

# Feature identification framework for back injury risk in repetitive work with application in sheep shearing

Mark Robinson, *Student Member, IEEE*, Lei Lu, Ying Tan, *Senior Member, IEEE*, Denny Oetomo, *Senior Member, IEEE*, and Chris Manzie, *Senior Member, IEEE*

**Abstract—Background:** Lower back injuries are a serious global problem. Most of these injuries occur over time with repeated sub-acute stresses. Neuromuscular control dysfunction could predict injury, however injuries are almost never observed alongside this data. No labels are available to identify important features that may be predictive of injury. While there are many individual differences in injury development, the population trend is that each individual's injury tolerance decreases over time with exposure, indicating a monotonic process.

**Methods:** This paper proposes a framework for identifying key features of injury using an unsupervised technique that exploits knowledge of injury aetiology by analysing which features contribute to the popular trend using weak monotonicity from data segmented by task repetitions. The feature selection also evaluates feature redundancy. The efficacy of the framework is demonstrated through data from on-site sheep shearers over one day using 17 wearable inertial measurement units and 16 surface electromyography (sEMG) sensors.

**Results:** Consistent with literature, the results demonstrate sEMG features derived from the erector spinae and multifidus muscles are the most important indicators for lower back injury.

To evaluate the performance of the proposed population-trend based unsupervised feature selection technique, the self-reported fatigue information is treated as some 'ground truth' information so that this proposed technique can compare with 5 existing unsupervised feature selection techniques. **Conclusion:** The proposed technique is shown to be the most consistent with the self-reported fatigue information, demonstrating the effectiveness of the proposed method.

**Index Terms**—Lower back injuries, weak monotonicity, population trend, activity recognition.

## I. INTRODUCTION

**M**OST people suffer from a back disorder at one time or another [1], with back pain recently becoming the leading cause of disability worldwide [2]. Epidemiological studies of workplace injury data indicate that repetitive stress is a major risk factor associated with lower back disorders [3], and these disorders are especially prevalent among occupations that require prolonged or repetitive spinal flexion [4].

It is proposed that time-dependent changes in kinematics and neuromuscular control are important in the aetiology of many back injuries [5]. It was demonstrated that the majority of back injuries do not occur from a single large exposure, but from repeated and prolonged sub-acute exposure to stresses [6], [7]. It has been suggested that motor control deficiency is the most important factor in predicting the development

of back disorders [8]. However, this has not been proven, and there is a clear need for studies that can confirm a link between motor control deficiencies to pain and injury in realistic occupational tasks [9].

Recent developments in advanced wearable sensor technology make it possible to collect biometric data outside of the lab. However, the collection of kinematic and neuromuscular control data over the time-period that would lead to injury (perhaps several months or more) is still very challenging as usually it is difficult to record injury or 'ground truth', making it difficult to link relate these factors to injury. Thus, the collected data is not able to be labelled.

Without labels, this paper aims to identify key features that are likely to contribute to injury through a unsupervised feature selection. Successfully identifying these features without labelled data can connect factors that are related to injury risk with data collected at the time-scale of the physiological changes that lead to injury, rather than the longer time-scale of injury development.

Extensive research has been performed on feature selection methods with applications in healthcare [10], [11]. Especially, recent advances of deep learning techniques, i.e. convolutional neural networks (CNN), long short-term memory (LSTM) networks, have demonstrated excellent performance on learning latent features from heterogeneous sensor data [12]. However, the problem with these machine learning techniques are that they either work as supervised learning, which needs labels, or the learned results are difficult to explain and can not provide explicit information to end-users such as clinicians or patients. In order to understand the role of features in back injury prevention in terms of physical variables and time-dependent physiological changes, feature selection method needs to be interpretable for end-users.

Typically, unsupervised feature selection defines cost functions to exploit intrinsic structural properties in the data [13]. In cases with a family of processes, knowledge of similarities between processes can be exploited [14]. In this context, a population of workers are exposed to similar stresses, and injuries develop in a similar way. The common information can form part of the cost function, where features that are similar among the population can be judged as relevant. In terms of lower back injury, it is evident in the literature that many injuries occur as a process where injury tolerance is reduced over time; injury risk continually 'gets worse' as this process evolves [3], [6]. When the relevant time-period of these changes are known, it's possible to incorporate the concept of monotonicity in injury development into the cost function to help determine feature relevance [11], though this monotonic trend may be corrupted by noises from sensors, human variations, and repetitions of tasks.

Manuscript received April 4, 2022. This work is supported by Australian Wool Innovation (project number ON-00639). M. Robinson, L. Lu, Y. Tan, D. Oetomo, and C. Manzie are with the Faculty of Engineering and Information Technology, The University of Melbourne, Parkville, VIC, Australia.

L. Lu is also with the Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK.

Corresponding authors: M. Robinson, email: mrobinson2@student.unimelb.edu.au; and L. Lu, email: lei.lu@eng.ox.ac.uk

In this paper, we present an unsupervised framework, where labelled injury data is absent, to identify important indicators of back injury risk in repetitive work from relevant data collected from a population of workers. By utilising the concept of weak monotonicity, we propose a new cost function that can exploit the commonality between the population of workers, as well as the domain knowledge of back injury aetiology to identify key indicators of back injury, while being robust to the periodicity inherent in repetitive work, and without requiring the long-term collection of neuromuscular data.

The effectiveness of the framework is demonstrated in this paper in an application identifying key features affecting lower back injury risk in the repetitive task of sheep shearing. In sheep shearing, injury rates are severe and injury risk worsens throughout the working day, with 68% more injuries occurring in the last two hours of work compared to the first two hours of work, with lower back injuries the biggest problem [15]. In the experiments, kinematic and surface electromyographic (sEMG) data are collected in real working conditions for Australian sheep shearers. The data collected for one complete day of sheep shearing from each subject is used in our analysis. Key indicators of injury are identified. In order to evaluate the performance of feature identification, self-reported fatigue levels from each sheep shearer is treated as a pseudo ‘ground truth’. The performance of the proposed technique is compared with 5 popular unsupervised feature selection techniques, showing a superior performance. The top selected features are also discussed in terms of the existing literature regarding the aetiology of lower back injury.

The presented framework contributes to the state-of-the-art as it improves the feasibility of the real-world collection of neuromuscular control data and its relation to injury. The application of the framework provides evidence linking neuromuscular control features to lower back injury without the need for data collection over the timescale of injury development. It cannot replace the gold standard of comprehensive longitudinal studies in occupational settings over many months. However, given the practical issues and the difficulty of identifying important features that relate kinematics and neuromuscular control to lower back injury, the proposed framework is a useful tool in the absence of these studies.

## II. PRELIMINARIES

This section introduces the key concepts used in the unsupervised feature selection part of the proposed framework.

### A. Weak Monotonicity

This subsection introduces a new concept of weak monotonicity, which is a relaxation of strict monotonicity [16], and we denote it as  $\Delta$ -weak monotonicity. Strict monotonicity of a collected signal indicates that each measurement point is strictly higher (or lower) than the previous. Weak monotonicity relaxes this requirement such that measurement points must be higher (or lower) than the measurement point minus (plus) a small margin, calculated from the previous measurement points. The relaxation is developed to improve the robustness of the trend evaluation, as we seek to identify population wide trends with the consideration of individual variations and in the presence of measurement noises.

For simplicity, this section focuses on a scalar signal only.

*Definition 1:* A collected tuple  $\{t_k, z_k\}_{k \in \mathbb{N}}$  is  $\Delta$ -weakly monotonically increasing for a constant  $\Delta \geq 0$  if the following condition holds,

$$\begin{aligned} t_{k+1} &> t_k, \quad t_k \xrightarrow{k \rightarrow \infty} \infty \\ z_{k+1} &> z_k - \Delta, \quad k = 1, 2, \dots, \end{aligned} \quad (1)$$

where,  $\Delta$  represents the level of uncertainties.  $\circ$

*Remark 1:* When  $\Delta = 0$ ,  $\Delta$ -weak monotonicity becomes strict monotonicity [14]. The value of  $\Delta$  plays an important role to characterize this weak trend. For example, a  $\Delta$ -weakly monotonically increasing signal might have a decreasing trend. In order to avoid such an error, statistical tools such as the Mann-Kendall (MK) test [17], for monotonic trend, can be used to identify the presence of a trend as a *prior* before analysing its strength. This leads to the definition of the trend indicator for a given data sequence  $\mathbf{z} = [z_1 \dots z_{N_p}]^T$  with  $N_p$  measurement points:

$$\omega(\mathbf{z}) = \begin{cases} 1 & \text{if trend } \mathbf{z} \text{ is positive by MK test} \\ -1 & \text{if trend } \mathbf{z} \text{ is negative by MK test} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where, the test significance level is typically set at  $p < 0.05$ . Other statistical tests, or other *prior* knowledge of the trend could be used to define  $\omega$  as well.

The following measure  $\Delta M$  is used to quantify the  $\Delta$ -weakly monotonic (increasing) trend of a finite duration signal with  $N_p$  measurement points,

$$\begin{aligned} \Delta M &= \frac{1}{N_p - 1} \left( \sum_{k=1}^{N_p} \alpha_k \right), \\ \alpha_{k+1} &= \begin{cases} 1 & \text{if } z_{k+1} \geq z_k - \Delta, k = 0, \dots, N_p - 1 \\ -1 & \text{otherwise} \end{cases} \end{aligned} \quad (3)$$

for  $k = 0, \dots, N_p - 1$ . The  $\Delta M$  measure has a range within  $[-1, 1]$ . When  $\Delta M$  is close to 1 or  $-1$ , it indicates that the monotonically increasing or decreasing trend of a signal is more pronounced.

It is usually hard to find an appropriate (or less conservative) upper bound  $\Delta$  of a signal. A possible solution of estimating such uncertainties is to use some statistical properties of the signal from off-line measurements or data segments through on-line moving windows. When a group of signals or processes or subjects is considered, it is possible to estimate averaged variations among them.

It is worthwhile to highlight that the definition of  $\Delta$ -weak monotonicity does not require the sampling periods to be fixed, i.e.,  $t_{k+1} - t_k \neq t_k - t_{k-1}, k = 1, 2, \dots$ , this allows for more general cases of data sampling.

