Accounting for spatial dependency in multivariate spectroscopic data

Prakash M[1], Sarin J. K[1,2], Rieppo L[1,3], Afara I. O[1], Töyräs J[1,2]

mithilesh.prakash@uef.fi

jaakko.sarin@uef.fi

lassi.rieppo@oulu.fi

isaac.afara@uef.fi

juha.toyras@uef.fi

[1] Department of Applied Physics, University of Eastern Finland, Kuopio, Finland
[2] Diagnostic Imaging Center, Kuopio University Hospital, Kuopio, Finland
[3] Research Unit of Medical Imaging, Physics and Technology, Faculty of Medicine, University of Oulu, Oulu, Finland

Running title: Spatial dependency in multivariate NIRS

Corresponding author:
Mithilesh Prakash, M.Sc. (Tech.)
Department of Applied Physics
University of Eastern Finland
Kuopio, Finland
Tel: +358 449204680
Fax: +358 17162131
Email: mithilesh.prakash@uef.fi

Abstract

We examine a hybrid multivariate regression technique to account for the spatial dependency in spectroscopic data due to adjacent measurement locations in the same joint by combining dimension reduction methods and linear mixed effects (LME) modeling. Spatial correlation is a common limitation (assumption of independence) faced in diagnostic applications involving adjacent measurement locations, such as mapping of tissue properties, and can impede tissue evaluations. Near-infrared spectra were collected from equine joints (*n*=5) and corresponding biomechanical (*n*=202), compositional (*n*=530), and structural (*n*=530) properties of cartilage tissue were measured. Subsequently, hybrid regression models for estimating tissue properties from the spectral data were developed in combination with principal component analysis (PCA-LME) scores and least absolute shrinkage and selection operator (LASSO-LME). Performance comparison of PCA-LME and principal component regression, and LASSO-LME and LASSO regression was conducted to evaluate the effects of spatial dependency. A systematic improvement in calibration models' correlation coefficients and a decrease in cross validation errors were observed when accounting for spatial dependency. Our results indicate that accounting for spatial dependency using a LME-based approach leads to more accurate prediction models.

Keywords: Linear mixed effects; articular cartilage; near infrared (NIR) spectroscopy; spectroscopic mapping; principal components; LASSO.

1. Introduction

Articular cartilage, a connective tissue covering the ends of bones in a joint, is susceptible to post-traumatic osteoarthritis (PTOA) due to focal injuries caused by sudden excessive impact loading. The injury, although initially localized, often spreads over time, resulting in altered functional performance of the whole joint. Arthroscopic evaluation of tissue properties around the injury site and assessing the spread of the injury could enable optimal surgical intervention, thereby minimizing the risk of PTOA. Currently, in clinical arthroscopies[1], cartilage is assessed visually through an endoscope and by palpating the tissue surface with a metal hook[2]. This method is qualitative, unreliable, and poorly reproducible[3,4], thus necessitating development of novel, quantitative, robust, and reliable methods.

Non-destructive diagnostic tools, such as near-infrared (NIR) spectroscopy, have shown potential for arthroscopic characterization of articular cartilage integrity[5]. NIR spectroscopy is a vibrational spectroscopic technique that has been utilized for spatial assessment of cartilage biomechanical, compositional, and structural properties[6–8]. In these studies, multivariate regression was utilized to relate cartilage NIR spectra with its tissue properties. However, conventional multivariate regression methods, such as partial least squares (PLS), are based on the underlying assumption of independent observations[9], whereas biomedical characterization of tissue integrity, for example in arthroscopy, often involves multiple measurement locations within close proximity in the same joint. This grouping effect of samples introduces spatial dependency and is likely to cause unreliable correlations if unaccounted for in regression modeling[10,11].

Linear mixed effects (LME) regression and its input parameters, namely fixed effects and random effects, can be designed for specific datasets to account for grouping effects. Since only a limited number of regressors (input variables) can be utilized in model creation using LME, adaptation for a large set of variables, such as NIR spectra, requires dimension reduction and/or variable selection methods. Hence, the input variables need to be methodically selected by retaining only the most important ones.

Application of dimension reduction methods, such as principal component analysis (PCA)[12] via PCA score, and variable selection and regularization methods like LASSO (least absolute shrinkage and selection operator) or *L1*-penalization[13], are effective approaches for reducing the high dimensionality of the data, such as NIR spectra. PCA finds a set of projections that maximizes the variance in the original dataset; hence, the data structure in the sample space is captured even in the low dimensional subspaces. LASSO[14] is a regularization method ideal for creating sparse models with high statistical accuracy in predictions.

In this study, we propose a hybrid technique, which combines dimension reduction methods and LME regression, to account for spatial dependency in analysis of multivariate dataset. This is based on the hypothesis that the hybrid regression technique can effectively model the relationship between cartilage NIR spectra and its properties while accounting for dependencies within the data.

