ARTICLE IN PRESS

Applied Ergonomics xxx (2017) 1–15

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

A data-driven approach to modeling physical fatigue in the workplace using wearable sensors

Zahra Sedighi Maman [a], Mohammad Ali Alamdar Yazdi [a], Lora A. Cavuoto [b], Fadel M. Megahed [c, *]

[a] Department of Industrial and Systems Engineering, Auburn University, AL 36849, USA
[b] Department of Industrial and Systems Engineering, University at Buffalo, NY 14260, USA
[c] Farmer School of Business, Miami University, OH 45056, USA

ARTICLE INFO

Article history:
Received 4 August 2016
Received in revised form
28 January 2017
Accepted 1 February 2017
Available online xxx

Keywords:
Analytics
Feature selection
Penalized regression
Physical fatigue

ABSTRACT

Wearable sensors are currently being used to manage fatigue in professional athletics, transportation and mining industries. In manufacturing, *physical fatigue* is a challenging ergonomic/safety “issue” since it lowers productivity and increases the incidence of accidents. Therefore, physical fatigue must be managed. There are two main goals for this study. First, we examine the use of wearable sensors to detect physical fatigue occurrence in simulated manufacturing tasks. The second goal is to estimate the physical fatigue level over time. In order to achieve these goals, sensory data were recorded for eight healthy participants. *Penalized logistic* and *multiple linear regression* models were used for physical fatigue detection and level estimation, respectively. Important features from the five sensors locations were selected using *Least Absolute Shrinkage and Selection Operator (LASSO)*, a popular variable selection methodology. The results show that the LASSO model performed well for both physical fatigue detection and modeling. The modeling approach is not participant and/or workload regime specific and thus can be adopted for other applications.

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 (Barker and Nussbaum, 2011; Yung, 2016). 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 (Gawron et al., ; Sharpe et al., 1991). 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 (Björklund et al., 2000; Côté et al., 2005; Huysmans et al., 2010). These effects might lead to reduced performance, lowered productivity, deficits in work quality, and increased incidence of accidents and human errors (Yung, 2016; Kumar, 2001; Visser and van Dieën, 2006; Looze et al., 2009). Physical fatigue can also result in longer term adverse health outcomes, including, e.g., *chronic fatigue syndrome* (Yung, 2016) and *reduced immune function* (Kajimoto, 2008). “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” (Yung, 2016).

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 (Williamson et al., 2011). The specific precursor(s) for physical fatigue and/or injury development often goes unidentified (Cavuoto and Megahed, 2016). However, researchers have postulated that through delineation of the quantitative details of relevant variables, appropriate interventions and injury control can be developed (Kumar, 2001). How to best

* Corresponding author.
E-mail addresses: zzs0016@auburn.edu (Z. Sedighi Maman), mza0052@auburn.edu (M.A. Alamdar Yazdi), loracavu@bu_alo.edu (L.A. Cavuoto), fmegahed@miamioh.edu (F.M. Megahed).

http://dx.doi.org/10.1016/j.apergo.2017.02.001

Please cite this article in press as: Sedighi Maman, Z., et al., A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied Ergonomics (2017), http://dx.doi.org/10.1016/j.apergo.2017.02.001

quantify workplace conditions, particularly physical exposures experienced by the worker, remains an open research question (Kim and Nussbaum, 2013). 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 (Gallagher and Heberger, 2013; Garg and Kapellusch, 2009; Marras et al., 2009). 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 (Cavuoto and Nussbaum, 2014). 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)) (Balkin et al., 2011) or changes in local muscle activation (using electromyography (EMG)) (Dong et al., 2014).

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 (Garg and Kapellusch, 2009; Marras et al., 2009). The increasing availability of pervasive sensing technologies, including wearable devices (Kim and Nussbaum, 2013; Poh et al., 2010; Nyan et al., 2008; Pantelopoulos and Bourbakis, 2010), 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* (Kim and Nussbaum, 2013; Waters et al., 1993; Vignais et al., 2013). 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 (Liu et al., 2009). Since the late 1990s, there have been large advances in the field of wearable technology (Bonato, 2010). Currently, wearable sensors are lower cost, easy to use and have minimal interference with the wearer (Liu et al., 2009). They have become the predominant devices for monitoring mobility and physical activity (Kang et al., 2010).

Evidence suggests the translation of these technologies to a work environment would be achievable and a potentially valuable addition for injury prevention (Kim and Nussbaum, 2013; Vignais et al., 2013; Battini et al., 2014; Cheng et al., 2012; Schall et al., 2015). While there has been an emphasis on research to develop occupational health and safety sensor systems and establish their potential in a lab environment (Kim and Nussbaum, 2013; Schall et al., 2015; Ray and Teizer, 2012), most current workplace applications have been limited to: a) posture analysis (Schall et al., 2015, 2016a, 2016b), b) task classification (Kim and Nussbaum, 2013), c) basic physiological monitoring (Cheng et al., 2012; Gatti et al., 2014), d) computerized application of traditional observational tools (Diego-Mas and Alcaide-Marzal, 2014), and e) specialized (industry-specific) physical fatigue detection/management systems. The specialized physical fatigue applications are limited to the following three domains (Cavuoto and Megahed, 2016):

- athletics, where the focus is primarily on monitoring athletes' performances (e.g., the *Catapult System* (Aughey, 2010; Catapult usa,), the *Viper Pod* by STATSports® (GPS unit,), etc.);
- sleep-induced fatigue, in mining (CAT's *Fatigue Risk Management System* (Cat,)); and
- driver drowsiness detection systems in transportation (see (Cavuoto and Megahed, 2016) 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 (Cavuoto and Megahed, 2016). 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 (Pantelopoulos and Bourbakis, 2010). 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 (Balkin et al., 2011). 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 paper 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 paper 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 paper, 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 paper is organized as follows. Section 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 3. We provide our results and discuss their ergonomic/safety implications in Section 4. We offer our conclusions and our opinions about future research directions in Section 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. 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 (Yung, 2016). 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., (Amick et al., 2003; Cook and Burgess-Limerick, 2004; Bosch et al., 2007; Kimura et al., 2007; Yung et al., 2014)). Generally speaking, increased discomfort is positively related to physical fatigue and reduced work capacity (de Oliveira Sato and Coury, 2009). In the literature, there are several different questionnaires and rating scales that are used for measuring physical fatigue. For a detailed review, please see (Yung, 2016; Neuberger, 2003). In our analysis, we use the *Borg Rating of Perceived Exertion* (Borg, 1998) 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) (Roja et al., 2006; Chang et al., 2009), *force variability* (Chang et al., 2009), *tremor* (see references within (Yung, 2016)), *changes in posture/gait* (Roja et al., 2006; Chang et al., 2009), and the *multi-joint coordination between different segments* (i.e. hip, knee, ankle, pelvis, trunk) (Burgess-Limerick et al., 1993; Seay et al., 2011; Hu and Ning, 2015). 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 (Rodacki et al., 2001). Therefore, with physical fatigue, movement coordination can change so as to keep up motor performance in terms of movement accuracy (Forestier and Nougier, 1998). 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 (Rodacki et al., 2001). 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 1, represent a range of physiological and movement parameters that can capture physical fatigue and stress from physical tasks. In Table 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 (Yung et al., 2014). An innovative aspect of this paper 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 3.

