PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND

*Dissertations in Forestry and Natural Sciences*


MITHILESH PRAKASH

# OPTIMIZATION OF MULTIVARIATE REGRESSION TECHNIQUES FOR NEAR-INFRARED SPECTROSCOPIC CHARACTERIZATION OF ARTICULAR CARTILAGE

UNIVERSITY OF
EASTERN FINLAND

PUBLICATIONS OF THE UNIVERSITY OF EASTERN FINLAND
DISSERTATIONS IN FORESTRY AND NATURAL SCIENCES

N:o 349

*Mithilesh Prakash*

# OPTIMIZATION OF MULTIVARIATE REGRESSION TECHNIQUES FOR NEAR-INFRARED SPECTROSCOPIC CHARACTERIZATION OF ARTICULAR CARTILAGE

ACADEMIC DISSERTATION

To be presented by the permission of the Faculty of Science and Forestry for public examination in the Auditorium SN201 at the University of Eastern Finland, Kuopio, on 20th September 2019, at 12 o'clock.

University of Eastern Finland
Department of Applied Physics
Kuopio 2019

Grano Oy
Jyväskylä, 2019
Editors: Pertti Pasanen, Jukka Tuomela,
Raine Kortet , and Matti Tedre

Distribution:
University of Eastern Finland Library / Sales of publications
julkaisumyynti@uef.fi
http://www.uef.fi/kirjasto

ISBN: 978-952-61-3160-3 (print)
ISSNL: 1798-5668
ISSN: 1798-5668
ISBN: 978-952-61-3161-0 (pdf)
ISSNL: 1798-5668
ISSN: 1798-5668

Author's address:
University of Eastern Finland
Department of Applied Physics
P.O. Box 1627
70211 Kuopio, Finland
email: mithilesh.prakash@uef.fi

Supervisors:
Professor Juha Töyräs
University of Eastern Finland
Department of Applied Physics
P.O. Box 1627
70211 Kuopio, Finland
email: juha.toyras@uef.fi

Academy Research Fellow Isaac O. Afara
University of Eastern Finland
Department of Applied Physics
P.O. Box 1627
70211 Kuopio, Finland
email: isaac.afara@uef.fi

Post-Doctoral Fellow Lassi Rieppo
University of Oulu
Research Unit of Medical Imaging,
Physics and Technology, Faculty of Medicine
P.O. Box 5000
90014 Oulu, Finland
email: lassi.rieppo@oulu.fi

Reviewers:
Professor Silvia Serranti
Sapienza Università di Roma
Department of Chemical Engineering Materials
and Environment,
Via Eudossiana, 18
00184 Roma, Italy
email: silvia.serranti@uniroma1.it

Scientist and Project Leader Rajesh Kumar
Norwegian University of Science and Technology
Biophysics and Medical Technology,
Department of Physics
N-7491, Trondheim, Norway
email: rajesh.kumar@ntnu.no

Opponent:
Professor Hugh J. Byrne
Head, FOCAS Research Institute
TU Dublin - FOCAS Research Institute
Kevin Street
D08 NF82, Dublin, Ireland
email: hugh.byrne@dit.ie

Mithilesh Prakash
OPTIMIZATION OF MULTIVARIATE REGRESSION TECHNIQUES FOR NEAR-INFRARED SPECTROSCOPIC CHARACTERIZATION OF ARTICULAR CARTILAGE
Kuopio: University of Eastern Finland, 2019
Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences
N:o 349

## ABSTRACT

Articular cartilage, the soft tissue covering the ends of articulating bones, facilitates smooth joint movements. This specialized connective tissue is *avascular* and *aneural*, with limited self-healing capabilities. Traumatic injuries to cartilage are often not apparent in current clinical diagnostics. The early detection of cartilage degeneration could potentially aid in the prescription of treatment measures that could halt the degenerative process. It could also prevent the development of post-traumatic osteoarthritis — an incurable condition associated with cartilage erosion, pain, and reduced joint mobility. Currently, the assessment and repair of joint injuries are performed during arthroscopy. However, arthroscopic evaluations rely on the visual assessment and manual palpation of the cartilage surface, so they are subjective and poorly reproducible. This feature necessitates the development of arthroscopic methods more quantitative in nature for the rapid assessment of cartilage integrity.

In the past decade, near-infrared spectroscopy (NIRS) has gained popularity as a nondestructive and rapid characterization tool for evaluating the integrity of cartilage and other joint tissues. NIRS-based evaluations rely heavily on multivariate regression analysis to relate spectroscopic measurements with tissue properties. These multivariate analysis techniques are often adopted from other spectroscopic applications and must be optimized, in terms of reliability and robustness, for cartilage data. The main limitation in the direct application of conventional regression techniques is the underlying assumption of the independence of observations. In tissue mapping and other protocols that involve repeated measurements, a spatial dependency is introduced in the data due to adjacent measurement locations. This thesis aims to provide insight into effective multivariate approaches for the analysis of cartilage spectral data, account for spatial dependency during the analysis of cartilage spectral data, and address the challenges in generalizing *in vitro* models for *in vivo* applications.

Study **I**, a comprehensive comparative study, sought to determine an optimal multivariate technique for predicting the properties of articular cartilage from its NIRS data. Partial least squares regression (PLSR), the most commonly applied technique in chemometrics, emerged as the optimal regression technique, with its performance further enhanced by variable (wavelength) selection methods. Study **II** addressed the limitations of the direct application of conventional regression techniques, such as PLSR, in experiments where adjacent measurement locations create spatial dependency. This was achieved by the development of a hybrid regression technique that accounts for repeated measures in NIRS and other spectroscopic techniques. Study **III** applied the hybrid regression technique developed in study **II** to arthroscopic evaluations of cadaveric human knee joints *ex vivo*. The hybrid models

trained on the *in vitro* measurements reliably modelled the relationship between cartilage NIRS data and its biomechanical properties. The trained models, assisted by k-nearest neighbours (kNN)-based classifiers, also reliably predicted biomechanical properties from arthroscopically acquired NIRS data. Thus, NIRS enables the quantitative evaluation during arthroscopy of the biomechanical properties of human cartilage.

In conclusion, the studies in this thesis explore and optimize regression techniques for the NIRS characterization of articular cartilage integrity. The limitations of conventional regression techniques were addressed via the development of a hybrid regression technique to aid the arthroscopic application of NIRS, thus paving the way for clinical applications in joint tissue diagnosis.

***National Library of Medicine Classification:*** *QT 34.5, QT 36, WB 288, WE 300, WE 348, WN 180*
***OCIS codes:*** *300.1030, 300.6340, 170.6510*
***Medical Subject Headings:*** *Cartilage, Articular; Osteoarthritis/diagnosis; Spectroscopy, Near-Infrared; Multivariate Analysis; Regression Analysis; Arthroscopy; Joints; Knee Joint; Collagen; Proteoglycans; Biomechanical Phenomena*
***Yleinen suomalainen asiasanasto:*** *nivelrusto; nivelrikko; lähi-infrapunaspektroskopia; monimuuttujamenetelmät; regressioanalyysi; nivelet; kollageenit; biomekaniikka*

## ACKNOWLEDGEMENTS

The studies in this thesis were conducted during 2016 to 2019 in the Biophysics of Bone and Cartilage (BBC) group at the Department of Applied Physics, University of Eastern Finland.

I am grateful for the opportunities Finland has bestowed on me to develop into a thinking individual. I have met some wonderful people (I continue to do so) who help me grow wiser by the day.

First and foremost, I would like to thank my main supervisor Prof. Juha Töyräs for trusting in my abilities to deliver on this challenging task of completing a doctoral degree. He has been a key force in educating, inspiring, and motivating, me from the early stages of my studies, enabling me to achieve the objectives of my thesis in a timely manner. He has coached me in downhill skiing, ice hockey, and even deep diving. I will cherish those moments for a long time.

I would like to thank my co-supervisors: Adjunct Prof. Isaac Afara and Senior Researcher Lassi Rieppo, both of whom have guided me in times of need and provided me with valuable feedback. I would like to commend their commitment to their supervision despite being in different cities and countries at the start of my PhD studies. I would like to extend my thanks to Jaakko Sarin for his guidance as a senior colleague and introducing me to the practicalities of research work, and for inspiring me at work and beyond.

I am grateful to the preliminary thesis Reviewers, Professor Silvia Serranti, Ph.D., and Scientist Rajesh Kumar, Ph.D., for their valuable feedback for improving this thesis. I thank Gerald G. Netto, Ph.D., for the language review.

The members of the BBC group have helped me keep my work-and-play balance. I would like to thank past and present BBC members for helping me to keep it together. A special mention to Aapo and Jari for accompanying me in navigating the bar streets of Kuopio.

I would like to thank the Academy of Finland; the Instrumentarium Science Foundation; the Research Committee of the Kuopio University Hospital Catchment Area for State Research Funding, Kuopio, Finland; and The Finnish Foundation for Technology Promotion for funding this research project.

Last but not least, I would like to thank my parents, my lovely sister, and my grandparents, for all their good wishes; you have been a blessing in my life. It would not have been easy without their support.

Kuopio, 20th September, 2019

*Mithilesh Prakash*

## LIST OF PUBLICATIONS

This thesis consists of the present review of the author's work in multivariate data analysis of near infrared spectroscopy for cartilage and the following selection of the author's publications:

**I** M. Prakash, J. K. Sarin, L. Rieppo, I. O. Afara, and J. Töyräs, "Optimal regression method for near-infrared spectroscopic evaluation of articular cartilage," *Applied spectroscopy* **71**, 2253–2262 (2017).

**II** M. Prakash, J. K. Sarin, L. Rieppo, I. O. Afara, and J. Töyräs, "Accounting for spatial dependency in multivariate spectroscopic data," *Chemometrics and Intelligent Laboratory Systems* **182**, 166–171 (2018).

**III** M. Prakash, A. Joukainen, J. Torniainen, M. Honkanen, L. Rieppo, I. O. Afara, H. Kroger, J. Töyräs, and J. K. Sarin, "Near-infrared spectroscopy enables quantitative evaluation of human cartilage biomechanical properties during arthroscopy," *Osteoarthritis and Cartilage* **27**, 1235–1243 (2019).

Throughout the thesis, these papers are referred to by Roman numerals.

## AUTHOR'S CONTRIBUTION

The publications selected in this dissertation are original research papers on multivariate analysis of near-infrared spectroscopic data of articular cartilage. In all the studies the author participated in the study design, analysis and was the principal author.

**I** The author carried out the comparison of different multivariate techniques on equine data collected from previous studies [1].

**II** The author developed the hybrid regression technique to solve spatial dependency on equine data collected from previous studies [1, 2].

**III** The author was involved in sample extraction, NIRS measurements, biomechanical measurements in collaboration with orthopaedic surgeons and other members of the research group.

# TABLE OF CONTENTS

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AI | Areas of interest |
| BiPLS | Backward interval partial least squares |
| CARS | Competitive adaptive reweighted sampling |
| CCD | Charge-coupled device |
| CT | Computed tomography |
| ECM | Extra cellular matrix |
| EDTA | Ethylenediaminetetraacetic acid |
| FCD | Fixed charged density |
| FTIR | Fourier transform infrared spectroscopy |
| GA | Genetic algorithm |
| ICRS | International Cartilage Repair Society |
| kNN | k-nearest neighbors |
| LASSO | Least absolute shrinkage and selection operator |
| LASSO-LME | LASSO based LME |
| LME | Linear mixed effects |
| LS-SVM | Least squares version of support vector machines |
| MC-UVE | Monte Carlo uninformative variable elimination |
| MIR | Mid-infrared |
| MRI | Magnetic resonance imaging |
| MSC | Multiplicative scatter correction |
| NaCl | Sodium chloride |
| NIR | Near-infrared |
| NIRS | Near-infrared spectroscopy |
| NW | Norris-Williams |
| OA | Osteoarthritis |
| OCT | Optical coherence tomography |
| PBS | Phosphate-buffered saline |
| PCA | Principal component analysis |
| PCA-LME | Principal component based LME |
| PCR | Principal component regression |
| PG | Proteoglycan |
| PLM | Polarized light microscopy |
| PLS | Partial least sqaures |
| PLSR | Partial least squares regression |
| RMSE | Root mean square error |
| RMSEC | RMSE of calibration |
| RMSECV | RMSE of cross-validation |
| RMSEP | RMSE of prediction |
| RPIQ | Ratio of performance to interquartile range |
| RSS | Residual Sum of Squares |
| SEC | Standard error of calibration |
| SEV | Standard error of validation |
| SEP | Standard error of prediction |
| SG | Savitzky-Golay |
| SNV | Standard normal variate |
| TTL | Transistor–Transistor Logic |
| UV | Ultra violet |

| | |
|---|---|
| VCPA | Variable combination population analysis |

## LIST OF SYMBOLS

| | |
|---|---|
| $A$ | Absorbance |
| $b_0$ | Scalar parameters |
| $b_{ref}$ | Scalar parameters |
| $c$ | Speed of light in a vacuum |
| $d$ | Diameter |
| $D$ | Dark spectrum |
| $\epsilon$ | Strain |
| $e$ | The error not modelled in original spectra |
| $E_{Inst}$ | Instantaneous modulus |
| $E_{Eq}$ | Equilibrium modulus |
| $E_{Dyn}$ | Dynamic modulus |
| $E_m$ | Measured/Uncorrected modulus |
| $F$ | Force |
| $h$ | Planck constant |
| $h_t$ | Tissue thickness |
| $I$ | Light intensity after traversing the sample |
| $I_0$ | Initial light intensity |
| $Ind.$ | Th number of independent variables |
| $\kappa$ | Theoretical correction factor |
| $k$ | Spring constant |
| $\lambda$ | Wavelength |
| $m$ | The number of points in the smoothing window |
| $\rho$ | Spearman's rank correlation |
| $\sigma$ | Stress |
| $R$ | Reflectance spectrum from standard |
| $R^2$ | Coefficient of determination |
| $\Re$ | Radius of the indenter |
| $S$ | Sample spectrum |
| $t_{relax}$ | Relaxation time in seconds |
| $T$ | Transmittance |
| $\mu$ | Absorption coefficient of the material |
| $v$ | Frequency |
| $\nu$ | Poisson's ratio |
| $x^{'}$ | First order derivative |
| $x^{''}$ | Second order derivative |
| $x_{avg\ spectra}$ | Average or mean spectra |
| $x_{corr}$ | Corrected spectra |
| $x_{org}$ | Original sample spectra |
| $x_s$ | Smoothed sample spectra |
| $x_{ref}$ | Reference spectrum |
| $x_T$ | Training spectral data |
| $x_V$ | Validation spectral data |
| $y_i$ | Measured value of the *i*th sample |
| $y_j$ | Measured value of the *j*th sample |
| $y_T$ | Training reference data |
| $y_V$ | Validation reference data |
| $\hat{y}_i$ | Predicted value of the *i*th sample |

| | |
|---|---|
| $\hat{y}_j$ | Predicted value of the $j$th sample |
| $y_s$ | Actual value of the $s$th observation |
| $\hat{y}_s$ | Predicted value of the $s$th observation |

# 1 INTRODUCTION

Articular cartilage is a unique connective tissue lining the ends of articulating bones. Together with synovial fluid, cartilage facilitates the smooth and near-frictionless movements of the diarthrodial joints. The main components of this tissue are water and the extracellular matrix (ECM). The ECM consists mainly of collagen, proteoglycans (PG), and chondrocytes; chondrocytes are the cells in cartilage [3]. Chondrocytes maintain and synthesize the tissue matrix, while all the other components contribute to the biomechanical properties of articular cartilage [4, 5]. Changes in the structure and composition of cartilage due to ageing (wear and tear) or trauma can upset tissue homeostasis, leading to degeneration [6, 7].

Osteoarthritis (OA), characterized by the erosion and loss of cartilage tissue, is a painful condition that results in reduced mobility, an overall reduction in the quality of life, and substantial socioeconomic burden globally [8]. Although the pathogenesis of OA is unclear, OA can stem from the injury of cartilage, subchondral bone, or the meniscus [6, 9, 10]. The biomechanical, structural, and biochemical, properties of cartilage change with the progression of OA [11–13]. Currently, joint defects are diagnosed via clinical examination followed by radiographic examination, and if required, MRI. Cartilage is invisible in native X-ray images, whereas CT and MRI lack the resolution to detect minor changes [14, 15]. During the arthroscopic examination of meniscal injuries or ligament tears, the cartilage surface is usually examined for lesions and defects. The severity of cartilage injury is diagnosed using the International Cartilage Repair Society (ICRS) scoring system, which is based on the relative depth of lesions. However, the outcome of ICRS-based diagnosis is subjective and poorly reproducible [16]. Although several repair techniques are available for the treatment of local osteochondral defects [17], knowledge of lesion severity and of spread from the site of a defect is critical for the success of repair arthroscopy [18]. Since arthroscopy is subjective, its inter- and intra-operator reproducibilities are poor [16, 19]. Therefore, arthroscopic techniques that are more quantitative could substantially facilitate the early detection of OA [20]. Earlier studies have demonstrated the potential of near-infrared spectroscopy (NIRS) for detecting changes in cartilage structure and morphology during degeneration [21–23].

NIRS is a promising technique for evaluating the integrity of cartilage [24, 25]. This method is sensitive to both micro- and macroscopic properties of the tissue. This rapid nondestructive method penetrates deep into biological tissues [26], beyond articular cartilage thickness [23, 27], permitting the simultaneous quantitative evaluation of both cartilage and subchondral bone [28, 29]. Several studies have successfully applied NIRS for the *in vitro* and *ex vivo* assessment of cartilage, via robust multivariate analytical techniques. These include principal component analysis (PCA) and partial least squares (PLS) regression, to estimate the composition, degenerative stage, and thickness of articular cartilage [21–23].

With overlapping and broad absorption bands, NIRS data is complex; therefore, multivariate regression techniques have become a standard method for analysing its spectroscopic data. Multivariate regression models can be improved through a process known as variable selection. Numerous multivariate regression techniques

and variable selection methods exist, but they are currently not optimized for the evaluation of cartilage spectral data. Furthermore, the utilization of these methods on spatially dependent data sets, such as the mapping of knee joint tissue properties, violates the assumptions of the independence of observations [30]. The adaptation of NIRS for joint diagnostics helps assess cartilage properties in animal and human joints *in vitro* and *ex vivo* [17, 21, 22, 31, 32]. However, these limited exploratory studies can only be considered as a proof of concept of the technique, necessitating the optimization of multivariate analysis methods for the accurate spectroscopic diagnosis of cartilage pathology.