2. Materials and Methods

To account for spatial dependency in the dataset, the contributing levels of dependency must first be identified. The levels of dependency are defined by the experimental design and the scope of the application. In our application on NIR-based characterization of cartilage, joint level (measurement locations grouped in one complete joint) and bone level (measurement locations grouped on one bone of a particular joint) were identified (Figure 1) as the two significant levels of dependency[15]. (Other application specific dependency levels can be accommodated in the design matrix)

[Proposed position for Fig. 1]

Subsequently, models were developed for relating the predictors ($\boldsymbol{X}$) to the response variables ($\boldsymbol{y}$) while accounting for the identified dependency levels (grouping effects). The adaptation of LME can be written in the equation form:

$$\boldsymbol{y_i} = \boldsymbol{X\beta} + \boldsymbol{Zu_1} + \boldsymbol{Mu_2} + \varepsilon, \qquad (1)$$

where $\boldsymbol{y_i}$is an N(number of observations)-by-1 response vector of reference values for the $\boldsymbol{i}^{\text{th}}$ tissue property, $\boldsymbol{X}$ is an *N*-by-*P* (dimension reduced NIR spectra) matrix of fixed effect regressors, $\boldsymbol{\beta}$ is a *P*-by-1 vector of fixed effects coefficients, $\boldsymbol{Z}$ is an *N*-by-*Q* (grouping count) random effects design matrix, $\boldsymbol{M}$ is an *N*-by-1 vector of additional random effects vector, $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ are the mixed effects coefficients of sizes *Q*-by-1 and 1-by-1 respectively, and ε is an *N*-by-1 vector representing the observation error. Restricted maximum likelihood method was employed for estimating LME[16].

2.1 Equine cartilage dataset

In this study, we utilized NIR spectral data from equine cartilage measured in earlier studies[17,18]. In summary, metacarpophalangeal joints (*n*=5) were acquired from a slaughterhouse, and specific areas of interest (AI, *n*=44) with cartilage lesions of varying severity were selected by a veterinary surgeon. Subsequently, a 15 x 15 mm grid consisting of 25 measurement locations was marked on each AI with a felt-tip pen (Figure 1). The measurement locations were equally spaced (*interdistance* = 2.5 mm), and locations with highly eroded cartilage were excluded, yielding a total of 869 measurement points. NIR spectral measurements and thickness values were acquired on each of the 869 measurements; however, biomechanical measurements were performed only on 202 locations and compositional analysis on 530 locations due to limitations set by sample preservation and geometry, respectively. NIR spectra was matched with corresponding tissue property based on location during regression analysis.

2.2 NIR spectral measurements

The NIR spectroscopy instrumentation consisted of a halogen light source (wavelength range: 360–2500 nm, power 5 W, optical power: 239 µW in a d_{fiber} = 600 µm, Avantes BW, Apeldoorn, Netherlands), and a spectrometer (wavelength range: 200–1160 nm, Avantes BW, Apeldoorn, Netherlands). A customized fiber optic probe (*d*=5 mm) consisting of seven fibers (d_{fiber}=600 µm) within the central window (*d*=2 mm), the six outer fibers for transmitting, and the central one for collecting the reflectance spectrum, was utilized. Prior to sample measurements, dark and reference

spectra were acquired. Dark spectrum was acquired with the spectrometer light source switched off in order to collect background noise. With the light source switched on, reference spectrum was acquired from a reflectance standard (Spectralon, SRS-99, Labsphere Inc., North Sutton, USA). The absorbance values of each sample spectra were scaled as per Beer-Lambert's law using the dark and reference spectra. In addition, signal acquisition time was optimized to maximize the signal to noise ratio. The average of three spectral measurements that each consisted of eight co-added spectral scans ($t_{eight\ scans}$=720 ms) was calculated. To preprocess the spectra (700-1050nm), Savitzky-Golay estimates of the second derivative using 41 points (or 25 nm) and a third-order polynomial for the smoothing were computed. This preprocessing not only removes baseline offset and dominant linear terms but also enhances subtle absorption peaks.

2.3 Cartilage thickness and biomechanical measurements

Cartilage thickness at all NIRS measurement locations was determined using optical coherence tomography (OCT) via the Ilumien PCI Optimization System, (St. Jude Medical, St. Paul, MN, USA) at an operating wavelength of 1305±55 nm, axial resolution <20 μm, and lateral resolution 25–60 μm. The samples were fully immersed in phosphate-buffered saline (PBS) during the measurements.

Biomechanical indentation measurements were performed at 202 locations using a customized material testing device consisting of a load cell (Sensotec, Columbus, OH, USA) with force resolution of 5 mN, an actuator (PM1A1798-1 A, Newport, Irvine, CA, USA) with displacement resolution of 0.1 μm (PM500-1 A, Newport, Irvine, CA, USA),

and a plane-ended cylindrical indenter (d=0.53 mm). Equilibrium modulus (E_{eq}) and dynamic modulus (E_{dyn}) were calculated using an indentation protocol detailed in Korhonen et al[19] and Sarin et al[17].

2.4 Reference measurements of cartilage composition and structure

The osteochondral samples were first decalcified in a solution containing formalin and ethylenediaminetetraacetic acid (EDTA), then fixed in paraffin blocks from which thin sections (n=4, thickness=5 μm) were cut using a microtome along the measurement line. The sections were then subjected to histological imaging, *i.e.*, Fourier transform infrared (FTIR) microspectroscopy (n=1) and polarized light microscopy (PLM, n=3).