3. Methods

Our approach, depicted in Fig. 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.

3.1. Data collection

3.1.1. 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 participants were in good health. In Table 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 3.3 can be applied when n (# of participants) is small.

3.1.2. Equipment

Each participant was instrumented with four inertial measurement units IMUs (see Fig. 2) while performing the task. Each IMU

Table 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

Please cite this article in press as: Sedighi Maman, Z., et al., A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied Ergonomics (2017), http://dx.doi.org/10.1016/j.apergo.2017.02.001

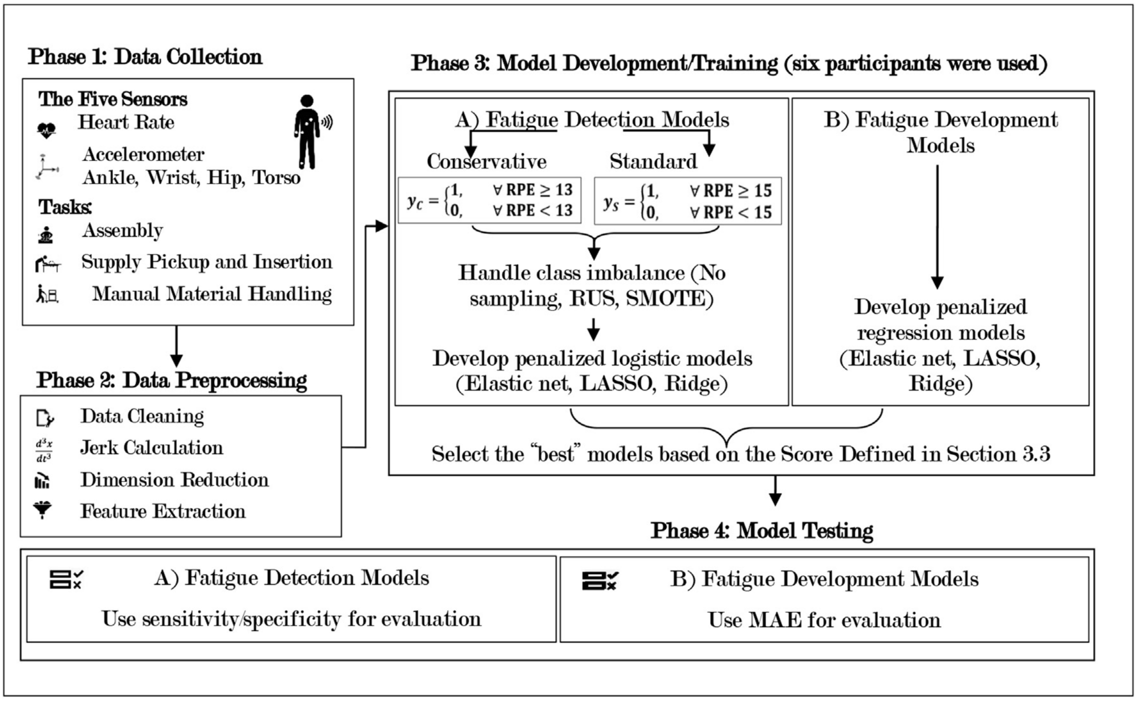


Fig. 1. An overview of the proposed method.

Table 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

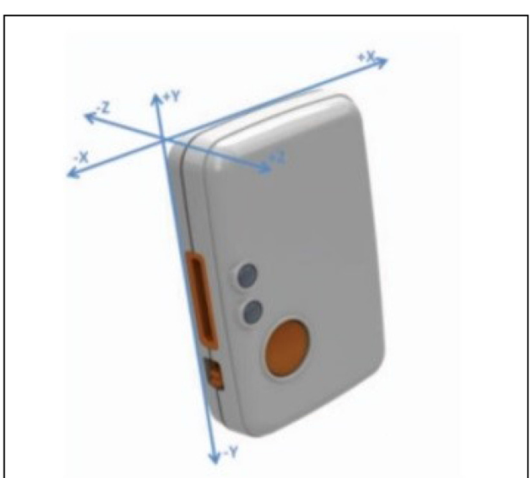


Fig. 2. A Shimmer3 device with its reference coordinate system.

was a Shimmer3 (Shimmer, Dublin, Ireland, www.shimmersensing.com), which is small-sized (51 mm × 34 mm x 14 mm), low-power using, 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. Fig. 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 1.

3.1.3. 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 psycho-motor 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 reason behind selecting this posture. Firstly, standing in this task is a widely adopted industrial working posture (Lehman et al., 2001). 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 (Dempsey, 1998). 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 (Quiros,). Also, usually the machine operators and assembly line workers required in such tasks characterized by extended standing reported these discomforts (Balasubramanian et al., 2009).

(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 (Bonato et al., 2003). Studies have shown that the risk of low back disorders significantly increases by repeated bending and lifting activities (Kelsey et al., 1984; Magora, 1973; Marras et al., 1993). The tasks cycle time was 2 min, 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 (Lu et al.,). This task simulated warehousing operations by picking cartons (whose weight is 10*kg*, 18*kg*, or 26*kg*), 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 min. 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 Fig. 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 (Atallah et al., 2010). In addition, determination of torso inclination during the task required the sensors on the chest (Bao and Intille, 2004). 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.