Predicting cartilage tissue properties during arthroscopy is a challenging task [33]. Narrow joint spaces restrict optimal probe-cartilage contact, resulting in noisy spectra during arthroscopic acquisition. Without outlier detection and optimization (noisy spectra), predictions with *in vitro* models could be unreliable. Additionally, certain wavelengths may be saturated due to increased water content between probe and cartilage and hence become unusable. This thesis aims to provide an optimal approach for the multivariate regression analysis of cartilage spectral data; by employing novel regression algorithms, it aims to address the challenges of generalizing *in vitro* prediction models for use in arthroscopy (*ex vivo*).

In this thesis, the study **I** focused on the determination of an optimal multivariate regression technique for estimating the properties of articular cartilage from its spectral data. This comparison study utilized conventional and advanced regression techniques. Additionally, it investigated the performance of variable selection methods. PLS regression emerged as the most optimal (with consistent performance) for our cartilage dataset. Its prediction performance was further enhanced by the Monte-Carlo uninformative variable elimination method. In study **II**, a hybrid regression technique was developed to account for spatial dependency in NIRS measurements and to improve the efficiency and reliability of the model. Principal component-based linear mixed effects (PCA-LME) and least absolute shrinkage and selection operator (LASSO)-based LME were compared against standard regression models. Accounting for spatial dependency resulted in improved performance over standard regression technique models, and PCA-LME performed consistently better than LASSO-LME. In study **III**, the hybrid model developed (PCA-LME) was utilized, and a protocol was designed for the selection of the best spectra from a series of possibly noisy spectra acquired during arthroscopy from the human knee joint cartilage *ex vivo*. In addition, our novel use of the classifier reliably discarded spectral outliers, enhancing prediction performance with arthroscopic spectra.

# 2 ARTICULAR CARTILAGE

## 2.1 STRUCTURE AND COMPOSITION

Articular cartilage is the connective tissue lining the ends of the articulating bones of a joint (Figure 2.1) [6, 34, 35]. Together with synovial fluid [36], articular cartilage enables near-frictionless movements of the joints; with the meniscus [37], articular cartilage helps in distributing the load to the underlying subchondral bone [38]. The thickness of cartilage varies with location, age, gender, and species [39–42]; the typical thickness in the human knee is between 1 to 6 mm [39, 40, 43].

Conceptually, articular cartilage can be considered as a biphasic material which has a fluid phase (water and dissolved electrolytes), and a solid phase; the latter comprises chondrocytes (cells), collagen fibres (type II), proteoglycans (PG), and other glycoproteins [44, 45]. Around 60% to 87% of articular cartilage consists of interstitial water, of which 30% is in the intrafibrillar space of the collagen network. The amount of water varies with the fixed charge density (FCD), collagen orientation, and the stiffness of the collagen network (i.e. resistance to swelling) [46, 47]. Chondrocytes represent 1% to 5% of the tissue volume and are involved in the synthesis and maintenance of the components of the cartilage matrix [44]. Collagen, the most abundant protein in the body, is the key structural element providing tensile properties to the tissue and offers minimal resistance to compression [48, 49].

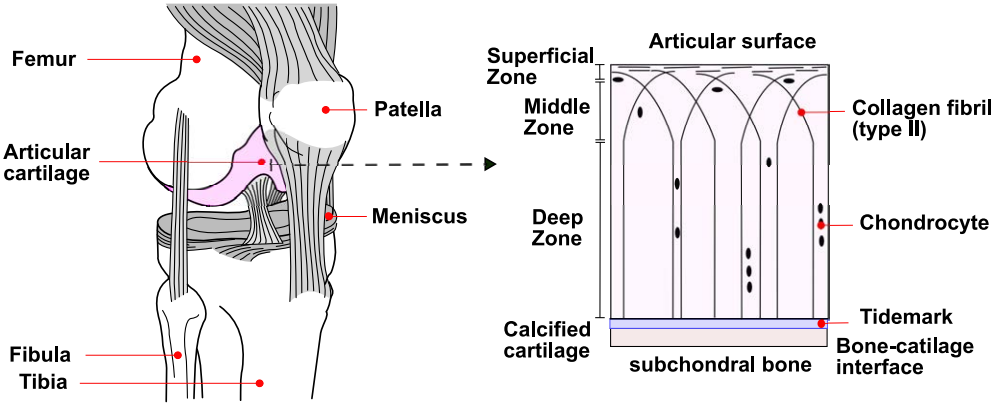


**Figure 2.1:** Human knee joint and schematic representation of articular cartilage.

## 2.2 TISSUE PROPERTIES

Mature articular cartilage is stratified into superficial, middle, deep, and calcified, zones, based on collagen orientation [50]. The structure and distribution of these constituents vary between these zones (Figure 2.1) [44]. PG (or fixed charge density, FCD) and collagen contents increase, while fluid fraction decreases from the articular surface to the bone cartilage interface. In the superficial zone (5%-10% of

cartilage thickness), collagen fibrils are densely packed (although the density is the lowest among all the zones) and are oriented parallel to the cartilage surface. In the middle zone (10%-20%), these fibrils (moderate density) bend towards the subchondral bone, with random orientation. In the deep zone (20%-90%), the fibrils (highest density) run perpendicular to the cartilage surface (Figure 2.1).

Articular cartilage can be considered in a mechanical sense as a poro-viscoelastic anisotropic material although numerous other models with varying degrees of complexity exist [51, 52]. Viscoelasticity and the interplay of fluid and solid components have time-dependent properties that can be observed as flow-dependent and flow-independent behaviours (Table 2.1) [53]. Flow-dependent behaviour is characterized by the frictional flow of the interstitial fluid [54], whereas flow-independent time-dependent behaviour is characterized by the intrinsic viscoelasticity of the matrix [55].

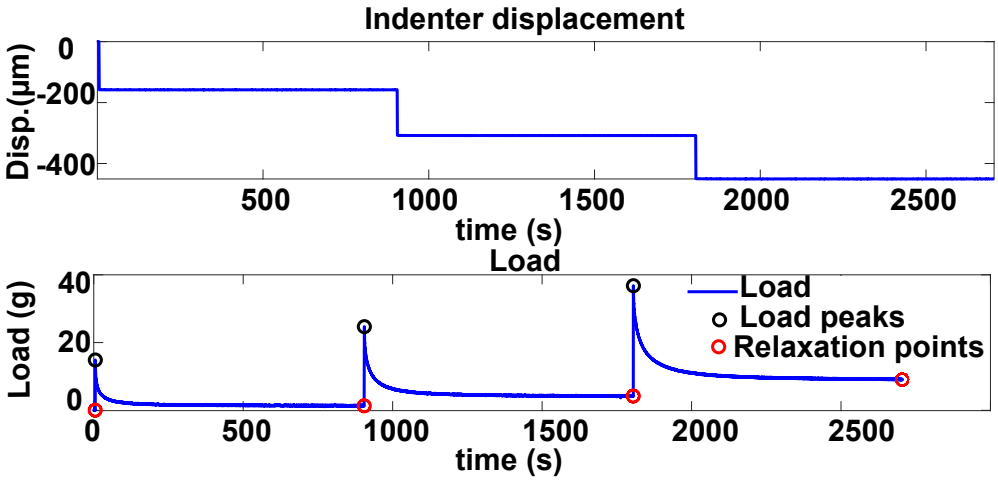


**Figure 2.2:** Stress-relaxation response of human knee articular cartilage

**Table 2.1:** Cartilage response.

| *Loading* | *Response* |
|---|---|
| No load | Tissue matrix swell due to osmotic pressure. The tissue expansion volume is controlled by collagen [53] . |
| Compression | A transient response where fluid flows out of tissue. The dynamic response is due to high fluid pressure and collagen tension. The static response is due to proteoglycans (PGs) [54, 55]. |

Under constant compressive force, articular cartilage behaves as a porous material and the interstitial fluid flows rather easily. The matrix [56], however, regulates the fluid flow; at mechanical equilibrium, the fluid flow ceases. At the same time, resistance starts to build up against further compression and is mainly controlled by the solid matrix [57]. When the tissue is unloaded, the fluid flows back into the matrix, regaining the original volume. This flow-dependent behaviour is observed

in the standing position (i.e. loading) for a certain time and then resting (i.e. unloading). This stress-relaxation response of articular cartilage can be measured in laboratory conditions, for example by indentation testing (Figure 2.2). However, under dynamic loading, fluid flow plays almost no role in the mechanical response. This is because the interstitial fluid has no time to escape, so the tissue becomes pressurized and carries the load applied. Therefore, the structure also plays an important role in the mechanical response of the tissue (Table 2.2).

**Table 2.2:** Structure and tissue property of cartilage [3].

| *Structure* | *Tissue property* |
|---|---|
| Collagen | Dynamic and tensile properties.<br>Poisson's ratio.<br>Permeability.<br>Non-linearity in compression-tension. |
| Proteoglycans (PGs) | Equilibrium modulus.<br>Permeability.<br>Tissue swelling. |
| Interstitial fluid | Permeability.<br>Instantaneous response.<br>Dynamic response. |

## 2.3 OSTEOARTHRITIS: DEVELOPMENT, DIAGNOSIS, AND TREATMENT

Osteoarthritis (OA) is a painful joint disease affecting hundreds of millions of people globally [58]. OA could arise due to ageing (wear and tear, fatigue) or trauma [59, 60]. The pathogenesis of OA is still unclear [61, 62]. However, injuries to cartilage, menisci or supporting ligaments, impair joint mechanics and can lead to post-traumatic OA [63]. Hence, the early detection of cartilage degeneration could enable clinicians to immediately prescribe appropriate treatment or therapy.

Articular cartilage can be considered as a primitive tissue as it is devoid of any connection to the nervous or circulatory system; hence, it has a limited self-healing capability [45]. In the early stages of OA [64], the superficial zone of cartilage experiences loss of PGs and disruption of the collagen network. This results in decreased tensile and compressive properties (i.e. softening) of the tissue, making it susceptible to further damage. Surface fibrillation and tissue swelling are noticeable at this stage. In the second stage, chondrocytes in the tissue matrix respond by clearing the damaged matrix and increasing the synthesis of PGs and collagen. If these chondrocytes fail to restore the homeostatic balance, the disease progresses to the third stage [65]. The decrease in the activity of chondrocytes results in the rapid loss of PGs and increased fibrillation of the articular surface [66]. Eventually, the cartilage layer is fully eroded, exposing the underlying bone.

Osteoarthritic changes begin to affect the whole joint within a timeline of a few months to a few decades [67, 68]. Bone remodelling occurs to compensate for changes in loading and increasing wear and tear in other joint tissues, thus leading to stiffer joint movements, swelling, and pain. OA most frequently occurs in

hand, knee, and hip, joints. Thus, OA reduces the quality of life and impacts a person's productivity and in turn that of the society [69].

Clinically, joint diagnosis is based on physical examinations; severity is gauged by swelling, pain, and impaired joint movements [70, 71]. The initial diagnosis is confirmed by either X-ray or MRI [72, 73]. Since cartilage is invisible in conventional X-ray images, diagnosis is based on the joint space narrowing and increased density of the subchondral bone. However, these indirect observations can only be detected in the later stages of OA [74]. While MRI has excellent tissue contrast that enables the evaluation of cartilage health [75], its drawbacks are the cost and relatively low image resolution [76]. Hence, the initial signs of OA, i.e. cartilage fibrillation, cannot be observed.

**Table 2.3:** Summary of ICRS scores [77].

| *Score* | *Criterion* |
|---|---|
| ICRS 0 | Normal intact cartilage and no surface defects. |
| ICRS 1 | Surface fibrillation and/or softening of the surface, tissue swelling and fissures. |
| ICRS 2 | Extended tissue defects up to <50% of cartilage thickness. |
| ICRS 3 | Defects extend >50% of cartilage thickness but does not reach the subchondral bone. |
| ICRS 4 | Defects expose the subchondral bone. |

The early diagnosis of cartilage degeneration is important to enable the initiation of appropriate repair or therapy; cartilage has limited self-healing properties [78]. The reduction of possible risk factors that lead to cartilage damage, such as being overweight, muscle weakness, and repetitive and intense loading, is possible [79]. The reduction of some others, such as hereditary, gender, or ageing, factors is unfeasible [80]. OA symptoms can be managed by taking anti-inflammatory and analgesic drugs to relieve pain [81], while intra-articular hyaluronic injections may be effective in aiding joint movements, but the overall results are unconvincing [73].

Currently, OA has no cure, but drugs for altering the disease and slowing the rate of progression are continuously under development [82]. Surgical interventions, such as mosaicplasty or autologous chondrocyte implantation, have been developed for the repair of injured cartilage [83, 84]. But surgical interventions are expensive and, hence, not affordable for all. Future developments of existing drugs and surgical intervention, and monitoring the progress of these therapies, would require effective diagnostic methods.

Clinically, joint tissue repairs for ligament and meniscal tears are conducted via arthroscopic surgeries. During such surgeries, cartilage surfaces are also examined for lesions. Unfortunately, this examination is highly qualitative and subjective, due to the use the of visual inspection and manual palpation of the cartilage surface [85, 86]. Furthermore, arthroscopic evaluation suffers from both intra-observer and inter-observer subjectivity [87–89]. Cartilage injuries observed in arthroscopy are graded according to the International Cartilage Repair Society (ICRS [77], Table 2.3) grading system. The current diagnostic scenario is ineffective, and 75% arthroscopists consider the application of quantitative techniques during arthroscopy [90]. Near-infrared spectroscopy (NIRS) has shown promising results for evaluating early

changes in cartilage structure and composition; it has also been utilized in monitoring the progress of surgical interventions in animal models [28, 91]. Although NIRS could aid orthopaedic surgeons during arthroscopy, it relies heavily on the predictions of multivariate regression techniques; these techniques are currently not optimized for cartilage data and perform sub-optimally outside laboratory conditions. This thesis focused on optimizing regression techniques for the accurate estimation of cartilage properties from its spectral data and designing protocols for transitioning the technique from *in vitro* to *ex vivo* applications.

# 3 NEAR-INFRARED SPECTROSCOPY

Near-infrared spectroscopy (NIRS) is a rapid and non-destructive tool for characterizing a wide range of homogeneous and heterogeneous materials [92–94]. NIRS is based on the vibrational and rotational transitions of atoms or molecules at ambient temperatures, in the wavelength region of 700 to 2500 nm (Figure 3.1). NIRS has good tissue penetration depth ($\geq$ 5mm [95]) and minimal need for sample preparation. With advancements in computational power and multivariate regression techniques, NIRS has been gaining popularity in several fields; examples include pharmaceutics, agriculture, and biomedical engineering [94, 96].

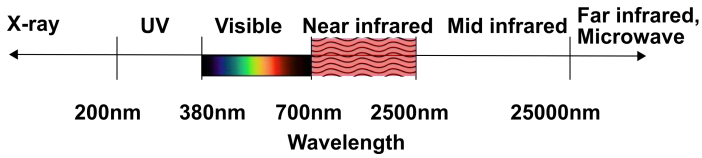


**Figure 3.1:** The electromagnetic spectrum.

## 3.1 THEORY

Optical spectroscopy utilizes light or photons to irradiate the molecular bonds of the sample causing a shift in energy levels from a ground state ($v = 0$) to an excitation state ($v \geq 1$) (Figure 3.2). The energy of a photon is given by the following equation:

$$Energy = hv = \frac{hc}{\lambda}, \tag{3.1}$$

where $h$ is Planck's constant, $v$ the frequency, $c$ the speed of light, and $\lambda$ the wavelength. A transition from $v = 0$ to $v = 1$ is called a fundamental transition, and transitions from $v = 0$ to $v > 1$ are overtones; at higher temperatures, transitions from $v = 1$ to $v > 1$ called hot bands occur. The mid-infrared (MIR) region excites fundamental molecular vibrations; and near-infrared (NIR) spectral incidence on a sample results in the stretching and bending of molecular bonds, giving rise to overtones (Figure 3.3 and Table 3.1). Molecular bonds can be approximated by a spring model, and any stretching or compression action results in an equal and opposite force. Hooke's law represents this force developed by the spring thus:

$$F = kx, \tag{3.2}$$

where $F$ is the spring force of the bond, $k$ the spring constant, and $x$ the distance between nuclei. Calculations of the vibration frequencies of overtones requires understanding quantum mechanics (harmonic and anharmonic oscillation models [97]) and are out of the scope of this thesis. Multiple overtones occurring simultaneously

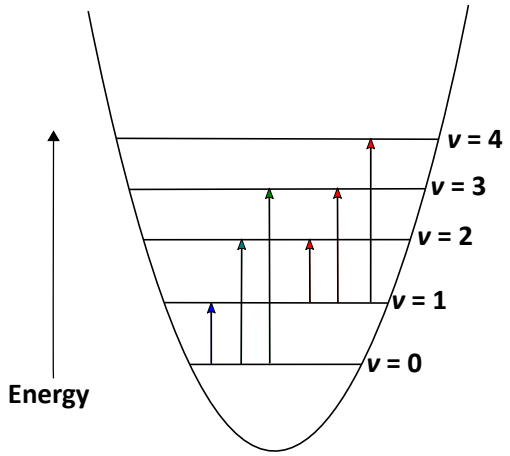


**Figure 3.2:** Energy levels for a vibrating molecule. Transition from $v = 0$ to $v = 1$ is fundamental, transitions from $v = 0$ to $v > 1$ are overtones and transitions from $v = 1$ to upper levels are called hot bands.

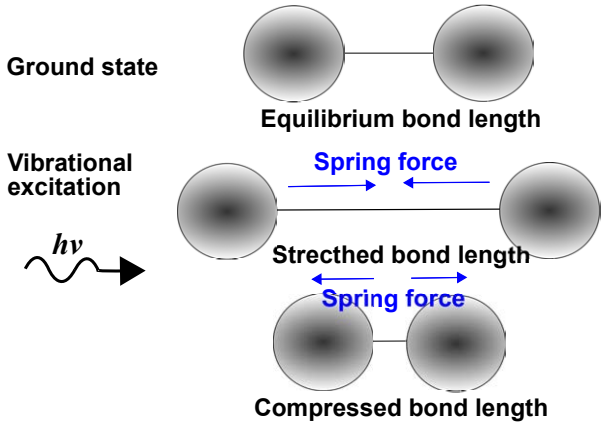


**Figure 3.3:** Vibrational stretching mechanisms. 2D diagram captures the change in bond length but also changes in bond angles occur.

give rise to combination bands. Hence unlike MIR, NIR spectral energies lack signature peaks relating to a particular molecular band, rendering them hard to analyse.

