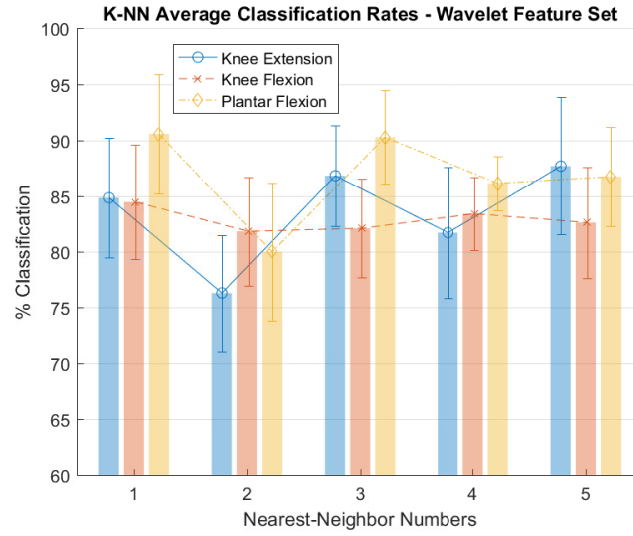
Figure 5. Average k-NN classification accuracy by experiment and nearest-neighbor value. Error bars indicate standard deviation of correct classifications. These include wavelet-derived features. A consistent performance improvement is demonstrated here on average.

Overall performance changes are indicated in Table 2. Average classification enhancement was as high as 12.4% for the two nearest-neighbor knee flexion case.

Table 2. Average performance changes for wavelet vs. non-wavelet derived feature sets. Positive values indicate improvement over models without wavelet-derived features. Largest gains seen in Knee Flexion cases. Values reported in the table are percents.

k	Extension	Flexion	Plantar Flexion
1	1.2	7.6	2.7
2	0.8	12.4	1.8
3	7.4	4.2	2.4
4	0.5	6.3	3.8
5	4.8	7.1	1.8

Classification accuracy variation is investigated by running a trial for k values 1 through 5, 100 times a piece for each experiment for a total of 1500 trials. Using the full feature set, optimal classification accuracy is obtained with one-nearest neighbor at using a 10-fold cross-validation scheme, however, higher nearest-neighbor numbers still yield reasonable accuracies. The partial feature set (excluding wavelet-derived features) exhibited less variation for the knee extension experiment and optimal classification accuracy with 5-nearest neighbors, as illustrated by Figure 6. Figure 7 through Figure 11 illustrate the model performance for varying k value and feature vector length in the same manner as Figure 6, for each muscle-feature set pairing.

Figure 6. Model performance for knee extension measurements. Blank spaces indicate cases that did not occur. These are results from the nonwavelet feature set.

30
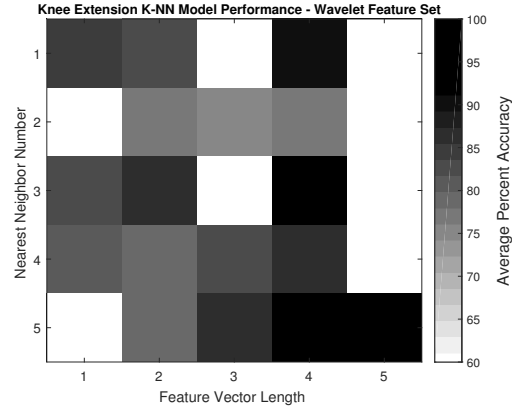
Figure 7. Model performance for knee extension measurements. Blank spaces indicate cases that did not occur.
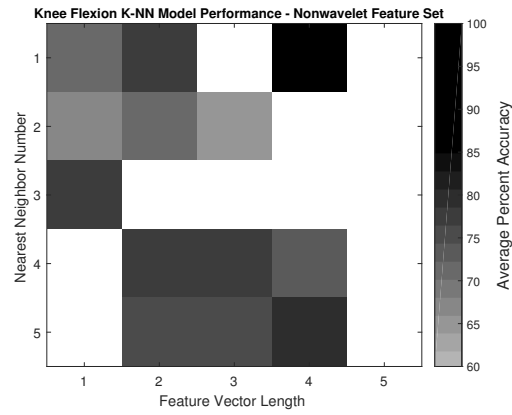
Figure 8. Model performance for knee flexion measurements. Blank spaces indicate cases that did not occur. These are results from the nonwavelet feature set.
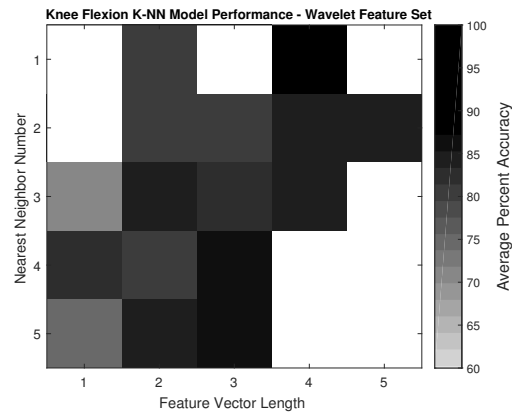
Figure 9. Model performance for knee flexion measurements. Blank spaces indicate cases that did not occur.

Figure 10. Model performance for plantar flexion measurements. Blank spaces indicate cases that did not occur. These are results from the nonwavelet feature set.
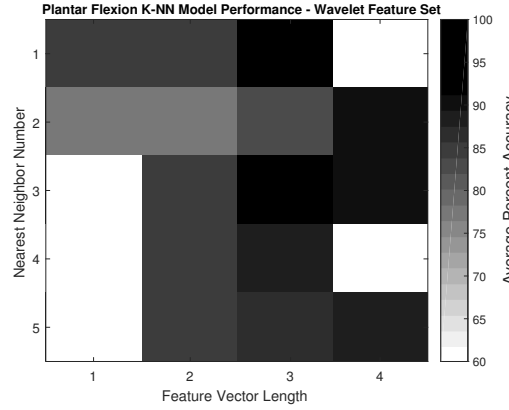
Figure 11. Model performance for plantar flexion measurements. Blank spaces indicate cases that did not occur.

3.3.2 SVM Classification Results and Model Instability

The SVM classifiers offered improved model stability as indicated by standard deviation of classification rate and the heatmaps that follow. Additionally, and particularly with the RBF case, the SVM models appear in many cases to be more robust – that is, the models offer reasonable classification rates, but with fewer features, as compared against the k-NN models.

The variation in classification accuracy was investigated in a manner otherwise identical to the k-NN model stability investigation. Curiously, the performance actually decreased for plantar flexion cases as indicated in Table 3. In general, the

performance gains were not as drastic as the k-NN case. However, the feature vector lengths only exceeded 4 in one case and the tighter bounds on classification performance deviation suggest that this approach may have greater utility developing robust models. Figures 12 and 13 show the average classification accuracy by experient and kernel type. Figure 14 through Figure 19 show the model instability for varying feature vector length for both kernel types.
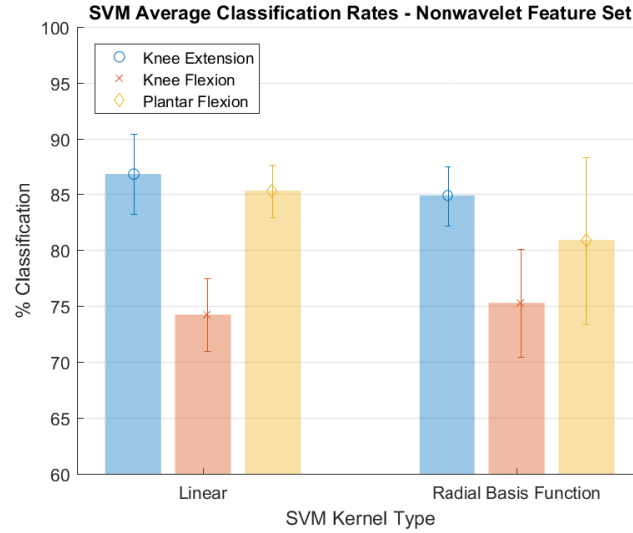
Figure 12. Average SVM classification accuracy by experiment and kernel type. Error bars indicate standard deviation of correct classifications. These results are from the nonwavelet derived features.

Table 3. Average performance changes for wavelet vs. non-wavelet derived feature sets. Positive values indicate improvement over models without wavelet-derived features. Note the substantial improvement for the knee flexion RBF case. Values indicated in the table are percents.

Kernel	Extension	Flexion	Plantar Flexion
Linear	0	3.1	-5.0
RBF	5.7	24.0	-1.2

3.4 Discussion

This approach to the bedrest study seeks to demonstrate the feasibility of longitudinal classification in the physiological domain and investigate the utility of wavelet-derived
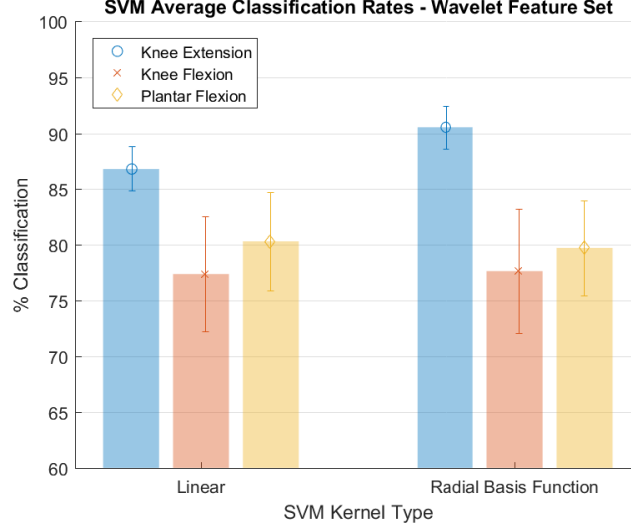
Figure 13. Average SVM classification accuracy by experiment and kernel type. Error bars indicate standard deviation of correct classifications. These results are from the wavelet-derived features.
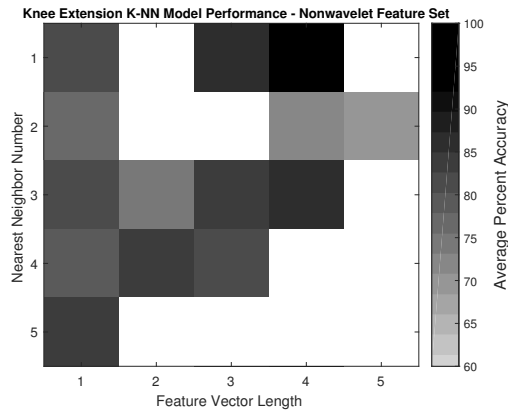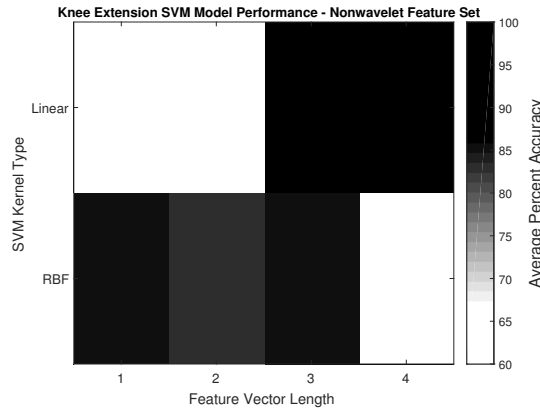

Figure 14. Model performance for knee extension measurements. Blank spaces indicate cases that did not occur. These are results from the nonwavelet feature set.

features to inform the classification algorithms. The observation size and subject-to-subject variation highlighted supervised classification efforts as an appropriate scope for this investigation. However, it is desirable that underlying physiological causes are identified to better connect this work to, for example, clinical research.

Here, k-NN and SVM classification accuracy suggests that sEMG signals can be used as an indicator of neuromuscular changes due to prolonged disuse in spite of some intrinsic challenges related to these types of longitudinal studies. Indeed, one
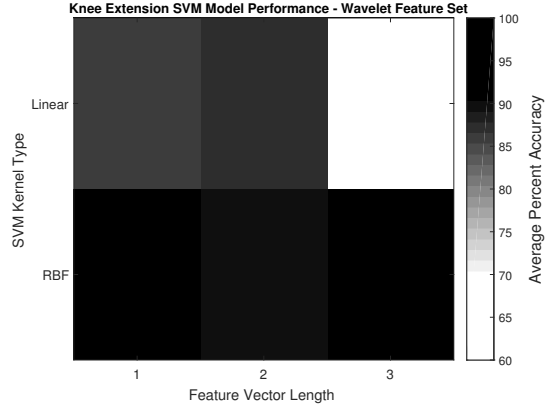
Figure 15. Model performance for knee extension measurements. Blank spaces indicate cases that did not occur.
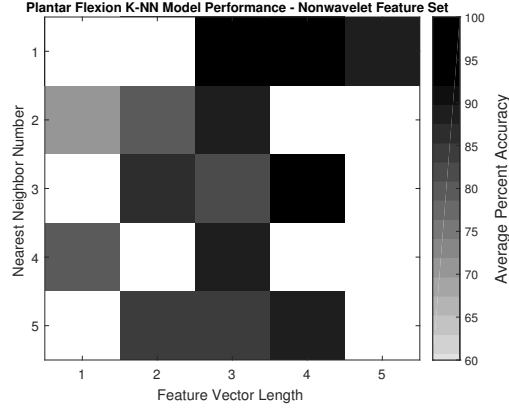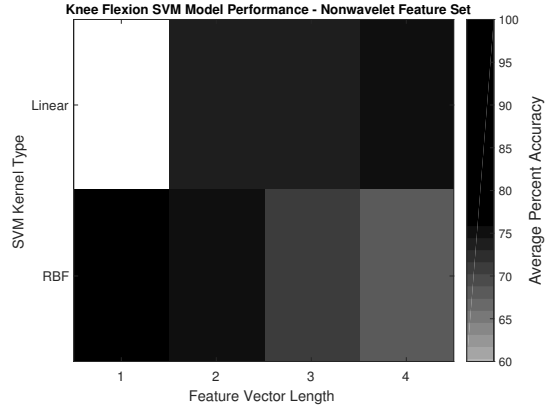

Figure 16. Model performance for knee flexion measurements. Blank spaces indicate cases that did not occur. These are results from the nonwavelet feature set.

of the issues related to the comparison of sEMG signals recorded weeks or months apart deals with the change in body composition, which may occur with some variety depending on the studied condition itself (disuse) as well as on dietary intake characteristics. As an example, subcutaneous fat reduces the amplitudes measured in sEMG so that the electrodes would be more sensitive if this layer decreased. The second issue is related to the fact that surface electrodes must be repositioned for post-bed rest measurements, thus adding more variability in terms of the relative contribution of the different motor units to the sEMG signal. However, the classification accuracy demonstrated here suggests that the effect of these two issues was – while perhaps
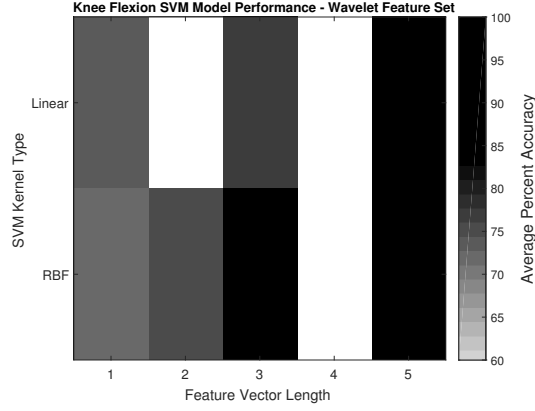
Figure 17. Model performance for knee flexion measurements. Blank spaces indicate cases that did not occur.
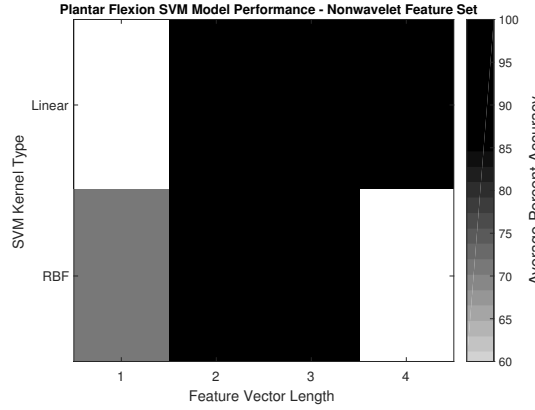
Figure 18. Model performance for plantar flexion measurements. Blank spaces indicate cases that did not occur. These are results from the nonwavelet feature set.

not negligible – still insufficient to overcome the measurement in a deleterious manner such that classification was impossible.

Generally, there are fewer features employed in the cases where the partial feature set is used. There are fewer features available overall, of course, and since the feature selection algorithm adds features based on improving classification accuracy, one concludes that the non-wavelet based features do not themselves constitute a robust feature vector, depending on the experiment and nearest-neighbor number in question. For example, a feature vector of length 1 with 5 nearest-neighbors still performs quite well for the knee extension case. However, the addition of wavelet-based fea-
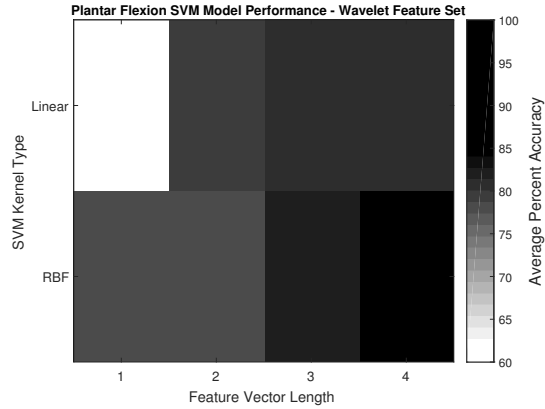
Figure 19. Model performance for plantar flexion measurements. Blank spaces indicate cases that did not occur.

tures increases accuracy in most cases (see, for example, Figure 13), and sometimes by a substantial margin (as much as 12% in one case).

While the SVM classifiers offered enhanced model stability evident from tighter bounds on error bars, and often a smaller feature set, the classification accuracy was not necessarily improved – it appeared that in some cases, for example plantar flexion under the linear kernel, classification rates actually *dropped* slightly. On the other hand, models with tighter classification rate error bounds and smaller feature vectors should be considered as more generalizable, so the SVM method holds some promise for further investigation.

Wavelet-derived features outperformed features extracted from the original conditioned signal and enabled enhanced class-state discriminatory capabilities in both classifier types. Further exploration in the use of wavelet-derived features on larger sample sizes may illuminate more precise frequency ranges that explain the variation in the data. As one traverses the wavelet cascade, the filter ranges become narrower. The bounds depend on the wavelet, sample size, and sampling rate in question, but one can broadly think of the frequency range being halved as more filters are applied, with the details coefficients representing higher frequency regions and the approximation coefficients representing lower frequency regions. Time-Frequency localization

then illuminates regions of importance in making the determination between the before and after states.

Additional efforts are needed to bridge the gap between causal mechanisms and the classification task. It is hoped that a trend of taking into account these types of approaches during the design phase of experiments becomes popular. For example, changes in MF after bed rest have been observed in other muscles such as the soleus [61] (decrease, lasting up to 10 days), the lumbo-pelvic musculature [62] (decrease, lasting up to 1 year), and an 8-week bed rest without countermeasures was demonstrated to induce a significant decrease in MF and muscle fiber conduction velocity indicated by sEMG measured from the VL during isometric knee extensions [63]. The fact that muscle atrophy is unevenly distributed in response to prolonged disuse (e.g., greater in postural than in non-postural muscles, greater in the extensors than in the other thigh muscle groups, and greater in the calf muscles than in the other muscle groups) [64–67] combined with subject-to-subject variations complicate posthoc interpretation for this study. A larger sample size and more stable classification results may assist in providing physiological interpretations for observed discriminatory capabilities of the features under study. The more stable models offered by SVM classification point to a direction in which enhanced SVM algorithms may be used to improve classification.

3.5   Bedrest Study Conclusions and Significance

This section demonstrates high classification accuracy of observations as before or after a bed rest period of 35 days and overall improvement on classification yield by utilizing wavelet-derived features. The versatility of this method can be tested by performing similar bed rest studies with time variation with a particular interest on shorter bed rest periods. Such a study would establish a method of tracking changes in the neuromuscular system, toward a quantitative means of describing

the stages of physical conditioning or deconditioning. Shortening bed rest periods and applying a similar method will establish a lower limit to the "resolution" on the measurement. Conclusive links between deconditioning evoked neuromuscular changes and mathematical features employed here are as yet open questions.

CHAPTER 4

SENSORIMOTOR ADAPTATION STUDY: THE NASA JUMP-DOWN TEST

We initiated a collaboration with NASA's Johnson Space Center (JSC) Neuroscience Laboratory near the end of 2015. The purpose of this collaboration is primarily to analyze one of their more dynamic datasets, the JDT. Experiments in this department are typically designed with simplicity of measurement in mind. The experiment may yield rather simple measurements that a researcher will then interpret in some context such as reaction times before and after spaceflight, or changes in fMRI scans, and similar physiological contexts. For a more specific example, the JDT is related to the Functional Task Test (FTT) protocol in which astronauts complete a variety of tasks at various intervals before and after spaceflight. These tasks might include drilling screws into a grid as quickly and efficiently as possible, walking a straight line with arms folded and eyes closed, running a small obstacle course, moving from a prone position to a standing position, and so on [8, 9, 68]. The measurement is often in dimensions of seconds or, in the case of the post-spaceflight field test in which an astronaut simply attempts to walk, it is some measure of "correct" steps taken. The design principle is clear: design an experiment such that the output is simple to understand and employ domain knowledge (neurophysiology and space adaptations) to characterize and deduce cause and effect.

These studies often serve to highlight an area for exploration using other tools such as fMRI [16, 69]. The present case fits into this body of work by investigating another method of examining a more complex experiment whose variables are not readily analyzed.

4.1 Motivation and Background

It is well-known that astronauts undergoing exposure to microgravity experience sensorimotor disorder [70], and upon return to a planetary surface this change in physiological state makes balance and coordination difficult. This physiological deterioration presents a particular risk for any future Mars mission in which astronauts must re-immerse themselves in a gravitational well after nearly a year of spaceflight in which the sensorimotor systems lack stimulus. Toward that end, algorithmic prediction of sensorimotor performance based on sEMG signals offers potential for a non-invasive, physically-compact (or at least light on equipment) method for assessing performance.

4.2 The Jump-down Test

The experiment is broad and has multiple components revolving around the JDT. There are bedrest and spaceflight cases. The spaceflight cases involve either the space shuttle (SS), which consists of short-term missions of varying length, and the International Space Station (ISS) consisting of between 3- and 6-month long missions. The bedrest cases consist of four subgroups – control, exercise, exercise with testosterone, and a flywheel exercise. The datasets are all anonymized so that only the time, with respect to launch and return (see Table 4), and the sEMG signals themselves are known.

At each point marked in the table, subjects are brought to JSC in Houston, Texas to perform the JDT. The JDT is performed by instructing a subject to jump, flat-footed with feet together (i.e., non-sequential) off a 30 cm tall platform equipped with

Table 4. The relationship between the labels given to flight or bedrest cases and their actual relationship to the Launch (L) date and Return (R) date.

Day Label	Bedrest Description	Flight Description
A	L - 12	L - 180
B	L - 6	L - 90
C	L - 3	L - 60
D	R + 0	R + 0
E	R + 1	R + 1
F	R + 6	R + 6
G	R + 12	R + 30

a load cell for each foot onto a force plate on the ground (see Figure 20). Electrodes (Bagnoli-8 EMG amplifier system, Delsys Inc.) are attached to the TA and MG (refer to Figure 20(b)) of both limbs of the subject. In this work, the abbreviation GM (Gasctrocnemius Medialis) is often used instead of MG (Medial Gastrocnemius) due to the different standards of communication in this and the work of Chapter 3. A sampling rate, F_s, of 4000 Hz is used.

The full output then includes signals from the two force plates, the four electrodes, and the three axes (X, Y, Z) of the force plate, as shown in Figure 21.

4.3   Experimental Parameters

Of the three portions of the JDT shown in Figure 21, the analytically important parameters are derived from the sEMG signals and the force plate outputs. The load cells are used to determine time of take off from the jump platform, marked by the time at which the last foot leaves the platform. In most cases the time differences are small as the subject is instructed to attempt the jump with both feet at the same time.

The sEMG electrodes are recording myogenic activity for the duration of the experiment. The measurements prior to takeoff and after achievement of stable posture are not used. The remaining measurement has been divided into the in-air portion

(a) Researcher standing on platform for JDT.
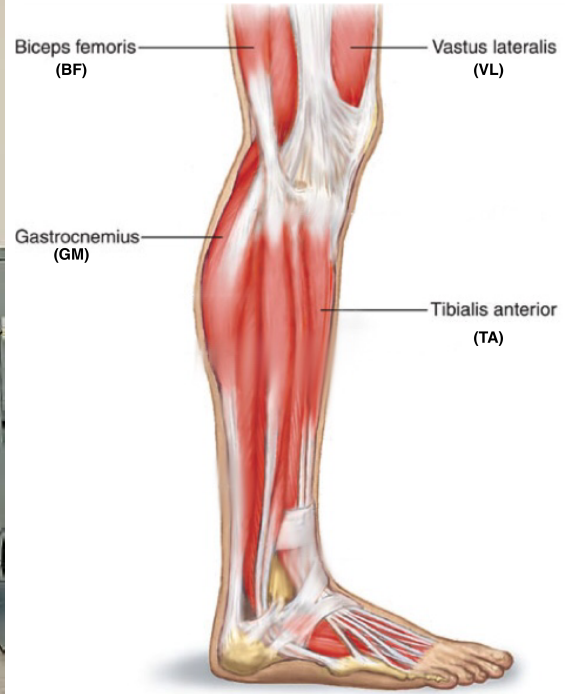
(b) Lower limb muscles.

