ULTRASONOGRAPHY FOR THE PREDICTION OF MUSCULOSKELETAL FUNCTION

D MIGUEZ

PhD 2017

ULTRASONOGRAPHY FOR THE PREDICTION OF MUSCULOSKELETAL FUNCTION

DIEGO MIGUEZ ABAD

A thesis submitted in partial fulfilment of the requirements of the Manchester Metropolitan University for the degree of Doctor of Philosophy

School of Healthcare Science

Manchester Metropolitan University

2017

Declaration

This thesis is submitted to Manchester Metropolitan University in support of my application for admission to the degree of Doctor of Philosophy. No part of this thesis has been submitted in support of an application for another degree or qualification of this or any other institution of learning. Work relating to this thesis has appeared in the following peer reviewed publication:

A technical note on variable inter-frame interval as a cause of non-physiological experimental artefacts in ultrasound.

Miguez, Diego, et al. 2017. Royal Society Open Science

Acknowledgements

I would like to thank my Director of Studies, Dr. Emma Hodson-Tole, for all her continuous support, help and guidance throughout this project. I would never be grateful enough for all the efforts you have also put on this.

I would also like to thank my supervisors, Dr Peter Harding and Prof. Ian Loram, for their important help in all the particularities of this tricky process and good advice that made me expand my understanding in all aspects.

I would also like to thank my family and close friends for all their support during this journey, I could feel you all were with me even in the distance and I have always tried to make you feel proud of me.

A special thank you to my girlfriend Rachel, you have always inspired me to do my best and given me support and love. I wouldn't have done it without you.

I would also want to thank all the ones who made feel like home since the first moment: Ryan, Ortega, Sonia, Lynne, David, Dani, Ricard, Jorge, Miguel, Old Bedians, Oscar, Matias, Nacho, Chueko, Keto, Sara, Chris, Alba and many more.

I want to also thank my guinea pigs of my experiments, without them I would have had to go and find some others: Aaron, Adrian, Elisabetta, Kate, Maria, Ortega, Phil, Rachel, Steve and Susann.

“Subindo polo Torrillon,

e baixando pola Santeira.

Na Rigueira,

fruta e pixa canta se queira.”

Cochomouro

Table of Contents

List of Tables

List of Tables

List of Figures

List of Figures

List of Figures

Abstract

Ultrasound (US) imaging is a well-recognised technique for studying *in vivo* characteristics of a range of biological tissues due to its portability, low cost and ease of use; with recent technological advances that increased the range of choices regarding acquisition and analysis of ultrasound data available for studying dynamic behaviour of different tissues. This thesis focuses on the development and validation of methods to exploit the capabilities of ultrasound technology to investigate dynamic properties of skeletal muscles *in vivo* exclusively using ultrasound data. The overarching aim was to evaluate the influence of US data properties and the potential of inference algorithms for prediction of net ankle joint torques.

A fully synchronised experimental setup was designed and implemented enabling collection of US, Electromyography (EMG) and dynamometer data from the *Gastrocnemius medialis* muscle and ankle joint of healthy adult volunteers. Participants performed three increasing complexity muscle movement tasks: passive joint rotations, isometric and isotonic contractions. Two frame rates (32 and 1000 fps) and two data precisions (8 and 16-bits) were obtained enabling analysis of the impact of US data temporal resolution and precision on joint torque predictions. Predictions of net joint torque were calculated using five data inference algorithms ranging from simple regression through to Artificial Neural Networks.

Results indicated that accurate predictions of net joint torque can be obtained from the analysis of ultrasound data of one muscle. Significantly improved predictions were observed using the faster frame rate during active tasks, with 16-bit data precision and ANN further improving results in isotonic movements. The speed of muscle activation and complexity of fluctuations of the resulting net joint torques were key factors underpinning the prediction errors recorded. The properties of collected US data in combination with the movement tasks to be assessed should therefore be a key consideration in the development of experimental protocols for *in vivo* assessment of skeletal muscles.

Abstract

1 Overview

The mammalian nervous and skeletal muscle systems must operate together to produce the mechanical forces required to support and move the skeleton. Their appropriate functioning is therefore required for survival and quality of life as any tasks requiring active force production, from breathing and swallowing through to posture and locomotion, require appropriate mechanical force generation and transmission. Throughout history skeletal muscles have therefore been the focus of much research, evidenced by the first application of EMG to the study of voluntary contractions by Du Bois-Raymond in 1848. Yet there is still much about the *in vivo* behaviour of skeletal muscles, in both health and disease, that has yet to be explained.

In more recent times ultrasound imaging has become a popular modality with which to non-invasively study skeletal muscle properties *in vivo*. Since the first B-mode medical ultrasound measurements by Joseph Holmes in 1950, ultrasound has been used in a wide range of medical applications (e.g. echocardiography and obstetric sonography). Despite the fact that the basic operation of ultrasound has not changed fundamentally in recent years, ultrasonic instrumentation and techniques have been constantly enhanced, raising the importance, value and impact of ultrasound in medical imaging. As technology has developed and become more widely available to research scientists and clinicians important questions with regard the optimal parameters for data collection and analysis are arising, but have yet to be addressed within the current literature.

This thesis is focused on the development and validation of methodological approaches to exploit the capabilities of ultrasound technology, and to assess dynamic properties in skeletal muscles *in vivo* exclusively using ultrasound data. An important part of the project was to study the impact of the capabilities of high-end ultrasound devices, such as the accessibility to raw data and to frame rates higher

than 1000 fps, enabled by the use of a data acquisition add-on hardware managed by customisable object-oriented interfaces.

Joint torque relates a force exerted and its perpendicular distance from the axis of motion, and is representative of the contribution of all synergist muscles for a precise task. This is traditionally measured using dynamometry, which restricts the range of movements that can be assessed and limit the ability to acquire realistic data during natural movements such as stairs climbing or cycling.

Due to its versatility, Ultrasound technology has the capability of recording images of single muscles during free moving conditions by either the use of sufficiently long cables or wireless probes. In order to validate its predicting power, US derived joint torque values should be compared with dynamometer obtained, assuming the latter as the ground truth.

The validation of US data from individual muscles for whole joint torque prediction would reduce the need of using multiple ultrasound probes for representing the performance of a joint, and would also enable the study of the individual contribution of synergist muscles to the overall joint torque production. This would be a significant step forward in the ability to study and understand human movement.

The work presented in this thesis is highly multi-disciplinary, spanning skeletal muscle anatomy and mechanics; ultrasound imaging technology; computer vision and data prediction. Chapter 2 provides detailed background information and a review of relevant literature relating to these key topics, including contextualisation of the use of ultrasound technology for the analysis of the neuro-musculoskeletal system and the new emergent methodologies, applicable to ultrasound image data that may be able to assess properties of muscles.

My work started with the development of custom-made interfaces for controlling ultrasound transmission and reception parametrisations, in order to establish and

implement a protocol for the collection and analysis of ultrasound data from human participants. Particular efforts were put in the study of the temporal behaviour of ultrasound devices in order to assure full synchronisation of the ultrasound data with other recorded signals (e.g. dynamometer based joint torque, EMG) (Miguez et al., 2017). Chapter 3 describes all the processes completed to extract dynamic features of muscle movement and shape change from recorded ultrasound data for subsequent use in the prediction of the mechanical output of muscle movements, namely joint torque. Analysis of pilot data indicated the potential of an active shape model (ASM) approach to image segmentation for automatically locating the muscle of interest in the image, and showed the suitability of global optical flow algorithms for quantification of recorded muscle movements. Five inference algorithms, namely: linear regression; quadratic regression; cubic regression; kernel support vector machine; and artificial neural networks were identified and implemented for prediction of net joint torque from ultrasound derived data and the principles underpinning their implementation are also detailed in Chapter 3.

The experimental setup and protocols used for acquisition and analysis of ultrasound data are detailed in Chapter 4. The focus of the study was to address four key research questions:

- Do US images of a single muscle (GM) contain enough information to accurately predict net ankle joint torque?

- Does the US raw data precision (16 bit) provide information that 8-bit compressed data does not?

- Is the temporal resolution (inverse of frame rate) an important factor for extracting dynamic information from US data? Does it depend on the movement task performed?

- Which data prediction techniques provide the most accurate predictions of net joint torque from US data?

In order to address these research questions the developed data acquisition protocol was implemented to collect US, EMG and dynamometer data from the lower leg of 9 healthy adult participants. In order to provide increasing complexity of muscle movement tasks, participants performed three types of muscle movements: passive ankle joint rotations; isometric contractions and isotonic contractions. Data were collected at two different frame rates, 32 and 1000 fps, in order to study the influence of temporal resolution on torque prediction. Two data precisions, 8 and 16-bits, were also obtained to enable analysis of the impact of data precision on torque prediction errors. All collected images were analysed to extract information pertaining to muscle shape, displacement and size and these measures were used as inputs for each of the five inference algorithms, implemented to predict the resulting net joint torque. The error of the predicted joint torques was calculated using root mean square error and the results from each of the movement tasks are reported in Chapters 5 – 7.

Chapter 8 reviews and interprets the findings of the presented work, answering the research questions in light of the differences in the results obtained from the different movement tasks studied. Suggestions for future work are proposed, including: strategies to compensate for the limitations identified in the presented work; improvements of the proposed methodology to enable implementation in real time; and consideration of additional techniques to facilitate joint torque prediction in more complicated functional tasks.

2 Literature Review and Background

The main objective of this chapter is to provide detailed background information of the main topics of this thesis. It is divided into four main sections: Section 2.1 introduces skeletal muscle physiology and anatomy and contextualises the use of US and EMG technology for their study. Section 2.2 details processes of US transmission and acquisition defining terms related to the use of raw data (RF), US image formation, wave types and temporal US features. Section 2.3 explores the post-processing of US data in terms of image segmentation and pixel tracking approaches and, finally, Section 2.4 considers techniques available for prediction of external data (e.g. joint torques) using US imaging derived information.

2.1 Skeletal Muscle Anatomy and Physiology

2.1.1 Introduction

The musculoskeletal system plays a key role in enabling voluntary movement, underpinning all aspects of daily living including locomotor and postural functions. Skeletal muscles produce movements by generating torques, acting about joints, to move the levers of the skeletal system. The torques are generated through active and passive properties of the contractile element of the muscle, the muscle fibre, transforming chemical energy into mechanical energy by producing tension which results in the generation of force.

From a functional perspective, muscles can be grouped into 3 categories depending on their role during a given movement.

- Agonist Muscles provide the main concentric force during a movement, therefore they are responsible for causing the desired joint motion. Biceps brachii during an elbow flexion constitutes an example of an agonist muscle.

- Synergist muscles: Muscles that assist and contribute to the agonist muscle performance but don't directly apply torque. Within this group we can distinguish between Fixators, that limit movements of the agonist muscle that don't contribute to the desired movement in specific joint movements; and Neutralizers, that also avoid undesired actions of the agonist muscle but unlike fixators they are usually associated with a pulling movement. Examples of these are when the trapezius and rhomboids prevent scapula movement towards the inner upper body during an elbow flexion or the triceps preventing the elbow to flex when the biceps rotates the forearm in order to face the palm of the hand up.

- Antagonist muscles: These muscles contract eccentrically when the agonist muscle contracts concentrically. Their functions include slowing joint motion in order to avoid injuries or helping to increase the precision of the movement. Triceps is the antagonist muscle during an elbow flexion.

A single muscle can fall into different categories depending on the task, for example triceps is agonist during elbow extensions but is an antagonist during elbow flexion (Figure 2.1).

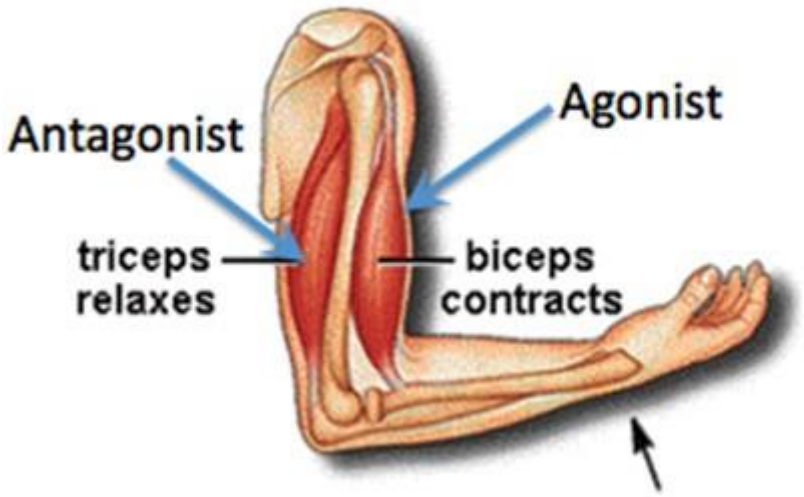

Figure 2-1 Illustration of muscle groups during an elbow flexion from (Sterling Hot Yoga 2014). Biceps and triceps are the agonist and antagonist muscles respectively.

Types of muscle contractions

There are five different ways of classifying muscle contractions, distinguishing between the direction of change in length of the muscle tendon unit, and whether velocity or tension is constant.

- Concentric contraction: Muscle tendon unit shortens while the fibres contract.

- Eccentric contraction: Muscle tendon unit lengthens while the fibres actively producing tension.

- Isometric contraction: Muscle tendon unit length stays constant, while the fibres actively generate tension. Isometric contractions may occur when the resisting tension is not overcome (e.g. pushing a wall) or when holding an object. Muscles involved in posture control typically perform isometric contractions.

- Isokinetic contraction: Muscle changes length during the contraction but velocity of movement is constant. Isokinetic contractions can be either concentric or eccentric.

- Isotonic contraction: The elicited torque and therefore the tension during the movement is constant. Isotonic contractions can also be eccentric or concentric.

2.1.2 Muscle structure

Skeletal muscles are encased by the epimysium, and connected to bones via elastic tissues called tendons (Figure 2.2). Within the epimysium, muscle fascicles can be found surrounded by the perimysium, another connective tissue. Muscle fascicles are composed of bundles of muscle fibres that are connected alongside each other and encased by the endomysium. Usually muscle fibres are as long as the muscle they are part of, ranging from less than 1 mm for the Stapedius (Seikel et al. 2009) to more than 500 mm for the Sartorius muscle (Harris 2005).

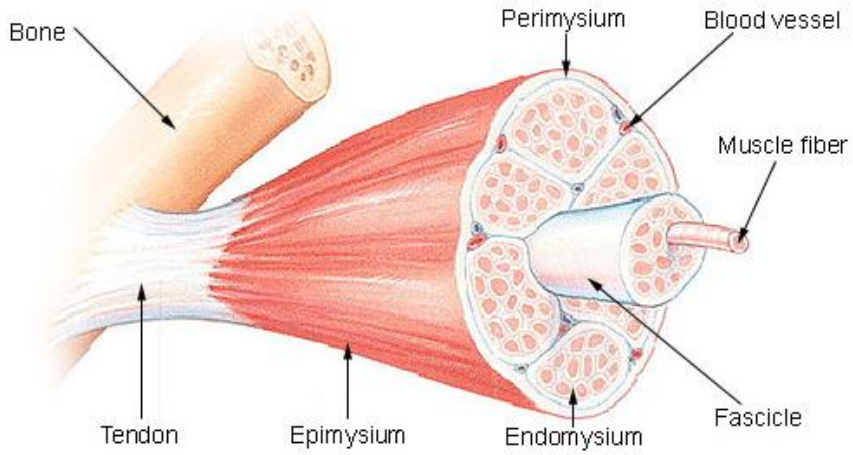

Figure 2-2 Architecture of a skeletal muscle showing its attachment to the bone via the tendon that at the same time is connected to the muscle through the epimysium from (National Cancer Institute n.d.).

Muscle fibres (also known as myofibers) are mostly made of contractile structures called myofibrils that consist of repetitive structures along their length known as sarcomeres, which are separated one from another by the Z disk. The sarcomere is

the fundamental unit of muscle structure and (alongside a number of other proteins) contain three protein myofilaments that are directly involved in the contraction of muscles: thin filaments (actin) and thick filaments (myosin) intercalated, and elastic filaments (titin) that connect myosin to the Z disk (Figure 2.3).

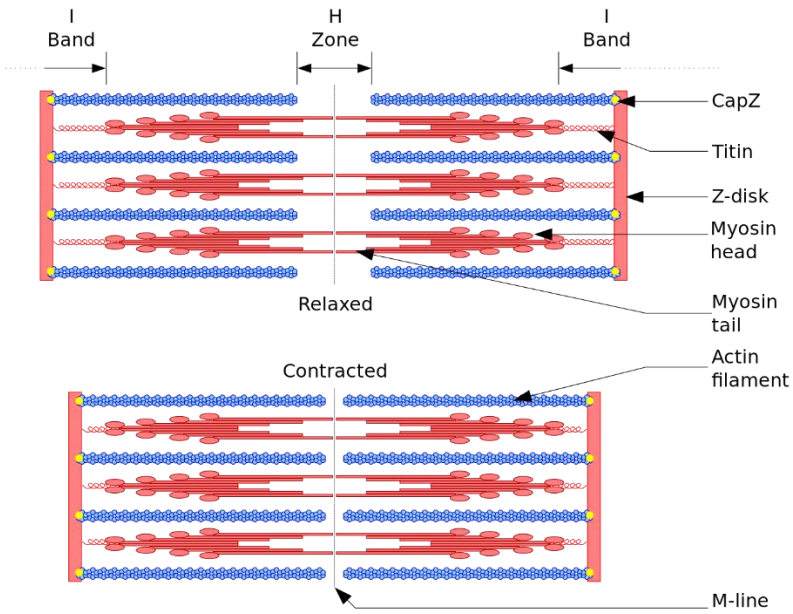

Figure 2-3 Sarcomere structure showing the position and binding of thick and thin filaments when contracted (bottom) and when relaxed (top) (Richfield 2014).

These elemental muscle structures contract when triggered by electrical impulses originated in the brain that travel via the nervous system. The motor unit (MU) is the fundamental unit of the neuromuscular system and is composed of a single motor neuron, its axon and all muscle fibres it innervates (Sherrington 1906). Obtaining information on properties of individual MUs provides a means of: i) studying how MUs are controlled by the neural system to produce force or coordinated movement patterns; ii) determining the degree of dysfunction caused by pathologies such as cerebral palsy and Parkinson's disease or as a consequence of brain insults e.g. stroke; iii) providing diagnostic insight into neurodegeneration associated with a range of neuropathies and myopathies; iv) estimating numbers of MUs in a muscle.

Traditionally MU properties are studied using EMG, a technique that measures the electrical potential of contracting muscles fibres via surface or intramuscular electrodes. The use of this technique however has some implicit limitations (e.g. surface EMG can only be used for superficial muscles; intramuscular is invasive, not representative of all MU fibres) that restrict the information EMG can provide.

2.1.3 Force production and factors that affect it

The *Sliding filament theory* (Huxley 1957) explains how force is produced by the contraction of muscle fibres. According to it, the heads of the myosin attach to the actin filaments forming a cross-bridge, and then pull them in order to return to the equilibrium state in a similar way to a spring. This way a contractile force is generated. This situation is replicated along the length of the different sarcomeres producing muscle fibre shortening and therefore a muscle contraction. Recently, Herzog and co-workers have shown that titin also plays a key role in this process by shortening its length and thus pulling the myosin against the z disks, resulting in an increase in the overall generated force (Herzog et al. 2015).

According to the sliding filament theory, a contraction is triggered when a nervous impulse arrives at the neuromuscular junction. When two or more pulses arrive two situations can occur depending on the time between each pulse: If the frequency of the nervous impulses is high enough (between 15 and 30 Hz depending on the muscle) the muscle is stimulated at the exact time relaxation was meant to occur and therefore the muscle maintains the initial contraction. When the period between consecutive impulses is shorter, the twitches overlap one another increasing the overall generated force. This phenomenon is known as tetanic contraction (Hill & Hartree 1921).

The contracting force of a muscle is proportional to the amount of motor units recruited within the muscle pool and on the frequency of the nervous impulses

(period between consecutive stimulations), and is transmitted to the bones via the aponeurosis in order to translate that contractile force into a movement.

As explained by (Gordon et al. 1966), muscle force depends on the number of cross-bridges formed by actin and myosin filaments during an activation, and on the elastic particles (titin) (Herzog et al. 2015) that increase the force during passive movements (Figure 2.4). When the myosin and actin totally overlap the maximum number of cross bridges are formed and therefore a maximum of the active component of the muscle tension is produced. This optimal length is also known as the resting length, and the active component of the muscle tension decreases when moving away from it (when stretching or compressing the muscle from the resting length). In parallel to this, when a relaxed muscle is pulled it offers a resisting tension that when close to the resting length is negligible but increases exponentially (becoming predominant) with greater stretching lengths (Keynes & Aidley 2001).

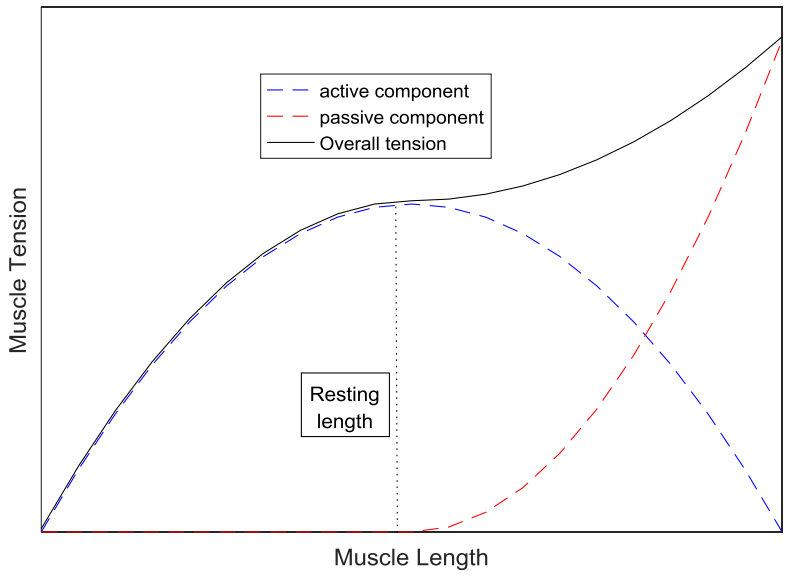

Figure 2-4. Simplified Active and Passive length-tension relationship for fibre bundles based on (Gordon et al. 1966).

The behaviour observed in sarcomeres can be extrapolated to the whole muscle. Thus, the length–tension relationship affects the joint angle-torque relationship, but

there are other factors to consider, such as the muscle-tendon interaction, the muscle cross sectional area and the moment arm length of the muscle; making the length tension relationship more complex to study *in vivo*. Figure 2.5 provides an illustrative representation of the joint torque exerted when varying joint angle within its range of motion. The joint angle for which the torque is maximum is known as the optimal angle or peak torque angle.

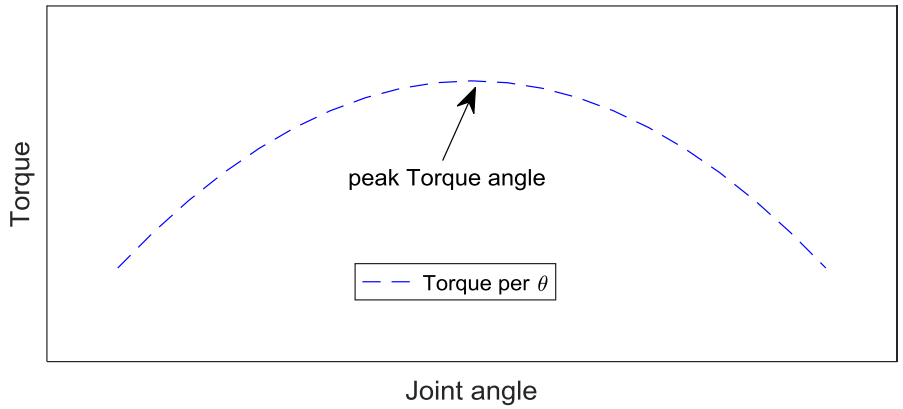

Figure 2-5. Simplified Torque vs. Angle relationship based on (Gordon et al. 1966).

2.1.4 Methods of Studying *in vivo* Skeletal Muscle Function

Ultrasound imaging was first used to study skeletal muscle mechanics *in vivo* in the 1990s where changes in the angle of fascicles relative to the long axis of the muscle (pennation angle) was measured during muscle contraction (Ito et al. 1998; Narici et al. 1996) and passive joint movements (Kawakami et al. 1998). Since then, work on the analysis of skeletal muscle via ultrasonography can be found in the literature, with manual efforts to estimate fascicle strain (Lopata et al. 2006), moment arm length (Maganaris 2004), tendon parameters (Li & Tong 2005), and sensitivity comparisons between electromyograms and ultrasonography (Hodges et al. 2003).

Relevant work was focused on the analysis of the contribution of the calf muscles to ankle torque in both isokinetic conditions (Arndt et al. 1998) and during natural movements (Hamner et al. 2010). The contribution of the triceps surae muscles (Gastrocnemius Medialis and Lateralis and Soleus) to plantarflexion was studied for different knee angles (Kawakami et al. 1998; Kawakami et al. 2000), observing an

increase in the GM muscle of 43% EMG activation levels during full extension as opposed to 90 degree flexion (Shinohara et al. 2006). The specific contribution of GM muscle during MVC for this work was the 33% of the overall EMG of the triceps surae. Related research showed the contribution of ankle plantar flexors during submaximal isometric tasks (Masood et al. 2014), concluding that GM and GL contribute 29% and 19% respectively to plantarflexion torque in terms of EMG activation, also corresponding to 29% and 25% respectively in terms of glucose uptake, during 30% MVC isometric contractions. Further work also compared the relationship between ankle torques during plantar flexion and inversion, suggesting its potential role outside the sagittal plane for lateral stabilization (Vieira et al. 2013), which resulted in a contribution of 13% of the plantarflexion peak torque during ankle inversion.

Manual image analysis methods are routinely used to quantify human-measurable parameters, such as muscle thickness, fascicle orientation/curvature and fascicle length. While these methods are manual, they have provided many significant insights into skeletal muscle function *in vivo* and therefore make the case for the breadth of information about dynamic skeletal muscle accessible via ultrasonography.

Early computational image analysis techniques were geared towards the analysis of motion fields generated from ultrasonography sequences during muscle contraction (Loram et al. 2006; Liu et al. 2007; Shi et al. 2007). After the establishment of manual analysis protocols the development of alternative computational approaches followed; these can make absolute estimates of muscle state, rather than tracking changes in state over time. There are many notable works which aim to estimate muscle fascicle orientation/curvature and/or length from single images (Rana et al. 2009). The most promising work on the quantification of fascicle orientation used a vessel enhancement filter to enhance the anisotropic (tube-like) textures (which correspond with, and are assumed to be fascicles) in images of human calf muscle with two different techniques. The Radon Transform was used to quantify global fascicle curvature and discrete oriented wavelets were used to quantify local fascicle

orientation from ultrasound images of the vastus lateralis and gastrocnemius muscles (Rana et al. 2009; Darby et al. 2012). To date all previous work applying US to study the musculoskeletal system has made use of US devices typically devised for the use in clinical environments. As interest grows in developing computational ultrasound image analysis approaches to facilitate study of the musculoskeletal system more work is required to understand the characteristics of US data that should be considered when designing new experimental protocols and how different US properties may influence study findings.

2.2 Ultrasound

US is one of the fastest and cheapest technologies in medical imaging. Key features of the processes involved acquiring ultrasound data (e.g. data precision, wave transmission and temporal features), and their influence on previous studies of the musculoskeletal system are detailed in this section in order to provide an initial vision of the full potential of US technology for the study of the musculoskeletal system. Definitions of key concepts and an explanation of the different processing steps for obtaining a B-mode US image from the raw data is therefore provided.

2.2.1 US Definitions

In the context of ultrasound imaging the term raw data refers to the original single element (including virtual elements) data without any transformation applied. These data still contain the carrier signal, defined as the higher frequency signal in which the input signal is modulated, for reducing wave deformation during propagation (Figure 2.6-left). Radio frequency (RF) data (a name that harks back to the origins of US imaging as an extension of RADAR), also known as beamformed raw data, is ordered by its correct element distribution in the x-axis and still contains the carrier signal (Figure 2.6-right). This transformation is automatically done in conventional US devices within the acquisition software. Conventional ultrasound devices only allow the use of a compressed version of the single element RF data for posterior analysis. Some devices more recently made commercially available to researchers

do however allow access to the original data before any filtering or modification is applied.

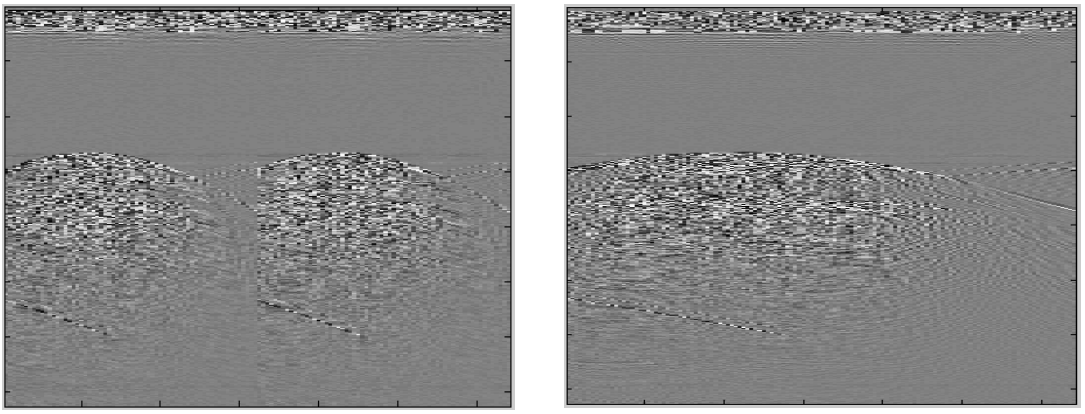

Figure 2-6 2D image of US raw data (left) and RF-data (right) data of a silicon ball in a water tank. Each image represents a 6x6 cm cross sectional area.

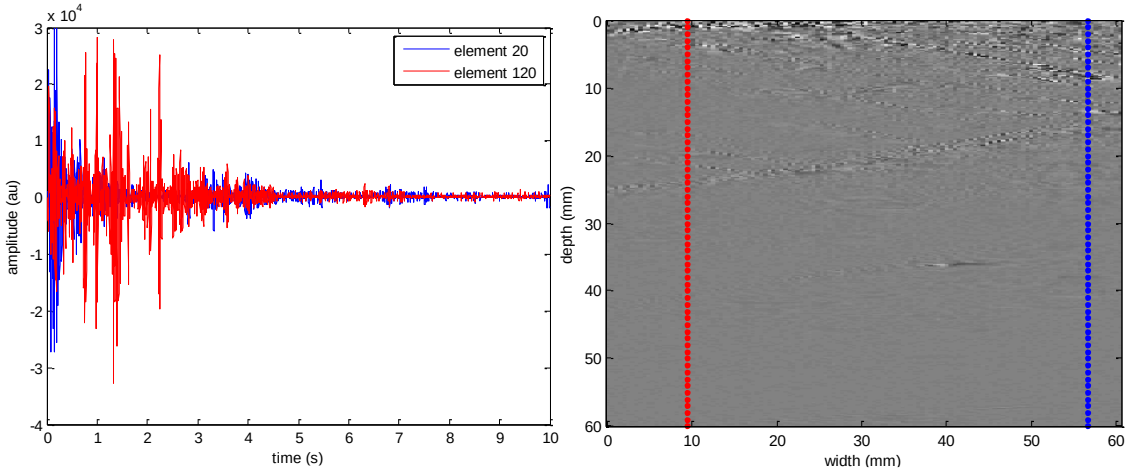

Figure 2-7 RF beamformed data of the GM muscle of a participant. Left: Signal amplitude vs Time of the US element number 20 (blue) and 120 (red). Right: 2D representation in width and depth, showing the relative position of the elements 20 and 120.

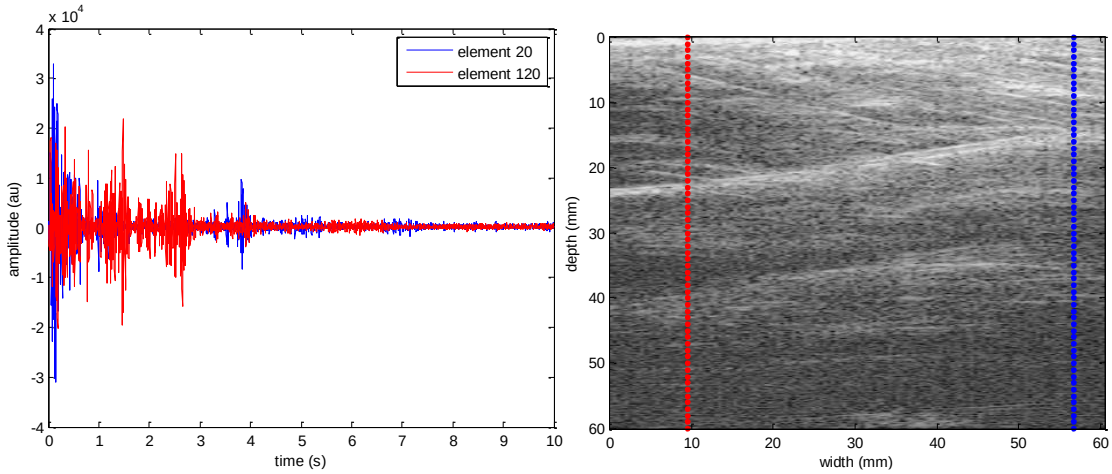

Figure 2-8 IQ demodulated data of the GM muscle of a participant. Left: Signal amplitude vs Time of the US element number 20 (blue) and 120 (red). Right: 2D representation in width and depth, showing the relative position of the elements 20 and 120.

The B-mode image is the video compressed IQ images. It is the two-dimensional representation of the IQ demodulated element data. Its values usually vary from 0 to 255 (8 bits). In conventional US devices, B-mode images can be stored and posteriorly accessed in different video formats, usually under the Digital Imaging and Communications in Medicine (DICOM) standard, that include JPEG lossy compressions. Explanations regarding the US data format and compressions can be found below and in chapter 4.2.

2.2.2 US data precision and compression

As the post-processing and compression and storage of US data are typically implemented in "black box" function, i.e. functions for which the implementation is impossible to access by the user, and because part of the data is lost during those processes; it is not possible to reverse them in order to obtain the initial data, therefore the compressed video data are typically the only available information for many who study musculoskeletal function. The availability of raw data from some US devices does, however, enable the use of significantly more precise data and personalised filtering. This allows users to not only increase the full precision of the data (from 8 to up to 16 bit) but to also apply their own post-processing, adapted to the specific purpose of the measurements. For example, a jump from 8 to 12-bit data precision would increase by 16 times the resolution.

To date, only four investigations have compared the differences between working with the original US raw data and the video compressed data. A resume of the methods, results and gaps of knowledge observed is detailed below:

In-vivo perfusion quantification by contrast ultrasound: Validation of the use of linearized video data vs. raw RF data (Rognin et al. 2008)

This paper used contrast ultrasound, which involves the application of a substance (usually containing bubbles) that increase the reflection of the US waves in order to increase the contrast of the received echoes. The technique was applied to a rabbit kidney and an estimation of the perfusion parameter (indicative of the concentration of contrast agent present in an US image region of interest) was made for both 16-bit RF-data and 8-bit greyscale images.

Their comparison of RF vs. video data consisted of an estimation of a model function based on the obtained data, with a significant deviation reported to occur. However, there is no image quality comparison between raw data and 8-bit video. Conclusions are mainly recommendations about what settings (e.g. dynamic range and gain) must be chosen to obtain US images that provide useful results for the techniques

used. The lack of comparison between image precisions is however a missed opportunity in this work.

High-speed Cineloop Ultrasound vs. Video Ultrasound Tongue Imaging: Comparison of Front and Back Lingual Gesture Location and Relative Timing (Wrench & Scobbie 2008)

A study of intergestural timing between tongue root retraction and tongue tip rising is provided in this paper. US tongue images were supposed to be collected using an Ultrasonix RP system, a device that can be triggered remotely via Ethernet (the delays in the US acquisition times may be quite variable and dependent on the internet traffic). The hardware trigger of this system consisted of a 25 ns duration pulse, almost impossible to acquire with conventional acquisition cards. Hardware for stretching the trigger impulses from 25 ns to 11 ms was built and fed into an analog-to-digital card on the controlling PC. Despite the fact that an 11 ms duration trigger pulse only allows to detect up to 45 frames per second (fps) without aliasing according to Nyquist (maximum detectable frame rate = 1/(2*11ms), the paper reports that the system operated at 98 fps. This suggests an overlap must have occurred between consecutive frame triggers and, if correct, would imply a serious flaw in the presented paper.

Their results showed the tongue movement during the pronunciation of the /ele/ phoneme. An important variation in the time of the observed events (tongue raise and fall) was observed between cineloop (RF) and video data due to the differences in frame rate, being 100 fps for the US raw data and 98 fps for the frame grabber card output; showing a lack of synchronisation in the measurements that can imply a non-direct comparison of the same phenomena in time. Furthermore, it was stated that for one of the systems (video), a tenth of the collected frames had unchanged image parts compared with consecutive frames, implying that at least 10% of the comparisons were performed using images representative of different events in time.

Comparison of Echo-Power Estimation Using Linearized Video Data and Raw Data for Dynamic Contrast-Enhanced Ultrasound (Payen et al. 2012)

In a similar way to the first paper (Rognin et al. 2008), contrast agent US was collected from a subcutaneous pancreatic adenocarcinoma model. Comparisons of echo-power estimations (pixel brightness observed in the US images) were made between three different software programmes: i) manufacturer specific (CHIQ, Toshiba) applied to raw data, and two applied to video compressed data: ii) a commercially available system called VueBox and iii) a system developed in-house. US data collection was first triggered for the compressed data by pressing one button, and then the raw data started by manually pressing another button, meaning that there was no synchronisation between images collected for each precision.

VueBox and in-house software were both used to estimate echo-power using the lower precision data and the results compared with those from the manufacturer specific software that used raw data precision (assumed as the ground truth data). Due to the fact that the echo-power values resulting from the two test software lacked any units, a so called "Linearization" process was applied to the video data in order to "minimise" errors between estimated powers; representing an attempt to equalize the results of both methods (RF and video). Despite the fact that this work concluded that no important differences were observed between raw and compressed US data results, the total lack of synchronisation together with the biasing involved in the linearization process diminish the certainty of any findings that could be extracted from this work.

Echo-power estimation from log-compressed video data in dynamic contrast-enhanced ultrasound imaging (Payen et al. 2013)

This work is a continuation of the previous paper (Payen et al. 2012) again using the three software programmes (one for raw data and two for video data) for estimations of the echo-power in US images. A deeper comparison of the three programmes across the different data precisions was made, although the problems

with the data synchronisation were stated not to have been resolved. In fact it was stated that "*sequences did not match spatially (pixel size and coordinates) or temporally (sequences' beginning and end, intervals between consecutive images)*".

An approach for synchronisation of the raw data and video compressed data was explained in the Appendix of the paper. An estimation of the time delay between RF-data and video data was obtain based on image registration, a process of aligning different images of the same scene using geometric transformations. This delay was assumed constant and applied to every video compressed image. This work concluded that both software that used video compressed data provided results as accurate as the one that used raw data.

US data precision Overview

From the reviewed literature in this section, it is concluded that none of the papers have a proper and accurate synchronisation of the frames with the external data (in fact it is an issue that might condition some results). Furthermore, results regarding RF vs. video US data comparisons are not rigorous enough to identify the advantages of using one or the other imaging method. According to these conclusions, a comparison of the advantages observed between using raw data and video data with a fully synchronised experimental setup is needed, and is a gap in current understanding that is aimed to be addressed within this thesis. In order to do this it is important to have an understanding of the processes involved with ultrasound data acquisition and these are summarised below.

2.2.3 Wave Transmission

There are different factors that affect the generated US wave deforming it in both amplitude and frequency. Attenuation is defined as the loss of amplitude in a signal due to different phenomena that can be classified into absorption, refraction and dispersion, distinguishing within the last one between reflection and scattering.

According to the Snell's law, when an US wave propagating strikes the intersection between two media with an angle, part of it will reflect to the same and another part of it will refract with an angle proportional to the propagation velocity ratio between both media. The propagation velocity also depends on elastic modulus and density of the medium and therefore both parameters impact the angle of the refracted ray.

Absorption is defined as the process of transforming vibrational energy into other type of energy (mainly thermal) during propagation in viscous materials and, together with scattering, contributes to the attenuation in Ultrasound propagation. The main factors that influence absorption are the viscosity of the medium, its relaxation time and the acoustic field frequency.

The temporal resolution of US imaging can be increased from the conventional 25 – 80 frames per second to ultrafast rates of 35,000 – 40,000 frames per second by using plane waves, together with parallel processing of the individual element data (beamforming). Nowadays measurements of blood motion, tissue motion, brain activity and contrast agent dynamics can be precisely acquired by using ultrafast ultrasound. The use of these techniques, together with advances in the development of graphical processing units (GPUs), may allow the implementation of real-time high-quality B-mode imaging with dynamic focusing in the near future (Tanter & Fink 2014).

Unlike conventional focused ultrasound, ultrafast ultrasound is based on the transmission of unfocused plane waves all along the transducer elements. This technique allows an important increase in the frame rate, but the quality of the obtained images is reduced due to the lack of axial resolution and contrast (Figure 2.9).

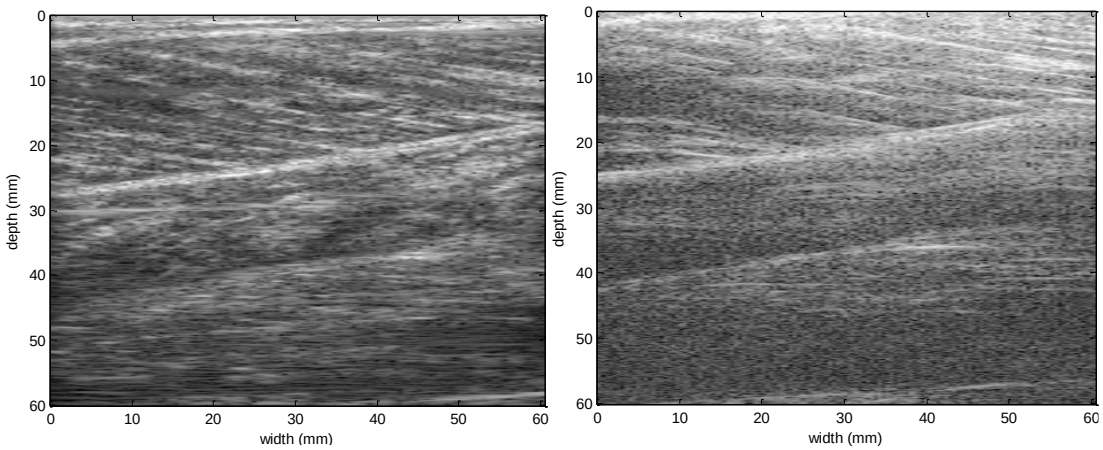

Figure 2-9 Conventional focused (left) (30mm focal depth, 128 elements) vs. single Plane Wave (right) US B-mode images of the GM muscle of one participant.

The latest state-of-the-art US devices are almost fully customisable in terms of beam transmission characteristics, including the application of individual time delays to the transmission from each element of the US array. Calculating the inclination and curvature of the wave, and its translation into time units, allows the transmission of a great variety of wave shapes (Montaldo et al. 2009). Using this approach, angled plane waves can be easily implemented to increase the axial resolution, signal to noise ratio and contrast of the obtained B-Mode images by sacrificing the frame rate (resulting frame rate ≈ maximum frame rate/number angled waves) (Salles et al. 2014). Figure 2.10 shows the resulting image quality increase and associated frame rate decrease as the number of angled plane waves are increased, compared with a conventional focused transmission in a human tissue-mimicking gelatine phantom used as reference.

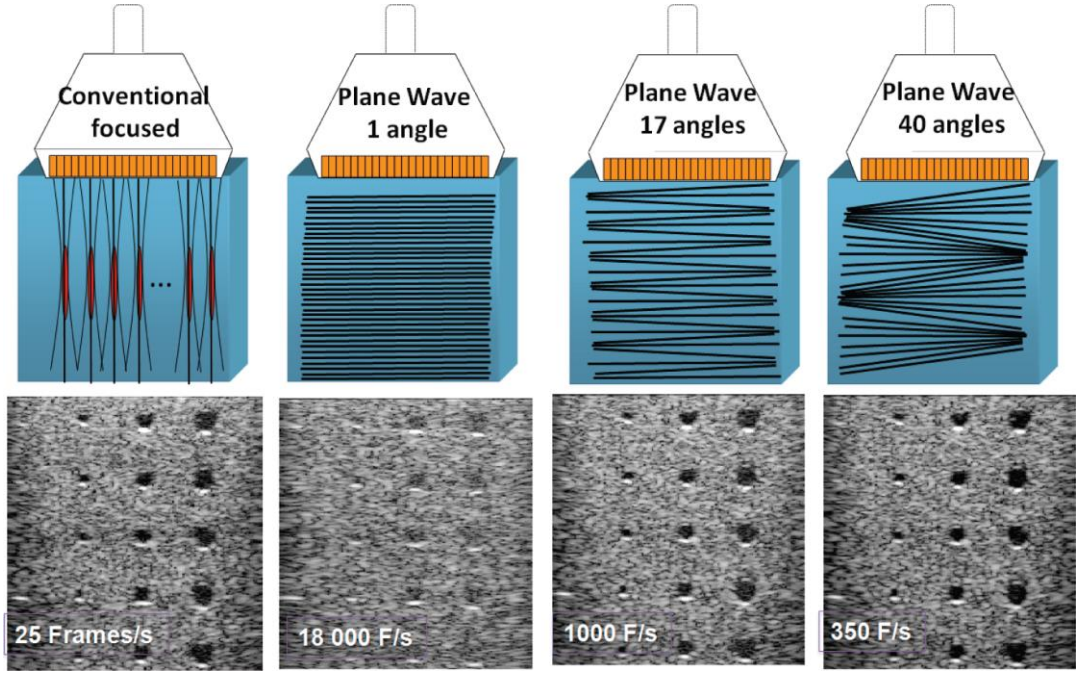

Figure 2-10 Comparison of Conventional focused, plane wave and 17 and 40 angled waves respectively US transmissions when imaging a calibration phantom Source:(Tanter & Fink 2014).

Taking advantage of plane wave imaging approaches, different reports have focused on the study of the musculoskeletal system, obtaining good quality images with the implicit increase in frame rate (compared with conventional focused waves). The work has included study of muscle-tendon unit shortening (Hauraix et al. 2013) and stiffness (Hauraix et al. 2015), gastrocnemius injury recovery monitoring in rats (Leineweber & Gao 2015) and accurate measurement of the delay between muscle activation and contraction (Nordez et al. 2009; Hauraix et al. 2013).

2.2.4 Temporal US features

Frame rates above 1,000 fps have permitted the enhancement of existing techniques such as shear wave elastography and Doppler and contrast imaging; and the emergence of new imaging tools. The use of plane-wave compounding implies a higher amplitude transmitted signal, allows an important increase of the image quality at the region of interest (ROI) and permits a faster generation of focused and shear waves (crucial in elastography applications) (Tanter & Fink 2014).

US frame rates greater than 1,000 fps provides temporal resolutions similar to EMG, allowing quantification of the alterations in the amplitude and bandwidth of surface EMG which occur with changes in fibre orientation and length quantified with US imaging. Additionally, ultra-fast US investigations have studied the delay between muscle activation and force production (electromechanical delay), used to assess muscle function (Nordez et al. 2009), and the quantification of microscopic muscle movements, such as involuntary contractions, making easier to achieve accurate results of 3-D tissue velocities (Deffieux et al. 2008).

A potential downside of the increases in US frame rates in recent years is that errors or inaccuracies in the time stamping of the collected US frames can bias the results. Due to the increase in the number of frames collected, systematic "per frame" errors would have a much higher impact on the overall synchronisation. Furthermore, the processing and memory requirements for the controlling PCs also increase exponentially. As many physiological studies involve US image analysis in combination with external inputs, such as isokinetic dynamometers (Hauraix et al. 2015) or instrumented footplates (Darby et al. 2012), accurate synchronisation of the experimental setup becomes critical. Despite this, relatively few studies which use higher frame rates make use of external "frame triggers" (electronic signals generated after every US transmission) for correct synchronisation of the acquisition system (Hauraix et al. 2013). In most studies, no explicit information regarding the timing accuracy of the collected images nor information on the frame rate variability are reported. This means that it is not stated if the US frames arrive with a constant or a variable delay or if any corrections have been applied to the obtained data for mitigating any time differences. Due to these gaps in the literature, and possible lack of awareness of the importance of accurate synchronisation, a detailed investigation including time stamping of US images for both conventional and ultrafast ultrasound is required to underpin appropriate application of such devices for the study of the neuromuscular system, work which has been completed and can be found in Section 3.1.

2.3 Ultrasound Image Processing

2.3.1 Segmentation Techniques

Many studies of the neuromuscular system focus on studying static biological properties (such as muscle volume or muscle fascicle tissue and muscle belly thickness) as well as tissue dynamics (e.g. changes in fascicle orientation, strain and strain rate) (Darby et al. 2013; Angelini & Gerard 2006). To do this, identification of the region or muscle of interest is required. One approach to completing this task computationally is to use image segmentation. This can be considered a tricky and sometimes subjective stage of the analysis due to the variability present in US images, which can be classified into two different groups:

- Acoustic behaviour in heterogeneous and homogenous media: Grainy or diffuse backscatter and signal loss that is usually present throughout the entire image (sometimes becoming more noticeable with increased depth or when encountering tissues with large stiffness differences). Most common issues include speckle noise, signal attenuation or electromagnetic background noise.

- Acquisition related variability: Experimental conditions, often operator dependent, that affect the repeatability of the ultrasound images when measuring the same conditions. These cannot be corrected or compensated after the experiment but are usually easy to identify. Deformation due to the pressure applied to the muscle/subcutaneous fat layer via the probe, probe angle, position and transmission frequency are characteristic examples.

A reliable segmentation method should take into account these inconsistencies, and be robust against them in order to be suitable for application to different participants and tasks. Initial work regarding the extraction of information from US data was based on the analysis of landmark locations identified manually from US images (Aggeloussis et al. 2010) for muscle tracking when lengthening, but these methods

are highly time consuming and include subjective factors associated with operator decisions. This led to the appearance of semi-automated methods, methods that needed Region of Interest (ROI) definition or "guidance" in their initialisation in order to improve their segmentation performance when tracking muscle thickness and pennation angle (Kawamoto et al. 2014; Mannion et al. 2012; Darby et al. 2012). Some of the ways of approaching this problem are therefore described below:

Active Shape Model (ASM) is an image processing technique based on the application of statistical models of shape and appearance to represent objects in images, and in recent years has become one of the most prolific segmentation techniques applied to medical images (Noble & Boukerroui 2006). It combines active contour optimization with statistical analysis using machine learning techniques.

The aim of segmentation is to find subsets of images that contain semantically important regions. ASM approaches make use of a prior model representing the shape of regions that are expected to be found in the image. They typically attempt to find the best position match between the model and the data contained within a new image. Once obtained the location of the best match, one can confirm if the actual target is present within that region.

As an initial step, sequences of collected ultrasound images are used as training data. These images should be representative of the different possible shapes and angles, likely to be found in other image sequences, in order to implement a method that accounts for a wide data variability. The multidimensional values of the position of landmarks that describe easily identifiable shapes are annotated in the images used for the training process of the ASM. For the purposes of studying skeletal muscle, the aponeurosis bounds the muscle borders and, in many cases, provides a clear feature which can be manually identified and annotated. The resulting landmarks are aligned (figure 2.11) and then superimposed condensing all the variability observed in the training data.

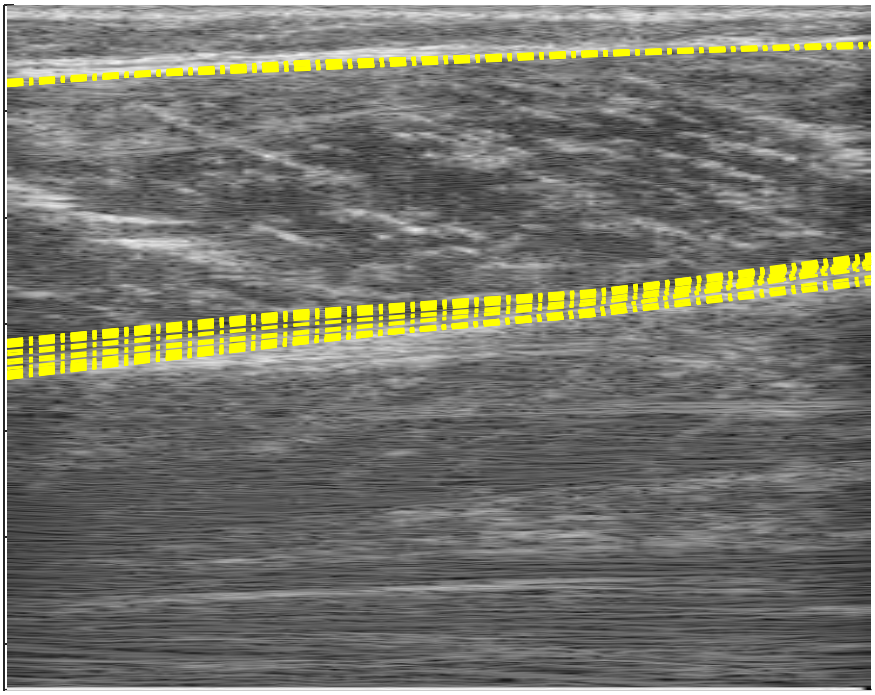

Figure 2-11. B-mode image of the GM showing the superposition of the superficial and deep aponeurosis landmarks in 26 different situations.