2.5 Collagen and proteoglycan distribution

Spatial collagen and proteoglycan (PG) distributions were measured by FTIR microspectroscopy using a Thermo iN10 MX FT-IR microscope (Thermo Nicolet Corporation, Madison, WI, USA). The microscope was operated in transmission mode at a spectral resolution of 4 cm^{-1} and a pixel size of 25 × 25 μm^2. 500-μm-wide regions including full cartilage thickness were mapped from each sample in the mid infrared region. The average of four scans per pixel were obtained. The collagen content for 530 sample locations was estimated from the amide I peak (1584-1720 cm^{-1}), and PG contents for these samples was obtained from the carbohydrate region (984-1140 cm^{-1}).

193 *2.6 Collagen orientation*

194 Collagen orientation was measured using an Abrio PLM system (CRi, Inc., Woburn,
195 MA, USA) on top of a conventional light microscope (Nikon Diaphot TMD, Nikon, Inc.,
196 Shinagawa, Tokyo, Japan). This system consists of a green bandpass filter, a circular
197 polarizer, and a computer-controlled analyzer comprising of two liquid crystal
198 polarizers and a charged couple device (CCD) camera. The specimens (n=530) were
199 imaged at identical orientation using a 4.0x objective, which resulted in a pixel size of
200 2.53 × 2.53 μm^2. The collagen fiber orientation in the resulting images show 0 degrees
201 for collagen aligned parallel to cartilage surface and 90 degrees for collagen aligned
202 perpendicular to cartilage surface.

203

204 Table 1 summarizes the dataset used in this study.

205

206 **[Proposed position for Table. 1**]

207

208 *2.7 Hybrid regression analysis approach*

209 *2.7.1 PCA-LME regression technique*

210 With the hybrid PCA-LME regression technique, the equation is:

211 $$\textbf{Tissue property} \sim \textbf{PCA scores} + (1 \mid \textbf{Joints}_{1\text{-}5}) + (1 \mid \textbf{Bones}_{\text{Upper-Lower}}). \quad (2)$$

212 Hence, PCA scores are used as fixed effects, and measurement grouping information
213 from joint level (Z matrix) and bone level (M matrix) are used as mixed effects (1 |
214 dependency levels) to predict the tissue property (Figure 1 and Equation 2).

The data was split into calibration and test datasets (Figure 2) at the AI level. Hence, the calibration and test datasets had no spectra from same AIs. With a 10:1 data split, 40 AIs were utilized in calibration of the PCA-LME model and 4 AIs to test the model. This split was repeated 11 times with each AI included in the test set only once.

During modelling, calibration dataset was subjected to a 10-fold cross-validation. In cross-validation, the number of PCA scores to build the PCA-LME model was varied from 1 to 15 and the model with the smallest *RMSECV* was retained. The retained model was used to predict the tissue property from the test dataset and the root mean square error of prediction (*RMSEP*) was calculated to assess model performance.

The PCA scores of the test dataset were calculated by adjusting the test spectra according to the mean of the calibration spectra ($\mu_{calibration}$). This was performed by first calculating $\mu_{calibration}$ and principal components coefficients (Coeff) of the calibration spectra. Subsequently, $\mu_{calibration}$ was subtracted from the test spectra, and then multiplied by Coeff (Figure 2).

[Proposed position for Fig. 2]

235 *2.7.2 LASSO-LME regression technique*

236 Similar to the PCA-LME regression approach, dimension reduction was performed
237 with LASSO. LASSO adds a penalty equivalent to the sum of absolute values of the
238 magnitude of the coefficients to each wavelength and a nonnegative regularization
239 parameter λ. Next, utilizing 10-fold cross-validation, the sum of squared errors (*SSE*)
240 of the LASSO fit is calculated and plotted for varying λ values. Wavelengths
241 corresponding to the minimum *SSE* were retained. The dimension-reduced NIR
242 spectra, based on the selected wavelengths, were used as input to LME regression
243 (according to equation 1).

244

245 The performance of models was evaluated as an average of 11 iterations. Spearman's
246 rank correlation (ρ), a distribution free and non-parametric statistic, was employed to
247 assess the regression models due to assumptions on normality of the dataset. The
248 commonly used coefficient of determination (R^2) was not used as its definition for
249 mixed model is unclear[20]. All spectral and regression analysis were done using
250 custom-made programs designed on MATLAB R2017b (Mathworks Inc, Natick, MA).

251

252 **3. Results**

253 The results (average of 11 iterations) of the hybrid regression models in comparison
254 to corresponding standard regression models (Table 2) showed consistently higher
255 correlation coefficient and lower *RMSECV*. Hence, the inclusion of spatial dependency
256 information in the models resulted in better performance. The changes in ρ relative to
257 *RMSEP* in test sets (Table 2 and Figure 4) were inconsistent for PG content, E_{eq} and
258 collagen content.

[Proposed position for Table. 2]

[Proposed position for Fig. 3]

[Proposed position for Fig. 4]

The optimal number of principal components were 7-9. The performance of LASSO-based variable selection varied depending on the predicted parameter. Prediction of cartilage thickness required the highest number of variables (Figure 3), while E_{eq} required the least.

4. Discussion

In this study, we propose a hybrid regression technique to account for the effect of spatial dependency commonly encountered in spectroscopic characterization of biological tissues. We first identified the dependency levels based on known groupings of the measurement locations. We then introduced hybrid techniques that combine variable reduction methods (based on PCA and LASSO) and LME regression, allowing incorporation of the identified dependency levels into the predictive models. The performance of PCA-LME and LASSO-LME were then compared with that of PCR and LASSO regression, respectively. The results highlighted the importance and benefits of accounting for spatial dependency.

