## Graduate Theses, Dissertations, and Problem Reports

2020

# An In-Depth Investigation of the Effects of Work-Related Factors on the Development of Knee Musculoskeletal Disorders among Construction Roofers

Amrita Dutta
amdutta@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Civil Engineering Commons, and the Construction Engineering and Management Commons

# AN IN-DEPTH INVESTIGATION OF THE EFFECTS OF WORK-RELATED FACTORS ON THE DEVELOPMENT OF KNEE MUSCULOSKELETAL DISORDERS AMONG CONSTRUCTION ROOFERS

**Amrita Dutta**

A DISSERTATION

Submitted to the Benjamin M. Statler College of Engineering and

Mineral Resources at West Virginia University in partial fulfillment of the

requirements for the Degree of Doctor of Philosophy in Civil and

Environmental Engineering

Fei Dai, PhD, Chair
John P. Zaniewski, Ph.D.
Roger Chen, Ph.D.
Scott Breloff, Ph.D.
Yoojung Yoon, Ph.D.

**Department of Civil and Environmental Engineering**

**Morgantown, West Virginia**
**2020**

# ABSTRACT

# AN IN-DEPTH INVESTIGATION OF THE EFFECTS OF WORK-RELATED FACTORS ON THE DEVELOPMENT OF KNEE MUSCULOSKELETAL DISORDERS AMONG CONSTRUCTION ROOFERS

**Amrita Dutta**

Construction roofers have the uppermost likelihood of developing knee musculoskeletal disorders (MSDs). Roofers spend more than 75% of their total working time being restricted to awkward kneeling postures and repetitive motions in a sloped roof setting. However, the combined effect of knee-straining posture, roof slope and their association to knee MSDs among roofers are still unknown. This dissertation aimed to provide empirical evidence of the effects of two roofing work-related factors namely, roof slope and kneeling working posture, on the development of knee MSDs among construction roofers. These two factors were assessed as potential to increase knee MSD risks in roofing by evaluating the awkward knee rotations and heightened activation of knee postural muscles that might occur in sloped-shingle installation. Moreover, a novel ranking-based ergonomic risk analysis method was developed to identify the riskiest working phase in the sloped-shingle installation operation. In addition, a data fusion method was developed for treating multiple incomplete experimental risk related datasets that would affect the accuracy of risk assessments due to human and technology-induced errors during experimental data collection.

The findings revealed that roof slope, working posture and their interaction have significant impacts on developing knee MSDs among roofers. Knees are likely to have increased exposure to MSD risks during placing and nailing shingles on sloped roof surfaces. The established data fusion method has been proven feasible in handling up to 40% missing data in MSD risk-related datasets.

The contributions lie in enhanced understanding of the physical risk exposures of roofers' knee MSDs and creation of the ranking-based ergonomic analysis method and the fusion method that will help improve the MSD risk assessment in construction. In the long run, these outcomes will help develop new knee joint biomechanical models, effective interventions, and education and training materials that will improve the workplace to promote health and safety of roofers.

# ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude and appreciation to my amazing thesis advisor, Dr. Fei Dai for the continuous support for my Ph.D. study and research, for his motivation, enthusiasm, encouragement and immense knowledge. His intellectual guidance, insightful comments, constructive criticism and mentorship in each step of my Ph.D. journey have helped me understand how to be organized, disciplined, how to proceed step by step with ideas in research, how to write and improve the academic writing skills.

Besides my advisor, I would also like to extend my immense gratitude to Dr. Roger Chen, Dr. Yoojung Yoon, Dr. Scott Breloff, and Dr. John P. Zaniewski for serving as my committee members and for providing valuable suggestions, assistance, constructive and useful comments regarding my research and dissertation.

I would also like to thank John Z. Wu, Erik W. Sinsel, Robert E. Carey, Christopher M. Warren from the National Institute for Occupational Safety and Health (NIOSH) for helping in the experimental setup, data collection and data processing. My sincere gratitude to them for all the insightful suggestions, comments, feedbacks that they gave me on my manuscripts as coauthors.

I would especially like to thank the department of Civil and Environmental Engineering of West Virginia University for the resources and the logistic supports for this study.

I am also thankful to my fellow lab mates, Mohammed Sujon and Linjun Lu, for the research related discussion that we used to have in the lab and for all the fun we did while working together in the lab.

I extend my deep appreciation to my father, Ajit Chandra Dutta, mother, Kalpana Deb and my younger brother, Arin Dutta, for their unconditional love, moral support, encouragement and

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1: INTRODUCTION

## 1.1. Background

Awkward kneeling posture is a common source of knee musculoskeletal disorders (MSDs) among construction roofers. As roofers spend more than 75% of their total working time being restricted to awkward kneeling postures and repetitive motions in a sloped roof setting, they become susceptible to knee MSDs (CPWR 2018). Shingle installation is a prolonged repetitive and awkward task that residential roofers commonly preform on the jobsite which may result in chronic knee pain, irritation in the knee joints (i.e., bursitis), and osteoarthritis (Dulay et al. 2015; Reid et al. 2010).

During shingle installation, roofers encounter awkward posture when their knees undergo significant amount of rotations beyond their tolerance limit. They are typically required to maintain an awkward posture due to deep flexed kneeling (>90°) with medial and lateral rotations in knee joints on the sloped rooftop during shingle installation. This awkward kneeling posture also makes the muscles of lower limb less efficient in producing the required amount of force to accomplish a task, which may result in higher muscle activation compared to a neutral posture and overload the muscle (Kaushik and Charpe 2008; Lunde et al. 2014). Roofers also experience greater discomfort and pain in their lower extremities with an increase in slope (Choi 2008). Statistics indicated that the median days away from work—a key indicator of severity of injuries and illnesses which is used to measure time loss resulting from injuries—due to knee injuries is 16 days, which is the second highest (after shoulder injuries) by body parts among all the industries (BLS 2019). In the construction industry, the days away from work due to knee injuries account for 7.8% of the total days away from work resulting from nonfatal injuries, the fourth highest next

to back, hand and shoulder (CPWR 2018). Research has indicated that occupations with high kneeling requirements have the uppermost likelihood for the development of knee injuries (Porter et al. 2010). According to the Nonfatal Occupational Injuries and Illness report 2016, 38% of the total injuries involving lower extremities were located in the knees (BLS 2019). According to Wang et al. (2015), roofers' MSD rate is 30% higher than the average MSD incident rate reported for the entire construction trades. In the state of Washington, the insurance premium composite base rate for roofers is the highest ($7.03) among all building construction trades (Washington DOL & I 2018). According to several survey studies done on musculoskeletal symptoms among construction workers in different construction trades, knee MSD incident rate is highest among roofers as their job involves frequent kneeling (Marras and Karwowski 2003). As a result, there is a pressing need to alleviate the ergonomic injuries and develop interventions for reducing knee MSD among construction roofers.

While a combination of biomechanical models and sensor technologies have been used to assess work-related MSDs by measuring body motions, tissue loading and knee joint contact stress (Chen et al. 2017; Lee et al. 2017; Valero et al. 2017), the slanted surface of rooftops diminishes direct application of these methods to roofing context. Safety and health organizations such as the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH) have promoted general ergonomic practices and guidelines to minimize the risk of developing work related MSDs among construction workers (Albers and Estill 2007; OSHA 2019). In spite of these developed ergonomic practices and guidelines, there is still a lack of proper practical intervention to prevent knee MSD among residential roofers because: 1) most of the guidelines are very generic and do not consider the variance of posture preferences in roofing operations and 2) the existing tools are designed for

tasks performed on flat ground and the consideration of tasks on slanted residential roof surfaces that may involve continuous kneeling and lead to development of knee injuries is not included.

Moreover, very limited research works assessed the MSD risk factors for residential roofers, let alone the knee MSD risk factors. Roof slope, working technique and working pace were assessed for low back disorders among roofers (Wang et al. 2017). Roof slope, visual cue, muscle fatigue and task performed were assessed for balance maintenance during roofing on sloped roof surface (Lu et al. 2015). Effects of cross-slope roof walking were measured on the lower extremity kinematics of roofers (Breloff et al. 2019b). Two interventions —knee pads and knee savers, were tested to see if they can minimize lower extremity kinematics when roofers work in a fully flexed kneeling posture on sloped roof surface (Breloff et al. 2019c). Although these studies have greatly advanced the understanding of roofers' MSDs and the associated risk factors, they are not either specific to knee MSDs or did not consider the postural variances of dynamic kneeling posture that is involved in shingle installation roofing work. More specifically, the combined effects of constrained kneeling postures during shingle installation coupled with the slanted rooftop surfaces, and their association to knee MSDs among construction roofers have not been studied yet. Therefore, it is crucial to examine the roofing work-related factors that lead to the development of knee MSDs among construction roofers so that effective interventions, training program, safety strategies and ergonomic tools can be designed to reduce knee MSDs among them.

The accuracy of MSD risk assessments is often affected by the incompleteness in the risk exposure data when some data become missing due to technology- or human-induced errors during data collection. So proper treatment of those incomplete data is inevitable to ensure accurate and reliable assessment of the effects of the potential MSD risk factors on the imposed MSD risks.

However, such a method that can extrapolate the latent relationship between the risk factors and the risk indicators for replacing those missing data is lacking in the literature.

## 1.2. Research Objectives

The overall objective of this research was to add new knowledge on the association of work-related physical hazards to the development of knee MSD among construction roofers by systematically evaluating the effects of roofing work-related factors on their biomechanical and psychophysical responses, so that, effective interventions can be developed to reduce knee MSDSs among roofers' community. This research hypothesized that roofer knee MSD may directly or indirectly result from the exposure to the work-related physical hazards including sloped roof surface, awkward kneeling posture, prolonged working, and forceful actions or overexertion.

To accomplish the overall objective of this research, the following five specific objectives were pursued:

**Objective 1: To assess the impacts of roofing-related factors, namely working posture and roof slope, on the knee joint rotation during repetitive sloped shingle installation task.** The main hypothesis is that roof slope, kneeling posture, and their interaction would significantly impact the knee joint rotation during shingle installation and may lead to knee MSDs among construction roofers.

**Objective 2: To assess the impact of two residential roofing work-related factors — roof slope and awkward kneeling posture — on the activation of knee flexor and extensor muscles during the repetitive sloped shingle installation task.** The main hypothesis is that roof slope, kneeling posture, and their interaction would significantly impact he activation of the knee postural muscles during performance of roofing, which may lead to knee MSDs among construction roofers.

**Objective 3: To develop a ranking-based ergonomic method for suggesting the potential risky phases in terms of awkward knee rotations and repetitive motions during shingle installation operations on slanted surfaces.** The research questions are --- (i) which phase is the riskiest in a shingle installation task that yields the most awkward knee rotations and repetitions? (ii) what is the impact of each knee rotation on the imposed MSD risks? (iii) what is the relative contribution of each knee rotation to potential knee MSD risk measurement?

**Objective 4: To develop a method for treating multiple imperfect MSD risk related experimental datasets with missing values by fusing those datasets for ensuring reliable risk assessment in construction.** The research questions are --- (i) how to treat multiple MSD risk related datasets that contain missing values? (ii) can data fusion be a feasible option for treating multiple MSD risk-related datasets with missing values? (iii) how to hold the latent relationships among multiple risk indicators and risk factors during the fusion so that a complete representation of the risk indicator values can be preserved in the fused data?

**Objective 5: To evaluate the numerical stability of Canonical Polyadic Decomposition (CPD) method in fusing multiple risk-related sparse datasets and hence the effectiveness of applying this technique for MSD risk assessment in construction when the quality of risk-related datasets becomes degraded due to the presence of missing data.** The research questions are --- (i) given a different portion of missing data in two sparse risk-related datasets, whether, at what portion of missing data, and to what extent, the fusion-based risk assessment outperforms the one with no fusion? (ii) Can CPD be useful as a data fusion technique for MSD risk assessment?

### 1.3. Research Undertakings

To test the hypothesis of the specific **objective 1**, the effects of two roofing related knee MSD risk factors (namely roof slope and kneeling posture) were evaluated as potential knee MSD

risk factors in a laboratory setting. To indicate knee MSD risk, five awkward knee rotations were measured. The five awkward knee rotations were: flexion, abduction, adduction, internal rotation and external rotation. At different roof slope and kneeling posture, the changes in the above-mentioned knee joint rotation angles were measured and compared. Using statistical methods, the effect of roof slopes and kneeling postures on the knee joint rotation angles were analyzed which can help understand if knees can be exposed to MSD risks during shingle installation on slanted rooftops.

To test the hypothesis of the specific **objective 2**, the effects of kneeling posture and roof slope on the activation of major knee postural muscles during shingle installation were evaluated via a laboratory assessment. Maximum normalized electromyography (EMG) data were collected from five knee flexor and extensor muscles. The postural muscles were: Rectus femoris (RF), Vastus lateralies (VL), Vastus medialis (VM), Biceps femoris (BF) and Semitendinosus (S). At different roof slope and kneeling posture, the activation of the above-mentioned postural muscles was measured and compared to evaluate if postural variation due to awkward kneeling posture in a sloped working environment can overload the muscle tissue which can potentially lead to muscle fatigue. A systematic and statistical evaluation of knee flexor and extensor muscle activations associated with awkward kneeling roofing tasks performed on a sloped roof surface will be important to develop effective interventions and training materials to minimize muscle overloading.

To answer to the research questions of the specific **objective 3**, time series knee joint rotation data (kinematics) were also analyzed for in-depth understanding of work-specific risk exposures to roofers' knees during shingle installation. A ranking-based ergonomic method for suggesting potentially risky phases that may increase knee MSD risk during shingle installation

operations was developed. The entire shingle installation task was first segmented into multiple phases and five knee rotation angles—flexion, abduction, adduction, internal and external rotation were measured as risk indicators from a laboratory-based shingle installation simulation. The risk indicators were aggregated applying scoring models and phases were ranked according to the imposed risk of knee MSDs in terms of repetition, forceful exertions and prolonged kneeling to identify the riskiest phase for knee MSDs and the impact of each knee rotation on knee MSD. Using Spearman's rank correlation test, the relative contribution of each knee rotation to the knee MSD risk measurement were statistically investigated. By assessing risk in an operation level, this study will reveal the unsafe working technique in roofing which can facilitate effective intervention strategies (education, training, and tools) for reducing knee exposure and hence minimizing knee injuries and disorders among construction roofers.

To answer to the research questions of the specific **objective 4**, a data fusion method was proposed that applies Canonical Polyadic Decomposition (CPD) tensor decomposition to fuse multiple sparse risk-related datasets and fill in missing data in MSD risk assessment studies. The effectiveness of CPD in fusing multiple sparse risk datasets by leveraging the correlation among multiple risk indicators and the risk factors within those datasets were also analyzed. For different proportions of missing values, the performance of the proposed fusion method was investigated. To validate the proposed method, reconstructed fused data were statistically compared to the original data obtained from two roofing experimental studies. The risk assessment results obtained from the fused dataset were also compared to those obtained from the experimental datasets. The proposed method will be a new contribution to the field of occupational safety and health which will enable accurate and reliable assessment of work-related risk factors on the imposed risks by treating multidimensional imperfect and incomplete risk exposure datasets with missing values.

To answer to the research questions of the specific **objective 5**, risk assessment results obtained with fusion and without fusion was compared for different proportion of missing vales from the risk datasets. For evaluation, two knee MSD risk-related datasets collected from roofing experimental studies were fused applying CPD for different proportion of missing values. Consistencies of the risk assessment results obtained from fused and unfused sparse datasets were then compared to those obtained from the original experimental datasets. The findings of this study will help understand the numerical stability of fusion through CPD, i.e., despite of using malformed data as input, to what extent of missing values, data fusion can provide more accurate and reliable MSD risk assessment results than without fusion of missing data and hence, can be used as an alternative approach of replacing missing data in MSD risk assessment studies.

### 1.4. Organization of Dissertation

In *Chapter 1*, the background, the overall objective, the specific objectives and the research undertakings have been discussed.

In *Chapter 2*, two roofing work-related factors namely roof slope and kneeling working posture have been assessed as potentials to increase knee MSD risks during roofing by evaluating the awkward knee joint rotations.

In *Chapter 3*, two roofing work-related factors namely roof slope and kneeling working posture have been assessed as potentials to increase knee MSD risks during roofing by evaluating the heightened activation of knee postural muscles.

In *Chapter 4*, a novel ranking-based ergonomic risk analysis method has been developed to identify the riskiest working phase of sloped-shingle installation operation that might yield the most awkward kneeling postures, prolonged working, forceful actions or overexertion, and repetitive motions.

In *chapter 5*, a data fusion method has been proposed and validated for treating multiple incomplete experimental risk-related datasets for robust risk assessment of MSDs in construction by applying Canonical Polyadic Decomposition (CPD) technique.

In *chapter 6*, the numerical stability of fusion through CPD in risk assessment has been analyzed and discussed.

In *chapter 7*, a general discussion based on the findings of the study and recommendations have been provided.  Limitations and future extensions are also discussed.

Finally, in *chapter 8*, conclusion and key contributions of the study are discussed.

# CHAPTER 2: ASSESSING WORK-RELATED RISK FACTORS FOR MUSCULOSKELETAL KNEE DISORDERS IN CONSTRUCTION ROOFING TASKS

**Abstract**

Roofers often suffer from musculoskeletal disorders (MSDs) to their knees due to spending a large amount of time kneeling while performing work-related roofing activities on sloped rooftops. Several ergonomic studies have identified kneeling as a potential risk factor for knee injuries and disorders. Existing biomechanical models and sensor technologies used to assess work-related risk factors for different construction trades are not applicable in roof work settings especially on slanted rooftop surfaces. This work assesses the impacts of work-related factors, namely working posture and roof slope, on the potential risk of developing knee MSDs due to residential roofing tasks in a laboratory setting. Nine human subjects participated in the experiment and mimicked shingle installation on a slope-configurable wooden platform. Maximum angles of right and left knee flexion, abduction, adduction, and axial rotation (internal and external) were measured as risk indicators using a motion capture system under different roof slope settings. The results demonstrated that roof slope, working posture and their interaction may have significant impacts on developing knee MSDs during roofing activities. Knees are likely to be exposed to increased risk of MSDs due to working in a dynamic kneeling posture during shingle installation. In our study, flexion in both knees and adduction in the right knee were found lower in high-pitched rooftops; however, abduction in the left knee and internal rotation in the right knee were found higher during shingle installation. Hence proper attention is needed for these situations. This study provides useful information about the impact of roof work settings on knee MSDs development, which may facilitate effective interventions such as education, training, and tools to prevent knee injuries in construction roofing tasks.

## 2.1. Introduction

Roofers often spend a large percentage of their daily working time in kneeling postures which has been considered a primary risk factor for knee disorders (Xu et al. 2017). While a combination of biomechanical models and sensor technologies have been used to assess work-related musculoskeletal disorders (MSDs) by measuring body motions, tissue loading and knee joint contact stress (Chen et al. 2017; Li et al. 2017; Valero et al. 2017), the slanted surface of rooftops diminishes direct application of these methods to this specific context. Little work has assessed the knee MSD risk factors, specifically among residential roofers, due to their constrained kneeling postures coupled with the slanted rooftop surfaces. This study examines the impact of constrained kneeling posture and slanted roof surface—two common residential roofing work-related factors, on knee MSDs.

## 2.2. Severity of Occupational Knee Injuries

Knee injuries are common musculoskeletal disorders that involve chronic knee pain, discomfort, tingling and numbness, irritation of the fluid filled sacs in the knee joints (bursitis), meniscal lesions or tears, and osteoarthritis (Dulay et al. 2015; Reid et al. 2010). Statistics indicated that the median days away from work—a key indicator of severity of injuries and illnesses which is used to measure time loss resulting from injuries—due to knee injuries is 16 days, which is the second highest (after shoulder injuries) by body parts among all the industries (BLS 2019). In the construction industry, the days away from work due to knee injuries account for 7.8 % of the total days away from work resulting from nonfatal injuries, the fourth highest next to back, hand and shoulder (CPWR 2018). Research has indicated that occupations with high kneeling requirements have the uppermost likelihood for the development of knee injuries (Porter et al. 2010). It is echoed in the study analyzing the MSD claims data of coal miners, conducted by Gallagher et al. (2009)

and Porter et al. (2010), which revealed that the frequency of knee injuries is 1.7 times greater than low back injuries, which has incurred an average cost of $13,121 per knee MSD, comparable to the average cost per low back MSD ($14,378). Kneeling is a regular activity of construction roofers (Tennant et al. 2015). Due to the unique work environment (slanted rooftops), residential roofers spend more than 75% of their working time in crawling, and squatting, stooping and kneeling postures, and the cumulative effects of these awkward postures combined with repetitive motions have caused a high incident rate of MSDs (Wang et al. 2015; Wang et al. 2017) among roofers. In addition, the insurance premium composite base rate for the construction roofers is the highest ($7.03) among all building construction trades, around 3 times as high as the average rate of all building construction trades (Washington DOL & I 2018). Knee awkward posture has been proven to be associated with knee MSDs (Hofer et al. 2011). As roofers encounter awkward posture while their knees undergo significant amount of rotations, sometimes even beyond their tolerance limit, during kneeling while performing roofing task; knee disorder is a common source of MSDs among construction roofers. In summary, construction roofers suffer from knee MSDs due to their unique harsh work conditions (Welch et al. 2009). There is an urgent need to examine the factors that lead to the development of knee MSDs among this population.

## 2.3. State of Ergonomic Practices for Occupational Knee MSD Prevention

While not specific for knee MSD prevention, safety and health organizations such as the Occupational Safety and Health Administration (OSHA) and the National Institute for Occupational Safety and Health (NIOSH) have promoted general ergonomic practices to minimize the risk of developing work-related MSDs among construction workers. For example, NIOSH published a booklet titled *Simple Solutions: Ergonomics for construction workers* that suggests simple and inexpensive tools (e.g., using less labor-intensive and less time-consuming materials

and building components, auto-feed screw gun with extension, kneeling creeper and knee pads) to minimize the stress on knees during construction tasks that involve continuous kneeling (Albers and Estill 2007). OSHA has provided training programs (10-hour and 30-hour programs) and tools (eTools, eMatrix, v-Tools) which are intended to promote workplace safety and make workers more knowledgeable about the potential hazards in jobsites. In the OSHA eTool, kneeling for repetitive activities has been described as a risk factor for knee injuries and potential solutions have been provided using protective measures such as kneepads and power stretchers during work (OSHA 2019). In addition, the Infrastructure Health and Safety Association (IHSA) in Canada has published three brochures that provide information on how to minimize MSD risks and develop hazard control plans based on threats associated with common residential roofing tasks (IHSA 2015a; IHSA 2015b; IHSA 2015c). These brochures provide general safety solutions to protect roofers from low back, shoulder and knee injuries (IHSA 2015a), heat stress (IHSA 2015b) and fall hazards (IHSA 2015c). In spite of these developed ergonomic practices and guidelines, there is still a lack of proper practical intervention to prevent knee MSD among residential roofers because: 1) most of the guidelines are very generic and do not consider the variance of posture preferences in roofing operations and 2) the existing tools are designed for tasks performed on flat ground and the consideration of tasks on slanted residential roof surfaces that may involve continuous kneeling and lead to development of knee injuries is not included.

## 2.4. To-Date Ergonomic Studies for Construction Roofers

Until recently, ergonomic studies on residential roofing construction have primarily focused on two areas: surveillance studies in prevalence of ergonomic injuries among roofers and work-related ergonomic MSD risks in the roof setting.

### 2.4.1. Surveillance studies in prevalence of ergonomic injuries among roofers

Fredericks et al. (2005) studied the statistical number of nonfatal occupational injuries involving days away from work by source of MSD. They reported that worker motion and position are two common sources of injuries for workers performing the residential roofing activities. The authors also reported that knee injuries contribute to 7% of the nonfatal injuries that involve days away from work – the third highest after low back injuries (21%) and finger injuries (9%). Repetition and awkward posture are common physical risk factors that may cause MSDs while performing roofing tasks (Jaffar et al. 2011). Musculoskeletal conditions among roofers are associated with work limitations, missed work and reduced physical functioning which result in early departure from the workforce and even disability (Welch et al. 2009; 2010; 2015). Roofing has the second highest work-related MSD incident rate (52%) among all construction tasks, which is more than 30% of the work-related MSD incidents reported in the whole construction industry (40%) (Wang et al. 2015). Roofers working for small employers often lack proper training (Moore and Wagner 2014). As a result, the MSD risks caused by the residential roofing tasks require more attention from safety and health regulators and researchers (Dai et al. 2016; Smith-Jackson et al. 2011).

### 2.4.2. Related works in assessing ergonomic MSD risks in construction roofing

In the research community, only a few efforts have assessed the work-related MSD risk factors for roofing activities. Wang et al. (2017) examined the influences of roof slope, working technique, and working pace in kneeling and stooped postures on development of low back disorders in roofing. However, the lower extremities were not systematically assessed in their study. It is unclear if knees are exposed to MSD risks due to their flexion, adduction, abduction, and/or rotation when roofers work on slanted rooftops with a posture preference of kneeling.

Breloff et al. (2019b) examined the lower extremity kinematics of roofers during traverse walking across sloped roof surfaces. Nevertheless, to date, no work has assessed roofing work-related knee MSD risk factors, such as, kneeling posture and roof slope, in construction roofing. Further assessment is necessary to identify the knee MSD risk factors involved in construction roofing activities so that efficient intervention tools can be developed to prevent knee disorders among roofers.

### 2.4.3. Potential work-related knee MSD risk factors in roofing activities

Based on the findings by Wang et al. (2015), site visits, and literature review; the following potential work-related knee MSD risk factors in roofing activities are identified and presented below.

### 2.4.3.1. Working posture

Prolonged static kneeling posture alters subsequent gait stance-phase knee flexion and adduction moments, which have been identified as risk factors for developing knee osteoarthritis (Kajaks and Costigan 2015). Frequent and dynamic kneeling activities, where the lower legs often rotate internally or externally, can cause anterior cruciate ligament (ACL) disorder of knee (Sharrard and Liddell 1962). These findings are relevant to roofing operations, as during shingle installation roofers are typically required to maintain an awkward posture due to deep flexed kneeling (>90°) with medial and lateral rotations on the sloped rooftop. In spite of this, the relationship between the working posture undertaken by the roofers during shingle installation and knee MSD risks is yet to be investigated.

### 2.4.3.2. Roof slope

Previous studies found evidence on association between surface slope and postural balance (Choi and Fredericks 2008; Wang et al. 2017). For kneeling posture, trunk flexion angle drops and

muscle activation increases as slope increases (Wang et al. 2017). Roofers also experience greater discomfort and pain in their lower extremities with an increase in slope (Choi 2008). This adds a motivation to quantitatively assess the impact of roof slope angles on lower extremities, specifically the knees, so as to identify risky work setting that may develop knee MSD during roofing construction.

## 2.5. Problem Statement and Research Objective

Based on the above review, a systematic evaluation of residential roofing working conditions on roofers' knee musculoskeletal MSD risks is lacking. More specifically, the combined effect of knee-straining posture of roofers, roof slopes and their association to knee MSDs resulting from awkward knee rotations is still unknown. Without such knowledge, it is difficult to develop effective interventions to prevent the awkward knee rotations and hence, MSDs among roofers. Therefore, the objective of this research is to assess the impacts of roofing-related factors, namely working posture and roof slope, as a potential to increase knee MSD risk during construction roofing tasks by evaluating the awkward knee rotations that might generate while kneeling on sloped roof surface.

## 2.6. Experiment Design and Implementation

### 2.6.1. Variables

This study evaluated two roofing related knee MSD risk factors in this experiment. They are: roof slope (0, 15 and 30 degree) and working posture (static kneeling and dynamic kneeling). These roof slopes were chosen to correspond to roof pitches that are commonly used in construction sites (Johnson 1976). Also, these slopes were used in a similar study on low back disorder of roofers at sloped roof surfaces (Wang et al. 2017).

To assess the impact of working posture on the knee MSD risk, two typical kneeling scenarios in roofing practices were considered. The first scenario was the static posture which referred to a certain motionless body posture that occurs just after placing and right before installing shingles during shingles installation work. The participants adopt this technique when they flex their knees and the trunk so that they can put their non-nailing hand on the roof surface and hold the nail gun with their nailing hand. This way, their trunk stayed more inclined towards the roof surface and thereby the thigh remained in upright position. This working technique is referred to as *static kneeling posture* in this paper. This posture represents the general body orientation of the roofers during shingles installation task. In this motionless working posture (just after placing and right before affixing the shingles), the risk indicators of knee MSDs were assessed so that the potential of increasing knee MSD risk due to performing shingle installation in dynamic working posture could be evaluated.

In the second scenario, the participants performed the entire shingle installation work. In specific, they lifted their trunk upward a little bit to reach for shingles first and placed shingles in front of them on the rooftop. Then, they grabbed the nail gun from aside and installed shingles side by side by nailing. After finishing the shingles nailing, they replaced the nail gun and returned to the upright position resting on their knees. This entire working technique is referred to as *dynamic kneeling posture* in this paper. This posture requires much lateral movement of the knees with respect to the midline of the body and the knee joints.

To reflect the effects of the factors and indicate the knee MSD risk in construction roofing activities, this study measured five response variables. They are: 1) Maximum knee flexion angle, 2) Maximum knee abduction angle, 3) Maximum knee adduction angle, 4) Maximum knee internal rotation, and 5) Maximum knee external rotation. The knee flexion angle is the angle between

thigh and lower lag due to bending of knee from the upright standing position, i.e. the rotation of the lower leg about the medio-lateral axis that runs from left to right of the leg through the knee joint (Grood and Suntay 1983) (Figure 1). Deep flexion activities generate significant net quadriceps moments at the knee and increase the stress on the patellar tendon (Nagura et al. 2002).

Knee abduction and adduction is the rotation of lower leg with respect to the anterior-posterior axis (Grood and Suntay 1983) (Figure 2.1). The anterior-posterior axis at the knee joint runs from the front of the leg to the back through the knee joint. Abduction is rotating away from the axis and adduction is rotating towards the axis. Higher knee adduction and abduction induces more stress and force on the medial compartment of the knee and increases the risk of developing osteoarthritis (Barrios et al. 2010).

The other knee rotations in a kneeling posture are internal and external rotation of knee about the longitudinal axis (Figure 1) (Grood and Suntay 1983). In upright standing position, this axis passes vertically along the leg. During deep-knee flexion, external rotation of the tibia relative to the femur places additional stress on the knee joint ligaments, particularly the posterior cruciate ligament (PCL) and medial collateral ligament (MCL) (Hofer et al. 2011). Internal rotation of tibia places stress on the knee joints (Coplan 1989).
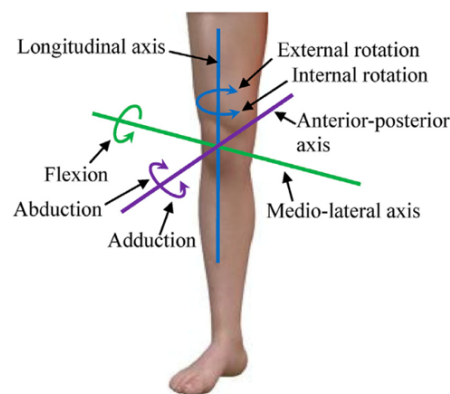


**Figure 2.1:** Knee flexion, abduction-adduction and internal-external rotation

### 2.6.2. Participants

A total of 9 healthy male volunteers [26.1 years (5.6 years), 180.2 cm (6.1 cm), and 99.7 kg (27.6 kg)] participated in this experiment. Although the participants were not actual roofers, but they had experience in home remodeling work. 8 of 9 participants were right-handed. As over 97% of roofers are male (BLS 2019), females were excluded from this study. No participant had any history of musculoskeletal or neurological disorders. The research protocol was approved by both the Institutional Review Board (IRB) of National Institute for Occupational Safety and Health (NIOSH) and West Virginia University. Before participating in the study, the subjects read and competed an informed consent form.

### 2.6.3. Instruments

The knee kinematic data (segment endpoint data from motion capture system) was collected using a VICON optical motion analysis system with 14 MX Vicon cameras (Oxford, U.K). 42 retroreflective motion capture markers were placed bilaterally on the participant's hip joints, thighs, knee joints, shanks, ankles, toes, heels and feet, following the approach used in a similar study done by Pollard et al. (2011). As the participants were shod, the markers of the heels, feet and toes were attached to the shoes of the participants. The collected 3D coordinates of these markers were used to calculate the knee angles and determine the kneeling postures in the point cloud. The mathematical calculations of the knee angles using the markers were elaborated in the following "Data Processing" section. The kinematic data was recorded at 100 Hz.

A 1.2 ×1.6 m custom-made adjustable wood platform was built to mimic the rooftop on which the simulated shingle installation was conducted. Two sets of wooden legs were bolted right next to each other to set the simulator at two different fixed slopes (15° and 30°) during the experiment. A battery powered lift with 25-inch adjustable forks would raise one side of the roof

simulator and depending on which angle was needed, one of the two sets of legs was chosen. Anti-skidding tape was attached to the platform surface to increase the friction and avoid slips or falls. Eight markers were placed on this wooden platform to construct a roof specific rigid body for post-processing. Figure 2.2 depicts the wooden platform along with a participant nailing shingles while kneeling on it at 0°, 15° and 30° slopes.
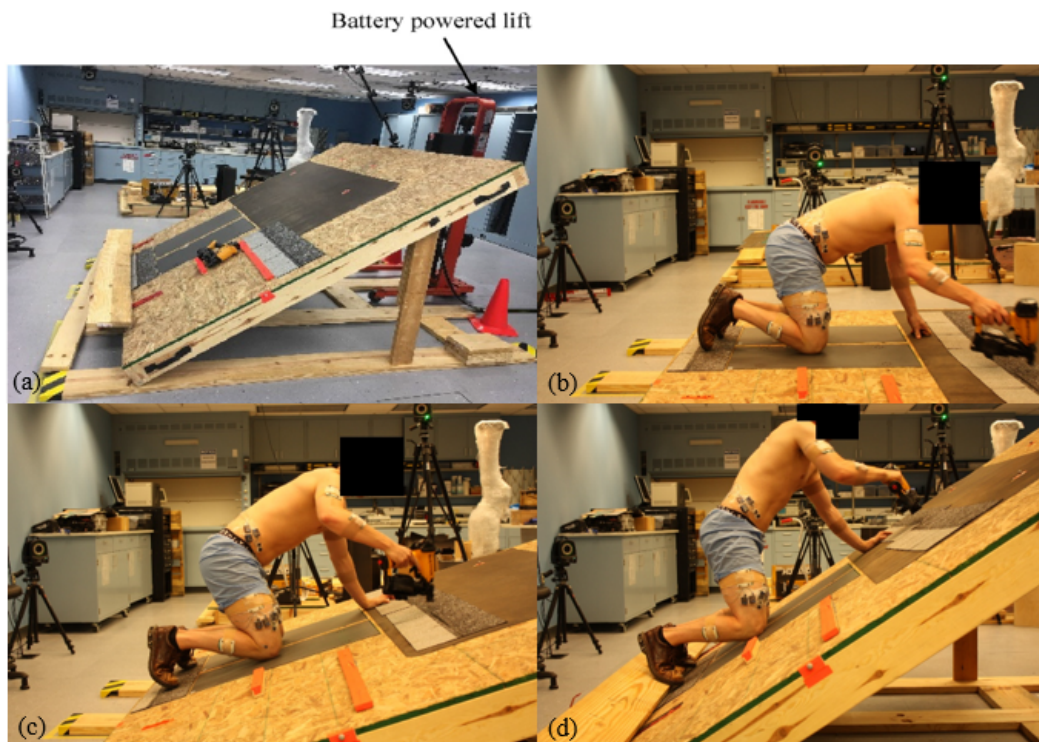


**Figure 2.2:** Experimental setting for the shingle installation simulation (a) wooden platform, (b) at 0° slope, (c) at 15° slope, and (d) at 30° slope.

### 2.6.4. Procedure

The experiment was performed in the NIOSH biomechanics laboratory. The experimental procedure was introduced to the participants upon their arrival to the lab. After the participants signed the informed consent forms, motion markers were placed on them for kinematic calibration and data collection. After that, the participants completed the static and dynamic trials. In static kneeling trials, the participants flexed the knee and trunk in a posture, so that they could rest their

'non-nailing' hand on the rooftop and hold a nail gun in the other hand. In dynamic kneeling trials, when the participants were instructed to start, they first reached for and placed the shingles and then picked up the nail gun from their right side and mimicked securing 6 nails (3 in each) into two shingles side by side on the roof segment. Once done, the participants replaced the nail gun and returned to their resting position. In summary, a complete dynamic trial consisted of 7 phases—reaching for shingles, placing shingles, grabbing the nail gun, moving to first nailing position, installing shingles by nailing, replacing the nail gun, and retuning to upright position. Figure 2.3 depicts the time series plot of knee angles across different phases in a dynamic trial.

To reduce inconsistencies with the demarcation of the phases of shingling, only one human expert manually identified each transition point. The expert used videos of each trial and motion capture data of the right wrist (changes in direction) to fine tune the definitions of the shingling phases. All static and dynamic kneeling trials were completed on the roof simulator for three slopes angles — 0°, 15°, and 30° which resulted in total 6 (2×3) tasks for 2 postures and 3 slope angles. All conditions were randomized and due to the number of possible combinations, no subject preformed the testing conditions in the same order. Each participant performed each trial 5 times.

### 2.6.5. Data processing

This study calculated 5 knee rotational angle variables in Visual 3D (C-Motion, Inc.) using the kinematic data captured by the VICON optical motion analysis system. Each marker has *x, y, z* positions based on a calibrated origin. Using these positions, the segment local coordinates i.e., the vector components, are calculated and then transformed to the global 3D (*XYZ*) coordinate system. The marker positions along with the local and global coordinate systems of the right leg is depicted in Figure 2.4.
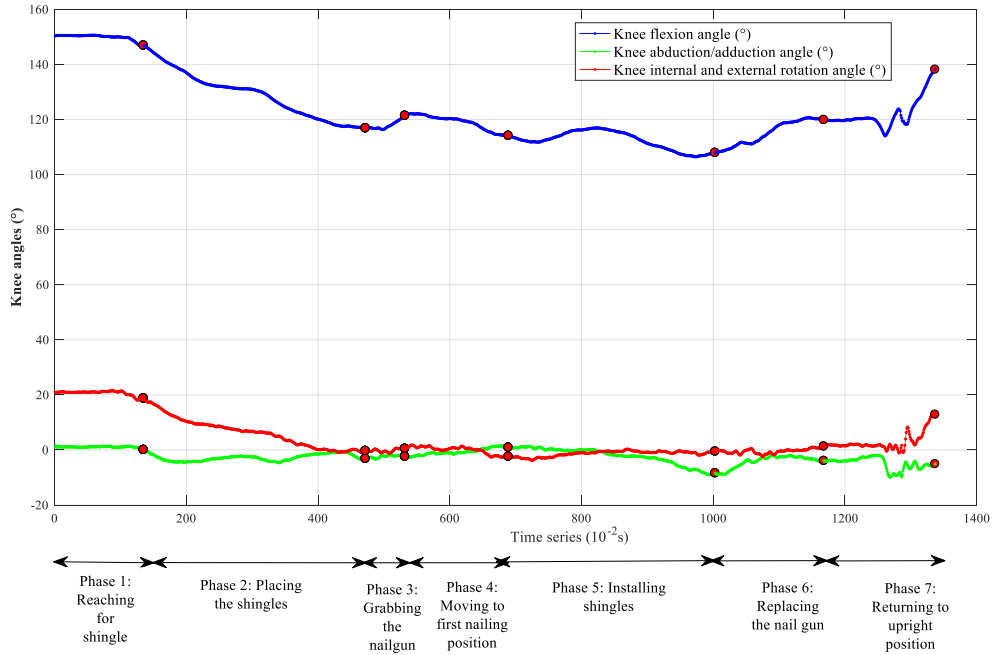
**Figure 2.3:** Time series plot of knee angles across different phases in a dynamic trial
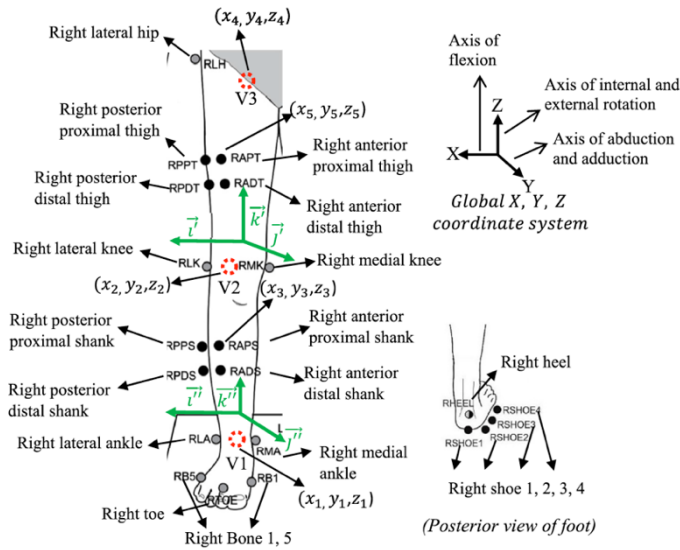


**Figure 2.4:** The markers set up in the hip joint, thigh, knee joint, shank, ankle, toes and foot for knee angles calculation. Only the markers of right leg are presented here. Total 42 markers were set up in both legs

The local coordinate system of a segment is a reference system that is fixed within a segment but moves as the segment moves. This local coordinate system of one segment is oriented

22

relative to the local coordinate system of a second segment and the joint angles are those between the two segments (Robertson et al. 2013). Thus, from the 3D coordinates of the markers placed on the thigh $(x_5, y_5, z_5)$, hip joint $(x_4, y_4, z_4)$ and knee joint $(x_2, y_2, z_2)$, the vector components, i.e., the local coordinate system $(\vec{i'}, \vec{j'}, \vec{k'})$, of thigh segment can be calculated. From the markers placed on the shank $(x_3, y_3, z_3)$, ankle joint $(x_1, y_1, z_1)$ and knee joint $(x_2, y_2, z_2)$, the vector components i.e., the local coordinate system $(\vec{i''}, \vec{j''}, \vec{k''})$ of shank (lower leg) segment can be calculated. The above markers were used to track the motion of the thigh and shank. In addition, several virtual markers were employed to calculate the segment coordinate system. Virtual marker 1 (V1) was considered in the midpoint of the distance between right medio ankle (RMA) and right lateral ankle (RLA) markers. Virtual marker 2 (V2) was considered in the midpoint of the distance between right lateral knee (RLK) and right medial knee (RMK) joint markers. Both of V1 and V2 markers were used to calculate the local coordinate system of the shank segment. Virtual marker 3 (V3) was considered in the hip joint center and calculated as ¼ of the intertrochanteric marker distance i.e., the distance between the left and right hip joint markers. Both V2 and V3 virtual markers were used to calculate the local coordinate system of the thigh segment. Combining these two local coordinate systems of the thigh and shank segments, a rotation transformation matrix was constructed which was further used to compute the knee abduction-adduction (about *Y*-axis) flexion (about *X*-axis) and axial rotation (internal-external) (about *Z*-axis) implementing the equations documented by Robertson et al. (2013). For dynamic trials, the 7 phases were segmented manually by watching the motion capture data and adding an event at the start and end of each phase in Visual 3D. The purpose of the study was to analyze the differences of knee angles in static and dynamic postures during shingle installation. As the static posture is a motionless work technique that roofers undertake just after placing and right before installing shingles, in the

23

dynamic trials, this study analyzed only the kinematics data of phase 3: grabbing the nail gun through phase 6: replacing the nail gun, excluding the data from phase 1 to phase 2 that involved placing shingles, and the data of phase 7 that involved returning to starting upright position for fair comparison. So, for dynamic trials, maximum knee angles were computed across phase 3 through phase 6. For static trial, the motion capture system was set to record the knee kinematics data for seven seconds. Data of the middle three seconds were extracted for analysis to avoid any artifact motion from the first and last two seconds. The maximum knee angles of the static trial were calculated from those extracted data. These maximum knee angles were used as the response variables for the statistical analysis described in the next section.