The superposed annotated aponeurosis landmarks are expressed as a mean value and a standard deviation, forming a Point Distribution Model (PDM). Information regarding the aponeurosis positioning within the PDM is then condensed and expressed in a lower dimensional expression by the use of Principal Component Analysis, and the resulting information (the provisional ASM) is applied to new ultrasound images, and can be compared with ground truth values in order to validate the model (e.g. manually labelled images).

The main disadvantages associated with the implementation of ASMs reside on the use of only shape constraints that lack information regarding textures across the target object and the presence of noise that can affect the definition of shapes. In order to compensate for these, and allow ASM to become a more robust tool, solutions such as increasing the width of the search profile to reduce the effect of noise; and grouping landmarks and altering the direction of the search profile to avoid shape distortion in the search procedure have been applied. Further to these Cunningham (2015), developed a PDM with the best pose (translation, scale and rotation) and shape parameters are found to match a model instance to a new set

of image points by minimising the sum of square distances between corresponding model and image points. This approach was applied to study skeletal muscles and is the implementation that has been used within the work presented in this thesis. Importantly, studies have shown ASM as a really useful tool for greyscale images (Cootes et al. 2000) and for segmentation of musculoskeletal muscles (Darby et al. 2012).

2.3.2 Dynamic Tracking Applied to Ultrasound Imaging

One of the main goals of using ultrasound imaging for the study of skeletal muscle is obtaining qualitative and quantitative information of the different characteristics of muscle movements, ranging from the computation of general flow velocities and accelerations in big sonicated areas, to analysis of localised regions associated with specific musculoskeletal elements such as muscle fascicles. To empirically account for velocities and acceleration at different scales is particularly difficult for the case of musculoskeletal imaging, in which tissues such as tendons or fascicles appear as granular backscatters of the ultrasound waves. Although the aponeuroses are usually well-defined components (Figure 2.9), structures within the muscle cross-section area are not trivial to distinguish, meaning two main options can be considered in order to obtain reliable dynamic information:

- Application of filtering techniques that convert granular contours into solid shapes. Examples of this are the use of Gaussian kernels and vessel enhancement filters (Frangi et al. 1998) for a better definition of muscle fascicles (Rana et al. 2009); and the application of Kalman filters for vessel tracking and calculation of transverse vessel area (Guerrero et al. 2007).

- Adaptation of existing image processing algorithms for the tracking of speckles or granular objects. Techniques based on Block Matching Algorithms and Optical Flow have been recently developed and adapted to a wide variety of ultrasound imaging data such as the tracking of tendon movements

(Mohamed 2015) or myocardial motion estimations (Angelini & Gerard 2006). These methods are also known as speckle tracking techniques.

Filtering techniques such as vessel enhancement modify collected images in order to "make them look better". Although they have been to have good performance (Rana et al. 2009), they can "force" unclear, defined structures to belong to pre-determined diameters and inclinations that in some cases are impossible to validate. In other words, they adapt the collected images to a preformed notion of what they should look like, rather than adapting the techniques to the inherit characteristics of the ultrasound data. The following work therefore focuses on the application of speckle tracking based techniques.

2.3.2.1 Block Matching Algorithms (BMA)

BMA were one of the first automatic techniques applied for speckle tracking in ultrasound images (Boukerroui et al. 2003). These methods assume that the brightness of a selected block is constant between consecutive frames and while an important number of different implementations of these algorithms exist, they usually follow a similar structure:

- Firstly, the initial frame is subsampled into smaller blocks, usually of the same size.

- Then an iterative search for blocks of similar intensity is performed within the entire or sub-portion of the following frame, typically known as the region of interest or window.

- A so called Motion Vector is then calculated as the displacement between the block's position in the initial frame and the position of its best fit in the following one. This measured motion is assigned to every pixel contained inside the block, and is usually measured in pixels per frame.

These steps are repeated for all non-overlapping blocks within the initial frame and for each of the remaining frames and provide two main outputs per block: the motion vector and the matching error function. A block considered best fit corresponds to the minimum matching error.

The principle difference between BMAs resides in the way the best fit is calculated (also known as distortion measure). Therefore, the great majority of BMAs can be classified into two different groups:

- *Mean Square Error (MSE):* This matching criterion computes the error function as the Euclidean distance between two consecutive frames. Besides being one of the most common cost functions used in BMA, its main disadvantage is that it penalises large errors more, therefore small errors are more assumable with this method. Another similar variant of MSE is to calculate the error function as a sum of the error values rather than squaring them. This method is commonly known as Sum of Absolute Distance (SAD).

- *Normalized Cross-Correlation (NCC):* This method shifts a reference block from one frame and calculates its correlation with equal sized blocks in another posterior frame. The outcome of NCC is a two-dimensional correlation map that defines how similar different patches (blocks) between images are. It differs from the MSE approach in the normalisation of the resulting differences, obtaining coefficients that vary between 0 (no correlation) and 1 (perfect correlation). Its main limitation is the increased computational cost compared with MSE making it less suitable for real-time purposes.

One of the biggest drawbacks of BMA resides in its high computational time when using exhaustive search methods. One of main strategies in order to solve this problem was the reduction of the number of computations per iteration, resulting in modalities generically known as "fast full search" BMA. These techniques perform faster by reducing the number of blocks used in the iterative search improving the

overall efficiency, however they showed low efficiency when computing small motion displacements or irregular flows (Lin & Tai 1996).

2.3.2.2 Optical Flow

Optical flow computes the pattern of apparent motion of image intensity in the time domain. The 2D velocities representing the horizontal and vertical movement of all visible surface points in a pair of images is known as the motion field, it is the geometrical concept that relates a 3-dimensional movement with its projection on a 2-dimensional plane. Despite the existence of a wide range of Optical Flow techniques, their processing usually follows a similar structure that can be divided into three main steps:

- In order to improve the signal to noise ratio of the image a low-pass or band-pass filter, commonly known as *smoothing*, is applied. Different filtering techniques are applied depending on the structures they are designed to be applied to (e.g. vessel enhancement filters in muscle, nerve or artery ultrasound imaging (Frangi et al. 1998; Namburete et al. 2011)).

- Basic dynamic information is extracted from local or global structures such as correlated shapes or position movements that provide individual two-dimensional pixel motions between frames.

- The individual information is compiled in order to build a two-dimensional matrix containing the velocity estimations in both X and Y Cartesian components (sometimes expressed in polar coordinates).

A useful classification was presented in (Barron et al. 1994) dividing the available techniques into four main groups depending on the approaches used for estimating motion vectors between frames:

- **Differential Techniques**: These techniques assume constant velocity during short durations and translational movements between consecutive frames, therefore the temporal and spatial derivatives are close to zero. As a result, they compute velocity from spatiotemporal derivatives of image intensity or filtered versions of the image:
    - Horn and Schunck: gradient constraint combined with a global smoothness term (Horn & Schunck 1981).
    - Lucas-Kanade: weighted least squares fit of local first-order constraints to a constant model for velocity in each small spatial neighbourhood (Lucas & Kanade 1981).
    - Nagel: second order derivatives to measure optical flow with a global smoothness constraint like Horn-Schunck (Nagel 1987).

- **Region-Based Matching**: Due to the fact that accurate numerical differentiation may be impractical because of factors such as noise, unavailability of a sufficient number of frames or the existence of sharp edges or aliasing in the image acquisition process; different approaches should be taken into account. Region-based matching assumes that the best fit between image regions at different times is obtained by using the spatial displacement contained in the velocity:
    - Anandan: Based on Laplacian pyramid and a coarse-to-fine Sum of Square Differences (SSD)-based matching strategy that allows computation of large displacements (Anandan 1987).
    - Singh: Two-stage matching method, consisting the first in the computation of SSD values, and the second in propagating velocity by using neighbourhood information (Singh 1990).

- **Energy-Based Methods**: The velocity vector is estimated using the output energy of velocity-tuned filters, particularly used in the frequency domain. Gabor filtering is typically used for correlating patches in the spatiotemporal frequency domain, thus these methods are similar to the Lucas-Kanade approach in both results and method. An example of this can be found in

(Heeger 1987), that computes velocity as a maximum of the distribution of velocity-tuned units, using a Least squares fit of spatiotemporal energy to a plane in frequency space.

- **Phase-Based Techniques**: Unlike Energy-Based methods, these techniques are focused on the phase of the velocity vector in the frequency domain. Zero-crossing techniques are included in this classification as they can also be interpreted as phase level crossings. Examples of this are described in (Waxman et al. 1988), which applies spatiotemporal filters to binary edge maps to track edges in real-time; and (Fleet & Jepson 1990), that defines component velocity in terms of the instantaneous motion normal to level phase con tours in the output of band-pass velocity-tuned filters.

Due to the great variety of available OF techniques, the better suitability of one technique before others is strongly related to the peculiarities of the images used in the process. For example, when estimating motion from ultrasound images of the musculoskeletal system pyramidal OF techniques (capable of the computation of high magnitude flows) would not be ideally suitable for the study of subtle muscle movements or for ultra-fast frame rates because the displacement between two consecutive frames would very small and therefore achievable without escalating it to bigger dimensions.

While BMAs need less iterations than OF algorithms, as they compare precise sized blocks instead of pixel by pixel calculations, the computations performed in OF methods are more efficient than BMA. This latter factor makes OF techniques more efficient overall and therefore more suitable for real-time information extraction of dynamic properties of muscle behaviour using ultrasound technology, which will be an important consideration for extending the use of ultrasound technology within the clinical setting.

2.3.2.3 Speckle tracking algorithms resume

There is a need to use automatic tracking techniques for the assessment of ultrasound images so that additional objective information that may be available, but undetectable to even expert operators, can be quantified. In order to improve healthcare and understanding of the neuromuscular function the full extent of the information contained in US imaging should be exploited.

Optical flow techniques have been previously shown as a valuable tool in Contrast-enhanced Ultrasound (Lee & Kim 2013) and for B-mode tracking of arterial wall motion (Stoitsis et al. 2005), which also highlighted the greater efficiency of optical flow than block matching techniques. An important benchmarking of different algorithms based in speckle tracking in ultrasound synthetic images was performed by (Alessandrini et al. 2016). They compared the displacement (and strain accuracy) using elastic registration (AFFD), two Block Matching algorithms (RF tracking and B-mode block matching) and two Optical Flow techniques (AAOF (Barbosa et al. 2014) and Sparse Demons (Somphone et al. 2013)). Limitations of the presented work were the application to synthetic data (no US artefacts) and the use of non-commercial algorithms (operator dependent). Besides these factors, conclusions clearly stated that block-matching methods performed significantly worse than OF and elastic registration. The two OF and the AFFD provided similar tracking results overall, but OF was shown to be more efficient and faster than elastic registration ((0.08 – 3.8s/ frame vs 10 min/frame for AFFD).

Another comparison of the performance between Optical Flow and Block Matching Algorithms was done in (Darby et al. 2012). An automatic feature detector using Lucas-Kanade tracker showed more consistent results due to its smoother performance as opposed to the jumps observed in a semi-automated (i.e. required manual initialisation) spatial cross-correlation technique used in (Loram et al. 2006).

The aforementioned studies have led me to conclude there was better suitability of Optical Flow instead of BMA for the extraction of dynamic information from

ultrasound skeletal muscle images. Information regarding the performance of different OF approaches for assessment of skeletal muscle has however not previously been completed. Therefore, a larger study of the performance of different OF algorithms was defined as a key objective of this thesis.

2.4 Physiological Data Prediction using Machine Learning

The following section provides a review of the techniques available for prediction of external data (e.g. joint torques) using US imaging derived information.

Optical flow assumes conservation of pixel brightness between consecutive frames and throughout the entire sequence of images. As this assumption is quite unrealistic for the case of ultrasound imaging it limits the use of this methods to a qualitative estimation of movements rather than a quantitative one (Duan et al. 2005). Furthermore, while Optical Flow algorithms can help understand relationships empirically they are not able to predict dynamic properties such as net joint torque, a skill that would have great value for both clinical and basic science research.

Ultrasound images have been used for the prediction of parameters such as Maximum Voluntary Contraction (MVC) torque (Cuesta-Vargas & Gonzalez-Sanchez 2014). MVC torque was estimated initially using EMG activations from medial and lateral gastrocnemii and soleus muscles and then corrected using the volumetric muscle information obtained from 2D Hill-model estimations of muscle cross sectional area from US images (DeOliveira & Luporini Menegaldo 2010). This showed that information extracted from US data can be successfully used for predicting physiological parameters such as MVC. However, the significant error found in the resulting predictions (bigger than 20% RMSE, (DeOliveira & Luporini Menegaldo 2010)) and the limited measures taken from the ultrasound data indicate scope for the exploration of the potential of application of data inference techniques such as Machine Learning.

2.4.1 Machine Learning Classification

Machine learning explores the theory and implementation of algorithms capable of extracting information from complex datasets, to identify similar patterns in new data and make predictions based on the presence and form of these patterns. In order to provide a good understanding of all the capabilities of this subfield of computer science, the different algorithms can be grouped depending on the "guidance" provided to the algorithm and on the desired outcome.

Depending on the pre-processing and the available information prior to the use of these systems, we can classify Machine Learning Algorithms into four major groups (Baştanlar & Ozuysal 2014):

- **Supervised Learning:** Finds patterns in data in order to predict accurately outputs using already given (labelled) inputs. Regression and Classification Algorithms are the most popular examples of supervised learning.

- **Unsupervised Learning:** Finds relationships within data, extracting patterns or structures from unlabelled data. Clustering algorithms (k-means) and dimensionality reduction systems (PCA) are examples of Unsupervised Learning.

- **Semi-supervised Learning:** Mix of the first two, meaning that data is partially labelled, for example providing estimations of future stock values in different markets.

- **Reinforcement Learning:** The algorithm interacts with a dynamic environment in which the only information available is whether the estimated output is successful. Unlike supervised learning, reinforcement learning inputs and outputs are available but not paired.

Machine Learning algorithms can also be classified depending on the type of problem they are meant to provide solutions to (Zhang & Ayodele 2010). While a great number of them can be used individually as an overall system, some of these algorithms can be combined in order to provide better insight, and therefore more predicting power, to complex relationships:

Regression Algorithms: Model transference functions relating different variables in order to use them to predict the outputs of new inputs. New inputs are defined as data points that are not part of the training set of the algorithm. Differences between algorithms of this type are mainly based on the error function used for measuring the prediction accuracy such as Least Squares or Linear Regression.

Classification: Distributes the available data into a predetermined number of classes. Depending on the number of categories or classes in which the data has to be grouped, we can distinguish between Binary and Multiclass Classification. Bayes and Support Vector Machine Classifiers are amongst the most commonly used algorithms for classification.

Clustering algorithms: Organizes the dataset into different groups by means of similarity. In other words, a set of inputs is divided into a number of different groups that are unknown beforehand (unsupervised learning). Clustering algorithms usually differ in the type of similarity used such as square distances or different statistical distributions. Examples of these are k-means, k-medians or Hierarchical clustering.

Dimensionality Reduction: An unsupervised learning algorithm that reduces the number of random variables in a process. The inputs are narrowed down by mapping them into a lower-dimensional space and therefore data is described using less information. One of the most used techniques for this is Principal Component Analysis (PCA).

Regularization: An extension made to another method (typically regression) that penalizes models based on their complexity, favouring simpler models that are also

better at generalizing. The two most used regularizations are Ridge Regression, Least Absolute Shrinkage and Selection Operator (LASSO), also known as L2 and L1 respectively.

Artificial Neural Networks (ANN): Mathematical models that mimic biologic neurons decision-making structure and behaviour. In their basic form, ANN assign weights to the different inputs and compute transference functions, that conform to the so called hidden layer, in order to obtain the different outputs. Some modalities of ANN include iterative calculations that feedback the error of the output estimations known as back propagation. Networks with big amounts of hidden layers are also known as *Deep Networks* and are capable of computing significantly more complex and/or non-linear input-output relationships.

Despite the great variety of Machine Learning algorithms, they usually follow a similar basic structure as showed in Figure 2.12. In order to evaluate the performance of the algorithm, data is firstly pre-processed and then separated in two parts: usually the biggest part of the data is used for training the algorithm, meanwhile a small portion is kept for testing the model. The desired algorithm is then selected and trained with the initial data, using a set of different parameters that modify its performance (algorithm tuning). The model then generates predictions of the output that are compared with the ground truth. In iterative implementations the error between the output and its estimation is sent back to the tuning stage in order to find a combination of parameters that better infer the desired outputs. Once reached the desired fit the model is validated for that precise purpose.

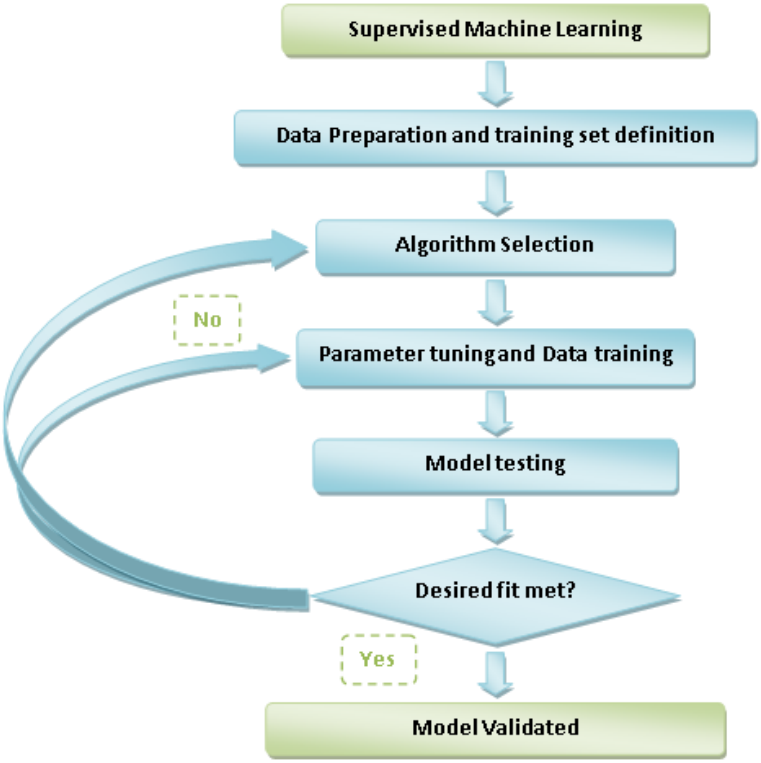

Figure 2-12. Basic block diagram of Supervised Machine Learning Algorithms.

As one of the main goals within this project is to establish relationships between different known inputs (i.e. measures from US data) and outputs (i.e. net joint torque), the study of physiological parameter inference methods has been focused on *supervised learning*. Within this category, both classification and regression algorithms can provide new insights to the functioning of the musculoskeletal system: the former being able to distinguish between different types of movements such as passive or active contractions (Cunningham et al. 2013); the latter for predicting dynamic properties such as wrist angles from US and EMG data (Xie et al. 2009) .

The recent advances in the area of pattern classification and machine learning research have triggered new applications for medical imaging. In order to directly provide high-level interpretation of results to aid clinical decisions, classification technologies are often used to reach more objective and statistically solid conclusions than can be achieved by human assessors. They offer a means to merge multiple features (e.g. spectral, texture and motion-based) to provide more reliable

interpretation and a growing interest in applying more sophisticated learning and classification algorithms has been observed in general ultrasound (Moon et al. 2012).

Classification algorithms such as Support Vector Machines have been previously applied to musculoskeletal ultrasound images to distinguish between active and passive muscle contractions during joint rotations (Cunningham et al. 2013), and the application of statistical method based algorithms such as Mutual Information (MI) for the detection of muscle activations has also been reported (Harding et al. 2012). As the focus of this thesis is to evaluate prediction of musculoskeletal function from US images, regression machine learning algorithms are the most appropriate to be applied and are hence described in fuller detail below.

2.4.2 Regression Machine Learning Algorithms

Unlike classification algorithms, where only qualitative outcomes can be inferred, regression algorithms are capable of quantitative predictions between inputs and outputs (Hastie et al. 2009). A resume of the most significant algorithms for regression is provided below, focusing on their suitability for the prediction of dynamic properties of skeletal muscles using ultrasound.

2.4.2.1 Multiple Linear Regression

Linear Regression is the most simple implementation of regression algorithms (Peckov 2012), and identifies scalar relationships between input-output pairs. It can be modelled as:

$$\mathrm{y} = offset + kx + \varepsilon \qquad (1)$$

Where y is the output value for the given input x, k is a scalar, and ε is the error component of the model (expressed as the difference between the predicted output

and the real output). It is possible to use more than one input for the prediction, which is known as *Multiple Regression:*

$$y = k_0 + \sum_{i=1}^{n} k_i x_i + \varepsilon \quad (2)$$

where *k* is a vector containing the parameters known as regression coefficients (predictors). This vector can be divided into k_0, the offset of the model, and k_i, the scalar predictor values for each one of the *n* inputs.

As the main goal of these algorithms is to estimate the set of outputs closest to the ground truth, the error vector (ε) is minimised using the Residual Sum of Squares function (RSS):

$$RSS(k) = \sum_{j=1}^{m} \left(y_j - \left(k_{o+} \sum_{i=1}^{n} x_i k_i \right) \right)^2 \quad (3)$$

By characterizing it into matrix notation, and assuming it is non-singular (the inverse of the matrix exists), the implementation of the algorithm can be simplified to:

$$\hat{k} = (X^T X)^{-1} X^T y \quad (4)$$

This equation is the unique solution of the minimization problem and therefore the *general linear model for regression.*

If a simple example with two inputs and one output is used, Linear Regression algorithms basically try to find an optimal line that best groups the input-output pairs in terms of least square error (or other goodness of fit indicator), as showed in Figure 2.13.

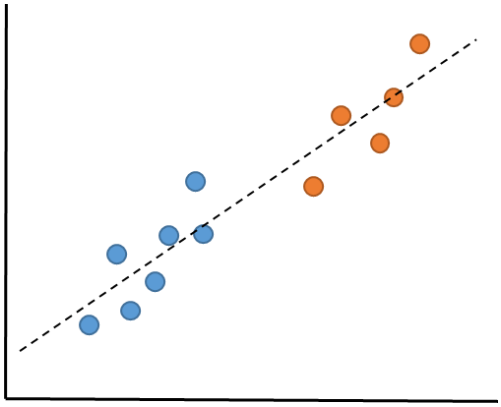

Figure 2-13. Linear regression (black dotted line) for two sets of paired inputs and outputs (blue and orange).

Overall, the linear regression model is the most efficient regression algorithm in terms computational costs and is particularly suitable for small datasets, however it is only capable of computing simple relationships between inputs and outputs.

2.4.2.2 Polynomial Regression

Polynomial regression differs from linear regression in the use of polynomials for modelling the relationship between inputs and outputs instead of linear coefficients. Polynomial models are one of the most used algorithms for fitting experimental data due to their simplicity and high efficiency (Peckov 2012), and can be expressed as:

$$polynom = k_0 + \sum_{i=1}^{n} k_i P_i \tag{5}$$

Where k is a vector of constant values, and P contains the "terms" of the polynomial expression over the input variable X_j:

$$P_i = \prod_{j=1}^{m} X_j^{\alpha_{i,j}} \ , \alpha_{i,j} > 0 \tag{6}$$

for which $\alpha_{i,j}$ defines the different degrees of the overall equation.

According to Weierstrass theorem, any continuous output can be estimated by a polynomial within an interval (Bishop 1963). In fact, different degree polynomial

approaches can accurately fit a desired output, constituting one of the biggest advantages of the use of Polynomial Regression algorithms.

Polynomial regression was applied to the same example data used to illustrate linear regression (Figure 2.13), with results shown in Figure 2.14. It can be observed that higher degree approaches (Figure 2.14, right) provide lines with a greater number of inflection points for reducing the deviation.

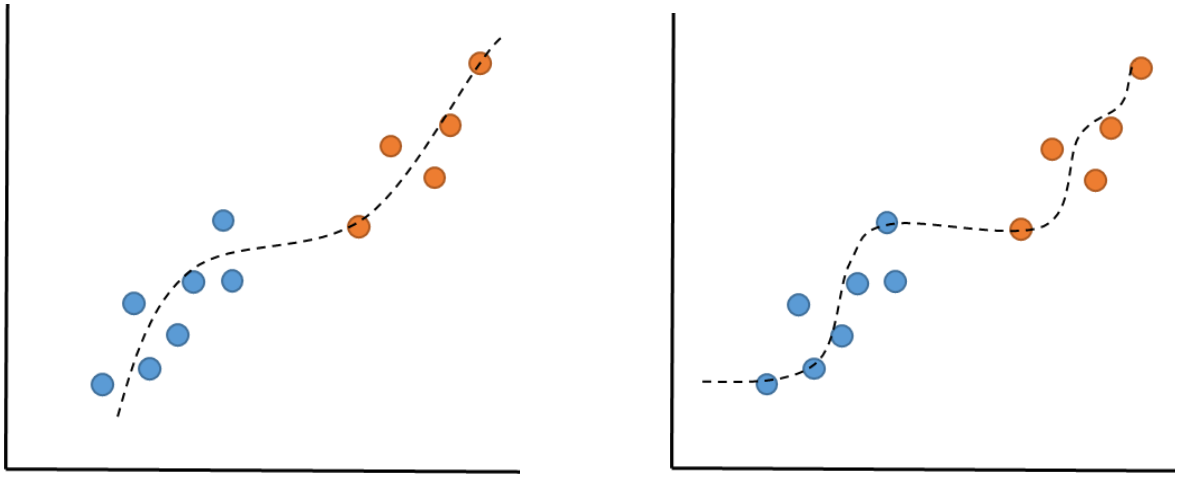

Figure 2-14. Low degree (left) vs high degree (right) Polynomial regression (black dotted line) for two sets of paired inputs and outputs (blue and orange).

Although capable of predicting significantly more complex relationships than linear approaches, polynomial approaches can produce important errors due to oscillation near the limit of the intervals in which they are applied. This "Runge phenomenon" was first discovered by Carl Runge in 1901 when studying the behaviour of errors using polynomial fits (Boyd & Xu 2009), and its effect is proportional to the degree of the polynomial used. Therefore, the use of high degree polynomial approaches can result in unstable models that also require higher computational time for their implementation.

2.4.2.3 Support Vector Machine (SVM)

SVM derives from the General Portrait algorithm and statistical learning theory, and was first developed by Vapnik and Chervonenkis more than 50 years ago (Vapnik & Chervonenkis 1971). Its main function is to calculate hyperplanes that spatially separate different training data points in order to create a classifier able to correctly differentiate new or unseen data, and has been further improved and developed since the 1990s (Smola & Schölkopf 2004). The optimal hyperplane in SVMs is defined as the one which provides maximum distance between the different input data points closest to it (also known as Support Vectors). Therefore, the main goal of SVMs is to minimise the classification error while maximising the hyperplane margin. Support Vectors are the only ones needed to define the hyperplane and thus the rest of the data points can be excluded from all calculations, simplifying the overall implementation of SVM.

Unlike in linear regression, SVMs try to find a hyperplane that maximizes the distance between the support vectors of each of the different inputs. This hyperplane is then used for classifying new inputs in either of the groups. In other words, regression algorithms try to best fit the training data whereas in Support Vector Machine the main objective is to find the biggest distance between groups of inputs. Figure 2.15 graphically explains the difference in the performance of SVM and linear regression.

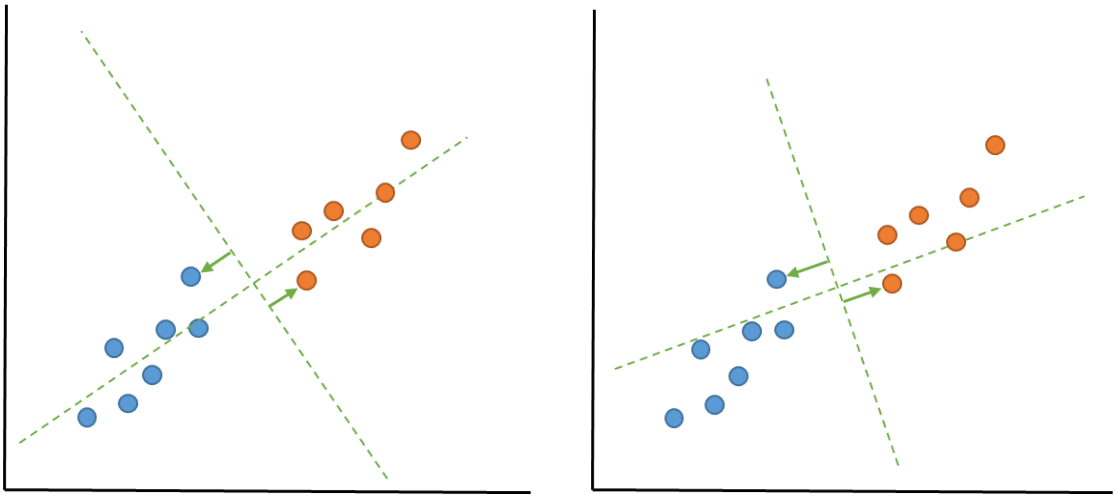

Figure 2-15. Linear regression (left) vs SVM linear classifier (right) hyperplanes (green dotted lines) for two sets of paired inputs and outputs (blue and orange circles). Green arrows represent the distance between the hyperplane and the support vectors for each of the two inputs.

Although it shows a worse fit than the linear regression in the example in Figure 2.15, on the whole SVM performs better as a generalised method being able to predict new inputs more accurately.

2.4.2.4 Support Vector Regression:

The particularity of providing better generalised solutions than statistical regressions makes SVM very suitable for defining rules that differentiate multiple inputs, hence it is a well-known tool for classification in supervised learning, and for clustering in unsupervised learning. Furthermore, these algorithms can be applied for regression by the addition of a loss function that includes a distance measure. These methods are also known as Support Vector Regression (SVR).

Unlike classification problems, where the main goal is to maximise the separation between support vectors and the hyperplane, in SVR the main objective is to minimise a cost function that measures the error between a target value and its model estimation. Among the different cost/error functions available, the most commonly found in the literature are the Square loss, Laplace and ε-insensitive (Brereton & Lloyd 2010). Figure 2.16 provides a graphical comparison between them.

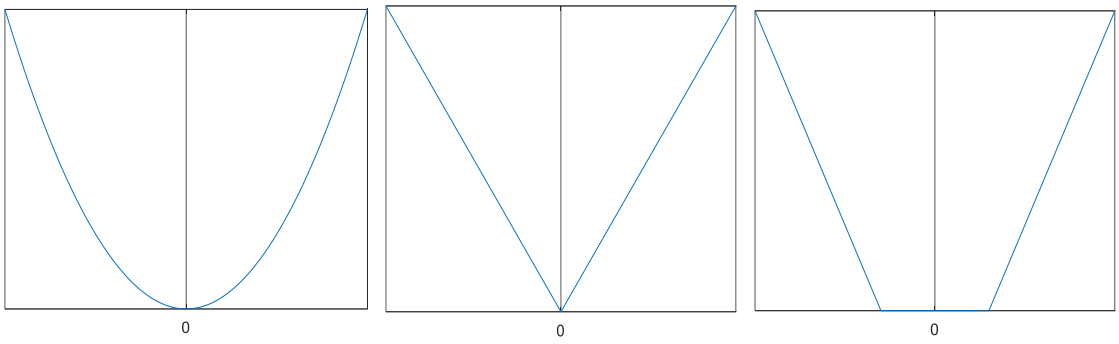

Figure 2-16. SVR Error function comparison between Quadratic (left), Laplacian (middle) and ε-insensitive (right). X axis represents the residual between target and predicted output and Y axis represents the penalization to be applied.

Laplacian loss function: Also known as least modulus loss or least absolute deviations, basically minimizes the absolute difference between the ground truth value and the estimations provided by the model used.

Quadratic: is also called Least squares error function or squared loss and its main goal is to minimize the sum of the square of the differences between the target value and the prediction of the regression model. Due to the fact that the error is squared quadratic is more sensitive than Laplacian especially when the residual is big (greater than 1), the latter is however more robust in terms of allowing the existence of outliers (in the case they can easily be ignored). Furthermore, least squares is more stable than Laplacian as it provides slight changes in the regression line to small changes in data whereas some big jumps can appear when using the Laplacian loss function.

ε-insensitive loss function was developed by Vapnik (1995) and has become one of the standards in SVR. It uses an error threshold parameter defined by ε that rules a piecewise function: The function to be minimized is 0 when the difference between the predicted output and its real value is within the range ± ε, and performs a Laplacian loss function (L1-norm) otherwise. Having a range in which no penalizations are applied to estimations close the real value leads to sparse support vectors, meaning that the amount of zero coefficients is reduced. Sparse SVM is

usually associated with a fast evaluation and optimization of the model (Gomez-Verdejo et al. 2011).

Analogously to the implementation of the linear SVM for classification shown in Figure 2.15, the chosen error function is minimized together with the maximization of the distance between the hyperplane and the support vectors. This way, linear and non-linear input-output relationships can be predicted with the use of these methods. Nevertheless, in some situations even high degree polynomial regressions are not capable of providing models that accurately predict new data points. In order to solve this (Cortes & Vapnik 1995) proposed a transformation of the data to a higher dimensional space for which a linear regression can be applied. This technique is known as "the kernel trick" and is explained below.

2.4.2.5 Kernel Trick:

Kernels are defined as symmetric functions that measure the similarity between pairs of data points, in other words, they use two data points as an input and a scalar that relates them as an output (Campbell 2001). These similarity measures are mapped to a high dimensional space that modifies their spatial distribution. Within this high dimensional feature space, the initial non-linear patterns can be explained using linear regression as it is graphically represented in Figure 2.17. The resulting linear pattern in the feature space has to be "unmapped" in order to obtain the model in the initial input space.

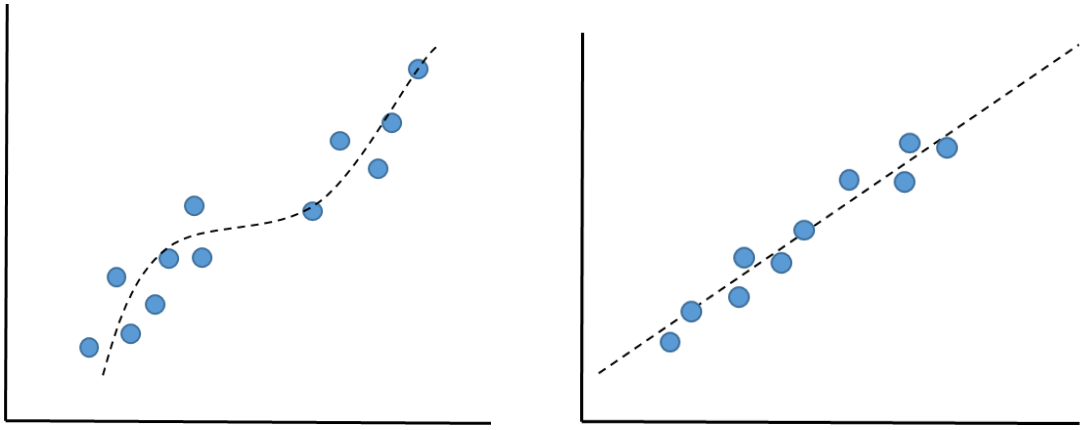

Figure 2-17. Example of application of the Kernel Transformation to a nonlinear input-output relationship. Left shows the dataset distribution in input space, and right shows the inner product distribution in the high-dimensional feature space.

The main advantage of the use of kernels for the distribution of data in high dimensional feature spaces is to reshape the data distribution when data is not (easily) separable in the input space. Furthermore, instead of having to calculate all the coordinates in the high dimensional space, kernel methods only compute the inner products between the corresponding input-output pairs in the feature space, reducing the computational requirements of the process.

2.4.2.6 Regression ANNs

As previously stated (Section 2.4.1), Artificial Neural Networks are capable of computing nonlinear complex input-output relationships and have been proven as a reliable prediction method for both Classification (Cunningham et al. 2013; Dreiseitl & Ohno-Machado 2002), and Regression (Papadopoulos & Haralambous 2011; Mantzaris et al. 2008). These techniques have the ability to acquire knowledge through the training phase, by assigning weights to the synapses that connect different neurons, and are good at discarding information that does not contribute enough to the model performance.

Neural Network architecture can be grouped into three different units: input layers containing the raw information; output layers, the predicted results that should be as close to the targets (ground truth values); and hidden layers, that assign different

weights to the neurons connecting inputs and outputs in order to compute patterns that relate them.

Activation functions determine whether a neuron is activated or not depending on the weights of the inputs connecting that precise neuron. Depending on the way the neuron activation is calculated two main groups can be distinguished (Klaus Debes, Alexander Koenig 2005):

Dot products: The neuron is activated depending on the sum of the weighted inputs.

$$neuron_i activation = \sum_{j=1}^{n} w_{ij} x_j \quad (7)$$

Where x is a vector containing all the inputs connected to the $neuron_i$, and w contains the weights of those synapses. When equalizing this equation to 0, a hyperplane is constructed defining the threshold for which that neuron should activate or not.

Distance measures: The activation of the neuron depends on a measure of the spatial distance between the inputs and their weights. Euclidean distance and Mahalanobis distance are examples of distance measures applied to this group of activation functions.

Output Functions: Within the hidden layer, weights of the outputs are calculated by applying a transfer or activation function to the weights of the inputs connected to the hidden layer neurons. The most common transfer functions used are the threshold function (or step function), which outputs a zero if the input is below the "threshold value" and outputs a 1 otherwise; and the logistic function, that outputs nonlinear continuous changes when the input varies. Special cases of the latter are sigmoid, which ranges between 0 and 1, and the hyperbolic tangent, which ranges between -1 and 1. Figure 2.18 provides a visual representation of the distribution of the different output functions including their equations.

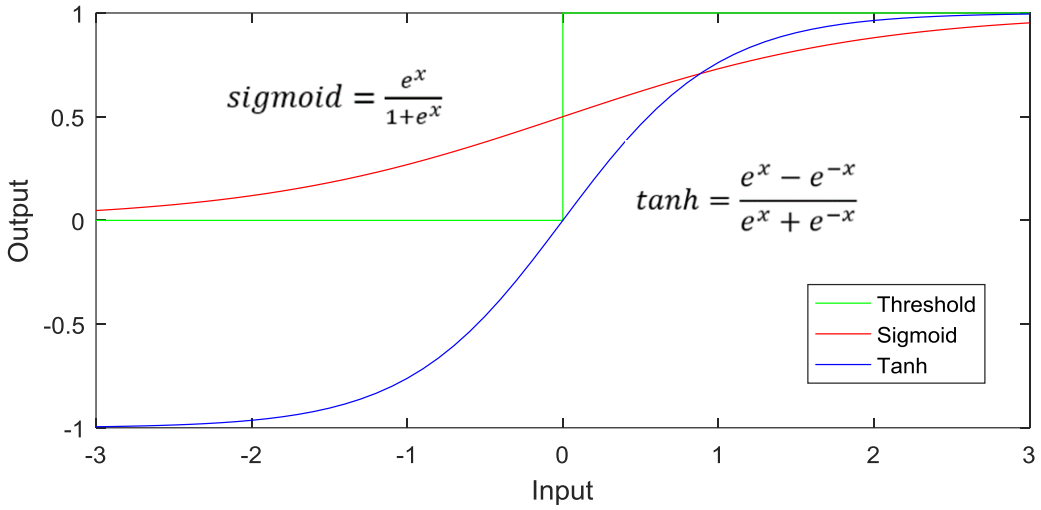

Figure 2-18. Comparative of threshold (green), sigmoid (red) and Tanh (blue) Output Functions used in ANN Hidden layer neurons.

Although the distributions in figure 2.18 have zero mean (are centred in the cross between X and Y axes), different mean values can be implemented by the addition of the so called *Bias neurons*. These nodes are added as another input with an invariable weight of 1 to that precise hidden neuron. They act as a shift equivalent to the "offset" of a polynomial expression, obtaining a non-zero centred output function. This way, the network would fit the data when all the other nodes are 0.

Training Algorithms: Analogous to Support Vector Regression, ANNs also use minimization of loss functions in order to learn from input-output relationships. The main difference between both in the use of these functions is that in SVM loss functions are applied to the target - prediction difference, meanwhile in ANNs they are used for readapting the weights of the synapses and biases between neurons during a process called back propagation.

Assuming an ANN consists of n hidden layers connecting inputs and outputs, thus the loss function would be n-dimensional, the different loss algorithms principally focus on the minimization of the first and second derivatives of that n-dimensional function. Methods that minimize the first derivative are known as gradient based or first order methods, and the Jacobian Matrix stores the partial derivatives in respect to each weight. Second derivatives of the loss functions are known as second order

methods and their partial quadratic derivatives are usually grouped in a Hessian matrix.

The amount of available algorithms for ANN training (back propagation) has exponentially increased in recent years. For the purposes the work presented in this thesis, two different training algorithms were implemented using the Matlab Neural Network Toolbox (Mathworks, USA) and used for parameter prediction due to their fast convergence and their good performance when accurate training is required (MathWorks 2016), and are therefore described in more detail here:

Levenberg-Marquardt (LM):
This algorithm was developed in the 1960s by Kenneth Levenberg and Donald Marquard to solve slow convergence when using the gradient descent algorithm (also known as steepest descent or error back propagation): when the step size chosen is small, the number of iterations increases a lot; and when the step size is large oscillation can occur.

In order to solve the gradient problem, the Gauss-Newton approach uses the second derivative of the loss function, but it only works (converges) when the quadratic error (Hessian) matrix is non-singular. Furthermore, the calculation of the Hessian matrix and its inverse can be computationally expensive (the size of it is n^2 compared with n for the gradient descent case).

Levenberg-Marquardt algorithm mixes both Newton and gradient descent by performing gradient descent approach until the error function curvature is approachable by the quadratic approximation (Marquardt 1963).

Scaled Conjugate Gradient (SCG):
This is a modified version of the Conjugate Gradient (CG) method, in which the gradient descent search for a minimum of the loss function is performed in the direction given by the Hessian matrix. Unlike the Gauss-Newton approach,

Conjugate Gradient does not store the Hessian matrix, reducing this way the computational requirements.

The CG can also be considered a hybrid between first and second order methods. It basically combines the model-trust region approach in LM with the gradient approach search within the Gauss-Newton. SCG was designed by Moller in 1990s for being applied to non-symmetric quadratic areas in order to enable its application to a wider range of problems (Møller 1993).

The training process shares a similar iterative structure for adjusting the weights of the synapses of the different neurons in order to improve the model and obtain results closer to the targets. This iterative structure is quite similar to the one showed in the previous section regarding Supervised Machine Learning, and a simple schematic of its structure is shown in Figure 2.19.

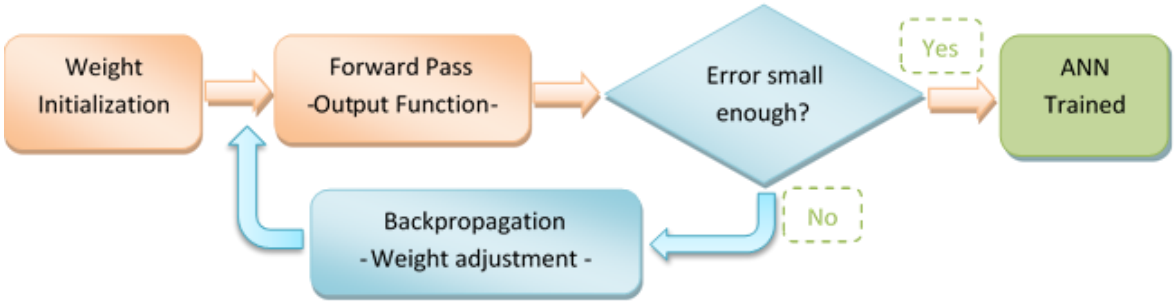

Figure 2-19. Artificial Neural Network iterative structure including Forward Pass and Backpropagation.

The synapses weights are initially assigned, either arbitrary or by following a precise rule, and first results are computed by the application of the output function selected. This initial result is compared with the final target, usually giving an important difference between both (as the algorithm has not been trained yet). As long as the difference between the target and the generated output is bigger than the model objective, backpropagation is applied. Within this stage, new weights for the different synapses are calculated by the use of a training algorithm. Once updated the new weights, the error is computed again and compared with the

desired target. When the error condition is satisfied the model is considered trained and can be used for the prediction of new inputs.

On some occasions, when the error is not reduced after a certain amount of iterations, the model is believed to have converged to a minimum and cannot be further trained. In order to obtain outputs closer to the target and therefore meet the error requirement, the implementation of a more complicated topology including additional hidden layers and neurons should be studied together with the use of different output and/or training functions.

Another parameter to take into account for the performance of training algorithms is the net topology and complexity: Models that are too simple usually provide inaccurate predictions when inferencing complicated data sets, therefore the model approaches need to have a certain complexity. On the other hand, the consequence of too complex models is the over-fitting of the data. Complex models try to match the given targets so accurately that they end up fitting data that does not provide information to the model (noise) (Mantzaris et al. 2008).

Regularization

Together with the loss functions, both SVM and ANN regression models include a so called regularization function (also known as shrinkage methods) to be minimized whose main objective is to avoid data over-fitting in order to obtain a more general solution, able to provide good predictions to new data. In other words, the main goal of regularization functions is to avoid the excessive contribution of one input over the rest (Smola & Schölkopf 2004). The two main algorithms used for regularization share the same distribution as the Quadratic and the Laplacian Support Vector Regression error functions explained in the previous section (Figure 2.16 left and centre respectively).

Regression algorithms that use the Quadratic shaped regularization are known as *Ridge Regression* and are equal to the sum of the squares of the weights. It is

computationally more efficient as it has an analytical solution, but it is unable of providing sparse coefficients. The algorithms that use the sum of the different weights as regularization are known as *LASSO (Least Absolute Shrinkage and Selection Operator) Regression*. Besides being less efficient than Ridge regression they have a built-in feature selection and allow sparse coefficients. Ridge Regression works better for dense/global models (meaning the ones that take into account all variables) meanwhile LASSO works better for local (it doesn't involve all the variables, just a subset of them).

Regression algorithms Resume:

There is an important controversy between the use of complex SVR algorithms and ANNs when talking about the better performance of one or another. It is important to keep in mind that in machine learning there is always a trade-off between capacity and generalization, as stated in (Cortes & Vapnik 1995). Algorithms that are capable of learning from more complex input-output relationships by the use of higher degree polynomials or kernel based (high capacity) are able to fit better the training data. On the other hand, nonlinear techniques such as Kernel Methods and ANNs have significantly bigger computational resource requirements, meaning that more memory should be allocated for their implementation and the computational time required would be also bigger.

2.5 Conclusions

A contextualisation of the use of ultrasound technology for the analysis of the neuromuscular system and the new emergent methodologies applicable to ultrasound image data that may able to assess properties of muscles has been the focus of this chapter. The literature has shown the need of further study of skeletal muscle properties with ultrasound, specifically highlighting current gaps in understanding of the influence of:

i. precision and temporal resolution of acquired ultrasound data;
ii. choice of analysis techniques for measurement of changes in muscle properties from ultrasound data;

iii. choice of analysis approaches that can learn from input-output relationships to make predictions based on new data.

In order to address, and move towards filling, these gaps in current knowledge it is necessary to collect experimental data from skeletal muscles of human participants. To ensure robustness and validity of findings it was therefore vital to develop, evaluate and establish data acquisition and analysis methodologies that were objective (without human related errors) and fully synchronised. The process to achieving this is therefore detailed in the next chapter.

3 Methodology

To address the research questions stated in Section 2.5, a standard protocol for the collection of ultrasound data has been established and implemented, from human participants, and its subsequent analysis. The different experimental conditions under which ultrasound data were collected and the specific parameters for the data acquisition are described in detail in following chapters. Here, details of the methods applied to capture and analyse the US data are provided. The aim of these analyses was to extract dynamic features of muscle movement and shape change and to subsequently use these measures to predict the output of the muscle movements studied, namely joint torque.

3.1 Temporal behaviour of US devices

The main goal for this first stage was to establish a comprehensive understanding of the operation of specific ultrasound machines available within my laboratory, and to write code that would allow collection and processing of ultrasound images together with other external data (e.g. dynamometer based net joint torque). This work has been published in Royal Society Open Science (REF), and both publication and data used can be accessed at http://rsos.royalsocietypublishing.org/content/4/5/170245?panels_ajax_tab_tab=roysoc_tab_art&panels_ajax_tab_trigger.

Problems due to synchronisation between ultrasound and other data acquisition devices (e.g. dynamometers and electromyography) have previously been reported (Seynnes et al. 2015). The focus of these discussions has however been on the potential offset between systems at initialisation of data acquisition. For example, (Finni et al. 2012) demonstrated how increasing the offset between ultrasound-derived tendon length and dynamometer measured force significantly altered measures of elastic hysteresis in human Achilles tendon, with implications for understanding the ability of human tendon to dissipate energy due to its material

viscosity. Synchronisation may however also result from variable inter-frame intervals (IFIs), a factor which, to our knowledge, has not previously been specifically investigated.

The work presented here therefore addresses two key questions:
Is IFI constant within a sequence of ultrasound images?
Can accounting for variation in IFI provide better synchronisation between experimental data measures?

3.1.1 Effects of Inaccurate Timing US data in physiological measurements

To demonstrate the effects of assuming a constant IFI, both ultrasound and surface myoelectric (EMG) data were recorded from a skeletal muscle whilst small involuntary muscle twitches were evoked using low-level electrical stimulation of the motor nerve. A constant IFI would provide a consistent, constant relationship between electrical stimulation and the resulting muscle activation and tissue displacement seen in the collected ultrasound sequence.

Experimental Setup:
Ultrasound (LogicScan 128, Telemed Ultrasound Medical Systems, Italy) and EMG (Trigno Wireless EMG, Delsys Inc, USA) data were collected from the medial gastrocnemius muscle of the left leg of a single participant. The participant provided written informed consent and the study was approved by the Faculty of Science and Engineering Ethics committee at Manchester Metropolitan University.

The ultrasound probe (linear array, 59 mm long, 7 MHz latent frequency) was aligned with the fascicle plane and held in position using elasticated bandage. The EMG electrode was placed immediately alongside the ultrasound probe, on a patch of skin that had been shaved and cleaned of surface debris. For electrical stimulation, an anode (DEL001, Biosense Medical LTD, USA) was placed over the

patella and cathode (Bluesensor SP, Ambu, Denmark) placed over the tibial nerve in the region of the popliteal fossa.

Ultrasound at both 42 fps and 82 fps (the frame rate stated on the ultrasound machine during collection), and EMG data were collected for 40 seconds, during low level electrical stimulation (DS7, Digitimer, UK, 12 mA, 200 μs pulse width, 1 pulse every 5 seconds) of the nerve (Figure 3.1). Ultrasound imaging, EMG and stimulation were synchronised using a common trigger, controlled via the MATLAB data acquisition toolbox (R2014a, Mathworks, USA). EMG was collected at a rate of 2 kHz through an analogue acquisition card (USB-6211, National Instruments, USA).

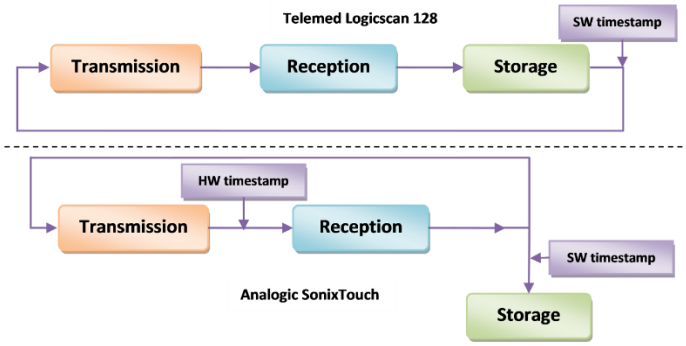

Figure 3-1. Block Diagram of the process of ultrasound transmission, reception and data storage proposed to occur in the two US machines analysed, Telemed Logicscan128 (top) and Analogic SonixTouch (bottom). Note the linear arrangement in the Telemed process, which underpins accumulation of time delays when a constant IFI is assumed. The approximate positions of software (SW) and hardware (HW) timestamp signals collected and studied are indicated at appropriate points in the cycle.

Analysis:

To identify the occurrence of the electrically evoked muscle twitches in the recorded ultrasound images a KLT feature tracking (Lucas & Kanade 1981) and mutual information based approach was applied, exactly as described by (Harding et al. 2016). Briefly, the KLT tracker estimates movement between consecutive images. It first classifies image regions by calculating a measure of their likelihood to be accurately identified in the subsequent frame of the sequence (Shi & Tomasi 1994). The most "trackable" image regions (known as *features*) are then tracked into the next frame of the image sequence. Each feature undergoes an iterative search procedure which attempts to minimise the differences between the feature selected

in frame one and the new position of this feature in frame two (Lucas & Kanade 1981). The result of this process is a set of vectors describing the magnitude and the direction of movement at multiple points across the image.

To identify the presence of a "muscle twitch" between two frames, the movement vectors were analysed using a mutual information analysis (Shannon 1948). This has been shown to differentiate between small, random, noise like movements of features within ultrasound images, and more coherent, structured movements; such as those of muscle twitches resulting from electrical stimulation (Harding et al. 2012). The result of this analysis is a frame-by-frame measure representing the likelihood that a muscle twitch is present at any given frame.

The EMG data were plotted against mutual information results to demonstrate the effects of assuming a constant IFI (Figure 3.2). This figure is plotted on the assumption that EMG data were sampled at a constant rate of 2 kHz (set via the data acquisition toolbox) and US was sampled at a constant rate of 82 fps (frames per second = 1/IFI), as stated by the US device display.