In this thesis, near-infrared absorption was measured in reflectance mode. The intensity of light beam traversing the sample is altered due to the scattering and absorption in the sample (equation 3.3). The absorbance values are calculated using Beer-Lambert's law as follows: (equation 3.4).

$$I = I_0 10^{-\mu l}, \tag{3.3}$$

$$A = -\log_{10} T = -\log(\frac{I}{I_0}). \tag{3.4}$$

where $I$ is the intensity of the light after traversing the sample, $I_0$ is the initial in-

**Table 3.1:** Near-infrared bands of interest [97].

| *Wavelength range(nm)* | *Assignment of bond stretch* |
|---|---|
| 780–850 | $3^{rd}$ overtone N–H |
| 850–950 | $3^{rd}$ overtone C–H |
| 950–1100 | $2^{nd}$ overtones of N–H and O–H |
| 1100–1230 | $2^{nd}$ overtone C–H |
| 1300–1420 | Combination C–H |
| 1400–1550 | $1^{st}$ overtones of N–H and O–H |
| 1650–1800 | $1^{st}$ overtone C–H |
| 1900–2000 | $2^{nd}$ overtones of O–H bending and C=O |
| 2000–2200 | Combination N–H stretching, combination O–H , $2^{nd}$ overtone N–H bending |

tensity of light, $\mu$ the absorption coefficient of the material, and $l$ is the distance traversed through the material. $I$, $I_0$, $\mu$, $A$, $T$, are wavelength ($\lambda$)-dependent parameters.

## 3.2 INSTRUMENTATION

The main instrumentation in a modern NIR reflectance spectroscopy consists of a light source, an optical probe, and spectrometer(s). Fibre-optic cables transmit light from the light source to the sample and conduct the scattered light back from the sample to the spectrometer.

The spectrometer (Figure 3.4) utilized in this thesis is commonly used in reflectance spectrometry and is based on the Czerny-Turner configuration. The scattered light (Figure 3.4) (from the sample) collected from the probe is passed through a narrow slit and then onto a collimating mirror, a reflection grating, a focusing mirror, and finally onto an electronic detector. The resolution of the system is controlled by the slit size and different grades of reflectance grating. The resolution of the system can be increased by narrowing the slit and increasing the number of lines on the grating. A narrowed slit increases signal loss; however, this can be compensated by longer exposure time. Hence, the resolution of the system is inversely proportional to exposure time. Real-time applications, such as *in vivo* arthroscopy, require fast acquisition times. As faster acquisition times increase noise in the spectra, finding a balance between the two is important.

Commonly employed light sources in NIRS instrumentation include tungsten-halogen and xenon lamps; they emit light covering ultraviolet (UV), visible, and NIR, spectra (Figure 3.1). Clinical applications, however, require the filtering out of harmful UV range, for safety reasons.

To minimize inherent instrumentation noise and to scale absorbance values, dark and reference spectra are acquired prior to sample measurements. A dark spectrum is acquired by blocking the light source to the spectrometer, enabling the measurement of background electrical noise. Next, the light source is unblocked, and a

reference spectrum is measured from a high reflectance standard (~100% e.g., Spectralon, SRS-99, Labsphere Inc., North Sutton, USA). The absorbance values are then calculated as:

$$A = -\log_{10}(\frac{S-D}{R-D}), \tag{3.5}$$

where $S$ is the sample spectrum, $D$ the dark spectrum, and $R$ the reference spectrum from a reflectance standard. The absorbance value at each wavelength creates a spectrum (Figure 3.4).

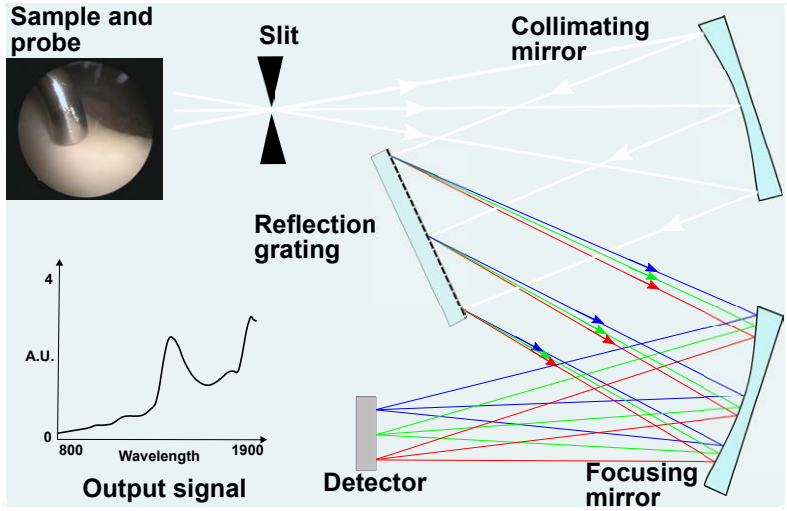


**Figure 3.4:** Schematic diagram of a conventional reflectance spectrometer. Optical fibres housed in the probe transmit light into a sample from which the scattered and reflected light is collected and transmitted to the spectrometer. Inside the spectrometer, light is channelled through a slit, followed by reflections from a collimating mirror, a reflection grating, a focusing mirror and on to a detector.

## 3.3 PREPROCESSING METHODS

Preprocessing of the absorbance spectrum is an important step in NIRS applications [98]. This step removes unwanted features of the signal, improving the performance of the calibration model. Most preprocessing techniques that are popular can be grouped into scatter-correction and spectral derivative methods.

Scatter-correction methods are designed to eliminate the variability in the spectra due to scattering effects. These methods include de-trending, standard normal variate (SNV), and normalization; and multiplicative scatter correction (MSC), inverse MSC, extended MSC, and other MSC variants. MSC is a two-step method. First, the correction coefficients from additive and multiplicative contribution are estimated (Equation 3.6). Then spectra correction is applied from the coefficients estimated (Equation 3.7).

$$x_{org} = b_0 + b_{ref,1} * x_{ref} + e, \tag{3.6}$$

$$x_{corr} = \frac{x_{org} - b_0}{b_{ref,1}} = x_{ref} + \frac{e}{b_{ref,1}} \tag{3.7}$$

where $x_{org}$ refers to original sample spectra, $x_{ref}$ is a reference spectrum utilized for preprocessing all the data, and $e$ the error not modelled in $x_{org}$; $x_{corr}$ is the corrected spectra, and $b_0$ and $b_{ref,1}$ are sample specific scalar parameters.

Similarly, SNV first centres and scales each spectrum to correct the interference from light scatter (Equation 3.8). The advantage of MSC is that the entire spectra are related to a common reference spectrum. If the reference is free from noise, then MSC is a good choice; otherwise, SNV should be considered.

$$x_{corr} = \frac{(x_{org} - x_{avg\ spectra})}{Standard\ Deviation\ of\ spectrum} \tag{3.8}$$

Spectral derivative methods are designed to eliminate both additive and multiplicative contributions of the noise in the spectra. The first derivatives eliminate the baseline effects while the second derivatives eliminate both additive and multiplicative effects. 'Finite differences' is a basic method in spectral derivation; the estimation of the first derivative (Equation 3.9) is based on two adjacent spectral measurement points, and that of the second derivative (Equation 3.10) is based on two measurement points on the first-order derivative of the spectra.

$$x_i^{'} = x_i - x_{i-1}, \tag{3.9}$$

$$x_i^{''} = x'_i - x'_{i-1} = x_{i-1} - 2 * x_i + x_{i+1} \tag{3.10}$$

where $x_i^{'}$ is the first order derivative and $x_i^{''}$ the second order derivative at an arbitrary point $i$.

Spectral derivation groups of interest include Norris-Williams (NW) derivatives and Savitzky-Golay (SG) polynomial derivatives. NW is a basic derivative method designed to avoid noise amplification experienced when using the 'finite differences' method. First, by means of a specific window size, it smoothes spectra (Equation 3.11). Next, the first-order derivative is performed on two smoothed values over a

defined interval size between the points (Equation 3.12). The second-order derivation is performed on two times the value at a point $i$ and smoothed values on either side of $i$ over a defined interval distance, as in Equation 3.13.

$$x_{s,i} = \frac{\sum_{j=-m}^{m} x_{org,i+j}}{2m+1}, \tag{3.11}$$

$$x_i^{'} = x_{s,i+interval} - x_{s,i-interval}, \tag{3.12}$$

$$x_i^{''} = x_{s,i-interval} - 2 * x_{s,i} + x_{s,i+interval} \tag{3.13}$$

where $m$ is the number of points in the smoothing window centred at the point $i$.

Similar to NW, SG includes a smoothing step and derivatives at the centre point $i$; a polynomial is fitted in a symmetric window on the raw data. SG is preferred when the peaks in the spectra are defined by a few points because these peaks are not smoothed out completely.

## 3.4 SPECTRAL CHARACTERIZATION OF CARTILAGE PROPERTIES

Absorption peaks in the NIR spectrum of cartilage arise from vibrations of different molecular bonds in the tissue matrix, including O-H, N-H, C-H, and S-H, bonds. These peaks are affected by the attenuation of light intensity as a result of energy loss in path length, which in turn is due to the thickness of the sample [27]. The aforementioned chemical bonds correspond to the main building blocks of cartilage (i.e., water, PG, and collagen); they can, therefore, characterize cartilage structure and composition [99]. However, the direct visual interpretation of cartilage NIR spectral data to determine tissue properties is not a straightforward task, due to multicollinearity and highly overlapping bands in the NIR spectral range.

Initial NIRS studies on cartilage showed correlation between mean absorbance values in the spectral ranges 1150–1220 nm and 1340–1475 nm and a degenerative condition in cartilage, such as the type of lesion or grade of injury (for example, OA or ICRS grade) [16, 17]. This technique of peak intensity analysis or analysing area under the curve is called univariate analysis. Univariate analysis fails to capture information from wider wavelength regions. The quantitative assessment of cartilage can be substantially improved by utilizing information from the broad spectral range, through techniques known as multivariate regression. For a summary of NIRS studies, see Table 3.2.

**Table 3.2:** Earlier NIRS studies on cartilage tissue relevant to this thesis.

| *Period* | *Study* | *Property* | *Source* | *Model* |
|---|---|---|---|---|
| 2008-13 | Spahn et al [31], Padalkar et al [100] | Water content | Ovine, bovine | Univariate, PLSR |
| 2008-18 | Marticke et al [101], Spahn et al [31], Afara et al [102–104] and Sarin et al [105] | Biomechanical | Ovine, bovine, hominine, equine | Univariate, PLSR |
| 2010 | Baykal et al [106] | Collagen content | Bovine | PLSR |
| 2012-15 | Afara et al [21, 27, 103] | Thickness | Murine, bovine, hominine | PCA, PLSR |
| 2015 | Afara et al [103, 107] | PG content | Bovine, hominine | PCA, PLSR |

# 4 MULTIVARIATE REGRESSION TECHNIQUES

Multivariate regression techniques comprise methods that establish the relationship between the property of a sample ($Y$- reference variable) and the spectral data of that sample ($X$- explanatory variable). The three main steps in analysing NIRS spectral data are preprocessing, qualitative analysis, and regression. Preprocessing aims to improve the signal by removing unwanted features and to simplify the spectral data (discussed in section 3.3). The qualitative analysis aims to classify spectra based on patterns recognized via supervised and unsupervised learning approaches. Quantitative or regression methods aim to estimate or predict the reference properties of a sample from its spectral data. (Figure 4.1).

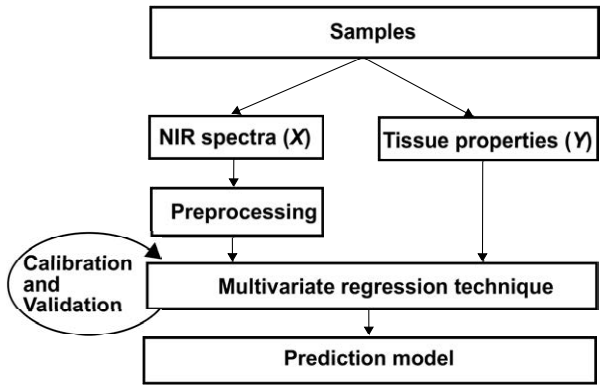


**Figure 4.1:** Typical workflow of multivariate regression analysis. [92]

## 4.1 CALIBRATION AND VALIDATION

NIRS prediction models are trained to replace destructive methods of determining sample properties ($Y$); it utilizes spectroscopic features ($Y$), such as spectral intensities or absorbances. The model development process comprises of the following steps [108]:

1. The calibration model is developed by using a training set ($x_T, y_T$) and validated by using a validation set ($x_V, y_V$). The standard error of validation ($SEV$) is used to further tune model parameters.

2. To assess the baseline model performance in ideal conditions, the standard error of calibration ($SEC$, equation 4.1) is computed from ($x_T, y_T$) and ($x_V, y_V$) datasets.

3. Finally, to evaluate the real-world performance of the model, the standard error of prediction ($SEP$, equation 4.2) is computed on an independent dataset.

In practice, the first and second steps are performed together using a cross-validation protocol (examples are k-fold, randomization, and bootstrap), so *SEC* and *SEV* are computed simultaneously. During cross-validation, spectral data ($X$) and reference properties ($Y$) are partitioned into calibration datasets (for example ~75%) and validation datasets (for example ~25%). The performance of a model is often assessed by using error metrics, such as the *SEC*, root mean square error of cross-validation (*RMSECV*), *SEP* and correlation coefficient ($R$). This coefficient and these ratios are calculated as follows:

$$SEC = \sqrt{\frac{1}{N_{cal.} - 1 - Ind.} \sum_{i=1}^{N_{cal.}} (\hat{y}_i - y_i)^2} \tag{4.1}$$

$$SEP = \sqrt{\frac{1}{N_{pred.} - 1} \sum_{j=1}^{N_{pred.}} (\hat{y}_j - y_j)^2} \tag{4.2}$$

$$RMSECV = \sqrt{\frac{1}{N_{cal.}} \sum_{i=1}^{N_{cal.}} (\hat{y}_i - y_i)^2} \tag{4.3}$$

$$R^2 = 1 - \frac{\sum_{s=1}^{n} (y_s - \hat{y}_s)^2}{\sum_{s=1}^{n} (y_s - \bar{y})^2} \tag{4.4}$$

where $\hat{y}_i$ and $\hat{y}_j$ are the predicted values of the $i$th sample in the calibration and $j$th sample in the prediction sets, $y_i$ and $y_j$ are the measured values of the $i$th sample in the calibration set and $j$th sample in the prediction sets, $N_{cal.}$, $N_{pred.}$ are the sample count in the calibration and prediction datasets, and *Ind.* is the number of independent variables in the regression. $y_s$ and $\hat{y}_s$ denote actual and predicted of the $s$th observation in the datasets including calibration, validation and testing, $\bar{y}$ denotes the mean value of the measured data. In general *SEC* decreases as $R^2$ increases, $R$ is always greater than $R^2$, $0 \leq R^2 \leq 1$, and $0 \leq RMSEC$.

## 4.2 MULTIVARIATE REGRESSION TECHNIQUES

Principal component regression (PCR) and partial least squares regression (PLSR) are common multivariate regression techniques in NIRS [92]. Both these techniques create new independent variables called components (or latent variables) that are linear combinations of the original data ($X$). PCR utilizes only spectral data ($X$) while PLSR utilizes both spectral data ($X$) and tissue property ($Y$).

Regularization regression methods, such as ridge regression and least absolute shrinkage and selection operator (LASSO), and the least squares version of support vector machines (LS-SVM) have also been recently used in NIRS applications [109–113]. Regularization solves the overfitting problem, by adding a penalty term to the objective function, thereby controlling overall model complexity. Both ridge regression and LASSO regression are applicable to multicollinear datasets; LASSO is more computationally efficient. While LASSO and ridge regression are linear models, non-linear relationships between ($X$) and ($Y$) can be modelled using support vector regression. For a summary of these regression techniques and their corresponding hyperspace, see Table 4.1 and Figure 4.2.

**Table 4.1:** Summary of regression techniques

| *Technique* | *Summary* | *Advantage(s)* | *Disadvantage(s)* |
|---|---|---|---|
| PCR | Linear projection method, reduces the dimensionality of the data using only explanatory data, X, into uncorrelated subspace. Latent variables are regressed using ordinary least squares. | Dimensionality reduction and handles multicollinearity in $X$. | Latent variables are based only on the variance of X while ignoring the variance of $Y$. |
| PLSR | Linear regression technique based on reducing dimensionality by projecting explanatory data, $X$, to a subspace of latent components maximizing covariance between $X$ and $Y$. | Dimensionality reduction and handles multicollinear in $X$. | The output is a linear combination of input. |
| Ridge | Shrinkage regression technique. Dimensions with the least variance are shrunk the most. | Stable when $P \gg N$ *. | Selects all predictors in the final model instead of a subset of variables. |
| LASSO** | Shrinkage regression technique based on minimizing the sum of squared error and setting some coefficient estimates to zero. | The solution is sparse so computationally efficient. | Covariate selection is arbitrarily done if the dataset is highly collinear. |
| LS-SVM** | Least squares version of the support vector variant. Creates a model based on newly formed support vectors from the training dataset. | Can also model non-linear relationships. | Lack of sparseness. |

* $P$ is the dimension of the dataset of $N$ observations.
** LASSO: least absolute shrinkage and selection operator, LS-SVM: least squares version of support vector machines.

## 4.3 OPTIMIZING REGRESSION MODELS

After the preprocessing step, regression model performance may be improved by variable (i.e., wavelength) selection and dimension reduction methods. Variable selection methods focus on selecting wavelengths that best predict tissue property while eliminating redundant wavelengths or by eliminating wavelengths that fail to improve the performance of the model.