Figure 20. Feet are situated on two independent load cells prior to jumping down onto a single force plate (a). Figure 20(b) shows the lower limb with muscles referenced throughout this dissertation labeled with bolded, parenthetical abbreviations [71].

and the touchdown portion.

Figure 22 demonstrates the nomenclature of the overall signal while Figure 23 illustrates some physiological parameters within the sEMG measurements of the JDT. Note that these are for the MG muscle in particular. The antagonist muscle, the TA, will have different behaviors.

The Preparatory Latency (PL) marks the time between leaving the platform (initiating the jump) and beginning the phase of the jump in which the subject initatiates anticipatory postural adjustments preparing for the landing. It is assumed that a subject whose sensation of movement and position is obfuscated by effects of spaceflight
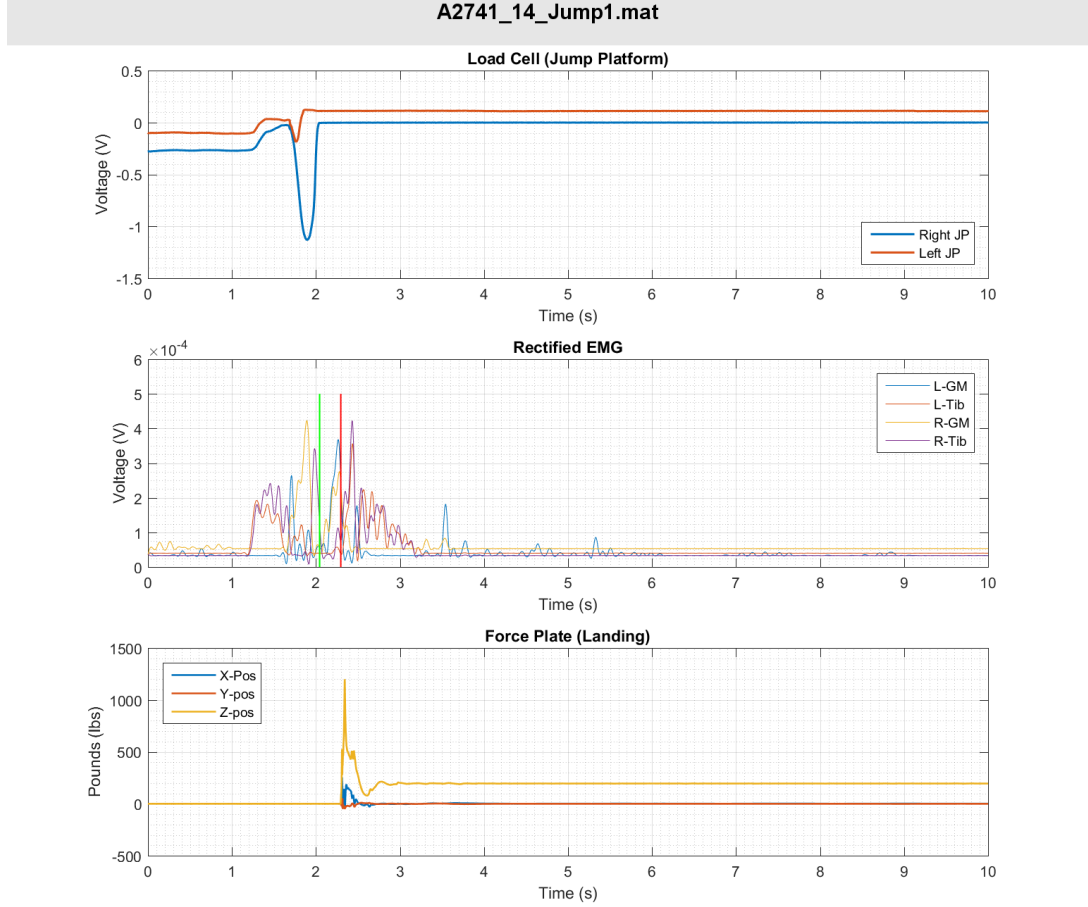
Figure 21. Overview of signals from JDT experiment. Top: output from right and left load cells. Middle: sEMG envelope for all four muscles with takeoff (green) and touchdown (red) marked. Bottom: Force plate output from X, Y, and Z (vertical) directions.

or bedrest would show a difference in either the latency itself or the quality of the adjustment or both.

The Monosynaptic Reflex (MSR) is a reflex intrinsic to the transmission of a signal from sensor neuron to motor neuron via the spinal cord. The knee-jerk response elicited by a tap on the patellar tendon is a commonly appreciated example. Reschke et al. (2009) found that the MSR latency actually increased after bedrest, demonstrating that head-down bedrest has some functionality as an analog for spaceflight [9].
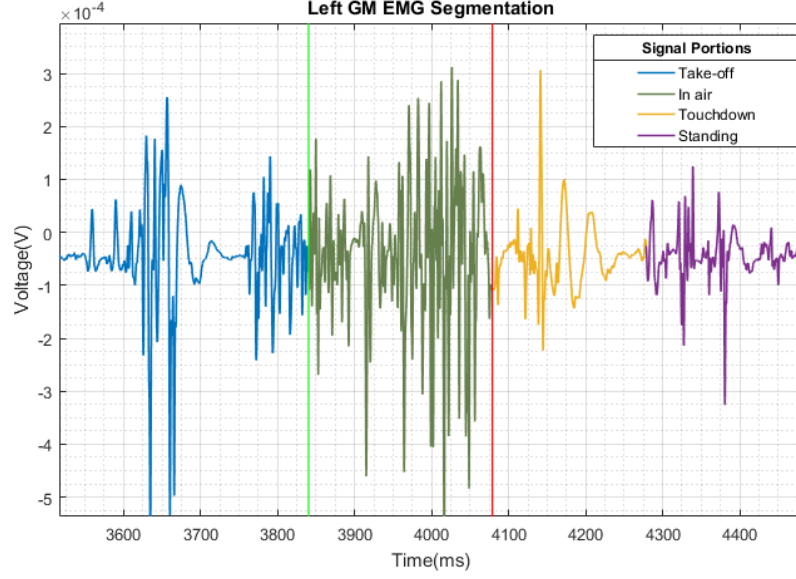
Figure 22. Segmented example of sEMG signal with labeled portions, from the JDT experiment. Green vertical line indicates takeoff (point at which last foot leaves the platform) while the red vertical line indicates landing (point at which first force is registered on the force plate).
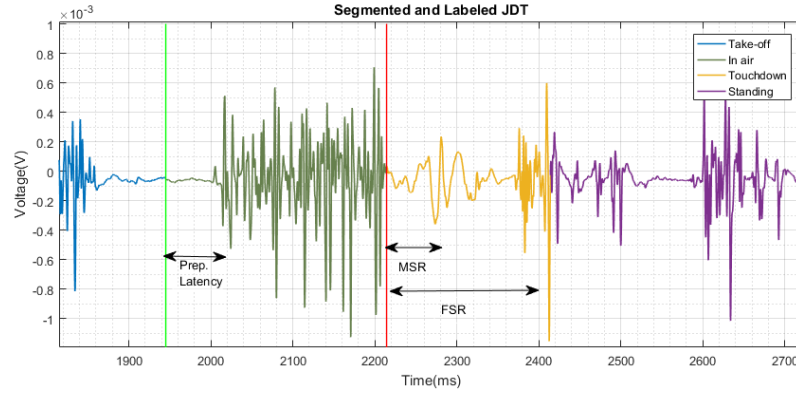
Figure 23. Illustration of physiological parameters from sample of JDT experiment.

The Functional Stretch Reflex (FSR) follows between 200 and 300 ms after touchdown and is responsible for final stabilization of posture after landing. The FSR takes places after tendon-stretch induced MSR, but before there is time for voluntary contractions to take place. Changes in FSR behavior could signal risk of injury as mechanisms to protect against injury are altered [72].

Now considering the force plate, the Z-force vector output is used to mark the

moment of touchdown and to characterize the biomechanics of the landing. Similar to the reasoning demonstrated by Ortega (2010), we divide the force output curve into multiple parts, each as a quantitative description of the biomechanical outcomes [73].
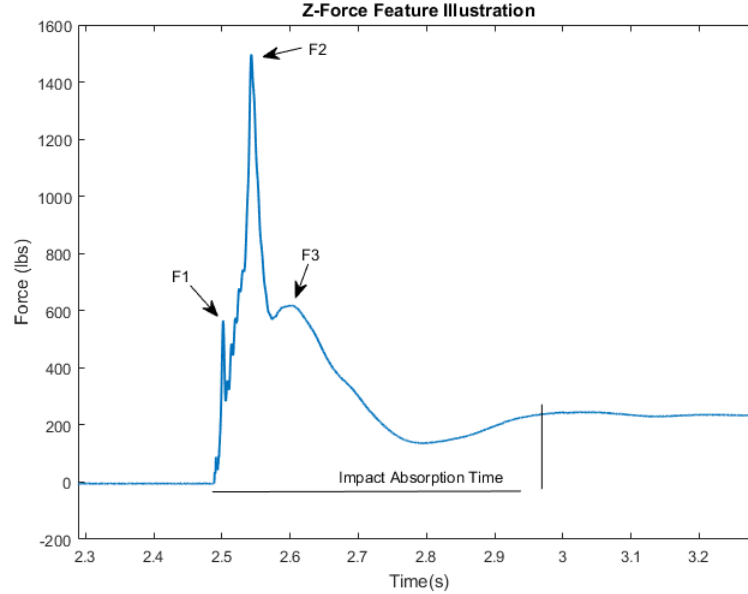
Figure 24. Example of force curve output from a JDT. A description of the biomechanics can be reasoned from these curves.

The annotations on Figure 24 describe points of interest on the force curve. The first point, F1, is the point at which the front portion of the foot (toes) begin the impact. It is from here that touchdown is marked and the in-air portion of the sEMG signals are concluded. As the landing continues, the subject shifts from toe-contact to heel-contact. The moment immediately after toe-contact, in which the transfer from toe to heel begins, is where the apparent force (impulse-momentum transfer) drops off and begins to rise to a peak. This is the point at which impulse-momentum transfer takes place most rapidly, registering a force far greater than the weight of the subject. The forces F1 and F2 together can give a sense of the quality of the initial landing – a flat-footed landing is one in which the ratio $\frac{F2}{F1}$ is higher.

Continuing along the force curve, F3 is representative of the adjustment of the trunk of the body as the subject stands upright following the landing, and the im-

pact absorption time (IAT) measures the total time to postural stability following touchdown. For each force measurement there is also a corresponding duration measurement (e.g., W1) which is obtained by estimating the full-width-at-half-maximum (FWHM) duration of the force peak.

When examining these curves, it is helpful to keep in mind the impulse-momentum theorem for an object with constant mass.

$$F\Delta t = m\Delta v \tag{8}$$

Where F indicates force, t indicates time, and v indicates velocity. As momentum is transferred from the subject to the force plate, the force curve features can be interpreted in this framework as illustrated in previous paragraphs.

$$F = m\frac{\Delta v}{\Delta T} \tag{9}$$

In this re-arranged form (Equation 9), for example, one can intuit that a subject landing onto a force plate must negotiate, via biomechanical adjustments before landing (e.g., APA) and after landing, the reduction of velocity to zero. It is these biomechanical adjustments that are ultimately responsible for altering the ΔT for both the overall impact absorption and for the FWHM estimated peak-time durations. As ΔT decreases, the momentum is transferred more rapidly and the force plate will register a larger value, as seen most clearly in the toe-to-heel transition resulting in the peak F2.

4.3.1   Signal Characteristics

The sEMG signals are each sampled at 4000 Hz and contain abrupt transitions making them unsuitable for analysis by Fourier Transform. In particular, the signals do not

meet the conditions for (weak-sense) stationarity,

$$E[x(t)] = m_x(t) = m_x(t + \tau) \text{ for all } \tau \in \mathbb{R} \qquad (10)$$

where E is the expectation value of some time-dependent (t) signal x and m_x is the time-dependent mean signal. Put plainly, the mean of the signal is not reliably constant. Whereas for a signal to be stationary in the weak sense, its mean and autocovariance must be independent of time. It is in this domain that WT approaches are well-suited for analysis. The line of inquiry presented in Chapter 3 and the fact outlined here then suggest the use of the WT.

In particular, the Daubechies 4 (or 'db4' in `MATLAB`) wavelet was ultimately decided upon based on previous work [11] and empirical observation with sEMG signals in the bedrest and JDT contexts.

4.4 Feature Engineering

While the force curves present obvious features for informing algorithms, the sEMG signals are more ambigious. Feature extraction (and subsequent feature engineering or transformation) is part trial-and-error, part domain knowledge, and part signal processing. In analyzing the JDT, we use a mix of domain-common approaches and wavelet-based approaches. In any case, we test a newly constructed, domain specific feature for interpretability purposes.

The thrust of any analysis here is to demonstrate some quantifiable difference in JDT performance over time to make observations about the effects of spaceflight on human physiology. In particular, the concern is recovery from the observed issues with the sensorimotor system – i.e., how quickly does one typically recover, can we glean any insights from this performance to how spaceflight changes the sensorimotor region of the nervous system, and, relevant to the proposal at hand, can we establish any algorithmic approaches to these problems?

Figure 25 shows green lines obtained from the load cells in the JDT platform to indicate the point at which the last foot left the platform. This is the initiation of the jump and the "in-air" portion. The red line is obtained from the point at which the force plate just begins to register a rapid increase in z-force (that is, when toes touch the plate). The in-air portion is naturally defined as the time between these two points. The landing portion includes an arbitrary-but-informed 300 ms after touchdown as an attempt to capture the entire signal consisting of a sEMG corresponding to ground reaction forces. The bottom plot is the CWT of the above signal. We define the following novel features based on the scalogram energies, C.

$$Airsum = \sum_{t=takeoff}^{touchdown} \sum_{scale=1}^{max} C_{t,scale} \tag{11}$$

$$Landsum = \sum_{t=touchdown}^{touchdown+300ms} \sum_{scale=1}^{max} C_{t,scale} \tag{12}$$

$$Bilateral\ Ratio = \frac{L_{Airsum} + R_{Airsum}}{L_{Landsum} + R_{Landsum}} \tag{13}$$

Here, the motivation is to compare the sEMG activity before and after landing under the working assumption that after spaceflight, the ability to anticipate the landing and therefore adapt a feedforward strategy is altered. The assumption is that the ratios of signal energy before and after touchdown will be different as compared to the baseline days (any day before R+0). The benefit of the CWT approach is that, should relevant sEMG activity be restricted to high or low frequency (low or high scale, respectively), we can alter the limits of the second summation in the Airsum and Landsum cases to reflect that.

The wavelet approach may additionally be recast in terms of iEMG. We would then write the sums as,
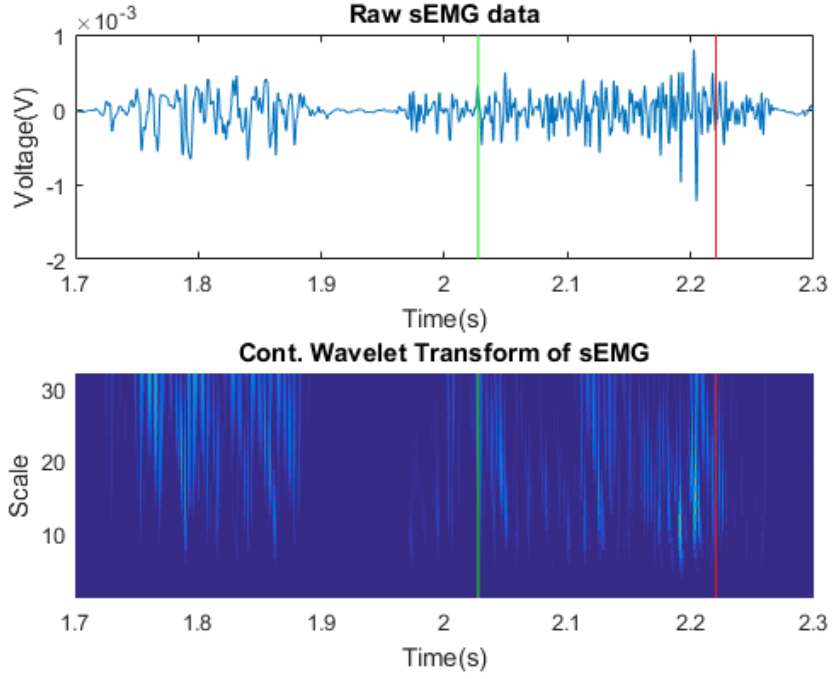
Figure 25. An illustration of the equations describing signal energy before and after landing. Top: Raw sEMG trace. Bottom: Continuous Wavelet Transform (CWT) of the top signal. Signal energy contribution is added up before and after landing to construct new features.

$$Airsum_{iEMG} = \sum_{t=takeoff}^{touchdown} |X_t| \qquad (14)$$

$$Landsum_{iEMG} = \sum_{t=touchdown}^{touchdown+300ms} |X_t| \qquad (15)$$

In the case of Equations 14 and 15, the cumulative, integrated iEMG is calculated by adding the value of the sEMG signal at each moment in time across whatever boundary is chosen. When plotted, this more traditional measure gives a sense of changes by communicating changes in the rate of accumulation in the sEMG signal. Likewise, the wavelet scalogram can also indicate these changes.

Figure 26 demonstrates the comparable utility of the wavelet by showing its performance alongside more traditional measures. Figure 26 also shows how the biome-

chanics might be teased out of the sEMG signals. What follows is a brief walk-through of the data presented in the figure.

First, examine the more easily interpreted cumulative iEMG signals in the right-most column. The TA and the GM constitute an agonist/antagonist pair. They are responsible for dorsiflexion and plantarflexion about the ankle respectively. Subjects stiffen their feet in anticipation of landing and this is seen prior to the red vertical line by an increasing slope for the RMG and LMG signals. Shortly after landing, the RMG and LMG signals taper off (black arrows) while the LTA and RTA experience a burst of activity (orange rectangles and orange circles), indicating a shift from plantarflexion to dorsiflexion, and thus concluding the process of stabilizing the landing as the LTA and RTA also taper off shortly after the burst is initated. This is also seen in the CWT case where the shift from GM activity to TA activity is seen before and after the red vertical line in the form of bright regions in a wavelet scalogram. Here, the difference lies in the extra dimension of scale. See Appendix 6.2 for more information on the wavelets and their scalograms.
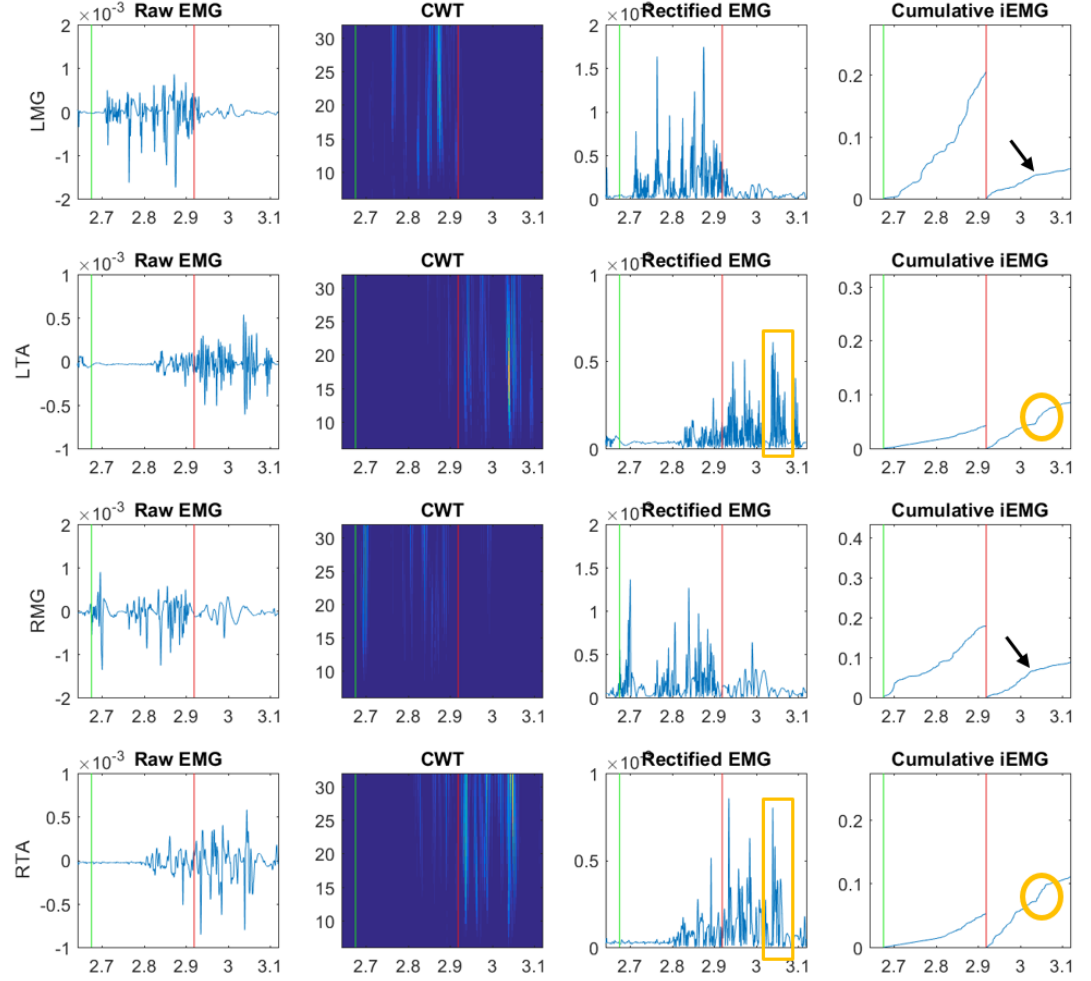
Figure 26. The left-most column contains raw sEMG signals for each of the four muscles measured during the JDT experiment. The next column over contains CWT scalograms showing the percentage of energy contribution at various times and scale values. The right-most column shows iEMG calculated over the take-off and touch-down boundaries. The green and red vertical lines indicate take-off and touchdown. All horizontal axes have the same scale.

In personal communications during collaborative visits with colleagues at JSC and in the work initated by Jones and Watt (1971) [72, 74] it has been suggested that the FSR tends to lie between 200 and 300 ms in falling (or JDT) subjects, with little variation outside this window. Early attempts at analyzing this problem had

the cutoff set at 200 ms prior to relocating it at 300 ms. The bilateral ratio – whether calculated by wavelet energy sums or iEMG – is resilient to effects of signal noise due to extension, as the signal energy contributions between landing and activation of the FSR are essentially nil in comparison. There are however other confounding factors that when taken into consideration with the idea of anticipatory postural adjustment, suggest to err on the side of a longer cut off (e.g., 300 ms).

As suggested from the results of the work illustrated in Chapter 3, flexible algorithmic approaches may be well-suited toward navigating experiments with a multitude of confounding factors. In the case of the JDT, pinning down the precise location of MSR, FSR, and PL can be difficult due to issues arising from the design of the JDT experiment.

In particular, difficulties arise from the fact that there are two independent load cells measuring lift-off of each foot whereas touchdown is measured from a single force-plate that both feed land on. The JDT protocol has researchers instructing the subject to jump off the platform with both feet together and onto the force plate. This instruction is not particularly natural and, strictly speaking, happens rarely. In fact, the subjects typically lead slightly with one foot or the other, and this lead-time can vary considerably. Consult the top pane of Figure 21 for an example of leading limbs.

The practical effect of this is that the starting time of the JDT marked by a green line may be close to the PL read for a particular muscle. Likewise, the red line, the point of touchdown, may be misleading with respect to a particular muscle and make estimation of the MSR and FSR locations difficult. For this reason, estimations of these factors is not included in the model at this time. See Section 6.2 for more details.

4.5 Exploratory Characteristics

With novel features defined and two primary aspects of the JDT experiment described, we now turn to some exploratory results in order to generate hypotheses, describe the nature of the experiment, and investigate the quantitative consequences of these novel features.

4.5.1 Parametric vs. Non-parametric Methods

First, we present a high-level view of the distribution of the bilateral ratios for the TA and MG muscles calculated by way of wavelet or iEMG statistics.

Each of the bilateral ratios, whose value distributions are presented in Section 4.5.2, can only loosely be approximated by probability distributions with fixed parameters. That is, we cannot say that modeling these values as gaussian or lognormal distributions, for example, is reasonable. Take the Cullen and Frey graph of Figure 27. Here, a range of statistical moments are presented and regions that various statistical distributions occupy are compared against the statistical moments of the observed data [75, 76]. In this case, it is difficult to support any claims of the particular statistical distribution for our engineered feature. Similar reasoning applies for other engineered features. Because we cannot assume knowledge of these statistical distributions, non-parametric methods are a good fit for examining the data.