Assumptions of sample independence can lead to unreliable correlations as elucidated by Ranstam et al[21]. Conforti et al suggested a viable approach for considering spatial variations in soil organic matter content[22], where dimension reduction was performed by combining PLS scores with LME modeling, thereby accounting for the dependency structure in the measured data. However, this approach is unsuitable when the reference parameters or independent parameters are blinded or unknown, such as in independent testing or real-time applications. As predictors and responses are required to obtain the scores, PLS cannot be performed only on the spectral data. However, this practical limitation during independent testing could be easily circumvented by employing PCA reduction of the spectra. We also examined the potential of regularization by LASSO as it effectively shrinks the input spectra[23]. Since a significant number of spectral variables are penalized with zero (or absolute) coefficients, the resulting models are sparse and hence provide feature selection. Although the resulting models based on LASSO-LME are efficient, the penalization process in itself is computationally time consuming in comparison to other regression methods[24], such as the adopted PCA-LME approach.

In our application, the main dependency levels for cartilage dataset were joint level and articulating bones (n=2 per joint). The levels, however, could be tailored to suit the experimental design in specific biomedical applications. The results of the comparison (Table 2) show that standard versions of the regression models (PCR and LASSO) have slightly lower correlations in comparison to the LME-based regression models. This can be attributed to spatial dependency within the dataset, consistent

with Singer et al[9] and Ranstam et al[10,21]. Furthermore, the correlation between NIR spectra and biomechanical properties were better than with cartilage composition or structure. The thickness of cartilage and its NIR spectrum are highly correlated, as the cartilage thickness is representative of the path length, which affects the absorption. The biomechanical parameters of cartilage are highly influenced by the superficial cartilage[25]. On the other hand, the matrix composition and structure are an average of superficial, middle and deeper layers of cartilage, thus providing information on the entire tissue cross-section.

The present results indicate that the hybrid regression technique can effectively account for dependency in NIRS data. Identifying and including relevant dependent levels in the experimental design could be a limiting factor in this study; hence, careful selection of the levels is important. In some iterations, the test set includes values outside the calibration model range, thus making the model perform poorly. This is especially observed in predictions of small datasets, such as equilibrium modulus ($n = 202$). In addition, the relatively lower number of observations in the test set has a drastic effect on *RMSEP* value but not so much on correlation (Figure 4B). The performance of PCA-LME can potentially be further improved with variable selection methods, such as genetic algorithm[26]. Other alternatives to account for dependency in data structures include (but are not limited to) multilevel modeling[27] for hierarchical or clustered dataset and kringing[28], commonly utilized in genetic studies and geostatistics, respectively.

The present results support our hypothesis and thus advocate the application of LME-based regression technique for NIR spectroscopic characterization of cartilage. Importantly, this method can be easily extended to other spectroscopic applications.

5. Acknowledgments

Tuomas Selander is acknowledged for statistical consultation. This study was funded by the Academy of Finland (project 267551, University of Eastern Finland), Research Committee of the Kuopio University Hospital Catchment Area for the State Research Funding (project PY210 (5041750, 5041744 and 5041772), Kuopio, Finland), Instrumentarium Science Foundation (170033) and Tekniikan edistämissäätiö (8193). Dr. Afara acknowledges grant funding from the Finnish Cultural Foundation (00160079).

6. Author contributions

Prakash M.: Algorithm design and analysis.

Sarin, J.K.: Acquisition of data.

Rieppo, L: Supervision of statistical analyses**.**

Afara I.O.: Supervision of statistical analyses.

Töyräs J.: Study conception and design**.**

All authors contributed in the preparation and approval of the final submitted manuscript.

7. Conflict of Interest

The authors have no conflicts of interest in the execution of this study and preparation of the manuscript.

8. References

[1] J.H. Ahn, J.H. Wang, J.C. Yoo, Arthroscopic all-inside suture repair of medial meniscus lesion in anterior cruciate ligament—deficient knees: Results of second-look arthroscopies in 39 cases, Arthrosc. J. Arthrosc. Relat. Surg. 20 (2004) 936–945. doi:10.1016/j.arthro.2004.06.038.

[2] C.S.M. Brittberg, Mats MD, PHD; Winalski, Evaluation of Cartilage Injuries and Repair, J. Bone Jt. Surg. 85 (2003) 58–69. doi: 10.2106/00004623-200300002-00008

[3] B.H. Brismar, T. Wredmark, T. Movin, J. Leanderson, O. Svensson, Observer reliability in the arthoscopic classification of osteoarthritis of the knee, J. Bone Jt. Surg. 84 (2002) 42–47. doi:10.1302/0301-620X.84B1.11660.

[4] G. Spahn, H.M. Klinger, G.O. Hofmann, How valid is the arthroscopic diagnosis of cartilage lesions? Results of an opinion survey among highly experienced arthroscopic surgeons, Arch. Orthop. Trauma Surg. 129 (2009) 1117–1121. doi:10.1007/s00402-009-0868-y.