### B. Correlation

For a group of subjects  $j = 1, 2, \dots, N$ , let  $\Sigma^j : \{t_k^j, \mathbf{x}_k^j\}$  represent the set of features  $\mathbf{x}_k^j \in \mathbb{R}^{N_f}$  for the  $j^{th}$  subject at time  $t_k$ , where  $k = 1, \dots, N_j$  with  $N_j \in \mathbb{N}$  indicates the samples. Here  $\mathbb{R}$  is the set containing all real numbers, and  $\mathbb{N}$  is the set containing all integers. The correlation between two

subjects  $p$  and  $q$ , where  $p, q = 1, 2, \dots, N$ , for the  $i^{th}$  feature is defined as follows,

$$\rho(\mathbf{x}_i^p, \mathbf{x}_i^q) = \frac{\sum_{k=1}^{N_s} (x_{k,i}^p - \bar{x}_i^p)(x_{k,i}^q - \bar{x}_i^q)}{\left( \sum_{k=1}^{N_p} (x_{k,i}^p - \bar{x}_i^p)^2 \right)^{1/2} \left( \sum_{k=1}^{N_q} (x_{k,i}^q - \bar{x}_i^q)^2 \right)^{1/2}} \quad (4)$$

where,  $\bar{x}_i^p = \frac{1}{N_p} \sum_{k=1}^{N_p} x_{k,i}^p$  and  $\bar{x}_i^q = \frac{1}{N_q} \sum_{k=1}^{N_q} x_{k,i}^q$  are the averaged signals with data truncation of  $N_s = \min\{N_p, N_q\}$ .

### C. Feature Orthogonality

In the context of regression, orthogonality is an important factor to investigate the existence of redundant information between two features. Intuitively, orthogonality with inner product is equivalent to ‘‘uncorrelatedness’’. If two features are orthogonal to each other, the second feature provides new information that the first feature does not have.

To this end, an ordinary least squares (OLS) regression is calculated of a new candidate feature in terms of the already selected features.

For convenience of notation, for the  $i^{th}$  feature of the  $j^{th}$  subject, we introduce the vector of  $N_p$  measurements  $\tilde{\mathbf{x}}_i^j = [x_{1,i}^j \ \dots \ x_{N_p,i}^j]^T \in \mathbb{R}^{N_p}$ , and define a matrix  $\tilde{\mathbf{X}}_{n_a}^j = [\tilde{\mathbf{x}}_1^j \ \dots \ \tilde{\mathbf{x}}_{n_a}^j]^T \in \mathbb{R}^{N_p \times |n_a|}$  to indicate the selected set of  $n_a$  features for the  $j^{th}$  subject.

The fraction of unexplained variance in (6), calculated from the OLS regression of a candidate feature  $\tilde{\mathbf{X}}_{n_a}^j$  with the existing feature set  $\tilde{\mathbf{X}}_{n_a}^j$ , seen in (5), will be used establish feature orthogonality.

$$\tilde{\mathbf{x}}_{n_a+1}^j = \tilde{\mathbf{X}}_{n_a}^j (\tilde{\mathbf{X}}_{n_a}^{jT} \tilde{\mathbf{X}}_{n_a}^j)^{-1} \tilde{\mathbf{X}}_{n_a}^{jT} \mathbf{x}_{n_a+1}^j, \quad (5)$$

provided that  $(\tilde{\mathbf{X}}_{n_a}^{jT} \tilde{\mathbf{X}}_{n_a}^j)$  is non-singular. This matrix is non-singular if  $|n_a|$  is smaller than the number of data points in each feature. The second step estimates the fraction of unexplained variance (i.e. the complement of the coefficient of multiple determination), which reflects the level of new information that  $\mathbf{x}_{n_a+1}^j$  brings to the existing feature set  $\tilde{\mathbf{X}}_{n_a}^j$  [18]. This fraction is calculated as,

$$u_{n_a+1}^j = \frac{\sum_{k=1}^{N_p} (x_{k,n_a+1}^j - \hat{x}_{k,n_a+1}^j)^2}{\sum_{k=1}^{N_p} (x_{k,n_a+1}^j - \bar{x}_{k,n_a+1}^j)^2}, \quad (6)$$

where  $\bar{x}_i = \frac{1}{N_p} \sum_{k=1}^{N_p} x_{k,i}$ .

With consideration of  $N$  subjects in the population, we have the following averaged orthogonality index

$$\bar{u}_{n_a+1} = \frac{1}{N} \sum_{j=1}^N u_{n_a+1}^j. \quad (7)$$

It is noted that  $\bar{u}_{n_a+1}$  takes values between  $[0, 1]$ , with 1 indicating that none of the new feature’s variance is explained

by the already selected features. In this paper, the averaged orthogonality defined in (7) will be the part of the cost function to select new features.

## III. THE PROPOSED FRAMEWORK

This section presents the proposed framework which identifies key features in the collected time-series data that relate to injury risk. The diagram is presented in Figure 1. Details of each block in Figure 1 are presented in the following subsections.

The overall goal is to select the desired feature set  $X^*$  from the full set of features  $X$  across all subjects that best represents the injury development process. Without the availability of labelled data, unsupervised feature selection or ranking algorithms are adapted here. A key part of any unsupervised feature ranking algorithm is to establish feature relevance without labels. In the proposed framework this is achieved by analysing the individual monotonic trends with a  $\Delta$ -weak monotonicity indicator, and the commonality of these trends across the population for the extracted features. It is hypothesised that this is a good measure of feature relevance under the assumption that there is an underlying monotonic process, i.e. the risk of injury is increasing over time. This assumption is accepted for the majority of back injuries [6], [7], and is confirmed in [15] within the context of the application presented in this paper.

However, there are challenges in identifying monotonic trends in time-series data collected from many people performing repetitive tasks. When performing repetitive tasks, the repetition will be present in the time-series or signals, showing a clear periodicity. This periodicity will interfere with the calculation of weak monotonicity. Therefore, the influence of task repetitions should be removed where possible. In this work, human activity recognition (HAR) techniques are employed to remove the influence of repetitions. Such techniques can identify the start and end point of one repetition. The features are then extracted from a window corresponding to one task repetition, reducing the periodicity in the extracted feature vectors. It is noted that though the proposed framework is general enough to handle various tasks, introducing HAR techniques leads to task specific feature extraction algorithms, which are consistent with the requirement of identifying motor dysfunctions. As elucidated in [19], motor control dysfunctions are task specific, and each task should be analysed differently with a potentially different set of features.

After feature relevance is established for the candidate features, features are selected iteratively using an indicator to minimise feature redundancy, i.e. having multiple features that yield the same information.

### A. Pre-processing

The input to the framework is data (R) containing information of kinematics and neuromuscular signals from a large number of sensors, collected from a number of subjects performing repetitive work. We can sample this set of signals at given uniform or non-uniform sampling instants, and in later analysis, an appropriate sampling period is selected.

In general, the pre-processing step involves normalisation, filtering, and/or data fusion to obtain a new representation

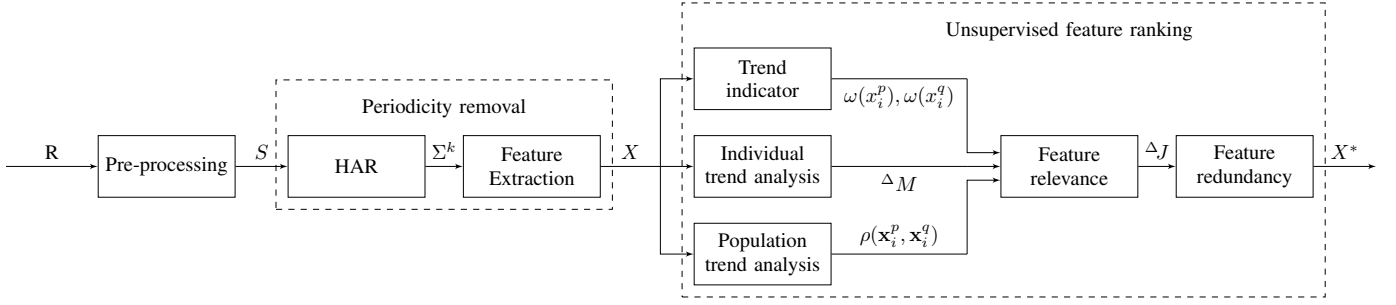


Fig. 1. The proposed framework consists of the periodicity removal through human activity recognition (HAR), before ranking the extracted features using a trend based unsupervised technique to connect kinematic and neuromuscular features to injury risk. The detail of each box in this diagram is presented throughout Section III.

of the data. The pre-processing can produce multiple streams (channels) of data from each sensor. For the  $\ell^{th}$  channel,

$$S_\ell := \{S_\ell(t_0), \dots, S_\ell(t_0 + m_\ell \delta_\ell)\}, \ell = 1, \dots, N_c, \quad (8)$$

where,  $S_\ell$  is sampled at  $m_\ell$  instants with sample rate  $\delta_\ell$ . Then, the pre-processed data-set across all channels is

$$S := \{S_1, \dots, S_{N_c}\}, \quad (9)$$

which allows for asynchronous sampling between channels.

### B. Periodicity Removal

1) *Human activity recognition (HAR)*: As the periodicity in the measured data should be removed, the first step is to solve a HAR problem to identify the start and finish points of the activities within the repetitive task. Once the data is segmented, candidate features can be extracted from each activity. It is assumed that this family of sensors is able to detect the activities of interest.

In order to remove the periodicity from task repetition, activity recognition can be used to identify the activities and segment the data into each iteration in the segmented sequence  $\Sigma^k$ . Features are then extracted from within each iteration to represent the behaviour of the task. It is assumed here that existing activity recognition techniques can well-approximate the ideal segmentation  $\Sigma^k$ . Such HAR techniques have been demonstrated with good accuracy in several contexts, including workplace tasks [20], [21]. Slight error in the HAR algorithm is generally inevitable and is captured with a larger  $\Delta$  parameter in the weak monotonicity calculation.

There are many possible approaches to solving the HAR classification problem itself, the most successful methods tend to incorporate a way to model temporal dependencies in the data [20]. A person-independent classifier is able to reduce the amount of data required for training, and to obtain a good result for these models, it is necessary to select features that minimise inter-individual differences [20], [22].

We represent task repetitions as a sequence, which can be measured by the pre-processed data:

$$\Sigma^k : \{(S_1, S_2, \dots, S_{N_c}) \mid \forall n_\ell : t_0 + n_\ell \delta_\ell \in [T_k, T_{k+1})\}, \quad (10)$$

Here the interval  $[T_k, T_{k+1})$  is the time during the  $k^{th}$  activity.

