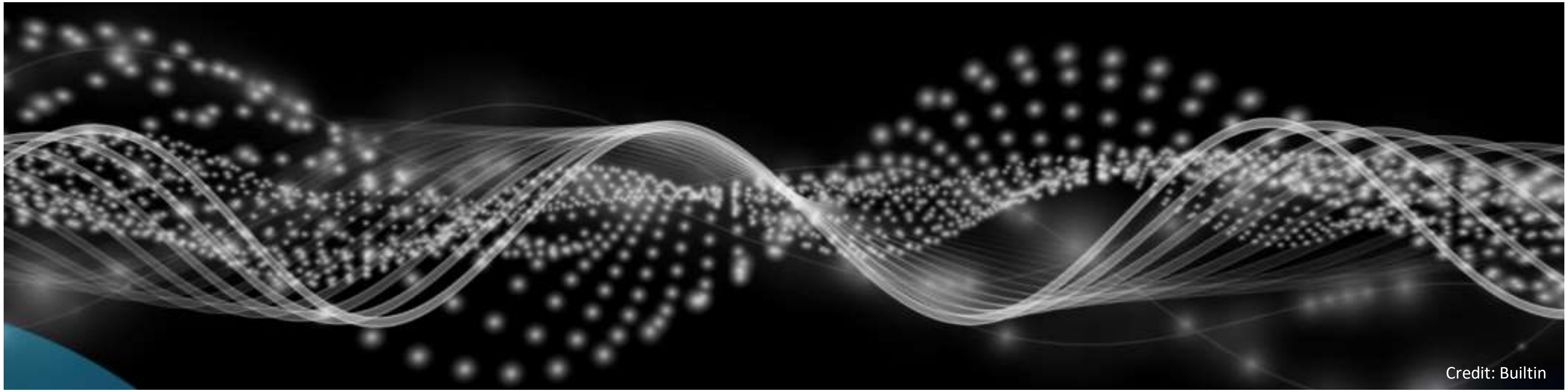
Another Pipeline in Local Partial Least Squares Regression (LPLS) Methods: Assessing the Impact of Wavelet Transform Integration

Antoine Deryck[a], Andreas Niemöller[b], Vincent Baeten[a], Juan Antonio Fernández Pierna[a]

[a] Walloon Agricultural Research Center (CRA-W), Knowledge and Valorization of Agricultural Products Department
[b] Bruker Optics GmbH & Co. KG.