### 2.6.6. Data analysis

Since there were more than one response variables in this study, multivariate ANOVA (MANOVA) was first applied to examine the impacts of the two factors (i.e., slope and posture) and their interaction on the all response variables. Then, univariate ANOVA was applied to measure the impacts of these factors and their interaction on each individual response variable. Prior to conducting the ANOVA, the assumptions of the ANOVA technique (normality of residuals, constant variance of residual and independence of observations) were evaluated using the graphical approach discussed in Freund et al. (2010). The normal probability plots and residual vs. fitted plots were drawn for each response variable to observe the residual's variabilities. Once the assumptions were evaluated and satisfied, the ANOVA was applied. Following these, Tukey Post-Hoc analysis was performed on the factor (i.e., slope) with three levels to further analyze where the significant difference existed among different levels. The $p$-value was set as 0.05 for all tests. All tests were performed with Minitab 18 (Minitab, Inc.) and RStudio (RStudio, Inc.) software tools.

## 2.7. Results

Tables 2.1 and 2.2 showed the effects of the independent variables on the response variables for left and right knees. Here, Flexion_L, Abduction_L, Adduction_L, Internal_Rotation_L and External_Rotation_L refer to flexion, abduction, adduction, internal and external rotation of the left knee. Flexion_R, Abduction_R, Adduction_R, Internal_Rotation_R and External_Rotation_R means flexion, abduction, adduction, internal and external rotation of the right knee. As the influence of the independent variables were different for the two knees, the results were presented and discussed separately.

**Table 2.1:** Effects of the factors and their interaction on the response variables for left knee

| Factors | MANOVA | ANOVA | | | | |
|---|---|---|---|---|---|---|
| | | Flexion_L | Abduction_L | Adduction_L | Internal_Rotation_L | External_Rotation_L |
| Slope | <0.001 (10.3) * | <0.001 (24.9) * | 0.11 (2.2) | 0.088 (2.5) | 0.5 (0.7) | 0.79 (0.2) |
| Posture | <0.001 (116.6) * | <0.001 (285.5) * | 0.33 (0.9) | 0.041 (4.2) * | 0.008 (7.3) * | 0.014 (6.2) * |
| Slope*Posture | <0.001 (4.2) * | 0.002 (6.2) * | 0.042 (3.1) * | 0.234 (1.5) | 0.86 (0.2) | 0.63 (0.5) |

*Note:* * indicates a significant p value. F-values are provided in the Parenthesis.

**Table 2.2:** Effects of the factors and their interactions on the response variables for right knee

| Factors | MANOVA | ANOVA | | | | |
|---|---|---|---|---|---|---|
| | | Flexion_R | Abduction_R | Adduction_R | Internal_Rotation_R | External_Rotation_R |
| Slope | <0.001 (14.2) * | <0.001 (29.9) * | 0.13 (2.1) | 0.002 (6.8) * | 0.194 (1.7) | 0.47 (0.8) |
| Posture | <0.001 (114.8) * | <0.001 (319.3) * | 0.004 (8.5) * | 0.008 (7.3) * | <0.001 (19.8) * | 0.75 (0.1) |
| Slope*Posture | 0.003 (3.3) * | <0.001 (9.1) * | 0.47 (0.8) | 0.59 (0.5) | 0.041 (3.1) * | 0.45 (0.8) |

*Note:* * indicates a significant p value. F-values are provided in the Parenthesis.

### 2.7.1. Interaction effects of roof slope and posture

As indicated in Tables 2.1 and 2.2, there were significant interactions effects of slope and posture on the mean change of the maximum flexion of both knees, maximum abduction of left knee and maximum internal rotation of right knee. Due to these findings, at each slope, maximum left and right knee flexion, maximum left knee abduction and maximum right knee internal rotation at static and dynamic postures were compared. For each posture, differences in between three slopes were compared as well. The interaction effects were plotted in Figs. 2.5, 2.6, 2.7 and 2.8. Here, 'S' indicates static and 'D' indicates dynamic posture.

At 0°, 15° and 30° slopes, maximum knee flexion in left and right knee in dynamic posture was significantly higher than that in static posture. In static posture, maximum flexion of left and right knee at 30° roof slope was significantly lower compared to 0° and 15° roof slopes. In dynamic posture, maximum left knee flexion at 0° slope was significantly higher than that in 30° slope. But, for right knee, no significant difference in maximum flexion was observed in between any roof slope for dynamic posture.

At 30° slope, maximum left knee abduction was significantly higher in dynamic posture than in static posture, but for 0° and 15° slopes, the difference in static and dynamic postures was not significant. In dynamic posture, maximum left knee abduction at 30° roof slope was significantly higher than that in 0° and 15° slope, but no significant difference was found in between any of the roof slopes for static posture.
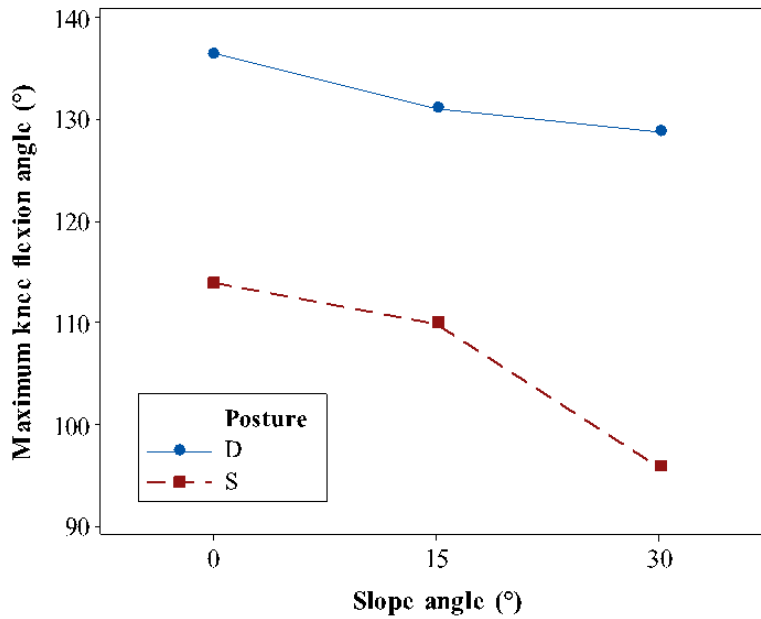
**Figure 2.5:** Interaction effect of slope and posture on left knee flexion
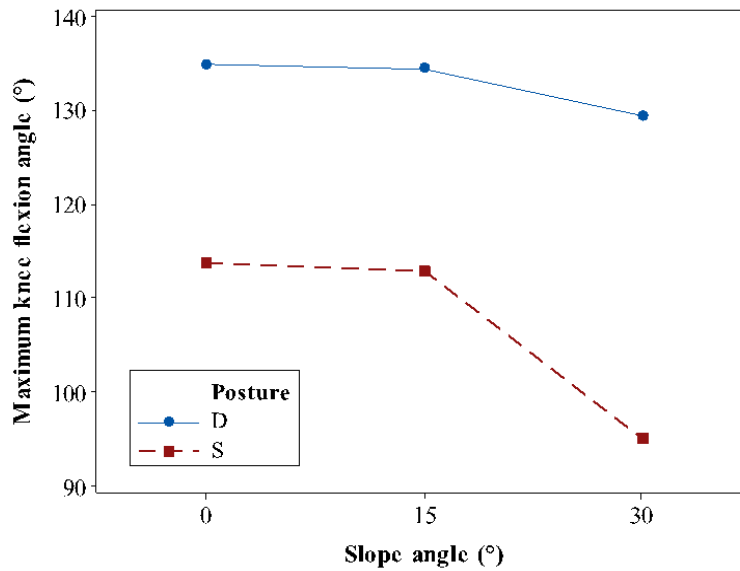


**Figure 2.6:** Interaction effect of slope and posture on right knee flexion

Similarly, at 30° slope, maximum right knee internal rotation was significantly higher in dynamic posture than in static posture, but for 0° and 15° slopes, the difference in static and dynamic posture was not significant. In addition, for static posture, right knee internal rotation at

30° slope was significantly lower than in 15° slope. But no significant difference in internal rotation was found in between any of the roof slopes in dynamic kneeling posture.
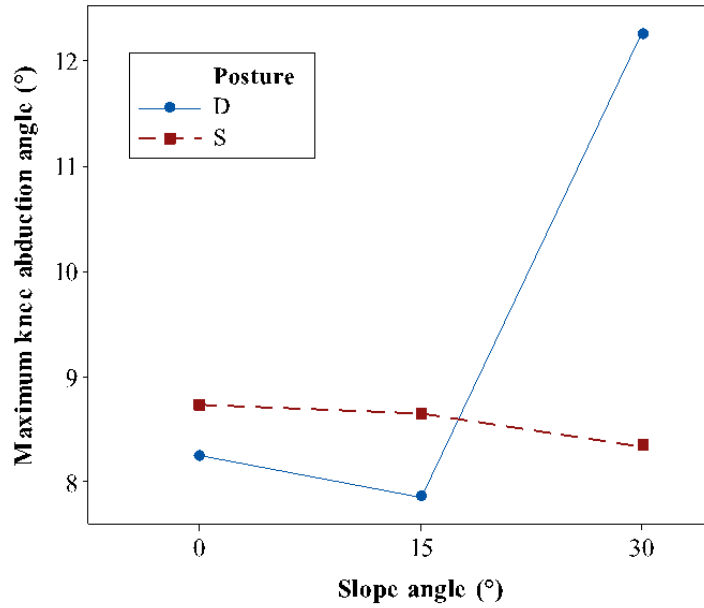


**Figure 2.7:** Interaction effect of slope and posture on left knee abduction
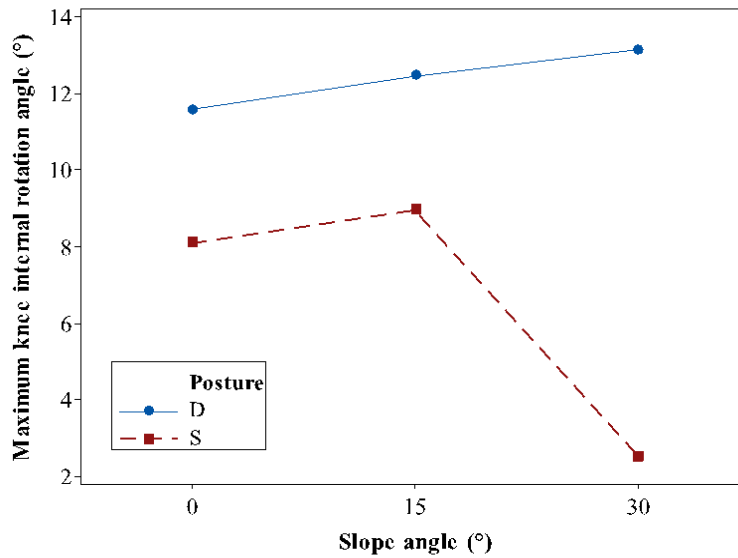


**Figure 2.8:** Interaction effect of slope and posture on right knee internal rotation

### 2.7.2. Main effects of roof slope

The significant main effect of slope was observed for the maximum knee flexion of left and right knee and adduction of right knee. The impact of slope on maximum knee flexion angle

implied that, with the increase in slope angle, the maximum knee flexion angle significantly decreased for left knee (Figure 2.9). From the Tukey post hoc analysis, we found significant differences in maximum knee flexion angles between three slopes for left knee. In Figure 2.9, we observed that, each of the three roof slopes (0°, 15° and 30°) was categorized in three different groups A, B and C. That means, each of these 3 roof slopes had significantly different impact from each other on the maximum knee flexion angle. On the other hand, for right knee, slopes 0° and 15° were categorized in the same group (A) whereas slope 30° was in another group (B), which means the maximum knee flexion angle at 30° slope was significantly lower compared to that at 0° and 15° slopes. However, there was no significant difference in maximum knee flexion between 0° and 15° slope.
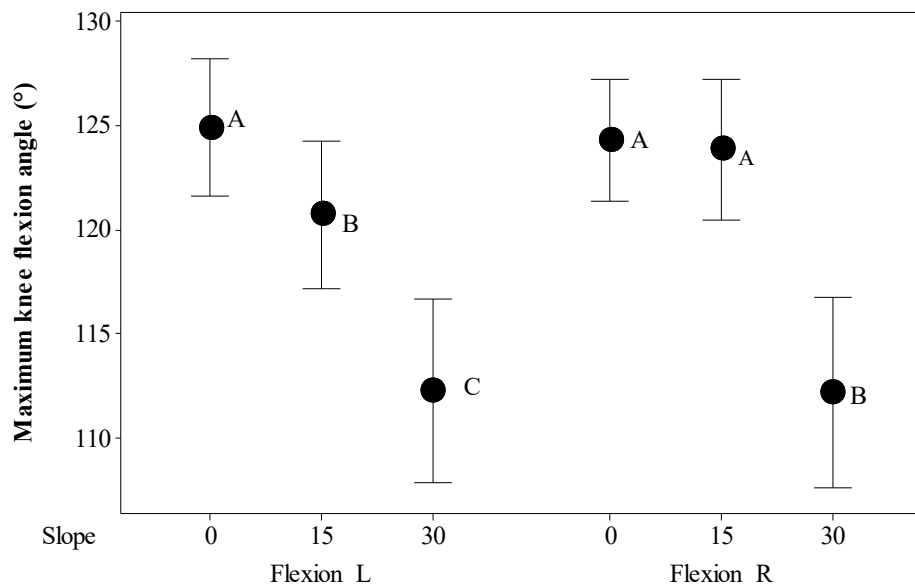


**Figure 2.9:** Effect of slope on maximum knee flexion angle (both for left and right knee). All observations are categorized into 3 groups by the 3 roof slopes; the mean at each roof slope is expressed by a dot, enclosed by the 95% confidence interval

Similarly, the maximum right knee adduction at 0° slope was significantly higher than that at 15° and 30° roof slopes (Figure 2.10). Here, negative value indicates adduction. For comparison, the absolute values were considered.
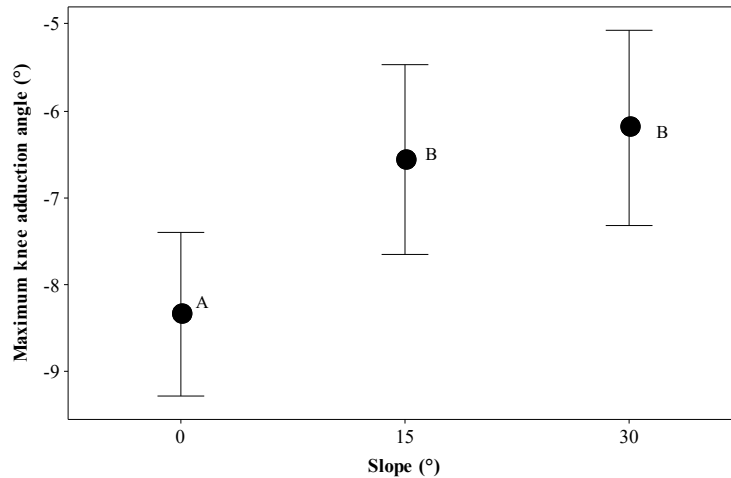


**Figure 2.10:** Effect of slope on maximum adduction of right knee

No significant main effect of slope was found for the maximum abduction angles in both knees, maximum adduction angle in left knee and maximum internal and external rotations in both knees.

### 2.7.3. Main effects of posture

The significant main effect of working posture was observed for the maximum knee flexions, adductions and internal rotations in both knees. Working posture was also found to be significantly associated with the abduction in right knee and external rotation in left knee.

From the group association of Tukey post hoc analysis in Figure 2.11, maximum left and right knee flexions were significantly higher in dynamic (D) posture than in static (S) posture.

Although no significant main effect of slope was observed in the maximum left knee abduction, for right knee, the maximum abduction was significantly higher in dynamic posture (Figure 2.12).
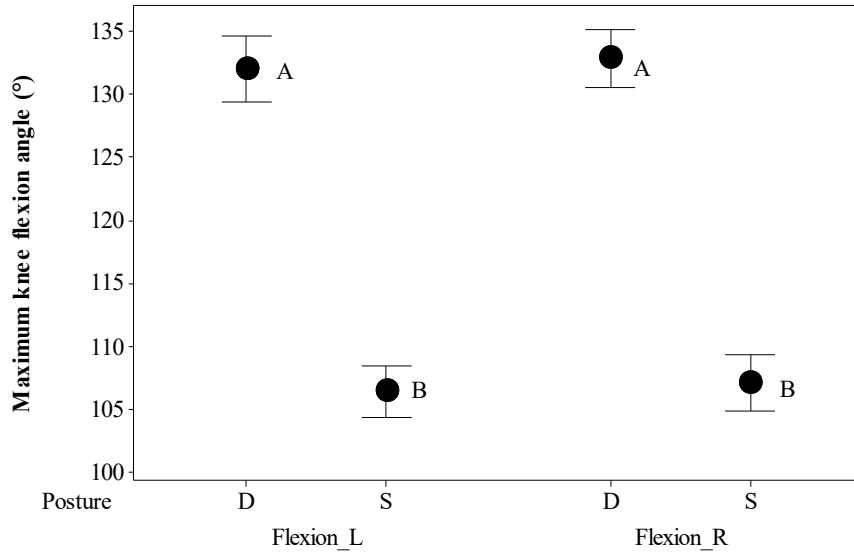
**Figure 2.11:** Effect of working posture in flexion angle
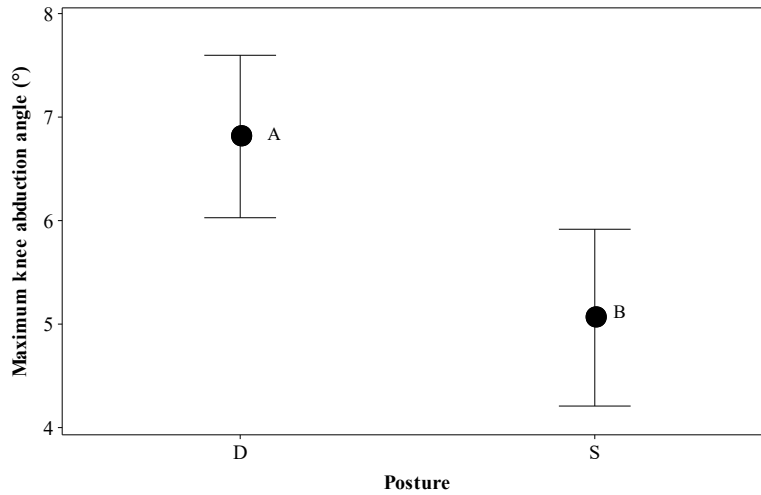


**Figure 2.12:** Effect of posture on the maximum abduction of right knee

On the other hand, maximum left and right knee adductions were significantly higher in dynamic posture (Figure 2.13) compared to the static posture (absolute values were considered for adduction).
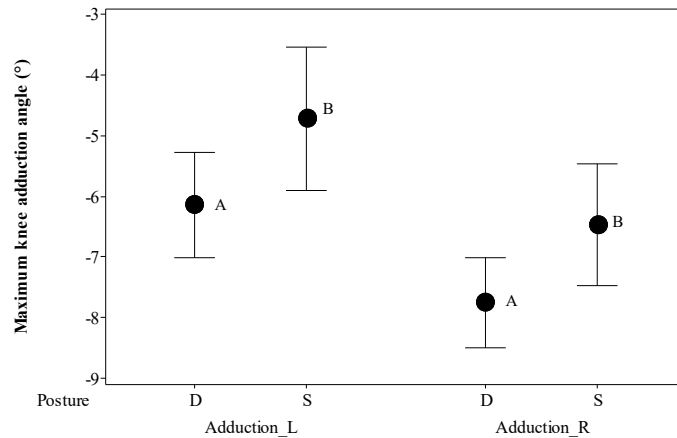
**Figure 2.13:** Effect of posture on the maximum knee adduction

For both knees, we found significantly higher maximum internal rotations in dynamic posture than in static posture (Figure 2.14). For left knee, external rotation was significantly higher in dynamic posture as well. Negative value indicates external rotation and for comparison, the absolute value was considered.
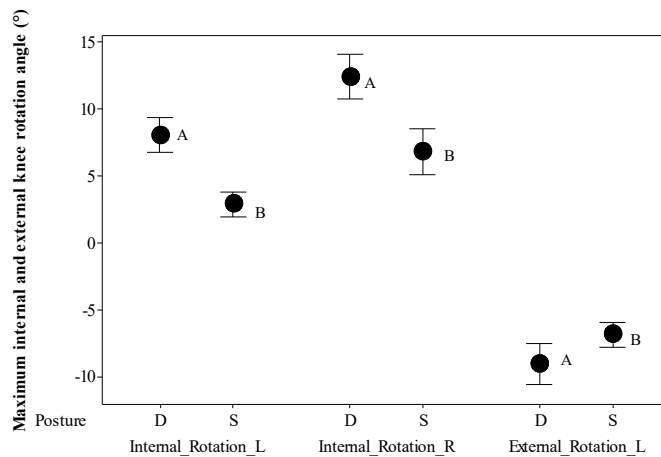


**Figure 2.14:** Effect of posture on knee internal and external rotation

## 2.8. Discussion

Our results demonstrated that roof slopes and working postures significantly impact knee rotational movement–flexion, abduction, adduction, internal and external rotation. As the slope

angle increases, the participants need to adjust their thighs to incline more towards the rooftop surface, which results in the decrease of the angle between the thigh and rooftop surface; that is, the knee flexion angle is decreased. This relationship is corroborated by our findings— as the slope increases, the knee flexion decreases. We found the lowest maximum flexion angle when the participants were working on 30° slope (left: 112.3°, right: 112.2°). Maximum flexion angles were greater in dynamic postures (left: 132.1°, right: 132.9°) compared to static posture scenarios (left: 106.5°, right: 107.1°). This may be explained by a subtle difference in the trunk alignment and working technique undertaken by the participants between static and dynamic postures. In the static posture in our experiment, the participants put their non-nailing hand on the roof surface and grabbed the nail gun in their nailing hand to install shingles—this caused the trunk to stay more inclined towards the roof surface and thereby the thigh to be in more upright position. But in the dynamic posture, the participants started simulated nailing of shingles, lifting their trunk upward which caused the thigh to bend and thereby the knee flexion angle to increase. Also, we observed an interaction effect of slope and posture on knee flexion of both knees. For all three roof slopes (0°, 15° and 30°), while the participants were in dynamic posture, both knees flexed significantly higher (left: 136.5°, 131.1°, 128.8°, right: 134.9°, 134.4°, 129.4° respectively) compared to when they were in static condition (left: 113.9°, 110°, 95.8°; right: 113.7°, 112.8°, 95.1° respectively).

According to our results, slope did not have a significant impact on left and right knee abduction. However, for different working postures the abduction in the right knee changed significantly. The reason could be due to the nature of the postures. In the static posture, the participants had negligible movement in the lower limbs as they were in a static position. But in dynamic posture, they simulated nailing of shingles which involved greater movement among the lower limbs. In the protocol, the participants simulated nailing 2 shingles side by side (3 nails in

each, from left to right). During the shingle installation phase, right shank moved laterally away and left shank moved towards the knee joint and caused abduction and adduction in the right and left knee respectively. Similarly, due to continuous lateral movement of the lower limb, right shank moved towards and left shank moved laterally away from the knee joint that caused adduction and abduction in right and left knees respectively. Although the impact of posture on the right knee abduction was significant, the impact was non-significant in the left knee abduction. One reason may be that 8 out of 9 participants were right-handed and they used their right hand for nailing activity which caused more lateral movement on the right knee compared to the left knee. Our analysis showed that, in dynamic posture, maximum right knee abduction was significantly higher (6.8°) compared to static posture (5.1°). Although there appeared to be no significant impact of roof slope and working posture on the left knee abduction, we found a significant interaction effect of slope and posture on the left knee abduction. For high-pitched roof (30° slope), our result showed that maximum left knee abduction became significantly higher in dynamic posture (12.3°) than in static posture (8.3°) which eventually increases knee MSD risk. For dynamic posture, the adduction in the left and the right knee were significantly higher (6.2°, 7.8°) than that in static posture (4.7°, 6.5°). In terms of roof slope, maximum knee adduction in the right knee was significantly higher in 0° roof slope (8.4°) than in 15° (6.6°) and 30° slope (6.2°), while no significant main effect of roof slope was found to cause adduction on the left knee.

Both knees internally rotated more in dynamic posture during shingle installation than in static posture. Statistically, significantly higher internal rotation in the right knee (13.2°) was observed at 30-degree slope during shingle installation than static condition (2.5°). There appeared to be no significant interaction effect of slope and posture on internal rotation of the left knee. In

addition, roof slope did not seem to have any significant impact on the internal rotation of either of the knees.

The left knee significantly rotated away in dynamic posture (9.1°) compared to static posture (6.9°). On the other hand, we did not find any significant effect of posture on the right knee external rotation during shingle installation. No significant interaction among slope and posture was observed to cause external rotation in any of the knees. Pollard et al. (2011) quantified and calculated knee angles during kneeling and squatting and revealed that posture has a significant impact on knee flexion and internal rotation. Our experiment also observed a significant impact of posture on flexion and internal rotation for both knees which echoes that finding.

Overall, this study demonstrated that the participants become exposed to potential knee MSD risks during shingle installation on rooftop. As higher knee flexion, abduction, adduction and rotation are required in dynamic kneeling, additional stress and force is induced on the knee joint ligaments. Conversely, a relatively lower knee flexion angle is required on a high-pitch roof than a low-pitch roof, and knee adduction in the right knee is also the minimum in high slope roof. But in dynamic posture, high-pitch roof causes higher abduction in left knee and internal rotation in right knee and hence, proper intervention is essential to minimize these awkward rotations while installing shingles on sloped roof surface. Some of the possible solutions might be using wearable or external devices, such as, knee pads, knee supporting devices and knee savers that have been proven to be useful in minimizing knee flexion during kneeling and squatting postures as suggested in previous studies (Dooley et al. 2019; Gray et al. 2015; Pollard et al. 2011). However, the focus of the current study was only the knee kinematics during a shingling process. Though several interventions are available, we did not want to alter the kinematics due to these interventions and hence were not utilized. In the future, both wearable and external interventions should be studied

to observe how their addition can influence the lower extremity kinematics during a shingling process on sloped roof surface. Also, the high knee muscles activation and knee joint contact stress- two common sources of knee MSDs, were out of scope of this study and hence were not analyzed.

This experimental study used nine participants to draw conclusions about the above findings. Typically, human musculoskeletal loadings are related to ergonomic injury risks and joint degeneration. The relationship between kinematics (i.e., joint angle measurements in this study) and the musculoskeletal loadings is modulated by biomechanics, not random. Such modulated relationship engenders that nine subjects are appropriate for this study and it is not necessary to have a large sample size to draw conclusions like those involving variables that are random. In fact, it is well accepted to validate biomechanical models using less than ten subjects in the research community. Besides, the kinematics data were collected from multiple trials for each subject to study the impact of slope, posture and their interactions on the response variables. Although from the statistical power analysis (G*Power, Franz Faul, Germany), some of the response variables statistical power might be low, but the statistical power does not diminish the clinical relevance. Moreover, the statistical power does not depend only on sample size but also on effect size of the factors which is calculated to understand the magnitude of the mean differences between different groups of the factors. If the effect size of a factor on the mean differences between two groups is large (Cohen's term $d > 0.2$), which was found in most of the mean differences in our study, it is possible to detect such an effect with small sample numbers (Sullivan and Feinn 2012).

In the process of selecting participants, this study did not control for handedness. 8 right-handed and 1 left-handed subjects were involved in this study. As lefties make up about 10% of

the population, this study coincidently represented the population. Considering that our long-term goal would be to develop some standardized recommendations, having ~10% of the participants as lefties is good practice. That being said, the results would probably have changed with a 50/50 split if there was more balance between left and right-handed individuals. Our first priority is to help with trainings and/or interventions. We could in the future develop an entire manuscript on the difference between the two handedness and determine if trainings/interventions need to be specialized depending on handedness.

**2.9. Limitations of the Study**

This study has several limitations. First, this study primarily focused on risks to knee injuries. As part of the lower extremity, the ankles may undergo excessive pressure while performing roofing activities in slanted roof settings. However, in this study associated risks to ankles were not examined. Second, maximum knee angles were used as a metric to assess the work-related risk factors of knee MSD in construction roofing works. Time series features of the knee angles data were not fully explored. Third, the participants in this study were not professional roofers. It is possible that professional roofers can adopt posture by optimizing the risks based on their years of experience and from the perspective of kinematics, there should be some differences between professional roofers and non-professional roofers. However, this study intended to identify the initial extreme posture individuals without any prior roofing experience encounter when they are first exposed to a sloped working environment, so that knowledge built upon the findings can be used to train new roofers and intervention development.

**2.10. Conclusion and Future Research**

This study analyzed the impact of the unique sloped work setting—commonly encountered by residential roofers during performing construction roofing works—and different working

techniques for kneeling posture on knee MSD risks in roofing activities. It provided useful information to statistically understand the effects of work-related risk factors on knee MSDs while performing shingle installation on a sloped rooftop. For example, dynamic kneeling posture during shingle installation may generate significantly higher knee flexion, abduction, adduction, internal and external rotation compared to the knee rotations required to maintain a static posture, which provides the evidence of increased knee MSD risk in sloped shingle installation task. High-pitched roof can reduce knee flexion and adduction, but it may induce high knee abduction and internal rotation during dynamic shingle installation. Proper intervention is required during shingle installation and roof slope should be such that they induce minimum flexion, abduction and internal rotation of knee which is important to reduce knee MSD risk caused by construction roofing works.

Future studies on safety interventions will involve experiments in construction sites and participation of a larger number of professional roofers to provide more comprehensive understanding of work setting-related knee disorder risks among construction roofers. Moreover, work-related factors, such as different slope and posture will be assessed for knee MSDs observing the EMG signals of knee flexor and extensor muscles and knee joint contact stress. Also, by considering the spatiotemporal nature of the data set, cumulative effects of the knee bending angles will be analyzed and relative risks at different phases of the entire shingle installation task will be examined. The findings of this study advance the construction knowledge by providing insights of roof work-related conditions on knee MSD development and thereby helping reducing knee MSD risks of roofers through developing training and ergonomic solutions.

# CHAPTER 3: EFFECTS OF WORKING POSTURE AND ROOF SLOPE ON ACTIVATION OF LOWER LIMB MUSCLES DURING SHINGLE INSTALLATION

**Abstract**

Awkward and extreme kneeling during roofing generates high muscular tension which can lead to knee musculoskeletal disorders (MSDs) among roofers. However, the combined impact of roof slope and kneeling posture on the activation of the knee postural muscles and their association to potential knee MSD risks among roofers have not been studied. The current study evaluated the effects of kneeling posture and roof slope on the activation of major knee postural muscles during shingle installation via a laboratory assessment. Maximum normalized electromyography (EMG) data were collected from knee flexor and extensor muscles of seven subjects, who mimicked the shingle installation process on a slope-configurable wooden platform. The results revealed a significant increase in knee muscle activation during simulated shingle installation on sloped rooftops. Given the fact that increased muscle activation of knee postural muscles has been associated with knee MSDs, roof slope and awkward kneeling posture can be considered as potential knee MSD risk factors.

## 3.1. Introduction

Awkward kneeling posture is considered as a primary risk factor for musculoskeletal disorders (MSDs) among occupations that require frequent kneeling (Xu et al. 2017). Due to the unique work condition of slanted rooftops, residential roofers spend more than 75% of their working time in crawling, squatting, stooping and kneeling postures. The cumulative effects of these awkward postures combined with repetitive motions have, in large part, led to a high incident rate of MSDs among residential roofers (Dulay et al. 2015; Wang et al. 2015). Awkward postures during a task can lead to less efficient force production in skeletal muscle. This decrease in muscle efficiency may result in higher muscle activation and muscle overloading compared to a neutral posture (Kaushik and Charpe 2008). Cumulative muscle overloading coupled with repetitive motions without adequate recovery time may cause MSDs due to overexertion or imbalance (Hofer et al. 2011; Kumar 2001). According to Marras and Karwowski (2003), the incident rate of knee MSDs is the highest among residential roofers in comparison to other workers in construction.

Studies have been conducted on assessments of MSD risk factors for roofers. Lu et al. (2015) investigated the effects of roof slope, visual cue, muscular fatigue, and task performance on the lower limb postural muscle activation associated with balance maintenance. Wang et al. (2017) examined the influences of roof slope, working technique, and working pace in kneeling and stooped postures on the development of low back disorders to roofers. The effects of roof slope and kneeling posture on the knee kinematics of roofers were investigated by Breloff et al. (2019a). The impacts of traverse walking across a sloped roof surface on lower extremity kinematics of roofers were examined by Breloff et al. (2019b). Risky phases in terms of awkward knee rotations and repetitive motions during shingle installation were assessed by Dutta et al. (2020a). These studies have greatly advanced understanding of MSD risks in roofing by examining

the association of the roofing work-related factors to the potential MSD development among roofers.

Working posture has been identified as a vital mechanical variable to account for occupational safety and used to compute muscle activation and joint loads (Tennant et al. 2018; Tennant et al. 2014). Residential roofers typically kneel on the sloped rooftop during roofing work. These awkward kneeling postures are associated with an increased risk of knee MSDs such as meniscal disorders and knee osteoarthritis (Canetti et al. 2020; Gallagher et al. 2011). However, the effect of such constrained awkward kneeling postures on the peak muscle activation of lower extremities in residential roofing and their association to knee MSDs has not been studied. Furthermore, increasing the inclination of working surface has been identified as a contributing factor for muscular loading in the lower limb muscles (Lu et al. 2015). In a kneeling posture, while it was revealed that lumber spine-muscle activation increases at a higher-pitched roof surface during roofing (Wang et al. 2017), the impact of roof slope on the activation of knee postural muscles during roofing is still unknown. Despite the advances in the previous studies on MSDs in the residential roofing work-related factors, there is limited knowledge regarding the activation of associated muscles of the lower limbs during kneeling in residential roofing on a sloped surface, which are associated to the osteoarthritis of knee joints.

The objective of the current study was to assess the impact of two residential roofing work-related factors —roof slope and awkward kneeling posture — on the activation of knee flexor and extensor muscles during the repetitive shingle installation task. This study hypothesized that roof slope, kneeling posture, and their interaction would significantly impact the activation of the knee postural muscles during performance of roofing, which may lead to knee MSDs among construction roofers.

41

### 3.2. Methods

### 3.2.1. Participants

The participants were the same as mentioned in Chapter 2. However, only valid EMG data from seven out of the nine participants were available and used for this analysis. The protocol was approved by both the Institutional Review Boards (IRBs) of West Virginia University (WVU) and National Institute for Occupational Safety and Health (NIOSH).

### 3.2.2. Instruments

Muscle activation signals were recorded using a surface electromyography (EMG) system (Noraxon Desktop Direct Transmission System with myoMUSCLE Master software, Arizona, USA). According to the instructions outlined in Reichert and Stelzenmueller (2011), surface EMG Ag/AgCl electrodes were placed bilaterally on the palpated muscle bellies of each of three extensor muscles [rectus femoris (RF), vastus lateralis (VL) and vastus medialis (VM)] and two flexor muscles [biceps femoris (BF) and semitendinosus (S)]. The EMG data were collected at a rate of 1,000 Hz.

Knee kinematics data (segment endpoints) were also collected using the VICON optical motion analysis system with 14 MX Vicon cameras (Oxford, UK). Retroreflective motion capture markers were placed in participants' hip and knee joints, thighs and lower legs to capture their three-dimensional (3D) coordinates and derive knee flexion angles (Breloff et al. 2019a).

Simulated shingle installation was conducted on the same wooded platform that was depicted in Figure 2.2 in Chapter 2.

### 3.2.3. Experimental design

The experimental design was comprised of two independent variables: roof slope and kneeling posture. The roof slope included three levels: 0°, 15° and 30°. The kneeling posture included two levels: static posture and dynamic posture.

The static posture is a deep-flexed kneeling position where the roofers flex their knees and trunk, place their non-nailing hand on the rooftop, and hold the nail gun with the opposite hand with negligible movement in their lower limbs. This posture represents the general body orientation of the roofers during shingle installation. It can be considered as a non-working condition considering that they neither place nor nail shingles in this posture.

The dynamic posture is a working technique that involves generating cyclic muscle contraction and relaxation as the residential roofers perform the entire shingle installation task. First the residential roofers reach for and grasp the shingles by lifting the trunk slightly upward. Then they place the shingles on the rooftop. Next, they pick up the nail gun from the side and begin nailing the shingles side by side. Following nailing, the nail gun is returned to its starting position; and finally, the roofers return to the upright position resting on their knees.

The dependent variables were peak normalized EMG of the aforementioned ten muscles.

### 3.2.4. Procedures

The experiment was performed at the biomechanics laboratory, NIOSH (Morgantown, WV). The experiment procedure was introduced to the participants upon their arrival and they read and signed the informed consent forms. The EMG electrodes and motion markers were then placed on the designated locations of the participants. The skin areas for the EMG electrodes were shaved and cleaned before the placement of the electrodes. After that, the participants mimicked the

shingling task on the wooden platform using different combinations of postures and slopes (2 postures × 3 slopes = 6 scenarios). In the static kneeling, all participants placed their left hand on the wooden platform and held the nail gun with their right hand. In the dynamic kneeling, they secured six nails (three in each) into two shingles side by side on the roof segment with a pneumatic nail gun (weight 5.8 pounds). The six scenarios were randomized by slope and then posture for each participant. For each combination of the randomized slope and posture, each participant performed five repeated trials.

### 3.2.5. Data processing

Collected EMG signals were preprocessed using Visual 3D (C-Motion, Inc). First, the raw EMG signals were filtered using a 2nd order Butterworth high-pass filter with a cutoff frequency of 20 Hz and a low-pass filter with a cutoff frequency of 500 Hz. Then, the signals were rectified by a full-wave rectifier. Next, short-term fluctuations were removed from the signals with a moving average filter that had a three data point window. The value of the resulting signal at each point was the average of the values in the window of three data points.

Following these, the preprocessed EMG signals were normalized with respect to a reference EMG measurement for each participant to ensure fair comparison. This was done using the method suggested by Chapman et al. (2010) and described as below.

Let the EMG signal of certain muscle of the left or right knee observed for participant $i$ at slope $s$, posture $p$, trial $j$, and time $t$ be represented as $E_{i,s,p,j}(t)$,

where $i \in \{1, 2, 3, 4, 5, 6, 7\}$ representing each one of the participants,

$p \in \{1, 2\}$ with 1 and 2 representing the static and dynamic postures, respectively,

$s \in \{1, 2, 3\}$ with 1, 2 and 3 representing 0°, 15° and 30° slopes, respectively, and

$j \in \{1, 2, 3, 4, 5\}$ representing each one of the trials.

For participant $i$, the reference EMG measurement was calculated using the following equation:

$$R_i = \frac{\sum_{s=1}^{3} \sum_{p=1}^{2} \sum_{j=1}^{5} \max_t (E_{i,s,p,j}(t))}{N} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (3.1)$$

where $N$ is the total number of trials for each participant. In this study, $N = 30$ for the three slopes and two kneeling postures with five repeated trials for each combination.

Consequently, the normalized EMG signal of certain muscle for participant $i$ at slope $s$, posture $p$, trial $j$, and time $t$ was computed using the following equation:

$$NE_{i,s,p,j}(t) = \frac{E_{i,s,p,j}(t)}{R_i} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (3.2)$$

For each participant, the maximum value of the normalized EMG signal within each trial was selected as the dependent variable.

### 3.2.6. Data analysis

The assumptions of analysis of variance (ANOVA) (i.e., constant variance of residuals, normality of residuals, and independence of observations) were examined before analyses using a graphical approach (Freund et al. 2010). Multivariate ANOVA (MANOVA) was then applied to evaluate the effects of the two independent variables (i.e., roof slope and posture) and their interaction on all the dependent variables. The independent variables that demonstrated significant effects were further analyzed using univariate ANOVA. Tukey post-hoc analysis was performed for any independent variable with three or more levels (i.e., slope) to explore the effect differences between levels. The $p$-value was set as 0.05 for all tests. All tests were performed in Minitab 19 (Minitab, Inc.)

### 3.3.Results

Tables 3.1 and 3.2 show the *p*-values illustrating the bilateral effects of the working conditions on the muscle activation. As the effects of the working conditions were different for both knees, the results were presented and discussed separately. Here, LBF, LRF, LS, LVL and LVM denoted the BF, RF, S, VL and VM muscles of the left knee, respectively. RBF, RRF, RS, RVL and RVM denoted the BF, RF, S, VL and VM muscles of the right knee, respectively.