2) *Feature extraction*: The feature extraction takes the segmented data ( $\Sigma^k$ ) as input. For each segment of the data, representing one task repetition, it calculates the candidate features that can represent the behaviour in each task. The feature extraction generates the following mapping  $\Phi$  such that it maps the measurements over one task as a set of vectors

$$\Phi : \Sigma^k \rightarrow \{t_k, \mathbf{x}_k\}_{k=1, \dots, N_p}, \quad (11)$$

where,  $t_k$  is representative of the time when the  $k^{th}$  iteration happens (e.g. time at the start of the task), with  $N_p$  total iterations. Here  $\mathbf{x}_k \in R^{N_f}$  is a vector of signals that are extracted from measurements of  $N_c$  pre-processed channels, with  $N_f$  the number of features. The features can be grouped across each task iteration

$$X^j := \{\mathbf{x}_1^j, \dots, \mathbf{x}_{N_p}^j\}, \quad (12)$$

and across all subjects,

$$X := \{X^1, \dots, X^j, \dots, X^{N_s}\} \quad (13)$$

to give the full feature-set.

The specific features to be extracted will be entirely dependent on the application; for the case study in sheep shearing, these will be presented in Section IV.

### C. Unsupervised feature ranking

1) *Individual trend analysis*: Most back injuries occur as a result of repeated sub-acute stresses which lead to injury over time. This has been clearly identified in the literature where injury tolerance decreases over time to the point where the risk of injury is elevated under the same workplace stresses [6]. Injury risk continually ‘gets worse’ as this process evolves.

The proposed framework exploits this knowledge by using it to determine feature relevance. The concept of weak monotonicity described in Section II is employed as a robust measure of the trend of the population in the presence of measurement noises and human variations. The individual trend analysis receives the extracted features ( $X$ ) as input and calculates the Weak Monotonicity indicator ( $\Delta M_i^p$ ) for each feature from each subject using (3).

2) *Population trend analysis*: The purpose of this function is to establish the similarity between subjects within the population with respect to each feature in order to further identify relevant features. It is noted in [3], [6], that the trend of increasing injury risk in repetitive work is shared across the population. Different people in the occupation will be exposed



to similar repetitive stresses, and injuries occur in similar ways even though there are inter-individual differences in the initial injury tolerance, and rate-of-onset of injury. Therefore, to evaluate the relevance of features with weakly monotonic trends it is useful to determine the commonality of the trends across the population for a given feature. A feature with a pronounced trend, that evolves similarly across the population should be considered relevant.

As summarised in Fig 1, this module takes the extracted features for subjects ( $X$ ) as input and outputs the between subject correlation for that feature across pairs of subjects ( $\rho(\mathbf{x}_i^p, \mathbf{x}_i^q)$ ). For each feature, this can be calculated for each pair of subjects using (4).

3) *Feature Relevance*: After calculating the  $\Delta$ -weak monotonicity indicator ( $\Delta M_i^p$ ) and the similarity between subjects in each feature ( $\rho(\mathbf{x}_i^p, \mathbf{x}_i^q)$ ), a cost function ( $\Delta J_i$ ) is proposed to evaluate the relevance of each feature in terms of characterizing the common trend of the group of subjects. The feature set  $X = \{X_i\}_{i=1, \dots, N_f}$  is defined, and the relevance of the  $i^{th}$  feature is defined as

$$\Delta J_i = \frac{2}{N(N-1)} \sum_{p=1}^{N-1} \sum_{q=p}^N \left[ \left( \omega(\mathbf{x}_i^p) \Delta M_i^p + \omega(\mathbf{x}_i^q) \Delta M_i^q \right) \cdot \rho(\mathbf{x}_i^p, \mathbf{x}_i^q) \right], \quad (14)$$

where  $\Delta M_i^p$  is the calculated  $\Delta$ WM indicator of the  $i^{th}$  feature for the  $p^{th}$  subject, the trend indicator  $\omega(\cdot)$  is defined in (2).

4) *Feature Redundancy*: The features are scored so far based on their monotonic trend and the commonality of that trend across the population. However, multiple highly relevant features may contain redundant information. Once the features are scored in terms of relevance, the goal is to select the a set of most suitable features, which have less overlapping (redundant) information. Features can be selected iteratively by evaluating the average orthogonality index described in (7).

Given a current selection of features,  $n_a$ , the next best feature is selected iteratively using a greedy approach until the desired number of features are selected. The feature with highest relevance score ( $X_{k,1}$ ) is initially added to  $n_a$ . Then to select each additional feature, the averaged orthogonality index is calculated for the features not in  $n_a$  using (7) to give the orthogonal part of the relevance score, and the next feature is selected as the solution to the following optimization problem,

$$\arg \max_{i \notin n_a} \Delta J_i \cdot \bar{u}_{n_a+1}, \quad (15)$$

where,  $\bar{u}_{n_a+1}$  is defined in (7).

This process of adding new features is repeated until  $X^*$  features have been selected. The feature selection process is summarised in Algorithm 1.

#### IV. EXPERIMENTAL SETTING

##### A. Participants

Nine male sheep shearers aged  $33.55 \pm 14.44$ , with a range between 21 and 61, with heights  $1.83\text{m} \pm 0.093\text{m}$  were recruited for the study. All shearers provided informed written consent and the experiment was approved by The University of Melbourne Human Ethics advisory group (Ethics ID 1853436). This represents a wide-ranging selection of shearers with

---

##### Algorithm 1: Feature Selection Algorithm Based on $\Delta$ WM and orthogonality index

---

**Input** :  $|X^*|$ , Features  $\Delta J_i$ , as in (14)  
**Output**: Selected  $X^*$  features, where  $|X^*| \leq |X|$

```

1 Initialise feature set  $n_a \leftarrow X_1$ 
2 while  $|n_a| < |X^*|$  do
3   for  $X_i \notin n_a$  do
4     for  $j \leftarrow 1$  to  $N$  do
5       Compute  $\hat{\mathbf{x}}_{n_a+1}^j$  according to Eq. (5).
6       Compute  $u_{n_a+1}^j$  according to Eq. (6).
7     end
8     Compute the averaged correlation index  $\bar{u}_{n_a+1}$  according to Eq. (7).
9     Compute  $\Delta J_i \cdot \bar{u}_{n_a+1}$ .
10  end
11   $n_a \leftarrow n_a + \arg \max_{i \notin n_a} \Delta J_i \cdot \bar{u}_{n_a+1}$ .
12 end
```

---

varying levels of skill, from two weeks experience to more than 40 years. Each shearer was recorded for the entire work day, which is eight hours split over four two-hour sessions (shearing runs). Interruptions occur in shearing, and so for three subjects only three of the runs took place. One shearer was recorded over three consecutive days, to give some insight into inter-day variations, and these data are included as two extra subject-days, giving 11 days of data collected from nine subjects. Some shearers use a ceiling-mounted back harness during their regular work, a common piece of safety equipment in sheep shearing. Five subjects made use of a harness and this was noted for the statistical analysis.

##### B. Experimental Setup

Kinematic data was collected with the Xsens Awinda portable motion capture system. The system consists of seventeen inertial measurement units (IMUs), and through sensor fusion of the accelerometer, gyroscope, and magnetometer channels in each combined with a scaled skeletal model, the joint kinematics were produced by the Xsens MVN Analyse software and sampled at 60Hz. The experimental setup can be seen in Figures 2, 3, 4, and 5. And the protocol of the observations is shown in Figure 6.

Subjects muscle activity was collected using Delsys Avanti wireless sEMG sensors. Sixteen sEMG sensors sampled at 2148Hz with a 10mm inter-electrode distance. The sensors were bilaterally placed on the following muscles: Erector Spinae at the level of the first and third lumbar vertebrae (L1 & L3 ES), Multifidus at the level of the fifth lumbar vertebrae (L5 MF), Rectus Abdominis (RA), External Oblique (EO), Gluteus Medius (GM), Vastus Lateralis (VL), and Biceps Femoris (BF). The EMG sensors were placed bilaterally in accordance with the SENIAM guidelines [23]. To hold all sensors in place for the full work day, each sensor was additionally secured with the kinesiology tape.

The sEMG sensors were calibrated using a series of maximum voluntary contractions (MVCs) [24]. For the L1 ES, L3 ES, and L5 MF sensors, a standing isometric back extension exercise was performed with the torso flexed at 45°, manually



Fig. 2. Catch and drag activity with



Fig. 3. Shearing activity with instrumented shearer

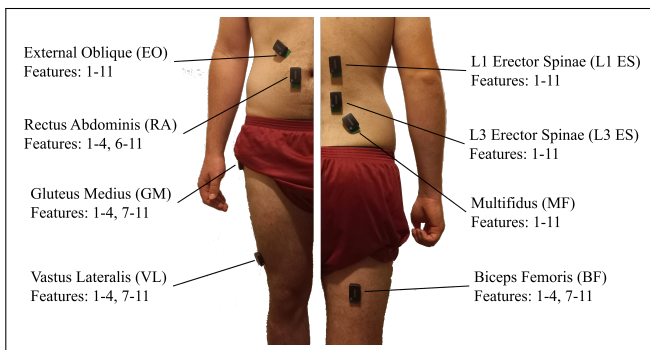


Fig. 4. EMG Sensor placement and features extracted from each sensor. Features - (1) Spectral indices from filtered EMG, (2) Shannon entropy filtered EMG, (3) Auto-correlation filtered EMG, (4) EMG envelope co-contraction ratio, (5) EMG envelope-Joint angle average inner product, (6) EMG envelope mean, (7) EMG envelope variance, (8) EMG envelope skew, (9) EMG envelope kurtosis, (10) EMG envelope auto-correlation, (11) EMG envelope percentiles ( $10^{th}$ ,  $25^{th}$ ,  $50^{th}$ ,  $75^{th}$ ,  $90^{th}$ ,  $100^{th}$ ).

braced by a researcher. For the RA, and EO muscles, a sit-up position was adopted, again with a  $45^\circ$  torso angle with manual bracing provided by a researcher. The subject was instructed to flex initially as if completing the sit-up, then attempt to twist right and left. For the GM muscle the subject was instructed to lie on each side and abduct their hip, against the provided bracing. For the VL and BF muscles the subject was instructed to sit in a chair or on the raised shearing boards, grasp the edge of the surface and sequentially flex and extend each knee against the bracing provided at the ankle. The sEMG activation was normalised to the highest activation for each muscle achieved in any of the exercises.

For the motion capture set-up, the subjects' body-segment lengths were measured. The IMUs were then placed using Xsens provided velcro straps and tight-fitting shirt; with the lower body sensors placed under the shearers' clothes, and additionally secured with the tape. The IMUs were placed according to the Xsens guidelines, with the exception of the lower leg sensors. The recommended placement on the shin-bone was not suitable for sheep shearers, as the inside of both legs are required to hold and manoeuvre sheep during the work. This would expose the IMUs to large external forces and cause significant discomfort for the shearers. Thus, these IMUs were re-positioned to the outside of the leg.

