A Data Driven Framework to Predict the Fatigue among Manufacturing Workers Using Wearable Sensors

by

Zahra Sedighi Maman

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 4, 2018

Keywords: Fatigue Management, Statistical Surveillance, Analytics, Feature Selection, Parallel Computation, American Society of Safety Engineers

Approved by

Fadel Megahed, Chair, Assistant Professor of Industrial and Systems Engineering
Lora Cavuoto, Assistant Professor of Industrial and Systems Engineering
Richard Sesek, Associate Professor of Industrial and Systems Engineering
Sean Gallagher, Associate Professor of Industrial and Systems Engineering
Aleksandr Vinel, Assistant Professor of Industrial and Systems Engineering

Abstract

Worker fatigue has been known as a significant phenomenon in the manufacturing occupations. In these occupations, physical fatigue is a challenging ergonomic/safety issue as it lowers productivity and boosts the incidence of injuries. The objective of this dissertation is to prevent the fatigue occurrence in the manufacturing occupations by monitoring the individual's body using the wearable sensors on the wrist, torso, ankle, and hip coupled with a heart rate sensor. Specifically, this research, 1) examines whether the commercially wearable sensors, extracting appropriate ergonomic-related metrics, can be used to detect the occurrence of fatigue on an individualized level for different occupational tasks, 2) proposes a comprehensive framework consisting of four phases including detection, identification, diagnosis, and recovery to manage fatigue in manufacturing occupations using wearable sensors. Overall, the goal of this research is to develop analytical models that present important findings for accident and injury prevention by managing fatigue in the manufacturing occupations.

Acknowledgments

Firstly, I would like to sincerely thank my advisors Dr. Fadel Megahed, and Dr. Lora Cavuoto for their continuous mentor-ship during my Ph.D. career. I learned to perform research because of its impact to help humanity have a much better life. I would like to say my special thanks to Dr. Ying-Ju Chen for her endless help and encouragement in this research. I like to acknowledge my committee members Dr. Richard Sesek, Dr. Sean Gallaghar, and Dr. Alexandre Vinel for their helpful advice, recommendations and continuous support during my Ph.D. career and when my advisor was not at Auburn University. Similarly, I would like to thank my university examiner, Dr. Mark Clark, who offered valuable insights towards my dissertation.

I would like to present my special thanks to Jonathan Heath for all his endless help, support, self-devotion, and inspiration all the time. I would like to reveal my thanks to Dr. Rassoul Noorossana, my master degree advisor for all the lessons he taught me. Last but not least, I would like to exhibit my thanks to my special teacher Bahar Nafez for endlessly supporting me, without having her in my life I would not have had a Ph.D. in the U.S.

Much appreciation goes towards faculty members in Industrial and Systems Engineering at Auburn University, as well as so much appreciation to the staff who were all paramount to the completion of my degree.

I would like to dedicate this dissertation to my family, my mother Mina, my father Bayanollah, and my brothers Mehrdad and Masoud for their loving support, valuable self-devotion, and encouragement.

This dissertation is submitted to the Graduate Faculty of Auburn University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Industrial and Systems Engineering. It involves the research work performed starting in late August 2015 until August 2018. Some of the ideas in this work, as in any research, are based on the analysis of others, and I have done my best to provide references to these sources. In the spring semester of 2015,

I started to work with Dr. Megahed. Since I was interested in data analysis, he decided to assign me to the collaborative research project between Auburn University and University at Buffalo which granted by American Society of Safety Engineers Foundation titled ASSIST: Advancing Safety Surveillance using Individualized Sensor Technology on August 2015 for three years. Dr. Cavuoto is a principal investigator of this project from University at Buffalo. The goal of this project was to find a solution for detecting fatigue in manufacturing occupations with wearable sensors. In the fall semester of 2015, we started to work with the pilot data collected beforehand by one of Dr. Cavuoto's Ph.D. students for manual material handling. Dr. Megahed suggested that I review the current literature as a starting point and boost my knowledge in fatigue-related researches. In the meanwhile, Dr. Cavuoto and her Ph.D. Student Amir Baghdadi outlined the design of manufacturing tasks for the lab-based experiment, and Dr. Cavuoto purchased the wearable sensors for this research. In the spring semester of 2016, I traveled to Buffalo to enhance my knowledge about working with wearable sensors, data collection and the experiments designed for this research. I was involved in collecting data with wearable sensors during manufacturing tasks.

Then, Amir began to collect data from the students and manufacturing workers who were currently working in manufacturing environments. After the data collection for eight participants, we started to examine this exciting topic on whether using wearable sensors can predict fatigue in manufacturing tasks or not. I wrote the original draft; however, since it was my first paper in this field, Dr. Cavuoto and Dr. Megahed used a lot of their expertise to develop this paper further. They led me to appropriate feature generation, model selection, and they assisted me in writing this paper. Also, Mohammad Alamdar Yazdi helped me with cleaning data and data visualization. It should be noted that the Applied Ergonomics reviewers and the editor have also provided valuable feedback, which was applied in the current form of the paper that is shown in Chapter 2. I presented this study at the 2016 INFORMS Annual Meeting in Nashville, Tennessee, the 2016 Graduate Research Showcase at Auburn University, This is Research: Symposia 2016 at Auburn University, and Academic Research Colloquium at University of Dayton 2017.

In spring 2017, while Amir was collecting data for more participants, we started to think about using more prediction models to improve our prediction accuracy. Dr. Cavuoto and Dr. Megahed advised me through several alternatives including predictive variables, prediction models, and modeling procedures. Dr. Megahed introduced me to Dr. Ying-Ju Chen to assist me in the use of the Ohio Supercomputer Center (OSC) [1] in order to speed up computation procedure. Dr. Chen helped us with efficient coding and parallelized modeling to submit to the OSC, where the time needed for modeling decreased significantly. Dr. Chen, with her statistical background, helped us to boost the performance of our developed models for fatigue prediction. After I finished the modeling for fatigue detection using data from 23 participants, I presented our modeling results at the 2017 INFORMS Houston, Texas and I used them in my dissertation proposal. Later on, we decided to go beyond only developing fatigue detection models. Therefore, we began to think about providing a Fatigue Management framework. I wrote the first draft of the paper, then Amir, Dr. Chen, Seamus Lombardo an undergraduate researcher at University at Buffalo, Dr. Cavuoto and Dr. Megahed provided their feedback which I employed to improve the paper further. This resulted in the current form of the paper presented in chapter 3.

Completing this dissertation was extremely challenging; however, in the process, I developed my knowledge not only in data analytics but also in the ergonomic and safety field. The multi-disciplinary focus of this dissertation has led to the incorporation of many statistical, ergonomic, and industrial engineering concepts, which made several valuable contributions to the literature.

The views and opinions of this research are the obligation of the authors exclusively and do not illustrate the views of the ASSE.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Problem Description and Significance

Occupational accidents and injuries stand for a challenging safety problem. The UNs International Labour Organization (ILO) estimates that there are 313 million occupational accidents causing injuries. In addition, it estimates 160 million occupational diseases annually, ensuing in 2.3 million worker related fatalities [3]. The ILO estimated that occupational accidents and diseases cost the global economy $3 trillion annually, which was 4% of global Gross Domestic Product in 2015. It is critical to note that workplace fatalities are only one indicator of occupational safety performance. The Bureau of Labor Statistics collects detailed non-fatal occupational injuries and diseases and found that the rates/amount of affected workers is much higher, with an occurrence rate of 9.3 per 100 full-time workers in 2015 [4]. Moreover, the Bureau of Labor Statistics estimated that there were over 1.5 million occupational accidents plus injuries, resulting in a median leave of 10 days from work in 2015 [4].

Worker fatigue is a common root-cause of occupational injuries since it results in impaired musculoskeletal function and increased risk-taking behavior. Physically demanding work subject to long duration, repetition, or high workload in the long term without proper recovery can result in fatigue [5]. In a study involving 606 construction workers, Zhang et al. (2015) [6] showed that there is a significant negative correlation between being fatigued and physical/cognitive function. Fatigue has adverse consequences on motor control and strength capability [7]. It can lead to the loss of balance [8], which is the main cause of slips and fall fatalities in the workplaces [9]. Additionally, fatigue impairs decision-making and concentration, causing a decrease in performance and alertness [10, 11]. These consequences result in increased injury risk, a reduction of work performance and job quality as well as an increment in the accidents related to human errors [12, 13]. Fatigue prevalence is estimated to be 57% in U.S. advanced manufacturing workers [14], 60% of the Japanese working populations [15], and 15-20% of the Canadian workforce [16]. Ricci et al. (2007) [17] reported that the health-related loss in productivity time for fatigued workers exceeds double their non-fatigued counterparts. The financial ramifications of fatigue outcomes are estimated to cost U.S. employers approximately $136 billion annually [17]. Due to its high prevalence and severe consequences on the well-being of the industrial workers, fatigue has turned into a critical research topic.

The goal of this dissertation is to investigate how to detect fatigue in the manufacturing occupations by monitoring the individual's body using wearable sensors. Although fatigue has been studied widely, there is not a standard/accepted definition of fatigue [18]. Lu et al. (2017) [14] used the term fatigue to express the lower level of strength and capacity resulting from work activities. In Figure 1.1 the fatigue mechanism is shown for occupational tasks.

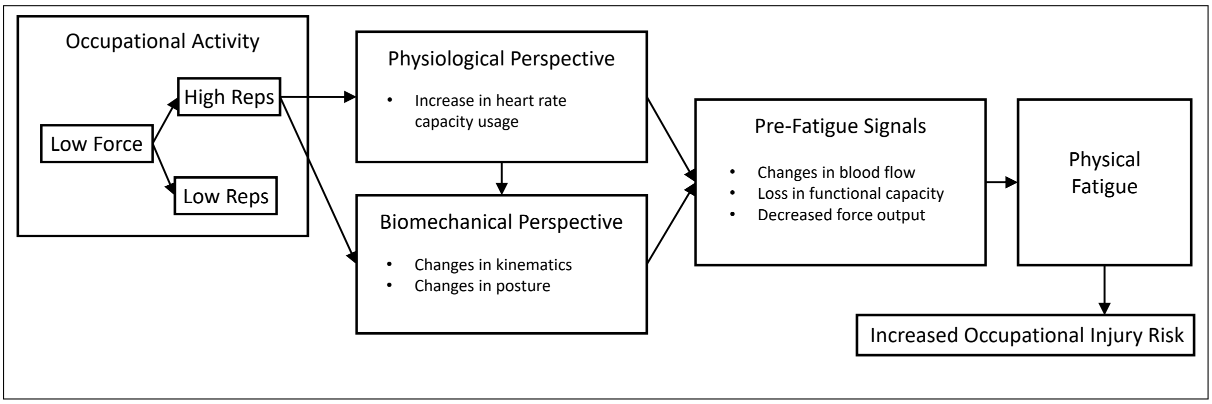


Figure 1.1: Fatigue process in occupational activities

I focus on manufacturing tasks which can be characterized by repetitive physical work with low force and extended working hours, high reliance on shift work, awkward postures, and insufficient recovery time which have been linked to fatigue and injury. The leading causes of fatigue in occupational activities can be classified into at least two categories: a) physiological, and b) biomechanical which are shown in Figure 1.1.

From a physiological point of view, the capacity of the body for a sustained physical activity depends on the heart's capacity [19] as indicated in Figure 1.1. The repetitive physical activities are commonly characterized by high energy demand and restricted blood flow during the activity [20], due to skeletal muscle "anaerobiosis" [21]. From a biomechnical point of view, the product between force and repetition is one of the main drivers in fatigue development [22, 23]. The repetitive occupational tasks will result in changes in kinematics [24, 25] and postures [26] (Figure 1.1).

There are several pre-signal indicators because of physiological and biomechanical changes including a decrease in blood flow [20], loss in functional capacity [27] and drop in force output [28], which result in physical fatigue and potential injury as shown in the last phases of the fatigue process in Figure 1.1.

Traditional approaches for exposure evaluation commonly rely on the visual inspection performed by a safety observer. However, due to limitations of an individual observer, sampling methods are utilized such that only a few workers are regularly observed and just for a comparatively short duration [29, 30]. Common observational methods concentrate on one main risk factor, such as posture or force, or a combined group of factors for a repetitive task as with the NIOSH Work Practices Guide [31]. These approaches are inadequate for fatigue detection because they fail to detect the interactive nature of many of the risk factors as well as the variability of the task performed [22, 32, 33]. Additionally, these approaches do not consider the characteristics of the individual, beyond general anthropometric and demographic attributes, such as height, age or body weight [34]. Current approaches to fatigue monitoring and detection rely either on fitness-for-duty tests to conclude if the worker has enough capacity before beginning work, monitoring of sleep, monitoring of brain activation [35] or monitoring local muscle fatigue [36]. The growing availability of sensing technologies, including wearable devices [37, 38, 39, 40], with health information has the potential to provide real-time monitoring, recording, and communication of individuals' physical and environmental exposures [31, 37, 41].

Researchers recommend that the adaptation of these technologies to a work environment would be a feasible and potentially valuable addition for injury prevention [37, 41]. Most of the current studies on the applications of wearable sensors at the workplace have concentrated on posture analysis [42] or task classification [37]. However, there is a lack in a) how to best combine the data from multiple wearable sensors, b) successful use of these data for fatigue detection, c) identifying successful measurement variables for fatigue detection, d) identifying more informative sensors, and e) using wearable sensors in practice.

This dissertation aims to develop and examine the use of wearable sensors for data-driven occupational exposure assessments in manufacturing occupations. It provides a measurement of fatigue status along with the identification of informative sensors that allows data-driven decisions to intervene in manufacturing work environments, which maximize both safety and performance.

1.1.1 Sensor Selection

The selection of accelerometer locations has a significant impact on the accuracy of results [43]. Sensor placement relies on the task and its corresponding risk. Most of the manufacturing tasks, such as manual material handling and supply insertion, include a high amount of lifting and walking. Studies show that the lifting strategies change during the activity [44], for example, individuals change their strategy from bending the knees to bending the trunk [26]. Therefore, the lifting task is characterized by an increase in trunk bending torque and trunk flexion. Lifting tasks are accurately described by tracking the upper body (torso) and wrist with wearable sensors [45, 43]. During repetitive lifting, individuals compensate upon being fatigued by decreasing the hip movement while increasing trunk motion [46, 47]. By using a sensor placed on the hip, the motion produced through lifting, walking and standing can be detected [48]. Additionally, the ankles play an important role in maintaining balance in the sagittal plane during walking and standing [49]. Studies have been conducted to monitor physical fatigue while walking through the use of an accelerometer placed on the ankle [50]. For practical use, it is favorable to use a smaller number and simpler set of sensors on non-intrusive body locations and approximate the kinematics of other target locations using estimation techniques [50]. This research uses data collected from four accelerometers placed on the hip, wrist, torso, and ankle, along with a heart rate sensor, to recognize motion and posture.

It should be noted that past studies do not address which overall combination of sensors and sensor locations provide the best results for fatigue detection.

1.1.2 Sensor Measurements

There are several non-invasive measurements for occupational fatigue detection that enable continuous monitoring of movement such as acceleration and jerk which are used to measure the intensity of a physical activity and offer motion analysis [51]. Some studies have observed the effects of physical activity on musculoskeletal pain among worker populations [52] and found that high intensity occupational physical activity has been associated with increased risks of several chronic health conditions and may be deleterious to the worker's health [53, 54, 55]. In addition, the other type of measurements is kinematic which reveal the structure and function of the foot, posture, and body movement [56, 57]. Evaluating the body kinematics can be used for the purpose of postural stability analysis to reduce the rate of fall risks [58, 59]. Studies show that in a lifting task, both movement kinematics and muscle activation patterns change during physical fatigue development [24] and individuals continually change their movement strategies to maintain task performance [24, 25]. These kinematic changes due to physical fatigue may increase the potential risk of injury [60, 46]. For instance, increases in torso kinematics can increase low back disorder risk [61].

Also, measurements of the body's spatial orientation (posture) are being used in field-based research for evaluating the occupational exposure to awkward postures of the low back and shoulder [62, 63]. Twisting is one of the main factors used for reporting the occupational low back pain [64, 65, 61]. Furthermore, when the combination of repetitive lifting and torso bending/twisting is performed, fatigue is most likely to happen [66].

Besides, physiological measures including heart rate have been used for fatigue detection [67, 68]. Heart rate may show changes besides physical strain, including the emotional stress [67]. It has been shown that as fatigue progresses, muscular tone increases, which may lead to blood occlusion [68].

Self-reported fatigue measures such as rating of perceived exertion (RPE) and subjective fatigue level have also been used as a popular measure in occupational fatigue detection [50, 69]. Generally, increased discomfort during physical activity is positively related to fatigue and reduced work capacity [70]. RPE scale is an effective method to quantify and monitor the intensity of the exercise [71]. There is also a strong relationship between RPE and physiological variables such as heart rate and lactic acid [72, 73].

1.2 Research Objectives

Two specific objectives are addressed in this dissertation:

- Examine whether the commercially wearable sensors with appropriate metrics extracted from these sensors can be used to detect the occurrence of fatigue on an individualized level for different operational tasks.
- Propose a framework consisting of four phases, including detection, identification, diagnosis, and recovery to manage fatigue in manufacturing occupations using less number of commercially wearable sensors.

Overall, the proposed approach in this dissertation provides a framework to develop predictive models by bringing together information from wearable sensors and data analytics. With the use of wearable sensor monitoring techniques, practitioners would be able to monitor the individualized body fatigue, which would consequently reduce the likelihood of work-related accidental injuries. It is hypothesized that this monitoring allows detection of at-risk individuals based on the identification of fatigue and behavior deviations to enable intervention before the injury. There exists a lack of consensus on how to best combine the data from multiple sensors and successfully evaluate risk from this data to allow multi-parametric monitoring and individualized detection of safety risk [39, 74, 75]. This kind of multi-sensor data monitoring approach is essential given the multi-factorial nature of fatigue development. Concerning effective technological approaches to fatigue measurement, it is necessary that the system be able to predict, measure, and monitor fatigue in the operational environment, and allow for intervention when deficits are identified or anticipated with appropriate interventions [35].

1.3 Dissertation Layout

The remainder of this dissertation is organized as follows: In chapter 2, a data-driven methodology is proposed to show the effectiveness of wearable sensors to detect fatigue in several manufacturing occupations. It should be noted that this research was published in Applied Ergonomics in March 2017. Chapter 3 describes a proposed framework for fatigue management in four phases: detection, identification, isolation, and intervention. Chapter 4 summarizes the contribution of these studies and pretenses a discussion and direction for future researches.

Chapter 2

A Data-Driven Approach to Modeling Physical Fatigue in the Workplace Using Wearable Sensors

2.1 Introduction

Fatigue in the workplace is a multidimensional construct that diminishes a worker's performance. It results from prolonged activity, and is associated with psychological, socioeconomic and environmental factors [76, 12]. From an occupational health and safety perspective, fatigue must be managed since it has significant short-term and long-term implications. As noted, occupational fatigue is comprised of multiple dimensions including mental and physical fatigue. Physical fatigue is characterized as a reduction in ability to perform a physical task resulting from preceding physical exertion [77, 78]. In manufacturing environments, physical fatigue may be most critical because in the short-term, physical fatigue can result in discomfort, diminished motor control, and reduced strength capacity [79, 80, 7]. These effects might lead to reduced performance, lowered productivity, deficits in work quality, and increased incidence of accidents and human errors [12, 5, 13, 81]. Physical fatigue can also result in longer term adverse health outcomes, including, e.g., *chronic fatigue syndrome* [12] and *reduced immune function* [15]. "These outcomes have been associated with future morbidity and mortality, work disability, occupational accidents, increased absenteeism, increased presenteeism, unemployment, reduced quality of life, and disruptive effects on social relationships and activities" [12].

Important parameters in the development of physical fatigue, and subsequent risk, include the length of time-on-task, work pace, and the timing of rest breaks [11]. The specific precursor(s) for physical fatigue and/or injury development often goes unidentified [82]. However, researchers have postulated that through delineation of the quantitative details of relevant variables, appropriate interventions and injury control can be developed [5]. How to best quantify workplace conditions, particularly physical exposures experienced by the worker, remains an open research question [37]. Traditional approaches to exposure assessment often rely on visual inspection performed by a trained observer. These approaches fail to capture the interactive nature of multiple risk factors as well as the variability of the work performed [22, 32, 33]. In addition, these methods do not take into account the characteristics of the individual, beyond general anthropometric and demographic attributes, such as height, age or body weight [34]. In particular, these methods fail to account for physical fatigue. Current approaches to physical fatigue monitoring and detection often rely on fitness-for-duty tests to determine whether the worker has sufficient capacity prior to starting work, diaries of sleep habits, or intrusive monitoring of brain activation (using electroencephalography (EEG)) [35] or changes in local muscle activation (using electromyography (EMG)) [36].

Accurate quantification of physical exposures is an important component of physical fatigue development. The traditional measurement approaches are able to capture *what* happened (through statistics and traditional surveillance) and a general overview of *how* it happened (video and observational data). The *why* it happened remains unclear. Improved instrumentation for data collection has been identified as a critical need for occupational safety and health [32, 33]. The increasing availability of pervasive sensing technologies, including wearable devices [37, 40, 38, 83], combined with the digitization of health information has the potential to provide the necessary *in situ* monitoring, recording, and communication of individuals' physical and environmental exposures to address the *why* [37, 31, 41]. Sensing technology can range from commercially available, wearable devices, to health monitoring devices (such as blood pressure monitors that collect and transmit data to health professionals), to dense sensor networks (including motion, video, RFID and pressure sensors), all of which provide a vast array of information regarding a person's activities. Since conventional camera-based systems require costly devices, large space in addition time-consuming adjustment experiments, utilizing the wearable sensor-based system is significantly less costly [84]. Since the late 1990s, there have been large advances in the field of wearable technology [85]. Currently, wearable sensors are lower cost, easy to use and have minimal interference with the wearer [84]. They have become the predominant devices for monitoring mobility and physical activity [86].

Evidence suggests the translation of these technologies to a work environment would be achievable and a potentially valuable addition for injury prevention [37, 41, 87, 88, 42]. While there has been an emphasis on research to develop occupational health and safety sensor systems and establish their potential in a lab environment [37, 42, 89], most current workplace applications have been limited to: a) posture analysis [42, 90, 91], b) task classification [37], c) basic physiological monitoring [88, 92], d) computerized application of traditional observational tools [93], and e) specialized (industry-specific) physical fatigue detection/management systems. The specialized physical fatigue applications are limited to the following three domains [82]:

- athletics, where the focus is primarily on monitoring athletes' performances (e.g., the *Catapult System* [94, 95], the *Viper Pod* by STATSports® [96], etc.);
- sleep-induced fatigue, in mining (CAT's *Fatigue Risk Management System* [97]); and
- driver drowsiness detection systems in transportation (see [82] for a detailed discussion).

In most physically-demanding occupations, e.g. construction, manufacturing, and agriculture, there have been minimal workplace applications that are directly related to physical fatigue detection [82]. To date, there is a lack of consensus on how to best combine the data from multiple sensors and successfully evaluate risk from this data *in situ* and *in real-time* to enable multi-parametric monitoring and individualized detection of safety risk [83]. Such a multi-sensor data monitoring approach is essential due to the multi-factorial nature of physical fatigue development. For effective technological approaches to physical fatigue measurement, it is essential that the system can predict physical fatigue (prior to a detrimental productivity/safety impact), measure and monitor physical fatigue in the operational environment, and allow for intervention when deficits are identified or anticipated with appropriate interventions [35]. Moreover, considerations of individualized baseline conditions are necessary but often ignored in a population-based approach to safety. Understanding individual variability in underlying physiological function and monitoring work-generated loading can ensure safety via early detection of risk and hazards.

This chapter sets the foundation for using *minimally-intrusive wearable sensors* for monitoring, detecting and diagnosing *whole-body fatigue* for physically demanding occupations.

The main objective of this chapter is to develop a data-driven, task-independent method that can be used to model physical fatigue through the use of inexpensive wearable sensors. To achieve this objective, We have addressed the following research questions:

(A) What are appropriate metrics for quantifying worker physical fatigue?

(B) Can commercially available *wearable sensors* be used to detect the occurrence of physical fatigue on an individualized level for different operational tasks?

(C) What information needs to be extracted from the sensors for different ergonomic targets? In this chapter, we only examine the detection of physical fatigue (i.e. has it happened or not) and the development of physical fatigue (i.e. based on Borg ratings, how physically fatigued is the worker).

The remainder of the chapter is organized as follows. Section 2.2 provides some necessary background on how *whole-body fatigue* is measured since this informs our bio-sensor selection. Then, we present our methodology for model development and evaluation in Section 2.3. WE provide our results and discuss their ergonomic/safety implications in Section 2.4. We offer our conclusions and our opinions about future research directions in Section 2.5. We present links for our de-identified data and code in the *Supplementary Materials* Section to allow researchers to replicate and build on this study.

2.2 Justification for measurement/sensor selection

There are several measures for evaluating whole-body fatigue. These measures can be classified into those that are used in a laboratory setting and those that can be used in field studies [12]. In this section, we focus on the measures that can be deployed in field studies since a primary aim of this study is to examine whether *inexpensive sensors* can be used to detect and diagnose physical fatigue in traditional occupational environments.

One logical measure for assessing physical fatigue in the workplace is to ask the worker to rate their perceived physical fatigue. Accordingly, *self-reported physical fatigue* is frequently assessed in several field-studies (see e.g., [98, 99, 100, 101, 102]). Generally speaking, increased discomfort is positively related to physical fatigue and reduced work capacity [70]. In the literature, there are several different questionnaires and rating scales that are used for measuring physical fatigue. For a detailed review, please see [12, 103]. In our analysis, we use the *Borg Rating of Perceived Exertion* [104] due to its simplicity and wide use within the *ergonomics* literature.

In addition to perceived ratings, there are several other measures that can be used to assess physical fatigue. These include *heart rate* (HR) [68, 67], *force variability* [67], *tremor* (see references within [12]), *changes in posture/gait* [68, 67], and the *multi-joint coordination between different segments* (i.e. hip, knee, ankle, pelvis, trunk) [105, 106, 107]. The sequential movement of the body segments is affected and finally controlled by muscular forces, which under physical fatigue may encounter a change; distinctive patterns of segment movements and/or muscle activation may develop [108]. Therefore, with physical fatigue, movement coordination can change so as to keep up motor performance in terms of movement accuracy [109]. One study showed that the ankle and hip joint angular displacements remained relatively unchanged between physical fatigue and non-physically fatigued conditions, but indicated that most changes in movement amplitude occurred at the knee joint level [108]. Note that, with the exception of force variability and the multi-joint coordination between different segments, these measures can be extracted from recreational activity monitors which now include a heart rate monitor in addition to accelerometers or inertial measurement units (IMUs).