To identify the exact IFIs, metadata were extracted from the proprietary files (Telemed Video Data files, extension: .tvd) using a utility made freely available by the US device manufacturer (b to lines, Telemed, Italy). Extracted data provide many specifics about the parameters of transmission, reception and, most importantly, the IFIs accurate to 100 ns (Figure 3.1).

Results:

Evidence of IFI variability is demonstrated with the simultaneous acquisition of ultrasound and EMG data recorded from medial gastrocnemius muscle during low-level electrical stimulation of the tibial nerve. Figure 3.2 illustrates the effects of assuming a constant IFI, and shows that the time difference between the application of a stimulus and the resulting muscle activation is 49 ms, greater than the IFI (12.2 ms), even from the very first stimulation at t = 5 s. The difference at t = 40 s increases to over 300 ms, equivalent to more than 24 times the IFI. This error is so

great that the muscle movement appears to *precede* the electrical stimulus, an electrophysiological impossibility.

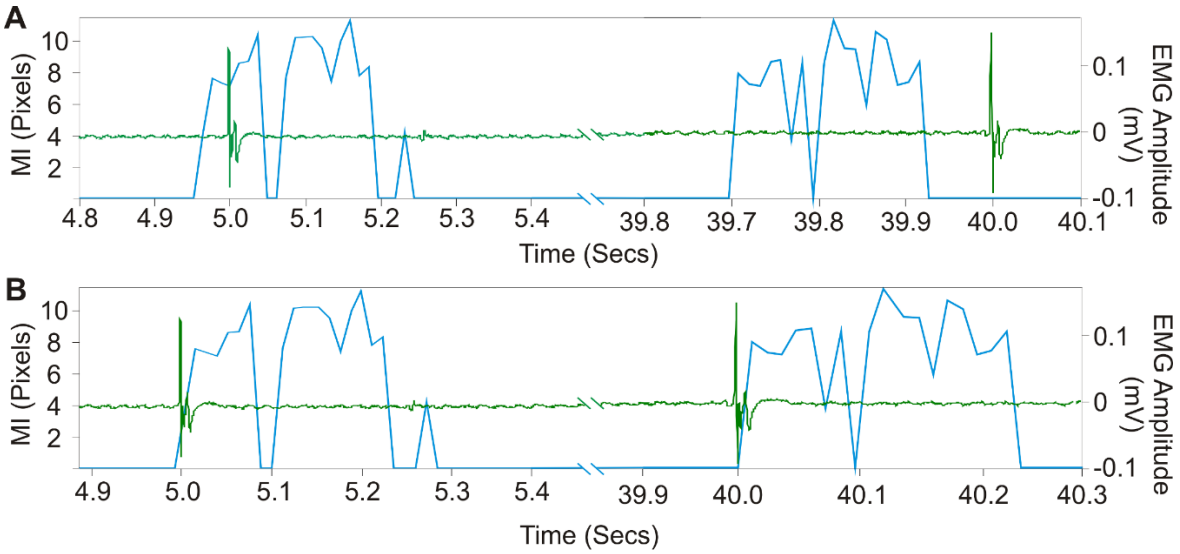

Figure 3-2. Illustration of data collected during low level electrical stimulation of MG in a healthy, adult volunteer: Surface EMG (green) is plotted together with the results of Mutual Information applied to the KLT features extracted from the ultrasound images (blue); The first and the last stimuli (at 5 and 40 seconds, respectively) are shown when: A) assuming a constant frame rate of 82 fps and B) using the time information contained in the frame headers provided by the manufacturer's proprietary software.

The IFI values extracted from metadata confirmed the experimental findings, and were actually found to follow a bi-modal distribution, regardless of the imaging parameters used. At 82 fps the majority (92.72 %) of the frames clustered around a frame rate of 83.5 fps, whilst the remaining frames clustered around 62.5 fps. When data were recorded at 42 fps the majority (85.62 %) of the frames were clustered around approximately 41.5 fps whereas the rest clustered at 36 fps. The resulting mean frame rates were 81.83 fps ± 5.38 (mean ± standard deviation, N = 17761 frames) and 40.81 ± 2.086 fps (N = 9984 frames), respectively.

The application of frame specific IFI values can be seen in Figure 3.2-B, which shows how the timing information extracted from the headers of the US image files corrects the synchronisation errors seen in Figure 3.2-A. This demonstrates that the timing correction performed did reduce the measurable delay between the electrical stimulation and mechanical activation, so that it is consistently shorter than the IFI (82 fps, or an IFI of 12.2ms), the greatest possible accuracy for this experiment.

This example justifies the need for the analysis of the timings involved in the different processes in US imaging, ranging from transmission and reception to US data storage; that will be the focus of the following subsections.

The US device used for most of the research presented in this thesis is the SonixTouch (Analogic, BK Ultrasound, USA). This device has been chosen due to its capability for very high temporal resolution (up to 40000 frames per second) and its versatility as it is almost fully customisable in terms of emission reception and post processing options. These factors allow almost total control of the different processes involved, ranging from single ultrasonic element transmission to the image formation of an entire muscle sectional area, by using C++ custom built solutions that have been programmed and adapted to the project needs.

3.1.2 Inter vs Intra frame time

The presented literature review revealed that there was a significant lack of information relating to the timing of collected image frames in many skeletal muscle studies. This lack of frame timing information can have significant implications for the quantification and interpretation of time varying physiological signals. As such, a reliable method of capturing timing information of each US frame has been developed within this research.

Clinical ultrasound devices sequencing using array probes are based on individual transmissions of a certain number of elements and the posterior acquisition of the RF data. In the simplest B-mode sequence, the ultrasound beam is moved across the face of the transducer array one line at a time. Typically, 96, 128, or 256 pulses are transmitted along the lateral direction of the array and RF data are collected for each transmission (Figure 3.3). The number of scan lines is termed the Line Density (Wikisonix (Ultrasonix) n.d.).

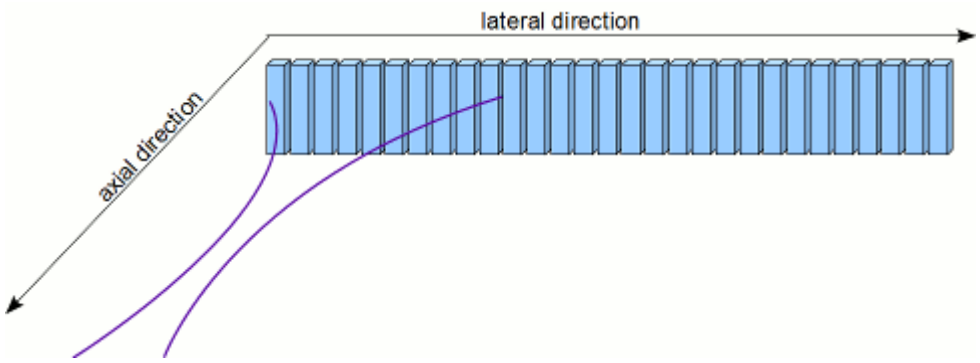

Figure 3-3. Representation of ultrasound array probe transmission. Source: (Wikisonix (Ultrasonix) n.d.).

The ultrasound propagates through most soft tissue at a speed of 1540 m/s. Therefore, to image deeper points in the tissue, the ultrasound system must wait in receive mode long enough for the echoes to return. Every ultrasound B-mode image is composed of a certain number of scan lines. The time between consecutive scanlines (defined as the intra-frame time) does not have to be constant, therefore the time between consecutive frames (inter frame interval, IFI) should not be considered constant either.

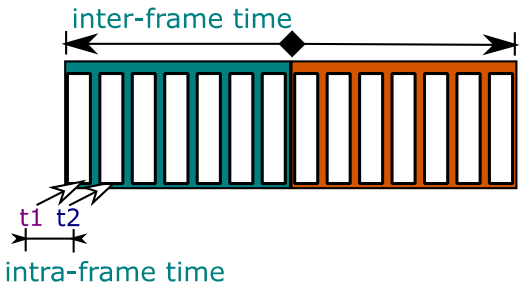

Figure 3-4. Graphic representation of the difference between intra and inter frame timings.

Temporal consistency is a term used in the field of image processing. It is defined as the lack of image information between time-points in homogeneous regions and tangentially to image gradients. This absence of temporal data can lead to variations that affect the overall optimality of the resulting deformation (Facciolo et al. 2011). Therefore, the assumption of constant inter-frame or intra-frame intervals can lead to important errors capable of invalidating a whole set of experiments, as demonstrated above (Section 3.1.1). With the mission of providing insight to these problems, an accurate study of the timings of the ultrasound device becomes

essential in order to be able to strengthen the impact of results derived from Ultrasound data.

3.1.3 Hardware Timing adaptation circuit design and methods

The most reliable and accurate way to obtain the time each US frame has been collected is by using hardware signals generated by the US machines. For the case of the equipment used in this project, the SonixTouch (Analogic, *previously Ultrasonix*, BK ultrasound, USA) provides a pulse-shaped hardware signal that is triggered every time an US transmission occurs. Figure 3.5 compares the shape of the trigger signal captured with an oscilloscope with the information provided by the manufacturer.

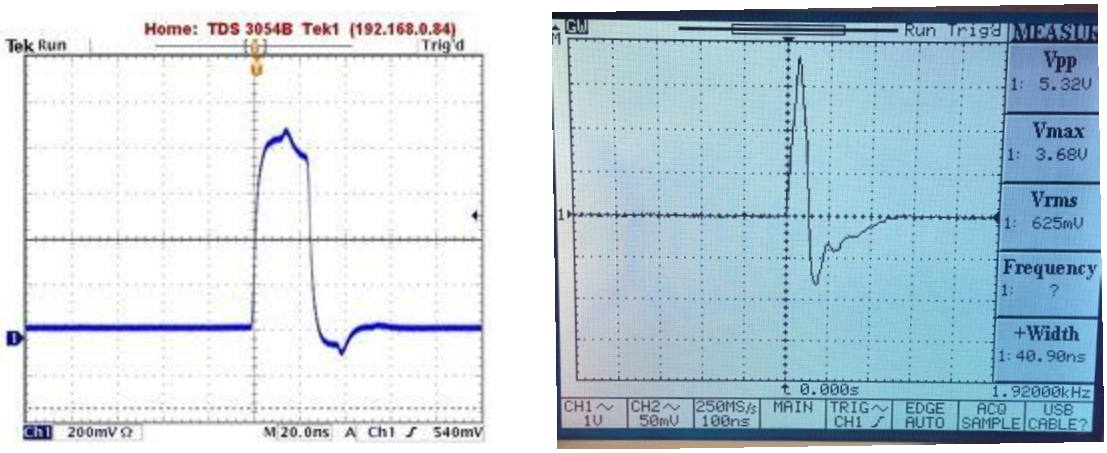

Figure 3-5. Left: Oscilloscope screen captures of the US hardware trigger provided by the manufacturer (Wikisonix (Ultrasonix) n.d.) . Right: and empirically acquired at the lab.

The hardware timing signal is a pulse signalling the end of the ultrasonic transmission of every frame; it consists of a differential TTL signal with a 25 ηs pulse length, equivalent to one cycle of the ultrasound hardware internal clock (40 MHz). The screen capture in Figure 3.5 left was taken from the manufacturer's online information (Wikisonix (Ultrasonix) n.d.). In contrast to this information, the empirically obtained trigger signal (right image), captured using a digital Oscilloscope (GWInstek GDS820-S, Taiwan), consists of a 5.32 V peak-to-peak pulse

with approximately the same length (25 ηs). Therefore, for a correct acquisition of this signal an acquisition card with a minimum acquisition rate of 80 Ms/s is required.

A signal of such a short length (25 ηs) could not be reliably acquired by our data acquisition card (USB-6212, National Instruments, USA), which operate at a maximum of 400 kS/s. As the maximum frame rate of the machine is 40,000 frames per second, the minimum interval between two consecutive frames will always be bigger than 25 μs. Therefore, the optimal solution was to design and implement a circuit capable of holding any input signal up to 12.5 μs to avoid overlap between consecutive frame triggers; resulting in an output signal easy to acquire with any conventional sample rate data acquisition (DAQ) card (80 kS/s).

The most suitable solution found, in terms of complexity and cost, was the implementation of an electronic circuit for extending the pulse width of the US hardware timing signal. The circuit consists of a 555 Timer, in a monostable configuration, along with a NOT logic gate for inversing the trigger signal (this monostable circuit triggers whenever the trigger pin is connected to ground). A schematic of the designed circuit can be found in Figure 3.6.

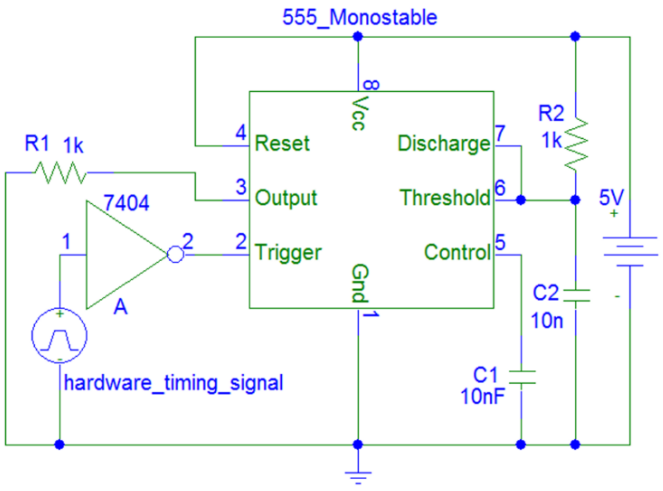

Figure 3-6. Schematic of the electronic circuit designed and implemented for extending the pulse width of the hardware timing signal to a rate achievable by the used DAQ card (USB-6212, National Instruments, USA), consisting of a 555 Monostable timer activated by the inverse (7404 NOT gate) of the hardware timing signal.

For this circuit, the time the trigger signal is held mainly depends on the Resistance (R2) and the capacitor (C2) used, following the equation: $\tau(s) = 1.1R2C2$. Where τ is the time the signal is held.

The designed circuit was tested using the software PSPICE (ORCAD, Oregon, US), a circuit simulator application for simulation and verification of analogue and mixed-signal circuits. A simulation was undertaken by generating a pulse similar to the one measured with the US machine (5.32 V, 25 ηs) obtaining the results showed in Figure 3.7.

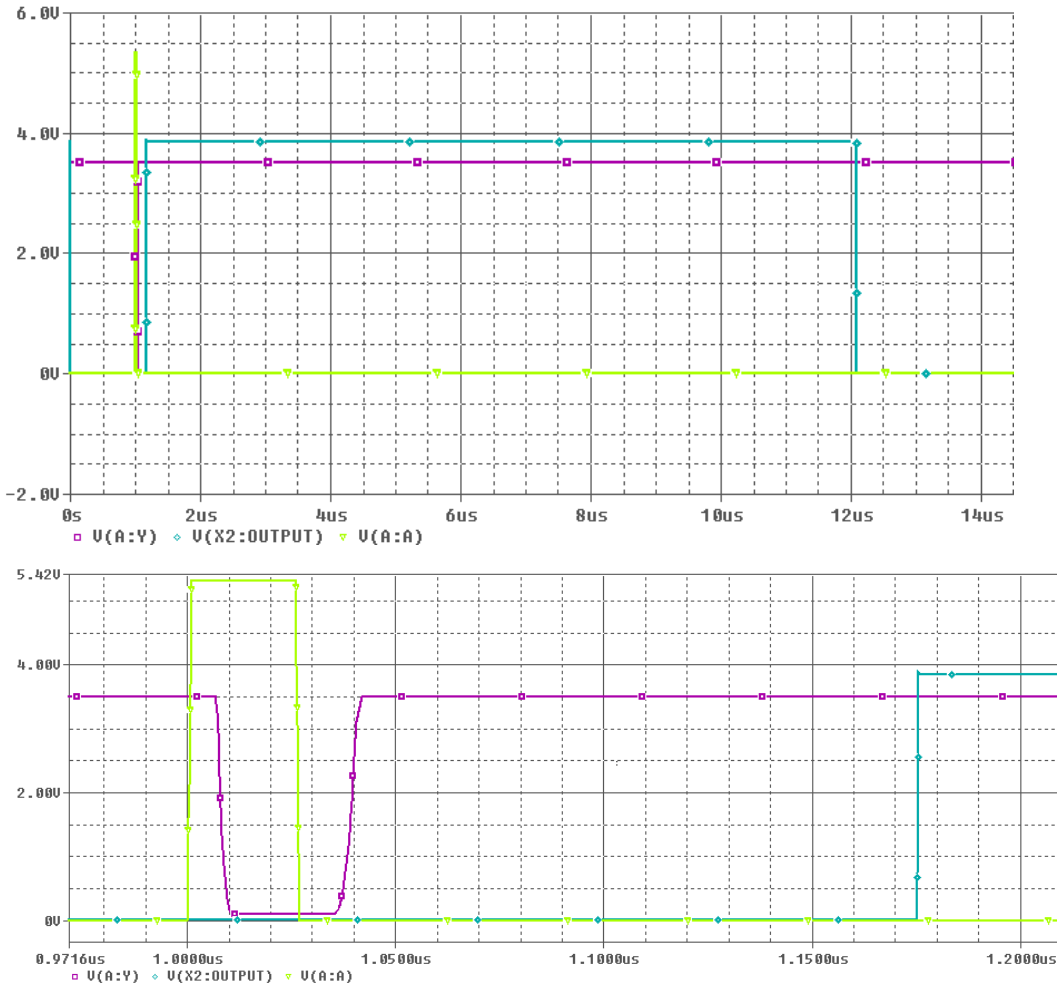

Figure 3-7. Simulation Results of the trigger conditioning circuit. Hardware trigger (light green) is inverted by a NOT gate (purple) and then held by the 555-timer until the capacitor discharges, obtaining the resulting longer trigger signal (blue). Bottom figure is a zoom of the initial one (top).

The simulation results show that the hardware trigger signal is successfully held for 11 μs after the initial impulse. The use of the NOT gate together with the 555 timer

introduces a 175 ns delay (constant) in the resulting signal. Although this delay does not affect the overall result, it should be taken into account when measuring absolute time values.

On building and implementing the circuit, empirical recording of the resulting trigger signal revealed it to have the predicted characteristics (Figure 3.8), enabling accurate identification of individual frame capture instance.

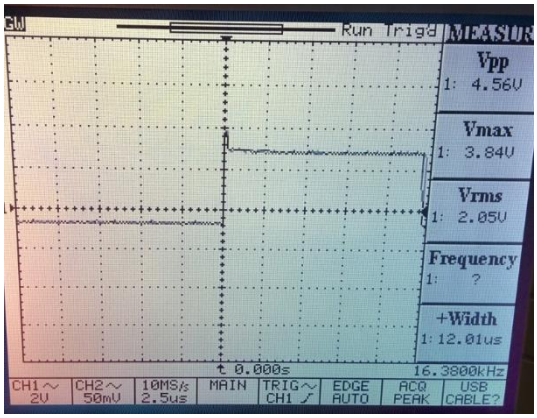

Figure 3-8. Empirical Resulting signal: R=1KΩ, C=10nF→ Time delay = 11µs.

Thus, the designed circuit both amplified and extended the duration (24 µs pulse width) of the hardware trigger making it measurable by the DAQ card and was therefore incorporated into the main data collection protocols for these studies.

3.1.4 Frame rate dependent timing in US devices

In order to understand the specific behaviour of the SonixTouch US device in terms of timing, a further study of trigger times depending on different factors was completed. For this, a comparison of the IFI at different frame rates was empirically tested using two different device configurations: i) Standard Clinical Ultrasound, Conventional US waves transmission at frame rate values below 75 fps and ii) Plane wave transmission at frame rates above 75 fps. Further explanations of these configurations are provided in Chapter 4. Frame triggers were collected, using the adapted electronic circuit explained in the previous section, and analysed in order to study their distributions.

All measurements were acquired using the same conditions in terms of depth, transmission and reception delay and the number of focal points was turned to zero. The number of ultrasound elements used was changed from 3 (4684 fps) to 256 (15 fps) in order to obtain different frame rate values. This way, frame rate was isolated from the other factors for this study

A visual example of the instant frame rate distributions during a 10 second acquisition of the slowest (15 fps) and the fastest (4684 fps) frame rates analysed can be seen in figure 3.9.

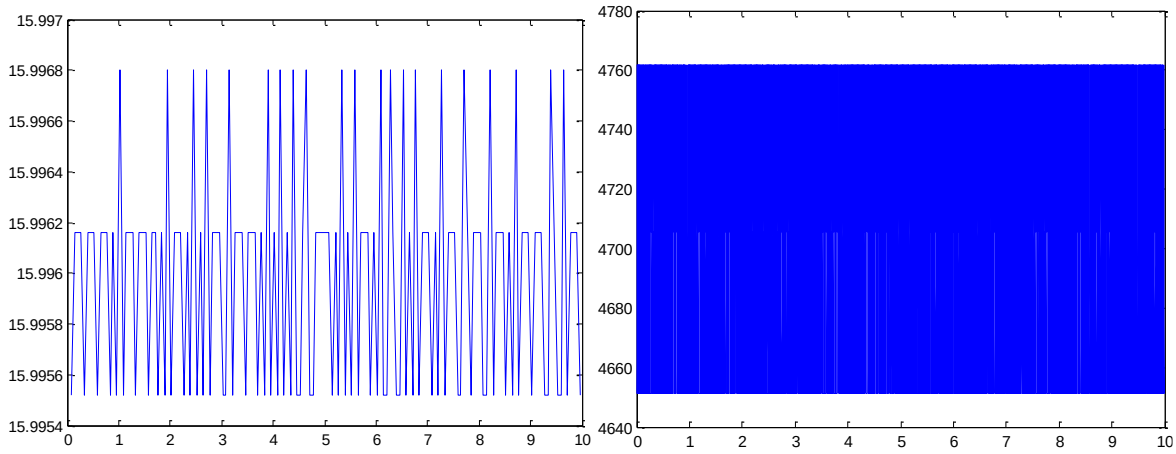

Figure 3-9. Frame rate variability between consecutives frames in 10 s duration acquisitions (X axis – time [secs]; Y axis – frame rate [fps]). Left image corresponds to 15 fps configuration i, and the right one to a 4684 fps measurements using configuration ii.

The individual IFIs for the different frame rates tested are compiled in Table 3.1.

Table 3-1 Temporal Results by analysing hardware time triggers at different frame rates.

Screened Frame Rate (fps)	#Peaks (10s)	IFI (s)	Average Frame Rate (fps)	Std. Dev (s)
15	159	62.5×10^{-3}	15.9	1.70×10^{-6}
31	318	31.3×10^{-3}	31.8	1.56×10^{-6}
54	548	18.2×10^{-3}	54.8	1.49×10^{-6}
109.8	1097	9.12×10^{-3}	109.7	1.40×10^{-6}
200	2008	5×10^{-3}	200.8	1.68×10^{-6}
501	5019	2×10^{-3}	501.9	1.22×10^{-6}
1003	10038	0.996×10^{-3}	1003.8	1.48×10^{-6}
2007	20077	0.498×10^{-3}	2007.7	1.58×10^{-6}
4684	46849	0.213×10^{-3}	4684.9	1.57×10^{-6}

No differences between frame triggers from two configurations were observed. Looking at the standard deviations in the table it is clear all the trigger signals seem to arrive equidistant, meaning that no variability on the IFI was observed for those measurements. The only remarkable difference is the one observed between the screened frame rate and the empirically collected one that differed in values below 1 fps.

Plane wave imaging

One of the requirements for achieving framerates above 1000 fps is the use of plane waves, a configuration that allows to select the amount of US piezoelectric elements triggered for the US transmission every frame. Different tests varying plane wave line duration, depth and number of plane waves were made in order to analyse the implementation time of individual processes involved in the ultrasound transmission and reception, and account for possible IFI inconsistencies to enable their use in fully synchronized setups.

In order to analyse the time execution depending on the acquired depth, acquisitions using only one element and one plane wave were performed. This way,

the frame rate was kept to the maximum (between 4014 to 7564 fps) while varying the depth from one to 10 cm. Results compiled in Table 3.2 show an IFI proportional to the depth, that appears linear when plotted (Figure 3.10).

Table 3-2 Plane wave mode timing results by increasing the acquisition depth.

Depth (cm)	IFI (s)	Time per Line (s)	Std. Dev (s)
1	1.32×10^{-4}	32.2×10^{-6}	1.39×10^{-6}
2	1.45×10^{-4}	45.2×10^{-6}	1.18×10^{-6}
4	1.71×10^{-4}	71.2×10^{-6}	1.48×10^{-6}
6	1.97×10^{-4}	97.1×10^{-6}	1.50×10^{-6}
8	2.23×10^{-4}	123.1×10^{-6}	1.72×10^{-6}
10	2.49×10^{-4}	149.1×10^{-6}	1.74×10^{-6}

"Time per Line" is calculated by subtracting the manual line duration (100 µs) from the obtained IFIs. The standard deviations obtained are below the acquisition system error (2.5 µs), therefore with the equipment employed here it is impossible to distinguish if the variations found are empiric errors or just the system accuracy.

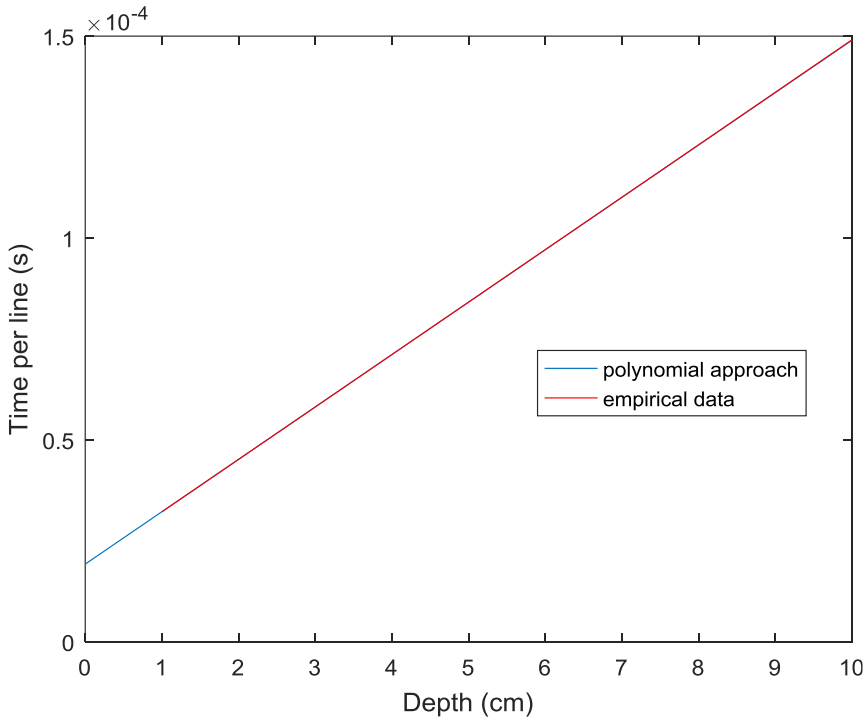

Figure 3-10. Time per line as a function of acquisition depth. A linear fit (blue) was applied to the acquired temporal data (red), and extended to the depth = 0cm.

The observed behaviour between scan line duration and imaging depth can be explained by applying a linear fit to the obtained results. This was carried out by using a Matlab function that applies a best fit in terms of least-squares to the temporal data, obtaining equation 8.

$$Scanline\ duration\ (\mu s) = 19.208\ (\mu s)\ + 12.987(\mu s/cm) * depth\ (cm) \quad (8)$$

Assuming the propagation of sound equal to 1540 m/s in human tissues, the constant multiplying the depth is equal to its inverse expressed in μs/cm as showed in equation 9.

$$\frac{1}{v_{propagation\ tissue}} = \frac{1\ s}{1540\ m} * \frac{1\ m}{100\ cm} * \frac{10^6 \mu s}{1\ s} * 2 = 12.9870\ \mu s/cm \quad (9)$$

Therefore, the independent value of the polynomial expression (offset) represents the time required for the system to process every frame and it is independent of the other transmission and reception parameters. This result relates empirically obtained hardware timings with a software process in the US device. Similar values have been previously reported (but not demonstrated) in (Wrench & Scobbie 2011) for a SonixRP machine, a different model by the same manufacturer, showing a slightly longer overhead (offset value) of 27μs. This means that different machines have different processing times that must be taken into account when analysing external phenomena in real time.

3.1.5 Software timings and comparison with hardware trigger

Hardware signals are not the only method for acquiring timing information from US devices. Other available sources include the generation of additional files containing the time stamps of the individual frames or their inclusion in the US file headers to be extracted using custom code. Due to the fact that the timing information contained in these solutions is associated with controlling system processes, for the purposes of this thesis they are termed *software timings.*

In order to study software timings and determine the amount of variability likely in experimental data using this US device, the ultrasound machine was controlled by customized software developed in C++ based on the Ulterius software development kit (Analogic, *previously Ultrasonix*, BK ultrasound, USA). Ulterius provided a method by which the number of frames transmitted since the start of acquisition could be obtained, and the time each US frame was made available to software was recorded from the processor clock (precision: 3.579 GHz →27.937 μs) of the PC. Using this approach, the behaviour of two different methods for logging US system frame timing (timestamp) was investigated in an experiment. The methods compared were the collection of hardware generated timing signals, and software timestamp data (HW and SW respectively in Figure 3.1 lower panel). Data were collected for 10 seconds per trial. No ultrasound images were analysed during this experiment, as the focus was solely to investigate the difference in time stamping methods.

In order to evaluate the time stamping performance at different frame rates, the number of elements was manually modified varying from 64 (approx. 150 fps) to 256 (approx. 37 fps). During data acquisition, the IFI from both the hardware timing signal and the software timestamps were recorded simultaneously. In addition, a control experiment was carried out, in which IFIs were recorded at three frame rates but no US images were copied to local memory. This provided a condition during which memory consumption and processor overhead were reduced to their absolute minimum, ensuring that the transfer and storage of data could not cause any variation in the IFIs recorded.

Comparison of the IFI values obtained from capturing the hardware timing signal and the software timestamps showed inconsistencies (Figure 3.11). Clear discontinuities in the IFIs from software timestamps were evident in all but the slowest frame rates tested, meaning that frames were "dropped", i.e. not able to be collected. When data were collected at the slowest frame rate (37 fps) although no frames were dropped (Figure 3.11), closer inspection of the software timestamp signal reveals inconsistency in the IFI did still occur (Figure 3.12). In contrast, IFIs

from hardware timing signals are shown to be constant across the length of each trial (Figure 3.11), which confirms that ultrasound transmission did take place, regardless of whether image data were made available for collection or not.

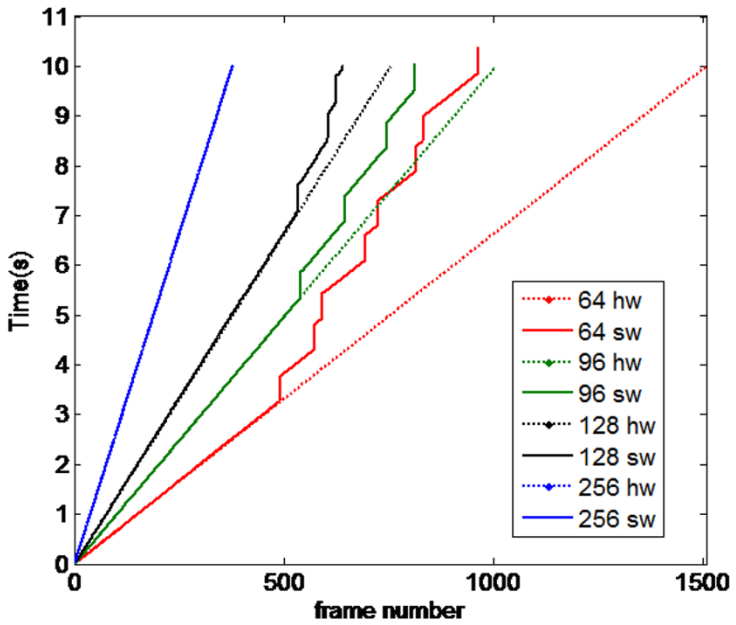

Figure 3-11. Hardware (dotted lines) vs. Software (flat lines) frame time stamps from 10 second long acquisitions using 64 (red), 96 (green), 128 (black) and 256 (blue) elements (resulting on 150.7, 100.5, 75.4 and 37.7 fps frame rates respectively) for the Analogic SonixTouch machine.

During tests when no data were transferred to the local memory, i.e. when processing time and memory consumption were reduced to the bare minimum, no frames were dropped for any of the analysed frame rates. The software timestamps and hardware timings signals collected during these experiments were still found to differ though. The distribution of the hardware timing signal exhibited a variance lower than the measurement accuracy at which the data were collected (25 μs) meaning the transmission was constant in time. The software timestamps however show a great deal of variance, proportionally increasing with frame rate. This shows that, even with minimal computational overhead, the IFIs recorded via software timestamps are not as consistent as those recorded by the hardware timing signal and the implications of these findings are discussed further below.

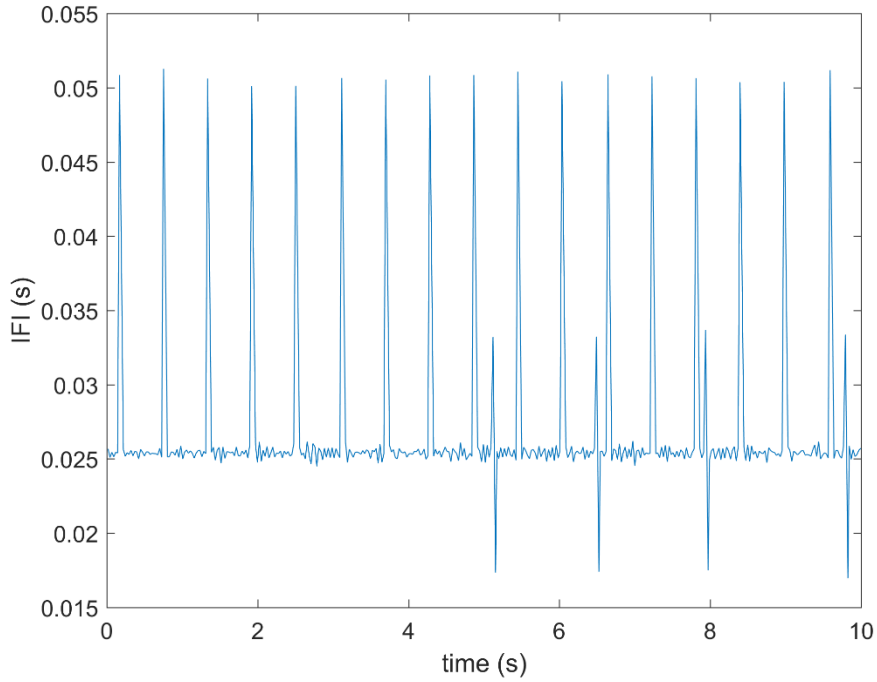

Figure 3-12. IFIs calculated from software timestamps from a 10 second duration acquisition using 256 elements (resulting on 37.7 fps frame rate) for the Analogic SonixTouch machine. A bimodal distribution of IFIs comprising ~0.025 and ~0.05 seconds is evident in the first 5 seconds of acquisition, with more variability occurring between 5 – 10 seconds.

3.1.6 US devices temporal behaviour discussion

This work addresses two key questions posed in relation to variations in inter-frame intervals (IFI) in recorded ultrasound image sequences. It has been shown from empirical data (Figure 3.2), for the second machine tested (Telemed Logicscan 128), that assuming a constant IFI can lead to dissociation of measurements; to such an extent that a muscle is shown to respond to an electrical stimulation prior to the delivery of the stimulus (Figure 3.2-A). Whilst this demonstration provides a somewhat extreme example, it illustrates the consequences of the inappropriate relationships (mechanical response of muscle preceding electrical activation) which can be observed if variations in IFI are not accounted for. One approach to mitigate these issues is the reduction of the acquisition duration; however, it has been observed that within an ultrasound sequence of just 5 seconds the error had accumulated to more than four times the IFI (49 ms). A second approach, which has been implemented, is to employ the accurate timing data provided by the

ultrasound device (Figure 3.2-B). This option however, depends on the provision of that utility by the manufacturer and therefore may not be available across all commercial devices. The facility to access accurate frame time information should therefore be a consideration when new devices are purchased for the laboratory, particularly if tissue dynamics are likely to be the study focus. For cases when equipment is already present in the laboratory, but accurate frame timing information is not available from the device, a Matlab utility has been provided for performing the analysis described in the methods, available from "http://uk.mathworks.com/matlabcentral/fileexchange/59006-ultrasound-find-best-framerate". The code enables users to assess whether or not data, collected using the stimulated muscle twitch protocol described, were acquired by an US machine with a non-constant IFI and allows the correction of constant delays present in the data collection.

In the case of the first machine tested (Analogic SonixTouch), the observed frame losses in the software timestamps show that, while ultrasound beam transmission was consistent (Figure 3.11), it was not possible to save all frames to memory. The frame loss is inversely proportional to the IFI, indicating that the system memory saturated at rates proportional to the amount of frames collected. Despite these issues, variations in the IFIs (frame loses) recorded by software time stamping were found to have zero mean (i.e., they were non-cumulative), therefore to assume a constant IFI was valid. This said, the fact that frames could be lost during an acquisition must be accounted for to ensure synchronisation is maintained throughout the duration of an experiment. In the work presented here, the identification of the frame losses was achievable via the control software, but this could also easily be achieved via the comparison of software timestamps and the hardware timing signal. The frame losses have wider implications when extracting dynamic information from ultrasound data. During tissue displacement measurements, these "jumps", if not accounted for, would result in a twofold increase in the resulting measurements (assuming constant movement).

In summary, it has been shown that the stated temporal consistency of ultrasound devices cannot necessarily be relied upon. In tests of two machines, it has been demonstrated that one machine (Telemed Logicscan 128) has highly variable IFIs, but provides a method by which reliable IFIs may be extracted from recorded ultrasound sequences. In contrast, the other machine (Analogic SonixTouch) was shown to have an IFI that was indeed constant (or at the least had a lower variability than our measurement accuracy) but where the most accessible method by which IFIs could be recorded (SW timestamp was found to reveal *false variability* (Figure 3.12). In the case of this machine, while transmission and reception were constant, hardware limitations resulted in some frames being irretrievable due to memory saturation (Figure 3.11).

Consideration of these findings in the design and implementation of experiments is critical for the reporting of reproducible data and interpretation of results in relation to physiological phenomena. My results show that synchronously starting data acquisition across multiple devices (e.g. ultrasound, electromyography, dynamometry) is not adequate for correct intra-trial synchronisation, a factor which may have contributed to the discrepancies in tendon hysteresis values reported between animal and human studies (Finni et al. 2012). Reduced intra-trial synchronisation may also be a contributing factor in the variation in muscle behaviour identified using other ultrasound modalities (e.g. large variations reported in the delay between EMG determined activation and tissue Doppler/motion mode based muscle movement reported across two recording days as: -20.3 ± 21.0 ms and 17.4 ± 27.2 ms, (Pulkovski et al. 2008)). The importance of intra-trial synchronisation is also critical when considering recent development of computational ultrasound image analysis techniques for applications such as diagnosis and monitoring of features of neurodegenerative diseases (Harding et al. 2016). In that report, computational analysis provided an automated means of detecting tissue displacements (twitches) resulting from involuntary activations (fasciculations), but the work should be extended to evaluate performance of the approach against recorded myoelectric signals (current clinical gold standard), which will need to be recorded in a fully synchronous manner. In addition, the work

could be extended to provide additional biomarkers of disease progression by quantifying the spatial and temporal features of recorded twitches where, again, intra-trial synchronisation will be an important factor to consider.

3.1.7 Conclusion

The work in this section provides the tools needed for correcting any temporal anomalies related with the US machine used. This ensures that any phenomena observed in the US data collected shall not be related with synchronisation problems, as the use of the hardware trigger timing has proven a reliable source for time stamping the collected US frames. Thus, all the US data empirically collected using the SonixTouch device will use the trigger timing circuit described above and include a text file logging the precise time every US frame was collected in order to provide reliable measures comparable with external data acquired in synchrony with US.

The next sections in this chapter will focus on the implementation of fully automated analysis methods able to extract dynamic properties of the muscular system from US data in a manner that is unbiased by human operator errors.

3.2 Active Shape Models

In order to remove the reliance on manual region of interest (ROI) identification for US based image segmentations, and fully automate the approach, an option is the use of techniques such as Active Shape Models (ASM). This technique is able to localise deformations of original shapes in order to fit new images, and has proven to be a reliable option for the identification and localisation of individual muscles in B-mode images (Darby et al. 2012; Cunningham, P. Harding, et al. 2017).

Within this section, an explanation of the ASM that was created by (Cunningham 2015) is detailed, following the steps which are summarised below. This model has been posteriorly validated with the participant ultrasound data shown in chapters 5,

6 and 7. The main objective of the implementation of ASM within this thesis is to find two bright lines within the US image (aponeurosis characterised as a strong echo) surrounded by low brightness regions (no media changes). In other words, it looks for three consecutive regions (low – high – low) with significant contrast in brightness within a certain depth in the B-mode images.

3.2.1 Creation of Active Shape Models from manually labelled data

B-mode images at 25 fps of the lower calf muscles of 20 participants (11 male, 9 female with ages between 18 and 50 years) were collected for passive, active and combined contractions during ankle rotations. Images were collected using an US device (Aloka ProSound SSD-5000, UST-5712 probe) in order to characterise the position and inclination of the superficial and deep aponeuroses of Soleus and Gastrocnemius Medialis (GM) muscles. The US images were manually labelled by an expert operator defining the x and y coordinates of 4 points per aponeurosis using a Matlab GUI designed for that purpose (Figure 3.13), making a total of 8 data points per frame, for 10 frames per participant acquisition (i.e. a total of 200 frames marked from the 20 participants).

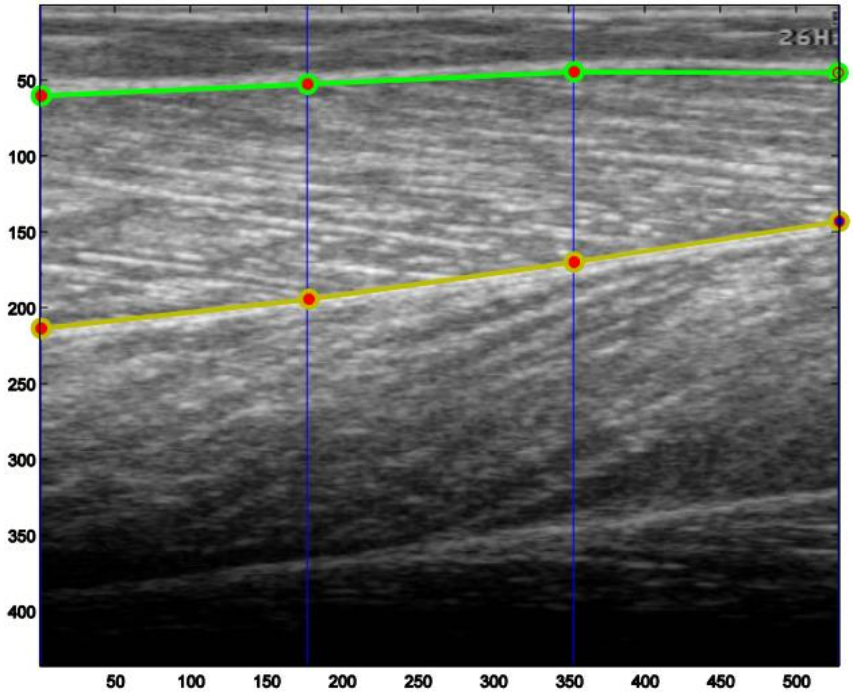

Figure 3-13. Image annotation GUI from (Cunningham 2015) over a B-mode image including the manual definition of 4 points (red) per aponeurosis (superficial in green and deep in yellow) of the GM muscle of a participant. Image represents a 5.94 x 5 cm cross section area.

Procrustes analysis was applied by aligning the coordinates of the upper and lower aponeurosis data in order to minimise the inter-point distances. For each aponeurosis, all the frame by frame curves of all the participants were spatially superimposed. The fact that the x coordinates of the delimiting points were the same in every labelled frame simplified this procedure to a one-dimension alignment (Y plane). Characteristic aponeurosis curves were selected and used as the training set in order to contain significantly different motions and shapes. As a result of this process, a matrix containing the variance of the different contour lines (also known as covariance matrix) was created.

After the alignment of the shape contours, PCA was applied in order to reduce the amount of variables involved in the prediction and enable a faster and more efficient segmentation process. The first principal component is the combination of variables (4 y coordinates per aponeurosis contour for this precise case) that accounts for the biggest part of the overall variance contained in the covariance matrix. The following principal components (e.g. second principal component) are the linear

combination of variables that, having a correlation of 0 with the previous component, explain the largest amount of the remaining variance. These principal components correspond with the eigenvectors (components) and the eigenvalues (variances) of the covariance matrix. Here, the principal components for the relative vertical distance between the two aponeurosis contours and for their position within the ultrasound image was calculated, and named as the location model.

3.2.2 Evaluation of ASM performance with new participant US data

The ASM approach explained above was tested with new participant data in order to evaluate its performance for the specific data that would be collected within the experimental protocols (detailed in Chapter 4). The parameters needed for the application of the model to new data were stored in a text file, containing the variance of the different aponeurosis boundaries in the training set. The different parameters were divided in three groups:

- Location model: information regarding the relative vertical distance (depth in the US image) between the two aponeuroses and their positions within the US image;
- Shape model: information of the mean position and standard deviation of each one of the points belonging to the two aponeuroses.
- General parameters: information regarding acquisition parameters such as height and width of the ultrasound image (in mm) or the number of iterations used for finding the best fit.

A Matlab function was programmed implementing the ASM in order to automatically segment the GM muscle. Firstly, two random (upper and lower aponeurosis) were generated within a region delimited by the eigenvectors and eigenvalues of the covariance matrix. Due to the fact that the GM aponeurosis constitutes a bright area in contrast with the rest of the B-mode image, the function compares the intensity of the pixels above and below the created line in order to quantify the contrast before and after the GM. If the brightness contrast was bigger than the current

maximum (being zero for the first iteration), the last generated line was stored as the new current maximum. These steps were repeated until the counter reached the maximum number of iterations (5000 default). The maximum contrast neighbourhood lines at the end of the loop (once reached the maximum number of iterations stated in the ASM text file) is assumed as the best fit for delimiting the GM muscle (figure 3.14) for the first frame of the ultrasound footage.

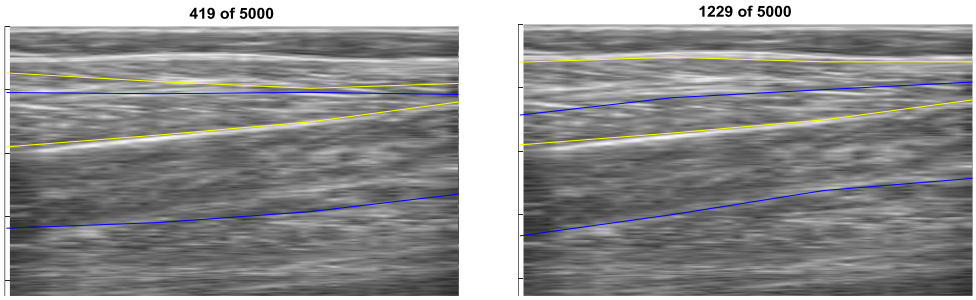

Figure 3-14. Superficial and deep aponeurosis located using ASM in a B-mode image of gastrocnemius medialis: Blue lines represent the generated shapes for that iteration and yellow lines represent the current maximum.

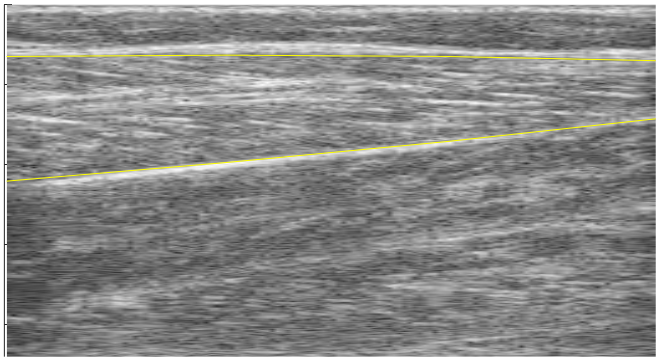

Figure 3-15. Best fit of the superficial and deep aponeurosis located using ASM in a B-mode image of GM.

A similar procedure is followed for locating the superficial and deep aponeurosis in the following frames. Due to the fact that movements of the aponeurosis between consecutive frames are reasonably small (below pixel range) during data collection using conventional frame rates (20-80 fps), both the variances and number of maximum iterations used for finding the best fit for the first image are reduced for calculating the remaining. This way, the algorithm becomes more efficient and faster.

Pilot data collected for one participant following the protocols described in chapter 4 was segmented for the GM muscle using this automatic procedure for one isokinetic dorsiflexion, one isometric contraction at 60% MVC and an isotonic movement at 30% MVC (muscle tasks detailed in Chapter4) collected at 32 frames per second. The resulting ASM segmented images were evaluated against manual annotations that consisted of marking seven points per aponeurosis.

ASM error was evaluated in terms of percentage accuracy as described in (Cunningham 2015):

$$P = \frac{S \cap A}{\mu(A)} \times 100 \tag{10}$$

, where P is defined as the proportion of the area of the expert annotation A that intersects with the area of the automatic segmentation S, expressed as a percentage, and μ(A) is the area of the expert annotation.

Figure 3.16 graphically shows the distribution of the percentage accuracy for the three muscle movements studied.

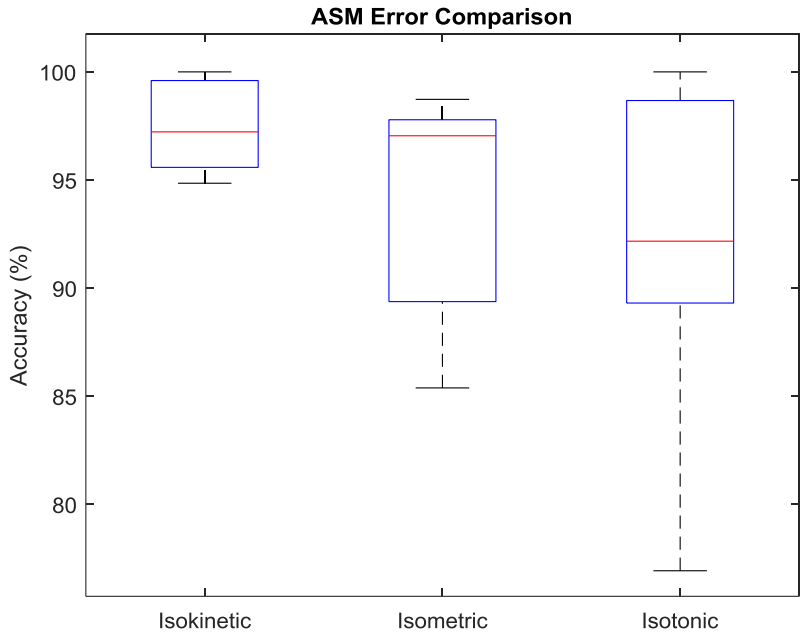

Figure 3-16. Boxplot comparison of the percentage accuracy distribution (equation 10) for the 3 muscle movements at a frame rate of 32fps. Manual annotations applied to 384 US images from 2 participants.

The resulting percentage averages and standard deviations consisted of 97.49 ± 2.05 % for isokinetic passive movements, and 94.15 ± 4.74 % and 92.13 ± 7.17 % for isometric and isotonic muscle contractions respectively.

Furthermore, the performance of this technique has been also applied to neck muscles (Cunningham, P. Harding, et al. 2017), but expressed in terms of Jaccard index (J), which gives the proportion of the combined areas of two arbitrary polygons that overlap; and the Dice similarity coefficient (D), which is more popular and less discriminating than the Jaccard index, and compares the similarity of two samples. These parameters are described as:

$$J = \frac{|S \cap H|}{|S \cup H|} \qquad D = \frac{2|S \cap H|}{|S| + |H|} \tag{11}$$

where S refers to the expert annotations (assumed as ground truth), and H is the automatic segmentation (ASM). The resulting values of Jaccard index and Dice coefficient for the three muscle movements are compiled in Table 3.3.

Table 3-3. Average and standard deviation results for the Jaccard index and Dice similarity coefficient between ASM and expert notations during isokinetic passive movements and isometric and isotonic contractions.

	Jaccard index		Dice coefficient	
	Mean	std	mean	std
Isokinetic	0.952	0.0121	0.9754	0.0064
Isometric	0.9246	0.0523	0.9601	0.0288
Isotonic	0.9146	0.072	0.954	0.0406

The average Dice similarity coefficient value for the GM muscle during three muscle movements (D = 0.9632 ± 0.0253) was 10% superior to its application in neck muscles (96% vs 86%). This good performance declined us to use it as the segmentation method for the empirically collected ultrasound data in this thesis.

3.3 Quantification of muscle movement from US image sequences with Optical Flow

This section details the approaches implemented to quantify muscle movements captured in the US data. As detailed in Chapter 2, previous studies indicate better

suitability of OF based algorithms for the analysis of ultrasound images compared to other approaches such as BMA (REFS). Furthermore, it was also concluded that differential OF algorithms were more adequate for the extraction of dynamic properties from ultrasound images and they therefore form the basis of work presented here.

Horn Schunck (HS) and Lucas-Kanade (LK) have previously been proven reliable for motion tracking in Synthetic Ultrasound images (Baraldi et al. 1996). While both techniques assume constant brightness patterns in time, they offer different approaches in both the smoothing term used and in whether they are applied to the local (LK) or global (HS) domain.