Credit: Builtin

A little bit of context

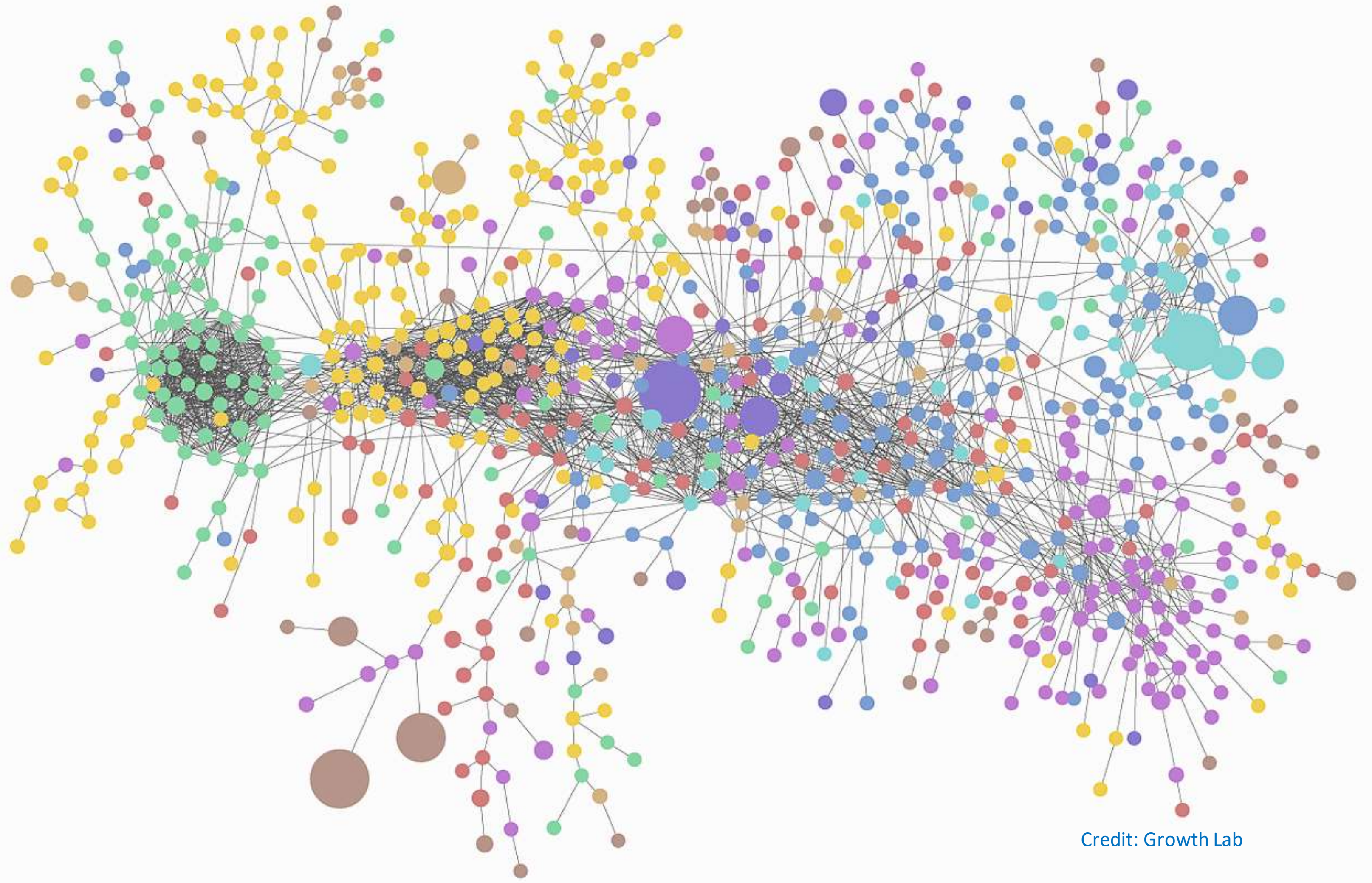
New analytical tools and technologies

↓

More and more data collected, including much more complexity

↓

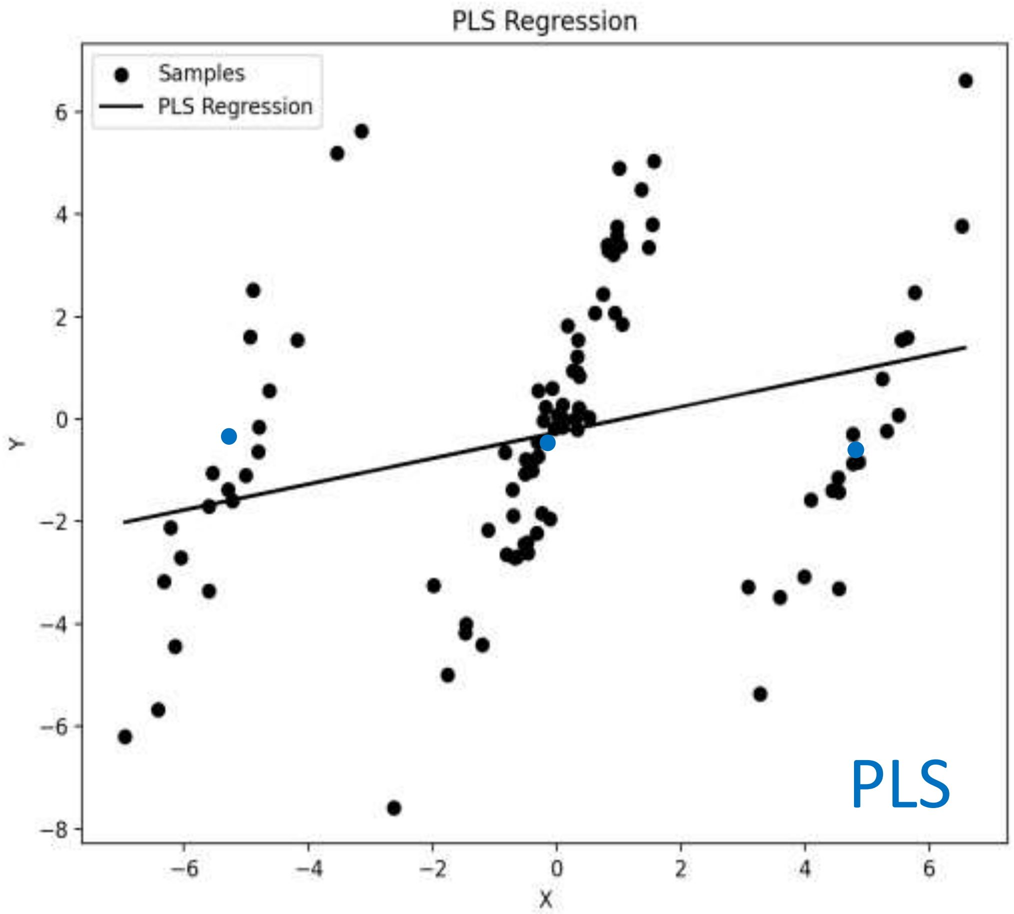
Limitations of classical methods such as PLS (based on the study of global variability)

