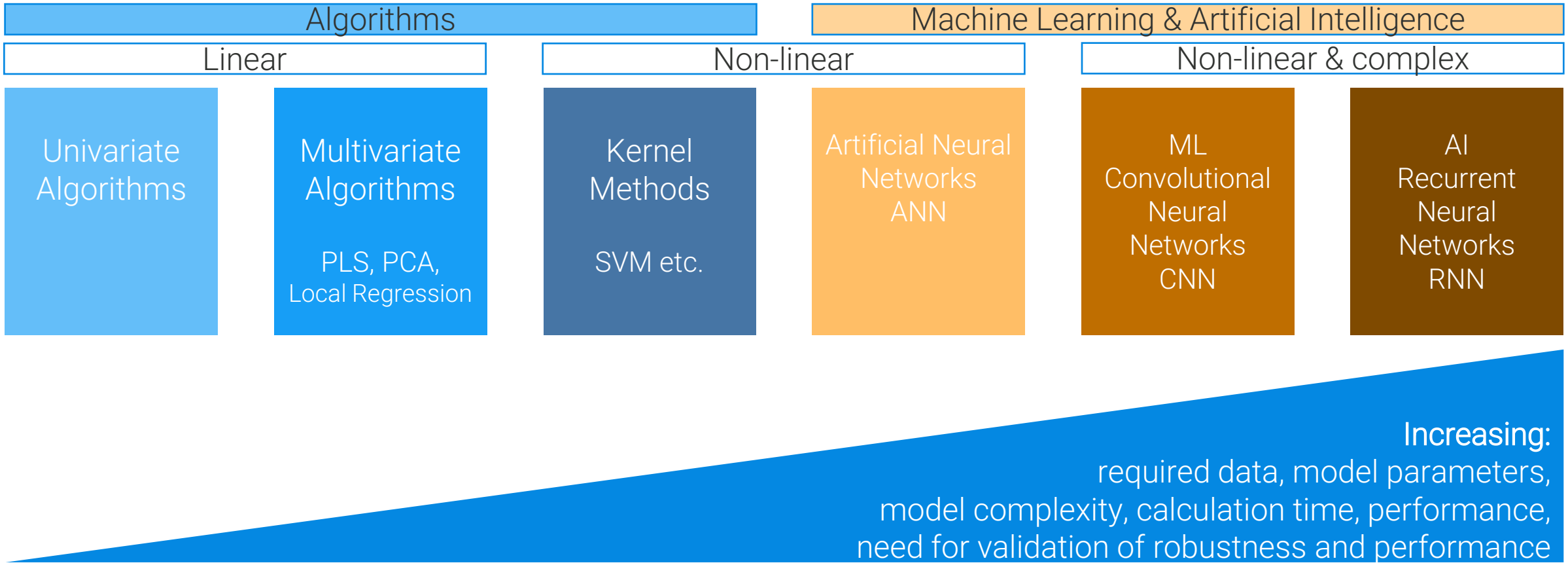
QUANT3

Are you prepared? Rediscovering wavelets and a new local regression approach

Andreas Niemöller

Algorithms: Increasing Demands but Decreasing Expertise



Algorithms implemented in QUANT3

Algorithms

Linear

- Univariate Algorithms
- Multivariate Algorithms
  - PLS
  - Local Regression

Non-linear

- Kernel Methods
  - SVR

Machine Learning & Artificial Intelligence

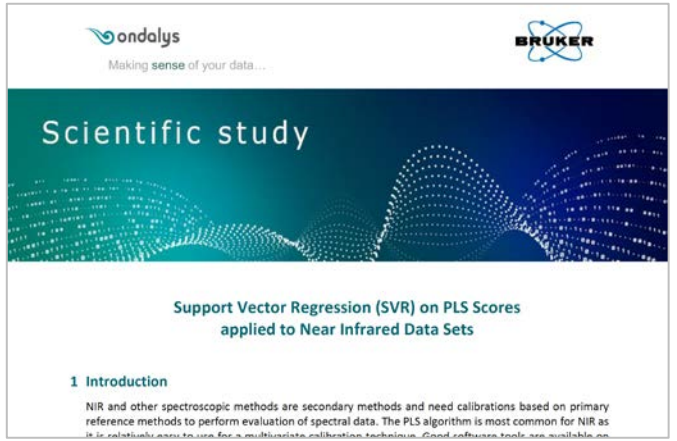
Non-linear

- Artificial Neural Networks
  - ANN

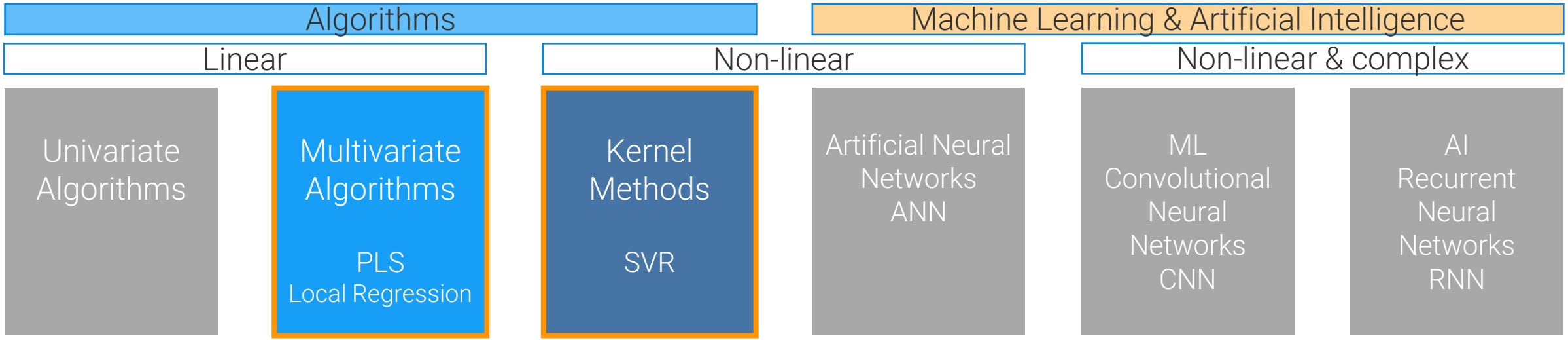
Non-linear & complex

- ML
  - Convolutional Neural Networks
  - CNN
- AI
  - Recurrent Neural Networks
  - RNN

- Implemented in QUANT3
