Data in Brief ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Performances of full cross-validation partial least squares regression models developed using Raman spectral data for the prediction of bull beef sensory attributes

Ming Zhao [a,*], Yingqun Nian [b,c], Paul Allen [b], Gerard Downey [a], Joseph P. Kerry [c], Colm P. O'Donnell [a]

[a] *School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4, Ireland*
[b] *Department of Food Quality and Sensory Science, Teagasc Food Research Centre, Ashtown, Dublin 15, Ireland*
[c] *School of Food and Nutritional Sciences, University College Cork, Cork, Ireland*

ARTICLE INFO

Article history:
Received 10 February 2018
Received in revised form 13 February 2018
Accepted 17 April 2018

Keywords:
Selected Raman shift ranges
Sensory attributes
Bull beef
Partial least squares regression models

ABSTRACT

The data presented in this article are related to the research article entitled "Application of Raman spectroscopy and chemometric techniques to assess sensory characteristics of young dairy bull beef" (Zhao et al., 2018) [1]. Partial least squares regression (PLSR) models were developed on Raman spectral data pre-treated using Savitzky Golay (S.G.) derivation (with 2nd or 5th order polynomial baseline correction) and results of sensory analysis on bull beef samples ($n = 72$). Models developed using selected Raman shift ranges (i.e. 250–3380 cm^{-1}, 900–1800 cm^{-1} and 1300–2800 cm^{-1}) were explored. The best model performance for each sensory attributes prediction was obtained using models developed on Raman spectral data of 1300–2800 cm^{-1}.

DOI of original article: https://doi.org/10.1016/j.foodres.2018.02.007

* Corresponding author.
E-mail address: ming.zhao@ucd.ie (M. Zhao).

https://doi.org/10.1016/j.dib.2018.04.056
2352-3409/

Please cite this article as: M. Zhao, et al., Performances of full cross-validation partial least squares regression models developed using Raman spectral data for the prediction of bull beef sensory attributes, Data in Brief (2018), https://doi.org/10.1016/j.dib.2018.04.056

Specifications Table

Subject area	*Spectroscopy, Chemometrics*
More specific subject area	*Performance of PLSR models developed using selected Raman shift ranges (i.e. 250–3380 cm^{-1}, 900–1800 cm^{-1} and 1300–2800 cm^{-1})*
Type of data	*Table*
How data was acquired	*Raman spectroscopy, Results of sensory analysis, Chemometrics*
Data format	*.doc*
Experimental factors	*Raman spectral data were pre-treated using Savitzky Golay (S.G.) derivation with 2nd or 5th order polynomial baseline correction.*
Experimental features	–
Data source location	*School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4, Ireland*
Data accessibility	*Data is with this article*

Value of the data

- To demonstrate PLSR models developed using Raman spectra in the 1300–2800 cm^{-1} range can give best prediction performance on sensory attributes of bull beef.
- Results of this work are in agreement with a previous study by [2] that the Raman frequency range of 1300–2800 cm^{-1} is the most suitable range for prediction of bull beef eating quality parameters.
- This data suggested other researchers to select an optimal Raman shift range for further meat science studies.

1. Data

PLSR models were developed on Raman data pre-treated using Savitzky Golay (S.G.) derivation with 2nd and 5th order polynomial baseline correction. Prediction performance of models developed using selected Raman shift ranges (i.e. 250–3380 cm^{-1}, 900–1800 cm^{-1} and 1300–2800 cm^{-1}) were summarized in Table 1. PLS models developed using S.G. derivation pre-treated Raman spectra in the 1300–2800 cm^{-1} range performed best (R^2CV values of 0.36–0.84) while spectra in the range 900–1800 cm^{-1} performed worst (R^2CV values of 0.03–0.66).

2. Experimental design, materials and methods

For the prediction of beef sensory attributes, partial least squares regression (PLSR) models were developed using pre-processed Raman spectroscopic data (X data) collected on the 21st day post-mortem using pre-selected frequency ranges (i.e. 250–3380 cm^{-1}, 900–1800 cm^{-1}, 1300–2800 cm^{-1}); these were selected on the basis of spectral signal intensities. Measured values of sixteen sensory attributes were used as individual Y variable for PLS regression. Leave-one-out cross-validation was performed to evaluate the performance of PLSR models using parameters such as root mean square error of calibration (RMSEC) and cross-validation (RMSECV), the coefficient of determination on calibration (R2C) and cross-validation (R^2CV) and the bias which is calculated as the difference between the average of actual and predicted values for each data set [3]. For a satisfactory prediction performance, the value of R^2 is expected to be close to 1 while values of RMSECV and bias are expected to be close to 0.

Please cite this article as: M. Zhao, et al., Performances of full cross-validation partial least squares regression models developed using Raman spectral data for the prediction of bull beef sensory attributes, Data in Brief (2018), https://doi.org/10.1016/j.dib.2018.04.056

Table 1
Summary of full cross-validation PLSR model performances for sensory attributes prediction using all beef samples ($n = 72$).