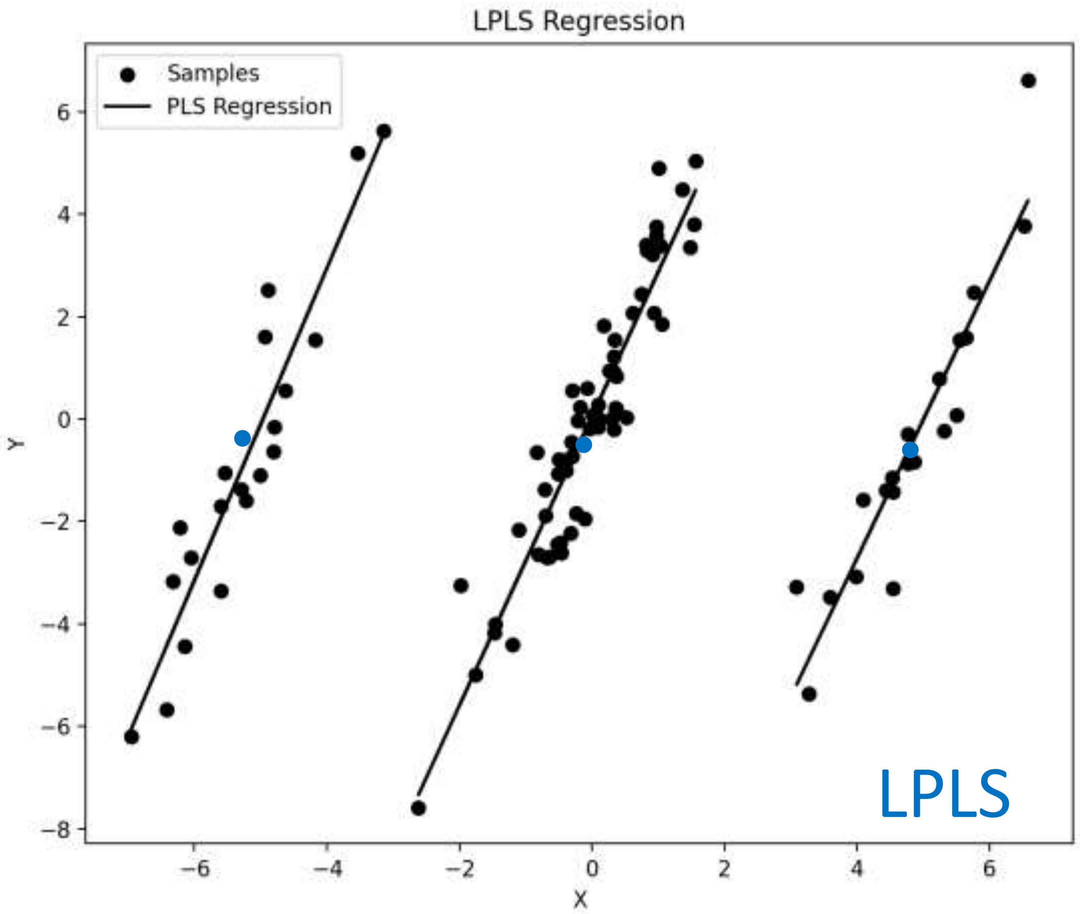


Credit: Growth Lab

Wallonie recherche CRA-W

LPLS Algorithms





LPLS Algorithms

- Better consideration of data complexity
- Robust to outliers
- Allow centralization of samples of different natures

- Difficult to interpret
- Increased computation time
- Technical constraints for implementation on handheld instruments

LPLS Algorithms

- Better consideration of data complexity
- Robust to outliers
- Allow centralization of samples of different natures

- Difficult to interpret
- Increased computation time
- Technical constraints for implementation on handheld instruments

Local partial least squares based on global PLS scores

Guanghui Shen, Matthieu Lesnoff, Vincent Baeten, Pierre Dardenne, Fabrice Davrieux, Hernan Ceballos, John Belalcazar, Dominique Dufour, Zengling Yang, Lujia Han, Juan Antonio Fernández Pierna

First published: 13 March 2019 | https://doi.org/10.1002/cem.3117 | Citations: 27

Regression models based on new local strategies for near infrared spectroscopic data

F. Allegrini, J.A. Fernández Pierna, W.D. Fragoso, A.C. Olivieri, V. Baeten, P. Dardenne

Development of soft-sensor using locally weighted PLS with adaptive similarity measure

Sanghong Kim, Ryota Okajima, Manabu Kano, Shinji Hasebe

Comparison of locally weighted PLS strategies for regression and discrimination on agronomic NIR data