**Table 3.1:** The resulting p-values for the left knee

| Independent variables | MANOVA | ANOVA | | | | |
|---|---|---|---|---|---|---|
| | | LBF | LRF | LS | LVL | LVM |
| Slope | **<0.001** | 0.343 | **0.001** | **0.011** | **<0.001** | **<0.001** |
| Posture | **<0.001** | **<0.001** | **<0.001** | **<0.001** | **<0.001** | **<0.001** |
| Slope*Posture | **<0.001** | 0.121 | 0.676 | 0.091 | 0.118 | **0.019** |

*Note: Bold font indicates a significant effect.*

**Table 3.2:** The resulting p-values for the left knee

| Independent variables | MANOVA | ANOVA | | | | |
|---|---|---|---|---|---|---|
| | | RBF | RRF | RS | RVL | RVM |
| Slope | **<0.001** | **0.007** | 0.684 | 0.776 | **0.007** | **<0.001** |
| Posture | **<0.001** | **<0.001** | **<0.001** | **<0.001** | **<0.001** | 0.056 |
| Slope*Posture | **<0.001** | **0.002** | 0.718 | **0.015** | 0.107 | 0.160 |

*Note: Bold font indicates a significant effect.*

### 3.3.1.  Main effects of roof slope

The *p*-values obtained from the ANOVA test implied that slope had a significant impact on the maximum normalized activation of the LRF, LS, LVL and LVM muscles and the RBF, RVL and RVM muscles (Tables 3.1 and 3.2). Tukey post-hoc analysis results plotted in Figure 3.1 illustrates that for the LRF muscle, the 0° slope was categorized in one group (denoted as group A) and the 15° and 30° slopes were categorized in another group (denoted as group B). This meant

that the effect at 0° roof slope (group A) on the maximum normalized muscle activation (MNMA) was significantly different from the effect at 15° and 30° slopes (group B). From the trend of the graph for the LRF muscle, the MNMA at 15° and 30° slopes was significantly higher than that at 0° slope. The levels of the independent variable sharing a common group (e.g., 15° and 30° slopes in group B for LRF) indicated that there was no significant difference in their effects on the dependent variables (i.e., MNMA). For the RBF muscle, the only significant difference that existed was between 15° (group A) and 30° slopes (group B) where the MNMA at 30° slope was significantly higher than that at 15° slope. Slope did not have any significant effect on the MNMA of the LBF and RRF muscles.
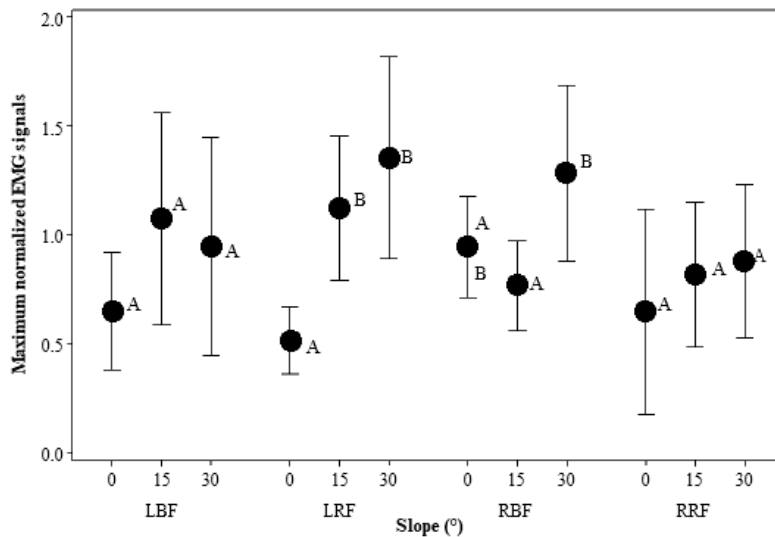


**Figure 3.1:** Average maximum normalized EMG signals of the LBF, LRF, RBF and RRF muscles with 95% confidence intervals. Results were averaged by the three roof slopes. Here, 'A' and 'B' illustrate significant and non-significant effect differences between levels of slope.

For the LVM and RVL muscles, the MNMA at 30° slope (group B) was significantly higher than that at 0° slope (group A) (Figure 3.2). For the LVL muscle, the MNMA at 30° slope (group B) was significantly higher than that at 0° and 15° slopes (group A). For the RVM muscle,

47

the MNMA at 15° and 30° slopes (group B) was significantly higher than that at 0° slope (group A).
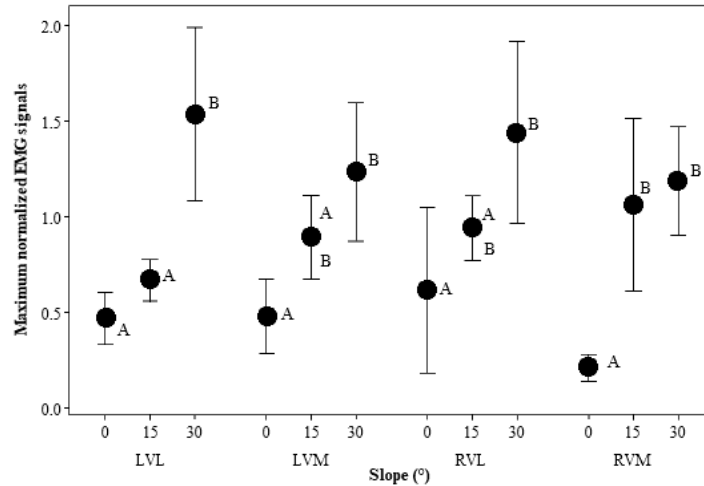


**Figure 3.2:** Average maximum normalized EMG signals of the LVL, LVM, RVL and RVM muscles with 95% confidence intervals. Results were averaged by the three roof slopes. Here, 'A' and 'B' illustrate significant and non-significant effect differences between levels of slope.
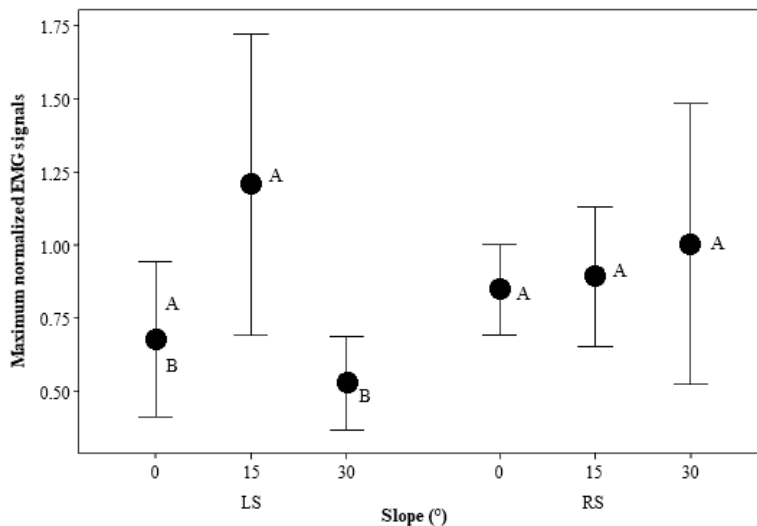


**Figure 3.3:** Average maximum normalized EMG signals of the LS and RS muscles with 95% confidence intervals. Results were averaged by the three roof slopes. Here, 'A' and 'B' illustrate significant and non-significant effect differences between levels of slope.

For the LS muscle, the MNMA at 30° slope (group B) was significantly lower than that at 15° slope (group A). For the RS muscle, no significant difference was observed between any two of the roof slopes (all in group A) (Figure 3.3).

### 3.3.2. Main effects of posture

For both knees, the MNMA of all muscles but RVM was significantly associated with posture. For all muscles, the MNMA was significantly higher in the dynamic posture than in the static posture. It indicates that higher muscle recruitment was required during shingle installation than what was required to maintain a static flexed kneeling posture. The effect of posture on different postural muscles are depicted in Figures. 3.4, 3.5 and 3.6.
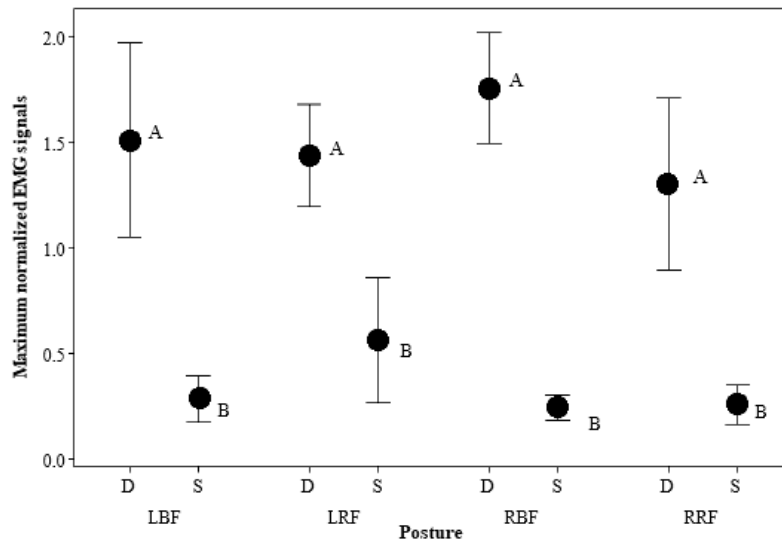


**Figure 3.4:** Average maximum normalized EMG signals of the LBF, LRF, RBF and RRF muscles with 95% confidence intervals. Results were averaged by the two kneeling postures. Here, 'A' and 'B' illustrate significant and non-significant effect differences between level of posture. 'D' and 'S' denote dynamic and static postures, respectively.
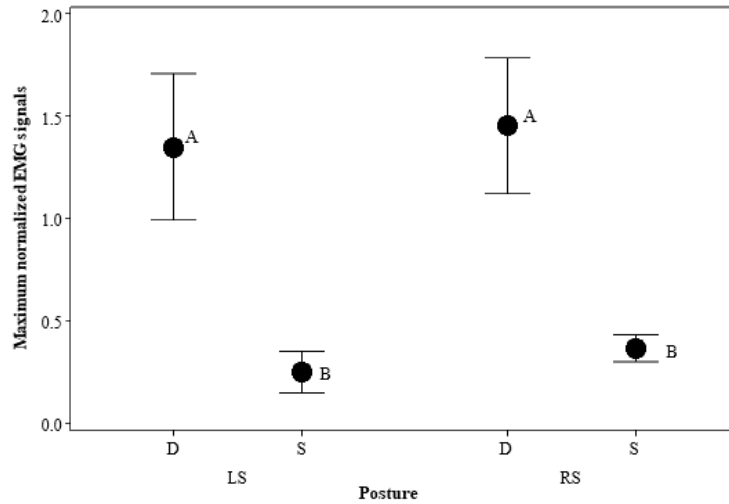
**Figure 3.5:** Average maximum normalized EMG signals of the LS and RS muscles with 95% confidence intervals. Results were averaged by the two kneeling postures. Here, 'A' and 'B' illustrate significant and non-significant effect differences between levels of posture. 'D' and 'S' denote dynamic and static postures, respectively.
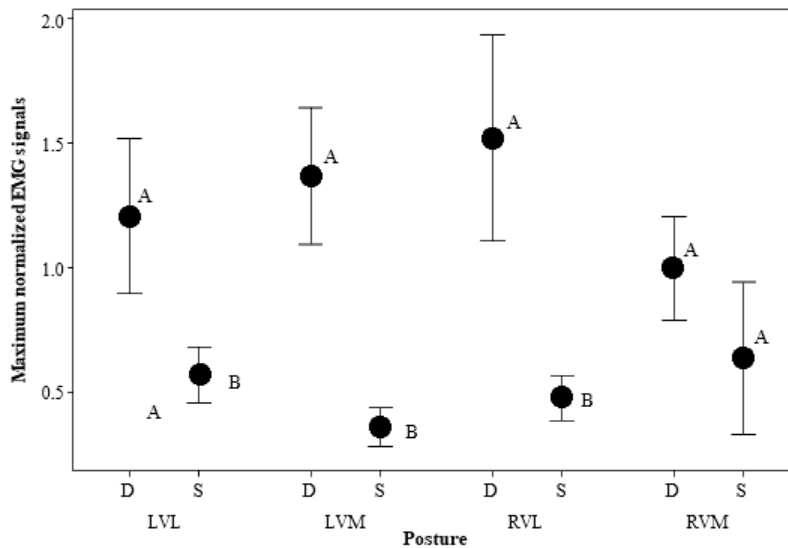


**Figure 3.6:** Average maximum normalized EMG signals of the LVL, LVM, RVL and RVM muscles with 95% confidence intervals. Results were averaged by the two kneeling postures. Here, 'A' and 'B' illustrate significant and non-significant effect differences between levels of posture. 'D' and 'S' denote dynamic and static postures, respectively.

### 3.3.3. Interaction effects of roof slope and posture

As indicated in Tables 3.1 and 3.2, there was significant interaction effects between slope and posture on the MNMA of the LVM, RBF and RS muscles. Due to these results, at each slope, the maximum normalized EMG signals in the static and dynamic postures were compared for these muscles. For each posture, differences between any two of the three slopes were compared as well.

In Figure 3.7, at 0°, 15° and 30° slopes, the MNMA in the dynamic posture was significantly higher than that in the static posture for the RBF muscle. In the dynamic posture, the MNMA at 30° slope was significantly higher than that at 0° and 15° slopes. No significant difference was observed between any two slopes in the static posture.
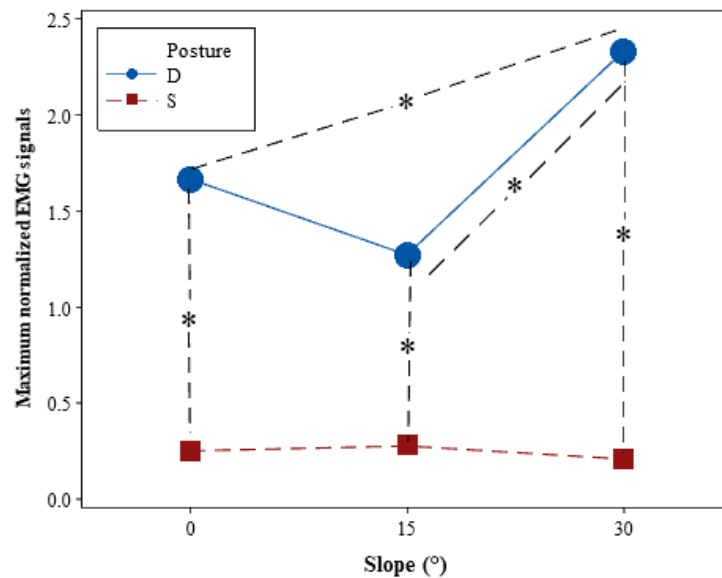


**Figure 3.7:** Interaction effect of slope and posture on the maximum normalized activation of the RBF muscle. Here, 'D' and 'S' denote dynamic and static postures, respectively. '*' indicates statistical significance.

In Figure 3.8, for the RS muscle, at 15° and 30° slopes, the MNMA in the dynamic posture was significantly higher than that in the static posture. In the dynamic posture, the MNMA at 30°

slope was significantly higher than that at 0° slope. No significant difference was observed between any two slopes in the static posture.
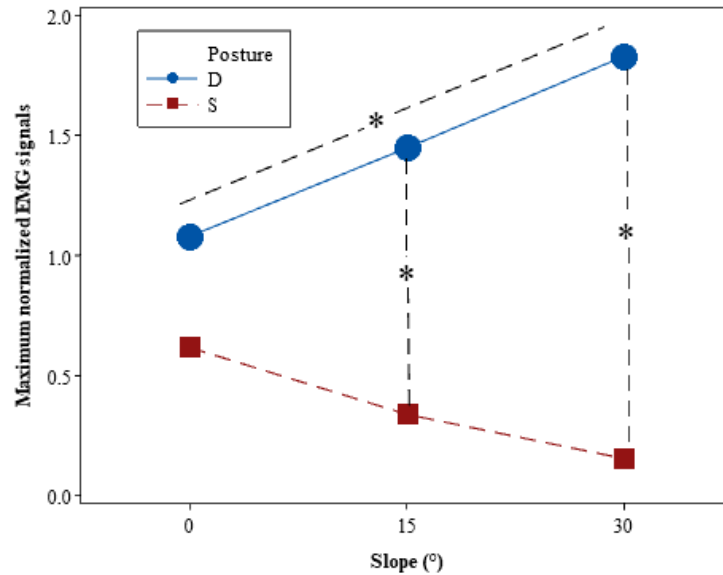


**Figure 3.8:** Interaction effect of slope and posture on the maximum normalized activation of the RS muscle. Here, 'D' and 'S' denote dynamic and static postures, respectively. '*' indicates statistical significance.
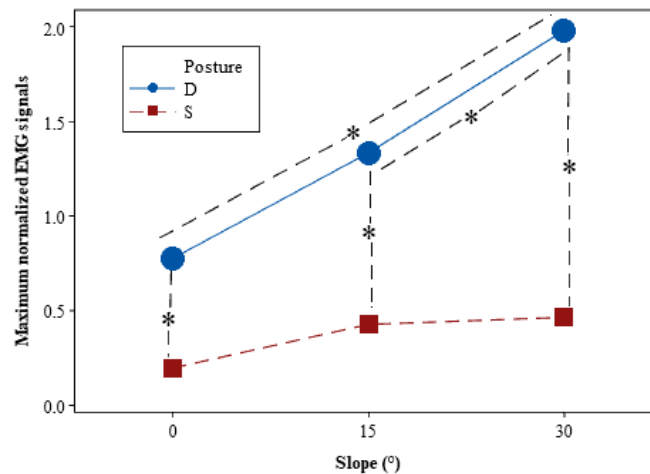


**Figure 3.9:** Interaction effect of slope and posture on the maximum normalized activation of the LVM muscle. Here, 'D' and 'S' denote dynamic and static postures, respectively. '*' indicates statistical significance.

In Figure 3.9, for the LVM muscle, at each slope, the MNMA in the dynamic posture was significantly higher than that in the static posture. During shingle installation in the dynamic posture, the MNMA at 30° slope was significantly higher than that at 0° and 15° slopes. No significant difference was observed between any two slopes in the static posture.

### 3.4. Discussion

The current study assessed the effects of roof slope and kneeling posture on the maximum normalized activation of knee postural muscles in sloped residential roof shingle installation. Both the static and dynamic kneeling postures were studied as potential knee MSD risk factors because the peak activation of the muscle can be obtained in either of these scenarios (Andriacchi and Favre 2014). The five bilateral muscles were included in the current study, as they are responsible for the flexion and extension of each knee joint during kneeling and the increased activation of these muscles is potentially associated with MSD developments (Kingston et al. 2016).

### 3.4.1.  Effects of roof slope

In general, at a high-pitched roof (30°), the knee extensor muscles of the left knee (LRF, LVL, LVM) and the right knee (RVL and RVM) exhibited a statistically significant increase in the maximum normalized activation. The possible reason for higher muscle activation during shingle installation on a sloped surface could be explained by the muscle length-tension relationship (Gordon et al. 1966), which relates the isometric contraction force to the muscle length, at which the contraction occurs. According to this relationship, muscles operate with high active tension when close to the optimal resting length (or a posture around 90° of knee flexion). But when the muscle is lengthened too much (e.g., flexion greater than 140°) or shortened too much (e.g., flexion less than 80°), the isometric active tension generated in the muscle decreases because the less than optimal force is produced due to the insufficiency of the actin-myosin

filament overlap and binding site availability within the muscle. When residential roofers knelt on the sloped surface during shingle installation, it was observed that the knee flexion angle, on average, ranged from ~110° to 125°, and with the increase in roof slope (between 0° to 30°), the flexion angle decreased by 12° to 15°. Due to the decrease in knee flexion, the knee extensor muscles might contract concentrically and the flexor muscles might contract eccentrically as residential roofers tended to incline a bit to the roof surface to maintain balance (Kennedy and Cresswell 2001). As the length of the muscles deviated from the optimal resting length during shingle installation on the sloped roof surface, the ability of producing the maximum active tension in the muscles decreased. In this situation, the muscle activation was triggered by the nerve simulation, which promoted the recruitment of more motor neurons that, in turn, stimulated the muscle fibers and resulted in the generation of the required muscle force to perform the shingle installation task.

For the knee flexor muscles, although the maximum normalized activation of the RBF muscle significantly increased at high-pitched roof, the maximum normalized activation of the LS muscle decreased. According to the length-tension relationship, tension development in a muscle increases as the resting length of the muscle increases up to an optimal length. Tension development then decreases with further increase in the muscle length. Although the length of the knee flexor muscle LS increased at high-pitched roof due to decrease in knee flexion, it might be possible that the muscle was still within the range of optimal resting length when the active tension generation capability of the muscle increased and hence the activation decreased (since the muscle had the required active tension available to perform the task). Another possible reason was that all the participants were right-handed and they used their right hands for grasping the nail gun, reaching for shingles, placing and nailing them. This might have caused more repetitive lateral

movement of the right knee compared to the left knee and contributed to higher cyclic muscle contraction and elongation of the right knee flexor muscle (RBF). Onishi et al. (2002) studied the differences in activities among three flexor muscles during maximum voluntary isometric and isokinetic knee flexion and reported that the activation of the S muscle decreased, and the BF muscle increased as flexion decreased, which corroborated our findings. However, further investigation is needed to substantiate the explanation.

### 3.4.2. Effects of posture

For most of the muscles tested in this study, higher muscle recruitment was required during shingle installation than what was required to maintain a static flexed kneeling posture. The possible reason could be explained by the knee flexion angles generated in the static and dynamic postures. It was observed that the knee flexion in the dynamic posture was higher (~132°) than that in the static posture (~106°). This might cause concentric contractions to the flexor muscles as the muscles were shortened due to increased flexion during shingle installation. Meanwhile, eccentric contractions were generated to the extensor muscles as they were lengthened. Thanks to this deviation from the ideal or optimal resting length, the ability to produce the maximum active tension by these muscles decreased leading to possible requirement for more muscle recruitment during shingle installation. Our findings matched with the findings from Kingston et al. (2016), who measured the peak activation of lower limb muscles during high flexion static kneeling and troweling movements in kneeling postures and reported significant increase in muscle activation during the movement task.

### 3.4.3. Interaction effects of roof slope and posture

For low to high-pitched rooftops (0°, 15° and 30°), the maximum normalized activation of the LVM, RBF and RS muscles was significantly higher in the dynamic kneeling posture than in

the static kneeling posture. In addition, the maximum normalized activation of these muscles was significantly higher at high-pitched roof slope (30°) compared to two other roof slopes (0° and 15°) at the dynamic kneeling posture. The possible reason for the higher muscle activation during shingle installation on high-pitched rooftop could be the extra effort required to maintain balance on the high-pitched rooftop by accounting for the postural variances involved in the dynamic posture. Also, the instability caused by the dynamic posture and the reduction of the base of support with increase of roof slope might be a potential reason for heightened muscle activation, because more work by the muscles was required to stabilize the trunk. In addition, simultaneous contraction of knee flexor and extensor muscles to help with knee joint instability on sloped-surface might contribute to the heightened muscle activation during sloped-shingle installation. Nevertheless, further studies are required to investigate the exact muscle physiology and their association to heightened muscle activation during sloped-shingle installation.

### 3.5. Limitation of Study

As with all laboratory studies, there are limitations which are important to note. First, the experiment was conducted with seven participants. Typically, muscle activation is associated with the generation of joint contact forces, and thereby is relatable to MSDs. There are biomechanical reasons behind the association between muscle activation and the musculoskeletal loadings. Based on this reasonable association, seven subjects are appropriate for this study and it is not necessary to have a large sample size to draw the conclusion like those from random variables. Historically, biomechanical models have been validated using less than ten subjects in literature (Ha and Han 2017; Lay et al. 2007; Li et al. 2017).

Second, all the shingles were initially placed at the right side of the participants. They reached for those shingles first and then placed them in the front. Next, they grabbed the nail gun

from their right side and nailed shingles. Once done, they replaced the nail gun to their right side. Therefore, it could be expected that the right knee muscles did more work. Since there is no standard procedure for shingle placement, and most people tend to be right hand dominant, shingles were placed on the dominant side on an account that it is what most roofers would do in the workplace.

Third, the participants were not professional roofers. From the perspective of biomechanics, there should be differences in the kinematics between professional roofers and non-professional roofers that can impact the muscle activation. While differences exist, knee MSD incident rates are the highest among construction roofers. By studying the effects of the work-related factors on the muscle activation of novice workers, this study intended to know the initial risk that a roofer without any prior experience could get exposed to during sloped shingle installation. The findings of this experimental framework can be used to develop investigations for training new roofers and/or interventions, so that the heightened muscle activation can be minimized, and MSD risks can be reduced in roofing activities. For example, knee flexion has impacts on the muscle contraction and thereby on the muscle activation. So, roofers should adjust their knee flexion (not less than 80° and not greater than 140°) during shingle installation on sloped roof surfaces. Using knee savers and knee pads might be helpful in this regard because they can reduce the peak lower extremity kinematics during sloped shingle installation (Breloff et al. 2019c). Moreover, knee savers can reduce cumulative muscular effort and fatigue during prolonged kneeling (Pejhan et al. 2019). A known level of knee-muscle activation at different work settings and kneeling postures will also help understand the mechanism related to the onset of knee MSDs (Nagura et al. 2006) and develop knee joint biomechanical models for computing in-vivo muscle and contact forces in different occupational tasks (Lin et al. 2010).

### 3.6.Conclusion and Future Research

This study examined the impact of roof slope and awkward kneeling posture — two common residential roofing work-related factors — on the peak normalized activation of knee flexor and extensor muscles as potential risk factors of knee MSDs among construction roofers. The findings suggested that roof slope, awkward kneeling posture, and their interaction all have an association with the peak normalized activation of the knee postural (flexor and extensor) muscles, implying that they are potential risk factors of knee MSDs. The findings also suggested that roofers become exposed to a greater risk of developing knee MSDs with the increase of roof slope during shingle installation as the dynamic kneeling posture during shingle installation on high-pitched rooftops requires significantly higher muscle loading for task performance compared to a static flexed kneeling posture. Therefore, the heightened muscle activation while kneeling during shingle installation on high-pitched rooftop should be given particular attention.

In the future, to provide more comprehensive understanding of knee MSD risks among construction roofers, assessments of roofing-related factors will be performed by observing the knee joint contact force and the ground reaction force captured by force plate. Future studies will also include testing of potential interventions such as knee pads and roofing footwear on the muscle activation with the participation of professional roofers in a real construction site.

# CHAPTER 4: IDENTIFYING POTENTIALLY RISKY PHASES LEADING TO KNEE MUSCULOSKELETAL DISORDERS DURING SHINGLE INSTALLATION OPERATION

**Abstract**

Repeated and prolonged awkward kneeling can result in musculoskeletal disorders (MSD) in construction roofers. However, task-specific risk assessment for roofers' knee injuries is still missing in the literature. This study identified a ranking-based ergonomic method for suggesting potentially risky phases that may increase knee MSD risk during shingle installation operations. On a slope-adjustable wooden platform in a laboratory setting, nine subjects performed shingle installations that included seven phases: i) reaching for shingles, ii) placing shingles, iii) grabbing the nail gun, iv) moving to first nailing position, v) nailing shingles, vi) replacing the nail gun, and vii) returning to upright position. Flexion, abduction, adduction, internal and external knee rotations were measured to assess relative risks of these phases by ranking them using a scoring model. The ranking results revealed that the phases of placing shingles and nailing shingles lead to the most knee MSD risk exposure, and awkward flexion, abduction and adduction involved in these phases can significantly contribute to the potential knee MSD risk measurement. Using the ranking-based method, this study suggested that certain phases of the shingle installation process may increase knee MSD risk, which is useful for developing effective interventions to reduce knee MSD risk exposures from roof shingle installation.

## 4.1. Introduction

Shingle installation is a prolonged repetitive and awkward task that residential roofers commonly preform on the jobsite. Shingle installation involves awkward crawling, stooping or kneeling postures and repetitive motions that result in roofers' knee MSD including chronic knee pain, knee joint irritation (i.e., bursitis), and osteoarthritis (Dulay et al. 2015). As roofers spend more than 75% of their total working time being restricted to awkward postures and repetitive motions in a sloped roof setting, they suffer from a high incidence rate of MSD (CPWR 2018). It has been shown that knee awkward postures and repetitive motions are associated with knee MSD (Hofer et al. 2011). Nevertheless, it is still unknown which phases of the shingle installation operation might yield the most awkward postures and repetitive motions that result in potentially the greatest knee MSD risk on slanted roof surfaces. A detailed ranking of the typical kneeling shingle installation phases based on the awkward postures and repetitive rotations can provide insights about the association of these phases with the knee MSD risk development. Although ranking-based methods were applied to study risks in construction, associated with schedule delays (Bagaya and Song 2016), life-cycle of green buildings (Qin et al. 2016), and highway construction projects (El-Sayegh and Mansour 2015), they are yet to be applied in investigations of work related MSD risks. It would be helpful to rank the phases in a sloped kneeling shingle installation process based on the MSD risk caused by the awkward postures and repetitive motions. Such a ranking could reveal the shingle installation phases which pose the greatest risk of knee MSD, improving the focus and development of new intervention methods.

## 4.2. Background

### 4.2.1. Prevalence of knee injuries among roofers

According to the Nonfatal Occupational Injuries and Illness report 2016, 38% of the total injuries involving lower extremities were located in the knees (BLS 2019). According to Wang et al. (2015), roofers' MSD rate is 30% higher than the average MSD incident rate reported for the entire construction trades. In the state of Washington, the insurance premium composite base rate for roofers is the highest ($7.03) among all building construction trades (Washington DOL & I 2018). It is apparent that there is a pressing need to alleviate the ergonomic injuries and develop interventions for reducing knee MSD among construction roofers.

### 4.2.2. State of practice in ergonomics to prevent knee MSDs among roofers

To protect the roofers from knee MSD, some generic solutions suggested in the existing literature include the use of powered mechanical caulk and seam welding equipment, and wearing knee pads while kneeling and installing new roofs (Spielholz et al. 2006). To reduce the risk of resulting MSD among construction workers, safety and health organizations have recommended general ergonomic practices and guidelines. For example, a booklet titled *Simple Solutions: Ergonomics for construction workers* published by the National Institute for Occupational Safety and Health (NIOSH) suggests using knee supporting devices such as kneeling creeper and knee pads that may prevent additional stress on knees in construction work that requires kneeling (Albers and Estill 2007). Similar protective devices such as knee pads and power stretchers for kneeling are suggested in e-tools promoted by the Occupational Safety and Health Administration (OSHA) (OSHA 2019). However, these protective measures are generic and even if they are able to prevent knee injuries, they are designed for work on flat surfaces. Guidelines are still lacking for tasks performed on slanted rooftops for knee MSD prevention for roofers.

### 4.2.3. State of ergonomics research on MSD among roofers

Previous ergonomic studies mainly focused on severity, prevalence, and causes of ergonomic injuries among roofers. In a survey study, motion/position and overexertion were revealed as the two most prevalent sources of nonfatal injuries to roofers (Fredericks et al. 2005). A laboratory assessment done by Choi (2008) found that roofers experience greater pain in their lower extremities during shingle installation on sloped roof surfaces compared to flat surfaces. Roofers' MSD has been identified as a possible cause of reduced physical functioning and disability in workplaces (Welch et al. 2010). In a pilot study conducted by Lee et al. (2017), feasibility of wearable sensors such as activity trackers and physiological monitors were assessed in facilitating data collection about roofers' heart rate, energy expenditure, metabolic equivalents, and sleep efficiency. Wang et al. (2017) examined the work-related risk factors of roof slope, working technique, working pace, and posture to the development of low back disorders among roofers and found significant association of these factors to cause low back disorders among roofers. A similar study assessed the effects of roof slope and kneeling posture as two potential risk factors to knee MSD development in sloped kneeling shingle installation task (Breloff et al. 2019a). Lower extremity kinematics of roofers at cross-slope roof walking were investigated by Breloff et al. (2019b), leading to establishment of their associations to MSD.

### 4.3. Problem Statement and Research Objective

Based on the review, in-depth insights on work-specific risk exposures for roofers' knees during shingle installation are still lacking. As roofers frequently perform shingle installation, the required constant awkward and extreme kneeling posture causes bending of their knees, sometimes exceeding their normal range of motion for a prolonged time. It has been proven that knee awkward rotations and repetitive motions can lead to knee MSD (Hofer et al. 2011). A ranking of the

kneeling shingle installation phases based on the MSD risk caused by the knee awkward rotations and repetitive motions might suggest the phases that potentially lead to the greatest knee MSD risk. These findings could be vital to the development of targeted interventions to prevent the knee MSD of roofers. However, such a ranking is still missing. Therefore, the objective of this study is to identify a ranking-based ergonomic method for suggesting the potential risky phases in terms of awkward knee rotations and repetitive motions during shingle installation operations on slanted surfaces.

## 4.4. Methodology

Figure 4.1 provides the schematic overview of the methodology, according to which time-series knee kinematics data in terms of five knee rotation angles—flexion, abduction, adduction, internal and external rotation are used to compute 15 risk indicators for each phase of the shingle installation process. Using these risk indicators, the phases are then ranked by all five and individual knee rotations, respectively. In ranking phases by all five knee rotations, total 15 risk indicators are combined to generate a risk score for each phase, which is considered as the measure of the risk to rank the relative risk of the phases based on awkward and extreme knee rotations. In ranking phases by individual knee rotation, only 3 risk indicators associated with each individual knee rotation (e.g., flexion) are combined to generate a risk score of each phase for risk ranking. Based on the resulting rankings, lastly, the relative contribution of each knee rotation to the knee MSD risk measurement is analyzed by correlating the rankings generated from each knee rotation with the ranking generated from all knee rotations. The following presents the methodology in detail. It first defines the shingle installation phases. This is followed by selection of the risk indicators. Next, the steps involved in the workflow of this methodology are elaborated.

**4.4.1. Shingle installation phases definition**

In a typical shingle installation process, roofers typically place and install shingles by nailing with a pneumatic nail-gun. Specifically, roofers start this process by reaching for shingles first and then placing them facing forward on the roof surface. Next, they grab the nail gun from aside and become ready to start nailing the shingles. After that, they nail the shingles side by side. They finish the shingle installation process by replacing the nail gun and returning to their initial upright position. As a result, seven phases were defined in this study (Figure 4.2) and the kinematics data were subsequently segmented for analysis according to these phase definitions.
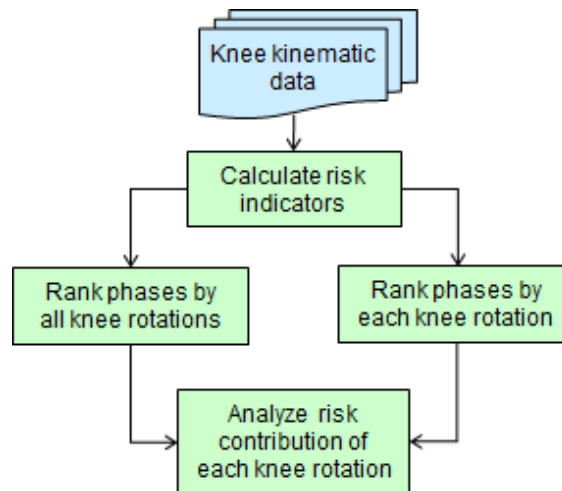


**Figure 4.1:** Methodology overview.

**4.4.2. Risk indicator selection**

During the shingle installation process, the repeated deep kneeling posture encountered by the roofers will increase the risk of knee MSD due to the extreme and awkward position of the knee joint. Repeated and high contact stress at the knee joint contributes to knee MSD such as knee osteoarthritis and damage of articular cartilage of the knee joint (Kajaks and Costigan 2015). Knee with a deep flexed position (beyond 90°) produces large forces and moments that may result in high contact stress in the knee joint surface (Nagura et al. 2002). A significant

increase (by over 80% ) in the knee joint contact stress was observed with an increase in knee flexion from 15.5° to 90° (Thambyah et al. 2005). These findings represent a strong association between knee rotations and knee joint contact stress which can relate to knee MSD. As the knee rotations undergo larger awkward posture during shingle installation and impose the knee joint to more awkward positions, thereby increasing knee joint contact force, knee MSD risk is considered to increase. Awkward knee posture is considered as a deep flexed position of the knee (>90°) accompanied by medial and lateral rotations that cause increased amount of joint contact stress. Hence, this study defines risk as an increase in awkward knee rotations that may increase the knee joint contact stress.
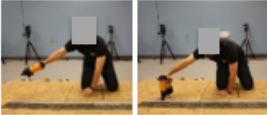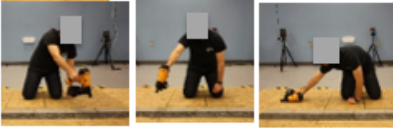


| Phase | Task performed | Illustration |
|-------|---------------|--------------|
| Phase 1 | Reaching for shingles | |
| Phase 2 | Placing the shingles | |
| Phase 3 | Grabbing the nail gun | |
| Phase 4 | Moving to first nailing position | |
| Phase 5 | Nailing shingles | |
| Phase 6 | Replacing the nail gun | |
| Phase 7 | Returning to upright position | |

**Figure 4.2:** Defined seven phases of a shingle installation trial.

To assess the impact of awkward kneeling postures and repetitive motions at different kneeling shingling installation phases, five knee rotations – flexion (about the medio-lateral axis), abduction and adduction (about the anterior-posterior axis), internal and external rotation (about the longitudinal axis) were measured (Figure 4.3). Deep flexion activities produce large force and moment on the knee joint (Nagura et al. 2006). High knee adduction and abduction impose stress on the knee joint and increase the risk of developing knee osteoarthritis (Barrios et al. 2010). Internal and external rotation of the tibia relative to the femur place additional stress on the knee joints ligaments (Coplan 1989; Hofer et al. 2011). As the shingle installation is a dynamic process, three metrics including maximum, cumulative, and average measurements of the above five knee rotations were extracted from the time series kinematic data as risk indicators. These metrics have been found to have significant associations with the MSD risk to upper and lower limbs (Gyemi et al. 2016; Hatfield et al. 2015; McClellan et al. 2009; Zampporri and Aguinaldo 2017). In a particular trial, the maximum knee rotation value for a phase refers to the peak knee rotation angle within that phase. As the extreme and awkward postures are associated with MSD, it is advantageous to evaluate the maximum knee rotation which could relate to the extreme rotation of knees due to forceful exertion in the shingle installation operation. Knee MSD risks can increase when roofers work in an awkward posture for a long period of time without adequate recovery time. The cumulative knee rotation value is the summation of the knee rotation angles across that phase. It represents the risk due to maintaining an awkward knee rotation angle over time. The average knee rotation value for a phase represents the mean knee rotation within that phase. It accounts for the risk due to knee repetitive motions. If the roofers perform the same task which involves rotations in their knees repeating every few seconds even under their tolerance limit, it can still develop knee MSD. Therefore, these metrics (maximum, cumulative, and average

measurements) of the five knee rotations (flexion, abduction, adduction, internal and external rotation) were selected, resulting in a number of 15 risk indicators at each phase.
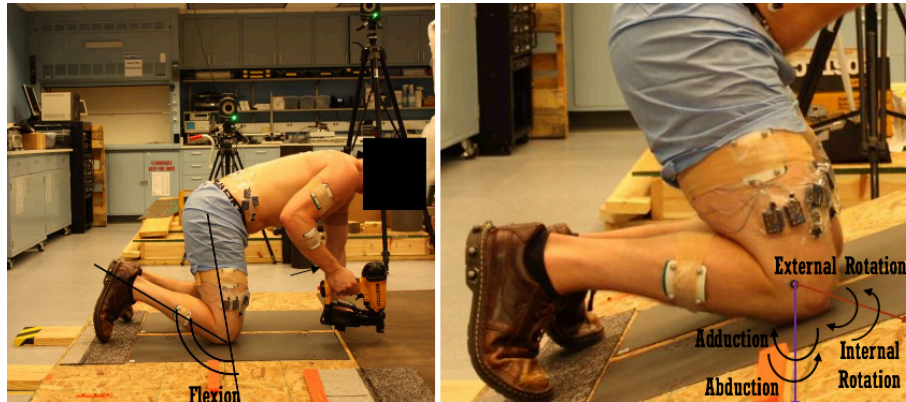


**Figure 4.3:** Illustration of five knee rotation angles.

### 4.4.3. Workflow details

### 4.4.3.1. Calculating risk indicators

The calculation of the risk indicators for each phase follows the below procedures. Since the calculation procedure for each knee rotation is the same, the flexion angle calculation is shown as an example.

Let the flexion angle of the certain knee (i.e., left or right) observed for subject $x$ at trial $t$, phase $p$, and time $i$ be represented as $F_{x,t,p,i}$.

Then, the maximum flexion for subject $x$ at trial $t$, and phase $p$ will be:

$$F_{x,t,p}^{max} = \max_{i}(F_{x,t,p,i}) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(4.1)}$$

The cumulative flexion for subject $x$ at trial $t$, and phase $p$ will be:

$$F_{x,t,p}^{cum} = \sum_i F_{x,t,p,i} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(4.2)}$$

The average flexion for subject $x$ at trial $t$, and phase $p$ will be:

$$F_{x,t,p}^{avg} = \underset{i}{\text{avg}}(F_{x,t,p,i}) \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(4.3)}$$

These values are then averaged over all the subjects and the trials to compute the risk indicators. So, using equations (4.1), (4.2), and (4.3), three risk indicators associated with the flexion at phase $p$ will be derived as:

$$\{F_p^{max}, F_p^{cum}, F_p^{avg}\} = \left\{\underset{x}{\text{avg}}[\underset{t}{\text{avg}}(F_{x,t,p}^{max})], \underset{x}{\text{avg}}[\underset{t}{\text{avg}}(F_{x,t,p}^{cum})], \underset{x}{\text{avg}}[\underset{t}{\text{avg}}(F_{x,t,p}^{avg})]\right\} \ldots\ldots\ldots\ldots(4.4)$$

Similarly, the above equations were used to compute the risk indictors at phase $p$ for the abduction, adduction, internal and external knee rotations. Given $p \in [1, 7]$, a total number of 105 risk indictors were obtained for all seven phases. Both the left and right knees' risk indictors were computed.

### 4.4.3.2. Ranking phases by all knee rotations

The seven shingle installation phases were first ranked considering the set of all knee rotations. To this end, the 15 risk indicators – maximum, cumulative, and average angles of all five knee rotations are combined by applying a scoring model to generate a risk score for each phase. The resulting risk scores were then compared for ranking the phases. The phase that has the highest score is considered potentially riskiest for the knee MSD. The resulting ranking is referred to as *multi-angle based ranking* in this study as it accounts for the contributions of all five knee rotations to the measurement of the risk scores, and the ranking represents the knee MSD risk associated with multiple awkward knee rotations. Due to the dynamic nature of the shingling process—the *multi-angle based rankings* are computed for the left and right knees separately.