Based on the above discussion and references, a series of physical fatigue indicators have been selected for this investigation. These indicators, summarized in Table 2.1, represent a range of physiological and movement parameters that can capture physical fatigue and stress from physical tasks. In Table 2.1, each physical fatigue indicator is accompanied by the measurement approach and relevant sensor that will be investigated for incorporation in this study. Deviations in these measures are commonly attributable to safety and health risk. In addition, they have been used successfully in lab and field environments for evaluating physical fatigue and risk [102]. An innovative aspect of this chapter is that we use a *data analytic* approach to generate features from these measures. We hypothesize that such a data-driven approach can be more powerful in detecting and modeling physical fatigue. More details on how we generate features from these sensors are provided in Section 2.3.

Table 2.1: Whole-body Physical Fatigue Indicators, Measures and Sensors

Physical Fatigue Indicator	**Measure**	**Sensor**
Physiological stress	Heart rate	Heart rate monitor
Change in posture/motion	Accelerations and inclination angles	IMU
Decrements in motor control and coordination	Movement variability	IMU
Physiological tremor	Movement variability	IMU
Changes in work output and task completion time	Movement durations and repetitions	IMU

2.3 Method

Our approach, depicted in Figure 2.1, consists of four phases. In Phase 1, the data is collected through two different types of sensors in 5 total locations. The second phase of data preprocessing consists of four sequential steps: a) *data cleaning*, where missing/erroneous data are detected, the data from all sensors are synchronized, and down sampling is applied to account for the variations in data collection frequency in the recorded datasets; b) *jerk calculation* from the raw acceleration data; c) application of *dimension reduction* techniques to reduce the size of the data without losing significant amount of information; and d) *feature extraction features* (i.e. potential predictors) from the multiple sensors. In Phase 3, several penalized regression models are applied to the data. *Penalized logistic regression* and *penalized regression* models were used for physical fatigue detection and development, respectively. The fourth phase involves model evaluation and testing to showcase the utility of our approach. Additional information on each of these phases is provided in the subsections below.

2.3.1 Data Collection

Participants

In this study, eight participants (3 female, 5 male; age 18 - 62 years) were recruited over a period of 3.5 months from the local community. Two of the participants were currently working in manufacturing and the remainder were students with differing levels of physical work experience. The experimental procedures were approved by the University at Buffalo Institutional Review Board and participants provided informed consent at the start of the experiment. All

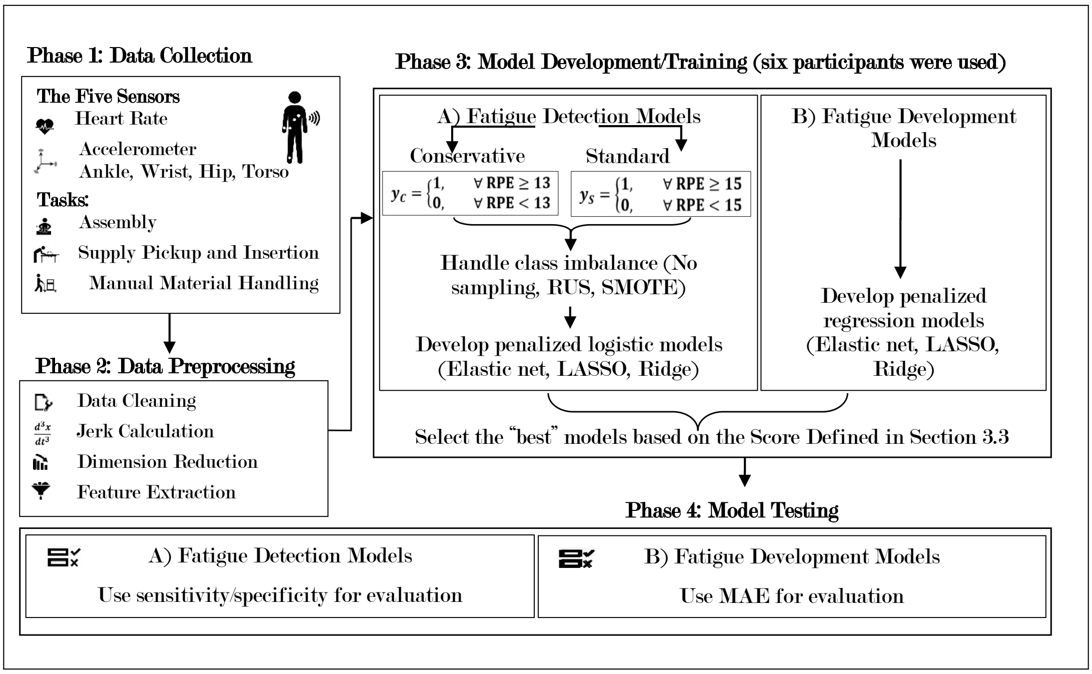


Figure 2.1: An Overview of the Proposed Method

participants were in good health. In Table 2.2, we present the demographic and relevant physical/medical characteristics of the study participants. Moreover, we highlight whether each participant was used for model development (i.e., training) or evaluation (i.e., testing). Assignment was based on the goal of developing a model that was independent of demographic information. First, participants P1, P4 and P6 (younger and older participants) were selected for the train set to cover the range of participant age. Second, in order to make the model independent from gender, participant P3 was assigned to the training set and P8 was assigned to the testing set. Third, to have consistent ages between groups, P2 and P5 were assigned to the training set and P7 was added to testing set. Due to the time commitment involved for each participant in the study, it was difficult to obtain a large sample size. Moreover, note that the small sample size for training mimics the standard deployment of new technology by industry. The proposed models in Section 2.3.3 can be applied when n (# of participants) is small.

Equipment

Each participant was instrumented with four inertial measurement units IMUs (see Figure 2.2) while performing the task. Each IMU was a Shimmer3 (Shimmer, Dublin, Ireland, www.shimmersensing.com), which is small-sized (51 mm x 34 mm x 14 mm), low-power using,

Table 2.2: Relevant Demographic, Physical and Physiological Measures for Participants. *Train/Test* describes the use of data in *model development/evaluation*, and *RHR* denotes the observed *resting heart rate* for the participant.

Participant	Train/Test	Gender	Handedness	Age	Height (m)	Weight (kg)	RHR
P1	Train	Female	Right	18	1.66	48.0	42 bpm
P2	Train	Male	Right	21	1.89	79.7	65 bpm
P3	Train	Female	Right	29	1.77	70.2	62 bpm
P4	Train	Male	Left	62	1.71	88.8	71 bpm
P5	Train	Male	Right	23	1.71	69.3	67 bpm
P6	Train	Male	Right	59	1.6	73.8	67 bpm
P7	Test	Male	Right	30	1.72	72.2	74 bpm
P8	Test	Female	Right	19	1.62	62.5	62 bpm

and equipped with wireless transmission capabilities. The sensor contains a low-noise analog accelerometer, a digital wide range accelerometer and magnetometer, and a digital gyroscope. Figure 2.2 shows a Shimmer3 device with the reference coordinate system. The 3-axial data of acceleration, angular velocity, and magnetic field, all in the sensors body frame (x y z), were recorded on an SD card at a sampling rate of 51.2 Hz. Each sensor was oriented with the internal *y*-axis directed along the segment. The sensors were attached by an elastic strap. A heart rate monitor chest strap was worn throughout the experiment (Polar CR800X, Polar). Our sensor selection was informed by Table 2.1.

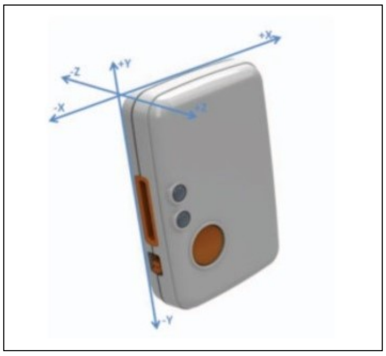


Figure 2.2: A Shimmer3 device with its reference coordinate system

Experimental Procedure

Participants completed three in-lab experimental sessions, one on each of three different days and lasting approximately four hours. Each experimental session involved the completion of one physically fatiguing task that lasted three hours. The session was divided into three one-hour periods representing a replicated task, with a one minute rest between each period to allow for subjective rating collection. At the start of the session participants completed a sleep quality questionnaire, a risk taking behavior task (Balloon Analogue Risk Task (BART)), and a psychomotor vigilance task (using PC-PVT). These measures served as a baseline of sleepiness and behavior. In addition, the subject was asked to lay in a supine position for five minutes to measure resting heart rate. After baseline measurements, the participant was provided with instructions on the relevant physically fatiguing task for the session. The physically fatiguing tasks were divided into:

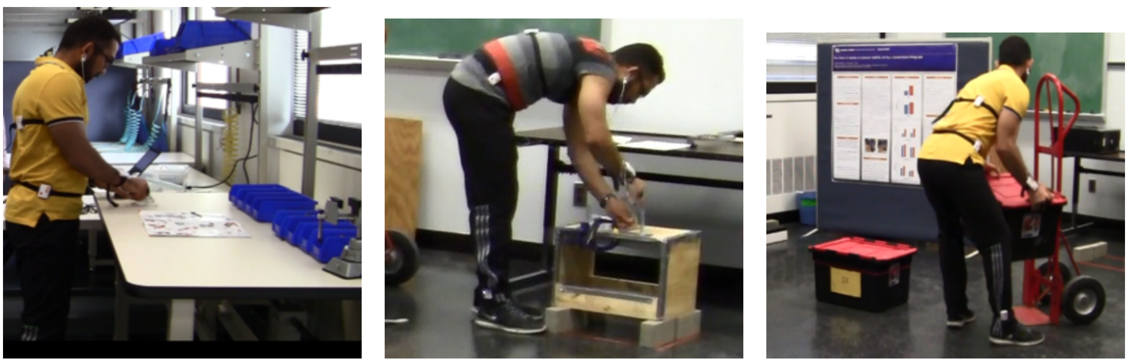
(A) Parts Assembly Task (PA-Light): In this task, part assembly operation requiring fine motor control was simulated. During the task the participants were asked to use *Erector Assembly Kits* to build (sub) assemblies based on visual work instructions. During performing the task they had stationary standing position throughout the three one-hour periods. There are several reasons behind selecting this posture. Firstly, standing in this task is a widely adopted industrial working posture [110]. Secondly, staying in this posture in daily working for the long periods can irritate physical fatigue, lower back pain and solidness in the neck/shoulders, and other health issues [111]. Thirdly, this posture decreases the blood flow to the muscles, quickens the onset of physical fatigue, and causes pain in the leg, back and neck muscles [112]. Also, usually the machine operators and assembly line workers required in such tasks characterized by extended standing reported these discomforts [113].

(B) Supply Pickup and Insertion Task (SPI-Moderate): The task included walking with supplies to a bolt box and bending forward for fastening and unscrewing the bolts. This task was selected to induce a common awkward posture held for a sustained duration used in many manufacturing processes. Deep torso flexion causes a high flexion moment on the lumbar spine [44]. Studies have shown that the risk of low back disorders significantly increases by repeated bending and lifting activities [114, 115, 116]. The tasks' cycle time was 2 minutes, which is representative of several high-volume manufacturing industries.

(C) Manual Material Handling (MMH-Most Difficult): Almost 45% of the industrial workers reported that the high levels of walking are one of the main sources of physical fatigue [14]. This task simulated warehousing operations by picking cartons (whose weight is

10kg, 18kg, or 26kg), loading them on a 2-wheeled dolly, transporting them to a destination, and then palletizing them at the destination. Participants palletized the cartons based on an order sheet provided to them. The median time across all participants used to move each carton in the cycle was 1 minutes. Participants completed two sets each of three scenarios. Each scenario involved 18 cartons for a total of 108 cartons moved during the three hours.

During performing these tasks, the aforementioned sensors one attached at each of the right ankle, right wrist, hip, and torso (see Figure 2.3). It is clear from the studies of physical fatigue detection that the position of the sensors plays a critical role in physical fatigue detection. Sensor placement depends on the task being monitored, and previous studies of activity monitoring for similar task components were used to guide placement. The MMH and SPI tasks included extended periods of walking combined with upper extremity movement. Thus, the hip, ankle, and wrist sensors were hypothesized to provide the best information [45]. In addition, determination of torso inclination during the task required the sensors on the chest [48]. These sensors were located only on one side of the body as the goal was to provide a simplified sensor approach for practical implementation in the workplace, rather than requiring a full set of sensors across the body.



(a) Part Assembly (b) Supply Pick up & Insertion (c) Manual Material Handling

Figure 2.3: Sensor Placement on a Subjects

The participants were given target performance levels for each task. For MMH, participants were asked to palletize 16 cartons in three different orders, and each order was repeated twice. In the SPI task, a cycle time of 2 minutes was given and for PA task there was a cycle time of 15 minutes for each subassembly. During each task participants were given instructions on how to perform the task.

Participants provided their subjective exertion using the Borg 6-20 RPE (Ratings of Perceived Exertions) scale [104] every 10 minutes. This was used to validate physical fatigue development.

2.3.2 Data Preprocessing

Data Cleaning

First, we used exploratory data analysis methods to check for erroneous data. Possible examples of erroneous data include: faulty sensor values (too high or too low), noisy data, and participants deviating from the experimental protocol. For the SPI task session for P1 and P5, the participants had around 10 min rest during the session and therefore returned to their baseline state, interrupting physical fatigue development. These two of the subject-task data sets were eliminated out of the 24 data sets in this phase; therefore, we ended up with 22 subject-task data sets. Then, we synchronized the sensor data and removed all observations that were captured prior to the beginning of the experimental procedure or post-procedure. We also removed the data corresponding to the first ten minutes of the session since it reflected: (a) the lack of familiarity of the participant with the task and (b) the pre-steady state for the participant's heart rate. Down sampling was applied in order to make the data consistent, with a frequency of 25 Hz. This procedure was performed on all sensors, tasks and participants' data.

Jerk Calculation

For each of the four IMUs, we calculated *jerk* from the acceleration data. Jerk is the rate of change of acceleration, and thus can be calculated by obtaining the derivative of acceleration with respect to time. We used a numerical approach to obtain the derivative. The inclusion of jerk in our methodology is motivated by its successful use in detecting physical fatigue in athletics [95].

Dimension Reduction

For each of the IMUs and the HR monitor, a massive amount of data is collected over time. *We hypothesize that the size of the data can be significantly reduced without losing much information related to physical fatigue detection and development.* Specifically, we assume that real-time in manufacturing, construction and similar occupations can be defined to be within a 10-minute time window. Thus, it is sufficient to collapse the data from each sensor into appropriate statistics (hereafter, features) that capture the variability within that time window. In addition, for the *acceleration* and *jerk* data, we hypothesize that the *magnitude* of the vector is sufficient for the purposes of physical fatigue detection and modeling. The *heart rate* data was normalized to the *resting heart rate* (RHR) and *age-predicted maximum heart rate* ($HR_{max} = 220 - age$) for the *percent heart rate reserve* (HRR; calculated as $(HR_{avg} - RHR)/(HR_{max} - RHR)$).

Feature Extraction/Generation

We propose six sets of features that may be predictive of physical fatigue occurrence and/or level. *Set 1* contains descriptive statistics for each sensor (summarized in Table 2.3), computed at 10-minute intervals. *Set 2* offers the percent change when compared to the baseline for each of the features of *Set 1*. The percent change for any feature, x, is calculated as: $(x_{current} - x_{baseline})/x_{baseline}$. We define the baseline for all features to correspond to the window where physical fatigue is likely to be minimum, i.e., the time window spanning minutes 11-20. *Set 3* contains the Cumulative Sum (CUSUM) for *Set 1* features. The CUSUM statistic for any feature x is defined as: $CUSUM(x)_i = (x_i - x_{baseline}) + CUSUM(x)_{i-1}$, where $CUSUM[x]_0 = 0$, i represents the current time window, and $CUSUM(x)_0 = 0$. The CUSUM statistic is widely used in the *change-point detection* literature (see [117] for a detailed introduction). Based on *Sets 1, 2 and 3*, we have 21 (7 statistics * 3 feature sets) HRR features, 84 (21*4 locations) *acceleration-related* (ACC) features and 84 *jerk-related* features. *Set 4* includes the elements of the correlation matrix between the four IMUs for each 10 minute window (as observed from the acceleration data), see Table 2.3. *Set 5* includes the percentage change in the defined correlation features when compared to the baseline (similar to *Set 2*). Based on *Sets 4, 5*, we have 12 total features (6 for each of *Sets 4* and *5*). *Set 6* includes statistics computed from the joint histogram of acceleration in the *current time window* and *one of two baselines*. For this feature set, the mean and standard deviation of the overlapping area in the joint histogram were calculated. There are two baselines in this case; the first as mentioned earlier, and the second is the previous 10-minute window in the experiment. This feature set is computed for acceleration only. Table 2.3 shows the elements for the joint histogram set of features. Based on *Set 6*, we can compute 16 features (i.e., 2 baselines * 2 features * 4 locations). Thus, in total, we have generated 217 features.

Table 2.3: Generated Feature Sets

Set	Features	Explanation
1	Descriptive Statistics	10th percentile, 25th percentile, 50th percentile, 75th percentile, 90th percentile, Trimmed mean, Std
2	Percent change in descriptive statistics	$(x_{current} - x_{baseline})/x_{baseline}$
3	CUSUM of descriptive statistics	$CUSUM(x)_i = (x_i - x_{baseline}) + CUSUM(x)_{i-1}$
4	Correlation between the different accelerometers	e.g. Correlation between wrist Acceleration and hip Acceleration
5	Percent change in correlation between the accelerometers	$(\rho_{current} - \rho_{baseline})/\rho_{baseline}$
6	Joint histogram	Mean and Std Dev of overlapping area considering the first 10 minutes and previous time window as baseline

2.3.3 Model Development

A primary objective of this chapter is to determine how to detect physical fatigue and understand its level based on information extracted from the wearable sensors. We distinguish between physical fatigue occurrence (a binary outcome reflecting whether physical fatigue has occurred or not) and development through our analysis of the RPE. Specifically, we define two binary decision rules for physical fatigue occurrence (y):

$$\mathrm{y}_C = \begin{cases} 1, & \forall\, \mathrm{RPE} \geq 13 \\ 0, & \forall\, \mathrm{RPE} < 13 \end{cases}, \tag{2.1}$$

and

$$\mathrm{y}_S = \begin{cases} 1, & \forall\, \mathrm{RPE} \geq 15 \\ 0, & \forall\, \mathrm{RPE} < 15 \end{cases}. \tag{2.2}$$

The reader should note that the first decision rule is more conservative (C), while the second can be considered as standard (S) based on the *Borg scale*.

The summary statistics for participants RPEs by task is shown in Table 2.4. It shows that the RPE for these recorded self-reported exertion ratings are consistent with the results from studies related to similar manufacturing tasks (see [118, 119]).

Table 2.4: Summary statistics for participants RPEs

Task	Average	Std Dev	Min	Max	Percentage of RPEs $>=$ 13	Percentage of RPEs $>=$ 15
PA	10.34	2.88	6	17	21 %	8 %
SPI	11.99	3.08	6	19	47 %	21 %
MMH	11.18	2.72	6	16	35 %	15 %

The goal of modeling physical fatigue occurrence is to determine: $y_a = f(X)$, where X is a vector containing the features x, and $a = C, S$. The actual value of the *Borg Scale* is used in modeling physical fatigue development, i.e. $y_{Borg} = f(X)$. If these models are predictive, one can use the appropriate vector of features X to determine whether a worker is physically fatigued or not, and to what extent that worker is physically fatigued. Thus, practitioners can replace the RPE with information extracted from the wearable sensors.

An intuitive approach to model physical fatigue development is to use regression methods to fit the function f in the above paragraph. Based on the experimental procedure, standard (i.e. ordinary least square) fitting of the regression model cannot be used since:

(A) We expect potentially significant correlation between the features generated from each IMU at each time window. From an ergonomics perspective, we expect these IMUs to offer some *overlapping* information.

(B) We are proposing a methodology that can allow researchers and practitioners to include more features (e.g., due to adding additional sensor types). Thus, the number of features (p) might be close to the number of observations (n).

For these scenarios, the use of *penalized regression* models is more appropriate (see [120, 121] for detailed explanations). Below, we provide an overview of how these models are incorporated for physical fatigue detection (i.e. penalized logistic regression) and physical fatigue development (i.e. penalized regression).

Penalized Logistic Regression for Physical Fatigue Detection

In this subsection, we will drop the subscript for y. The reader should note that we perform this analysis on both the conservative and standard binary outcomes defined in Equations (1-2). The first stage in penalized logistic regression is to standardize the entire dataset so that each feature has a mean of zero and a unit standard deviation. The general function to find the coefficients in penalized logistic regression models is given by:

$$\max \quad l(\beta) - \lambda[\alpha \sum_{j=1}^{m} \mid \beta_j \mid + \frac{1}{2}(1-\alpha) \sum_{j=1}^{m} \beta_j^2], \tag{2.3}$$

β_j represents the regression coefficients, and m is the total number of features (m is equal to 217). The $\lambda \geq 0$ is a tuning parameter that controls the strength of the penalty. More specifically, λ shrinks each β_j toward the origin and enforces sparse solutions. The value of α represents different popular parameterizations of penalized regression. In our analysis, we consider the three parameterizations: *LASSO* [120], *ridge regression* [122], *elastic-net* [123] for the two different binary outcomes. For a general discussion on these approaches, the reader is referred to Hesterberg et al. (2008) [121].

Penalized Regression for Understanding Physical Fatigue Development

In this subsection, we utilize penalized regression for an assumed continuous *Borg rating*, i.e. y_{Borg}. Consider a standard linear regression between the predicted *Borg rating* ($\widehat{y}_{Borg}$ or alternatively $\widehat{RPE}$) and the m features:

$$\widehat{y}_{Borg} = \widehat{RPE} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_m x_m + \varepsilon, \tag{2.4}$$

where x_s, β_s, and ε represent the features, coefficients, and residuals for the model, respectively. It is assumed that $\varepsilon \sim N(0, \sigma^2)$. Now, let us consider the penalized regression model. The general function to find the coefficients for the different parameterizations of the model is:

$$\min \quad (RPE - \widehat{RPE})^T (RPE - \widehat{RPE}) + \lambda[\alpha \sum_{j=1}^{m} \mid \beta_j \mid + \frac{1}{2}(1-\alpha) \sum_{j=1}^{m} \beta_j^2], \tag{2.5}$$

where RPE is the actual *Borg rating*, $\widehat{RPE}$ is the prediction, and $\lambda \geq 0$ is the tuning parameter. The value of α represents different popular parameterizations of penalized regression as explained earlier. In Section 2.4, we examine which method is more suitable for modeling physical fatigue development based on our data.

2.3.4 Model Evaluation (Testing)

The purpose of this phase is to first compare the performance of *ridge regression*, *LASSO*, and *elastic-net* for our two *penalized logistic regression* models and our *penalized regression model* based on the six participants identified as "train" in Table 2.2. There are several important methodological "issues" that need to be addressed in order to identify the "best" model from the *training phase*.

First, the number of models examined in any *penalized (logistic) regression* model can be narrowed down according to the amount of variation explained by the model (hereafter *fraction deviance*). In this chapter, we limited our choice of models to ones whose *fraction deviance* is $\sim$70%. This value was selected since, in our experience, higher values would result in *over-fitting*. In addition, the selection of a narrow range for the values (i.e $\sim$70%) reduces the number of models to be evaluated.

The second "issue" relates to how to measure the performance of these models. For the two binary models, we use *sensitivity* and *specificity* to compute the effectiveness of the models in predicting the *physically fatigued* and *non-physically fatigued* states, respectively. Note that *accuracy* is not appropriate since the data is imbalanced (i.e. $n_{fatigued} << n_{non-fatigued}$). For the continuous model, we compute the *mean absolute error (MAE)*. In our estimation, MAE is the most appropriate performance measure since its interpretation is straightforward. For example, $MAE = 2$ means that the prediction is, on average, 2 units on the *Borg Scale* from the participants' reported exertion values.

Third, our modeling approach should result into a practical and applicable model. An important aspect to, therefore, consider is the portion of possible features used in the model. For example, a model using all 217 features is not likely to be of practical use. Therefore, we have avoided on defining those models with the highest performance measure(s) to be "best".

Instead, we would like to "reward" models with high performance measure(s) and small *proportion of features* selected since their interpretation is much easier. Accordingly, we proposed a scoring system that combines *fraction deviance*, *traditional performance measure (TPM)*, and *proportion of the factors* selected in a single measure. The score (S) for a given model, k, is defined as:

$$S_k = FractionDeviance * TPM * \frac{m}{r_k}, \tag{2.6}$$

where $TPM = Sensitivity * Specificity$ for the *physical fatigue detection* case, $TPM = MAE$ for the *physical fatigue development* case, m is the # of possible features, and r_k is the # of features selected by model k. Based on Eq. 2.6, a higher value of S would mean that the model is likely to be *good* (from *fraction deviance* and *TPM* perspectives), and practical (smaller values for r are encouraged).

One of the potential pitfalls of modeling physical fatigue in well-designed jobs in industry is that the $n_{fatigued} << n_{non-fatigued}$. This is especially significant with limited amounts of training data. In such cases, *random sampling* for the conservative and standard logistic regression models would not be appropriate [124, 125]. Thus, the fourth question is to examine how to handle this problem. Recent approaches in the *data analytics* community address this issue through re-sampling [124].Re-sampling is typically applied by either over-sampling the minority class and/or under-sampling the majority class [125]. In this chapter, we utilize *Random Under Sampling (RUS)* and *Synthetic Minority Over-sampling Technique* (SMOTE). RUS is a systematic process where some of the cases from the majority class are randomly removed from the training dataset until the remaining number of cases in the two classification categories becomes approximately equal. In SMOTE, the minority class is over sampled using synthetic examples (see [124, 125] for details).

Once the "best" models are selected from the *training* stage (based on maximizing S_k), it is important to evaluate how these models perform on *subjects* that were not included in the *model development* step. We *test* the *best* models on two participants whose data was not included in the *training* step. This step with our dataset since it ensures that the *hidden* effects of the physiological, demographic and other individualized characteristics on physical fatigue occurrence/development are not considered. For the *physical fatigue detection* models, we use *sensitivity* and *specificity* to compute the effectiveness of the models in predicting the *physically fatigued* and *non-physically fatigued* states in the two participant identified as "test" in Table 2.2. The *MAE* is used to evaluate the *testing* performance for the physical fatigue development models. The results for the *training* and *testing* phases of our models are presented in Section 2.4.

2.4 Results

In this section, we present experimental results for the selected *physical fatigue detection* and *development* models. The results correspond to the different *penalized logistic regression* and *penalized regression* models (with SMOTE, RUS, and *random sampling*). We divide this section into three main subsections: (1) *training* results of the *physical fatigue detection models*, (2) *training* results of the *physical fatigue development models*, and (3) results of the evaluation/testing for both *physical fatigue detection* and *physical fatigue development* models.