Due to their nature, one of the most important drawbacks of HS and LK techniques is their performance when computing large motions. Pyramidal approaches (down-sampling of a frame by a factor of two per level in both height and width, in order to handle larger displacements) have proven reliable for computing larger optical flows (Bouguet 1999), but their use implies much higher computational requirements therefore their efficiency is substantially reduced. Region-based methods work better at lower signal to noise ratios and important movements but assume constant motion in neighbouring areas (not suitable for local motions) when compared with differential methods (Duan et al. 2005). Therefore, details of the implementation and application to US data of these two differential OF algorithms is provided below, including an analysis of their performance compared with external dynamic data.

3.3.1 OF Background Information

In order to have a better understanding of the behaviour of different OF algorithms, some of the basics that rule these techniques are explained here. The function that computes the position shift between two frames is:

$$f(x, y, t) = f(x + dx, y + dy, t + dt) \tag{12}$$

where dx and dy are the displacements in both X and Y axis, and dt is the time interval between frames. When applying Taylor Series Expansion to the equation above, we obtain the following:

$$f(x,y,t) = f(x,y,t) + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial t}dt$$

$$\rightarrow \ \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial t}dt = \ f_x dx + f_y dy + f_t dt = 0. \tag{13}$$

If we divide every one of the resulting terms by the derivative of time (dt) we obtain the so called *Optical Flow Constraint Equation*:

$$f_x u + f_y v + f_t = 0 \tag{14}$$

where u and v are the velocities of a pixel in the X and Y axis respectively. Therefore, both velocities can be expressed as a function of the other and of the time:

$$v = \frac{f_x}{f_y}u - \frac{f_t}{f_y} \quad and \quad u = \frac{f_y}{f_x}v - \frac{f_t}{f_x} \tag{15}$$

3.3.2 Horn and Schunck

Horn and Schunck were pioneers in the computation of motion between frames (Horn & Schunck 1981). In the algorithm they developed two main constraints are assumed:

- *Brightness Constancy Assumption*: The intensity of the same point is constant throughout the different frames that compose a video, meaning that their brightness does not vary over time;

- *Smoothness Constraint*: All the pixels that belong to the same moving object move in the same magnitude and direction.

However, both constraints would have small errors, meaning that the brightness of a precise pixel will slightly vary between frames and that the motion of pixels that belong to the same object would vary in real images. Therefore, the main strategy in HS is the minimisation of these two assumptions as described in the equation below:

$$\iint \left\{ \left(f_x u + f_y v + f_t\right)^2 + \lambda\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right) \right\} dxdy \quad , \tag{16}$$

Data Term Smoothness Term

where the data term is the Optical Flow Constraint Equation squared. This equation can be solved by differentiating with respect to the two line velocities u and v:

$$\begin{aligned} \left(f_x u + f_y v + f_t\right) f_x + \lambda(\Delta^2 u) = 0 \\ \left(f_x u + f_y v + f_t\right) f_x + \lambda(\Delta^2 v) = 0 \end{aligned} \tag{17}$$

As a result, this system is solved by applying the Laplacian to the resulting equations:

$$\Delta^2 u = u_{xx} + u_{yy} \tag{18}$$

Derivative masks are then applied in order to solve the equations and obtain the resulting flow in both X and Y components.

In order to use HS optical flow techniques within the data analysis previously developed, customised Matlab (Mathworks, USA) code was implemented with the use of some functions of the Computer Vision package in order to apply a version of Horn and Schunck technique to ultrasound images collected from the Gastrocnemius Medialis (GM) during different contractions and passive movements.

A block diagram of the iterative process is shown in Figure 3.16, detailing the different temporal and spatial kernels applied.

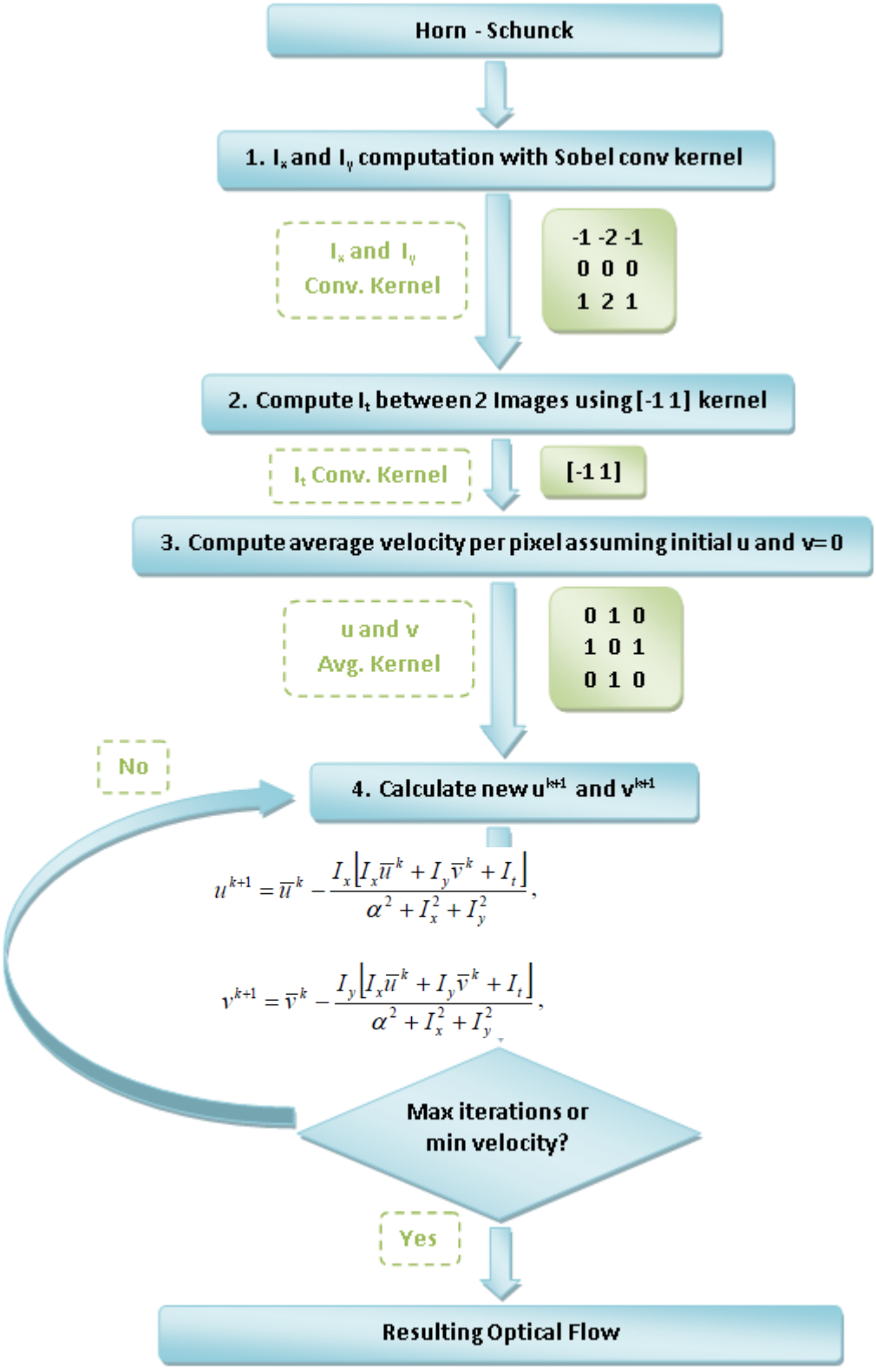

Figure 3-17. Block Diagram of Horn and Schunck Optical Flow Algorithm applied in my research. I_x, I_y and I_t represent the image brightness derivatives for the x and y directions and time respectively.

3.3.3 Lucas-Kanade

The OF constraint equation can also be written in relation to the time derivative as:

$$f_x u + f_y v = -f_t \tag{19}$$

Within a small neighbour of n by n pixels, and assuming smoothness constraint, every pixel would become one out of n equations, providing an over determined system (more equations than unknowns) that would look like this in Matrix form:

$$\begin{bmatrix} f_{x1} & f_{y1} \\ \dots & \dots \\ f_{xn} & f_{yn} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -f_{t1} \\ \dots \\ -f_{tn} \end{bmatrix} \Leftrightarrow Mu = f_t \quad , \tag{20}$$

where ***M*** is the n by 2 Matrix of the X and Y axis displacement in every pixel within the neighbourhood, and u is the 2 by 1 matrix containing the flow in both directions.

To compute the optical flow in every pixel the approach would be as simple as putting the matrix M to the right side of the equal, but due to the fact that M is not a square matrix the inverse of M cannot be calculated. In order to solve this, the pseudoinverse of the matrix is calculated by multiplying both terms by the transpose of M:

$$M^T M u = M^T f_t \Rightarrow u = (M^T M)^{-1} M^T f_t \tag{21}$$

In the Lucas-Kanade algorithm, the position shift of a point or pixel between consecutive frames is assumed smaller than a pixel (assuming abrupt changes as assumable outliers). With a frame rate big enough supposed, a way of simplifying this equation would be to assume that the time interval is almost zero. In order to do this, we apply a minimisation of the square error of the displacement of the pixel movement within a neighbourhood, that is, the same "Data Term" used as the first constraint of the HS algorithm, also known as Least Squares Fit:

$$min \sum_{i}^{n} \left(f_{xi}u + f_{yi}v + f_t\right)^2 \tag{22}$$

When minimising we approach the derivatives in x and y to 0 and then write them relative to the temporal derivative f_{ti}:

$$\begin{aligned} \sum_{i}^{n} \left(f_{xi}u + f_{yi}v + f_{ti}\right) f_{xi} = 0 \Rightarrow \sum_{i}^{n} {f_{xi}}^2 u + \sum_{i}^{n} f_{xi}f_{yi}v = -\sum_{i}^{n} f_{xi}f_{ti} \\ \sum_{i}^{n} \left(f_{xi}u + f_{yi}v + f_{ti}\right) f_{yi} = 0 \Rightarrow \sum_{i}^{n} f_{xi}f_{yi}u + \sum_{i}^{n} {f_{yi}}^2 v = -\sum_{i}^{n} f_{yi}f_{ti} \end{aligned} \tag{23}$$

Unlike the previous example with the M matrix, there is no need to use the pseudoinverse matrix, as our minimisation matrix is square. Therefore, the motion flows u and v can be expressed in the following form:

$$\begin{aligned} u = \frac{-\sum f_{yi}^2 \sum_i^n f_{xi}f_{ti} + \sum_i^n f_{xi}f_{yi} \sum_i^n f_{yi}f_{ti}}{\sum_i^n {f_{xi}}^2 \sum_i^n {f_{yi}}^2 - \left(\sum_i^n f_{xi}f_{yi}\right)^2} \\ v = \frac{\sum_i^n f_{xi}f_{ti} \sum_i^n f_{xi}f_{yi} - \sum f_{xi}^2 \sum_i^n f_{yi}f_{ti}}{\sum_i^n {f_{xi}}^2 \sum_i^n {f_{yi}}^2 - \left(\sum_i^n f_{xi}f_{yi}\right)^2} \end{aligned} \tag{24}$$

Analogously to the HS technique, a version of Lucas Kanade Algorithm has been programmed in Matlab code following the block diagram in Figure 3.17. The performance of this OF algorithm in the analysis of real Ultrasound B-mode images of GM is shown in the following section.

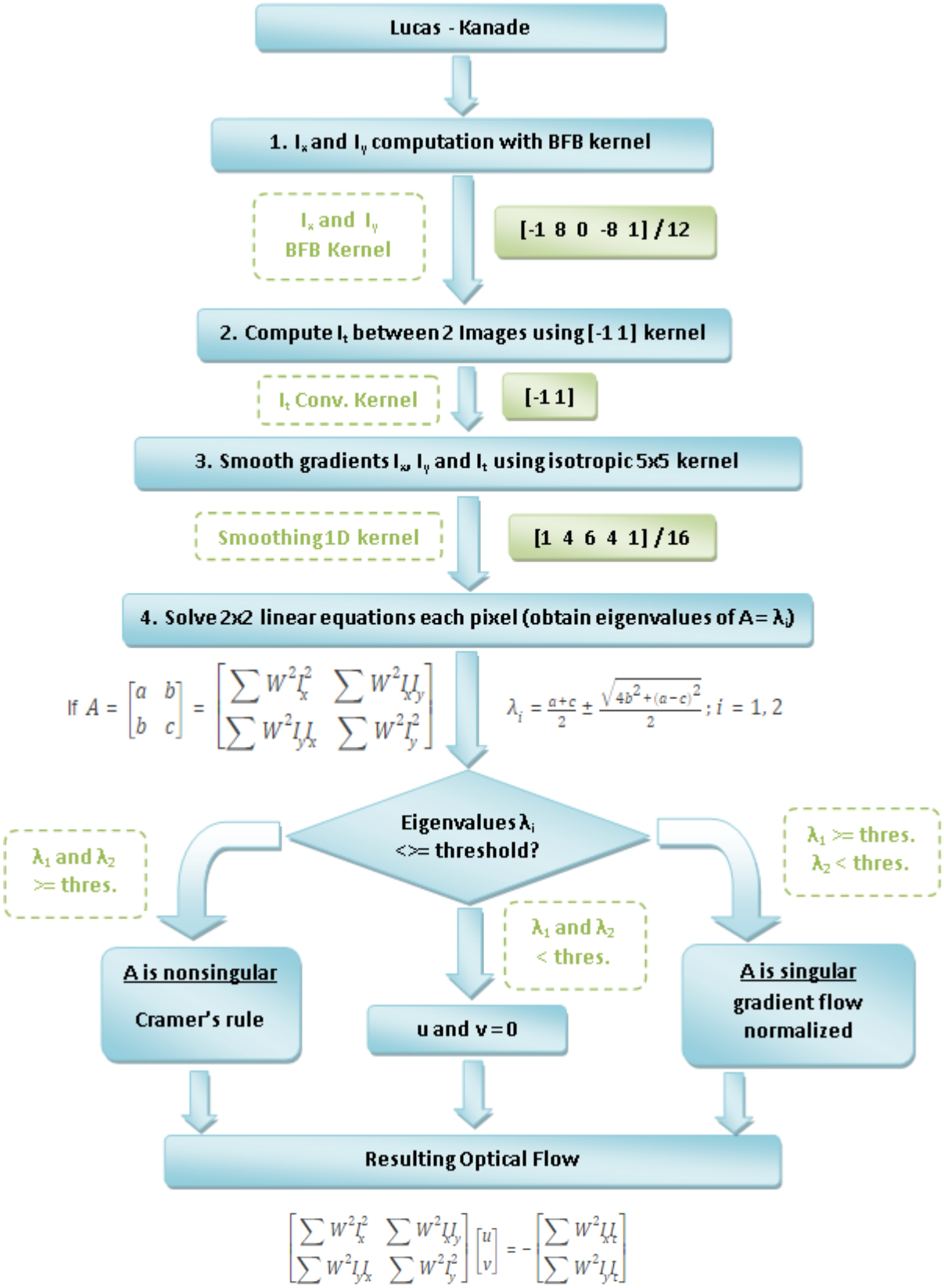

Figure 3-18. Block Diagram of Lucas and Kanade Optical Flow Algorithm applied for our research. W represents a window function that accentuates the constraints at the centre of each section.

3.3.4 Optical Flow Code Implementation and Analysis

A pilot study of one participant was undertaken in order to provide a preliminary evaluation of the performance of the chosen OF algorithms. Ultrasound data were collected for eight seconds from the medial gastrocnemius muscle (GM) of the left leg at a framerate of 32 fps when performing one isometric contraction at 30% of the Maximum Voluntary Contraction (MVC) for six seconds before going back to the relaxation state. Joint angle and generated torque were acquired using an isokinetic dynamometer chair. The ultrasound probe (linear array, 59 mm long, 7 MHz latent frequency) was aligned with the fascicle plane and held in position using elasticated bandage.

As explained in Section 2.1, parameters such as fascicle orientation and fascicle length influence the force generating capacity of a muscle. For ankle movements, even considering force contributions from other muscles (Bojsen-Møller & Magnusson 2015), the force generated by the GM muscle should relate to the net joint torque (Landin et al. 2015). Therefore, movement vectors extracted from the GM muscle belly using OF algorithms should also be related with the muscle fascicle length and angle and thus to the joint torque acquired with the dynamometer.

Lucas-Kanade and Horn-Schunck optical flow were calculated for every pixel located inside the GM. The muscle was isolated from the skin layers and other muscles, using the Active Shape Model explained in the previous section. The OF techniques output an estimation of the direction and speed of moving objects in both polar and Cartesian coordinates (Figure 3.19).

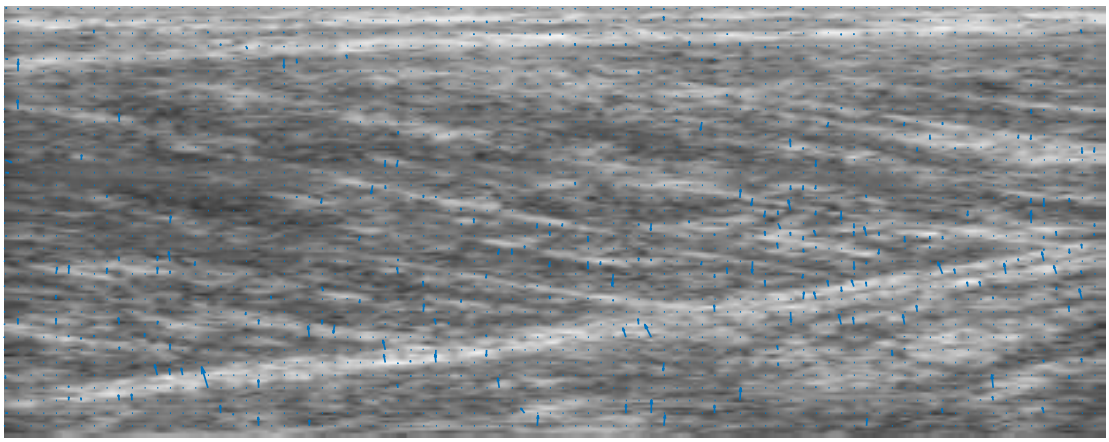

Figure 3-19. Frame to frame Optical Flow. Blue "→" represent the vectors connecting the pixel shift between two consecutive frames during an isometric contraction. The image represents a 6x2.5 cm muscle cross section.

The resulting movement vectors were decomposed into X and Y Cartesian coordinates, and their relative "per frame" contribution stored as a cumulative sum (in order to account for the overall displacement rather than relative between consecutive frames). The resulting X and Y absolute displacement coordinates were then compared with the generated torque acquired with the dynamometer (Figure 3.20).

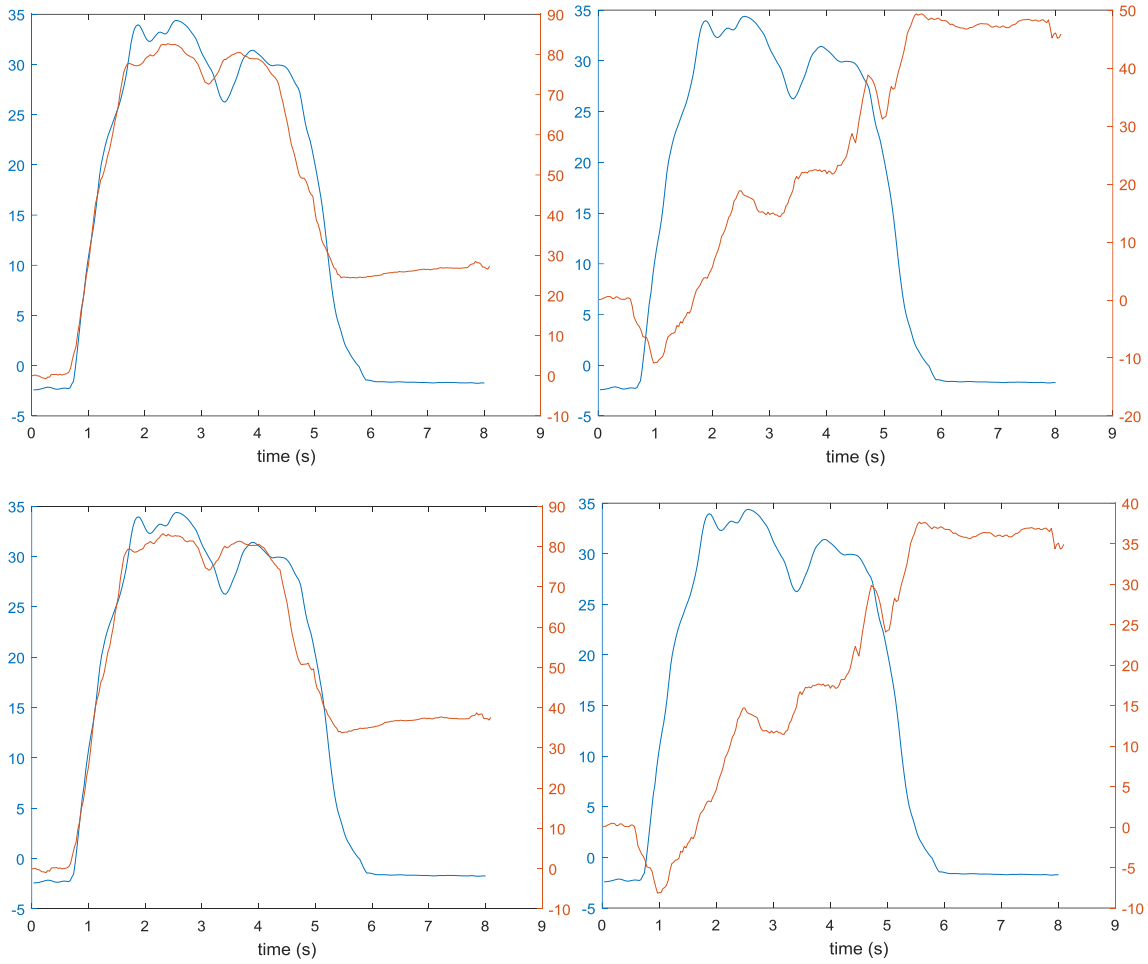

Figure 3-20. Comparison of the acquired joint Torque (blue) with the X axis (left) and Y axis (right) Cartesian components of HS (up) and LK (down) Optical Flow displacements (red) during isometric contraction at 50% MVC.

OF results showed proportional variations to the obtained torque that were particularly representative of the joint torque in the X axis direction. In order to evaluate the goodness of fit between OF motion and generated torque, a fitting approach was calculated for each resulting Torque vs. US motion binomial. Initially, a polynomial of order one that best fit the resulting torque and the OF results in a least square values sense was estimated for each trial. Next, two fitness indicators were used to describe how similar the OF approach is to the joint Torque:

- Normalised Root Mean Square Error (NRMSE) was calculated to provide an indicator of the fit between the muscle movement and the torque values:

$$NRME\ fit = \frac{\|OFx - Torque\|}{\|OFx - mean(Torque)\|} \quad (25)$$

where OFx is the OF result for the x coordinate and Torque is the generated torque acquired by the dynamometer chair. NRMSE fit values vary between -infinity (worst fit) to 1 (perfect fit). If the cost function is equal to zero, it means that x is no better than a straight line at matching Torque mean.

- Root Mean Square Error (RMSE) outputs an average error in the Torque estimation per sample. This error value is expressed in the same units as the Joint Torque (N m)

$$RME\ fit = \sqrt{\frac{\sum_{i=1}^{n}(Torque_i - OFx_i)^2}{n}} \quad (26)$$

, where n is the number of samples.

As a result of this process, Figure 3.21 graphically shows that the performance of Horn-Schunck and Lucas Kanade Optical Flow algorithms during isometric contractions provide a very good fit with the generated net joint torque, therefore demonstrating an association between the OF motion results and the joint torque.

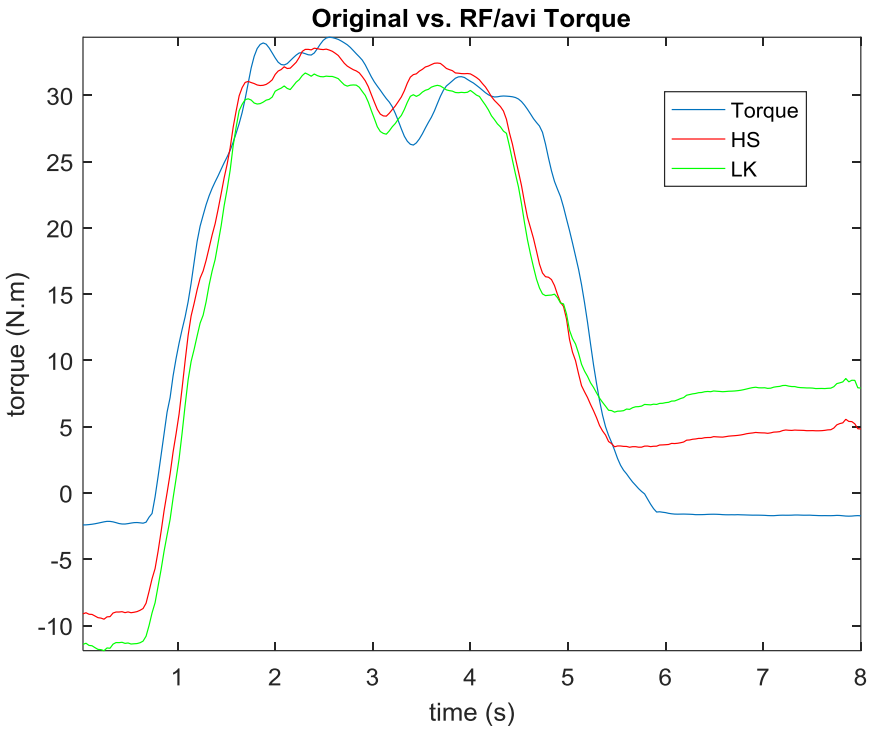

Figure 3-21. Lucas-Kanade (green) and Horn-Schunck (red) Optical Flow polynomial fit results vs Generated Torque (blue) during a 30% MVC isometric contraction.

From the graphic results in Figure 3.21 it can be observed that LK appears more sensitive to small motions than HS due to the appearance of small jumps in smooth areas probably related to local motions within the muscle. The resulting RMSE values were 6.8886 N m for the LK approach and 4.838 N m for Horn Schunck that translates to a NRMSE goodness of fit of 0.4838 and 0.6523 respectively. These results show a better performance of Horn Schunck for the estimation of joint torque collected from a dynamometer chair. Due to this finding, motion tracking analysis in following chapters will include HS as the technique used.

4 Data Collection and Analysis

This chapter describes the experimental setup and protocols used for acquisition and analysis of ultrasound data to address the research questions:

- Do US images of a single muscle (GM) contain enough information to accurately predict net ankle joint torque?

- Does the US raw data precision (16 bit) provide information that 8-bit compressed data does not?

- Is the temporal resolution (inverse of frame rate) an important factor for extracting dynamic information from US data? Does it depend on the movement task performed?

- Which data prediction techniques provide the most accurate predictions of net joint torque from US data?

Firstly, section 4.1 provides an overview of the data acquisition setup, including the specifications of the different devices used and the way they were connected to ensure simultaneous data acquisition. Section 4.2 details the post-processing of the data collected in order to obtain the input and output pairs needed for modelling and testing the different prediction algorithms. Finally, section 4.3 details the implementation of four different data prediction approaches, including evaluations of their performance.

4.1 Experimental Setup

In order to be able to test and compare the different post-processing algorithms needed for accurate predictions of net joint torque, ultrasound (32 or 1000 fps) and surface EMG (2000 Hz) data were collected from the *Gastrocnemius medialis* muscle of the left leg of 10 healthy, adult participants(5 male and 5 female, 28.9 ± 4.5 years old). Participants completed a range of passive ankle joint rotations and active contractions in an isokinetic dynamometer chair from which net joint torque signals (4000 Hz) were also collected. All participants provided informed written consent and ethical approval for the work had been provided by the local ethics committee in the Faculty of Science & Engineering at Manchester Metropolitan University.

4.1.1 Experimental Protocol

Participants laid prone, legs fully extended and the left foot firmly fixed to the footplate of an isokinetic dynamometer (Cybex Humac, CSMI, Stoughton MA – USA, Figure, 4.1-left). The input axis of the dynamometer was adjusted to align with the ankle joint centre of rotation. The ultrasound probe was placed over the mid-belly region of MG, aligned to the fascicle plane and held securely in position using a custom-built holder as shown in Figure 4.1-left. GL and soleus surface EMG electrodes were placed on the muscle belly, and the GM electrode was positioned immediately alongside the ultrasound probe, on patches of skin that had previously been shaved and cleaned of surface debris (Figure 4.1-right). EMG data was mainly used for making sure no muscle activations occurred during the passive trials. Ultrasound RF data were collected at 32 and 1000 fps and consisted of 128 (number of US piezoelectric elements) x 2200 (time delay related depth) pixel element data collected with 16-bit precision.

Once all equipment was attached and the participant was comfortably positioned on the dynamometer bed they were asked to perform a short warm up period, consisting of trying the three different passive and active tasks included in the protocol (further explained below), in order to also familiarise themselves with the

instrumentation and ensure they understood the requirements for each task. Once the warm up was completed, the participant was asked to rapidly push the footplate applying maximum strength and hold it for a second before relaxing, and repeat this procedure three times; in order to record the maximum isometric voluntary contraction (MVC) torque value as the average of the three attempts. This information was used to define the torque representing 30 % and 60 % of MVC in the subsequent experimental trials. After that, it was checked that the participant was able to perform an overall rotational movement range of 15 degrees from the equilibrium position in both directions, resulting on a 30-degree movement overall (from -15 to +15 degrees).

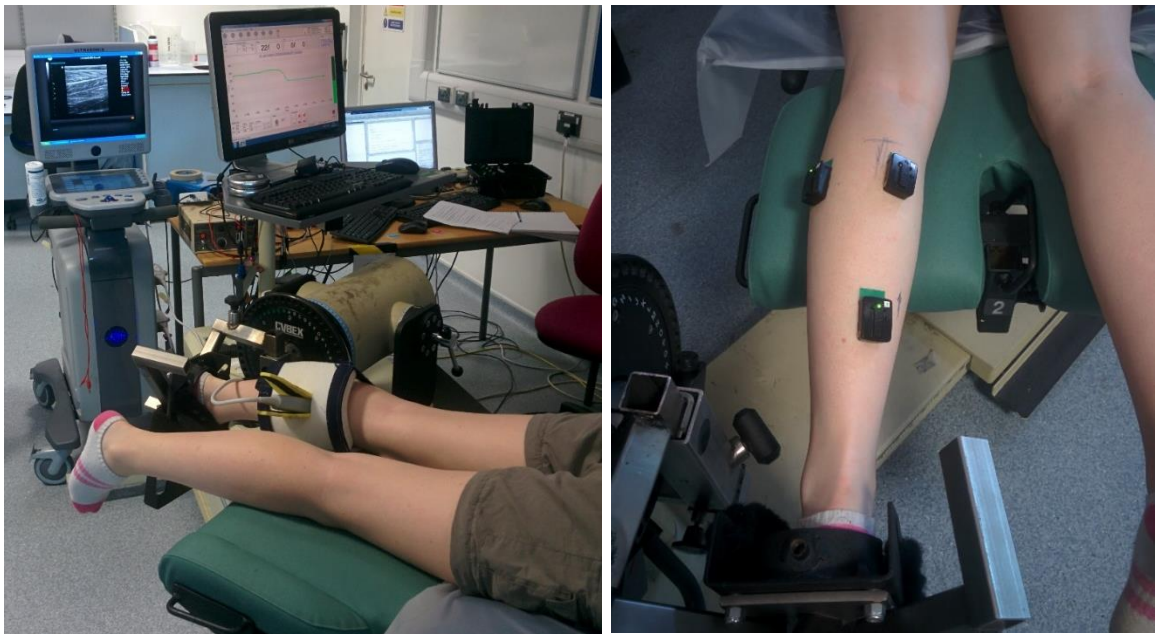

Figure 4-1. Experimental setup view including participant and US probe positioning on the dynamometer chair(left); and showing the Surface EMG electrode placement covering GM, GL and Soleus muscles of the left leg (right).

During the experiment the participant was asked to perform one of three tasks:
Isokinetic measurements refer to the rotation of the footplate, controlled by the dynamometer chair, imposing a passive ankle joint plantar or dorsiflexion. The footplate moved at a constant velocity of 4 degrees per second from -15 to +15 degrees, fast enough to assure a representative range of movement of the ankle in 8 seconds and slow enough to avoid involuntary contributions of other muscles (Hoang et al. 2009). Passive trials were acquired for dorsiflexion and plantar flexion

separately, and participants were asked to remain as relaxed as possible during the trial refraining from any voluntary contraction.

Isometric measurements involved the footplate being locked at the neutral angle (90^{o} for the ankle) whilst the participant steadily increased the force against the footplate by attempting to dorsiflex the foot. Visual output of the torque produced was provided to the participant on a screen set in their line of sight. They were instructed to steadily ramp up torque to either 30 % or 60 % MVC, holding this effort for two seconds before slowly relaxing. The whole trial lasted 8 seconds.

Isotonic measurements refer to eccentric and concentric contractions, whereby the dynamometer moves the foot pedal against the resisting torque produced by the participant, aiming to maintain a constant output torque. In other words, the participant first pushes a resistance applied by the dynamometer to the pedal equal to a percentage of the MVC torque. As with the isometric contractions, torque feedback was provided to participants overlaying a plot of the desired torque output which they were encouraged to follow. The trial consisted of both an eccentric and a concentric phase and lasted eight seconds in total. Due to limitations with the equipment only eight participants were able to perform these trials.

Each of the three muscle contraction types included two different conditions (i.e. dorsiflexion or plantarflexion for isokinetic and 30 or 60 % MVC for both isometric and isotonic contractions) and two trials per condition were recorded. In addition, the entire protocol was completed for each of two different US frame rates (32 and 1000 fps), the total number of trials from each participant was therefore 24 (3 contraction types × 2 conditions × 2 repetitions × 2 frame rates). The order in which tasks were presented to participants was randomised in order discriminate issues related with fatigue or task order related biasing.

The data collected from one of the participants had to be discarded due to a positioning of the probe too close to the tendon area, therefore the final sample size

consisted of nine participants for the passive movements and isometric contractions, and seven participants for the isotonic trials.

4.1.2 Experimental Distribution and Data Collection Systems

The data acquisition process was fully synchronised, and supervised using custom written Matlab code (Version 2014a) running on a controlling PC as described in the block diagram in Figure 4.2. The Matlab script in the controlling PC defined the DAQ card acquisition conditions in terms of sample rate and number of inputs, together with the US parameters, and also provided visualization of results after every trial in order to assure all systems connected were performing correctly.

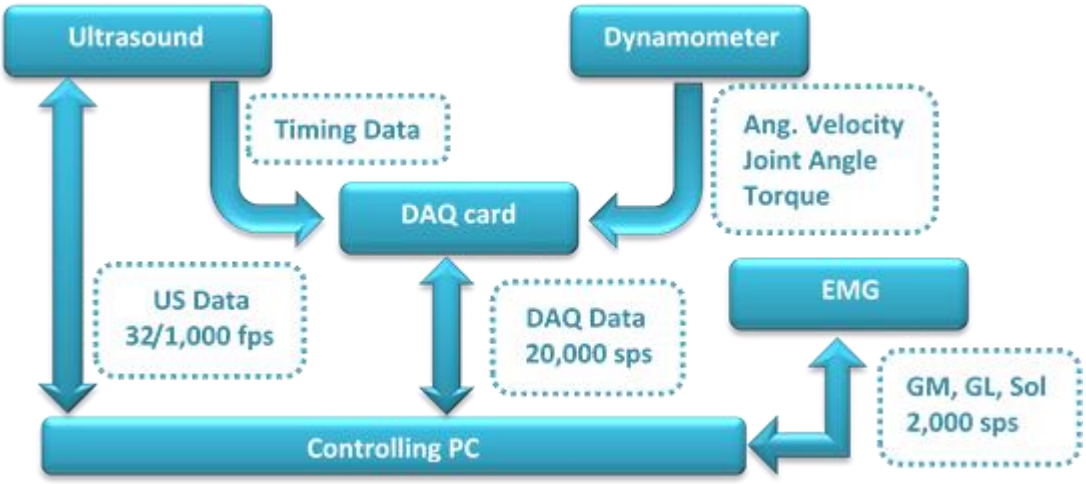

Figure 4-2. Block diagram of the different data collection systems included in the experimental setup, including sampling rates.

The ultrasound data were transferred to the controlling PC via USB connection. The frame transmission times were recorded using the trigger timing circuit described in Section 3.1, which was connected to a data acquisition card (NI-PCI-6229, National Instruments, Texas, USA) via BNC connector. The joint angular velocity, joint angle and net joint torque measured by the isokinetic dynamometer were also recorded via the data acquisition card. The card was installed in the controlling desktop PC and acquired these data at 20000 samples per second. Ultrasound data were collected at two different frame rates by using the conventional interface (32 fps)

and the research interface (1000 fps), and EMG was collected at a rate of 2 kHz via the same data acquisition card. The whole setup, including ultrasound imaging and timing, EMG and dynamometer data, was synchronised using a common trigger, controlled via the MATLAB data acquisition toolbox (R2014a, Mathworks, USA). Figure 4.3 shows the distribution of the different instrumentation used.

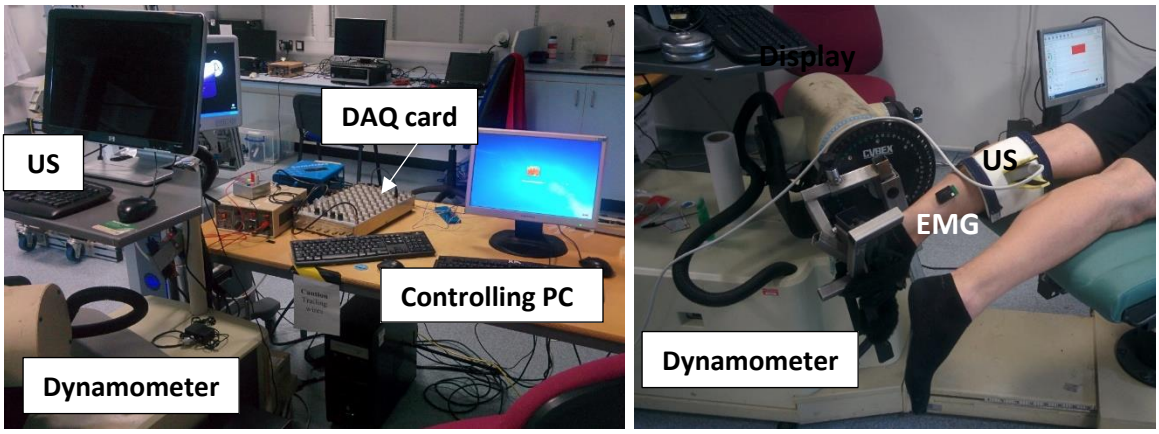

Figure 4-3. View of the different instrumentation used in the setup. Left: Acquisition devices used. Right: Probes position and distribution.

4.2 US and Dynamometer Data Analysis

All collected data were stored on a high-end specification desktop PC (6-core 3.2GHz processor, 16GB RAM, GPU 4GB). All data post-processing was performed with this PC, using the computer vision system, signal, control and identification toolboxes of Matlab (R2016a version, Mathworks, USA) applied in custom written.

4.2.1 US Raw data processing

The 16-bit raw data collected from the Clinical and Research interfaces were firstly beamformed, defined as the organisation of the US individual element data into time/space order still containing the carrier signal (see Section 2.2). In order to filter the carrier signal off, the In Phase Quadrature (IQ) Demodulation is used (Figure 4.4). This process included four main steps. Firstly, the original RF-signal was decomposed into its cosine and sine components obtaining the In-phase (I) and the

Quadrature (Q) signals that still contain the carrier signal. Secondly, a Finite Impulse Response (FIR) low pass filter was applied to both I and Q signals in order to remove unwanted noise and negative frequency spectrum. Once I and Q were filtered, the envelope detection of the overall signal was determined by calculating the module of the I-Q vector, in other words, by calculating the square root of the sum of the squares of I and Q. Finally, a logarithmic compression was applied to data in order to approximate the obtained envelope to values within the human visual perception range. The resulting 2D representation of the outcome of this process is the B-mode image (Figure 4.4).

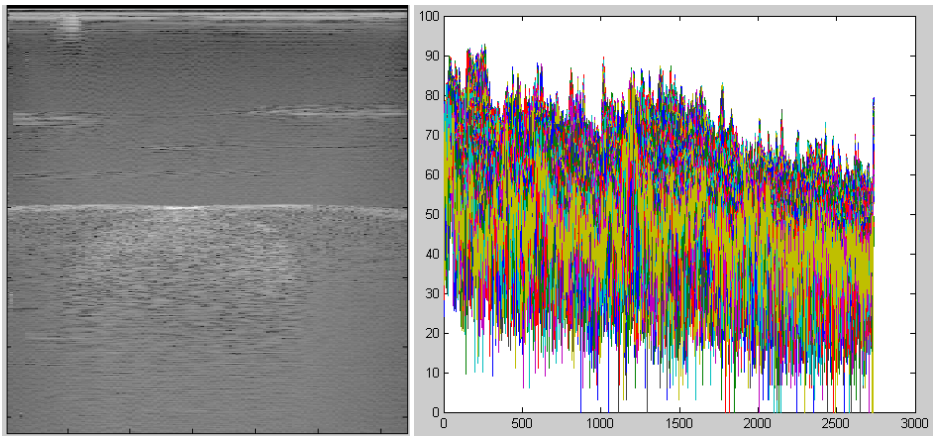

Figure 4-4. IQ resulting data of a silicon sphere suspended in a water tank. Left: 2D image. Right: Individual US element Signal amplitude in Time (128 elements overall).

After applying the IQ-demodulation to the US data the resulting RF-data were copied and stored as a separate parameter, and an 8-bit compression was applied (using the uint8 Matlab built-in function) to provide two versions with different precisions: i) 8 bit and ii) 16 bit. The post processing applied to the US data detailed in the following subsections were applied to each version of the RF-data, to enable precision dependent performance of the prediction techniques for different motor tasks recorded to be made.

4.2.2 Image Segmentation and Muscle Cross Sectional Area Calculation

The GM muscle was segmented for all the frames and trials per participant by applying an ASM consisting of 5000 shape pairs generated for localising the superficial and deep aponeuroses for the first frame, and 50 shape pairs for computing the shift of the aponeuroses between consecutive frames, using the approach detailed in Sections 3.2 and 3.3. The outcome of this process was the location, in pixels, of the superficial and deep aponeurosis within each US image. The distances between the extreme points delimiting both aponeuroses were transformed to distance natural units using the US acquisition information (60 × 60 mm sonicated area) in order to enable the calculation of the area encapsulated within the segmented region (Figure 4.5). This measure provides information related with the muscle cross sectional area.

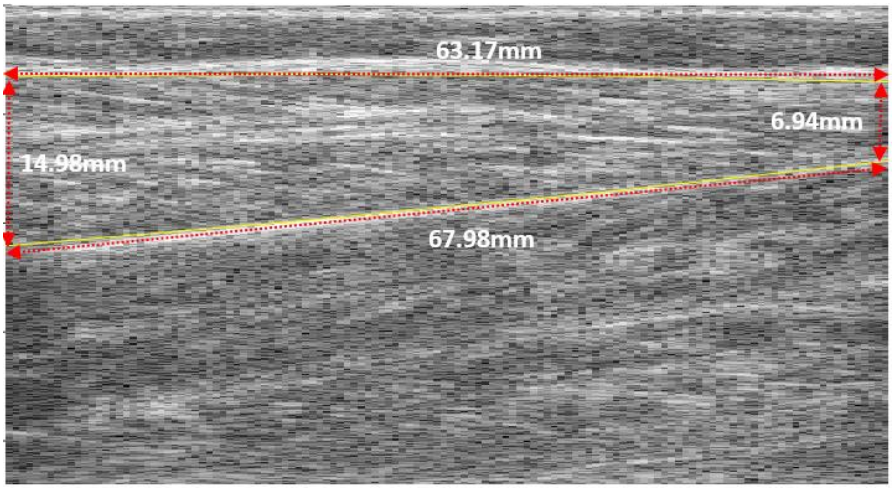

Figure 4-5. Representation of the distances between the extreme points of the two aponeurosis delimiting the GM muscle. The resulting cross sectional area for this frame is 701.72mm^2.

The muscle cross-sectional area (CSA) was calculated for every frame in mm^2 units, and the resulting array subtracted by the previous frame contribution:

$$area\_diff(n) = area(n) - area(n-1) \quad (27)$$

As a result, the "per frame" relative differences in cross-sectional area between consecutive frames was obtained (Figure 4.6) and used as an input for the prediction process explained in the following section.

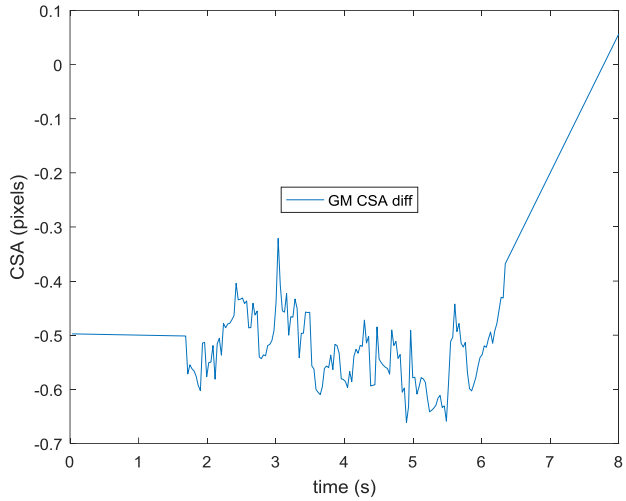

Figure 4-6. Muscle cross-sectional area diff results extracted from the ASM during a passive dorsiflexion movement.

4.2.3 Optical Flow

In order to obtain information related to motion within the GM muscle, OF motion vectors were calculated from the ultrasound data using the implementation detailed in Section 3.3.4. The cross-sectional area flow vectors were calculated between consecutive frames and within the area delimited by the two aponeuroses located by the ASM, obtaining a mean flow value in the x axis direction as an outcome (Figure 4.7).

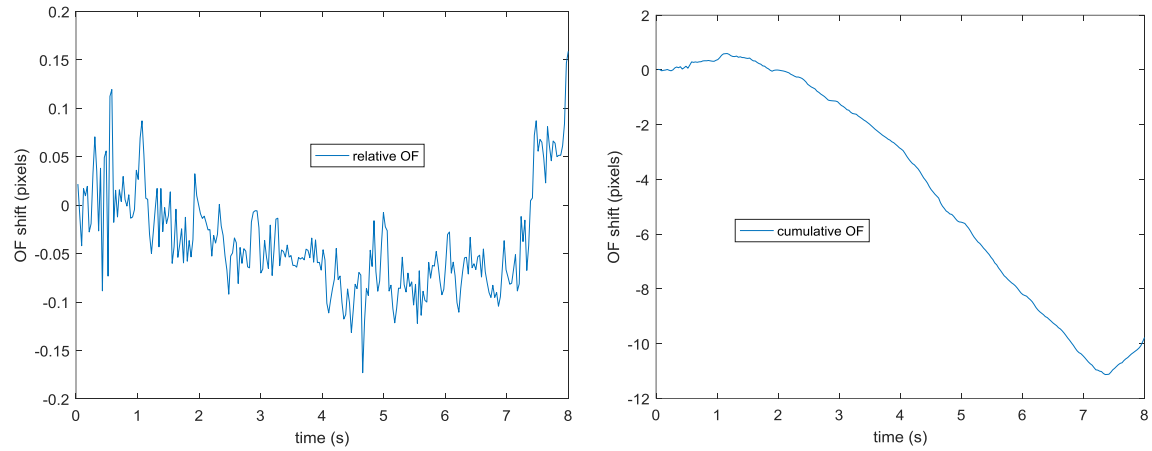

Figure 4-7 Horn-Schunck OF results when applied to the whole GM muscle cross-sectional area during a passive dorsiflexion movement. Left: OF X axis relative "between frames" results per frame, right: cumulative sum of it. Y axis Values in pixels.

4.2.4 Joint Torque

The voltage values corresponding to the joint torque data collected from the dynamometer were translated into units of torque (N m) using the results of a custom calibration process, completed using the same experimental equipment used for data acquisition. In order to reduce noise observed in the voltage signal (Figure 4.8), a Savitzky–Golay filter of degree 3 and 5000 points (out of the 200000 collected) was applied. This filter applies a least square fit that reduces the signal to noise ratio without compromising the original shape of the signal (Savitzky & Golay 1964), by applying a polynomial interpolation of a degree chosen, and using a subsample of the original data (window). As a result, a smooth response comparable with the US derived results mentioned above was obtained.

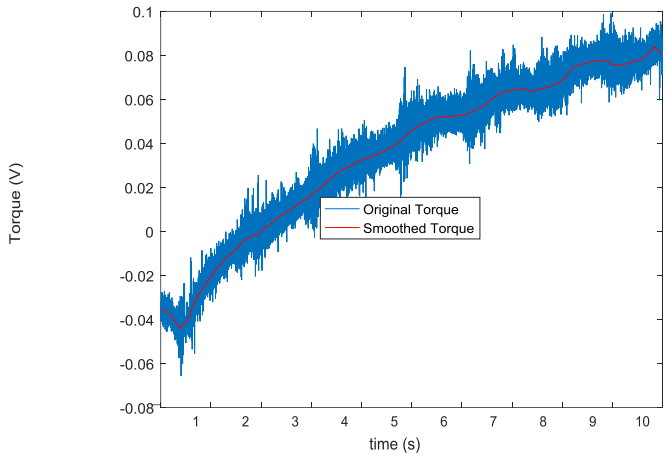

Figure 4-8. Original (blue) vs smoothed (red) joint Torque during a passive joint movement using Savitzky–Golay filter.

The resulting filtered torque signal was considered the target signal to be achieved by transformation of the calculated US derived input values: muscle thickness and OF derived motion.

4.3 Prediction of Joint Torque with US-based Data

The signals obtained from processing the US and dynamometer data from every trial and participant were gathered into two different matrices (Figure 4.9): "Big Input", containing the GM cross-sectional area together with the OF x axis motion from the collected frames; and "Big Output", containing the filtered torque data from the dynamometer expressed in N m.

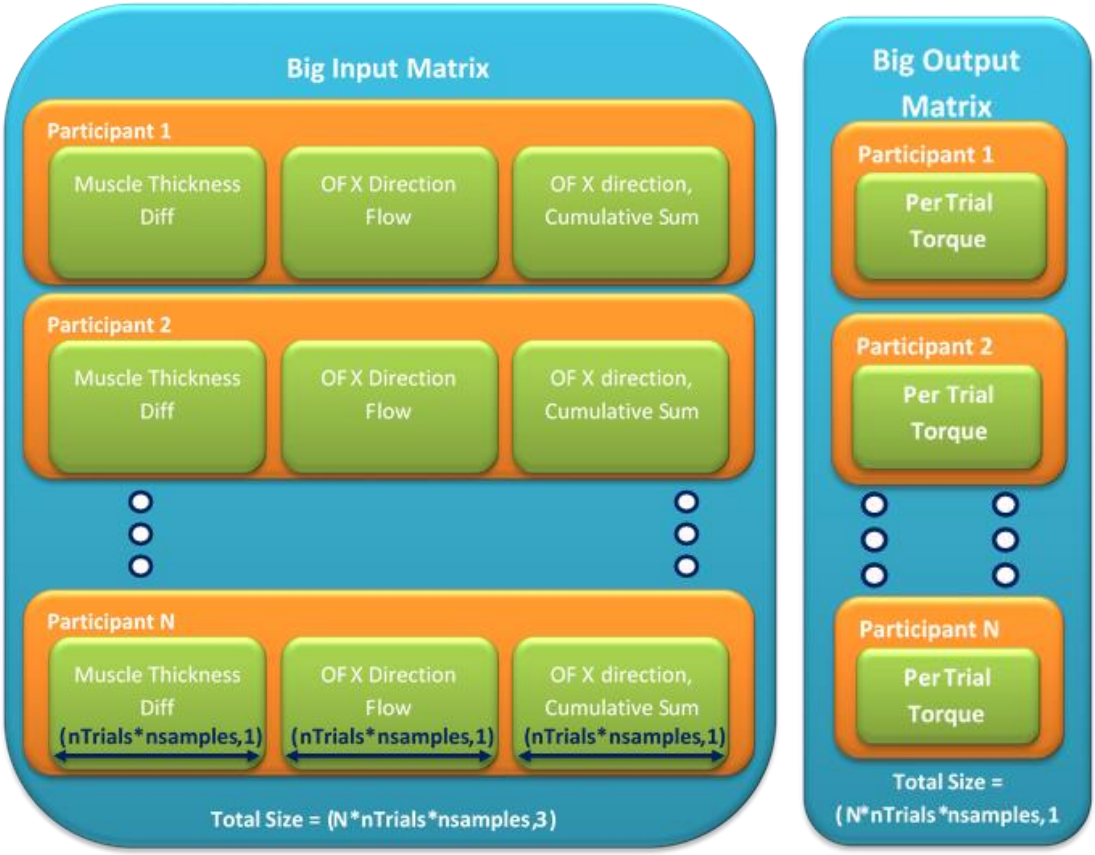

Figure 4-9 Block diagram of the contents of the input-output pairs gathered per participant and trial.