Matthieu Lesnoff, Maxime Metz, Jean-Michel Roger

Global or Local? A Choice for NIR Calibrations in Analyses of Forage Quality

G. Sinnaeve, P. Dardenne, and R. Agneessens View all authors and affiliations

Volume 2, Issue 3 | https://doi.org/10.1255/jnirs.43

Averaging and Stacking Partial Least Squares Regression Models to Predict the Chemical Compositions and the Nutritive Values of Forages from Spectral Near Infrared Data

by Mathieu Lesnoff [1,2,3], Donato Andueza [4], Charlène Barotin [5], Philippe Barre [5], Laurent Bonnal [1,2], Juan Antonio Fernández Pierna [6], Fabienne Picard [4], Philippe Vermeulen [6] and Jean-Michel Roger [3,7,*]

Investigation of a LOCAL Calibration Procedure for near Infrared Instruments

John S. Shenk, Mark O. Westerhaus, and Paolo Berzaghi View all authors and affiliations

Volume 5, Issue 4 | https://doi.org/10.1255/jnirs.115

ARTICLE | September 3, 1998

Optimization in Locally Weighted Regression

Vitezslav Centner, and D. Luc Massart

Wallonie recherche CRA-W

LPLS Algorithms

- Better consideration of data complexity
- Robust to outliers
- Allow centralization of samples of different natures

- Difficult to interpret
- Increased computation time
- Technical constraints for implementation on handheld instruments

Dataset reduction techniques

Local partial least squares based on global PLS scores

Guanghui Shen, Matthieu Lesnoff, Vincent Baeten, Pierre Dardenne, Fabrice Davrieux, Hernan Ceballos, John Belalcazar, Dominique Dufour, Zengling Yang, Lujia Han, Juan Antonio Fernández Pierna

First published: 13 March 2019 | https://doi.org/10.1002/cem.3117 | Citations: 27

Development of soft-sensor using locally weighted PLS with adaptive similarity measure

Sanghong Kim, Ryota Okajima, Manabu Kano, Shinji Hasebe

Regression models based on new local strategies for near infrared spectroscopic data

F. Allegrini, J.A. Fernández Pierna, W.D. Fragoso, A.C. Olivieri, V. Baeten, P. Dardenne

Comparison of locally weighted PLS strategies for regression and discrimination on agronomic NIR data

Matthieu Lesnoff, Maxime Metz, Jean-Michel Roger

Global or Local? A Choice for NIR Calibrations in Analyses of Forage Quality

G. Sinnaeve, P. Dardenne, and R. Agneessens View all authors and affiliations

Volume 2, Issue 3 | https://doi.org/10.1255/jnirs.43

Averaging and Stacking Partial Least Squares Regression Models to Predict the Chemical Compositions and the Nutritive Values of Forages from Spectral Near Infrared Data

by Mathieu Lesnoff [1,2,3], Donato Andueza [4], Charlène Barotin [5], Philippe Barre [5], Laurent Bonnal [1,2], Juan Antonio Fernández Pierna [6], Fabienne Picard [4], Philippe Vermeulen [6] and Jean-Michel Roger [3,7,*]

Investigation of a LOCAL Calibration Procedure for near Infrared Instruments

John S. Shenk, Mark O. Westerhaus, and Paolo Berzaghi View all authors and affiliations