2.4.1 Selected Models for Physical Fatigue Detection

As discussed earlier in Section 2.3.3, three different penalized logistic regression models were recommended for physical fatigue detection. We applied these three models to data from six participants and the results showed that the LASSO model performed better than the *ridge regression* and *elastic-net* models. The LASSO model exceeded the others in two important areas: (1) the model included fewer features and (2) explained a larger portion of the variation. Therefore, we decided to use only the LASSO model going forward.

The results for the training step of *physical fatigue detection* models are shown in Table 2.5, which is separated into two groups of models according to the conservative vs. standard approach. For each group, we label the highest score in bold. In groups with multiple similar high scores, we use the top 2 models. Practitioners can pick the model with the higher sensitivity, higher specificity, or the smallest number of selected features. Since our *scoring system* is a heuristic, there is no guarantee that favoring one metric would lead to "better" results in the deployment phase. Two important results can be observed from Table 2.5:

(A) Performance using standard/conservative scenarios: It is obvious from Table 2.5 that scores for models when standard scenario was used exceeded those when the conservative scenario was used. This may suggest that the conservative models are over-fit, i.e. when these models are used on the testing subjects, their performance will be much worse. Next to, for each of the *conservative* and *standard* scenarios, a comparison of the models reveal that the number of features selected (out of 217) is consistently larger in conservative scenarios.

(B) Performance using LASSO modeling with the RUS sampling technique: The models in bold show that the LASSO model with RUS sampling technique is the only model that performs well in the two groups. Therefore, it is reasonable to select the *LASSO model with RUS sampling* as the best option for *modeling physical fatigue detection*.

Table 2.5: Training Performance of the Different LASSO Penalized Logistic Regression Models

Conservative/Standard	Sampling Tech.	Sensitivity	Specificity	# Features	Score
Standard	No sampling	0.80	0.99	21	5.73
Standard	**RUS**	**0.95**	**0.89**	**16**	**8.03**
Standard	**SMOTE**	**0.93**	**0.89**	**15**	**8.38**
Conservative	**No sampling**	**0.96**	**0.88**	**22**	**5.83**
Conservative	**RUS**	**0.96**	**0.88**	**22**	**5.83**
Conservative	SMOTE	0.95	0.92	23	5.77

Predictive Features for the Standard Physical Fatigue Detection Scenario

Table 2.6 shows the 16 selected features and their corresponding coefficients for the LASSO model using RUS sampling for standard physical fatigue detection. The resulting coefficients are sorted from largest to smallest. The absolute value of those coefficients show the relative contribution of each feature since the entire dataset was standardized in the first stage of penalized modeling. The larger the absolute value of the coefficient, the greater the influence of the feature on differentiating between a *physically fatigued* and a *non-physically fatigued* state. The three most important features all correspond to the wrist movement. These were followed by features that relate to torso movement. Changes in movement patterns in *wrist* and *torso* are important in detecting physical fatigue. For example, the *standard deviation of* *joint histogram*$_1$ (i.e., index 1 refers to the joint histogram using the previous time window as a baseline) in wrist acceleration (*Wrist ACC: standard deviation of joint histogram1*) shows that when the acceleration in 2 consecutive time windows are distributed differently from each other, the body did not follow the same movement as in the previous time window. Therefore a large standard deviation here means that the participant is feeling more physically fatigued. A visual depiction of the location of the features of Table 2.6 is provided in Figure 2.4. The figure highlights how much the wrist is involved in standard physical fatigue detection. The wrist and torso account for almost 62.5% of the selected features in the model. Interestingly, none of the heart rate features were selected in the *standard physical fatigue detection model*.

Definition of the Selected Features	Coefficient
Wrist ACC: standard deviation of joint histogram1	0.88
Torso Jerk: CUSUM of standard deviation	0.78
Hip ACC: median	0.39
Wrist Jerk: standard deviation	0.37
Hip Jerk: percentage change of 75th percentile	0.05
Wrist ACC: percentage change of 25th percentile	0.03
Wrist Jerk: 90th percentile	0.001
Ankle Jerk: CUSUM of 10th percentile	-0.03
Wrist Jerk: CUSUM of standard deviation	-0.05
Hip & Ankle ACC: percentage change of correlation	-0.09
Wrist Jerk: percentage change of 10th percentile	-0.19
Ankle ACC: percentage change of median	-0.27
Ankle ACC: median	-0.31
Torso ACC: CUSUM of median	-0.78
Wrist Jerk: CUSUM of 75th percentile	-1.10
Wrist ACC: CUSUM of trimmed mean	-1.20
(Intercept)	-1.28

Table 2.6: Selected Features for the Logistic LASSO Model with RUS Sampling (Standard Scenario)

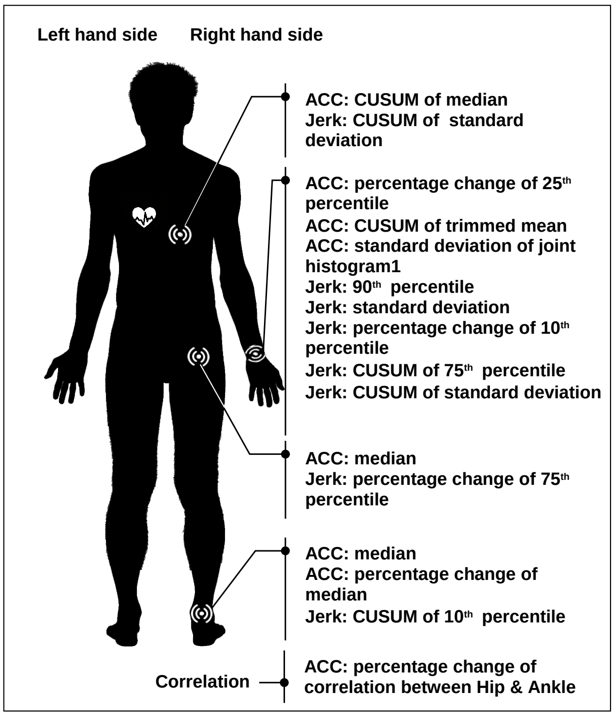


Figure 2.4: The Location of Selected Features on Body for the Standard Fatigue Detection Model

Predictive Features for the Conservative Physical Fatigue Detection Scenario

Similar to the standard approach, the 22 selected features and their corresponding coefficients are shown in Table 2.7. Figure 2.5 likewise shows the placement of sensors and features on the body. The five most important features were related to hip, wrist, and ankle movements. These important features show that left tail of the hip movement distribution (*Hip ACC: 25th* *percentile*) and right tail of the wrist movement distribution (*Wrist ACC: 90th percentile*) are associated with physical fatigue in participants. As participants feel or identify physical fatigue, they change their hip movement. On the other hand, greater variability in ankle movement (*Wrist Jerk: CUSUM of 75th percentile*) and the wrist (*CUSUM of 75th percentile*) during the experiment was associated with a lack of physical fatigue in the body. In contrast with the standard scenario, two features that relate to the heart rate have been selected by the model. In this case, *heart rate* was involved in detecting physical fatigue, but its importance was not as significant as the mentioned sensors.

Definition of the Selected Features	Coefficient
Hip ACC: 25th percentile	0.88
Wrist ACC: 90th percentile	0.87
Hip Jerk: percentage change of median	0.55
Hip ACC: mean of joint histogram1	0.49
Ankle ACC: mean of joint histogram1	0.43
Ankle Jerk: percentage change of 75th percentile	0.39
Ankle ACC: CUSUM of 25th percentile	0.27
Torso ACC: mean of joint histogram1	0.18
Wrist & Torso ACC: correlation	0.13
Wrist ACC: percentage change of median	0.10
Torso Jerk: CUSUM of standard deviation	0.10
Wrist Jerk: 90th percentile	0.08
Hip & Ankle ACC: percentage change of correlation	-0.05
Torso & Ankle ACC: correlation	-0.20
HRR: percentage change of standard deviation	-0.23
Hip ACC: CUSUM of median	-0.24
(Intercept)	-0.49
HRR: percentage change of median	-0.55
Ankle ACC: trimmed mean	-0.57
Torso ACC: percentage change of trimmed mean	-0.64
Hip ACC: percentage change of trimmed mean	-0.71
Wrist Jerk: CUSUM of 75th percentile	-0.76
Ankle ACC: CUSUM of 10th percentile	-0.82

Table 2.7: Selected Features for the Logistic LASSO Model with RUS Sampling (Conservative Scenario)

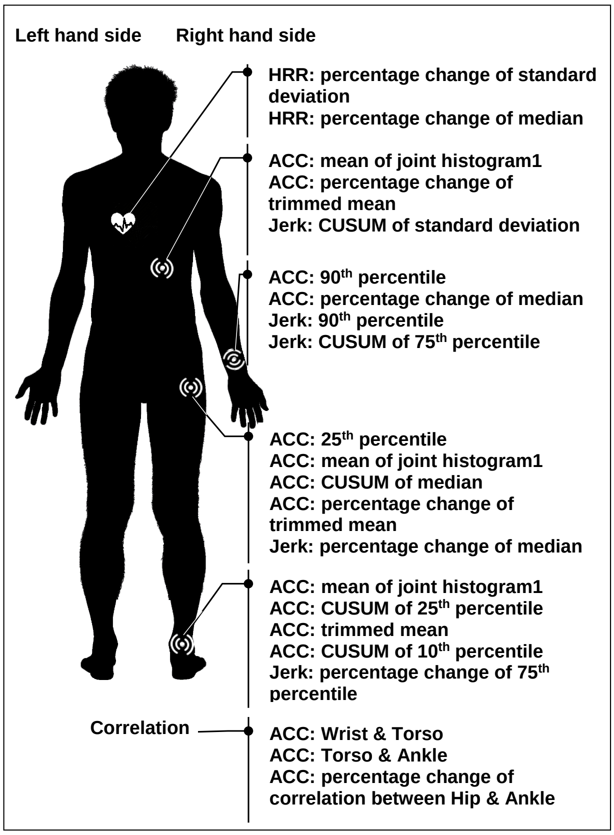


Figure 2.5: The Location of Selected Features on Body for the Conservative Physical Fatigue Detection Model

2.4.2 Selected Model for Physical Fatigue Development

The procedure for selecting the best model for physical fatigue development was similar to the approach for physical fatigue detection. The LASSO model performed better than the other penalized regression methods. The performance of using this model had an MAE = 1.34, r = 21, and score = 5.40 for the training data. The 21 selected features and their corresponding coefficients are presented in Table 2.8 and visualized in Figure 2.6.

Definition of the Selected Features	**Coefficient**
(Intercept)	11.33
Ankle Jerk: percentage change of 75th percentile	0.76
Hip ACC: median	0.61
Hip ACC: mean of joint histogram1	0.49
Wrist ACC: standard deviation of joint histogram1	0.29
Torso ACC: mean of joint histogram1.	0.26
Ankle ACC: CUSUM of standard deviation	0.18
Wrist & Torso ACC: correlation	0.17
Torso Jerk: CUSUM of standard deviation	0.13
Wrist ACC: 90th percentile	0.01
Torso & Ankle ACC: correlation	-0.02
Hip ACC: CUSUM of median	-0.02
HRR: percentage change of 10th percentile	-0.03
Torso ACC: percentage change of trimmed mean	-0.09
Ankle ACC: percentage change of 25th percentile	-0.10
Torso ACC: CUSUM of median	-0.10
Ankle ACC: median	-0.13
Torso ACC: percentage change of median	-0.14
Ankle ACC: percentage change of median	-0.18
Ankle ACC: CUSUM of 10th percentile	-0.33
Hip ACC: percentage change of trimmed mean	-0.77
Wrist Jerk: CUSUM of 75th percentile	-0.99

Table 2.8: Selected Features for the LASSO Model with RUS Sampling

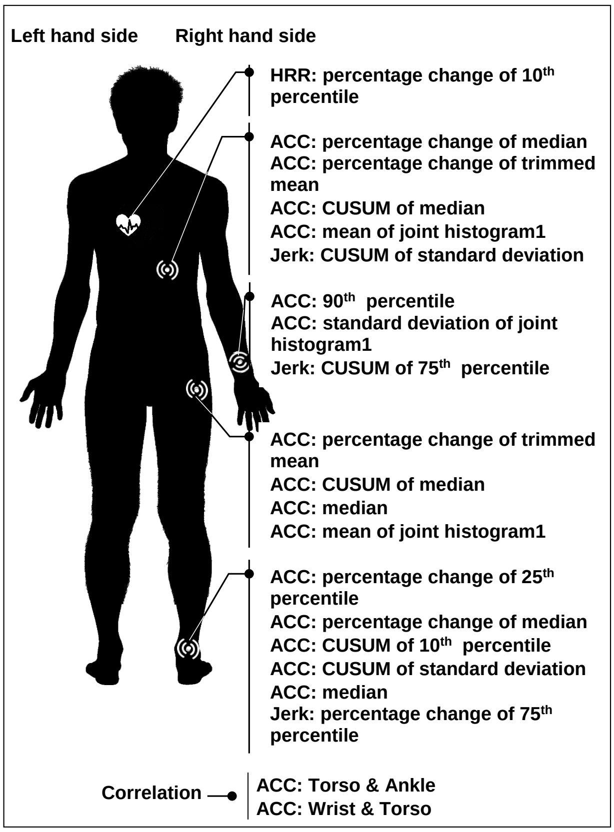


Figure 2.6: The Location of Selected Features on Body for the Physical Fatigue Development Model

Table 2.8 shows that the first five important features correspond to the wrist, hip and ankle movement. The wrist was the body part common to both physical fatigue detection and development models. Table 2.8 highlights that the percent change of third quartile in ankle movements (*Ankle Jerk: percentage change of 75th percentile*) in each time window were correlated with increasing physical fatigue. Additionally, large hip movements (*Hip ACC: median*) are strongly indicative of a participants having physical fatigue. In contrast, greater variability in the movement of the wrist (*Wrist Jerk: CUSUM of 75th percentile*) indicates that the participant is not physically fatigued. The selected variables yielded by these models provide strong evidence that the wrist status has a profound effect on physical fatigue detection and development, followed by the hip, ankle, and torso which also play an important role in physical fatigue modeling, while heart rate is less of an indicator.

2.4.3 Testing of the Models

Implementation of Selected Physical Fatigue Detection Model on Two Test Participants

The selected models from section 2.4.1 were used to test their implementation for physical fatigue detection on the two participants (*P7* and *P8*) who were not used during the training stage.Table 2.9 shows the performance of the *standard physical fatigue detection model* corresponded similarly to its training performance. The performance of the *conservative physical fatigue detection model* was slightly worse than its training performance. Thus, our *scoring heuristic* provides some insight into selecting "suitable" models from the training step. An additional interesting observation from Table 2.9 is that the *standard LASSO model* detected all *physically fatigued* states for the two test participants since its $sensitivity = 1$.

Table 2.9: Performance of Selected Physical Fatigue Detection Models on "Test" Participants

Conservative/Standard	Sampling technique	Penalized logistic model	Sensitivity	Specificity
Conservative	RUS	LASSO	0.65	0.70
Standard	**RUS**	**LASSO**	**1.00**	**0.79**

Implementation of Selected Physical Fatigue Development Model on Two Test Participants

Similar to the testing of the physical fatigue detection models, the physical fatigue development model from Section 2.4.2 was tested on the two participants, resulted in an MAE = 2.16. An MAE of 2.16 indicates that the RPE prediction was on average 2 units off from the recorded RPE for test participants. This result is particularly good since RPEs are perceived measures of physical fatigue (i.e., highly variable and not very accurate).

Figure 2.7 shows the time series plots for the recorded (black) and predicted (gray) RPE for the 2 test participants. These plots provide insights for how each model performs for a given task and a given "test" subject. Both models perform "better" for the *Supply Pick up and Insertion* and *Manual Material Handling* tasks when compared to the less physically fatiguing *Parts Assembly* task. This observation is based on the similarity of trends between the predicted and recorded values of the RPE.

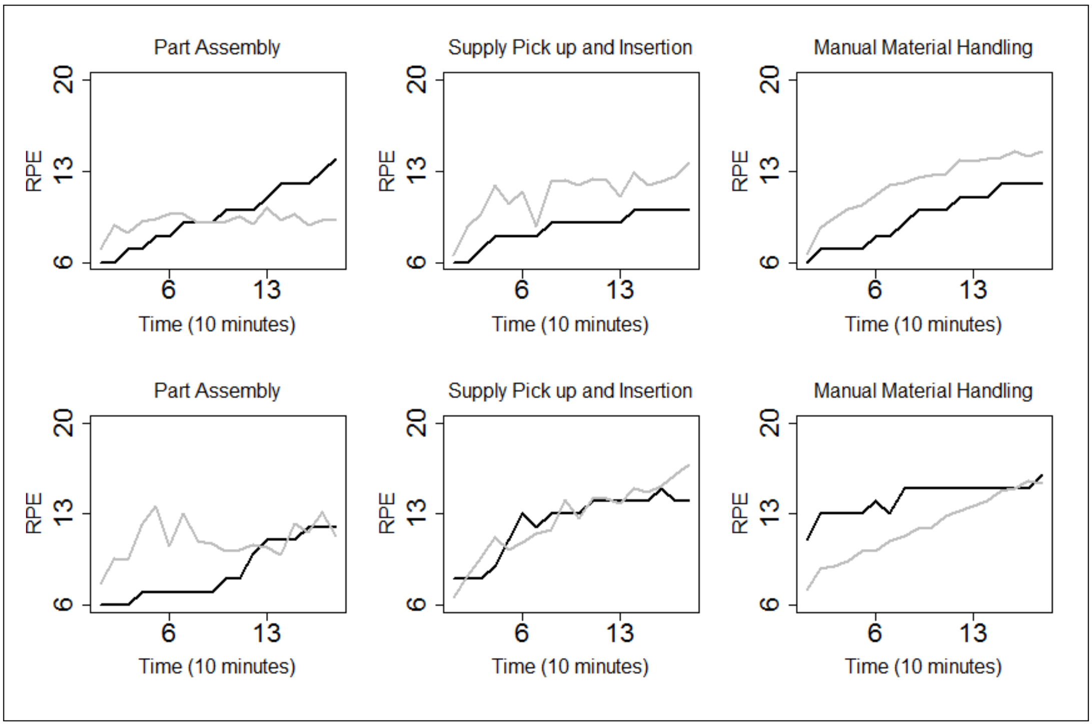


Figure 2.7: Predicted RPE (Gray) vs. Actual (Black) - *P7* (top) and *P8* (bottom)

Overall, the attached sensor to the wrist played an important role in detecting the occurrence of physical fatigue and estimating its level. The "best" model for physical fatigue detection was the *standard LASSO model* with RUS sampling. The "best" model for physical fatigue development was the LASSO. In addition, the *physical fatigue detection model* performed better than the *physical fatigue development model* in the testing step. Therefore, if we were to recommend a single model of those developed in this study, we would recommend using the standard approach for physical fatigue detection that corresponds to the RUS-based Standard Logistic Regression LASSO model.

2.5 Discussion and conclusion

2.5.1 Summary of the resulting models

Worker physical fatigue is an important safety concern in manufacturing environments and monitoring physical fatigue is essential to prevent accident and injury occurrences. From a hypothetical point of view, the utilization of predictive models for physical fatigue modeling can provide a chance to better incorporate the understanding of the physiology and psychology of fatigue. Model predictions can be tested and the results can be utilized to refine the model and the understanding of the basic phenomena [126]. In the majority of physical fatigue models the attention is on the impact of circadian rhythmic, sleep loss, and the resultant sleepiness on becoming physically fatigued [11, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135]. These models are typically confined to workrest and/or sleep-wake information as the inputs rather than the nature of the manufacturing work. For industrial workers, the nature of the work and the work setting can impact the utilization of non-work periods in forming sleep-wake behavior [126]. Only a few of the available physical fatigue models join the work processes into the assessment of the physical fatigue-related risk connected with a work schedule [11]. In this study we attempted to model physical fatigue by considering responses to the nature of the work. Therefore, we simulated three basic manufacturing tasks (MMH, SPI, PA) so as to induce physical fatigue. Then we developed a data-driven approach to dealing with the occurrence of physical fatigue and estimating its level. While the models were built from data on three tasks performed for four hours each, the outputs provided are independent of task, an important consideration for practical implementation. To our knowledge, no other studies have examined physical fatigue modeling for three common, but disparate, manufacturing tasks performed for an extended duration.

The selected features and their coefficients for the best model for physical fatigue detection were shown in Table 2.6. In previous studies that have used IMUs for monitoring physical activity and work tasks, although not necessarily physical fatigue, popular features computed from the acceleration signal are the mean [48, 136, 137, 138, 139, 140, 141, 142], variance or standard deviation [137, 138, 140, 142, 143], and the correlation between acceleration axes

[48, 142]. In this study, six feature sets were developed to capture the physical fatigue-related information while reducing the size of the data and accounting for the time dependency. The features which corresponded to the wrist, torso and hip sensors had the main contribution to physical fatigue detection. These features are shown in Table 2.10. For the three simulated manufacturing tasks, upper extremity movement was a main component of the task. As seen in Table 11, the largest absolute coefficients were associated with the wrist sensor as well as the highest number of features. This is consistent with previous studies that showed the wrist as a significant location for sensor placement [48, 136, 137].

Table 2.10: Dominant features for the standard physical fatigue detection model

Definition of the Selected Features	**Coefficient**
Wrist Acceleration: standard deviation of joint histogram1	0.88
Torso Jerk: CUSUM of standard deviation	0.78
Hip Acceleration: median	0.39
Wrist Jerk: standard deviation	0.37
Torso Acceleration: CUSUM of median	-0.78
Wrist Jerk: CUSUM of 75th percentile	-1.10
Wrist Acceleration: CUSUM of trimmed mean	-1.20

In understanding the resulting model, it is important to understand the specific features that were selected in the penalized regression process. Those features with positive coefficients are contributing to the determination of physical fatigue, while those with negative coefficients are mitigating factors. Significant features contributing to the determination of physical fatigue in this model include:

- Wrist Acceleration Standard Deviation of the Joint Histogram 1: This represents the variation in the overlapped distribution area between two consecutive time windows. Therefore, a higher similarity in the wrist acceleration between two consecutive periods is an indication the participant is not physically fatigued. Whereas, high variability from one time period to the next would indicate physical fatigue.

- Torso Jerk CUSUM of Standard Deviation: A measurement of the deviation in torso jerk or smoothness compared to the baseline state. A larger variability from baseline suggests a change in the smoothness of the body movements, contributing to physical fatigue.

- Hip Acceleration Median: A measure of the central tendency of the hip acceleration distribution, if the participant maintains a high level of hip acceleration, then they are more likely to report feeling physically fatigued.

- Wrist Jerk Standard Deviation: This feature measures the variation in the wrist jerk, or variation in the change in acceleration. Higher variability in the smoothness of the movement led to detection of physical fatigue.

On the other hand, the following features had a negative relationship with physical fatigue detection:

- Wrist Acceleration CUSUM of Trimmed Mean: The trimmed mean is the mean after removing 5% from the beginning and end of the time window. If the participant maintained a higher wrist acceleration compared to the first time window, then they were less likely to report physical fatigue.

- Wrist Jerk CUSUM of 75th Percentile: This feature shows the sum of the deviation in locations of the right tail of the wrist acceleration distribution over the time windows compared to the first time window. If the distribution of the wrist jerk skewed to the right, indicating a larger value of high jerk, then physical fatigue would not be present. As with the previous feature, if the participant maintains a high wrist acceleration, not slowing down, then they are not physically fatigued.

- Torso Acceleration CUSUM of Median: Representing the sum of the variation in the median of the torso acceleration. Slowing down, a negative deviation in torso acceleration, from the beginning of the experiment would indicate the participant was physically fatigued.

As these results demonstrate, the features with negative coefficients were related to the CUSUM comparing the overall level and variation of movement from the start of the task. Whereas those with positive coefficients were more often related to movement variability. Previous work has shown that wrist acceleration, which is used to characterize the hand movement pattern, is useful for estimating biomechanical demands and physical fatigue for exercise tasks such as swimming [144, 145]. Industrial workers who need to move their hands and wrist repeatedly and/or powerfully are at greater risk for cumulative trauma disorders (CTD) [136]. The results of the current models are consistent with these findings and suggest that monitoring wrist movements is important. In current bio-mathematical models of physical fatigue prediction, population variability is not accounted for in determining the final physical fatigue score, thus 50% of individuals will have a predicted physical fatigue score greater than the average value predicted by developed models and half will be underneath that [126]. Our study gives solid confirmation that in the lab-based experiment the penalized logistic LASSO model with RUS sampling in standard approach performs better than the current generation of bio-mathematical model. Its estimation for physical fatigue state (0 or 1) was 100%.

In the best model for physical fatigue level prediction the features which have larger coefficients corresponded to the wrist, hip and ankle sensors. These features are shown in Table 2.11.

Table 2.11: Dominant features for the physical fatigue level prediction

Definition of the Selected Features	**Coefficient**
Ankle Jerk: percentage change of 75th percentile	0.76
Hip Acceleration: median	0.61
Hip Acceleration: mean of joint histogram 1	0.49
Ankle Acceleration CUSUM of 10th percentile	-0.33
Hip Acceleration: percentage change of trimmed mean	-0.77
Wrist Jerk: CUSUM of 75th percentile	-0.99

- Ankle Jerk Percentage Change of 75th Percentile: This feature shows the percentage change in right tail of the ankle jerk distribution over the time window compared to the first time window. If the participant maintains a high level of ankle jerk in right tail of its distribution, then physical fatigue is developing.

- Hip Acceleration Median: As described above.

- Hip Acceleration Mean of Joint Histogram 1: The mean in the overlapped distribution area between two consecutive time windows. If the participant maintains a high level of hip acceleration then the participant would indicate physical fatigue.

The features with a negative relationship with predicting the physical fatigue level included:

- Wrist Jerk CUSUM of 75th Percentile: As described above.
- Hip Acceleration Percentage Change of Trimmed Mean: This represents the percentage change of trimmed mean of hip acceleration over the time window compared to the first time window. This feature shows that if participants kept consistent hip movement over all time windows, then it likely corresponds to their walking behavior and less likely to reported physical fatigue.
- Ankle Acceleration CUSUM of 10th percentile: This feature shows the sum of the deviation in location of the left tail of the ankle acceleration distribution over the time window. Slowing down, a negative deviation in ankle acceleration, from the beginning of the experiment would indicate the participant was physically fatigued.

Similar to the physical fatigue detection model, these results show the features with negative coefficients were related to the CUSUM features, whereas those with positive coefficients were more related to the distribution of the movement. The results of this model confirmed the previous results that monitoring wrist movement is imperative.