Cullen and Frey graph
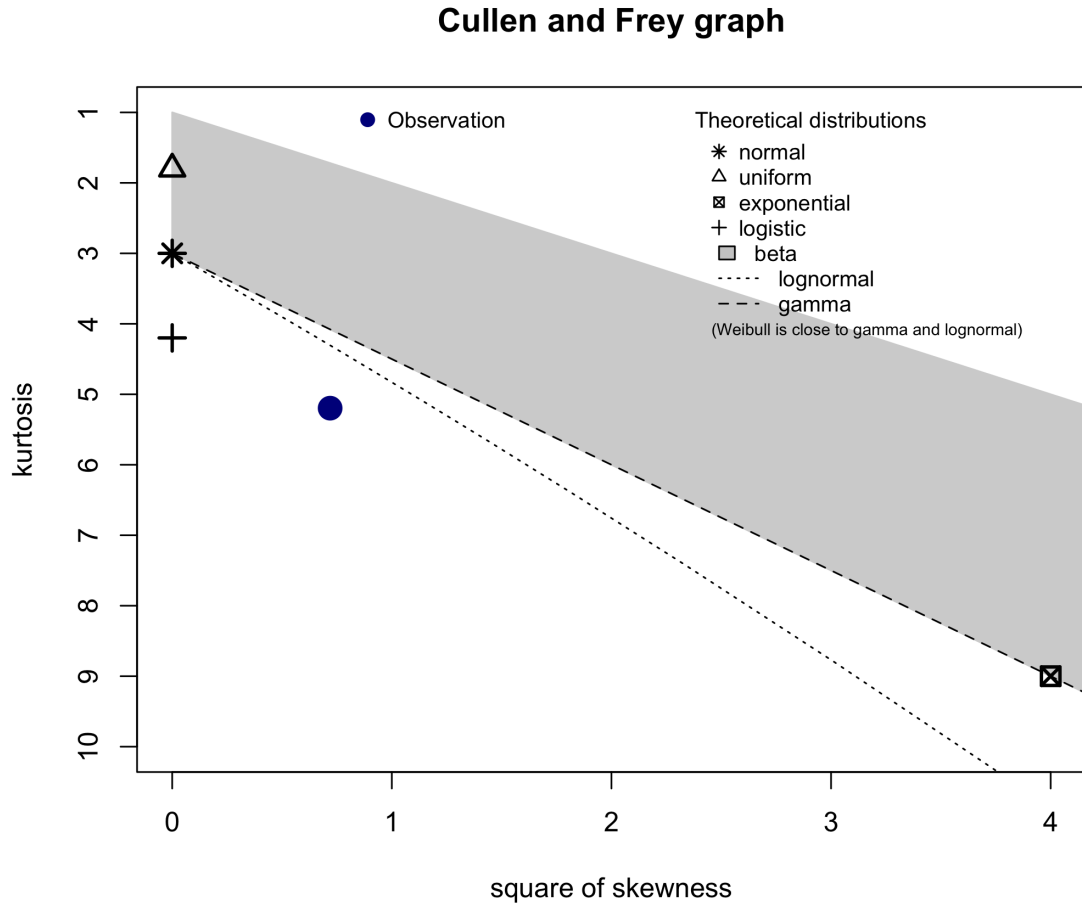
Figure 27. The closest statistical distribution for the iEMG-derived MG bilateral ratio (calculated on JDT data) appears to be log-normal – but it is yet far off the mark.

Figure 28 illustrates further comparison to the log-normal distribution, demonstrating the ill-fitting comparison. In each subfigure, the empirical distribution is plotted against its theoretical comparison, in their various representations.
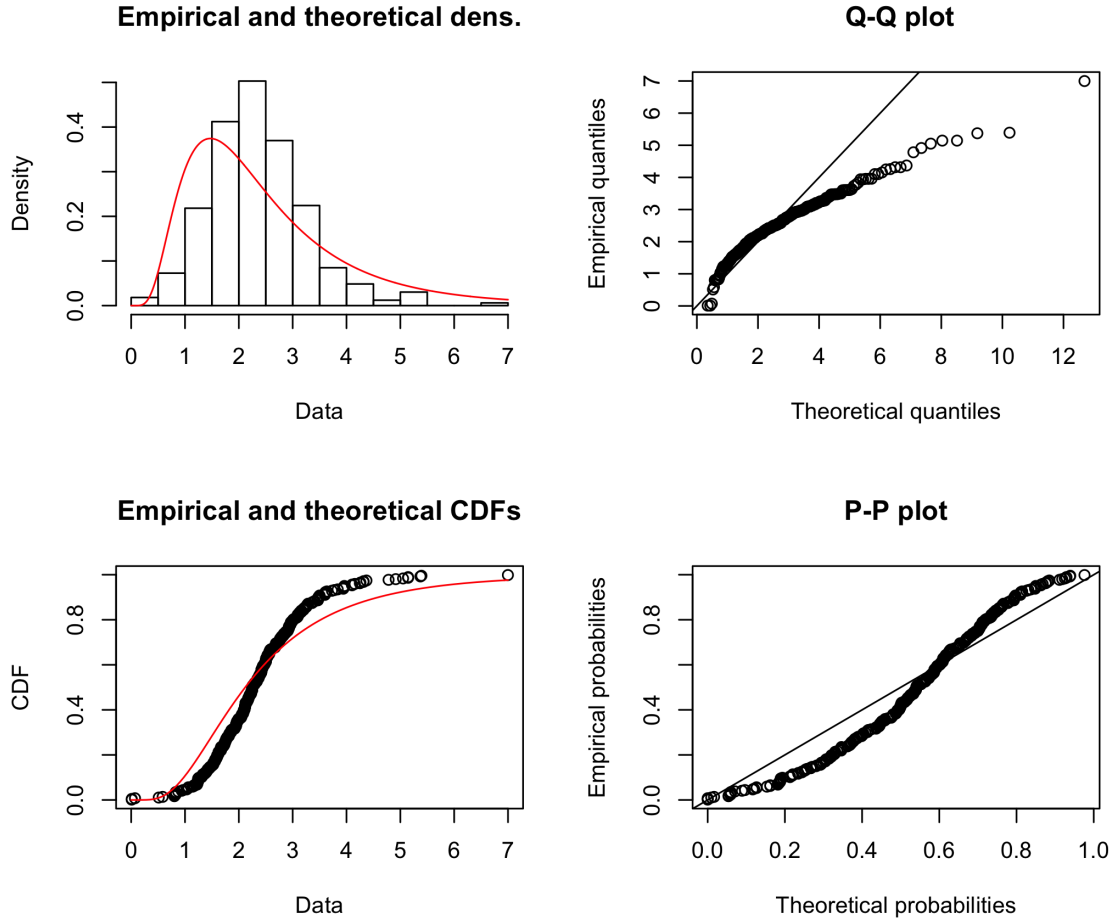
Figure 28. Horizontal "Data" axes refer to the ratio value itself. The theoretical plotted distribution is that of the log-normal.

4.5.2 Bilateral Ratios

Here, the feature behavior of the subject population is investigated. Violin plots of the novel features defined by Equations 11 through 15 are shown to highlight the distribution of JDT behavior. The violin plots are produced by performing a kernal density calculation on the composite feature, the bilateral ratio, to produce a probability density function. In short, the violin plots communicate a sense of the distribution of all the data, and their curves are formed by mirroring the probability density function (or more crudely, the histogram).

In all violin plots below, the median is indicated by a red diamond and in some cases extreme outliers are truncated. Note that the scale does change from plot to plot. Note also, that in rare cases there are methodological outliers in which the subject actually performs more than three jumps. These cases are not considered in any analysis. Recall also that, as ISS measurements for day D are not possible, the transition is immediately from day C to day E.
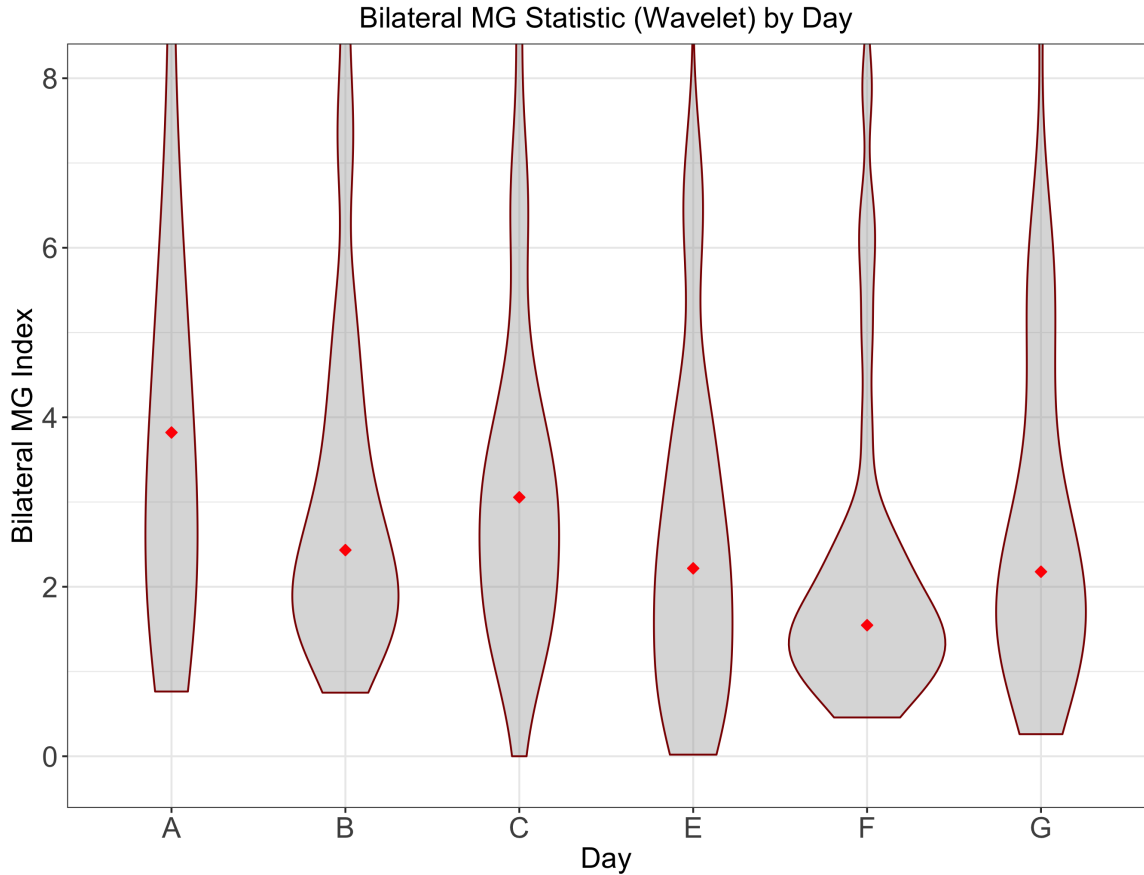
Figure 29. Distribution of Wavelet-derived MG Bilateral Ratio by Day. The MG statistic tends to occupy a wider range of values.

The violin plot of Figure 29 shows stronger consistency (tighter grouping) after Day A. Day A is expected to have a higher variance as it is the first experience of the JDT that subjects have. Day's B and C are intended to establish a baseline.
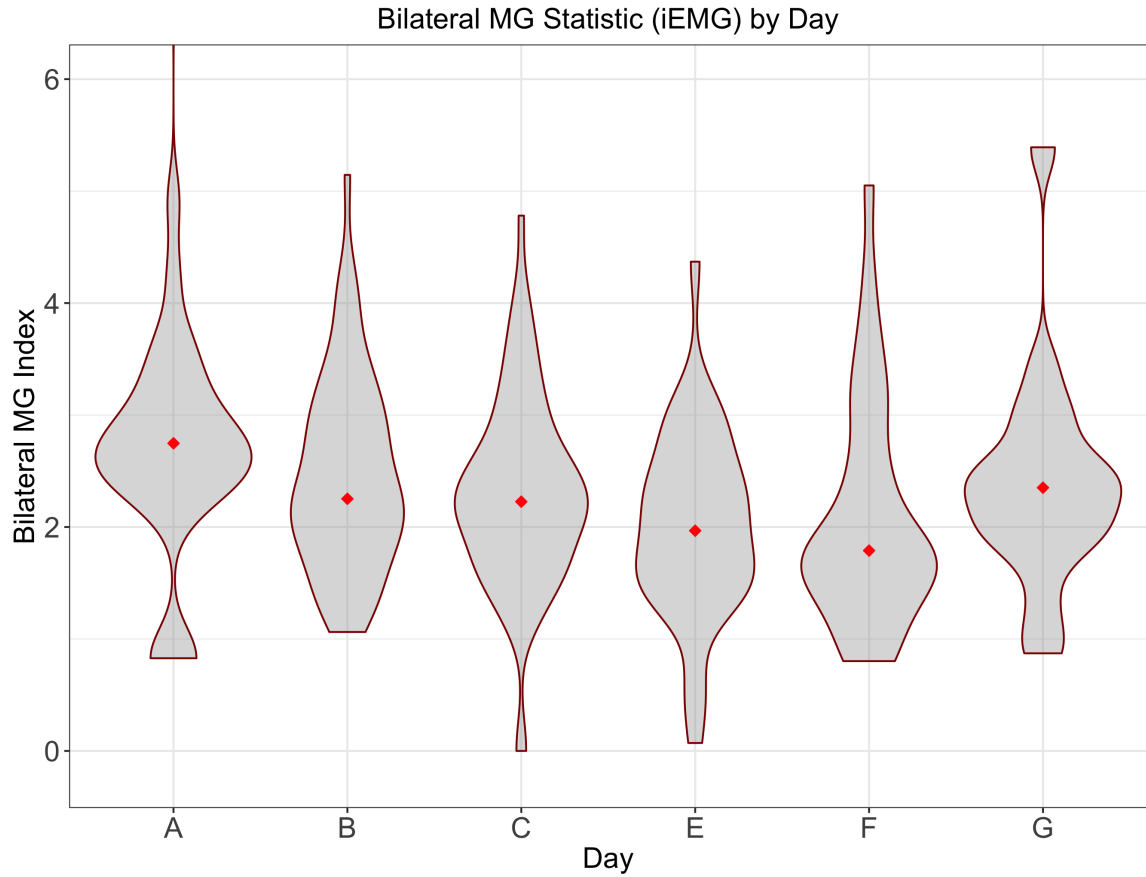
Figure 30. Distribution of iEMG-derived Bilateral Ratio by Day. The iEMG-derived statistic exhibits similar behavior as the wavelet-derived statistic here, but has tighter grouping.

Figure 30 illustrates a common tendency: while the behavior of the wavelet-derived statistics and the iEMG-derived statistics is quite similar, the grouping of the iEMG-derived statistics tend to be tighter.
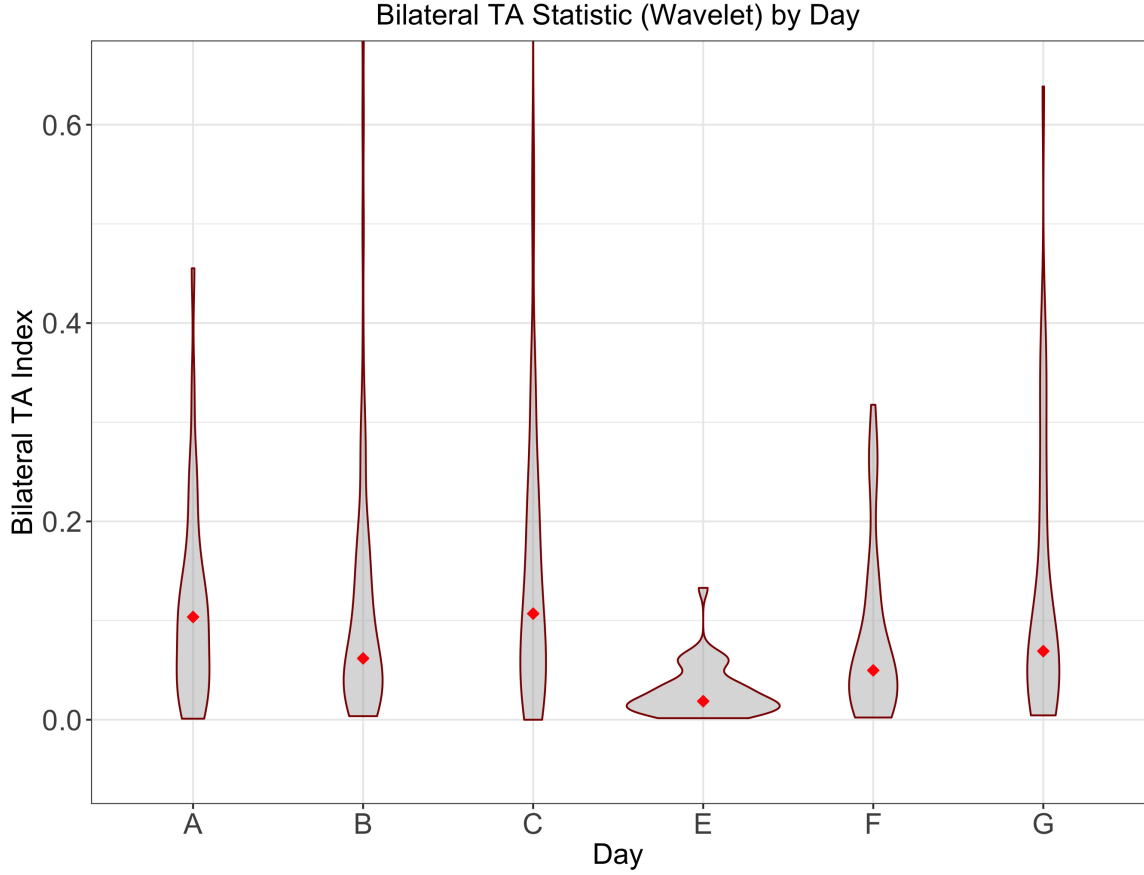
Figure 31. Distribution of Wavelet-derived TA Bilateral Ratio by Day.

The TA statistics, as illustrated in Figure 31, in general occupy a smaller range of values than the MG statistics. The difference in *Landsum* and *Airsum* values is smaller, so their ratios will be too. Additionally, as the TA exhibits its greatest activity after touchdown (see Figure 26), the ratio will exhibit behavior inverse to that of the MG. Therefore, the ratio will often be less than 1.
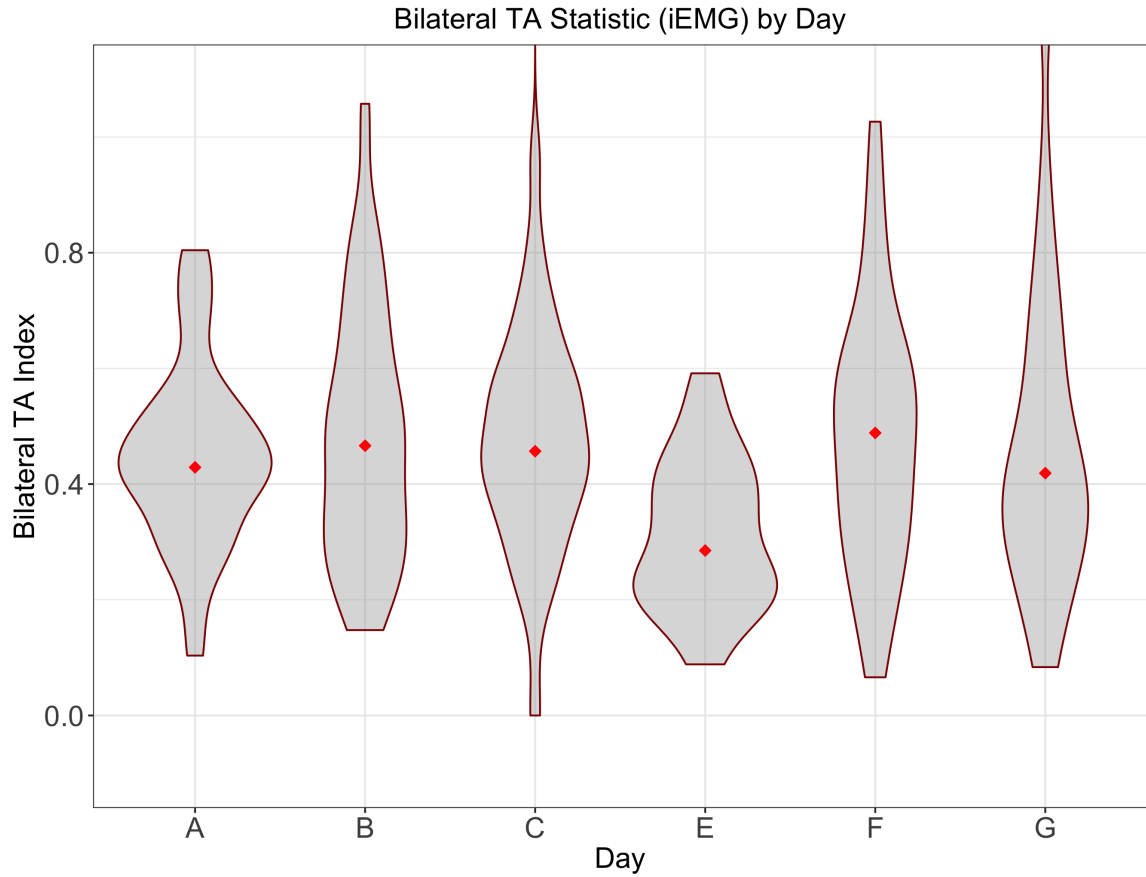
Figure 32. Distribution of iEMG-derived TA Bilateral Ratio by Day. As in the MG case, the iEMG-derived bilateral ratio exhibits tighter grouping.

Figure 32 exhibits similar behavior as before with its tighter grouping and ratios less than unity.
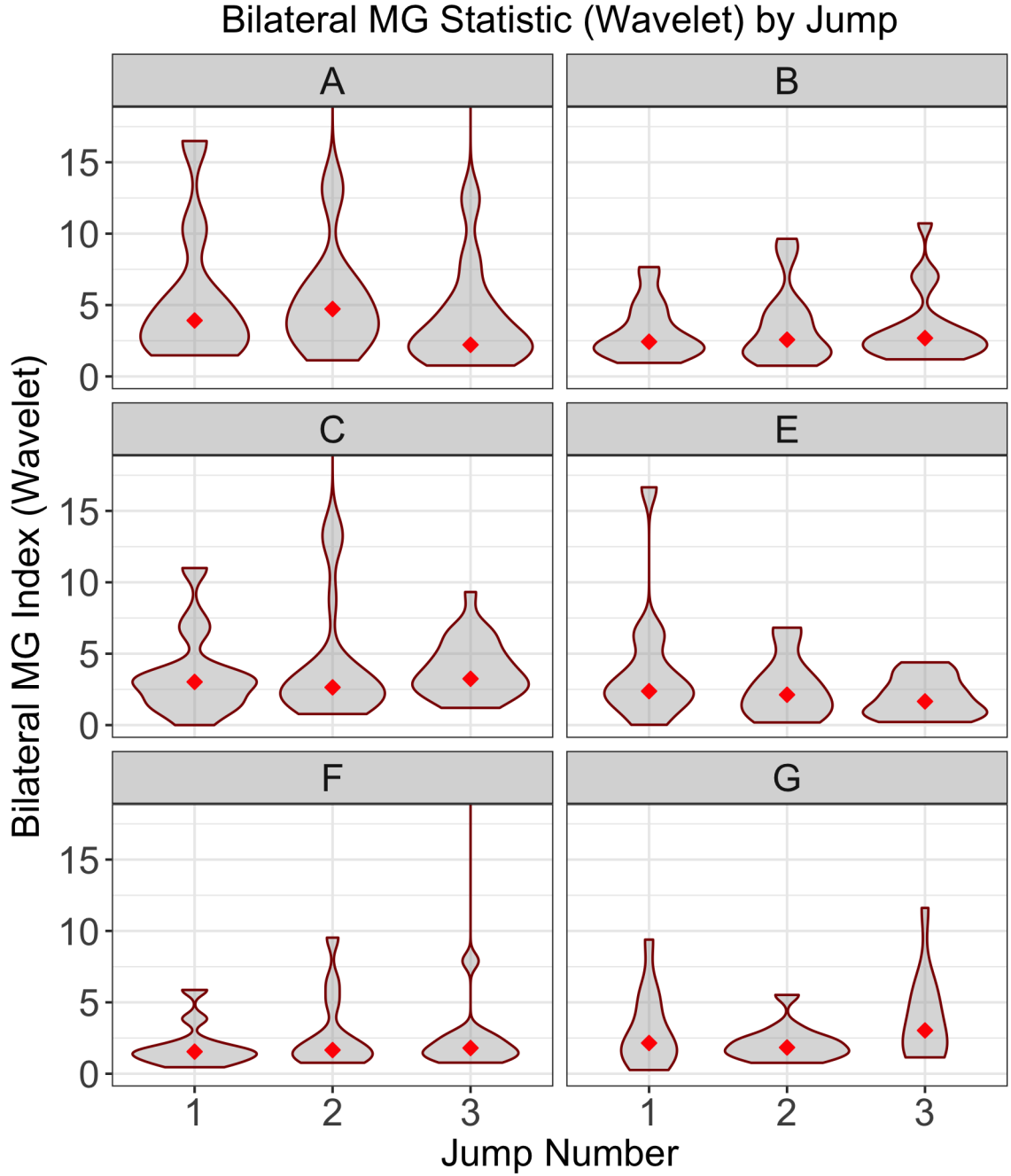
Figure 33. Distribution of Wavelet-derived MG Bilateral Ratio by Jump. The center row highlights results before and immediately after spaceflight.