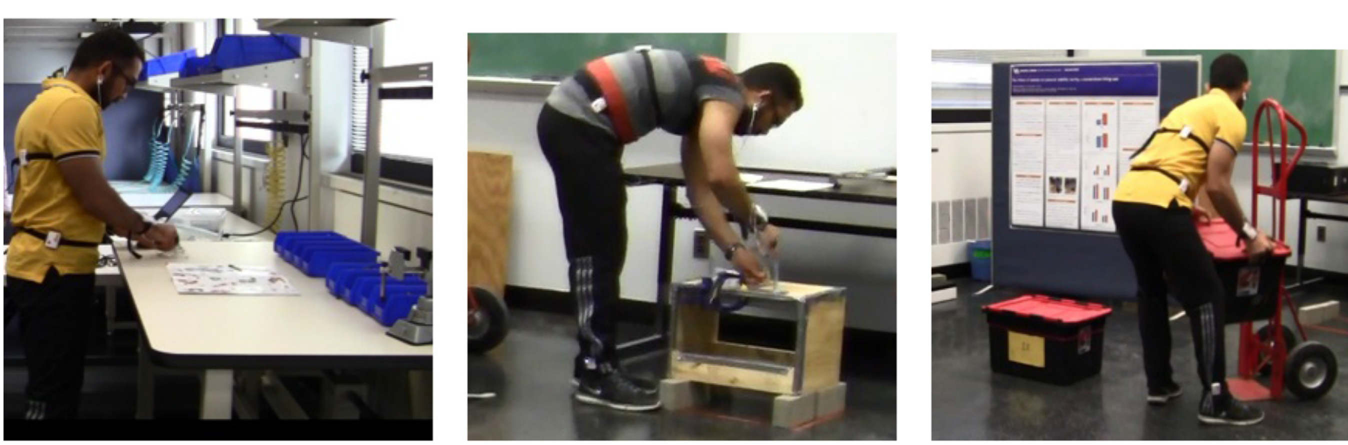
The participants were given target performance levels for each task. For MMH, participants were asked to palletize 16 cartons in three different orders, and each order was repeated twice. In the SPI task, a cycle time of 2 min was given and for PA task there was a cycle time of 15 min for each subassembly. 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 (Borg, 1998) every 10 min. This was used to validate physical fatigue development.

3.2. Data preprocessing

3.2.1. 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



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

Fig. 3. Sensor placement on a subjects.

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.

3.2.2. 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 (Catapult usa,).

3.2.3. 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-min 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 $(HR_{avg} - RHR)/(HR_{max} - RHR)$).

3.2.4. 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 3), computed at 10-min 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 (Montgomery, 2014) 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 min window (as observed from the acceleration data), see Table 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-min window in the experiment. This feature set is computed for acceleration only. Table 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.

3.3. Model development

A primary objective of this paper 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):

$$y_C = \begin{cases} 1, & \forall \text{ RPE } \geq 13 \\ 0, & \forall \text{ RPE } < 13 \end{cases}, \tag{1}$$

and

$$y_S = \begin{cases} 1, & \forall \text{ RPE } \geq 15 \\ 0, & \forall \text{ RPE } < 15 \end{cases}. \tag{2}$$

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 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 (Mital et al., 1994; Straker et al., 1996)).

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.

Table 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 min and previous time window as baseline

Please cite this article in press as: Sedighi Maman, Z., et al., A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied Ergonomics (2017), http://dx.doi.org/10.1016/j.apergo.2017.02.001

Table 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%

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:

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. 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 (Tibshirani, 1996; Hesterberg et al., 2008) 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).

3.3.1. 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) and (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:

$$\max \quad l(\beta) - \lambda\left[\alpha \sum_{j=1}^{m}|\beta_j| + \frac{1}{2}(1-\alpha)\sum_{j=1}^{m}\beta_j^2\right], \tag{3}$$

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* (Tibshirani, 1996), *ridge regression* (Hoerl and Kennard, 1970), *elastic-net* (Zou and Hastie, 2005) for the two different binary outcomes. For a general discussion on these approaches, the reader is referred to (Hesterberg et al., 2008).

3.3.2. 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:

$$\widehat{y}_{Borg} = \widehat{RPE} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \varepsilon, \tag{4}$$

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:

$$\min \quad \left(RPE - \widehat{RPE}\right)^T\left(RPE - \widehat{RPE}\right) + \lambda\left[\alpha\sum_{j=1}^{m}|\beta_j| + \frac{1}{2}(1-\alpha)\sum_{j=1}^{m}\beta_j^2\right], \tag{5}$$

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 4, we examine which method is more suitable for modeling physical fatigue development based on our data.

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. 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 paper, 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. $n_{fatigued} \ll n_{non-fatigued}$). 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:

$$S_k = Fraction\ Deviance * TPM * \frac{m}{r_k}, \tag{6}$$

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. (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 $n_{fatigued} \ll n_{non-fatigued}$. 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 (Dag et al., 2016; Kotsiantis et al., 2006). 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 (Dag et al., 2016). Re-sampling is typically applied by either over-sampling the minority class and/or under-sampling the majority class (Kotsiantis et al., 2006). In this paper, 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 (Dag et al., 2016; Kotsiantis et al., 2006) for details).

Once the "best" models are selected from the *training* stage (based on maximizing S_k), it is important to evaluate how these model performs 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. 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 4.

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.

4.1. Selected models for physical fatigue detection

As discussed earlier in Section 3.3.1, 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 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 5:

Performance using standard/conservative scenarios: It is obvious from Table 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.

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*.

4.1.1. Predictive features for the standard physical fatigue detection scenario

Table 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

Table 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

Please cite this article in press as: Sedighi Maman, Z., et al., A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied Ergonomics (2017), http://dx.doi.org/10.1016/j.apergo.2017.02.001

Table 6
Selected features for the logistic LASSO model with RUS sampling (standard scenario).

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

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 6 is provided in Fig. 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*.