Considering both of the best models for detecting physical fatigue and predicting its level, the primary sensors in modeling physical fatigue were located at the wrist and hip. The selected features in these models showed that monitoring the distributions of the movements (Set 1), percent change of movements (Set 2), movement variability (Set 3), and similitude between current movements and the first time window (Set 6) were generally powerful. In this study the heart rate features were not as critical as the movement features. While heart rate has traditionally served as an indicator of physical fatigue, the changes in movement were stronger predictors. This may have resulted from the specific manufacturing tasks simulated and the development of a single model that covers the different tasks. In both the SPI and PA tasks, the task did not impose a high cardiovascular load and the heart rate remained relatively consistent throughout the three-hour period. More variability was observed in the movement. This is consistent with a previous study on monitoring activity in activities of daily living, where heart rate was not a main element of the resulting model, while IMUs on the wrist and thigh were included [138].

2.5.2 Implementing physical fatigue modeling in the workplace

It is critical for the reader to note that there is currently very little published data on how models are being utilized as a part of work environment settings (see [129, 130, 131, 133, 134, 135, 146]). Typical implementations take the sleep-wake history to develop work schedules and shift work rotations [126]. However, these models do not coordinate the nature of the work being attempted and its potential impact on physical fatigue and safety. For instance, in aviation it is well established that the level of workload and exposure to risk is not steady over a flight, but this is not considered by models [147]. While the models developed in this study may be applicable to a number of tasks, the more important outcome of this study is the description of a modeling approach that could be applied. To implement the proposed modeling approach from this study in a manufacturing workplace, the following aspects must be considered.

First, the safety critical workers in workplaces, who are inclined to end up physically fatigued while doing their jobs should be distinguished [147]. These workers will be a suitable sample to develop the predictive physical fatigue model since the difference between their physically fatigued and not physically fatigued states during their jobs will be distinguishable when modeling. Therefore, in order to identify these critical workers, meeting with the workers in workplace is viewed as essential to the process [147]. Second, in developing a valid physical fatigue model the data need to cover a variety of conditions and should be recorded over an extended period for different workers. It should be noted that more data in this stage will help to develop a strong predictive physical fatigue model. Third, the assigned tasks for those critical workers who will be monitored with the same model should be standardized. The result of the predictive physical fatigue model is highly dependent on the manufacturing task condition (i.e. amount of workload, work condition, works' speed, trained workers). Consistent task circumstances for the workers with the same tasks would help to avoid variation in the predictive physical fatigue model. The purpose of physical fatigue modeling to detect the occurrence of physical fatigue or its level should be defined, since each of them need differing sensors and predictive model. Based on the results in this chapter, the wrist sensor was the fundamental sensor which should be utilized. Therefore, for physical fatigue detection the sensors on wrist, torso and hip are required. For physical fatigue level prediction, the wrist, hip and ankle sensors are required. Next, the working hours of the critical workers needed to be divided to short intervals (i.e. in this study we selected 10 minutes). Then, after recording the data, the model can be developed by defining the statistically important features (Sets 1, 2, 4, 6) proposed in this study. The R programming language (https://www.r-project.org/) and MATLAB which were used to generate the results in this study can be accessed through the following Github Repository: https://github.com/zahrame/Fatigue-modeling.git in order to implement predictive physical fatigue models in the workplace. After developing the model, the predictive model can be used to predict the occurrence or level of physical fatigue in the other workers in the workplace.

2.5.3 Study limitations and future research

There are a few main limitations that must be acknowledged for this study. First, the sample size is small as a result of the long time commitment required for each participant. However, the sample size is consistent with other studies that have focused on lab-based modeling of physical fatigue [148]. In addition, since the study was completed as a within-subjects design some of the variability in responses across tasks was minimized. Each participants 180 minutes of data was also divided into 18, 10-minute segments, allowing the models to be built from a larger number of data points. Future studies are needed to investigate whether the current models are valid when applied to a larger sample. Second, all of the participants were physically healthy and most were young adults. Some of the physical and health characteristics may be different from a standard industrial population. There is limited evidence on the role of demographic and other individual differences in the development, recognition, tolerance, and accommodation of physical fatigue [149]. The effect of different demographic variables (i.e. age, sex, economic status, race, and marital status, personality traits, and circadian rhythms) needs to be explored in future models of physical fatigue. Third, the participants had limited training time (10 minutes) to become familiar with the task. Therefore prior experience and physical fitness may have affected the initial results recorded at the start of the experiment. Fourth, the modeling in this study was assessed only based on the worker's self-reported Borg rate level, which may be biased based on other factors outside the physical fatigue level, including motivation and discomfort. This perception would affect the results of physical fatigue modeling, therefore, for the future research this issue should be considered. This study focused on the work process to model physical fatigue, so in order to enrich the model from a human work performance perspective, further analyses concerned with quantitative performance measures (e.g., number of defects in a time window and average task completion time over a time window) should also be examined. Moreover to evaluate the effect of job conditions on physical fatigue, measures including shift work, job/task rotation, pace constraints and repetitiveness for tasks need to be investigated. The results in this study showed that participants' physical fatigue level were close to those of the predicted values in MMH and SPI task (Figure 2.7), but it performed poorly with the PA task. Therefore, it is necessary to identify what type of tasks are most amenable to physical fatigue modeling in order to implement preventive actions for exposed workers. We recommend that future ergonomic studies consider different sensor combinations to improve modeling physical fatigue for tasks concentrated primarily on the upper extremity. This would allow for a more precise analysis of the relationships between this task and related factors.

Acknowledgments

This research was partially supported by the American Society for Safety Engineers (ASSE) Foundation grant titled ASSIST: Advancing Safety Surveillance using Individualized Sensor Technology. The authors would like to acknowledge two anonymous reviewers, whose feedback, comments and suggestions have greatly improved this chapter. The authors would like to acknowledge Amir Baghdadi, a Mechanical Engineering Ph.D. Student in the University at Buffalo, who collected the data for this research. We are also grateful for Gabriel Gazetta de Araujo, Julia Bitencourt and Tamires Dias, undergraduate researchers at Auburn University, whose insights on wearable sensors contributed to this work.

Supplementary materials

In this study, the R programming language (https://www.rproject.org/) and MATLAB were used to generate the results. Our raw data and code can be accessed through the following Github Repository: https://github.com/zahrame/Fatigue-modeling.git.

Chapter 3

A data analytic framework for physical fatigue management using wearable sensors

3.1 Introduction

The advancements in automation, computation, information, sensing and software systems led to a new paradigm of *advanced manufacturing* systems [150]. An important managerial goal in *advanced manufacturing* is to investigate how workers can be better integrated in the evolving cyber-physical infrastructure so that the impact of their skills can be maximized [151]. Thus, the role of labor within the manufacturing firm has transformed, where the following changes have been observed: (a) a reduction in mundane tasks [152], (b) an increased dependency on highly-trained workers [153], (c) an increase in worker's autonomy and responsibility [154], and (d) the introduction of new job duties [154]. From a worker's perspective, this expanded role has significant financial rewards [155]. However, there is a growing body of evidence, which suggests that the associated workloads result in high levels of fatigue and other negative health outcomes [156, 157, 153, 158, 14].

As discussed in chapter 2, an important first step in managing fatigue is the rapid and accurate detection of its occurrence. *Fatigue detection* techniques can be divided into two categories: qualitative and quantitative. Qualitative methods are centered around the use of fatigue surveys [14]. From a practical perspective, the utility of such methods is limited to investigations aiming to assess workloads and/or redesign jobs. However, they are not suitable for real-time, shop-floor-wide *fatigue detection*, since they are not scalable and are potentially disruptive. For example, consider a situation where there are 70 workers on the shop-floor and their fatigue ratings are measured every 5 minutes. The administration of surveys in this situation would require a large number of surveyors, and would disrupt production (reducing the productivity of workers [159]). The quantitative approaches, of the second category, rely on using one or more sensor technologies to model changes in human performance, that we showed in chapter 2.

This chapter proposes a framework for using *wearable sensors* to manage worker fatigue in manufacturing environments. A framework is proposed instead of a model to allow for the detection/diagnosis of multiple fatigue modes. The main premise is that *advanced manufacturing* firms require specialized labor [153, 14]. Thus, the jobs can then be grouped by the type of activities. This is reasonable since the main tasks performed by a CNC, computer numerical control, machinist are different from those done by a welder. The proposed framework is made of four phases: (a) *detection*, where the goal is to detect if/when a worker has become fatigued, (b) *identification*, where the most important variables for diagnosing fatigue are identified, (c) *diagnosis*, where the information captured from phases (a-b) is used to pinpoint the fatigue mode, and (d) *recovery*, where a suitable intervention is applied to return to a non-fatigued state. The phases are adapted from the structured methodology used by quality engineers for fault detection and diagnosis [160]. Note that none of the existing quantitative approaches for fatigue modeling present information on the *identification, diagnosis and recovery* stages needed for managing fatigue.

The proposed framework utilizes *wearable sensors* for three main reasons. First, based on a survey of U.S. manufacturing safety professionals, 54.1% of the respondents were "in favor of using wearable technologies at work to track [occupational safety and health] risk factors" [161]. From the responses, Schall et al. (2018) [161] estimated that U.S. manufacturing firms would spend, on average, an estimated $68.67 per worker for a wearable device. Second, the use of *wearables* presents a unified benchmark of performance that does not depend on the cycle time of the process. The third, and perhaps the most important reason, *wearables* present an individualized view of the performance of the worker. Unlike other outcomes, e.g., work quality which may be affected by upstream performances.

There are several differences regarding this chapter when compared to chapter 2. First, we focus on the whole fatigue process from fatigue detection to fatigue diagnosis, however, in chapter 2, we only focused on fatigue detection. In chapter 2 we did not consider the modeling to a specific manufacturing task since there was no focus on diagnosis. Second, in this chapter, we would show that fatigue can be predicted when the number of the sensors is limited. Third, isolation and diagnosis phases are proposed in this chapter, where they have not been discussed in chapter 2.

The remainder of the chapter is organized as follows. In Section 3.2, an overview of the relevant literature on fatigue management in manufacturing environments is presented. Our proposed framework for detecting, identifying and diagnosing fatigue root-causes is discussed in Section 3.3. In Section 3.4 two case studies are investigated to evaluate the utility of the framework in managing fatigue during two manufacturing tasks. Our concluding remarks and future research suggestions are presented in Section 3.5. We offer our code and data as *supplementary materials* to encourage adoption in practice and further investigations by researchers.

3.2 Literature review

The literature on physical fatigue detection in manufacturing environments can be classified into: (a) *exhaustion detection*, and (b) *occupational fatigue detection*. In the first group, studies attempt to identify *extreme fatigue*, i.e. exhaustion, which results in an inability to generate muscle forces and consequently, a worker's inability to perform the job [162]. Since exhaustion in the manufacturing workplace is often on the muscle level (*localized fatigue*), the associated literatures [163, 44, 164, 165, 166, 167, 168, 169, 170, 50] is characterized by: (i) primarily utilizing invasive EMG and EEG sensors, (ii) focusing on one task element only (e.g., lifting or walking), and (iii) no attempt to generalize the developed models to focus on a more complex task. In the second group, the studies focused on detecting *occupational fatigue*, which is less extreme than exhaustion, where the workers are still able to perform their job at a diminished level. Those studies, e.g. [166, 171, 167, 50], have often utilized pervasive sensors including IMUs and heart rate monitors . In addition, in chapter 2 we have developed a generalized model for detecting fatigue across multiple manufacturing tasks. However, the model involved over 20 predictors and lacked the interpretability that makes it effective for the consequent phases of fatigue identification, diagnosis and recovery. Table 3.1 summarizes the literature in the two groups. In this chapter, we focus on *occupational fatigue* since it is: (i) a precedent to exhaustion, and (ii) more aligned to the working environment in advanced manufacturing environments. Moreover, our proposed framework is evaluated using multiple complex manufacturing tasks in an attempt to showcase its potential generalizability.

Table 3.1: A summary of the two major research streams of fatigue modeling

Category	**Paper**	**Tasks**	**Sensors**	**Method**
Exhaustion	[172]	Walking	EMG	Statistical test
	[173]	Walking	3D optical tracking	LDA, SVM, KNN, NB
	[170]	Walking	IMUs	SVM
	[163]	Lifting	EMG	Time frequency analysis
	[44]	Lifting	EMG	Statistical test
	[164]	Lifting	EMG	Statistical test
	[169]	Squat	Infrared cameras	Linear regression, HMM
Occupational fatigue	[166]	Walking	EMG	Linear regression
	[171]	Walking	Accelerometer	Statistical test
	[167]	Walking	Reflective markers	Statistical test, LDA
	[50]	Material handling	IMUs	SVM
	[174]	Material handling, supply insertion & part assembly	IMUs, HR	Penalized logistic regression

* where LDA=linear discriminant analysis, SVM=support vector machines, KNN= k-nearest neighbors, and HMM=hidden Markov models.

From a detailed review of the literature, we could not identify any papers that discuss the identification and diagnosis of fatigue. This may be attributed to the implicit assumption in the literature that management or the individual worker can handle those stages once fatigue has been identified. However, as indicated in Levenson (2017) [175], “workplace fatigue is a systems problem”, and there needs to be a systematic approach to identify its root-causes. This is a critical gap since the end goal is intervening to prevent the unwanted negative consequences on the worker and the production process.

3.3 Methodology

Figure 3.1 presents an overview of the four phases of the proposed framework for managing physical fatigue. The first phase is comprised of five main steps: (a) *sensor selection*, where practitioners should identify appropriate sensors for fatigue detection; (b) *data preprocessing and feature generation*, where the sensors' data are prepared for analysis; (c) *model construction and validation*, where statistical and data analytic models are trained for distinguishing between fatigued and non-fatigued states; (d) *measuring usefulness*, where models are evaluated based on accuracy, sensitivity, specificity, etc.; and (e) *ease of use analysis*, where the best model in step (d) is evaluated by constraining the number of sensors used. Note that steps (d) and (e) are based on the *Technology Acceptance Model (TAM)* [176]. The outcome from

Phase 1 is the selection of an appropriate model for prospective analysis. In Phase 2, the subset of features/predictors that are most frequently used in predicting the fatigue state is identified. This subset presents insights into what features are most predictive, which is an important input to the following phase. Phase 3 utilizes visual analytic methods (specifically an interactive parallel coordinates plot) to help management understand how the variation in the values of the predictors impact the fatigue state (i.e. from 0 to 1). Based on the insights gained from the *fault diagnosis* phase, a suitable evidence-based intervention can be selected in Phase 4.

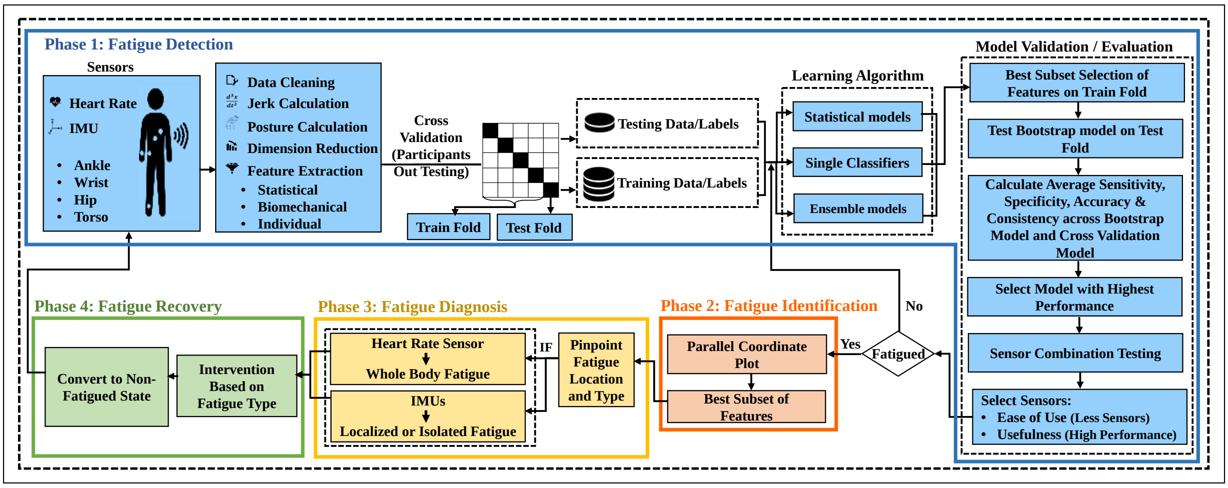


Figure 3.1: An overview of proposed method

3.3.1 Phase 1: Fatigue Detection

Sensor Selection

Cavuoto & Megahed (2017) [177] discussed several fatigue indicators, which included heart rate, heart rate variability, tremor and performance. They suggested that these indicators can be monitored using pervasive wearable sensors. In a follow-up work, in chapter 2 we showed that four IMU sensors (located at the ankle, hip, torso and wrist) coupled with a heart rate sensor can be used to detect fatigue in different manufacturing tasks. Similar to chapter 2, we suggest using these wearable sensors for fatigue detection in this chapter. More importantly, our framework presents a systematic approach to answer the question: "what are the gains associated with wearing an extra sensor?" In essence, this question attempts to quantify whether the hassle and cost associated with wearing an extra sensor can be justified with a significant/practical improvement in fatigue detection. This question, which has not been addressed in the literature, is tackled in the *usability analysis* in Phase 1.

Data Preprocessing

Cleaning

The first step in analyzing data is to ensure that the data is correct and cleaned. For *wearable sensors* data, four main cleaning steps are proposed. First, a low-pass filter should be applied on the acceleration data for noise removal. Second, collected data should be visualized to check for any additional erroneous data, i.e. data that were not corrected through the automated filtering in step 1. Possible examples of erroneous data include faulty sensor values (too high and/or too low), and participants who had not experienced fatigue based on their subjective fatigue ratings. Third, the data from the different sensors should be synchronized and any observations that were captured outside of the experimental window should be eliminated. The fourth step involves the normalization of the heart rate data through the computation of: *percent heart rate reserve (%HRR)*. Note that %HRR accounts for both an individual's *resting heart rate (RHR)* and his/her *age-predicted maximum heart rate* $HR_{max} = 220 - \text{age}$. The %HRR can be computed as:

$$\%HRR = \frac{\text{Heart Rate} - RHR}{HR_{max} - RHR} \times 100. \tag{3.1}$$

The interpretation of the % HRR is a percentage of an individual's heart rate capacity being used. Since it accounts for both their resting and maximum heart rates it allows for standardizing the heart rate data. For example, if the %HRR =50, this means that the person is using 50% of their heart rate capacity, i.e. is half way between his/her resting and maximum heart rates.

Jerk and Posture Calculation

The four IMUs (attached at the ankle, wrist, hip and torso) measure the acceleration associated with a person's dynamic motion. From the acceleration profile, other components of motion can be computed. Jerk, which is the derivative of acceleration with respect to time, should be computed since it has been shown to be effective in detecting fatigue in several occupational settings (see e.g., Catapult Sports (2018) [178] for several applications in professional sports). In addition, changes in work posture are also indicative of fatigue [177]. In this chapter, the approach of Baghdadi et al. (2018) [50] is used for posture calculation, where: (a) a Kalman filter is first used to calculate position in the three (*xyz*) directions, and then (b) posture is estimated from the positional data. The reader is referred to Baghdadi et al. (2018) [50] for more details on posture calculation.

Dimension Reduction and Feature Extraction

Based on the aforementioned data preprocessing steps, one would have 12 acceleration profiles (4 IMUs $\times$ 3 directions [x y z]) and 4 jerk profiles (rate of change of the magnitude of the acceleration profile for each IMU) each sampled at 25 Hz. In addition, there is a %HRR profile sampled at 1000 Hz. These profiles cannot be directly used in predictive models and thus, features summarizing these profiles need to be generated. In this article, we propose utilizing features that would summarize the profiles based on a non-overlapping time window of the 17 profiles. The selection of the length of the time window should depend on: (a) length of the cycle for task, (b) consequences of fatigue on the worker and production, and (c) managing the trade-off between false alarms and early detection.

To capture the changes within the profile and provide insights to the later isolation and diagnosis phases, three sets of features are generated from the 17 profiles. The first set corresponds to statistical features from the acceleration, jerk, posture and %HRR. For each of these profiles, the mean and coefficient of variation (CV) are computed for each time-window to capture the intensity and variation changes. Features capturing the intensity and spread are commonly used in the fatigue detection literature (see e.g., [48, 138, 179]). The second set corresponds to biomechanical features, which allow for identifying and diagnosing the type of fatigue. This set includes features such as: *number of steps in the time interval*, *mean step time and length*, and *mean foot/hip oscillations*. The biomechanical features used in our framework are depicted in Figure 3.2. Note that these features are calculated for each time window. Those features are computed based on the code provided by Baghdadi et al. (2017) [50]. The third, and last feature set contains both *age* and *gender*, which may be used to explain performance differences across different individuals (see [180, 181] for more details). A description of the proposed features for each of the three sets is provided in Table 3.2.

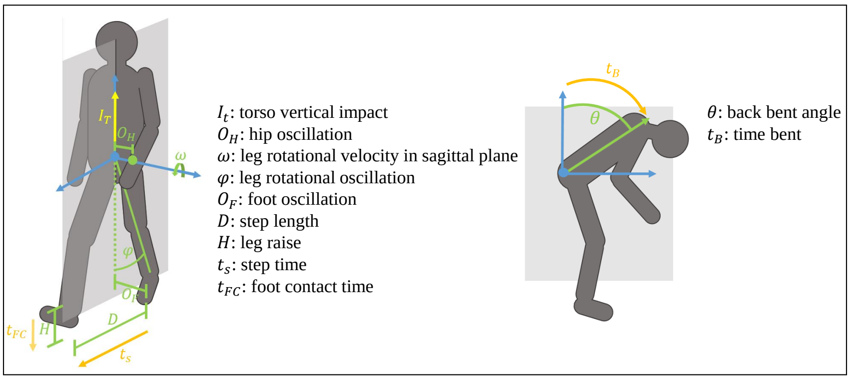


Figure 3.2: Biomechanical features illustration

Model Construction and Validation

Cross validation

A leave p-participants out cross validation approach can be used to split the preprocessed dataset into training and testing sets. Cross validation is commonly used to avoid overfitting [192]. A typical approach to cross validation is dividing the dataset into 10 folds, where the models are selected based on the average/median prediction performance across 10 non-overlapping test datasets. The literature suggests that 10-fold cross validation may reduce the variation between the train and test performance [124]. Note that in *fatigue detection* studies such as ours, each participant's data maybe autocorrelated. Thus, the plain k-fold cross validation approach is not suitable since the train and test datasets are not independent. To alleviate this problem, we recommend leaving p participants out for the cross validation, where the value of p corresponds to approximately 10% of the participants in the data analytic study.

Feature Selection and Dimension Reduction

When the number of potential features/predictors is large, the computational complexity for training a classification algorithm increases. Feature reduction is typically applied to reduce the computational burden. More importantly, it leads to: (a) an improved prediction performance, and (b) an increased generalization capability. Algorithms for feature selection/reduction can

Table 3.2: Generated feature sets

Category	#	Feature	Definition	Justification
Statistical	1	%HRR.Mean	Average percent of heart rate reserve	[182] [183] [48] [44] [47] [184] [185] [186] [170]
	2	Wrist.jerk.Mean	Average wrist jerk or smoothness magnitude	
	3	Wrist.ACC.Mean	Average wrist acceleration magnitude	
	4	Wrist.xposture.Mean	Average wrist angular position in sagittal plane	
	5	Wrist.yposture.Mean	Average wrist angular position in transverse plane	
	6	Wrist.zposture.Mean	Average wrist angular position in coronal plane	
	7	Hip.jerk.Mean	Average hip jerk magnitude	
	8	Hip.ACC.Mean	Average hip acceleration magnitude	
	9	Hip.xposture.Mean	Average hip angular position in coronal plane	
	10	Hip.yposture.Mean	Average hip angular position in transverse plane	
	11	Hip.zposture.Mean	Average hip angular position in sagittal plane	
	12	Torso.jerk.Mean	Average torso jerk magnitude	
	13	Torso.ACC.Mean	Average torso acceleration magnitude	
	14	Torso.xposture.Mean	Average torso angular position in sagittal plane (bending)	
	15	Torso.yposture.Mean	Average torso angular position in transverse plane	
	16	Torso.zposture.Mean	Average torso angular position in coronal plane	
	17	Ankle.jerk.Mean	Average ankle jerk magnitude	
	18	Ankle.ACC.Mean	Average ankle acceleration magnitude	
	19	Ankle.xposture.Mean	Average ankle angular position in coronal plane	
	20	Ankle.yposture.Mean	Average ankle angular position in transverse plane	
	21	Ankle.zposture.Mean	Average ankle angular position in sagittal plane	
	22	%HRR.CV	Coefficient of variation in %HRR	
	23	Wrist.jerk.CV	Coefficient of variation in the wrist jerk	
	24	Wrist.ACC.CV	Coefficient of variation in the wrist acceleration magnitude	
	25	Wrist.xposture.CV	Coefficient of variation in the wrist angular position in sagittal plane	
	26	Wrist.yposture.CV	Coefficient of variation in the wrist angular position in transverse plane	
	27	Wrist.zposture.CV	Coefficient of variation in the wrist angular position in coronal plane	
	28	Hip.jerk.CV	Coefficient of variation in the hip jerk magnitude	
	29	Hip.ACC.CV	Coefficient of variation in the hip acceleration magnitude	
	30	Hip.xposture.CV	Coefficient of variation in the hip angular position in coronal plane	
	31	Hip.yposture.CV	Coefficient of variation in the hip angular position in transverse plane	
	32	Hip.zposture.CV	Coefficient of variation in the hip angular position in sagittal plane	
	33	Torso.jerk.CV	Coefficient of variation in the torso jerk magnitude	
	34	Torso.ACC.CV	Coefficient of variation in the torso acceleration magnitude	
	35	Torso.xposture.CV	Coefficient of variation in the torso angular position in sagittal plane	
	36	Torso.yposture.CV	Coefficient of variation in the torso angular position in transverse plane	
	37	Torso.zposture.CV	Coefficient of variation in the torso angular position in coronal plane	
	38	Ankle.jerk.CV	Coefficient of variation in the ankle jerk magnitude	
	39	Ankle.ACC.CV	Coefficient of variation in the ankle acceleration magnitude	
	40	Ankle.xposture.CV	Coefficient of variation in the ankle angular position in coronal plane	
	41	Ankle.yposture.CV	Coefficient of variation in the ankle angular position in transverse plane	
	42	Ankle.zposture.CV	Coefficient of variation in the ankle angular position in sagittal plane	
Biomechnical	43	Number of steps	Number of gait cycles during the fixed time interval	[187] [50] [26] [188] [189] [51] [190] [191]
	44	Mean step time	Average duration of each gait cycle	
	45	Mean step length	Average length of each gait cycle	
	46	Time bent	The duration spent in bent posture	
	47	Mean back bent angle	Average angle of torso in bent posture w.r.t vertical axis	
	48	Mean hip oscillation	Average side-to-side range of motion in hip	
	49	Mean foot oscillation	Average side-to-side range of motion in foot	
	50	Mean leg rotational velocity in sagittal plane	Average angular velocity of leg in sagittal plane	
	51	Mean leg rotational oscillation in sagittal plane	Average angular range of motion for leg in sagittal plane	
	52	Mean torso vertical impact	Average value of peak vertical acceleration in torso	
	53	Mean back rotational position in sagittal plane	Average range of bending posture while doing the task	
Individual	54	Age	-	[180]
	55	Gender		[181]

be categorized into three main groups [193]: (1) *filter methods*, where univariate statistical approaches are typically used to select features based on their relationship to the response, (2) *wrapper methods*, where the important features are kept based on their prediction performance, and (3) *embedded methods*, which involve the use of methods such as LASSO for selecting the most predictive features. We refer the reader to the discussion in Feng & George Shanthikumar (2017) [194] for a detailed overview of the application of feature selection methods in production and operations management.