### 4.4.3.3. Ranking phases by individual knee rotation

The seven phases were then ranked based on each individual knee rotations to help understand the impact of each individual knee rotation on the imposed risks. For this, instead of using the 15 risk indicators, only the 3 risk indicators (i.e., maximum, cumulative and average angles) of certain knee rotations are considered. For example, to assess the risk due to flexion, only

the 3 risk indicators of the flexion rotation were used to generate the flexion-based phase ranks by applying the scoring model. These rankings computed based on individual knee rotations are referred to as *single-angle based rankings* as they account for the contribution of individual knee rotations to the measurement of the risk scores and represent the knee MSD risk associated with that individual knee rotation of interest. A total of five single-angle based rankings were generated for the seven phases. Similar to the *multi-angle based ranking*, the *single-angle based rankings* are analyzed for each knee separately.

**4.4.3.4. Assessing the relative contribution of each knee rotation to potential knee MSD risk measurement**

To analyze which angle significantly contributed to the knee MSD risk measurement, associations between the multi-angle based ranking and the five single-angle based rankings of the phases were further analyzed by computing Spearman's rank correlation coefficient using the equation: $1 - \frac{6\sum d^2}{N^3 - N}$, where $d$ is the difference in rankings and $N$ is the number of variables (phases = 7) (Rosso 1997). The Spearman's rank correlation coefficient indicates the strength of association between two rankings. The higher the value (approaching to 1), the stronger the association between the rankings. The rotation angle with the highest coefficient value is deemed to have the strongest influence on the multi-angle based ranking that relates to the potential knee MSD risk measurement associated with the multiple knee rotations.

**4.5. Implementation and Results**

**4.5.1. Participants**

This study included the same nine participants as detailed in Chapter 2. The research protocol was approved by both the Institutional Review Boards (IRB) of the National Institute for Occupational Safety and Health and West Virginia University.

### 4.5.2. Instruments

A VICON optical motion capture system equipped with 14 MX Vicon cameras (Oxford, UK) was used to collect the segment endpoint data of the participants. Forty-two (42) retroreflective markers for motion capture were placed bilaterally on the lower extremities of the participants including feet, heels, toes, ankles, shanks, knee joints, thighs, and hip joints, as illustrated in Figure 4.4. The collected three-dimensional (3D) coordinates of these markers by VICON were used to calculate the knee angles. The kinematic data were recorded at 100 Hz.
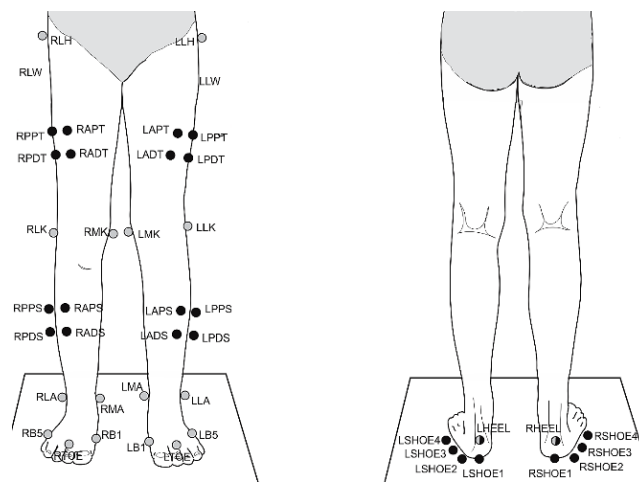


**Figure 4.4:** Marker positions.

To mimic the roof surface for shingle installation, the same wooden platform displayed in Figure 2.2 in Chapter 2 were used.

### 4.5.3. Procedure

The experiment was conducted in the biomechanics laboratory at NIOSH. Upon arrival to the lab, the participants were placed with motion markers for kinematic calibration and data collection. They were in a deep kneeling posture on the roof simulator prior to the start of data collection. When instructed to begin, the participants first reached for and placed two shingles in front of them. Then they picked up the nail gun from their right side and mimicked affixing six nails (three in each) into the two shingles side by side on the roof simulator, starting on the left

and moving to the right of the shingle. Once finishing, the participants replaced the nail gun and returned to their resting/starting position. To sum, the participants performed the seven phases of the shingle installation task as one continuous activity. Each participant performed the simulated shingle installation task on the roof simulator at three slopes — 0°, 15°, and 30°. At each slope, the task was performed five times by each one of the participants. This resulted in a total of 45 trials data (5 trials × 9 participants) at each of the slope angles.

**4.5.4. Data processing**

Using the coordinates of the markers captured by VICON, the five knee rotations were computed in Visual 3D (C-Motion, Inc., Germantown, MD) using the method provided by Robertson et al. (2013). These knee rotations data were collected at a rate of 100 data points per second during each trial resulting in a huge set of data points. From these data points, the maximum, cumulative and average of the five knee rotations for each phase were computed for a certain knee (left/right) on each slope. This way, for a certain knee at each slope, 315 (9 participants × 5 trials × 7 phases) data points of maximum, cumulative, and average of the five knee rotations were obtained, respectively. These knee rotations were then used to formulate the risk indicators necessary for the subsequent analysis.

**4.5.5. Risk indicators**

Tables 4.1 and 4.2 present the 15 risk indicators— the maximum (Max), cumulative (Cum.), and average (Avg.) of the five rotations averaged over all subjects and trials, across seven phases on three slopes for the left and right knees, respectively.

**Table 4.1:** Risk indicators computed for the left knee

| Indicator | Slope / Phase | 0° Mean | 15° Mean | 30° Mean |
|---|---|---|---|---|
| Max Flex. (°) | 1 | 137.0 | 133.6 | 137.6 |
| | 2 | 135.1 | 130.8 | 128.4 |
| | 3 | 133.7 | 128.7 | 120.3 |
| | 4 | 135.2 | 129.6 | 122.1 |
| | 5 | 131.9 | 123.5 | 117.5 |
| | 6 | 132.0 | 127.5 | 126.9 |
| | 7 | 137.4 | 133.1 | 137.0 |
| Max Abd. (°) | 1 | 9.6 | 8.5 | 9.5 |
| | 2 | 7.2 | 7.9 | 10.0 |
| | 3 | 7.6 | 8.7 | 7.8 |
| | 4 | 7.7 | 8.5 | 11.9 |
| | 5 | 8.1 | 8.6 | 12.3 |
| | 6 | 8.3 | 8.3 | 8.0 |
| | 7 | 10.8 | 8.3 | 9.5 |
| Max Add. (°) | 1 | -5.2 | -3.1 | -4.8 |
| | 2 | -6.1 | -3.3 | -4.7 |
| | 3 | -5.1 | -2.8 | -4.0 |
| | 4 | -4.4 | -3.0 | -4.2 |
| | 5 | -8.2 | -5.5 | -4.3 |
| | 6 | -8.5 | -5.4 | -4.4 |
| | 7 | -5.4 | -3.7 | -4.8 |
| Max Int. Rot. (°) | 1 | 5.1 | 10.4 | 10.4 |
| | 2 | 5.2 | 10.4 | 9.5 |
| | 3 | 4.7 | 8.4 | 8.1 |
| | 4 | 4.4 | 8.6 | 8.5 |
| | 5 | 5.7 | 7.4 | 8.3 |
| | 6 | 5.8 | 7.4 | 8.3 |
| | 7 | 5.4 | 10.2 | 10.0 |
| Max Ext. Rot. (°) | 1 | -8.8 | -8.5 | -7.3 |
| | 2 | -8.9 | -9.2 | -9.5 |
| | 3 | -7.8 | -9.9 | -8.1 |
| | 4 | -8.6 | -9.7 | -8.8 |
| | 5 | -10 | -8.8 | -8.7 |
| | 6 | -7.7 | -7.9 | -7.8 |
| | 7 | -9.4 | -8.4 | -7.9 |

| Indicator | Slope / Phase | 0° Mean | 15° Mean | 30° Mean |
|---|---|---|---|---|
| Cum. Flex. (°) | 1 | 22688 | 18979 | 20714 |
| | 2 | 58899 | 52254 | 61009 |
| | 3 | 9249 | 9389 | 8896 |
| | 4 | 16074 | 15015 | 15412 |
| | 5 | 37125 | 32926 | 31460 |
| | 6 | 13553 | 13686 | 12243 |
| | 7 | 20461 | 20789 | 19589 |
| Cum. Abd. (°) | 1 | 1348 | 1113 | 1179 |
| | 2 | 2453 | 2687 | 3342 |
| | 3 | 509 | 667 | 500 |
| | 4 | 701 | 832 | 1133 |
| | 5 | 1731 | 1900 | 2410 |
| | 6 | 752 | 774 | 689 |
| | 7 | 1267 | 1136 | 1092 |
| Cum. Add. (°) | 1 | -623 | -275 | -509 |
| | 2 | -1618 | -720 | -1202 |
| | 3 | -263 | -170 | -265 |
| | 4 | -337 | -168 | -409 |
| | 5 | -1223 | -712 | -699 |
| | 6 | -664 | -424 | -297 |
| | 7 | -511 | -365 | -472 |
| Cum. Int. Rot. (°) | 1 | 560 | 1228 | 1350 |
| | 2 | 1155 | 3110 | 3265 |
| | 3 | 227 | 546 | 552 |
| | 4 | 311 | 742 | 827 |
| | 5 | 967 | 1240 | 1726 |
| | 6 | 438 | 555 | 537 |
| | 7 | 412 | 1165 | 1170 |
| Cum. Ext. Rot. (°) | 1 | -985 | -865 | -579 |
| | 2 | -2022 | -2605 | -3034 |
| | 3 | -436 | -648 | -503 |
| | 4 | -639 | -778 | -832 |
| | 5 | -1616 | -1615 | -1743 |
| | 6 | -564 | -765 | -682 |
| | 7 | -885 | -935 | -716 |

| Indicator | Slope / Phase | 0° Mean | 15° Mean | 30° Mean |
|---|---|---|---|---|
| Avg. Flex. (°) | 1 | 127.7 | 130.5 | 133.9 |
| | 2 | 128.8 | 126.9 | 118.3 |
| | 3 | 131.7 | 126.5 | 116.8 |
| | 4 | 131.9 | 124.9 | 118.6 |
| | 5 | 126.1 | 117.3 | 112.9 |
| | 6 | 125.9 | 120.1 | 117.8 |
| | 7 | 130.5 | 130.1 | 132.0 |
| Avg. Abd. (°) | 1 | 7.6 | 7.5 | 7.9 |
| | 2 | 5.6 | 6.1 | 6.5 |
| | 3 | 6.8 | 8.2 | 6.9 |
| | 4 | 5.9 | 6.3 | 8.4 |
| | 5 | 6.0 | 6.2 | 8.6 |
| | 6 | 7.7 | 7.4 | 6.8 |
| | 7 | 7.9 | 6.5 | 7.6 |
| Avg. Add. (°) | 1 | -3.2 | -1.9 | -2.9 |
| | 2 | -3.9 | -1.7 | -1.9 |
| | 3 | -4.4 | -2.3 | -3.0 |
| | 4 | -2.9 | -1.4 | -3.1 |
| | 5 | -4.3 | -2.4 | -2.3 |
| | 6 | -5.6 | -3.5 | -2.8 |
| | 7 | -3.2 | -2.4 | -3.4 |
| Avg. Int. Rot. (°) | 1 | 3.2 | 8.8 | 8.2 |
| | 2 | 2.5 | 7.6 | 6.0 |
| | 3 | 3.4 | 7.6 | 7.2 |
| | 4 | 2.5 | 5.9 | 6.1 |
| | 5 | 3.3 | 4.1 | 6.4 |
| | 6 | 3.7 | 4.6 | 5.7 |
| | 7 | 2.8 | 7.9 | 7.9 |
| Avg. Ext. Rot. (°) | 1 | -5.6 | -6.5 | -4.6 |
| | 2 | -5.3 | -6.9 | -6.2 |
| | 3 | -6.3 | -8.9 | -6.9 |
| | 4 | -5.8 | -7.2 | -7.0 |
| | 5 | -6.0 | -6.5 | -6.5 |
| | 6 | -5.9 | -6.2 | -6.5 |
| | 7 | -5.7 | -6.0 | -5.3 |

**Table 4.2:** Risk indicators computed for the right knee

| Indicator | Slope Phase | 0° Mean | 15° Mean | 30° Mean |
|---|---|---|---|---|
| Max Flex. (°) | 1 | 138.0 | 138.0 | 139.3 |
| | 2 | 135.0 | 135.0 | 130.0 |
| | 3 | 133.0 | 132.0 | 121.0 |
| | 4 | 134.0 | 133.0 | 122.0 |
| | 5 | 131.0 | 126.0 | 117.0 |
| | 6 | 132.0 | 131.0 | 128.0 |
| | 7 | 137.0 | 137.5 | 139.0 |
| Max Abd. (°) | 1 | 9.2 | 9.2 | 8.0 |
| | 2 | 7.1 | 6.3 | 6.0 |
| | 3 | 5.5 | 5.3 | 5.5 |
| | 4 | 5.8 | 4.9 | 5.9 |
| | 5 | 6.0 | 6.4 | 6.7 |
| | 6 | 5.7 | 6.4 | 6.8 |
| | 7 | 8.8 | 8.0 | 7.9 |
| Max Add. (°) | 1 | -5.7 | -62 | -7.0 |
| | 2 | -8.0 | -6.8 | -7.5 |
| | 3 | -6.8 | -6.6 | -4.0 |
| | 4 | -8.9 | -6.5 | -6.8 |
| | 5 | -8.8 | -6.9 | -7.0 |
| | 6 | -5.9 | -6.5 | -5.8 |
| | 7 | -5.5 | -6.9 | -6.9 |
| Max Int. Rot. (°) | 1 | 10.0 | 13.1 | 14.4 |
| | 2 | 11.0 | 11.6 | 11.8 |
| | 3 | 9.9 | 10.0 | 12.6 |
| | 4 | 10.7 | 11.7 | 12.2 |
| | 5 | 9.4 | 11.0 | 12.4 |
| | 6 | 9.6 | 10.0 | 12.0 |
| | 7 | 10.4 | 13.0 | 14.4 |
| Max Ext. Rot. (°) | 1 | -9.9 | -10.1 | -8.4 |
| | 2 | -6.0 | -6.4 | -6.9 |
| | 3 | -4.7 | -5.4 | -6.7 |
| | 4 | -4.9 | -6.0 | -6.4 |
| | 5 | -7.9 | -7.3 | -9.7 |
| | 6 | -7.3 | -7.2 | -11.0 |
| | 7 | -9.1 | -9.9 | -7.1 |

| Indicator | Slope Phase | 0° Mean | 15° Mean | 30° Mean |
|---|---|---|---|---|
| Cum. Flex. (°) | 1 | 23509 | 19630 | 20111 |
| | 2 | 62190 | 53776 | 61015 |
| | 3 | 9269 | 9689 | 8866 |
| | 4 | 16149 | 15299 | 15241 |
| | 5 | 37608 | 33956 | 31075 |
| | 6 | 13841 | 14126 | 12210 |
| | 7 | 20284 | 21434 | 18355 |
| Cum. Abd. (°) | 1 | 1245 | 1185.1 | 924 |
| | 2 | 1782 | 1946 | 1822 |
| | 3 | 295 | 374 | 341 |
| | 4 | 446 | 407 | 521 |
| | 5 | 834 | 1091 | 1103 |
| | 6 | 415 | 485 | 528 |
| | 7 | 925 | 1071 | 896 |
| Cum. Add. (°) | 1 | -598 | -724 | -729 |
| | 2 | -2935 | -1999 | -2227 |
| | 3 | -405 | -429 | -270 |
| | 4 | -858 | -583 | -504 |
| | 5 | -1787 | -1215 | -1089 |
| | 6 | -446 | -572 | -351 |
| | 7 | -481 | -800 | -641 |
| Cum. Int. Rot. (°) | 1 | 1299 | 1517 | 1771 |
| | 2 | 4120 | 3947 | 4870 |
| | 3 | 585 | 660 | 849 |
| | 4 | 1071 | 1138 | 1434 |
| | 5 | 1979 | 2400 | 2843 |
| | 6 | 834 | 728 | 897 |
| | 7 | 1053 | 1560 | 1453 |
| Cum. Ext. Rot. (°) | 1 | -1506 | -1285 | -1000 |
| | 2 | -1526 | -1740 | -1792 |
| | 3 | -258 | -352 | -419 |
| | 4 | -428 | -414 | -562 |
| | 5 | -1273 | -1165 | -1806 |
| | 6 | -567 | -500 | -840 |
| | 7 | -1033 | -1260 | -719 |

| Indicator | Slope Phase | 0° Mean | 15° Mean | 30° Mean |
|---|---|---|---|---|
| Avg. Flex. (°) | 1 | 131.8 | 134.3 | 130.7 |
| | 2 | 130.2 | 130.3 | 118.1 |
| | 3 | 131.8 | 130.5 | 116.4 |
| | 4 | 129.5 | 126.9 | 117.1 |
| | 5 | 125.1 | 120.3 | 110.9 |
| | 6 | 128.3 | 124.0 | 116.4 |
| | 7 | 129.3 | 133.9 | 124.9 |
| Avg. Abd. (°) | 1 | 6.6 | 7.6 | 5.5 |
| | 2 | 4.3 | 4.4 | 3.5 |
| | 3 | 4.6 | 4.7 | 4.5 |
| | 4 | 4.1 | 3.2 | 3.9 |
| | 5 | 3.2 | 3.6 | 4.1 |
| | 6 | 4.3 | 4.6 | 5.1 |
| | 7 | 5.8 | 5.9 | 5.8 |
| Avg. Add. (°) | 1 | -3.6 | -5.2 | -4.9 |
| | 2 | -6.0 | -5.1 | -3.7 |
| | 3 | -5.8 | -6.0 | -3.1 |
| | 4 | -6.7 | -5.1 | -3.9 |
| | 5 | -5.8 | -4.6 | -3.8 |
| | 6 | -4.2 | -5.0 | -3.3 |
| | 7 | -3.2 | -5.5 | -4.7 |
| Avg. Int. Rot. (°) | 1 | 7.5 | 10.7 | 11.5 |
| | 2 | 8.1 | 9.1 | 8.1 |
| | 3 | 8.2 | 8.9 | 10.6 |
| | 4 | 8.0 | 8.8 | 10.2 |
| | 5 | 6.2 | 7.6 | 9.8 |
| | 6 | 7.2 | 6.7 | 8.7 |
| | 7 | 7.1 | 10.6 | 10.7 |
| Avg. Ext. Rot. (°) | 1 | -8.0 | -8.9 | -6.1 |
| | 2 | -3.5 | -4.4 | -4.1 |
| | 3 | -3.8 | -4.3 | -5.7 |
| | 4 | -3.6 | -3.8 | -4.8 |
| | 5 | -4.6 | -4.8 | -7.1 |
| | 6 | -5.7 | -4.9 | -8.7 |
| | 7 | -6.2 | -6.7 | -4.8 |

73

### 4.5.6. Multi-angle based ranking

This study implemented two types of scoring models – an *aggregation-based scoring model* and a *multiplication scoring model* for the multi-angle based ranking, both of which have been proven to be useful in ranking tasks with multiple criteria (El-Sayegh and Mansour 2015; Tofallis 2012; Tofallis 2014).

Typically, the *aggregation-based scoring model* includes three steps: 1) data normalization, 2) indicator weighing, and 3) weighted indictor aggregation to generate an overall risk score for ranking. For the data normalization, the following methods were tested for each risk indicator in this study:

1) <u>Dividing by maximum:</u> It converted the value of a risk indicator (e.g., Max Flex.) for a certain phase as proportion of the maximum (e.g., maximum of Max Flex.) among the seven phases as:

$$Y = \frac{X}{X_{max}}$$ ……………………………………………….(4.5)

2) <u>Dividing by sum:</u> It converted the value of a risk indictor (e.g., Max Flex.) for a certain phase as proportion of the sum (e.g., sum of Max Flex.) across the seven phases as:

$$Y = \frac{X}{\sum X}$$ ………………………………………………… (4.6)

3) <u>Range normalization:</u> It scaled the value of a risk indicator for a certain phase to [0, 1] in regard to the maximum and minimum values of the risk indicator among the seven phases as:

$$Y = \frac{X - X_{min}}{X_{max} - X_{min}}$$ ………………………..……….………………… (4.7)

In the above equations, $X$ represents the value of a risk indicator for a certain phase; $\sum X$ is the sum of the values of that risk indictor across all the seven phases; $X_{max}$ and $X_{min}$ are the

highest and lowest values of that risk indicator among the seven phases, respectively; and $Y$ is the normalized value of $X$.

After normalization, weights were assigned to each risk indicator. As no literature that exists reveals the relative contribution of each knee rotation to the knee MSD risk measurement and all the five knee rotations observed in this study are biomechanically responsible to cause knee MSD, this study assumed that all risk indicators have equal contribution to the knee MSD risk and hence have been assigned equal weights. Finally, all the weighted indicators were summed to yield an overall risk score for each phase by:

$$\text{Risk Score (Aggregation-based)} = w_1 Y_1 + w_2 Y_2 + w_3 Y_3 + \ldots + w_n Y_n \quad \ldots\ldots\ldots\ldots(4.8)$$

where $Y_i$ $(i = 1, 2, \ldots n)$ is the normalized value of a risk indicator for a phase and $w_i$ $(i = 1, 2, \ldots n)$ is the weight assigned to each indicator.

Alternatively, the *multiplication scoring model* does not require normalization. The risk score of a phase for a set of risk indicators was calculated by:

$$\text{Risk Score (Multiplication)} = X_1^{W_1} X_2^{W_2} X_3^{W_3} \ldots X_n^{W_n} \quad \ldots\ldots\ldots\ldots\ldots (4.9)$$

where $X_i$ $(i = 1, 2, \ldots n)$ is the value of a risk indicator for a phase and $w_i$ $(i = 1, 2, \ldots n)$ is the weight assigned to each indicator. In this study, $w_i$ $(i = 1, 2, \ldots n)$ was set to 1 for both of the scoring models and $n = 15$, the total number of the risk indicators.

In both aggregation-based and multiplication scoring models, the phase with the highest score was ranked first (1). Applying these scoring models, this study generated the risk scores for multi-angle based ranking at different roof slopes, which was provided in Table 4.3.

Based on Table 4.3, the multi-angle based ranking of the phases were computed at each slope as shown in Figure 4.5.

**Table 4.3:** Risk scores of seven phases

| Phase | Left Knee | | | | Right Knee | | | | Slope |
|---|---|---|---|---|---|---|---|---|---|
| | Multiplication (×10²⁶) | Aggregation | | | Multiplication (×10²⁶) | Aggregation | | | |
| | | Divide by max | Range normalization | Divide by sum | | Divide by max | Range normalization | Divide by sum | |
| 1 | 1.70 | 10.80 | 6.60 | 2.10 | 47.00 | 11.60 | 9.50 | 2.40 | |
| 2 | 56.00 | 13.00 | 8.00 | 3.00 | 1070.00 | 13.00 | 11.00 | 3.00 | |
| 3 | 0.02 | 9.40 | 4.70 | 1.70 | 0.04 | 8.50 | 4.00 | 1.60 | |
| 4 | 0.04 | 9.00 | 3.00 | 1.60 | 1.00 | 9.50 | 5.80 | 1.80 | 0° |
| 5 | 41.00 | 12.70 | 9.00 | 2.60 | 49.00 | 11.00 | 5.40 | 2.30 | |
| 6 | 0.80 | 11.00 | 6.65 | 2.06 | 0.30 | 8.80 | 3.50 | 1.70 | |
| 7 | 1.10 | 10.70 | 7.50 | 2.04 | 6.00 | 10.00 | 7.00 | 2.00 | |
| 1 | 2.10 | 10.40 | 7.00 | 2.05 | 170.00 | 12.00 | 10.30 | 2.50 | |
| 2 | 192.00 | 13.00 | 9.50 | 3.00 | 740.00 | 13.00 | 9.80 | 2.90 | |
| 3 | 0.20 | 9.80 | 6.70 | 1.80 | 0.10 | 8.65 | 4.00 | 1.60 | |
| 4 | 0.10 | 9.00 | 4.50 | 1.70 | 0.30 | 8.70 | 3.50 | 1.65 | 15° |
| 5 | 15.00 | 11.40 | 6.00 | 2.40 | 41.00 | 10.65 | 5.00 | 2.25 | |
| 6 | 0.60 | 10.00 | 4.60 | 2.00 | 0.50 | 9.00 | 3.00 | 1.70 | |
| 7 | 3.10 | 10.50 | 6.60 | 2.10 | 108.00 | 11.80 | 10.00 | 2.40 | |
| 1 | 5.60 | 10.70 | 7.80 | 2.10 | 57.00 | 11.50 | 10.00 | 2.30 | |
| 2 | 722.00 | 13.50 | 10.00 | 3.00 | 530.00 | 12.60 | 7.60 | 3.00 | |
| 3 | 0.10 | 9.00 | 3.00 | 1.60 | 0.06 | 8.00 | 2.00 | 1.50 | |
| 4 | 3.00 | 10.40 | 6.00 | 2.00 | 1.00 | 9.00 | 3.70 | 1.80 | 30° |
| 5 | 67.00 | 11.60 | 7.00 | 2.50 | 157.00 | 11.60 | 7.00 | 2.50 | |
| 6 | 0.20 | 9.30 | 3.50 | 1.70 | 2.00 | 10.00 | 5.40 | 1.90 | |
| 7 | 6.00 | 10.80 | 8.00 | 2.05 | 15.00 | 10.80 | 8.60 | 2.00 | |
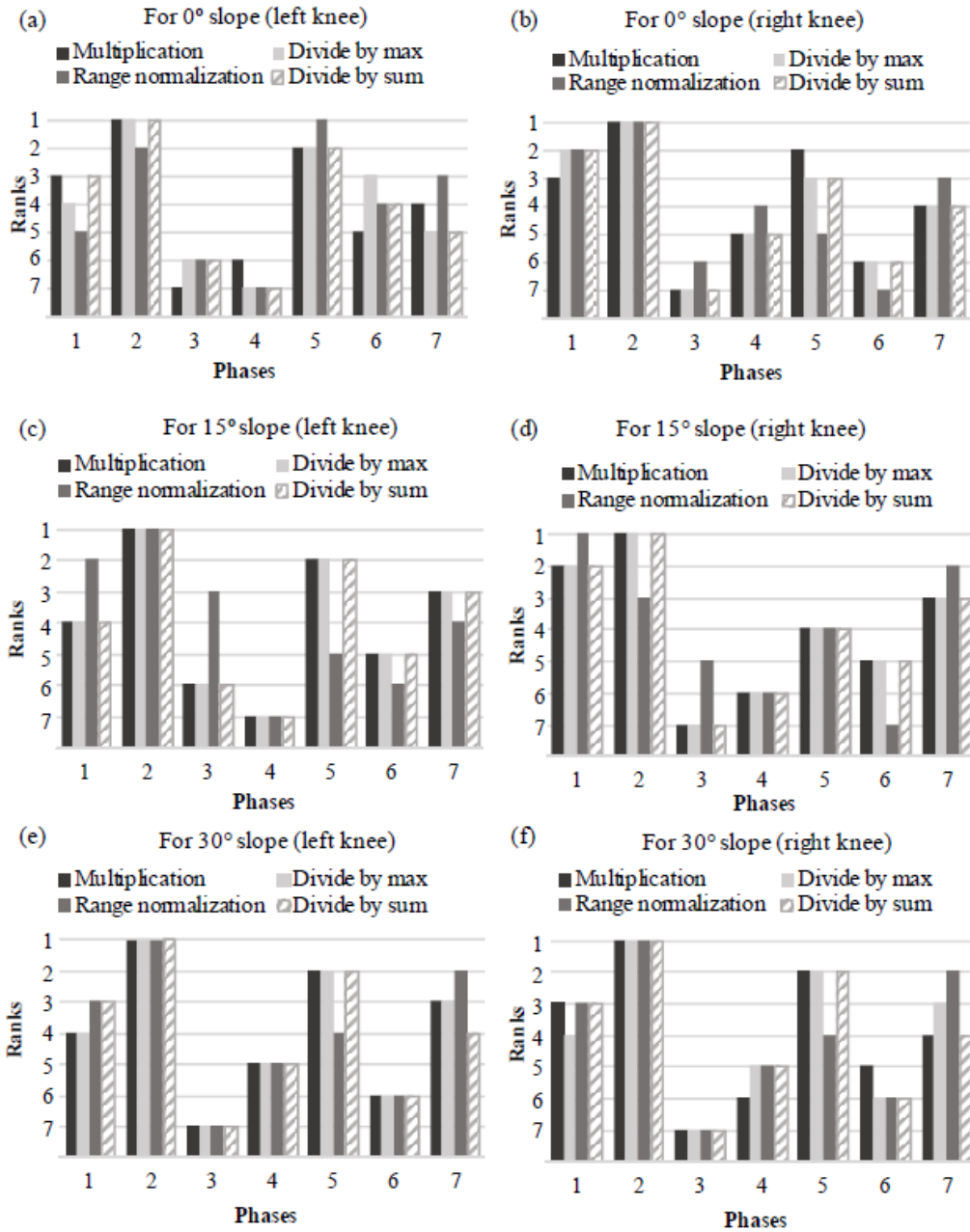
**Figure 4.5:** Multi-angle based ranks of the seven phases (a) 0° slope (left knee); (b) 0° slope (right knee); (c) 15° slope (left knee); (d) 15° slope (right knee); (e) 30° slope (left knee); (f) 30° slope (right knee).

### 4.5.6.1. Consistency analysis for scoring model selection

As this study was to identify the riskier phases that might involve comparatively higher awkward rotations of knees and repetitions in any roof setting and therefore could potentially contribute to the knee MSD, it was important to select the scoring model that would provide the most consistent rankings across the three roof slopes. For that reason, the scoring model to be used in the subsequent risk assessment was selected based on the consistency of the multi-angle based rankings across different roof slopes. Spearman's rank correlation test was used to analyze the association of the multi-angle based rankings between every two slopes. This approach was used to identify which scoring model provides the maximum number of strongest associations and hence consistency of rankings across different roof slopes. A coefficient close to one (1) suggests a strong association.

**Table 4.4:** Spearman's correlation coefficients between 2 slopes by different ranking models

| Slope | Multiplication | | Divide by max | | Range normalization | | Divide by sum | |
|---|---|---|---|---|---|---|---|---|
| | Left | Right | Left | Right | Left | Right | Left | Right |
| 0°-15° | **0.929** | 0.857 | 0.857 | **0.929** | 0.286 | 0.786 | 0.893 | **0.929** |
| 0°-30° | **0.929** | **0.964** | 0.679 | 0.929 | 0.571 | 0.714 | 0.821 | 0.929 |
| 15°-30° | **0.893** | **0.893** | **0.893** | **0.893** | 0.536 | 0.857 | 0.857 | **0.893** |

The results of Spearman's correlation for testing the rank association between slopes are provided in Table 4.4. For the left knee, the strongest association of ranks between 0° and 15° and between 0° and 30° slopes were obtained by the multiplication scoring model (0.929). Between 15° and 30° slopes, the strongest association was obtained by both the multiplication and aggregation-based scoring models when normalization was done via dividing by max (0.893). For the right knee, the strongest rank association between 0° and 15° slopes was obtained by the aggregation-based model when normalization was done via dividing by max and dividing by sum (0.929). Between 15° and 30° slopes, the strongest associations were observed for both the

multiplication and aggregation-based scoring models while normalizing by dividing by maximum and dividing by sum (0.893). Finally, between 0° and 30° slopes, only the multiplication scoring model provided the strongest rank association (0.964). Based on these results it is noticeable that the multiplication scoring model could compute the most consistent ranks among different roof slopes and hence was used for the subsequent risk analysis. For both knees, the achieved powers of the consistency analysis between different slopes ranged from 72% to 99%, which was acceptable to demonstrate an association or causal relationship between two variables (Cohen 2013).

### 4.5.6.2. Multi-angle based risks

According to Table 4.3 and Figure 4.5, for the left knee, at all three slopes, Phase 2 scored the highest, Phase 5 scored the second highest, and thus were ranked first and second respectively. The next riskiest phases were Phase 7 (ranked fourth at 0°, third at 15° and 30° slopes) and Phase 1 (ranked third at 0°, fourth at 15° and 30° slopes).

For the right knee, Phase 2 scored the highest and was ranked first at all three slopes. The next riskiest phases were Phase 5 (ranked second at 0° and 30°, fourth at 15° slopes) and Phase 1 (ranked second at 15°, third at 0° and 30° slopes). The least risky phases for both knees were Phases 4, 6, and 3.

### 4.5.7. Single-angle based rankings by the multiplication scoring model

Table 4.5 shows the phase rankings for each individual knee angle. The resulting single-angle based rankings demonstrated a certain pattern of risk at the seven phases across three slopes, but not quite consistent. Therefore, averaging the ranks across three slopes were performed, and the results are presented in Table 4.6. From Table 4.6, it is evident that Phase 2 involved the highest

flexion, abduction, adduction, internal, and external rotations for the left knee at all three roof slopes and hence, could be deemed as the riskiest phase in terms of awkward knee rotations during slope shingle installation. As to flexion, abduction and external rotation, the next riskiest phase was Phase 5, when the participants faced extreme adduction as well. According to the ranking results, the next riskiest phases were Phases 1 and 7 because Phase 1 was ranked third in terms of flexion, abduction and adduction; and Phase 7 was ranked fourth in terms of flexion, internal and external rotation, second in terms of adduction and third in terms of abduction. Similarly, for the right knee, at Phase 2, the participants experienced the highest amount of flexion, adduction and internal rotation which was turned out to be riskiest. The next riskiest phases were Phases 1 and 5. Phase 1 generated the highest amount of abduction and external rotation and Phase 5 was ranked second in terms of flexion, adduction, internal and external rotations. After Phases 2, 5 and 1, the next riskiest phase for the right knee was Phase 7 that was ranked second in terms of abduction and third in terms of external rotation. The ranking results revealed that, in terms of awkward knee rotations, Phases 4, 6, and 3 were the least risky phases for both knees.

### 4.5.8. Spearman's correlation test result to assess relative contributions of each angle

Table 4.7 presents the level of association between the multi-angle and single-angle based rankings at each slope. For the left knee, the flexion based ranking was most significantly associated with the multi-angle based rankings at 0° and 30°slopes ($r = 0.964$, $p$ value= 0.003), whereas the abduction based ranking was most significantly associated at 15° slope ($r = 0.964$, $p$ value= 0.003). For the right knee, the flexion based ranking was most significantly associated with the multi-angle based ranking at 0° slope ($r = 1$, $p$ value= 0.000). At 15° slope, the abduction based ranking ($r = 0.929$, $p$ value= 0.007) and at 30° slope, both the flexion and adduction based rankings were most significantly associated with the multi-angle based ranking ($r = 0.964$, $p$ value= 0.003).

These results implied that, flexion, abduction and adduction most significantly contributed to the MSD risk measurement.

**Table 4.5:** Ranks of the 7 phases for each knee angle

| Angle | Phases | Slopes (Left knee) | | | Slopes (Right knee) | | |
|---|---|---|---|---|---|---|---|
| | | 0° | 15° | 30° | 0° | 15° | 30° |
| Flex. | 1 | 3 | 4 | 3 | 3 | 4 | 3 |
| | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 3 | 7 | 7 | 7 | 7 | 7 | 7 |
| | 4 | 5 | 5 | 5 | 5 | 5 | 5 |
| | 5 | 2 | 2 | 2 | 2 | 2 | 2 |
| | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| | 7 | 4 | 3 | 4 | 4 | 3 | 4 |
| Abd. | 1 | 3 | 3 | 4 | 1 | 1 | 2 |
| | 2 | 2 | 1 | 2 | 2 | 2 | 3 |
| | 3 | 7 | 6 | 7 | 7 | 6 | 7 |
| | 4 | 6 | 7 | 3 | 5 | 7 | 6 |
| | 5 | 4 | 2 | 1 | 4 | 4 | 4 |
| | 6 | 5 | 5 | 6 | 6 | 5 | 5 |
| | 7 | 1 | 4 | 5 | 3 | 3 | 1 |
| Add. | 1 | 4 | 5 | 3 | 5 | 4 | 3 |
| | 2 | 2 | 3 | 1 | 1 | 1 | 1 |
| | 3 | 6 | 6 | 7 | 4 | 7 | 7 |
| | 4 | 7 | 7 | 5 | 3 | 5 | 5 |
| | 5 | 1 | 1 | 4 | 2 | 2 | 2 |
| | 6 | 3 | 2 | 6 | 6 | 6 | 6 |
| | 7 | 5 | 4 | 2 | 7 | 3 | 4 |
| Int. Rot. | 1 | 4 | 2 | 2 | 3 | 3 | 3 |
| | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| | 3 | 6 | 6 | 6 | 7 | 6 | 6 |
| | 4 | 7 | 4 | 5 | 4 | 5 | 5 |
| | 5 | 1 | 5 | 4 | 2 | 4 | 2 |
| | 6 | 3 | 7 | 7 | 6 | 7 | 7 |
| | 7 | 5 | 3 | 3 | 5 | 2 | 4 |
| Ext. Rot. | 1 | 3 | 5 | 7 | 1 | 1 | 3 |
| | 2 | 2 | 1 | 1 | 4 | 3 | 4 |
| | 3 | 7 | 3 | 6 | 7 | 7 | 7 |
| | 4 | 5 | 4 | 3 | 6 | 6 | 6 |
| | 5 | 1 | 2 | 2 | 3 | 4 | 1 |
| | 6 | 6 | 7 | 4 | 5 | 5 | 2 |
| | 7 | 4 | 6 | 5 | 2 | 2 | 5 |

**Table 4.6:** Averaged single-angle based phase ranks over three slopes

| Rank | Phases | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Left | | | | | Right | | | | |
| | Flex. | Abd. | Add. | Int. Rot | Ext. Rot | Flex. | Abd. | Add. | Int. Rot | Ext. Rot |
| 1 | 2 | 2 | 2, 5 | 2 | 2 | 2 | 1 | 2 | 2 | 1 |
| 2 | 5 | 5 | 6, 7 | 1 | 5 | 5 | 2, 7 | 5 | 5 | 5 |
| 3 | 1 | 1, 7 | 1 | 5 | 4 | 1 | 5 | 1 | 1 | 7 |
| 4 | 7 | 4, 6 | 3, 4 | 7 | 1, 7 | 7 | 6 | 4 | 7 | 2 |
| 5 | 4 | 3 | | 4 | 3 | 4 | 4 | 7 | 4 | 6 |
| 6 | 6 | | | 6 | 6 | 6 | 3 | 6, 3 | 3 | 4 |
| 7 | 3 | | | 3 | | 3 | | | 6 | 3 |

## 4.6. Discussion

The goal of this study was to apply ranking methods to a shingle installing process to determine if the methods could provide useful risk rankings for potential knee MSD. Fifteen (15) risk indicators of knee MSD associated with awkward knee rotational angles were combined using a scoring model to generate risk scores of the phases that led to a ranking of the phases.

Among the two scoring models explored in this study for ranking, the aggregation-based scoring model showed some inconsistencies in the ranking depending upon the selection of the normalization process. The multiplication scoring model generally overcomes this kind of inconsistency as this model does not need any normalization. For datasets where indicators have different numerical scales, the multiplication scoring model presents better performance, grounded that rescaling of any particular indicator generates no impacts on the ranking outcome in this model (Tofallis 2012). Among the three metrics considered in this study, the cumulative angle values were much greater than the maximum and average angle values. The multiplication scoring model thus minimized the chances of the cumulative angles, even without any normalization, influencing the ranks. To assure the consistencies of the ranks across different roof slopes, Spearman's rank correlation test was used. The result showed that the multiplication scoring model was more

consistent in generating ranks across different roof slopes and hence was found more suitable for this study. At all three roof slopes, though the measured individual knee rotation angles varied, the ranking results presented similar risk patterns. The variation in the individual knee rotations due to slope change could slightly change the resulting risk scores of the phases, but it did not influence their overall risk pattern across different slopes. In other words, the comparative risks among the phases were not substantially affected across different slopes. Nevertheless, such observation should be further confirmed.

**Table 4.7:** Level of association between the multi-angle and single-angle based phase ranks

| Angle | Spearman correlation (r) | p value | Slope (°) | Knee |
|---|---|---|---|---|
| Flex. | **0.964** | **0.003** | | |
| Abd. | 0.75 | 0.750 | | |
| Add. | 0.821 | 0.034 | 0° | |
| Int. Rot. | 0.821 | 0.034 | | |
| Ext. Rot. | 0.929 | 0.007 | | |
| Flex. | 0.893 | 0.012 | | |
| Abd. | **0.964** | **0.003** | | |
| Add. | 0.714 | 0.088 | 15° | Left |
| Int. Rot. | 0.536 | 0.236 | | |
| Ext. Rot. | 0.429 | 0.354 | | |
| Flex. | **0.964** | **0.003** | | |
| Abd. | 0.821 | 0.034 | | |
| Add. | 0.893 | 0.012 | 30° | |
| Int. Rot. | 0.821 | 0.034 | | |
| Ext. Rot. | 0.607 | 0.167 | | |
| Flex. | **1.00** | **0.000** | | |
| Abd. | 0.821 | 0.034 | | |
| Add. | 0.536 | 0.234 | 0° | |
| Int. Rot. | 0.964 | 0.003 | | |
| Ext. Rot. | 0.643 | 0.139 | | |
| Flex. | 0.821 | 0.034 | | |
| Abd. | **0.929** | **0.007** | | |
| Add. | 0.821 | 0.034 | 15° | Right |
| Int. Rot. | 0.857 | 0.023 | | |
| Ext. Rot. | 0.893 | 0.012 | | |
| Flex. | **0.964** | **0.003** | | |
| Abd. | 0.679 | 0.109 | | |
| Add. | **0.964** | **0.003** | 30° | |
| Int. Rot. | 0.893 | 0.012 | | |
| Ext. Rot. | 0.643 | 0.138 | | |

*Note: Bold numbers indicate the most significant r and p values*

Based on awkward knee rotations, the multi-angle based ranking suggested both knees were exposed to the greatest potential knee MSD risk while placing shingles (Phase 2). The next suggested riskiest phase was nailing shingles (Phase 5). One possible reason for this result might be that these two phases of the shingle installation operation required more repetition of extreme and awkward movement of the knees for placing and installing shingles compared to the other phases. In these phases, the participants encountered larger awkward knee rotations multiple times while leaning forward to grasp the shingles and moving backward for placing and nailing them. These repetitive motions along with awkward rotations might lead to extra stress and force sustained by the knee joint ligaments; these operations turned out to be potentially risky phases for knee MSD. Another reason could be that the duration of these two phases were also relatively longer which contributed to the higher cumulative awkward knee rotations for extended periods of time. In this study, the risk scores computed using the scoring model took into account the maximum, cumulative and average of the knee rotation angles (risk indicators in this study) which represented the risk of extreme knee rotation due to forceful exertion, prolonged and repeated awkward kneeling respectively — the knee MSD risk factors. As these two phases involved more repetition of extreme and awkward knee rotations for a longer duration compared to the other phases, the risk scores were also probably higher in these two phases.