  - ε-SVR on PLS scores



Algorithms implemented in QUANT3



- Implemented in QUANT3
  - ε-SVR on PLS scores
  - Local Regression based on libraries of wavelet transformed spectra

Local Regression: New Approach but what for?

- LOCAL method by Shenk and Westerhaus (1997)
  Pioneer work on spectra, no data compression, no transformation, no model(s) behind

Shenk J.S., Westerhaus M.O. and Berzaghi P., Investigation of a LOCAL calibration procedure for NIR instruments, J. Near Infrared Spectrosc, 5, 223–232, 1997

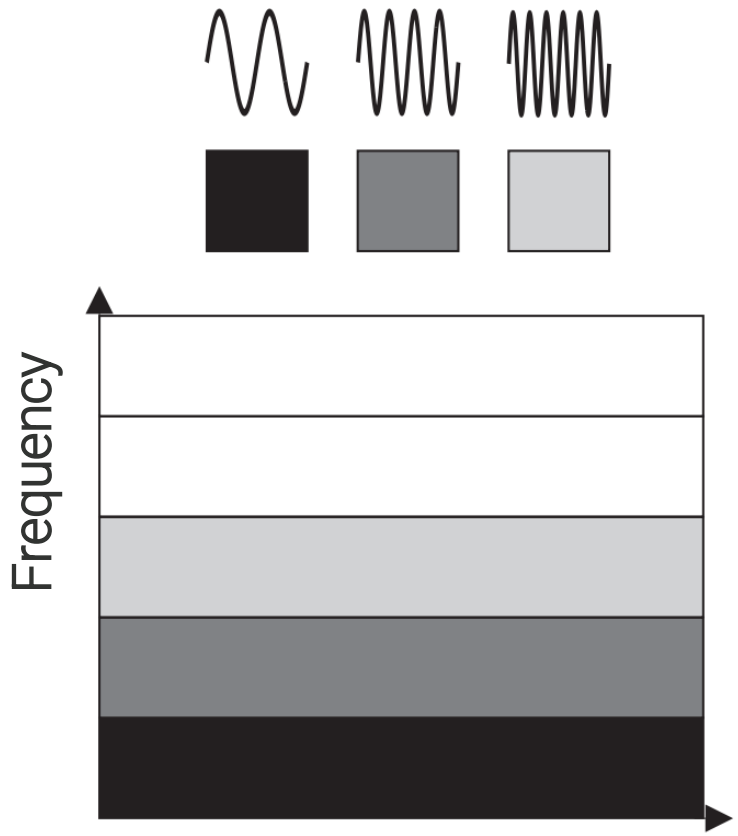
- PLS based Local Regression (LCPS or LCCRS, 2016)
  PLS score-based library utilizes specific information, data compression, model based