4.1.2. Predictive features for the conservative physical fatigue detection scenario

Similar to the standard approach, the 22 selected features and

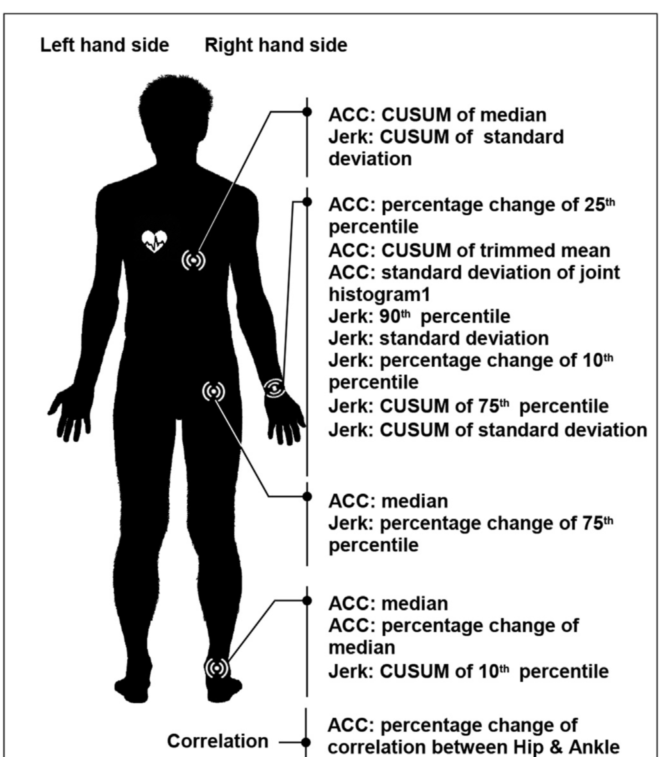


Fig. 4. The location of selected features on body for the standard fatigue detection model.

Table 7
Selected features for the logistic LASSO model with RUS sampling (conservative scenario).

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

their corresponding coefficients are shown in Table 7. Fig. 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.

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 8 and visualized in Fig. 6.

Table 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 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.

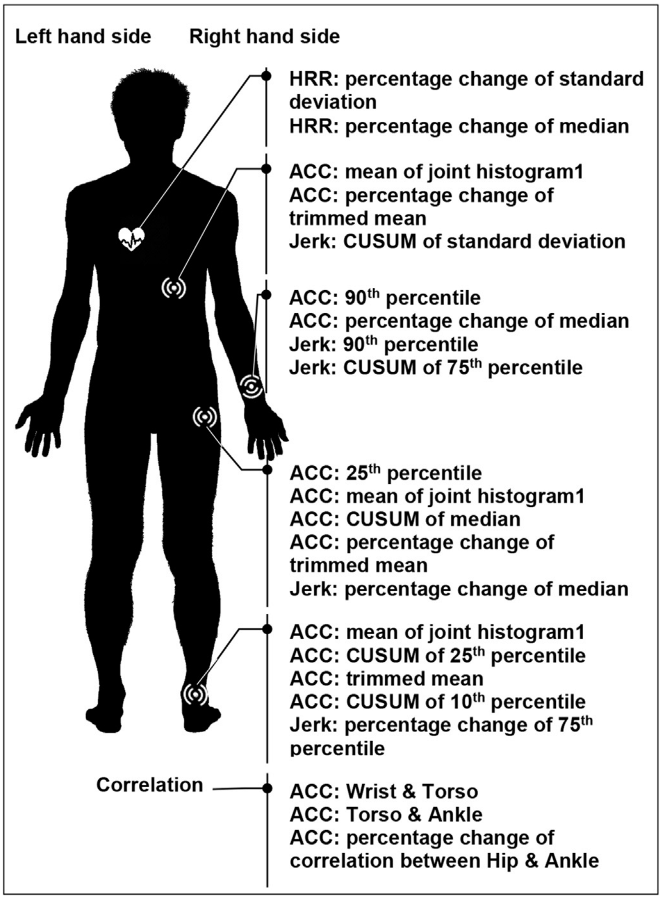


Fig. 5. The location of selected features on body for the conservative physical fatigue detection model.

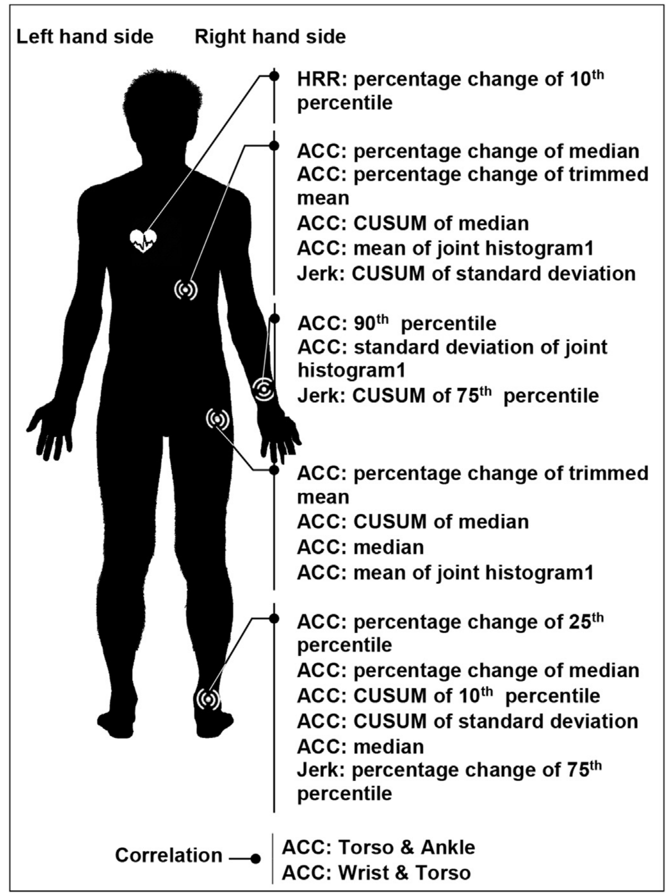


Fig. 6. The location of selected features on body for the physical fatigue development model.

Table 8
Selected features for the LASSO model with RUS sampling.

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

4.3. Testing of the models

4.3.1. Implementation of selected physical fatigue detection model on two test participants

The selected models from section 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 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 9 is that the *standard LASSO model* detected all *physically fatigued* states for the two test participants since its *sensitivity* = 1.

4.3.2. 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 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).

Fig. 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.