Typical variable selection methods used for analysing NIRS data include Monte Carlo uninformative variable elimination (MC-UVE), competitive adaptive reweighted sampling (CARS), variable combination population analysis (VCPA), backward interval PLS (BiPLS), genetic algorithm (GA), and jackknife [114–119]. For a summary of these methods, see Table 4.2. Alternatively, to reduce high dimensional data, PCA and LASSO can also be highly effective. Dimension reduction can also improve the computational efficiency of a model and make the results easier to interpret via

visualizations (Figure 4.2).

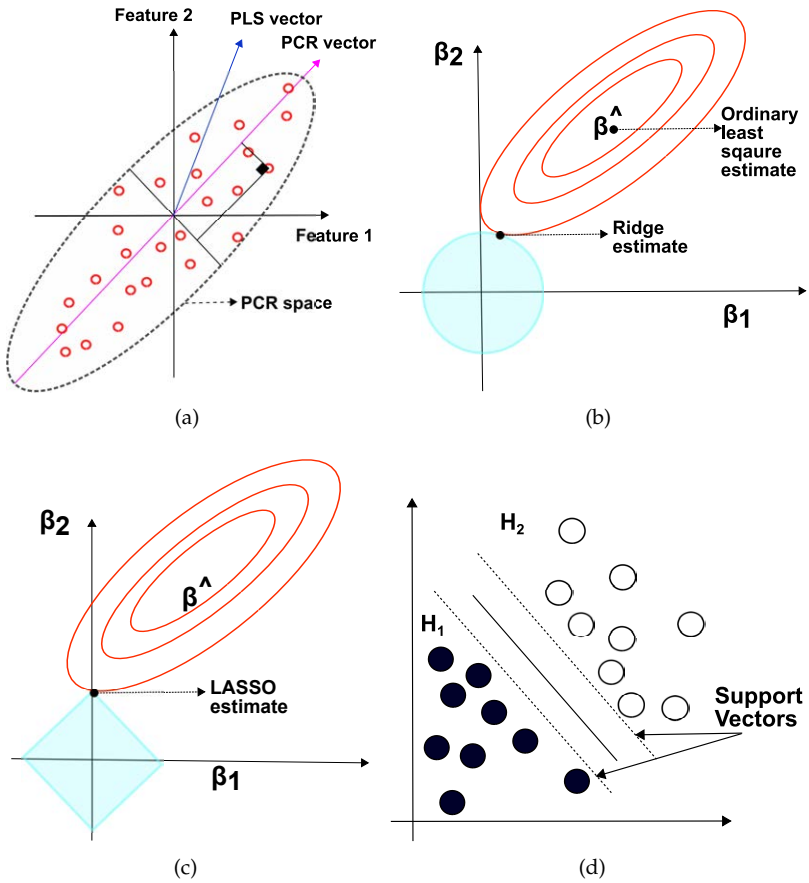


**Figure 4.2:** Hyperspace of regression methods: (a) Difference between PCR and PLS vectors in orthonormal transformed space [120]; (b) Ridge and (c) LASSO estimates with contours of error and constraint functions. The circle and rhombus shaded regions around the origin are $\beta_1^2 + \beta_2^2 \leq t^2, |\beta_1| + |\beta_2| \leq t$;, (d) SVM boundaries in $H_1$ and $H_2$ hyperplane space [121].

**Table 4.2:** Summary of variables selection methods

| *Method* | *Summary* | *Advantage(s)* | *Disadvantage(s)* |
|---|---|---|---|
| GA | A method based on genetics and natural selection principles. An initial population (or wavelengths) are randomly chosen ($t$), the fitness of the population is measured. New parents from the population are selected, crossover to the existing population ($t+1$), perform mutation ($t+1$) and determine the fitness of the population. This process is repeated and the best population is chosen. | Automation possible, a combination of GA and PLS performs better than standalone PLS regression. | Complicated parameters, stochastic and computationally heavy. |
| CARS | Follows 'survival of the fittest' theory, removing unused variables and reduces collinear effects of modelling variables. Each variable is treated as individual and variables with large PLS coefficients are retained by adaptive reweighted sampling technique. | Simple and robust. | Selection based on regression coefficients may not be an optimal method. |
| VCPA | VCPA utilizes exponentially decreasing function to eliminate variables with little or no contribution thereby shrinking the variable space. | Sparse variable selection method. | Biased towards retaining very few variables. |
| BiPLS | The data is subdivided into non-overlapping intervals and PLS models are created by leaving one interval out. Poorest performing interval in RMSECV values is omitted. | Simple to implement. | Low impact on modelling performance. |
| MC-UVE | A large number of models are calibrated with a random selection of wavelengths. The coefficients of the models are utilized to assign a stability index to each wavelength. Wavelengths above certain stability index are chosen. | Easy to use, computationally fast, decreased overfitting. | Performance improvements are not significant. Selection based on regression coefficients. |
| Jack-knife | Student's t-test statistics is applied for selecting variables within determined a threshold (e.g. $t = 0.5$). | Simple method resembling cross-validation. | Ineffective method. Selection based on regression coefficients may not be an optimal method. |

* GA: Genetic algorithm; CARS: Competitive adaptive reweighted sampling;
VCPA: Variable combination population analysis; BiPLS: Backward interval selection methods;
MC-UVE: Monte Carlo uninformative variable selections.

## 4.4 LIMITATIONS OF CURRENT REGRESSION PROTOCOLS

Although multiple options exist for multivariate regression and variable selection, they are not all optimal for the analysis of cartilage NIR spectral data (Tables 4.1, 4.2). Hence, a comparative study is necessary to investigate techniques most suited for cartilage spectral data. Furthermore, the adaptation of NIR spectroscopy for the evaluation of cartilage integrity, such as in tissue mapping applications, often involves, due to experimental design, adjacent measurement locations. Measurements from adjacent locations would enable the assessment of the spread of lesion on the cartilage surface; this evaluation can be used to determine the area of damaged cartilage that needs replacement. Closely spaced measurements can also be useful for the post-surgery evaluation of repair success and during follow-up after reconstructive surgeries. NIRS spectral data collected in this manner are spatially dependent [30]. Conventional regression techniques, such as PLSR and PCR, assume independence of observations; hence, models developed using these methods are less reliable unless spatial dependency is accounted for. Additionally, models developed in a controlled environment (*in vitro*) must be modified to work with the real-time collection of NIRS spectra data (*ex vivo* or *in vivo*) where suboptimal probe-cartilage contact often results in noisy spectral data. This thesis aims to address these challenges.

# 5 AIMS OF THE THESIS

Near-infrared spectroscopy-based evaluation of articular cartilage health has gained popularity in recent years. However, current multivariate regression techniques utilized in the development of prediction models need to be optimized for the reliable evaluation of cartilage properties for both *in vitro* and *in vivo* applications. To address these challenges in the clinical applications of this optical technique, the following are the aims of this thesis:

- Determining an optimal multivariate regression technique for the analysis of cartilage properties, using NIRS spectral data.
- Accounting for spatial dependency to satisfy statistical assumptions in cartilage mapping during arthroscopy.
- Optimizing the real-time selection of NIRS spectral data by employing spectral classifiers and hybrid regression techniques on NIRS spectral data from *ex vivo* human cadaver knees.

# 6 MATERIAL AND METHODS

This thesis consists of three independent studies. In studies, **I** and **II**, equine data (collected from an earlier study) consisting of NIRS spectral data, thickness, and biomechanical, structural, and compositional, values, were utilized [1, 2]. The samples were obtained from a slaughterhouse in Utrecht, Netherlands; hence, no ethical permissions were necessary. For study **III**, human cadaver samples were utilized, and ethical permission was obtained from the local research ethics committee (Decision number 150/2016, Research Ethics Committee of the Northern Savo Hospital District, Kuopio University Hospital, Kuopio, Finland), prior to the commencement of the study.

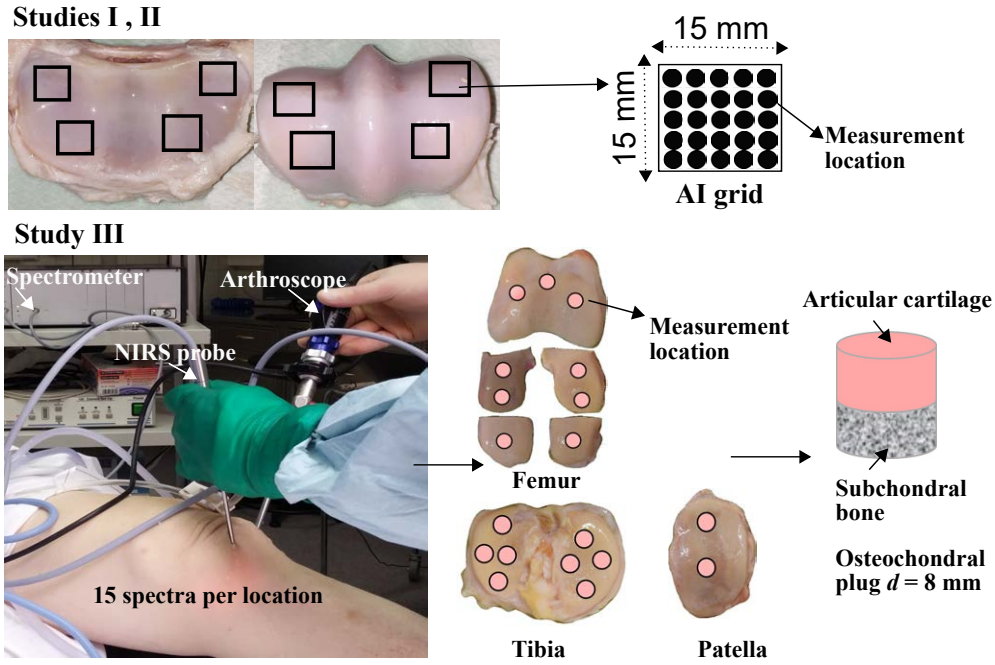


**Figure 6.1:** Samples used in this thesis and the corresponding measurement locations. Studies **I** and **II** utilized equine fetlock joints with the specific areas of interest (AI) marked with a grid. In study **III**, NIRS arthroscopy was performed on human cadaveric knee joints prior to sample extraction.

In studies **I** and **II**, equine fetlock joints ($n = 5$) were extracted and several areas of interest (AIs, $n = 44$) of varying lesion severity identified; by two veterinary orthopaedic surgeons. Each AI ($15 \times 15$ mm) consisted of 25 measurement locations arranged in an equispaced $5 \times 5$ grid pattern. Some locations in the AIs were omitted due to complete loss of cartilage matrix; hence, totally, 869 locations were measured.

In study **III**, anatomical locations on the surfaces of the tibia, femur, and patella,

of both knee joints of human cadavers ($n = 8$ males and 1 female, $age = 68 \pm 7$), was arthroscopically probed, visually assessed, and scored, all by an experienced orthopaedic surgeon at Kuopio University Hospital (Kuopio, Finland). Cartilage surface integrity was visualized with a conventional arthroscope (4 mm, 30° inclination Karl Storz GmbH Co, Tuttlingen, Germany) and scored using a novel optical probe in accordance with the ICRS grading system (Table 2.3). For a summary of the data in this thesis, see Table 6.1.

**Table 6.1:** Summary of materials and methods utilized in studies **I-III**.

| *Study* | *Joint* | *Number of joints* | *Measurement locations* | *Methods* |
|---|---|---|---|---|
| **I** | Equine fetlock | $N$ = 5 | $n$ = 869[a], 202[b] | *in vitro* NIRS, OCT, mechanical testing |
| **II** | Equine fetlock | $N$ = 5 | $n$ = 869[a], 202[b], 530[c] | *in vitro* NIRS, OCT, mechanical testing |
| **III** | Human knee | $N$ = 18 | $n$ = 265[a,b] | *ex vivo* NIRS, *in vitro* NIRS arthroscopy, mechanical testing |

[a] NIRS.
[a] Tissue thickness and modulus measurements.
[c] PG and collagen measurements.

## 6.1 NEAR-INFRARED SPECTROSCOPY

The spectral data and reference properties obtained from the equine samples were used in studies **I** **II**. The instrumentation for diffuse reflectance spectroscopy comprised of a light source (Avalight-HAL-S, Avantes BV, Apeldoorn, Netherlands), a spectrometer (AvaSpec-ULS2048XL, Avantes BV, $\lambda$ = 200–1160 nm, $resolution = 0.4$ nm), and a fibre-optic probe. The optical probe consists of seven fibres ($d = 600$ $\mu$m), with the central fibre collecting diffuse reflected light back to the spectrometer and the six peripherally-positioned fibres irradiating the samples. Spectral data within the 700–1150 nm resgion was utilized during analysis to enable a direct comparison with earlier studies.

In study **III**, the instrumentation consisted of a light source (AvaLight-HAL-(S)-Mini, $\lambda$ = 360–2500 nm, Avantes BV), two spectrometers (AvaSpec-ULS2048L, $\lambda$ = 350–1100 nm, $resolution = 0.6$ nm and AvaSpec-NIR256-2.5-HSC, $\lambda = 1000 - 2500$ nm, $resolution = 6.4$ nm, Avantes BV, Apeldoorn, Netherlands), and a custom-made arthroscopic fibre-optic probe. The design of the stainless-steel probe resembles a conventional arthroscopic hook (Figure 6.2), and it can withstand the autoclave sterilization process. The probe ($d = 3.25$ mm) consists of 114 fibre-optic cables ($d = 100$ $\mu$m) for the transmission of light to the sample, and seven for either spectrometer to collect the light scattered and reflected from the sample. In studies **I** to **III**, tissue structure plays a substantial role in the scattering and reflection of light, while the tissue composition mainly contributes to the magnitude of light absorption.

Before NIRS measurement, reference measurements consisting of background and reference spectra were acquired in order to calculate absorbance values ac-

cording to Equation 3.5. In studies **I** and **II**, NIRS was measured *in vitro* under laboratory conditions with probe-tissue contact in a perpendicular orientation with cartilage surface. The cartilage surface was periodically hydrated by placing PBS-soaked cloths over other locations. Each measurement consisted of an average of three spectra. NIRS (700-1050 nm) preprocessing consisted of smoothing and filtering by an SG filter (25 nm window) followed by second derivative pre-treatment.


**Figure 6.2:** Similarities between conventional arthroscopic hook used by surgeons (**A**) and custom designed NIRS optical probe (**B**) used in study **III** of this thesis.

In study **III**, NIRS measurements were performed arthroscopically (*ex vivo*) and repeated under laboratory conditions (*in vitro*). During arthroscopy, NIR spectra ($n = 15$ per location) were acquired while the knee joint was distended with saline solution (0.9% NaCl concentration), at room temperature (25°C). During *in vitro* measurements, NIR spectra were acquired at the same temperature (25°C) as in *ex vivo* measurements. After establishing optimal (perpendicular) probe contact with the cartilage surface, spectral data (average of three spectra per location) were acquired. To preprocess the spectra (710–1850 nm), a third order SG filter was applied for smoothing the spectral data. We had a window size of 17.40 nm (or 29 data points) for the spectrometer with a 0.6 nm resolution, and a window size of 108.8 nm (or 17 data points) for the spectrometer with a 6.4 nm resolution.

## 6.2 CARTILAGE THICKNESS AND BIOMECHANICAL TESTING

In studies **I** and **II**, cartilage thickness was determined using optical coherence tomography (OCT) with an Ilumien PCI Optimization System (St. Jude Medical, St. Paul, MN, USA), a 1305±55 nm scanning wavelength, an axial resolution of $\leq 20$ $\mu$m, and a lateral resolution of 25–60 $\mu$m. OCT was utilized because of the thinness of the cartilage (~0.8 mm, Table 7.1). Subsequently, biomechanical measurements (Figure 6.3) were conducted via indentation testing, with samples immersed in phosphate-buffered saline (PBS) [122, 123]. The material testing device consisted of a plane-ended indenter ($d = 530$ $\mu$m), a load cell (1000 g, $sensitivity \pm 0.25\%$, Model 303 31, Honeywell Sensotec Sensors, Columbus, OH, USA), and an actuator (displacement resolution was 0.1 $\mu$m, PM500-1 A, Newport, Irvine, CA, USA). Instantaneous modulus was measured on all 869 measurement locations; however, due to longer acquisition times, the dynamic and equilibrium moduli were measured for only 202 points.

In study **III**, vernier calipers ($resolution = 0.01$ mm) was used to determine cartilage thickness; the cartilage layer was much thicker than in studies **I** and **II** (Table 7.5). The thickness was estimated as the average of four longitudinal measurements

equidistant around the perimeter of the 8 mm plugs. As in studies **I** and **II**, the biomechanical properties of the cartilage were determined via indentation testing ($d$ = 667 or 728 $\mu$m plane-ended indenter), using the same material testing device. For the summary on biomechanical measurements protocol, see Table 6.2 and Figure 6.3.

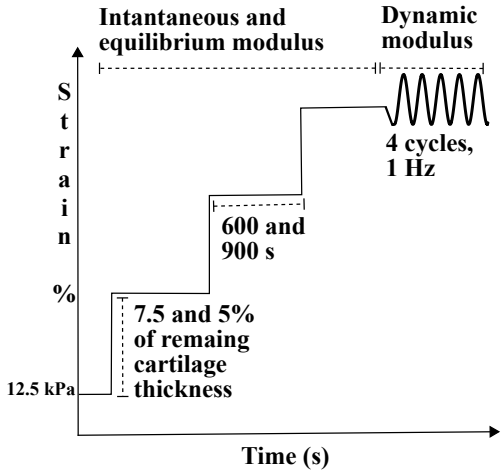


**Figure 6.3:** Summary of the biomechanical measurements protocol. In studies, **I** and **II**, 7.5% indentation thickness with 600 seconds of relaxation time was utilized, while in study **III**, 5% indentation thickness with 900 seconds of relaxation time were utilized.

**Table 6.2:** Summary of the biomechanical protocol in studies **I-III**.

| *Study* | *Instrumentation* | *Prestress* ($\sigma$) | *Protocol* | *Parameters* |
|---|---|---|---|---|
| **I/II** | Indentation material tester | 12.5 kPa | Stress-relaxation protocol (3 x 7.5%, $t_{relax}$ = 10 min) + sinusoidal loading (*frequency*=1.0 Hz, $\epsilon^{*}$=1%) | $E_{Inst}$, $E_{Eq}$ $E_{Dyn}$ |
| **III** | Indentation material tester | 12.5 kPa | Stress-relaxation protocol (3 x 5%, $t_{relax}$ = 15 min) + sinusoidal loading (*frequency*=1.0 Hz, $\epsilon$=2%) | $E_{Inst}$, $E_{Eq}$ $E_{Dyn}$ |

* Sinusoidal strain amplitude w.r.t to the tissue thickness.