Since the end goal of our proposed framework is to enable the diagnosis of fatigue and the recommendation of an appropriate intervention, we recommend a two-step approach for feature selection. In the first step, simple filter approaches (e.g., information gain or correlation analysis) should be combined with visualizations (e.g., time series charts, parallel coordinates plot, and scatter diagrams). The goal of the first step is to provide practitioners with an understanding of how fatigue affects and/or is associated with changes in the potential predictors. From this step, any features that are unchanged in the fatigued and non-fatigued states should be removed. The reader should note that the insights gained from the visualization will also be utilized in diagnosing the root-causes of fatigue. In the second step, several structured wrapper and/or embedded methods (e.g., best subset selection and LASSO) should be examined. Preference should be given to techniques that result in a small number of features (i.e. more interpretable) and a relatively large prediction performance (i.e. good fatigue detection with a low false alarm rate).

Bootstrapping

To further prevent over-fitting and the bias associated with selecting a training dataset, we recommend the application of *bootstrapping* [195], which is a computational procedure that uses intensive re-sampling with replacement. An important assumption behind *bootstrapping* is that the sample distribution is a good approximation to the population's distribution. Recent studies have shown an improved performance of analytical models when bootstrapping is deployed (e.g., see [196, 197]).

Analytical modeling

The analytical classification models can be categorized into: (a) Statistical models, (b) Single classifiers, and (c) Ensemble models. The pros and cons of using these methods [198, 2] are shown in Table 3.3. Note that we do not include more advanced *deep learning models* since they often require special computing resources (i.e. graphical processing units, GPUs) and would be quite difficult to implement for a large number of workers.

Table 3.3: Comparing the three different analytical categories. Table is adapted from Wang (2016) [2]

	Statistical models	Single classifiers	Ensemble models
High accuracy in general			✓
High speed of learning against # of variables and samples	✓		
High tolerance to redundant variables		✓	✓
High tolerance to collinearity		✓	✓
High dealing with overfitting		✓	
Less complexity and easy parameter handling	✓		

Several classification methods, i.e. statistical models, single classifiers, and ensemble models, are viable candidates for utilization in fatigue prediction. From our framework's perspective, it is impossible to predetermine which methods will work best for a given application. This is due to the fact that these methods are data-driven and thus, are application-dependent. In the following paragraphs, we highlight some commonly used methods within each category.

Statistical models attempt to build a relationship between the input variables and response through the use of parametric methods. Examples include: *logistic regression* and *penalized logistic regression*. Those are classification techniques where the probability of a dichotomous outcome is a function of the predictors/features [199, 200]. A key difference between the two aforementioned approaches lies in how they handle sparse datasets. Specifically, logistic regression's performance can vary significantly with sparse data [201]. On the other hand, based on the result in chapter 2 the penalized logistic regression approach usually provides a more consistent performance.

In the single classifier category, some commonly used classifiers include: *decision trees (DT)*, *naive Bayes (NB)*, *artificial neural networks (ANN)*, *k-nearest neighbors (kNN)*, and *support vector machines (SVM)*. Those non-parametric approaches are commonly used in several production and operations management applications. The reader is referred to [200, 202, 203, 204] for examples of those applications. We recommend exploring one or more of those models for fatigue classification.

For the third category, ensemble models are comprised of several single classifiers, where the final classification of the response is based on some voting or weighting procedure [205]. The premise for these methods is that combining a large number of single classifiers allows for a more diverse representation of the data and consequently, a more accurate prediction. Commonly used ensembles include [204]: (a) *random forests (RFs)*, which are ensemble classification algorithms that utilize trees as base classifiers to generate many classifiers and aggregate their results via voting [206]; (b) *bagging* [207], where bootstrapping is used to generate a new training dataset, and combine several base learners to fit a weak learner to the data; and (c) *boosting* [208], which creates different base learners by sequentially reweighing the instances in the training set. *Boosting* gives different weights to the base learners based on their accuracy. The final model obtained by the boosting algorithm is a linear combination of several base learners weighted by their own performance. For a more detailed introduction on the aforementioned analytical models, the reader is referred to [209, 210].

Measuring Usefulness

To evaluate the performance of the analytical models, we recommend using four performance measures: (a) *accuracy*, which presents the percentage of correct classifications made by a given model, (b) *sensitivity*, which captures the ability to detect the fatigued cases, (c) *specificity*, which measures the correct classification of non-fatigued cases, and (d) a newly proposed *consistency* metric, which is a simple metric that captures the absolute difference between the metrics in (b) and (c). This metric can be used by practitioners to gage whether a model is equally capable of predicting both the fatigued and non-fatigued states. The mathematical formula below show how each of these metrics is computed first for each fold, and then averaged across all folds:

$$Accuracy_j = \frac{1}{n}\sum_{i=1}^{n}\frac{TP_{ij}+TN_{ij}}{TP_{ij}+TN_{ij}+FP_{ij}+FN_{ij}}. \tag{3.2}$$

$$Mean\ Accuracy = \frac{1}{m}\sum_{j=1}^{m} Accuracy_j. \tag{3.3}$$

$$Sensitivity_j = \frac{1}{n}\sum_{i=1}^{n}\frac{TP_{ij}}{TP_{ij}+FN_{ij}}. \tag{3.4}$$

$$Mean\ Sensitivity = \frac{1}{m}\sum_{j=1}^{m} Sensitivity_j. \tag{3.5}$$

$$Specificity_j = \frac{1}{n}\sum_{i=1}^{n}\frac{TN_{ij}}{TN_{ij}+FP_{ij}}. \tag{3.6}$$

$$Mean\ Specificity = \frac{1}{m}\sum_{j=1}^{m} Specificity_j. \tag{3.7}$$

$$Consistency_j = |Sensitivity_j - Specificity_j|. \tag{3.8}$$

$$Mean\ Consistency = \frac{1}{m}\sum_{j=1}^{m} Consistency_j. \tag{3.9}$$

where TP, TN, FP, FN denote the number of true positives, true negatives, false positives, and false negatives, respectively. i denotes the number of the bootstrapping samples, j is the number of the training or testing data sets, n is the number of bootstrapped samples, and m is the number of folds in the leave p-participants-out cross validation.

Ease of Use Analysis

In addition to evaluating its usefulness, an important aspect for technology adoption is usability. In the context of our framework, usability can be measured using two metrics: (a) *total number of features selected*, and (b) *total number of sensors needed to generate these features*. In general, models are more interpretable if the number of features are smaller (assuming no significant differences in prediction capabilities). Workers and more practitioners will also be more inclined to adopt the framework if it requires less sensors since it will: (i) be much cheaper; for example, requiring one IMU instead of four, would reduce the cost by a factor of four; (ii) make the process less invasive to the worker; and (iii) reduce the time needed for the worker to wear and strap all the sensors. Therefore, our framework will not only consider prediction performance, but it will also evaluate how the prediction performance varies while restricting the number of sensors that can be used. At this stage, one would have a model that can accurately predict the fatigue state (based on the leave p-participants out cross validation approach), while having a relatively small number of features. This model can now be deployed for near real-time prediction.

3.3.2 Fatigue Identification

Once the model is deployed and fatigue is identified, it is important to understand how the predictors' change when an individual becomes fatigued. Typically, machine learning models are thought of as "black boxes", where it is difficult to understand how the predictors affect the response. However, an important aspect of recovering from fatigue is being able to diagnose its root-causes. Since we favor having a lower number of features in our model selection (see Section 3.3.1), we hypothesize that the chosen prediction model will have a relatively low number of features. Thus, one can use a *parallel coordinates plot* to depict how the chosen features vary with the dichotomous response. The use of such a plot will enhance the interpretation of the model and assist practitioners in diagnosing the type of fatigue in the next phase.

3.3.3 Fatigue Diagnosis

In this phase, one would determine which type of fatigue occurred. Since this framework focuses only on physical fatigue, there are two main types of fatigue that are possible [177]: (a) whole body fatigue, and (b) localized muscle fatigue. Based on the parallel coordinates plot from the previous phase, one would identify the important features for prediction. If the features are derived from only one IMU (as in our first case in Section 3.4.1), one would conclude that the worker is experiencing localized muscle fatigue, near that IMU's location. Alternatively, if the features are derived only from the heart rate sensor (see Section 3.4.2), this implies that the worker is experiencing *whole body fatigue*. The last possibility would include features selected from one or more IMU and the heart rate sensor. In this case, the individual is experiencing a combination of whole-body fatigue (i.e. respiratory related) and localized fatigue. Based on the diagnosis, one can assign appropriate interventions in the next stage.

3.3.4 Fatigue Recovery

From a management perspective, it is important to prescribe interventions that eliminate/reduce the safety hazards. In essence, "safety does not happen by accident" [211] and thus, it is important to intervene to eliminate/mitigate the sources of fatigue. We recommend utilizing the *safety design hierarchy* [212] from safety engineering. This hierarchy presents a structured approach for interventions, where practitioners should consider six actions in order of effectiveness. Since this is a well-known concept to safety professionals, we do not detail this further.

In our estimation, the fatigue diagnosis stage allows practitioners to directly pinpoint the hazard (i.e. type of fatigue). Practitioners can then prescribe interventions from a large number of options, including: (a) redesigning the task (which can eliminate the development of fatigue), (b) assigning rest breaks (which can reduce the level of fatigue before it reaches potentially dangerous levels), and (c) job rotation (where workers would essentially cycle between harder and easier jobs). The type of intervention assigned will depend on the resources available to safety practitioners and the constrains of their production operations. For this reason, we only recommend the adoption of the *safety design hierarchy* without providing a recommendation for the type of interventions to be assigned. The reader is referred to the survey of [14] for a discussion of the type of interventions used by advanced manufacturing workers and safety professionals in combating physical fatigue at the workplace.

3.4 Case Studies

To evaluate the performance of the proposed framework, we examine two case studies. The first case study involves a simulated manual material handling (MMH) task, and the second is a supply pick-up and insertion (SI) task. Both case studies replicate typical fatiguing manufacturing tasks (see the survey in [14] for details) in a controlled lab environment in order to facilitate the data collection process. Since the data collection, data preprocessing and model construction steps are the same for the two tasks, we only explain them in detail in Sections 3.4.1 and 3.4.1.

3.4.1 Case Study 1: Manual Material Handling

Data Collection, Preprocessing and Feature Generation

Twenty four participants (9 females, 15 males; mean age 36.37 years with the standard deviation of 16.67 years) were recruited over a period of 11 months from the local community. Five of the participants were manufacturing workers, and the remainder represented a convenience sample of students with varying degrees of physical work experience. All participants reported that they were in good physical and mental health. In addition, they were screened by completing the Physical Activity Readiness Questionnaire (PAR-Q) [213] at the start of the session to assess their eligibility to participate. They also provided informed consents at the start of the experiment. All study procedures were approved by the university's institutional review board (IRB).

Participants completed one three-hour experimental session for the simulated MMH task and another for the SI task. The order of the two experiments was randomized and participants had to complete the experiments in different days. The MMH task involved palletizing and transporting several weighted containers (see Figure 3.3). Each participant was asked to perform the task at a set pace for three hours continuously (without breaks) to induce fatigue. Per the discussion in Section 3.3.1, four IMUs placed at the ankle, hip, wrist and torso, and a heart rate monitor on the chest were used for data collection. Furthermore, participants provided their subjective exertion (RPE) using the Borg Scale [104] every ten minutes.

Figure 3.3: A participant carrying out the MMH task, adopted from chapter 2

The four step data cleaning procedure discussed in Section 3.3.1 was deployed for our case studies. After using the low pass filter for de-noising the IMU data, we used $RPE \geq 13$ as a cutoff for fatigue in step 2 per the analysis in chapter 2. Based on step 2, a total of nine participants were removed from the data for the following reasons: (a) three participants did not get fatigued by the end of the experiment; (b) three reported being fatigued within the first half an hour of the experiment (i.e. they may have been fatigued prior to conducting the experiment); (c) the IMUs failed to record data for two of the participants during the experiment; and (d) one of the participants deviated from the experimental protocol by taking two 10-minute bathroom breaks. As a result, we ended up with 15 participants whose data were deemed reliable for analysis. After synchronizing the data from the sensors in step 3, we removed the first 10 minutes of experimental data to avoid capturing the learning effect [58]. Then, the % HRR was computed in step 4 as explained in the methodology section. After step 4, the jerk and posture profiles were generated based on the procedure of Baghdadi et al. (2018) [50] which was highlighted in Section 3.3.1.

To reduce the computational burden and to maintain a balanced dataset for training, we have only kept 20% of the data for each participant. These 20% corresponded to: (a) 10% (i.e. $10\% \times 180$ minutes $= 18$ minutes) at the beginning of the experiment, after the first 10 minutes are removed, where the participants are not fatigued, and (b) 10% at the end, where the participants are fatigued. The rationale for removing the 80% of the data is two-fold. First, the separation ensures that the differences between the fatigued and non-fatigued data for each participant are maximized, while the differences within each group are minimal. Second, based on chapter 2, we can assume that the size of the data can be decreased without losing much information related to fatigue detection. For each participant, we coded the response as 0 (for the first 18 minutes) and 1 for the latter 18 minutes to reflect the non-fatigued and fatigued states, respectively. Recall that our data cleaning procedure ensured that these values reflect the estimated RPEs by each participant.

Based on the discussion in Section 3.3.1, it is important to set the size of the time window prior to generating the features in Table 3.2. In our case studies, we have used a non-overlapping time window of 2 minutes. This means that each of the 18 minutes was divided into nine fractions of two-minute periods. The rationale for selecting two-minutes for the time window was mainly based on the observation that the average cycle time for MMH was approximately one minute. Therefore, each two-minute time interval is guaranteed to include at least one cycle of the task. Based on this decision, we generated the proposed features from each sensor for each two-minute time window. The reader can replicate our analysis by consulting our data and code (see the *Supplementary Materials* Section).

Model Construction and Validation

As a first step for feature selection, time series plots of all features were constructed to evaluate which features were virtually unchanged from the non-fatigued to fatigued states. Based on the visualizations, 15 (of the 55 candidate) features were dropped. The second step (where wrapper or embedded methods are used) of feature selection is applied after the training and test samples are generated using the leave p-participants out cross validation approach. Based on the discussion in Section 3.4.1, we had 15 participants with reliable data for this case study. Thus, $p = 2$ (i.e. $2/15 = 13\%$) was used for the leave p-participants-out cross validation approach to split the data into training and test sets. This resulted into 105 possible training/test sets ($15!/((15-2)! \times 2!) = 105$), which we would evaluate to obtain an estimate of the variation in the performance of our analytical models.

Prior to deploying the analytical models, two additional tasks were carried out. First, the last step of variable selection was deployed using two popular methods: *best subset selection* and *LASSO* (refer to Section 3.3.1 for details). Second, to reduce the bias from model training and improve the performance of the predictive models bootstrap resampling with replacement was applied to the training data. The sample size for each bootstrap sample was $n = 234$, which was based on 13 participants $\times$ 18 samples per participant. For our analysis, we used 200 bootstrap samples (each having $n = 234$) based on the recommendation of Pattengale et al. (2009) [214].

To develop the fatigue prediction models, several methods were applied during our preliminary analysis of the data. The models evaluated included: *logistic regression, penalized* *logistic regression, decision trees (DT), naive Bayes (NB), k-Nearest Neighbors (kNN), support vector machines (SVM)*, and *three ensemble models* (*random forest (RF), bagging,* and *boosting*). Due to their relatively poor performance, DT, NB and kNN were eliminated. In addition, models using *best subset selection* typically had better prediction performance with less features than their LASSO counterparts. Therefore, our case study focused on using the *best subset selection* with the following five analytical models: (a) logistic regression, (b) SVM, (c) RF, (d) RF with bagging (hereafter bagging), and (e) RF with boosting (hereafter boosting). In addition, we compared these five models to the approach in chapter 2 since it was the only research that considered multiple tasks in the context of occupational fatigue (see Table 3.2). To ensure that the comparison is fair, we considered two different variants of the penalized logistic regression approach with LASSO proposed in chapter 2. The first is utilizing their approach and features (on our data), and the second involves using their methodology with our features and data. In our estimation, this allows us to better evaluate whether our proposed method is superior to theirs. The reader should note that they did not consider model interpretation in their feature generation and thus we expect that our features are easier to interpret by practitioners.

Fatigue Detection Results

In Table 3.4, the predictive performance of our five models is compared with the two variants from chapter 2. The table shows the mean (and standard deviation in parentheses) for each of our four metrics. In addition, the average number of features selected by each model is also presented. The reported results are based on 105 constructed test datasets from the two-participants-out cross validation. For the first three numeric columns, a higher value is desired since it reflects a better prediction performance. The consistency column captures the average absolute difference between the sensitivity and specificity for each model, evaluated on the 105 test datasets. It is noted that the smaller the consistency is, the similar performance in detecting fatigued and no-fatigued states simultaneously would be. Moreover, a smaller number of features facilitates the interpretation of the model, which is important in the fatigue identification and diagnosis phases.

Table 3.4: Mean performance and the corresponding standard deviation of the classification methods for fatigue detection in MMH task, (the recommended model is **in bold**)

Category	Model	Sensitivity	Specificity	Accuracy	Consistency	# of Features
	Bagging	**0.872 (0.13)**	**0.869 (0.15)**	**0.870 (0.09)**	**0.143 (0.17)**	**5.35**
	Boosting	0.871 (0.13)	0.872 (0.15)	0.870 (0.08)	0.147 (0.17)	5.352
BSS	Random Forest	0.879 (0.14)	0.879 (0.15)	0.879 (0.09)	0.152 (0.18)	5.352
	Support Vector Machine	0.811 (0.18)	0.828 (0.17)	0.820 (0.11)	0.198 (0.19)	5.352
	Logistic Regression	0.790 (0.17)	0.766 (0.20)	0.778 (0.11)	0.227 (0.20)	5.352
LASSO	Penalized Logistic Regression*	0.802 (0.20)	0.916 (0.11)	0.859 (0.11)	0.175 (0.20)	18.943
	Penalized Logistic Regression	0.810 (0.13)	0.775 (0.17)	0.793 (0.08)	0.197 (0.16)	11.133

* features used in the model are only those generated in chapter 2

Four main observations from Table 3.4 need to be highlighted. First, as expected from the preliminary analysis, the number of features selected with the *best subset selection* are much less than those selected by the LASSO model. This means that the usability of the analytical models with the BSS model is much higher than that with LASSO since practitioners' need to monitor and understand approximately five features (instead of 11 or 19). Second, the performance of all seven models is relatively high with an overall average accuracy greater than 0.77. Third, the performance of the three ensembles is better than the remaining models. Fourth, the penalized logistic regression in chapter 2 outperforms its variant with our features from a prediction perspective. However, this comes at the cost of adding eight features to the model (i.e. $\approx 70\%$ increase in the variables used). Based on these observations and this case study, one can conclude that our framework has shown higher detection performance (with less features) when compared to competing models from the literature.

The next logical research question is to examine how the prediction performance varies while limiting the number of sensors used. To evaluate this question, we utilize the bagging model since Table 3.4 showed that it had the lowest consistency and had similar prediction performance to the two other ensembles. Table 3.5 reports the prediction results, when features are limited to those from one, two, three, four and all sensor combinations. Note that the values that are not shown in the table (e.g. Ankle, Hip, Wrist and HR sensors) reflect scenarios when a prediction was not possible. This means that the main features that detected the fatigue were eliminated with the added constraints on which possible features to select from.

From the results in Table 3.5, one can see that the prediction performance does not vary significantly as the number of sensors' are changed. For example, the average accuracy varies from 0.850 to 0.871 (with a standard deviation $\approxeq 0.09$) as the number of sensors vary. This is

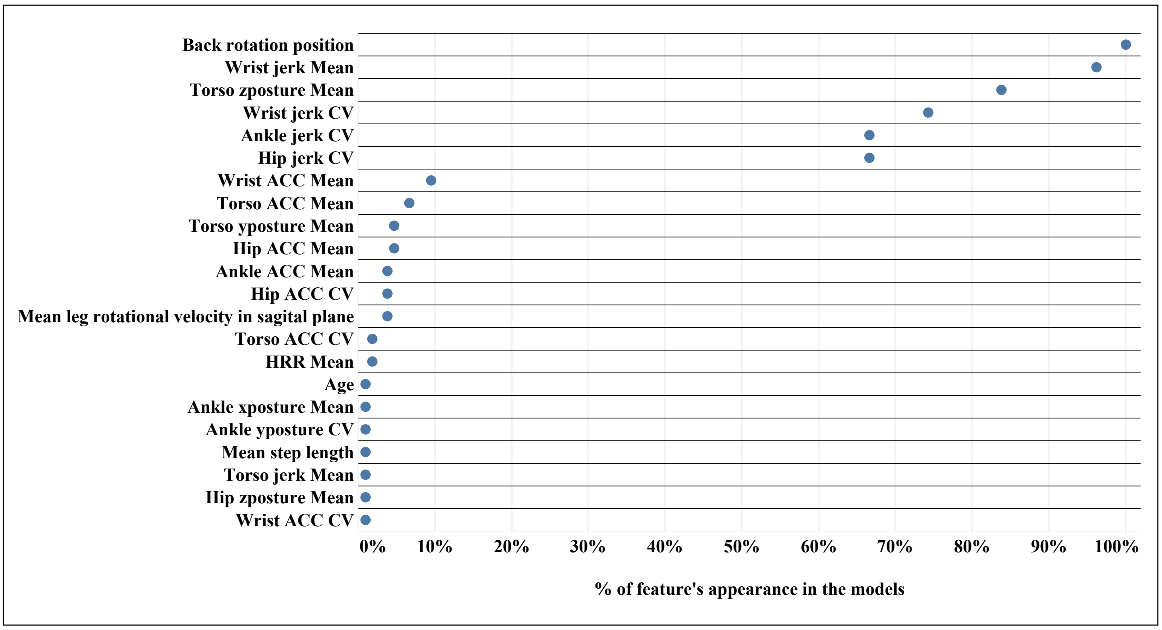
Table 3.5: Mean performance and the corresponding standard deviation of the Bagging model for fatigue detection using different sensor combinations for the MMH task (the recommended model is **in bold**)

# sensors	Sensor Combination					Sensitivity	Specificity	Accuracy	Consistency
5	Ankle	Hip	Wrist	Torso	HR	0.872 (0.13)	0.869 (0.15)	0.870 (0.09)	0.143 (0.17)
4	Ankle	Hip	Wrist	Torso		0.875 (0.13)	0.868 (0.15)	0.871 (0.09)	0.141 (0.17)
	Ankle	Hip		Torso	HR	0.850 (0.15)	0.875 (0.13)	0.862 (0.09)	0.142 (0.16)
		Hip	Wrist	Torso	HR	0.877 (0.12)	0.863 (0.15)	0.870 (0.08)	0.144 (0.15)
	Ankle		Wrist	Torso	HR	0.872 (0.12)	0.864 (0.15)	0.868 (0.08)	0.146 (0.16)
	Ankle	Hip	Wrist		HR	-	-	-	-
3			Wrist	Torso	HR	0.877 (0.12)	0.864 (0.15)	0.870 (0.08)	0.141 (0.15)
	Ankle			Torso	HR	0.844 (0.15)	0.874 (0.13)	0.859 (0.09)	0.141 (0.16)
	Ankle	Hip		Torso		0.850 (0.15)	0.875 (0.13)	0.862 (0.09)	0.142 (0.16)
		Hip	Wrist	Torso		0.877 (0.12)	0.863 (0.15)	0.870 (0.08)	0.143 (0.16)
		Hip		Torso	HR	0.859 (0.15)	0.874 (0.14)	0.866 (0.10)	0.143 (0.16)
	Ankle		Wrist	Torso		0.873 (0.12)	0.863 (0.15)	0.868 (0.08)	0.145 (0.16)
	Ankle	Hip			HR	-	-	-	-
	Ankle	Hip	Wrist			-	-	-	-
	Ankle		Wrist		HR	-	-	-	-
		Hip	Wrist		HR	-	-	-	-
2			Wrist	Torso		0.877 (0.12)	0.864 (0.15)	0.870 (0.08)	0.141 (0.15)
	Ankle			Torso		0.844 (0.15)	0.874 (0.13)	0.859 (0.09)	0.141 (0.16)
		Hip		Torso		0.859 (0.15)	0.875 (0.14)	0.867 (0.10)	0.143 (0.16)
				Torso	HR	0.842 (0.15)	0.859 (0.14)	0.850 (0.10)	0.144 (0.16)
	Ankle	Hip				-	-	-	-
	Ankle				HR	-	-	-	-
	Ankle		Wrist			-	-	-	-
		Hip	Wrist			-	-	-	-
		Hip			HR	-	-	-	-
			Wrist		HR	-	-	-	-
1				**Torso**		**0.842 (0.15)**	**0.859 (0.14)**	**0.850 (0.10)**	**0.144 (0.16)**
	Ankle					-	-	-	-
		Hip				-	-	-	-
			Wrist			-	-	-	-
					HR	-	-	-	-

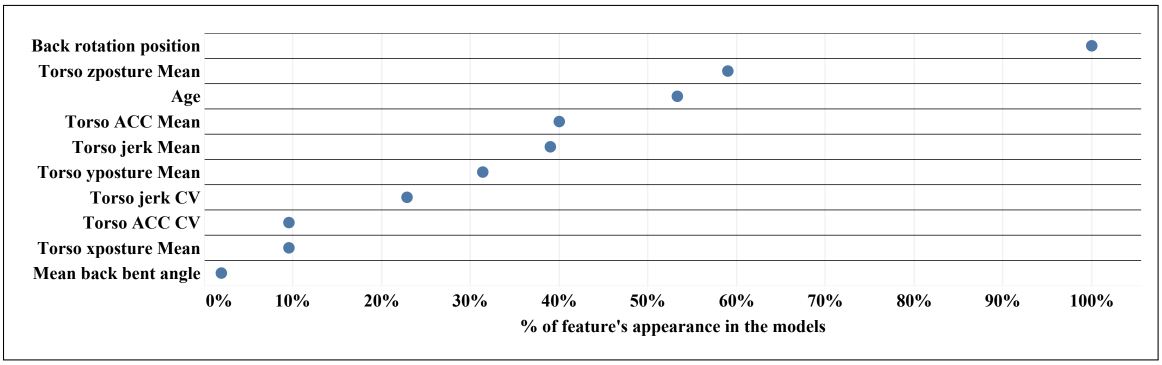
only true if the torso IMU is included in the analysis. Based on this observation, we recommend only using the torso IMU sensor for detecting fatigue in manual material handling environments (that are similar to those analyzed in our case study). While the prediction performance is almost the same, the costs incurred by the firm are much lower, and the usability of the system by using only one sensor is significantly improved. This is an important practical takeaway, which has not been reported in previous studies investigating fatigue in MMH tasks (see the references in Table 3.1).