Table 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**

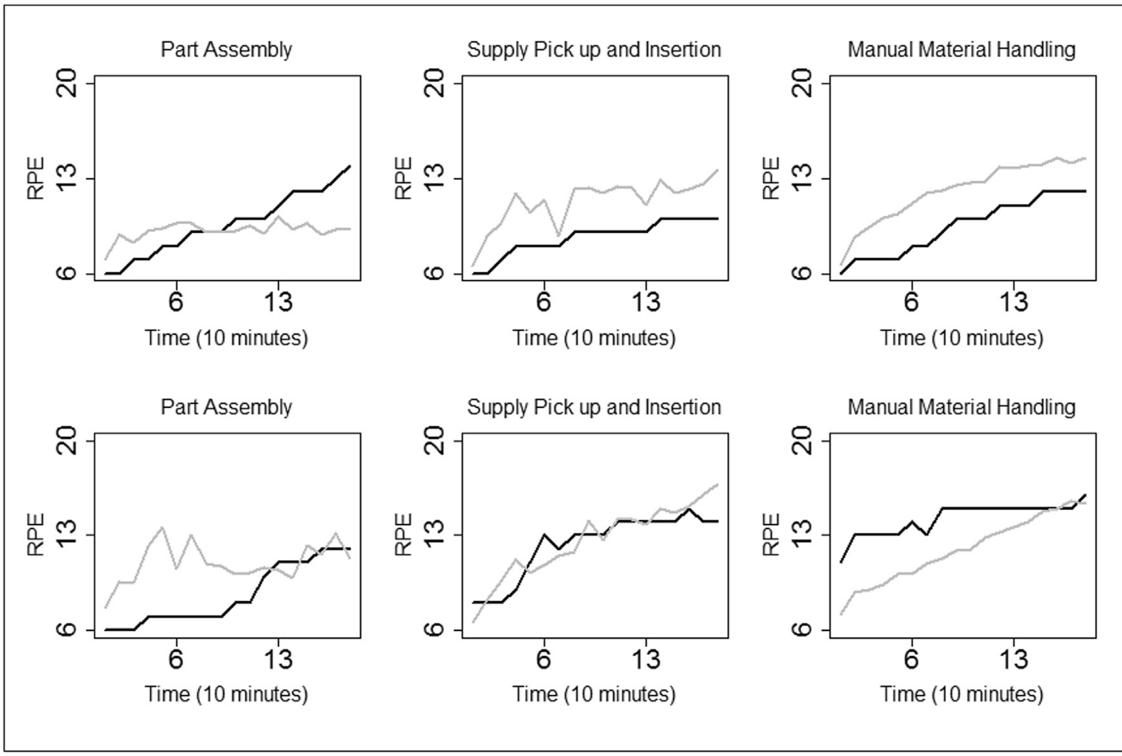


Fig. 7. Predicted RPE (Gray) vs. Actual (Black) - *P7* (top) and *P8* (bottom).

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.

5. Discussion and conclusions

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 (Dawson et al., 2011). 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 (Williamson et al., 2011; Dawson et al., 2011; Krueger, 1989; Achermann, 2004; Borbély, ; Borbély and Achermann, 1992; Mallis et al., 2004; Hursh et al., 2004; Belyavin and Spencer, 2004; Jewett and Kronauer, 1999; Roach et al., 2004). 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 (Dawson et al., 2011). 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 (Williamson et al., 2011).

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 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

Please cite this article in press as: Sedighi Maman, Z., et al., A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied Ergonomics (2017), http://dx.doi.org/10.1016/j.apergo.2017.02.001

acceleration signal are the mean (Bao and Intille, 2004; Marras and Schoenmarxlin, 1993; Kern et al., 2003; Pirttikangas et al., 2006; Huynh and Schiele, 2005; Heinz et al., 2003; Krause et al., 2003; Ravi et al., 2005), variance or standard deviation (Kern et al., 2003; Pirttikangas et al., 2006; Heinz et al., 2003; Ravi et al., 2005; Lee and Mase, 2002), and the correlation between acceleration axes (Bao and Intille, 2004; Ravi et al., 2005). 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 10. For the three simulated manufacturing tasks, upper extremity movement was a main component of the task. 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 (Bao and Intille, 2004; Marras and Schoenmarxlin, 1993; Kern et al., 2003). The reader should note that we placed the sensor on the participant's right wrist, irrespective of handedness (P4 was left handed). Our hypothesis was based on the fact that our tasks required equal loading for both hands. Thus, this placement allows for consistent relationships/angles between the different sensors (across participants). To quantify the effect of handedness, we reran our analyses by removing P4 from the training set. The resulting models for fatigue detection and fatigue development were similar in terms of the features selected and their performance in training/testing. Thus, we concluded that for our tasks our choice of sensor location is appropriate. Note that we did not include these results here for conciseness; however, we make them available on Github (see Supplementary Materials).

To utilize our 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 (Ohgi, 2002; Ohgi et al., 2000). Industrial workers who need to move their hands and wrist repeatedly and/or powerfully are at greater risk for cumulative trauma disorders (CTD) (Marras and Schoenmarxlin, 1993). 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 (Dawson et al., 2011). 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 models. Its estimation for physical fatigue state (see standard model) 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 11. The features with positive coefficients included the following:

Table 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

Table 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:** The median value for hip acceleration for this time window.
- **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.

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

- **Wrist Jerk CUSUM of 75th Percentile:** See the above explanation of ankle jerk percentage change for an interpretation of this feature.
- **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 (Pirttikangas et al., 2006).

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 (Borbély, ; Borbély and Achermann, 1992; Mallis et al., 2004; Belyavin and Spencer, 2004; Jewett and Kronauer, 1999; Roach et al., 2004; Åkerstedt et al., 2004)). Typical implementations take the sleep-wake history to develop work schedules and shift work rotations (Dawson et al., 2011). 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 (Gander et al., 2011). 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, several 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 (Gander et al., 2011). 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 (Gander et al., 2011). 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.

In addition, 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 paper, 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 min). 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 codes, which were used to generate the results in this study, can be accessed through the link provided in the Supplementary Materials Section. These codes can be used to develop predictive physical fatigue models in the workplace.

5.3. Study limitations and future research

There are a few 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 (Murata, 2016). 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 min of data was also divided into 18, 10-min 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 (Di Milia et al., 2011). 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 min) 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 (Fig. 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 paper. 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.r-project.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.