Instantaneous modulus ($E_{Inst}$) was determined at the ramp phase of either the $2^{nd}$ or $3^{rd}$ step and equilibrium modulus ($E_{Eq}$) from the fit to the last three equilibrium points (Figure 2.2). Dynamic modulus ($E_{Dyn}$) was calculated as a ratio of the stress and strain amplitudes. Moduli were computed using the Hayes solution (Equation 6.1 [124]) for indentation geometry assuming the following Poisson's ratios: $\nu$ = 0.5 for $E_{Inst}$, 0.1 for $E_{Eq}$ , and 0.5 for $E_{Dyn}$[125]:

$$E = E_m \frac{(1-\nu^2)\pi\Re}{2\kappa h_t}, \tag{6.1}$$

where $E_m$ is the measured modulus, $\nu$ the Poisson's ratio, $\Re$ the radius of the indenter, $\kappa$ the theoretical correction factor [126], and $h_t$ the cartilage thickness.

## 6.3 HISTOLOGY

In study **II**, the cartilage composition and structural properties from the equine samples were determined. Osteochondral samples were subjected to decalcification in a solution of formalin and ethylenediaminetetraacetic acid (EDTA). After decalcification, the samples were embedded in paraffin blocks and four sections of 5 $\mu$m in thickness were extracted using a microtome. Histological imaging of the sections utilized Fourier transform infrared – FTIR ($n = 1$) micro-spectroscopy and polarized light microscopy – PLM ($n = 3$).

### 6.3.1 Collagen and proteoglycan distribution

FTIR imaging in FTIR microscopy was performed with a Thermo iN10 MX FT-IR microscope (Thermo Nicolet Corporation, Madison, WI, USA) in transmission mode. The images were acquired at a spectral resolution of 4 $cm^{-1}$ and $25 \times 25$ $\mu m^2$ pixel size. Each image was 2500 $\mu$m wide, covering the full thickness of cartilage in the MIR region. With an average of four scans per pixel and 530 sample locations, collagen content was estimated from amide $I$ peak in the 1584-1720 $cm^{-1}$ region and PG content from the carbohydrate area in the 984-1140 $cm^{-1}$ region.

### 6.3.2 Collagen orientation

Collagen orientation was measured using an Abrio PLM system (CRi, Inc., Woburn, MA, USA) and a conventional light microscope (Nikon Diaphot TMD, Nikon, Inc., Shinagawa, Tokyo, Japan). The main components of the set-up included a green band-pass filter, a circular polarizer, and a computer-regulated analyser consisting of dual liquid crystal polarizers and a charge-coupled device (CCD) camera. The imaging was performed at a 4x objective and a $2.53 \times 2.53$ $\mu m^2$ pixel size. The resulting images displayed $0^\circ$ for collagen orientation running parallel to cartilage surface and $90^\circ$ orientation for perpendicularly oriented fibrils.

## 6.4 MULTIVARIATE REGRESSION MODELLING AND STATISTICAL ANALYSES

In the study **I**, a comparison of different regression techniques was conducted to determine the optimal multivariate regression technique for the analysis of cartilage data. It compared PLSR, PCR, ridge, LASSO, and LS-SVM, regression techniques. To optimize the regression techniques the hyperparameters, namely the number of components for PLSR and PCR, shrinkage penalty factor for ridge regression, step size for LASSO and lambda for LS-SVM, were iterated from minimum to maximum values. The resulting series of models each technique yielded was subjected to a k-fold cross-validation ($k = 10$) and tested using an independent test set. The model

performing the best, in terms of root mean square error of calibration (*RMSEC*), $R^2_{Train}$, and $R^2_{Test}$, from each series, was retained. Finally, the performance metrics and the hyperparameters of the models retained were compared. This protocol ensured optimal settings for models under each regression technique. Variable selection methods namely, MC-UVE, CARS, VCPA, BiPLS, GA and jack-knife, were used to further optimize the best regression technique.

Study **II** identified spatial dependency due to adjacent measurement locations as a valid limitation that needed to be addressed in our regression analysis. Hence, the development of hybrid regression techniques (Figure 6.4), by combining dimension reduction methods and linear mixed effects (LME) modelling, became necessary. We focused on exploring the spatial dependency levels by grouping samples measured in the same joint (i.e., joint dependency) and samples measured on a particular bone of the joint (i.e., bone dependency) [127]. These dependency levels are accounted as mixed effects and dimension-reduced spectra as fixed effects, in LME models (equation 6.2).

$$Tissue\ property \approx NIR_{Dimension\ reduced\ spectra} + (1|Joints_{1-5}) + (1|Bones_{Upper-Lower}) \quad (6.2)$$

Briefly, PCA scores for PCA-LME and LASSO selected wavelengths in LASSO-LME were first computed and LME models were calibrated using information from 40 AIs (Figure 6.1, equation 6.2). The dependency levels containing sample grouping information are stored in $Z$ and $M$ design matrices 6.4. Hence, tissue property was predicted by utilizing dimension reduced spectra and dependency levels as shown in the equation in 6.2.

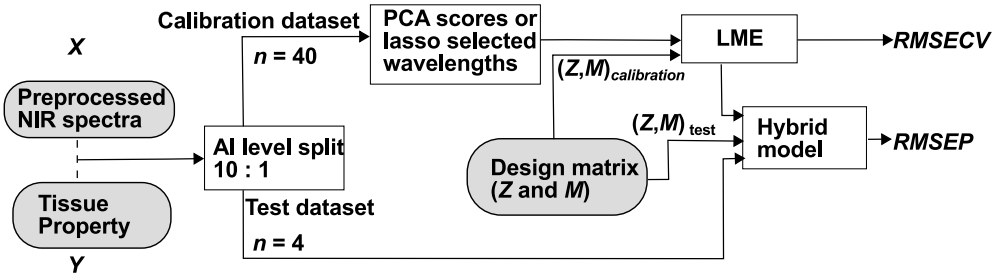


**Figure 6.4:** The hybrid regression technique.

To evaluate the model, 40 AIs were used in training set; the remaining four AIs were used for testing (Figure 6.4). This process was repeated eleven times until all the AIs were included in the test set exactly once. To determine the effect of spatial dependency, on model performance, the prediction models from the standard versions, namely PCR and LASSO regression, were subjected to a similar data split, and their performance was compared using *RMSECV* and *RMSEP* values.

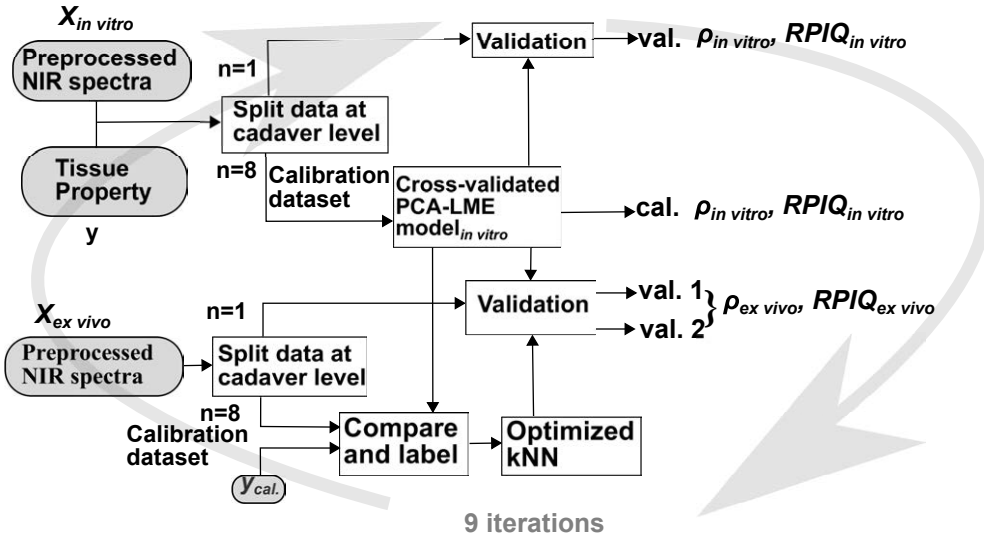


**Figure 6.5:** Hybrid regression technique and kNN classifier in the process flow.

In study **III**, the PCA-LME regression technique developed in study **II**, was utilized to create models for estimating cartilage spectra *in vitro* and biomechanical properties, while accounting for spatial dependency. Additionally, a k-nearest neighbours (kNN) classifier was employed to reject outlying *ex vivo* spectra resulting from sub-optimal probe-cartilage contact (Figure 6.5, equation 6.3).

$$\begin{aligned} Tissue\ property \approx\ & PCA\ Scores + (1|Location_{1-19}) \\ & + (1|Bones_{Tibia,Femur,Patella}) + (1|Joints_{Left\ or\ right}) + (1|Cadaver_{1-19}) \end{aligned} \quad (6.3)$$

The prediction model performances were evaluated for both *in vitro* and *ex vivo* NIRS data; we used Spearman's rank correlation ($\rho$) and the ratio of performance to interquartile range (RPIQ). The RPIQ was calculated as a ratio of the RMSE to the interquartile range.

# 7 RESULTS

This chapter presents a synopsis of the main results and the relationship between studies **I** to **III** (Figure 7.1). For detailed results, kindly refer to the original publications included in the supplementary material.

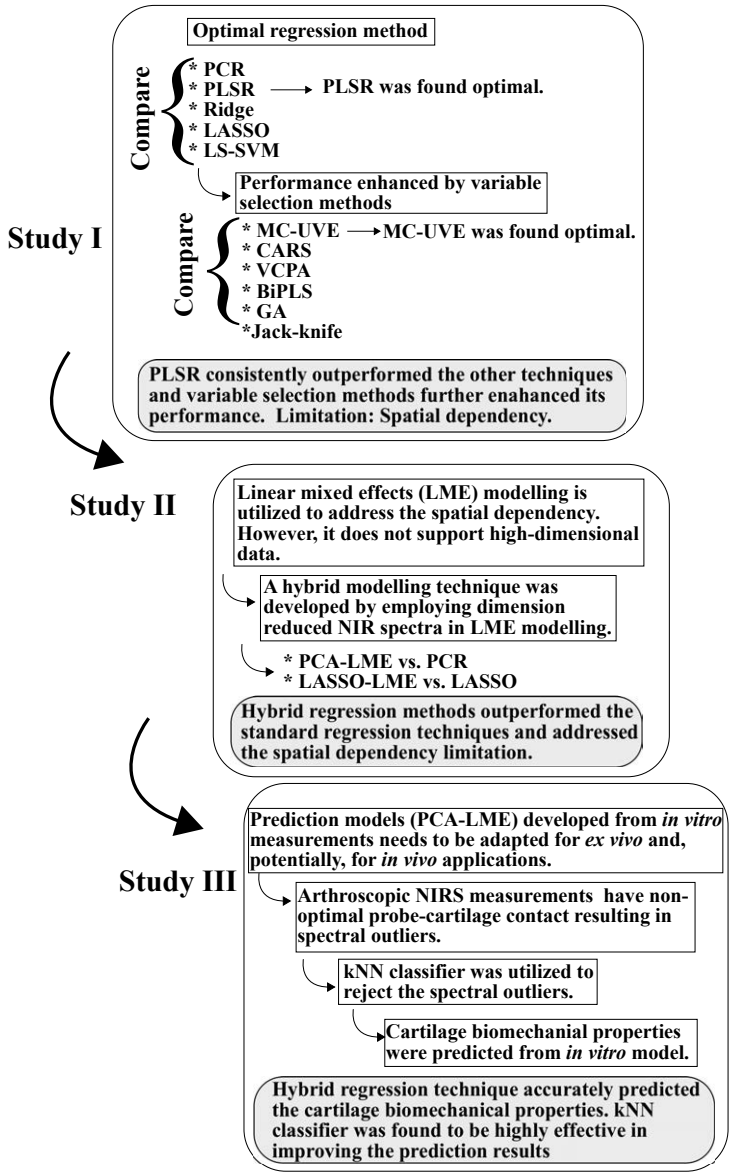


**Figure 7.1:** Summary of the main findings referred in this thesis.

## 7.1 COMPARISON OF REGRESSION TECHNIQUES

Study **I** using multivariate regression models, predicted from NIR spectra equine cartilage thickness and biomechanical properties (Table 7.1), namely instantaneous modulus, equilibrium modulus and dynamic modulus. The performance of different regression methods was evaluated using an independent test set. Compared to other regression models, PLSR performed consistently well, with a higher $R^2_{Testset}$, lower $RMSEP$, and faster computational times (Table 7.2). However, Zou's significance test revealed that performance differences amongst the top three techniques (PLSR, LASSO, and ridge regression) were non-significant (Table 7.2) [128].

**Table 7.1:** Summary of equine cartilage properties determined in studies **I** and **II**.

| *Property* | *Mean* | *Range* |
|---|---|---|
| Thickness (mm) | 0.89 | 0.32 - 1.81 |
| Instantaneous Modulus (MPa) | 4.76 | 0.11 - 20.88 |
| Equilibrium Modulus (MPa) | 2.00 | 0.03 - 5.38 |
| Dynamic Modulus (MPa) | 9.43 | 0.24 - 23.30 |
| PG content (AU)* | 6.31 | 0.60 - 14.71 |
| Collagen Content (AU)* | 33.35 | 12.16 - 64.39 |
| Collagen Orientation angle (°)* | 71.12 | 37.13 - 83.75 |

* An average of superficial, middle and deeper layers of the cartilage.

Variable selection methods were applied to optimize PLSR performance. MC-UVE was observed to not only improve PLSR results but also simplify the model, with a reduced number of components required for modelling the relationship between spectra and cartilage properties (Table 7.3).

**Table 7.2:** Test set results of the regression techniques comparison.

| Technique* | $R^2$ | RMSEP | Time** (secs) | Technique | $R^2$ | RMSEP | Time (secs) |
|---|---|---|---|---|---|---|---|
| Thickness (mm) | | | | Equilibrium Modulus (MPa) | | | |
| $PLSR^5$ | 75.5 | 0.11 | 2.50 | $PLSR^5$ | 68.5 | 0.94 | 1.00 |
| RIDGE | 74.0 | 0.11 | 1200 | LASSO | 60.1 | 1.06 | 96 |
| LASSO | 68.6 | 0.12 | 170 | LS-SVM | 54.2 | 1.14 | 0.16 |
| LS-SVM | 67.8 | 0.13 | 0.30 | RIDGE | 54.1 | 1.15 | 505 |
| $PCR^{13}$ | 67.3 | 0.13 | 1.00 | $PCR^{15}$ | 32.6 | 1.38 | 0.37 |
| Instantaneous Modulus (MPa) | | | | Dynamic Modulus (MPa) | | | |
| $PLSR^6$ | 51.0 | 2.46 | 2.0 | LASSO | 66.3 | 3.56 | 102 |
| RIDGE | 49.7 | 2.49 | 1100 | LS-SVM | 65.9 | 3.58 | 0.15 |
| LASSO | 48.7 | 2.52 | 5800 | $PLSR^2$ | 64.8 | 3.63 | 1.0 |
| $PCR^5$ | 44.2 | 2.62 | 0.60 | RIDGE | 61.3 | 3.82 | 470 |
| LS-SVM | 42.8 | 2.66 | 0.25 | PCR | 60.3 | 3.86 | 0.33 |

* Number of components is indicated in the power for PLSR and PCR.
** The computation times were calculated based on using Intel(R) Core(TM) i5-2400 CPU at 3.10 GHz with 64-bit Operating System.

**Table 7.3:** Test set results of variable selection methods comparison on PLSR models.

| Method | N | $R^2$ | RMSEP | Method | N | $R^2$ | RMSEP |
|---|---|---|---|---|---|---|---|
| Thickness (mm) | | | | Equilibrium Modulus (MPa) | | | |
| MC-UVE | 4 | 75.9 | 0.10 | None | 5 | 68.5 | 0.94 |
| None** | 5 | 75.5 | 0.11 | MC-UVE | 5 | 65.5 | 0.99 |
| GA | 8 | 74.8 | 0.11 | VCPA | 6 | 54.7 | 1.13 |
| JK | 1 | 74.0 | 0.11 | CARS | 2 | 54.1 | 1.14 |
| BiPLS | 12 | 70.0 | 0.12 | GA | 3 | 51.9 | 1.17 |
| CARS | 5 | 69.7 | 0.12 | JK | 6 | 43.0 | 1.27 |
| VCPA | 5 | 65.5 | 0.12 | BiPLS | 6 | 38.2 | 1.32 |
| Instantaneous Modulus (MPa) | | | | Dynamic Modulus (MPa) | | | |
| CARS | 3 | 51.8 | 2.44 | CARS | 3 | 77.8 | 2.89 |
| None | 5 | 51.0 | 2.46 | MC-UVE | 3 | 73.9 | 3.13 |
| VCPA | 4 | 49.0 | 2.51 | GA | 3 | 72.7 | 3.50 |
| MC-UVE | 2 | 48.3 | 2.53 | JK | 5 | 67.4 | 3.50 |
| GA | 4 | 48.3 | 2.53 | None | 2 | 64.8 | 3.63 |
| BiPLS | 7 | 45.5 | 2.59 | BiPLS | 7 | 55.9 | 4.07 |
| JK | 7 | 45.4 | 2.60 | VCPA | 1 | 50.5 | 4.31 |

** Variable selection not applied.

## 7.2 ACCOUNTING FOR SPATIAL DEPENDENCY IN SPECTROSCOPIC DATA

In study **II**, hybrid regression models, namely PCA-LME and LASSO-LME, performed better than standard regression algorithms (Table 7.4). PCA scores and wavelengths selected by LASSO were utilized together with LME, for regression modelling. For PCA-LME, the optimal number of components ranged from seven to nine. We observed that PCA-LME performed consistently better in modelling equine cartilage properties from spectral data than LASSO-LME did, across all properties. Furthermore, a mismatch in the proportionality between LASSO-LME's correlation values and the errors observed was detected.