Here, the breakdown by jump and by day is presented. Figure 33 shows the distribution of the wavelet-derived MG bilateral ratio by jump for each day. The behavior of loose grouping and larger range of values for wavelet and MG statistics

holds in this view as well. The center row of this chart (and other by-jump figures) highlights the difference between the final trial day of experiments before spaceflight, and the first day of experiments after spaceflight.
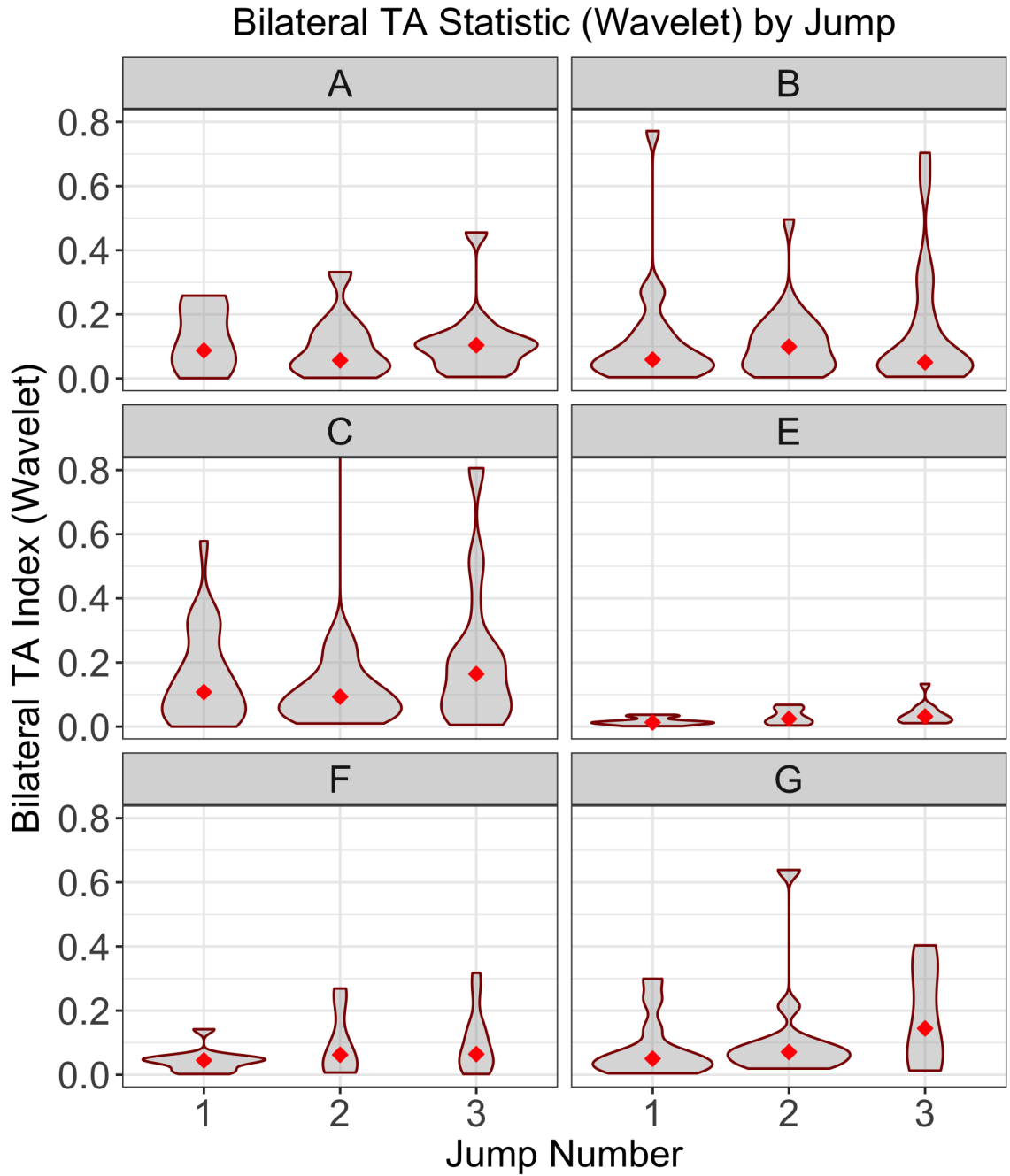
Figure 34. Distribution of Wavelet-derived TA Bilateral Ratio by Jump. The center row highlights results before and immediately after spaceflight.

Figure 34 shows particularly interesting, and potentially insightful behavior. The TA bilateral ratio practically collapses post-spaceflight for the entire subject population. Further, post-spaceflight experiment days F and G are tightly grouped, with these values becoming more loosely distributed as time goes on and more jumps are performed. This could reflect that the subject population exhibits a range of adaptation (or readaptation) to the JDT after reintroduction to the gravitational well.
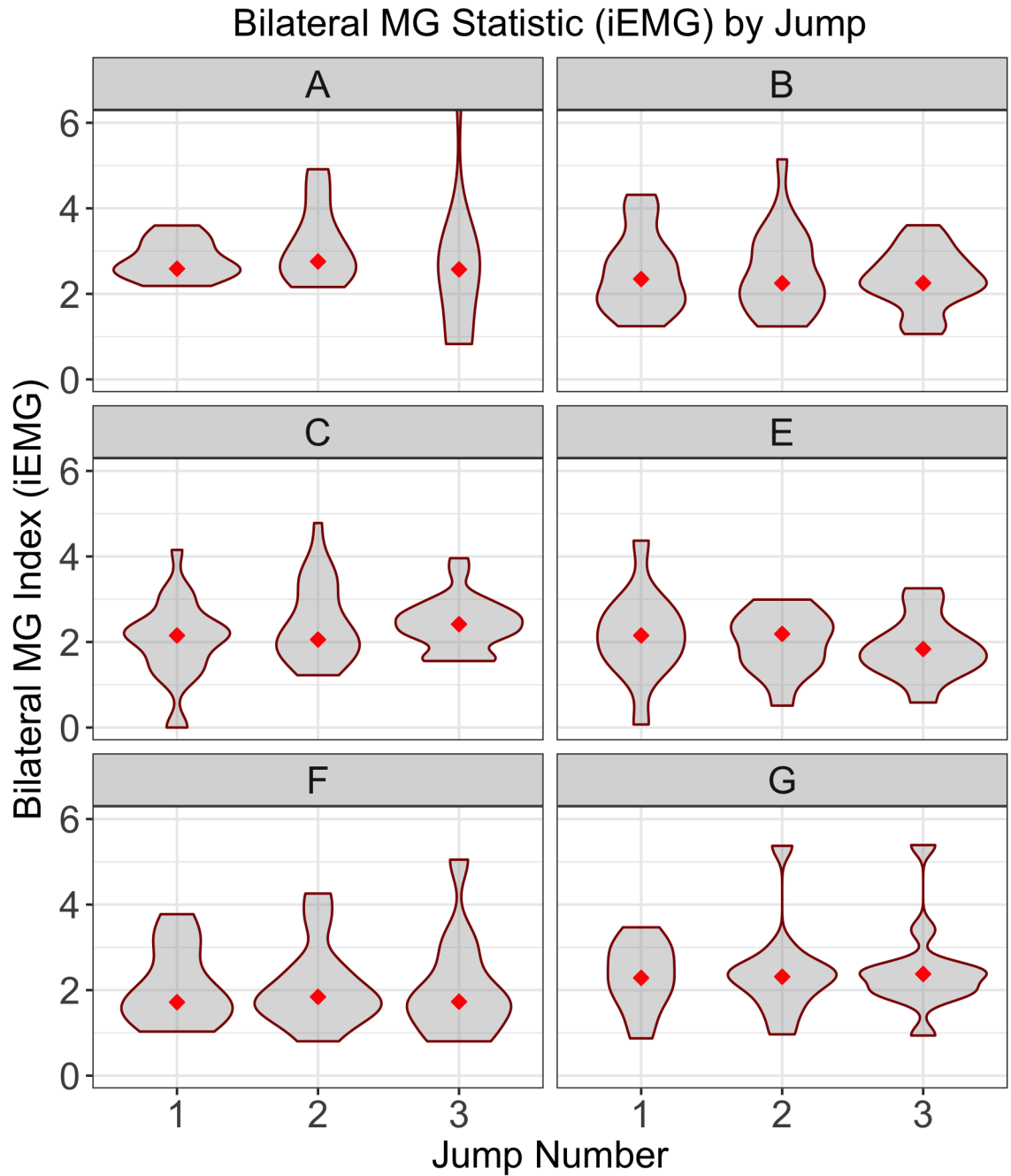
Figure 35. Distribution of iEMG-derived MG Bilateral Ratio by Jump. The center row highlights results before and immediately after spaceflight.

Figure 35 breaks down the distributions shown in Figure 30 by day. In this view, it is difficult to judge, by eye, any trends that are more apparent in the agglomerated view of Figure 30. The contrast in these two figures serve to highlight the inter-subject

variability as well as the intra-subject variability, when looking at the breakdown by jump.
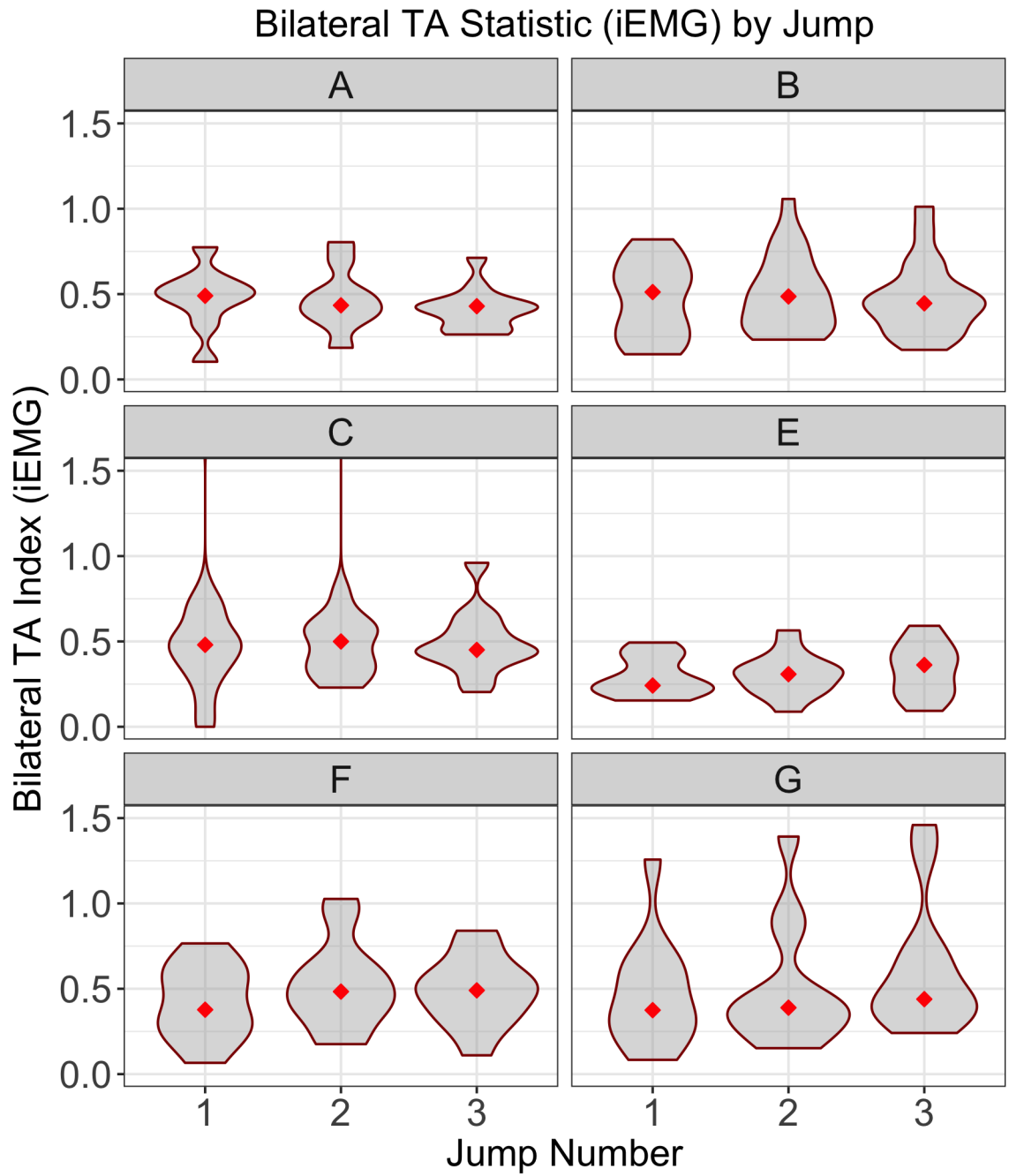
Figure 36. Distribution of iEMG-derived TA Bilateral Ratio by Jump. The center row highlights results before and immediately after spaceflight.

Finally, Figure 36 shows a similar trend in that iEMG statistics tend to have

tighter grouping and the TA tends to be less than unity.

4.5.3 Force Characteristics

In Figure 24, the force features were introduced. On their own, F1, F2, and F3 exhibit variability that can be attributed, in part, to the weight of individual subjects. However, as reasoned in Section 4.3, it is rather the ratios that are of interest. In particular, the ratio of $\frac{F2}{F1}$, presented in Figure 37, is interesting as it shows a distinctive post-spaceflight recovery.
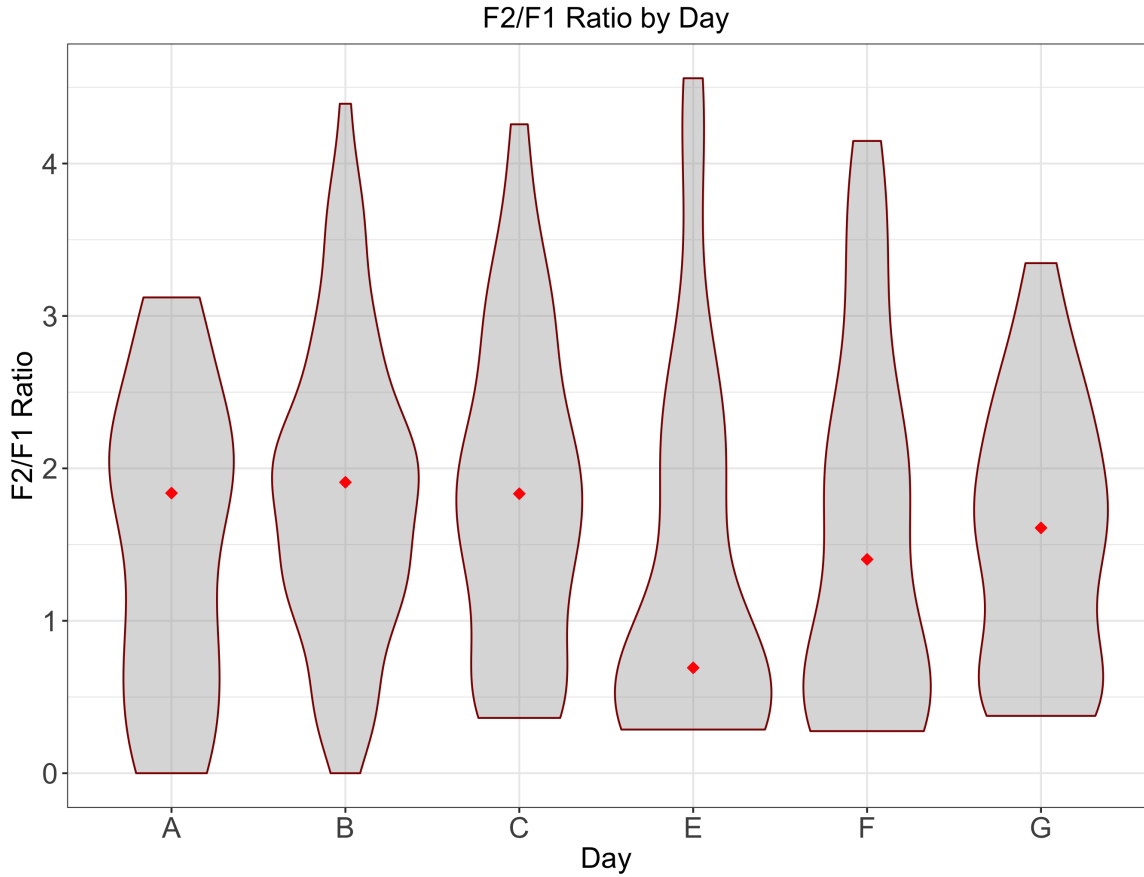
Figure 37. Distribution of key force ratio by day.

The zoomed-in view of Figure 38, showing the breakdown of the $\frac{F2}{F1}$ ratios by jump, shows that while there is more variance than apparent, the trend still holds: post-spaceflight $\frac{F2}{F1}$ ratios exhibit a trajectory of recovery toward the higher values

established by baseline measurements on days B and C, but with more variance. This trend is consistent with trends observed in sEMG contexts, particularly in figures describing the behavior of the bilateral ratios calculated for the TA muscle in Section 4.5.2.
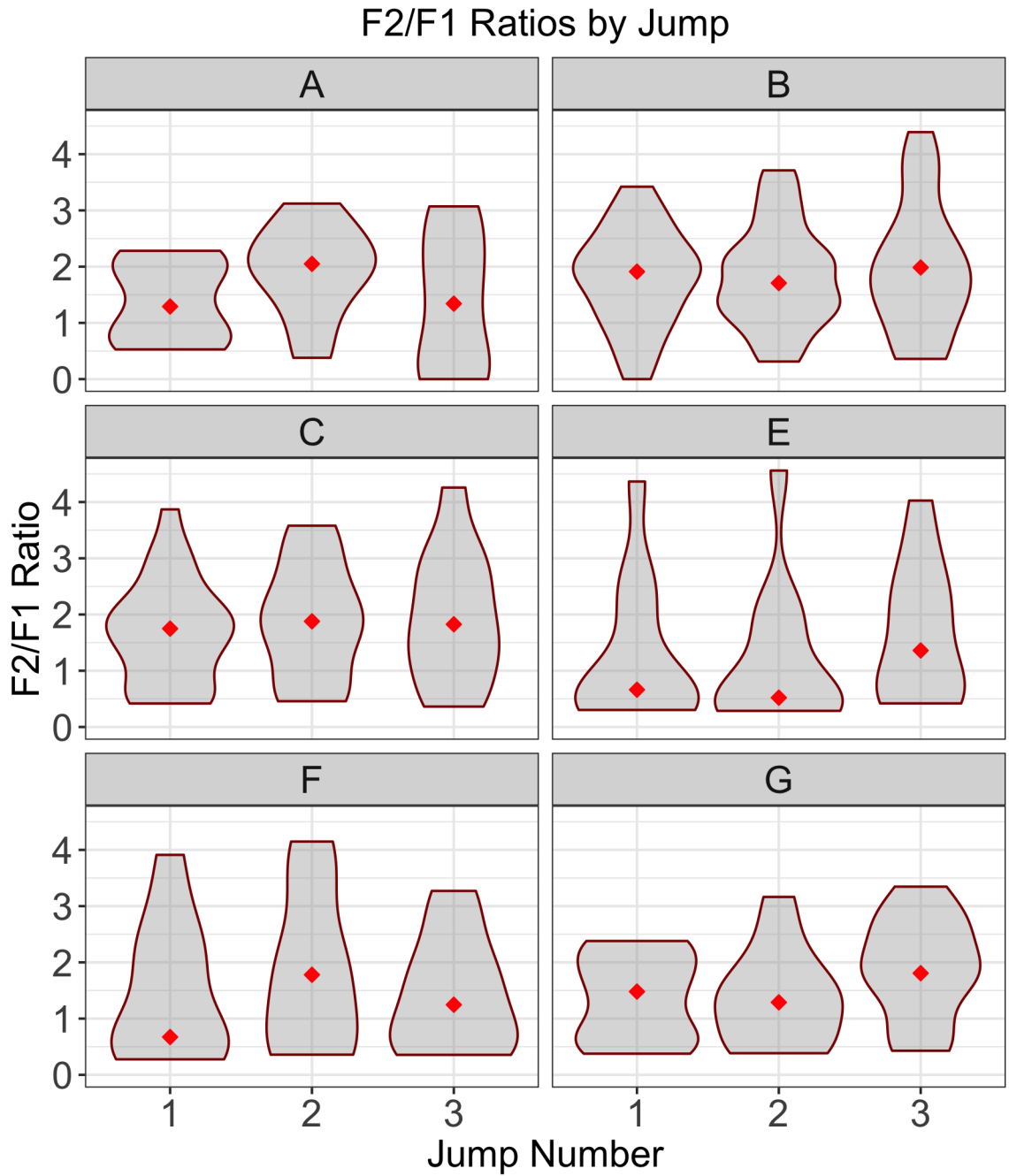
Figure 38. Distribution of key force ratio by jump.

Continuing the description of biomechanical outcomes, we also examine the FWHM duration of the third force peak, F3, interpreted as the force exerted on the force plate as the subject stands straight after landing the jump. Figure 39 shows a steady median with similar distributions for baseline days, with a sudden post-spaceflight spike. The median value of W3 eventually settles on a value similar to that of the baseline days, but still exhibits a larger distribution of values suggesting that different subjects may take longer to recover, or recover in different ways. That is, it is quite possible that subjects develop new muscular recruitment strategies to deal with the fall and landing post-spaceflight.
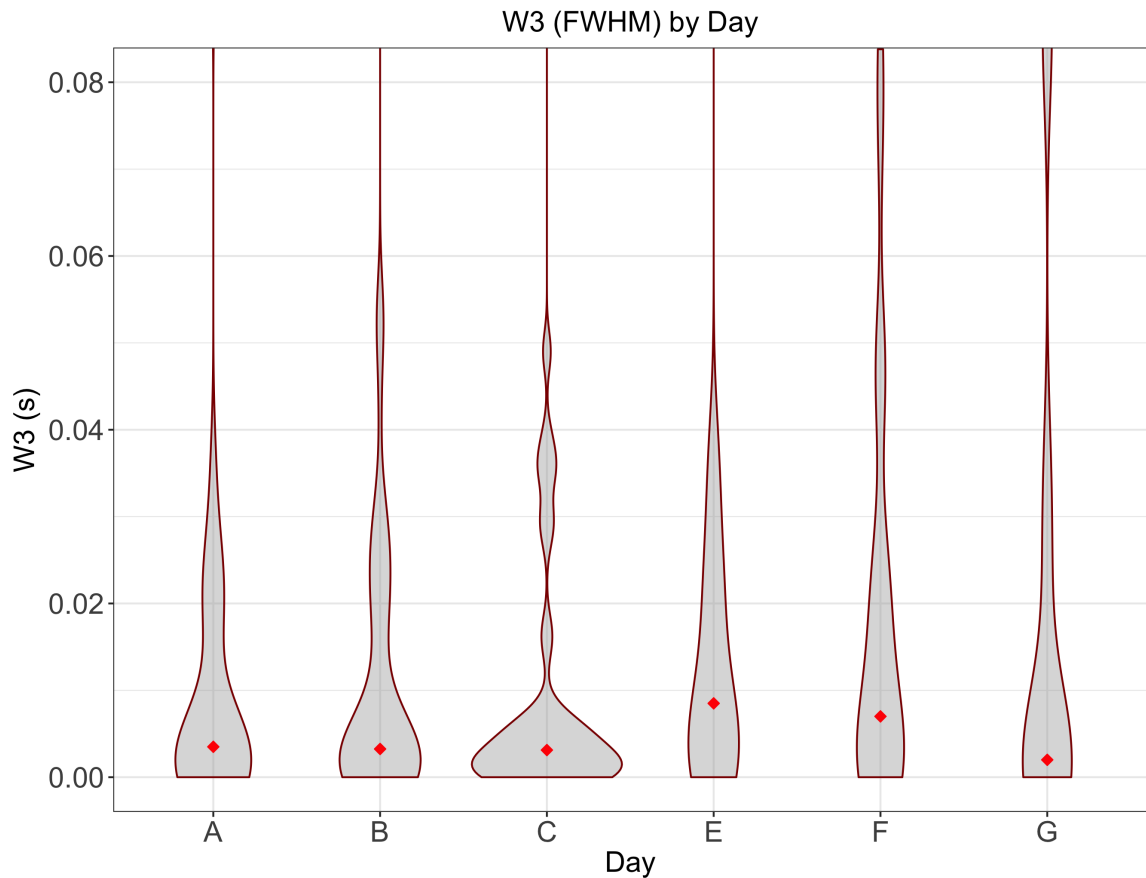
Figure 39. Distribution of FWHM duration of F3 (W3) by day, illustrating changes in the time it takes for subjects to stand straight after landing.

4.6 Pairwise Statistical Significance Testing between JDT experiment days

In addition to the statistical distributions presented above, we also present some select tables showing significance testing between JDT experiment days (e.g., day A, day B, etc.) for various features. A pairwise Wilcoxon test is performed, as this method is non-parametric and thus suitable for the distributions observed in the previous section. Statistical distribution differences are reported in the form of p-values. In general, one should be cautious [77] in drawing conclusions from p-values, but here they are presented to further develop some intuitive sense of the data.

The following selected tables are captioned with their variable comparisons. First, we present Table 5 which has what are traditionally considered rather large p-values. This particular table is shown to warn against overly general interpretations of p-values. While these values are larger than what is commonly (and arbitrarily [78]) accepted as "statistically significant", they indicate that there are differences in the subject populations along the selected variables. Day A and day B, for example, are likely to be more different than, say, day B and day C, which comports with our expectation that subjects generally improve as they acclimate to the JDT.

Table 5. Pairwise testing p-value comparison for lta_iemg_air and normTime by day.

	A	B	C	E	F
B	0.2473				
C	0.6897	0.6897			
E	0.7498	0.6897	0.8842		
F	0.9795	0.3406	0.7241	0.7498	
G	0.6897	0.8842	0.7498	0.7498	0.6897

Table 6 shows stronger differences, suggesting that the iEMG of the LTA measured after touchdown is a feature with higher explanatory power than that of Table 5. It is particularly comforting that days B and C are very different from day E. We should expect, then, that an algorithm should have much less difficulty distinguishing these

classes when informed by features yielding similar p-value behavior.

Equally important, we should recognize the possibility that different features might better offer variance for different day comparisons. For example, in Table 6, days E and F test favorably compared to those days as described in Table 7.

Table 6. Pairwise testing p-value comparison for lta_iemg_lnd and normTime by day.

	A	B	C	E	F
B	0.1951				
C	0.3851	0.4637			
E	0.0104	0.0697	0.0104		
F	0.7195	0.0839	0.2043	0.0087	
G	0.3162	0.9413	0.7195	0.2043	0.2043

Table 7 demonstrates the quality of the bilateral ratio cast in terms of iEMG as calculated on the MG muscles. Notably, this feature appears to have better discriminatory power across a broader range of pairwise comparisons than the standalone features. The comparison may not be favorable for a single pair, though, as demonstrated by examining the p-value for the day C and E comparison of Table 6 (i.e., 0.1124 vs 0.0104). At this point we expect that a broader feature base should have greater classification accuracy – particularly for a multi-classification problem. This shall be seen in Chapter 5.