References

Achermann, P., 2004. The two-process model of sleep regulation revisited. Aviat. Space, Environ. Med. 75 (Suppl. 1), A37–A43.
Åkerstedt, T., Folkard, S., Portin, C., 2004. Predictions from the three-process model of alertness, aviation. Space, Environ. Med. 75 (3), A75–A83.
Amick III, B.C., Robertson, M.M., DeRango, K., Bazzani, L., Moore, A., Rooney, T., Harrist, R., 2003. Effect of office ergonomics intervention on reducing musculoskeletal symptoms. Spine 28 (24), 2706–2711.
Atallah, L., Lo, B.P., King, R.C., Yang, G.-Z., 2010. Sensor Placement for Activity Detection Using Wearable accelerometers., in: BSN, pp. 24–29.
Aughey, R.J., 2010. Australian football player work rate: evidence of fatigue and pacing. Int. J. Sports Physiol. Perform. 5 (3), 394–405.
Balasubramanian, V., Adalarasu, K., Regulapati, R., 2009. Comparing dynamic and stationary standing postures in an assembly task. Int. J. Industrial Ergonomics 39 (5), 649–654.
Balkin, T.J., Horrey, W.J., Graeber, R.C., Czeisler, C.A., Dinges, D.F., 2011. The challenges and opportunities of technological approaches to fatigue management. Accid. Analysis Prev. 43 (2), 565–572.
Bao, L., Intille, S.S., 2004. Activity Recognition from User-annotated Acceleration Data, in: International Conference on Pervasive Computing. Springer, pp. 1–17.
Barker, L.M., Nussbaum, M.A., 2011. The effects of fatigue on performance in simulated nursing work. Ergonomics 54 (9), 815–829.
Battini, D., Persona, A., Sgarbossa, F., 2014. Innovative real-time system to integrate ergonomic evaluations into warehouse design and management. Comput. Industrial Eng. 77, 1–10.
Belyavin, A.J., Spencer, M.B., 2004. Modeling performance and alertness: the qinetiq approach. Aviat. Space, Environ. Med. 75 (3), A93–A103.
Björklund, M., Crenshaw, A.G., Djupsjöbacka, M., Johansson, H., 2000. Position sense acuity is diminished following repetitive low-intensity work to fatigue in a simulated occupational setting. Eur. J. Appl. Physiol. 81 (5), 361–367.
Bonato, P., 2010. Wearable sensors and systems. IEEE Eng. Med. Biol. Mag. 29 (3), 25–36.
Bonato, P., Ebenbichler, G.R., Roy, S.H., Lehr, S., Posch, M., Kollmitzer, J., Della Croce, U., 2003. Muscle fatigue and fatigue-related biomechanical changes during a cyclic lifting task. Spine 28 (16), 1810–1820.
A. A. Borbély, A two process model of sleep regulation., Human Neurobiology.
Borbély, A.A., Achermann, P., 1992. Concepts and models of sleep regulation: an overview. J. Sleep Res. 1 (2), 63–79.
Borg, G.A., 1998. Borg's Perceived Exertion and Pain Scales. Human Kinetics, Champaign, IL.
Bosch, T., De Looze, M., Van Dieen, J., 2007. Development of fatigue and discomfort in the upper trapezius muscle during light manual work. Ergonomics 50 (2), 161–177.
Burgess-Limerick, R., Abernethy, B., Neal, R.J., 1993. Relative phase quantifies interjoint coordination. J. Biomechanics 26 (1), 91–94.
Cat — fatigue risk management system — caterpillar, http://www.cat.com/en_US/support/operations/frms.html, [Online; accessed 1-June-2016].
Catapult usa — wearable technology for elite sports, http://www.catapultsports.com/, [Online; accessed 1-June-2016].
L. A. Cavuoto, F. M. Megahed, Understanding fatigue and the implications for worker safety, Proceedings of the ASSE Professional Development Conference & Exposition, Atlanta, GA, June 26–29 (2016).
Cavuoto, L.A., Nussbaum, M.A., 2014. Influences of obesity on job demands and worker capacity. Curr. Obes. Rep. 3 (3), 341–347.
Chang, F.-L., Sun, Y.-M., Chuang, K.-H., Hsu, D.-J., 2009. Work fatigue and physiological symptoms in different occupations of high-elevation construction workers. Appl. Ergon. 40 (4), 591–596.
Cheng, T., Migliaccio, G.C., Teizer, J., Gatti, U.C., 2012. Data fusion of real-time location sensing and physiological status monitoring for ergonomics analysis of construction workers. J. Comput. Civ. Eng. 27 (3), 320–335.
Cook, C., Burgess-Limerick, R., 2004. The effect of forearm support on musculoskeletal discomfort during call centre work. Appl. Ergon. 35 (4), 337–342.
Côté, J.N., Raymond, D., Mathieu, P.A., Feldman, A.G., Levin, M.F., 2005. Differences in multi-joint kinematic patterns of repetitive hammering in healthy, fatigued and shoulder-injured individuals. Clin. Biomech. 20 (6), 581–590.
Dag, A., Topuz, K., Oztekin, A., Bulur, S., Megahed, F.M., 2016. A probabilistic data-driven framework for scoring the preoperative recipient-donor heart transplant survival. Decis. Support Syst. 86, 1–12.
Dawson, D., Noy, Y.I., Härmä, M., Åkerstedt, T., Belenky, G., 2011. Modelling fatigue and the use of fatigue models in work settings. Accid. Analysis Prev. 43 (2), 549–564.
de Oliveira Sato, T., Coury, H.J.C.G., 2009. Evaluation of musculoskeletal health outcomes in the context of job rotation and multifunctional jobs. Appl. Ergon. 40 (4), 707–712.
Dempsey, P.G., 1998. A critical review of biomechanical, epidemiological, physiological and psychophysical criteria for designing manual materials handling tasks. Ergonomics 41 (1), 73–88.
Di Milia, L., Smolensky, M.H., Costa, G., Howarth, H.D., Ohayon, M.M., Philip, P., 2011. Demographic factors, fatigue, and driving accidents: an examination of the published literature. Accident Analysis & Prevention 43 (2), 516–532.
Diego-Mas, J.A., Alcaide-Marzal, J., 2014. Using kinect sensor in observational methods for assessing postures at work. Appl. Ergon. 45 (4), 976–985.
Dong, H., Ugaldey, I., El Saddik, A., 2014. Development of a fatigue-tracking system for monitoring human body movement. In: Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2014 IEEE International. IEEE, pp. 786–791.
Forestier, N., Nougier, V., 1998. The effects of muscular fatigue on the coordination of a multijoint movement in human. Neurosci. Lett. 252 (3), 187–190.
Gallagher, S., Heberger, J.R., 2013. Examining the interaction of force and repetition on musculoskeletal disorder risk a systematic literature review. Hum. Factors J. Hum. Factors Ergonomics Soc. 55 (1), 108–124.
Gander, P., Hartley, L., Powell, D., Cabon, P., Hitchcock, E., Mills, A., Popkin, S., 2011. Fatigue risk management: organizational factors at the regulatory and industry/