**Table 7.4:** Comparison of standard and hybrid regression techniques.

| Property | Standard Regression | | Hybrid Regression | |
|---|---|---|---|---|
| | $\rho$ | RMSEP % | $\rho$ | RMSEP % |
| Thickness (mm) | 0.67 | 18.54 | 0.74 | 16.40 |
| | 0.57* | 20.17 | 0.65 | 18.36 |
| Dynamic Modulus (MPa) | 0.46 | 37.80 | 0.56 | 34.71 |
| | 0.29 | 42.73 | 0.27 | 39.58 |
| PG content (AU) | 0.34 | 22.42 | 0.42 | 22.54 |
| | 0.34 | 22.26 | 0.37 | 21.89 |
| Equilibrium Modulus (MPa) | 0.32 | 37.50 | 0.48 | 35.02 |
| | 0.38 | 33.90 | 0.46 | 34.84 |
| Collagen Content (AU) | 0.35 | 23.34 | 0.27 | 24.90 |
| | 0.29 | 22.90 | 0.32 | 23.14 |
| Collagen Orientation angle (°) | 0.27 | 25.01 | 0.23 | 25.35 |
| | 0.25 | 23.54 | 0.27 | 24.29 |

PCR and PCA-LME are shown in white rows while LASSO and LASSO-LME are shown in grey rows. The performance parameters $\rho$ (Spearman's rank correlation) and root mean square errors of prediction (RMSEP) were computed.

## 7.3 HYBRID REGRESSION TECHNIQUE EMPLOYED ON SPECTRAL DATA ACQUIRED FROM HUMAN CADAVERIC KNEE JOINTS

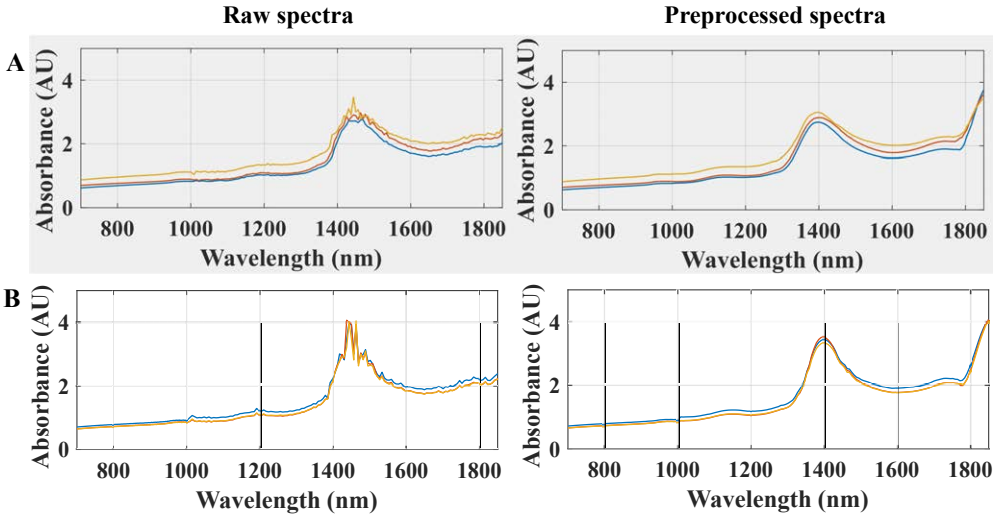


**Figure 7.2:** The representative raw and preprocessed ($3^{rd}$ order SG smoothed) spectra for *in vitro* (A) and *ex vivo* (B) measurements.

**Table 7.5:** Summary of reference human cartilage properties analysed in study **III**.

| *Property* | *Mean* | *Range* |
|---|---|---|
| Thickness (mm) | 2.58 | 1.22 - 5.90 |
| Instantaneous Modulus (MPa) | 13.81 | 0.12 - 51.82 |
| Equilibrium Modulus (MPa) | 0.89 | 0.02 - 3.67 |
| Dynamic Modulus (MPa) | 5.53 | 0.09 - 19.76 |

The hybrid regression method developed, PCA-LME, when applied for modelling the relationship between preprocessed NIR spectra (Figure 7.2) and biomechanical properties of human cartilage *in vitro* (Table 7.5), provided results with good performance metrics (Figure 7.3). To reject spectral outliers, we introduced the kNN classifier in the regression workflow; we observed it improved the accuracy of predictive models on arthroscopic spectra (Figure 7.3). The effect of the kNN filter is visualised in Figure 7.4 (ii) vs. (iii).

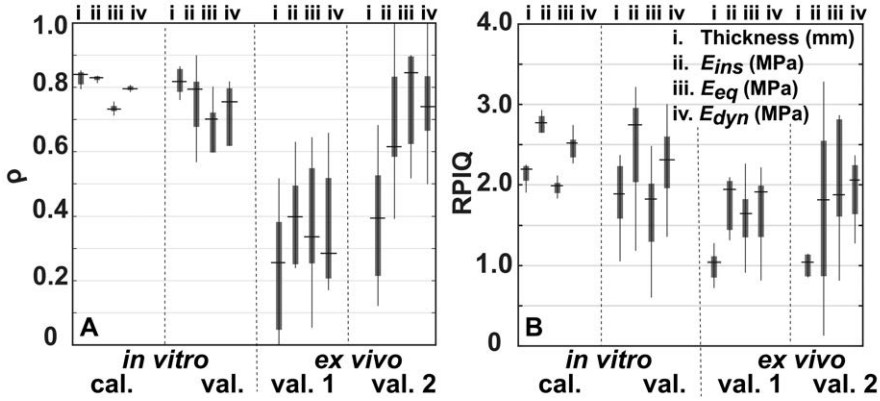


**Figure 7.3:** Performance of the prediction models on human knee cartilage data. Results of *in vitro* data show calibration (cal.) and validation (val.) on an independent dataset performance, while the results under *ex vivo* show validation model performance without (val. 1) kNN classifier and with classifier (val. 2) employed for rejecting spectral outliers.

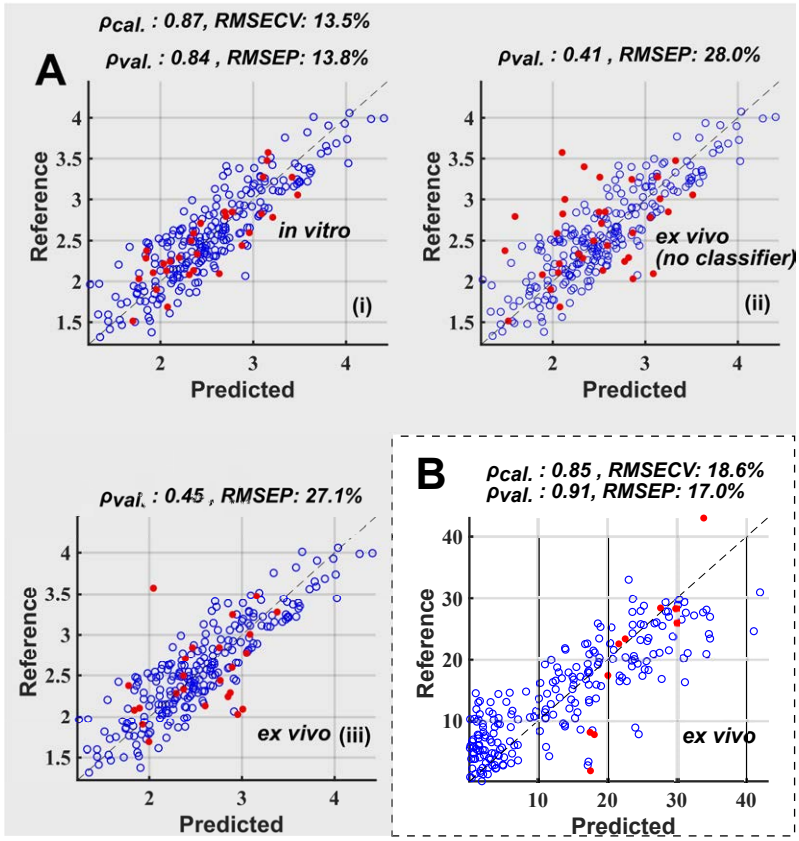


**Figure 7.4:** Predicted vs. reference plots for thickness (mm, **A**) and instantaneous modulus (MPa, **B**). The plots depict calibration data (blue, unfilled) and test set (red, filled). The plots in A show the model performance on *in vitro* **(i)** and *ex vivo* measurements, without **(ii)** and with **(iii)** the classifier. The plot in B shows the effect of classifier in predicting the instantaneous modulus.

# 8 DISCUSSION

This thesis explored various approaches for optimizing multivariate regression techniques for the NIRS characterization of articular cartilage integrity. The thesis also focused on adapting *in vitro* models for potential clinical applications. Addressing the limitations of spatial dependency with respect to adjacent measurement locations (as identified by Singer et al [129] and Ranstam et al [30, 130]) resulted in the development of a hybrid regression technique (study **II**). Study **III** applied this multivariate technique developed in study **II**, for the arthroscopic *ex vivo* evaluation of human cartilage biomechanical properties.

The studies included in this thesis encompass the development of NIRS-based characterization of articular cartilage. Firstly, a shift from employing classical regression techniques in their native form to employing optimized regression techniques suitable for dependent datasets improved prediction model performance (from study **I** to **II** and **III**). Secondly, the novel arthroscopic NIRS probe recently developed by our research group enabled the assessment of cartilage integrity during arthroscopy (from studies **I** and **II** to **III**). Thirdly, the regression techniques developed were first tested on animal models and then successfully adapted for the arthroscopic evaluation of human cartilage *ex vivo*, demonstrating the robustness of the approach (from studies **II** to **III**). Lastly, prediction models developed (using spectral data collected under *in vitro* laboratory conditions) were employed in making predictions based on noisy and real-world arthroscopic spectra, improving model robustness. The following sections touch upon the results of each individual study and their importance.

## 8.1 OPTIMAL REGRESSION TECHNIQUE

NIRS-based characterization of cartilage integrity relies on the performance of a multivariate model, with particular interest in regression models for estimating physico-chemical and material properties from the spectral data. With numerous regression techniques available, it is necessary to evaluate the performance of all applicable methods and options, to improve prediction performance. The high dimensionality and multicollinearity nature of NIRS data coupled with limited sample size potentially creates challenges during regression analysis; study **I** was formulated to address these concerns. The primary aim was to undertake a comprehensive comparison study consisting of classical regression techniques and variable selection methods. This was conducted, to the best of our knowledge, for the first time, to determine the optimal method for the analysis of cartilage NIRS spectra. The secondary aim was to observe how different regression techniques set weights or estimators in relating spectral data to the biomechanical properties measured. PLSR and shrinkage regression techniques exhibited the best calibration and test performance. The ability of PLSR to capture the variances of both predictor variables (NIRS) and response variables (tissue properties) enabled superior calibration performance [131]. This is unlike PCR (Table 7.2) [132, 133], which is based only on variances observed in NIR spectra.

Shrinkage regression techniques (ridge and LASSO) are good at optimizing the cost function or residual sum of squares (RSS), by adding a penalty term in the linear regression cost function [121, 134, 135]. The inherent resistance of ridge regression to over-fitting is probably why it outperformed LASSO. LASSO, despite its sparse nature, yielded good prediction performance on a test set; however, it was computationally demanding [136]. Although PLSR performed consistently better across all tissue properties, Zou's significance test revealed that the performance difference between PLSR and shrinkage regression techniques was statistically non-significant. We observed LS-SVM to be prone to overfitting possibly due to a data piling (coinciding projections of data in subspace) issue resulting from high data dimensionality [137, 138].

The approach adopted in this thesis to address the issue of the high dimensionality of NIRS data in regression analysis involved the application of variable selection methods [114, 139, 140]. This involves restricting cartilage NIRS wavelength regions prone either to saturation due to water in the tissue or to outliers due to water between the probe and cartilage surface during arthroscopic spectral measurements. Hence, variable selection methods could eliminate least-performing spectral regions, effectively reducing spectral variables for regression that were available and focusing regression computations on the spectral variables retained [141]. MC-UVE was observed to retain spectral data in the wavelength range corresponding to C-H (725-780 nm) and O-H (925-980 nm) while the N-H region centred at 850 nm (Table 3.1), suggesting that these regions are important in the assessment of cartilage properties. The narrow NIR range of spectral data 750-1100 nm in study **I** could have restricted the performance of the regression models; we observed the variable selection methods displayed marginal improvements over PLSR performance (Table 7.3).

The percentage of error in predicting the test values for equilibrium or dynamic moduli was in the 11% - 13% range for PLSR. The acceptable range of errors in predicting biomechanical parameters from NIRS data, is however, unclear. The results of the regression analysis could have benefited from a larger pool of data, which may improve prediction outcomes. The limitations of study **I** were the restriction of spectral data to a narrow band due to restrictions imposed by instrumentation and the lack of accounting for grouping effects resulting from adjacent measurement locations. These limitations were addressed in subsequent studies.

## 8.2 NEED FOR A HYBRID REGRESSION TECHNIQUE

NIRS-based characterization of cartilage integrity is undoubtedly a promising technique, but classical regression techniques in their current form are unsuitable for arthroscopic applications [21, 91]. The arthroscopic assessment of cartilage tissue properties to determine the severity or spread of a lesion could enable orthopaedic surgeons to map areas of damaged cartilage for optimal repair and to evaluate the success of the intervention [142–146]. These steps involve adjacent measurement locations, and classical regression techniques fail to handle spatial dependency resulting from repeated measures during arthroscopic procedures.

Numerous applications have utilized LME modelling to handle grouping effects in data [147–150]. Since LME cannot handle high dimensional data [151], it needs to be combined with dimension reduction methods [152], such as PCA or PLS. Recently, Conforti et al combined LME with PLS to account for spatial dependency in

the analysis of spectral data collected from soil samples [153]. However, since PLS requires both spectral data and tissue properties to compute PLS latent variables, this presents a limitation during independent testing or live arthroscopy measurements when tissue properties are unavailable. Hence, this study utilized PCA as the dimension reduction method, which was then used to calculate principal component scores as a fixed effect input to LME modelling. Next, LASSO was employed due to its ability to create sparse models on high dimensional data, without significant loss in prediction performance [110]. LASSO combined with LME also performed reasonably well. However, errors and correlations were inconsistent, due to the possibility of higher variances in the test set than in the training data. This inconsistency was unavoidable without introducing bias in the results; the division of data into training and test datasets was performed at the smallest level (AI) of the hierarchy. However, this inconsistency could be absent in other applications.

The improvements in the performance of prediction models when accounting for spatial dependency became apparent in a comparison of the results of the hybrid regression techniques with those of the standard regression technique. We thus addressed the practical limitation imposed by the study design in tissue mapping applications.

## 8.3 THE APPLICATION OF THE HYBRID REGRESSION TECHNIQUE IN KNEE ARTHROSCOPY

In developing prediction models for the cartilage assessments using arthroscopically recorded spectra, the challenges of handling imperfect probe-cartilage contact needed to be addressed [154, 155]. Achieving perfect contact between the probe and articular surface was a problem when manoeuvring the inner contours of the knee joint. First, PCA-LME regression technique was used to build *in vitro* models from data collected in a controlled laboratory environment. This ensured that the models were trained on spectra that were collected as accurately as possible. Next, a protocol employing a kNN classifier trained to reject noisy NIR spectra (resulting from poor probe-cartilage contact) from arthroscopic measurements was developed. Finally, these *in vitro* models were applied on the NIR spectra measured during arthroscopy to predict cartilage biomechanical properties while the kNN classifier was used to reject spectral outliers.

The integration of the hybrid regression technique and spectral classifier enabled, to our knowledge, for the first time, the prediction of the functional properties of cartilage from the human knee joint on arthroscopically acquired spectra. This integrated approach could allow real-time assessment of cartilage conditions well beyond the superficial layer and uncover conditions indiscernible during a conventional arthroscopic procedure [20]. The use of a wider spectral bandwidth (Study **III**, 710 to 1850 nm) in the model as compared to earlier studies (Studies **I** and **II**, 700 to 1050 nm) led to improved correlations. The limited number of cadavers ($n$ = 9) and the lack of diversity in terms of age, gender, and cartilage conditions, are a shortcoming of this proof-of-concept study.

The optical power rating of the light source (AvaLight-HAL-S-Mini, Avantes BV) ranges from 3.2 to 7 mW in the 350 - 1100 nm wavelength region. The photochemical damage from UV rays and the thermal damage due to the rise in tissue temperature are potential effects to be considered in clinical applications. These effects can be minimized by using a UV filter to eliminate the UV wavelength band, and a transis-

tor–transistor logic (TTL) shutter to limit the overall tissue exposure time; acquisition time was 2.4 secs per measurement in Study **III**. The tissue damage threshold for an NIR light source within a 700-1064 nm wavelength range and an exposure time of two seconds ranges between 119.8 to 454.8 $W/cm^3$ [156]. Thus, NIRS operating within the 710-1850 nm wavelength range (Study **III**) is non-ionizing, and the heating effect is minimal as the power per area for the optical probe is 0.68 $W/cm^2$. However, during arthroscopies, the knee joint is continuously irrigated by saline at room temperature. Thus, the heating effect if any in cartilage, (due to its nearly white colour), is minimized.

## 8.4 FUTURE STUDIES

The prediction models developed in study **III** would benefit from the inclusion of structural and compositional properties as these can aid the surgeon's assessments of cartilage integrity. Furthermore, the models used in study **III** were based on samples collected from mostly male (eight out of nine) and older cadavers. The dataset could be substantially diversified by including variations in gender, age, and knee conditions, aiding in the creation of a generalized cartilage model. These variations could be modelled as mixed effects in the hybrid regression technique, further enhancing the applicability of the model. The models developed could be utilized in predicting properties of cartilage *in vivo*, possibly in individuals undergoing total knee replacements. This *in vivo* assessment could serve as a validation for the model, and verification could be performed by subjecting the cartilage collected after knee replacement to *in vitro* measurements.

The probe-cartilage contact could further be improved with electronic sensors, adjustable or flexible probe heads, or other similar advancements. These design enhancements would enable the use of the probe on difficult surface contours of the joint, without problematic spectral outliers. These improvements could be utilized and combined with other imaging modalities, such as OCT and ultrasound, yielding a comprehensive and quantitative multimodal tool for arthroscopy [157, 158].

# 9 SUMMARY AND CONCLUSIONS

The recent improvements in computational power have led to the utilization of powerful multivariate regression techniques in NIRS, enabling accurate and reliable assessments of articular cartilage conditions. The studies in this thesis have supplemented these developments in the application of NIRS by optimizing multivariate regression techniques, enabling arthroscopic characterization of cartilage integrity in the human knee joint.