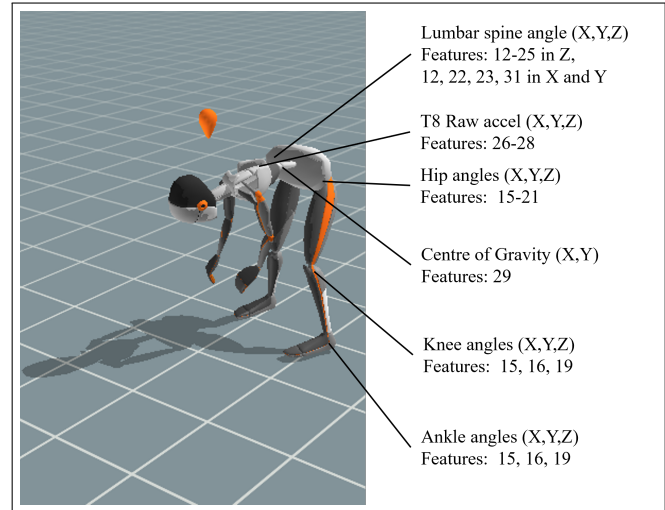


Fig. 5. Shearing motion capture output and features extracted from full body kinematic model. Features - (12) Joint angle (JA) approximate entropy, (13) JA Discrete relative phase, (14) EMG envelope-JA average inner product, (15) JA Mean, (16) JA Variance, (17) JA Skew, (18) JA Kurtosis, (19) JA percentiles ( $10^{th}$ ,  $25^{th}$ ,  $50^{th}$ ,  $75^{th}$ ,  $90^{th}$ ,  $100^{th}$ ), (20) JA Shannon Entropy, (21) Joint angular velocity (JAV) SPARC, (22) Dimensionless jerk from JAV, (23) JAV RMS, (24) Dimensionless jerk from joint angular acceleration (JAA), (25) JAA RMS, (26) JAA Mean, (27) JAA Variance, (28) JAA Percentiles, (29) Position XY Norm, (30) JV Approximate entropy, (31) JA RMS, (32) shearing time.

The motion capture system was calibrated using the procedure recommended by Xsens. This involves the subject standing in a static-pose before naturally walking a short distance forward, before returning to the original location and position. The calibration is achieved by minimising errors over the dynamic movement.

The motion signals and sEMG data were synchronised using the Delsys Trigger Module, with data recording initiated from the commercial Xsens MVN Analyse software.

### C. Protocol

As the study involved instrumenting shearers who were undertaking their regular work activities, the regular shearing rest-work periods were followed. The standard shearing day starts at 7:30am, and shearing takes place in four two-hour sessions (shearing runs), with a one hour lunch break between run 2 and 3, and two 30-minute breaks between runs 1 and 2, and runs 3 and 4. This is illustrated in Figure 6. At the start and end of each run, shearers were asked to rate their perceived level of fatigue using a modified Borg CR10 scale, seen in Appendix B.

Shearers were asked to be at their shearing stand 45 minutes early at the start of the day to allow for sensor placement and calibration of the motion capture and sEMG sensors. Before the start of the other runs, shearers were required at their stand 15 minutes before the run to allow for re-calibration of the motion capture system. The motion capture IMUs were recharged during the lunch break, and were re-placed and re-calibrated in the 15 minutes prior to run 3.

One shearers was instrumented each day for the full working period. Some disruptions occur during regular shearing (e.g. running out of sheep), and these disruptions meant that for three subjects the final shearing run did not occur. A minimum of six hours of data was collected for each subject.



Fig. 6. The rest-work period for sheep shearing: four two-hour shearing sessions ('runs') starting at 7:30am, 10:00am, 1:00pm, and 3:30pm, with two short 30 minute breaks (SB1, and SB2), and a one-hour long break (LB) in the middle of the day. Shearing consists of the catch & drag and shearing activities (shown in Figures 2 and 3) which are repeated throughout each of the runs.

#### D. Pre-processing

Two sets of data were collected from each subject, sEMG sampled at 2148 Hz and motion capture sampled at 60 Hz. The data were synchronised at the collection stage, and a different pre-processing step was required for each to produce the pre-processed data  $S$ , as shown in Figure 1.

1) *Signal Pre-processing for sEMG Measurements:* The raw sEMG signals were filtered with a 2<sup>nd</sup> order Butterworth filter with a pass-band between 20 – 450 Hz. Further to this, the sEMG envelope (ENV) was calculated by rectifying the filtered sEMG signal and low-pass filtering with a cutoff frequency of 6 Hz [25]. This envelope signal was normalised to the maximum value of the envelope from the MVCs collected for that days shearing. This produced the filtered signal and the EMG envelope from each sEMG sensor.

2) *Signal Pre-processing for Kinematic Measurements:* The signal pre-processing for the kinematic data is largely achieved with MVN Analyse software, which produces full-body joint angles, body-segment poses, and velocities. Additionally, joint angular accelerations were derived from this data, and sensor raw accelerations were also extracted. This produced several channels of data from each pair of sensors.

#### E. Periodicity Removal

To allow for the task based sampling, and provide context to the feature extraction, the periodicity in the pre-processed data  $S$  is removed by solving an HAR problem to efficiently identify the different shearing activities.

1) *Human Activity Recognition:* During sheep shearing, there are two main activities that make up the shearing task, (i) the catch and drag phase (CD), where a shearer walks into the holding pen and gains control of a sheep which typically weighs around 70 kg before dragging it backwards to the shearing stand; (ii) and the shearing activity (SH), where the shearer must stand with a fully flexed spine holding the sheep between their legs and manoeuvre the sheep into several positions while removing the wool with the handpiece. The kinematic and sEMG data were segmented into non-overlapping sequences of uneven length corresponding to the shearing (SH) and catch and drag (CD) activities of the task for each sheep, as seen in Figure 6. The combination of two adjacent segments represents the full shearing task which is repeated over the day. Each shearer will shear at a different level of skill, and each will therefore potentially complete a different number of task iterations.

The data was segmented into the shearing, and catch and drag phases of the task using a person-independent sheep

shearing human activity recognition algorithm requiring only two IMUs for implementation developed in [21]. A Hidden Markov Model (HMM) was chosen as it is a simple classification method that is capable of modelling temporal dependencies in the data and is less sensitive to human variations. An HMM based person-independent activity recognition model was trained using half of the data from 6 subjects that was manually labelled from the motion capture video output, seen in Figure 5.

The task was modelled using the two activity states (CD, and SH), the emission probability distributions for each state were assumed to be jointly Gaussian over the features. The initial emission distribution parameters were calculated without temporal information, and the state transition matrix was initialised with uniform probabilities. The performance of the classification can be evaluated by the  $F_1$  score [20], which incorporates both the standard precision and recall metrics. A higher  $F_1$  score indicates a better performance. The  $F_1$  score for identifying the shearing cycle was established from 6 subjects as 96.47%, using a 75/25 train test split [20]. Leave-one-person-out cross validation was also performed, which evaluates the model's ability to generalize to unseen subjects as required for a person-independent model. For this, the model achieved an  $F_1$  score of 95.10%.

2) *Feature Extraction:* To generate the mapping  $\Phi$ , and determine  $\Sigma^k$  for each shearer, data analysis is done for each task iteration. Assuming that there are  $N_s$  numbers of segments corresponding to the number of sheep shorn, and each segmentation has  $N_{p,s}$  measurements, then we compute one point of the feature signal, leading to a time-series data. A search of the literature was performed to identify a broad range of candidate features that are suspected to relate to injury. In addition, a number of common statistical features were also extracted as candidates. The feature signals used in this work are listed in Figure 4 for the sEMG sensors, and Figure 5 for the kinematic features; the calculation and literature basis for each feature can be found in Appendix A.

Most of the features were extracted from the shearing activity. Some features were extracted from the last 5 seconds of the catch and drag activity. It is noted that the last 5 seconds of the catch and drag is important as it represents the drag. This is the most significant manual handling portion of the task, where the highest muscle forces occur. However, none of these features were represented in the top 10.

The candidate features are extracted from the segmented data channels as per Figures 4 and 5 for the sEMG and kinematic features respectively.

### F. Unsupervised Feature Ranking

Candidate features that may contribute to an increased risk of lower back injury in sheep shearing are identified from the literature, and included in the analysis. Additionally, common statistical features are also included to widen the search. Altogether, there are 376 features from the measurements of 33 wearable sensors.

With the candidate features ( $\Sigma^k$ ) extracted,  $\omega(\mathbf{x}_i^j)$  for each feature was calculated using the Mann-Kendall test. Then the  $\Delta J$  indicator ( $\Delta M_i^p$ ) was calculated using (3) for each subject and each feature. For each feature the population trend analysis was performed by calculating the correlation between pairs of subjects using (4) to establish  $\rho(\mathbf{x}_i^p, \mathbf{x}_i^q)$ .

With input from the individual trend analysis and the population trend analysis, the relevance ( $\Delta J$ ) was calculated for each feature using the cost-function presented in (15). The top 10 features by relevance ( $\Delta J$ , prior to accounting for redundancy) for  $\Delta = 0.15$  and  $\Delta = 0.10$  are presented in Table I. It can be seen from Table I, that lower back sEMG features from the shearing activity are dominant.

Then using Algorithm 1, the final  $X^*$  features are iteratively selected. The top 10 selected features are presented in Table II for  $\Delta = 0.15$  and  $\Delta = 0.10$ . It can be seen from Table II, that lower back sEMG features are again dominant, with the addition of sEMG spectral features and some lumbar kinematic features present.

### G. Statistical Analysis of Identified Features

The features proposed by Algorithm 1 were statistically analysed using the procedure in [26] (which evaluated hand-picked features from the same data-set) with a Linear Mixed Model (LMM), shown in (16). The LMM was chosen to model the hierarchical, and unbalanced nature of the observational study. For the  $j^{th}$  subject, the  $m^{th}$  run, and the  $k^{th}$  measurement, the data points in the  $i^{th}$  feature  $x_{jmk}^i$  can be predicted as

$$\hat{x}_{jmk}^i = \beta_0 + \alpha h_j + \delta r + \gamma h_j r + a_j + b_{jm} + e_{jmk} \quad (16)$$

where  $i \in X^*$  is the feature,  $r$  is the run number ( $r \in \{1, 2, 3, 4\}$ ), and  $h_j \in \{0, 1\}$  indicates the  $j^{th}$  shearer's harness usage. Here  $\beta_0$  is the model intercept and  $\delta$  is the run fixed effect. The  $\alpha$  parameter is the harness fixed effect and  $\gamma$  is the harness/run interaction fixed effect. The REs ( $a_j, b_{jm}, e_{jmk}$ ) are Gaussian distribution, characterized by zero mean and their corresponding variances. The random variable  $a_j \sim \mathcal{N}(0, \sigma_j^2)$  is subject dependent and the random variable  $b_{jm} \sim \mathcal{N}(0, \sigma_m^2)$  is run-dependent with variance which is assumed to be same across subjects. The residual  $e_{jmk} \sim \mathcal{N}(0, \sigma_k^2)$  is assumed to have the same variance across subjects and runs. By fitting the data to this model in (16), we can compute these parameters ( $\beta_0, \alpha, \delta, \gamma, \sigma_j^2, \sigma_m^2, \sigma_k^2$ ), and a P value for each fixed effect is returned via a Wald test. The  $\delta$  parameter is the time-of-day effect and so  $P < 0.05$  here indicates that the fitted linear trend over the day is significant across the sampled population.