Fatigue Identification Results

A first step in understanding fatigue is to examine how frequently a feature is selected all of the 105 *two-participants-out cross validation bagging model* test sets. In this section, we limit our analysis to two cases: (a) when all five sensors are utilized, and (b) when only the torso sensor is used. The results for these analyses are shown in Figures 3.4a and 3.4b, respectively. From both figures, one can see that all three categories of features (i.e. statistical, biomechanical, and individual features) are selected in our models. For the five sensor case, one biomechanical feature (*mean back rotational position*, i.e. feature #53 in Table 3.2) and five statistical features appeared in more than 65% of the models. All other remaining features appeared in less than 10% of the models. On the other hand, age becomes a much more predictive factor if we only rely on the torso sensor. In that case, *back rotational position* is still selected in 100% of the models.



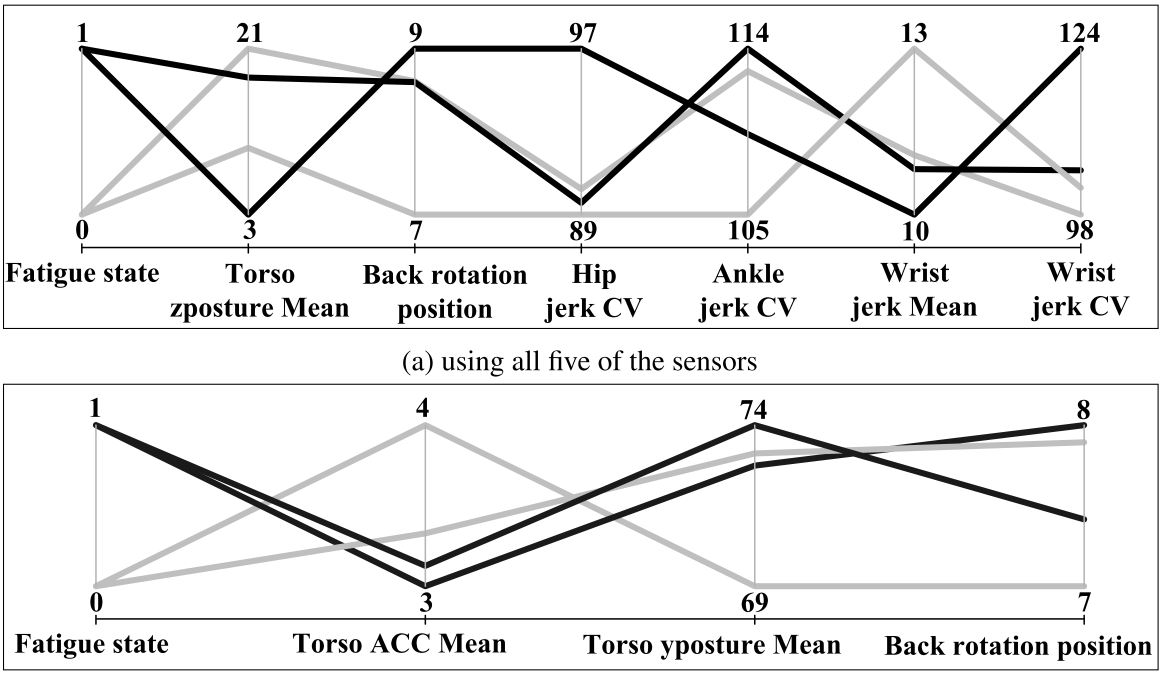
(a) using all five of the sensors



(b) using the torso sensor only

Figure 3.4: Important features visualization in the MMH task using the Bagging model

Once a list of predictive/important features is established, we then investigate how those features vary as the participant transition from the non-fatigued to fatigued states. As highlighted in Section 3.3.2, this analysis can be done visually using a *parallel coordinates plot*. Figure 3.5 depicts this analysis (using the median model sorted by accuracy) for the five sensors and one sensor cases. Note that the lines graphed in these plots represent the average values per variable for each of the two participants in the test set examined by the median model.



(a) using all five of the sensors

(b) using the torso sensor only

Figure 3.5: Features visualization for the median Bagging model in the MMH task

From Figure 3.5a, one can see that all of the six features highlighted in 3.4a are present in the median model. It is interesting to note that only the wrist features exhibited a consistent pattern across both participants when examining the fatigued cases (**black line**) and the non-fatigued cases (**gray line**). Specifically, the *coefficient of variation for wrist jerk* tended to be higher, and the *mean wrist jerk* tended to be lower in the fatigued cases. For the remaining four features, there were not any consistent patterns for both test subjects. Similarly from Figure 3.5b, one can see that only the *torso ACC mean* feature showed a clear separation between the fatigued and non-fatigued states for both participants. We hypothesize that these two figures may provide justification for why the ensemble models outperformed the logistic regression models. Specifically, these plots may suggest interactive and non-linear effects that can be trained for and captured using the ensemble models.

Fatigue Diagnosis Results

From the fatigue identification results, one can conclude that the type of fatigue is localized at the back. This conclusion is supported by: (a) the prediction performance is almost unchanged (and high) when only the features from the torso sensor are used for prediction, and (b) the *mean back rotational position* was selected as an important feature in 100% of the models. This was the only feature that was selected in 100% of the models. Our results are consistent with findings in the ergonomics literature, which suggest that *manual material handling* may lead to a higher prevalence of back injuries [215].

3.4.2 Case Study 2: Supply Pick up and Insertion

Task Description and Data Preparation

Similar to the task in chapter 2, we examined supply pickup and insertion task. The task involved walking while carrying supplies, and then bending forward to unscrew and fasten bolts at the supply box (destination). A snapshot of the experiment is provided in Figure 3.6. The task's cycle time was set for two minutes to mimic the activity in chapter 2. By design, this activity should be less fatiguing than the MMH task of the first case study.

Figure 3.6: Sensor placement on a participant for SI task, adopted from chapter 2

The mechanism used to collect and preprocess data is similar to that used in case study 1. The four step data cleaning procedure suggested in Section 3.3.1 resulted in having 13 participants (instead of 15 for the first case study) with reliable/clean data. Then, the sensor data were synchronized after removing the initial ten minutes of the experiment. After down sampling, the jerk, posture, and % HRR profiles were computed. Similar to the MMH task, the first 18 minutes of the data (after removing the learning period) were labeled as *not fatigued* and the last 10 minutes were marked as *fatigued*.

From those two-eighteen minute periods, we generated the list of features in Table 3.2. Based on the visual feature selection procedure, 41 of those features were retained for further analysis. The leave-two-participants-out cross validation resulted in 78 training and test datasets. This is smaller than the datasets used in the first case study since the number of participants with reliable data was smaller. Two hundred bootstrap samples with fixed sample size (11 participants $\times$ 18 samples per participant = 198) were used to evaluate the stability of proposed models.

To reduce the computational burden, we only examined the seven models analyzed in case study 1. This means that we did not examine whether the kNN, NB or decision trees performed adequately for this task. The results for using these seven models for fatigue detection are presented in the following subsection.

The predictive performance of the seven models is summarized in Table 3.6. Similar to Table 3.4, this table shows the mean (and standard deviation in parentheses) for each of the four performance measures as well as the average number of features selected by each model. The reader should note the reported results are based on 78 constructed test datasets from the two-participants-out cross validation.

Table 3.6: Mean performance and the corresponding standard deviation of the classification methods for fatigue detection in SI task, (the recommended model is **in bold**)

Category	Model	Sensitivity	Specificity	Accuracy	Consistency	# of Features
	Bagging	0.863 (0.12)	0.910 (0.10)	0.886 (0.08)	0.097 (0.13)	6.346
	Random Forest	**0.876 (0.12)**	**0.918 (0.10)**	**0.897 (0.08)**	**0.100 (0.13)**	**6.346**
BSS	Boosting	0.868 (0.12)	0.893 (0.12)	0.880 (0.09)	0.118 (0.13)	6.346
	Support Vector Machine	0.728 (0.19)	0.847 (0.16)	0.787 (0.12)	0.226 (0.16)	6.346
	Logistic Regression	0.525 (0.28)	0.723 (0.21)	0.624 (0.12)	0.391 (0.27)	6.346
LASSO	Penalized Logistic Regression*	0.674 (0.19)	0.925 (0.15)	0.800 (0.14)	0.257 (0.23)	16.179
	Penalized Logistic Regression	0.748 (0.22)	0.824 (0.06)	0.786 (0.10)	0.151 (0.16)	20.868

* features used in the model are only those generated in chapter 2

There are two main observations to be made pertaining to the results in Table 3.6. First, the number of features selected with the *best subset selection* are much less than those selected by the LASSO model. This means that the usability of the analytical models with the BSS model is much higher than that with LASSO since practitioners' need to monitor and understand approximately six features (instead of 16 or 21). Second, the prediction performance of the three ensembles is much higher than all other models. Note that the performance gap is much larger in this task than in the MMH task. Based on this case study, our framework has shown higher detection performance (with less features) when compared to competing models from the literature.

Next, we examine how the prediction performance varies while restricting the number of sensors used when performing SI task. To gage this question, we utilize the *random forest model* since Table 3.6 showed that it had the highest prediction performance. Table 3.7 shows the prediction results when features are limited to those from one, two, three, four and all sensor combinations. Similar to the earlier example, the values, which are not shown reflect scenarios when a prediction was not possible.

From the results in Table 3.6, one can observe that the prediction performance does not vary significantly as the number of sensors' are changed. For instance, the average accuracy varies from 0.854 to 0.897 (with standard deviations $\approxeq$ 0.05) as the number of sensors vary. Note that this observation only holds if the heart rate sensor is included in the analysis. Accordingly, using solely the heart rate sensor is appropriate for detecting fatigue in supply pick up and insertion environments (that are similar to those analyzed in our case study). Similar to the earlier case study, this is a novel contribution (showcasing that one sensor can present similar performance to multiple sensors with a much higher usability).

Fatigue Identification Results

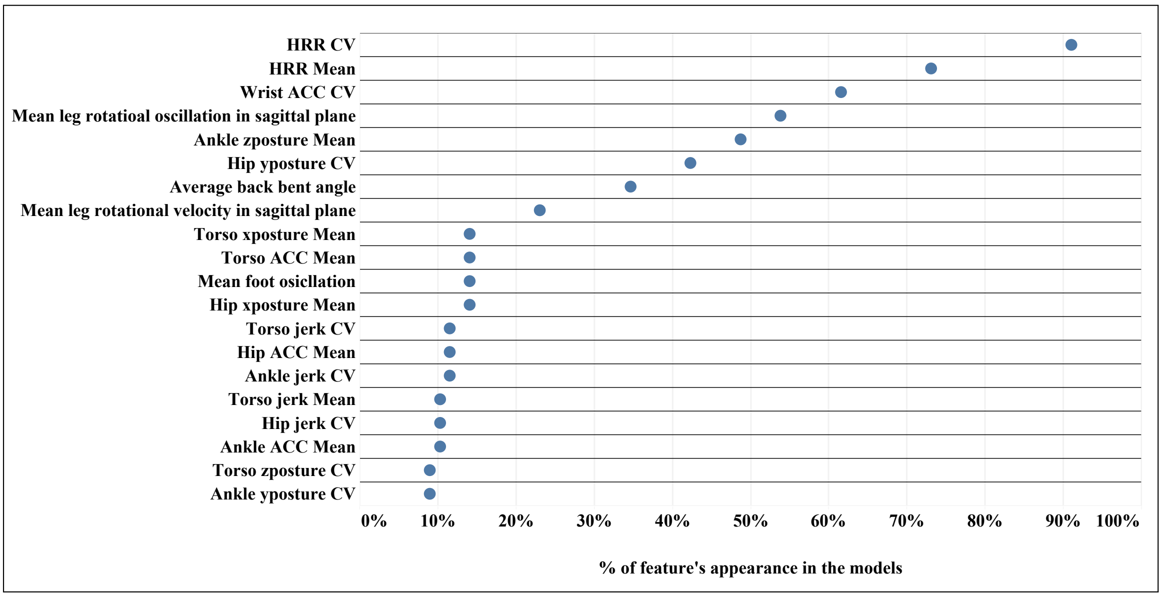
As in case study 1, we follow the two-step approach for fatigue identification. First, we examine how frequently a feature is selected from all of the 78 *two-participants-out cross validation random forest model* test sets. We limit the analysis to two cases: (a) when all five sensors are utilized, and (b) when only the heart rate sensor is used. The results from these analyses are

Table 3.7: Mean fatigue detection performance (and the corresponding standard deviation) of the random forest model using different sensor combinations for the SI task (the recommended approach is **in bold**)

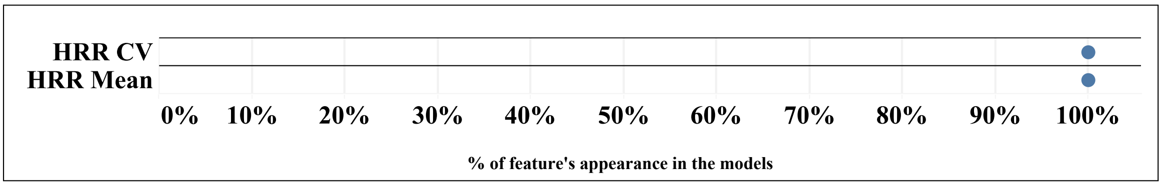
# sensors	Sensor Combination					Sensitivity	Specificity	Accuracy	Consistency
5	Ankle	Hip	Wrist	Torso	HR	0.876 (0.12)	0.918 (0.10)	0.897 (0.08)	0.100 (0.13)
4	Ankle	Hip	Wrist	Torso		0.863 (0.13)	0.911 (0.05)	0.887 (0.07)	0.097 (0.12)
	Ankle	Hip		Torso	HR	0.853 (0.13)	0.893 (0.12)	0.873 (0.09)	0.107 (0.13)
		Hip	Wrist	Torso	HR	0.853 (0.17)	0.911 (0.14)	0.882 (0.13)	0.121 (0.14)
	Ankle	Hip	Wrist		HR	0.834 (0.16)	0.921 (0.10)	0.877 (0.10)	0.132 (0.16)
	Ankle		Wrist	Torso	HR	0.826 (0.19)	0.955 (0.05)	0.890 (0.04)	0.138 (0.19)
3		Hip	Wrist	Torso		0.867 (0.08)	0.904 (0.08)	0.885 (0.05)	0.100 (0.09)
	Ankle	Hip			HR	0.856 (0.07)	0.887 (0.07)	0.871 (0.04)	0.117 (0.15)
		Hip	Wrist		HR	0.831 (0.07)	0.923 (0.06)	0.877 (0.04)	0.124 (0.14)
	Ankle			Torso	HR	0.825 (0.08)	0.935 (0.06)	0.880 (0.04)	0.131 (0.14)
			Wrist	Torso	HR	0.818 (0.07)	0.927 (0.06)	0.872 (0.04)	0.131 (0.15)
		Hip		Torso	HR	0.852 (0.07)	0.874 (0.08)	0.863 (0.05)	0.132 (0.16)
	Ankle		Wrist		HR	0.820 (0.06)	0.957 (0.04)	0.888 (0.04)	0.147 (0.19)
	Ankle	Hip	Wrist			-	-	-	-
	Ankle	Hip		Torso		-	-	-	-
	Ankle		Wrist	Torso		-	-	-	-
2				Torso	HR	0.823 (0.08)	0.896 (0.07)	0.859 (0.05)	0.109 (0.12)
	Ankle				HR	0.828 (0.08)	0.917 (0.07)	0.872 (0.04)	0.114 (0.13)
		Hip			HR	0.837 (0.07)	0.904 (0.07)	0.870 (0.05)	0.117 (0.14)
			Wrist		HR	0.818 (0.08)	0.920 (0.06)	0.869 (0.05)	0.123 (0.14)
		Hip		Torso		-	-	-	-
	Ankle		Wrist			-	-	-	-
	Ankle	Hip				-	-	-	-
	Ankle			Torso		-	-	-	-
		Hip	Wrist			-	-	-	-
			Wrist	Torso		-	-	-	-
1					**HR**	**0.820 (0.08)**	**0.889 (0.07)**	**0.854 (0.05)**	**0.102 (0.13)**
	Ankle					-	-	-	-
		Hip				-	-	-	-
			Wrist			-	-	-	-
				Torso		-	-	-	-

shown in Figures 3.7a and 3.7b, respectively. Neither cases included any individual features (which is different from the earlier case when age appeared in both). Only statistical features were selected in the one sensor model, which is perhaps not surprising since none of the biomechanical features can be generated if only the heart rate sensor is used. For the five sensor case, one biomechanical feature (*mean leg rotational oscillation in sagittal plane*), i.e. feature #51 in Table 3.2 and three statistical features appeared in more than 50% of the models. On the other hand, in the single sensor case, all the statistical features (*HRR CV*, *HRR Mean*) created using the heart sensor were selected in 100% of the models.

Second, we investigate how those features range as participants transition from the non-fatigued to fatigued states. Figure 3.8 illustrates this analysis (using the median model sorted by accuracy) for: (a) the five sensors, and (b) the one sensor cases. Recall that the lines graphed in these plots represent the average values per variable for each of the two participants in the test set examined by the median model. The conclusion is similar to that of case study 1, where



(a) using all five of the sensors



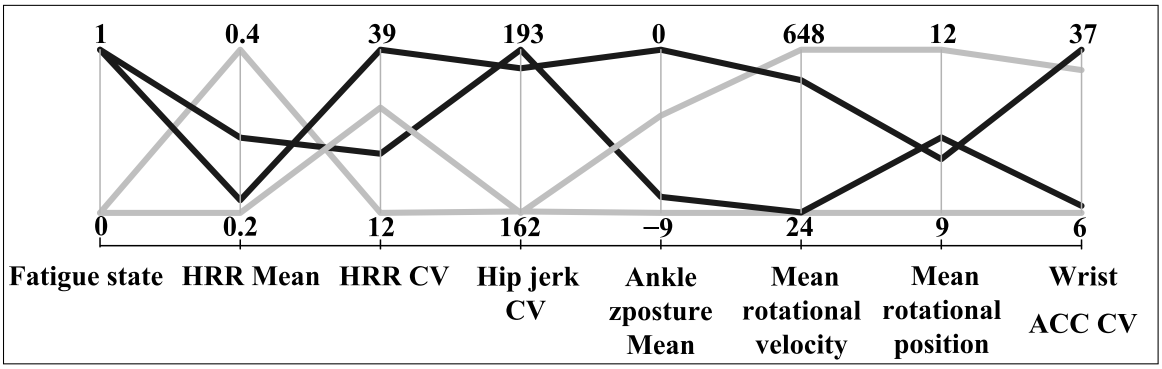
(b) using the heart rate sensor only

Figure 3.7: Important features visualization in the SI task using the Random Forest model

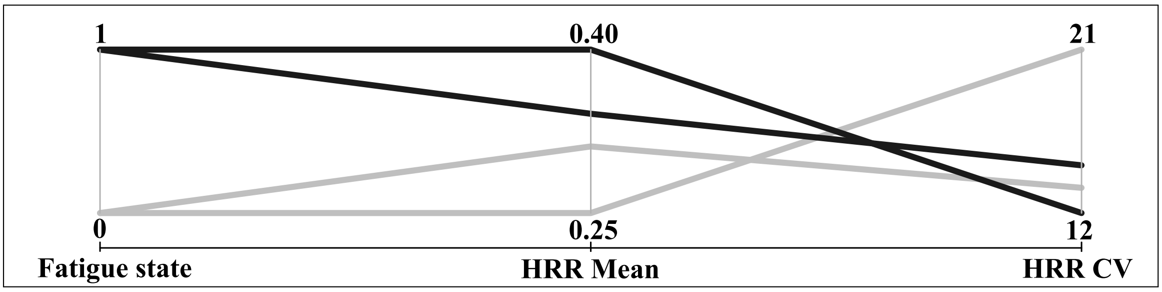
only one feature had different values for the non-fatigue (gray line) and fatigue cases (**black line**) across the two test participants. However, here, this effect is only observed for the one sensor case. Specifically, in Figure 3.8b, the mean HRR is higher in the fatigued state. This result makes sense since an increased heart rate is a fatigue symptom (see Cavuoto & Megahed (2017) [177] for more details).

Fatigue Diagnosis Results

From the fatigue identification results, one can conclude that the participants experience whole-body fatigue in the SI task. This conclusion is based on the ability to accurately detect the non-fatigue and fatigue states through the use of only the heart rate sensor. The elevated mean percent HRR shown for both participants in Figure 3.8b supports this conclusion.



(a) using all five of the sensors



(b) using HR sensor only

Figure 3.8: Features visualization for the median Random Forest model in the SI task

3.5 Discussion and Conclusion

3.5.1 Summary of the Main Contributions

In this chapter, we proposed an integrated framework for managing fatigue in manufacturing workplaces using minimally-intrusive wearable sensors. Based on the case studies in Section 3.4, this study makes three main contributions. First, we demonstrated the capability of using a unified modeling approach for managing physical fatigue in different occupational tasks/settings. The case studies show the ability to detect, identify, and diagnose fatigue in multiple occupationally-relevant settings. The ability to identify/diagnose fatigue through the use of wearable sensors has not been shown prior in the literature. Second, the insights from the *fatigue identification* phase of our framework can be used to inform sensor placement and selection. We demonstrated that the prediction performance using one sensor is equivalent to that of using all sensors for our two case studies. Third, we showed that the importance of different types of features varies with different manufacturing tasks. Significant features contributing to the determination of physical fatigue in the manual material handling using the torso sensor includes:

- Torso Acceleration Mean: This feature represents the average of the torso acceleration. Slowing down the activity when compared to the corresponding activity pace at the beginning of the experiment would suggest that the participant was physically fatigued. Results regarding the torso acceleration during the MMH task is consistent with those reported the literature. Since torso flexion is identified as a critical factor to increase the risk of low back pain [61], increases in torso kinematics would increase the overall low back disorder risk [61].

- Torso Y Posture Mean: This feature measures the average angular position of the torso in the transverse plane. It is associated with the twisting and lateral bending. The increased twisting angle has been observed for the fatigued participants during the task. In the MMH task, taking a step instead of twisting the torso may reduce the risk of low back pain [61]. Despite twisting is one of the factors for reporting the occupational low back pain [64, 65, 61], there is no well-known prevention strategies for torso twisting and lateral bending [61].

- Back Rotation Position: This feature measures the average range of torso bending in the sagittal plane while doing the MMH. Results in this study show that a larger bending angle contributes to physical fatigue. When the combination of repetitive lifting and torso bending is performed, fatigue is most likely to happen [66]. Fatigued torso muscles can affect the stability of spine [216]. In addition, the increase motion in different planes will increase the loading on the spine [217], and will increase the risk of low back pain [61, 64].

On the other hand, the important features for the supply insertion task are:

- HRR Mean: this feature measures the average percentage of an individuals heart rate capacity. Results suggested that the fatigued participants used the higher capacity of their heart. Our result is consistent with the literature that high heart rate could be an indicator of fatigue [218]; fatigue is associated with the increase in the percentage of maximum capacity [219, 220].

- HRR CV: this feature measures the variability in using heart rate capacity. Results show that largest variability in maximum heart rate capacity may indicate physical fatigue. This is consistent with the literature; heart rate variability is an indicator of physical fatigue [221, 222, 223].

Thus, researchers and practitioners should consider this finding when developing models for detecting/managing fatigue in other production settings.

3.5.2 Relevance to Operations Management Practice and Research

In our estimation, the proposed framework and the case study findings have significant implications for both practice and research. From a practical perspective, we have shown that changes in a worker's physical performance can be detected and modeled using wearable sensors. Utilizing the principles behind the technology adoption model, we have shown that fatigue associated specialized jobs can be detected using one sensor (without a loss in prediction performance). The emphasis on fatigue identification and diagnosis through visual analytical approaches allows practitioners to identify the risks, which are to be tackled through an appropriate intervention strategy. In essence, our framework can provide near real-time insights into the well-begin of shop-floor workers and their associated productivity levels. This information can be incorporated into the safety and productivity components of the SQDCM (safety, quality, delivery, cost, and morale) lean production effectiveness dashboard.

Our framework attempts to bridge the gaps between predictive and prescriptive analytics in the context of human performance modeling. The sequential nature of our framework attempts to overcome the "black box" nature of machine learning algorithms. We have shown that the sequential application of predictive models when combined with visual analytic tools can provide insights for prescriptive interventions. Furthermore, this study demonstrates that futuristic production environments can capture in real-time the well-being of their workers in addition to the data typically captured on the equipment. This can allow for more dynamic operational interventions (e.g., work-rest scheduling models).

3.5.3 Limitations and Suggestions for Future Research

There are a few limitation that may influence the interpretation of our results. First, the sample sizes are small as a result of time committed by each participant. Second, the participants for our two case studies varied in age and experience. Some of them represented a convenience sample of undergraduate and graduate students who may have a limited experience with manufacturing operations. Others were recruited from industry, and as such, are much more experienced/trained. Thus, our 10 minute training window may not be sufficient for some participants, i.e. the baseline performance for the non-fatigued state may not reflect their true steady-state performance. Third, the fatigue detection models are based on the participants' perceived ratings of exertion. Different participants may have varying levels of pain tolerance. Thus, we implicitly assume that the aliasing of perception and fatigue will have the same effect on performance as fatigue alone. This assumption is reasonable based on the ergonomics literature. Specifically, Mehta and Cavuoto (2015) ([224] [p. 94]) state that "... muscle activation, perception of discomfort, and/or motivation, might have a greater contribution to fatigue development than peripheral factors". Fourth, the evaluation of our framework's performance was limited to focused lab experiments. Future studies should evaluate how this framework performs in the field.

In our estimation, there are three main streams of research that can capitalize on our framework and findings. First, studies should investigate how our framework can be extended to simultaneously monitor and manage fatigue for hundreds of workers. While our current prediction performance is excellent for an individual worker (and for typical predictive modeling applications in the literature), it will suffer from a high false alarm rate if implemented across the shop-floor. To alleviate this issue, future research should consider: (a) reducing the frequency of data collection, which would increase the average time (but not samples) between false alarms; and (b) controlling the false discovery rate [225], which is designed for testing multiple hypotheses. Second, there are several *information systems*, *ethical* and *legal* implications that arise from collecting workers' performance data. Policies that account for these implications are needed. Third, there is an excellent opportunity for *operations research* models that can optimize recovery (or alternatively minimize fatigue development) while meeting the demands of the production schedule and the resource constraints. Such models will benefit from the data-driven/real-time nature of our framework.