The following are the prominent findings of this thesis:

1. PLSR effectively handled multicollinearity in spectral data and related these spectra to cartilage tissue properties. Variable selection methods simplified PLSR prediction models, improving the performance of the model.

2. The hybrid regression technique was capable of simultaneously accounting for relevant levels of dependencies and taking into consideration the assumption of sample independence.

3. The combination of hybrid regression modelling and a spectral classifier enabled the NIRS-based arthroscopic evaluation of the biomechanical properties of articular cartilage in the human knee joint.

Thus, the findings reported in this thesis and further work in this direction (see Discussion) could revolutionize the arthroscopic assessment of cartilage health and improve our understanding of osteoarthritis.

# BIBLIOGRAPHY


[1] J. K. Sarin, H. Brommer, D. Argüelles, P. Puhakka, S. Inkinen, I. O. Afara, S. Saarakkala, and J. Töyräs, "Multimodality scoring of chondral injuries in the equine fetlock joint ex vivo," *Osteoarthritis and cartilage* **25**, 790–798 (2017).

[2] J. K. Sarin, L. Rieppo, H. Brommer, I. O. Afara, S. Saarakkala, and J. Töyräs, "Combination of optical coherence tomography and near infrared spectroscopy enhances determination of articular cartilage composition and structure," *Scientific reports* **7**, 10586 (2017).

[3] A. J. Sophia Fox, A. Bedi, and S. A. Rodeo, "The basic science of articular cartilage: structure, composition, and function," *Sports health* **1**, 461–468 (2009).

[4] R. F. Loeser, "Aging and osteoarthritis: the role of chondrocyte senescence and aging changes in the cartilage matrix," *Osteoarthritis and Cartilage* **17**, 971–979 (2009).

[5] V. C. Mow, M. H. Holmes, and W. M. Lai, "Fluid transport and mechanical properties of articular cartilage: a review," *Journal of biomechanics* **17**, 377–394 (1984).

[6] J. Buckwalter and H. Mankin, "Articular cartilage: degeneration and osteoarthritis, repair, regeneration, and transplantation.," *Instructional course lectures* **47**, 487–504 (1998).

[7] V. C. Mow, A. Ratcliffe, and A. R. Poole, "Cartilage and diarthrodial joints as paradigms for hierarchical materials and structures," *Biomaterials* **13**, 67–97 (1992).

[8] E. Sebbag, R. Felten, F. Sagez, J. Sibilia, H. Devilliers, and L. Arnaud, "The world-wide burden of musculoskeletal diseases: a systematic analysis of the world health organization burden of diseases database," *Annals of the rheumatic diseases* **78**, 844–848 (2019).

[9] J. Day, J. Van Der Linden, R. Bank, M. Ding, I. Hvid, D. Sumner, and H. Weinans, "Adaptation of subchondral bone in osteoarthritis," *Biorheology* **41**, 359–368 (2004).

[10] J. Christoforakis, R. Pradhan, J. Sanchez-Ballester, N. Hunt, and R. K. Strachan, "Is there an association between articular cartilage changes and degenerative meniscus tears?" *Arthroscopy: The Journal of Arthroscopic & Related Surgery* **21**, 1366–1369 (2005).

[11] C. Armstrong and V. Mow, "Variations in the intrinsic mechanical properties of human articular cartilage with age, degeneration, and water content.," *The Journal of bone and joint surgery. American volume* **64**, 88–94 (1982).

[12] M. Hakulinen, S. Saarakkala, J. Töyräs, H. Kröger, and J. Jurvelin, "Dual energy x-ray laser measurement of calcaneal bone mineral density," *Physics in Medicine & Biology* **48**, 1741 (2003).

[13] J. Rieppo, M. Hyttinen, E. Halmesmaki, H. Ruotsalainen, A. Vasara, I. Kiviranta, J. Jurvelin, and H. Helminen, "Changes in spatial collagen content and collagen network architecture in porcine articular cartilage during growth and maturation," *Osteoarthritis and Cartilage* **17**, 448–455 (2009).

[14] A. Horng, E. Brun, A. Mittone, S. Gasilov, L. Weber, T. Geith, S. Adam-Neumair, S. D. Auweter, A. Bravin, M. F. Reiser, et al., "Cartilage and soft tissue imaging using x-rays: propagation-based phase-contrast computed tomography of the human knee in comparison with clinical imaging techniques and histology," *Investigative radiology* **49**, 627–634 (2014).

[15] U. M. Døhn, B. J. Ejbjerg, M. Hasselquist, E. Narvestad, M. Szkudlarek, J. M. Møller, H. S. Thomsen, M. Østergaard, et al., "Are bone erosions detected by magnetic resonance imaging and ultrasonography true erosions? a comparison with computed tomography in rheumatoid arthritis metacarpophalangeal joints," *Arthritis research & therapy* **8**, R110 (2006).

[16] G. Spahn, G. Felmet, and G. O. Hofmann, "Traumatic and degenerative cartilage lesions: arthroscopic differentiation using near-infrared spectroscopy (nirs)," *Archives of orthopaedic and trauma surgery* **133**, 997–1002 (2013).

[17] G. O. Hofmann, J. Marticke, R. Grossstück, M. Hoffmann, M. Lange, H. K. Plettenberg, R. Braunschweig, O. Schilling, I. Kaden, and G. Spahn, "Detection and evaluation of initial cartilage pathology in man: a comparison between mrt, arthroscopy and near-infrared spectroscopy (nir) in their relation to initial knee pain," *Pathophysiology* **17**, 1–8 (2010).

[18] P. Orth and H. Madry, "Cartilage repair: arthroscopic," *Arthroscopy: Basic to Advanced*, 189 (2016).

[19] J. Malda, J. C. de Grauw, K. E. Benders, M. J. Kik, C. H. van de Lest, L. B. Creemers, W. J. Dhert, and P. R. van Weeren, "Of mice, men and elephants: the relation between articular cartilage thickness and body mass," *PloS one* **8**, e57683 (2013).

[20] R. Thein and M. Eichenblat, "Concealed knee cartilage lesions: is arthroscopic probing therapeutic?" *The American journal of sports medicine* **27**, 495–499 (1999).

[21] I. Afara, I. Prasadam, R. Crawford, Y. Xiao, and A. Oloyede, "Non-destructive evaluation of articular cartilage defects using near-infrared (nir) spectroscopy in osteoarthritic rat models and its direct relation to mankin score," *Osteoarthritis and cartilage* **20**, 1367–1373 (2012).

[22] I. O. Afara, I. Prasadam, H. Moody, R. Crawford, Y. Xiao, and A. Oloyede, "Near infrared spectroscopy for rapid determination of mankin score components: a potential tool for quantitative characterization of articular cartilage at surgery," *Arthroscopy: The Journal of Arthroscopic & Related Surgery* **30**, 1146–1155 (2014).

[23] I. O. Afara, I. Prasadam, R. Crawford, Y. Xiao, and A. Oloyede, "Near infrared (nir) absorption spectra correlates with subchondral bone micro-ct parameters in osteoarthritic rat models," *Bone* **53**, 350–357 (2013).

[24] A. Sakudo, "Near-infrared spectroscopy for medical applications: current status and future perspectives," *Clinica Chimica Acta* **455**, 181–188 (2016).

[25] I. A. Olumegbon, A. Oloyede, and I. O. Afara, "Near-infrared (nir) spectroscopic evaluation of articular cartilage: a review of current and future trends," *Applied Spectroscopy Reviews* **52**, 541–559 (2017).

[26] F. Faris, M. Thorniley, Y. Wickramasinghe, R. Houston, P. Rolfe, N. Livera, and A. Spencer, "Non-invasive in vivo near-infrared optical measurement of the penetration depth in the neonatal head," *Clinical Physics and Physiological Measurement* **12**, 353 (1991).

[27] I. Afara, S. Singh, and A. Oloyede, "Application of near infrared (nir) spectroscopy for determining the thickness of articular cartilage," *Medical engineering & physics* **35**, 88–95 (2013).

[28] J. K. Sarin, N. C. te Moller, I. A. Mancini, H. Brommer, J. Visser, J. Malda, P. R. van Weeren, I. O. Afara, and J. Töyräs, "Arthroscopic near infrared spectroscopy enables simultaneous quantitative evaluation of articular cartilage and subchondral bone in vivo," *Scientific reports* **8**, 13409 (2018).

[29] I. O. Afara, C. Florea, I. A. Olumegbon, C. T. Eneh, M. K. Malo, R. K. Korhonen, and J. Töyräs, "Characterizing human subchondral bone properties using near-infrared (nir) spectroscopy," *Scientific reports* **8**, 9733 (2018).

[30] J. Ranstam and J. A. Cook, "Considerations for the design, analysis and presentation of in vivo studies," *Osteoarthritis and cartilage* **25**, 364–368 (2017).

[31] G. Spahn, H. Plettenberg, H. Nagel, E. Kahl, H. M. Klinger, T. Mückley, M. Günther, G. O. Hofmann, and J. A. Mollenhauer, "Evaluation of cartilage defects with near-infrared spectroscopy (nir): an ex vivo study," *Medical engineering & physics* **30**, 285–292 (2008).

[32] Y. Kobrina, L. Rieppo, S. Saarakkala, J. S. Jurvelin, and H. Isaksson, "Clustering of infrared spectra reveals histological zones in intact articular cartilage," *Osteoarthritis and cartilage* **20**, 460–468 (2012).

[33] P. Kiviranta, J. Töyräs, M. Nieminen, M. Laasanen, S. Saarakkala, H. Nieminen, M. Nissi, and J. Jurvelin, "Comparison of novel clinically applicable methodology for sensitive diagnostics of cartilage degeneration," *Eur Cell Mater* **13**, 46–55 (2007).

[34] E. Kheir and D. Shaw, "Hyaline articular cartilage," *Orthopaedics and Trauma* **23**, 450–455 (2009).

[35] K. R. Flik, N. Verma, B. J. Cole, and B. R. Bach, "Articular cartilage," in *Cartilage repair strategies* (Springer, 2007), pp. 1–12.

[36] H. Forster and J. Fisher, "The influence of continuous sliding and subsequent surface wear on the friction of articular cartilage," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* **213**, 329–345 (1999).

[37] D. C. Fithian, M. A. Kelly, and V. C. Mow, "Material properties and structure-function relationships in the menisci.," *Clinical orthopaedics and related research*, 19–31 (1990).

[38] J. M. Mansour, "Biomechanics of cartilage," *Kinesiology: the mechanics and pathomechanics of human movement*, 66–79 (2003).

[39] G. Ateshian, L. Soslowsky, and V. Mow, "Quantitation of articular surface topography and cartilage thickness in knee joints using stereophotogrammetry," *Journal of biomechanics* **24**, 761–776 (1991).

[40] S. Faber, F. Eckstein, S. Lukasz, R. Mühlbauer, J. Hohe, K.-H. Englmeier, and M. Reiser, "Gender differences in knee joint cartilage thickness, volume and articular surface areas: assessment with quantitative three-dimensional mr imaging," *Skeletal radiology* **30**, 144–150 (2001).

[41] G. Meachim, "Effect of age on the thickness of adult articular cartilage at he shoulder joint.," *Annals of the rheumatic diseases* **30**, 43 (1971).

[42] R. Stockwell, "The interrelationship of cell density and cartilage thickness in mammalian articular cartilage.," *Journal of anatomy* **109**, 411 (1971).

[43] F. M. Hall and G. Wyshak, "Thickness of articular cartilage in the normal knee.," *JBJS* **62**, 408–413 (1980).

[44] J. Buckwalter and H. Mankin, "Articular cartilage: part i tissue design and chondrocyte-matrix interactions," *JBJS* **79**, 600–611 (1997).

[45] J. A. Buckwalter and J. Martin, "Degenerative joint disease.," in Clinical symposia (summit, nj: 1957), Vol. 47, 2 (1995), pp. 1–32.

[46] F. G. Donnan, "The theory of membrane equilibria.," *Chemical Reviews* **1**, 73–90 (1924).

[47] A. Maroudas, "Distribution and diffusion of solutes in articular cartilage," *Biophysical journal* **10**, 365–379 (1970).

[48] G. Kempson, H. Muir, C. Pollard, and M. Tuke, "The tensile properties of the cartilage of human femoral condyles related to the content of collagen and glycosaminoglycans," *Biochimica et Biophysica Acta (BBA)-General Subjects* **297**, 456–472 (1973).

[49] R. K. Korhonen, P. Julkunen, W. Wilson, and W. Herzog, "Importance of collagen orientation and depth-dependent fixed charge densities of cartilage on mechanical behavior of chondrocytes," *Journal of biomechanical engineering* **130**, 021003 (2008).

[50] R. Minns and F. Steven, "The collagen fibril organization in human articular cartilage.," *Journal of anatomy* **123**, 437 (1977).

[51] W. Wilson, C. Van Donkelaar, B. Van Rietbergen, and R. Huiskes, "A fibril-reinforced poroviscoelastic swelling model for articular cartilage," *Journal of biomechanics* **38**, 1195–1204 (2005).

[52] A. Mak, "The apparent viscoelastic behavior of articular cartilage—the contributions from the intrinsic matrix viscoelasticity and interstitial fluid flows," *Journal of biomechanical engineering* **108**, 123–130 (1986).

[53] W. Hayes and L. Mockros, "Viscoelastic properties of human articular cartilage.," *Journal of applied physiology* **31**, 562–568 (1971).

[54] M. Holmes, W. Lai, and V. Mow, "Singular perturbation analysis of the nonlinear, flow-dependent compressive stress relaxation behavior of articular cartilage," *Journal of biomechanical engineering* **107**, 206–218 (1985).

[55] W. C. Hayes and A. Bodine, "Flow-independent viscoelastic properties of articular cartilage matrix," *Journal of biomechanics* **11**, 407–419 (1978).

[56] M. Wong, M. Ponticiello, V. Kovanen, and J. Jurvelin, "Volumetric changes of articular cartilage during stress relaxation in unconfined compression," *Journal of biomechanics* **33**, 1049–1054 (2000).

[57] R. Shirazi, A. Shirazi-Adl, and M. Hurtig, "Role of cartilage collagen fibrils networks in knee joint biomechanics under compression," *Journal of biomechanics* **41**, 3340–3348 (2008).

[58] M. Cross, E. Smith, D. Hoy, S. Nolte, I. Ackerman, M. Fransen, L. Bridgett, S. Williams, F. Guillemin, C. L. Hill, et al., "The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study," *Annals of the rheumatic diseases* **73**, 1323–1330 (2014).

[59] R. F. Loeser, "Aging and osteoarthritis," *Current opinion in rheumatology* **23**, 492 (2011).

[60] E. M. Roos, "Joint injury causes knee osteoarthritis in young adults," *Current opinion in rheumatology* **17**, 195–200 (2005).

[61] M. B. Goldring and S. R. Goldring, "Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis," *Annals of the New York Academy of Sciences* **1192**, 230–237 (2010).

[62] D. S. Howell, "Pathogenesis of osteoarthritis," *The American journal of medicine* **80**, 24–28 (1986).

[63] D. D. Anderson, S. Chubinskaya, F. Guilak, J. A. Martin, T. R. Oegema, S. A. Olson, and J. A. Buckwalter, "Post-traumatic osteoarthritis: improved understanding and opportunities for early intervention," *Journal of Orthopaedic Research* **29**, 802–809 (2011).

[64] N. Ishiguro, T. Kojima, and A. R. Poole, "Mechanism of cartilage destruction in osteoarthritis.," *Nagoya journal of medical science* **65**, 73–84 (2002).

[65] T. M. Griffin and F. Guilak, "The role of mechanical loading in the onset and progression of osteoarthritis," *Exercise and sport sciences reviews* **33**, 195–200 (2005).

[66] W. R. Trickey, G. M. Lee, and F. Guilak, "Viscoelastic properties of chondrocytes from normal and osteoarthritic human cartilage," *Journal of Orthopaedic Research* **18**, 891–898 (2000).

[67] M. Pattrick, E. Hamilton, R. Wilson, S. Austin, and M. Doherty, "Association of radiographic changes of osteoarthritis, symptoms, and synovial fluid particles in 300 knees.," *Annals of the rheumatic diseases* **52**, 97 (1993).

[68] M. A. Risberg, B. E. Oiestad, R. Gunderson, A. K. Aune, L. Engebretsen, A. Culvenor, and I. Holm, "Changes in knee osteoarthritis, symptoms, and function after anterior cruciate ligament reconstruction: a 20-year prospective follow-up study," *The American journal of sports medicine* **44**, 1215–1224 (2016).

[69] D. J. Hunter, D. Schofield, and E. Callander, "The individual and socioeconomic impact of osteoarthritis," *Nature Reviews Rheumatology* **10**, 437 (2014).

[70] F. M. Hall, "Radiographic diagnosis and accuracy in knee joint effusions," *Radiology* **115**, 49–54 (1975).

[71] J. Bekkers, L. Creemers, W. Dhert, and D. Saris, "Diagnostic modalities for diseased articular cartilage-from defect to degeneration: a review," *Cartilage* **1**, 157–164 (2010).

[72] N. J. Manek and N. E. Lane, "Osteoarthritis: current concepts in diagnosis and management.," *American family physician* **61**, 1795–1804 (2000).

[73] K. Sinusas, "Osteoarthritis: diagnosis and treatment.," *American family physician* **85** (2012).

[74] A. Guermazi, F. W. Roemer, D. Burstein, and D. Hayashi, "Why radiography should no longer be considered a surrogate outcome measure for longitudinal assessment of cartilage in knee osteoarthritis," *Arthritis research & therapy* **13**, 247 (2011).

[75] J. Olive, M.-A. D'ANJOU, K. Alexander, S. Laverty, and C. Theoret, "Comparison of magnetic resonance imaging, computed tomography, and radiography for assessment of noncartilaginous changes in equine metacarpophalangeal osteoarthritis," *Veterinary radiology & ultrasound* **51**, 267–279 (2010).

[76] F. Eckstein, F. Cicuttini, J.-P. Raynauld, J. C. Waterton, and C. Peterfy, "Magnetic resonance imaging (mri) of articular cartilage in knee osteoarthritis (oa): morphological assessment," *Osteoarthritis and cartilage* **14**, 46–75 (2006).