Volume 5, Issue 4 | https://doi.org/10.1255/jnirs.115

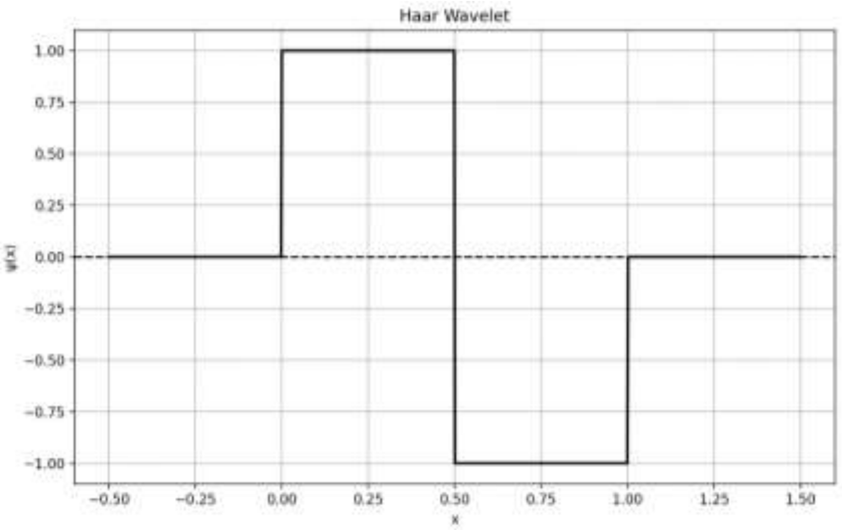
ARTICLE | September 3, 1998

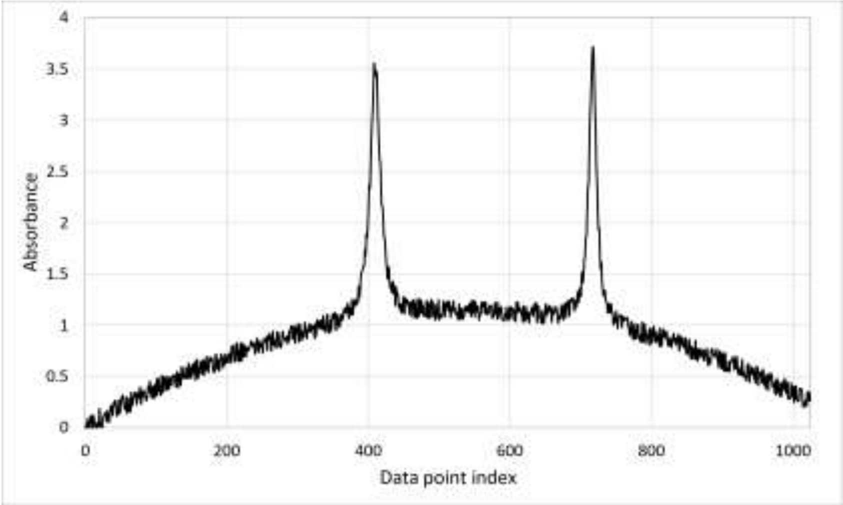
Optimization in Locally Weighted Regression