company level. Accid. Analysis Prev. 43 (2), 573–590.
Garg, A., Kapellusch, J.M., 2009. Applications of biomechanics for prevention of work-related musculoskeletal disorders. Ergonomics 52 (1), 36–59.
Gatti, U.C., Schneider, S., Migliaccio, G.C., 2014. Physiological condition monitoring of construction workers. Automation Constr. 44, 227–233.
V. J. Gawron, J. French, D. Funke, An overview of fatigue.
GPS unit — performance monitoring & analysis — STATSports, http://statsports.com/technology/viper-pod/, [Online; accessed 1-June-2016].
Heinz, E.A., Kunze, K.S., Sulistyo, S., Junker, H., Lukowicz, P., Tröster, G., 2003. Experimental evaluation of variations in primary features used for accelerometric context recognition. In: European Symposium on Ambient Intelligence. Springer, pp. 252–263.
Hesterberg, T., Choi, N.H., Meier, L., Fraley, C., et al., 2008. Least angle and 1 penalized regression: a review. Stat. Surv. 2, 61–93.
Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12 (1), 55–67.
Hu, B., Ning, X., 2015. The influence of lumbar extensor muscle fatigue on lumbar–pelvic coordination during weightlifting. Ergonomics 58 (8), 1424–1432.
Hursh, S.R., Redmond, D.P., Johnson, M.L., Thorne, D.R., Belenky, G., Balkin, T.J., Storm, W.F., Miller, J.C., Eddy, D.R., 2004. Fatigue models for applied research in warfighting. Aviat. space, Environ. Med. 75 (Suppl. 1), A44–A53.
Huynh, T., Schiele, B., 2005. 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. ACM 159–163.
Huysmans, M.A., Hoozemans, M.J., van der Beek, A.J., de Looze, M.P., van Dieën, J.H., 2010. Position sense acuity of the upper extremity and tracking performance in subjects with non-specific neck and upper extremity pain and healthy controls. J. Rehabilitation Med. 42 (9), 876–883.
Jewett, M.E., Kronauer, R.E., 1999. Interactive mathematical models of subjective alertness and cognitive throughput in humans. J. Biol. Rhythms 14 (6), 588–597.
Kajimoto, O., 2008. Fatigue science for human health, springer Japan, tokyo. Ch. Dev. a Method Eval. Fatigue Its Econ. Impacts 33–46. http://dx.doi.org/10.1007/978-4-431-73464-2_3. URL. http://dx.doi.org/10.1007/978-4-431-73464-2_3.
Kang, H.G., Mahoney, D.F., Hoenig, H., Hirth, V.A., Bonato, P., Hajjar, I., Lipsitz, L.A., 2010. In situ monitoring of health in older adults: technologies and issues. J. Am. Geriatrics Soc. 58 (8), 1579–1586.
Kelsey, J.L., Githens, P.B., White, A.A., Holford, T.R., Walter, S.D., O'Connor, T., Ostfeld, A.M., Weil, U., Southwick, W.O., Calogero, J.A., 1984. An epidemiologic study of lifting and twisting on the job and risk for acute prolapsed lumbar intervertebral disc. J. Orthop. Res. 2 (1), 61–66.
Kern, N., Schiele, B., Schmidt, A., 2003. Multi-sensor activity context detection for wearable computing. In: European Symposium on Ambient Intelligence. Springer, pp. 220–232.
Kim, S., Nussbaum, M.A., 2013. Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. Ergonomics 56 (2), 314–326.
Kimura, M., Sato, H., Ochi, M., Hosoya, S., Sadoyama, T., 2007. Electromyogram and perceived fatigue changes in the trapezius muscle during typewriting and recovery. Eur. J. Appl. Physiology 100 (1), 89–96.
Kotsiantis, S., Kanellopoulos, D., Pintelas, P., et al., 2006. Handling imbalanced datasets: a review. GESTS Int. Trans. Comput. Sci. Eng. 30 (1), 25–36.
Krause, A., Siewiorek, D.P., Smailagic, A., Farringdon, J., 2003. Unsupervised, Dynamic Identification of Physiological and Activity Context in Wearable computing., in: ISWC, vol. 3, p. 88.
Krueger, G.P., 1989. Sustained work, fatigue, sleep loss and performance: a review of the issues. Work & Stress 3 (2), 129–141.
Kumar, S., 2001. Theories of musculoskeletal injury causation. Ergonomics 44 (1), 17–47.
Lee, S.-W., Mase, K., 2002. Activity and location recognition using wearable sensors. IEEE pervasive Comput. 1 (3), 24–32.
Lehman, K., Psihogios, J., Meulenbroek, R., 2001. Effects of sitting versus standing and scanner type on cashiers. Ergonomics 44 (7), 719–738.
Liu, T., Inoue, Y., Shibata, K., 2009. Development of a wearable sensor system for quantitative gait analysis. Measurement 42 (7), 978–988.
Looze, M. d., Bosch, T., Dieën, J. v., 2009. Manifestations of shoulder fatigue in prolonged activities involving low-force contractions. Ergonomics 52 (4), 428–437.
L. Lu, R. F. Sesek, F. M. Megahed, L. N. Cavuoto, A survey of the prevalence of fatigue among workers at u.s. manufacturers, Applied ErgonomicsUnder Review.
Magora, A., 1973. Investigation of the relation between low back pain and occupation. iv. physical requirements: bending, rotation, reaching and sudden maximal effort. Scand. J. Rehabilitation Med. 5 (4), 186.
Mallis, M.M., Mejdal, S., Nguyen, T.T., Dinges, D.F., 2004. Summary of the key features of seven biomathematical models of human fatigue and performance, aviation. Space, Environ. Med. 75 (3), A4–A14.
Marras, W.S., Schoenmarxlin, R.W., 1993. Wrist motions in industry. Ergonomics 36 (4), 341–351.
Marras, W.S., Lavender, S.A., Leurgans, S.E., Rajulu, S.L., Allread, W.G., Fathallah, F.A., Ferguson, S.A., 1993. 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.