On average, the single-angle based rankings demonstrated a risk pattern similar to the multi-angle based ranking. Among all the 7 phases, the placing shingle (Phase 2) and nailing shingle (Phase 5) phases were found to be potentially the riskiest phases with extreme awkward knee rotations. The ranking results demonstrated that these two phases were ranked either first or second for a majority of the individual knee rotation angles. For example, for the left knee all five knee rotations and for the right knee, flexion, adduction and internal rotation were extreme in

Phase 2. Although during reaching for shingles (Phase 1) the subjects seemed to experience extreme abduction and external rotation in the right knee (ranked first for abduction and external rotation in right knee), overall, the awkward knee rotations imposed by all other lower ranked phases were less extreme than that imposed by placing and nailing shingles phases. One possible reason could be the higher postural sway, required during placing and nailing shingles on sloped rooftops. The subjects needed to maintain the balance on the inclined roof surface by raising the body's center of mass which might force the knees to rotate more awkwardly and repeatedly, sometimes more than the normal tolerance limit. In all other phases, the participants either leaned forward to reach for shingles and moved backward to return to the resting position or used their nailing hand to grab and replace the nail gun. These phases required very few rotation movements in the knees and hence imposed less postural awkwardness.

Spearman's correlation test result to assess the relative contributions of each knee angle demonstrated that, for the left knee, flexion and abduction, and for the right knee, flexion, abduction and adduction significantly contributed to the potential knee MSD risk measurement. However, further biomechanical studies are needed to investigate the impacts of these knee rotations on knee MSD. To the best of the authors' knowledge, so far, no publications available have measured the knee MSD risk by exploiting the relative contribution of the knee rotation angles. This finding can be used as a guide in identifying which knee rotations might be prioritized for developing knee MSD risk assessment tools and effective interventions to prevent knee injuries among roofers.

Based on the above discussion, it is evident that special considerations such as knee interventions and protective measures are essential so that knee flexion, abduction and adduction can be minimized during awkward kneeling while placing and nailing shingles at slanted roof

surfaces. Possible interventions include wearing knee protecting devices such as knee pads or knee savers while kneeling which can minimize the impact on knees. However, these interventions were not tested in this study. As the goal was to suggest the risky phases in terms of awkward knee rotation, this study did not use any interventions (wearable or external) as that could have altered the knee rotations. Also, the participants were not professional roofers and there might be some postural variations between professional and non-professional roofers based on their working techniques, which might impact the knee rotations. Hence, further biomechanical assessment is needed to understand those variations in the knee kinematics between professional and non-professional roofers, and the impact of knee protecting devices are needed to be examined.

This study used nine participants' knee kinematics data to compute the risk indicators. Typically, the relationship between kinematics (i.e., knee rotational angles) and musculoskeletal loadings, which is associated to MSD, is influenced by biomechanics, not random. This relationship warrants the sample size of nine subjects appropriate for this study and a large sample size is not necessary. In the research community, it is well accepted to use less than ten subjects for risk analysis with biomechanical models.

## 4.7. Study Limitations

First, the experiment was completed in a controlled laboratory setting rather than a construction site to minimize the possible risk. Although the experiment was designed based on the common site practices and the participants simulated the shingle installation task to capture the real scenario of the construction site, further assessment in a real work setting is still needed to justify the findings. Second, none of the study participants were experienced roofers. But they were physically active and had working experience with home remodeling projects. From the biomechanical perspective, there should be some difference between novice and experienced

roofers, but in this study novice roofers were chosen to identify the risk that an individual without any prior roofing experience may face when they first get exposed to slanted roof top during shingle installation. This study presumed considerably similar biomechanical reactions of the participants and professional roofers. However, if professional roofers were used, it could have altered the individual knee rotations to some extent. Therefore, further assessment is necessary to justify such premise. The goal of this study was to identify an ergonomic method for suggesting the risky phases in the shingle installation process, and the subject tests only served to demonstrate the procedure and the framework. Third, this study only focused on risks to knee MSD, excluding associated risks to other lower extremities such as ankles that may undergo excessive pressure during shingle installation. Finally, this study identified risky phases only based on knee rotations. Other indicators (e.g., muscle activations and joint loadings) for knee MSD risks would be complementary but were not assessed in this study.

## 4.8. Conclusion and Future Extension

This study identified an ergonomic method for ranking the relative risks of the shingle installation phases to knee MSD development. The results suggested that the phases including placing and nailing shingles are the riskiest phases in terms of awkward rotation and repetition; and the awkward flexion, abduction and adduction can potentially contribute the most to the knee MSD risk measurement.

All of the findings have been drawn based on the experimental study done with non-professional roofers in this study and hence might not be generalizable. Future work will include professional roofers to study knee MSD risk in shingle installation including starter and ridge cap shingles. Moreover, effective knee interventions (wearable and external) will be tested and developed with their participation in real work settings. EMG signals of thigh muscles and knee

joint loadings will be observed and analyzed to assist such process. This study provided a method that may generate useful information on the unsafe condition of shingle installation operations to facilitate effective intervention strategies (education, training, and tools) for reducing knee exposure and hence minimizing knee injuries and disorders among construction roofers.

# CHAPTER 5: FUSING IMPERFECT EXPERIMENTAL DATA FOR RISK ASSESSMENT OF MUSCULOSKELETAL DISORDERS IN CONSTRUCTION USING CANONICAL POLYADIC DECOMPOSITION

**Abstract**

Field or laboratory data collected for work-related musculoskeletal disorders (MSDs) risk assessment in construction often becomes unreliable as a large amount of data go missing due to technology-induced errors, instrument failures or sometimes at random. Missing data can adversely affect the assessment conclusions. This study proposes a method that applies Canonical Polyadic Decomposition (CPD) tensor decomposition to fuse multiple sparse risk-related datasets and fill in missing data by leveraging the correlation among multiple risk indicators within those datasets. Two knee work-related MSD risk-related datasets—3D knee rotation (kinematics) and electromyography (EMG) of five knee postural muscles—collected from previous studies were used for the validation and demonstration of the proposed method. The analysis results revealed that for a large portion of missing values (40%), the proposed method can generate a fused dataset that provides reliable risk assessment results highly consistent (70%-87%) with those obtained from the original experimental datasets. This signified the usefulness of the proposed method for use in work-related MSD risk assessment studies when data collection is affected by a significant amount of missing data, which will facilitate reliable assessment of work-related MSD risks among construction workers. In the future, findings of this study will be implemented to explore whether, and to what extent, the fused dataset outperforms the datasets with missing values by comparing consistencies of the risk assessment results obtained from these datasets for further investigation of the fusion performance.

## 5.1. Introduction

Work-related musculoskeletal disorders (MSDs) are one of the most common causes of days away from work and physical disabilities in the construction industry (BLS 2019). An increased exposure to risk factors in the workplace can enhance the likelihood of work-related MSDs; hence, proper identification of possible risk exposures and developing MSD prevention strategies are essential to alleviate work-related MSDs.

Collecting risk exposure data with human subject involvement is most often accepted as the gold standard for understanding risky behaviors and conditions that may expose workers to MSD risks on construction sites (David 2005). Generally, these data are collected in laboratory settings or real construction sites by technologies such as optical motion capture systems or surface electromyography sensors. However, data collected from these technologies often suffer from 'drop out', a phenomenon in which data is missing due to technology-induced errors (e.g., disconnection of sensors, errors in communicating with the database server, instrument failures), human-induced errors (e.g., accidental human omission) or other unknown reasons (Data 2016). The result is incompleteness of the collected risk exposure data that may lead to invalid conclusions on the effects of the potential work-related MSD risk factors. Missing data is a common problem associated with data collection in ergonomic risk assessment using technologies, regardless of the quality of the research design (Young et al. 2011). Therefore, it should be carefully handled. In doing so, reserving the interrelation among the potential risk factors and the risk indicators across multiple datasets is necessary. Among several benefits of data fusion, one is revealing the latent pattern of the data and leveraging collaborative relationships among various factors within multiple datasets based on that pattern. This benefit can be utilized to reserve the interrelation among different factors and the risk indicators across the datasets to fill in the missing

data. This study proposes a method for dealing with multiple imperfect and incomplete datasets by applying a Canonical Polyadic Decomposition (CPD) technique to treat the imperfect data for work-related MSD risk assessment. CPD decomposes the incomplete datasets based on the latent relationship among different risk factors and the risk indicators, then reconstructs a new dataset through fusion as a high-order tensor (Rabanser et al. 2017). This newly reconstructed dataset is referred to as fused dataset, which can then be used for assessing the risk of work-related MSDs. To validate the effectiveness of the CPD-based method in assessing work-related MSDs, two work-related MSD risk-related datasets collected from prior experimental studies (original datasets) were intentionally modified to represent incomplete datasets. Then CPD was applied for fusion and to reconstruct the fused datasets. The risk assessment results obtained using the fused datasets were further compared to those obtained by using the original datasets to evaluate the performance of the fusion treatment.

### 5.2. Background

### 5.2.1. Importance of research

Missing data is a common problem in research studies that involve human subjects and technologies for data collection. They can reduce statistical power of a study and lead to erroneous conclusions (Lin et al. 2012). To potentially mitigate this issue, the sample size for data collection is typically increased. However, this is not always possible due to research design, limitations in budget and human resources. It is not always feasible to regenerate the data by repeating the experiment, as it can be costly and time-consuming. There are existing methods of handling missing data in literature such as simply omitting observations with the missing data (Kang 2013), regression imputation (Button et al. 2013), mean substitution (Malhotra 1987), replacing the missing data with the last observed values (Hamer and Simpson 2009), and estimating the missing

data using conditional distribution of the other variables (Gelman and Raghunathan 2001). While useful, these methods may not be optimal for tackling data missing in work-related MSD risk assessment. Simply omitting observations with missing data typically decreases the sample size and may adversely affect the statistical power (Kang 2013). Regression imputation predicts a missing value from other variables, but it barely adds any new information except increasing the sample size and compromising the standard error (Button et al. 2013). If there is a great inequality in the proportion of missing values of different variables, mean substitution may lead to inconsistent biases and underestimation of errors (Malhotra 1987). Replacing missing data with the last observed values assumes that there will be no changes in the outcome (Hamer and Simpson 2009). Using conditional distribution of the other variables becomes inappropriate when a large amount of data is missing, as in this situation, the relationship among the variables needs to be properly computed (Gelman and Raghunathan 2001; Kang 2013).

Moreover, in real life scenarios, a phenomenon can be described and characterized with a set of factors. Each of these factors is referred to as the 'dimension' of that phenomenon (Lahat et al. 2015). When a research problem includes multiple dimensions, it is considered 'multidimensional'. This is often the case in work-related MSD risk studies, in which risk exposure data can be collected in terms of different measurements, often referred to as risk indicators, and working conditions for the purpose of in-depth understanding of risks. The working conditions may include various work settings and postures, and the risk indicator measurements may include kinematics, kinetics and electromyography (EMG) measurements. Quite often, multiple risk indicators can provide a more detailed insight into the worker's risk-inducing behavior, compared to what a single risk indicator can provide as these multiple risk indicators may be correlated. In work-related MSD risk studies that involves analyzing the effects of working conditions on risk

indicators, the working conditions and the risk indicators can be termed as a 'dimension' of the risk phenomenon. This type of data can be considered as high-dimensional data. When data is missing in multiple high-dimensional risk-related datasets, the existing methods will not be suitable as the interrelation among different risk indicators and potential risk factors across multiple datasets may not be well captured. Fuzzy modeling has been used in multidimensional missing data imputation (Amiri and Jensen 2016). However, for high-dimensional datasets where capturing the latent relationship among various dimensions is essential during imputation, the performance of Fuzzy algorithms degrades (Luengo et al. 2012; Winkler et al. 2011). Besides, the performance of Fuzzy models in dealing with multiple datasets is yet to be tested. In contrast, data fusion can be an efficient method in this regard that integrates information based on the interrelation among the factors and risk indicators in such a way that it can produce a more complete representation of the measurements of the risk indicators, compared to that of an incomplete dataset (Castanedo 2013).

### 5.2.2. State of research on data fusion in construction

In the construction industry, data fusion methods have been used for automated identification, location estimation, dislocation detection of construction materials in jobsites (Razavi and Haas 2010; Razavi and Haas 2011), automated progress tracking of construction projects (Shahi et al. 2012), structural health monitoring (Soman et al. 2018), and damage identification of civil structures (Anaissi et al. 2018). In construction ergonomics, a computationally efficient approach using data fusion was developed to recognize construction workers' awkward postures (Chen et al. 2017). Fusion of data from continuous remote monitoring of construction workers' location and physiological status was used to identify safe and unsafe behaviors of construction workers (Cheng et al. 2012) . Fusion of spatiotemporal and workers'

thoracic posture data has been done for understanding the worker's activity type for productivity assessment (Cheng et al. 2013). A position and posture data fusion method was proposed for evaluation of construction workers' behavioral risks (Chen et al. 2019) . However, no study has been done that uses data fusion of incomplete risk-related datasets for assessing work-related MSD risks.

### 5.2.3. State of research on tensor decomposition for data fusion

To integrate multidimensional data, it is essential to learn the relationships between those dimensions to understand the multidimensional nature of a phenomenon (Lahat et al. 2015). There are currently several existing approaches for analyzing multiple datasets. Zheng et al. (Zheng 2015) summarized the methods of fusing multiple datasets into three categories. The first category is stage-based fusion methods, where different datasets are used at different stages of data mining tasks (Zheng et al. 2011 ). The second category is deep learning-based fusion methods that use a neural network to extract original features from multiple datasets and learn a new representation of those features for classification and prediction purposes (Wen et al. 2016). The third category is semantic meaning-based fusion methods that identify the association between different features across multiple datasets and carry the semantic meaning, i.e., the latent relationship between the features during fusion (Zheng 2015). Similarity-based data fusion is one type of the semantic meaning-based fusion methods, in which similarity between the features are measured from multiple datasets. These similarities can leverage the correlation among the features collectively and help obtain any missing information of a feature based on the information available for another feature. Based on the similarities, different datasets can be fused (Zheng 2015). This study is interested in fusion of risk-related datasets containing different features that represent work-related MSD risk factors in construction and are inherently related to each other. Understanding the latent

relationship between different features across multiple datasets plays a vital role in proper fusion. Also, learning the similarities between those features is useful to impute the missing values. As a result, the similarity-based data fusion methods lend themselves well to tackling the fusion problem in this study.

Tensor decomposition is one type of similarity-based data fusion methods (Sorber et al. 2015). Tensors are generalizations of matrices to higher dimensions, and are powerful to model multidimensional data (Papalexakis et al. 2017). Tensors are multidimensional arrays of numerical values that are widely used in different applications in data analysis and machine learning, including filling in missing values (Dauwels et al. 2012), anomaly detection (Xie et al. 2017), object profiling (Charlier et al. 2018), discovering patterns (Xiong et al. 2010), and predicting evolution (Dunlavy et al. 2011). Matrix factorization was previously used in filling out missing data (Žitnik and Zupan 2014); however, it only works with two-dimensional data and may not be applicable to high-dimensional data like those used in work-related MSD risk studies. Tensor decomposition is a useful tool to deal with high-dimensional datasets. It accurately extracts the correlations among various dimensions from different datasets and learns the latent structures and collaborative relationships among the dimensions to approximate the pattern of the data (Kolda and Bader 2009).

With increases in dimensions of the datasets, the number of the elements in a tensor also increases. This increase in elements make the dataset difficult to deal with in terms of computational and memory requirements. By decomposing the high-dimensional data presented in form of a tensor, the problems associated with high-dimensionality can be alleviated or even removed (Vervliet et al. 2014). Also, tensor decomposition can be used as an efficient tool for missing value prediction (Acar et al. 2011; Asif et al. 2016). Though tensor decomposition has

been abundantly applied in areas such as scientific computing, signal processing, and social media data analysis, its use in construction industry-related datasets is limited. One application of tensor decomposition with construction datasets was concerned with awkward posture recognition in which worker's motion data was presented as high order tensor (Chen et al. 2017). However, this technique has yet to be established as an alternative method of treating imperfect experimental data to provide meaningful and accurate risk assessment results in general work-related MSD risk assessment studies.

The two most widely used tensor decomposition algorithms are *canonical polyadic decomposition* (CPD) and *Tucker decomposition* (TD). CPD is generally advised for use with latent parameter estimation, while the TD is suggested in use of subspace estimation, compression, and dimensionality reduction (Rabanser et al. 2017). As the objective of this study is to fuse two incomplete datasets based on the latent correlations among all dimensions of those datasets, CPD is more suitable than TD in fulfilling the purpose of this study. Moreover, CPD is very effective in capturing the interactions among various dimensions of a high-dimensional datasets and therefore, it can effectively be used for imputing missing data (Dauwels et al. 2012). Considering many other tensor decomposition techniques are based on CPD and TD and CPD is more suitable for this study, the following subsection will only discuss CPD for tensor decomposition.

**5.2.3.1. Canonical polyadic decomposition (CPD)**

In order to impute missing data, CPD extrapolates the latent structures and collaborative relationships among different dimensions, such as rows and columns of a tensor. The CPD method factorizes a tensor into a sum of component rank-one tensors (Kolda and Bader 2009). For example, in CPD, a given three-dimensional tensor $Y \in \mathbb{R}^{I \times J \times K}$ can be written as:

$$Y \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r \equiv [\![A, B, C]\!] \quad \text{............................................................ (5.1)}$$

Where $R$ is a positive integer, which is referred to as the rank of a tensor. The rank of tensor $Y = R$ can be defined as the smallest number of rank-1 tensor that is required to represent the tensor $Y$ as their sum (Rabanser et al. 2017) (Figure 1); Rank-1 tensor is defined as a decomposition of an $N$-dimensional tensor into one outer product of $N$ vectors. $R$ is also called latent factor. $I$, $J$ and $K$ are the sizes of the dimensions of the tensor $Y$ where, $\boldsymbol{a}_r \in \mathbb{R}^I$, $\boldsymbol{b}_r \in \mathbb{R}^J$ and $\boldsymbol{c}_r \in \mathbb{R}^K$ for $r = 1...R$. '∘' indicates the vector outer product. $\boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r \in \mathbb{R}^{I \times J \times K}$ is an outer product of three vectors $\boldsymbol{a}_r$, $\boldsymbol{b}_r$ and $\boldsymbol{c}_r$ and is referred to as a rank-1 tensor. $A$, $B$ and $C$ are the factor matrices obtained for each dimension after decomposition of the tensor $Y$ and are referred to as the combination of the vectors from the rank-1 components. Here, the factors matrices $A = [\boldsymbol{a}_1, \boldsymbol{a}_2,..., \boldsymbol{a}_R] \in \mathbb{R}^{I \times R}$, $B = [\boldsymbol{b}_1, \boldsymbol{b}_2,..., \boldsymbol{b}_R] \in \mathbb{R}^{J \times R}$ and $C = [\boldsymbol{c}_1, \boldsymbol{c}_2,..., \boldsymbol{c}_R] \in \mathbb{R}^{K \times R}$. Their column vectors are expressed explicitly as: $\boldsymbol{a}_r = [a_{1r}, a_{2r}, a_{3r}, ..., a_{Ir}]^T$, $\boldsymbol{b}_r = [b_{1r}, b_{2r}, b_{3r}, ..., b_{Jr}]^T$, $\boldsymbol{c}_r = [c_{1r}, c_{2r}, c_{3r}, ..., c_{Kr}]^T$.

Elementwise, equation (5.1) can be expressed as:

$$Y_{ijk} \approx \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} \text{ for } i = 1...I; j = 1...J; k = 1...K. \ldots\ldots\ldots (5.2)$$

CPD of a 3-dimensional tensor can be formalized as follows:

$$\min_{\widehat{Y}} \| Y - \widehat{Y} \|; \text{Where, } \widehat{Y} = \sum_{r=1}^{R} \boldsymbol{a}_r \circ \boldsymbol{b}_r \circ \boldsymbol{c}_r \equiv [\![A, B, C]\!] \ldots\ldots\ldots (5.3)$$

Where $Y$ is the original tensor and $\widehat{Y}$ is a low rank approximation of tensor $Y$. CPD of a tensor is typically computed using the alternating least square (ALS) algorithm where each factor matrix is iteratively solved using a least square method. This algorithm fixes all factor matrices, except one, to optimize for the non-fixed matrix and then repeats this procedure for each matrix repeatedly until it converges (Kolda and Bader 2009; Kolda and Sun 2008).

The CPD of a three-dimensional tensor is illustrated in Figure 5.1.
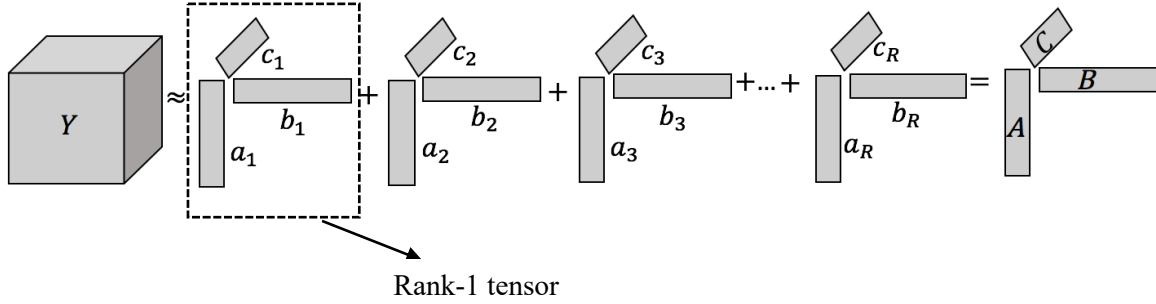
**Figure 5.1:** CPD tensor decomposition

## 5.3. Problem Statement and Research Objective

For assessing work-related MSD risks among construction workers, human-based data can potentially suffer from missing data points due to dropout from the data collection technology. However, an in-depth method in handling multiple imperfect datasets for assessing work-related MSD risks is missing in the existing literature. Tensor decomposition-based data fusion can be potentially useful in this regard. It may help understand the data distribution of each dataset and consider the correlation among risk indicators captured at multiple work settings or conditions to fill in the missing data points. Therefore, the objective of this research is to develop a method that applies CPD tensor decomposition techniques to fuse multiple imperfect datasets and replace missing data by considering the correlation among the risk indicators for assessing the work-related MSD risks among construction workers.

## 5.4. Proposed Method

Figure 5.2 provides a schematic overview of the proposed method. First, multiple risk-related incomplete datasets captured in multiple experimental settings are represented as high-dimensional tensors. Then these tensors are fused by applying the CPD tensor decomposition. The CPD first decomposes these tensors into factor matrices that represent the latent structures and collaborative relationships among all dimensions. Based on the correlation among different

dimensions, the CPD then reconstructs a new dataset with all missing values filled. This fused new dataset is ready to be used for subsequent risk assessments. In the following subsections, the steps of the proposed method are described in detail.
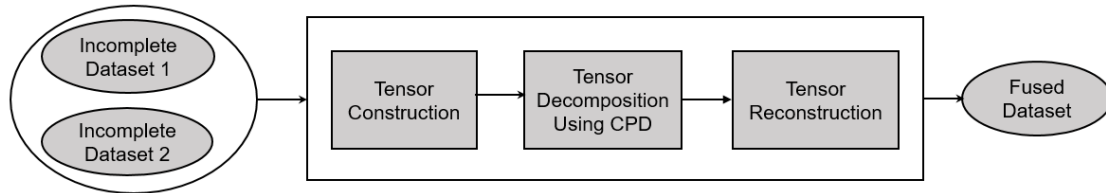


**Figure 5.2:** Overview of Proposed Method

### 5.4.1. Incomplete datasets for MSD risk assessment

For work-related MSD risk assessment in construction, the effects of various work-related factors on specific risk indicators are often measured. Therefore, a typical work-related MSD risk-related dataset contains information such as work settings/experimental conditions and specific risk indicators. Often, a work-related MSD risk may not be characterized with a single type of risk indicator and in that case, multiple types of work-related MSD risk indicators are measured, e.g., kinematics (flexion, abduction/adduction, internal/external rotation), muscle activity data recorded by EMG, and kinetic data recorded by force plates, and can be available in multiple datasets.

To describe the proposed method, consider two datasets DS1 and DS2 (denoted for *Incomplete Dataset 1* and *Incomplete Dataset 2* in Figure 5.2). Each dataset contains two different types of work-related MSD risk indicators collected from several subjects, investigating various work-induced risks for a number of work settings (e.g., working technique, working posture, working condition). This is common practice of work-related MSD risk assessment studies in construction, where different types of risk indicators are often measured from the same group of people and the same work settings for an in-depth understanding of a specific type of risk. In this case, a data value in DS1 or DS2 represents a value of a risk indicator measured from a subject at

certain work settings. For example, to measure the knee work-related MSD risk, two common types of risk indicators are measured – awkward knee rotations and maximum EMG of knee postural muscles. Awkward knee rotations may include flexion, abduction/adduction, and internal/external rotations. Maximum EMG of knee postural muscles may include measurements of maximum flexor and extensor muscle activations. The datasets can be represented in a tabular form as shown in Table 5.1.

**Table 5.1:** Formats of DS1 and DS2

| Dataset (DS) | Tabular Format |
|---|---|
| DS1 | $subject, setting_1, setting_2, …, setting_q, ind_1^1, ind_1^2, …, ind_1^m$ |
| DS2 | $subject, setting_1, setting_2, …, setting_q, ind_2^1, ind_2^2, …, ind_2^n$ |

In Table 5.1, each dataset contains a type of risk indicator measurements collected for several subjects at a number of work settings. The column *subject* contains subjects' IDs. The columns $setting_1, setting_2, …, setting_q$ contain arrangements of $q$ work settings at which measurements of $m$ risk indicators in columns $ind_1^1, ind_1^2, …, ind_1^m$ are collected in DS1, and $n$ risk indicators are collected in columns $ind_2^1, ind_2^2, …, ind_2^n$ in DS2. For example, the column $setting_1$ can represent working postures that have two arrangements – static and dynamic. The column $setting_2$ can represent slopes of the working site that have three arrangements – 0°,15° and 30°. So, in the column $setting_1$, the values are either 'static' or 'dynamic' and in the column $setting_2$, the values are either '0°', '15°' or '30°'. If a study involves only these two settings, then the value of $q$ will be 2. Similarly, the columns $ind_1^1, ind_1^2, …, ind_1^m$ can contain kinematics measures (flexion, abduction/adduction, internal/external rotation) and $ind_2^1, ind_2^2, …,$ $ind_2^n$ can contain EMG measures of several muscles. To be more specific with an example, in DS1, a value of any of the $m$ risk indicator columns can be a certain measurement value of a risk indicator (e.g., flexion) captured for certain subject at certain arrangements of $q$ work settings.

Note that in both datasets, the values of columns of *subject* and $setting_1, setting_2, ..., setting_q$ are identical; the only difference is the measurements of the risk indicators. These datasets become incomplete when a significant amount of risk indicator measurements is not captured during data collection process.

### 5.4.2. Tensor construction

Since the datasets contain multiple risk indicator measurements for the same combination of subjects and settings, the risk indicator measurements from both datasets are concatenated, yielding a new dataset (DS3) as illustrated in Table 5.2. This process is called early integration data fusion (Žitnik and Zupan 2014).

**Table 5.2:** Format of DS3

| Dataset (DS) | Format |
|---|---|
| DS3 | $subject, setting_1, setting_2, ..., setting_q, ind_1^1, ..., ind_1^m, ind_2^1, ..., ind_2^n$ |

In Table 5.2, each risk indicator measurement in DS3 is collected for certain subject at different arrangements of $q$ work settings. Thus, 'subject' and each 'work setting' are considered as dimensions of each risk indicator, resulting in a total of (*1+q)* dimensions. For multiple risk indicators, this method allows datasets to be represented as a 'multi-dimensional' or 'multi-modal' tensor, where each cell of the tensor contains a value of a certain risk indicator, measured for a subject at certain arrangements of $q$ work settings. In this study, the 'subject', each 'work setting', and the entire set of 'risk indicators' can be considered as dimensions of the tensor which has a total of *(1+ q + 1)* dimensions.

To illustrate this tensor representation with a simplified example, consider only one work setting (e.g., working postures) denoted by '*setting*'. In this case, DS3 can be represented as a three-dimensional tensor (*subject × risk indicators × setting*) $(M_{i,j,k}) \in \mathbb{R}^{s \times r \times t}$ (Figure 5.3), where *i*, *j* and *k* correspond to dimensions *'subject', 'risk indicators'* and *'setting'* respectively; and *s*, *r* and

*t* are their sizes where $i \in \{1,2,\ldots s\}$, $j \in \{1,2,\ldots r\}$ and $k \in \{1,2,\ldots t\}$. Each slice along the *setting* axis in Figure 5.3 represents a matrix $(M_{i,j}) \in \mathbb{R}^{s \times r}$ given a setting arrangement *k*.
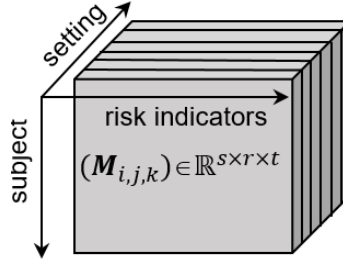


**Figure 5.3:** Schematic representation of DS3 as a 3D tensor $M$

### 5.4.3. Tensor decomposition

Following tensor construction, CPD is applied to decompose the tensor into factor matrices. As mentioned earlier, CPD learns the latent structures and collaborative relationships among the dimensions of a tensor. Thus, each factor matrix represents the latent relationships between the corresponding dimension and all the other dimensions captured by latent factor $R$ *(*previously discussed in Section 5.2.3.1). Before applying CPD, the latent factor $R$ should be computed properly to ensure unique mapping of the original tensor to the decomposed tensor in the form of factor matrices. In the above example, $(M_{i,j,k}) \in \mathbb{R}^{s \times r \times t}$ is associated with three dimensions – '*subject*', '*risk indicators*' and *'setting'*; hence, three factor matrices are obtained after decomposition, denoted as $FM_1 \in \mathbb{R}^{s \times R}$, $FM_2 \in \mathbb{R}^{r \times R}$, $FM_3 \in \mathbb{R}^{t \times R}$, as shown in Figure 5.4. Here, the factor matrix $FM_1$ obtained for the dimension '*subject*' represents the latent relationships between the subject and the risk indictors collected at different arrangements of the setting. $FM_2$ is obtained for the dimension '*risk indicators*' and represents the latent relationships between the risk indicators and the subjects at different arrangements of the setting for which those risk indicators were collected. $FM_3$ is obtained for the dimension '*setting*' and represents the latent relationships between different arrangements of the setting and the risk indicators collected for the subjects at

those arrangements of the setting. CPD leverages all the inter-dimensional correlations among all the dimensions of a tensor and performs data fusion based on those correlations to create a reconstructed tensor. The structures of the factor matrices and the method of computing $R$ will be discussed further in "Implementation and Results".
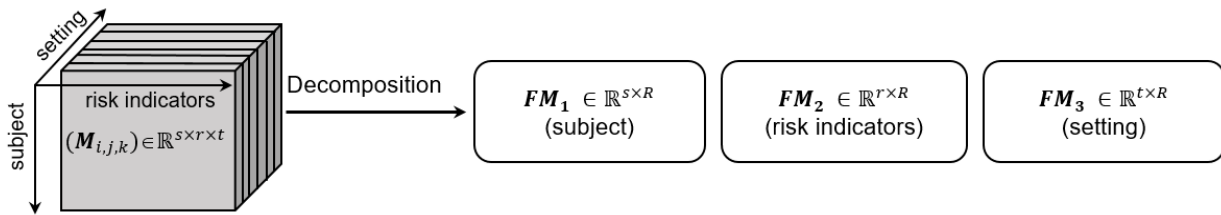


**Figure 5.4:** Schematic illustration of CPD decomposition of the 3D tensor $\boldsymbol{M}$

### 5.4.4. Tensor reconstruction

After tensor decomposition, the resulting factor matrices are used to reconstruct a new fused tensor where the missing values are filled, and values of the MSD risk indicators are fused based on the inter-dimensional correlations. In the prior example, such tensor can be denoted as $(\boldsymbol{M'}_{i,j,k}) \in \mathbb{R}^{s \times r \times t}$, which is computed using the following equation:

$$\boldsymbol{M'}_{i,j,k} = \sum_{r=1}^{R} \boldsymbol{FM_1}_r \circ \boldsymbol{FM_2}_r \circ \boldsymbol{FM_3}_r \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (5.4)$$

Where $\boldsymbol{FM_1}_r$, $\boldsymbol{FM_2}_r$, $\boldsymbol{FM_3}_r$ are the column vectors of $\boldsymbol{FM_1}$, $\boldsymbol{FM_2}$ and $\boldsymbol{FM_3}$ respectively, $R$ is the latent factor and '∘' is the vector outer product. Using the newly reconstructed tensor, values of the MSD risk indicators for different subjects at different settings are extracted and reorganized into a tabular form, readily available for risk assessment.

## 5.5. Implementation and Results

### 5.5.1. Original datasets collected from previous experiments

The current study considered two risk-related datasets that were collected from the authors' prior human subject laboratory experimental studies. These studies assessed work-related factors

for knee MSDs among residential construction roofers who work on sloped environments. One dataset contains calculated knee rotation (kinematics) data, representing five knee rotational angles – flexion, abduction, adduction, internal and external rotation (Breloff et al. 2019a). The second data set contains EMG data, describing muscle activation of five knee postural muscles – biceps femoris, rectus femoris, semitendinosus, vastus lateralis and vastus medialis (Dutta et al. 2020b). Awkward knee rotations and maximum muscle activation have been identified as two quantitative risk indicators of work-related MSDs in low extremities (Hofer et al. 2011; Kingston et al. 2016). The five knee rotation angles represent the MSD risk associated with awkward and extreme kneeling postures, while the EMG data represent the MSD risk associated with heightened activation of knee postural muscles that might cause knee joint overloading. Both types of data were collected from the previous experimental study in which 9 subjects simulated the shingle installation roofing task at three roof slopes (0°, 15° and 30°) with two kneeling postures (static and dynamic). Each participant performed the task for five trials. Although the experiment was designed and implemented carefully, both datasets initially collected from the experiment had missing values. About 2% data were missing in the knee rotation dataset and about 20% data were missing in the EMG dataset at random. In this study, observations with missing data were intentionally removed from the initial datasets for the sake of obtaining complete datasets for evaluation. In the knee rotation dataset, each data point represents the maximum angle value of a knee rotation for a certain subject at a specific slope, posture and trial. In the EMG dataset, each data point represents the maximum normalized EMG value of a knee postural muscle of a certain subject at a specific slope, posture and trial. Since only maximum measurements in both datasets were collected as risk indicators, the two datasets could be easily synchronized.

### 5.5.2. Imperfect (or incomplete) datasets construction (datasets with missing values)

To evaluate the performance of the proposed method in dealing with missing data, this study was designed to fuse knee rotation and EMG datasets for different proportions of missing values using the CPD method. A proportion of data was randomly removed from both knee kinematics and EMG datasets with percentages of missing data as shown in Table 5.3. Therefore, fourteen incomplete datasets—seven for knee rotations and seven for EMG—were constructed with different portions of missing values.

**Table 5.3:** Proportion of missing data

| Datasets | Missing proportion (%) | | | | | | |
|---|---|---|---|---|---|---|---|
| Knee kinematics data | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| Knee EMG data | 10 | 20 | 30 | 40 | 50 | 60 | 70 |

### 5.5.3. Tensor construction

The incomplete rotation and EMG datasets were then used for tensor construction. As the knee rotation and EMG data were collected from the same subjects at the same work settings (i.e., roof slope, working posture, and trial), they could be combined by the early integration method (Pavlidis et al. 2002). In this method, the columns of these two types of risk indicators were concatenated to form a single set of length ten vectors which was referred as 'risk indicators' in this study. For tensor construction, forty-nine incomplete datasets as shown in Table 5.4 were considered. One sparse rotation dataset and one sparse EMG dataset were selected according to the combinations of missing proportions specified in Table 5.4.

An $N$ dimensional tensor $T_N \in \mathbb{R}^{I_1 \times I_2 \times \cdots \cdots \times I_N}$ has size $(T_N) = I_1 \times I_2 \times \ldots \ldots \times I_N$, where, $I_i$ is the size of its $i^{th}$ dimension (Dauwels et al. 2012). As the risk indicators were measured from multiple subjects at different working postures, roof slopes and multiple trials, the collected data

could be represented as a five-dimensional sparse tensor $T \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$, where $I_1 =$ No. of

subjects (total 9), $I_2 =$ No. of risk indicators [EMG and kinematics; total 10 (5 knee rotation angles

and 5 knee muscle EMG)], $I_3 =$ No. of trials (total 5), $I_4 =$ No. of postures [total 2 (static and

dynamic)], $I_5 =$ No. of slopes [total 3 (0°,15°,30°)]. In short, the tensor $T$ maps the knee rotations

and EMG measurements of nine subjects, obtained from five trials, performed on three roof slopes,

using two postures. Tensor $T$ was referred to as 'incomplete tensor' in this study.

**Table 5.4:** Incomplete datasets for tensor construction

| Missing proportion from kinematics data (%) | Missing proportion from EMG data (%) | | | | | | |
|---|---|---|---|---|---|---|---|
| 10 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| 20 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| 30 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| 40 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| 50 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| 60 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
| 70 | 10 | 20 | 30 | 40 | 50 | 60 | 70 |

For a certain posture and slope, the sparse tensor $T$ is denoted as $T_{i,j}$ which represents the

values of the risk indicators obtained from five trials performed by each subject at that roof slope

and posture. In $T_{i,j}$ , $i \in \{1,2\}$, with 1 and 2 representing the static and dynamic postures

respectively; $j \in \{1,2,3\}$, with 1, 2 and 3 representing the 0°, 15° and 30° slopes respectively.

Then, the five-dimensional tensor $T$ can be illustrated in Figure 5.5.

Similarly, the original datasets collected from the previous experiments were constructed

to a five-dimensional tensor $X \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$ as the benchmark data, which was referred to as

the "original tensor" in this study and was used for validation of performance of the fusion method.
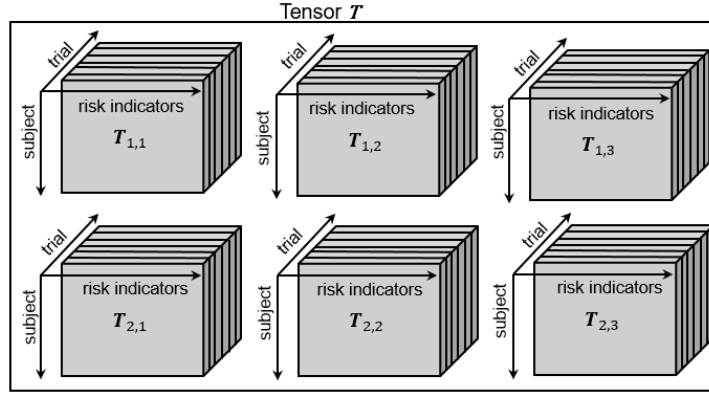
**Figure 5.5:** Representation of the five-dimensional tensor $T$

## 5.5.4. Tensor decomposition

Once a tensor $T$ was constructed, the CPD tensor decomposition was applied to decompose the tensor $T$ into five factor matrices, representing information in each dimension. The factor matrices were represented as matrices $A$, $B$, $C$, $D$, and $E$, providing latent information in dimensions of subjects, risk indicators, trials, postures, and slopes, respectively.

CPD factorizes the five-dimensional sparse tensor $T \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$ as:

$$T \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r \circ d_r \circ e_r \equiv [\![A, B, C, D, E]\!] \quad \dots\dots\dots\dots\dots\dots\dots\dots (5.5)$$

Where $a_r \in \mathbb{R}^{I_1}$, $b_r \in \mathbb{R}^{I_2}$, $c_r \in \mathbb{R}^{I_3}$, $d_r \in \mathbb{R}^{I_4}$, $e_r \in \mathbb{R}^{I_5}$, for $r = 1, \dots, R$. Factor matrix $A = [\,a_1, a_2 \,\dots\dots\dots\, a_R] \in \mathbb{R}^{I_1 \times R}$, $B = [\,b_1, b_2 \,\dots\dots\dots\, b_R] \in \mathbb{R}^{I_2 \times R}$, $C = [\,c_1, c_2 \,\dots\dots\dots\, c_R] \in \mathbb{R}^{I_3 \times R}$, $D = [\,d_1, d_2 \,\dots\dots\dots\dots\, d_R] \in \mathbb{R}^{I_4 \times R}$, and $E = [\,e_1, e_2 \,\dots\dots\dots\dots\, e_R] \in \mathbb{R}^{I_5 \times R}$ are collections of all $R$ vectors of $a_r, b_r, c_r, d_r, e_r$ respectively.

Figure 5.6 illustrates the process of the CPD tensor decomposition.
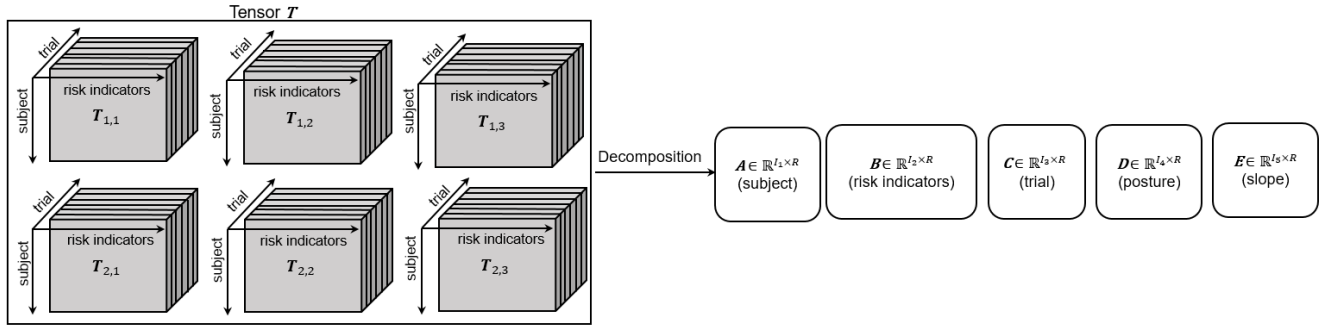
**Figure 5.6:** CPD tensor decomposition of tensor $\boldsymbol{T}$

Equation (5.5) can be represented as:

$$\boldsymbol{T} \approx \sum_R \boldsymbol{A}_r \circ \boldsymbol{B}_r \circ \boldsymbol{C}_r \circ \boldsymbol{D}_r \circ \boldsymbol{E}_r \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(5.6)}$$

Factor matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ and $\boldsymbol{E}$ are computed by solving the following optimization equation:

$$\arg\min \| \boldsymbol{T} - \sum_R \boldsymbol{A}_r \circ \boldsymbol{B}_r \circ \boldsymbol{C}_r \circ \boldsymbol{D}_r \circ \boldsymbol{E}_r \| \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(5.7)}$$

To solve Equation (5.7), the alternating least squares algorithm was applied using an optimization-based program *cpd_als* in Tensor Toolbox from MATLAB (Bader and Kolda 2019).