Once processed and organised the resulting Big Input and Big Output matrices were iteratively subdivided in order to cross validate the different prediction algorithms. The information of one participant is taken out of the matrices to be used as validation of the generated prediction models. Five different regression models were applied to the rest of the participants paired input-output data: Linear Regression, Squared Regression, Cubic Regression, Support Vector Regression using Radial Basis Function Kernels, and Artificial Neural Networks ranging from two to three hidden layers with up to 10 neurons each, all of which are detailed below (also see Section 2.4).

4.3.1 Linear Regression

This polynomial regression of degree one is the simplest and therefore fastest prediction method to implement. It is only able to identify scalar relationships that can be modelled in the following manner:

$$\text{linear} = \ offset + kx \tag{28}$$

where *linear* is the prediction output value for the given input *x* and *k* is a scalar.

4.3.2 Quadratic Regression

One option for solving non-linearly related input-output relationships is the use of quadratic regressions that basically consist of polynomial regression of degree two. This approach is capable of predicting curved input-output relationships at a slightly higher computational cost, being able to express them as:

$$quadratic = \ k_0 + k_1 P_1 + k_2 P_2^2 \tag{29}$$

where quadratic is the predicted output and k_n are constants and P_n are the "n" polynomial terms.

4.3.3 Cubic Regression

These "degree three" polynomial prediction algorithms are an optimal solution when the input-output relationships are not accurately predicted by quadratic approaches, together with low implementation costs as a key factor.

$$\text{cubic} = \ k_0 + k_1 P_1 + k_2 P_2^2 + k_3 P_3^3 \tag{30}$$

where cubic is the predicted output and k_n are constants and P_n are the "n" polynomial terms.

4.3.4 Support Vector Regression using Radial Basis Function Kernels

In order to be able to account for more complicated input-output relationships, SVR kernel algorithms can move them to a feature space in which the initial non-linear relationship can be explained by simpler inner product relationship (Section 2.4.2); and they can provide better generalisations than statistical regressions for predicting new inputs. The RBF Kernel matrix can be defined as:

$$K(x, x') = exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \tag{31}$$

where x is the input matrix and σ is the radius of influence of support vectors.

For the case of Gaussian kernels, the number of observations is much smaller than the feature space dimensionality implicating the possibility of finding a solution that fits the training data perfectly. In order to avoid overfitting, the use of a regularisation term (ridge/quadratic in our case) becomes essential in order to provide a generalised solution able to predict new data, obtaining the resulting equation:

$$E(w) = \frac{1}{2}\sum_{i=1}^{N}\{f(x_i, w) - y_i\}^2 + \frac{\lambda}{2}\|w\|^2 \tag{32}$$

where x and y are the input and output matrices respectively, w is the matrix containing the kernel weights for the training examples, and λ the regularization constant. The first term of the sum is the loss function, and the second (the one containing λ) the ridge regularization term.

For the implementation of the RBF Kernel SVM prediction algorithm, the Matlab km_krr function developed by (Vaerenbergh 2009) was used, which belongs to the "Kernel Methods" toolbox. This function outputs a prediction of a new input when given an input matrix associated with an output matrix using kernel ridge regression, together with the regression weights. The input and output matrices correspond with the Big Input and Big Output defined above (Figure 4.9), without the validation values separated during the cross-validation iterative process. The remaining data needed for the implementation of this function were the ridge regularization constant ($\lambda = 0.3$); and the kernel parameter ($\sigma = 3$), which represents the magnification of the kernel distance of the Gaussian distance. The values of these parameters were chosen as they corresponded to the best fits to sample data during an initial parametrisation that tested regularization values from 10^{-5} to 1; and kernel parameters from 1 to 20.

4.3.5 Artificial Neural Network

The ANN was created as a standalone function using the Neural Network Toolbox of Matlab, and applied to the empirically collected data in order to predict new inputs. The implemented functions takes as input arguments the training input and output matrices, and the regularization parameter ($r = 0.01$). Furthermore, another variable was parametrised, checking the performance of the torque prediction using different complexity ANNs that ranged from [1 1], two hidden layers with one neuron each; to [10 10 10], 3 hidden layers with 10 neurons per layer.

The different topology complexity ANNs were created and trained with the training input and output matrices (See Section 4.3) using the Levenberg-Marquard convergence algorithm (Section 2.4.2). Once trained, the resulting ANNs were used for predicting new data (cross validation process); and the best prediction result from all of the studied ANN topologies was noted and outputted as the overall ANN best fit per condition studied.

4.3.6Prediction Algorithm Training and Cross-Validation

The training data was used by the different algorithms applied to estimate a model that both provides a good fit to the training data and is a good estimator of new inputs (validation data - red), in a leave-one-out process. The result of this was a predicted torque signal (size = nsample × ntrials, Figure 4.9) from one participant estimated from the data of the other participants. The next step was to separate the data of a second participant for validation, using the rest for training the different regression models again. This whole process was repeated "N" times in order to cross-validate the models with "new" participant data. In other words, the different algorithms generate an estimation of the joint torque of a new participant, whose data were excluded from the training process, based on the information provided by the post processing of ultrasound data of the GM.

4.3.7 Prediction Algorithm Performance Indicators

In order to evaluate how well the predicted signal matched the real torque acquired by the dynamometer, the Root Mean Square Error was used as a "goodness of fit" indicator:

$$RMSE\ fit = \sqrt{\frac{\sum_{i=1}^{n}(Torque_i - OFx_i)^2}{n}} \quad (33)$$

where n is the number of samples.

4.4 Methods Summary

This chapter has provided details of the experimental set-up, data acquisition, analysis and evaluation. The following chapters describe the results obtained from the different prediction algorithms when applied to data of different precision (8-bit vs 16 bit) and frame rates (32 fps vs 1000 fps), enabling an evaluation of the performance of each of the algorithms for prediction of net ankle joint torque and the influence of ultrasound data properties. The results have been separated so that each of the muscle contraction types are considered independently, an over-arching discussion reviewing the results provided in Chapter 8.

5 Predicting Ankle Torque during Passive Joint Rotation

5.1 Introduction

This chapter deals with metricising the performance of inference methods (detailed in Section 4.3) for prediction of ankle joint torque from US processed data of the GM muscle during passive ankle joint movements. The data used for the work presented corresponds to the passive measurements described in Section 4.1.1, and addresses four main questions:

- Is there a difference in the prediction error of ankle joint torque generated during passive joint rotation using image analysis of data collected at different frame rates?

- Is there a difference in the prediction error of ankle joint torque generated during passive joint rotation using image analysis of data collected with different precision?

- Is error in the prediction of ankle joint torque generated during passive joint rotation affected by the prediction technique used?

- Is there an important interaction between two or more of these features?

In order to provide background information regarding the value of investigating passive muscle movements, the next section contextualises previous work related to the study of muscles during passive ankle rotations. Furthermore, error measures between predicted joint torque and its empirically acquired value are analysed for the three main factors studied in this chapter: temporal resolution, data precision

and prediction algorithm used. Finally, a discussion (section 5.7) is provided highlighting any interdependence identified between the aforementioned factors.

5.2 Ankle joint in Passive movements

Knowledge of muscle passive properties is required for modelling studies such as Hill models (Hill 1952). These techniques can benefit from individualised measures to help inform aspects of personalise medicine; and for conditions such as post-stroke symptoms (Li & Tong 2016) or cerebral palsy (Brandenburg et al. 2016), for which spasticity can be a problem when trying to relearn motor coordination. This condition (spasticity) is also problematic in neurodegenerative diseases (e.g. Motor Neuron Disease), for which measures of muscle stiffness could be useful as a biomarker of changing upper motor neuron/muscle properties in clinical trials of different interventions.

During passive movement muscle-tendon units behave like a spring-mass system and, on surpassing a certain threshold length (also known as slack length) they develop a tensile force (Woittiez et al. 1984). Muscle fascicles can be considered as arranged in series with tendons and lengthen relative to their compliance and proportionate to the overall muscle-tendon tension (Zajac 1989). Later work showed buckling of the muscle fascicles at very short lengths, suggesting some fascicle recruitment when passively lengthening the muscle. This study also reported that the greater part of muscle-tendon unit length changes occur in the tendon area (Herbert et al. 2011).

Previous research focused on the study of gastrocnemius fascicles and muscle-tendon unit during passive movements using US (Wakahara et al. 2005; Hoang et al. 2007; Hodson-Tole et al. 2016). The second work used a methodology developed in (Hoang et al. 2005) for measuring passive length-tension properties of the gastrocnemius *in vivo*, that included the inference of nine unknown parameters in order to estimate length-tension curves. A modification of this methodology can be found in (Nordez et al. 2010), in which they concluded that muscle force observed

during passive rotations can be related to two main factors: muscle resting length and changes in tissue stiffness. This way they reduced the number of parameters required to infer length-tension curves from nine to just two. This initial work relied on measures of the fascicle length and angle using the cross-section area imaged by the US probe.

Furthermore, US based measures of cross sectional area strongly depend on the probe angle and positioning, (Strasser et al. 2013). They can vary between experimental protocols making them a potential source of variability impossible to correct unless using probe orientation tracking systems. Therefore, a more general methodology for the study and inference of muscle force parameters (joint torque) during passive movement should be developed, one which addresses the limitations of US technology.

5.3 Torque Predictions using Standard Clinical Ultrasound

As introduced in Section 3.1, Standard Clinical Ultrasound defines two different conditions: the use of conventional focused waves for the US transmission, and an acquisition rate of 32 fps, resulting in a temporal resolution of 3.125 ms between frames. Variability in the IFIs was present between frames but here this was accounted for by the use of the hardware trigger circuit (Section 3.1). The time logs generated during data collection were checked and it was found that no frames were lost during any of the acquisitions. This likely reflects the low frame rate in with the standard clinical ultrasound implementation (implying a lower resource demand on the controlling PC) and means that any IFI variability in the collected data can be accounted for and corrected, therefore any discrepancies between actual and predicted net joint torque should not be due US timing phenomena.

The cross-validation results of the five prediction methods for the passive trials within this section are compiled in tables in which "Participant #" represents the

accuracy of the prediction made by each technique when data from that participant was removed from the training process and used as the validation data.

5.3.1 Dorsiflexion Movements

The following section reports the results from the analysis of 8-bit and 16-bit data relating to the dorsi-flexion phase of the passive ankle joint rotation movements. Table 5.1 presents individual participant and mean group data from the 8-bit data.

Table 5-1. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during passive dorsiflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	1.778	1.779	1.760	1.810	2.022
Participant 2	0.918	0.845	0.889	0.936	0.743
Participant 3	0.788	0.691	0.643	0.766	0.704
Participant 4	0.749	0.729	0.683	0.779	0.801
Participant 5	2.002	1.862	2.045	1.921	1.829
Participant 6	0.455	0.466	0.447	0.465	0.450
Participant 7	0.750	0.715	0.729	0.705	0.468
Participant 8	1.625	1.420	1.324	1.551	1.650
Participant 9	1.602	1.792	1.902	1.559	1.870
Average	1.185	1.144	1.158	1.166	1.171
Stand. Dev.	0.562	0.562	0.611	0.543	0.655

Only small differences in the performance of the five different inference methods can be observed. The lowest average error was observed in quadratic regression, but all techniques presented results that were very similar. The lowest average result (1.144 ± 0.562 N m) corresponds to just 7.62 % of the maximum torque during passive trials (15 N m).

In contrast individual and group data from the analysis 16-bit data are shown in Table 5.2. Again the performance of the five different inference methods showed quite similar results in terms of both average and standard deviation. The lowest average error was observed in Quadratic regression and the lowest deviation for

ANN. The lowest average result (1.220 ± 0.577 N m) corresponds to just 8.13 % of the maximum torque during passive trials (15 N m).

Table 5-2. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during passive dorsiflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	1.668	1.642	1.586	1.765	1.688
Participant 2	0.840	0.920	1.015	0.864	0.768
Participant 3	0.821	0.699	0.645	0.773	0.575
Participant 4	0.874	0.840	0.778	0.923	0.820
Participant 5	2.071	1.990	2.245	1.979	1.980
Participant 6	0.525	0.471	0.424	0.535	1.000
Participant 7	0.849	0.832	0.864	0.802	0.859
Participant 8	2.146	1.868	1.845	2.069	1.906
Participant 9	1.690	1.721	1.736	1.637	1.669
Average	1.276	1.220	1.237	1.261	1.252
Stand. Dev.	0.614	0.577	0.629	0.593	0.550

To enable direct comparison of the differences in prediction error resulting from the different data precision levels investigated, the results obtained are shown in graphical form (Figure 5.1). To aid data precision comparison the graphs share the same Y axis range (that represents joint torque RMSE) for the 8- and 16-bit data analysed. It is clear that the prediction errors from the two data precision levels are very similar.

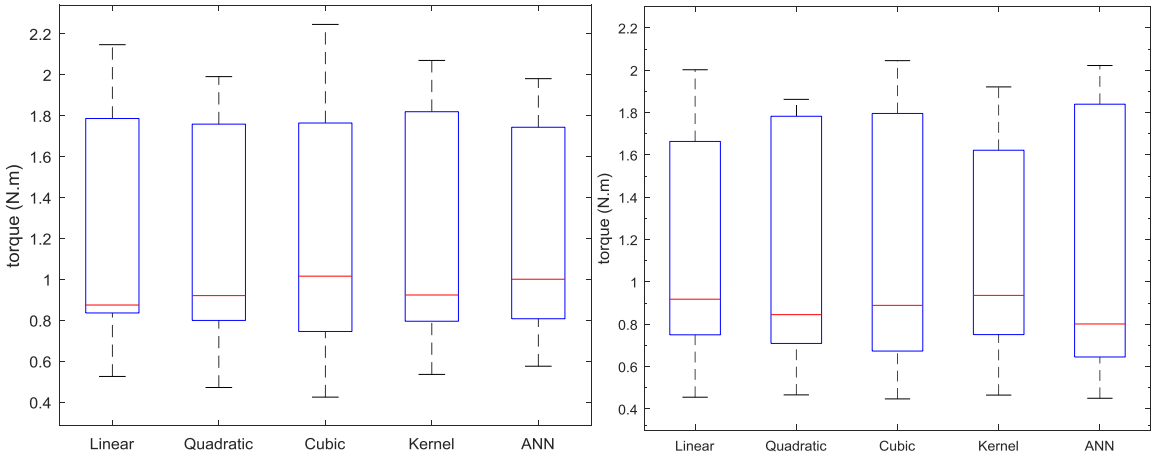

Figure 5-1. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Passive measurements during dorsiflexion from nine participants.

5.3.2 Plantar Flexion Movements

Table 5.3 presents individual participant and mean group data, resulting from analysis of 8-bit data from the plantar flexion phase of the passive ankle joint rotations.

Table 5-3. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during passive plantarflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.838	5.229	10.214	1.313	2.028
Participant 2	1.073	1.056	1.115	0.935	1.172
Participant 3	1.367	1.324	1.347	1.392	1.575
Participant 4	1.064	1.061	1.053	0.967	1.011
Participant 5	1.604	1.382	1.398	1.580	1.282
Participant 6	2.261	2.368	2.405	2.673	1.384
Participant 7	0.931	0.768	0.733	0.913	0.895
Participant 8	2.238	2.346	2.323	2.858	2.198
Participant 9	1.687	1.672	1.667	1.917	1.844
Average	1.674	1.912	2.473	1.617	1.488
Stand. Dev.	0.652	1.362	2.956	0.731	0.456

The performance of the inference algorithms presented was similar for all cases but cubic regression, for which an inaccurate prediction when validating the first participant resulted in a notable increase of both the average and the standard deviation calculated for the group. The best prediction result for this condition in both average and deviation was found for the case of ANN, showing a 1.488 ± 0.456 N m that corresponded to 9.92% of the maximum torque in passive trials.

In contrast Table 5.4 presents individual participant and mean group data, resulting from analysis of the plantar flexion phase of the passive ankle joint rotations.

Table 5-4. RMSE Results for the 5 inference methods used for Joint Torque predictions. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.756	3.786	11.164	1.308	2.331
Participant 2	1.232	1.345	1.388	0.692	1.148
Participant 3	1.581	1.567	1.572	1.548	1.515
Participant 4	1.242	1.284	1.286	1.154	1.228
Participant 5	1.703	1.551	1.578	1.755	1.321
Participant 6	2.452	2.476	2.477	2.825	2.383
Participant 7	1.112	1.030	1.012	1.221	1.202
Participant 8	3.500	3.916	3.865	4.657	5.006
Participant 9	1.897	1.869	1.880	2.092	1.856
Average	1.942	2.091	2.913	1.917	1.999
Stand. Dev.	0.808	1.078	3.210	1.197	1.223

Again, the five prediction methods obtained similar results for all cases but cubic regression, whose average and standard deviation were again increased mainly due to a really bad fit when validating the first participant. The best predicting power was observed for the linear regression that, despite having a slightly worse average than Kernel SVR, showed the lowest standard deviation for this condition. This average (1.941 ± 0.808 N m) corresponded to 12.94 % of the maximum torque in passive joint rotation tests.

To enable direct comparison of any differences in torque prediction error resulting from differences in data precision, results obtained are shown in graphical form (Figure 5.2). To aid data precision comparison the graphs share the same Y axis range (that represents joint torque RMSE) for the 8- and 16-bit data analysed. Again it is clear that the prediction errors from the two data precision levels are very similar.

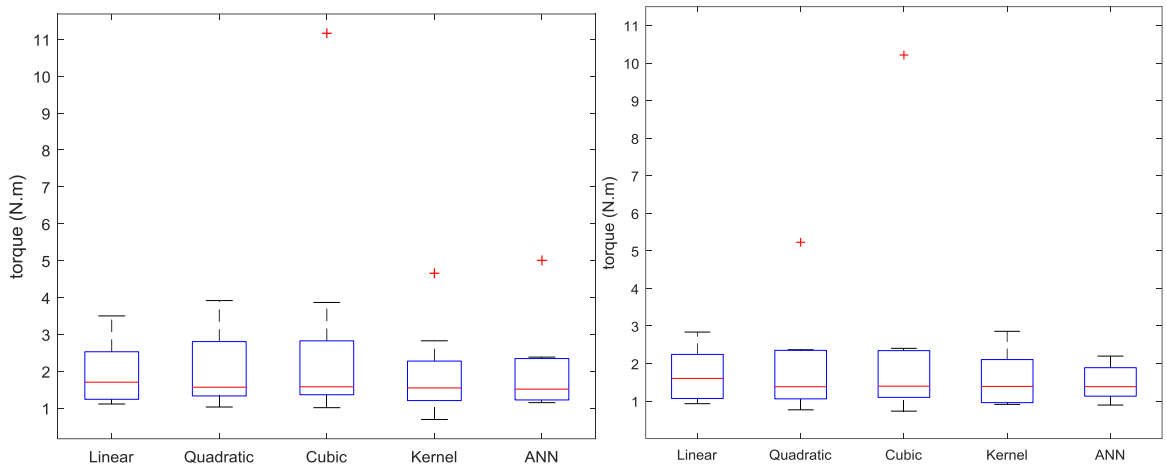

Figure 5-2. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Passive measurements during plantar flexion from nine participants.

5.4 Torque Predictions using Ultra-Fast frame rate Ultrasound

The US data used to obtain the results presented in this section consisted of 8000 frames per trial (8 seconds long each) that exponentially increased the computational overhead of the processing applied to the RF-data. One important limitation was found for one of the prediction methods. In order to be able to apply the Kernel SVM based prediction, assuming all the collected frames were to be used in training, the processing PC would require a minimum 2TB of RAM. Due to computational (and financial) considerations, only 1000 points per trial were used for the Kernel SVM results as this was the maximum possible with the available RAM of the system (16GB).

5.4.1 Dorsiflexion Movements

The following section reports the results from the analysis of 8-bit and 16-bit data relating to the dorsi-flexion phase of the passive ankle joint rotation movements. Table 5.5 presents individual participant and mean group data from the 8-bit, fast frame rate data.

Table 5-5. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during passive dorsiflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.246	2.219	2.340	2.253	1.493
Participant 2	1.049	0.974	0.927	0.959	0.945
Participant 3	1.630	1.640	1.640	1.651	1.697
Participant 4	1.229	1.188	1.170	1.213	1.076
Participant 5	3.165	3.150	3.135	3.155	3.114
Participant 6	0.853	0.664	0.672	0.703	0.690
Participant 7	0.631	0.564	0.520	0.591	0.539
Participant 8	1.363	1.269	1.282	1.274	1.274
Participant 9	1.673	1.616	1.597	1.649	1.722
Average	1.538	1.476	1.476	1.494	1.394
Stand. Dev.	0.777	0.812	0.832	0.809	0.768

Similar results were observed for all the inference methods but linear regression, which showed a slightly worse performance. As observed in conventional framerate results, the worst predictions observed were associated with two participants (1 and 5) for all algorithms. The best prediction in both terms of average and standard deviation was observed in ANN, with a result (1.394 ± 0.768 N m) that corresponds to 9.29 % of the maximum torque during passive trials (15 N m).

Table 5.6 presents individual participant and mean group data from the 16-bit, fast frame rate data. The performance of the five different inference methods showed quite similar results in terms of both average and standard deviation. The lowest average error was observed in Kernel SVR with an RMSE of (1.394 ± 0.716 N m) corresponding to 9.29 % of the maximum torque during passive trials (15 N m).

Table 5-6. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during passive dorsiflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.470	2.528	2.666	2.540	2.565
Participant 2	1.253	1.230	1.197	1.217	0.969
Participant 3	1.503	1.513	1.515	1.520	1.783
Participant 4	1.157	1.128	1.119	1.157	1.157
Participant 5	2.451	2.410	2.377	2.409	2.353
Participant 6	0.583	0.551	1.455	0.520	0.943
Participant 7	0.509	0.470	0.438	0.483	0.824
Participant 8	1.137	1.532	1.506	1.174	1.193
Participant 9	1.533	1.514	1.501	1.524	1.503
Average	1.400	1.431	1.530	1.394	1.477
Stand. Dev.	0.697	0.709	0.660	0.716	0.632

To enable direct comparison of any differences in torque prediction error resulting from differences in data precision results obtained are shown in graphical form (Figure 5.3). To aid data precision comparison the graphs share the same Y axis range (that represents joint torque RMSE) for the 8- and 16-bit data analysed. Here different trends between 16-bit and 8-bit are difficult to distinguish, with the 16-bit data predictions providing similar errors to 8-bit in both median and variance.

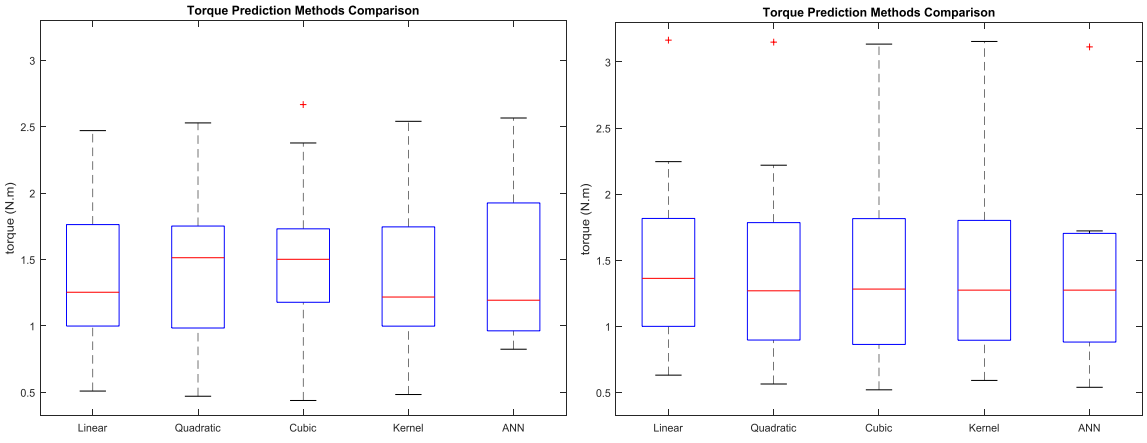

Figure 5-3. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Passive measurements during dorsiflexion from nine participants collected at 1kS/s.

5.4.2 Plantar Flexion Movements

The following section reports the results from the analysis of 8-bit and 16-bit data relating to the plantar flexion phase of the passive ankle joint rotation movements. Table 5.7 presents individual participant and mean group data from the 8-bit, fast frame rate data.

Table 5-7. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during passive plantarflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	1.917	1.729	1.741	1.820	1.600
Participant 2	1.323	0.981	1.062	1.124	0.948
Participant 3	1.969	2.068	2.109	1.951	1.771
Participant 4	1.205	1.161	1.239	1.116	1.235
Participant 5	3.591	3.464	3.418	3.569	3.396
Participant 6	2.623	2.573	2.580	2.608	2.742
Participant 7	1.205	0.997	1.086	1.094	1.074
Participant 8	4.416	4.677	4.706	4.606	4.500
Participant 9	2.825	2.568	3.057	2.679	2.598
Average	2.342	2.246	2.333	2.285	2.207
Stand. Dev.	1.125	1.237	1.233	1.212	1.198

The performance of the inference algorithms presented was similar for the five different prediction algorithms used. ANN showed the best prediction result for this condition in average resulting in a RMSE of 2.207 ± 1.198 N m that corresponded to 14.71% of the maximum torque in passive trials.

Table 5.8 presents individual participant and mean group data from the 8-bit, fast frame rate data. The five prediction methods obtained similar results except linear regression, which showed a slightly bigger error in terms of average. The best predicting power was observed for the cubic regression, with an average of 1.854 ± 0.668 N m that corresponded to 12.36 % of the maximum torque in passive joint rotation tests.

Table 5-8. RMSE Results for the five inference methods used for Joint Torque predictions. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.193	2.148	2.184	2.120	1.603
Participant 2	1.655	1.432	1.385	1.390	0.931
Participant 3	1.777	1.799	1.788	1.771	1.700
Participant 4	1.091	0.946	0.931	0.952	0.859
Participant 5	3.154	3.053	3.036	2.982	2.956
Participant 6	1.882	1.715	1.740	1.722	2.726
Participant 7	1.319	1.114	1.065	1.154	1.228
Participant 8	2.700	2.360	2.287	2.449	2.807
Participant 9	2.363	2.252	2.266	2.293	1.923
Average	2.015	1.869	1.854	1.870	1.859
Stand. Dev.	0.658	0.662	0.668	0.655	0.807

To enable direct comparison of any differences in torque prediction error resulting from differences in data precision results obtained are shown in graphical form (Figure 5.4). To aid data precision comparison the graphs share the same Y axis range (that represents joint torque RMSE) for the 8-bit and 16-bit data analysed. Different trends between 16-bit and 8-bit results are again apparent. Unlike the dorsiflexion results (Figure 5.3), for which no differences were appreciated, here the 16-bit data provide slightly smaller median errors than the 16-bit data.

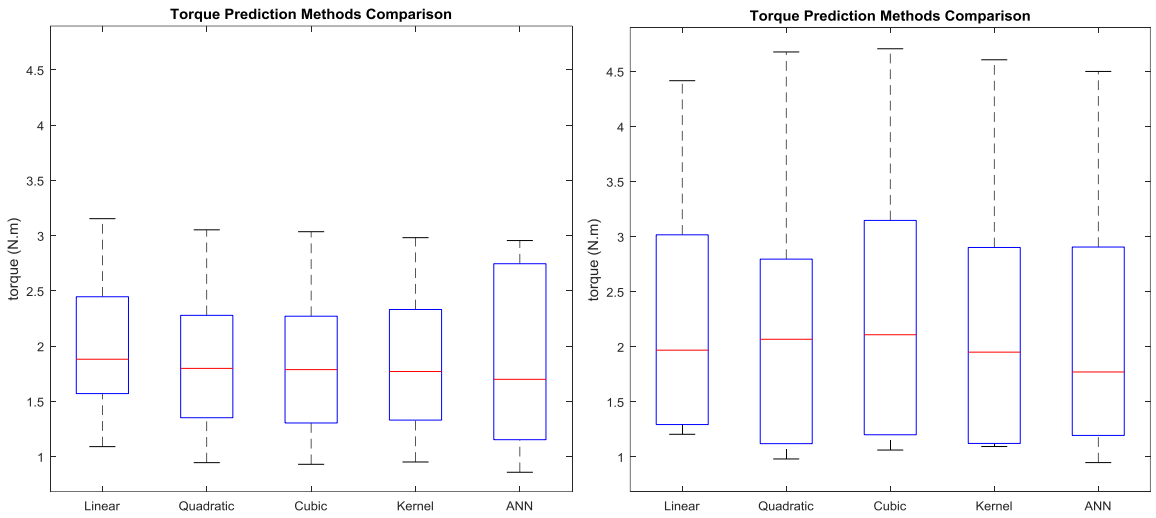

Figure 5-4. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Passive measurements during Plantarflexion from nine participants collected at 1kS/s.

5.5 Features Underlying Algorithm Performance

In order to provide a better understanding of the reasons behind better or worse fits of the predicted torque, Figure 5.5 shows the performances of the different prediction methods (polynomial regression, Kernel Methods and ANN) for an example of a bad fit of the original torque.

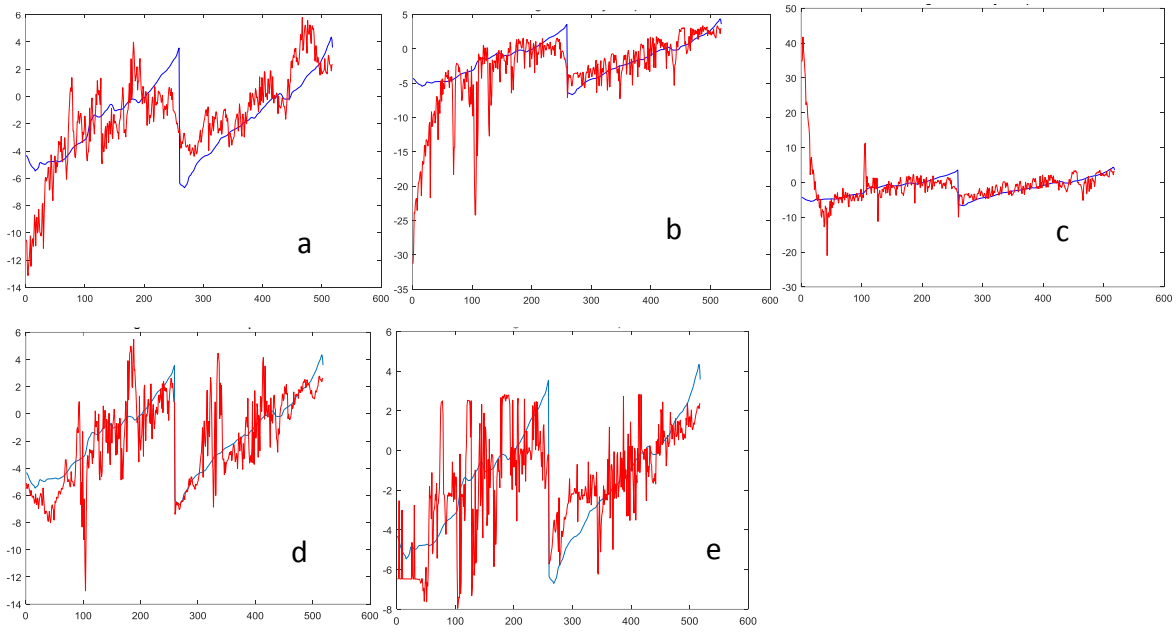

Figure 5-5. Acquired Joint Torque (blue) vs. Predicted Torque (red) graphical results in N m for the Dorsiflexion trials (2 trials, 260 frames each) for 5 prediction methods: Linear (a), Quadratic (b), Cubic (c), Kernel SVM (d) and ANN (e). Resulting RMSE Torque values were 2.31, 4.54, 6.01, 1.97 and 3.55 N m respectively. Positive values of Torques mean eccentric resistance to the pedal while negative mean concentric resistance.

Three different trends can be appreciated concerning the bad performance of the different prediction algorithms. The polynomial regression based, namely Linear, Quadratic and Cubic (Figure 5.5-a, -b and -c respectively), showed a clear initial offset that seems to be necessary for a better fit in posterior frames but that produces errors during the initial part of the trial. Both Kernel SVM and ANN have a similar behaviour for the case showed in Figure 5.5-d and -e respectively. Torque fit showed a noisy response with important jumps (especially in the ANN case) when small disturbances can be observed in the ground torque (blue line). In order to quantify the influence of the prediction on the initial frames for the different inference algorithms used, Figure 5.6 shows the distribution of the differences

between the prediction and the empirically measured value of the joint torque for the five methods for the example above.

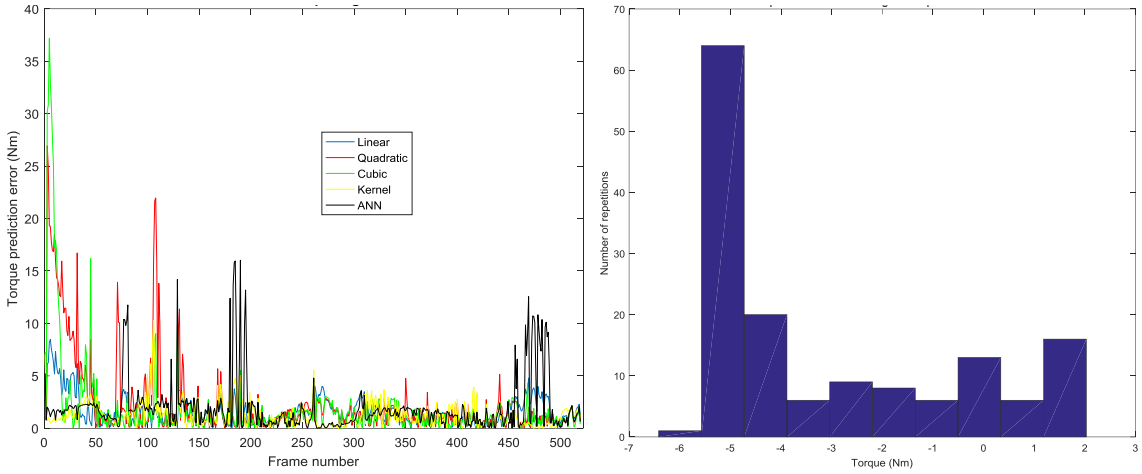

Figure 5-6. Left: Acquired Joint Torque vs. Predicted Torque difference in N m for the Dorsiflexion trials (2 trials, 260 frames each) in absolute values for the 5 prediction methods: Linear (blue), Quadratic (red), Cubic (green), Kernel (yellow) and ANN (black), X axis expresses the frame number and Y axis the Torque difference in N m. Right: Histogram of the Torque values associated with prediction errors bigger than 5 N m. X axis Torque values in N m.

The numerical results were analysed and showed that the first 52 frames (10% of the whole trial) accounted for the 21.25%, 38.37%, 47.20%, 12.59% and 23.25% of the overall error respectively for the Linear, Quadratic, Cubic, Kernel and ANN algorithms. Therefore, the significant errors observed in the initial frames in Figure 5.5 were more pronounced in the case of Quadratic and Cubic regressions.

The histogram (Figure 5.6-right) shows a concentration of the errors between real values and predictions when the torque is around -5 N m, indicating that the algorithms struggled with providing accurate predictions when the GM was totally plantarflexed.

In order to provide insights regarding the performance of the prediction algorithms with different torques, a comparison of the torque values associated with prediction errors bigger than 5 N m (28 % of the maximum torque for passive trials) for both Dorsiflexion (Figure 5.7) and Plantarflexion (Figure 5.8) trials is provided.

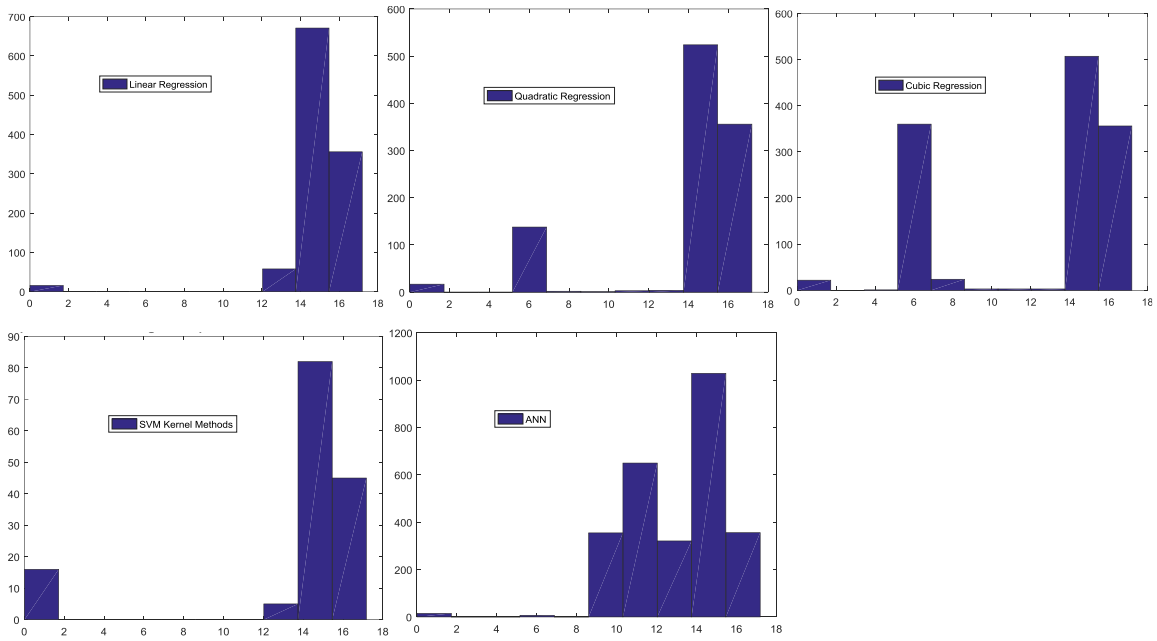

Figure 5-7. Histogram of the Torque values associated with prediction errors bigger than 5 N m during passive dorsiflexion for Linear, Quadratic, Cubic, Kernel and ANN methods. X axis Torque values in N m. Due to computational limitations values for Kernel SVR were collected from 1000 samples per trial only (8000 per trial for the other methods).

Results show that important errors in passive dorsiflexion predictions are associated with high torque values for all the techniques studied. Linear and Kernel have a similar distribution of the torque values with a small contribution of the error in the slowest torque values. An important contribution is observed in the mid values for quadratic and cubic regressions, associated with the relaxed state of GM region. Finally, ANN shows a more spread distribution of the errors through all the torque values.

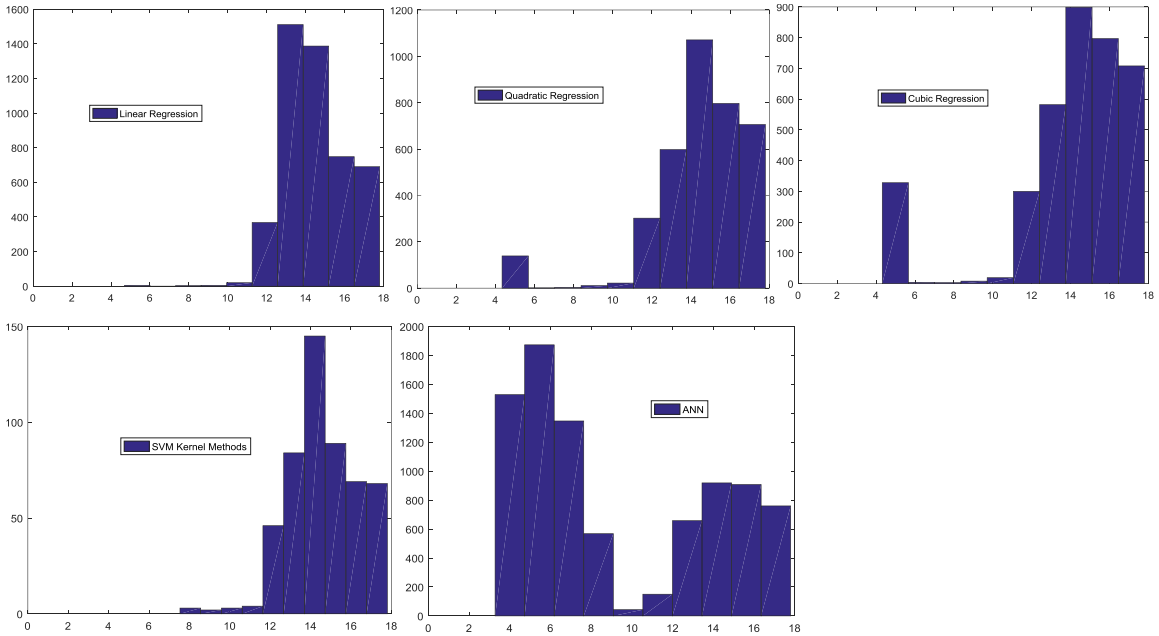

Figure 5-8. Histogram of the Torque values associated with prediction errors bigger than 5 N m during passive plantarflexions for Linear, Quadratic, Cubic, Kernel and ANN methods. X axis Torque values in N m. Due to computational limitations values for Kernel SVR were collected from 1000 samples per trial only (8000 per trial for the other methods).

Regarding the relationship between torque values and prediction errors bigger than 5 N m (28 % of maximum torque) during plantarflexion passive trials, similar conclusions as the ones extracted for dorsiflexion trials can be observed here for the different prediction algorithms.

5.6 Statistical Analysis

To determine if there were any statistically significant differences between the conditions studied in this chapter, a non-parametric ANOVA (Kruskal-Wallis) test was applied to the data contained in Tables 5.1- 5.8 using the "kuskalwallis" Matlab function. This test determines if the samples come from the same population, or from populations with the same distribution, by comparing the medians of the different groups. In Kruskal-Wallis test, the resulting p-value measures the significance of a Chi-square statistic, as opposed to the F-statistic used in parametric analysis of variance.

The torque RMSE values contained in the tables were compiled into wider groups by prediction algorithm, data precision, frame rate, and a combination of the latter two respectively. Despite the use of 0.05 as the critical value used for determining significance, Bonferroni correction was also applied when performing multiple comparisons dividing the critical value by the number of null hypothesis.

5.6.1 Prediction algorithms statistical influence

Kruskal-Wallis test was applied to all the data presented in this chapter divided into the five algorithms used for joint torque prediction.

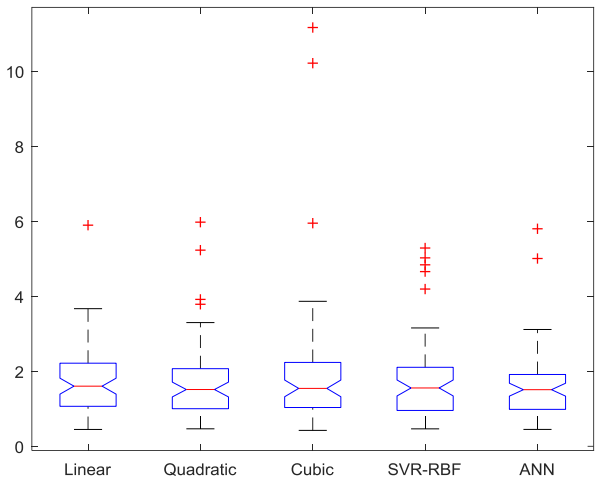

Figure 5-9. Kruskal-Wallis test results for the five prediction algorithms results during Passive movements

The analysis resulted in p = 0.9582 ($\chi 2$ = 0.37, Deg. of freedom = 4), which indicates that the null hypothesis that the five inference methods come from the same non-parametric distribution cannot be rejected at a 5% significance level.

As a result of this, differences appreciated between different prediction algorithms in the present chapter lack statistical significance and therefore strong conclusions should not be extracted from them.

5.6.2 Statistical assessment of the influence of data precision and Frame Rate

In order to analyse the influence of data precision and frame rate in the prediction of joint torque during passive movements, RMSE values were split into two groups per condition for every inference method. Kruskal-Wallis test was applied to the resulting two-column matrix obtaining the overall p-value per inference method, as detailed in Table x.x. Conditions comprised the comparison of 8-bit versus 16-bit US data precision, slow (32 fps) versus fast (1000 fps) frame rates, and the combination of 8-bit precision and slow frame rate versus 16-bit precision and fast frame rate respectively.

Table 5-9 Kruskal-Wallis p-values obtained per inference method using RMSE results comparing data precision, frame rate and a combination of the two.

Method	8bit vs 16bit	32fps vs 1kfps	32fps & 8bit vs 1kfps & 16bit
Linear	0.7697	0.1802	0.2293
Quadratic	0.7101	0.2847	0.2417
Cubic	0.8482	0.166	0.1946
SVR RBF	0.5064	0.0409	0.3113
ANN	0.7697	0.1802	0.2172

In order to compensate for the increased chance of incorrectly rejecting a false hypothesis when performing multiple tests to the same data sample, Bonferroni correction is applied to the critical P value, becoming 0.0167 (0.05/3) instead of the original 0.05.

No significant differences were seen for any of the conditions shown in Table 7.9, meaning that any differences appreciated in the RMSE results in previous sections should not be taken into account, due to the fact that the null hypothesis of all the conditions being part of the same distribution cannot be rejected at a 5% significance level.

5.7 Discussion

The work described here addressed the four questions posed at the beginning of the chapter. Firstly, no significant differences were observed in the prediction error of ankle joint torque generated during passive joint rotation between US data collected at 32 fps and 1000 fps. Error predictions were slightly worse (bigger in value) for the faster frame rate for both dorsiflexion (Figure 5.3) and plantar flexion (Figure 5.4) passive trials, but still not significant. As the angular velocity during passive movements was four degrees per second, for every degree the joint was rotated an overall of 8 and 250 frames were collected respectively for the two frame rates used (32 and 1000 fps). For the fastest frame rate, it could be possible that the movement between two consecutive frames would be too small for the ASM and for the OF algorithms to be able to account for a muscle displacement. Therefore if very small passive muscle movements are the main focus of any future experimental work, further consideration of the image analysis approach to extract small inter-frame differences would be required.

When considering the potential influence of data precision, opposite trends were observed between the use of raw 16-bit data and 8-bit (Figure 5.4 and Figure 5.5). Results were almost identical between the two precisions at conventional frame rate (32 fps); whereas at 1000 fps results were slightly better using 16-bit precision for the dorsiflexion trials, with the opposite outcome for the plantarflexion passive rotations (lower errors for the 8-bit precision). This fact could be associated with differences in the resulting images related to the type of US transmitted wave. Standard Clinical Ultrasound images usually show well defined contours, therefore a change in precision would not have an important impact in the performance of both ASM and OF algorithms. On the other hand, the better precision in the US data would probably have a bigger impact in granular looking images as the ones obtained using plane waves. The differences in predictions between the plantar and dorsiflexion conditions is a surprising and the reasons for this are unclear. Predictions during dorsiflexion (Figure 5.1 and Figure 5.3) were more accurate than plantar flexions (Figure 5.2 and Figure 5.4), which also produced higher variances

in the cross-validation of the nine participants. This could be related with the implications associated to both movements: During the dorsiflexion trials, the rate at which the muscle fascicles are being pulled is probably more gradual than the rate at which muscle fascicles slack during plantarflexion movements that could be related with the muscle region imaged (muscle belly), during which some small jumps can occur that might not impact the overall joint torque acquired with the dynamometer. This appreciation would support the findings in (Herbert et al. 2011) that refer to a small contribution of the GM muscle to the overall muscle-tendon unit lengthening (due to a bigger influence of the tendon) and specially when the muscle fascicles slack at short muscle-tendon lengths. Other factors that may have contributed to these results include consideration of possible torque contributions from other structures playing a larger role to net joint torque in plantar versus dorsiflexion. This is an aspect of the work that requires further investigation.

Section 5.5 graphically demonstrated how the five prediction techniques (Linear, Quadratic and Cubic Regressions, Kernel SVR and ANN) provide different approaches when analysing their response for a particular bad case. Further analysis showed an interesting trend when observing the torque at which errors bigger than 5 N m occurred. All prediction algorithms showed an important concentration of errors in the biggest torque values, with some content in areas of relaxed state region for the case of quadratic and cubic regressions. ANN provides a more spread distribution of errors in torque, showing even bigger concentrations in values of torque between 4 and 8 N m (equivalent to 22% to 44% of maximum torque). This indicates that the greater part of the input-output relationships that the algorithms were not capable to accurately predict were related with the extreme positions where the muscle tension had maximums and minimums.

No important interaction between two or more of the features explained above was observed for the case of joint torque predictions using US data during passive joint rotations. For the other effects seen two main factors can explain the reasons for not observing significant differences in the torque predictions for the experimental conditions of frame rate, data precision and inference algorithm evaluated here. The

first point relates to the nature of the movement observed. The behaviour of GM under passive movements is quite smooth and representative of the ankle joint overall torque. Secondly, due to the simplicity of the function to predict as it consists (Figure 5.5 torques bad fit) of a straight line that progressively goes up or down in time (dorsiflexion or plantarflexion).

From the results presented here it is now possible to recommend that, for the study of joint torques resulting from passive joint movements, 32 fps 8-bit ultrasound data is recommended, as it provides similar prediction results with improved efficiency as a result of having to process a smaller number of frames with reduced precision. Passive muscle movements however only make up a small portion of the conditions under which skeletal muscle function may be tested. As such, in order to conclude whether the use of the 16-bit precision raw data or a better temporal resolution can make a difference for the prediction of joint torque, the study of more complex muscle movements including voluntary activations is required. Chapter 6 therefore reports the results of analysis of data collected during isometric contractions, where the net joint torque produced is the result of voluntary muscle activation.

6 Predicting Ankle Joint Torque during Isometric Contractions

6.1 Introduction

This chapter reports the performance of inference algorithms namely Linear, Quadratic and Cubic Regressions, Kernel SVR and ANN; for the prediction of ankle joint torque produced by the performance of isometric contractions, defined as voluntary contractions with no changes in the muscle-tendon unit length. Evaluation of torque prediction during passive ankle joint rotations (Chapter 5) provided insight into torque prediction capabilities within relaxed, non-activated muscles. This necessarily involved evaluation of a relatively simple relationship between muscle shape change and resulting torque. Movement tasks however require skeletal muscle activation for the production of mechanical force and power. The influence of US data characteristics and differences between torque prediction algorithms may therefore not hold for active contractions, meaning that specific assessment of different forms of muscle activation are required to more fully identify the relative importance in terms of the *in vivo* study of skeletal muscle behaviour.

Analogous to chapter 5, this chapter aims to address four main questions related with the main features analysed:

- Is there a difference in the prediction error of ankle joint torque generated isometric contractions using image analysis of data collected at different frame rates?

- Is there a difference in the prediction error of ankle joint torque generated isometric contractions using image analysis of data collected with different precision?

- Is error in the prediction of ankle joint torque generated isometric contractions affected by the prediction technique used?

- Is there an important interaction between two or more of these three features?

This chapter describes the performance of Inference methods (detailed in Section 4.3) for predicting joint Torque from US processed data of the GM muscle during isometric contractions (See Section 4.1.1).

In order to validate the use of prediction techniques in ultrasonography, and evaluate their performance during isometric contractions this chapter firstly describes previous work and phenomena associated with the study of isometric voluntary contractions (Section 6.2). After that, the prediction results during isometric contractions at two submaximal activations (30 and 60% MVC) using 8-bit and 16-bit US data precision are showed for conventional framerates of 32 fps (Section 6.3), and for 1000 fps (Section 6.4), followed by a study of the performance of different prediction techniques (Section 6.5). Finally, Section 6.6 provides an overview of the results obtained in previous sections and discusses the potential application of the methodology for isometric contractions.

6.2 Isometric contractions in lower limb muscles

Isometric contractions refer to muscle activations that do not produce a change in the length of the muscle-tendon unit. During these contractions, the joint angle remains constant, with the tension produced always smaller than the load. Nevertheless and due to the elastic properties of tendinous tissues, shortening of the muscle fascicles can occur, as shown by (Ito et al. 1998) for the case of the tibialis anterior muscle during isometric dorsiflexions. As a result, muscle fascicles shorten and, in pennate muscles, vary the pennation angle while generating forces.

The total muscle-tendon unit length remains unchanged as the compliant tendons are stretched by the fascicle displacement. More recent studies have shown a similar behaviour for the case of GM muscle during submaximal isometric contractions (Héroux et al. 2016) under similar experimental conditions to those recorded here (Chapter 4). In that work, no significant changes were observed between proximal (close to the knee joint) and distal (ankle joint area) fascicle shortening, including also identical onset times in both ends of the GM muscle (Héroux et al. 2016).

Previous work regarding parameter inference using US data during isometric contractions has been focused on estimating the maximum voluntary torque possible, through measurement of muscle volumes (Miyatani et al. 2004) using EMG driven Hill-models (DeOliveira & Luporini Menegaldo 2010) or Magnetic Resonance imaging (Baxter & Piazza 2014). These predictions were based on the estimation of an overall parameter non-dependant of temporal variances; therefore, no previous attempts of data inference of instantaneously generated net joint torque have been found in literature. Such information is however important for understanding *in vivo* muscle function both in healthy populations and those with pathologies affecting skeletal muscle function and mobility.

6.3 Torque Predictions using Standard Clinical Ultrasound

The performance of prediction algorithms estimating the ankle torque during isometric contractions is studied in this Section using Standard Clinical US, which consisted of the use of US focused waves at a framerate of 32 fps. Isometric trials were acquired at two different submaximal contraction levels (30% and 60%), and results were processed using the cross-validation methodology explained in the previous section.

The prediction of the joint torque using the Linear, Quadratic and Cubic Regressions, Kernel Methods and ANN was subtracted to the empirically obtained with the

dynamometer, obtaining the RMSE that provides an indicator of their goodness of fit. These results are compiled in Tables in which "Participant #" represents the participant whose data was separated from the validation process and used as validation.

6.3.1 Isometric Contractions at 30% MVC

The following section reports the results from the analysis of different precision data (8-bit and 16-bit) relating to isometric contractions at 30% of the MVC value. The results corresponding to the use of 8-bit data are shown in Table 6.1 and include the individual participant and the overall average and standard deviation per prediction algorithm used.