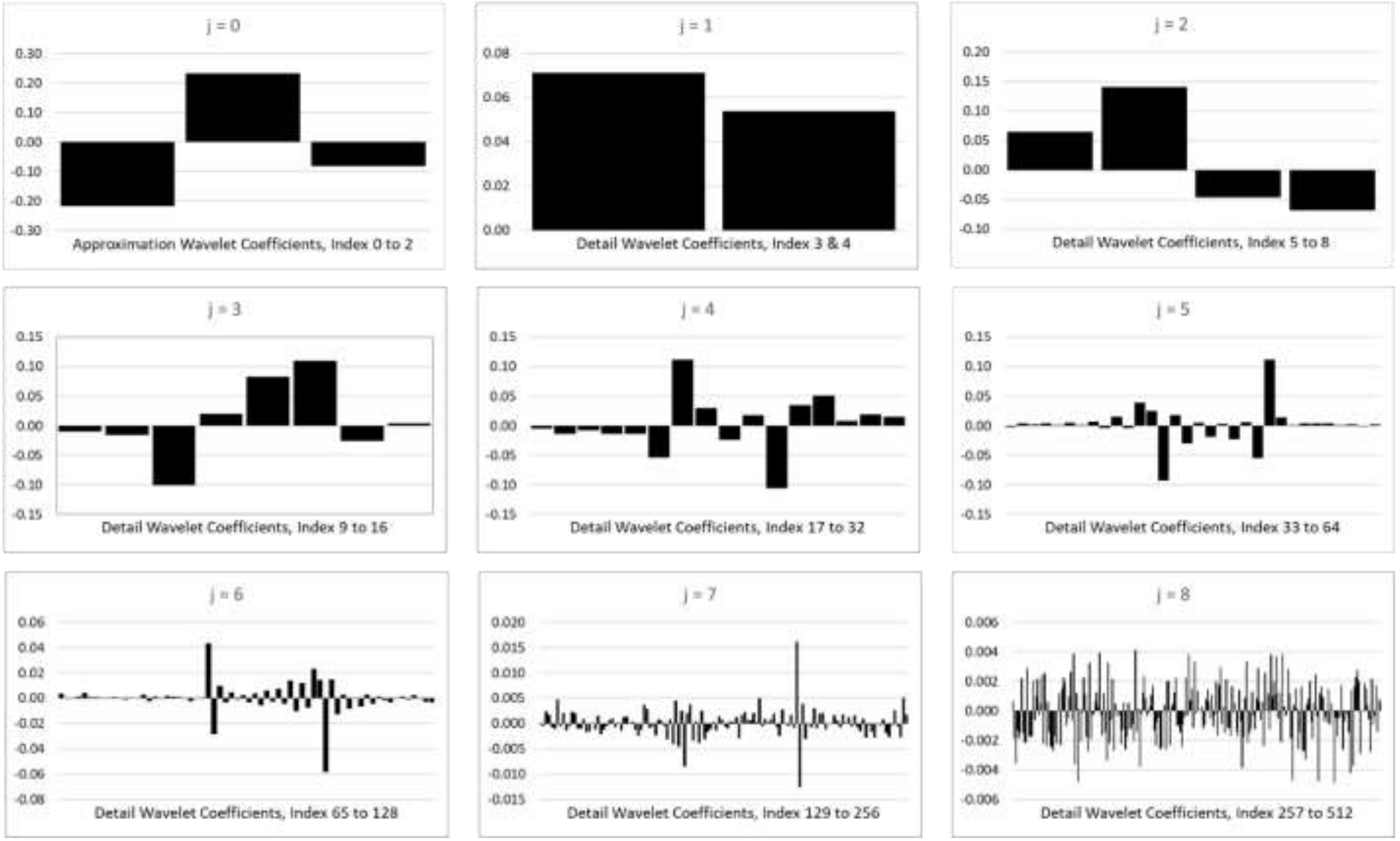
Vitezslav Centner, and D. Luc Massart

Wallonie recherche CRA-W

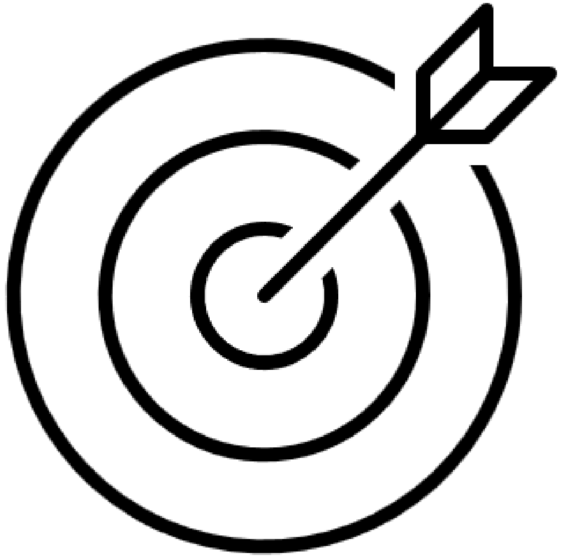
Wavelet Transform







Objectives



Assessing the impact of wavelet transform within a LPLS pipeline (WLPLS)

Comparing its performance with the traditional LPLS pipeline (LPLS)

Comparing its performance with another data reduction technique within a LPLS pipeline: the LPLS on global PLS scores (LPLS-S)

Methodology

- 24,644 spectra of feed formulations (CRA-W database)
- Measured with FOSS XDS spectrometer (1100-2498 nm by 2 nm)
- Reference analytical methods for constituents (% total weight of the sample)

Constituents	N	Min	Max	Mean	Std
ASH	20,065	0.80	37.00	7.32	3.26
MOIST	23,392	2.04	16.70	11.36	1.82
PROT	22,371	6.80	62.30	20.47	8.10

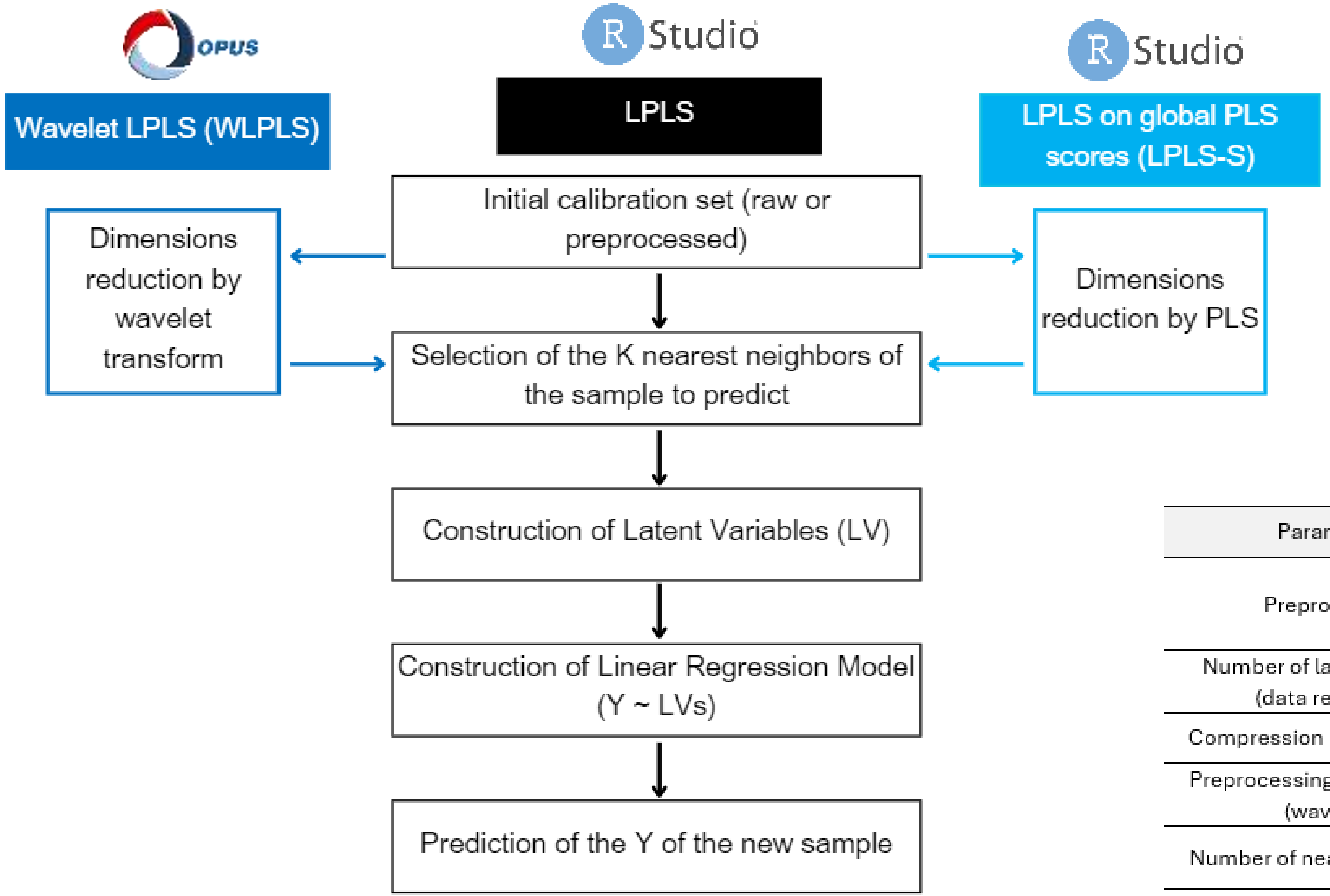
Parameters measured in % total weight of the sample