### 5.5.4.1. Selection of $R$ (number of rank-1 tensors) for CPD

Before applying CPD, a proper $R$ needs to be selected to guarantee a unique mapping of the original tensor to the decomposed tensor. To determine the smallest number of rank-1 tensors required for CPD, the core consistency diagnostic method (CORCONDIA) described in (Bro and Kiers 2003) was applied with the N-way toolbox of MATLAB (Andersson and Bro 2000). In this method, a Tucker core is used for assessing the appropriateness of a CPD model. Tucker core array, when presenting the coordinates of the parts of the multi-dimensional array within the subspaces, is considered as the regression of a multi-dimensional array onto the subspaces characterized by the factor matrices (Bro and Kiers 2003). In the CONCORDIA method, to signify an optimal representation of the multi-dimensional array with respect to the subspaces defined by

the factors, a parameter called "core consistency" is measured. A core consistency value close to 100% indicates an appropriate CPD model that can properly represent the variability of the factors. With the increase of the number of components ($R$), the core consistency value typically increases up to a certain point and then decreases. To apply CPD in this study, one component model ($R=1$) was first attempted and then different values of $R$ were tried by gradually increasing the number of components until the core consistency value reached close to 100%. Those values of $R$ with the highest core consistencies were considered for CPD application.

### 5.5.5. Tensor reconstruction (fused dataset for risk assessment)

Once the tensor was decomposed into factor matrices, these factor matrices were used to reconstruct a new tensor from its decomposed state. This reconstructed tensor is an approximate mapping of the sparse tensor $\boldsymbol{T}$, which was referred to as 'fused tensor' and represented as $\boldsymbol{T}' \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$, calculated by:

$$\boldsymbol{T}' = \sum_{r=1}^{R} \boldsymbol{A}_r \circ \boldsymbol{B}_r \circ \boldsymbol{C}_r \circ \boldsymbol{D}_r \circ \boldsymbol{E}_r \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{(5.8)}$$

In this tensor, all the missing values were imputed, and the risk indicator values were reconstructed. Tensorlab (a MATLAB package for tensor computation) was used for the imputation and reconstruction (Debals et al. 2019) . Note that the factor matrices computed by Equation (5.7) ensured the minimal reconstruction error during the imputation and reconstruction of the risk indicators. The reconstructed risk indicators were then extracted to a spread sheet, readily for use in work-related MSD risk assessment.

To allow for reliable observations, the procedures in subsections 5.5.2 to 5.5.5 (except for construction of the five-dimensional tensor $\boldsymbol{X}$ from the original datasets) were repeated three times. In each repeat, the random removal of certain portion of data as designed earlier led to different

preservation in each dataset. Three repeats ensured a two-thirds (67%) probability that the averaged results were more accurate than a single experiment while maintaining time-efficiency in implementation.

### 5.5.6. Performance analysis of the data fusion

The performance of the data fusion method was evaluated by comparing the reconstructed data to the original data using two statistical measures—root mean square error (RMSE) and mean absolute error (MAE). RMSE and MAE were used to measure the standard error and average magnitude of the error of the fused tensor, respectively (Alimuddin et al. 2012; Ng and Kantor 2000).

The *RMSE* was calculated using the following equation:

$$RMSE = \sqrt{\frac{1}{N}\left[\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\sum_{i_3=1}^{I_3}\sum_{i_4=1}^{I_4}\sum_{i_5=1}^{I_5}\left(X_{i_1,i_2,i_3,i_4,i_5} - T'_{i_1,i_2,i_3,i_4,i_5}\right)^2\right]} \quad \dots\dots\dots\dots\dots (5.9)$$

Here, $X_{i_1,i_2,i_3,i_4,i_5}$ is an element (a value of a risk indicator of a certain subject at a certain slope, posture and trial) in the original tensor $X$ and $T'_{i_1,i_2,i_3,i_4,i_5}$ is an element in the re-constructed tensor $T'$ (fused dataset). $N$ is the total number of cells of the tensor which was computed as $N = I_1 \times I_2 \times I_3 \times I_4 \times I_5$.

The *MAE* was calculated using the following equation:

$$MAE = \frac{1}{N}\left[\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\sum_{i_3=1}^{I_3}\sum_{i_4=1}^{I_4}\sum_{i_5=1}^{I_5}\left|\left(X_{i_1,i_2,i_3,i_4,i_5} - T'_{i_1,i_2,i_3,i_4,i_5}\right)\right|\right] \quad \dots\dots\dots\dots\dots (5.10)$$

The performance of the data fusion method was also evaluated by comparing the risk assessment results obtained from tensor $T'$ to those obtained from tensor $X$ collected in previous experimental studies. In this comparison, consistency of the effects of roof slope and working

posture on five knee rotation angles and five knee postural muscles' EMG between the original and fused datasets were considered as a performance metric. In previous studies, the authors explored if the work-related factors (i.e., roof slope and working posture) are associated with knee MSDs among construction roofers (Breloff et al. 2019a; Dutta et al. 2020b). Those studies considered the original tensor *(X)* for risk assessments, whereas in this study, the fused tensor $(T')$ was considered for assessing the similar risk to see if the risk assessment results were consistent. This way, the performance of the fusion method could be evaluated.

The consistency (Con) was computed using the following equation:

$$\text{Con} = \frac{\sum_{i=1}^{P} f(m_i, m'_i)}{P} \times 100 \quad \text{...........................................................} \quad (5.11)$$

Where, $P$ = Total number of possible effects of roof slope, working posture and their interactions on five knee rotation angles and EMG of five knee postural muscles [total 30 = 3 factors (slope, posture, slope-posture interaction) ×10 response variables (5 knee rotations and 5 EMG measurements)].

$m_i$ denotes the effects of the factors (slope and posture) on the response variables obtained from the original tensor *X*.

$m'_i$ denotes the effects of the factors (slope and posture) on the response variables obtained from the fused tensor $T'$.

$$f(m_i, m'_i) = \begin{pmatrix} 1, & if \ m_i = m'_i \\ 0, & if \ m_i \neq m'_i \end{pmatrix}.$$

All the resulting measurements of the performance metrics were averaged over the three trial repetitions.

### 5.5.7. Results

Table 5.5 summarizes the results, including the RMSE, MAE, consistency (Con), number of rank-1 tensors $R$ (latent factor) and core consistency values computed for the knee rotation and EMG datasets with different proportions of missing values.

**Table 5.5:** Summary of Results

| Missing Proportion (%) | | RMSE | MAE | Con (%) | R | Core consistency (%) |
|---|---|---|---|---|---|---|
| Knee Rotation | EMG | | | | | |
| 10 | 10 | 1.39 | 0.69 | 80 | 17 | 98 |
| 10 | 20 | 1.38 | 0.8192 | 80 | 14 | 96.5 |
| 10 | 30 | 1.39 | 0.818 | 80 | 15 | 97 |
| 10 | 40 | 1.4 | 0.8412 | 80 | 14 | 96 |
| 10 | 50 | 1.44 | 0.805 | 77 | 14 | 96.5 |
| 10 | 60 | 1.57 | 0.9597 | 77 | 12 | 95 |
| 10 | 70 | 2.34 | 1.47 | 74 | 10 | 94.67 |
| 20 | 10 | 1.67 | 0.9791 | 80 | 12 | 97 |
| 20 | 20 | 1.73 | 1.04 | 80 | 11 | 97 |
| 20 | 30 | 1.78 | 1.04 | 80 | 11 | 95.6 |
| 20 | 40 | 1.75 | 1.06 | 77 | 12 | 94.8 |
| 20 | 50 | 1.79 | 1.08 | 74 | 13 | 94 |
| 20 | 60 | 1.83 | 1.09 | 74 | 13 | 93 |
| 20 | 70 | 2.45 | 1.25 | 70 | 10 | 92.7 |
| 30 | 10 | 1.82 | 1 | 77 | 13 | 97 |
| 30 | 20 | 2 | 1.14 | 74 | 11 | 95.6 |
| 30 | 30 | 2.14 | 1.17 | 74 | 11 | 96 |
| 30 | 40 | 2.19 | 1.25 | 74 | 10 | 95 |
| 30 | 50 | 2.26 | 1.29 | 74 | 10 | 95.7 |
| 30 | 60 | 2.32 | 1.24 | 74 | 10 | 93.5 |
| 30 | 70 | 5.6 | 3.4 | 57 | 11 | 94 |
| 40 | 10 | 2.11 | 1.11 | 87 | 12 | 96.5 |
| 40 | 20 | 2.95 | 1.32 | 74 | 10 | 94 |
| 40 | 30 | 2.98 | 1.31 | 74 | 10 | 95 |
| 40 | 40 | 2.52 | 1.24 | 74 | 12 | 93.7 |
| 40 | 50 | 2.83 | 1.3 | 74 | 12 | 92.76 |
| 40 | 60 | 2.7 | 1.39 | 74 | 11 | 93 |
| 40 | 70 | 2.77 | 1.51 | 70 | 13 | 92 |
| 50 | 10 | 4.9 | 1.69 | 80 | 22 | 93 |
| 50 | 20 | 5.11 | 1.81 | 80 | 31 | 93 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 50 | 30 | 5.28 | 1.99 | 80 | 29 | 92.6 |
| 50 | 40 | 6.19 | 2.24 | 77 | 34 | 92 |
| 50 | 50 | 6 | 2.4 | 77 | 43 | 91.5 |
| 50 | 60 | 6.5 | 2.7 | 77 | 40 | 92 |
| 50 | 70 | 6.8 | 3.8 | 60 | 10 | 91 |
| 60 | 10 | 3.38 | 1.45 | 84 | 12 | 93 |
| 60 | 20 | 4.4 | 1.76 | 77 | 15 | 93.6 |
| 60 | 30 | 4.63 | 1.8 | 77 | 13 | 92 |
| 60 | 40 | 5 | 2 | 77 | 13 | 92.8 |
| 60 | 50 | 5.32 | 2.08 | 74 | 10 | 93 |
| 60 | 60 | 5.39 | 2.16 | 70 | 16 | 92 |
| 60 | 70 | 5.6 | 2.2 | 70 | 11 | 91 |
| 70 | 10 | 6.7 | 2.7 | 77 | 20 | 90 |
| 70 | 20 | 7.7 | 3.2 | 74 | 17 | 89 |
| 70 | 30 | 7.5 | 3 | 74 | 12 | 91.5 |
| 70 | 40 | 7.54 | 3.38 | 70 | 10 | 90 |
| 70 | 50 | 8 | 3.3 | 70 | 11 | 92.4 |
| 70 | 60 | 8.29 | 3.4 | 70 | 11 | 90 |
| 70 | 70 | 8.8 | 3.67 | 67 | 12 | 88 |

**5.5.7.1.RMSE curves for datasets with different proportions of missing values**

Figure 5.7 displays the RMSE curves for different proportions of missing data from the knee rotation and EMG datasets. In Figure 5.7, rot-DS and EMG-DS means the rotation dataset and EMD dataset respectively.

Generally, the RMSE curve demonstrated an increasing trend with the increase of missing values in both rot-DS and EMG-DS. It was graphically observed that, for up to 40% missing data from rot-DS and up to 60% missing data from EMG-DS, the RMSE values displayed a small variation from 1 to 3. When the proportion of the missing data from rot-DS increased to 50%, the RMSE value increased drastically for different proportions of missing values in EMG-DS. The RMSE curves for 50%, 60% and 70% missing data from the rot-DS shifted slightly upward, indicating that with the increase of missing values in rot-DS, the RMSE values generally increased. From 50% to up to 70% missing values from rot-DS and up to 70% missing values from EMG-

DS, the RMSE values varied considerably from 3.5 to 9. These results suggested that with the increase in missing values in both datasets, the performance of the fusion method decreased. The variation in RMSE values was small for up to 40% missing values in rot-DS and up to 60% missing values in EMG-DS. However, with further increase of missing values in both datasets, RMSE values varied considerably.
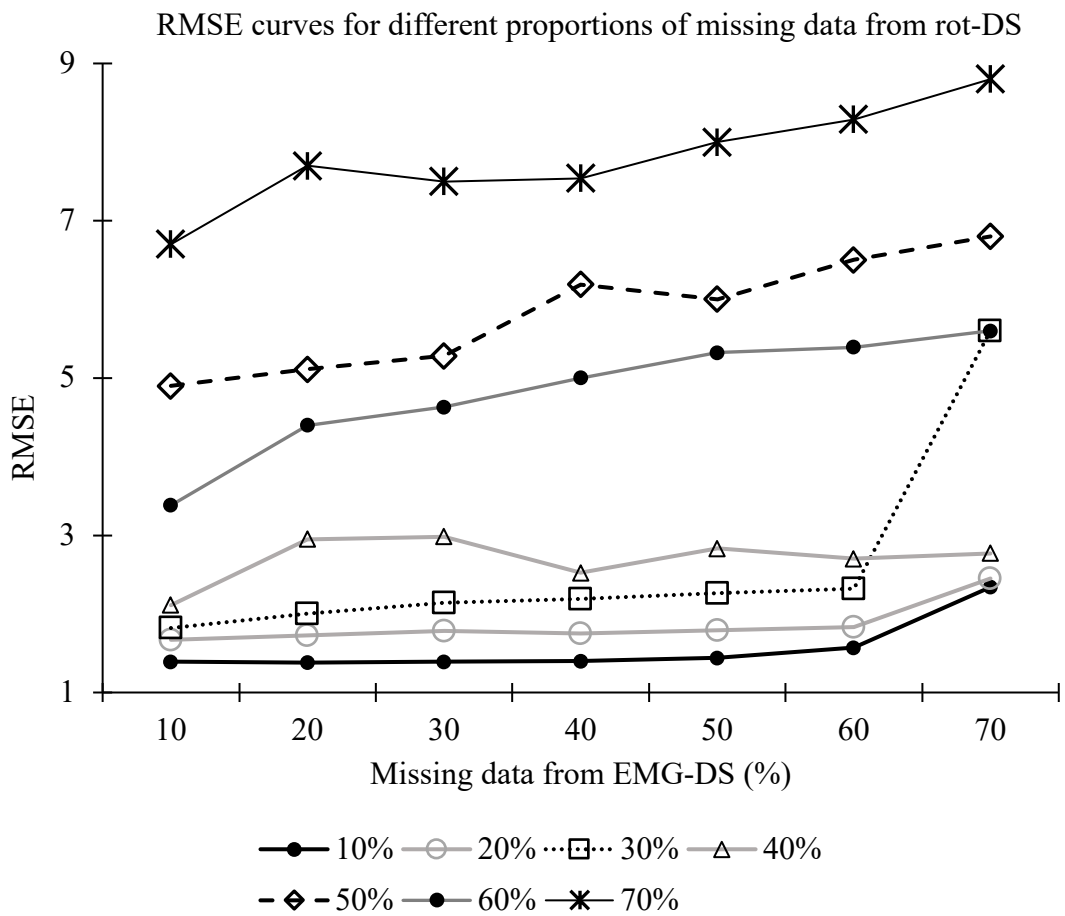


**Figure 5.7:** RMSE curves for 10%, 20%, 30%, 40%, 50%, 60% and 70% missing data from rotation dataset

**5.5.7.2.MAE curves for datasets with different proportions of missing values**

Figure 5.8 shows the MAE curves for different proportions of missing data from the knee rotation and EMG datasets. In Figure 5.8, rot-DS and EMG-DS means the rotation dataset and EMD dataset respectively.
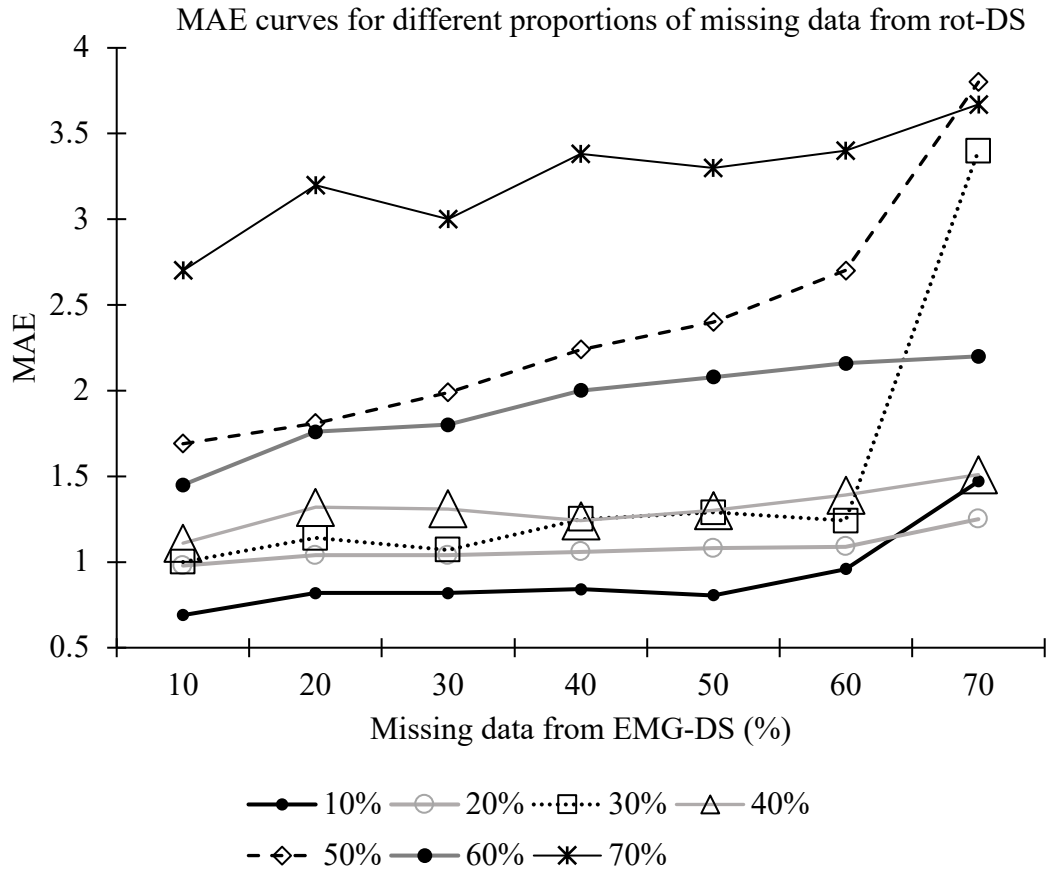


**Figure 5.8:** MAE curves for 10%, 20%, 30%, 40%, 50%, 60% and 70% missing data from rotation dataset

The MAE curves demonstrated similar patterns to the RMSE curves. In general, the MAE curves demonstrated an increasing trend with the increase of missing values in both datasets. It was observed that, for up to 40% missing values from rot-DS and up to 60% missing values from

EMG-DS, the variation of the MAE values was between 0.6 to 1.5. When missing values in EMG-DS were further increased (up to 70%), a sharp increase in MAE curve was observed when 30% data were missing in rot-DS. When the proportion of missing values in rot-DS was further increased to 50%, MAE values also tended to increase. Similar to the RMSE curves, for 50%, 60% and 70% missing data in rot-DS, the MAE curves shifted slightly upward, indicating that with the increase of missing values in rot-DS, MAE values also increased. In general, from 50%-70% missing values in rotation dataset and up to 70% missing values in EMG dataset, the values of MAE varied considerably from 1.5-4. These results indicated that with the increase in missing values in both datasets, the performance of the fusion method tended to decease as MAE values increased. The variation in MAE values was small for up to 40 % missing values in rot-DS and up to 60 % missing values in EMG-DS. However, with further increase of missing values in both datasets, the MAE values varied considerably.

### 5.5.7.3. Consistency curves for datasets with different proportions of missing values

Figure 5.9 shows the consistency curves for different proportions of missing data from the knee rotation and EMG datasets. In Figure 5.9, rot-DS and EMG-DS means the rotation dataset and EMD dataset respectively.

The consistency curves generally demonstrated a decreasing trend with the increase of missing values in both datasets. It was observed that for up to 70% missing values from rot-DS and up to 60% missing values from EMG-DS, the consistency values varied to a small extent from 87% to 70%. However, for 70% missing values in EMG-DS, the consistency value decreased considerably for several configurations, particularly for 30%, 50% and 70% missing values in rot-DS. These results indicated that with the increase in missing values in both datasets, the performance of the fusion method decreased in general.
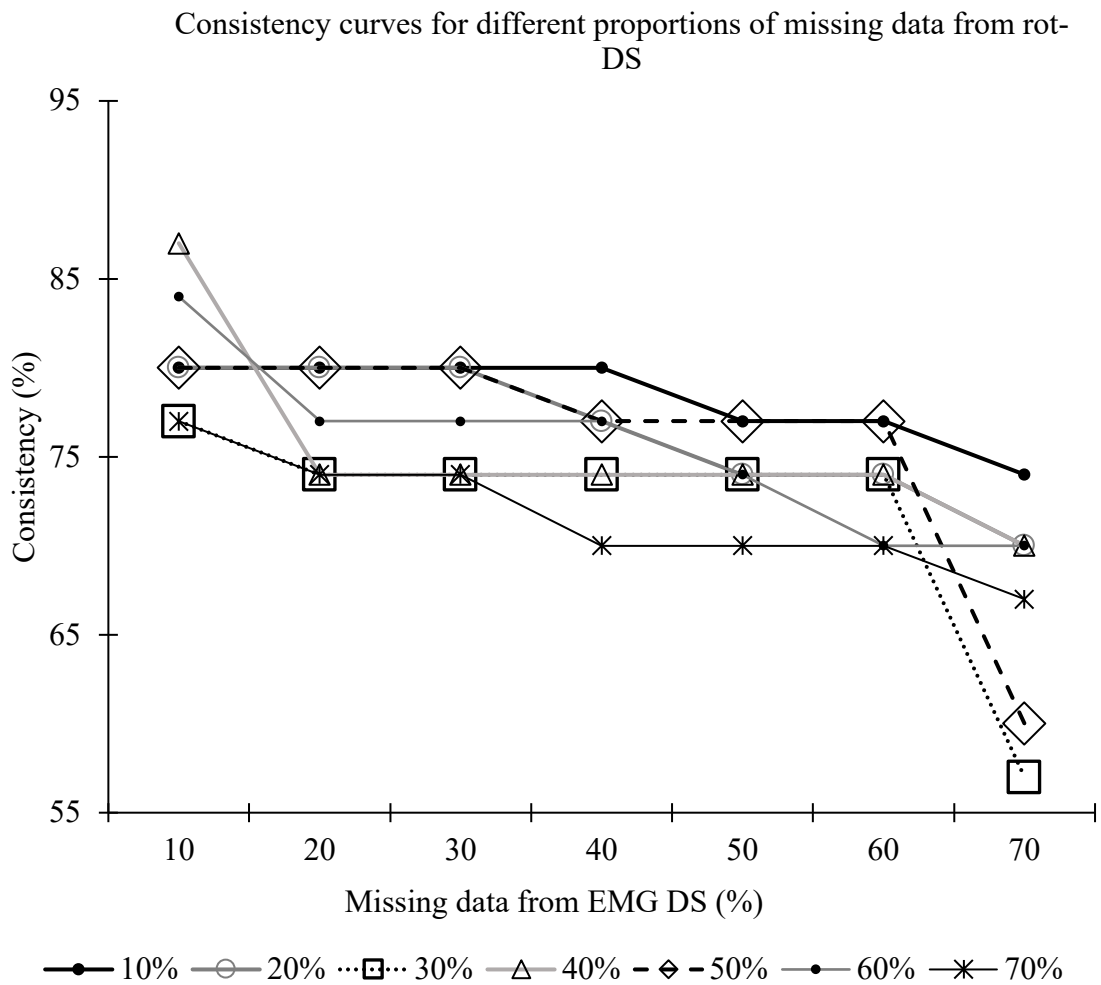
**Figure 5.9:** Consistency curves for 10%, 20%, 30%, 40%, 50%, 60% and 70% missing data from rotation DS

## 5.6.Discussion and Study Limitations

To ensure reliable risk assessment of work-related MSD, proper data collection is crucial. However, human-based data can potentially have missing data points due to dropout from the data collection technology. Moreover, risks sometimes cannot be fully quantified with a single risk indicator and thus multiple heterogeneous risk indicators are often collected for risk assessment.

As a result, a viable method is needed that reserves the interrelation among multiple risk indicators and potential risk factors during rectification of these multiple imperfect and incomplete datasets.

This research proposed the use of a CPD tensor decomposition-based data fusion method, which fuses multiple imperfect datasets in order to replace missing data. The proposed method generates a fused dataset by learning the latent structures and collaborative relationships among the potential risk factors and the risk indicators captured in multiple work settings. By learning their correlations, this proposed method replaces the missing values in the multiple incomplete datasets with minimum artifacts. The sample analysis results demonstrated that for up to a significant portion of missing values from both datasets, the proposed method can generate a fused dataset, by which reliable risk assessment results can be obtained consistent to those obtained from the original datasets to a great extent (87% to 70%). The possible reason for some inconsistencies in the risk assessment results may be attributed to the presence of significant amount of missing values in the individual datasets. In spite of this, for the specific datasets used in this study, the proposed method performed consistently up to 40% of missing data in the original datasets and could be reasonably applied to those datasets. The RMSE and MAE values demonstrated that the reconstruction errors in fusion were very small up to this range of missing data. This suggests that the reconstructed fused values of the risk indicators in the new dataset are close to those in the original, unaltered experimental datasets and hence, a good fusion performance. Although there is no standard threshold for good or bad RMSE and MAE values, low values of RMSE and MAE (within 10% of the range of the reference values) are preferable. RMSE should be less than 10% of the range of target property value (Alexander et al. 2015). In the current study, for different portions of missing values, the obtained RMSE values ranged from 0.74% to 5% relative to the difference between the minimum and maximum values of the risk indicators in the original dataset.

Additionally, for up to 40% missing values from the knee rotation dataset and 60% missing values from the EMG dataset, the RMSE values were within 15% of the average value of the risk indicators in the original dataset. Similarly, the MAE values were within 10% of the average value of the risk indicators in the original dataset. However, for the specific datasets used in this study, the RMSE and MAE values varied considerably as the proportion of missing data increased to 50% and 70% in the rotation and EMG datasets respectively. The possible reason might be the presence of significant amount of missing values in the individual datasets that drastically deteriorated the performance of the fusion method in reconstruction of the data and therefore negatively affected the RMSE and MAE values.

Apart from the previous three performance metrics, the performance of the present method can be further validated by the 'core consistency' values obtained when measuring number of rank-1 tensors $R$ (latent factor) for CPD. For different $R$ values presented in Table 5.5, the 'core consistency' values ranged from 88% - 98%, indicating that the CPD tensor decomposition is able to represent the variability of the multi-dimensional EMG and rotation data in the subspace substantially, thus providing a good fusion result. Overall, the findings suggest that the proposed method can handle up to 40% missing data and can reasonably be applied in work-related MSD risk assessment studies. The performance of this method was evaluated based on: a) how close a reconstructed fused dataset was to the original dataset, and b) how consistent the risk assessment results from a fused dataset were to those from the original dataset.

Although the proposed method has been proven successful in generating a fused dataset for used in risk assessment studies, there are still some limitations that are worth further investigation. First, in this study, only the maximum knee rotation angles and maximum normalized muscle activity measurements were considered for fusion. It would be interesting to

explore a detailed time series dataset instead of an aggregated metric such as maximum. Although data fusion is known to provide deeper and more accurate insights than those obtained from individual single data source (Khaleghi et al. 2013), another limitation of this work is the performance of the data fusion was only assessed by comparing the risk assessment results obtained from the fused datasets to the original dataset. The fused datasets were proven to be statistically similar to the original dataset, indicating that the fusion method was successful in reconstruction of a dataset when a significant amount of data is missing. Yet, how a fused dataset outperforms the individual incomplete datasets by which the fused dataset is constructed in terms of performance of risk assessment has not been measured.

## 5.7. Conclusion and Future Extension

The current research proposed a method that applies the CPD tensor decomposition technique to fuse multiple imperfect and incomplete datasets, as well as replacing missing data for assessing work-related MSD risks among construction workers. The proposed method helps not only in replacing missing values, but also holds the correlation among the potential risk factors and the risk indicators during replacement. The method was validated by comparing the risk assessment results obtained from the fused datasets to those from the original experimental dataset. RMSE and MAE values were also computed to compare the performance of the fusion method with the increase of missing values. The validation results suggest that the reconstructed fused datasets were statistically similar to the original experimental datasets and therefore able to provide consistent risk assessment results up to a significant amount of missing values. This method can be used as an alternative risk assessment method when data collected from experimental studies or real construction sites are incomplete due to missing data points. The findings will help enable accurate assessment of work-related risk factors for work-related MSDs among construction

workers. Although the proposed method only used two sparse datasets for fusion, it can be extended to fuse more datasets. In future, this method will be explored in fusing imperfect and incomplete multiple time series datasets. The performance of the fusion will also be further tested by comparing consistencies of the risk assessment results to those obtained from datasets with missing values.

# CHAPTER 6: APPLICATION OF DATA FUSION VIA CANONICAL POLYADIC DECOMPOSITION IN RISK ASSESSMENT OF MUSCULOSKELETAL DISORDERS IN CONSTRUCTION: PROCEDURE AND STABILITY EVALUATION

**Abstract**

Missing data is a common problem in data collection for work-related musculoskeletal disorder (MSDs) risk assessment studies. It can cause incompleteness of risk indicators leading to erroneous conclusion on potential risk factors. Previous study suggested that data fusion is potentially a way to solve this issue. This research evaluated the effectiveness of a data fusion technique using canonical polyadic decomposition (CPD) for work-related MSD risk assessment in construction. Two knee MSD risk-related datasets—3D knee rotation (kinematics) and electromyography (EMG) of five knee postural muscles—collected from previous studies were fused for the evaluation. By comparing the consistency performance with and without data fusion, it revealed that, for all low to high proportion of missing data (10%-70%) from both kinematics and EMG datasets, the work-related MSD risk assessment using fused datasets outperformed using unfused kinematics datasets. For large proportion of missing data (>50%) from both kinematics and EMG datasets, better performance was observed by using fused datasets in comparison with unfused EMG datasets. These findings suggested that data fusion using CPD generates a more reliable risk assessment and therefore is an effective approach for remedying missing data in work-related MSD risk evaluation.

### 6.1.Introduction

Missing data is a common problem associated with data collection in work-related musculoskeletal disorder (MSD) risk assessment studies (Young et al. 2011). Data collected for assessing risk, either from the laboratory or construction site, often suffer from irregular spatial and temporal resolution. Among multiple trials of experiments, data of a specific trial can be missing due to an instrument failure, human-induced error or at random. Therefore, missing data can occur in almost all research that involves data collection even the study is well-designed and conducted in a controlled environment (Brown and Kros 2003). Missing data can cause incompleteness and misrepresentation of the risk indicators, leading to less statistical power, biased and invalid conclusions on the effects of the potential risk factors on the assessed work-related MSD risk. Data fusion has been found to be potential in dealing with incomplete datasets used in work-related MSD risk assessment by replacing the missing data reserving the interrelation among different factors and the risk indicators across multiple datasets. In the authors' previous study, a data fusion method was created to treat imperfect MSD risk assessment data using a canonical polyadic decomposition (CPD) tensor decomposition (Dutta et al. 2020c). However, the numerical stability of the CPD fusion has not been evaluated comprehensively. When MSD risks need to be characterized and quantified by multiple risk indicators, how these risk indicators will interact with each other for risk assessment if some of their proportions become missing is still unknown. Also, whether the CPD fusion is numerically stable to reconstruct a new dataset by fusing those risk indicators and imputing the missing values to provide more accurate and reliable risk assessment results compared to the unfused sparse datasets has not been studied. This research evaluated the numerical stability of the data fusion via CPD for MSD risk assessment in construction when the quality of risk-related datasets becomes degraded due to the missing data in

multiple inputs. To this end, this research compared the risk assessment results resulting from the fused datasets to those resulting from multiple unfused sparse datasets. CPD was applied to fuse the sparse datasets with different portions of subsets intentionally removed from two real work-related MSD risk-related datasets, referred to as "original experimental datasets" that have been collected from previous experimental studies. The risk assessment results obtained from the fused and the unfused sparse datasets were then compared to those obtained from the original experimental datasets to elucidate the performance of the present technique.

## 6.2. Background

### 6.2.1. Problems with missing data

The problem associated with missing data is relatively common in almost all research. Studies that involve human subjects and technologies for data collection can have missing data for various reasons including technology-induced errors (e.g., disconnection of sensors, errors in communicating with the database server, instrument failures), human-induced errors (e.g., accidental human omission) or other unknown reasons (Data 2016). Missing data can have significant effects on the conclusions that can be drawn from the data. The absence of data reduces statistical power which refers to the possibility that the analysis results will reject the null hypothesis when it is false (Kang 2013). In work-related MSD risk studies, missing data can cause incompleteness and misrepresentation of the risk indicators, leading to inconsistent, biased estimation of the effects of the potential risk factors on the risk indicators which can lead to erroneous conclusion. For this reason, missing data should be dealt with carefully to ensure reliable risk assessment.

### 6.2.2. Research works in addressing missing data problems

There is a diversity of methods to handling missing data in existing literature. Some common methods include: simply omitting observations with the missing data (Kang 2013), regression imputation (Button et al. 2013), mean substitution (Malhotra 1987), replacing the missing data with the last observed values (Hamer and Simpson 2009), and estimating the missing data using conditional distribution of the other variables (Gelman and Raghunathan 2001). These methods have been proven useful in handling missing data in many research works. However, they have some limitations for application to impute missing data in work-related MSD risk assessment studies. In methods of omitting observations, that the observations with missing data are simply removed decreases the sample size and adversely affects the statistical power. In regression imputation methods, a missing value is predicted from other variables, but it does not add any new information except increasing the sample size and compromising the standard error. Mean substitution methods may cause inconsistent biases if there is a great inequality in the proportion of missing values of different variables. Methods by replacing missing data with the last observed values assume that there will be no changes in the outcome. For a large proportion of missing data, methods using conditional distribution of the other variables may provide inconsistent bias as the relationship among the variables needs to be properly computed. Moreover, the "high-dimensionality" or "multimodality" nature of work-related MSD risk-related data where risk exposures can be obtained in form of different risk indicators collected at different working settings (often referred to as risk factors) and from different subjects limit the applicability of the above-mentioned methods in missing values imputation in that the latent relationship among the risk indicators and the factors cannot be well captured (Gebregziabher and DeSantis 2010). Quite often, multiple risk indicators data are collected for in-depth understanding of risks and workers'

risk-taking behaviors. Data fusion can be an efficient method in this regard as by this method risk-related information are integrated from multiple datasets based on the latent relationship among the risk factors and risk indicators such that a more complete representation of the measurements of the risk indicators can be produced, as compared to what can be represented in incomplete and unfused datasets (Lahat et al. 2015).

### 6.2.3. Tensor decomposition-based data fusion

Tensor decomposition can accurately extract the interrelations among various features from multiple datasets and learn the latent structures and collaborative relationships among those features to approximate the pattern of the data during fusion. Therefore, this can be a useful tool for fusion of high-dimensional datasets (Kolda and Bader 2009). Tensors are generalizations of matrices to higher dimensions and are a very powerful tool to model multidimensional data (De Lathauwer et al. 2000). Tensor decomposition was previously used for anomaly detection (Xie et al. 2017), object profiling (Charlier et al. 2018), pattern recognition (Xiong et al. 2010), filling in missing values (Dauwels et al. 2012), and evolution prediction (Dunlavy et al. 2011), and posture recognition (Chen et al. 2017).

Canonical polyadic decomposition (CPD) is one of the most widely used tensor decomposition techniques, which is generally suggested to estimate latent parameters among different features within a dataset (Rabanser et al. 2017). CPD can effectively be used for imputing missing data as it captures the interactions among various features of a high-dimensional datasets during imputation (Dauwels et al. 2012). CPD has also been found to be effective in fusing multiple incomplete datasets and filling out missing data in signal processing and machine learning (Sidiropoulos et al. 2017).

CPD factorizes a tensor into a sum of component rank-one tensors (Carroll and Chang 1970; Hitchcock 1927). For example, in CPD, a given three-dimensional tensor $X \in \mathbb{R}^{I \times J \times K}$ can be written as:

$$X \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r \equiv [\![A, B, C]\!]$$ ................................................................ (6.1)

Where $R$ is a positive integer, which is referred to as the rank of a tensor. The rank of tensor $X = R$ can be defined as the smallest number of rank-1 tensor that is required to represent the tensor $X$ as their sum (Rabanser et al. 2017); Rank-1 tensor is defined as a decomposition of an $M$-dimensional tensor into one outer product of $M$ vectors. $R$ is also called latent factor. $I$, $J$ and $K$ are the sizes of the dimensions of the tensor $X$ where, $a_r \in \mathbb{R}^I$, $b_r \in \mathbb{R}^J$ and $c_r \in \mathbb{R}^K$ for $r = 1...R$. '∘' indicates the vector outer product. $a_r \circ b_r \circ c_r \in \mathbb{R}^{I \times J \times K}$ is an outer product of three vectors $a_r$, $b_r$ and $c_r$ and is referred to as a rank-1 tensor. $A$, $B$ and $C$ are the factor matrices obtained for each dimension after decomposition of the tensor $X$, and are referred to as the combination of the vectors from the rank-1 components (Kolda and Bader 2009). Here, the factors matrices $A \in \mathbb{R}^{I \times R}$, $B \in \mathbb{R}^{J \times R}$ and $C \in \mathbb{R}^{K \times R}$.

More details on CPD were previously discussed in section 5.2.3.1 in Chapter 5.

### 6.2.4. Applications of data fusion in construction

In construction, data fusion has been used to improve the quality of building cooling load measurement of building automation systems and for efficient and effective management knowledge discovery (Huang et al. 2009; Soibelman et al. 2004). Some of other applications of data fusion in construction included automated identification, location estimation, dislocation detection of construction materials in jobsites (Razavi and Haas 2010; Razavi and Haas 2011), automated progress tracking of construction projects (Shahi et al. 2012), structural health

monitoring (Soman et al. 2018), damage identification of civil structures (Anaissi et al. 2018) and constriction productivity monitoring (Pradhan et al. 2011). In construction ergonomics, a computationally efficient approach using tensor decomposition was developed to recognize construction workers' awkward postures (Chen et al. 2017). Fusion of data from continuous remote monitoring of construction workers' location and physiological status was used to identify safe and unsafe behaviors of construction workers (Cheng et al. 2012). Fusion of spatiotemporal and workers' thoracic posture data has been done for understanding the worker's activity type for productivity assessment (Cheng et al. 2013). A position and posture data fusion method was proposed for evaluation of construction workers' behavioral risks (Chen et al. 2019). A framework for location data fusion and pose estimation of excavators have been developed by Soltani et al. (2018). Dutta et al. (2020c) proposed a method of fusing multiple sparse risk-related datasets that applies CPD to fuse and fill in missing data by leveraging the correlation among multiple risk indicators within those datasets.

## 6.3. Problem Statement and Research Objective

Missing data can lead to misrepresentation of the risk indicators and underpowered analyses which can provide erroneous conclusions on the contribution of potential risk factors on the imposed work-related MSD risk. Representation of high dimensional MSD risk-related data as a multimodal tensor and use of the tensor for fusion of those data available in multiple risk-related datasets can fill in the missing values by reserving the interrelationship among different risk indicators and the risk factors across the datasets and hence, ensure complete representation of risk-related data and accurate risk assessment. In the authors' prior work, CPD has been found to be promising in fusing multiple incomplete high-dimensional risk-related datasets and filling out missing data within those datasets based on the latent correlations among the risk factors and risk

indicators (Dutta et al. 2020c). Yet, when applying the technique of data fusion with CPD, one additional question has not been answered. That is, given a different portion of missing data in two sparse risk-related datasets, it is still unknown whether, at what portion of missing data, and to what extent, the fusion-based risk assessment outperforms the one with no fusion. This knowledge is important to understand the effect of such tensor decomposition-based data fusion to enhance risk assessment when data collected from experimental studies or real construction sites are incomplete due to missing data points. Therefore, the objective of this research was to evaluate the usefulness of applying CPD for work-related MSD risk assessment in construction by comparing the consistencies of the risk assessment results obtained from fused and unfused datasets to those obtained from the original experimental datasets.

**6.4. Experimental Design and Implementation**

The experimental design included the procedures depicted in Figure 6.1. First, two risk-related datasets were collected from authors' previous human subject laboratory experimental studies. These datasets were considered as original experimental datasets. One experimental dataset (DS1) contained calculated knee rotation (kinematics) data, representing five knee rotational angles – flexion, abduction, adduction, internal and external rotation collected at different roof slopes and working postures. The second experimental dataset (DS2) contained EMG data, describing muscle activation of five knee postural muscles – biceps femoris, rectus femoris, semitendinosus, vastus lateralis and vastus medialis collected at the same roof slopes and working postures. As awkward knee rotations and heightened muscle activation have been identified as two quantitative risk indicators of work-related MSDs in lower extremities and the roof slope and the working posture are two work-related MSD risk factors, these datasets were representatives of work-related MSD risk-related datasets and therefore used in this study (Breloff

et al. 2019a; Dutta et al. 2020b; Kingston et al. 2016). Then, different proportions of missing values were intentionally removed from both experimental datasets to construct two sparse unfused datasets (SDS1 and SDS2). After that, both original experimental datasets and sparse unfused datasets were constructed to high dimensional tensors. The tensors constructed from the original experimental datasets and sparse unfused datasets were referred as "original tensor" and "sparse unfused tensor" respectively in this study. Next, CPD tensor decomposition was applied to the sparse unfused tensor for fusion of the knee kinematics and EMG data and filling out the missing data within those datasets. CPD first decomposed the sparse unfused tensor into factor matrices based on the latent relationship among different risk factors and the risk indicators. The factor matrices were then used to reconstruct the kinematics and EMG data through fusion as a high-order tensor. This newly reconstructed tensor was referred to as "fused tensor". The data represented in the fused tensor and the sparse unfused tensor were subsequently analyzed to produce the risk assessment results. Finally, the risk assessment results obtained from the fused and sparse unfused tensors were compared to those obtained from the original tensor to evaluate whether the fused datasets outperformed the sparse unfused datasets in risk assessment.
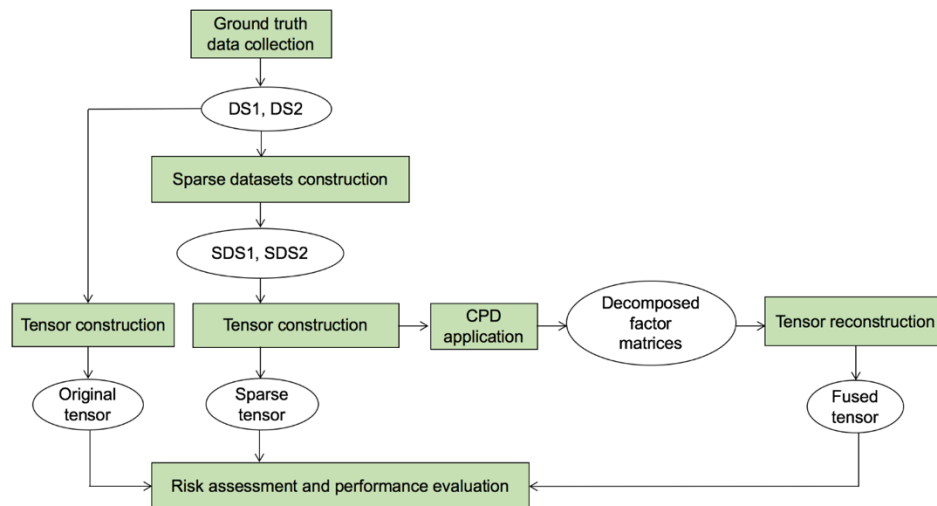


**Figure 6.1:** The procedures of the experimental design

130

### 6.4.1. Risk-related original experimental data collection

The knee rotation (kinematics) and the EMG datasets collected from the authors' prior human subject laboratory experimental studies described in Chapters 2 and 3 were used as risk-related data (original experimental datasets). The instrumentation, experimental set up, data collection procedures and formats of the datasets are discussed in Chapters 2 and 3 in more details.