[5] G. Spahn, H.M. Klinger, M. Baums, M. Hoffmann, H. Plettenberg, A. Kroker, G.O. Hofmann, Near-infrared spectroscopy for arthroscopic evaluation of cartilage lesions: results of a blinded, prospective, interobserver study., Am. J. Sports Med. 38 (2010) 2516–21. doi:10.1177/0363546510376744.

[6] M. V. Padalkar, N. Pleshko, Wavelength-dependent penetration depth of near infrared radiation into cartilage, Analyst. 140 (2015) 2093–2100. doi:10.1039/C4AN01987C.

[7] G. Spahn, H. Plettenberg, H. Nagel, E. Kahl, H.M. Klinger, T. Mückley, M. Günther, G.O. Hofmann, J.A. Mollenhauer, Evaluation of cartilage defects with near-infrared spectroscopy (NIR): An ex vivo study, Med. Eng. Phys. 30 (2008) 285–292. doi:10.1016/j.medengphy.2007.04.009.

[8] I.O. Afara, H. Moody, S. Singh, I. Prasadam, A. Oloyede, Spatial mapping of proteoglycan content in articular cartilage using near-infrared (NIR) spectroscopy, Biomed. Opt. Express. 6 (2015) 144. doi:10.1364/BOE.6.000144.

[9] M. Singer, T. Krivobokova, A. Munk, B. de Groot, Partial least squares for dependent data, Biometrika. 103 (2016) 351–362. doi:10.1093/biomet/asw010.

[10] J. Ranstam, Repeated measurements, bilateral observations and pseudoreplicates, why does it matter?, Osteoarthr. Cartil. 20 (2012) 473–475. doi:10.1016/J.JOCA.2012.02.011.

[11] J.A. Cook, J. Ranstam, Statistical models and confounding adjustment, Br. J. Surg. 104 (2017) 786–787. doi:10.1002/bjs.10245.

[12] H. Abdi, L.J. Williams, Principal component analysis, Wiley Interdiscip. Rev. Comput. Stat. 2 (2010) 433–459. doi:10.1002/wics.101.

[13] J. Schelldorfer, P. Bühlmann, S. Van De Geer, Estimation for High-Dimensional Linear Mixed-Effects Models Using ℓ1-Penalization, Scand. J. Stat. 38 (2011) 197–214. doi:10.1111/j.1467-9469.2011.00740.x.

[14] R. Tibshirani, Regression shrinkage and selection via the lasso: a retrospective, J. R. Stat. Soc. Ser. B (Statistical Methodol. 73 (2011) 273–282. doi:10.1111/j.1467-9868.2011.00771.x.

[15] M. Prakash, A. Joukainen, J.K. Sarin, L. Rieppo, I.O. Afara, J. Töyräs, Near-infrared Spectroscopy Based Arthroscopic Evaluation of Human Knee Joint Cartilage, Through Automated Selection of an Anatomically Specific Regression Model, in: Biophotonics Congr. Biomed. Opt. Congr. 2018, OSA, Washington, D.C., 2018: p. OF4D.3. doi:10.1364/OTS.2018.OF4D.3.

[16] M.G. Kenward, J.H. Roger, Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood, Biometrics. 53 (1997) 983. doi:10.2307/2533558.

[17] J.K. Sarin, M. Amissah, H. Brommer, D. Argüelles, J. Töyräs, I.O. Afara, Near Infrared Spectroscopic Mapping of Functional Properties of Equine Articular Cartilage, Ann. Biomed. Eng. 44 (2016) 3335–3345. doi:10.1007/s10439-016-1659-6.

[18] J.K. Sarin, L. Rieppo, H. Brommer, I.O. Afara, S. Saarakkala, J. Töyräs, Combination of optical coherence tomography and near infrared spectroscopy enhances determination of articular cartilage composition and structure, Sci. Rep. 7 (2017) 10586. doi:10.1038/s41598-017-10973-z.

[19] R.K. Korhonen, M.S. Laasanen, J. Töyräs, R. Lappalainen, H.J. Helminen, J.S. Jurvelin, Fibril reinforced poroelastic model predicts specifically mechanical behavior of normal, proteoglycan depleted and collagen degraded articular cartilage, J. Biomech. 36 (2003) 1373–1379. doi:10.1016/S0021-9290(03)00069-1.

[20] S. Nakagawa, H. Schielzeth, A general and simple method for obtaining R^2 from generalized linear mixed-effects models, Methods Ecol. Evol. 4 (2013) 133–142. doi:10.1111/j.2041-210x.2012.00261.x.

[21] J. Ranstam, J.A. Cook, Considerations for the design, analysis and presentation of in vivo studies, Osteoarthr. Cartil. 25 (2017) 364–368. doi:10.1016/j.joca.2016.06.023.

[22] M. Conforti, A. Castrignanò, G. Robustelli, F. Scarciglia, M. Stelluti, 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 (2015) 60–67. doi:10.1016/j.catena.2014.09.004.

[23] E. Andries, S. Martin, Sparse Methods in Spectroscopy: An Introduction, Overview, and Perspective, Appl. Spectrosc. 67 (2013) 579–593. doi:10.1366/13-07021.

[24] M. Prakash, J.K. Sarin, L. Rieppo, I.O. Afara, J. Töyräs, Optimal Regression Method for Near-Infrared Spectroscopic Evaluation of Articular Cartilage, Appl. Spectrosc. 71 (2017) 2253–2262. doi:10.1177/0003702817726766.