Table 7. Pairwise testing p-value comparison for bmg_iemg and normTime by day.

	A	B	C	E	F
B	0.0966				
C	0.0286	0.4119			
E	0.0151	0.0509	0.1124		
F	0.0265	0.0286	0.0772	0.8236	
G	0.0561	0.7133	0.8236	0.1112	0.0772

Lastly, we present the pairwise comparisons for the bilateral ratios for the TA as derived from the wavelet transformed statistics in Table 8. This statistic is interesting because much of the literature involving the FSR with lower-limb dominant motions

(e.g., jumping and landing) has focused on quantifying the behaviors of the soleus and the MG [9, 72]. In this view, all comparisons with day E appear to be very different. It should be noted that these p-values are not mathematically zero, but merely truncated as with the previous tables.

Table 8. Pairwise testing p-value comparison for bta_wav and normTime by day.

	A	B	C	E	F
B	0.8360				
C	0.4035	0.1713			
E	0.0001	0.0000	0.0000		
F	0.2041	0.1713	0.0220	0.0156	
G	0.8848	0.9012	0.3409	0.0000	0.1903

Table 8 again supports the notion that different features offer different explanatory power. This feature dimension offers excellent discriminatory power for day E comparisons, but does relatively poorly (compared to other feature dimensions) when comparing day A and day B. The comparison between days B and C on this feature dimension also suggests, more so than other dimensions, that there is yet some population differences for the baseline days – possibly due to continued learning, or adjustment of neural strategies.

The takeaway from this is that each feature dimension offers a different (some more than others) view into the changing dynamics of the JDT. It is difficult to make general statements based off of, say, some metric measured from the MG muscles, when another view of the same experiment might tell a different story. The JDT, being a gestalt composition of various moving parts in a dynamic setting, is too complex to be reduced to such analysis. However, we may yet find some insight into underlying physiology by examining changes in feature dimension behavior over time.

4.7 Summary of Hypothesis

In contrast to the work of Chapter 3, the JDT involves a more dynamic experiment with transient events. Researchers typically try to design their experiments for simplistic measures and then interpret the results in their appropriate domain context. As previously described, the FTT measures coordination capabilities of astronauts by having them complete obstacles (pylon course or the pegboard drill test, for example) and measuring time to completion.

In this case, however, there are some domain specific requirements that demand a more involved experiment. The goal of the JDT is to evoke reflexes in response to ground reaction forces and the anticipation of landing in order to observe how sensorimotor strategies might change, or how an astronaut will perform upon return to a planetary surface.

Similarly to Chapter 3, we propose that a classifier model can be constructed to differentiate between pre- and post-spaceflight cases on the basis of sEMG alone. This may prove valuable in future deployment of predictive systems for gauging astronaut readiness prior to exiting their vehicle. Other, more Earthly possibilities exist in gauging sensorimotor capabilities in the elderly, and longitudinal physiological tracking in general.

In addition to, and as part of building the classifier model, it is important to build an analysis that describes the outcomes in a manner conducive to interpretration by researchers in the physiological domain. This is particularly important because the JDT is by its very nature a gestalt composition of moving parts acting (and reacting) to transient events. Therefore, we also examine regression techniques that lend themselves to such multi-faceted interpretation.

Hypotheses not addressed in this work are described in Chapter 6.

CHAPTER 5

LASSO REGRESSIONS AND ALGORITHMIC CLASSIFICATION

5.1   Motivation and Ground Truth

When assessing or improving algorithmic performance it is important to understand the labels one uses. In the JDT case, the labels of pre- and post-spaceflight or the day of experiment with respect to launch date are obvious. What is perhaps less obvious is what these labels represent in the biomechanical realm. Actual JDT performance differs from subject to subject, though there are broad trends as illustrated in Chapter 4. Therefore, we attempt to make connections between the sEMG measurements and the force-plate outcomes, where the force-plate outcomes effectively become a ground truth.

The observed variables, or features, are derived from the time-region prior to and just after touchdown. The force outcomes are measured immediately following touchdown, where they are produced. The approach of Section 5.2 then treats these force outcomes as measures to be correlated with sEMG variables, and thus open some path toward interpretation of the experiment. In a loose sense, one can think of this as an attempt at mapping from sEMG measurement (input to anticipatory strategy and ground reaction forces) to force outcome.

In any case, a classification algorithm successfully discriminating between pre- and post-spaceflight is necessary to demonstrate some capability of tracking longitudinal

changes on some minimum of non-invasively retreived data. This work is presented in Section 5.3.

5.2 The LASSO Regression

The data set contains many sEMG-derived features whose relationship with force outcomes will be examined for the ultimate purpose of giving a succint yet more in-depth look at the changes in JDT behavior over time. We focus here on the force ratio $\frac{F_2}{F_1}$ due to its ease of interpretation.

First, we discuss the shrinkage estimator approach to regularizing the data in order to avoid overfitting. The terms LASSO, ridge, and elastic net regression refer to fitting a generalized linear model with a penalty term (or two) to obtain regularized fit coefficients.

$$\hat{\beta}_{lasso} = argmin\{\frac{1}{2}\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p}x_{ij}\beta_j)^2 + \lambda\sum_{j=1}^{p}|\beta_j|\} \tag{16}$$

$$\hat{\beta}_{ridge} = argmin\{\frac{1}{2}\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p}x_{ij}\beta_j)^2 + \lambda\sum_{j=1}^{p}\beta_j^2\} \tag{17}$$

Equations 16 and 17 are the LASSO and ridge regressions, respectively. On the left-hand side of these equations are the fit coefficients for the respective approach. On the right-hand side, we have the regularization (or shrinkage) parameter, λ, the response y_i, the predictor elements x_{ij}, and fit coefficients β_j, where β_0 is constant. The calculations are run over p features for N observations. In our case, $N > p$, so we are not dealing with a strictly high-dimensional problem. As an aside, for related reading, Chen et al. (1998) explore this approach, referred to in the signal processing context as basis pursuit [79], which is sometimes employed in a wavelet context.

The LASSO implements an ℓ^1-norm penalty while the ridge approach implements an ℓ^2-norm penalty. These penalties are the second term on the right-hand side of

their respective equations. Qualitatively, the LASSO results in sparse solutions in which many of the minor contributors within the predictor set are forced to zero. While the ridge approach will not yield sparse solutions, it is better able to cope with correlated variables, shrinking their coefficients toward one another effectively averaging their contributions.

Zou and Hastie (2005) introduced a compromise penalty term referred to as the elastic net penalty [80], which is a combination of the two penalty norms used in equations 16 and 17.

$$\sum_{j=1}^{p}(\alpha|\beta_j| + (1-\alpha)\beta_j{}^2) \tag{18}$$

where,

$$\alpha = \frac{\lambda_1}{\lambda_1 + \lambda_2} \tag{19}$$

By equation 19 and $\lambda_i \geq 0$ (that is, only non-negative regularization parameters are sensible), we have $\alpha \in [0, 1]$, and casting the statement in this manner introduces a convenient intuition whereby α ranges between LASSO-penalty behavior ($\alpha = 0$) to ridge-penalty behavior ($\alpha = 1$). The conceptual framework for the glmnet [81] package in R, employed in the analytical efforts of this section, is as follows. Specifying a maximum or minimum value of α reduces us to one free parameter λ_i, while using the elastic nets approach demands that we choose two parameters. Optimal values of either parameter can be chosen via cross-validation, further illustrated in Section 5.2.2.

5.2.1 Standardization of Predictor Variables

Predictor variables are standardized such that they are expressed in the same dimensions (for interpretability purposes) and so that their respective coefficients are not

75

penalized in an uneven or biased manner. As described by G. James et al. (2013) [82] in the chapter on regularization, the feature space is standardized by,

$$\tilde{x}_{ij} = \frac{x_{ij}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_{ij} - \bar{x}_j)^2}} \tag{20}$$

where the symbols follow from equations 16 and 17 and $\bar{x}_j$ is the mean of the observations on feature j.

5.2.2 Regularization of the JDT Feature Space

Recall that the goal is to exploit any relationship between the sEMG features and force outcomes to examine JDT dynamics. Ideally, then, we want a sparse solution which is interpretable, but also reliable and not obfuscated by covariate data. This places the LASSO and ridge approaches in tension, and thus suggests that we examine correlations between variables and compare some results for a range of α and λ values.

Figure 40 shows a correlogram of standardized predictor variables. Here, it is clear that most features are not correlated, and thus of no concern with regard to interpretability on a LASSO regression. However, variables such as `bta_wav` and `bta_iemg` are clearly, and perhaps obviously, correlated. On the other hand, they represent the same notion of the bilateral ratio of sEMG activity on the TA, so for interpretation purposes, the point is moot. In the final analysis, cross-validated standard errors are used to assess model efficacy. This more qualitative approach can be used to check if substitutes may be used in the final analysis, or if interpretation of the analysis is obfuscated in some way. For this particular case, it is encouraging the strongest correlated and anti-correlated variables are those that constitute parts of the bilateral ratio calculation, or reflect sEMG activity within the same muscle and time-period (e.g., the in-air and post-touchdown portions of the signal).

The desired outcome is a sparse matrix of coefficients representing the efficacy of a predictor (or feature) in describing force outcomes. Incorporation of the ℓ^1-norm
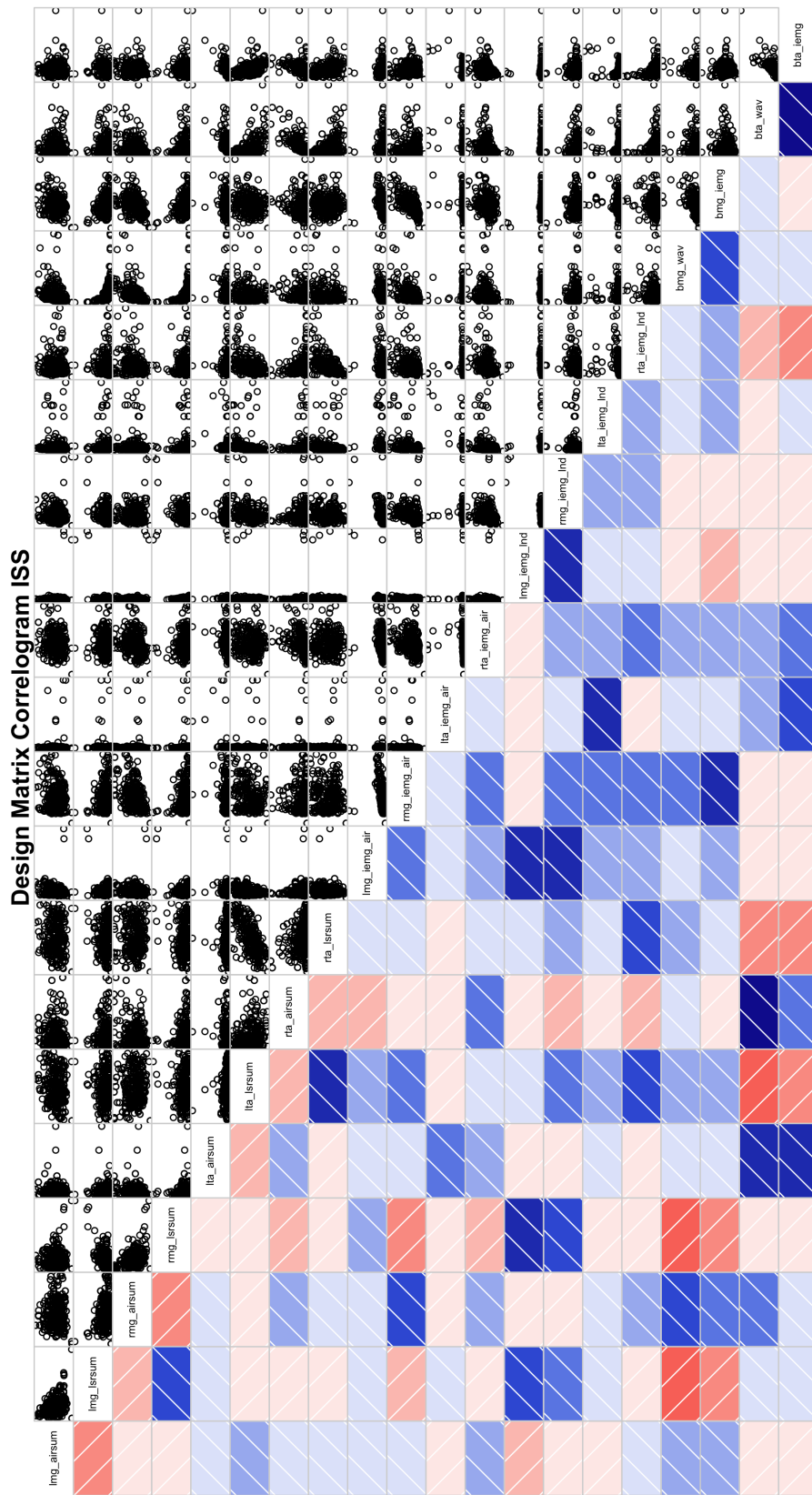
Figure 40. Corellogram illustrating correlated variables. Depth of color indicates stronger correlation (or anti-correllation) while the color indicates an inverse (red) or direct (blue) relationship. The top portion shows scatterplots of the respective variables. The tow-column intersection between two variables indicates that it is these two variables whose correlation is being illustrated.

is then most prudent as this penalty forces insignificant model contributions to zero, but we must also examine the possibility of correlated predictor variables playing a role. Therefore, we will examine the effect of an elastic net penalty using a grid search on both α and λ. Note that rather than a single specific value of λ, a sequence of λ's are utilized to show the regularization path for the predictor variables.

Figure 41 shows the Mean-Squared Error (MSE) of fit for varying penalty terms (α) and free regularization parameter (λ). Performance is not greatly improved over the LASSO ($\alpha = 1$), but rather the local optimum is merely shifted. As the regularization becomes more ridge-like ($\alpha = 0$), we observe that the algorithm does not converge in a satisfactory manner – the regularization term grows much larger before being competitive with larger α, and may reduce the degrees of freedom to unrealistically small values while introducing extreme bias to the model. For purposes of building generalizable and interpretable models, the LASSO is both sufficient and optimal.
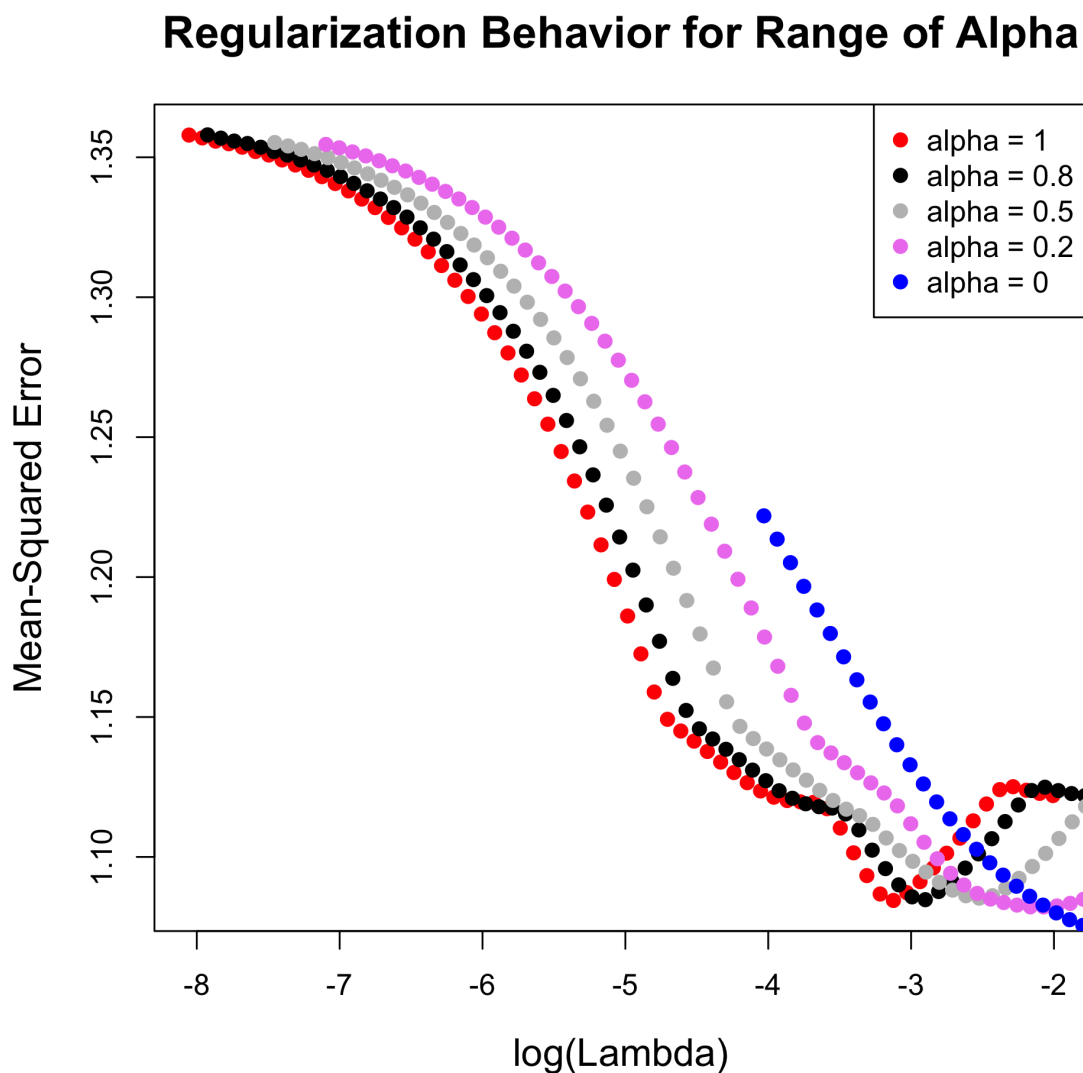
Figure 41. Fit performance of approaches with varying penalization ranging from the ridge penalty to the LASSO penalty.

Optimal values of λ are determined by cross-validation. It is standard practice to choose a value of λ one standard deviation away from the value giving the lowest MSE [81, 83], and that standard is followed here with the underlying rationale that a regularization parameter too restrictive may be overly optimistic and yield a less generalizable model. Exceptions are made if the regularization is such that the model is overly restrictive. Figure 42 illustrates such a case for the response variable $\frac{F_2}{F_1}$

under LASSO penalization. The cross-validated sequence of λ's terminates when the penalty reaches a point at which all predictor coefficients go to zero. For the purpose of this work, however, it is the sequence of λ's that we are interested in.
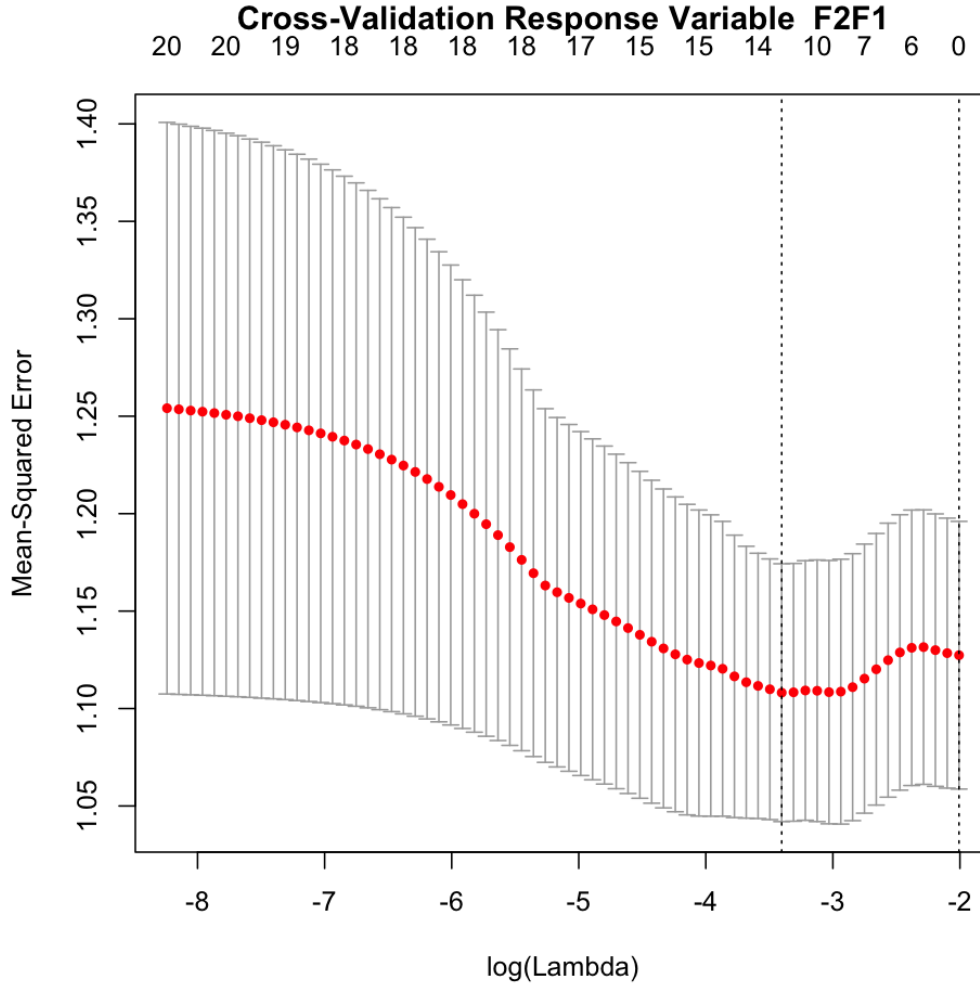
Figure 42. The k-fold cross-validated MSE is reported for varying regularization parameter λ. The error bars indicate the MSE standard deviation. Two dotted lines indicate λ values that minimize the MSE and the value that is one standard deviation greater. The top axis shows the degrees of freedom of the model – that is, the quantity of nonzero coefficients.

With a good sequence of λ's in hand, we now examine Figure 43, showing the regularization paths for the feature space correlating to the response variable, the force outcome $\frac{F_2}{F_1}$. The top axis indicates the degrees of freedom (nonzero coefficients)

present in the model, while the bottom axis indicates the size of the ℓ^1-norm error. Intuitively, then, reading the graph from left to right, the LASSO coefficients are allowed to grow along with the penalty term, as the regularization parameter, λ, grows smaller. Put another way, as the penalty is relaxed, more features enter into the description of the model.

Conveniently, however, the trends largely hold as the penalty is relaxed. The dominant predictors' coefficients simply grow while minor contributors hover about the zero-line. Figure 43 shows that the dominant predictors are `rmg_iemg_lnd` (i.e., the integrated sEMG calculated from the right MG after touchdown), `rta_iemg_air` (i.e., the integrated sEMG calculated from the right TA during the in-air portion of the JDT), and `rmg_iemg_air`.

Figure 43. LASSO regularization paths for features describing the $\frac{F_2}{F_1}$ outcome. Top axis indicates the degrees of freedom in the model from the number of nonzero coefficients, showing the number of variables in play. The bottom axis shows growing penalty term as we relax the regularization parameter and allow more features in the model. Finally, the higher a variable's coeffient, the stronger the positive, direct relationship with the outcome is. Likewise, the more negative a coefficients indicate a stronger inverse relationship with the outcome.

5.2.3   JDT and LASSO Interpretability

One might conclude from this portion of the analysis that the ratio $\frac{F_2}{F_1}$, understood as describing the smoothness of transition from toe (initial force-plate contact) to heel, is best described by the reaction to the landing (`rmg_iemg_lnd`) and actuation (in

anticipation) prior to landing. It is particularly noteworthy that the `rta_iemg_air` and the `rmg_iemg_air` features are anti-correlated with one another, as these two perform opposite functions as agonist and antagonist co-contracting pairs (see Figure 20).