Table 6-1. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during Isometric contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.821	6.266	6.251	6.468	3.307
Participant 2	7.109	7.798	7.460	6.405	6.267
Participant 3	6.220	4.875	4.794	5.268	5.822
Participant 4	4.361	4.438	4.329	4.507	3.935
Participant 5	4.121	4.341	5.195	3.700	4.756
Participant 6	5.192	3.815	3.727	3.534	3.495
Participant 7	8.740	7.846	8.190	8.675	9.080
Participant 8	11.247	8.422	9.432	9.032	8.319
Participant 9	5.302	5.053	4.973	4.725	4.242
Average	6.568	5.873	6.039	5.813	5.469
Stand. Dev.	2.276	1.752	1.932	2.006	2.088

Small differences can be appreciated between the different prediction algorithms, with slightly better results found for ANN in terms of average. This best fit result for this condition (5.813 ± 2.006 N m) corresponds to 19.39% of the average participant maximum torque during 30% MVC passive trials (29.97 N m).

Same results were obtained using the 16-bit precision data and compiled in table 6.2. Small differences can be appreciated between prediction algorithms, finding slightly better results for quadratic regression and Kernel Methods in terms of

standard deviation and average respectively. The average best result (6.399 N m) corresponds to a 21.3% of the average participant maximum torque during passive trials (29.97 N m).

Table 6-2. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during Isometric contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.799	6.441	6.588	6.562	6.137
Participant 2	8.111	8.592	8.010	7.542	6.533
Participant 3	6.661	5.147	4.965	5.612	6.066
Participant 4	4.885	5.028	4.825	5.198	4.947
Participant 5	3.958	4.395	5.464	3.867	6.054
Participant 6	5.701	4.429	4.360	3.941	5.131
Participant 7	8.832	8.092	8.461	8.998	8.867
Participant 8	12.268	9.670	10.848	10.316	11.324
Participant 9	5.739	6.089	5.918	5.551	6.228
Average	6.995	6.432	6.604	6.399	6.810
Stand. Dev.	2.488	1.930	2.124	2.198	2.028

The prediction results using 16 and 8-bit data for 30% MVC isometric contractions were compiled and graphically expressed (Figure 6.1) using boxplots, in order to account for the distribution of the obtained errors per prediction algorithm used in further detail.

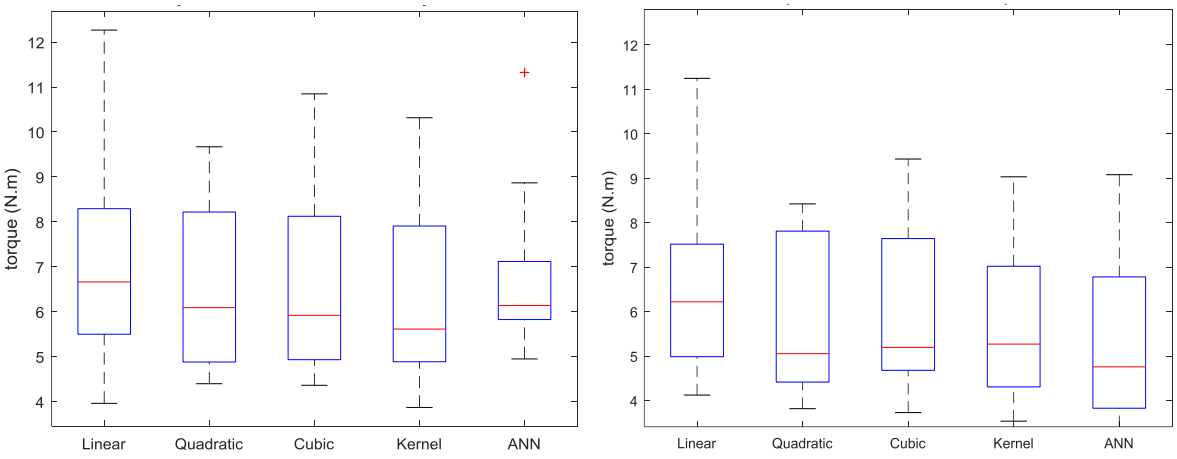

Figure 6-1. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Isometric measurements at 30% MVC from nine participants.

The distribution of the results was quite similar when comparing 16-bit and 8-bit results for the 30% MVC isometric contractions in both median and variance.

6.3.2 Isometric Contractions at 60% MVC

In order to enable the study of how stronger contractions can affect the prediction of joint torque in conventional framerates, acquisitions at 60% MVC were performed using the same experimental setup and conditions.

A resume of the errors obtained when using the different prediction algorithms can be found in Table 6.3, including the resulting averages and standard deviations in Torque units (Nm).

Table 6-3. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during Isometric contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.365	7.133	8.120	6.826	6.692
Participant 2	21.274	16.278	15.895	16.270	11.716
Participant 3	9.618	6.161	6.508	7.376	5.623
Participant 4	13.679	9.949	9.646	11.166	12.484
Participant 5	8.036	8.267	9.294	7.355	6.973
Participant 6	17.229	15.165	15.078	13.508	12.689
Participant 7	13.838	10.587	10.534	12.107	10.064
Participant 8	34.559	33.074	32.801	33.807	33.693
Participant 9	6.152	5.141	5.036	4.583	3.472
Average	14.528	12.417	12.546	12.555	11.490
Stand. Dev.	9.071	8.630	8.392	8.794	8.942

From the table it can be observed that Linear Regression provides the worse fit of all methods. Quadratic and Cubic Regressions and Kernel Methods provide almost identical results in both terms of average and deviation. The best fit was observed with ANN, with a RMSE of 11.490 ± 8.942 N m equivalent to a 19.17% of the average maximum torque during 60% MVC isometric trials (59.94 N m).

In contrast, Table 6.4 presents the individual performance together with overall average and deviation values when processing 16-bit data precision.

Table 6-4. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during Isometric contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	5.599	7.953	9.046	7.147	7.460
Participant 2	23.482	19.601	19.330	19.051	17.083
Participant 3	12.114	8.606	8.730	10.161	5.640
Participant 4	15.214	11.226	10.694	12.447	12.241
Participant 5	8.914	9.537	10.679	8.178	10.094
Participant 6	18.405	16.186	16.094	14.682	13.742
Participant 7	14.571	11.686	11.555	13.121	13.217
Participant 8	33.756	31.800	31.233	32.605	33.429
Participant 9	6.508	4.815	4.625	5.157	4.270
Average	15.396	13.490	13.554	13.617	13.020
Stand. Dev.	8.944	8.167	7.875	8.275	8.694

Similar values can be found among the different algorithms in both average and standard deviation except in linear regression, which predicted worse than the rest. Best performance was found for the ANN, resulting in an average of 13.020 N m difference from the empirically acquired torque which corresponds to a 21.72% of the average maximum torque during 60% MVC isometric tasks (59.94 N m).

Despite the fact that average results were proportionally similar to the ones at 30% MVC (values between 13 and 15 N m at 60% MVC corresponding to values between 6.5 and 7.5 N m at 30% MVC), important differences can be appreciated for the standard deviation values, which were much bigger in the case of 60% MVC.

The prediction results using 16 and 8-bit data for 60% MVC isometric contractions showed in tables 6.3 and 6.4 were compiled and graphically expressed (Figure 6.2) using boxplots, in order to account for the distribution of the obtained errors per prediction algorithm used in further detail.

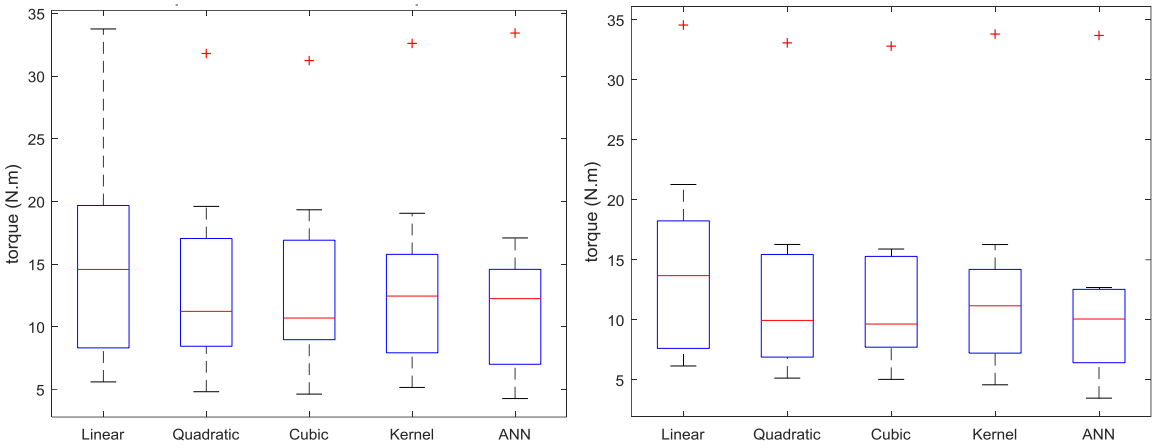

Figure 6-2. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Isometric measurements at 60% MVC from nine participants.

The distribution of the results was quite similar when comparing 16-bit and 8-bit results for the 60% MVC isometric contractions in both median and variance.

Kernel Methods and Quadratic and Cubic regressions showed more consistent and accurate predictions for the 4 conditions studied in this section (8 and 16-bit precision at 30% and 60% MVC isometric contractions). On the other hand, linear regression provided the worse fits to the empirically collected joint torque, showing its limited application for isometric contractions at conventional framerates.

6.4 Torque Predictions using Ultra-Fast frame rate Ultrasound

In order to evaluate the effect of temporal resolution in the prediction of generated torque from US individual muscle data, the same experiment showed in the previous section for standard clinical US was undergone at a framerate of 1000 fps. This increase in the temporal resolution comes at a cost, the use of plane waves for the US transmission that lack lateral resolution as explained in Chapter 2.

As a result, 8000 US images were collected per trial (8 seconds long) as opposed to the 256 obtained for the measurements in the previous section. This increase in the size of the collected data exponentially increased the computational time of the

post-processing of the US data used for the training and validation of the prediction algorithms. Because of these, results regarding Kernel methods with ultra-fast framerates only used 1000 data points (8 times smaller temporal resolution) per trial due to memory and computational limitations.

The following sections detail the results of the prediction of the different algorithms for isometric contractions at two different activation levels (30 and 60% MVC), also showing the influence of the data precision in their performance.

6.4.1 Isometric Contractions at 30% MVC

The following section reports the results from the analysis of 8-bit and 16-bit data relating to the isometric contractions at a submaximal contraction level of 30% MVC. Table 6.5 presents individual participant and mean group data from the 8-bit, fast frame rate data.

Table 6-5. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during Isometric contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.323	6.260	6.206	6.225	6.482
Participant 2	4.302	3.843	3.837	4.043	3.964
Participant 3	7.405	7.274	7.232	7.316	7.792
Participant 4	3.953	3.643	3.797	3.847	3.881
Participant 5	11.298	11.057	11.220	11.114	10.959
Participant 6	9.582	11.638	11.608	10.913	11.853
Participant 7	2.168	1.958	1.942	2.126	2.140
Participant 8	4.801	4.554	4.564	4.678	4.532
Participant 9	3.188	2.928	3.745	2.345	3.582
Average	5.891	5.906	6.017	5.845	6.132
Stand. Dev.	3.041	3.483	3.417	3.362	3.422

All the inference techniques provided similar results for 8-bit data in both terms of average and standard deviation for this conditions. Kernel SVR showed the best fit (5.845 ± 3.362 N m) equivalent to a 19.5% of the average participant maximum torque during 30% MVC passive trials (29.97 Ns m).

Furthermore, the results for the different prediction algorithms using 16-bit US data precision at fast frame rate can be found in Table 6.6, including individual participant results and global mean and deviation values.

Table 6-6. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during Isometric contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.099	6.133	6.140	6.019	6.222
Participant 2	4.286	3.925	3.899	4.285	7.159
Participant 3	7.286	7.195	7.200	7.222	6.967
Participant 4	3.668	3.498	3.522	3.738	4.433
Participant 5	10.626	10.614	10.614	10.574	6.411
Participant 6	10.811	12.850	12.974	11.971	11.818
Participant 7	1.783	1.677	1.682	1.962	1.733
Participant 8	4.408	4.270	4.293	4.501	4.219
Participant 9	1.409	1.385	1.619	1.253	1.736
Average	5.597	5.727	5.771	5.725	5.633
Stand. Dev.	3.437	3.914	3.909	3.649	3.105

From the table 6.6 it can be appreciated a quite similar behaviour in all the prediction algorithms in both average and deviation, finding slightly better results for ANN and Linear Regression. The latter showed the best average result (5.597 N m) corresponding to 18.67% of the average participant maximum torque during passive trials.

The different data precision information contained in tables 6.5 and 6.6 at ultra-fast framerates was compiled and showed in Figure 6.3 for the 30% MVC submaximal isometric contraction level.

This graphical results provide information regarding the median and variance of the distribution of the errors per prediction algorithm used.

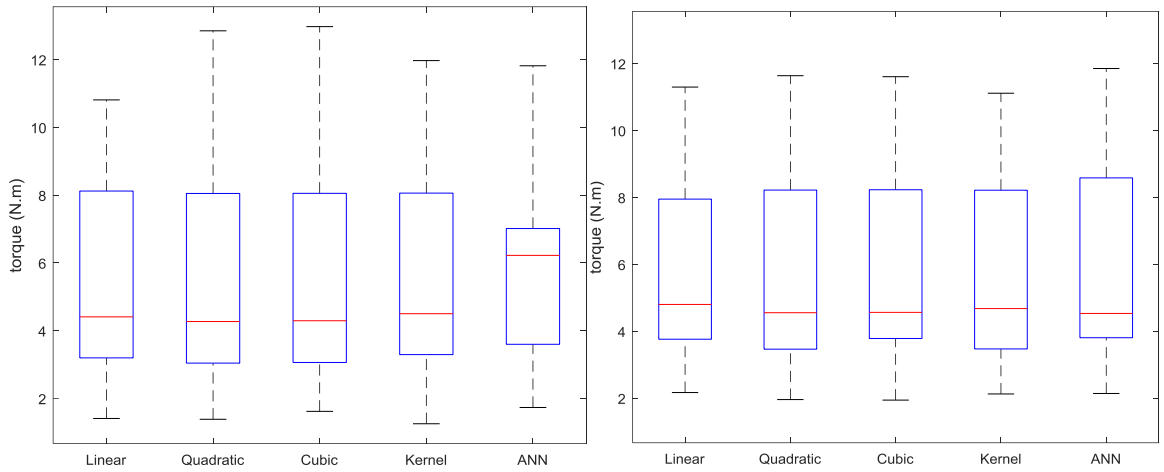

Figure 6-3. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Isometric measurements at 30% MVC from nine participants collected at 1kS/s.

Small differences can be observed between prediction algorithms for the case of the median, showing similar variances for both 16 and 8-bit data precision. Only ANN provided worse average results but smaller deviation for the bigger precision results.

6.4.2 Isometric Contractions at 60% MVC

Similar to previous section with 30% MVC isometric contractions, equivalent results for the biggest submaximal activation level studied in this work (60% MVC) when acquiring US data at 1000 fps are present in this section using two different US data precision. Table 6.7 presents individual participant and mean group data from the 8-bit, fast frame rate data.

Table 6-7. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during Isometric contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	12.597	10.737	11.362	11.454	11.557
Participant 2	20.350	26.235	26.385	23.803	21.263
Participant 3	6.037	4.584	4.173	4.854	4.979
Participant 4	6.315	4.802	5.713	5.428	5.633
Participant 5	18.037	15.372	16.033	16.320	14.263
Participant 6	9.225	7.250	7.464	7.260	8.053
Participant 7	5.713	4.100	4.194	4.733	4.628
Participant 8	9.161	6.535	6.198	7.339	6.000
Participant 9	3.904	5.246	5.262	3.724	6.924
Average	10.149	9.429	9.643	9.435	9.255
Stand. Dev.	5.749	7.275	7.374	6.697	5.532

Linear Regression provided the worst fit in terms of average, however its standard deviation was one of the lowest. Quadratic and Cubic Regressions showed almost identical results in both terms of average and deviation. The best fits were observed with Kernel SVR and ANN, being slightly better for the latter one, resulting in a RMSE of 9.255 ± 5.532 N m equivalent to a 15.44% of the average maximum torque during 60% MVC isometric trials (59.94 N m).

In contrast Table 6.8 presents individual participant and mean group data, resulting from analysis of 16-bit precision data during isometric contractions at 60 % MVC for a frame rate of 1000 fps.

Table 6-8. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during Isometric contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	9.867	6.545	6.442	7.719	5.147
Participant 2	10.306	6.152	6.003	7.448	6.838
Participant 3	5.155	2.577	2.503	3.168	2.340
Participant 4	5.686	3.518	3.548	4.311	3.119
Participant 5	13.441	7.999	7.858	9.998	7.389
Participant 6	7.653	5.316	5.675	4.169	6.002
Participant 7	4.830	3.286	3.271	3.546	3.157
Participant 8	6.694	4.516	4.757	4.793	4.130
Participant 9	3.152	3.781	3.902	2.587	3.776
Average	7.420	4.854	4.884	5.304	4.655
Stand. Dev.	3.244	1.780	1.739	2.501	1.778

Linear Regression showed worst results in both terms of average and standard deviation, followed by Kernel SVR. Quadratic and Cubic Regressions together with ANN provided the best fits to the empirically acquired joint torque, being slightly better for the last one. ANN RMSE value (4.655 ± 1.778) corresponded to 7.77% of the average maximum torque during 60% MVC isometric tasks (59.94 N m). This value represents the most accurate prediction of Joint torque observed.

The different data precision information for the 60% MVC submaximal isometric contraction level at ultra-fast framerates was compiled and showed in Figure 6.4, These graphical results provide information regarding the median and variance of the distribution of the errors per prediction algorithm used.

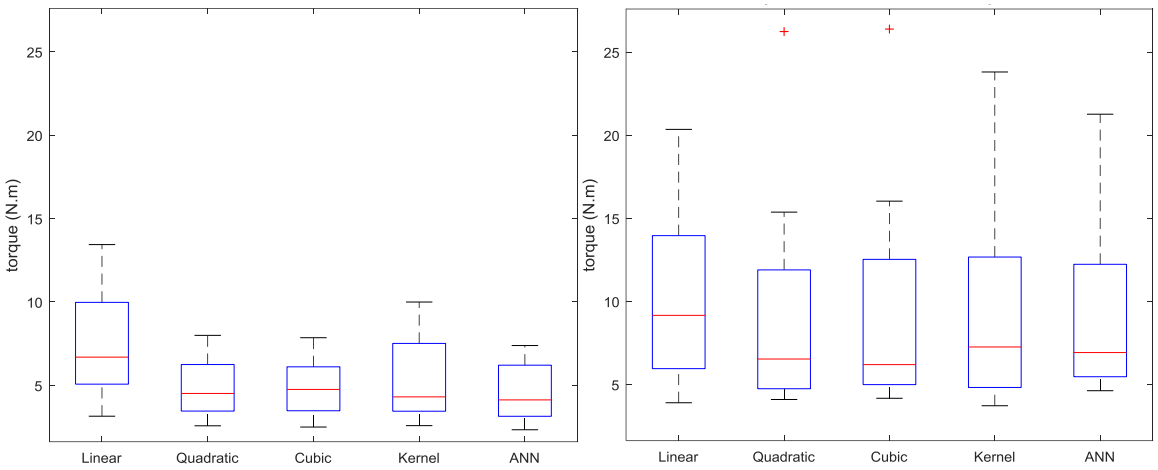

Figure 6-4. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Isometric measurements at 60% MVC from nine participants collected at 1kS/s.

Important differences can be appreciated for the case of 60% MVC isometric contractions. The use of higher precision data improved notably the fit of the predicted torque to the empirically acquired values in both median and variance, moving from an average for the five algorithms of 9.58 N m for the 8-bit precision, and 5.424 for the case of 16-bit; that correspond to a 15.98% and 9.05% of the average maximum torque for 60% MVC isometric contractions respectively.

6.5 Features Underlying Algorithm Performance

As showed in Chapter 5, an effective way of understanding how the 5 inference methods used in this thesis provide different approaches to solve the prediction of joint torque from US related information, is to analyse their behaviour when they provide bad fits of the desired data. In order to provide insight into this matter, Figure 6.5 shows the performances of the different prediction methods (Linear, Quadratic and Cubic regressions, SVM Kernel Methods and ANN) during an isometric trial at 30% MVC that numerically provided a fit of the original torque with an error bigger than the 30% of the maximum Torque.

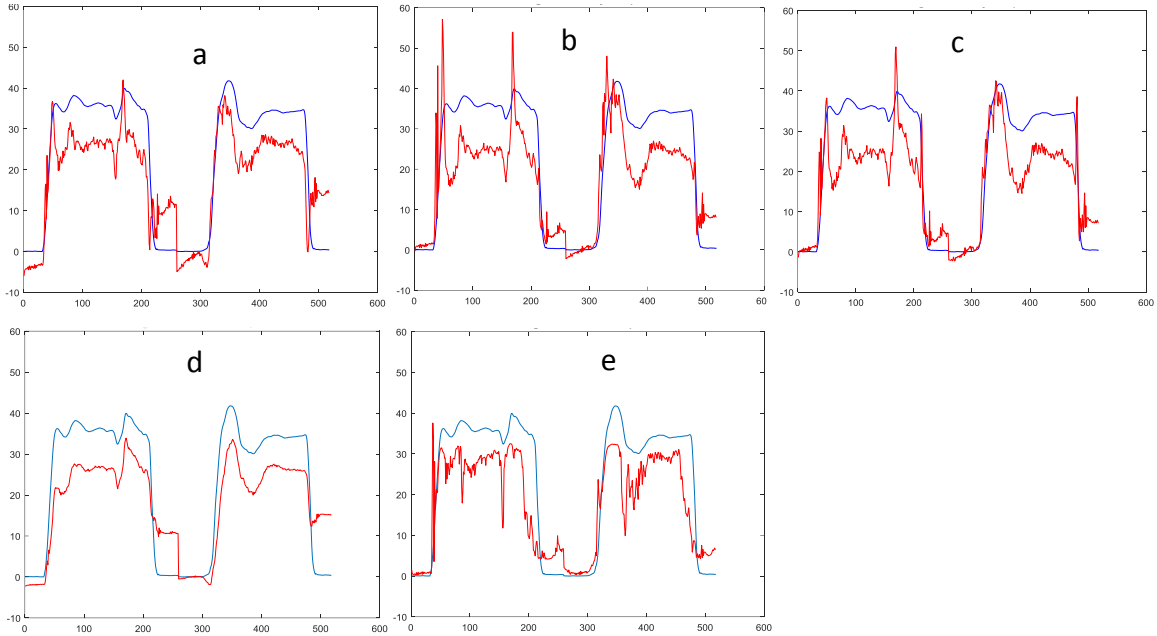

Figure 6-5. Acquired Joint Torque (blue) vs. Predicted Torque (red) graphical results in N m for the Isometric trials at 30%MVC for 5 prediction methods: Linear (a), Quadratic (b), Cubic (c), Kernel SVR (d) and ANN (e). Resulting RMSE Torque values were 9.11, 9.67, 9.39, 9.18 and 9.28 N m respectively.

As opposed to the equivalent analysis for passive trials (Section 5.6), in which the different algorithms clearly showed different behaviours during bad predictions, their performance is quite similar during isometric trials. Only Kernel SVR (Figure 6.5-d) showed a behaviour slightly different than the rest of inference algorithms, having a smoother response throughout the entire trial.

Great part of the mismatches observed between the predicted and the empirically acquired torque values consist of unproportionate increases and decreases during fast changes of the torque associated with the beginning and end of activations. Another important factor to consider is the appearance of different torque levels during joint torque relaxing states, showing contraction levels in the muscle that are not associated with any joint torque.

In order to further study the behaviour of the different inference algorithms used in this chapter, the distribution of torque values associated with prediction errors bigger than 5Nm was studied for Isometric trials at 30% (Figure 6.6) and at 60% MVC (Figure 6.7).

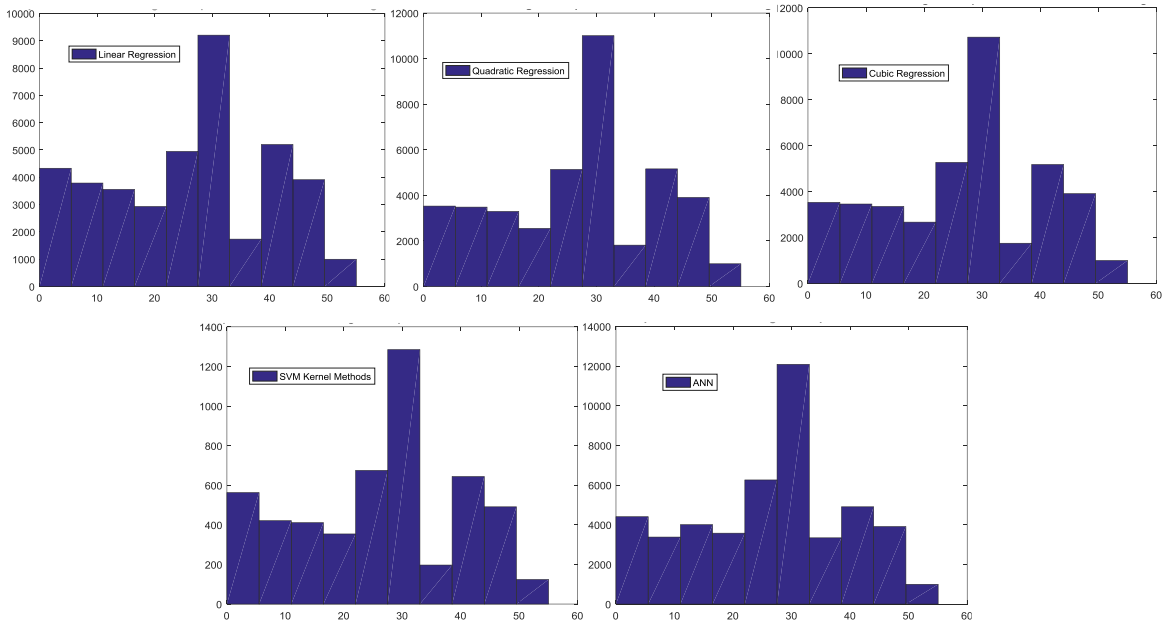

Figure 6-6. Histogram of the Torque values associated with prediction errors bigger than 5 N m during Isometric 30% MVC isometric trials for Linear, Quadratic, Cubic, Kernel and ANN methods. X axis Torque values in N m. Due to computational limitations values for Kernel SVR were collected from 1000 samples per trial only (8000 per trial for the other methods).

Due to the nature of the isometric tasks, the greatest part of the data is concentrated in torques close to both the relaxed state and to the desired activation level. Therefore, for the case of 30% MVC isometric tasks, and if the errors were evenly distributed, two peaks would be observed at torque equal to 0 and 30 N m respectively. The 5 prediction algorithms show a similar distribution of errors, consisting of an important peak in the 30% MVC area and an evenly spread error in the rest of torque values. However, as opposed to described before, no important errors were observed around the repose state (O N m), meaning that all inference algorithms provided good predictions of the ankle joint torque with low or non-existent activations in the GM muscle.

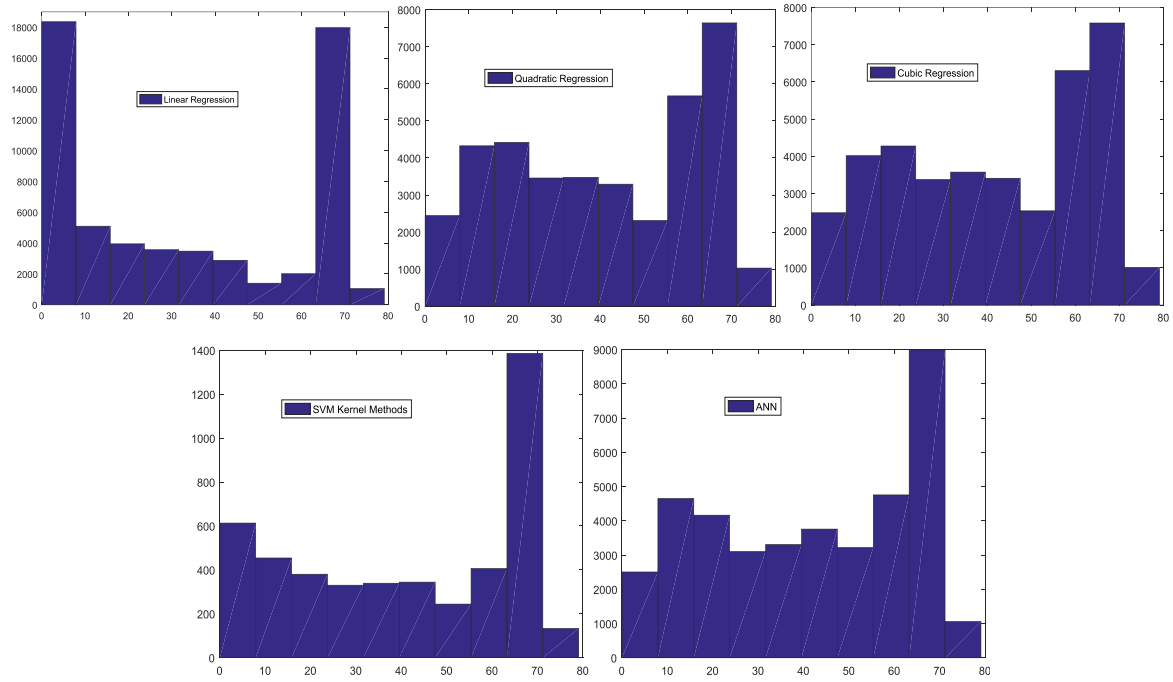

Figure 6-7. Histogram of the Torque values associated with prediction errors bigger than 5 N m during 60% MVC isometric trials for Linear, Quadratic, Cubic, Kernel and ANN methods. X axis Torque values in N m. Due to computational limitations values for Kernel SVR were collected from 1000 samples per trial only (8000 per trial for the other methods).

Regarding the torque distribution associated with errors bigger than 5Nm for isometric tasks at 60% MVC, only the Linear Regression showed an evenly error distribution as explained above. Nevertheless, this technique also showed the worst results for this conditions, therefore a more spread torque error was not linked with a lower error in the prediction in this case. On the other side, the inference technique that provided best results (ANN) had an evenly distribution of the errors outside the 60% MVC torque values, as also observed for the Quadratic and Cubic Regressions and SVM Kernel Methods.

6.6 Statistical Analysis

Analogously to Section 5.6, statistical analyses were performed in order to be able to distinguish if the results in tables 6.1 – 6.8 show statistical differences between the different factors studied in this chapter, namely data precision, frame rate, muscular activation level and inference method.

Kruskal-Wallis test p-values are used as a descriptor using the null hypothesis that the factors compared come from the same distribution. In other words, significant p-values (smaller than 0.05) mean that this hypothesis can be rejected and therefore there is evidence that the conditions compared are different.

Statistical tests were applied to the data shown in previous sections of this chapter rearranged in order to obtain the significance of five different factors: inference algorithm, percentage of MVC, data precision, frame rate and a combination of the last two.

6.6.1 Prediction algorithms statistical influence

Firstly, the data from the tables 6.1 to 6.8 was divided into five different columns of a matrix containing the RMSE values of each of the corresponding inference algorithms: Linear, Quadratic and Cubic regressions, Support Vector Regression and Artificial Neural Network. Kruskal-Wallis was applied to the resulting matrix obtaining the distribution shown in Figure 6.9.

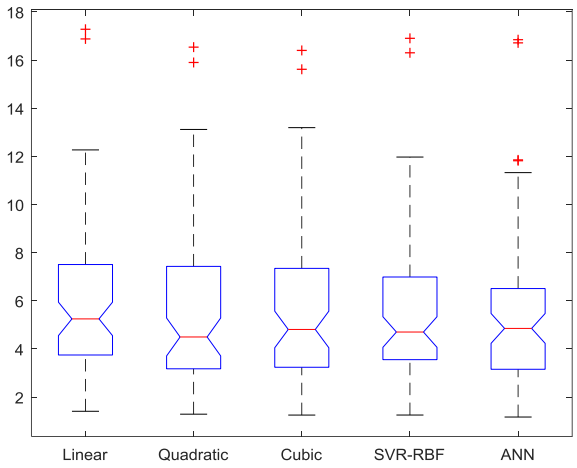

Figure 6-8. Kruskal-Wallis test results for the five prediction algorithms results during Isometric contractions

According to the results obtained ($p = 0.5653$, $\chi 2 = 2.96$, Deg. of freedom = 4)), the hypothesis that the data from the different inference methods comes from the same distribution cannot be rejected, therefore there is no statistical power to

determine if one inference method performs better or worse than other for isometric contractions.

6.6.2 Data precision, frame rate and activation level statistical influence

RMSE values obtained during isometric contractions were firstly split 8-bit and 16-bit data precision for one of the inference methods. Kruskal-Wallis test was applied to the resulting two-column matrix obtaining a p-value as the main outcome. Iteratively, this process was firstly repeated for the remaining inference methods and then splitting the data in terms of frame rate, frame rate together with data precision, and activation level. Table 6.9 contains the p-values corresponding to the four conditions analysed and the five inference methods. Error values at 60% MVC were halved in order to provide comparisons using equivalent magnitudes to the 30% MVC without affecting the outcome of the statistical analyses.

Table 6-9. Kruskal-Wallis p-values obtained per inference method using RMSE results comparing data precision, frame rate, a combination of the two and activation levels during isometric contractions.

Method	8bit vs 16bit	32fps vs 1kfps	32fps&8bit vs 1kfps&16bit	30 vs 60 %MVC
Linear	0.7018	0.0041	0.0192	0.246
Quadratic	0.581	0.0015	0.0065	0.025
Cubic	0.5581	0.0013	0.0044	0.0315
SVR RBF	0.7525	0.0087	0.0397	0.0438
ANN	0.835	0.0261	0.0689	0.007

As explained in Chapter 5, the critical P-value needed for being able to reject the null hypothesis becomes 0.0125 (0.05/4) after applying the Bonferroni correction.

No significance was obtained for effort level (30 vs 60 %MVC) or data precision, meaning that the differences appreciated in previous sections regarding data precision and activation level could change if performing again the same experiment under the same conditions.

However, the null hypothesis should be rejected for all inference methods but ANN when comparing frame rates, and for quadratic and cubic regressions when comparing 8-bit data at low frame rate versus 16-bit data at fast frame rate. For these cases, the statistical analysis indicates that the described differences in RMSE between conditions are statistically different.

6.7 Discussion

The work described in this chapter addressed the validation of different inference methods able to accurately predict ankle net joint torques from US data of the GM during Isometric contractions at two activation levels. The influence of using different data precision and temporal resolution on the accuracy of the predictions was also studied, to identify whether there was a more suitable algorithm depending on the task and conditions involved. As with consideration of the passive data (Chapter 5), four main questions were addressed (Section 6.1). The results however provide different answers to those from the passive data and will be reviewed in more detail below.

From the results presented here, temporal resolution proved to be an important factor for an accurate prediction of joint torque. There was an improvement in the average RMSEs of the five inference algorithms of 16.8% for 30% MVC isometric trials, and of 155% for the 60% MVC ones when comparing the two frame rates studied (Figures 6.1vs 6.3 and Figures 6.2 vs 6.4). Due to the fact that, in the experimental set-up here (i.e. ramping to 30% or 60% MVC was completed over the same time frame), stronger contractions would lead to faster muscle fascicle movements, this finding indicates that the algorithms performance is related with the rate of change in the muscle during activation and relaxation phases. To my knowledge however, no previous work has assessed how contraction velocity affects algorithm performance in US data so it is difficult to put this finding in the context of the current literature. However, these results suggest that frame rate is important if faster movement tasks are to be assessed.

The use of 16 and 8-bit data precision did not impact the performance of prediction in slow frame rate acquisitions (Figures 6.1 and 6.2). However, a significant influence of the data precision could be observed for the case of 60% MVC isometric trials at 1000 fps framerate, which provided a 76.8% reduction on the average error (expressed as RMSE) for the five prediction algorithms used(Figure 6.4). This finding links with the fourth question proposed regarding the interaction between two or more features studied, as the most accurate predictions were observed when using the highest data precision (16-bit) and the fastest frame rate (1000 fps), associated with a higher temporal resolution. These conditions resulted in an RMSE value equivalent to only 7.75% of the maximum torque during those trials. This information reinforces the discussion stated for passive contractions in Section 5.6 regarding the need of better precision in grainy looking US images such as the ones collected with the use of plane waves for transmission. The association between high US data precision and high temporal resolution seems to strengthen when studying muscle movements with increasing velocity in activation and relaxation phases and it is predicted should be also observed in other muscle behaviours such as muscle twitches (Harding et al. 2015) or faster contractions associated with control tasks.

With regard to the differences between the prediction techniques used, no significant differences were appreciated for the five inference methods (Figure 5.9). Nevertheless, it should be noted that ANN seemed to provide more consistent predictions in terms of RMSE average in 5 out of the 8 conditions studied in the present chapter (Tables 6.1 to 6.8). This finding indicates that the higher capacity of this prediction technique to understand complex input-output relationships could be able to make a difference in terms of predictive power. This could be related to the increased complexity of the torque function to be predicted. In comparison to the linear behaviour observed for passive movements (Section 5.6), the isometric torque profile consisted of a hill with small fluctuations occurring at time points throughout the trial, as can be seen in figure 6.6. Further differences between prediction algorithms studied in this research could be found when looking at more

complex movements that both include muscle activations and joint rotations and is an aspect that will be addressed in Chapter 7.

With regard the behaviour of the inference algorithms when they provided worse predictions, a torque level mismatch during 0 joint torque levels was observed in Section 6.5. These discrepancies could be related with the hysteretic behaviour of the muscle-tendon unit, defined as inconsistencies in the length-tension curve during shortening and lengthening of muscles that cannot be attributed to force-velocity characteristics (Gillard et al. 2000). As a result, nonlinearities are found on the muscle behaviour, as previously reported (Finni et al. 2012; Seynnes et al. 2015) who compared different methodologies for the assessment of hysteresis values *in vivo*, and discussed the analysis of mechanical properties of tendons with US respectively evaluating whether the hysteresis seen were attributable to the experimental setup. According to the conclusions extracted from my work, where experimental signals were synchronously collected and US data inter-frame intervals quantified (Chapter 4), the differences observed suggest that hysteresis occurs at the level of the muscle fascicles therefore it is not attributable to the experimental setup and is a physiological phenomenon which could be further investigated.

As opposed to passive movement results (Chapter 5), for which no important differences were observed, the prediction results for isometric contractions significantly improved when using bigger data precision and better temporal resolution. The better temporal resolution seemed a key factor for being able to account for fast changes in the generated torque associated with rapid contractions of the muscle fascicles of the GM muscle. The need of using plane waves for the US transmission in order to achieve ultra-fast frame rates, resulted in grainy looking US images that lacked well-defined muscle fascicles. Due to this, the use of the higher data precision proved to have an important impact on the way that inference algorithms accurately predict joint torque from US images of the GM. As a result, and despite the fact that the resulting plane wave images can seem to be of "bad quality" to the human eye, the work undergone in this chapter proves that the grainy-looking US images contain more valuable information for the prediction of

torque in isometric contractions. Therefore, for future studies of isometric contractions experimenters should consider whether implementation of plane-wave imaging modalities is required.

Once showed the successful performance of prediction algorithms with US data during isometric contractions, the flow of this research should lead to the study of even more complex muscular movements. These movements should combine ankle joint movements and muscular activations at different MVC levels, and might show important differences regarding the prediction algorithms able to provide accurate predictions. Thus, isotonic contractions are the main focus of the last experimental chapter of this thesis, detailed in Chapter 7.

7 Torque Prediction during Isotonic Contractions

7.1 Introduction

The results presented here report the performance of data inference algorithms for the prediction of net ankle joint torque from US processed data collected during eccentric and concentric isotonic muscle contractions. Evaluation of torque prediction during passive ankle joint rotations (Chapter 5) and isometric contractions (Chapter 6) provided insight into torque prediction capabilities within relaxed, non-activated muscles, and with activations without changes in the muscle-tendon unit length respectively. Isotonic movement tasks require skeletal muscle activation for the production of mechanical force and power. The influence of US data characteristics and differences between torque prediction algorithms may therefore not hold for active contractions where muscle tendon unit length changes also occur, meaning that specific assessment of different forms of muscle activation are required to more fully identify the relative importance in terms of the *in vivo* study of skeletal muscle behaviour.

Analogous to chapters 5 and 6, this chapter aims to address four main questions related with the main features analysed:

- Is there a difference in the prediction error of ankle joint torque generated during isotonic contractions using image analysis of data collected at different frame rates?

- Is there a difference in the prediction error of ankle joint torque generated isotonic contractions using image analysis of data collected with different precision?

- Is error in the prediction of ankle joint torque generated isotonic contractions affected by the prediction technique used?

- Is there an important interaction between two or more of these features?

This chapter describes the performance of inference methods (detailed in Section 4.3) for predicting net ankle joint torque from US processed data of the GM muscle during isotonic contractions (See Section 4.1.1).

Firstly, prediction results are studied for conventional frame rates using focused waves for the US transmission (Section 7.2), including a study on the influence of US data precision in the prediction accuracy. These same conditions are posteriorly studied for ultrafast frame rates (1000 fps, Section 7.3). These results are then compared (section 7.4) providing insight regarding the influence of temporal resolution in US data based predictions. Posteriorly, the temporal response of the prediction algorithms when providing erratic fits is analysed (Section 7.5). Finally, conclusions are extracted from the overall analysis of the obtained results (Section 7.6), discussing the reasons behind better or worse predictions.

Prior to the mentioned structure, phenomena associated with isotonic contractions related muscle physiology is introduced in the following section, containing previous research and related work.

7.2 Isotonic contraction

The previous results have focused on how torque changes in relation to either a change in length of the muscle tendon unit or to voluntary activation but with no change in muscle tendon unit length. During the majority of movement tasks however torque is generated in the face of both active contraction and muscle tendon unit length changes. Therefore for wider application of ultrasound data to be explored it is important to evaluate such conditions and here these are represented with isotonic contractions.

Muscle tendon unit length change can be either concentric (shortening) or eccentric (lengthening). Eccentric activity is associated with injury risks (Lieber and Friden, 2000), associated with excessive strains leading to disruption of the sarcomere structure. In addition, muscle has been shown to respond differently to eccentric and concentric training interventions, with eccentric training (e.g. downhill running) leading to an increase in the number of sarcomeres in series (Lynn et al., 1994). This adaptation is thought to enable the muscle to continue operating on the desirable portion of the force-length curve (Lieber and Friden, 2000).

Differences in the mechanical demands of eccentric versus concentric contractions also occur when considering the force-velocity relationship of muscles, with more force producing capability occurring during eccentric conditions (Hill, 1938). This suggests benefits in terms of economy of locomotion for muscle fibres to operate eccentrically, as fewer muscle fibres would need to be recruited to generate the same force output if the muscle operated in the eccentric portion of the force-velocity relationship (Fukunaga et al. 2001; Roberts et al. 2016). Despite the eccentric portion of the force-velocity relationship having been well published in the literature, there is however little understanding of the factors that influence force output under eccentric conditions (Tomalka et al. 2017). This partly reflects challenges in studying mechanical muscle properties using *in situ* preparations (e.g. damage to samples). To be able to study neuromechanical aspects of eccentric (and concentric) contractions *in vivo* would therefore provide a valuable tool for researchers. As such analysis of ultrasound data for prediction of torque during concentric and eccentric contractions could have great value for improving understanding of *in vivo* skeletal muscle mechanical properties.

7.3 Torque Predictions using Standard Clinical Ultrasound

This section reports the performance of prediction algorithms estimating the ankle torque during isotonic contractions using Standard Clinical US, which transmits focused waves at a frame rate of 32 fps. Isotonic trials were acquired at two

different submaximal contraction levels (30% and 60% MVC), and a cross-validation process was applied to the resulting data. RMSE values were used as the goodness of fit indicator for the difference between the empirically acquired joint torque using a dynamometer chair, with predictions using Linear, Quadratic and Cubic Regressions, Kernel SVR and ANN techniques.

7.3.1 Isotonic Contractions at 30% MVC

This section reports differences between the use of 8-bit and 16-bit US data during isotonic contractions at 30% MVC. Table 7.1 compiles the data regarding 8-bit US data compression including overall mean and deviation values together with individual participant performance for each inference algorithm studied.

Table 7-1. RMSE results for the five inference methods used for Joint Torque predictions using 8-bit precision US data during isotonic contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	7.705	6.787	6.905	6.693	6.215
Participant 2	7.313	6.244	6.215	6.499	4.910
Participant 3	7.242	6.400	6.372	7.340	5.584
Participant 4	12.175	9.763	9.943	12.378	8.705
Participant 5	10.010	7.254	7.583	8.389	4.152
Participant 6	7.268	7.080	7.046	7.237	6.242
Participant 7	10.571	9.532	9.668	10.533	7.401
Average	8.898	7.580	7.676	8.439	6.173
Stand. Dev.	2.005	1.457	1.525	2.213	1.525

From the data it can be observed a similar behaviour in all the prediction algorithms except for the case of ANN. ANN provided similar deviation values but showed an important reduction in the RMSE values. This overall best results (6.17 ± 1.53 N m) corresponded to 11.92% of the maximum torque during 30% MVC isotonic trials (51.8 N m).

The individual and group RMSE results using 16-bit precision data are shown in Table 7.2. The same trends as seen in the 8-bit data can be appreciated, with ANN being the best prediction method in terms of both average and deviation with a

result (6.283 ± 0.997 N m) equivalent to 12.13% of the maximum joint torque value for this condition.

Table 7-2. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit precision US data during isotonic contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	8.105	7.292	7.336	7.068	6.531
Participant 2	7.694	6.778	6.818	6.688	5.384
Participant 3	7.238	6.749	6.817	7.454	5.542
Participant 4	12.525	11.168	11.090	14.190	7.430
Participant 5	10.110	7.503	7.376	8.773	5.259
Participant 6	7.128	7.053	7.060	7.332	6.091
Participant 7	10.362	9.541	9.680	10.031	7.748
Average	9.023	8.012	8.025	8.791	6.283
Stand. Dev.	2.026	1.689	1.677	2.646	0.997

In order to allow a direct comparison of data precision in the prediction error during 30% MVC isotonic contraction, the information contained in tables 7.1 and 7.2 is shown (Figure 7.1), using the same Y axis range in order to aid the direct comparison.

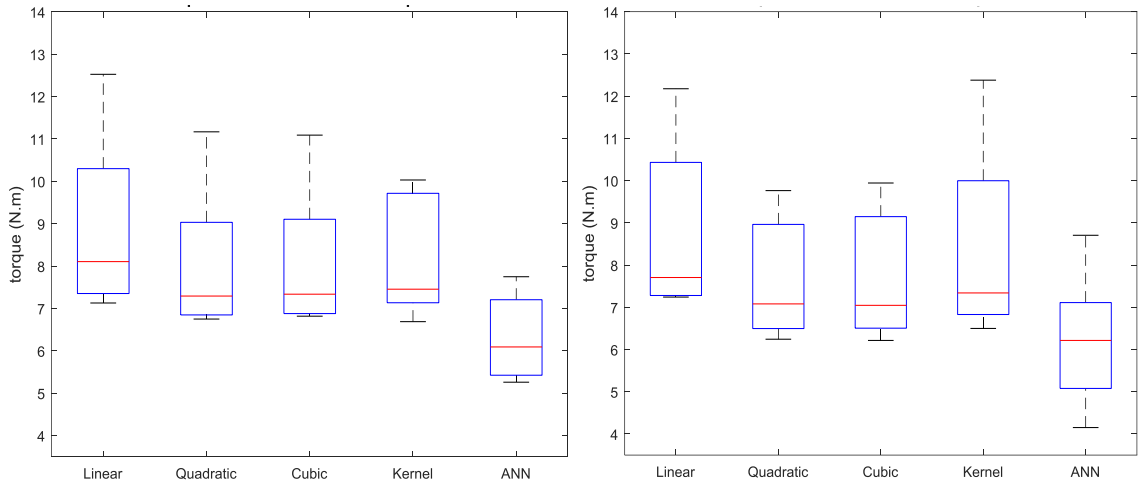

Figure 7-1. Boxplot comparison of the RMSE values distribution for the 5 Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of isotonic measurements at 30% MVC from 9 participants.

Similar trends are observed at both precisions, being linear regression the worse prediction values in terms of variance meanwhile ANN showed an important improvement in terms of average. Quadratic and Cubic Regressions together with

Kernel SVR showed similar intermediate performances when predicting isotonic contractions at 30% MVC collected at 32 fps.

7.3.2 Isotonic Contractions at 60% MVC

Table 7.3 shows the individual and group data resulting from analysing 8-bit precision US data during isotonic contractions at a submaximal level of 60% MVC.

Table 7-3. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	17.467	15.161	15.020	14.306	16.242
Participant 2	12.981	12.648	12.754	12.286	11.011
Participant 3	13.662	11.110	11.308	10.046	12.720
Participant 4	22.648	20.727	19.601	23.986	14.422
Participant 5	17.302	14.889	14.476	16.679	13.084
Participant 6	10.134	10.582	9.639	9.711	9.323
Participant 7	20.668	18.104	16.363	20.304	14.614
Average	16.409	14.746	14.166	15.331	13.059
Stand. Dev.	4.428	3.701	3.313	5.343	2.333

According to the results a similar performance can be observed for the different prediction algorithms in terms of RMSE, appreciating a better fit for the case of ANN especially in the average. This mean error (13.06 ± 2.33 N m) corresponded to 15.38% of the maximum joint torque for isotonic 60% MVC trials (84.88 N m). In contrast, Table 7.4 presents results during the same trials but using 16-bit precision US data collected at a frame rate of 32 fps.

Table 7-4. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	17.821	17.107	16.758	14.784	12.045
Participant 2	13.357	12.731	12.840	12.762	11.349
Participant 3	12.839	11.195	11.044	10.288	12.165
Participant 4	21.882	23.654	24.812	24.977	12.230
Participant 5	17.389	15.843	15.837	18.449	12.661
Participant 6	9.984	10.888	10.248	9.293	9.116
Participant 7	19.610	17.729	16.646	19.086	12.631
Average	16.126	15.593	15.455	15.663	11.742
Stand. Dev.	4.202	4.501	4.903	5.551	1.238

Prediction techniques showed similar trends in both terms of average and deviation in all cases but ANN that showed a reduction in the measured error in average and an almost three times smaller deviation. The best fit result was equivalent to 13.83% the maximum joint torque for this conditions.

A direct comparison between the two data precisions was enabled by showing graphically the prediction results contained in tables 7.3 and 7.4 (Figure 7.2). No differences between precisions can be observed in all different techniques except ANN, that showed an important improvement in both median and variance when using 16-bit US data.

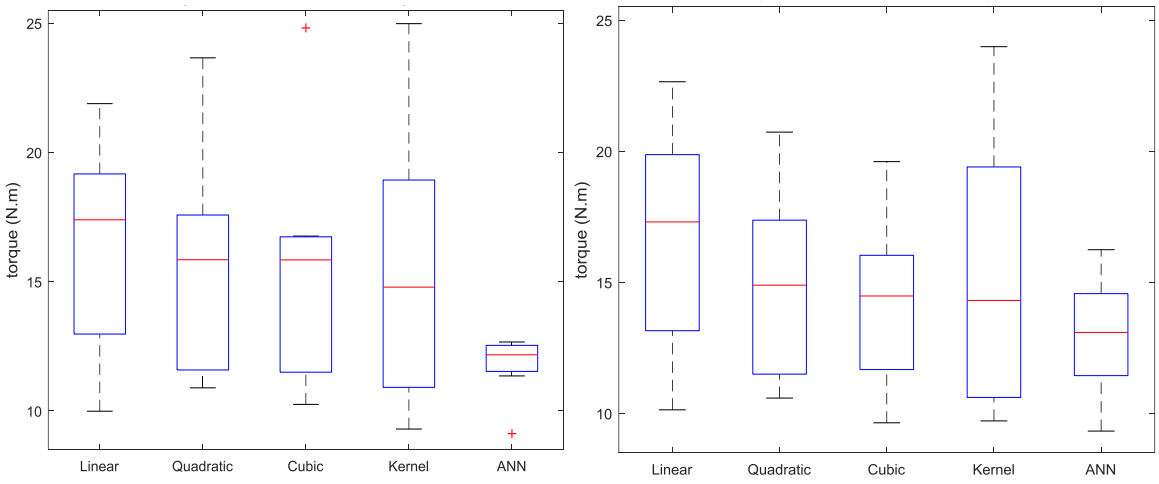

Figure 7-2. Boxplot comparison of the RMSE values distribution for the 5 Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of isotonic measurements at 60% MVC from 9 participants.

7.4 Torque Predictions using Ultra-Fast frame rate Ultrasound

This section shows joint torque prediction results using plane wave transmission at a frame rate of 1000 fps. These conditions led to an overall of 8000 frames per trial (8 second duration each), that exponentially increased the required resources for data processing. Due to these limitations, Kernel SVR results were calculated using only 1000 data points per trial (applying one-dimensional interpolation).

7.4.1 Isotonic Contractions at 30% MVC

The present section reports the analysis of 8-bit and 16-bit precision US data during isotonic contractions at 30% MVC. Table 7.5 shows the individual participant and mean group predictions using 8-bit data precision at a frame rate of 1000 fps.