[25] R.. Korhonen, M. Wong, J. Arokoski, R. Lindgren, H.. Helminen, E.. Hunziker, J.. Jurvelin, Importance of the superficial tissue layer for the indentation stiffness of articular cartilage, Med. Eng. Phys. 24 (2002) 99–108. doi:10.1016/S1350-4533(01)00123-0.

[26] A.S. Barros, D.N. Rutledge, Genetic algorithm applied to the selection of principal components, Chemom. Intell. Lab. Syst. 40 (1998) 65–81. doi:10.1016/S0169-7439(98)00002-1.

444 [27] H. Goldstein, W. Browne, J. Rasbash, Multilevel modelling of medical data,
445 Stat. Med. 21 (2002) 3291–3315. doi:10.1002/sim.1264.

446 [28] D. Allard, J.-P. Chilès, P. Delfiner: Geostatistics: Modeling Spatial Uncertainty,
447 Math. Geosci. 45 (2013) 377–380. doi:10.1007/s11004-012-9429-y.

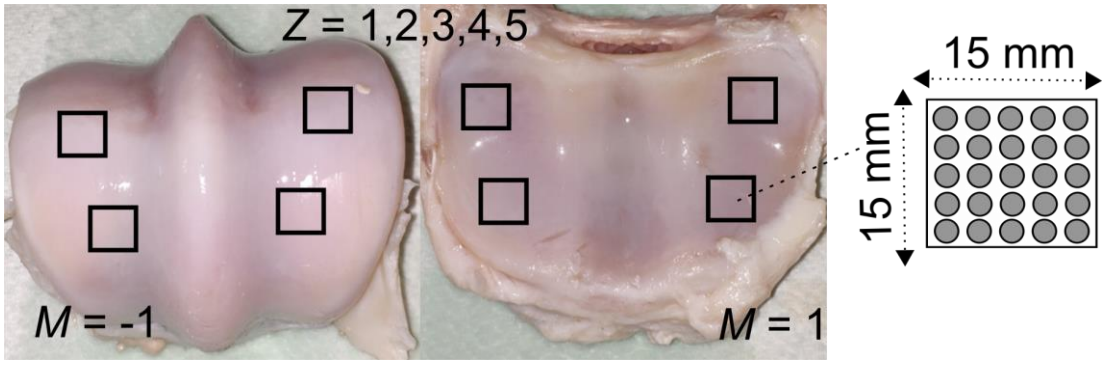


Figure 1: Areas of interest (AI) marked (black squares) on the articulating surfaces of equine metacarpophalangeal joint. In this study, grouping information is on two dependency levels, i.e. joint level and bone level, which is held in ***Z*** (sample count×5) and ***M*** (sample count×1) design matrices. Each AI (15mm **×** 15mm) has 25 equidistant measurement locations.

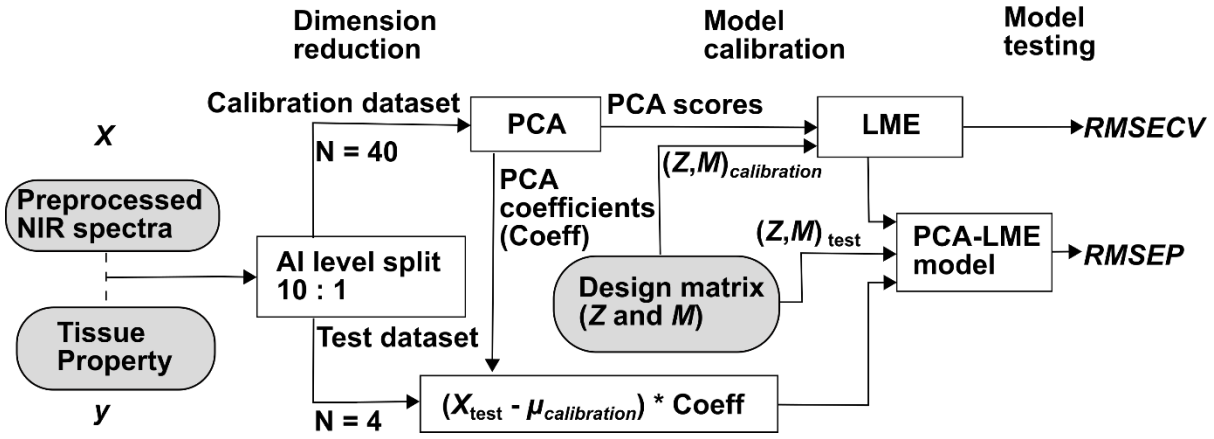


Figure 2: Schematic chart of PCA-LME regression technique, which combines principal component analysis (PCA) and linear mixed effects (LME) regression technique. Preprocessed near infrared (NIR) spectra (***X***), tissue property (***y***), and design matrices ***Z*** and ***M*** are the inputs. Design matrices ***Z*** and ***M*** are the mixed effects in LME modeling. Model performance was evaluated using the root mean square errors of cross validation (*RMSECV*) and prediction (*RMSEP*).