## V. RESULTS AND DISCUSSION

The purpose of this work is to present a population-trend based data-driven unsupervised feature selection framework

TABLE I  
TOP 10 FEATURES BY RELEVANCE SCORES ( $\Delta J$ )  
FOR  $\Delta = 0.15$  AND  $\Delta = 0.10$   
† INDICATES  $P < 0.05$ , ‡ INDICATES  $P < 0.01$  FOR  $\delta$  PARAMETER

Feature	$\Delta J$ $\Delta=0.15$	Feature	$\Delta J$ $\Delta=0.10$
L3 ES-R env $P_{10\%}$ †	0.4970	L3 ES-R env $P_{10\%}$ ‡	0.4364
L3 ES-R env $P_{25\%}$ ‡	0.4252	L3 ES-R env $P_{25\%}$ ‡	0.3680
L3 ES-R SHEN‡	0.3574	L3 ES-R SHEN‡	0.3359
L5 MF-L env $P_{10\%}$	0.3481	L5 MF-L env $P_{10\%}$	0.2971
L1 ES-R env $P_{10\%}$ †	0.3279	L1 ES-R env $P_{10\%}$ †	0.2812
L3 ES-R env $P_{50\%}$ ‡	0.3037	L5 MF-L SHEN	0.2665
L5 MF-R env $P_{10\%}$ †	0.3006	L5 MF-R env $P_{10\%}$ †	0.2620
L5 MF-L env $P_{25\%}$	0.2881	L3 ES-R env $P_{50\%}$ ‡	0.2503
L5 MF-L SHEN	0.2774	L5 MF-L env $P_{25\%}$	0.2476
L5 MF-R SHEN	0.2546	L5 MF-R SHEN	0.2373

TABLE II  
TOP 10 SELECTED FEATURES ( $X^*$ ) USING THE PROPOSED METHOD  
FOR  $\Delta = 0.15$  AND  $\Delta = 0.10$   
† INDICATES  $P < 0.05$ , ‡ INDICATES  $P < 0.01$  FOR  $\delta$  PARAMETER

Feature	Score $\Delta=0.15$	Feature	Score $\Delta=0.10$
L3 ES-R env $P_{10\%}$ †	0.4970	L3 ES-R env $P_{10\%}$ ‡	0.4364
L5 MF-L env $P_{10\%}$	0.2553	L5 MF-L SHEN	0.2300
L1 ES-R $\mu$ Freq.	0.1621	L1 ES-R Auto-C	0.1632
L5 MF-R SHEN	0.1535	L5 MF-R env $P_{10\%}$ †	0.1292
Pel-T8 flex-vel RMS†	0.1375	L3 ES-R $\mu$ Freq.‡	0.1179
L3 ES-R $\mu$ Freq.‡	0.1329	Pel-T8 Flex. RMS	0.0983
RA-L env $P_{10\%}$	0.1174	RA-L EMG Auto-C	0.0956
RA-L env Auto-C†	0.1017	L1 ES-R env $P_{10\%}$ †	0.0884
L1 ES-R env $P_{10\%}$ †	0.0948	EO-R SHEN	0.0806
Pel-T8 flex. Auto-C	0.0760	Pel-T8 Flex. Auto-C	0.0737

to identify key indicators of lower back injury for repetitive work where injury risk is known to worsen with repeated exposure to stresses. This is achieved by first segmenting the data into separate tasks by solving a HAR problem, and then selecting features using a filter based unsupervised approach. The feature selection method incorporates domain knowledge of injury aetiology to determine feature relevance, and a measure of feature orthogonality to eliminate redundancy, as described in Section III.

The results of the feature relevance module are displayed in Table I, presenting  $\Delta J$  without considering the overlapping information between features. These are shown for  $\Delta$  values of 0.1, and 0.15. Table II presents the selected features  $X^*$  after the iterative selection procedure to minimise redundancy. The top selected features were analysed statistically, with the significance of the trend parameter ( $\delta$ ) annotated in Tables I and II. In order to evaluate the effectiveness of the proposed feature selection technique, we first discuss the importance of the selected features from the existing literature, followed by the discussion of information overlapping. Finally, we use the self-reported fatigue levels collected after each run as a “ground truth” to compare the proposed method with the existing 5 methods in literature in terms of consistency with respect to the self-report fatigue level.

### A. Validation against the Literature

Table I identifies features selected by the proposed method with the highest relevance scores ( $\Delta J$ ), which occur before accounting for redundancy. Six of the ten features for both reported  $\Delta$  values are found to have a significant trend across the population in the time-of-day variable ( $\delta$ ). In Table II, five and four of the ten features have a significant trend over the day for  $\Delta$  values of 0.1 and 0.15 respectively after feature redundancy is considered.

The 10<sup>th</sup> percentile of the sEMG envelope magnitude at the L3 ES muscle is identified as the most suitable feature. The ES muscles are responsible for lumbar extension, and are very relevant in stooped work. This trend is more pronounced for the lower percentiles, compared to the higher percentiles. In sheep shearing, higher level contractions are required to manoeuvre the sheep, or counter-act disturbances from a struggling sheep, while less muscle activation is required to counter gravity. A decrease in lower level muscle activity here likely indicates that some load shifting is occurring to other muscles or passive tissues for the consistent low level contractions, while intermittent higher level active contractions are still required. By itself, this load sharing could indicate muscle fatigue, and as this feature decreases, this would indicate increased stresses on passive tissues or deeper muscles that are not observed with sEMG sensors; evidence of this is seen in [26]. Both of these processes have evidence to suggest they lead towards injury as pointed out in [27]. This indicates our result is consistent with previous evidence around possible indications of lower back injury. Example data of this feature is shown in Figure 7 from an expert, intermediate, and beginner shearer. The similarities between subjects, and trend over the day can be seen. There is a noted ‘warm-up’ period where the feature initially increases, before starting to decrease, and periods of recovery coinciding with the rest-periods in shearing can be seen for some subjects, which aligns with expected recoveries in fatigue and visco-elastic creep effects [5], [7].

Consistent with the literature, the multifidus (MF) muscle is also identified in Tables I and II. It was pointed out in [28] that it is a key muscle in the development of lower back pain, attracting a significant clinical focus [29]. The MF muscle is known to be responsible for stabilizing the lumbar spine, and is shown to be affected by prolonged static lumbar flexion [5]. A pronounced trend was identified for this muscle bilaterally, with the left side having higher relevance. These key features are consistent with the literature, showing that the proposed method, even though it is data dominant, can contribute to the understanding of lower back injuries. Example data of this feature is also presented in Figure 8, again with similarities and decreasing trend present. Periods of ‘warm-up’ and ‘recovery’ can also be seen in this feature.

The mean frequency of the L1 and L3 ES muscles also features in Table II. Mean frequency in the sEMG signal is commonly used to measure muscle fatigue, and muscle fatigue is well known to lead to injury [30]. This is interesting as it occurs despite the flexion-relaxation phenomenon being observed in the task [31].

Interestingly, our results suggest some statistical features that are not commonly discussed in the literature around low back injuries feature should be used to characterize low back injuries for sheep shearing. For example, Shannon entropy of

sEMG features is ranked 3<sup>rd</sup> in Table I. Shannon entropy is an indication of the level of information encoded in the signal, and a reduction would indicate a loss of complexity of neuromuscular control. This might be related to dysfunction of muscles. For example, in [32], Shannon entropy is found to differentiate between healthy and LBP individuals, and in [33] it was found to be influenced by muscle fatigue. Shannon entropy at multiple time-scales is investigated in [34], showing its ability of classifying healthy individuals from those with neuromuscular disorders.

The results in Table I provide evidence that sEMG data should play an important role in wearable sensors for the prevention of occupational lower back injuries in the context of stooped work. In particular, features extracted from the shearing part of the segmented data also outperform the features from the catch and drag phase of the task. This is consistent with literature that highlights the importance of prolonged and repetitive poor postures as an avenue to injury and poor performance [5], [11], [35]. However, it was also noted that sEMG features that measured low level muscle contractions performed well, and the shearing task contains lower level muscle contractions overall when compared to the catch and drag activity.

### B. Validation against Existing Unsupervised Feature Ranking Techniques using Self-Reported Fatigue

As the data was collected from shearers for one day, there are no observed injuries and therefore no proper ground truth available. This makes an evaluation of the proposed method difficult. Self-reported perceived fatigue data (annotated 1-10 scale) was collected from the subjects at the start and end of each shearing run. However, this is not a proper ground truth as this is perception based, and each individual may perceive the ‘same amount’ of fatigue differently. Additionally, while it is known that fatigue influences lower back injury, self-reported fatigue has not been directly linked to back injury risk; e.g. a change in reported fatigue from 5 to 6 is likely to increase injury risk, but by an unknown amount.

It is assumed that the self-reported fatigue is related to the ground-truth for injury risk. The feature selection methods are then evaluated based on this self-reported fatigue as a pseudo-ground-truth. More precisely, self-reported fatigue labels are used to train an ordinal classifier with sets of selected features. The performance of the feature selection will determine the accuracy of this classifier. This type of evaluation has the benefit of utilising the relative information of the self-reported fatigue scores, and better aligns with domain knowledge around self-reported fatigue and injury risk.

1) *Model for Evaluation:* The model used to evaluate the selected feature is seen in (17). It takes the selected  $X^*$  features as input, and predicts the self-reported fatigue score  $Y_k \in \{1, 2, \dots, 10\}$  using a nonlinear mapping  $f(\cdot)$  to have the predicted fatigue score  $\hat{Y}_k$ ,

$$\hat{Y}_k = f(X^*), \quad (17)$$

where the nonlinear mapping  $f$  is to be approximated from the measured data.