Table 7-5. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	4.052	4.071	4.300	3.550	3.821
Participant 2	7.061	5.830	6.028	6.103	4.199
Participant 3	5.659	5.293	5.175	5.265	6.900
Participant 4	9.714	6.975	7.023	7.717	6.981
Participant 5	8.048	5.287	5.511	5.758	4.820
Participant 6	4.871	3.510	3.431	3.869	2.990
Participant 7	10.052	10.574	10.487	10.093	9.185
Average	7.065	5.934	5.994	6.051	5.557
Stand. Dev.	2.337	2.338	2.294	2.267	2.198

Results show a slight worse fit of the empirically collected joint torque data for linear regression and a similar response for quadratic and cubic regressions and Kernel SVM. The best prediction was observed when applying ANN with an average (5.56 ± 2.20 N m) equivalent to 10.73% of the maximum joint torque for 30% MVC isotonic contractions (51.8 N m).

Opposite to these, results using 16-bit precision US data are compiled in Table 7.6, showing the RMSE values per participant together with the overall mean and deviation. Again, ANN showed the best prediction of all methods, that also showed similar trends consisting of a worst response for linear regression and a similar fit for kernel SVR and quadratic and cubic regressions. The best fit (5.07 ± 1.19) corresponded to 9.79% of the maximum acquired torque for these trials.

Table 7-6. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	4.592	4.934	6.721	3.417	4.073
Participant 2	6.880	5.300	5.327	5.560	5.074
Participant 3	6.356	5.490	5.509	5.650	4.815
Participant 4	9.889	7.038	7.030	7.722	6.692
Participant 5	9.024	6.194	6.248	6.363	5.481
Participant 6	5.093	3.492	3.489	3.860	3.201
Participant 7	9.863	8.855	8.997	8.868	6.161
Average	7.385	5.900	6.189	5.920	5.071
Stand. Dev.	2.216	1.701	1.701	1.952	1.193

A direct comparison was established for the two data precisions (8-bit and 16-bit) by grouping the RMSE values from tables 7.5 and 7.6 (Figure 7.3). Equal Y axis range representation was used to aid the precision comparison, which showed different trends depending on the prediction algorithm. No differences can be observed for the two precisions in the performance of Kernel SVR and linear regression, which provided the worse fit to the collected joint torque. Quadratic and Cubic regressions and ANN (best fit) showed similar prediction medians, with an important reduction on the variance when using the highest precision (16-bit).

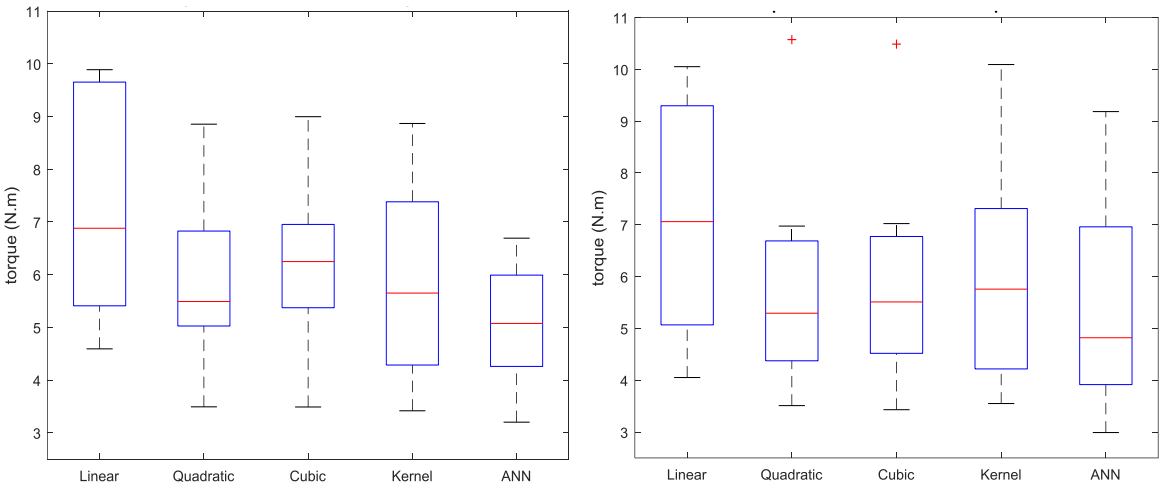

Figure 7-3. Boxplot comparison of the RMSE values distribution for the 5 Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of isotonic measurements at 30% MVC from 9 participants collected at 1kS/s.

7.4.2 Isotonic Contractions at 60% MVC

Following section reports the results from the analysis of 8-bit and 16-bit data during isotonic contractions at 60% MVC. Results for the predictions at 1000 fps for isotonic using 8-bit data precision were compiled in Table 7.7, showing both individual and mean group values at 1000 fps frame rate.

Table 7-7. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	8.415	9.739	8.240	8.975	7.822
Participant 2	11.670	10.625	10.001	10.841	9.346
Participant 3	6.697	6.749	5.200	6.350	5.697
Participant 4	17.274	17.730	14.123	17.472	14.329
Participant 5	14.534	12.476	12.466	12.743	11.212
Participant 6	6.227	5.800	4.378	5.948	4.348
Participant 7	9.314	9.238	7.606	9.012	8.499
Average	10.590	10.337	8.859	10.192	8.751
Stand. Dev.	4.122	3.965	3.593	3.991	3.346

Results show similar predictions for Linear and Quadratic regressions and Kernel SVR. On the other hand, Cubic regression and ANN provided better fits to the joint torque collected with the dynamometer, showing similar errors in terms of mean

and deviation. The best fit observed (8.75 ± 3.35 N m) corresponded to 10.31% of the maximum torque observed for isotonic contractions at 60% MVC (84.88 N m).

In contrast, Table 7.8 presents the individual participant and mean and deviation group data resulting from the analysis of 16-bit US data during isotonic contractions at 60% MVC collected at 1000 fps.

Table 7-8. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	7.797	7.955	6.847	7.341	5.545
Participant 2	11.418	8.729	8.677	9.442	7.381
Participant 3	7.096	6.742	5.842	6.164	6.159
Participant 4	15.405	11.281	10.399	12.536	10.727
Participant 5	16.149	12.356	12.754	12.982	12.163
Participant 6	7.884	8.543	7.123	8.001	5.946
Participant 7	11.621	8.192	6.848	8.889	6.001
Average	11.053	9.114	8.356	9.336	7.703
Stand. Dev.	3.689	1.979	2.450	2.568	2.651

Linear regression provided the worst predictions for these conditions, followed by Quadratic regression and Kernel SVR. The best results were observed with the use of ANN in terms of average, followed by Cubic regression. The best fit (7.70 ± 2.65) was equivalent to 9.07% of the maximum torque registered during these isotonic trials.

To enable a direct comparison between 8-bit and 16-bit data precisions in the predictions of joint torque data at 1000 fps, results obtained are graphically shown (Figure 7.4). In order to provide proportionate information, the same joint torque value range (Y axis) was used for both plots. Again, linear regression showed the worst results for both precisions in terms of deviation, and also in the median value for the case of 16-bit. All inference methods showed a similar performance when using 8-bit US data. However, different behaviours were observed for the case of 16-bit data precision, with an important improvement when using ANN, which showed the best prediction of all studied in this chapter regarding isotonic movements.

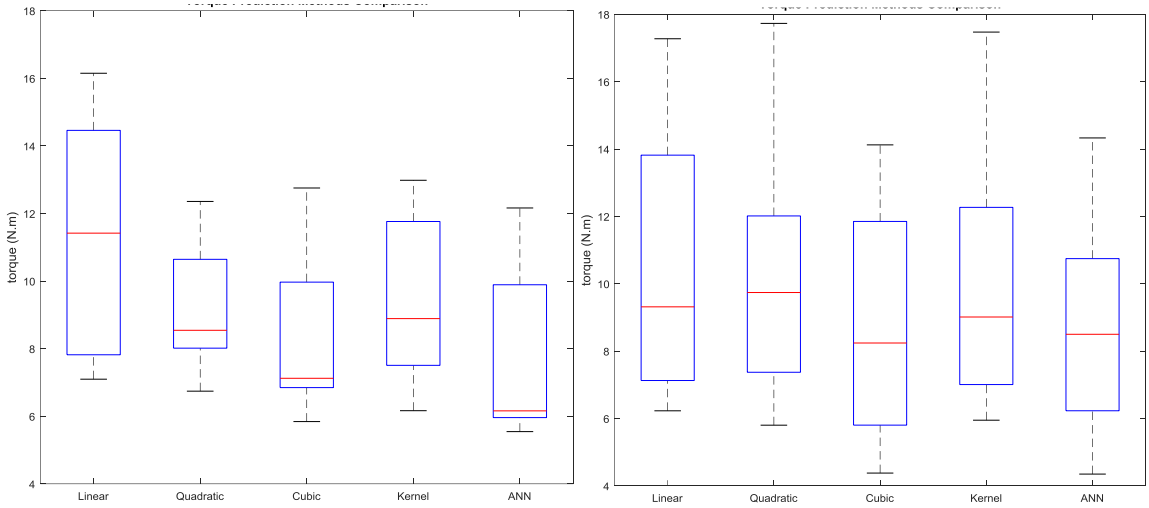

Figure 7-4. Boxplot comparison of the RMSE values distribution for the 5 Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of isotonic measurements at 60% MVC from 9 participants collected at 1kS/s.

7.5 Features Underlying Algorithm Performance

In order to provide a better understanding of the reasons behind the worst fits of the predicted joint torque, Figure 7.5 shows the torque predictions during two isotonic trials of the different prediction methods (polynomial regression, Kernel Methods and ANN) during a bad fit of the empirically collected torque.

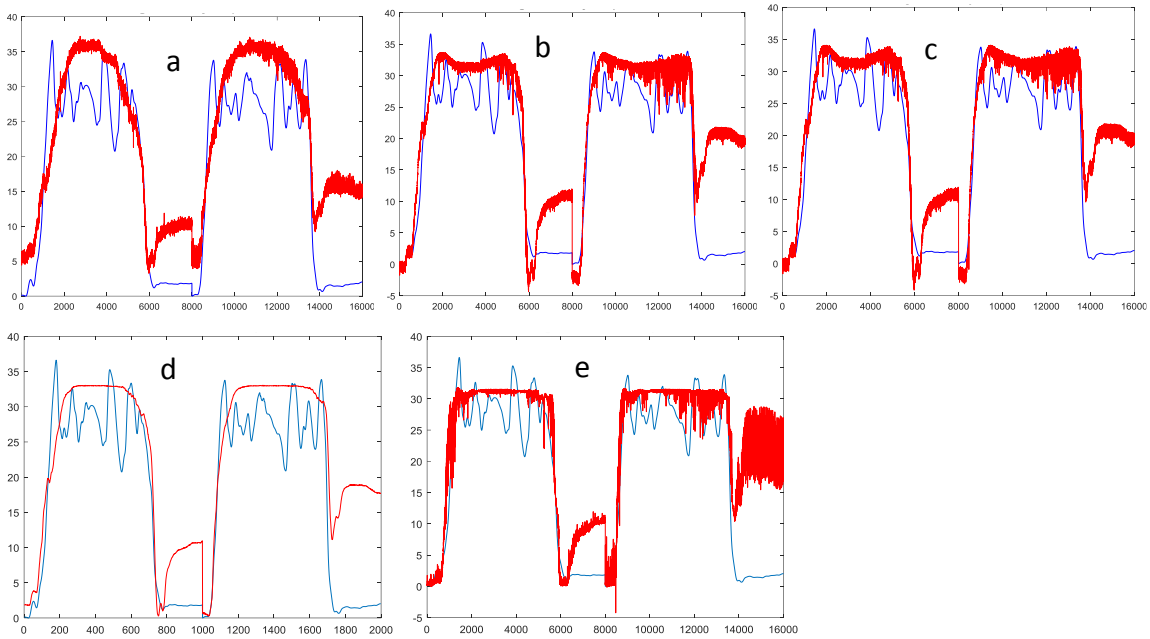

Figure 7-5. Acquired Joint Torque (blue) vs. Predicted Torque (red) graphical results in N m for the isotonic 30% MVC trials (2 trials, 8000 frames each) for 5 prediction methods: Linear (a), Quadratic (b), Cubic (c), Kernel SVM (d) and ANN (e). Resulting RMSE Torque values were 2.31, 4.54, 6.01, 1.97 and 4.34 N m respectively.

From the graphs, very similar trends can be observed for the different prediction algorithms, consisting of a lack of contour fit for the small activations related with muscle firing control during joint angle movements; and an incorrect offset value during muscle relaxation states. Kernel SVR showed a smoother response compared with the other algorithms that could probably be related with the use of eight times less data points due to computer limitations.

In order to provide insights regarding joint torque levels associated with bad prediction performances, a comparison of the torque values associated with errors in predictions bigger than 5 N m are provided for isotonic contractions at 30% MVC (Figure 7.6) and 60% MVC (Figure 7.7) submaximal activation levels.

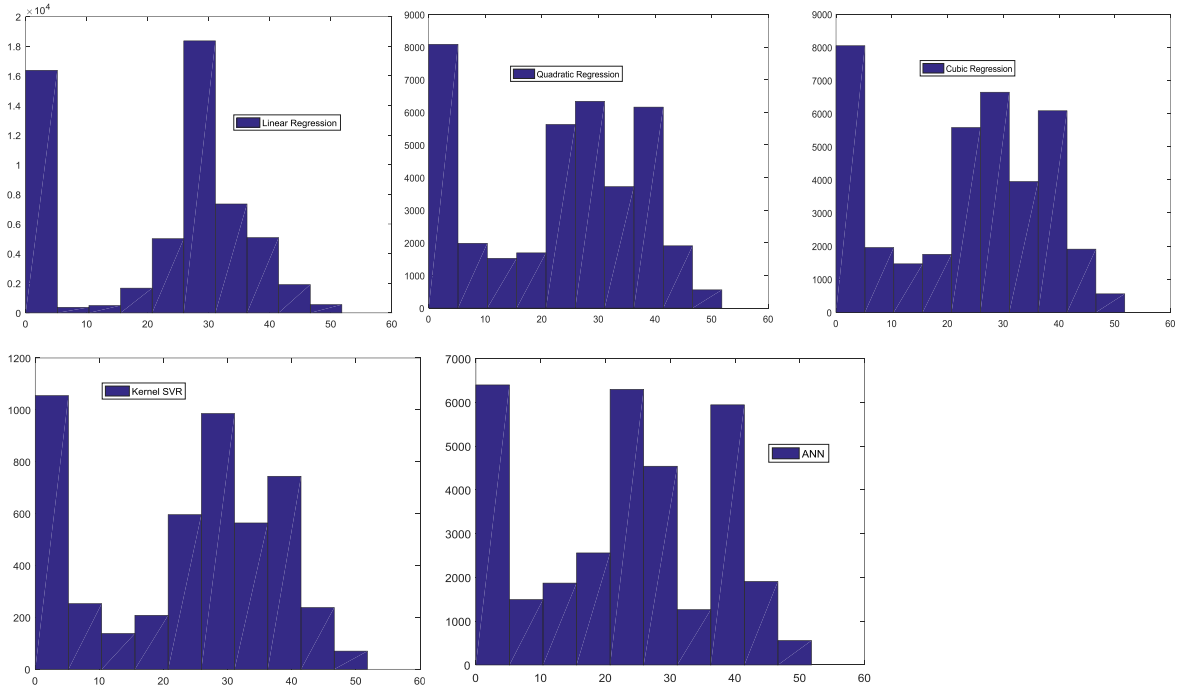

Figure 7-6. Histogram of the Torque values associated with prediction errors bigger than 5 N m during 30% MVC isotonic trials for Linear, Quadratic, Cubic, Kernel and ANN methods. X axis Torque values in N m. Due to computational limitations values for Kernel SVR were collected from 1000 samples per trial only (8000 per trial for the other methods).

Due to the nature of isotonic trials, the greatest part of the torque values were concentrated in the submaximal activation levels, corresponding to 30 N m in this case. As opposed to what was observed in Section 6.5 (Figure 6.6), the contribution of 30 N m torque is similar to the neighbouring values (except for linear regression case). This is due to the fluctuations during activation phases observed in Figure 7.5, proving that the prediction algorithms were not able to provide accurate fits of the torque in these situations.

Figure 7.7 provides insight regarding higher submaximal activations (60% MVC) during joint rotations in order to check errors associated with faster muscle displacements.

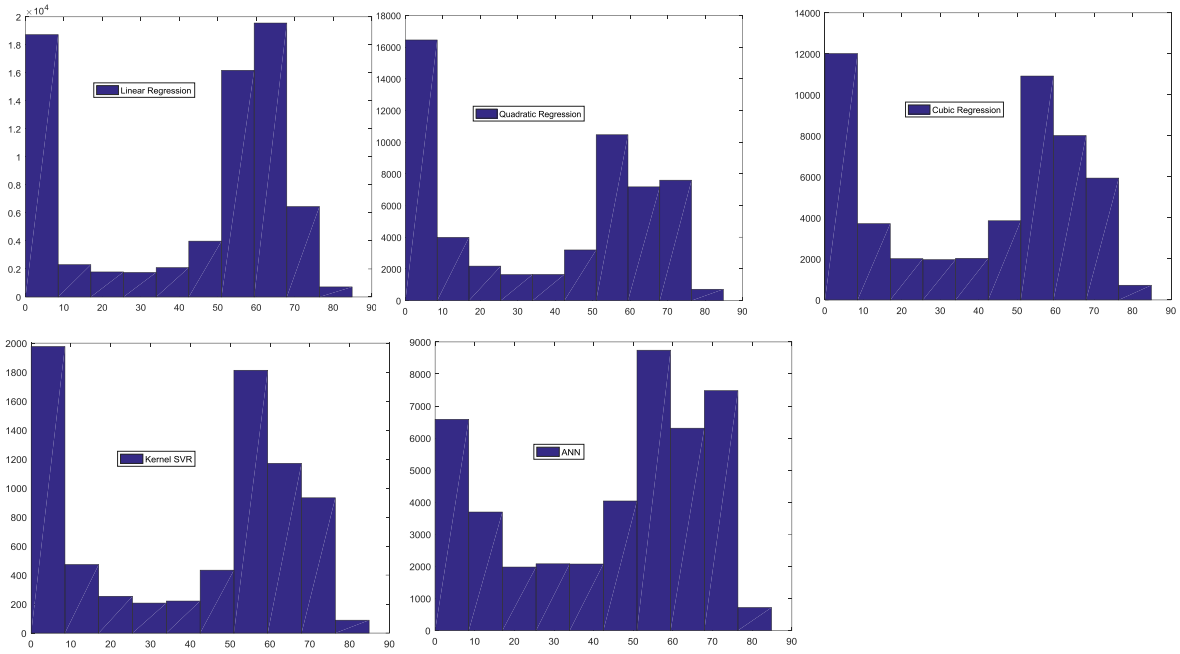

Figure 7-7 Histogram of the Torque values associated with prediction errors bigger than 5 N m during 60% MVC isotonic trials for Linear, Quadratic, Cubic, Kernel and ANN methods. X axis Torque values in N m. Due to computational limitations values for Kernel SVR were collected from 1000 samples per trial only (8000 per trial for the other methods).

Similar trends were observed compared with 30% MVC trials for all algorithms but ANN, which presented a well spread error distribution as the contribution of errors at the lowest torque were important but always smaller than the ones associated with higher torque values.

Compared to the equivalent results during isometric tasks showed in Section 6.5, important differences can be observed at the highest torque values. Due to the smoother plateau described during isometric contractions, the 60 N m contribution was much bigger than neighbouring values. In contrast, errors during isotonic trials were more spread at high torques as a result of the bigger fluctuations observed during muscle activations.

7.6 Statistical Analysis

Non-parametric ANOVAs were applied to the data contained in tables 7.1 to 7.8 in order to determine whether the RMSE values resulting from the different studied conditions were significantly different. In order to achieve this, Kruskal-Wallis was

applied to different arrangements of the obtained RMSE values differencing between inference algorithm applied, data precision, frame rate and muscular activation level, as explained previously in chapters 5 and 6.

In order to account for statistical differences in the predictions of joint torque during isotonic movements, Kruskal-Wallis tests were performed to the RMSE values shown in Tables 7-1 to 7-8, that at the same time were rearranged in order to compare the different values of data precision, frame rate, muscular activation and inference algorithm used.

7.6.1 Statistical differences in performance of different prediction algorithms

Firstly, a five-column matrix was built with all the RMSE torque values obtained using the five different inference algorithms for predicting joint torque from US data. Kruskal-Wallis non parametric ANOVA was applied to the resulting matrix obtaining the statistical significance between inference methods shown in Figure 7.9.

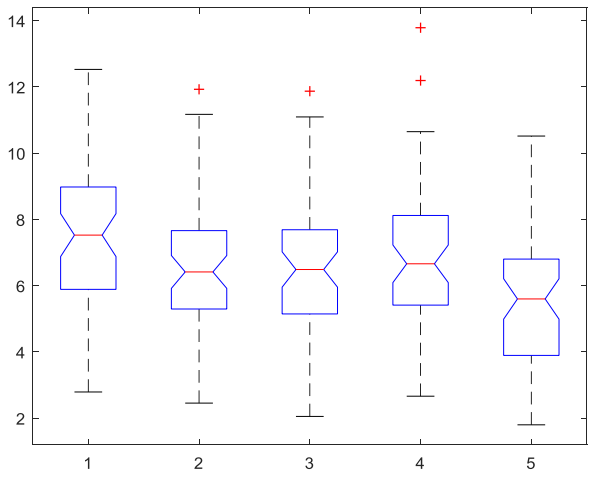

Figure 7-8. Kruskal-Wallis test results for the five prediction algorithms results during Isotonic contractions.

According to the resulting values of p = 0.0011 and , $\chi 2$ = 18.21, the hypothesis that the data from the different inference methods come from the same distribution can be rejected, meaning that the differences observed in the RMSE values between different prediction algorithms are statistically significant for isotonic contractions.

7.6.2 Data precision, frame rate and activation level statistical influence

In a similar fashion to the previous section, RMSE values obtained during isotonic movements were rearranged into two-column matrices in terms of data precision, frame rate, a combination of the two, and activation level. This process was performed using the RMSE values for every inference method studied, and the resulting significance p-values were compiled in Table 7.9.

Error values at 60% MVC were halved in order to provide comparisons with equivalent magnitudes to 30% MVC without affecting the outcome of the statistical analyses.

Table 7-9. Kruskal-Wallis p-values obtained per inference method using RMSE results comparing data precision, frame rate, a combination of the two and activation levels during isotonic contractions.

Method	8bit vs 16bit	32fps vs 1kfps	32fps&8bit vs 1kfps&16bit	30 vs 60 MVC
Linear	0.8058	0.0012	0.0432	0.0715
Quadratic	0.6464	3.15E-05	0.0101	0.4033
Cubic	0.4913	2.36E-05	0.0169	0.1403
SVR RBF	0.6941	1.04E-05	0.0077	0.1447
ANN	0.7743	0.0003	0.0131	0.2687

Analogously to section 6.6, the critical value after applying Bonferroni correction is p = 0.0125. As a result, no significance was obtained for either activation level or data precision, meaning that any differences appreciated in previous sections regarding data precision and activation level during isotonic movements could not

be observed if performing the same experiment again maintaining the same conditions

On the other hand, when comparing frame rate isolated for all inference methods, or together with data precision for the quadratic regression and the Support Vector Regression, the p-values obtained are lower than 0.0125. This means that the null hypothesis can be rejected and confirms that the differences observed in RMSE values in sections 7.3 to 7.5 should be similar if the same experiment is performed again for the mentioned cases.

7.7 Discussion

The work described in this chapter addressed the validation of different inference methods able to predict ankle net joint torques from US data of the GM during isotonic contractions that aimed for a constant output torque of 30% and 60% of the MVC while rotating the ankle within a 30 degree movement range. Important focus was applied to studying the impact of different data precisions and temporal resolutions in the prediction accuracy, together with the analysis of which inference algorithms provided better fits to the empirically collected joint torque.

Regarding data precision, the use of 16-bit data showed an important reduction of the variance and the average values for the fastest frame rate (1000 fps). However, no significant differences were observed at the two different precisions during isotonic trials collected at 32 fps. These findings can be related to the fact that isotonic contractions at 60% MVC were the fastest muscle contractions registered in this thesis and therefore the ones that should benefit more from the higher temporal resolution.
Therefore, temporal resolution proved to be an important factor in the performance of prediction algorithms during isotonic contractions as it implied a better fit to the collected torque for both 8-bit and 16-bit data and at the two different MVC levels studied.

The best inference algorithm for these trials was ANN, as it outperformed the other methodologies when applied to US data with different precisions and temporal resolutions (Figure 7.2). This conclusion could be due to the increased complexity of the isotonic trials, for which only ANN provided consistent predictions of the joint torque. However, it was clear that none of prediction techniques were able to follow the fluctuations in torque which occurred in this condition (Figure 7.5). This could be the result of a number of influencing factors. For example, the predictions are based on US data collected only from GM and this muscle may not provide representative behaviour to torque generated in the isotonic movements as it seems to for isometric contractions (Figure 6.5). There may be important additional contributions from the GL muscle (as shown when checking the EMG activation levels) and the elastic properties of the Achilles tendon. Secondly, the experimental data collected may not have provided enough examples of torque fluctuations for appropriate predictions, with participant compliance in achieving the task adding additional variability to the data set than seen in isometric or passive trials. For this particular case the use of more complex ANNs could provide better fits to the data, as deep neural networks are able to learn from more complicated input-output relationships. An important drawback from the use of these type of solutions is the exponential increase in the implementation time that could imply weeks of continuous processing for the case of 8000 data points per trial, one of the reasons this approach was not attempted within the context of this thesis.

It is interesting to note that all the predicted torques seemed to fit the larger contraction elements better than the periods in the trial associated with the relaxation phase between the transitions from eccentric to concentric (Figure 7.6). This trend was also observed in Chapter 6 when studying isometric contractions and could be related with not having enough quiet data represented in the data set, which could lead to the algorithms over-interpreting small changes in muscle shape and predicting a larger torque than required.

In summary, isotonic trials proved to be the most challenging muscle movement to predict with the proposed methodology. Besides providing good fits to the

empirically collected net joint torque, using ANN, 16-bit US data collected at 1000 fps; important errors were observed during fluctuations of torque at submaximal MVC levels and relaxation phases. These disagreements between predictions and empirically collected data could be related with how representative GM is of these muscle tasks, or limitations of the prediction algorithms of relating such complicated input-output relationship, for which complementary methodology such as the use of deep learning algorithms should be taken into account.

The work presented in this chapter considers the final movement condition included in the experimental protocol. From this and the previous two chapters (Chapters 5 and 6), it is clear that the effects of US data properties and the prediction algorithms perform to different degrees across the movement tasks studied. The next chapter (Chapter 8) will therefore consider the key research questions defined in the context of the full set of results, comparing performance and highlighting limitations of the work and proposals for future avenues of investigation.

8 Discussion

8.1 Summary and Contributions

This thesis focused on the validation of methodology to accurately predict net ankle joint torque exclusively using US data. The methodology has enabled consideration of experimental factors identified as not being fully understood from the literature review (Chapter 2). These included the study of the influence of both precision and temporal resolution of experimentally collected US data, the measurement of changes in muscle properties using automatic analysis techniques, and the choice of different algorithms to make predictions based on learning from input-output relationships.

In order to collect appropriate experimental data to address the research questions, a fully synchronised setup was successfully implemented collecting dynamometer, EMG and US data from participants performing passive and active muscle tasks. This setup also enabled the study of temporal inconsistencies using US devices that materialized in a peer-reviewed journal contribution (Miguez et al. 2017).

Three muscle movements were studied in this research. These tasks added different challenges to the processes involved in the prediction of joint torque using US data. They consisted of passive ankle joint rotations (Chapter 5), which provided insight for relaxed and non-activated muscles; isometric contractions (Chapter 6), that produced the condition of muscle activations without changes in the muscle-tendon unit length; and isotonic contractions (Chapter 7), which implied a constant overall muscle activation during ankle rotations. The results obtained validated my methodology as a reliable predictor of joint torque using US derived data, obtaining average errors as small as 7.62%, 7.77% and 9.07% of the maximum torque respectively for passive, isometric and isotonic tasks.

8.2 Influence of data precision on joint torque predictions

The first key research question proposed to be addressed within this work was:

- Is there a difference in the prediction error of ankle joint torque generated during passive movements and active contractions using image analysis of data collected with different precision?

One of the distinctive characteristics of this research was the capability of directly comparing US data that only differed in the dynamic range of possible values (bit precision) and shared the exact same post-processing, meaning that no additional compressions or "black-boxed" processes were applied. Literature in Section 2.2 documented previous attempts to provide direct comparisons of the influence of data precision in US studies that were unsuccessful due to a lack of synchronisation of the experimental setup (Wrench & Scobbie 2008; Payen et al. 2012; Payen et al. 2013), or a lack of direct comparison between precision of collected data (Rognin et al. 2008)

Data precision played an important role in the predictive power of the implemented algorithms, and was closely related with two key factors: the appearance of the resulting US images, and the velocity of the muscle movements studied. The good lateral resolution obtained when using focused waves for the US transmission (Figure 2.9) provided well-defined muscle fascicle shapes in the resulting US data. For these US transmission conditions, the use of bigger data precision did not imply an improvement on the joint torque predictions (Figures 5.1-2, Figures 6.1-2, and Figures 7.1-2).

Contrary to this, grainy looking US images are the result of the use of plane waves, an approach which is required in order to achieve fast frame rates. Despite not having an impact during passive trials, data precision provided an improvement in the prediction of joint torque during isometric and isotonic contractions, particularly important for the biggest activation levels (60% MVC). This finding provides a link with the second key factor related with the influence of data precision in joint torque

predictions. Due to the fact that 60% MVC submaximal contraction levels were reached over the same time period as 30% MVC, the associated muscle displacement velocity was also higher for 60% MVC. The fact that the use of 16-bit precision US data provided better predictions during 60% MVC contraction levels compared to 30% MVC (left Figures 6.3 vs 6.4, and 7.3 vs 7.4) for both isometric and isotonic tasks proved the value of higher data precision at faster muscle fascicle velocities. In conclusion, data precision greatly influences predictions of joint torque from US data when using plane waves as transmission and at faster frame rates. Therefore, data precision should be an important consideration for any one establishing new experimental protocols for the assessment of skeletal muscle and torque generation, but only if the resulting US images lack well-defined shapes, and when greater activation levels or faster movements are to be studied.

8.3 Influence of temporal resolution on joint torque predictions

The second key research question proposed to be addressed within this work was:

- Is there a difference in the prediction error of ankle joint torque generated at passive movements or active contractions using image analysis of data collected with different precision?

While the literature has shown how the use of fast frame rates enabled new studies of skeletal muscle *in vivo* behaviour (Deffieux et al. 2008; Nordez et al. 2009) and enhanced existing technologies (Tanter & Fink 2014), to my knowledge no ultrasound studies have included direct comparisons between conventional and fast frame rates and evaluated the influence of this parameter on subsequent assessment of muscle properties. The data analysed in this thesis was therefore collected at two different frame rates, 32 and 1000 fps, in order to provide examples of temporal resolutions similar to conventional US imaging devices and also explore the capabilities of high-end machines.

My work, presented in (Miguez et al. 2017) and detailed in Chapter 3, provided the certainty of a fully synchronised data acquisition with an accuracy below 40 μs (DAQ sample rate of 25 ksps). This is a requirement of the experimental protocol that is essential in order to be able to conclude that all IFI variabilities had been accounted for and therefore did not affect any of the results or the conclusions subsequently extracted from them.

The better predictions observed during higher submaximal activation levels associated with faster muscle contractions were not only related with the bigger precision of the US data discussed previously (Section 8.2, also Figure 6.4-left vs Figure 6.4-right), as frame rate was seen to also play an important role (Figure 7.2-right vs Figure 7.4-right). Under ideal conditions, it may be predicted that a muscle movement collected at a certain frame rate would give the exact same output as a movement undergone at twice the velocity and recorded at twice the original frame rate. The capability of optical flow algorithms for accurately extracting motion information from US images is closely related with the motion shift between consecutive frames, therefore a so called "optimal dynamic range" would be defined as a muscle shift big enough to be detected by the algorithms, and small enough to not saturate the maximum allowed shift between two images.

The results presented in Chapter 5 showed that when data were recorded at 1000 fps the algorithms were not able to predict the smooth movements that occurred during passive trials at rotational velocities of four degrees per second. This suggests the motion shift between frames may not have been large enough to be detected by the algorithms used. Furthermore, important improvements were observed when using the highest temporal resolution during the 60% MVC submaximal levels compared with 30% MVC, that were all achieved in less than one second during both isometric and isotonic trials. Overlaps in the distribution of errors obtained for both frame rates from prediction algorithms were observed for the case of 30% MVC trials, suggesting that the muscle shifts to achieve 30% MVC over one second were not fast enough to provide useful information at 1000 fps and at the same time too fast for 32 fps frame rate.

In conclusion, muscle movements associated with the achievement of 30% MVC submaximal ankle torque levels over the period of one second appears to mark the threshold for the beneficial use of frame rates of 1000 fps for predicting ankle joint torque using US data. 60% MVC levels achieved within a second using the highest temporal resolution provided a reduction of the mean error of up to 2.4 times (7.77% vs 18.67% RMSE during isometric contractions) compared with the same results at 30% MVC. There is therefore an interplay between muscle movement velocity, torque generation and the optimal US frame rate for data acquisition, which should be an important consideration for researchers wishing to benefit from application of more advanced ultrasound devices.

8.4 Influence of prediction algorithm on joint torque predictions

The third key research question proposed to be addressed within this work was:

- Is error in the prediction of ankle joint torque generated during passive movements or active contractions affected by the prediction technique used?

In order to provide a good representation of data inference techniques and study their suitability for prediction of joint torque using US data, five different algorithms were implemented for this thesis namely Linear, Quadratic and Cubic Regressions, Kernel SVR and ANN; listed in order of complexity of input-output relationships they are able to learn from (Section 2.4).

The selected algorithms provide different ways of treating the data in order to predict output data. However, they all provided bad fits to empirically collected net ankle torque during passive movements (Figure 5.5) and also for isometric movements (Figure 6.5) , also lacking statistical significance (Figures 5.9 and 6.9). The results in Chapters 5 and 6 therefore concluded that the type of algorithm used for the prediction did not have an impact on the accuracy of joint torque predictions. In contrast, ANN proved to be the most reliable prediction technique for isotonic

trials (Chapter 7), outperforming in both average and variance Kernel SVR and the three regression algorithms (Figures 7.1 to 7.4).

As concluded in Section 5.6, the differences in torque prediction error between the three movement tasks studied could be related with the complexity of the contour described by the torque values. For the passive movements it consisted of a straight line (Figure 8.1); for isometric contractions it consisted of an increasing and decreasing signal including a plateau located at the submaximal contraction level (Figure 8.2); and for isotonic tasks the hill contained important and fast fluctuations imposed by both voluntary activation and changes in ankle joint angle (Figure 8.3).

In conclusion, ANN proved to be the most consistent prediction technique for net joint torque using US data across the conditions studied. This nonlinear technique is capable of learning from complex input-output relationships with the use of hidden neuron layers, of high computational requirements together with a non-repeatability of results due to its internal "black-boxed" processes. Therefore, if joint torque prediction is the main focus of future work, researchers should consider application of ANN based methods to provide optimal predictions from collected data.

8.5 Recommendations depending on muscle task.

The last key research question proposed to be addressed within this work was:

- Is there an important interaction between two or more of these three features (precision, temporal resolution and prediction algorithm)?

Two key relationships between the different conditions studied can be extrapolated from this thesis, the use of 16-bit precision data together with 1000 fps frame rate, and the

If taking into account all the findings derived from this research, recommendations regarding data precision, temporal resolution and prediction algorithms can be extracted depending on the muscle task to predict joint torque from.

For passive movements, the use of linear or quadratic regression together with 8-bit precision data and 32 fps frame rate (Figure 8.1). These options were selected as no significant differences were observed in the predictive power of any combinations, and these are the most resource efficient available options.

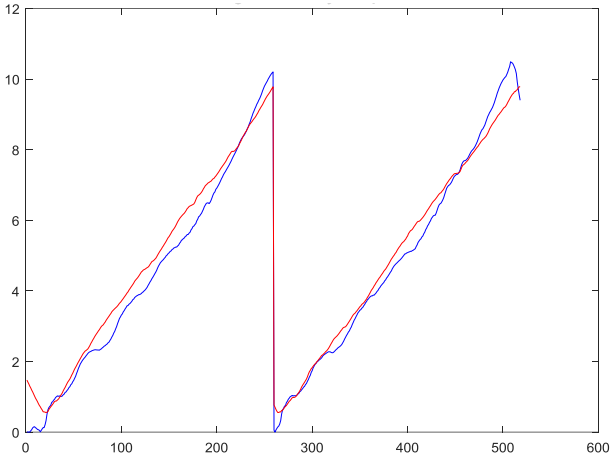

Figure 8-1 . Acquired Joint Torque (blue) vs. Predicted Torque (red) graphical results in N m for passive dorsiflexion trials using 8-bit data precision at 32 fps and Linear Regression. Resulting RMSE Torque value of 0.44 N m.

Linear regression provides good fits when the relationship between inputs and outputs is simple, suitable for smooth data with subtle changes, as graphically shown above. This best result for passive movements was able to fit the smooth response that was able fit the smooth changes observed during dorsiflexion. Whether this recommendation would hold for faster passive joint rotations cannot be identified from the work presented here and does warrant further investigation. However, as passive muscle properties are best studied at slower joint angular velocities (Hoang et al., 2007) the recommendation made here should hold true for investigations of passive tissue movements.

Regarding isometric contractions, the use of 16-bit precision data together with 1000 fps frame rate provides the least errors (Figure 8.2). For these contractions ANN is recommended as the prediction algorithm to use as it showed the most consistency during the different conditions.

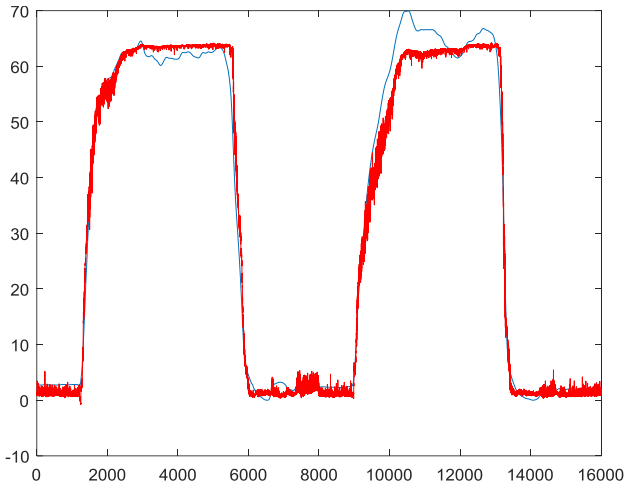

Figure 8-2. Acquired Joint Torque (blue) vs. Predicted Torque (red) graphical results in N m for isometric trials at 60% MVC using 16-bit data precision at 1000 fps ANN. Resulting RMSE Torque value of 3.08 N m.

ANN is capable of adjusting the weights of the different inputs in order to give a bigger significance to the ones that better explain the outputs. This way it is capable of avoiding the amplification of noisy inputs when the output is stable such as the case of stable activations and relaxation phases. However, this best prediction of isometric trials (Fig. 8.2) had a noisy response during relaxing phases and was also unable to predict accurately the bigger contraction registered at the beginning of the second voluntary contraction. Whether this could be improved upon with the use of a larger data set or more complex ANN warrants further research. The potential value of ANN is however clear from the work presented here.

Finally, the biggest precision and fastest frame rate are also recommended when predicting joint torque from US data recorded during isotonic contractions. In addition, ANN is also recommended for data analysis (Figure 8.3). This technique outperformed significantly the inference algorithms in both terms of mean values and variances for the different conditions analysed (Figure 7.3-left and Figure 7.4-left).

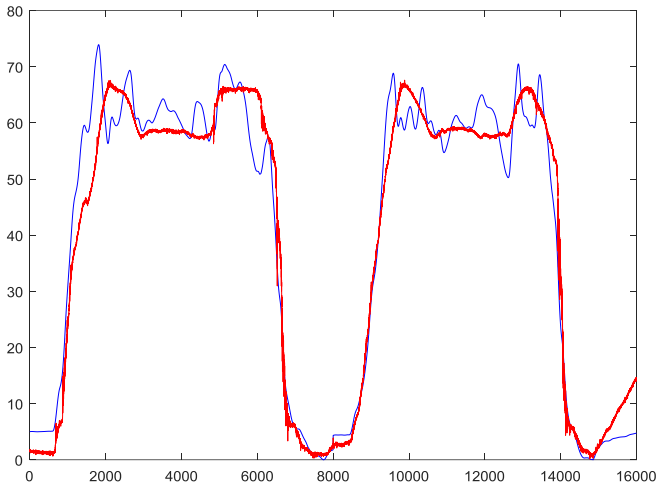

Figure 8-3 Acquired Joint Torque (blue) vs. Predicted Torque (red) graphical results in N m for isotonic trials at 60% MVC using 16-bit data precision at 1000 fps and ANN. Resulting RMSE Torque value of 4.9 N m.

As explained above regarding isometric contractions, ANN is also capable of adjusting the weights of the different inputs in order to match fluctuations of the torque such as the ones observed during the submaximal activations in isotonic trials. Despite being the best fit to the empirically collected torque during isotonic trials, the algorithm was not able to predict all the small activations part of the 60% MVC activation level plateau.

8.6 Limitations

The principle reasons behind errors in predictions of ankle joint torque during isotonic contractions were wrong estimations of the torque during relaxation states, particularly after strong contractions; and to accurately detect fluctuations in the torque signal during the period of activation. The inability of the inference algorithms to provide better predictions in these situations could be due to two main limitations: the small number of participants, which could not be representative enough of the entire population; and the use of information from only one muscle for studying tasks that involve other complementary muscles together with elastic components such as tendons.

8.7 Future Work

This last limitation could be tackled not only with the use of data from other muscles and/or the Achilles tendon, but also with the use of prediction algorithms capable of learning from more complex input-output relationships such as the use of deep learning algorithms. These techniques have been proven reliable and very accurate for the estimation of muscle fascicle curvatures using US, with errors as small as two degrees having been seen during dynamic standing tasks (Cunningham, P. J. Harding, et al. 2017). The application of these approaches for predictions of torques during more complex skeletal muscle behaviours or movement tasks is therefore an area of future work that may prove highly valuable in terms of both improving understanding of *in vivo* muscle behaviour and for provision of analysis tools for future investigation of muscle properties.

Regarding the US data post-processing, the inference algorithms were fed as inputs data derived from two main processes: ASM and OF. Different parameterisations were applied to the ASM properties, such as the number of iterations for finding the best fit, initial height of the two aponeuroses, and the distance between them (Section 3.2); in order to provide the most accurate location of the muscle boundaries. However, the ASM was prone to inaccuracies when structures such as blood vessels appeared near any of the aponeuroses. These problems necessitated high operator supervision, review and reassessment of the segmentation processes. Despite this some inaccuracies could still exist in the data set and could have affected results from this part of the methodology. The accuracy of the ASM for this specific data set was not specifically evaluated here, but previous work has shown high degrees of accuracy for similar approaches, with errors less than 1 mm (Darby et al., 2012). This evidence, combined with the high degree of manual supervision of the approach mean the segmentation results obtained here should be robust and with the small possible errors unlikely to significantly affect the prediction errors reported. However, it is pertinent to appreciate that some of the

challenges encountered with the ASM applied here have and could be solved with the use of alternative segmentation methods (Noble 2010).

Different parameterizations were also applied to the Horn-Schunck OF implementation (Section 3.3), which ranged from the maximum number of iterative flow calculations, to the definition of smaller regions to extract flow from within the muscle cross-sectional area. Preliminary pilot work, which has not been specifically included in the thesis involved trying different combinations of overall flow resulting from differences between superficial and deep muscle regions and consideration of flow statistics such as variance and standard deviation. The data that provided the best prediction results was however found to be the use of the overall mean of the entire muscle area, delimited by the two aponeuroses. The use of local OF methods, such as Lucas-Kanade or Bigun's structure tensor method, together with local segmentation techniques capable of locating structures contained in the muscle cross-sectional area such as muscle fascicles, could potentially provide useful information regarding contractions of small volumes within the muscle and be added to the inputs of the inference algorithms to improve their predictive power during subtle muscle contractions or for prediction of the torque fluctuations which proved challenging for all the inference algorithms assessed (Chapter 7). How much these additional approaches could improve torque prediction is currently unknown, but does warrant further investigation particularly if information on changes in architecture muscle features (e.g. fascicles) is the interest of the researcher.

One advantage of the approaches that were employed here is that the different methodologies applied to the collected US data have the possibility of being implemented in real time. Optical Flow using Horn Schunck has been proven achievable in real time for tracking respiratory motion in magnetic resonance guided high-intensity focused ultrasound (Zachiu et al. 2015). ASM has also been implemented online when visualising deep neck muscles (Cunningham, P. Harding, et al. 2017). Having real-time feedback would enable the use of this methodology by Clinical physiologist for assessing patient's joint torque with the only use of an US device in one muscle and without the need of waiting for results. Such a tool

could be valuable for informing interventions for a range of conditions, such as surgical decision making in the case of Cerebral palsy or requirement and location of Botox injections for the treatment of muscle spasticity.

In the results presented in Chapter 5, different errors in torque prediction were observed depending on the direction of the associated muscle movement. Specifically, this was especially accentuated for the case of passive dorsiflexion versus plantar flexion ankle joint rotations (Figures 5.1 vs 5.2, and Figures 5.3 vs 5.4). The length at which the muscle fibres are activated could have an influence on the forces produced, due to the force-length relationship (Figure 2.4). In any cyclical movement task there are periods when the muscle reaches a state of inactivity (quiescence) and will undergo passive lengthening and shortening. The muscle properties could therefore affect the length the fibre achieves just prior to activation. Results showing that shortening does not seem to be equal to lengthening is interesting as it may reflect the muscle properties or the tracking capabilities and such observations have been reported previously (Hodson-Tole et al. 2016). Further work should investigate what the case is (i.e. physiological or technical), and enable a fuller understanding of the interactions between passive and active muscle length changes during dynamic tasks.

Future work could be focused on the use of classification algorithms capable of firstly detecting the type of task performed, distinguishing between active and passive movements (Cunningham et al. 2013) and concentric and eccentric direction; and then automatically applying the prediction algorithms using the recommended parameters to facilitate better prediction. This would be valuable because it would improve the data predictions by allowing the use of specialized data processing, and would enable the study of individual muscle behaviour associated with motor control during free movement or when performing functional tasks. Current limitations in technology mean that such processes do however require extended processing times, meaning their current value may be limited for clinical application. However, the on-going developments in both hardware and

algorithmic approaches to deep learning are likely to improve this situation in coming years.

In conclusion, this research evaluated the influence of ultrasound data properties with a focus on the development and validation of methodological approaches to exploit the capabilities of ultrasound technology, and to predict dynamic properties in skeletal muscles *in vivo* exclusively using ultrasound data. Results indicated that accurate predictions of net joint torque from US data of one muscle are achievable, and showed different optimal parameters depending on the muscle movement to study. Suggestions for future work are proposed, including strategies to compensate the limitations observed in this work, improvements of the proposed methodology for its implementation in real time and its combination with other techniques for predicting more complicated functional tasks.

Bibliography

Aggeloussis, N. et al., 2010. Reproducibility of fascicle length and pennation angle of gastrocnemius medialis in human gait in vivo. *Gait and Posture*, 31(1), pp.73–77.

Alessandrini, M. et al., 2016. Detailed Evaluation of Five 3D Speckle Tracking Algorithms using Synthetic Echocardiographic Recordings. , pp.1–13.

Anandan, P., 1987. *Measuring Visual Motion from Image Sequences*. University of Massachusetts.

Angelini, E.D. & Gerard, O., 2006. Review of myocardial motion estimation methods from optical flow tracking on ultrasound data. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp.1537–1540.

Arndt, A.N. et al., 1998. Individual muscle contributions to the in vivo achilles tendon force. *Clinical Biomechanics*, 13(7), pp.532–541.

Baraldi, P. et al., 1996. Evaluation of differential optical flow techniques on synthesized echo images. *IEEE Transactions on Biomedical Engineering*, 43(3), pp.259–272.

Barbosa, D. et al., 2014. Fast Tracking of the Left Ventricle Using Global Anatomical Affine Optical Flow and Local Recursive Block Matching. *Midas Journal*, 10. Available at: http://hdl.handle.net/10380/3481.

Barron, J.L., Fleet, D.J. & Beauchemin, S.S., 1994. Performance of Optical Flow Techniques. *International journal of computer …*, 1(12), pp.43–77. Available at: http://link.springer.com/article/10.1007/BF01420984.

Baştanlar, Y. & Ozuysal, M., 2014. Introduction to machine learning. *Methods in molecular biology (Clifton, N.J.)*, 1107, pp.105–28. Available at: http://alex.smola.org/drafts/thebook.pdf%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/24272434.

Baxter, J.R. & Piazza, S.J., 2014. Plantar flexor moment arm and muscle volume predict torque-generating capacity in young men. *Journal of Applied Physiology*, 116(December 2013), pp.538–44. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/24371016.

Bishop, E., 1963. A generalization of the Stone–Weierstrass theorem. *Pacific Journal of Mathematics*, 13(4).

Bojsen-Møller, J. & Magnusson, S.P., 2015. Heterogeneous loading of the human Achilles tendon in vivo. *Exercise and Sport Sciences Reviews*, 43(4), pp.190–197.

Bouguet, J., 1999. Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm. *In Practice*, 1(2), pp.1–9. Available at: http://trac.assembla.com/dilz_mgr/export/272/doc/ktl-tracking/algo_tracking.pdf.

Boukerroui, D., Noble, J.A. & Brady, M., 2003. Velocity estimation in ultrasound images: {A} block matching approach. *Proc. International Conference on Information Processing in Medical Imaging*, pp.586–598.

Boyd, J.P. & Xu, F., 2009. Divergence (Runge Phenomenon) for least-squares polynomial approximation on an equispaced grid and Mock-Chebyshev subset interpolation. *Applied Mathematics and Computation*, 210(1), pp.158–168. Available at: http://dx.doi.org/10.1016/j.amc.2008.12.087.

Brandenburg, J.E. et al., 2016. Quantifying passive muscle stiffness in children with and without cerebral palsy using ultrasound shear wave elastography. *Developmental Medicine and Child Neurology*, 58(12), pp.1288–1294.

Brereton, R.G. & Lloyd, G.R., 2010. Support vector machines for classification and regression. *The Analyst*, 135(2), pp.230–267.

Campbell, C., 2001. An Introduction to Kernel Methods Kernel Methods. *Studies in Fuzziness and Soft Computing*, 66, pp.155–192.

Cootes, T., Baldock, E. & Graham, J., 2000. An introduction to active shape models. *Image Processing and Analysis*, pp.223–248. Available at: http://person.hst.aau.dk/lasse/teaching/IACV/doc/asm_overview.pdf.

Cortes, C. & Vapnik, V., 1995. Support-Vector Networks. *Machine Learning*, 20(3), pp.273–297.

Cuesta-Vargas, A.I. & Gonzalez-Sanchez, M., 2014. Prediction of Maximal Surface Electromyographically Based Voluntary Contractions of Erector Spinae Muscles From Sonographic Measurements During Isometric Contractions. *Journal of*

Ultrasound in Medicine, 33(3), pp.399–404. Available at: http://doi.wiley.com/10.7863/ultra.33.3.399 [Accessed June 10, 2017].

Cunningham, R., 2015. *The application of b-mode ultrasonography for analysis of human skeletal muscle*. PhD thesis. Manchester Metropolitan University.

Cunningham, R., Harding, P. & Loram, I., 2013. Automated Measurement of Human Skeletal Calf Muscle Contraction via B-Mode Ultrasound Imaging. *17th International Conference on Medical Imaging, Understanding and Analysis*.

Cunningham, R., Harding, P. & Loram, I., 2017. Real-Time Ultrasound Segmentation, Analysis and Visualization of Deep Cervical Muscle Structure. *Transactions on Medical Imaging*, 62(c), pp.653–665.

Cunningham, R., Harding, P.J. & Loram, I.D., 2017. Deep Residual Networks for Quantification of Muscle Fiber Orientation and Curvature from Ultrasound Images Deep Residual Networks for Quantification of Muscle. , (July).

Darby, J. et al., 2012. Automated regional analysis of B-mode ultrasound images of skeletal muscle movement. *Journal of Applied Physiology*, 112(2), pp.313–327.

Darby, J. et al., 2013. Estimating skeletal muscle fascicle curvature from B-mode ultrasound image sequences. *IEEE Transactions on Biomedical Engineering*, 60(7), pp.1935–1945.

Deffieux, T. et al., 2008. Assessment of the mechanical properties of the musculoskeletal system using 2-D and 3-D very high frame rate ultrasound. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 55(10), pp.2177–2190.

DeOliveira, L.F. & Luporini Menegaldo, L., 2010. Individual-specific muscle maximum force estimation using ultrasound for ankle joint torque prediction using an EMG-driven Hill-type model. *Journal of Biomechanics*, 43(14), pp.2816–2821. Available at: http://dx.doi.org/10.1016/j.jbiomech.2010.05.035.

Dreiseitl, S. & Ohno-Machado, L., 2002. Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, 35(5–6), pp.352–359.

Duan, Q. et al., 2005. Evaluation of Optical Flow Algorithms for Tracking

Endocardial Surfaces on Three-Dimensional Ultrasound Data. *Proc. SPIE Medical Imaging*, 5750, pp.159–169. Available at: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=861226.