465 **Table 1**: Summary of datasets

466

Measurements	**N**	**Additional details**
Equine cartilage dataset	5 joints, 44 AIs	Surgical extraction
NIR spectral measurements	869	Absorbance spectroscopy (700-1050 nm)
Thickness (mm)	869	Optical coherence tomography
Equilibrium Modulus (MPa)	202	Indentation testing
Dynamic modulus (MPa)	202	Indentation testing
PG content (AU)	530	FTIR microspectroscopy
Collagen content (AU)	530	FTIR microspectroscopy
Collagen orientation (º)	530	Polarized light microscopy

467

468 **Table 2**: The mean and range of different cartilage properties. A comparison of the assessment statistics of standard regression
469 techniques and introduced hybrid regression technique. The white rows represent PCR and PCA-LME, whereas grey rows represent
470 LASSO and LASSO-LME. *ρ* (Spearman's rank correlation), root mean square errors of cross validation (*RMSECV*) and prediction
471 (*RMSEP*) are shown for all predicted parameters.

Property	Mean (Range)	Standard regression				Hybrid regression			
		Calibration		Test		Calibration		Test	
		ρ	*RMSECV %*	*ρ*	*RMSEP %*	*ρ*	*RMSECV %*	*ρ*	*RMSEP %*
Thickness (mm)	0.89 (0.32 – 1.81)	0.78	13.94	0.67	18.54	0.85	12.02	0.74	16.40
		0.86	11.20	0.57	20.17	0.87	11.22	0.65	18.36
Dynamic Modulus (MPa)	9.43 (0.24 – 23.3)	0.49	28.07	0.46	37.80	0.69	22.75	0.56	34.71
		0.61	24.34	0.29	42.73	0.67	23.29	0.27	39.58
PG content (AU)	6.31 (0.60 – 14.71)	0.45	18.17	0.34	22.42	0.52	17.50	0.42	22.54
		0.51	17.10	0.34	22.26	0.53	17.52	0.37	21.89
Equilibrium Modulus (MPa)	2.00 (0.03 – 5.38)	0.44	28.91	0.32	37.50	0.63	24.28	0.48	35.02
		0.59	24.58	0.38	33.90	0.65	24.72	0.46	34.84
Collagen Content (AU)	33.35 (12.16 – 64.39)	0.41	19.79	0.35	23.34	0.43	19.21	0.27	24.90
		0.40	19.67	0.29	22.90	0.41	20.09	0.32	23.14
Collagen Orientation angle (º)	71.12 (37.13 – 83.75)	0.31	22.45	0.27	25.01	0.37	21.79	0.23	25.35
		0.41	21.39	0.25	23.54	0.41	21.91	0.27	24.29