It is noted that our proposed method selected the representative features shown in Table II, the evaluation procedure in (17) is only used to show that the selected features are



consistent with the self-reported fatigue. Therefore, we do not use other powerful machine learning methods such as deep learning [12], to learn latent features from the datasets. Instead, we employed the widely used random forest, modified as an ordinal classifier (RFOC) to find the approximate mapping.

A random forest classifier was developed to incorporate the ordered nature of the categories using the method in [36]. Given the ordered nature of the categories, the mean absolute error (MAE) was used to score the classifier's performance,

$$MAE = \frac{\sum_{k=1}^{N_k} |Y_k - \hat{Y}_k|}{N_k}, \quad (18)$$

where,  $Y_k$  is the reported fatigue score,  $\hat{Y}_k$  is the predicted value, and  $N_k$  is the number of segments. A small MAE value indicates that the selected features match well with the self-reported fatigue. Therefore, such an index will be used to evaluate various unsupervised feature ranking techniques.

A subset of 75% of the data was randomly sampled for model training, with the remaining 25% hold-out for testing. In order to compare the best model for each set of features, a grid search over the random forest hyper-parameters was performed with 5-fold cross-validation over the training dataset, using available high performance computing resources [37]. The hyper-parameters with the best average MAE were then used to train another RFOC using all of the training data.

In order to verify the effectiveness of the proposed framework, it is compared with 5 existing unsupervised feature ranking techniques using MAE score using the test data-set, as reported in Table III for different numbers of selected features  $N_f^* = 1, 2, \dots, 10$ . It is noted that unsupervised methods usually define cost functions to learn local or global structures of the datasets, and then select representative features. Here, we compare the developed method with the following 5 popular unsupervised feature selections methods.

Distance Correlation (DC) [38]

The Distance Correlation measure evaluates the similarity of two random variables, that is capable of capturing non-linear dependencies. This can be used to rank features based on their similarity across the family of processes.

Laplacian Score (LS) [39]

The Laplacian Score is an unsupervised feature selection method that ranks features based on their ability to preserve the local manifold structure in the data.

Strict Monotonicity (SM) [14]

Strict Monotonicity ranks features based on how well they exhibit a strict monotonically increasing or decreasing trend across a family of processes.

SPEC [40]

This method presents a framework for selecting features based on how consistent the feature is with the structure of the spectral graph created from the data.

Spearman's Rho Correlation (SRC) [41]

The Spearman's Rho correlation is a non-parametric assessment of statistical dependence between two variables, measuring both linear and non-linear relationships using a monotonic function. This can be used to rank features based on their similarity across the family of processes.

The comparison results are presented in Table III, the mean and standard deviation are calculated from the absolute error

across the data-points in the test-set. It can be seen from the table that the proposed weak-monotonicity based method has small prediction errors, with the minimum MAE values in almost all cases compared to other unsupervised techniques. There are three techniques choosing the best initial feature, and the distance correlation (DC) based method outperforms when exactly 8 features are chosen. This supports the claim that when an underlying monotonic process is present, feature selection techniques that exploit this knowledge will outperform in the unsupervised case. Additionally, the relaxation to weak monotonicity is able to capture important features missed by the SM method and improve performance. The averaged MAE scores for 10 cases of the proposed method outperforms the other 5 methods in literature, showing effectiveness of the proposed method.

### C. Discussion

This work proposes a framework to utilize a common trend in a population to provide guidance in unsupervised feature ranking techniques. Hence, this technique is not limited to lower back pain, or sheep shearing, but any family of processes that have a common trend. As a special class of cases, this technique is applicable to repetitive activities where injuries are typically not brought on by acute musculoskeletal stress, and the injury risk is known to increase over time.

It is understood that recent advances of artificial intelligence demonstrate the promising performance of machine learning methods for data regression analysis, especially deep learning methods [12], [42]. However, these methods typically require a large amount of training datasets, and the learned features are a combination or transformation from the original data, making the results hard to interpret. In particular, if we want to find a minimal number of wearable sensors to capture the injury information, the identified features should be measurable from sensors. Thus, deep learning methods are difficult to satisfy the requirements in this application. As demonstrated in Table II, our proposed method efficiently identified representative features from a large set of feature candidates. More importantly, the identified features have clear physical meaning with corresponding muscle locations, which is helpful to provide explicit implication to practitioners.

The framework has limitations in that it is not a general method for feature selection, and relies on the assumption of an underlying shared monotonic process. The calculation of weak monotonicity also does not allow for the use of categorical features. The accuracy of the feature selection also relies on the success of the activity recognition.

In the context of a wearable device, identifying important features is not the end of the discussion, because features do not exhibit a one-to-one mapping with the number of sensors required. For example, to extract joint angles and velocities from IMUs a sensor on the two adjacent body segments is required. But once the signal is derived, many features can be extracted from the same signal. It is also worth mentioning that in this case, there are two sensors required for the HAR process; IMUs on the pelvis and ribcage. The inclusion of these sensors increases the minimum complexity of a wearable device, but also allows for kinematic features to be extracted from the lumbar flexion angle. Any chosen human activity recognition method should consider the

number of sensors required, as this can impact the practicality in the wearable device context. Table II indicates that it may be possible to further reduce sensors if the activities could be recognised using only the sEMG signals. This has been previously demonstrated in recognising some repetitive activities [43], [44].

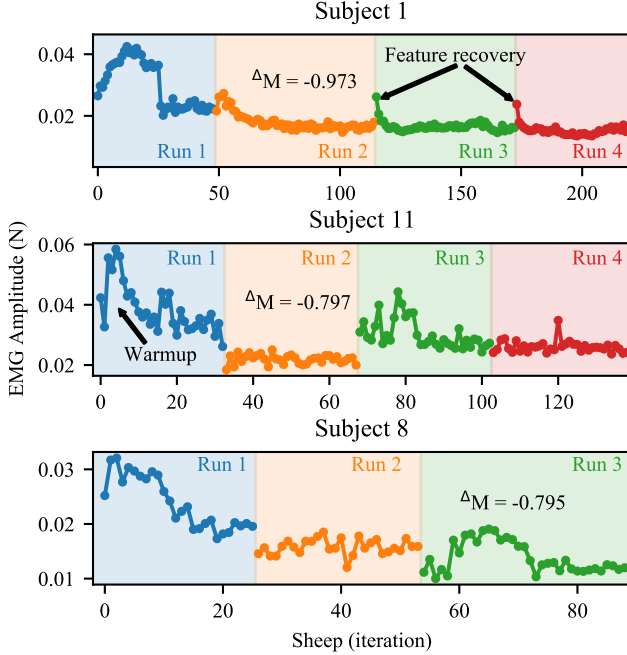


Fig. 7. Illustration of top selected feature - L3 ES Right Envelope 10th Percentile

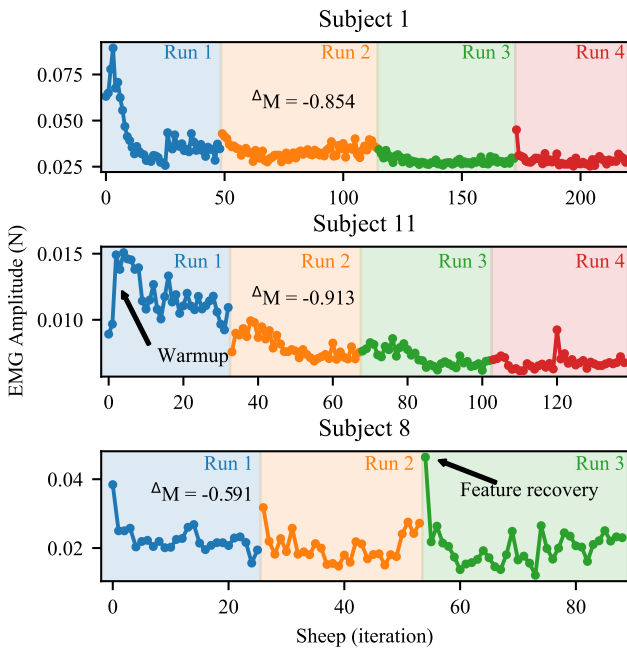


Fig. 8. Illustration of second selected feature - L5 MF Left Envelope 10th Percentile

## VI. CONCLUSIONS

This paper presents a novel population-trend based unsupervised framework to rank the key features that contribute to lower back injury in repetitive stooped work occupations from a large amount of collected data, with application in on-site sheep shearing. Overall, the proposed feature selection method allows for identifying indicators of injury risk by exploiting specific problem domain knowledge in terms of the population, which provides guidance of unsupervised feature ranking and improves the effectiveness of feature ranking. With the practical consideration of reducing the number of wearable sensors, the redundancy between features is also considered as a part of the performance index. The identified key features are consistent with the understanding of back injuries reported in literature, showing the effectiveness of the proposed technique. Furthermore, comparing with five popular unsupervised feature selection techniques in literature, the proposed technique shows better performance on the regression analysis of sheep shearers' self-reported fatigue levels.

## ACKNOWLEDGMENT

This project was funded by Australian Wool Innovation (AWI) research contract, project number ON-00639.

## REFERENCES

- [1] Z. Luo and Y. Yu, "Wearable stooping-assist device in reducing risk of low back disorders during stooped work," *IEEE ICMA 2013*, no. Ld, pp. 230–236, 2013.
- [2] J. Hartvigsen *et al.*, "What low back pain is and why we need to pay attention," *The Lancet*, vol. 391, no. 10137, pp. 2356–2367, 2018.
- [3] W. S. Marras *et al.*, "Biomechanical risk factors for occupationally related low back disorders," *Ergonomics*, vol. 38, no. 2, pp. 377–410, 1995.
- [4] F. A. Fathallah, "Musculoskeletal disorders in labor-intensive agriculture," *Applied Ergonomics*, vol. 41, no. 6, pp. 738–743, 2010.
- [5] M. Solomonow *et al.*, "Flexion-relaxation response to static lumbar flexion in males and females," *Clinical Biomechanics*, vol. 18, no. 4, pp. 273–279, 2003.
- [6] S. McGill, "The biomechanics of low back injury: Implications on current practice in industry and the clinic," *Journal of Biomechanics*, vol. 30, no. 5, pp. 465–475, 1997.
- [7] W. S. Marras, "Occupational low back disorder causation and control," *Ergonomics*, vol. 43, no. 7, pp. 880–902, 2000.
- [8] S. McGill, "Opinions on the links between back pain and motor control: the disconnect between clinical practice and research," in *Spinal Control*. Elsevier, 2013, pp. 75–87.
- [9] P. Hodges *et al.*, "Motor control of the spine and changes in pain: debate about the extrapolation from research observations of motor control strategies to effective treatments for back pain," Churchill Livingstone, 2013.
- [10] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16–28, 2014.
- [11] L. Lu *et al.*, "Evaluating rehabilitation progress using motion features identified by machine learning," *IEEE Transactions on Biomedical Engineering*, pp. 1–1, 2020.
- [12] Z. Chen *et al.*, "An attention based CNN-LSTM approach for sleep-wake detection with heterogeneous sensors," *IEEE Journal of Biomedical and Health Informatics*, 2020.
- [13] A. Yuan *et al.*, "Convex Non-Negative Matrix Factorization With Adaptive Graph for Unsupervised Feature Selection," *IEEE Transactions on Cybernetics*, pp. 1–13, 2020.
- [14] P. Baraldi *et al.*, "Differential evolution-based multi-objective optimization for the definition of a health indicator for fault diagnostics and prognostics," *Mechanical Systems and Signal Processing*, vol. 102, pp. 382–400, 2018.
- [15] P. Pal *et al.*, "A review of risk factors associated with low back injury in wool harvesting," *Journal of Occupational Health and Safety - Australia and New Zealand*, vol. 24, no. 5, pp. 435–453, 2008.
- [16] P. Baraldi, G. Bonfanti, and E. Zio, "Differential evolution-based multi-objective optimization for the definition of a health indicator for fault diagnostics and prognostics," *Mechanical Systems and Signal Processing*, vol. 102, pp. 382–400, 2018.