Marras, W.S., Cutlip, R.G., Burt, S.E., Waters, T.R., 2009. National occupational research agenda (nora) future directions in occupational musculoskeletal disorder health research. Appl. Ergon. 40 (1), 15–22.
Mital, A., Foononi-Fard, H., Brown, M.L., 1994. Physical fatigue in high and very high frequency manual materials handling: perceived exertion and physiological indicators. Hum. Factors J. Hum. Factors Ergonomics Soc. 36 (2), 219–231.
Montgomery, D.C., 2014. Introduction to Statistical Quality Control. John Wiley & Sons.
Murata, A., 2016. Proposal of a method to predict subjective rating on drowsiness using physiological and behavioral measures. IIE Trans. Occup. Ergonomics Hum. Factors 1–13.
Neuberger, G.B., 2003. 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 & Res. 49 (S5), S175–S183.
Nyan, M., Tay, F.E., Murugasu, E., 2008. A wearable system for pre-impact fall detection. J. Biomechanics 41 (16), 3475–3481.
Ohgi, Y., 2002. Proceedings of IEEE. Microcomputer-based Acceleration Sensor Device for Sports Biomechanics-stroke Evaluation by Using Swimmer's Wrist Acceleration, in: Sensors, vol. 1. IEEE, pp. 699–704, 2002.
Ohgi, Y., Yasumura, M., Ichikawa, H., Miyaji, C., 2000. Analysis of Stroke Technique Using Acceleration Sensor Ic in Freestyle Swimming. The Engineering of Sport, pp. 503–512.
Pantelopoulos, A., Bourbakis, N.G., 2010. A survey on wearable sensor-based systems for health monitoring and prognosis, systems, man, and cybernetics, Part C: applications and reviews. IEEE Trans. 40 (1), 1–12.
Pirttikangas, S., Fujinami, K., Nakajima, T., 2006. Feature selection and activity recognition from wearable sensors. In: International Symposium on Ubiquitious Computing Systems. Springer, pp. 516–527.
Poh, M.-Z., Swenson, N.C., Picard, R.W., 2010. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. Biomedical Eng. IEEE Trans. 57 (5), 1243–1252.
L. Quiros, Bnl ergonomics bulletin, Brookhaven National Laboratory, New York 4.
Ravi, N., Dandekar, N., Mysore, P., Littman, M.L., 2005. Activity Recognition from Accelerometer Data, in: AAAI, vol. 5, pp. 1541–1546.
Ray, S.J., Teizer, J., 2012. Real-time construction worker posture analysis for ergonomics training. Advanced Eng. Informatics 26 (2), 439–455.
Roach, G.D., Fletcher, A., DAWsON, D., 2004. A model to predict work-related fatigue based on hours of work. Aviat. Space, Environ. Med. 75 (Suppl. 1), A61–A69.
Rodacki, A.L., Fowler, N.E., Bennett, S.J., 2001. Multi-segment coordination: fatigue effects. Med. Sci. Sports Exerc. 33 (7), 1157–1167.
Roja, Z., Kalkis, V., Vain, A., Kalkis, H., Eglite, M., 2006. Assessment of skeletal muscle fatigue of road maintenance workers based on heart rate monitoring and myotonometry. J. Occup. Med. Toxicol. 1 (1), 1–9.
Schall, M.C., Fethke, N.B., Chen, H., Gerr, F., 2015. A comparison of instrumentation methods to estimate thoracolumbar motion in field-based occupational studies. Appl. Ergon. 48, 224–231.
Schall, M.C., Fethke, N.B., Chen, H., 2016. Working postures and physical activity among registered nurses. Appl. Ergon. 54, 243–250.
Schall, M.C., Fethke, N.B., Chen, H., 2016. Evaluation of four sensor locations for physical activity assessment. Appl. Ergon. 53, 103–109.
Seay, J.F., Van Emmerik, R.E., Hamill, J., 2011. Low back pain status affects pelvis-trunk coordination and variability during walking and running. Clin. Biomech. 26 (6), 572–578.
Sharpe, M., Archard, L., Banatvala, J., Borysiewicz, L., Clare, A., David, A., Edwards, R., Hawton, K.E., Lambert, H., Lane, R., et al., 1991. A report–chronic fatigue syndrome: guidelines for research. J. R. Soc. Med. 84 (2), 118.
Straker, L., Stevenson, M., Twomey, L., 1996. A comparison of risk assessment of single and combination manual handling tasks: 1. maximum acceptable weight measures. Ergonomics 39 (1), 128–140.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Methodol. 267–288.
Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D., Marin, F., 2013. Innovative system for real-time ergonomic feedback in industrial manufacturing. Appl. Ergon. 44 (4), 566–574.
Visser, B., van Dieën, J.H., 2006. Pathophysiology of upper extremity muscle disorders. J. Electromyogr. Kinesiol. 16 (1), 1–16.
Waters, T.R., Putz-Anderson, V., Garg, A., Fine, L.J., 1993. Revised niosh equation for the design and evaluation of manual lifting tasks. Ergonomics 36 (7), 749–776.
Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., Connor, J.L., 2011. The link between fatigue and safety. Accid. Analysis Prev. 43 (2), 498–515.
Yung, M., 2016. Fatigue at the Workplace: Measurement and Temporal Development. PhD Dissertation. University of Waterloo. Located at. https://uwspace.uwaterloo.ca/handle/10012/10119.
Yung, M., Bigelow, P.L., Hastings, D.M., Wells, R.P., 2014. Detecting within-and between-day manifestations of neuromuscular fatigue at work: an exploratory study. Ergonomics 57 (10), 1562–1573.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B Stat. Methodol. 67 (2), 301–320.

Please cite this article in press as: Sedighi Maman, Z., et al., A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied Ergonomics (2017), http://dx.doi.org/10.1016/j.apergo.2017.02.001