Facciolo, G. et al., 2011. Temporally consistent gradient domain video editing. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6819 LNCS, pp.59–73.

Finni, T. et al., 2012. On the hysteresis in the human Achilles tendon. *Journal of Applied Physiology*, (October 2012), pp.515–517.

Fleet, D.J. & Jepson, A.D., 1990. Computation of Component Image Velocitiy from Local Phase Information. , 5(1), pp.77–104.

Frangi, A.F. et al., 1998. Multiscale vessel enhancement filtering. *Medial Image Computing and Computer-Assisted Invervention - MICCAI'98. Lecture Notes in Computer Science, vol 1496*, 1496, pp.130–137.

Fukunaga, T. et al., 2001. In vivo behaviour of human muscle tendon during walking. *Proceedings of the Royal Society B: Biological Sciences*, 268(1464), pp.229–233. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2000.1361.

Gillard, D.M. et al., 2000. Isometric muscle length-tension curves do not predict angle-torque curves of human wrist in continuous active movements. *Journal of Biomechanics*, 33(11), pp.1341–1348.

Gomez-Verdejo, V. et al., 2011. Support vector machines with constraints for sparsity in the primal parameters. *IEEE Transactions on Neural Networks*, 22(8), pp.1269–1283.

Gordon, a. M., Huxley, a. F. & Julian, F.J., 1966. The variation in isometric tension with sarcomere length in vertebrate muscle fibres. *The Journal of Physiology*, 184(1), pp.170–192. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1357553&tool=pmcentrez&rendertype=abstract.

Guerrero, J. et al., 2007. Real-time vessel segmentation and tracking for ultrasound imaging applications. *IEEE Trans Med Imaging*, 26(8), pp.1079–1090. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17695128.

Hamner, S.R., Seth, A. & Delp, S.L., 2010. Muscle contributions to propulsion and support during running. *Journal of Biomechanics*, 43(14), pp.2709–2716. Available at: http://dx.doi.org/10.1016/j.jbiomech.2010.06.025.

Harding, P.J. et al., 2012. Automated detection of skeletal muscle twitches from B-mode ultrasound images: An application to motor neuron disease. *Pattern Recognition (ICPR), 2012 21st International Conference on*, (Icpr), pp.2630–2633.

Harding, P.J. et al., 2016. Ultrasound Based Detection of Fasciculations in Healthy and Diseased Muscles. *IEEE Transactions on Biomedical Engineering*, 63(3), pp.512–518.

Harding, P.J. et al., 2015. Ultrasound Based Detection of Fasciculations in Healthy and Diseased Muscles. *IEEE Transactions on Biomedical Engineering*, Accepted f, p.

Harris, A.J., 2005. Muscle Fiber and Motor Unit Behavior in the Longest Human Skeletal Muscle. *Journal of Neuroscience*, 25(37), pp.8528–8533. Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0923-05.2005.

Hastie, T., Tibshirani, R. & Friedman, J., 2009. The Elements of Statistical Learning. *Elements*, 1, pp.337–387. Available at: http://www.springerlink.com/index/10.1007/b94608.

Hauraix, H. et al., 2015. Muscle and tendon stiffness assessment using the alpha method and ultrafast ultrasound. *European Journal of Applied Physiology*, pp.1393–1400. Available at: http://link.springer.com/10.1007/s00421-015-3112-1.

Hauraix, H., Nordez, A. & Dorel, S., 2013. Shortening behavior of the different components of muscle-tendon unit during isokinetic plantar flexions. *Journal of applied physiology (Bethesda, Md. : 1985)*, 115(7), pp.1015–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23887903.

Heeger, D.J., 1987. Model for the extraction of image flow. *Journal of the Optical Society of America. A, Optics and image science*, 4(8), pp.1455–1471.

Herbert, R.D. et al., 2011. In vivo passive mechanical behaviour of muscle fascicles and tendons in human gastrocnemius muscle-tendon units. *The Journal of physiology*, 589(Pt 21), pp.5257–67. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3225678&tool=pmcentrez&rendertype=abstract.

Héroux, M.E.H., Stubbs, P.W. & Herbert, R.D., 2016. Behavior of human gastrocnemius muscle fascicles during ramped submaximal isometric contractions. *Physiol Rep*, 4(17), pp.1–10.

Herzog, W. et al., 2015. A new paradigm for muscle contraction. *Frontiers in Physiology*, 6(MAY), pp.1–11.

Hill, A. V & Hartree, W., 1921. The Thermo-Elastic Properties of Muscle. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 210, pp.153–173. Available at: http://www.jstor.org/stable/92098.

Hill, A.V., 1952. The thermodynamics of elasticity in resting striated muscle. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 139(897), pp.464–497.

Hoang, P. et al., 2009. Passive mechanical properties of gastrocnemius in people with multiple sclerosis. *Clinical Biomechanics*, 24(3), pp.291–298. Available at: http://dx.doi.org/10.1016/j.clinbiomech.2008.12.008.

Hoang, P.D. et al., 2005. A new method for measuring passive length-tension properties of human gastrocnemius muscle in vivo. *Journal of Biomechanics*, 38(6), pp.1333–1341.

Hoang, P.D. et al., 2007. Passive mechanical properties of human gastrocnemius muscle tendon units, muscle fascicles and tendons in vivo. *The Journal of experimental biology*, 210, pp.4159–4168.

Hodges, P.W. et al., 2003. Measurement of muscle contraction with ultrasound imaging. *Muscle and Nerve*, 27(6), pp.682–692.

Hodson-Tole, E.F., Wakeling, J.M. & Dick, T.J.M., 2016. Passive muscle-tendon unit gearing is joint dependent in human medial gastrocnemius. *Frontiers in Physiology*, 7(MAR), pp.1–8.

Horn, B. & Schunck, B., 1981. "Determining optical flow." *Artificial Intelligence*, 17(1–2), pp.185–203.

Huxley, A., 1957. Muscle structure and theories of contraction. *Prog Biophys Biophys Chem.*, 7, pp.257–318.

Ito, L. et al., 1998. Nonisometric behavior of fascicles during isometric contractions of a human muscle. *Journal of Applied Physiology*, 85(4), pp.1230–1235.

Kawakami, Y. et al., 2000. Fatigue responses of human triceps surae muscles during repetitive maximal isometric contractions. *Journal of applied physiology (Bethesda, Md. : 1985)*, 88(6), pp.1969–1975.

Kawakami, Y., Ichinose, Y. & Fukunaga, T., 1998. Architectural and functional features of human triceps surae muscles during contraction. *Journal of applied physiology (Bethesda, Md. : 1985)*, 85(2), pp.398–404. Available at: http://0-jap.physiology.org.wam.leeds.ac.uk/content/85/2/398 [Accessed June 9, 2017].

Kawamoto, S. et al., 2014. Ultrasound imaging and semi-automatic analysis of active muscle features in electrical stimulation by optical flow. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2014, pp.250–253.

Keynes, R.D. & Aidley, D.J., 2001. *Nerve and Muscle*, Cambridge: Cambridge University Press. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=120242&site=ehost-live.

Klaus Debes, Alexander Koenig, H.-M.G., 2005. Transfer Functions in Artificial Neural Networks. *Neuroinformatics*, pp.1–11.

Landin, D., Thompson, M. & Reid, M., 2015. Knee and Ankle Joint Angles Influence the Plantarflexion Torque of the Gastrocnemius. *Journal of clinical medicine research*, 7(8), pp.602–606.

Lee, J.H. & Kim, S.M., 2013. Estimating contrast agent motion from ultrasound images using an anisotropic diffusion-based optical flow technique. *Computers in Biology and Medicine*, 43(11), pp.1853–1862. Available at: http://dx.doi.org/10.1016/j.compbiomed.2013.09.004.

Leineweber, M. & Gao, Y., 2015. Quantifying skeletal muscle recovery in a rat injury model using ultrasound imaging. *Journal of Biomechanics*, 48(2), pp.379–382. Available at:

http://linkinghub.elsevier.com/retrieve/pii/S0021929014006472.

Li, L. & Tong, K., 2005. Musculotendon parameters estimation by ultrasound measurement and geometric modeling: application on brachialis muscle. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 5, pp.4974–4977.

Li, L. & Tong, R.K.Y., 2016. Combined Ultrasound Imaging and Biomechanical Modeling to Estimate Triceps Brachii Musculotendon Changes in Stroke Survivors. *BioMed Research International*, 2016.

Lin, Y. & Tai, S., 1996. Fast Full-Search Hock- atching Algorithm for Motion-Compensated.

Liu, Y.H. et al., 2007. Motion tracking on elbow tissue from ultrasonic image sequence for patients with lateral epicondylitis. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp.95–98.

Lopata, R.G.P. et al., 2006. In Vivo 3D Cardiac and Skeletal Muscle Strain Estimation. *2006 IEEE Ultrasonics Symposium*, pp.744–747. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4152054.

Loram, I.D., Maganaris, C.N. & Lakie, M., 2006. Use of ultrasound to make noninvasive in vivo measurement of continuous changes in human muscle contractile length. *Journal of Applied Physiology*, 100(4), pp.1311–1323. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16339341.

Lucas, B.D. & Kanade, T., 1981. An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, 2, pp.674–679. Available at: http://www.ic.unicamp.br/~rocha/teaching/2013s1/mc851/aulas/additional-material-lucas-kanade-tracker.pdf.

Maganaris, C.N., 2004. Imaging-based estimates of moment arm length in intact human muscle-tendons. *European Journal of Applied Physiology*, 91(2–3), pp.130–139. Available at:

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14685871.

Mannion, A.F. et al., 2012. Spine stabilisation exercises in the treatment of chronic low back pain: A good clinical outcome is not associated with improved abdominal muscle function. *European Spine Journal*, 21(7), pp.1301–1310.

Mantzaris, D. et al., 2008. A non-symbolic implementation of abdominal pain estimation in childhood. *Information Sciences*, 178(20), pp.3860–3866.

Marquardt, D.W., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), pp.431–441.

Masood, T. et al., 2014. Differential contributions of ankle plantarflexors during submaximal isometric muscle action: A PET and EMG study. *Journal of Electromyography and Kinesiology*, 24(3), pp.367–374. Available at: http://dx.doi.org/10.1016/j.jelekin.2014.03.002.

MathWorks, 2016. Choose a Multilayer Neural Network Training Function. *mathworks*. Available at: https://uk.mathworks.com/help/nnet/ug/choose-a-multilayer-neural-network-training-function.html [Accessed June 10, 2017].

Miguez, D., Loram, I. & Harding, P.J., 2017. A technical note on variable inter-frame interval as a cause of non-physiological experimental artefacts in ultrasound Subject Category : Subject Areas :

Miyatani, M. et al., 2004. The accuracy of volume estimates using ultrasound muscle thickness measurements in different muscle groups. *European Journal of Applied Physiology*, 91(2–3), pp.264–272.

Mohamed, A.S.A., 2015. *Automated Speckle Tracking in Ultrasound Images of Tendon Movements*. PhD thesis. University of Salford.

Møller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4), pp.525–533.

Montaldo, G. et al., 2009. Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(3), pp.489–506.

Moon, W.K. et al., 2012. Computer-Aided Diagnosis Based on Speckle Patterns in Ultrasound Images. *Ultrasound in Medicine & Biology*, 38(7), pp.1251–1261.

Nagel, H.-H., 1987. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33(3), pp.299–324.

Namburete, A.I.L., Rana, M. & Wakeling, J.M., 2011. Computational methods for quantifying in vivo muscle fascicle curvature from ultrasound images. *Journal of Biomechanics*, 44(14), pp.2538–2543. Available at: http://dx.doi.org/10.1016/j.jbiomech.2011.07.017.

Narici, M. V et al., 1996. In vivo human gastrocnemius architecture with changing joint angle at rest and during graded isometric contraction. *The Journal of physiology*, (1), pp.287–97. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8910216 [Accessed June 9, 2017].

National Cancer Institute, muscle structure - Wikimedia Commons. Available at: https://commons.wikimedia.org/wiki/File:Illu_muscle_structure.jpg [Accessed June 10, 2017].

Noble, J. a, 2010. Ultrasound image segmentation and tissue characterization. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine*, 224, pp.307–316.

Noble, J. a & Boukerroui, D., 2006. Ultrasound image segmentation: A survey. *IEEE Transactions on Medical Imaging*, 25(8), pp.987–1010. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-33746885878&partnerID=40&md5=35f0db39716e5d7fb5691d63138c7d29%5Cnhttp://ieeexplore.ieee.org/ielx5/42/34770/01661695.pdf?tp=&arnumber=1661695&isnumber=34770.

Nordez, A. et al., 2009. Electromechanical delay revisited using very high frame rate ultrasound. *Journal of applied physiology (Bethesda, Md. : 1985)*, 106(6), pp.1970–1975.

Nordez, A. et al., 2010. Improvements to Hoang et al.'s method for measuring passive length-tension properties of human gastrocnemius muscle in vivo. *Journal of Biomechanics*, 43(2), pp.379–382.

Papadopoulos, H. & Haralambous, H., 2011. Reliable prediction intervals with regression neural networks. *Neural Networks*, 24(8), pp.842–851. Available at: http://dx.doi.org/10.1016/j.neunet.2011.05.008.

Payen, T. et al., 2012. Comparison of Echo-Power Estimation Using Linearized

Video Data and Raw Data for Dynamic Contrast-Enhanced Ultrasound. , pp.1335–1338.

Payen, T. et al., 2013. Echo-power estimation from log-compressed video data in dynamic contrast-enhanced ultrasound imaging. *Ultrasound in medicine & biology*, 39(10), pp.1826–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23879926.

Peckov, A., 2012. *A Machine Learning Approach to Polynomial Regression*. PhD thesis. Jozef Stefan International Postgraduate School.

Pulkovski, N. et al., 2008. Tissue Doppler imaging for detecting onset of muscle activity. *Muscle & nerve*, 37(5), pp.638–649.

Rana, M., Hamarneh, G. & Wakeling, J.M., 2009. Automated tracking of muscle fascicle orientation in B-mode ultrasound images. *Journal of Biomechanics*, 42(13), pp.2068–2073. Available at: http://dx.doi.org/10.1016/j.jbiomech.2009.06.003.

Richfield, D., 2014. Medical gallery of David Richfield 2014. *WikiJournal of Medicine*, 1(2). Available at: https://en.wikiversity.org/wiki/WikiJournal_of_Medicine/Medical_gallery_of_David_Richfield_2014 [Accessed June 8, 2017].

Roberts, T.J. et al., 2016. Muscular Force in Running Turkeys : The Economy of Minimizing Work Author (s): Thomas J . Roberts , Richard L . Marsh , Peter G . Weyand and C . Richard Taylor Published by : American Association for the Advancement of Science Stable URL : http://www.js. , 275(5303), pp.1113–1115.

Rognin, N.G. et al., 2008. In-vivo perfusion quantification by contrast ultrasound:Validation of the use of linearized video data vs. raw RF data. *Proceedings - IEEE Ultrasonics Symposium*, pp.1690–1693.

Salles, S. et al., 2014. Experimental Evaluation of Spectral-Based Plane Wave Compounding. , 61(11), pp.1824–1834.

Savitzky, A. & Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8), pp.1627–1639. Available at: http://dx.doi.org/10.1021/ac60214a047.

Seikel, J.A., King, D.W. & Drumright, D.G., 2009. *Anatomy & Physiology for*

Speech, Language, and Hearing, Cengage Learning. Available at: https://books.google.co.uk/books?id=LFBOhaD1JHwC.

Seynnes, O.R. et al., 2015. Ultrasound-based testing of tendon mechanical properties: a critical evaluation. *Journal of Applied Physiology*, 118(2), pp.133–141. Available at: http://jap.physiology.org/lookup/doi/10.1152/japplphysiol.00849.2014.

Shannon, C.E., 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27(July 1928), pp.379–423.

Sherrington, C.S., 1906. *The integrative action of the nervous system.*,

Shi, J. et al., 2007. Assessment of muscle fatigue using sonomyography: Muscle thickness change detected from ultrasound images. *Medical Engineering & Physics*, 29(4), pp.472–479. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1350453306001469.

Shi, J. & Tomasi, C., 1994. Good features to track. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*, pp.593–600.

Shinohara, M. et al., 2006. The medial gastrocnemius muscle attenuates force fluctuations during plantar flexion. *Experimental Brain Research*, 169(1), pp.15–23.

Singh, A., 1990. An estimation-theoretic framework for image-flow computation. *Proceedings of Third International Conference on Computer Vision*, pp.168–177.

Smola, A.J. & Schölkopf, B., 2004. A Tutorial on Support Vector Regression. *Statistics and Computing*, 14(3), pp.199–222.

Somphone, O. et al., 2013. Fast myocardial motion and strain estimation in 3D cardiac ultrasound with Sparse Demons. *Proceedings - International Symposium on Biomedical Imaging*, (1), pp.1182–1185.

Sterling Hot Yoga, 2014. Agonist-Antagonist-Muscles. *http://shywmobile.com*. Available at: http://i1.wp.com/shywmobile.com/wp-content/uploads/2014/11/Agonist-Antagonist-Muscles_Nov-2014.jpg [Accessed June 9, 2017].

Stoitsis, J. et al., 2005. Analysis and quantification of arterial wall motion from B-

mode ultrasound images - comparison of block-matching and optical flow. *Conf Proc IEEE Eng Med Biol Soc*, 5, pp.4469–4472. Available at: http://dx.doi.org/10.1109/IEMBS.2005.1615459.

Strasser, E.M. et al., 2013. Association between ultrasound measurements of muscle thickness, pennation angle, echogenicity and skeletal muscle strength in the elderly. *Age*, 35(6), pp.2377–2388.

Tanter, M. & Fink, M., 2014. Ultrafast imaging in biomedical ultrasound. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(1), pp.102–119.

Tomalka, A. et al., 2017. The active force–length relationship is invisible during extensive eccentric contractions in skinned skeletal muscle fibres. *Proceedings of the Royal Society of London B: Biological Sciences*, 284(1854). Available at: http://rspb.royalsocietypublishing.org/content/284/1854/20162497.article-info [Accessed July 10, 2017].

Vaerenbergh, S. Van, 2009. Kernel Methods for Nonlinear Identification , Equalization and Separation of Signals. *Science*.

Vapnik, V.N. & Chervonenkis, A.Y., 1971. On the uniform convergence of relative frequencies of events to their probabilities. *Measures of Complexity: Festschrift for Alexey Chervonenkis*, XVI(2), pp.11–30.

Vieira, T.M.M. et al., 2013. How much does the human medial gastrocnemius muscle contribute to ankle torques outside the sagittal plane? *Human Movement Science*, 32(4), pp.753–767. Available at: http://dx.doi.org/10.1016/j.humov.2013.03.003.

Wakahara, T. et al., 2005. Effects of Passive Ankle and Knee Joint Motions on the Length of Fascicle and Tendon of the Medial Gastrocnemius Muscle. *International Journal of Sport and Health Science*, 3, pp.75–82. Available at: http://joi.jlc.jst.go.jp/JST.JSTAGE/ijshs/3.75?from=Google%5Cnhttp://japanlinkcenter.org/JST.JSTAGE/ijshs/3.75?from=Google.

Waxman, A.M., Wu, J. & Bergholm, F., 1988. Convected activation profiles and the measurement of visual motion. *Proceedings CVPR '88: The Computer Society Conference on Computer Vision and Pattern Recognition*, (60).

Wikisonix (Ultrasonix), WikiSonix. Available at:

http://www.ultrasonix.com/wikisonix/index.php/Main_Page [Accessed August 11, 2015].

Woittiez, R.D. et al., 1984. A three-dimensional muscle model: A quantified relation between form and function of skeletal muscles. *Journal of Morphology*, 182(1), pp.95–113. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6492171 [Accessed June 23, 2017].

Wrench, A. a & Scobbie, J.M., 2008. High-speed Cineloop Ultrasound vs. Video Ultrasound Tongue Imaging: Comparison of Front and Back Lingual Gesture Location and Relative Timing. , pp.57–60. Available at: http://eresearch.qmu.ac.uk/2012/.

Wrench, A. a & Scobbie, J.M., 2011. Very high frame rate ultrasound tongue imaging. Available at: http://eresearch.qmu.ac.uk/2505/.

Xie, H.-B. et al., 2009. Estimation of wrist angle from sonomyography using support vector machine and artificial neural network models. *Medical engineering & physics*, 31(3), pp.384–391.

Zachiu, C. et al., 2015. An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs. *Physics in Medicine and Biology*, 60(23), pp.9003–9029. Available at: http://stacks.iop.org/0031-9155/60/i=23/a=9003?key=crossref.8e9f5a4034d646a52689c66d47c1d7b1.

Zajac, F.E., 1989. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Critical reviews in biomedical engineering*, 17(4), pp.359–411.

Zhang, Y. & Ayodele, T.O., 2010. Types of Machine Learning Algorithms. In *New Advances in Machine Learning*. pp. 103–134.

A Appendix - Ankle Torque Prediction using down sampled US derived inputs

Data are presented here to provide insight regarding potential biasing implicit on the use of an eight times smaller resolution for the SVR prediction error results showed in chapters 5 to 8. Results of all torque prediction algorithms at the fastest frame rate (1000 fps), but decimated to an eighth part (from 8000 data points per 8 second duration trial to 1000) in order to directly compare them with the SVR results are therefore shown.

The results contained in this appendix were obtained from the same data collection detailed in chapter 4, and analysed using the methods described in chapter 3, with the addition of the decimation process. The three sections that form this appendix correspond to the RMSE values obtained when performing passive rotation, isometric and isotonic contractions respectively.

A.1 Predicting Ankle Torque during Passive Joint Rotations

A.1.1 Dorsiflexion Movements

This section reports the results from the analysis of 8-bit and 16-bit precision data relating to the dorsiflexion phase of the passive ankle joint rotation movements. Table A.1 presents individual participant and mean group data from the 8-bit, fast frame rate data decimated to 1000 data points per trial, together with the average and standard deviation using 8000 samples per trial (last two rows).

Table A-1. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during passive dorsiflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.270	2.189	2.295	2.253	1.989
Participant 2	1.045	1.001	0.950	0.959	0.438
Participant 3	1.633	1.623	1.625	1.651	1.641
Participant 4	1.233	1.168	1.153	1.213	1.167
Participant 5	3.160	3.231	3.215	3.155	3.402
Participant 6	0.852	0.675	0.677	0.703	0.743
Participant 7	0.635	0.543	0.503	0.591	0.602
Participant 8	1.361	1.291	1.309	1.274	1.245
Participant 9	1.673	1.612	1.594	1.649	1.640
Avg. 1ks	1.540	1.482	1.480	1.494	1.430
STD. 1ks	0.778	0.829	0.846	0.809	0.904
Avg. 8ks	1.538	1.476	1.476	1.494	1.394
STD. 8ks	0.777	0.812	0.832	0.809	0.768

Similar results were observed for all the inference methods with a slightly worse performance of Linear Regression in terms of RMSE mean. The obtained values were very similar to the ones obtained using 8000 points per trial (Table 5.5 and last two rows of this table). The best prediction in terms of average was observed in ANN, with a result (1.430 ± 0.904 N m) that corresponds to 9.53 % of the maximum torque during passive trials (15 N m).

Table A.2 presents individual participant and mean group data from the 16-bit, fast frame rate data. The performance of the five different inference methods showed quite similar results in terms of both average and standard deviation, with the worst performance found in Cubic Regression. The lowest average error was observed in Quadratic Regression with an RMSE of (1.389 ± 0.728 N m) corresponding to 9.26 % of the maximum torque during passive trials (15 N m).

Table A-2. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during passive dorsiflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.505	2.583	2.711	2.540	2.553
Participant 2	1.259	1.243	1.218	1.217	1.121
Participant 3	1.507	1.496	1.521	1.520	1.514
Participant 4	1.161	1.115	1.110	1.157	1.339
Participant 5	2.437	2.406	2.392	2.409	1.543
Participant 6	0.576	0.499	1.750	0.520	1.647
Participant 7	0.511	0.487	0.493	0.483	0.453
Participant 8	1.148	1.155	1.131	1.174	1.155
Participant 9	1.542	1.513	1.509	1.524	1.499
Avg. 1ks	1.405	1.389	1.537	1.394	1.425
STD. 1ks	0.702	0.728	0.680	0.716	0.555
Avg. 8ks	1.400	1.431	1.530	1.394	1.477
STD. 8ks	0.697	0.709	0.660	0.716	0.632

To enable direct comparison of any differences in torque prediction error resulting from differences in data precision results obtained are shown in graphical form (Figure A.1). No important trends between 16-bit and 8-bit results are apparent in both median and variance.

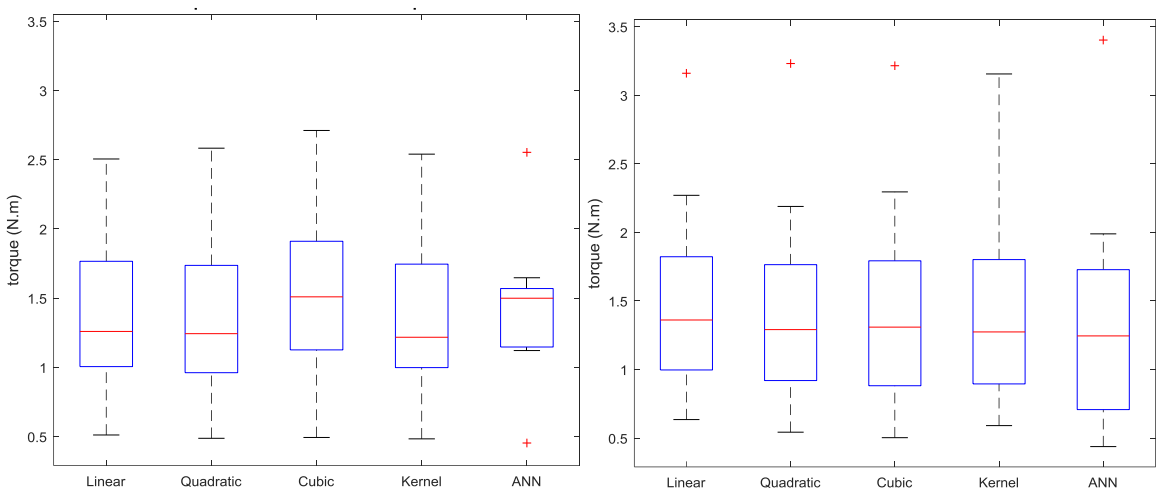

Figure A-1. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Passive measurements during dorsiflexion from nine participants collected at 1kS/s.

A.1.2 Plantar Flexion Movements

This section reports the results from the analysis of 8-bit and 16-bit data relating to the plantar flexion phase of the passive ankle joint rotation movements. Table A.3 presents individual participant and mean group data from the 8-bit, fast frame rate data.

Table A-3. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during passive plantarflexion. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	1.862	1.758	1.810	1.820	1.716
Participant 2	1.188	0.928	1.001	1.124	0.980
Participant 3	2.026	2.056	2.085	1.951	2.128
Participant 4	1.171	1.126	1.187	1.116	1.119
Participant 5	3.420	3.415	3.351	3.569	2.228
Participant 6	2.419	2.374	2.273	2.608	2.320
Participant 7	1.108	0.967	1.049	1.094	0.973
Participant 8	4.154	4.435	4.377	4.606	4.079
Participant 9	2.699	2.631	2.609	2.679	2.886
Avg. 1ks	2.227	2.188	2.194	2.285	2.047
STD. 1ks	1.062	1.183	1.128	1.212	1.013
Avg. 8ks	2.342	2.246	2.333	2.285	2.207
STD. 8ks	1.125	1.237	1.233	1.212	1.198

The performance of the inference algorithms presented was similar for the five different prediction algorithms used. ANN showed the best prediction result for this condition in both average and deviation resulting in a RMSE of 2.047 ± 1.013 N m that corresponded to 9.29% of the maximum torque in passive trials.

Table A.4 presents individual participant and mean group data from the 16-bit, fast frame rate data. The five prediction methods obtained similar results except linear regression and Kernel SVR, which showed a slightly bigger error in terms of average. The best predicting power was observed for the cubic regression, with an average of 1.794 ± 0.627 N m that corresponded to 11.96 % of the maximum torque in passive joint rotation tests.

Table A-4. RMSE Results for the five inference methods used for Joint Torque predictions. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	2.125	2.143	2.217	2.120	2.105
Participant 2	1.536	1.394	1.371	1.390	1.265
Participant 3	1.905	1.897	1.809	1.771	1.125
Participant 4	1.019	0.948	0.955	0.952	0.885
Participant 5	2.985	2.876	2.837	2.982	3.537
Participant 6	1.542	1.421	1.400	1.722	2.023
Participant 7	1.244	1.083	1.024	1.154	0.953
Participant 8	2.496	2.237	2.150	2.449	2.305
Participant 9	2.242	2.144	2.136	2.293	2.159
Avg. 1ks	1.899	1.794	1.766	1.870	1.817
STD. 1ks	0.629	0.627	0.626	0.655	0.854
Avg. 8ks	2.015	1.869	1.854	1.870	1.859
STD. 8ks	0.658	0.662	0.668	0.655	0.807

Results obtained are shown in graphical form (Figure A.2) sharing the same Y axis range (that represents joint torque RMSE) for the 8- and 16-bit data analysed. Besides having similar mean values, 16-bit results show lower variance of the RMSE values. When compared to the dorsiflexion results (Figure A.1), algorithms show bigger errors when predicting platarflexions.

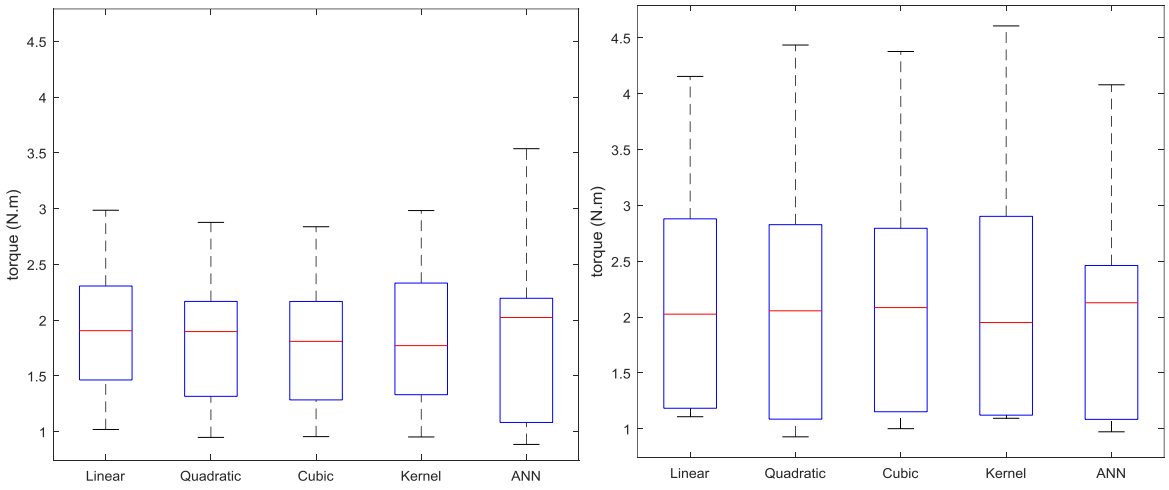

Figure A-2. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Passive measurements during Plantarflexion from nine participants collected at 1kS/s.

A.2 Predicting Ankle Joint Torque during Isometric Contractions

This section details the results of the prediction of the different algorithms for isometric contractions at two different activation levels (30 and 60% MVC) using 1000 data points per trial, also showing the influence of the data precision in their performance.

A.2.1 Isometric Contractions at 30% MVC

Table A.5 presents individual participant and mean group data from the 8-bit, fast frame rate data, including the average and standard deviation results when using 8000 samples per trial from table 6.5.

Table A-5. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during Isometric contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.344	6.274	6.210	6.225	6.415
Participant 2	4.081	3.651	3.702	4.043	3.689
Participant 3	7.465	7.350	7.316	7.316	7.326
Participant 4	3.822	3.571	3.793	3.847	4.625
Participant 5	11.222	11.054	11.125	11.114	10.134
Participant 6	9.625	11.824	11.827	10.913	11.778
Participant 7	1.929	1.812	1.809	2.126	2.237
Participant 8	4.626	4.430	4.516	4.678	4.142
Participant 9	2.494	2.776	3.794	2.345	2.870
Avg. 1ks	5.734	5.860	6.010	5.845	5.913
STD. 1ks	3.186	3.588	3.478	3.362	3.297
Avg. 8ks	5.891	5.906	6.017	5.845	6.132
STD. 8ks	3.041	3.483	3.417	3.362	3.422

All the inference techniques provided similar results for 8-bit data in both terms of average and standard deviation for this condition. Linear regression showed the best fit (5.734 ± 3.186 N m) equivalent to a 19.13% of the average participant maximum torque during 30% MVC passive trials (29.97 N m).

Furthermore, the results for the different prediction algorithms using 16-bit US data precision at fast frame rate can be found in Table A.6, including individual participant results and global mean and deviation values.

Table A-6. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during Isometric contractions at 30% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	6.128	6.153	6.209	6.019	6.309
Participant 2	4.178	3.866	3.813	4.285	3.411
Participant 3	7.298	7.208	7.213	7.222	7.301
Participant 4	3.602	3.434	3.443	3.738	3.534
Participant 5	10.607	10.566	10.534	10.574	11.024
Participant 6	10.859	12.925	13.076	11.971	12.960
Participant 7	1.665	1.567	1.539	1.962	1.623
Participant 8	4.380	4.258	4.383	4.501	4.251
Participant 9	1.262	1.302	1.832	1.253	1.732
Avg. 1ks	5.553	5.698	5.782	5.725	5.794
STD. 1ks	3.494	3.959	3.920	3.649	4.005
Avg. 8ks	5.597	5.727	5.771	5.725	5.633
STD. 8ks	3.437	3.914	3.909	3.649	3.105

From the table A.6 it can be appreciated a quite similar behaviour in all the prediction algorithms in both average and deviation, finding slightly better results for Linear Regression, which showed the best average result (5.553 N m) corresponding to 18.53% of the average participant maximum torque during passive trials.

The different data precision information contained in tables A.5 and A.6 at ultra-fast framerates using 1000 data points per trial was compiled and showed in Figure A.3 for the 30% MVC submaximal isometric contraction level.

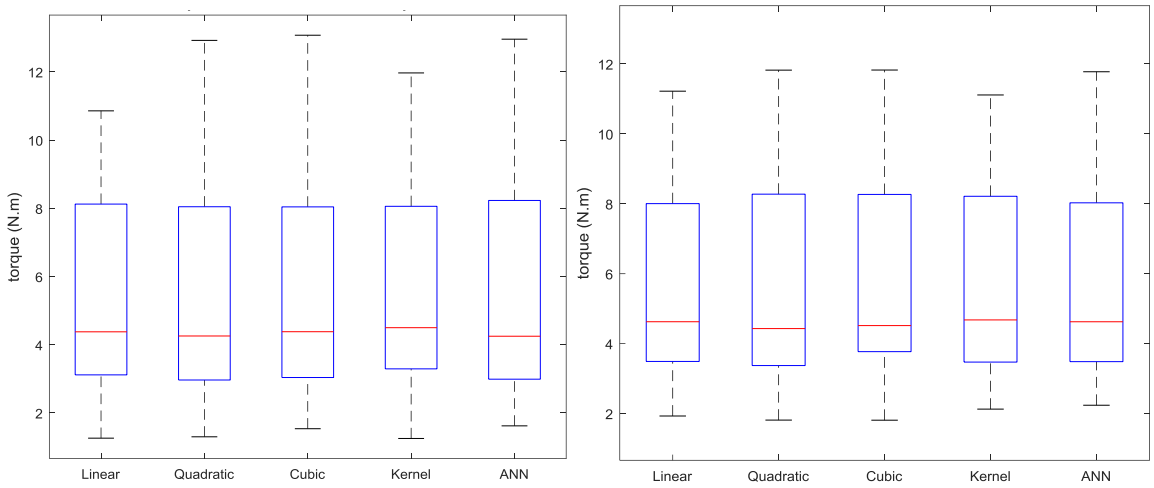

Figure A-3. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Isometric measurements at 30% MVC from nine participants collected at 1kS/s.

No differences can be observed between prediction algorithms for the case of the median, also showing similar variances for both 16 and 8-bit data precision.

A.2.2 Isometric Contractions at 60% MVC

Similar to previous section with 30% MVC isometric contractions, equivalent results for the biggest submaximal activation level studied in this work (60% MVC) when acquiring US data at 1000 fps are present in this section using two different US data precision. Table A.7 presents individual participant and mean group data from the 8-bit, fast frame rate data.

Table A-7. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit compressed US data during Isometric contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	12.544	10.914	11.443	11.454	11.168
Participant 2	20.627	26.506	26.718	23.803	28.163
Participant 3	6.137	4.773	4.367	4.854	6.699
Participant 4	6.243	4.707	5.667	5.428	5.686
Participant 5	17.919	15.439	16.011	16.320	15.386
Participant 6	9.420	7.302	7.609	7.260	7.880
Participant 7	5.605	3.985	4.099	4.733	5.323
Participant 8	9.145	6.787	6.632	7.339	5.939
Participant 9	3.549	4.237	4.557	3.724	4.982
Avg. 1ks	10.132	9.406	9.678	9.435	10.136
STD. 1ks	5.842	7.433	7.490	6.697	7.563
Avg. 8ks	10.149	9.429	9.643	9.435	9.255
STD. 8ks	5.749	7.275	7.374	6.697	5.532

Linear Regression provided one of the worst fit in terms of average, however its standard deviation was the lowest. Quadratic and Cubic Regressions and Kernel SVR showed almost identical results in both terms of average and deviation. The best fit was observed with quadratic regression, resulting in a RMSE of 9.406 ± 7.433 N m equivalent to a 15.69% of the average maximum torque during 60% MVC isometric trials (59.94 N m).

In contrast, Table A.8 presents individual participant and mean group data, resulting from analysis of 16-bit precision data during isometric contractions at 60 % MVC for a frame rate of 1000 fps.

Table A-8. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit compressed US data during Isometric contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	9.808	6.443	6.274	7.719	5.924
Participant 2	10.263	6.110	6.113	7.448	6.052
Participant 3	5.185	2.615	2.498	3.168	2.392
Participant 4	5.660	3.468	3.507	4.311	3.415
Participant 5	13.370	7.870	7.639	9.998	7.170
Participant 6	7.804	5.539	6.143	4.169	6.894
Participant 7	4.773	3.205	3.129	3.546	3.119
Participant 8	6.728	4.556	4.720	4.793	4.015
Participant 9	3.091	3.744	3.883	2.587	3.568
Avg. 1ks	7.409	4.839	4.879	5.304	4.728
STD. 1ks	3.234	1.755	1.743	2.501	1.784
Avg. 8ks	7.420	4.854	4.884	5.304	4.655
STD. 8ks	3.244	1.780	1.739	2.501	1.778

Linear Regression showed worst results in both terms of average and standard deviation, followed by Kernel SVR. Quadratic and Cubic Regressions together with ANN provided the best fits to the empirically acquired joint torque, being slightly better for the last one. ANN RMSE value (4.728 ± 1.784) corresponded to 7.89% of the average maximum torque during 60% MVC isometric tasks (59.94 N m). This value represents the most accurate prediction of Joint torque observed for this condition.

The different data precision information for the 60% MVC submaximal isometric contraction level at ultra-fast framerates was compiled and showed in Figure A.4, These graphical results provide information regarding the median and variance of the distribution of the errors per prediction algorithm used.

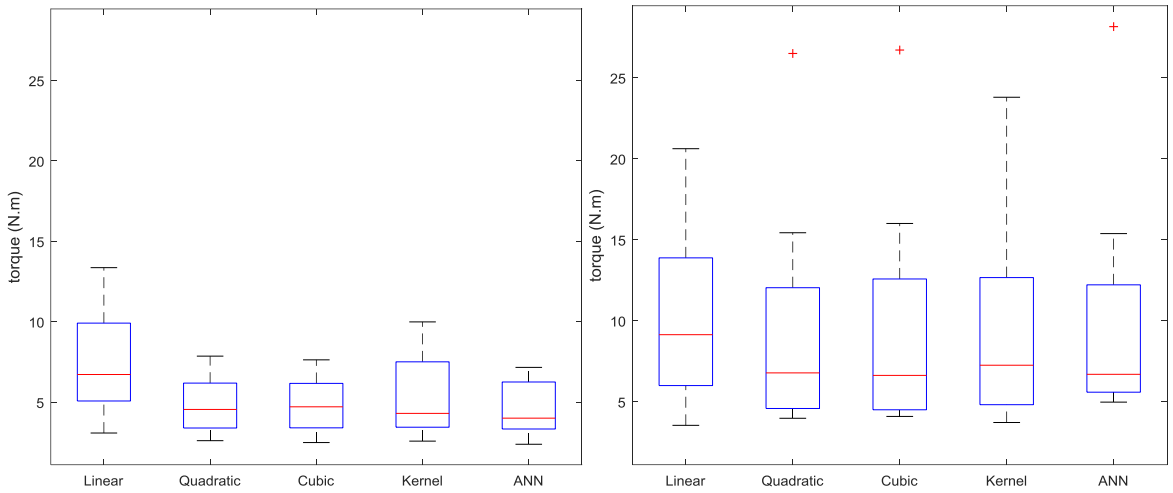

Figure A-4. Boxplot comparison of the RMSE values distribution for the five Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of Isometric measurements at 60% MVC from nine participants collected at 1kS/s.

Important differences can be appreciated for the case of 60% MVC isometric contractions. The use of higher precision data improved notably the fit of the predicted torque to the empirically acquired values in both median and variance, moving from an average for the five algorithms of 9.757 N m for the 8-bit precision, and 5.431 for the case of 16-bit; that correspond to a 16.28% and 9.06% of the average maximum torque for 60% MVC isometric contractions respectively.

A.3 Torque Prediction during Isotonic Contractions 1k

This section shows joint torque prediction results using plane wave transmission at a frame rate of 1000 fps during eight second duration isotonic contractions, but only using 1000 data points per trial.

A.3.1 Isotonic Contractions at 30% MVC

Table A.9 shows the individual participant and mean group predictions using 8-bit data precision at a frame rate of 1000 fps, together with the average and standard deviation results contained in Table 7.5 (last 2 rows) using 8000 samples per trial.

Table A-9. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	4.066	4.960	6.993	3.550	4.696
Participant 2	11.091	9.167	9.493	6.103	9.665
Participant 3	6.205	5.794	5.656	5.265	5.755
Participant 4	5.247	3.768	3.790	7.717	3.778
Participant 5	8.380	5.524	5.756	5.758	5.528
Participant 6	8.176	5.883	5.752	3.869	4.987
Participant 7	6.009	6.346	6.317	10.093	5.505
Avg. 1ks	7.025	5.921	6.251	6.051	5.702
STD. 1ks	2.356	1.656	1.730	2.267	1.871
Avg. 8ks	7.065	5.934	5.994	6.051	5.557
STD. 8ks	2.337	2.338	2.294	2.267	2.198

Results show a slight worse fit of the empirically collected joint torque data for linear regression and a similar response for quadratic and cubic regressions and Kernel SVM. The best prediction was observed when applying ANN with an average (5.702 ± 1.871 N m) equivalent to 11% of the maximum joint torque for 30% MVC isotonic contractions (51.8 N m).

Opposite to these, results using 16-bit precision US data are compiled in Table A.10, showing the RMSE values per participant together with the overall mean and deviation. Again, ANN showed the best prediction of all methods, that also showed

similar trends consisting of a worst response for linear regression. The best fit (5.189 ± 1.036) corresponded to 10.01% of the maximum acquired torque for these trials.

Table A-10. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	4.579	5.360	7.996	3.417	4.791
Participant 2	6.888	5.309	5.341	5.560	5.077
Participant 3	6.354	5.482	5.513	5.650	4.694
Participant 4	9.873	7.008	6.996	7.722	6.680
Participant 5	9.044	6.232	6.303	6.363	5.111
Participant 6	5.080	3.472	3.470	3.860	3.623
Participant 7	9.849	8.838	8.989	8.868	6.349
Avg. 1ks	7.381	5.957	6.373	5.920	5.189
STD. 1ks	2.218	1.666	1.831	1.952	1.036
Avg. 8ks	7.385	5.900	6.189	5.920	5.071
STD. 8ks	2.216	1.701	1.701	1.952	1.193

A direct comparison was established for the two data precisions (8-bit and 16-bit) by grouping the RMSE values from tables A.9 and A.10 (Figure A.5). Equal Y axis range representation was used to aid the precision comparison, which showed different trends depending on the prediction algorithm. No important differences can be observed for the two precisions in the performance of Kernel SVR and linear regression, which provided the worse fit to the collected joint torque. Quadratic and Cubic regressions and ANN (best fit) showed similar prediction medians, with an important reduction on the variance when using the highest precision (16-bit).

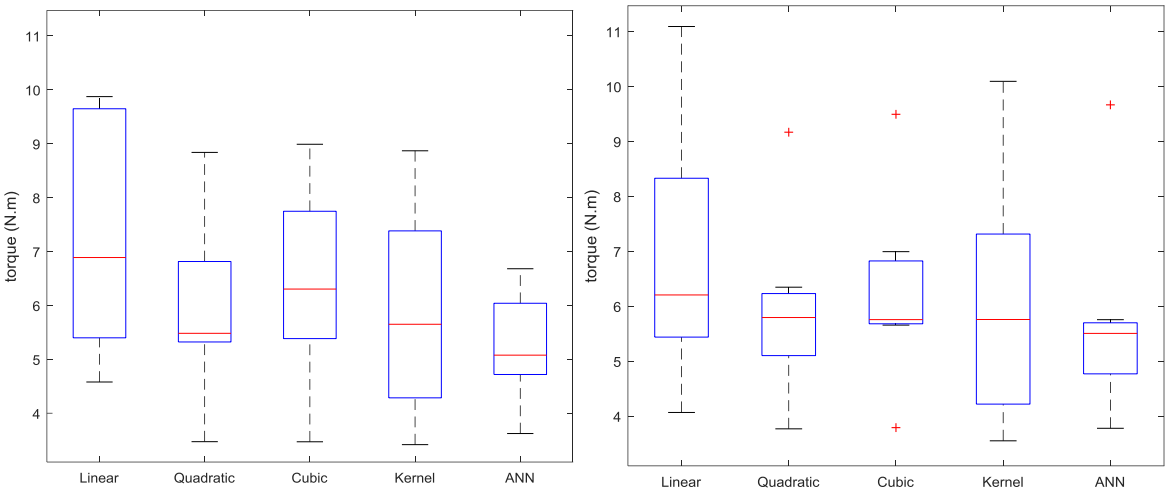

Figure A-5. Boxplot comparison of the RMSE values distribution for the 5 Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of isotonic measurements at 30% MVC from 7 participants collected at 1kS/s.

A.3.2 Isotonic Contractions at 60% MVC

Results for the predictions at 1000 fps for isotonic using 8-bit data precision during 60% MVC isotonic contractions were compiled in Table A.11, showing both individual and mean group values when using 1000 data points per trial.

Table A-11. RMSE Results for the five inference methods used for Joint Torque predictions using 8-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	7.952	9.195	8.009	8.975	7.061
Participant 2	18.297	16.599	15.608	10.841	13.944
Participant 3	7.362	7.469	5.674	6.350	5.283
Participant 4	9.289	9.634	7.840	17.472	7.151
Participant 5	15.142	12.979	12.536	12.743	10.007
Participant 6	10.338	9.589	7.355	5.948	7.394
Participant 7	5.565	5.534	4.493	9.012	5.333
Avg. 1ks	10.564	10.143	8.788	10.192	8.025
STD. 1ks	4.555	3.642	3.921	3.991	3.050
Avg. 8ks	10.590	10.337	8.859	10.192	8.751
STD. 8ks	4.122	3.965	3.593	3.991	3.346

Results show similar predictions for Linear and Quadratic regressions and Kernel SVR. On the other hand, Cubic regression and ANN provided better fits to the joint torque collected with the dynamometer, showing similar errors in terms of mean

and deviation. The best fit observed (ANN, 8.025 ± 3.05 N m) corresponded to 9.45% of the maximum torque observed for isotonic contractions at 60% MVC (84.88 N m).

In contrast, Table A.12 presents the individual participant and mean and deviation group data resulting from the analysis of 16-bit US data during isotonic contractions at 60% MVC collected at 1000 fps.

Table A-12. RMSE Results for the five inference methods used for Joint Torque predictions using 16-bit precision US data during isotonic contractions at 60% MVC. All values expressed in N m.

	Linear	Quadratic	Cubic	Kernel	ANN
Participant 1	7.522	7.465	6.563	7.341	4.849
Participant 2	17.904	13.614	13.560	9.442	11.150
Participant 3	7.857	7.446	6.736	6.164	6.858
Participant 4	8.329	6.068	5.633	12.536	5.804
Participant 5	16.810	12.796	13.255	12.982	12.157
Participant 6	13.233	14.386	12.085	8.001	10.389
Participant 7	6.936	4.847	4.083	8.889	4.262
Avg. 1ks	11.227	9.517	8.845	9.336	7.924
STD. 1ks	4.685	3.946	3.975	2.568	3.238
Avg. 8ks	11.053	9.114	8.356	9.336	7.703
STD. 8ks	3.689	1.979	2.450	2.568	2.651

Linear regression provided the worst predictions for these conditions, followed by Quadratic regression and Kernel SVR. The best results were observed with the use of ANN in terms of average, followed by Cubic regression. The best fit (7.924 ± 3.238) was equivalent to 9.33% of the maximum torque registered during these isotonic trials.

To enable a direct comparison between 8-bit and 16-bit data precisions in the predictions of joint torque data at 1000 fps, results obtained are graphically shown (Figure A.6). In order to provide proportionate information, the same joint torque value range (Y axis) was used for both plots. Again, linear regression showed the worst results for both precisions in terms of median and deviation. All inference methods showed a similar performance between precisions, with a reduction in the

variance when using 8-bit data for ANN (the best algorithm), which showed the best prediction of all studied regarding isotonic movements.

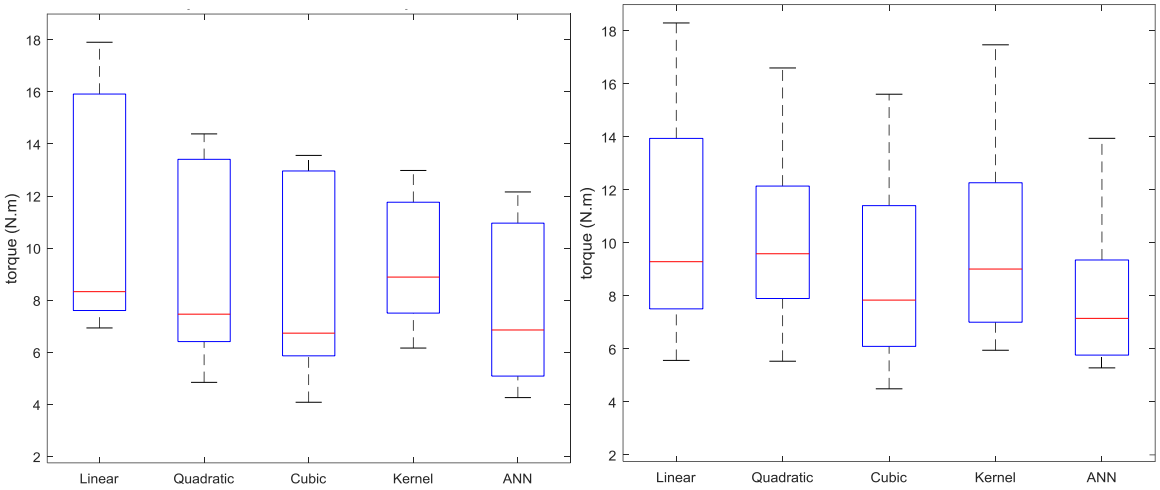

Figure A-6. Boxplot comparison of the RMSE values distribution for the 5 Inference methods studied using 16-bit raw data (left) vs 8-bit data (right). Inference values are the results of the cross-validation of isotonic measurements at 60% MVC from 9 participants collected at 1kS/s.

A.4 Discussion

The work presented in this appendix considers the influence of using a decimated version of the original data, due to computing memory limitations, in the prediction of muscle torque from US derived data.

Sections A.1 to A.3 showed the RMSE values between predictions and empirically collected data for passive movements, isometric and isotonic contractions which corresponded with the conditions studied in sections 5.4, 6.4 and 7.4 respectively.

In order to determine if there were any statistically significant differences between using 8000 or 1000 data points per trial, Wilcoxon signed rank test was applied to the resulting tables contained in this appendix. This non-parametric test assesses whether two data samples, as paired observations, are related by comparing their population mean ranks.

The resulting p-values from the application of this test to the RMSE values in tables A.1 to A.12 are compiled in Table A.13 showing the similarity in predictions between

using 1000 and 8000 samples per trial for the three muscle movements studied and distinguishing the different prediction algorithms used.

Table A-13. Wilcoxon sign rank test p-values obtained per inference methods using RMSE results from the three muscle movements studied.

	Linear	Quadratic	Cubic	Kernel	ANN
Passive	0.54006	0.887554	0.819794	1	0.832032
Isometric	0.020918	0.322281	0.949891	1	0.875162
Isotonic	0.592537	0.452378	0.33887	1	0.954601

Kernel SVR p-values were 1 in every condition as only 1000 samples per trial data was used due to memory limitations when attempting 8000 samples.

As a result, no significant differences were observed between the use of 1000 data points per trial compared with the original 8000 points for any of the conditions studied in this appendix but one, linear regression for isometric tasks. Therefore, the discussion and conclusions extracted from chapters 5 to 8 could also be extrapolated to this appendix.