Note however that Figure 43 shows the *overall* predictive trends of the experiment with respect to the $\frac{F_2}{F_1}$ ratio, regardless of day. The question of day-to-day (with respect to landing and launch dates) is also of interest, however. Therefore, we wish to also examine the regularization paths for the $\frac{F_2}{F_1}$ ratio under the JDT experiment day grouping. Figure 44 illustrates a slice of the process. Basically, since each set of regularization paths is similar in behavior to that of Figure 43, and we want to develop a heuristic method of estimating feature behaviors by day, we simply select the coefficients from some stable region of the graph – the right-hand side.
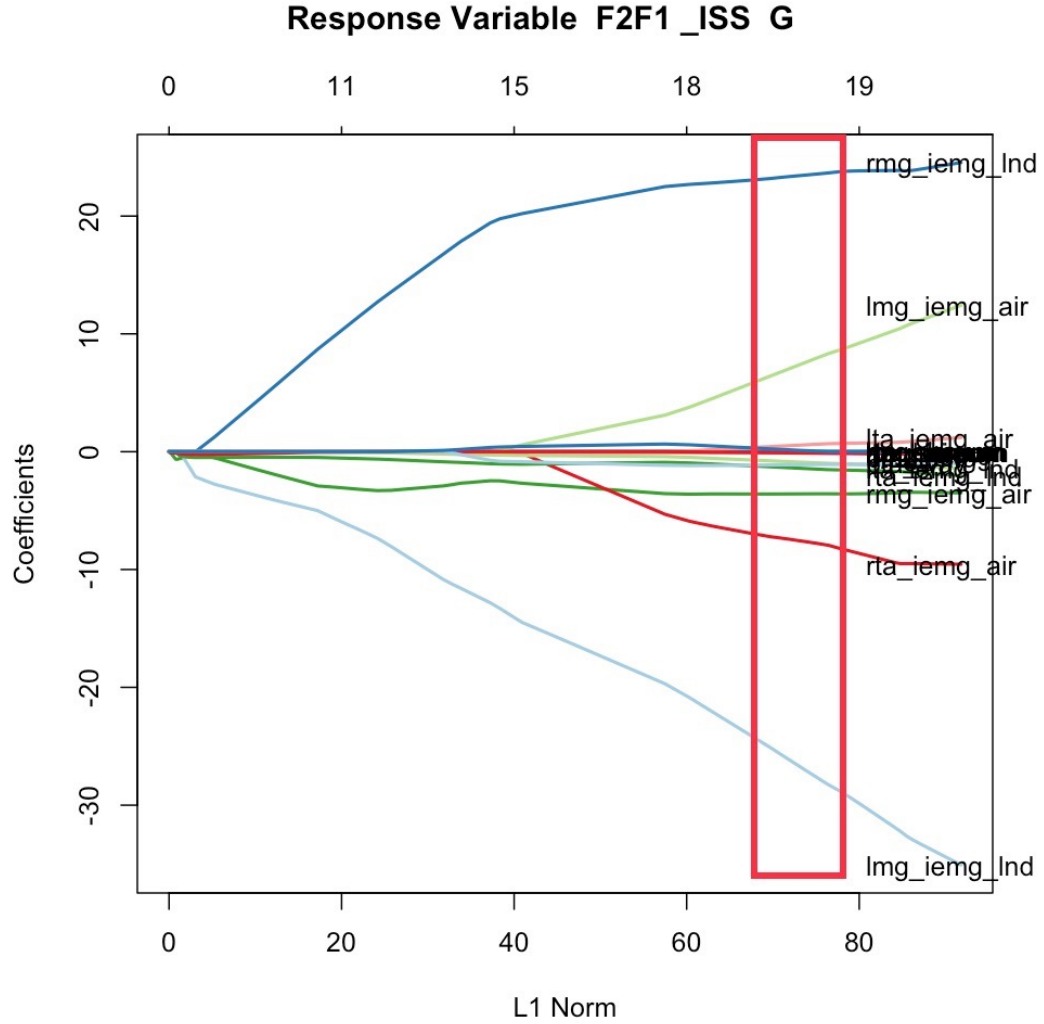
Figure 44. LASSO regularization path for $\frac{F_2}{F_1}$ response for Day G on ISS JDT experiment. Red selection indicates a slice from which coefficients may be selected.

This procedure is performed for each day. Within a given day the relative fractional contributions of the coefficients are calculated and reported in the stacked bar chart form shown in Figure 45.
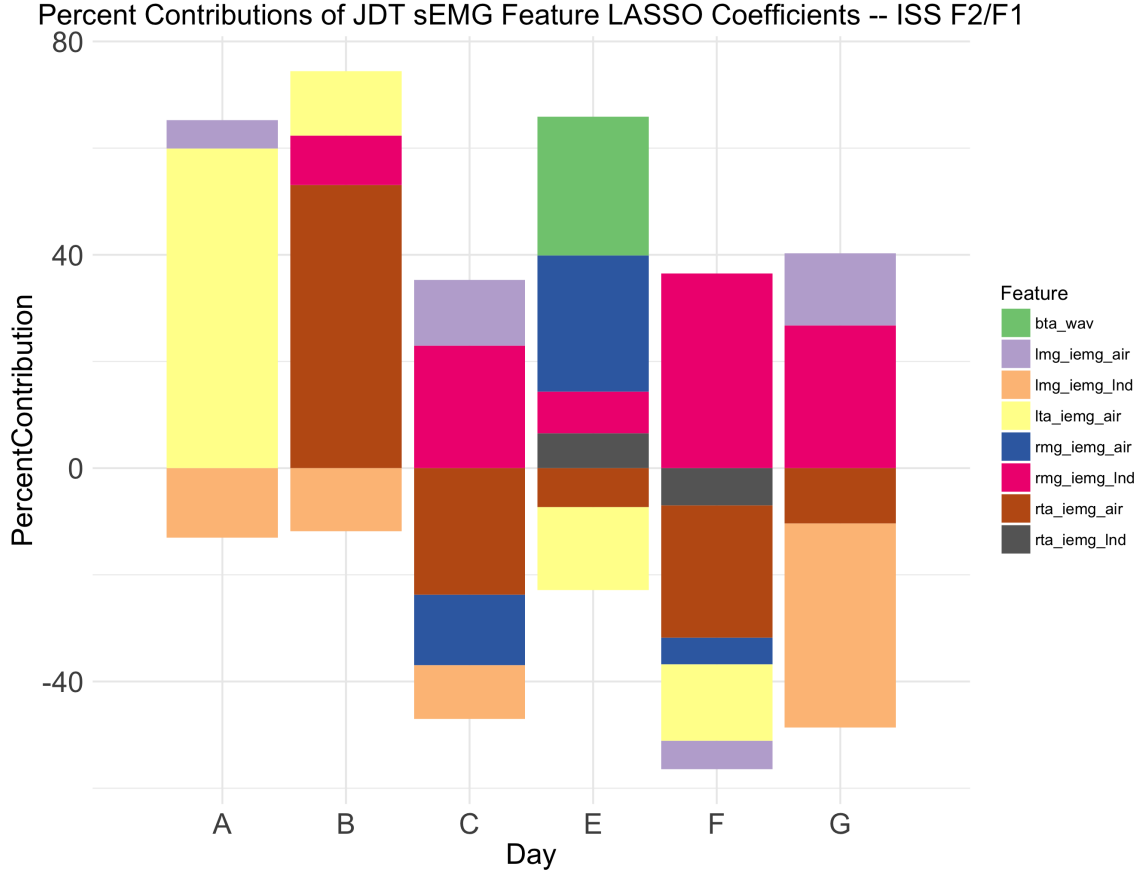
Figure 45. Fractional contributions of features to total LASSO coefficient space for each JDT experiment day expressed as percentages, for response variable $\frac{F_2}{F_1}$. Negative percentage contributions indicate features that are inversely related to the response variable. Positive contributions indicate direct relationships. Contributions less than 5% are discarded for visual clarity, so contributions may not total 100%.

As the subject population moves through the JDT experiment, we expect changes in neuromuscular activation strategies due to acclimation to the nature of the jump (establishing a baseline) and re-adaptation after return to planetary surface from pro-longed exposure to microgravity environment encountered during spaceflight.

When considering the transition from Day A to Day B, note that significant in-air activation of the TA occurs. This is consistent with the need for a stronger co-contraction of the agonist/antagonist pair (MG/TA) in stabilizing the foot for a steady landing. This is in turn consistent with the anecdotal experience of the author of this dissertation that the JDT experimental parameters are initially uncomfortable

due to the requirement that the subject "hops" rather than steps off the platform, which is already 30cm high. This discomfort has the nearly uniform result (judging by video evidence) that subjects rock slightly forward on their toes backward on their heels, exhibiting an overcorrection.

In general, across the subject population, Day A exhibits the greatest volatility when examining the distribution of behaviors by jump (e.g., Figure 33 and baseline behavior becomes more uniform on Day C. It should be noted that the difference between Day A and B is greater than the difference between Day B and C, on average, judging both by assessment video evidence and force feature outcomes (e.g, $\frac{F_2}{F_1}$).

The post-spaceflight trend suggests that on Day E the strategy is completely different. The bilateral ratio of the TA is present as a strong, directly correlated predictor only in this case. This fact along with the behavior of that same feature in Figures 31 and 32 suggest that TA behavior is worth a closer look. Much of the research – especially that involving stretch reflex investigations – has primarily focused around the MG or the soleus, both of which serve a similar function [7, 9, 68, 72, 84]

Further along the post-spaceflight (recovery) trajectory we see the fractional feature contributions evolving toward baseline (i.e., resembling Day C), but in a non-identical manner. The spirit of this is consistent with suggestions that while JDT performance eventually, and apparently, recovers to baseline performance, the strategies utilized are not quite the same. Video observations are insufficient to support any claims beyond "subjects return to reasonably high JDT performance by Day G," as the JDT occurs rather quickly for the human eye, and the muscle activation patterns are not visibly apparent. Here we claim that this method may provide greater insight. Naturally, however, caveats apply and are discussed in the next section.

5.2.4 Experimental and Analytical Limitations

Good models require that predictors adequately explaining the outcomes (response variables) are in place. In this work, we have employed a supervised regression method with ℓ^1-norm penalty (i.e., the LASSO) to link the only two sources of data that are relevant and present – the sEMG features and the features that are registered on the force plate. The assumption that sEMG features calculated during the in-air portion have some effect on the outcome is a key assumption, and is certainly a reliable one considering the agreement between the exploratory figures of Chapter 4 and the regularization paths presented in this chapter. Some issues remain however, that a more complete description of the JDT may be able to address.

No direct measurement of a "ground truth" is taken during the in-air portion. We do not know, for example, if there are trunk adjustments taking place during this period, or the particular manner in which the foot is plantar or dorsiflexed. We know that actuation occurs, but have no ground truth for which to compare it. An accelerometer placed at key points of the body would provide some way of comparing ankle flexion to lower limb actuation, and thus more directly assess APA, for example.

Additionally, as seen in Figure 43, the overall data seem to indicate a dominance of the right-limb. The particular role of the limbs does change (especially post-spaceflight, see Figure 44); but the idea that over the entire dataset, that one limb dominates, is a curiousity, and sparks a desire for knowledge of the subject's "handedness" in the data.

At this stage it should be noted that, despite relatively low MSE, no single feature could be used to completely capture the dynamics of the JDT. Instead this approach is best used, at this point, to obtain qualitative descriptions of neural strategies so that one can reason about neural adaptation due to spaceflight and recovery exhibited by the JDT.

5.3 Algorithmic Classification

The final part of the overarching narrative of this work is the classification of JDT labels by sEMG-derived features. Specifically, in this section, "labels" refers to the normalized days of the JDT experiment, measured with respect to launch and return dates, as outlined in Table 4. One can imagine a use-case wherein a classifier model is developed on an ever-growing subject population and, using a JDT analog suitable for microgravity environments, subject performance on a planetary surface is predicted based on what labels the JDT measurements are assigned.

While the previous sections address the construction of an analytical tool for enhancing reasoning about first principles, this portion of the dissertation addresses an engineering concern outlined in the use-case described in the previous paragraph. In reality, machine learning schema range from overly simplistic linear separations, to endlessly complex voter-fusion ensembles consisting of many models stacked or integrated with one another. We shall err on the side of the simple and generalizable in this work, and discuss more possibilities in Chapter 6. This approach has merit on three fronts. First, simpler approaches leave some room for connecting results to physical processes (interpretability) rather than the model resembling a black box performing a mathematical mapping. Second, this problem represents a new domain in which the more fundamental models should be employed to establish some baseline performance. Third, the actual deployment of complex models into production can be unrealistic depending on computational resource constraints, e.g., using radiation-hardened processors places a more severe upper limit on computational resources in space, were it to come to that. To further illustrate this third point, consider the Netflix prize in which a goal of increasing recommendation accuracy by 10% was crowd-sourced, successfully, but not implemented because the solution was too costly to scale-up [85].

A variety of classification schemes may be employed, but the Support Vector Ma-

chine (SVM) approach and results are described in detail. The SVM is chosen for this exposition of the work due to its flexibility [86] with hyperparameters, especially in the case of the Radial Basis Function (RBF) kernel. Equally important, the metrics for gauging classifier performance, including Receiver Operator Characteristics (ROC) and grid search validation will be addressed.

5.3.1 Support Vector Machine Classification

The SVM algorithm has enjoyed a storied history of development moving from the nonlinear-capable maximal margin classifier [87] into the more complex soft margin classifier [88], a scheme employed here in this work. The full theoretical development of the SVM will not be presented here, but it is strongly recommended for those curious to investigate its inner workings by way of examining the text by Shawe-Taylor and Cristianini, *Kernel Methods for Pattern Analysis* [86]. The intriguing result popularly referred to as the "kernel trick", a method by which a linear learning algorithm may learn nonlinear decision boundaries, and the relatively quick training time, are what inform the decision to focus primarily on the SVM in this work.

Intuitively, an earlier formulation of the SVM classification problem seeks solutions that place a hyperplane in such a manner as to maximize the distance between clusters of data and the drawn hyperplane. Figure 46 illustrates the basic principle of the SVM classification scheme. In a binary classification case, observations would be classified opposite to those observations located on the opposing side.
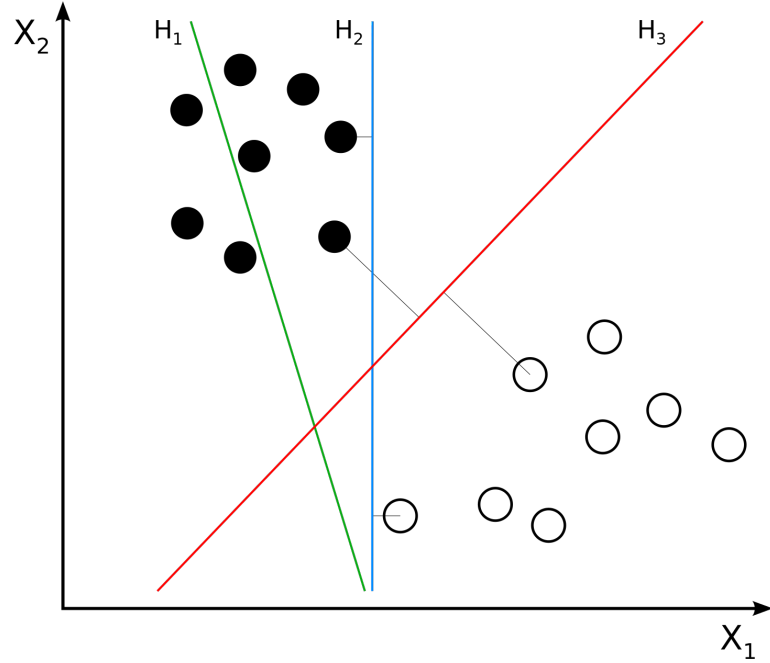
Figure 46. Illustration of maximal margin placement of hyperplane surrounded by data clusters. Illustration by Zack Weinberg (2012) [89], by way of the Wikimedia Commons.

In practice, data may not be so easily separated as in Figure 46, and we may also be operating in a space that is not easily visualized. This would also be an accurate description of the work presented here. The work done by Cortes et al. (1995) in pioneering the soft margin scheme [88] is the extension of the approach described by Figure 46.

For the JDT, then, we begin with a simple linear kernel SVM classifier with only two features – the bilateral ratios – by examining the ROC. The ROC is a representation of classifier performance drawn by comparing the true positive rate to the false positive rate. A true positive is a classification attempt that accurately classifies an observation, while a false positive is an attempt that incorrectly attributes one class to another. Intuitively, the case where the two rates are equal to one another indicates that the classifier is performing as good as pure guessing. Better classifiers are indicated by a high area under the curve (AUC), as would be the case when the ROC curve climbs rapidly, hugging the top and upper-left corners of the graph.

While this less formal approach serves this dissertation well, there is an excellent introductory paper to the nuances of ROC analysis by Tom Fawcett (2006) [90].

Before presenting the ROC, however, we will briefly visualize the decision surface for this classifier. In general, this is not possible due to the dimensionality of the predictor sets, and in any case it is not always particularly informative. Figure 47 illustrates such a case for a classification between days C (blue) and E (yellow), using the two bilateral ratios cast in terms of iEMG as the predictor features. A glance at the data suggests that post-spaceflight, JDT behavior described by the bilateral ratios is less varied, and smaller in magnitude in general. This claim would of course be weaker without the violin plots of Chapter 4 showing the various statistical distributions for the JDT measures. That is to say, Figure 47 says more about the kernels employed in the SVM algorithm than the data itself.
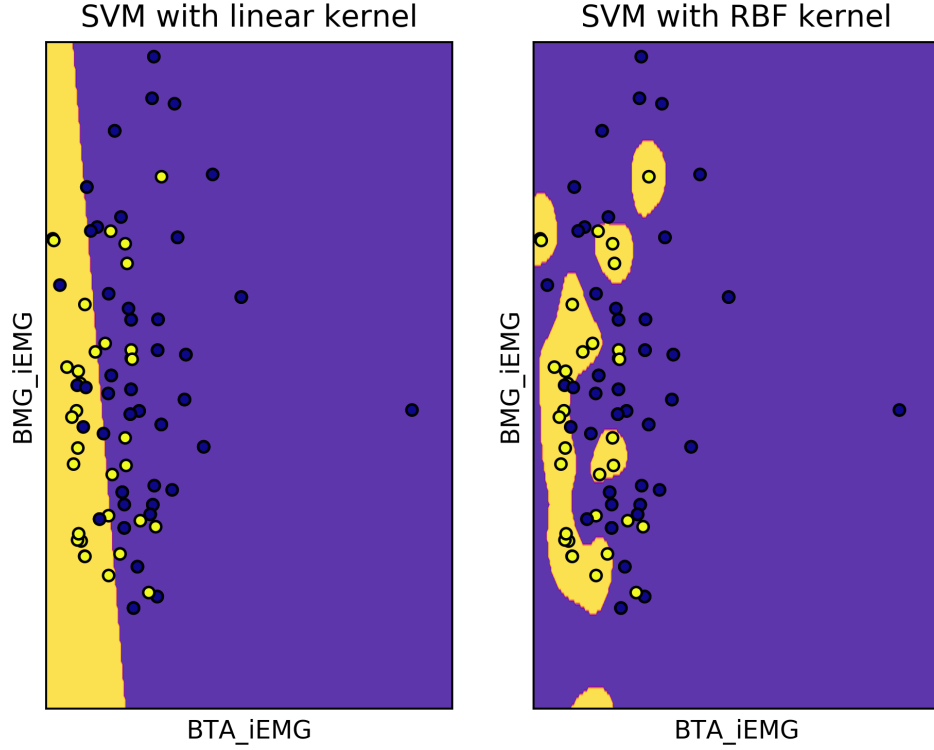
Figure 47. Left panel shows an example of a linear SVM kernel decision surface and the right panel shows an example of an RBF kernel decision surface. The RBF is clearly better suited to nonlinear observations. Yellow points indicate day E classification while blue points indicate day C classification.

The RBF kernel is the obvious choice for data such as this, though some comparisons will still be made throughout the rest of this section. The RBF handles nonlinear relationships better, and as will be shown, generally performs better when describing these data. Additionally, for real-valued data, the SVM is known to be a fast learner with added value as feature-space dimensions increase.

Figure 48 shows the ROC curve for a classification between JDT day C and day E using the linear kernel SVM. That is, the experiments that take place prior to launch and the experiment that takes place a day after return to the planetary surface. The predictor data consists only of the two bilateral ratio features. The classification is performed using a 5-fold cross-validation procedure with 80% of the samples allocated

to the training set and 20% of the samples allocated to the test set. All predictor variables, in all classification efforts shown below, are standardized in the manner described in Section 5.2.1.
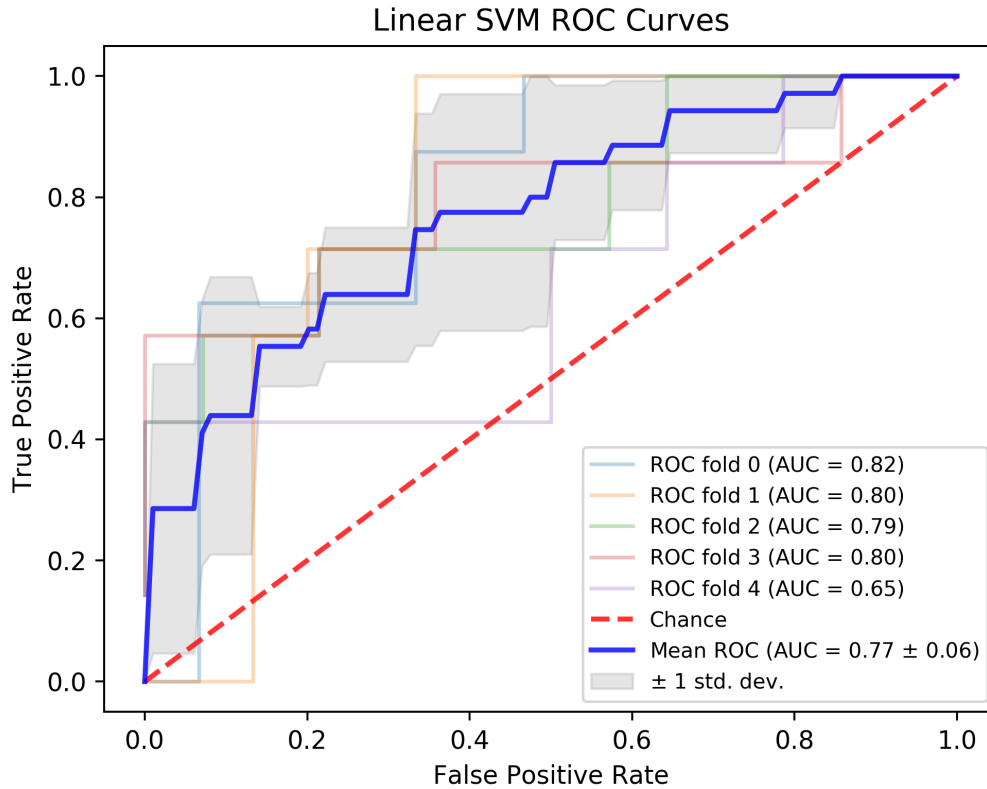
Figure 48. SVM classifier with linear kernel and predictor set populated only by the bilateral ratios. Performed with 5-fold cross-validation.

The red line of Figure 48 indicates the ROC trajectory at which the classifier performance would be equivalent to guessing. The thick blue line is the average performance across the folds, while the thinner pastel-colored lines are the performances for specific folds. The grey area represents the region through which any ROC trajectories are within one standard deviation of the average performance. This graph tells us that, for this particular configuration of predictor variables and kernel, it is possible, though improbable, to divide the test and training set up in a way that produces a much less effective classifier.

The previous two figures suggest that using the RBF kernel method and larger feature population may yield a better result. Figure 49 illustrates such an excellent outcome for classifying observations on day C versus day E.
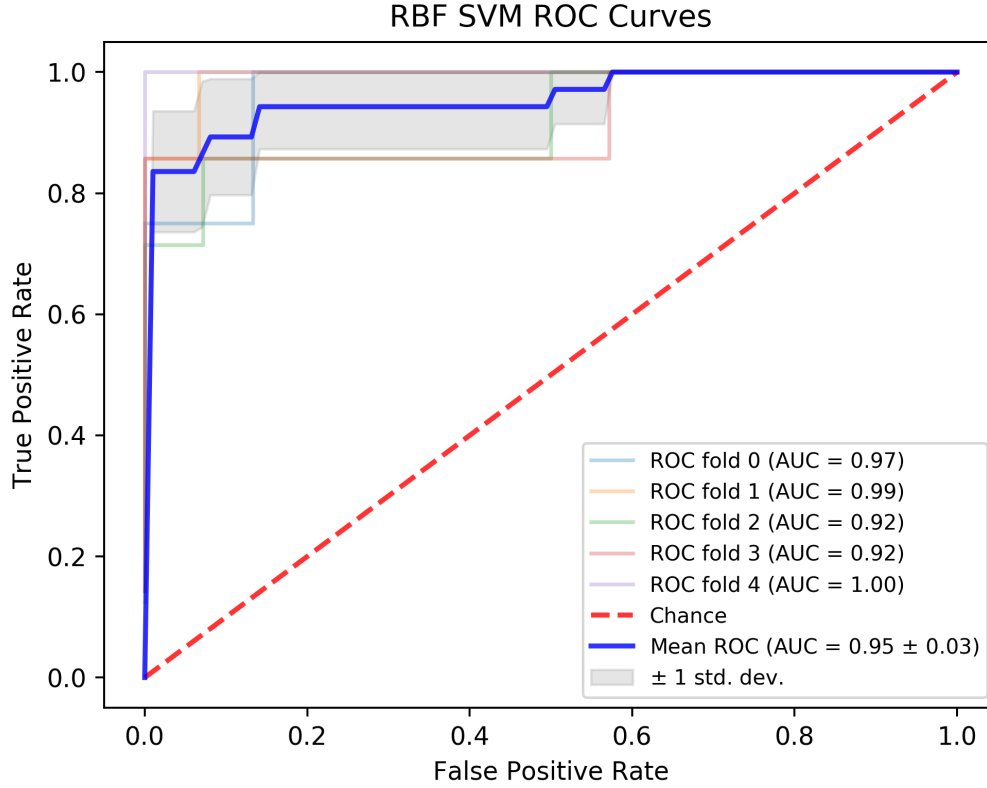
Figure 49. SVM classifier with RBF kernel and predictor set using all available features. Performed with 5-fold cross-validation.

The AUC is quite high compared to the linear kernel, two-feature case presented in Figure 48. The model is trained with 80% of the samples and tested on the remaining 20%, using a 5-fold cross-validation approach. The thinner pastel-colored lines and the grey one standard deviation region indicate a fairly tight grouping of behavior across the folds, indicating stable and accurate classification behavior.

At this point it is prudent to address some basic points regarding kernel methods. First off, a kernel represents a correspondence to an inner product between representations of all pairs of data in the feature space. This method avoids operating

in the actual data coordinate space, and thus reduces computational cost and eases the creation of the hyperplane. This implicit mapping is popularly referred to as the "Kernel Trick". A more satisfying investigation for the curious is to be had, again, by consulting with the text *Kernel Methods for Pattern Analysis* [86]. Recall that the inner product, for which the Kernel is a method of calculating, effectively computes the similarity between two vectors.

The key point, and underlying reason for introducing it here, is that there are additional hyperparameters that must be determined for optimal classification results. The RBF kernel, also referred to as the Gaussian kernel, is of the form

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\gamma \|\mathbf{x} - \mathbf{y}\|^2\right) \tag{21}$$

Where $\mathbf{x}$ and $\mathbf{y}$ are vectors in real space and $\gamma > 0$ is a free parameter that influences the effect of the distance between $\mathbf{x}$ and $\mathbf{y}$. Low values of γ yield a wide kernel distribution, meaning the influence of a training example has a far reach, while a higher value of γ indicates shorter decision influence, yielding a more finely-grained decision surface.