[77] M. Brittberg and C. S. Winalski, "Evaluation of cartilage injuries and repair," *JBJS* **85**, 58–69 (2003).

[78] P. Jüni, S. Reichenbach, and P. Dieppe, "Osteoarthritis: rational approach to treating the individual," *Best practice & research Clinical rheumatology* **20**, 721–740 (2006).

[79] C. Cooper and D. Coggon, "Physical activity and knee osteoarthritis," *The Lancet* **353**, 2177–2178 (1999).

[80] D. T. Felson, R. C. Lawrence, P. A. Dieppe, R. Hirsch, C. G. Helmick, J. M. Jordan, R. S. Kington, N. E. Lane, M. C. Nevitt, Y. Zhang, et al., "Osteoarthritis: new insights. part 1: the disease and its risk factors," *Annals of internal medicine* **133**, 635–646 (2000).

[81] A.-M. Malfait and T. J. Schnitzer, "Towards a mechanism-based approach to pain management in osteoarthritis," *Nature Reviews Rheumatology* **9**, 654 (2013).

[82] M. Fransen, S. McConnell, and M. Bell, "Exercise for osteoarthritis of the hip or knee.," *The Cochrane database of systematic reviews*, CD004286–CD004286 (2003).

[83] G. Bentley, L. Biant, R. Carrington, M. Akmal, A. Goldberg, A. Williams, J. Skinner, and J. Pringle, "A prospective, randomised comparison of autologous chondrocyte implantation versus mosaicplasty for osteochondral defects in the knee," *The Journal of bone and joint surgery. British volume* **85**, 223–230 (2003).

[84] C. Clar, E. Cummins, L. McIntyre, S. Thomas, J. Lamb, L. Bain, P. Jobanputra, and N. Waugh, "Clinical and cost-effectiveness of autologous chondrocyte implantation for cartilage defects in knee joints: systematic review and economic evaluation," (2005).

[85] M. S. Burman, "Arthroscopy or the direct visualization of joints: an experimental cadaver study," *JBJS* **13**, 669–695 (1931).

[86] F. R. Noyes and C. L. Stabler, "A system for grading articular cartilage lesions at arthroscopy," *The American journal of sports medicine* **17**, 505–513 (1989).

[87] B. Brismar, T. Wredmark, T. Movin, J. Leandersson, and O. Svensson, "Observer reliability in the arthroscopic classification of osteoarthritis of the knee," *The Journal of bone and joint surgery. British volume* **84**, 42–47 (2002).

[88] A. Javed, M. Siddique, M. Vaghela, and A. Hui, "Interobserver variations in intra-articular evaluation during arthroscopy of the knee," *The Journal of bone and joint surgery. British volume* **84**, 48–49 (2002).

[89] R. W. Wright, J. R. Ross, A. K. Haas, L. J. Huston, E. A. Garofoli, D. Harris, K. Patel, D. Pearson, J. Schutzman, M. Tarabichi, et al., "Osteoarthritis classification scales: interobserver reliability and arthroscopic correlation," *The Journal of bone and joint surgery. American volume* **96**, 1145 (2014).

[90] G. Spahn, H. M. Klinger, and G. O. Hofmann, "How valid is the arthroscopic diagnosis of cartilage lesions? results of an opinion survey among highly experienced arthroscopic surgeons," *Archives of orthopaedic and trauma surgery* **129**, 1117–1121 (2009).

[91] G. Spahn, H. Plettenberg, E. Kahl, H. M. Klinger, T. Mückley, and G. O. Hofmann, "Near-infrared (nir) spectroscopy. a new method for arthroscopic evaluation of low grade degenerated cartilage lesions. results of a pilot study," *BMC musculoskeletal disorders* **8**, 47 (2007).

[92] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, and N. Jent, "A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies," *Journal of pharmaceutical and biomedical analysis* **44**, 683–700 (2007).

[93] M. Blanco and I. Villarroya, "Nir spectroscopy: a rapid-response analytical tool," *TrAC Trends in Analytical Chemistry* **21**, 240–250 (2002).

[94] D. Delpy and M. Cope, "Quantification in tissue near–infrared spectroscopy," *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **352**, 649–659 (1997).

[95] M. Padalkar and N. Pleshko, "Wavelength-dependent penetration depth of near infrared radiation into cartilage," *Analyst* **140**, 2093–2100 (2015).

[96] C.-W. Chang, D. A. Laird, M. J. Mausbach, and C. R. Hurburgh, "Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties," *Soil Science Society of America Journal* **65**, 480–490 (2001).

[97] D. A. Burns and E. W. Ciurczak, *Handbook of near-infrared analysis* (CRC press, 2007).

[98] Å. Rinnan, F. van den Berg, and S. B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra," *TrAC Trends in Analytical Chemistry* **28**, 1201–1222 (2009).

[99] U. P. Palukuru, C. M. McGoverin, and N. Pleshko, "Assessment of hyaline cartilage matrix composition using near infrared spectroscopy," *Matrix Biology* **38**, 3–11 (2014).

[100] M. Padalkar, R. Spencer, and N. Pleshko, "Near infrared spectroscopic evaluation of water in hyaline cartilage," *Annals of biomedical engineering* **41**, 2426–2436 (2013).

[101] J. K. Marticke, A. Hösselbarth, K. L. Hoffmeier, I. Marintschev, S. Otto, M. Lange, H. K. Plettenberg, G. Spahn, and G. O. Hofmann, "How do visual, spectroscopic and biomechanical changes of cartilage correlate in osteoarthritic knee joints?" *Clinical Biomechanics* **25**, 332–340 (2010).

[102] I. Afara, S. Singh, and A. Oloyede, "Load-unloading response of intact and artificially degraded articular cartilage correlated with near infrared (nir) absorption spectra," *Journal of the mechanical behavior of biomedical materials* **20**, 249–258 (2013).

[103] I. O. Afara, M. Hauta-Kasari, J. S. Jurvelin, A. Oloyede, and J. Töyräs, "Optical absorption spectra of human articular cartilage correlate with biomechanical properties, histological score and biochemical composition," *Physiological measurement* **36**, 1913 (2015).

[104] I. O. Afara, S. Singh, H. Moody, L. Zhang, and A. Oloyede, "Characterization of articular cartilage recovery and its correlation with optical response in the near-infrared spectral range," *Cartilage* **8**, 307–316 (2017).

[105] J. K. Sarin, M. Amissah, H. Brommer, D. Argüelles, J. Töyräs, and I. O. Afara, "Near infrared spectroscopic mapping of functional properties of equine articular cartilage," *Annals of biomedical engineering* **44**, 3335–3345 (2016).

[106] D. Baykal, O. Irrechukwu, P.-C. Lin, K. Fritton, R. G. Spencer, and N. Pleshko, "Nondestructive assessment of engineered cartilage constructs using near-infrared spectroscopy," *Applied spectroscopy* **64**, 1160–1166 (2010).

[107] I. O. Afara, H. Moody, S. Singh, I. Prasadam, and A. Oloyede, "Spatial mapping of proteoglycan content in articular cartilage using near-infrared (nir) spectroscopy," *Biomedical optics express* **6**, 144–154 (2015).

[108] Z. Xiaobo, Z. Jiewen, M. J. Povey, M. Holmes, and M. Hanpin, "Variables selection methods in near-infrared spectroscopy," *Analytica chimica acta* **667**, 14–32 (2010).

[109] R. M. Balabin and E. I. Lomakina, "Support vector machine regression (svr/ls-svm)—an alternative to neural networks (ann) for analytical chemistry? comparison of nonlinear methods on near infrared (nir) spectroscopy data," *Analyst* **136**, 1703–1712 (2011).

[110] R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**, 273–282 (2011).

[111] A. Bellino, C. Colombo, P. Iovieno, A. Alfani, G. Palumbo, and D. Baldantoni, "Chemometric technique performances in predicting forest soil chemical and biological properties from uv-vis-nir reflectance spectra with small, high dimensional datasets," *iForest-Biogeosciences and Forestry* **9**, 101 (2015).

[112] E. Vigneau, M. Devaux, E. Qannari, and P. Robert, "Principal component regression, ridge regression and ridge principal component regression in spectroscopy calibration," *Journal of Chemometrics: A Journal of the Chemometrics Society* **11**, 239–249 (1997).

[113] H. Öjelund, H. Madsen, and P. Thyregod, "Calibration with absolute shrinkage," *Journal of Chemometrics: A Journal of the Chemometrics Society* **15**, 497–509 (2001).

[114] W. Cai, Y. Li, and X. Shao, "A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra," *Chemometrics and intelligent laboratory systems* **90**, 188–194 (2008).

[115] H. Li, Y. Liang, Q. Xu, and D. Cao, "Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration," *Analytica chimica acta* **648**, 77–84 (2009).

[116] Y.-H. Yun, W.-T. Wang, B.-C. Deng, G.-B. Lai, X.-b. Liu, D.-B. Ren, Y.-Z. Liang, W. Fan, and Q.-S. Xu, "Using variable combination population analysis for variable selection in multivariate calibration," *Analytica chimica acta* **862**, 14–23 (2015).

[117] X. Zou, J. Zhao, and Y. Li, "Selection of the efficient wavelength regions in ft-nir spectroscopy for determination of ssc of based on bipls and fipls models," *Vibrational spectroscopy* **44**, 220–227 (2007).

[118] R. Leardi and A. L. Gonzalez, "Genetic algorithms applied to feature selection in pls regression: how and when to use them," *Chemometrics and intelligent laboratory systems* **41**, 195–207 (1998).

[119] F. Westad and H. Martens, "Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression," *Journal of Near Infrared Spectroscopy* **8**, 117–124 (2000).

[120] B. Wise, *Properties of partial least squares (pls) regression, and differences between algorithms. eigenvector research*.

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

[122] R. Korhonen, M. Laasanen, J. Töyräs, J. Rieppo, J. Hirvonen, H. Helminen, and J. Jurvelin, "Comparison of the equilibrium response of articular cartilage in unconfined compression, confined compression and indentation," *Journal of biomechanics* **35**, 903–909 (2002).

[123] J. Töyräs, J. Rieppo, M. Nieminen, H. Helminen, and J. Jurvelin, "Characterization of enzymatically induced degradation of articular cartilage using high frequency ultrasound," *Physics in Medicine & Biology* **44**, 2723 (1999).

[124] W. Hayes, L. Keer, G. Herrmann, and L. Mockros, "A mathematical analysis for indentation tests of articular cartilage," *Journal of biomechanics* **5**, 541–551 (1972).

[125] J. Jurvelin, M. Buschmann, and E. Hunziker, "Optical and mechanical determination of poisson's ratio of adult bovine humeral articular cartilage," *Journal of biomechanics* **30**, 235–241 (1997).

[126] J. Jurvelin, I. Kiviranta, A.-M. Säämänen, M. Tammi, and H. Helminen, "Indentation stiffness of young canine knee articular cartilage—influence of strenuous joint loading," *Journal of biomechanics* **23**, 1239–1246 (1990).

[127] M. Prakash, A. Joukainen, J. K. Sarin, L. Rieppo, I. O. Afara, and J. Töyräs, "Near-infrared spectroscopy based arthroscopic evaluation of human knee joint cartilage, through automated selection of an anatomically specific regression model," in Biophotonics congress: biomedical optics congress 2018 (2018), OF4D.3.

[128] G. Y. Zou, "Toward using confidence intervals to compare correlations.," *Psychological methods* **12**, 399 (2007).

[129] M. Singer, T. Krivobokova, A. Munk, and B. De Groot, "Partial least squares for dependent data," *Biometrika* **103**, 351–362 (2016).

[130] J. Ranstam, *Repeated measurements, bilateral observations and pseudoreplicates, why does it matter?* 2012.

[131] O. Yeniay and A. Goktas, "A comparison of partial least squares regression with other prediction methods," *Hacettepe Journal of Mathematics and Statistics* **31**, 99–101 (2002).

[132] T. Næs and B. Mevik, "Understanding the collinearity problem in regression and discriminant analysis," *Journal of Chemometrics: A Journal of the Chemometrics Society* **15**, 413–426 (2001).

[133] J. Rodrigues, A. Alves, H. Pereira, D. da Silva Perez, G. Chantre, and M. Schwanninger, "Nir plsr results obtained by calibration with noisy, low-precision reference values: are the results acceptable?" *Holzforschung* **60**, 402–408 (2006).

[134] R. B. Gramacy, E. Pantaleo, et al., "Shrinkage regression for multivariate inference with missing data, and an application to portfolio balancing," *Bayesian Analysis* **5**, 237–262 (2010).

[135] J. B. Copas, "Regression, prediction and shrinkage," *Journal of the Royal Statistical Society. Series B (Methodological)*, 311–354 (1983).

[136] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "A sparse-group lasso," *Journal of Computational and Graphical Statistics* **22**, 231–245 (2013).

[137] J. Ahn and J. Marron, "The maximal data piling direction for discrimination," *Biometrika* **97**, 254–259 (2010).

[138] M. H. Lee, J. Ahn, and Y. Jeon, "Hdlss discrimination with adaptive data piling," *Journal of Computational and Graphical Statistics* **22**, 433–451 (2013).

[139] L. S. Aucott, P. H. Garthwaite, and J. Currall, "Regression methods for high dimensional multicollinear data," *Communications in Statistics-Simulation and Computation* **29**, 1021–1037 (2000).

[140] I. M. Johnstone and D. M. Titterington, *Statistical challenges of high-dimensional data*, 2009.

[141] C. Abrahamsson, J. Johansson, A. Sparén, and F. Lindgren, "Comparison of different variable selection methods conducted on nir transmission measurements on intact tablets," *Chemometrics and Intelligent Laboratory Systems* **69**, 3–12 (2003).

[142] B. R. Mandelbaum, J. E. Browne, F. Fu, L. Micheli, J. B. Mosely, C. Erggelet, T. Minas, and L. Peterson, "Articular cartilage lesions of the knee," *The American journal of sports medicine* **26**, 853–861 (1998).

[143] K. Hjelle, E. Solheim, T. Strand, R. Muri, and M. Brittberg, "Articular cartilage defects in 1,000 knee arthroscopies," *Arthroscopy: The Journal of Arthroscopic & Related Surgery* **18**, 730–734 (2002).

[144] T. A. Ahmed and M. T. Hincke, "Strategies for articular cartilage lesion repair and functional restoration," *Tissue Engineering Part B: Reviews* **16**, 305–329 (2010).

[145] A. Watrin-Pinzano, J.-P. Ruaud, Y. Cheli, P. Gonord, L. Grossin, I. Bettembourg-Brault, P. Gillet, E. Payan, G. Guillot, P. Netter, et al., "Evaluation of cartilage repair tissue after biomaterial implantation in rat patella by using t2 mapping," *Magnetic Resonance Materials in Physics, Biology and Medicine* **17**, 219–228 (2004).

[146] J. R. Slauterbeck, P. Kousa, B. C. Clifton, S. Naud, T. W. Tourville, R. J. Johnson, and B. D. Beynnon, "Geographic mapping of meniscus and cartilage lesions associated with anterior cruciate ligament injuries," *JBJS* **91**, 2094–2103 (2009).

[147] B. M. Bolker, M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White, "Generalized linear mixed models: a practical guide for ecology and evolution," *Trends in ecology & evolution* **24**, 127–135 (2009).

[148] A. Cnaan, N. M. Laird, and P. Slasor, "Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data," *Statistics in medicine* **16**, 2349–2380 (1997).

[149] R. H. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," *Journal of memory and language* **59**, 390–412 (2008).

[150] A. Agresti, J. G. Booth, J. P. Hobert, and B. Caffo, "2. random-effects modeling of categorical response data," *Sociological Methodology* **30**, 27–80 (2000).

[151] J. Schelldorfer, P. Bühlmann, and S. V. DE GEER, "Estimation for high-dimensional linear mixed-effects models using 21131-penalization," *Scandinavian Journal of Statistics* **38**, 197–214 (2011).

[152] P. Cunningham, "Dimension reduction," in *Machine learning techniques for multimedia* (Springer, 2008), pp. 91–112.

[153] M. Conforti, A. Castrignanò, G. Robustelli, F. Scarciglia, M. Stelluti, and G. Buttafuoco, "Laboratory-based vis–nir spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content," *Catena* **124**, 60–67 (2015).

[154] B. A. Lakin, B. D. Snyder, and M. W. Grinstaff, "Assessing cartilage biomechanical properties: techniques for evaluating the functional performance of cartilage in health and disease," *Annual review of biomedical engineering* **19**, 27–55 (2017).

[155] E. C. Quan, "Design of an in-vivo probe to detect cartilage degeneration," PhD thesis (Massachusetts Institute of Technology, 1998).

[156] J. N. Bixler, B. H. Hokr, M. L. Denton, G. D. Noojin, A. D. Shingledecker, H. T. Beier, R. J. Thomas, B. A. Rockwell, and V. V. Yakovlev, "Assessment of tissue heating under tunable near-infrared radiation," *Journal of biomedical optics* **19**, 070501 (2014).

[157] N. Te Moller, H. Brommer, J. Liukkonen, T. Virén, M. Timonen, P. Puhakka, J. Jurvelin, P. Van Weeren, and J. Töyräs, "Arthroscopic optical coherence tomography provides detailed information on articular cartilage lesions in horses," *The Veterinary Journal* **197**, 589–595 (2013).

[158] J. Liukkonen, P. Lehenkari, J. Hirvasniemi, A. Joukainen, T. Virén, S. Saarakkala, M. T. Nieminen, J. S. Jurvelin, and J. Töyräs, "Ultrasound arthroscopy of human knee cartilage and subchondral bone in vivo," *Ultrasound in medicine & biology* **40**, 2039–2047 (2014).

**MITHILESH PRAKASH**

*The studies in this thesis explore and optimize regression techniques for near-infrared spectroscopy (NIRS)-based characterization of articular cartilage integrity. The limitations of conventional regression techniques were addressed via the development of a hybrid regression technique to aid the arthroscopic application of NIRS, thus paving the way for clinical applications in joint tissue diagnosis.*

UNIVERSITY OF
EASTERN FINLAND


*uef.fi*

**PUBLICATIONS OF**
**THE UNIVERSITY OF EASTERN FINLAND**
*Dissertations in Forestry and Natural Sciences*

ISBN 978-952-61-3160-3
ISSN 1798-5668