472

Figure 3

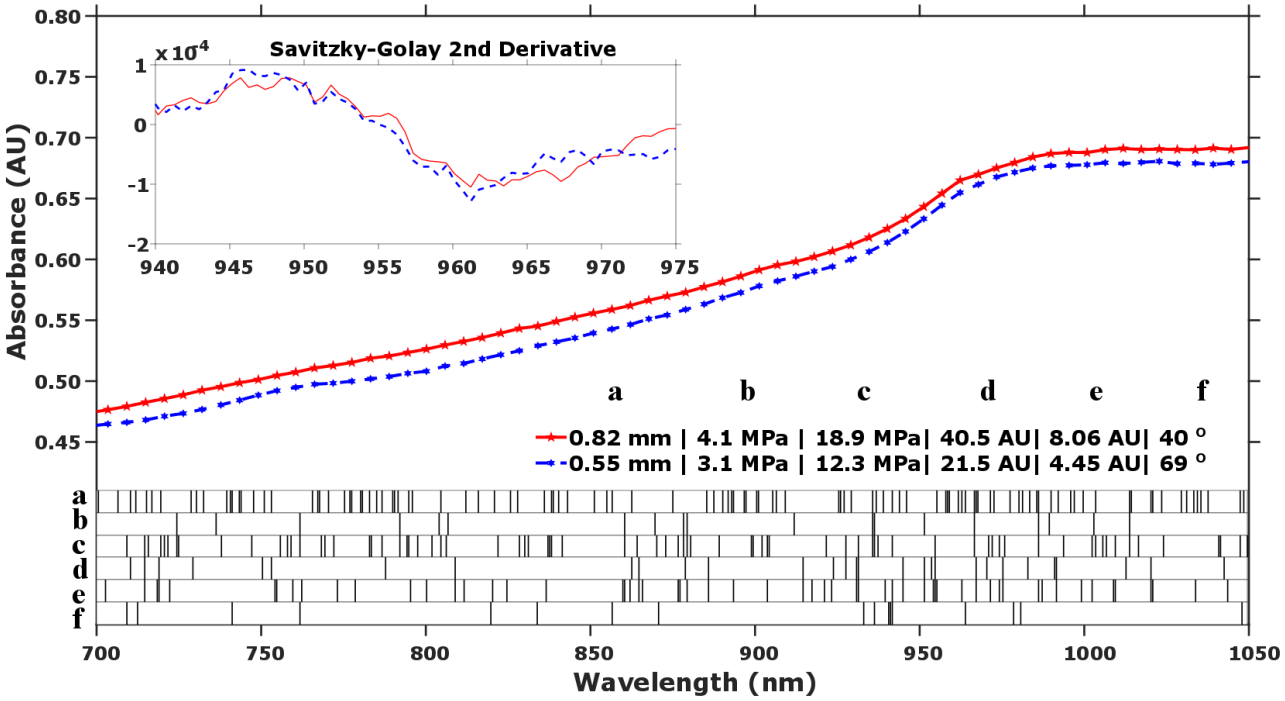
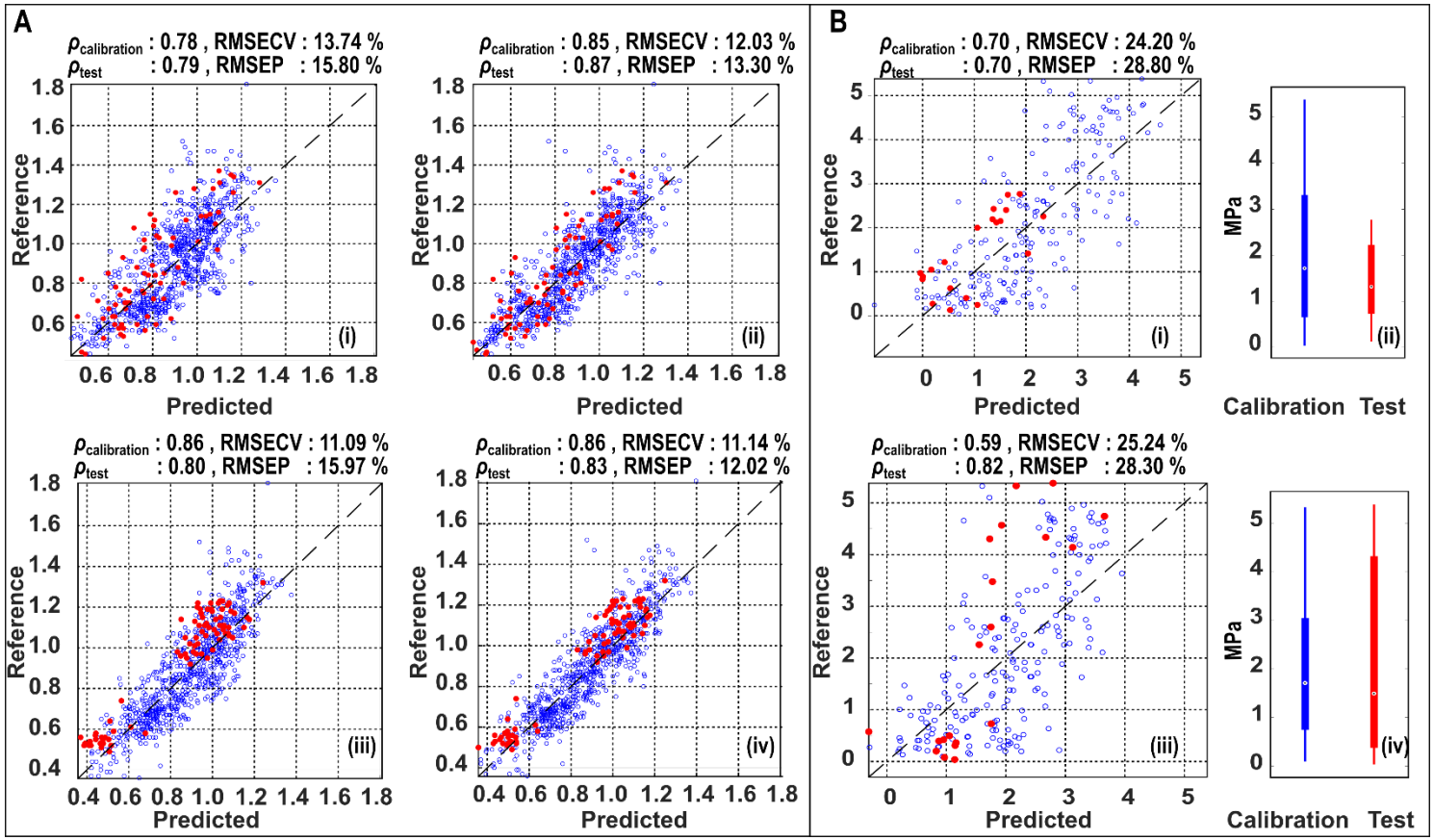


Figure 3: Representative NIR spectra (700 to 1050 nm) for samples with different cartilage (a) thickness (mm), (b) equilibrium modulus (MPa), (c) dynamic modulus (MPa), (d) collagen content (AU), (e) proteoglycan content (AU), and (f) collagen orientation angle (º). The top inset shows second derivative Savitzky-Golay preprocessed spectra in 940 to 975 nm. Bottom inset shows the least absolute shrinkage and selection operator (LASSO) based feature selection of the spectra.



488

489 **Figure 4**: Predicted (x axis) vs. reference (y axis) values of cartilage thickness [mm, A] and equilibrium modulus [MPa, B].
490 Performance on calibration (blue, unfilled) and test set (red, filled) as predicted by PCR [A,(i)], PCA-LME[A,(ii)], LASSO[A,(iii)] and
491 LASSO-LME[A,(iv)] regression models. Effect of outliers on LASSO-LME [B] model performance when the test set range is within
492 [(i),(ii)] and outside [(iii),(iv)] the calibration range.