International Workshop on Spectroscopy & Chemometrics 2023

Poster Abstracts

Bayesian partial least squares regression for milk trait predictions from spectral data

Szymon Urbas, Donagh Berry, Isobel Claire Gormley
School of Computer Science, University College Dublin, Belfield, Dublin 4.
Teagasc, Animal and Grassland Research and Innovation Centre, Moorepark Fermoy, Co. Cork, Ireland

Abstract:

Partial least squares (PLS) regression is a popular method for making predictions from spectral data in the agri-food sector. For example, traits such as protein content or heat stability can be predicted from high-dimensional mid-infrared spectral data of milk samples. The PLS regression approach is based on maximising the covariance between the predictor and response vectors through a latent-space decomposition. However, the method does not arise from the statistical inference of any particular probabilistic model. Thus it cannot account for parameter uncertainty and does not easily lend itself to modifications. We introduce a Bayesian latent-variable model which emulates the desirable properties of PLS. The modelling framework is very flexible and allows for modifications inspired by sparse regression approaches such as spike-and-slab or Bayesian LASSO. We compare the method to standard approaches when making predictions from mid-infrared spectral data and surface-enhanced Raman spectroscopy data. In addition to accurate point predictions, the statistical model provides prediction intervals with the correct coverage and offers potential extensions based on different correlation structures or domain-informed priors.

Prediction of soil bulk density in agricultural soils using mid-infrared spectroscopy

Longnan Shi[a,b], Sharon O'Rourke[b], Felipe Bachion de Santana[a], Karen Daly[a]
[a]*Environment Soils and Land Use Department, Teagasc, Johnstown Castle,Wexford, Ireland.*
[b]*School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4.*

Abstract

Traditionally, the laboratory analysis of soil bulk density (BD) is time-consuming, labour

intensive and expensive, especially for a national-scale soil assessment. Hence, how to fill the omissions of this key physical parameter for all or some records in soil databases is widely discussed. This study employed different chemometric and machine learning algorithms to estimate BD in Irish soil from 671 horizon-based samples from mid-infrared (MIR) spectral libraries by partial least square regression (PLSR), random forest, Cubist and support vector machine (SVM). The best performance was observed for the SVM model with higher ratio of performance to interquartile distance (RPIQ = 3.61) and R^2 (0.81) values and lower root mean square error of prediction (RMSEP = 0.132). Moreover, leave-one-out cross-validation was employed on SVM, and a generalization error of predicting unknown samples with a spectral soil BD model was calculated and compared to the published traditional BD pedotransfer functions (PTFs). Either on overall or specific depth categories, the spectral soil BD model is significantly better rather than traditional PTFs. Hence, high accuracy and the homogeneity of performance on different depth layers could be noteworthy strengths of spectral modelling techniques when carrying out national soil surveys and large scale carbon stock assessments.

Keywords: Soil; Bulk density; Mid-infrared; Spectroscopy; Chemometrics; Machine learning

NIR spectra of faeces for the prediction of feed efficiency in dairy cows

[1]Silvia Ampuero Kragten, [2,3]Thorsten Haak, [3]Karl-Heinz Südekum, [2]Fredy Schori

[1]Agroscope, Method Development and Analytics, 1725 Posieux, Switzerland, [2]Agroscope, Ruminant Nutrition and Emissions, 1725 Posieux, Switzerland, [3]University Bonn, Institute of Animal Science, Bonn, Germany

Abstract

While optimizing feed costs and economical use of resources, improving feeding efficiency of dairy cows should help limiting the nitrogen and greenhouse emissions associated with milk production. However, a wider evaluation of feed efficiency is hindered by the need for individual animal feed intake information, which for a large number of animals is too expensive to obtain and time consuming with today's tools. The present study, as part of a bigger project, aims at the assessment of the potential of NIRS to contribute to the prediction of feed efficiency (FE) in dairy cows including organic grazing cows. FE traits were evaluated for 86 lactating cows (60 Holstein, 16 Swiss Fleckvieh), (grazing and stall-fed), primiparous and multiparous, in mid and late lactation. Freeze-dried faecal samples were scanned in diffuse reflectance mode with a NIRFlex N-500 (Büchi, CH-Flawil), with 3 replicate spectra per sample, 21 scans per replicate, in the range of 4000 cm^{-1} to 10000 cm^{-1}. 132 samples were a 7-days mass ratio pool of a one-week measurement period. NIRS models were developed with PLS algorithms including: snv (normalization standard normal variate), mf (normalization MSC Full) and db1g2 (1st derivative BCAP Gap 2). NIRS models (independent calibration (cal) and validation (val) sets,

~70 and 30 % of samples resp.) showed moderate to good predictive potential for FE traits [R^2 (cal/val): 0.75/0.76, 0.58/0.58 and 0.88/0.82; residual prediction deviation (SD calibration-set/standard error of prediction): 2.2, 1.6 and 2.8 for FCE, RFI and RNI respectively]. These models should be incremented with samples from diverse origins (breeds, diets, production and environmental conditions) to improve their robustness.