### 6.4.2. Multiple sparse datasets construction (Data removal)

A proportion of data was intentionally removed from both knee kinematics and EMG datasets collected from the previous studies to represent them as sparse datasets. The data was removed at random. The proportions of the missing data from both datasets were provided in Table 6.1. This way, fourteen sparse datasets, seven for knee kinematics and EMG data each, were constructed with different portions of missing values.

**Table 6.1:** Proportion of missing data

| Datasets | Knee kinematics | knee EMG |
|---|---|---|
| | 10 | 10 |
| | 20 | 20 |
| | 30 | 30 |
| **Missing proportion (%)** | 40 | 40 |
| | 50 | 50 |
| | 60 | 60 |
| | 70 | 70 |

### 6.4.3. Fusion of the datasets using CPD tensor decomposition

The resulting kinematics and EMG datasets were then used for fusion. As the knee kinematics and EMG data were collected from the same subjects at the same work settings (i.e., roof slope, working posture, and trial), they were first combined by the early integration data fusion method (Žitnik and Zupan 2014). In this method, the column vectors of these two types of risk indicators were concatenated to form a single set of length ten vectors which was referred as "risk

indicators" in this study. For tensor construction, a sparse rotation dataset and a sparse EMG dataset were selected according to the combinations of missing proportions specified in Table 6.2.

**Table 6.2:** Datasets with missing values for tensor construction

| Missing proportion from kinematics data (%) | 10 | 20 | 30 | 40 | 50 | 60 | 70 |
|---|---|---|---|---|---|---|---|
| **Missing proportion from EMG data (%)** | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| | 40 | 40 | 40 | 40 | 40 | 40 | 40 |
| | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| | 60 | 60 | 60 | 60 | 60 | 60 | 60 |
| | 70 | 70 | 70 | 70 | 70 | 70 | 70 |

As the risk indicators were measured from different subjects at different working postures, roof slopes and multiple trials, the data were then represented as a five-dimensional sparse tensor $T \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$, where, $I_1$ = No. of subjects (total 9), $I_2$ = No. of risk indicators [kinematics and EMG: total 10 (5 knee rotation angles and 5 knee muscle EMG)], $I_3$ = No. of trials (total 5), $I_4$ = No. of postures [total 2 (static and dynamic)], $I_5$ = No. of slopes [total 3 (0°, 15°, 30°)]. In short, the tensor $T$ maps the knee rotation and EMG measurements of nine subjects obtained from five trials performed on three roof slopes using two postures. Tensor $T$ was referred to as sparse unfused tensor in this study.

In the similar way, the original datasets were also constructed to a five-dimensional tensor $Y \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$ as ground truth, which was referred to as original tensor in this study and was used later for validation of the fusion method.

Once a tensor $T$ was constructed, CPD tensor decomposition was applied to decompose the tensor $T$ into five factor metrics, representing information in each mode or dimension. The factor matrices were represented as matrices $A, B, C, D$ and $E$, providing latent information in modes

of subjects, risk indicators, trials, postures, and slopes respectively, where factor metrics $A \in \mathbb{R}^{I_1 \times R}$, $B \in \mathbb{R}^{I_2 \times R}$, $C \in \mathbb{R}^{I_3 \times R}$, $D \in \mathbb{R}^{I_4 \times R}$, $E \in \mathbb{R}^{I_5 \times R}$. Each factor matrix represented the latent relationship between the corresponding dimension and all the other dimensions captured by latent factor $R$. A proper $R$ values ensures a unique mapping of the original tensor to the decomposed tensor. In this study, $R$ value was computed applying the core consistency diagnostic method (CORCONDIA) method described in Bro and Kiers (2003).

Factor matrices $A$, $B$, $C$, $D$ and $E$ were then computed by solving the following optimization equation:

$$\arg \min \| T - \sum_R A_r \circ B_r \circ C_r \circ D_r \circ E_r \| \quad \text{..................................................................... (6.2)}$$

To solve Equation (2), the alternating least squares algorithm was applied using an optimization-based program *cpd_als* in Tensor Toolbox from MATLAB (Bader and Kolda 2019).

The reconstructed tensor $T' \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$ was then computed using the following equation:

$$T' = \sum_{r=1}^{R} A_r \circ B_r \circ C_r \circ D_r \circ E_r \text{.................................................................... (6.3)}$$

This reconstructed tensor was an approximate mapping of the sparse tensor $T$ which was referred to as fused reconstructed tensor and represented as $T' \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4 \times I_5}$. In this tensor, CPD imputed missing values and reconstructed the risk indicator values minimizing the reconstruction error. For this reconstruction, Tensorlab (a MATLAB package for tensor computation) was used (Debals et al. 2019). The reconstructed risk indicators were then extracted to a spread sheet to generate the risk assessment results.

### 6.4.4. Performance evaluation of data fusion

To examine the performance of fused and unfused sparse datasets in risk assessment, the following four consistency measures were computed as performance metrics: a) consistency of the effects of roof slope and working posture on five knee rotation angles between the original and fused datasets ($Con_{Rot\_f}$), b) consistency of the effects of roof slope and working posture on five knee postural muscles' EMG between the original and fused datasets ($Con_{EMG\_f}$), c) consistency of the effects of roof slope and working posture on five knee rotation angles between the original and sparse unfused datasets ($Con_{Rot\_s}$), d) consistency of the effects of roof slope and working posture on five knee postural muscles' EMG between the original and sparse unfused datasets ($Con_{Rot\_s}$).

$Con_{Rot\_f}$ was computed using Equation (6.4):

$$Con_{Rot\_f} = \frac{\sum_{i=1}^{P} f(m_i, m\prime_i)}{P} \times 100 \text{...........................} (6.4)$$

Where, $P$ is the total number of possible effects of roof slope, working posture and their interactions on five knee rotation angles [total 15 = 3 factors (slope, posture, slope-posture interaction) ×5 knee rotation angles]. $m_i$ denotes the effects of the factors (slope and posture) on the response variables (five knee rotation angles) obtained from the original experimental dataset; $m\prime_i$ denotes the effects of the factors (slope and posture) on the response variables (five knee rotation angles) obtained from the fused dataset; and $f(m_i, m\prime_i) = \begin{pmatrix} 1, & if \ m_i = m_i' \\ 0, & if \ m_i \neq m_i' \end{pmatrix}$.

$Con_{EMG\_f}$ was computed using Equation (6.5):

$$Con_{EMG\_f} = \frac{\sum_{i=1}^{Q} f(n_i, n\prime_i)}{Q} \times 100 \text{..........................} (6.5)$$

Where, $Q$ is the total number of possible effects of roof slope, working posture and their interactions on EMG values of five knee postural muscles [total $15 = 3$ factors (slope, posture, slope-posture interaction $\times 5$ knee postural muscles]. $n_i$ denotes the effects of the factors (slope and posture) on the response variables (EMG values of five knee postural muscles) obtained from the original experimental dataset; $n'_i$ denotes the effects of the factors (slope and posture) on the response variables (EMG values of five knee postural muscles) obtained from the fused dataset;

and $f(n_i, n'_i) = \begin{pmatrix} 1, & if \ n_i = n'_i \\ 0, & if \ n_i \neq n'_i \end{pmatrix}$.

$Con_{Rot\_s}$ was computed using Equation (6.6):

$$Con_{Rot\_s} = \frac{\sum_{i=1}^{P} f(m_i, k_i)}{P} \times 100 \ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (6.6)$$

Where, $P$ is the total number of possible effects of roof slope, working posture and their interactions on five knee rotation angles [total $15 = 3$ factors (slope, posture, slope-posture interaction) $\times 5$ knee rotation angles]. $m_i$ denotes the effects of the factors (slope and posture) on the response variables (five knee rotation angles) obtained from the original dataset; $k_i$ denotes the effects of the factors (slope and posture) on the response variables (five knee rotation angles) obtained from the sparse unfused dataset; and $f(m_i, k_i) = \begin{pmatrix} 1, & if \ m_i = k_i \\ 0, & if \ m_i \neq k_i \end{pmatrix}$.

$Con_{EMG\_s}$ was computed using Equation (6.7):

$$Con_{EMG\_s} = \frac{\sum_{i=1}^{Q} f(n_i, l_i)}{Q} \times 100 \ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. (6.7)$$

Where, $Q$ is the total number of possible effects of roof slope, working posture and their interactions on EMG values of five knee postural muscles [total $15 = 3$ factors (slope, posture, slope-posture interaction $\times 5$ knee postural muscles]. $n_i$ denotes the effects of the factors (slope

and posture) on the response variables (EMG values of five knee postural muscles) obtained from the original experimental dataset. $l_i$ denotes the effects of the factors (slope and posture) on the response variables (EMG values of five knee postural muscles) obtained from the sparse unfused dataset; and $f(n_i, l_i) = \begin{pmatrix} 1, & if\ n_i = l_i \\ 0, & if\ n_i \neq l_i \end{pmatrix}$.

## 6.5. Results

### 6.5.1. Consistency curves of rotation for different proportions of missing values

Figure 6.2 shows the consistency curves of rotation for different proportions of missing values from both datasets. Each plot shows two consistency curves. One (continuous line) is for the fused rotation values which were in the fused dataset generated by the fusion of rotation and EMG datasets for different proportion of missing values. The other (dash line) is for the unfused rotation values, i.e., the rotation values that were initially in the datasets generated by removing different proportions of data from the kinematics dataset before fusion (sparse unfused datasets described in Table 6.2). In general, for different proportions of missing data in both datasets, values in the fused datasets provided more consistent risk assessment results than those in the datasets with missing values, although for some proportions of missing values (e.g., 10% and 20% missing data from the rotation dataset with 60% and 70% missing data from the EMG dataset, 30% missing data from the rotation dataset with 30% to 50% missing data from the EMG dataset), both unfused and fused-based assessment results demonstrated the similar consistencies. These findings suggested better performance in risk assessment by the fused datasets that might provide more consistent risk assessment results than the unfused datasets with missing values.

(a) Consistency curve of rotation for 10% missing values from rotation dataset
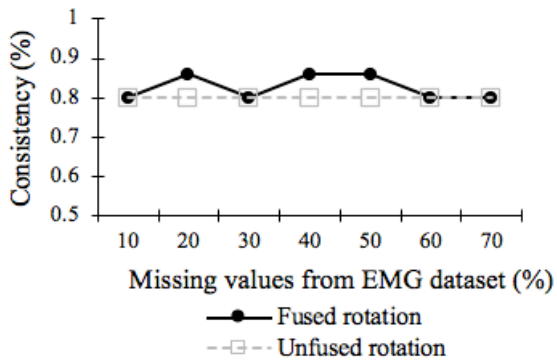
(b) Consistency curve of rotation for 20% missing values from rotation dataset

(c) Consistency curve of rotation for 30% missing values from rotation dataset

(d) Consistency curve of rotation for 40% missing values from rotation dataset

(e) Consistency curve of rotation for 50% missing values from rotation dataset

(f) Consistency curve of rotation for 60% missing values from rotation dataset

(g) Consistency curve of rotation for 70% missing values from rotation dataset

**Figure 6.2:** Consistency curves of rotation for different proportions of missing values from rotation (kinematics) dataset: (a) 10%; (b) 20%; (c) 30%; (d) 40%; (e) 50%; (f) 60%; (g) 70%.

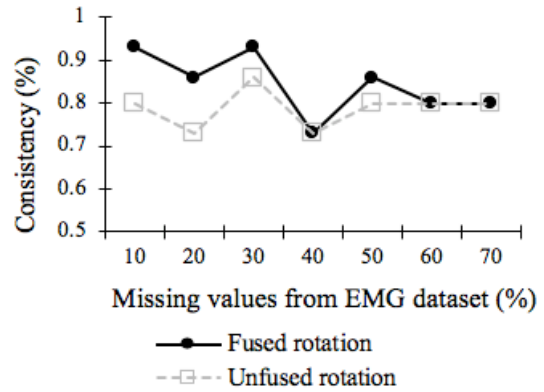### 6.5.2. Consistency curves of EMG for different proportions of missing values

Figure 6.3 shows the consistency curves of EMG for different proportions of missing values from both datasets. Similarly, each plot shows two consistency curves. One (continuous line) is for the fused EMG values which were in the fused dataset generated by fusion of rotation and EMG datasets for different proportions of missing values. The other (dash line) is for the unfused EMG values, i.e., the EMG values that were initially in the datasets generated by removing different proportion of data from the EMG dataset before fusion (sparse unfused datasets described in Table 6.2). For 10% to 40% missing values in rotation datasets and 10% to 60% missing values in EMG datasets, the performance of the fusion method was poor where the unfused EMG values in the datasets with missing values provided better consistency results than the fused EMG values in fused datasets. However, with the increase of the missing data in rotation datasets (from 50% to 70%) and for a large percentage of missing values in EMG datasets (50% - 70%), the fusion method demonstrated better performance. In general, for EMG data, the fused datasets demonstrated higher consistencies when large proportion of data (>50%) were missing in both

138

datasets. These findings indicated that the fused datasets performed better in risk assessment than

the unfused sparse datasets even when a large proportion of data were missing.



(a) Consistency curve of EMG for 10% missing values from rotation dataset

(b) Consistency curve of EMG for 20% missing values from rotation dataset

(c) Consistency curve of EMG for 30% missing values from rotation dataset

(d) Consistency curve of EMG for 40% missing values from rotation dataset
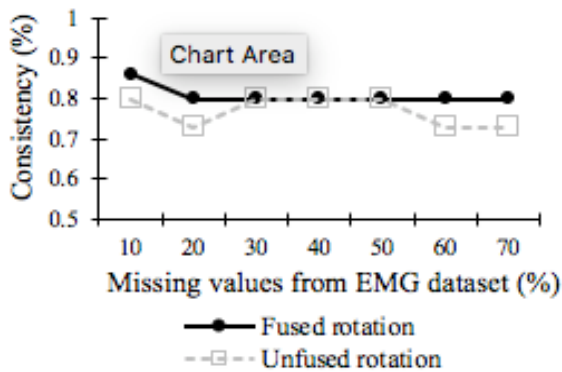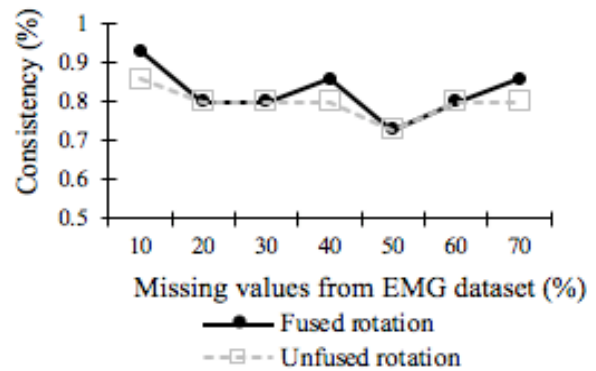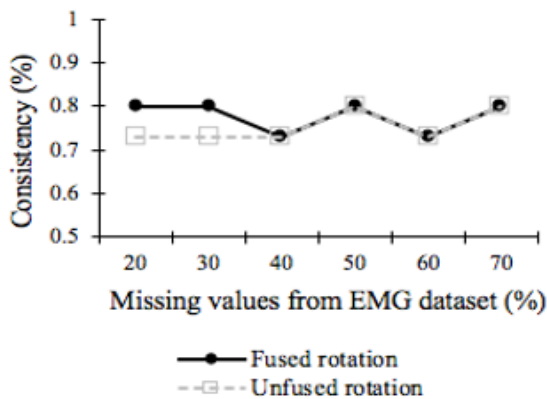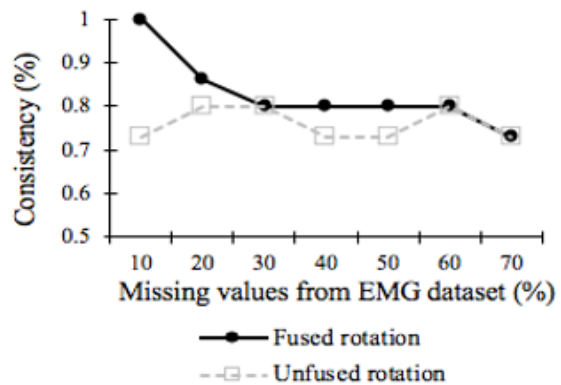
(e) Consistency curve of EMG for 50% missing values from rotation dataset

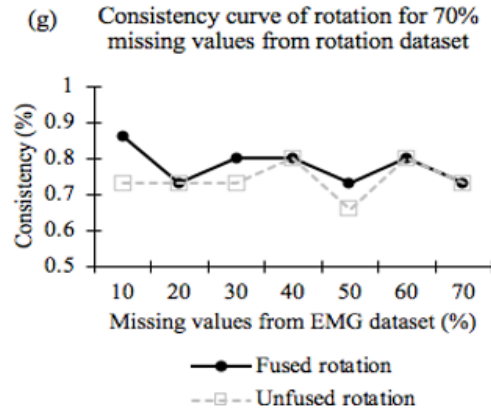(f) Cocnsistency curve of EMG for 60% missing values from rotation dataset

**Figure 6.3:** Consistency curves of EMG for different proportions of missing values from rotation (kinematics) dataset: (a) 10%; (b) 20%; (c) 30%; (d) 40%; (e) 50%; (f) 60%; (g) 70%

## 6.6. Discussion

Data collection for work-related MSD risk assessment either from laboratory experiments or from construction sites can potentially have missing data due to dropout – a phenomenon when data get missing due to any technology or human induced error. Quite often, risks cannot be fully quantified with a single risk indicator and thus multiple heterogeneous risk indicators data are collected for risk assessment. Our previous study indicated that data fusion by applying CPD can be efficient to fill in missing data in such datasets as the latent relationship among different risk factors and risk indicators available in those datasets are reserved. This research evaluated the numerical stability of data fusion in work-related MSD risk assessment that applies CPD to fuse multiple risk-related datasets and fill in missing data. More specifically, in spite of using malformed experimental data with missing value as input, if and to what extent fusion through CPD was numerically stable to provide more reliable risk assessment result compared to unfused sparse datasets was analyzed in this study. For that purpose, two risk-related datasets containing roofers' knee kinematics and EMG data collected from the authors' previous experimental studies

140

were used for fusion. The effects of roof slope and working posture on five knee rotation angles and five knee postural muscles' EMG were measured as risk assessment results. Different proportions of data were intentionally removed from those datasets to make them sparse. CPD was thus applied to fuse those sparse datasets. Risk assessment results were computed from the original, sparse, and fused datasets. The performance of the fusion was evaluated based on how consistent the risk assessment results obtained from the fused and unfused sparse datasets were to those obtained from the original experimental dataset. For this purpose, four consistency measures – consistency of the effects of roof slope and working posture on five knee rotation angles between the original and fused datasets, consistency of the effects of roof slope and working posture on five knee postural muscles' EMG between the original and fused datasets, consistency of the effects of roof slope and working posture on five knee rotation angles between the original and sparse datasets and consistency of the effects of roof slope and working posture on five knee postural muscles' EMG between the original and sparse datasets were computed as performance metrics.

The consistency results demonstrated that for significant portions of missing values from both datasets, the fused datasets could provide more reliable risk assessment results than the sparse unfused datasets. The risk assessment results obtained from the fused datasets were more consistent to those obtained from the original experimental datasets when compared to unfused datasets with missing values. For example, for 10%-70% missing data in both kinematics and EMG datasets, more consistent risk effects on the knee rotations were obtained when rotation and EMG datasets were fused rather than from the unfused sparse rotation dataset. The better consistency by the fused datasets is an indicator of good fusion performance of the CPD tensor decomposition-based data fusion. Previous study done by the authors also found that fusion

through CPD could provide reliable risk assessment results 87% to 70% consistent to those provided by the original experimental datasets for up to 40% missing data (Dutta et al. 2020c).

For 10%-40% missing data from kinematics dataset and 10%-60% missing data from EMG dataset, the fused datasets did not outperform the unfused sparse datasets when the risk effects were analyzed on muscles' EMG. However, with further increase of the missing data from 50% to 70% in both datasets, the fused dataset started showing better consistency than the unfused sparse EMG dataset, although the authors' previous study showed that, for more than 50% missing data, the performance of the fusion method started to decrease (Dutta et al. 2020c). Despite the degradation of the fusion performance, in this study the fused datasets performed better compared to unfused sparse datasets with 50% to 70% of missing data. The possible reason can be attributed to the significant amount of missing data in the unfused datasets which could not capture the effects of the risk factors on the imposed risk, but due to comparatively small reconstruction errors in the fusion, the fused dataset performed better in risk assessment. Overall, for the specific datasets used in this study, the fusion-based risk assessment outperformed the one with no fusion, even when a large proportion of data (>50%) were missing. The findings suggested that, applying data fusion through CPD can be a better alternative and effective approach for work-related MSD risk assessment in construction rather than using datasets with missing values when the data collection is badly affected due to the presence of missing data.

Nevertheless, there are some limitations in this study. First, ground truth data were collected from an experimental study performed in a controlled laboratory setting, not from real construction site. Second, only an aggregated metric, maximum knee rotation and normalized muscle activity measurement were considered for fusion, detailed time series features were not analyzed. The synchronization of multiple time series datasets with different time stamps and

sampling rates and the corresponding data processing during fusion are worthwhile for further investigation.

## 6.7. Conclusion and Future Extension

This research evaluated the numerical stability of fusion through canonical polyadic decomposition (CPD) for work-related MSD risk assessment in construction when the data collected suffer from missing data. The findings of this study suggested that CPD was numerically stable to generate a fused dataset based on the interrelations among the risk factors and the risk indicators that outperformed the unfused datasets with missing values in work-related MSD risk assessment. This signifies the effectiveness of using data fusion through CPD as an alternative approach of filling out missing data in work-related MSD risk assessment studies. The fused dataset was found to be more consistent to the original experimental dataset than the unfused sparse dataset in terms of generating risk assessment results. Even for a large proportion (>50%) of missing data, the fused dataset provided more reliable risk assessment results than the unfused sparse dataset. These findings ensure that data fusion through CPD can help enable accurate and reliable assessment of work-related risk factors for work-related MSDs among construction workers when the data collection is affected by missing data. In the future, the method of fusion of multiple MSD risk-related datasets by representing those datasets as high-dimensional tensor and then decomposing them to reconstruct a fused dataset can be expanded to explore similar studies related to other body parts, such as low back, feet and shoulder injuries of construction workers including roofers. Moreover, the experiment will be conducted in a real construction site involving the participation of professional roofers and data will be collected onsite using an inertial measurement units (IMU) based motion capture system. This will enable to explore the performance of the proposed fusion method to handle possible challenges involved in data

collection in a real construction site including missing data, noise, occlusion and sunlight issue for accurate and reliable risk assessment.

# CHAPTER 7: DISCUSSION AND FUTURE EXTENSION

## 7.1. General Discussion, Recommendations and Limitations

This research provides empirical evidence that unique sloped work setting and different working techniques for kneeling posture can significantly contribute to the development of knee MSDs among construction roofers. Roofers get exposed to potential knee MSD risks when they undertake shingle installation in a dynamic posture on rooftop. The findings of this research show that higher knee flexion, abduction, adduction and rotation is required in dynamic kneeling during sloped shingle installation. Activation of the knee extensor muscles are also highly increased during shingles installation on sloped rooftops which indicates muscle overloading and hence, represents knee MSD risks. The ranking-based ergonomic risk analysis method proposed in this research is implemented to analyze different phases of the dynamic shingle installation operation. Findings suggest that roofers experience the greatest risk of developing knee MSDs during placing and nailing shingle phases as in these phases they encounter the knee MSD risk factors including repetitions, forceful exertions and prolonged kneeling. Awkward flexion, abduction, adduction, internal and external rotations are also the highest in these two phases. Spearman's correlation test result to assess the relative contributions of each knee angle demonstrate that awkward flexion and abduction and adduction significantly contribute to the potential knee MSD risk measurement. These awkward and extreme rotations should be prioritized for developing knee MSD risk assessment tools and effective interventions to prevent knee injuries among roofers.

To minimize knee MSDs among roofers, awkward postures should also be modified with ergonomic modifications that maintains joint range of motion to accomplish work tasks within the mid-range of motion positions for vulnerable joints. For example, knee flexion can impact the

cyclic contraction and elongation of the lower limb postural muscles and thereby the muscle activation. So, roofers should adjust their knee flexion (not less than 80° and not greater than 140°) during shingle installation on sloped roof surfaces so that cumulative muscular effort and fatigue during prolonged kneeling can be reduced. Employers must protect roofers from the recognized work-related hazards by implementing other types of controls such as engineering or administrative controls and by ensuring that roofers wear appropriate personal protective equipment (PPE). Using knee savers might be helpful to reduce extreme posture because they can minimize the peak lower extremity kinematics during sloped shingle installation. Knee pads can also help protect the knees during prolonged kneeling by distributing body weight over a larger surface area and reducing the force passed on to the knee cap. Knee creepers can reduce stress on the knees and can help prevent serious muscle and joint problems. Roofers can use nail guns with extension for securing shingles. It will allow roofers to work in standing upright position and hence reduce awkward postures, minimize strain and muscle fatigue. However, further research is necessary to investigate the impact of prolonged standing on lower extremity disorders such as ankle and foot pain. Proper ergonomic tools should be utilized that allow workers to maintain optimal joint positions. Eliminating excessive force and awkward posture requirements will reduce worker fatigue and allow high repetitive tasks to be performed without a significant increase in MSD risk for residential roofers. Using mechanical assists, counterbalance systems, powered equipment and ergonomic tools will reduce work effort and muscle exertions. An ergonomic control to reduce roofers MSDs and fall from height due to instability caused by muscle fatigue can be that, roof slope had better be less than 30°. It will reduce both falls from height, low back pains and muscle overloading during roofing task as suggested by Choi and Fredericks (2008) and

(Wang et al. 2017). However, this will need approval from architects, engineers, building owners and general contractors.

As data collection for MSD risk assessment involving human subjects often get affected by the incompleteness in the risk indicator datasets due to missing data, this research has proposed a data fusion method to treat those malformed and imperfect datasets which has been found useful to provide reliable risk assessment results consistent to those obtained from experimental datasets. The numerical stability of CPD tensor decomposition which was used for method development was also analyzed. Results suggest that in spite of using malformed datasets as inputs, fusion through CPD is numerically stable to generate a fused dataset based on the interrelations among the risk factors and the risk indicators that outperformed the unfused datasets with missing values in work-related MSD risk assessment.

Although this study provided empirical evidence for the association of certain work-related factors to knee MSDs among construction roofers, a several limitations are noteworthy. First, the participants in this study were not professional roofers. From the perspective of kinematics, which have proven to be in correlation with MSDs, there should be differences between professional roofers and non-professional roofers. Nevertheless, all subjects involved in this study were physically active and have experience in works such as home remodeling. It is postulated that their biomechanical reactions in the experiments are to a large extent similar to those of professional roofers. However, scientific evidence still needs to be explored to justify such premise. It is possible that professional roofers, based on years of experience, have adopted postures that would mitigate or reduce the risk of knee injuries. However, given all of their training and experience, roofers still have the 2nd highest rate of MSDs in the construction sector. Moreover, this study employed young participants although the working roofers have a larger age range and according

to BLS (2019), roofers suffer most of the injuries at the age range of 18-24. This study chose novice individuals because it intended to observe what initial extreme postures individuals encounter when first exposed to a sloped working environment and without much roofing experience, individuals take on what works best for them. While the study is not there yet, this type of approach with novice individuals will allow us to use this experimental framework to develop new ideas for training and/or interventions to alleviate these awkward postures and hopefully reduce MSD risk in roofing activities. Moreover, several biomechanical studies were found in the previous literature where nonprofessional and novice participants were employed for risk assessment purposes.

Second, three slopes ($0^\circ$, $15^\circ$ and $30^\circ$) were considered for assessing MSD risk in this study which were selected based on the common site practice in the United States residential construction and for the sake of safety of the participants. Although any roof slope higher that $18^\circ$ can be considered as 'high-pitched roof' according to the definition provided by OSHA, roof slope can be higher than $30^\circ$. Hence, further assessment of MSD risks via observation of prolonged static and dynamic shingle installation trials performed on steep rooftops exceeding $30^\circ$ is required.

Third, knee joint rotation and EMG signals of muscle activation data were collected from the experiment as time series data. But only the maximum measurements within a trial from the knee joint rotation data; and the maximum, cumulative and average measurements of different phases within a trial from the EMG data were extracted for risk assessment. The time series features were not fully analyzed. Although these metrics are well established to represent MSD risk in the area of biomechanics, further analysis on the time series features of the risk-related dataset can provide more in-depth understanding of the relevant risk, specifically the changes of

knee joint rotation and postural muscle activation with different working conditions and duration that can be associated with prolonged kneeling and repetitions.

Finally, all the assessments and validation in this research were done based on an experimental study performed in a controlled laboratory setting due to strict protocol in human subject research. For the sake of the safety of the human subjects, a controlled setting is always needed to be maintained. For this reason, the environmental impact has not been extensively studied in this study. Only the work-related factors including roof slope and wring posture have been assessed. Temperature is one of the most encountered stressful factors in the environment (Ahmed 2006). Long term exposure to high and low temperatures affects the human skeletal (joints, muscles and bone) and biological system. This research has identified the potential effects of high and low temperature on human skeletal and biological system and their association to knee MSDs and fatal injuries (FI) among residential construction roofers based on the existing literature review.

## 7.2. Effects of Temperature on Human Skeletal and Biological System

In the following section, the effects of external temperature on human skeletal and biological system and their association to roofers MSDs and FI are discussed.

### 7.2.1. Effects of high temperature on the bones, joints, muscles

High temperatures for a long duration and a frequent period of time of heating may result in greater occupational heat stress which can lead to heat related illness. According to workers' compensation claims data, roofers have the highest rates of heat related illness (BLS 2019; Bonauto et al. 2007). Occupational heat increases the risk of heat stroke, which can be fatal (Bouchama and Knochel 2002) and also can reduce productivity. In this situation, body becomes unable to control its temperature and the body's temperature rises rapidly, the sweating mechanism

fails, and the body is unable to cool down (Hifumi et al. 2018). When human bones are subjected to gradual or sudden heating up to different temperatures repeatedly, bones become brittle, although without any considerable changes in the surface features (Hiller et al. 2003).

For proper physiological and cellular function, human body is designed to maintain a core temperature of 98°F (Pušnik and Miklavec 2009). But during summer, sometimes roofers are subjected to work for 6~8 hours in high temperature up to 120°F. Exposure to this increased temperature can also result in reduced cognitive function of the body and increased risk of MSD such as fall from height. Due to an excessive loss of the water and salt from the body, roofers may face heat exhaustion through excessive sweating (Kenny et al. 2018). In this situation, roofers may suddenly feel cramps in their postural muscles of upper and lower extremity due to sudden and involuntary contraction. Muscle cramps can make it temporarily impossible to use the affected muscle and decreases the productivity of the work. Muscles are the greatest contributor to peak joint contact forces (Kingston et al. 2016) which can contribute to musculoskeletal disorders in the joint. For example, high knee joint contact force is associated with knee joint degeneration or osteoarthritis (Andriacchi and Favre 2014). Muscle cramps can cause severe pain while being used for shingling during roofing task and repetitive exertion with a painful muscle may produce high contact force in the joint the muscle is passing through and overload the joint.

Heat can also increase the risk of injuries in workers as it may result in sweaty palms, fogged-up safety glasses, and dizziness. In roof setting, when roofer sweat a lot during strenuous and repetitive activity, they can be affected by heat cramp. This sweating decreases the body's salt and moisture levels. Low salt levels in muscles causes painful cramps. Sweating may also affect in roofing activities requiring grip such as placing shingles, grabbing nailing guns and installing shingles-- which requires manual performance by roofers. Heat cramps may also be a symptom of

150

heat exhaustion. Painful muscle cramp can alter the force that the muscles are normally able to generate, which can further overload the muscle and lead into MSDs generated by heightened muscular activation. Cumulative muscle overloading coupled with repetitive motion without adequate recovery time may cause MSDs due to overexertion or even imbalance (Kumar 2001).

Another form of medical condition associated with heat stress and prolonged physical exertion is Rhabdomyolysis (Yoshizawa et al. 2016). In this condition, roofers face rapid breakdown, rupture, and death of muscle cells, tissues and fibers. When muscle tissue dies, electrolytes and large proteins are released into the bloodstream that can cause irregular heart rhythms and seizures and damage the kidneys as well. They can cause unconsciousness to the roofers and induce fall related FIs to them.

Heat exposure has also been shown to produce other effects on workers, such as physical discomfort, reduced vigilance and increased fatigue, which could lead to accidents and injuries (Varghese et al. 2018).

In human muscle, optimal muscle performance (or less fatigue) can be obtained at 77–86°F and the extent of fatigue becomes greater and faster at around 100°F (Roots et al. 2009). Without adequate heat dissipation, short-term acute heat exposure can cause a rise in core body temperature and may result in direct heat illnesses including cardiovascular diseases (Vangelova et al. 2006). High increase in temperature and dehydration might have had negative behavioral effects such as physical fatigue, irritability, lethargy, impaired judgment, vigilance decrement, loss of dexterity, coordination and concentration among roofers which can potentially lead to a compromise of occupational safety (Kjellstrom et al. 2009; Xiang et al. 2013).

**7.2.2. Effects of high temperature on human biological system**

Generally, in hot weather during burning sunlight, as body temperature increases, the chance of observing abnormalities in the biological effects also increases (Ahmed 2006). Extended temperature and sunlight may cause hyperthermia—a condition of having a body temperature greatly above normal level (Simon 1993). Hyperthermia cause defects in central nervous system in all species and cell death (Upfold et al. 1989). This may cause pain in the skeletal muscles, makes the muscle rigid or sometimes disable to perform tasks (Rock and Kono 2008). In residential roofing, roofers spend most of their working time in awkward kneeling posture which involves cyclic contraction and relaxation of the knee postural muscles. More workload in a rigid knee flexor or extensor muscle and frequent change in knee joint rotation during shingle installation enhance the muscular demand and contribute more to the knee joint compressive force causing knee MSDs, such as knee joint degeneration, patellofemoral pain and osteoarthritis (Xu et al. 2017).

Heat causes a transient depression in overall protein synthesis (Plesofsky-Vig and Brambl 1985) -- a biological process at which cell produces protein. Moreover, heat shock can reduce normal protein synthesis of mammalian brain Edwards et al. (2002). Insufficient protein in cell reduces lean body mass (the difference between total body weight and body fat weight), muscle strength, and muscle function (Cava et al. 2017), which in turn causes cramping, soreness and fatigue to the muscles. Muscle fatigue is also dependent on temperature (Roots et al. 2009). When roofers work in real sites, they may be subjected to work at a temperature up to 120 degrees and have to work up to 6~8 hours per day which can easily cause muscle fatigue. In a fatigued condition roofer may lose their balance which can cause fall from height and death.

Heat stress results in alterations in the metabolic process of the human body (Kay and Marino 2000). Higher environmental temperature is also associated with decreased metabolic rate (Wilmore et al. 1975). Nilsen (1984) recorded that the hyperthermia distorts the small vessels and also produces irregularities of the microvascular pattern. The distortion in blood vessels may cause the impairment of the metabolic processes (carbohydrates genesis) and thus, may reflex some delaying in the growth of the cellular processes. Poor body metabolism causes soreness and weakness in the muscles which may be associated to metabolic muscle disorders (Phillips and Haas 2008). In shingle installation, roofers face the most awkward rotation, forceful exertion and repetition during placing and installation shingle phases (Dutta et al. 2020a). During these phases, using a sore and week muscle may further deteriorate the muscle condition and may overload the muscle tissues, at times exceeding the maximum strength that the muscle is able to generate, which can potentially lead to muscle fatigue. Cumulative muscle overloading coupled with repetitive motion without adequate recovery time may cause MSDs due to overexertion or even imbalance (Kumar 2001).

## 7.2.3. Effects of low temperature on the bones, joints, muscles

Working in cold temperature can put strain on the body. Temperature can also affect muscles contraction. In colder weather, the rate of oxygen release from hemoglobin to the muscle becomes slower which makes it difficult for muscles to contract in cold weather as opposed to warmer conditions. In this situation, less oxygen is available in the muscle, which causes muscle stiffness and as a result, roofers can overly exert to perform the shingle installation and can cause MSDs such as low back pain, knee disorders, shoulder MSD, muscles numbness etc.

Low temperatures cause blood vessels and arteries to narrow, restricting blood flow and reducing oxygen to the heart. In this situation, heart pumps harder to circulate blood through the

constricted blood vessels. As a result, the blood pressure and heart rate increase. A poor or less blood circulation through the joints in the body such as, knee and shoulder joint causes numbness and tingling in the parts connected to the joints, muscle cramps and swellings (Nunes and Bush 2012). When the blood does not circulate correctly, oxygen and nutrients cannot reach tissues effectively, which can result in stiffness and cramping (Parisi et al. 2003). In roofing task that involves dynamic motion of the lower limbs, especially knee, during placing and nailing shingles; stiff thigh muscles can contribute to excessive knee joint compressive load and overload due to forceful exertions (Kurita et al. 2012). Restricted blood flow may cause bones ache in cold weather (Robertson 2017).

**7.2.4. Effects of low temperature on human biological system**

Excessive cool temperature may cause Hypothermia, which is a disorder where body core temperature in human decreases below 35 °C, and the body fails to produce the required amount of heat to maintain thermal balance, due to an excess heat loss (Brown et al. 2012). As hypothermia progresses, there is an increasing decrease in cardiac output, heart rate and blood pressure (Beker et al. 2018).

As hypothermia continues, there is a progressively decrease in oxygen consumption by the brain which can cause sudden brain MSD and death. The blood consumption by spinal cord might decrease in low temperature. In residential roofing task, roofer most often have to maintain a stooped posture during shingle installation. Due to low blood circulation in the spinal cord, the low back muscles may become overloaded and provide feeling of numbness and tingling as the contraction will be limited due to less oxygen availability (Richardson et al. 2006).

**7.3. Future Extension**

From the above discussions, it can be understandable that there are certain areas in the risk assessment studies for roofers MSD that need further attention. First, high and low outdoor temperature can have some unavoidable impacts on human musculoskeletal and biological system. But these impacts cannot be observed during experiments in the lab setting. Although, it is well accepted to validate biomechanical models based on laboratory based experimental studies in the research community, for in-depth understanding and characterization of MSDs and FIs among roofers, conducting the roofing experiment in real construction site is necessary. This will allow to analyze the effects of different work-related factors such as roof setting, working posture on the change of skeletal joint movement with varying temperatures. The effect of the variation of temperature on muscle physiology and the activation of the postural muscles should also be explored further to study muscle fatigue in roofing. These types of studies will provide knowledge on the correlation of temperature variation to MSDs and FIs among residential roofers so that, effective interventions, training and ergonomic solutions can be developed for the sake of reducing MSDs and traumatic injuries in roofing community. Interventions, such as, thermal safety vest, water resistant insulated coat/jacket, insulated gloves and boots, ear covering hats, cooling vests, moisture-absorbing vest-- should be developed and tested to understand their effectiveness in protecting residential roofers during hot and cold weather.

Second, there is no study in the existing literature that represents the variation between the biomechanical reactions of experienced and novice roofers for understanding the effect of experience on posture optimization and risk-taking behavior. Roofing experiments should also be conducted involving both professional and novice roofers to see if experienced roofers can optimize their posture and hence can be less susceptible to MSDs. The variation in the

biomechanical reactions and postural preferences between the professional and novice roofers should also be studied to quantify the postural optimization with and without experience in roofing which will allow to identify the correlation between experience and MSD risk perception in roofing.

Third, as the work environment exposes roofers to both FIs and MSD risks, the imbalance measurement can be studied along with knee joint rotation and muscle activation for developing risk assessment methods and interventions. To facilitate the risk measurement for imbalance, tasks performed on steep rooftops exceeding 30°, which was not included in this study, can be studied with fall protection measures in the future.

Finally, all the assessments in this study were performed based on the general definition of awkward posture provided in the existing literature which suggests that the postures that result in increased risk of fatigue, pain or skeletal joint degeneration when are used repetitively or for prolonged periods. Generally, awkward posture is defined as the deviation of a body part from its natural or 'neutral' position while job tasks are being performed. However, there is no threshold of the deviation from the neutral position that can conclude if a posture is awkward or not. Frequent and high contact stress at knee joint is associated with knee osteoarthritis and damage of articular cartilage of the knee joint — two common forms of knee disorders. A previous study showed that, with an increase in knee flexion from 15.5° during walking to 90° during squatting posture, the contact stress in knee joint increased significantly by over 80% (Thambyah et al. 2005).  Knee flexion beyond 90° generates larger moment and forces which results in high stress in the knee joint (Nagura et al. 2002). These indicate a strong association between knee rotational angle and knee joint contact stress which relates to knee MSDs. Therefore, in this study, potential knee MSD risk is defined as an increase in knee rotations that creates high contact stress in the

knee joint. Knee MSD risk is considered to increase as the knee rotational angles encounters larger awkward postures. Awkward posture is thus considered as a deep flexed posture of knee (>90°) with medial and lateral rotation leading to increased amount of stress in the knee joint. However, further biomechanical studies on the correlation between the knee joint contact stress and the physical structures of knee joint skeletal components, such as, tibia, femur, patella, tendon and ligaments with varying knee joint rotation and the knee postural muscle activation are required to measure the threshold of knee joint rotation and muscle activation that can be defined as 'awkward' and hence, can be identified as 'risky' for knee MSDs such as knee osteoarthritis and knee joint degeneration.