TABLE III  
COMPARISON OF FEATURE SELECTION METHODS WITH UP TO 10 SELECTED FEATURES. BOLD TYPEFACE INDICATES THE LEAST ERROR.

No. Feat.	Mean Absolute Error $\pm$ Standard Deviation					
	WM ( $\Delta=0.15$ )	DC	LS	SM	SPEC	SRC
1	<b>1.8854 <math>\pm</math> 1.7577</b>	<b>1.8854 <math>\pm</math> 1.7577</b>	1.9785 $\pm$ 1.8608	<b>1.8854 <math>\pm</math> 1.8413</b>	2.2745 $\pm$ 1.7772	1.9737 $\pm$ 1.7332
2	<b>1.0072 <math>\pm</math> 1.5549</b>	1.5967 $\pm$ 1.7209	1.3986 $\pm$ 1.6323	1.5274 $\pm$ 1.7779	2.0358 $\pm$ 1.8059	1.4200 $\pm$ 1.5294
3	<b>0.7900 <math>\pm</math> 1.3710</b>	1.1002 $\pm$ 1.5416	1.3389 $\pm$ 1.5874	0.9117 $\pm$ 1.5431	1.5609 $\pm$ 1.7744	1.2459 $\pm$ 1.5554
4	<b>0.7184 <math>\pm</math> 1.2803</b>	0.7494 $\pm$ 1.2500	1.2697 $\pm$ 1.6014	0.7924 $\pm$ 1.4809	1.0143 $\pm$ 1.5073	1.1002 $\pm$ 1.4832
5	<b>0.5346 <math>\pm</math> 1.1460</b>	0.7017 $\pm$ 1.2756	1.2983 $\pm$ 1.5276	0.7184 $\pm$ 1.3910	0.6611 $\pm$ 1.2492	0.9212 $\pm$ 1.4213
6	<b>0.5394 <math>\pm</math> 1.1726</b>	0.6874 $\pm$ 1.2184	1.1623 $\pm$ 1.5502	0.6516 $\pm$ 1.3965	0.6444 $\pm$ 1.2125	0.7470 $\pm$ 1.2619
7	<b>0.3914 <math>\pm</math> 0.9841</b>	0.4654 $\pm$ 1.0157	1.1217 $\pm$ 1.5159	0.5251 $\pm$ 1.1811	0.6444 $\pm$ 1.2494	0.7136 $\pm$ 1.2336
8	0.4845 $\pm$ 1.1485	<b>0.4726 <math>\pm</math> 1.0159</b>	1.1909 $\pm$ 1.4904	0.5728 $\pm$ 1.2479	0.6038 $\pm$ 1.1439	0.7017 $\pm$ 1.2298
9	<b>0.3866 <math>\pm</math> 0.9515</b>	0.4773 $\pm$ 1.0090	1.0430 $\pm$ 1.4469	0.4177 $\pm$ 0.9547	0.5609 $\pm$ 1.0827	0.6325 $\pm$ 1.1840
10	<b>0.4272 <math>\pm</math> 0.9947</b>	0.4821 $\pm$ 1.0067	0.8854 $\pm$ 1.3613	0.4320 $\pm$ 1.0510	0.6014 $\pm$ 1.1524	0.5179 $\pm$ 1.0688
Mean	<b>0.7165 <math>\pm</math> 1.2608</b>	0.8618 $\pm$ 1.3114	1.2687 $\pm$ 1.5625	1.3114 $\pm$ 1.4135	1.0602 $\pm$ 1.4227	0.9974 $\pm$ 1.3839

- [17] Sheng Yue *et al.*, "Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series," *Journal of Hydrology*, vol. 259, p. 254-271, 2002.
- [18] N. J. Nagelkerke, "A note on a general definition of the coefficient of determination," *Biometrika*, vol. 78, no. 3, pp. 691-692, 1991.
- [19] P. W. Hodges, "Adaptation and rehabilitation from motoneurons to motor cortex and behaviour," in *Spinal Control: The Rehabilitation of Back Pain*, 1st ed., P. W. Hodges, J. Cholewicki, and J. H. van Dieen, Eds. Churchill Livingstone, 2013, ch. 6, pp. 59-74.
- [20] A. Bulling *et al.*, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys*, vol. 46, no. 3, pp. 1-33, 2014.
- [21] M. Robinson *et al.*, "Enabling context aware data analysis for long-duration repetitive stooped work through human activity recognition in sheep shearing," *2020 Annu Int Conf IEEE Eng Med Biol Soc.*, pp. 87-90, 2020.
- [22] Ó. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192-1209, 2013.
- [23] H. J. Hermens *et al.*, "Development of recommendations for SEMG sensors and sensor placement procedures," *J Electromyogr Kinesiol*, vol. 10, no. 5, pp. 361-374, 2000.
- [24] S. M. McGill *et al.*, "Muscle activity and spine load during pulling exercises: Influence of stable and labile contact surfaces and technique coaching," *J Electromyogr Kinesiol*, vol. 24, no. 5, pp. 652-665, 2014.
- [25] J. Cholewicki and S. M. McGill, "Mechanical stability of the in vivo lumbar spine: Implications for injury and chronic low back pain," *Clinical Biomechanics*, vol. 11, no. 1, pp. 1-15, 1996.
- [26] M. Robinson *et al.*, "Evaluation of ceiling-supported back harnesses in preventing injury in sheep shearing," *2021 Annu Int Conf IEEE Eng Med Biol Soc.*, pp. 6957-6961, 2021.
- [27] M. Solomonow, "Neuromuscular manifestations of viscoelastic tissue degradation following high and low risk repetitive lumbar flexion," *J Electromyogr Kinesiol*, vol. 22, no. 2, pp. 155-175, 2012.
- [28] M. D. Freeman *et al.*, "The Role of the Lumbar Multifidus in Chronic Low Back Pain: A Review," *PM and R*, vol. 2, no. 2, pp. 142-146, 2010.
- [29] D. A. MacDonald, G. Lorimer Moseley, and P. W. Hodges, "The lumbar multifidus: Does the evidence support clinical beliefs?" *Manual Therapy*, vol. 11, no. 4, pp. 254-263, 2006.
- [30] G. Shin *et al.*, "Creep and fatigue development in the low back in static flexion," *Spine*, vol. 34, no. 17, pp. 1873-1878, 2009.
- [31] M. Robinson *et al.*, "Effects of varying the rest period on the onset angle of lumbar flexion-relaxation in simulated sheep shearing: a preliminary study," in *IEEE Int Conf Rehabil Robot*, 2019, pp. 83-88.
- [32] M. Kaufman *et al.*, "Entropy of electromyography time series," *Physica A: Statistical Mechanics and its Applications*, vol. 386, no. 2, pp. 698-707, 2007.
- [33] J. G. Cashaback *et al.*, "Muscle fatigue and contraction intensity modulates the complexity of surface electromyography," *J Electromyogr Kinesiol*, vol. 23, no. 1, pp. 78-83, 2013.
- [34] R. Istenič *et al.*, "Multiscale entropy-based approach to automated surface EMG classification of neuromuscular disorders," *Medical and Biological Engineering and Computing*, vol. 48, no. 8, pp. 773-781, 2010.
- [35] M. Solomonow, B. H. Zhou, Y. Lu, and K. B. King, "Acute repetitive lumbar syndrome: A multi-component insight into the disorder," *Journal of Bodywork & Movement Therapies*, vol. 16, no. 2, pp. 134-147, 2012.
- [36] E. Frank and M. Hall, "A Simple Approach to Ordinal Classification," in *European Conference on Machine Learning 2001*, 2001, pp. 145-156.
- [37] L. Lafayette *et al.*, "Spartan Performance and Flexibility: An HPC-Cloud Chimera," in *OpenStack Summit, Barcelona*, 2016.
- [38] D. Edelmann *et al.*, "An updated literature review of distance correlation and its applications to time series," *International Statistical Review*, vol. 87, no. 2, pp. 237-262, 2019.
- [39] X. He *et al.*, "Laplacian Score for feature selection," *Advances in Neural Information Processing Systems*, pp. 507-514, 2005.
- [40] Z. Zhao and H. Liu, "Spectral feature selection for supervised and unsupervised learning," in *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 1151-1157.
- [41] M. Pedersen *et al.*, "On the relationship between instantaneous phase synchrony and correlation-based sliding windows for time-resolved fmri connectivity analysis," *Neuroimage*, vol. 181, pp. 85-94, 2018.
- [42] Y. LeCun *et al.*, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [43] G. Biagetti *et al.*, "Analysis of the EMG Signal during Cyclic Movements Using Multicomponent AM-FM Decomposition," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1672-1681, 2015.
- [44] S. S. Bangaru *et al.*, "Data quality and reliability assessment of wearable emg and IMU sensor for construction activity recognition," *Sensors (Switzerland)*, vol. 20, no. 18, pp. 1-24, 2020.
- [45] M. González-Izal *et al.*, "EMG spectral indices and muscle power fatigue during dynamic contractions," *J Electromyogr Kinesiol*, vol. 20, no. 2, pp. 233-240, 2010.
- [46] S. P. Silfies *et al.*, "Trunk muscle recruitment patterns in specific chronic low back pain populations," *Clinical Biomechanics*, vol. 20, no. 5, pp. 465-473, 2005.
- [47] B. Hu and X. Ning, "The influence of lumbar extensor muscle fatigue on lumbar-pelvic coordination during weightlifting," *Ergonomics*, vol. 58, no. 8, pp. 1424-1432, 2015.
- [48] S. Balasubramanian *et al.*, "On the analysis of movement smoothness," *Journal of NeuroEngineering and Rehabilitation*, vol. 12, no. 1, pp. 1-11, 2015.
- [49] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 6, pp. 2297-2301, 1991.
- [50] K. P. Granata and S. A. England, "Stability of dynamic trunk movement," *Spine*, vol. 31, no. 10, pp. 1-13, 2006.
- [51] J. J. Liddy *et al.*, "The efficacy of the Microsoft KinectTM to assess human bimanual coordination," *Behavior Research Methods*, vol. 49, no. 3, pp. 1030-1047, 2017.