A "support vector" is a training point that lies on the contours of the decision boundary. Choosing more samples as support vectors allows for more complex decision boundary contours, while fewer support vectors make for a smoother decision surface. This regularization parameter will be referred to as C, here, and is given its proper mathematical consideration by Guyon et. al (1993) [91] and Cortes and Vapnik (1995) [88].

At this point then, RBF kernel classifiers have two hyperparameters which must be tuned to yield an optimal classifier. The classifiers and related methods employed in this section are produced using the excellent Python machine learning library, Scikit-Learn [92], which also offers a grid search method for tuning the aforementioned hyperparameters, C, and γ. Figure 50 presents a visualization of such a grid search

used to optimize the classifier whose ROC metrics are presented in Figure 49.
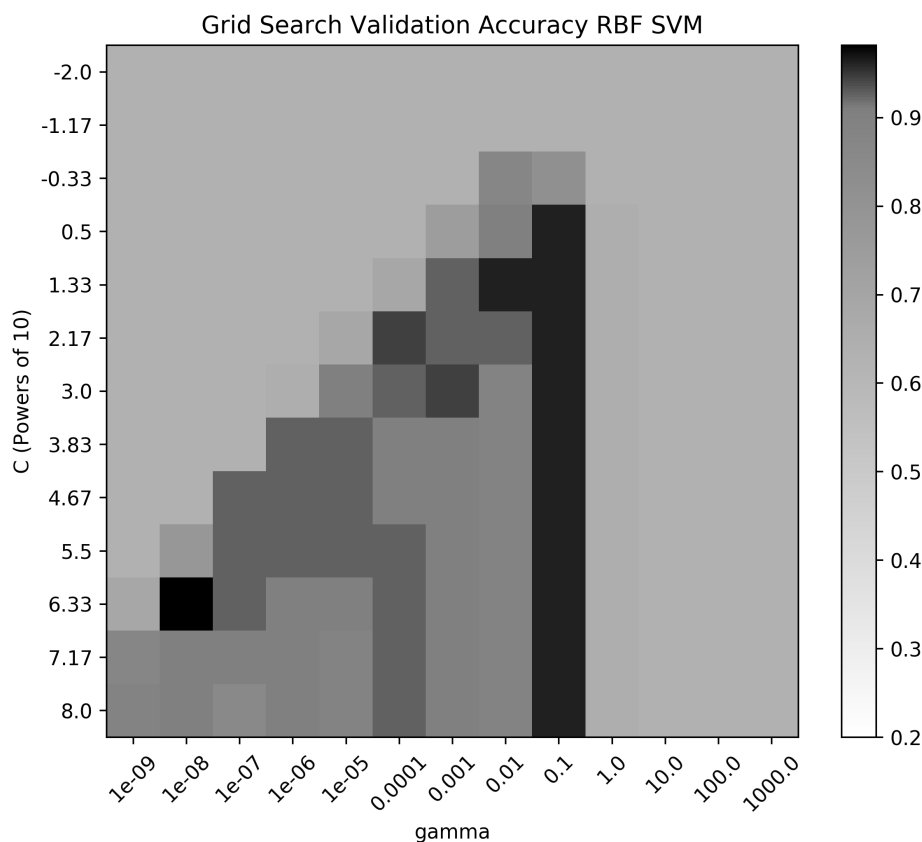
Figure 50. Grid search on hyperparameters C and γ assessed by 5-fold cross-validation classifier accuracy. The vertical axis is expressed in powers of 10.

In the case of the RBF kernel SVM classifier assigning labels of day C or day E to observations with the entire predictor set, a C of 21.54 and γ of 0.1 are optimal. The colorbar is normalized such that accuracy values near the optimal value are clearly visible, while hyperparameter combinations yielding poor accuracies are indistinguishable (or nearly so).

Clearly, then we can separate observations out before and after spaceflight with signficiant accuracy and with the use of only sEMG-derived features. Now we turn to the multi-classification problem. That is, we attempt to classify observations from JDT experiments beyond the day after landing and before the day of launch.

The SVM is inherently a binary classifier. To work around this, a one-against-one

96

approach is used [93] in the library method employed within Scikit-Learn. In this scheme, for K classes, $\frac{K(K-1)}{2}$ different binary SVM classifiers are trained on all possible class pairs. Test points are then classified by which class has the highest number of votes. Naturally, one can expect overall performance to drop when 'diluting' the problem space with more classes, and in general, this performance hit can be alleviated by having a large number of samples representing each class. Figure 51 shows the ROC performance metric for each class.

Note that for the multi-classification cases, we neglect day A as it is intended to give subjects a chance to acclimate to the nature of the JDT.
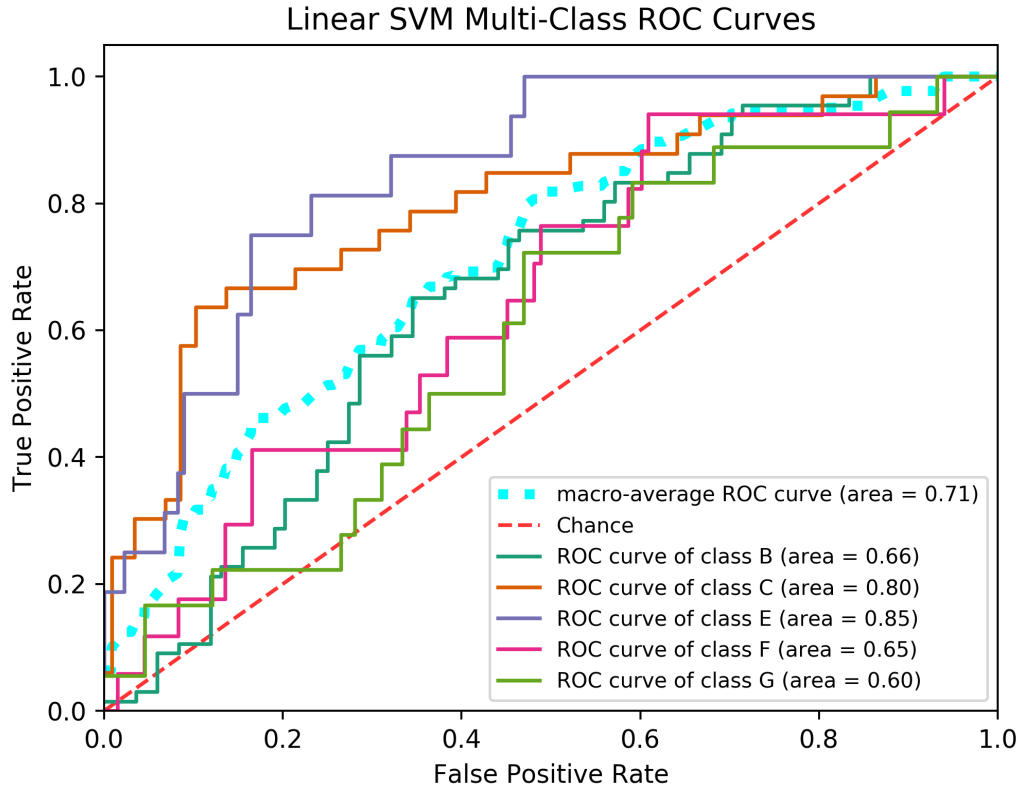
Figure 51. ROC performance metric for each class with associated AUC, and overall average performance (dashed cyan line). The individual lines represent class behaviors, rather than fold behaviors. The classification accuracy was checked with 5-fold cross-validation.

The linear SVM classifier shown in Figure 51 employs the entire feature set as

predictors, and performs better than chance overall, with days C and E performing particularly well. Recall, from Chapter 4, that the statistical distributions of days C and E had the most drastic differences on most sEMG features, suggesting that classification accuracy should be higher for these cases.

Finally, we examine the same multi-classification problem but with grid search-optimized RBF kernel. Figure 52 shows the ROC metrics of this classification scheme with a higher average AUC as compared to the metrics reported in Figure 51. The previous trend holds, however, as days C and E are classified most accurately. Note that the classification accuracies for these two days is still less than if we had trained only on days C and E as in Figure 49.
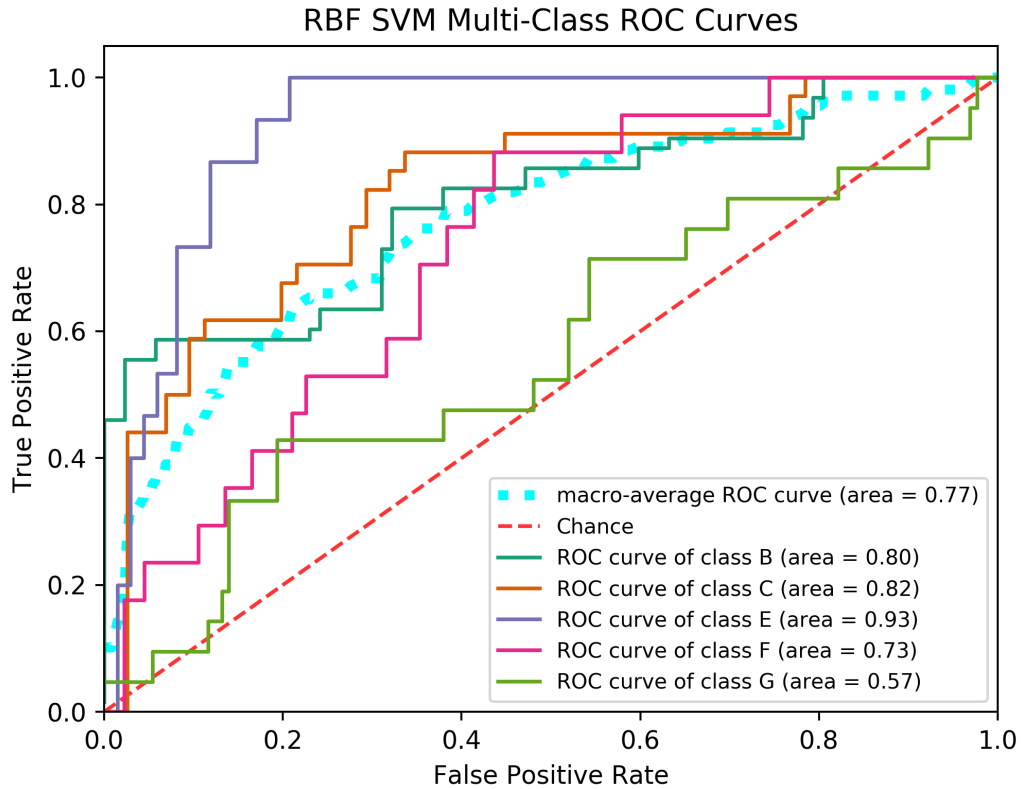
Figure 52. ROC Performance metric for multi-class JDT classificiation with RBF kernel with 5-fold cross-validation. Dashed cyan line indicates average classifier performance.

The SVM RBF kernel is clearly a strong performer, even if performance suffers for the multi-class problem. The grid search hyperparameter optimization procedure introduces some computational expense depending on how fine the grid is. It is promising that these results were obtained with hyperparameters that took approximately seventy-seven seconds to calculate on a late 2016 Macbook pro with a 2.9 GHz Intel i5 processor. These classification performance metrics may be thought of as the baseline from which further investigations into either more exotic algorithms, feature space transformations, or new domain-specific feature engineering (i.e., extensions or replacements for the bilateral ratio) can be compared.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this brief final chapter we take stock of what has been presented here, describe some next possible steps, and related explorations.

6.1 Conclusions

Chapter 3 introduces the first significant outcome presented in this work. There, a dataset consisting of sEMG signals taken during isometric contraction before and after a 35-day bedrest was analyzed by, put simply, extracting a variety of features from the signals and running classification algorithms to distinguish between signals measured on either side of the bedrest period. That work demonstrates feasibility in the construction of models for handling (classifying, here) longitudinal changes as indicated strictly by electrophysiological data – sEMG in this case. This is an interesting result because maximal use is made of the sEMG data, rather than using such signals as second-order evidence, as is often the case.

The JDT experiment itself is introduced in Chapter 4. The issues surrounding spaceflight-induced sensorimotor conditions are described to motivate the experiment. Domain-specific knowledge of the sensorimotor system and the mechanics of landing a jump is used in the engineering of two new features meant to convey physiological significance. Violin plots showing the statistical distribution of a number of these

sEMG-derived features, including the bilateral ratios (Equation 13), are shown to build some intuition regarding the evolving nature of the JDT statistics.

Chapter 5 presents two significant outcomes. First, Section 5.2 demonstrates the use of a shrinkage estimator as a new analytical tool for determining significant contributors to a ground truth variable. In Figure 45, this relationship was observed to change over time and thus highlight adaptative JDT strategies induced by spaceflight. In concert with the exploratory figures and statistical distributions of Chapter 4, one can see that a return to baseline JDT performance is achieved, but with the exciting possibility that a different JDT strategy has yielded this improvement.

The second outcome, presented in Section 5.3, is an extension of the study shown in Chapter 3 and applied to the JDT problem. Beginning with a simple two-class case, where the JDT is performed immediately before and after spaceflight, we demonstrate high classification accuracy using an RBF kernel SVM. This classification accuracy is commensurate with expectations developed by the statistical distributions of Chapter 4. The classification problem is expanded to include all but the first set of JDT experiments, labeled "Day A", as these jumps are highly variable and intended to help subjects acclimate to the experiment. While performance does suffer, the classification scheme performs well-above chance overall.

6.2 Future Work

There are multiple obvious extensions to the work presented in this dissertation. It is quite plausible that more complex ensemble methods could be used to improve classification accuracy for the multi-class case. There are likely also different ways in which we can engage in feature engineering for both intuitive purposes, and for enhancement of classification accuracy.

This author is particularly interested in the notion of unsupervised learning avenues of investigation. This approach may yield as-yet hidden insights into the struc-

ture of the data, were more subject data available. Variables like height and age, for example, are not included in the data as a result of the procedures taken to anonymize the data. Indeed, while ISS missions are known to last between three and six months, we do not know the specific durations of any of these. One would expect mission duration to have some relation to, say, recovery times.

In Section 5.2.4, some additional concerns and possiblities for their remedy are offered. The LASSO method employed in the manner described in Chapter 5 is intriguing in their use of a particular force outcome for discerning a ground truth that is otherwise not entirely obvious. Accelerometer data, for example, may be used to obtain some ground truth during the in-air portion, and aid the description of biomechanics post-touchdown.

There was some speculation about related investigations. When this work was proposed, discussion with a research group investigating neuromuscular fatigue [94,95] in the hand during spacesuit glove operation had recently taken place. While our research interests certainly aligned, there was a reasonable and just concern that the data were not quite ready to be released for study by outside groups. Once the general aims of the experiment are satisfied and data more freely released, some future investigation of the experiment using methods outlined in this work may yield insight into either the data or the method.

Lastly, there exist coherence studies for examining the relationships between EEG and sEMG features during experiments [96–98]. The use of EEG presents its own technical issues regarding noise sensititivity and motion artifacts, but the combined use of the EEG and sEMG measurements with, for example, the JDT, may give more information on underyling physiological changes driving neural strategies. It would be an amazing opportunity to observe correlated changes in the brain with the lower limbs measurements.

This work has relied on generalizable regression and classification methods. We

are confident that similar analyses can be performed with altered, domain-specific (e.g., hand fatigue experiments) feature engineering approaches.

REFERENCES

[1] K. Liao, M. F. Walker, A. C. Joshi, M. Reschke, M. Strupp, J. Wagner, and R. J. Leigh, "The linear vestibulo-ocular reflex, locomotion and falls in neurological disorders," *Restorative Neurology and Neuroscience*, vol. 28, no. 1, pp. 91–103, 2010.

[2] T. Kamali, R. Boostani, and H. Parsaei, "A multi-classifier approach to MUAP classification for diagnosis of neuromuscular disorders," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 191–200, 2014.

[3] C. Neal and C. Shearer, "Strategic knowledge gaps for the "moon first" human exploration scenario: Analysis and findings of the lunar exploration analysis group (LEAG) GAP-specific action team (SAT)," 2012. [Online]. Available: http://www.lpi.usra.edu/leag/GAP_SAT_03_09_12.pdf

[4] D. Beaty and M. Carr, "Analysis of strategic knowledge gaps associated with potential human missions to the martian system: Final report of the precursor strategy analysis group (p-SAG)," 2012. [Online]. Available: http://mepag.nasa.gov/reports/P-SAG_final_report_06-30-12_main_v26.pdf

[5] W. H. Paloski, C. M. Oman, J. J. Bloomberg, M. F. Reschke, S. J. Wood, D. L. Harm, B. T. Peters, A. P. Mulavara, J. P. Locke, and L. S. Stone, "Risk of sensory-motor performance failures affecting vehicle

control during space missions: a review of the evidence," *Journal of Gravitational Physiology*, vol. 15, no. 2, pp. 1–29, 2008. [Online]. Available: http://humanresearchroadmap.nasa.gov/evidence/reports/Sensorimotor.pdf

[6] M. Cermack, "Monitoring and telemedicine support in remote environments and in human space flight," *British Journal of Anaesthesia*, vol. 97, no. 1, pp. 107–114, 2006. [Online]. Available: https://academic.oup.com/bja/article-lookup/doi/10.1093/bja/ael132

[7] W. H. Paloski, S. J. Wood, A. H. Feiveson, F. O. Black, E. Y. Hwang, and M. F. Reschke, "Destabilization of human balance control by static and dynamic head tilts," *Gait & Posture*, vol. 23, no. 3, pp. 315–323, 2006. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S096663620500038X

[8] E. S. Tomilovskaya, M. F. Reschke, J. M. Krnavek, and I. Kozlovskaya, "Effects of long-duration space flight on target acquisition," *Acta Astronautica*, vol. 68, no. 9, pp. 1454–1461, 2011. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0094576509005554

[9] M. F. Reschke, J. J. Bloomberg, W. H. Paloski, A. P. Mulavara, A. H. Feiveson, and D. L. Harm, "Postural reflexes, balance control, and functional mobility with long-duration head-down bed rest," *Aviation, Space, and Environmental Medicine*, vol. 80, no. 5, pp. 45–54, 2009. [Online]. Available: http://openurl.ingenta.com/content/xref?genre=article&issn=0095-6562&volume=80&issue=S1&spage=45

[10] J. Bloomberg and A. Mulavara, "Changes in walking strategies after spaceflight," *IEEE Engineering in Medicine and Biology Magazine*, vol. 22, no. 2, pp. 58–62, 2003. [Online]. Available: http://ieeexplore.ieee.org/document/1195697/

[11] R. Stallard, E. Rejc, and K. C. Welch, "Wavelet-derived features as indicators of physiological changes induced by bed rest," *Expert Systems with Applications*, vol. 90, pp. 232–240, 2017.

[12] R. Stallard, I. Kofman, J. Cerisano, M. Rosenberg, M. Reschke, and K. Welch, "Algorithmic Classification Based on Astronaut Sensorimotor Performance Before and After Spaceflight," Nov. 2016, presented at the Kentucky Academy of Sciences at the University of Louisville, Louisville, KY.

[13] R. Stallard, I. Kofman, M. Rosenberg, J. Cerisano, M. Reschke, and K. Welch, "Feature Extraction, Transformation, and Algorithmic Classification of Astronaut Sensorimotor Performance," Galveston, Texas, Jan. 2017, poster presentation at the 2017 NASA Human Research Projects and International Workshop Conference, Galveston, TX.

[14] R. Seidler, A. Mulavara, V. Koppelmans, K. Cassady, P. Yuan, I. Kofman, Y. De Dios, R. Riascos-Castaneda, S. Wood, and J. Bloomberg, "Long duration head down tilt bed rest and spaceflight effects on neurocognitive performance: Extent, longevity and neural bases," *NASA HRP Galveston, TX*, 2017.

[15] A. Van Ombergen, S. Laureys, S. Sunaert, E. Tomilovskaya, P. M. Parizel, and F. L. Wuyts, "Spaceflight-induced neuroplasticity in humans as measured by MRI: what do we know so far?" *npj Microgravity*, vol. 3, no. 1, 2017. [Online]. Available: http://www.nature.com/articles/s41526-016-0010-8

[16] K. Cassady, V. Koppelmans, P. Reuter-Lorenz, Y. De Dios, N. Gadd, S. Wood, R. R. Castenada, I. Kofman, J. Bloomberg, A. Mulavara, and R. Seidler, "Effects of a spaceflight analog environment on brain connectivity and behavior," *NeuroImage*, vol. 141, pp. 18–30, 2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1053811916303378

[17] G. E. Loeb and C. Gans, *Electromyography for experimentalists.* University of Chicago Press, 1986.

[18] J. V. Basmajian and C. J. De Luca, *Muscles alive: their functions revealed by electromyography*, 5th ed. Williams & Wilkins, 1985.

[19] S. H. Nawab, R. P. Wotiz, and C. J. De Luca, "Decomposition of indwelling EMG signals," *Journal of Applied Physiology*, vol. 105, no. 2, pp. 700–710, 2008. [Online]. Available: http://jap.physiology.org/cgi/doi/10.1152/japplphysiol.00170.2007

[20] E. Kupa, S. Roy, S. Kandarian, and C. De Luca, "Effects of muscle fiber type and size on emg median frequency and conduction velocity," *Journal of Applied Physiology*, vol. 79, no. 1, pp. 23–32, 1995.

[21] S. Nair, R. French, D. Laroche, and E. Thomas, "The Application of Machine Learning Algorithms to the Analysis of Electromyographic Patterns from Arthritic Patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 174–184, 2009.

[22] H. Writer and R. Arora, "Vestibular rehabilitation: An overview," *Otorhinolaryngology Clinics - An International Journal*, vol. 4, pp. 54–69, 2012. [Online]. Available: http://www.jaypeejournals.com/eJournals/ShowText.aspx?ID=3564&Type=FREE&TYP=TOP&IN=~/eJournals/images/JPLOGO.gif&IID=280&isPDF=YES

[23] D. E. Parker, "Human vestibular function and weightlessness," *The Journal of Clinical Pharmacology*, vol. 31, no. 10, pp. 904–910, 1991.

[24] M. F. Reschke, J. J. Bloomberg, D. L. Harm, W. H. Paloski, C. Layne, and V. McDonald, "Posture, locomotion, spatial orientation, and motion sickness as

a function of space flight," *Brain Research Reviews*, vol. 28, no. 1, pp. 102–117, 1998.

[25] D. E. Parker, "Spatial perception changes associated with space flight: implications for adaptation to altered inertial environments," *Journal of Vestibular Research*, vol. 13, no. 4, pp. 331–343, 2003.

[26] H. Scherer, K. Helling, A. Clarke, and S. Hausmann, "Motion sickness and otolith asymmetry," *Biological Sciences in Space*, vol. 15, no. 4, pp. 401–404, 2001.

[27] M. Heer and W. H. Paloski, "Space motion sickness: Incidence, etiology, and countermeasures," *Autonomic Neuroscience*, vol. 129, no. 1, pp. 77–79, 2006. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1566070206002141

[28] S. Nair, R. French, D. Laroche, and E. Thomas, "The application of machine learning algorithms to the analysis of electromyographic patterns from arthritic patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, 2010. [Online]. Available: http://leadserv.u-bourgogne.fr/rfrench/french/Nair_French_Laroche_Thomas.pdf

[29] A. Subasi, M. Yilmaz, and H. R. Ozcalik, "Classification of EMG signals using wavelet neural network," *Journal of Neuroscience Methods*, vol. 156, no. 1, pp. 360–367, 2006. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0165027006001440

[30] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*, 4th ed. McGraw-Hill, 2009.

[31] Y. C. Pati and P. S. Krishnaprasad, "Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations," *IEEE Transactions on Neural Networks*, vol. 4, no. 1, pp. 73–85, 1993.

[32] O. Wainwright, "The lego prosthetic arm that children can create and hack themselves," *The Guardian*, 2015. [Online]. Available: http://www.theguardian.com/artanddesign/architecture-design-blog/2015/jul/22/lego-prosthetic-arm-that-kids-can-hack-themselves

[33] "Touch bionics: Technology that touches lives," 2017. [Online]. Available: http://www.touchbionics.com/

[34] A. Gijsberts and B. Caputo, "Exploiting accelerometers to improve movement classification for prosthetics," in *Rehabilitation Robotics (ICORR), Conference on.* IEEE, Jun. 2013, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/document/6650476/

[35] D. Bright, A. Nair, D. Salvekar, and S. Bhisikar, "EEG-based brain controlled prosthetic arm," in *Advances in Signal Processing (CASP), Conference on.* IEEE, Jun. 2016, pp. 479–483. [Online]. Available: http://ieeexplore.ieee.org/document/7746219/

[36] J. M. Fontana and A. W. L. Chiu, "Analysis of Electrode Shift Effects on Wavelet Features Embedded in a Myoelectric Pattern Recognition System," *Assistive Technology*, vol. 26, no. 2, pp. 71–80, Apr. 2014. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10400435.2013.827138

[37] E. Criswell and J. R. Cram, Eds., *Cram's introduction to surface electromyography*, 2nd ed. Sudbury, MA: Jones and Bartlett, 2011.