60%
Spectral library

15%
Optimisation dataset

25%
Validation dataset

Methodology



Parameters	LPLS	LPLS-S	WLPLS
Preprocessing	1st derivative - window size of 13 / SNV / None	1st derivative - window size of 13 / SNV / None	1st derivative - window size of 13 / SNV / None
Number of latent variables (data reduction)	10 to 20	/	/
Compression level (wavelets)	/	/	5 / 6 / 7
Preprocessing of coefficients (wavelets)	/	/	SNV / None
Number of nearest neighbors	100 / 200 / 300 / 400 / 500 / 1000	100 / 200 / 300 / 400 / 500 / 1000	100 / 200 / 300 / 400 / 500 / 1000
Number of latent variables (rank model)	1 to 20	1 to 20	1 to 20

Wallonie recherche CRA-W

LPLS: V. Centner, L. Massart, 1998, Optimization in Locally Weighted Regression, Analytical Chemistry 70 (1998), 19:4206-4211, https://doi.org/10.1021/ac980208r

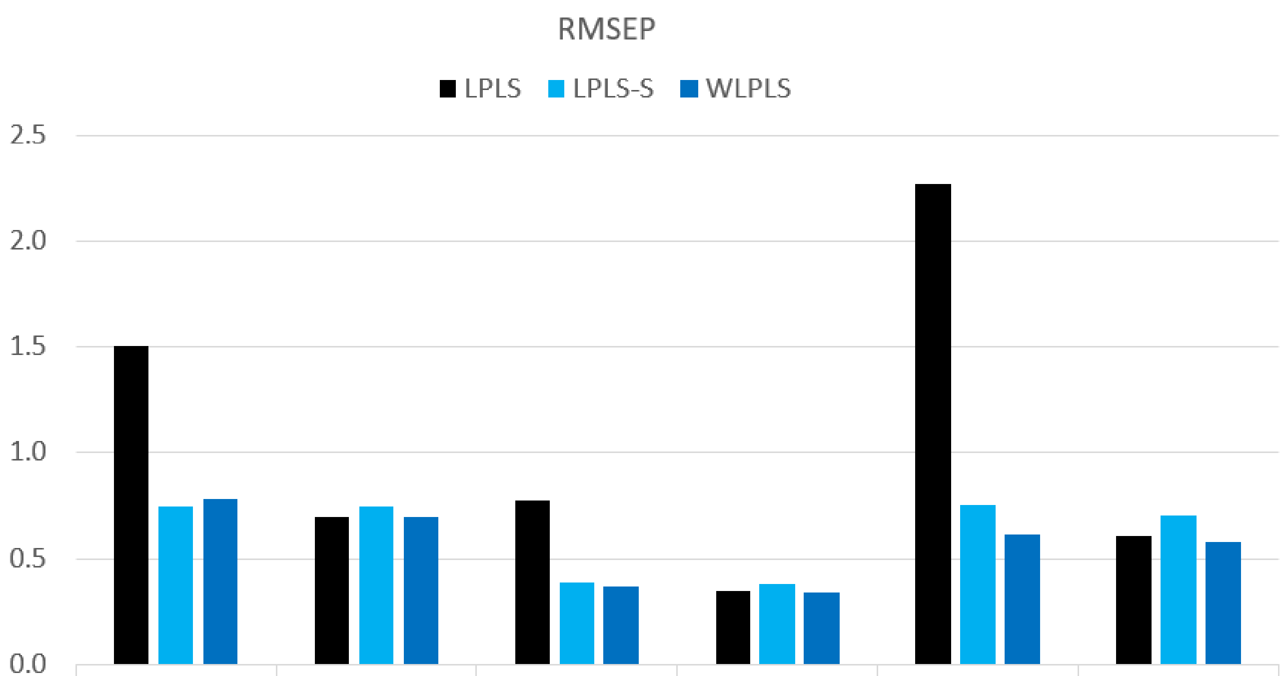
LPLS-S: G. Shen et al., 2019, Local Partial Least Squares Based on Global PLS Scores, Journal of Chemometrics 33 (2019), 5:e3117, https://doi.org/10.1002/cem.3117