Supplementary material

In this study, the R programming language (https://www.rproject.org/) and MATLAB were used to generate the results. Our modeling raw data, code and result files are available in the following repository: https:// www.dropbox.com/sh/i884pib7gb4n3u9/AACCdaonvdcysXcpMw9a0DYca?dl=0.

Chapter 4

Conclusion and Summary of Dissertation Contributions

4.1 Dissertation Contributions

Occupational fatigue has been known to be a potential cause of accidents and injuries in manufacturing occupations. However, the management of occupational fatigue has been limited by a dependence on subjective methods such as surveys or personal interventions for managing individuals' accidents or injuries. Development of wearable technology allows for better monitoring of manufacturing workers. The combination of these technologies with advancements in data science is leading to various research opportunities on the application of the Internet of Things (IoT) as a robust method for the fatigue management in occupational environments. These developments provide a path towards the implementation of fatigue management in practice. However, it is essential to study what equipment, sensors, type of data collection, and which analyses are more efficient and user-friendly for fatigue management. This study aimed to provide a solution for these research issues which are vital in manufacturing occupations. This dissertation has several contributions to the literature related to creating safe manufacturing workplaces through data-driven approaches, which includes: 1) exploring the effectiveness of wearable sensors in fatigue detection, 2) comprehensive framework for fatigue management, 3) exploring technology adoption models for user-friendly sensor selection, 4) developing useful metrics for fatigue measurement.

In chapter 2, we have the first model (Penalized Regression Model) which proves that wearables can be used for detecting fatigue on the individual level in manufacturing occupations. We showed that combining accelerometers with a heart rate monitor provides optimal results since we can capture multiple features of fatigue development in the tested manufacturing tasks. Then, in chapter 3, we demonstrated that the unified modeling approach consisting of data collection, cross-validation, feature selection, bootstrapping and model development with the wearable sensors can detect fatigue with high performance in the complex tasks such as manual material handling and supply insertion, which are common tasks in most of the occupational environments. In this framework, the proposed modeling approach performs well as long as the torso or heart rate sensor is present for the manual material handling and supply insertion tasks, respectively. We were able to satisfy the constraints regarding ease of use and usefulness by reducing the number of sensors to two while achieving the desired accuracy and financial benefits. Regarding practicality, we recommend using a BioHarness sensor in which the embedded accelerometer is used to monitor both the heart rate and torso simultaneously. We showed how both statistical (e.g., HRR mean) and task-related features (e.g., back rotation position) generated from these sensors are essential in predicting the fatigue status of the workers and determining the fatigue type in manufacturing tasks, which would allow for improved intervention strategies for specific workers after fatigue diagnosis.

4.2 Future Work

Common methods for fatigue detection in occupational tasks rely on using IMUs and heart rate sensors. The current work showed fatigue management in occupational tasks based on using the torso and heart rate sensors. Going forward, it is imperative to explore the performance of the proposed fatigue management framework when put into practice. Therefore, future work can test the proposed framework for fatigue detection/diagnosis within workforces such as warehousing or construction.

In order to implement the developed analytical model into practice, the output from the analytical model can be monitored over time. In chapter 3, we showed that the state of an individual whether he/she is fatigued or not can be obtained from the analytical model. The output of the analytical model was a binary variable (fatigued or non-fatigued); however, the probability of being fatigued can also be obtained from the model. Therefore, to a) properly implement the analytical model into practice and b) manage the trade-off between false alarms and early detection of fatigue, the popular quality tool called control chart should be used to monitor the probability of being fatigued over time. Through the use of a control chart, it is possible to measure the performance of the model in practice in terms of generating false signals when the individual is not fatigued as well as the performance of the model in detecting the fatigued individual. We can design the control chart considering practitioners' criteria. For example, what should be the performance of the analytical model if we aim to detect the fatigued state for an individual with the probability of more than 90% when the individual is fatigued? By using the standard control chart, we can detect large shifts in the average probability of fatigue. However, there are several control charts such as EWMA or CUSUM in the literature of statistical process control (SPC) that detect small shifts in the process parameter. Also, several sensitizing rules are recommended in the SPC literature to detect an out of control situation. For example, if we observe 2 out of 3 observations more than 0.4 for the probability of being fatigued, given an upper control limit of 0.5, we may classify the individual as feeling fatigued. Therefore, an overall control chart can be used to monitor the probability of being fatigued in practice using the output of the analytical model, and it will inform a practitioner to select the appropriate analytical model that satisfies particular fatigue detection criteria.

In addition, while developing a fatigue detection model, other factors may have a significant effect on the development of physical fatigue such as a) Task-related factors (task duration, amount of walking, rest duration, and task difficulty) and b) Individual related factors (experience, history of injury, amount of exercise). By investigating how these factors influence fatigue development, better intervention strategies may be established. Based on the current findings for fatiguing body type, the interventions such as modifying work-rest scheduling, job rotation, task redesign should be further developed and evaluated.

4.3 Overall Conclusion

This dissertation contributes to the increasing, but yet limited, research on the occupationally related fatigue detection using wearable sensors. The findings suggest that occupational fatigue can be measured by using torso and heart rate sensors, particularly for the complex tasks such as manual material handling and supply insertion. Substantially, the low-cost sensors and fatigue management framework recommended in this research may have implications for determining the work-rest scheduling process. Furthermore, the visual analytics representation of important features may signify specific fatiguing body parts while the output of the analytical models can be used for monitoring the probability of fatigue over time to detect fatigued workers in general. The results of this dissertation can contribute to determining whether specific recovery accommodations may be needed after fatigued workers have been identified in order to improve their quality of life, increase their workplace productivity, and reduce work-related musculoskeletal disorders with the goal of achieving both productivity and safety.

Appendix A: Applied Ergonomic Copyright Permission

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jun 18, 2018

This Agreement between Auburn Univeristy -- Zahra Sedighi Maman ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4371970449875
License date	Jun 18, 2018
Licensed Content Publisher	Elsevier
Licensed Content Publication	Applied Ergonomics
Licensed Content Title	A data-driven approach to modeling physical fatigue in the workplace using wearable sensors
Licensed Content Author	Zahra Sedighi Maman,Mohammad Ali Alamdar Yazdi,Lora A. Cavuoto,Fadel M. Megahed
Licensed Content Date	Nov 1, 2017
Licensed Content Volume	65
Licensed Content Issue	n/a
Licensed Content Pages	15
Start Page	515
End Page	529
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Title of your thesis/dissertation	A Data Driven Framework to Predict the Fatigue among Manufacturing Workers Using Wearable Sensors
Expected completion date	Aug 2018
Estimated size (number of pages)	120
Requestor Location	Auburn Univeristy 257 S Gay St, Apt B301 AUBURN, AL 36830 United States Attn: Auburn Univeristy
Publisher Tax ID	98-0397604
Billing Type	Invoice

Billing Address	Auburn Univeristy 257 S Gay St, Apt B301 AUBURN, AL 36830 United States Attn: Auburn Univeristy
Total	0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If

full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.
10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:
15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at

http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.
Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.
17. **For journal authors:** the following clauses are applicable in addition to the above:
Preprints:
A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).
Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).
If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.
Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.
Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do

- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.
Policies for sharing publishing journal articles differ for subscription and gold open access articles:
Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.
Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.
19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.
Terms & Conditions applicable to all Open Access articles published with Elsevier:
Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or

reputation. If any changes have been made, such changes must be clearly indicated.
The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.
If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.
Additional Terms & Conditions applicable to each Creative Commons user license:
CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.
CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.
CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0.
Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.
Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

References

[1] Ohio Supercomputer Center. Ohio supercomputer center. http://osc.edu/ark:/19495/f5s1ph73, 1987.

[2] Lidong Wang. Machine learning in big data. *International Journal of Advances in Applied Sciences*, 4(4):117–123, 2016.

[3] Shengli Niu. Ergonomics and occupational safety and health: An ilo perspective. *Applied ergonomics*, 41(6):744–753, 2010.

[4] BLS. Nonfatal occupational injuries and illnesses requiring days away from work. ⟨http://www.bls.gov/news.release/osh2.nr0.htm⟩. Accessed (November 10, 2016).

[5] Shrawan Kumar. Theories of musculoskeletal injury causation. *Ergonomics*, 44(1):17–47, 2001.

[6] Mingzong Zhang, Emily H Sparer, Lauren A Murphy, Jack T Dennerlein, Dongping Fang, Jeffrey N Katz, and Alberto J Caban-Martinez. Development and validation of a fatigue assessment scale for us construction workers. *American journal of industrial medicine*, 58(2):220–228, 2015.

[7] Maaike A Huysmans, Marco JM Hoozemans, Allard J van der Beek, Michiel P de Looze, and Jaap H van Dieën. Position sense acuity of the upper extremity and tracking performance in subjects with non-specific neck and upper extremity pain and healthy controls. *Journal of Rehabilitation Medicine*, 42(9):876–883, 2010.

[8] Antonio Nardone, Jessica Tarantola, Andrea Giordano, and Marco Schieppati. Fatigue effects on body balance. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, 105(4):309–320, 1997.

[9] Hongwei Hsiao and Petre Simeonov. Preventing falls from roofs: a critical review. *Ergonomics*, 44(5):537–561, 2001.

[10] Roger A Haslam, Sophie A Hide, Alistair GF Gibb, Diane E Gyi, Trevor Pavitt, Sarah Atkinson, and AR Duff. Contributing factors in construction accidents. *Applied ergonomics*, 36(4):401–415, 2005.

[11] Ann Williamson, David A Lombardi, Simon Folkard, Jane Stutts, Theodore K Courtney, and Jennie L Connor. The link between fatigue and safety. *Accident Analysis & Prevention*, 43(2):498–515, 2011.

[12] Marcus Yung. Fatigue at the workplace: Measurement and temporal development. PhD Dissertation, University of Waterloo. Located at: https://uwspace.uwaterloo.ca/handle/10012/10119, 2016.

[13] Bart Visser and Jaap H van Dieën. Pathophysiology of upper extremity muscle disorders. *Journal of Electromyography and Kinesiology*, 16(1):1–16, 2006.

[14] Lin Lu, Fadel M Megahed, Richard F Sesek, and Lora A Cavuoto. A survey of the prevalence of fatigue, its precursors and individual coping mechanisms among us manufacturing workers. *Applied Ergonomics*, 65:139–151, 2017.

[15] Osami Kajimoto. *Fatigue Science for Human Health*, chapter Development of a Method of Evaluation of Fatigue and Its Economic Impacts, pages 33–46. Springer Japan, Tokyo, 2008.

[16] Peter C Winwood, Arnold B Bakker, and Anthony H Winefield. An investigation of the role of non–work-time behavior in buffering the effects of work strain. *Journal of Occupational and Environmental Medicine*, 49(8):862–871, 2007.

[17] Judith A Ricci, Elsbeth Chee, Amy L Lorandeau, and Jan Berger. Fatigue in the us workforce: prevalence and implications for lost productive work time. *Journal of Occupational and Environmental Medicine*, 49(1):1–10, 2007.

[18] Robert Hockey. *The psychology of fatigue: Work, effort and control*. Cambridge University Press, 2013.

[19] Archibald Vivian Hill, CNH Long, and H Lupton. Muscular exercise, lactic acid, and the supply and utilisation of oxygen.parts -. *Proc. R. Soc. Lond. B*, 96(679):438–475, 1924.

[20] F Moscatelli, G Messina, A Valenzano, A Petito, AI Triggiani, MAP Ciliberti, V Monda, A Messina, D Tafuri, L Capranica, et al. Relationship between rpe and blood lactate after fatiguing handgrip exercise in taekwondo and sedentary subjects. *Biol Med (Aligarh)*, 1(2), 2015.

[21] TD Noakes and A St Clair Gibson. Logical limitations to the catastrophe models of fatigue during exercise in humans. *British Journal of Sports Medicine*, 38(5):648–649, 2004.

[22] Sean Gallagher and John R Heberger. Examining the interaction of force and repetition on musculoskeletal disorder risk a systematic literature review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 55(1):108–124, 2013.

[23] Sean Gallagher and Mark C Schall Jr. Musculoskeletal disorders as a fatigue failure process: evidence, implications and research needs. *Ergonomics*, 60(2):255–269, 2017.

[24] Kyle R Voge and Jonathan B Dingwell. Relative timing of changes in muscle fatigue and movement coordination during a repetitive one-hand lifting task. In *Proceedings of the 25th Annual international conference of the IEEE Engineering in Medicine and Biology Society*, volume 17, page 2, 2003.

[25] Jonathan B Dingwell, Jason E Joubert, Fernando Diefenthaeler, Joel D Trinity, et al. Changes in muscle activity and kinematics of highly trained cyclists during fatigue. *IEEE Transactions on Biomedical Engineering*, 55(11):2666–2674, 2008.

[26] P Dolan and MA Adams. Repetitive lifting tasks fatigue the back muscles and increase the bending moment acting on the lumbar spine. *Journal of biomechanics*, 31(8):713–721, 1998.

[27] Jason R Fuller, Karen V Lomond, Joyce Fung, and Julie N Côté. Posture-movement changes following repetitive motion-induced shoulder muscle fatigue. *Journal of Electromyography and Kinesiology*, 19(6):1043–1052, 2009.

[28] B Bigland-Ritchie, R Johansson, OC Lippold, and JJ Woods. Contractile speed and emg changes during fatigue of sustained maximal voluntary contractions. *Journal of neurophysiology*, 50(1):313–324, 1983.

[29] Svend Erik Mathiassen, Per Liv, and Jens Wahlström. Cost-efficient measurement strategies for posture observations based on video recordings. *Applied Ergonomics*, 44(4):609–617, 2013.

[30] Victor L Paquet, Laura Punnett, and Bryan Buchholz. Validity of fixed-interval observations for postural assessment in construction work. *Applied Ergonomics*, 32(3):215–224, 2001.

[31] Thomas R Waters, Vern Putz-Anderson, Arun Garg, and Lawrence J Fine. Revised niosh equation for the design and evaluation of manual lifting tasks. *Ergonomics*, 36(7):749–776, 1993.

[32] Arun Garg and Jay M Kapellusch. Applications of biomechanics for prevention of work-related musculoskeletal disorders. *Ergonomics*, 52(1):36–59, 2009.

[33] William S Marras, Robert G Cutlip, Susan E Burt, and Thomas R Waters. National occupational research agenda (nora) future directions in occupational musculoskeletal disorder health research. *Applied Ergonomics*, 40(1):15–22, 2009.

[34] Lora A Cavuoto and Maury A Nussbaum. Influences of obesity on job demands and worker capacity. *Current Obesity Reports*, 3(3):341–347, 2014.

[35] Thomas J Balkin, William J Horrey, R Curtis Graeber, Charles A Czeisler, and David F Dinges. The challenges and opportunities of technological approaches to fatigue management. *Accident Analysis & Prevention*, 43(2):565–572, 2011.

[36] Haiwei Dong, Izaskun Ugaldey, and Abdulmotaleb El Saddik. Development of a fatigue-tracking system for monitoring human body movement. In *Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2014 IEEE International*, pages 786–791. IEEE, 2014.

[37] Sunwook Kim and Maury A Nussbaum. Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. *Ergonomics*, 56(2):314–326, 2013.

[38] MN Nyan, Francis EH Tay, and E Murugasu. A wearable system for pre-impact fall detection. *Journal of Biomechanics*, 41(16):3475–3481, 2008.

[39] Alexandros Pantelopoulos and Nikolaos Bourbakis. A health prognosis wearable system with learning capabilities using nns. In *Tools with Artificial Intelligence, 2009. ICTAI'09. 21st International Conference on*, pages 243–247. IEEE, 2009.

[40] Ming-Zher Poh, Nicholas C Swenson, and Rosalind W Picard. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *Biomedical Engineering, IEEE Transactions on*, 57(5):1243–1252, 2010.

[41] Nicolas Vignais, Markus Miezal, Gabriele Bleser, Katharina Mura, Dominic Gorecky, and Frédéric Marin. Innovative system for real-time ergonomic feedback in industrial manufacturing. *Applied ergonomics*, 44(4):566–574, 2013.

[42] Mark C Schall, Nathan B Fethke, Howard Chen, and Fred Gerr. A comparison of instrumentation methods to estimate thoracolumbar motion in field-based occupational studies. *Applied Ergonomics*, 48:224–231, 2015.

[43] Di Wang, Fei Dai, and Xiaopeng Ning. Risk assessment of work-related musculoskeletal disorders in construction: State-of-the-art review. *Journal of Construction Engineering and management*, 141(6):04015008, 2015.

[44] Paolo Bonato, Gerold R Ebenbichler, Serge Henri Roy, Stephan Lehr, Martin Posch, Josef Kollmitzer, and Ugo Della Croce. Muscle fatigue and fatigue-related biomechanical changes during a cyclic lifting task. *Spine*, 28(16):1810–1820, 2003.

[45] Louis Atallah, Benny PL Lo, Rachel Christina King, and Guang-Zhong Yang. Sensor placement for activity detection using wearable accelerometers. In *BSN*, pages 24–29, 2010.

[46] Patrick J Sparto, Mohamad Parnianpour, Tom E Reinsel, and Sheldon Simon. The effect of fatigue on multijoint kinematics and load sharing during a repetitive lifting test. *Spine*, 22(22):2647–2654, 1997.

[47] Julie N Côté, Pierre A Mathieu, Mindy F Levin, and Anatol G Feldman. Movement reorganization to compensate for fatigue during sawing. *Experimental brain research*, 146(3):394–398, 2002.

[48] Ling Bao and Stephen S Intille. Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing*, pages 1–17. Springer, 2004.

[49] Kevin M Pline, Michael L Madigan, Maury A Nussbaum, and Robert W Grange. Lumbar extensor fatigue and circumferential ankle pressure impair ankle joint motion sense. *Neuroscience letters*, 390(1):9–14, 2005.

[50] Amir Baghdadi, Fadel M Megahed, Ehsan T Esfahani, and Lora A Cavuoto. A machine learning approach to detect changes in gait parameters following a fatiguing occupational task. *Ergonomics*, (just-accepted):1–21, 2018.

[51] Christina Strohrmann, Holger Harms, Cornelia Kappeler-Setz, and Gerhard Troster. Monitoring kinematic changes with fatigue in running using body-worn sensors. *IEEE transactions on information technology in biomedicine*, 16(5):983–990, 2012.

[52] VH Hildebrandt, PM Bongers, J Dul, FJH Van Dijk, and HCG Kemper. The relationship between leisure time, physical activities and musculoskeletal symptoms and disability in worker populations. *International archives of occupational and environmental health*, 73(8):507–518, 2000.

[53] Gil Harari, Manfred S Green, and Shira Zelber-Sagi. Combined association of occupational and leisure-time physical activity with all-cause and coronary heart disease mortality among a cohort of men followed-up for 22 years. *Occup Environ Med*, 72(9):617–624, 2015.

[54] Andreas Holtermann, Hermann Burr, Jørgen V Hansen, Niklas Krause, Karen Søgaard, and Ole S Mortensen. Occupational physical activity and mortality among danish workers. *International archives of occupational and environmental health*, 85(3):305–310, 2012.

[55] Hans Heneweer, Filip Staes, Geert Aufdemkampe, Machiel van Rijn, and Luc Vanhees. Physical activity and low back pain: a systematic review of recent literature. *European Spine Journal*, 20(6):826–845, 2011.

[56] Marilyn Lord. Foot pressure measurement: a review of methodology. *Journal of Biomedical Engineering*, 3(2):91–99, 1981.

[57] Yoshimasa Sagawa Jr, Katia Turcot, Stéphane Armand, Andre Thevenon, Nicolas Vuillerme, and Eric Watelain. Biomechanics and physiological parameters during gait in lower-limb amputees: a systematic review. *Gait & posture*, 33(4):511–526, 2011.

[58] Amir Baghdadi, Lora A Cavuoto, and John L Crassidis. Hip and trunk kinematics estimation in gait through kalman filter using imu data at the ankle. *IEEE Sensors Journal*, PP(99):1–8, 2018.

[59] P Simeonov, H Hsiao, J Powers, D Ammons, T Kau, and A Amendola. Postural stability effects of random vibration at the feet of construction workers in simulated elevation. *Applied ergonomics*, 42(5):672–681, 2011.

[60] Mary M Rodgers, Kevin J McQuade, Elizabeth K Rasch, Randall E Keyser, and Margaret A Finley. Upper-limb fatigue-related joint power shifts in experienced wheelchair users and nonwheelchair users. *Journal of Rehabilitation Research and Development*, 40(1):27–38, 2003.

[61] Michael J Jorgensen, Amit Handa, Prabaharan Veluswamy, and Manish Bhatt. The effect of pallet distance on torso kinematics and low back disorder risk. *Ergonomics*, 48(8):949–963, 2005.

[62] GC David. Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders. *Occupational medicine*, 55(3):190–199, 2005.

[63] Kay Teschke, Catherine Trask, Pete Johnson, Yat Chow, Judy Village, and Mieke Koehoorn. Measuring posture for epidemiology: comparing inclinometry, observations and self-reports. *Ergonomics*, 52(9):1067–1078, 2009.

[64] Wilhelmina E Hoogendoorn, Paulien M Bongers, Henrica CW de Vet, Marjolein Douwes, Bart W Koes, Mathilde C Miedema, Geertje AM Ariëns, and Lex M Bouter. Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. *Spine*, 25(23):3087–3092, 2000.

[65] Panel on Musculoskeletal Disorders, , the Workplace, and Institute of Medicine. *Musculoskeletal disorders and the workplace: Low back and upper extremities*. National Academy Press, 2001.

[66] Sean Gallagher, William S Marras, Alan S Litsky, and Deborah Burr. Torso flexion loads and the fatigue failure of human lumbosacral motion segments. *Spine*, 30(20):2265–2273, 2005.

[67] Fu-Lin Chang, Yih-Min Sun, Kao-Hsing Chuang, and Der-Jen Hsu. Work fatigue and physiological symptoms in different occupations of high-elevation construction workers. *Applied Ergonomics*, 40(4):591–596, 2009.

[68] Zenija Roja, Valdis Kalkis, Arved Vain, Henrijs Kalkis, and Maija Eglite. Assessment of skeletal muscle fatigue of road maintenance workers based on heart rate monitoring and myotonometry. *Journal of Occupational Medicine and Toxicology*, 1(1):1–9, 2006.

[69] Ling Hui, Gabriel YF Ng, Simon SM Yeung, and Christina WY Hui-Chan. Evaluation of physiological work demands and low back neuromuscular fatigue on nurses working in geriatric wards. *Applied ergonomics*, 32(5):479–483, 2001.

[70] Tatiana de Oliveira Sato and Helenice Jane Cote Gil Coury. Evaluation of musculoskeletal health outcomes in the context of job rotation and multifunctional jobs. *Applied Ergonomics*, 40(4):707–712, 2009.

[71] A Soriano-Maldonado, L Romero, P Femia, C Roero, JR Ruiz, and A Gutierrez. A learning protocol improves the validity of the borg 6-20 rpe scale during indoor cycling. *Int J Sports Med*, 35(5):379–84, 2014.

[72] Michael J Chen, Xitao Fan, and Sondra T Moe. Criterion-related validity of the borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *Journal of sports sciences*, 20(11):873–899, 2002.

[73] Alain Groslambert and Anthony D Mahon. Perceived exertion. *Sports medicine*, 36(11):911–928, 2006.

[74] Alexandros Pantelopoulos and Nikolaos Bourbakis. A formal language approach for multi-sensor wearable health-monitoring systems. In *BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on*, pages 1–7. IEEE, 2008.

[75] Alexandros Pantelopoulos and Nikolaos Bourbakis. Design of the new prognosis wearable system-prototype for health monitoring of people at risk. *Advances in Biomedical Sensing, Measurements, Instrumentation and Systems*, pages 29–42, 2010.

[76] Linsey M Barker and Maury A Nussbaum. The effects of fatigue on performance in simulated nursing work. *Ergonomics*, 54(9):815–829, 2011.

[77] Valerie J Gawron, Jonathan French, and Doug Funke. An overview of fatigue. 2001.

[78] MC Sharpe, LC Archard, JE Banatvala, LK Borysiewicz, AW Clare, A David, RH Edwards, Keith E Hawton, HP Lambert, RJ Lane, et al. A report–chronic fatigue syndrome: guidelines for research. *Journal of the Royal Society of Medicine*, 84(2):118, 1991.

[79] Martin Björklund, Albert G Crenshaw, Mats Djupsjöbacka, and Håkan Johansson. Position sense acuity is diminished following repetitive low-intensity work to fatigue in a simulated occupational setting. *European Journal of Applied Physiology*, 81(5):361–367, 2000.

[80] Julie N Côté, Denis Raymond, Pierre A Mathieu, Anatol G Feldman, and Mindy F Levin. Differences in multi-joint kinematic patterns of repetitive hammering in healthy, fatigued and shoulder-injured individuals. *Clinical Biomechanics*, 20(6):581–590, 2005.

[81] Michiel de Looze, Tim Bosch, and Jaap van Dieën. Manifestations of shoulder fatigue in prolonged activities involving low-force contractions. *Ergonomics*, 52(4):428–437, 2009.

[82] Lora A Cavuoto and Fadel M Megahed. Understanding fatigue and the implications for worker safety. Proceedings of the ASSE Professional Development Conference & Exposition, Atlanta, GA, June 26-29, 2016.

[83] Alexandros Pantelopoulos and Nikolaos G Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(1):1–12, 2010.

[84] Tao Liu, Yoshio Inoue, and Kyoko Shibata. Development of a wearable sensor system for quantitative gait analysis. *Measurement*, 42(7):978–988, 2009.

[85] Paolo Bonato. Wearable sensors and systems. *IEEE Engineering in Medicine and Biology Magazine*, 29(3):25–36, 2010.

[86] Hyun Gu Kang, Diane F Mahoney, Helen Hoenig, Victor A Hirth, Paolo Bonato, Ihab Hajjar, and Lewis A Lipsitz. In situ monitoring of health in older adults: technologies and issues. *Journal of the American Geriatrics Society*, 58(8):1579–1586, 2010.

[87] Daria Battini, Alessandro Persona, and Fabio Sgarbossa. Innovative real-time system to integrate ergonomic evaluations into warehouse design and management. *Computers & Industrial Engineering*, 77:1–10, 2014.

[88] Tao Cheng, Giovanni C Migliaccio, Jochen Teizer, and Umberto C Gatti. Data fusion of real-time location sensing and physiological status monitoring for ergonomics analysis of construction workers. *Journal of Computing in Civil Engineering*, 27(3):320–335, 2012.