RGB-2-Hyper-Spectral Image Reconstruction for Food Science Using Encoder/Decoder Neural Architectures

Robert Williamson[1], Jesus Martinez del Rincon[1], Anastasios Koidis[1], Carlos Reano[2]

[1] Queens University, Belfast, UK
[2] Universitat de Valencia, Spain

Abstract:

Hyper-spectral imaging captures spatial and spectral information of a subject. This is used for the identification of substances within a scene, and food analysis. Presented is an investigation into the capabilities of encoder/decoder deep learning architectures for hyper-spectral image reconstruction from RGB images. For this analysis state-of-the-art (SOTA) techniques for hyper-spectral image reconstruction and other architectures from different fields have been used. Our approach examines a food science case study, using a CPU-based server and different accelerators. An in-house multi-sensor setup was used to capture the dataset which contains hyper-spectral images of twenty slices of different Spanish Ham in the range of 400-100~nm and their analogous RGB images. The results show no degradation in the output when moving outside of the visual range. This study shows that the SOTA methods for reconstructing from RGB do not produce the most accurate reconstruction of the spectral domain within the range of 400-1000~nm.

Exploration of NIR region to quantify macro-components of raw milk

H M Hussain Khan[a,b], Ultan McCarthy [b], Norah O'Shea [a]

[a,]*Food Chemistry & Technology Department, Food Research Center, Teagasc Moorepark, Fermoy, Ireland.*
[b]*School of Science and Computing, Department of Land Sciences, South East Technolological University, Waterford, Ireland.*

Abstract

Broad overlapping bands arise due to multiple overtones are the characteristic property of NIR spectroscopy. Conventionally, the NIR spectral range (800-2500 nm) is divided into four regions:

combination (2000-2500 nm), 1st overtone (1500-2100), 2nd overtone (1000-1600) and 3rd overtone (700-1100), which differ in penetration power and signal strength. Therefore, the study compared NIR regions with full spectral range to quantify raw milk macro-components (fat and protein). NIR spectra of 150 raw milk samples were acquired in transmission mode (1-mm path length), and calibration models were developed using PLS regression with the complete spectral range and individual regions mentioned above. It was observed that the prediction performance of models developed using 1st overtone region of NIR outperforms other regions and complete spectral range, which can be linked to the trade-off between signal strength and penetration power. Although the penetration power increases in the short-wavelength region, the signal strength decreases due to a lower probability of occurrence. Conversely, the scattering of light increases at longer wavelengths, reducing the penetration power and increasing noise in the signal. Moreover, the influence of animal physiology on water absorption regions increases the complexity of models developed using the complete spectral range. Therefore, limiting the spectral range to 1st overtone region can improve the prediction performance and robustness of models developed for the macro-components of raw milk.

A comprehensive study using Hyperspectral Imaging (HSI) to evaluate the quality of meat products

Wenyang Jia [a], Alessandro Ferragina [b], Jyoti Prakash Mishra [c], Joseph Peller [d], Sara Erasmus [e], Saskia van Ruth [a,e,f], Anastasios Koidis [a,*]

[a] *Institute for Global Food Security, School of Biological Sciences, Queen's University, 19 Chlorine Gardens, Belfast BT9 5DL, Northern Ireland, UK*
[b] *Teagasc Food Research Centre, Food Quality and Sensory Science Department, Dublin, Ireland*
[c] *Teagasc Food Research Centre, Ashtown, Dublin , Ireland*
[d] *Greenhouse Horticulture, Wageningen University & Research, Wageningen, The Netherlands*
[e] *Food Quality and Design Group, Wageningen University and Research, P.O. Box 17, 6700 AA, Wageningen, the Netherlands*
[f] *UCD School of Agriculture and Food Science, University College Dublin, 4 Dublin, Ireland*
[*] *Corresponding author at: Institute for Global Food Security, School of Biological Sciences. Queen's University Belfast, 19 Chlorine Gardens, Belfast, BT9 5DJ. UK.*

Abstract

In response to the growing demand for rapid and real-time assessment amidst mounting concerns regarding meat quality, this study employed HSI using novel approaches. At Queen's University Belfast (QUB), HSI tools were utilised for adulteration identification, myoglobin evaluation, and

penetration study. NIR-HSI (900 - 1700 nm) was used to identify multi-adulterants (beef liver, beef heart and pork) in ground beef, sensitivity and specificity exhibit a value of 1.00 using the chemometrics method for the calibration and validation sets. Hyperspectral Imaging-Root Mean Square (HSI-RMS) value was proposed to clarify the complexity of HSI data. VIS-HSI (400 - 1000 nm) and NIR-HSI were used to assess the changes in myoglobin content of packaged veal products during storage following three levels of data fusion: data-level, feature-level, and decision-level. Using decision-level fusion, relative oxymyoglobin content (%) was predicted (R^2CV=0.71 and RMSECV=1.11%) for packaged products. VIS-HSI and NIR-HSI were also used to evaluate the penetration level using ham products, self-organising map analysis and HSI-RMS value were used to explore the influence of thickness for HSI data. Following a series of analyses and summaries conducted at QUB, this study highlights the effectiveness of HSI coupled with chemometrics and statistical methods in identifying adulteration in beef products. These data fusion approaches improve the accuracy of the analysis, which helps develop strategies for tracking the shelf life and quality of veal products for veal producers. The penetration study proposes an exploratory analysis for evaluating the penetration using HSI tools, which provides a perspective on obtaining more representative HSI data in the future. This comprehensive study advances innovative methods for addressing meat quality, and developing novel strategies to ensure the quality of meat products when using spectral-imaging related techniques.