# CHAPTER 8: CONCLUSION AND KEY CONTRIBUTIONS

## 8.1. Conclusion

This empirical study analyzed the effects of the unique sloped work setting and different working techniques for kneeling posture on knee MSD risks in roofing activities. A set of programed experiments were conducted to investigate the association of these factors to the development of knee MSDs in residential roofing. The participants were asked to simulate shingle installation task on a slope-configurable wooden platform. They mimicked placing and nailing shingles for three different roof slopes (0, 15 and 30 degree) and two different kneeling posture (static kneeling and dynamic kneeling) which were selected based on site practices, experimental safety and feasibility. To indicate knee MSD risk, knee kinematics data and bilateral knee postural muscle activation signals were measured using VICON optical motion analysis and surface electromyography (EMG) system respectively.

Research findings demonstrated that roof slopes, working postures and their interactions significantly impact knee rotational movement–flexion, abduction, adduction, internal and external rotation. Roofers become exposed to higher knee MSD risks during shingle installation compared to while in a static kneeling posture on high-pitched rooftop as higher knee flexion, abduction, adduction and rotation are required in dynamic kneeling posture. They have also been found to be significantly associated with the peak normalized activation of the knee postural (flexor and extensor) muscles. Specifically, roofers become exposed to a greater risk of developing knee MSDs with the increase of roof slope during shingle installation as the dynamic kneeling posture during shingle installation on high-pitched rooftops requires significantly higher muscle loading for task performance compared to a static flexed kneeling posture. Therefore, the awkward

knee joint rotation and the heightened muscle activation while kneeling during shingle installation on high-pitched rooftop should be given particular attention for effective intervention development.

A ranking-based ergonomic method has been developed and applied to identify the shingle installation phase with the greatest potential knee MSD risk. Aggregation-based and multiplication scoring models have been used to develop the risk assessment method. Roofers' knees are exposed to the greatest potential knee MSD risk while placing shingles and nailing shingles as roofers face forceful exertion, prolonged and repeated awkward kneeling movements in these phases. As these phases involve extreme awkward rotations in knee joint, the can potentially be considered as the riskiest phases in shingle installation operation. Awkward flexion, abduction and adduction should be prioritized for developing knee MSD risk assessment tools as they significantly contribute to the potential knee MSD risk measurement.

Although, risk exposure data collection with human subject involvement is essential for understanding construction worker's behaviors and conditions that may expose them to work-related MSD risks on construction sites, the accuracy of risk assessments is often affected by incompleteness in the risk exposure datasets when some data becomes missing due to human and technology-induced errors. To treat these imperfect MSD risk-related datasets, a data fusion method has been proposed using canonical polyadic decomposition which can help not only in replacing missing values, but also holds the correlation among the potential risk factors and the risk indicators during replacement. Numerical stability evaluation of fusion through CPD for work-related MSD risk assessment also suggested that in spite of using malformed experimental data with missing value as input, CPD is numerically stable to provide more reliable risk assessment result compared to unfused sparse datasets. To enable accurate assessment of work-related risk

factors, the proposed method can therefore be used as an alternative risk assessment method when data collected from experimental studies or real construction sites are incomplete due to missing data points.

## 8.2. Key Contributions

The outcome resulting from this research will greatly enhance the understanding of the physical hazard exposures, improve their risk assessment methods, and help develop more effective intervention methods to minimize the risks of knee MSDs in roofing construction industry substantially reducing work efficiency or productivity. The findings of this research provide useful information to statistically understand the effects of work-related risk factors: roof slopes and kneeling working postures, on knee MSDs while performing shingle installation on a sloped rooftop. For example, by knowing the changes in knee joint rotation at different roof slope and working kneeling posture, interventions such as, wearable or external devices, knee pads, knee supporting devices and knee savers can be tested and developed for preventing awkward knee rotation in sloped-shingle installation. A known level of knee-muscle activation at different work settings and kneeling postures also help understand the mechanism contributing to the onset of knee MSDs (Nagura et al. 2006) and develop knee joint biomechanical models for computing in-vivo muscle and contact forces in different occupational tasks (Lin et al. 2010). The ergonomic method proposed for detailed risk analysis of the shingle installation task may generate useful information on the unsafe conditions of shingle installation operations to facilitate effective intervention strategies (education, training, and tools) for reducing knee exposure and hence minimizing knee injuries and disorders among construction roofers. This research will advance the construction knowledge by providing insights of roofing work-related conditions on knee MSD development and thereby helping reducing knee MSD risks of roofers through developing training

and ergonomic solutions. Also, the data fusion method proposed in this research can be used for treating imperfect risk related datasets which can also be used as an alternative risk assessment method when data collected from experimental studies or real construction sites are incomplete due to missing data points. The findings will help enable reliable assessment of work-related risk factors for work-related MSDs among construction workers.

# PUBLICATIONS FROM THIS RESEARCH

*Journals:*

**Amrita Dutta,** Scott P. Breloff, Fei Dai, Erik W. Sinsel, Christopher M. Warren, Robert E. Carey and John Z. Wu, "Application of Data Fusion via Canonical Polyadic Decomposition in Risk Assessment of Musculoskeletal Disorders in Construction: Procedure and Stability Evaluation**", Journal of Construction Engineering and Management,** 2020 (In review).

**Amrita Dutta,** Scott P. Breloff, Fei Dai, Erik W. Sinsel, Christopher M. Warren, Robert E. Carey and John Z. Wu, "Fusing Imperfect Experimental Data for Risk Assessment of Musculoskeletal Disorders in Construction Using Canonical Polyadic Decomposition**", Automation in Construction,** 2020, Volume 119, 103322**.**

**Amrita Dutta**, Scott P. Breloff, Fei Dai, Erik W. Sinsel, Christopher M. Warren, Robert E. Carey and John Z. Wu. "Effects of Working Posture and Roof Slope on Activation of Lower Limb Muscles during Shingle Installation ", **Ergonomics**, 2020, 1-12.

**Amrita Dutta**, Scott P. Breloff, Fei Dai, Erik W. Sinsel, Christopher M. Warren and John Z. Wu, "Identifying Potentially Risky Phases Leading to Knee Musculoskeletal Disorders during Shingle Installation Operations", **Journal of Construction Engineering and Management,** 2020**,** Volume 146(3), 04019118 (**Selected as Editor's choice**).

Scott P. Breloff, **Amrita Dutta,** Fei Dai, Xiaopeng Ning, Erik W. Sinsel, Christopher M. Warren, and John Z. Wu, "Assessing Work-Related Risk Factors for Musculoskeletal Knee Disorders in Construction Roofing Tasks ", **Applied Ergonomics**, 2019, Volume 81,10290.

*Conference Proceedings:*

**Amrita Dutta,** Scott P. Breloff, Fei Dai, Erik W. Sinsel, Christopher M. Warren, Robert E. Carey and John Z. Wu, "Electromyography Signal Analysis of Knee Flexor and Extensor Muscles in Potential Knee Musculoskeletal Disorders during Roofing", Proceedings of the **ASCE Construction Research Congress,** 2020 (**Best Paper Award in the Health, Safety, and Workforce Issues track**).

**Amrita Dutta**, Scott P. Breloff, Fei Dai, Erik W. Sinsel, Christopher M. Warren, and John Z. Wu, "Identifying A Ranking Method for Assessing the Potential Risk of Knee Musculoskeletal Disorders among Roofers in Shingle Installation", Proceedings of The **ASCE International Conference on Computing in Civil Engineering** 2019: Data, Sensing, and Analytics (pp. 445-453), Reston, VA, American Society of Civil Engineers.

# ABOUT THE AUTHOR

**Amrita Dutta** is a PhD student in the department of Civil and Environmental Engineering at West Virginia University. Currently she is working as a graduate research assistant with Dr. Fei Dai. She received her master's degree in construction engineering and management from Virginia Tech and a bachelor's degree in civil engineering from Bangladesh University of Engineering and Technology. Her research interest includes statistical data analysis, computational modeling and data analytics with applications to work-related musculoskeletal disorders, occupational safety and health, construction ergonomics and risk exposures analysis. She has been awarded the Dianne Dubetz Anderson Ph. D Fellowship from the department of Civil and Environmental Engineering at West Virginia University for the academic year of 2019-2020. She is also a recipient of a Best Paper Award in the Health, Safety, and Workforce Issues track in the Construction Research Congress 2020.

# REFERENCES

Acar, E., Dunlavy, D. M., Kolda, T. G., and Mørup, M. (2011). "Scalable tensor factorizations for incomplete data." *Chemometrics and Intelligent Laboratory Systems*, 106(1), pp. 41-56.

Ahmed, R. (2006). "The relation between biological consequences and high temperature in mammals." *Int. J. Zool. Res*, 2, 48-59.

Albers, J., and Estill, C. F. (2007). "Simple solutions; ergonomics for construction workers." *Department of Health & Human Survices (DHHS), Cincinnati, OH. The National Institute for Occupational Safety and Health (NIOSH)*, Publicaion Number: 2007-22.

Alexander, D. L., Tropsha, A., and Winkler, D. A. (2015). "Beware of R 2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models." *Journal of Chemical Information and Modeling*, 55(7), pp. 1316-1322.

Alimuddin, I., Sumantyo, J. T. S., and Kuze, H. (2012). "Assessment of pan-sharpening methods applied to image fusion of remotely sensed multi-band data." *International Journal of Applied Earth Observation and Geoinformation*, 18, pp. 165-175.

Amiri, M., and Jensen, R. (2016). "Missing data imputation using fuzzy-rough methods." *Neurocomputing*, 205, pp. 152-164.

Anaissi, A., Makki Alamdari, M., Rakotoarivelo, T., and Khoa, N. (2018). "A tensor-based structural damage identification and severity assessment." *Sensors*, 18(1), 111.

Andersson, C. A., and Bro, R. (2000). "The N-way toolbox for MATLAB." *Chemometrics and intelligent laboratory systems*, 52(1), pp. 1-4.

Andriacchi, T. P., and Favre, J. (2014). "The nature of in vivo mechanical signals that influence cartilage health and progression to knee osteoarthritis." *Current rheumatology reports*, 16(11), 463.

Asif, M. T., Mitrovic, N., Dauwels, J., and Jaillet, P. (2016). "Matrix and tensor based methods for missing data estimation in large traffic networks." *IEEE Transactions on Intelligent Transportation Systems*, 17(7), pp. 1816-1825.

Bader, B., and Kolda, T. (2019). "MATLAB Tensor Toolbox Version 2.4." <https://www.sandia.gov/~tgkolda/TensorToolbox/index-2.6.html>. (May 14, 2019).

Bagaya, O., and Song, J. (2016). "Empirical study of factors influencing schedule delays of public construction projects in Burkina Faso." *Journal of Management in Engineering*, 32(5), 05016014.

Barrios, J. A., Crossley, K. M., and Davis, I. S. (2010). "Gait retraining to reduce the knee adduction moment through real-time visual feedback of dynamic knee alignment." *Journal of biomechanics*, 43(11), 2208-2213.

Beker, B., Cervellera, C., De Vito, A., and Musso, C. (2018). "Human Physiology in Extreme Heat and Cold." *Int Arch Clin Physiol*, 1(001).

BLS (2019). "Nonfatal Occupational Injuries and Illnesses: Cases with days away from work." <https://www.bls.gov/news.release/pdf/osh2.pdf>. (4 March, 2019).

Bonauto, D., Anderson, R., Rauser, E., and Burke, B. (2007). "Occupational heat illness in Washington State, 1995–2005." *American journal of industrial medicine*, 50(12), 940-950.

Bouchama, A., and Knochel, J. P. (2002). "Heat stroke." *New England journal of medicine*, 346(25), 1978-1988.

Breloff, S. P., Dutta, A., Dai, F., Sinsel, E. W., Warren, C. M., Ning, X., and Wu, J. Z. (2019a). "Assessing work-related risk factors for musculoskeletal knee disorders in construction roofing tasks." *Applied Ergonomics*, 81, 102901.

166

Breloff, S. P., Sinsel, E. W., Dutta, A., Carey, R. E., Warren, C. M., Dai, F., Ning, S., and Wu, J. Z. (2019c). "Are knee savers and knee pads a viable intervention to reduce lower extremity musculoskeletal disorder risk in residential roofers?" *International Journal of Industrial Ergonomics*, 74, 102868.

Breloff, S. P., Wade, C., and Waddell, D. E. (2019b). "Lower extremity kinematics of cross-slope roof walking." *Applied Ergonomics*, 75, 134-142.

Bro, R., and Kiers, H. A. (2003). "A new efficient method for determining the number of components in PARAFAC models." *Journal of Chemometrics: A Journal of the Chemometrics Society*, 17(5), pp. 274-286.

Brown, D. J., Brugger, H., Boyd, J., and Paal, P. (2012). "Accidental hypothermia." *New England Journal of Medicine*, 367(20), 1930-1938.

Brown, M. L., and Kros, J. F. (2003). "Data mining and the impact of missing data." *Industrial Management & Data Systems*.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., and Munafò, M. R. (2013). "Power failure: why small sample size undermines the reliability of neuroscience." *Nature Reviews Neuroscience*, 14(5), pp. 365-376.

Canetti, E. F., Schram, B., Orr, R. M., Knapik, J., and Pope, R. (2020). "Risk factors for development of lower limb osteoarthritis in physically demanding occupations: A systematic review and meta-analysis." *Applied Ergonomics*, 86, 103097.

Carroll, J. D., and Chang, J.-J. (1970). "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition." *Psychometrika*, 35(3), 283-319.

Castanedo, F. (2013). "A review of data fusion techniques." *The Scientific World Journal*, 2013.

Cava, E., Yeat, N. C., and Mittendorfer, B. (2017). "Preserving healthy muscle during weight loss." *Advances in nutrition*, 8(3), 511-519.

Chapman, A. R., Vicenzino, B., Blanch, P., Knox, J. J., and Hodges, P. W. (2010). "Intramuscular fine-wire electromyography during cycling: repeatability, normalisation and a comparison to surface electromyography." *Journal of Electromyography and Kinesiology*, 20(1), 108-117.

Charlier, J., State, R., and Hilger, J. (2018). "Non-negative paratuck2 tensor decomposition combined to lstm network for smart contracts profiling." *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 74-81.

Chen, H., Luo, X., Zheng, Z., and Ke, J. (2019). "A proactive workers' safety risk evaluation framework based on position and posture data fusion." *Automation in Construction*, 98, pp. 275-288.

Chen, J., Qiu, J., and Ahn, C. (2017). "Construction worker's awkward posture recognition through supervised motion tensor decomposition." *Automation in Construction*, 77, pp. 67-81.

Cheng, T., Migliaccio, G. C., Teizer, J., and Gatti, U. C. (2012). "Data fusion of real-time location sensing and physiological status monitoring for ergonomics analysis of construction workers." *Journal of Computing in Civil Engineering*, 27(3), pp. 320-335.

Cheng, T., Teizer, J., Migliaccio, G. C., and Gatti, U. C. (2013). "Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data." *Automation in Construction*, 29, pp. 24-39.

Choi, S., and Fredericks, T. (2008). "Surface slope effects on shingling frequency and postural balance in a simulated roofing task." *Ergonomics*, 51(3), 330-344.

Choi, S. D. (2008). "Postural balance and adaptations in transitioning sloped surfaces." *International Journal of Construction Education and Research*, 4(3), 189-199.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*, Routledge.

Coplan, J. A. (1989). "Rotational motion of the knee: a comparison of normal and pronating subjects." *Journal of Orthopaedic & Sports Physical Therapy*, 10(9), 366-369.

CPWR (2018). "The construction chart book : The United States construction industry and its workers, 6th Edition,." Silver Spring, MD.

Dai, F., Yoon, Y., and GangaRao, H. V. (2016). "State of practice of construction site safety in the USA." *Frontiers of engineering management*, 3(3), 275-282.

Data, M. C. (2016). *Secondary analysis of electronic health records*, Springer International Publishing.

Dauwels, J., Garg, L., Earnest, A., and Pang, L. K. (2012). "Tensor factorization for missing data imputation in medical questionnaires." *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2109-2112.

David, G. (2005). "Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders." *Occupational Medicine*, 55(3), pp. 190-199.

De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000). "On the best rank-1 and rank-(r 1, r 2,..., rn) approximation of higher-order tensors." *SIAM journal on Matrix Analysis and Applications*, 21(4), 1324-1342.

Debals, O., Van Eeghem, F., Vervliet, N., and De Lathauwer, L. (2019). "Tensor computations using Tensorlab." <https://www.tensorlab.net/demos/tutorial.pdf>. (May 19, 2019).

Dooley, E., Carr, J., Carson, E., and Russell, S. (2019). "The effects of knee support on the sagittal lower-body joint kinematics and kinetics of deep squats." *Journal of biomechanics*, 82, 164-170.

Dulay, G. S., Cooper, C., and Dennison, E. (2015). "Knee pain, knee injury, knee osteoarthritis & work." *Best Practice & Research Clinical Rheumatology*, 29(3), 454-461.

Dunlavy, D. M., Kolda, T. G., and Acar, E. (2011). "Temporal link prediction using matrix and tensor factorizations." *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(2), 10.

Dutta, A., Breloff, S. P., dai, F., Sinsel, E. W., Carey, R. E., Warren, C. M., and Wu, J. Z. (2020c). "Fusing Imperfect Experimental Data for Risk Assessment of Musculoskeletal Disorders in Construction Using Canonical Polyadic Decomposition." *Automation in Construction*.

Dutta, A., Breloff, S. P., Dai, F., Sinsel, E. W., Warren, C. M., Carey, R. E., and Wu, J. Z. (2020b). "Effects of working posture and roof slope on activation of lower limb muscles during shingle installation." *Ergonomics*, 1-12.

Dutta, A., Breloff, S. P., Dai, F., Sinsel, E. W., Warren, C. M., and Wu, J. Z. (2020a). "Identifying Potentially Risky Phases Leading to Knee Musculoskeletal Disorders during Shingle Installation Operations." *Journal of Construction Engineering and Management*, 146(3), 04019118.

Edwards, M. J., Walsh, D. A., and Li, Z. (2002). "Hyperthermia, teratogenesis and the heat shock response in mammalian embryos in culture." *International Journal of Developmental Biology*, 41(2), 345-358.

El-Sayegh, S. M., and Mansour, M. H. (2015). "Risk assessment and allocation in highway construction projects in the UAE." *Journal of Management in Engineering*, 31(6), 04015004.

Fredericks, T. K., Abudayyeh, O., Choi, S. D., Wiersma, M., and Charles, M. (2005). "Occupational injuries and fatalities in the roofing contracting industry." *Journal of Construction Engineering and Management*, 131(11), 1233-1240.

Freund, J., Rudolf, Mohr, L., Donna, and Wilson, J., William (2010). "Statistical Methods, 3rd edition." Academic Press, ELSEVIER, Cambridge, MA.

Gallagher, S., Moore, S., and Dempsey, P. G. (2009). "An analysis of injury claims from low-seam coal mines." *Journal of safety research*, 40(3), 233-237.

Gallagher, S., Pollard, J., and Porter, W. L. (2011). "Electromyography of the thigh muscles during lifting tasks in kneeling and squatting postures." *Ergonomics*, 54(1), 91-102.

Gebregziabher, M., and DeSantis, S. M. (2010). "Latent class based multiple imputation approach for missing categorical data." *Journal of Statistical Planning and Inference*, 140(11), 3252-3262.

Gelman, A., and Raghunathan, T. E. "Using conditional distributions for missing-data imputation." *Proc., Statistical Science*, 268-269.

Gordon, A., Huxley, A. F., and Julian, F. (1966). "The variation in isometric tension with sarcomere length in vertebrate muscle fibres." *The Journal of physiology*, 184(1), 170-192.

Gray, A. M., Ellison, K. T., and Smith, B. A. (2015). "Effects of Using Knee Savers on Lower Extremity Kinematics in Male Collegiate Baseball Catchers During Squatting." *Texas Orthopeadic Journal* 1, 51-59.

Grood, E. S., and Suntay, W. J. (1983). "A joint coordinate system for the clinical description of three-dimensional motions: application to the knee." *Journal of biomechanical engineering*, 105(2), 136-144.

Gyemi, D. L., van Wyk, P. M., Statham, M., Casey, J., and Andrews, D. M. (2016). "3D peak and cumulative low back and shoulder loads and postures during greenhouse pepper harvesting using a video-based approach." *Work*, 55(4), 817-829.

Ha, M., and Han, D. (2017). "The relationship between knee joint angle and knee flexor and extensor muscle strength." *Journal of physical therapy science*, 29(4), 662-664.

Hamer, R. M., and Simpson, P. M. (2009). "Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials." *American Journal of Psychiatry*, 166(6), pp. 639-641.

Hatfield, G. L., Stanish, W. D., and Hubley-Kozey, C. L. (2015). "Three-dimensional biomechanical gait characteristics at baseline are associated with progression to total knee arthroplasty." *Arthritis care & research*, 67(7), 1004-1014.

Hifumi, T., Kondo, Y., Shimizu, K., and Miyake, Y. (2018). "Heat stroke." *Journal of intensive care*, 6(1), 30.

Hiller, J. C., Thompson, T. J., Evison, M. P., Chamberlain, A. T., and Wess, T. J. (2003). "Bone mineral change during experimental heating: an X-ray scattering investigation." *Biomaterials*, 24(28), 5091-5097.

Hitchcock, F. L. (1927). "The expression of a tensor or a polyadic as a sum of products." *Journal of Mathematics and Physics*, 6(1-4), 164-189.

Hofer, J. K., Gejo, R., McGarry, M. H., and Lee, T. Q. (2011). "Effects on tibiofemoral biomechanics from kneeling." *Clinical Biomechanics*, 26(6), pp. 605-611.

Huang, G., Wang, S., Xiao, F., and Sun, Y. (2009). "A data fusion scheme for building automation systems of building central chilling plants." *Automation in Construction*, 18(3), 302-309.

IHSA (2015a). "MSD hazards and controls for low-slope roofers." <https://www.ihsa.ca/PDFs/Products/Id/IHSA024.pdf>. (10th August, 2017).

IHSA (2015b). "Heat Stress Prevention for Roofers." <https://www.ihsa.ca/PDFs/Products/Id/IHSA025.pdf>. (15 September, 2017).

IHSA (2015c). "Working at Heights and Fall Protection." <https://www.ihsa.ca/PDFs/Products/Id/IHSA026.pdf>. (20 August, 2017).

Jaffar, N., Abdul-Tharim, A., Mohd-Kamar, I., and Lop, N. (2011). "A literature review of ergonomics risk factors in construction industry." *Procedia Engineering*, 20, 89-97.

Johnson, W. E. (1976). "Roofers Handbook." *Carlsbad, CA: Craftsman Book Company*.

Kajaks, T., and Costigan, P. (2015). "The effect of sustained static kneeling on kinetic and kinematic knee joint gait parameters." *Applied Ergonomics*, 46, 224-230.

Kang, H. (2013). "The prevention and handling of the missing data." *Korean Journal of Anesthesiology*, 64(5), pp. 402-406.

Kaushik, V., and Charpe, N. A. (2008). "Effect of body posture on stress experienced by worker." *Studies on Home and Community Science*, 2(1), 1-5.

Kay, D., and Marino, F. E. (2000). "Fluid ingestion and exercise hyperthermia: implications for performance, thermoregulation, metabolism and the development of fatigue." *Journal of sports sciences*, 18(2), 71-82.

Kennedy, P., and Cresswell, A. (2001). "The effect of muscle length on motor-unit recruitment during isometric plantar flexion in humans." *Experimental brain research*, 137(1), 58-64.

Kenny, G. P., Wilson, T. E., Flouris, A. D., and Fujii, N. (2018). "Heat exhaustion." *Handbook of clinical neurology*, Elsevier, 505-529.

Khaleghi, B., Khamis, A., Karray, F. O., and Razavi, S. N. (2013). "Multisensor data fusion: A review of the state-of-the-art." *Information Fusion*, 14(1), pp. 28-44.

Kingston, D. C., Tennant, L. M., Chong, H. C., and Acker, S. M. (2016). "Peak activation of lower limb musculature during high flexion kneeling and transitional movements." *Ergonomics*, 59(9), pp. 1215-1223.

Kjellstrom, T., Gabrysch, S., Lemke, B., and Dear, K. (2009). "The 'Hothaps' programme for assessing climate change impacts on occupational health and productivity: an invitation to carry out field studies." *Global health action*, 2(1), 2082.

Kolda, T. G., and Bader, B. W. (2009). "Tensor decompositions and applications." *SIAM Review*, 51(3), pp. 455-500.

Kolda, T. G., and Sun, J. "Scalable tensor decompositions for multi-aspect data mining." *Proc., 2008 Eighth IEEE International Conference on Data Mining*, IEEE, 363-372.

Kumar, S. (2001). "Theories of musculoskeletal injury causation." *Ergonomics*, 44(1), 17-47.

Kurita, M., Tomita, T., Yamazaki, T., Fujii, M., Futai, K., Shimizu, N., Yoshikawa, H., and Sugamoto, K. (2012). "In vivo kinematics of high-flex mobile-bearing total knee arthroplasty, with a new post-cam design, in deep knee bending motion." *International orthopaedics*, 36(12), 2465-2471.

Lahat, D., Adali, T., and Jutten, C. (2015). "Multimodal data fusion: an overview of methods, challenges, and prospects." *Proceedings of the IEEE*, 103(9), pp. 1449-1477.

Lay, A. N., Hass, C. J., Nichols, T. R., and Gregor, R. J. (2007). "The effects of sloped surfaces on locomotion: an electromyographic analysis." *Journal of biomechanics*, 40(6), 1276-1285.

Lee, W., Lin, K.-Y., Seto, E., and Migliaccio, G. C. (2017). "Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction." *Automation in Construction*, 83, 341-353.

Li, X., Komeili, A., Gül, M., and El-Rich, M. (2017). "A framework for evaluating muscle activity during repetitive manual material handling in construction manufacturing." *Automation in Construction*, 79, 39-48.

Lin, J. Y., Lu, Y., and Tu, X. (2012). "How to avoid missing data and the problems they pose: design considerations." *Shanghai Archives of Psychiatry*, 24(3), pp. 181-184.

Lin, Y.-C., Walter, J. P., Banks, S. A., Pandy, M. G., and Fregly, B. J. (2010). "Simultaneous prediction of muscle and contact forces in the knee during gait." *Journal of biomechanics*, 43(5), 945-952.

Lu, M.-L., Kincl, L., Lowe, B., Succop, P., and Bhattacharya, A. (2015). "Muscular activity of lower limb muscles associated with working on inclined surfaces." *Ergonomics*, 58(2), 278-290.

Luengo, J., Sáez, J. A., and Herrera, F. (2012). "Missing data imputation for fuzzy rule-based classification systems." *Soft Computing*, 16(5), pp. 863-881.

Lunde, L.-K., Koch, M., Knardahl, S., Wærsted, M., Mathiassen, S. E., Forsman, M., Holtermann, A., and Veiersted, K. B. (2014). "Musculoskeletal health and work ability in physically demanding occupations: study protocol for a prospective field study on construction and health care workers." *BMC public health*, 14(1), 1075.

Malhotra, N. K. (1987). "Analyzing marketing research data with incomplete information on the dependent variable." *Journal of Marketing Research*, 24(1), pp. 74-84.

Marras, W. S., and Karwowski, W. (2003). "Occupational Ergonomics: Design and Management of Work Systems, *1st edition*." CRC Press, Boca Raton, FL.

McClellan, A. J., Albert, W. J., Fischer, S. L., Seaman, F. A., and Callaghan, J. P. (2009). "Shoulder loading while performing automotive parts assembly tasks: a field study." *Occupational Ergonomics*, 8(2, 3), 81-90.

Moore, J. R., and Wagner, J. P. (2014). "Fatal events in residential roofing." *Safety science*, 70, 262-269.

Nagura, T., Dyrby, C. O., Alexander, E. J., and Andriacchi, T. P. (2002). "Mechanical loads at the knee joint during deep flexion." *Journal of Orthopaedic Research*, 20(4), 881-886.

Nagura, T., Matsumoto, H., Kiriyama, Y., Chaudhari, A., and Andriacchi, T. P. (2006). "Tibiofemoral joint contact force in deep knee flexion and its consideration in knee osteoarthritis and joint replacement." *Journal of applied biomechanics*, 22(4), 305-313.

Ng, K. B., and Kantor, P. B. (2000). "Predicting the effectiveness of naive data fusion on the basis of system characteristics." *Journal of the American Society for Information Science*, 51(13), pp. 1177-1189.

Nilsen, N. Ø. (1984). "Endothelial changes and microvascular leakage due to hyperthermia in chick embryos." *Virchows Archiv B*, 46(1), 165-174.

Nunes, I. L., and Bush, P. (2012). "Work-related musculoskeletal disorders assessment and prevention." *Ergonomics-A Systems Approach*.

Onishi, H., Yagi, R., Oyama, M., Akasaka, K., Ihashi, K., and Handa, Y. (2002). "EMG-angle relationship of the hamstring muscles during maximum knee flexion." *Journal of Electromyography and Kinesiology*, 12(5), 399-406.

OSHA (2019). "(Occupational Safety and Health Administration), (https://www.osha.gov/SLTC/ergonomics/training.html)." (7/8/2019).

OSHA (2019). "Training and Assistance,." <https://www.osha.gov/SLTC/ergonomics/training.html>. (10 th April, 2019).

Papalexakis, E. E., Faloutsos, C., and Sidiropoulos, N. D. (2017). "Tensors for data mining and data fusion: Models, applications, and scalable algorithms." *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(2), 16.

Parisi, L., Pierelli, F., Amabile, G., Valente, G., Calandriello, E., Fattapposta, F., Rossi, P., and Serrao, M. (2003). "Muscular cramps: proposals for a new classification." *Acta neurologica scandinavica*, 107(3), 176-186.

Pavlidis, P., Weston, J., Cai, J., and Noble, W. S. (2002). "Learning gene functional classifications from multiple data types." *Journal of Computational Biology*, 9(2), pp. 401-411.

Pejhan, S., Denroche, S. K., Frew, G. J., and Acker, S. M. (2019). "Effects of Knee Savers on the quadriceps muscle activation across deep knee bending postures." *Applied Ergonomics*, 80, 193-199.

Phillips, P. S., and Haas, R. H. (2008). "Statin myopathy as a metabolic muscle disease." *Expert review of cardiovascular therapy*, 6(7), 971-978.

Plesofsky-Vig, N., and Brambl, R. (1985). "Heat shock response of Neurospora crassa: protein synthesis and induced thermotolerance." *Journal of bacteriology*, 162(3), 1083-1091.

Pollard, J. P., Porter, W. L., and Redfern, M. S. (2011). "Forces and moments on the knee during kneeling and squatting." *Journal of applied biomechanics*, 27(3), 233-241.

Porter, W. L., Mayton, A. G., and Moore, S. M. (2010). "Pressure distribution on the anatomic landmarks of the knee and the effect of kneepads." *Applied Ergonomics*, 42(1), 106-113.

Pradhan, A., Akinci, B., and Haas, C. T. (2011). "Formalisms for query capture and data source identification to support data fusion for construction productivity monitoring." *Automation in Construction*, 20(4), 389-398.

Pušnik, I., and Miklavec, A. (2009). "Dilemmas in measurement of human body temperature." *Instrumentation Science and Technology*, 37(5), 516-530.

Qin, X., Mo, Y., and Jing, L. (2016). "Risk perceptions of the life-cycle of green buildings in China." *Journal of Cleaner Production*, 126, 148-158.

Rabanser, S., Shchur, O., and Günnemann, S. (2017). "Introduction to tensor decompositions and their applications in machine learning." *arXiv preprint*.

Razavi, S. N., and Haas, C. T. (2010). "Multisensor data fusion for on-site materials tracking in construction." *Automation in construction*, 19(8), pp. 1037-1046.

Razavi, S. N., and Haas, C. T. (2011). "Reliability-based hybrid data fusion method for adaptive location estimation in construction." *Journal of Computing in Civil Engineering*, 26(1), pp. 1-10.

Reichert, B., and Stelzenmueller, W. (2011). *Palpation Techniques: Surface Anatomy for Physical Therapists*, Thieme, New York, NY.

Reid, C. R., Bush, P. M., Cummings, N. H., McMullin, D. L., and Durrani, S. K. (2010). "A review of occupational knee disorders." *Journal of occupational rehabilitation*, 20(4), 489-501.

Richardson, R. S., Duteil, S., Wary, C., Wray, D. W., Hoff, J., and Carlier, P. G. (2006). "Human skeletal muscle intracellular oxygenation: the impact of ambient oxygen availability." *The Journal of physiology*, 571(2), 415-424.

Robertson, B. (2017). "Q: Does the Weather Affect Your Body?" *Science and Children*, 54(7), 62.

Robertson, G., Caldwell, G., Hamill, J., Kamen, G., and Whittlesey, S. (2013). *Research methods in biomechanics, 2E*, Human Kinetics.

Rock, K. L., and Kono, H. (2008). "The inflammatory response to cell death." *Annu. Rev. Pathol. Mech. Dis.*, 3, 99-126.

Roots, H., Ball, G., Talbot-Ponsonby, J., King, M., McBeath, K., and Ranatunga, K. (2009). "Muscle fatigue examined at different temperatures in experiments on intact mammalian (rat) muscle fibers." *Journal of Applied Physiology*, 106(2), 378-384.

Rosso, R. (1997). *Statistics, probability and reliability for civil and environmental engineers*, Mc-Graw-Hill Publishing Company.

Shahi, A., Cardona, J. M., Haas, C. T., West, J. S., and Caldwell, G. L. (2012). "Activity-based data fusion for automated progress tracking of construction projects." *Construction Research Congress 2012: Construction Challenges in a Flat World*, pp. 838-847.

Sharrard, W., and Liddell, F. (1962). "Injuries to the semilunar cartilages of the knee in miners." *Occupational and Environmental Medicine*, 19(3), 195-202.

Sidiropoulos, N. D., De Lathauwer, L., Fu, X., Huang, K., Papalexakis, E. E., and Faloutsos, C. (2017). "Tensor decomposition for signal processing and machine learning." *IEEE Transactions on Signal Processing*, 65(13), 3551-3582.

Simon, H. B. (1993). "Hyperthermia." *New England Journal of Medicine*, 329(7), 483-487.

Smith-Jackson, T., Artis, S., Hung, Y.-H., Kim, H. N., Hughes, C., Kleiner, B., and Nolden, A. (2011). "Safety critical incidents among small construction contractors: A prospective case study." *Open Occupational Health & Safety Journal*, 3, 39-47.

Soibelman, L., Liu, L. Y., and Wu, J. "Data fusion and modeling for construction management knowledge discovery." *Proc., International Conference on Computing in Civil and Building Engineering, Weimar, Germany*.

Soltani, M. M., Zhu, Z., and Hammad, A. (2018). "Framework for location data fusion and pose estimation of excavators using stereo vision." *Journal of Computing in Civil Engineering*, 32(6), 04018045.

Soman, R., Kyriakides, M., Onoufriou, T., and Ostachowicz, W. (2018). "Numerical evaluation of multi-metric data fusion based structural health monitoring of long span bridge structures." *Structure and Infrastructure Engineering*, 14(6), pp. 673-684.

Sorber, L., Van Barel, M., and De Lathauwer, L. (2015). "Structured data fusion." *IEEE Journal of Selected Topics in Signal Processing*, 9(4), pp. 586-600.

Spielholz, P., Davis, G., and Griffith, J. (2006). "Physical risk factors and controls for musculoskeletal disorders in construction trades." *Journal of construction engineering and management*, 132(10), 1059-1068.

Sullivan, G. M., and Feinn, R. (2012). "Using effect size—or why the P value is not enough." *Journal of graduate medical education*, 4(3), 279-282.

Tennant, L., Kingston, D., Chong, H., and Acker, S. (2015). "The effect of work boots on knee mechanics and the center of pressure at the knee during static kneeling." *Journal of applied biomechanics*, 31(5), 363-369.

Tennant, L. M., Chong, H. C., and Acker, S. M. (2018). "The effects of a simulated occupational kneeling exposure on squat mechanics and knee joint load during gait." *Ergonomics*, 61(6), 839-852.

Tennant, L. M., Maly, M. R., Callaghan, J. P., and Acker, S. M. (2014). "Analysis of muscle activation patterns during transitions into and out of high knee flexion postures." *Journal of Electromyography and Kinesiology*, 24(5), 711-717.

Thambyah, A., Goh, J. C., and De, S. D. (2005). "Contact stresses in the knee joint in deep flexion." *Medical engineering & physics*, 27(4), 329-335.

Tofallis, C. (2012). "A different approach to university rankings." *Higher Education*, 63(1), 1-18.

Tofallis, C. (2014). "Add or multiply? A tutorial on ranking and choosing with multiple criteria." *INFORMS Transactions on Education*, 14(3), 109-119.

Upfold, J., Smith, M., and Edwards, M. (1989). "Quantitative study of the effects of maternal hyperthermia on cell death and proliferation in the guinea pig brain on day 21 of pregnancy." *Teratology*, 39(2), 173-179.

Valero, E., Sivanathan, A., Bosché, F., and Abdel-Wahab, M. (2017). "Analysis of construction trade worker body motions using a wearable and wireless motion sensor network." *Automation in Construction*, 83, 48-55.

Vangelova, K., Deyanov, C., and Ivanova, M. (2006). "Dyslipidemia in industrial workers in hot environments." *Central European journal of public health*, 14(1), 15.

Varghese, B. M., Hansen, A., Bi, P., and Pisaniello, D. (2018). "Are workers at risk of occupational injuries due to heat exposure? A comprehensive literature review." *Safety science*, 110, 380-392.

Vervliet, N., Debals, O., Sorber, L., and De Lathauwer, L. (2014). "Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis." *IEEE Signal Processing Magazine*, 31(5), pp. 71-79.

Wang, D., Dai, F., and Ning, X. (2015). "Risk assessment of work-related musculoskeletal disorders in construction: State-of-the-art review." *Journal of Construction Engineering and management*, 141(6), 04015008.

Wang, D., Dai, F., Ning, X., Dong, R. G., and Wu, J. Z. (2017). "Assessing Work-Related Risk Factors on Low Back Disorders among Roofing Workers." *Journal of Construction Engineering and Management*, 143(7), 04017026.

Washington DOL & I (2018). "Rates for Worker's Compensation: 2018 base rates by business type and classification code." <http://www.lni.wa.gov/ClaimsIns/Files/Rates/2018RatesBusTypeClassCode.pdf>. (7/10/2019).

Welch, L., Haile, E., Boden, L. I., and Hunting, K. L. (2009). "Musculoskeletal disorders among construction roofers—physical function and disability." *Scandinavian journal of work, environment & health*, 56-63.

Welch, L. S., Haile, E., Boden, L. I., and Hunting, K. L. (2010). "Impact of musculoskeletal and medical conditions on disability retirement—a longitudinal study among construction roofers." *American journal of industrial medicine*, 53(6), 552-560.

Welch, L. S., Russell, D., Weinstock, D., and Betit, E. (2015). "Best practices for health and safety technology transfer in construction." *American journal of industrial medicine*, 58(8), 849-857.

Wen, W., Wu, C., Wang, Y., Chen, Y., and Li, H. "Learning structured sparsity in deep neural networks." *Proc., Advances in Neural Information Processing Systems*, 2074-2082.

Wilmore, D. W., Mason Jr, A. D., Johnson, D. W., and Pruitt Jr, B. (1975). "Effect of ambient temperature on heat production and heat loss in burn patients." *Journal of applied physiology*, 38(4), 593-597.

Winkler, R., Klawonn, F., and Kruse, R. (2011). "Fuzzy c-means in high dimensional spaces." *International Journal of Fuzzy System Applications*, 1(1), pp. 1-16.

Xiang, J., Bi, P., Pisaniello, D., and Hansen, A. (2013). "Health impacts of workplace heat exposure: an epidemiological review." *Industrial health*, 2012-0145.

Xie, K., Li, X., Wang, X., Xie, G., Wen, J., Cao, J., and Zhang, D. (2017). "Fast tensor factorization for accurate internet anomaly detection." *IEEE/ACM Transactions on Networking (TON)*, 25(6), pp. 3794-3807.

Xiong, L., Chen, X., Huang, T.-K., Schneider, J., and Carbonell, J. G. (2010). "Temporal collaborative filtering with bayesian probabilistic tensor factorization." *Proceedings of the 2010 SIAM International Conference on Data Mining*, pp. 211-222.

Xu, H., Jampala, S., Bloswick, D., Zhao, J., and Merryweather, A. (2017). "Evaluation of knee joint forces during kneeling work with different kneepads." *Applied Ergonomics*, 58, 308-313.

Yoshizawa, T., Omori, K., Takeuchi, I., Miyoshi, Y., Kido, H., Takahashi, E., Jitsuiki, K., Ishikawa, K., Ohsaka, H., and Sugita, M. (2016). "Heat stroke with bimodal rhabdomyolysis: a case report and review of the literature." *Journal of intensive care*, 4(1), 71.

Young, W., Weckman, G., and Holland, W. (2011). "A survey of methodologies for the treatment of missing values within datasets: Limitations and benefits." *Theoretical Issues in Ergonomics Science*, 12(1), pp. 15-43.

Zampporri, J. A., and Aguinaldo, A. (2017). "THE EFFECTS OF COMPRESSION TIGHTS ON DYNAMIC KNEE MOTION DURING A DROP VERTICAL JUMP IN FEMALE COLLEGE ATHLETES." *35th Conference of the International Society of Biomechanics in Sports,* Cologne, Germany, 243.

Zheng, Y. (2015). "Methodologies for cross-domain data fusion: An overview." *IEEE Transactions on Big Data*, 1(1), pp. 16-34.

Zheng, Y., Liu, Y., Yuan, J., and Xie, X. (2011 ). "Urban computing with taxicabs." *Proceedings of the 13th International Conference on Ubiquitous Computing*, pp. 89-98.

Žitnik, M., and Zupan, B. (2014). "Data fusion by matrix factorization." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(1), pp. 41-53.