	Data type	Wavelength range(cm-1)	Spectral variables	PLS loadings	R^2C	RMSEC	R^2CV	RMSECV	Bias
Aroma	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	83	2	0.85	3.92	0.76	5.01	-0.003
		900-1800	45	1	0.64	6.06	0.43	7.73	-0.015
		250-3580	150	4	0.88	3.46	0.66	6.02	0.031
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	81	3	0.82	4.26	0.64	6.18	-0.03
		900-1800	36	1	0.48	7.25	0.41	7.86	-0.033
		250-3580	176	6	0.9	3.21	0.66	6	0.043
Tenderness	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	41	1	0.54	8.84	0.45	9.83	0.083
		900-1800	44	1	0.57	8.58	0.46	9.74	-0.132
		250-3580	33	2	0.39	10.18	0.26	11.36	0.004
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	45	1	0.63	7.91	0.55	8.9	0.001
		900-1800	40	2	0.67	7.5	0.49	9.42	0.051
		250-3580	69	3	0.64	7.79	0.41	10.19	-0.006
Juciness	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	121	2	0.69	5.41	0.47	7.1	0
		900-1800	87	1	0.43	7.28	0.32	8.06	-0.098
		250-3580	298	6	0.96	2	0.61	6.07	0.109
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	109	2	0.74	4.95	0.55	6.55	-0.094
		900-1800	95	2	0.73	5	0.48	7.07	-0.069
		250-3580	90	4	0.8	4.37	0.63	5.92	-0.014
Cohensiveness	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	85	2	0.72	5.15	0.61	6.12	-0.018
		900-1800	32	2	0.46	7.11	0.22	8.71	0.309
		250-3580	104	2	0.42	7.4	0.25	8.52	0.011
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	93	2	0.76	4.74	0.64	5.93	-0.003
		900-1800	84	2	0.69	5.41	0.45	7.32	0.035
		250-3580	150	5	0.88	3.4	0.64	5.93	0.033
E.D.	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	40	2	0.55	8.49	0.39	9.98	-0.061
		900-1800	52	1	0.49	9.77	0.41	10.63	0.126
		250-3580	86	7	0.8	5.58	0.26	10.97	0.058
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	59	2	0.66	7.4	0.5	9.07	-0.034
		900-1800	49	1	0.55	8.44	0.46	9.37	-0.095
		250-3580	94	3	0.56	8.32	0.33	10.45	0.012
Chewiness	S.G. 2nd der. Using 2nd polynomial	1300-2800	51	2	0.68	7.79	0.55	9.33	0.031

Table 1 (*continued*)

	with 9 smoothing points								
		900-1800	52	1	0.49	9.77	0.41	10.63	0.126
		250-3580	61	2	0.37	10.82	0.22	12.25	0.021
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	69	2	0.81	5.99	0.63	8.48	0.182
		900-1800	44	1	0.56	9.04	0.48	9.98	0.03
		250-3580	80	2	0.53	9.39	0.36	11.08	0.12
Fattiness	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	38	2	0.67	2.68	0.48	3.38	0.032
		900-1800	19	3	0.39	3.63	0.17	4.28	-0.049
		250-3580	59	4	0.67	2.66	0.43	3.57	0.038
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	55	2	0.79	2.13	0.62	2.88	0.025
		900-1800	61	2	0.71	2.5	0.48	3.39	0.042
		250-3580	50	3	0.55	3.1	0.41	3.6	0.103
Stringy	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	79	3	0.78	2.27	0.52	3.37	-0.059
		900-1800	48	2	0.59	3.08	0.38	3.85	-0.032
		250-3580	144	3	0.68	2.7	0.47	3.54	0.006
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	81	1	0.7	2.63	0.63	2.98	0.005
		900-1800	26	1	0.37	3.81	0.29	4.12	-0.011
		250-3580	140	6	0.87	1.75	0.55	3.27	0.079
Astringent	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	61	2	0.63	2.66	0.52	3.37	-0.059
		900-1800	69	1	0.69	2.44	0.61	2.76	0.045
		250-3580	92	5	0.78	2.04	0.51	3.09	0.036
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	90	3	0.89	1.44	0.76	2.16	-0.001
		900-1800	50	1	0.71	2.32	0.66	2.57	0.022
		250-3580	128	6	0.85	1.67	0.63	2.68	0.017
Beef flavour	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	100	3	0.91	3.01	0.8	4.65	0.128
		900-1800	85	2	0.67	5.82	0.46	7.57	-0.043
		250-3580	364	4	0.88	3.46	0.53	7.05	0.035
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	126	3	0.88	3.46	0.77	4.98	0.046
		900-1800	96	2	0.71	5.47	0.51	7.21	0.032
		250-3580	323	6	0.96	2.05	0.66	6	0.186
Metallic	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	64	2	0.55	2.47	0.36	3.01	-0.01
		900-1800	48	2	0.49	2.63	0.36	3.01	-0.014
		250-3580	125	4	0.64	2.23	0.27	3.2	-0.0007
	S.G. 1st der. Using 5th polynomial	1300-2800	67	2	0.72	1.95	0.54	2.55	-0.01