## APPENDIX A CANDIDATE FEATURES

1) *sEMG signals*: There are a few features coming from sEMG measurements.

**A: Spectral indices** For measured sEMG signal  $x_{sEMG}(k), k = 0, \dots, N_{p,s} - 1$ , by using appropriate Fourier Transform, we can obtain its frequency components with the frequency range  $[f_0, f_1]$ . We denote the power density spectrum of this signal is  $P_{sEMG}(f)$ , then the  $s^{th}$  spectral moment is defined as:

$$M_s = \int_{f_0}^{f_1} f^s \cdot P_{sEMG}(f) df, \quad s = \dots, -2, -1, 0, 1, 2, \dots \quad (19)$$

The mean frequency is defined as  $F_{mean} := \frac{M_1}{M_0}$ , which is the ratio between  $M_1$  and  $M_0$ . The  $FI_{nmsk} := \frac{M_{-1}}{M_5}$  is ratio of  $M_{-1}$  and  $M_5$ . The mean frequency is commonly used as a measurement of muscle fatigue, with the ratio of higher spectral indices developed to better indicate fatigue in dynamic contractions [45].

**B: Muscle co-contraction ratios** Muscle co-contraction ratios were also calculated for agonist-antagonist muscle pairs. These ratios have been used to assess trunk muscle synergies, and are altered under lower back pain [46]. This is defined as the average ratio of the sEMG envelope for the two muscles over the segment. For the amplitude of the EMG envelope for a muscle  $\eta^a(k) > 0, k = 0, \dots, N_{p,s} - 1$  with an antagonist muscle  $\eta^b(k) > 0$  we define the CCR:

$$CCR = \frac{1}{N_{p,s}} \sum_{k=0}^{N_{p,s}-1} \frac{\eta^a(k)}{\eta^b(k)}. \quad (20)$$

2) *Kinematic measurements*: There are several features related to Kinematic measurements.

**A: Mean lumbar-hip Ratio** During forward bending, both the hip and lumbar spine can contribute to the movement. The relative contribution of these joints is known as the lumbopelvic rhythm, and this is known to vary between individuals, and with pain and fatigue, making it potentially useful as an indicator of injury in stooped work [47].

For measured the angle of spine and angle of hip, denoted as  $\theta^{spine}(k), \theta^{hip}(k), k = 0, \dots, N_{p,s} - 1$  respectively, we define LHR as

$$LHR = \frac{1}{N_{p,s}} \sum_{k=0}^{N_{p,s}-1} \frac{\theta^{spine}(k)}{\theta^{hip}(k)}. \quad (21)$$

**B: Modified Spectral Arc Length (SPARC)** The modified measure of spectral arc length with an adaptive choice of cutoff frequency  $\omega_c$  are used to describe the smoothness of the movement indicative of good neuromuscular control, which is not biased by the length of the movement [48]. For a given joint angle  $\theta(k)$  and its angular velocity  $v(k)$ , where  $k = 0, \dots, N_{p,s} - 1$ , the appropriate Fourier Transform is  $V(\omega)$  with the cut-off frequency  $\omega_c$ .

In this context, it is important as each sheep may require a different length of time to shear for a single shearer, and each

shearer will shear sheep at different speeds. Hence the cut-off frequency will be adaptive based on each shearer's speed.

$$SPARC = - \int_0^{\omega_c} \left[ \left( \frac{1}{\omega_c} \right)^2 + \left( \frac{dV(\omega)}{d\omega} \right) \right]^{\frac{1}{2}} d\omega, \quad (22)$$

where the cut-off frequency is selected as

$$\omega_c \triangleq \min \{ \omega_c^{max}, \min \{ \omega, |V(r)| < \bar{V} \ \forall \ r > \omega \} \}, \quad (23)$$

where  $\omega_c^{max}$  is a given bound for the cut-off frequency. Here  $\bar{V}$  is the threshold on the magnitude of the Fourier Transform at the frequency when the magnitude permanently drops below this threshold.

**C: Dimensionless Jerk** Another smoothness measure is dimensionless jerk [48]. For a movement starting at  $t_1$  and finishing at  $t_2$  with the maximum angular velocity  $v_{peak}$  during this period, it is defined as joint

$$DLJ = \frac{(t_2 - t_1)^5}{v_{peak}^2} \int_{t_1}^{t_2} \left| \frac{d^2v}{dt^2} \right|^2 dt, \quad (24)$$

where  $v(\cdot)$  is the smooth or smooth approximations of the angular velocity of the movement.

**D: Approximate entropy (ApEn)**

An approximation of Kolmogorov-Smirnov (KS) entropy (the sum of all positive Lyapunov exponents) for a finite duration signal that is practical for short duration data [49]. A related measure, the calculation of the largest positive Lyapunov exponent has been used as a measure of spinal stability for repetitive dynamic movements [50], and is expected to relate to injury risk.

This calculation is more involved and the details can be found in [49, p. 2298-2299].

**E: Discrete relative phase** This is the time (samples) delay between the same event in two synergistic joint angles (here the peak angle during shearing is used) [51]. For example, for two joint angles  $\theta^a(k), \theta^b(k), k = 0, 1, 2, \dots, N_{p,s} - 1$ , it is defined as

$$DRP_{a,b} = \frac{1}{N_{p,s}} \left( \arg \max_{k=0,1,2,\dots,N_{p,s}-1} \theta^a(k) - \arg \max_{k=0,1,\dots,N_{p,s}-1} \theta^b(k) \right), \quad (25)$$

which can be either positive or negative.

3) *Combined sEMG and kinematic measurements (lumbar spine angle and lumbar muscles)*: The interaction between the lumbar kinematics and lumbar muscles in flexion has been shown to have potential implications for injury [5]. In the lab the flexion-relaxation phenomenon has been extensively studied. This is difficult to measure in working environments due to numerous confounding factors, such as movement speed, and foot placement and orientation [47]. As the EMG-off angle moves further into flexion, this measure would be expected to increase.

This is applied to a combination of a joint angle  $x^a(k), k = 0, \dots, N_{p,s} - 1$  of lumbar spine and sEMG envelope  $\eta^b(k), k = 0, \dots, N_{p,s} - 1$  measured for lumbar muscles.

$$Cross_{k,m} = \frac{1}{N_p} \sum_{k=0}^{N_{p,s}-1} x^a(k) \eta^b(k), \quad (26)$$

where appropriate re-sampling techniques are used to ensure these two signals have the same sampling rate.

4) *Statistics of signals*: Other than a measured signal, statistics of the signal can be also used as features. For generality, the following statistic features of a finite duration signal  $x(k) \in \mathcal{R}, k = 0, \dots, N_{p,s} - 1$ , (the kinematic, or the filtered sEMG signal, or the sEMG envelope signal) are defined. The mean and variance are represented as  $\bar{x}$ , and  $\sigma$  respectively, and the standard definitions are not reprinted here.

#### A: Skew

Fisher-Pearson coefficient of skewness of the signal  $x(k), k = 0, \dots, N_{p,s} - 1$  is defined

$$b = \frac{\frac{1}{N_{p,s}} \sum_{k=0}^{N_{p,s}-1} (x(k) - \bar{x})^3}{\sigma^{\frac{3}{2}}}. \quad (27)$$

#### B: Kurtosis

Fisher kurtosis is defined as follows  $\kappa = \frac{\frac{1}{N_{p,s}} \sum_{k=0}^{N_{p,s}-1} (x(k) - \bar{x})^4}{\sigma^4} - 3$ .

C: Auto-correlation Auto-correlation is the correlation of a sequence with the same sequence delayed by a time. In this case, this is calculated with a one-step delay:

$$r = \rho(x_k, x_{k-1}), \quad (28)$$

where  $\rho(\cdot, \cdot)$  is defined in (4).

#### D: Percentiles

A notion of percentile can be used to characterize a random signal satisfying standard normal distribution. A percentile is a value in the distribution that holds a specified percentage of the population below it. The general definition is that the  $p^{th}$  percentile is the value that holds  $p\%$  of the values below it. The  $10^{th}$ ,  $25^{th}$ ,  $50^{th}$ ,  $75^{th}$ ,  $90^{th}$ , and  $100^{th}$  (max) percentiles are calculated.

#### E: RMS and Norm

$$RMS = \frac{1}{N_{p,s}} \sqrt{\sum_{k=0}^{N_{p,s}-1} x(k)^2}. \quad (29)$$

For a vector of signals  $\{\mathbf{x}(k)\}_{k=0, \dots, N_{p,s} \in \mathcal{R}^{m_0}}$ , we can define its norm as

$$norm = \frac{1}{N_{p,s}} \sqrt{\sum_{k=0}^{N_{p,s}-1} \sum_{i=1}^{m_0} x_i(k)^2}. \quad (30)$$

#### F: Shannon entropy

For a random process  $x(k)$ , its Shannon entropy can be computed as

$$H(x) = - \sum_{k=0}^{N_{p,s}-1} P(x(k)) \log_2 (P(x(k))), \quad (31)$$

where  $P(x(k))$  is the probability of  $x(k)$ , which can be approximated by the relative frequency.

## APPENDIX B SELF-REPORTED FATIGUE SCALE

TABLE IV  
MODIFIED BORG CR10 SCALE FOR MUSCLE FATIGUE

Score	Explanation
1	Very weak
2	Weak - Fatigue is barely noticeable
3	Some fatigue
4	Some fatigue - Muscle fatigue is notable, it would be nice to slow down a bit
5	Strong fatigue - Tired and hard. It would be nice to take a rest, but you could shear another run
6	Strong fatigue - Tired and hard. You would like a rest, but you still do not have difficulties going on
7	Very strong fatigue - The muscle fatigue is so strong that you wish to stop and rest
8	Very strong fatigue - You have a strong desire to rest
9	Very strong fatigue - The muscle fatigue is nearly the worst you've experienced
10	Extremely strong fatigue - The muscle fatigue is the worst you have ever experienced before