[38] E. Park and S. Meek, "Fatigue compensation of the electromyographic signal for prosthetic control and force estimation," *IEEE Transactions on Biomedical*

Engineering, vol. 40, no. 10, pp. 1019–1023, Oct. 1993. [Online]. Available: http://ieeexplore.ieee.org/document/247800/

[39] A. Phinyomark, C. Limsakul, and P. Phukpattaranont, "Application of wavelet analysis in EMG feature extraction for pattern classification," *Measurement Science Review*, vol. 11, no. 2, 2011. [Online]. Available: http://www.degruyter.com/view/j/msr.2011.11.issue-2/v10048-011-0009-y/v10048-011-0009-y.xml

[40] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, and Y. Laurillau, "EMG feature evaluation for improving myoelectric pattern recognition robustness," *Expert Systems with Applications*, vol. 40, no. 12, pp. 4832–4840, 2013. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0957417413001395

[41] A. Phinyomark, A. Nuidod, P. Phukpattaranont, and C. Limsakul, "Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification," *Electronics and Electrical Engineering*, vol. 122, no. 6, 2012. [Online]. Available: http://www.eejournal.ktu.lt/index.php/elt/article/view/1816

[42] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0957417412001200

[43] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, and Y. Laurillau, "Feature extraction of the first difference of EMG time series for EMG pattern recognition," *Computer Methods and Programs in Biomedicine*, 2014. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0169260714002478

110

[44] B. Hudgins, P. Parker, and R. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993. [Online]. Available: http://ieeexplore.ieee.org/document/204774/

[45] Thalmic Labs Inc. (2013) Myo SDK manual. [Online]. Available: https://developer.thalmic.com/docs/api_reference/platform/index.html

[46] F. Riillo, L. Quitadamo, F. Cavrini, E. Gruppioni, C. Pinto, N. C. Past, L. Sbernini, L. Albero, and G. Saggio, "Optimization of EMG-based hand gesture recognition: Supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees," *Biomedical Signal Processing and Control*, vol. 14, pp. 117–125, 2014. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S174680941400113X

[47] M. V. Narici and M. D. de Boer, "Disuse of the musculo-skeletal system in space and on earth," *European Journal of Applied Physiology*, vol. 111, no. 3, pp. 403–420, Mar. 2011.

[48] A. Subasi, "Classification of emg signals using pso optimized svm for diagnosis of neuromuscular disorders," *Computers in biology and medicine*, vol. 43, no. 5, pp. 576–586, 2013.

[49] E. Lamounier, A. Soares, A. Andrade, and R. Carrijo, "A virtual prosthesis control based on neural networks for EMG pattern classification," in *Proceedings of the Artificial Intelligence and Soft Computing, Canada*, 2002. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.6816&rep=rep1&type=pdf

[50] R. Boostani and M. H. Moradi, "Evaluation of the forearm EMG signal features for the control of a prosthetic hand," *Physiological*

measurement, vol. 24, no. 2, p. 309, 2003. [Online]. Available: http://iopscience.iop.org/0967-3334/24/2/307

[51] Z. Liu and Z. Luo, "Hand motion pattern classifier based on emg using wavelet packet transform and lvq neural networks," in *IT in Medicine and Education, 2008. ITME 2008. IEEE International Symposium on*, vol. 18, no. 2, Dec 2008, pp. 28–32.

[52] M. Oskoei and Huosheng Hu, "Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1956–1965, Aug. 2008.

[53] W.-L. Zuo, Z.-Y. Wang, T. Liu, and H.-L. Chen, "Effective detection of parkinson's disease using an adaptive fuzzy k-nearest neighbor approach," *Biomedical Signal Processing and Control*, vol. 8, no. 4, pp. 364–373, 2013.

[54] B. Karlik, "Machine Learning Algorithms for Characterization of EMG Signals," *International Journal of Information and Electronics Engineering*, vol. 4, no. 3, pp. 189–194, 2014.

[55] E. Rejc, P. E. di Prampero, S. Lazzer, B. Grassi, B. Simunic, R. Pisot, G. Antonutto, and M. Narici, "Maximal explosive power of the lower limbs before and after 35 days of bed rest under different diet energy intake," *European Journal of Applied Physiology*, 2014. [Online]. Available: http://link.springer.com/10.1007/s00421-014-3024-5

[56] E. Rejc, S. Lazzer, G. Antonutto, M. Isola, and P. E. di Prampero, "Bilateral deficit and EMG activity during explosive lower limb contractions against different overloads," *European Journal of Applied Physiology*, vol. 108, no. 1, pp. 157–165, 2010. [Online]. Available: http://link.springer.com/10.1007/s00421-009-1199-y

[57] C. J. De Luca, L. Donald Gilmore, M. Kuznetsov, and S. H. Roy, "Filtering the surface EMG signal: Movement artifact and baseline noise contamination," *Journal of Biomechanics*, vol. 43, no. 8, pp. 1573–1579, 2010. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0021929010000631

[58] I. Daubechies, *Ten lectures on wavelets*, ser. CBMS-NSF regional conference series in applied mathematics. Philadelphia, Pa: Society for Industrial and Applied Mathematics, 1992, no. 61.

[59] A. Soares, A. Andrade, E. Lamounier, and R. Carrijo, "The Development of a Virtual Myoelectric Prosthesis Controlled by an EMG Pattern Recognition System Based on Neural Networks," *Journal of Intelligent Information Systems*, vol. 21, no. 2, pp. 127–141, 2003. [Online]. Available: http://dx.doi.org/10.1023/A:1024758415877

[60] Han-Pang Huang and Chun-Yen Chen, "Development of a myoelectric discrimination system for a multi-degree prosthetic hand," vol. 3. IEEE, 1999, pp. 2392–2397. [Online]. Available: http://ieeexplore.ieee.org/document/770463/

[61] D. G. Ruegg, T. H. Kakebeeke, J.-P. Gabriel, and M. Bennefeld, "Conduction velocity of nerve and muscle fiber action potentials after a space mission or a bed rest," *Clinical Neurophysiology*, vol. 114, no. 1, pp. 86–93, 2003.

[62] D. L. Belavy, J. K.-F. Ng, S. J. Wilson, G. Armbrecht, D. F. Stegeman, J. Rittweger, D. Felsenberg, and C. A. Richardson, "Influence of prolonged bed-rest on spectral and temporal electromyographic motor control characteristics of the superficial lumbo-pelvic musculature," *Journal of Electromyography and Kinesiology*, vol. 20, no. 1, pp. 170–179, 2010.

[63] E. Mulder, K. Gerrits, B. Kleine, J. Rittweger, D. Felsenberg, A. de Haan, and D. Stegeman, "High-density surface EMG study on the time course of central nervous and peripheral neuromuscular changes during 8weeks of bed rest with or without resistive vibration exercise," *Journal of Electromyography and Kinesiology*, vol. 19, no. 2, pp. 208–218, Apr. 2009. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1050641107000648

[64] G. Ferretti, H. E. Berg, A. E. Minetti, C. Moia, S. Rampichini, and M. V. Narici, "Maximal instantaneous muscular power after prolonged bed rest in humans," *Journal of Applied Physiology*, vol. 90, no. 2, pp. 431–435, 2001.

[65] M. D. de Boer, O. R. Seynnes, P. E. di Prampero, R. Piot, I. B. Mekjavi, G. Biolo, and M. V. Narici, "Effect of 5 weeks horizontal bed rest on human muscle thickness and architecture of weight bearing and non-weight bearing muscles," *European Journal of Applied Physiology*, vol. 104, no. 2, pp. 401–407, Sep. 2008. [Online]. Available: http://link.springer.com/10.1007/s00421-008-0703-0

[66] B. A. Alkner and P. A. Tesch, "Knee extensor and plantar flexor muscle size and function following 90 days of bed rest with or without resistance exercise," *European Journal of Applied Physiology*, vol. 93, no. 3, pp. 294–305, 2004.

[67] D. L. Belav, T. Miokovic, G. Armbrecht, C. A. Richardson, J. Rittweger, and D. Felsenberg, "Differential atrophy of the lower-limb musculature during prolonged bed-rest," *European Journal of Applied Physiology*, vol. 107, no. 4, pp. 489–499, Nov. 2009.

[68] J. J. Bloomberg, C. D. Batson, R. E. Buxton, A. H. Feiveson, I. S. Kofman, S. M. Lee, C. A. Miller, A. P. Mulavara, B. T. Peters, T. Phillips, and others, "UNDERSTANDING THE EFFECTS OF

LONG-DURATION SPACE FLIGHT ON ASTRONAUT FUNCTIONAL TASK PERFORMANCE," in *American Astronautical Society - ISS Research and Development Conference*, 2014. [Online]. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20140005022.pdf

[69] V. Koppelmans, B. Erdeniz, Y. E. De Dios, S. J. Wood, P. A. Reuter-Lorenz, I. Kofman, J. J. Bloomberg, A. P. Mulavara, and R. D. Seidler, "Study protocol to examine the effects of spaceflight and a spaceflight analog on neurocognitive performance: extent, longevity, and neural bases," *BMC Neurology*, vol. 13, no. 1, 2013. [Online]. Available: http://bmcneurol.biomedcentral.com/articles/10.1186/1471-2377-13-205

[70] M. Shelhamer, "Trends in sensorimotor research and countermeasures for exploration-class space flights," *Frontiers in Systems Neuroscience*, vol. 9, p. 115, Aug. 2015. [Online]. Available: http://journal.frontiersin.org/Article/10.3389/fnsys.2015.00115/abstract

[71] I. S. Kofman, M. Reschke, J. M. Cerisano, E. Fisher, E. Lawrence, B. T. Peters, and J. Bloomberg, "Changes in Jump-Down Performance After Space Flight: Short- and Long-Term Adaptation," in *18th IAA Humans in Space Symposium, Conference on*, Houston, TX, Jan. 2010. [Online]. Available: https://ntrs.nasa.gov/search.jsp?R=20100038359

[72] G. M. Jones and D. G. D. Watt, "Melvill jones muscular control of landing from unexpected falls in man," *Journal of Physiology*, vol. 219, no. 3, pp. 729–737, 1971.

[73] D. R. Ortega, E. C. Rodrguez Bes, and F. J. Berral de la Rosa, "Analysis of the Vertical Ground Reaction Forces and Temporal Factors in the Landing Phase of a Countermovement Jump," *Journal of Sports Science*

& *Medicine*, vol. 9, no. 2, pp. 282–287, Jun. 2010. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3761745/

[74] G. M. Jones and D. G. D. Watt, "Observations on the control of stepping and hopping movements in man," *The Journal of Physiology*, vol. 219, no. 3, pp. 709–727, 1971. [Online]. Available: http://doi.wiley.com/10.1113/jphysiol. 1971.sp009684

[75] A. Cullen and H. Frey, *Probabilistic Techniques in Exposure Assessment*, 1st ed. Plenum Press, 1999.

[76] M. L. Delignette-Muller and C. Dutang, "fitdistrplu : An r package for fitting distributions," *Journal of Statistical Software*, vol. 64, no. 4, 2015. [Online]. Available: http://www.jstatsoft.org/v64/i04/

[77] M. Baker, "Statisticians issue warning over misuse of P values," *Nature*, vol. 531, no. 7593, pp. 151–151, Mar. 2016. [Online]. Available: http://www.nature.com/doifinder/10.1038/nature.2016.19503

[78] T. Dahiru, "P - value, a true test of statistical significance? A cautionary note," *Annals of Ibadan Postgraduate Medicine*, vol. 6, no. 1, pp. 21–26, Jun. 2008.

[79] S. S. Chen and D. L. Donoho, "Application of basis pursuit in spectrum estimation," in *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, vol. 3. IEEE, 1998, pp. 1865–1868.

[80] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.

[81] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010. [Online]. Available: http://www.jstatsoft.org/v33/i01/

[82] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, ser. Springer Texts in Statistics. Springer New York, 2013, vol. 103, DOI: 10.1007/978-1-4614-7138-7. [Online]. Available: http://link.springer.com/10.1007/978-1-4614-7138-7

[83] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees.* CRC press, 1984, ch. 3.

[84] S. J. Wood, M. F. Reschke, and F. Owen Black, "Continuous equilibrium scores: Factoring in the time before a fall," *Gait & Posture*, vol. 36, no. 3, pp. 487–489, 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0966636212001427

[85] C. Johnston, "Netflix never used its $1 million dollar algorithm due to engineering costs," *Wired Magazine*, 2012. [Online]. Available: https://www.wired.com/2012/04/netflix-prize-costs/

[86] J. Shawe-Taylor and N. Cristianini, *Kernel methods for pattern analysis.* Cambridge, UK ; New York: Cambridge University Press, 2004.

[87] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers." ACM Press, 1992, pp. 144–152. [Online]. Available: http://portal.acm.org/citation.cfm?doid=130385.130401

[88] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. [Online]. Available: http://link.springer.com/10.1007/BF00994018

[89] Z. Weinberg, "SVM separating hyperplanes," 2012. [Online]. Available: https://en.wikipedia.org/wiki/Support_vector_machine#/media/File:Svm_separating_hyperplanes_(SVG).svg

[90] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, Jun. 2006. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S016786550500303X

[91] I. Guyon, B. Boser, and V. Vapnik, "Automatic capacity tuning of very large vc-dimension classifiers," in *Advances in Neural Information Processing Systems*. Morgan Kaufmann, 1993, pp. 147–155.

[92] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[93] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-layer learning revisited: a stepwise procedure for building and training a neural network," in *Neurocomputing*, F. F. Souli and J. Hrault, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, pp. 41–50. [Online]. Available: http://www.springerlink.com/index/10.1007/978-3-642-76153-9_5

[94] K. E. Madden, A. D. Deshpande, B. J. Peters, J. M. Rogers, E. A. Laske, and E. R. McBryan, "The influence of robotic assistance on reducing neuromuscular effort and fatigue during extravehicular activity glove use," in *International Conference on Environmental Systems (ICES)*, 07/2017 2017.

[95] K. Madden, B. Peters, and A. Deshpande, "The Influence of Robotic Grip Augmentation in Reducing Spacesuit Glove-Induced Muscle Fatigue during a Repetitive Gripping Task," Jan. 2017, poster presentation at the 2017 NASA Human Research Projects and International Workshop Conference, Galveston, TX.

[96] R. Kristeva-Feige, C. Fritsch, J. Timmer, and C.-H. Lcking, "Effects of attention and precision of exerted force on beta range EEG-EMG synchronization during a maintained motor contraction task," *Clinical Neurophysiology*, vol. 113, no. 1, pp. 124–131, Jan. 2002. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245701007222

[97] P. Grosse, "EEG-EMG, MEG-EMG and EMG-EMG frequency analysis: physiological principles and clinical applications," *Clinical Neurophysiology*, vol. 113, no. 10, pp. 1523–1531, Oct. 2002. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245702002237

[98] M. D. Luciw, E. Jarocka, and B. B. Edin, "Multi-channel EEG recordings during 3,936 grasp and lift trials with varying weight and friction," *Scientific Data*, vol. 1, p. 140047, Nov. 2014, doi:10.1038/sdata.2014.47. [Online]. Available: http://www.nature.com/articles/sdata201447

[99] MATLAB Documentation, "Wavelet Scalogram Using 1d Wavelet Analysis," note that this document appears to only be available in the MATLAB help file itself.

[100] A. Subasi, "Classification of EMG signals using combined features and soft computing techniques," *Applied Soft Computing*, vol. 12, no. 8, pp. 2188–2198, 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1568494612001330

[101] G. Strang and T. Nguyen, *Wavelets and filter banks*. Wellesley, MA: Wellesley-Cambridge Press, 1997.

Appendix A: Acronyms Used

Acronyms employed throughout this work are defined here for convience.

ANN — Artificial Neural Network

APA — Anticipatory Postural Adjustment

AR — Autoregression (coefficient)

AUC — Area Under Curve

CWT — Continuous Wavelet Transform

DWT — Discrete Wavelet Transform

EEG — Electroencephalogram

EMG — Electromyography

FSR — Functional Stretch Reflex

fMRI — Functional Magnetic Resonance Imaging

FT — Fourier Transform

FTT — Functional Task Test

FWHM — Full-Width-at-Half-Maximum

GM — Gastrocnemius Medialis

IAT — Impact Absorption Time

iEMG — integrated Electromyography

ISS — International Space Station

JDT — Jump-Down Test

JSC — Johnson Space Center

k-NN — k-Nearest Neighbors

LASSO — Least Absolute Shrinkage and Selection Operator

LDA — Latent Dirichlet Algorithm

MAV — Mean Absolute Value

MF — Median Frequency (also, F_M)

MG – Medial Gastrocnemius

MSR — Monosynaptic Stretch Reflex

MUAP — Motor Unit Action Potential

MVC — Maximal Voluntary Contraction

PCA — Principal Component Analysis

PL — Preparatory Latency

RBF — Radial Basis Function

RF — Rectus Femoris

RMS — Root-Mean-Square

ROC — Receiver Operator Characteristics

SDK — Software Developers Kit

sEMG — Surface Electromyography

STD — Standard Deviation

SVM — Support Vector Machine

TA — Tibialis Anterior

VL — Vastus Lateralis

WL — Waveform Length

WNN — Wavelet Neural Network

WPTE — Wavelet Packet Transform Energy

WT — Wavelet Transform

ZC — Zero Crossings

Appendix B: Wavelet Basics

Many of the signals analyzed in the physiological domain are not necessarily stationary. This is especially, and perhaps more obviously, the case for the sEMG signals of the JDT experiments, as there are abrupt, transient changes in the sEMG signal surrounding the touchdown point. Wavelet analysis offers an advantage over frequency analysis (as with, e.g., the Fourier Transform) in that it these transient elements can be investigated.

For the sake of comparison, two equations are presented here.

$$\hat{f}(\omega) = \int_{\mathbb{R}^n} f(x) e^{-i\omega \cdot x} \, dx \tag{22}$$

$$X(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \Psi^* \left(\frac{t - b}{a} \right) x(t) \, dt \tag{23}$$

Equation 22 is a definition of the familiar Fourier Transform, with the angular frequency ω, provided here for sake of comparison. The wavelet transform of Equation 23 shows a similar input/output mapping scheme. However, the wavelet transform is a convolution of the input signal with a set of functions consisting of dilations and scales of a mother wavelet. In Equation 23, a and b represent scale (dilation of mother wavelet) and time (translation across signal length) respectively. The Ψ^* indicates the complex conjugate of the wavelet function.

Figure 1 illustrates the immediate advantage of the wavelet transform in analyzing nonstationary signals (top, red). The signal has an abrupt change that a Fourier

transform would not capture. The scalogram (bottom) indicates the signal's energy coefficients which give measures of correlation between a chosen mother wavelet (a Gaussian in this case) which is scaled and translated across the signal. The scale values correspond inversely to the frequency for the associated signal energy. The particular conversion between scale and frequency is dependent on the chosen wavelet. Broadly, bright patches indicate a strong correlation between signal and wavelet. One advantage of the wavelet method is that we can selectively eliminate portions of the signal by removing its frequency contribution if noise is determined to be present at some particular scale/frequency region.
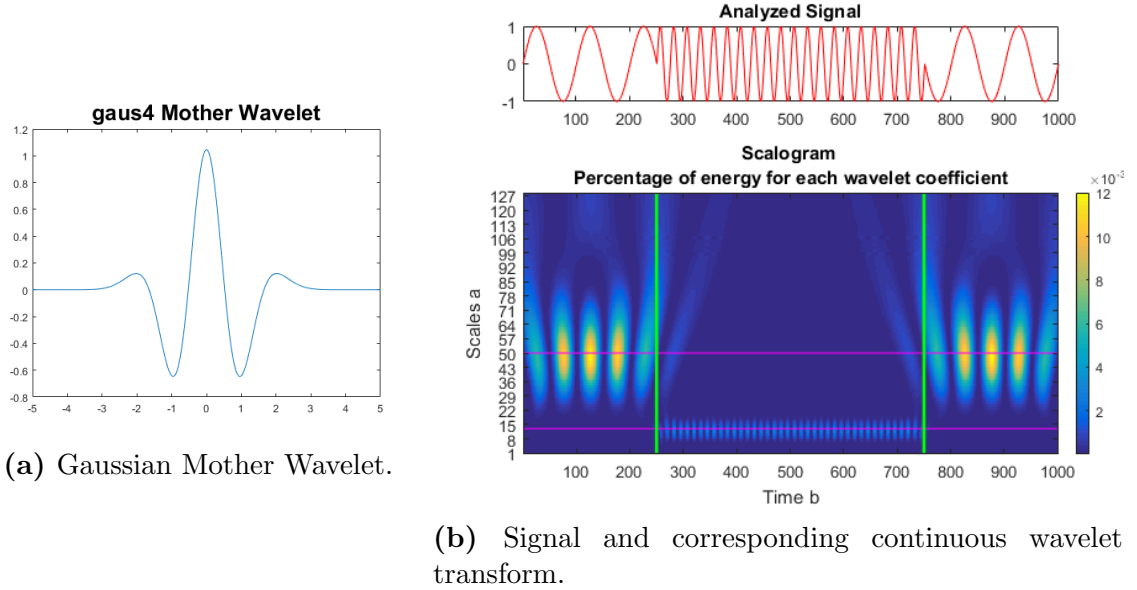

(a) Gaussian Mother Wavelet.

(b) Signal and corresponding continuous wavelet transform.

Figure 1. Left: Mother Wavelet "gaus4". Right: Signal and corresponding continuous wavelet transform. Adapted from Mathworks [99]

However, in practice, it is often the Discrete Wavelet Transform (DWT) that is used. In the bedrest work described in Chapter 3 it is the DWT that is employed. This technique is also relied upon in a number of other publications cited in this work [31, 39, 100].

The DWT is wavelet transform in which the wavelet is discretely sampled. The DWT of some signal, x, is passed through high-pass filters to acquire high-frequency

(or detail) coefficients and low-pass filters to acquire low-frequency (or approxima-tion) coefficients. This action is performed in a cascaded manner to acquire a set of discretely sampled wavelet transformed signals. Figure 2 lays out an example of the filtering scheme and its corresponding wavelet cascade. Equations 24 and 25 describe the low-pass and high-pass filter convolutions respectively.

$$y_{\text{low}}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \tag{24}$$

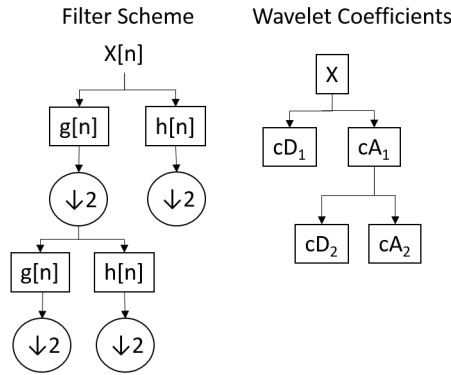$$y_{\text{high}}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \tag{25}$$

Figure 2. Filter and Coefficient Scheme for the Discrete Wavelet Transform

In the bedrest work of Chapter 3, a 3-level cascade was used. Each of these wavelet coefficient vectors (e.g., cD_1) resulting from the cascade can be inverted such that the original signal, x, can be reconstructed as

$$x = D_1 + D_2 + D_3 + A_3 \tag{26}$$

Where, for example, D_2 is the inverse wavelet transform of cD_2 and the lowercase "c" indicates that we are discussing the wavelet coefficients, while the "D" without the lowercase "c" indicates that we are discussing the reconstruction of that signal element from the coefficients. Put another way, D_2 has the same units as the original signal, whereas the coefficients, cD_2 are an abstract representation.

More detailed overviews of wavelet analysis and the two transforms described briefly here can be found in texts written by Ingrid Daubechies [58], for a more historical work, and Strang and Nguyen [101] for a more course-appropriate textbook.

CURRICULUM VITAE

Robert Stallard

Education

Ph.D., University of Louisville

Department of Electrical and Computer Engineering

Louisville, KY, May 2018

M.S., University of Louisville,

Department of Physics and Astronomy

Louisville, KY, May 2013

B.S., University of Louisville

Department of Physics and Astronomy

Louisville, KY, May 2010

Professional Positions

Graduate Research Fellow, Department of Electrical and Computer Engineering, 2013-2018

Graduate Teaching Assistant, Department of Physics and Astronomy, 2011-2013

Research Assistant, Conn Center for Renewable Energy Research, 2010-2011

Undergraduate Research Assistant, Department of Physics and Astronomy, 2008-2010

Honors and Awards

2016 NASA KY Graduate Fellowship

2015 NASA KY RIDG Grant

2013 University Graduate Fellowship

2011 Graduate Teaching Assistantship

2010 Bennet Award for Outstanding Physics Undergraduate

Journal and Conference Publications

R. Stallard, E. Rejc, and K.C. Welch, Wavelet-derived features as indicators of physiological changes induced by bed rest, Expert Systems with Applications, vol. 90, pp. 232240, 2017.

R. Stallard and K.C. Welch, Algorithmic Investigation of Sensorimotor Performance Using Biopotentials, presented at the American Society for Gravitational and Space Research, Seattle, Washington, 2017

R. Stallard and K.C. Welch, Algorithmic Classification Based on Astronaut Sensorimotor Performance Before and After Spaceflight, presented at the Kentucky Academy of Sciences, University of Louisville, 2016

R. Stallard, Feature Extraction, Transformation, and Algorithmic Classification of Astronaut Sensorimotor Performance, presented at the NASA Human Research Program and International Workshop, Galveston, Texas, 2017

A. N. Sidorov., K. Gaskill, M. Buongiorno Nardelli, J.L. Tedesco, R.L. Myers-Ward, C.R. Eddy, T. Jayasekera, K.W. Kim, R. Jayasingha, A. Sherehiy, R. Stallard, and G.U. Sumanasekera, Charge transfer equilibria in ambient-exposed epitaxial graphene on (0001) 6 H-SiC, Journal of Applied Physics, vol. 111, no. 11, p. 113706, 2012

A. N. Sidorov A. Sherehiy, R. Jayasinghe, R. Stallard, D.K. Benjamin, Q. Yu, Z. Liu, W. Wu, H. Cao, Y. P. Chen, Z. Jiang, and G. U. Sumanasekera, Thermoelectric power of graphene as surface charge doping indicator, Applied Physics Letters, vol. 99, no. 1, p. 013115, 2011