Table 1 (*continued*)

	with 7 smoothing points +nor.u.v.								
		900-1800	62	2	0.61	2.31	0.38	2.96	0.025
		250-3580	178	7	0.92	1.04	0.53	2.58	-0.014
Rancid	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	84	2	0.61	1.17	0.45	1.41	-0.005
		900-1800	24	2	0.35	1.51	0.2	1.7	0.001
		250-3580	207	4	0.77	0.89	0.4	1.48	0.029
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	98	2	0.64	1.13	0.52	1.31	0.016
		900-1800	38	2	0.43	1.42	0.25	1.65	-0.002
		250-3580	206	5	0.84	0.76	0.55	1.27	0.035
Flavour length	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	99	3	0.94	2.66	0.84	4.21	0.012
		900-1800	69	2	0.65	6.22	0.42	8.14	-0.126
		250-3580	372	3	0.78	4.97	0.5	7.5	0.038
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	122	3	0.87	3.81	0.64	6.35	-0.175
		900-1800	83	2	0.68	5.95	0.46	7.83	0.07
		250-3580	325	2	0.66	6.12	0.49	7.58	0.086
Res-metalic	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	77	2	0.66	2.12	0.49	2.64	-0.0003
		900-1800	25	1	0.11	3.43	0.03	3.63	-0.06
		250-3580	89	4	0.72	1.92	0.43	2.79	0.018
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	76	2	0.83	1.49	0.61	2.31	-0.027
		900-1800	36	2	0.52	2.52	0.28	3.12	0.019
		250-3580	136	6	0.86	1.38	0.58	2.38	-0.005
Res-fattiness	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	99	3	0.71	1.89	0.51	2.48	-0.005
		900-1800	49	1	0.48	2.53	0.39	2.77	-0.009
		250-3580	199	4	0.82	1.48	0.51	2.48	0.047
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	70	2	0.66	2.05	0.54	2.42	0.01
		900-1800	63	2	0.75	1.77	0.54	2.42	-0.017
		250-3580	141	7	0.9	1.08	0.59	2.29	0.056
Dryness	S.G. 2nd der. Using 2nd polynomial with 9 smoothing points	1300-2800	137	2	0.67	2.54	0.48	3.25	-0.023
		900-1800	76	2	0.49	3.17	0.35	3.63	0.023
		250-3580	282	4	0.81	1.95	0.49	3.2	-0.049
	S.G. 1st der. Using 5th polynomial with 7 smoothing points +nor.u.v.	1300-2800	101	3	0.79	2.04	0.59	2.89	0.06
		900-1800	53	2	0.63	2.79	0.41	3.44	0.033
		250-3580	231	6	0.9	1.41	0.45	3.34	0.109

PLSR, partial least squares regression models; S.G., Savitzky Golay; der., derivatives; nor.u.v., normalisation on unit vectors; # PLS loadings, number of PLS loadings; R^2C, coefficient determination of calibration; RMSEC, root mean square error of calibration; R^2CV, correlation coefficient of determination in cross-validation; RMSECV, root mean square error of cross-validation; IT, Initial Tenderness; ED, Ease of Disintegration; Res-, Residual (after effects); *n*, numbers of samples. (Note: The best performed PLS models developed in the Raman shifts of 1300–2800 cm^{-1} were highlighted in yellow).

Acknowledgments

The authors wish to thank Carol Griffin and Carmel Farrell (Teagasc Ashtown) for training of the beef sensory panel.

Funding sources

This work was funded by the Teagasc Walsh Fellowship Scheme. We wish to acknowledge Teagasc Johnstown Castle Research Centre for rearing the dairy bulls and Dawn Meats for slaughtering the dairy bulls.

Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.04.056.

References

[1] M. Zhao, Y. Nian, P. Allen, G. Downey, J.P. Kerry, C.P. O'Donnell, Application of Raman spectroscopy and chemometric techniques to assess sensory characteristics of young dairy bull beef (In press), Food Res. Int. (2018), http://dx.doi.org/10.1016/j.foodres.2018.02.007.

[2] Y. Nian, M. Zhao, C.P. O'Donnell, G. Downey, J.P. Kerry, P. Allen, Assessment of physico-chemical traits related to eating quality of young dairy bull beef at different ageing times using Raman spectroscopy and chemometrics, Food Res. Int. 99 (2017) 778–789.

[3] T. Næs, T. Isaksson, T. Fearn, T. Davis, A User-friendly Guide to Multivariate Calibration and Classification. Chapter 13. Validation, NIR Publications, UK (2002) 163–167.