Results

- The three LPLS pipelines gave high-quality predictions
- Small differences between pipelines for preprocessed data
- BUT significant differences for raw data (LPLS < LPLS-S and WLPLS)

RPD

	ASH	MOIST	PROT
PLS	2.733	3.699	9.109
LPLS	4.514	5.294	13.882
LPLS-S	4.370	4.849	11.444
WLPLS	4.504	5.358	13.935



	ASH RAW	ASH PREPRO	MOIST RAW	MOIST PREPRO	PROT RAW	PROT PREPRO
LPLS	1.505	0.698	0.772	0.346	2.269	0.608
LPLS-S	0.745	0.745	0.390	0.378	0.754	0.702
WLPLS	0.784	0.700	0.370	0.342	0.611	0.577

Discussion

LPLS	LPLS-S	WLPLS
• Significant computational power and time (not ideal for real-time analyses) • Usually requires preprocessing • Easily modifies library compositions without recalculations	• Reduces dataset to a few latent variables (e.g., 10-20 for a 700-point spectrum) • First PLS can remove noise (no further preprocessing required) • Relies on a simple data reduction technique • Changes in library composition require recalculation, affecting performance	• Reduces dataset to a few wavelet bands (e.g., 60-100 coefficients for a 700-point spectrum) • No need for spectral range selection or preprocessing • Less simple to understand and optimize • Adaptive through weighting of orthogonal coefficients • Easily modifies library compositions without recalculations

wallonie recherche CRA-W

Conclusion

Another Pipeline in Local Partial Least Squares Regression (LPLS) Methods: Assessing the Impact of Wavelet Transform Integration

Antoine Deryck[a], Andreas Niemöller[b], Vincent Baeten[a], Juan Antonio Fernández Pierna[a]

?

Local partial least squares based on global PLS scores

Guanghui Shen, Matthieu Lesnoff, Vincent Baeten, Pierre Dardenne, Fabrice Davrieux, Hernan Ceballos, John Belalcazar, Dominique Dufour, Zengling Yang, Lujia Han, Juan Antonio Fernández Pierna

First published: 13 March 2019 | https://doi.org/10.1002/cem.3117 | Citations: 27

ARTICLE | September 3, 1998

Optimization in Locally Weighted Regression

Vitezslav Centner, and D. Luc Massart

Conclusion

Another Pipeline in Local Partial Least Squares Regression (LPLS) Methods: Assessing the Impact of Wavelet Transform Integration

Antoine Deryck[a], Andreas Niemöller[b], Vincent Baeten[a], Juan Antonio Fernández Pierna[a]

?

Local partial least squares based on global PLS scores

Guanghui Shen, Matthieu Lesnoff, Vincent Baeten, Pierre Dardenne, Fabrice Davrieux, Hernan Ceballos, John Belalcazar, Dominique Dufour, Zengling Yang, Lujia Han, Juan Antonio Fernández Pierna

First published: 13 March 2019 | https://doi.org/10.1002/cem.3117 | Citations: 27

ARTICLE | September 3, 1998

Optimization in Locally Weighted Regression

Vitezslav Centner, and D. Luc Massart

No method is universally better (dataset dependant)

WLPLS offers specific advantages BUT similar benefits may apply to LPLS-S and other methods not explored here

Pipeline selection should be based on dataset characteristics and research goals

Wallonie recherche CRA-W

Further research and discussions

- Comparison of more algorithms applied to more datasets
- Why are local methods rarely used?
- Deep learning vs Local Methods?

Acknowledgments

BRUKER

OPUS

Wallonie recherche
CRA-W

Thank you for your attention

Antoine DERYCK

Walloon Agricultural Research Centre

Knowledge and valorization of agricultural products Department

Quality and authentication of agricultural products Unit

a.deryck@cra.wallonie.be | www.cra.wallonie.be

Wallonie recherche CRA-W