[89] Soumitry J Ray and Jochen Teizer. Real-time construction worker posture analysis for ergonomics training. *Advanced Engineering Informatics*, 26(2):439–455, 2012.

[90] Mark C Schall, Nathan B Fethke, and Howard Chen. Working postures and physical activity among registered nurses. *Applied Ergonomics*, 54:243–250, 2016.

[91] Mark C Schall, Nathan B Fethke, and Howard Chen. Evaluation of four sensor locations for physical activity assessment. *Applied Ergonomics*, 53:103–109, 2016.

[92] Umberto C Gatti, Suzanne Schneider, and Giovanni C Migliaccio. Physiological condition monitoring of construction workers. *Automation in Construction*, 44:227–233, 2014.

[93] Jose Antonio Diego-Mas and Jorge Alcaide-Marzal. Using kinect sensor in observational methods for assessing postures at work. *Applied Ergonomics*, 45(4):976–985, 2014.

[94] Robert J Aughey. Australian football player work rate: evidence of fatigue and pacing. *Int J Sports Physiol Perform*, 5(3):394–405, 2010.

[95] Catapult usa — wearable technology for elite sports. http://www.catapultsports.com/. [Online; accessed 1-June-2016].

[96] GPS unit — performance monitoring & analysis — STATSports. http://statsports.com/technology/viper-pod/. [Online; accessed 1-June-2016].

[97] Cat — fatigue risk management system — caterpillar. http://www.cat.com/en_US/support/operations/frms.html. [Online; accessed 1-June-2016].

[98] Benjamin C Amick III, Michelle M Robertson, Kelly DeRango, Lianna Bazzani, Anne Moore, Ted Rooney, and Ron Harrist. Effect of office ergonomics intervention on reducing musculoskeletal symptoms. *Spine*, 28(24):2706–2711, 2003.

[99] Catherine Cook and Robin Burgess-Limerick. The effect of forearm support on musculoskeletal discomfort during call centre work. *Applied Ergonomics*, 35(4):337–342, 2004.

[100] T Bosch, MP De Looze, and JH Van Dieen. Development of fatigue and discomfort in the upper trapezius muscle during light manual work. *Ergonomics*, 50(2):161–177, 2007.

[101] Mitsutoshi Kimura, Hirotaka Sato, Mamoru Ochi, Satoshi Hosoya, and Tsugutake Sadoyama. Electromyogram and perceived fatigue changes in the trapezius muscle during typewriting and recovery. *European Journal of Applied Physiology*, 100(1):89–96, 2007.

[102] Marcus Yung, Philip L Bigelow, Darnell M Hastings, and Richard P Wells. Detecting within-and between-day manifestations of neuromuscular fatigue at work: an exploratory study. *Ergonomics*, 57(10):1562–1573, 2014.

[103] Geri B Neuberger. Measures of fatigue: The fatigue questionnaire, fatigue severity scale, multidimensional assessment of fatigue scale, and short form-36 vitality (energy/fatigue) subscale of the short form health survey. *Arthritis Care & Research*, 49(S5):S175–S183, 2003.

[104] Gunnar Borg. *Borg's perceived exertion and pain scales.* Human kinetics, 1998.

[105] Robin Burgess-Limerick, Bruce Abernethy, and Robert J Neal. Relative phase quantifies interjoint coordination. *Journal of Biomechanics*, 26(1):91–94, 1993.

[106] Joseph F Seay, Richard EA Van Emmerik, and Joseph Hamill. Low back pain status affects pelvis-trunk coordination and variability during walking and running. *Clinical Biomechanics*, 26(6):572–578, 2011.

[107] Boyi Hu and Xiaopeng Ning. The influence of lumbar extensor muscle fatigue on lumbar–pelvic coordination during weightlifting. *Ergonomics*, 58(8):1424–1432, 2015.

[108] André LF Rodacki, Neil E Fowler, and Simon J Bennett. Multi-segment coordination: fatigue effects. *Medicine and Science in Sports and Exercise*, 33(7):1157–1167, 2001.

[109] Nicolas Forestier and Vincent Nougier. The effects of muscular fatigue on the coordination of a multijoint movement in human. *Neuroscience Letters*, 252(3):187–190, 1998.

[110] KR Lehman, JP Psihogios, and RGJ Meulenbroek. Effects of sitting versus standing and scanner type on cashiers. *Ergonomics*, 44(7):719–738, 2001.

[111] Patrick G Dempsey. A critical review of biomechanical, epidemiological, physiological and psychophysical criteria for designing manual materials handling tasks. *Ergonomics*, 41(1):73–88, 1998.

[112] L Quiros. Bnl ergonomics bulletin. *Brookhaven National Laboratory, New York*, 4, 2001.

[113] Venkatesh Balasubramanian, K Adalarasu, and Rahul Regulapati. Comparing dynamic and stationary standing postures in an assembly task. *International Journal of Industrial Ergonomics*, 39(5):649–654, 2009.

[114] Jennifer L Kelsey, Penny B Githens, Augustus A White, Theodore R Holford, Stephen D Walter, Theresa O'Connor, Adrian M Ostfeld, Ulrich Weil, Wayne O Southwick, and John A Calogero. An epidemiologic study of lifting and twisting on the job and risk for

acute prolapsed lumbar intervertebral disc. *Journal of Orthopaedic Research*, 2(1):61–66, 1984.

[115] A Magora. Investigation of the relation between low back pain and occupation. iv. physical requirements: bending, rotation, reaching and sudden maximal effort. *Scandinavian Journal of Rehabilitation Medicine*, 5(4):186, 1973.

[116] William S Marras, Steven A Lavender, Sue E Leurgans, Sudhakar L Rajulu, W Gary Allread, Fadi A Fathallah, and Sue A Ferguson. The role of dynamic three-dimensional trunk motion in occupationally-related low back disorders: The effects of workplace factors, trunk position, and trunk motion characteristics on risk of injury. *Spine*, 18(5):617–628, 1993.

[117] Douglas C Montgomery. *Introduction to statistical quality control*. John Wiley & Sons, 2014.

[118] Anil Mital, Hamid Foononi-Fard, and Max L Brown. Physical fatigue in high and very high frequency manual materials handling: perceived exertion and physiological indicators. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 36(2):219–231, 1994.

[119] LM Straker, MG Stevenson, and LT Twomey. A comparison of risk assessment of single and combination manual handling tasks: 1. maximum acceptable weight measures. *Ergonomics*, 39(1):128–140, 1996.

[120] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[121] Tim Hesterberg, Nam Hee Choi, Lukas Meier, Chris Fraley, et al. Least angle and 1 penalized regression: A review. *Statistics Surveys*, 2:61–93, 2008.

[122] Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

[123] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.

[124] Ali Dag, Kazim Topuz, Asil Oztekin, Serkan Bulur, and Fadel M Megahed. A probabilistic data-driven framework for scoring the preoperative recipient-donor heart transplant survival. *Decision Support Systems*, 86:1–12, 2016.

[125] Sotiris Kotsiantis, Dimitris Kanellopoulos, Panayiotis Pintelas, et al. Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, 30(1):25–36, 2006.

[126] Drew Dawson, Y Ian Noy, Mikko Härmä, Torbjorn Åkerstedt, and Gregory Belenky. Modelling fatigue and the use of fatigue models in work settings. *Accident Analysis & Prevention*, 43(2):549–564, 2011.

[127] Gerald P Krueger. Sustained work, fatigue, sleep loss and performance: A review of the issues. *Work & Stress*, 3(2):129–141, 1989.

[128] Peter Achermann. The two-process model of sleep regulation revisited. *Aviation, Space, and Environmental Medicine*, 75(Supplement 1):A37–A43, 2004.

[129] Alexander A Borbély. A two process model of sleep regulation. *Human Neurobiology*, 1982.

[130] Alexander A BORBÉLY and PETER Achermann. Concepts and models of sleep regulation: an overview. *Journal of Sleep Research*, 1(2):63–79, 1992.

[131] Melissa M Mallis, Sig Mejdal, Tammy T Nguyen, and David F Dinges. Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviation, Space, and Environmental Medicine*, 75(3):A4–A14, 2004.

[132] Steven R Hursh, Daniel P Redmond, Michael L Johnson, David R Thorne, Gregory Belenky, Thomas J Balkin, William F Storm, James C Miller, and Douglas R Eddy.

Fatigue models for applied research in warfighting. *Aviation, space, and environmental medicine*, 75(Supplement 1):A44–A53, 2004.

[133] Andrew J Belyavin and Michael B Spencer. Modeling performance and alertness: the qinetiq approach. *Aviation, Space, and Environmental Medicine*, 75(3):A93–A103, 2004.

[134] Megan E Jewett and Richard E Kronauer. Interactive mathematical models of subjective alertness and cognitive throughput in humans. *Journal of Biological Rhythms*, 14(6):588–597, 1999.

[135] Gregory D Roach, Adam Fletcher, and DREW DAWsON. A model to predict work-related fatigue based on hours of work. *Aviation, Space, and Environmental Medicine*, 75(Supplement 1):A61–A69, 2004.

[136] William S Marras and Richard W Schoenmarxlin. Wrist motions in industry. *Ergonomics*, 36(4):341–351, 1993.

[137] Nicky Kern, Bernt Schiele, and Albrecht Schmidt. Multi-sensor activity context detection for wearable computing. In *European Symposium on Ambient Intelligence*, pages 220–232. Springer, 2003.

[138] Susanna Pirttikangas, Kaori Fujinami, and Tatsuo Nakajima. Feature selection and activity recognition from wearable sensors. In *International Symposium on Ubiquitious Computing Systems*, pages 516–527. Springer, 2006.

[139] Tâm Huynh and Bernt Schiele. Analyzing features for activity recognition. In *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, pages 159–163. ACM, 2005.

[140] Ernst A Heinz, Kai S Kunze, Stefan Sulistyo, Holger Junker, Paul Lukowicz, and Gerhard Tröster. Experimental evaluation of variations in primary features used for accelerometric context recognition. In *European Symposium on Ambient Intelligence*, pages 252–263. Springer, 2003.

[141] Andreas Krause, Daniel P Siewiorek, Asim Smailagic, and Jonny Farringdon. Unsupervised, dynamic identification of physiological and activity context in wearable computing. In *ISWC*, volume 3, page 88, 2003.

[142] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. Activity recognition from accelerometer data. In *AAAI*, volume 5, pages 1541–1546, 2005.

[143] Seon-Woo Lee and Kenji Mase. Activity and location recognition using wearable sensors. *IEEE pervasive computing*, 1(3):24–32, 2002.

[144] Yuji Ohgi. Microcomputer-based acceleration sensor device for sports biomechanics-stroke evaluation by using swimmer's wrist acceleration. In *Sensors, 2002. Proceedings of IEEE*, volume 1, pages 699–704. IEEE, 2002.

[145] Y Ohgi, M Yasumura, H Ichikawa, and C Miyaji. Analysis of stroke technique using acceleration sensor ic in freestyle swimming. *The Engineering of Sport*, pages 503–512, 2000.

[146] Torbjörn Åkerstedt, Simon Folkard, and Christian Portin. Predictions from the three-process model of alertness. *Aviation, Space, and Environmental Medicine*, 75(3):A75–A83, 2004.

[147] Philippa Gander, Laurence Hartley, David Powell, Philippe Cabon, Edward Hitchcock, Ann Mills, and Stephen Popkin. Fatigue risk management: Organizational factors at the regulatory and industry/company level. *Accident Analysis & Prevention*, 43(2):573–590, 2011.

[148] Atsuo Murata. Proposal of a method to predict subjective rating on drowsiness using physiological and behavioral measures. *IIE Transactions on Occupational Ergonomics and Human Factors*, pages 1–13, 2016.

[149] Lee Di Milia, Michael H Smolensky, Giovanni Costa, Heidi D Howarth, Maurice M Ohayon, and Pierre Philip. Demographic factors, fatigue, and driving accidents: An

examination of the published literature. *Accident Analysis & Prevention*, 43(2):516–532, 2011.

[150] Advanced Manufacturing: A Snapshot of Priority Technology Areas Across the Federal Government, Product of the Subcomittee for Advanced Manufacturing OF THE National Science and Technology Council, Executive Office of the President of the United States. https://www.whitehouse.gov/sites/whitehouse.gov/files/images/Blog/NSTC\%20SAM\%20technology\%20areas\%20snapshot.pdf, 2016. Online; last accessed 6-April-2018.

[151] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zhlke. Human-machine-interaction in the industry 4.0 era. In *The 12th IEEE International Conference on Industrial Informatics (INDIN)*, pages 289–294, July 2014.

[152] W. Yakowicz. Robots are coming for dirty, dangerous, and dull jobs. Inc., http://www.inc.com/will-yakowicz/robots-replacing-dirty-dangerous-jobs-is-good.html, 2016. Online; accessed 6-April-2018.

[153] Aicha Ferjani, Achraf Ammar, Henri Pierreval, and Sabeur Elkosantini. A simulation-optimization based heuristic for the online assignment of multi-skilled workers affected by fatigue in manufacturing systems. *Computers & Industrial Engineering*, 112:663–674, 2017.

[154] N.E. Waldeck. *Advanced Manufacturing Technologies and Workforce Development*. Advanced Manufacturing Technologies and Workforce Development. Taylor & Francis, 2014.

[155] Report to the President on Ensuring American Leadership in Advanced Manufacturing. President's Council of Advisors on Science and Technology. Executive Office of the President of the United States. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-advanced-manufacturing-june2011.pdf, 2011. Online; last accessed 6-April-2018.

[156] Francisco Brocal and Miguel A Sebastián. Identification and analysis of advanced manufacturing processes susceptible of generating new and emerging occupational risks. *Procedia Engineering*, 132:887–894, 2015.

[157] David Romero, Peter Bernus, Ovidiu Noran, Johan Stahre, and Åsa Fast-Berglund. The operator 4.0: Human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems. In *APMS (Advances in Production Management Systems)*, 2016.

[158] Peter Gust, Ulf Müller, Nico Feller, and Michael Schiffmann. *Field Study on the Application of a Simulation-Based Software Tool for the Strain-Based Staffing in Industrial Manufacturing*, pages 3–12. Springer International Publishing, Cham, 2017.

[159] Xiqian Cai, Jie Gong, Yi Lu, and Songfa Zhong. Recover overnight? work interruption and worker productivity. *Management Science*, Preprint, 2018.

[160] Leo H Chiang, Evan L Russell, and Richard D Braatz. *Fault detection and diagnosis in industrial systems*. Springer-Verlag, 2000.

[161] Mark C Schall, Richard F Sesek, and Lora A Cavuoto. Barriers to the adoption of wearable sensors in the workplace: A survey of occupational safety and health professionals. *Human Factors*, page Online First, 2018.

[162] Andrea Ceschi, Evangelia Demerouti, Riccardo Sartori, and Joshua Weller. Decision-making processes in the workplace: How exhaustion, lack of resources and job demands impair them and affect performance. *Frontiers in psychology*, 8:313, 2017.

[163] Gerold R Ebenbichler, Paolo Bonato, Serge H Roy, Stefan Lehr, Martin Posch, Josef Kollmitzer, and U Croce Della. Reliability of emg time-frequency measures of fatigue during repetitive lifting. *Medicine and science in sports and exercise*, 34(8):1316–1323, 2002.

[164] Daniel HK Chow, Joseph WK Man, Andrew D Holmes, and John H Evans. Postural and trunk muscle response to sudden release during stoop lifting tasks before and after fatigue of the trunk erector muscles. *Ergonomics*, 47(6):607–624, 2004.

[165] Bradley S Davidson, Michael L Madigan, and Maury A Nussbaum. Effects of lumbar extensor fatigue and fatigue rate on postural sway. *European journal of applied physiology*, 93(1-2):183–189, 2004.

[166] Kohzoh Yoshino, Tomoko Motoshige, Tsutomu Araki, and Katsunori Matsuoka. Effect of prolonged free-walking fatigue on gait and physiological rhythm. *Journal of biomechanics*, 37(8):1271–1280, 2004.

[167] Minhyung Lee, Michael Roan, Benjamin Smith, and Thurmon E Lockhart. Gait analysis to classify external load conditions using linear discriminant analysis. *Human movement science*, 28(2):226–235, 2009.

[168] EB Fontes, BPC Smirmaul, FY Nakamura, G Pereira, AH Okano, LR Altimari, JL Dantas, AC de Moraes, et al. The relationship between rating of perceived exertion and muscle activity during exhaustive constant-load cycling. *International journal of sports medicine*, 2010.

[169] Michelle Karg, Gentiane Venture, Jesse Hoey, and Dana Kulic. Human movement analysis as a measure for fatigue: A hidden markov-based approach. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):470–481, 2014.

[170] Jian Zhang, Thurmon E Lockhart, and Rahul Soangra. Classifying lower extremity muscle fatigue during walking using machine learning and inertial sensors. *Annals of biomedical engineering*, 42(3):600–612, 2014.

[171] Jorunn L Helbostad, Sara Leirfall, Rolf Moe-Nilssen, and Olav Sletvold. Physical fatigue affects gait characteristics in older persons. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(9):1010–1015, 2007.

[172] Justin J Kavanagh, Steven Morrison, and Rod S Barrett. Lumbar and cervical erector spinae fatigue elicit compensatory postural responses to assist in maintaining head stability during walking. *Journal of Applied Physiology*, 101(4):1118–1126, 2006.

[173] Michelle Karg, Kolja Kühnlenz, Martin Buss, Wolfgang Seiberl, Ferdinand Tusker, Maren Schmeelk, and Ansgar Schwirtz. Expression and automatic recognition of exhaustion in natural walking. In Katherine Blashki, editor, *International Conference Interfaces and Human Computer Interaction*, pages 165–172, 2008.

[174] Zahra Sedighi Maman, Mohammad Ali Alamdar Yazdi, Lora A Cavuoto, and Fadel M Megahed. A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. *Applied Ergonomics*, 65:515–529, 2017.

[175] Alec Levenson. Workplace fatigue is a systems problem. *Consulting Psychology Journal: Practice and Research*, 69(2):130, 2017.

[176] Nikola Marangunić and Andrina Granić. Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, 14(1):81–95, 2015.

[177] Lora Cavuoto and Fadel Megahed. Understanding fatigue: Implications for worker safety. *Professional Safety*, (December):16–19, 2017.

[178] Catapult Sports. World Leader in Sports Technology for Elite Sports. http://www.catapultsports.com/, 2018. Online; accessed 1-June-2018.

[179] Zahra Sedighi Maman, Amir Baghdadi, Fadel Megahed, and Lora Cavuoto. Monitoring and change point estimation of normal (in-control) and fatigued (out-of-control) state in workers. In *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages V003T11A011–V003T11A011. American Society of Mechanical Engineers, 2016.

[180] Jane A Kent-Braun, Alexander V Ng, Julie W Doyle, and Theodore F Towse. Human skeletal muscle responses vary with age and gender during fatigue due to incremental isometric exercise. *Journal of Applied Physiology*, 93(5):1813–1823, 2002.

[181] Laura A Wojcik, Maury A Nussbaum, Dingding Lin, Peggy A Shibata, and Michael L Madigan. Age and gender moderate the effects of localized muscle fatigue on lower extremity joint torques used during quiet stance. *Human movement science*, 30(3):574–583, 2011.

[182] Amir Baghdadi, Zahra S Maman, Lin Lu, Lora A Cavuoto, and Fadel M Megahed. Effects of task type, task duration, and age on body kinematics and subjective fatigue. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 61, pages 1040–1040. SAGE Publications Sage CA: Los Angeles, CA, 2017.

[183] Ling Bao. *Physical activity recognition from acceleration data under semi-naturalistic conditions*. PhD thesis, Massachusetts Institute of Technology, 2003.

[184] CARL Foster. Monitoring training in athletes with reference to overtraining syndrome. *Medicine and science in sports and exercise*, 30:1164–1168, 1998.

[185] L Quagliarella, N Sasanelli, and G Belgiovine. An interactive fall and loss of consciousness detector system. *Gait & posture*, 28(4):699–702, 2008.

[186] Justin G Young, Matthieu B Trudeau, Dan Odell, Kim Marinelli, and Jack T Dennerlein. Wrist and shoulder posture and muscle activity during touch-screen tablet use: effects of usage configuration, tablet type, and interacting hand. *Work*, 45(1):59–71, 2013.

[187] Marc Bächlin, Kilian Förster, and Gerhard Tröster. Swimmaster: a wearable assistant for swimmer. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 215–224. ACM, 2009.

[188] Ann Hallemans, Sofie Beccu, Kelly Van Loock, Els Ortibus, Steven Truijen, and Peter Aerts. Visual deprivation leads to gait adaptations that are age-and context-specific: I. step-time parameters. *Gait & posture*, 30(1):55–59, 2009.

[189] Christian Larivière, Denis Gagnon, and Patrick Loisel. The comparison of trunk muscles emg activation between subjects with and without chronic low back pain during flexion–extension and lateral bending tasks. *Journal of electromyography and kinesiology*, 10(2):79–91, 2000.

[190] John D Willson and Thomas W Kernozek. Plantar loading and cadence alterations with fatigue. *Medicine and science in sports and exercise*, 31(12):1828–1833, 1999.

[191] Xiaoping Yun, Eric R Bachmann, Hyatt Moore, and James Calusdian. Self-contained position tracking of human movement using small inertial/magnetic sensor modules. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 2526–2533. IEEE, 2007.

[192] Galit Shmueli and Inbal Yahav. The forest or the trees? tackling simpson's paradox with classification trees. *Production and Operations Management*, 27(4):696–716, 2018.

[193] Avrim L Blum and Pat Langley. Selection of relevant features and examples in machine learning. *Artificial intelligence*, 97(1):245–271, 1997.

[194] Qi Feng and J George Shanthikumar. How research in production and operations management may evolve in the era of big data. *Production and Operations Management*, 2017.

[195] Bradley Efron and Robert J Tibshirani. An introduction to the bootstrap: Monographs on statistics and applied probability, vol. 57. *New York and London: Chapman and Hall/CRC*, 1993.

[196] Nilay Tanık Argon and Serhan Ziya. Priority assignment under imperfect information on customer type identities. *Manufacturing & Service Operations Management*, 11(4):674–693, 2009.

[197] Fredrik Ødegaard and Pontus Roos. Measuring the contribution of workers' health and psychosocial work-environment on production efficiency. *Production and Operations Management*, 23(12):2191–2208, 2014.

[198] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160:3–24, 2007.

[199] Joris M Keizers, J Will M Bertrand, and Jaap Wessels. Diagnosing order planning performance at a navy maintenance and repair organization, using logistic regression. *Production and Operations Management*, 12(4):445–463, 2003.

[200] Alan S Abrahams, Weiguo Fan, G Alan Wang, Zhongju John Zhang, and Jian Jiao. An integrated text analytic framework for product defect discovery. *Production and Operations Management*, 24(6):975–990, 2015.

[201] Gary King and Langche Zeng. Logistic regression in rare events data. *Political analysis*, 9(2):137–163, 2001.

[202] Aleda Roth, Jaya Singhal, Kalyan Singhal, and Christopher S Tang. Knowledge creation and dissemination in operations and supply chain management. *Production and Operations Management*, 25(9):1473–1488, 2016.

[203] Jonathan E Helm, Adel Alaeddini, Jon M Stauffer, Kurt M Bretthauer, and Ted A Skolarus. Reducing hospital readmissions by integrating empirical prediction with resource optimization. *Production and Operations Management*, 25(2):233–257, 2016.

[204] Ruomeng Cui, Santiago Gallino, Antonio Moreno, and Dennis J Zhang. The operational value of social media information. *Production and Operations Management*, 2017.

[205] Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer, 2000.

[206] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[207] Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

[208] Robert E Schapire. The boosting approach to machine learning: An overview. In *Nonlinear estimation and classification*, pages 149–171. Springer, 2003.

[209] Jiawei Han, Micheline Kamber, and Jian Pei. *Data mining: concepts and techniques*. Morgan Kaufmann Publishers, 2011.

[210] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[211] Jelle Vries, René Koster, and Daan Stam. Safety does not happen by accident: antecedents to a safer warehouse. *Production and operations management*, 25(8):1377–1390, 2016.

[212] Fred A Manuele. Risk assessment & hierarchies of control. *Professional Safety*, 50(5):33–39, 2005.

[213] Scott Thomas, Jeff Reading, and Roy J Shephard. Revision of the physical activity readiness questionnaire (par-q). *Canadian journal of sport sciences*, 17(4):338–345, 1992.

[214] Nicholas D Pattengale, Masoud Alipour, Olaf RP Bininda-Emonds, Bernard ME Moret, and Alexandros Stamatakis. How many bootstrap replicates are necessary? In *Annual International Conference on Research in Computational Molecular Biology*, pages 184–200. Springer, 2009.

[215] Anil Mital. *Guide to manual materials handling*. CRC Press, 2017.

[216] Kevin P Granata, Greg P Slota, and Sara E Wilson. Influence of fatigue in neuromuscular control of spinal stability. *Human factors*, 46(1):81–91, 2004.

[217] Fadi A Fathallah, William S Marras, and Mohamad Parnianpour. An assessment of complex spinal loads during dynamic lifting tasks. *Spine*, 23(6):706–716, 1998.

[218] Laurent Schmitt, Jacques Regnard, Maxime Desmarets, Fréderic Mauny, Laurent Mourot, Jean-Pierre Fouillot, Nicolas Coulmy, and Grégoire Millet. Fatigue shifts and scatters heart rate variability in elite endurance athletes. *PloS one*, 8(8):e71588, 2013.

[219] Walter M Fletcher and F Gowland Hopkins. Lactic acid in amphibian muscle 1. *The Journal of physiology*, 35(4):247–309, 1907.

[220] AV Hill and Hartley Lupton. Muscular exercise, lactic acid, and the supply and utilization of oxygen. *QJM: Quarterly Journal of Medicine*, (62):135–171, 1923.

[221] Vincent Pichot, Frederic Roche, Jean-michel Gaspoz, Franck Enjolras, Anestis Antoniadis, Pascal Minini, Frederic Costes, Thierry Busso, Jean-renÉ Lacour, and Jean Claude BarthÉlÉmy. Relation between heart rate variability and training load in middle-distance runners. *Medicine and science in sports and exercise*, 32(10):1729–1736, 2000.

[222] Conrad P Earnest, R Jurca, TSe Church, JL Chicharro, J Hoyos, and A Lucia. Relation between physical exertion and heart rate variability characteristics in professional cyclists during the tour of spain. *British journal of sports medicine*, 38(5):568–575, 2004.

[223] Miteshkumar Patel, Sara KL Lal, Diarmuid Kavanagh, and Peter Rossiter. Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert systems with Applications*, 38(6):7235–7242, 2011.

[224] Ranjana K Mehta and Lora A Cavuoto. The effects of obesity, age, and relative workload levels on handgrip endurance. *Applied ergonomics*, 46:91–95, 2015.

[225] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300, 1995.