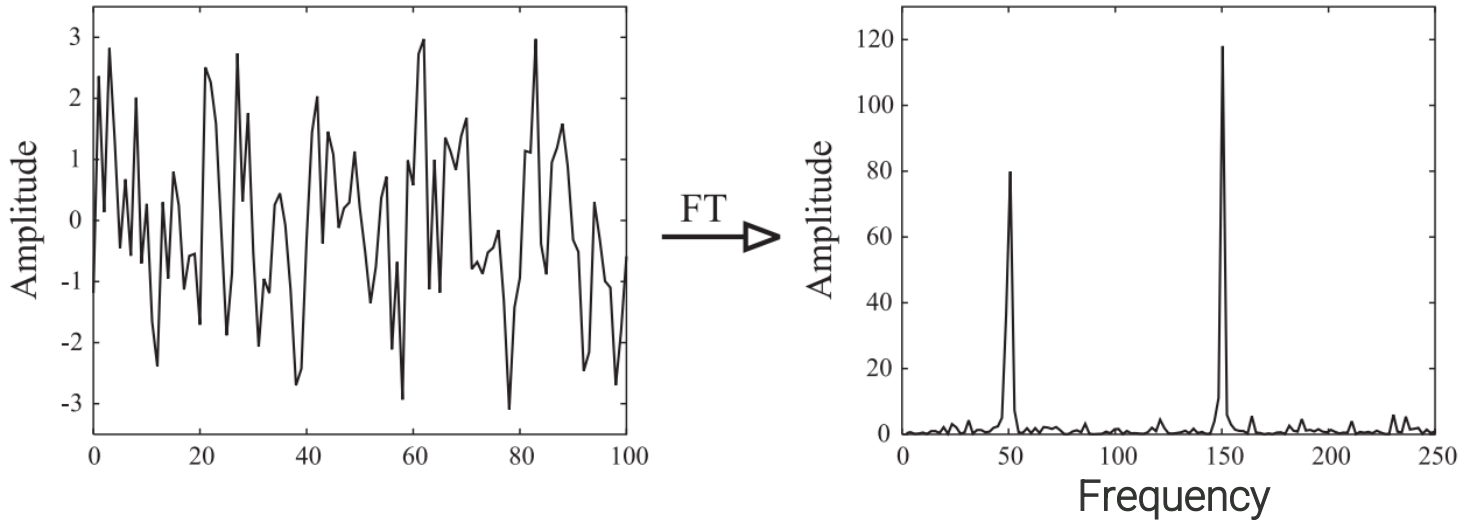
Allegrini F., Fernandez Pierna J.A., Fragoso W.D., Olivieri A.C., Baeten V., Dardenne P., Regression models based on new local strategies for near infrared spectroscopic data, Analytica Chimica Acta, 933, 50-58, 2016

- Requirements for new approach:
  - Sparse data representation (data compression), orthogonal latent variables
  - Model free library for updating and expanding without re-calculation of existing data

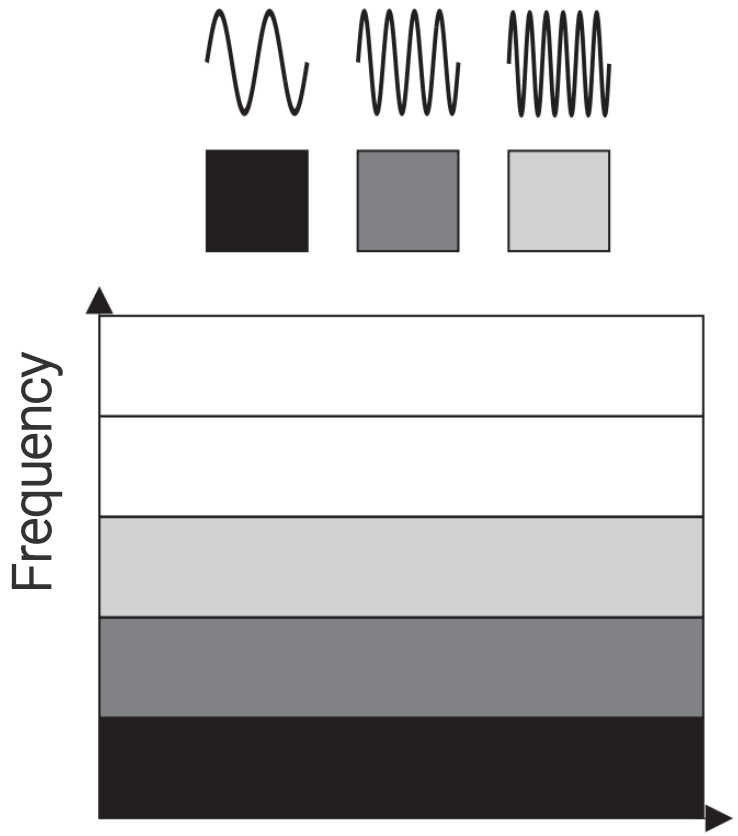
➡ Local Regression on wavelet transformed spectra

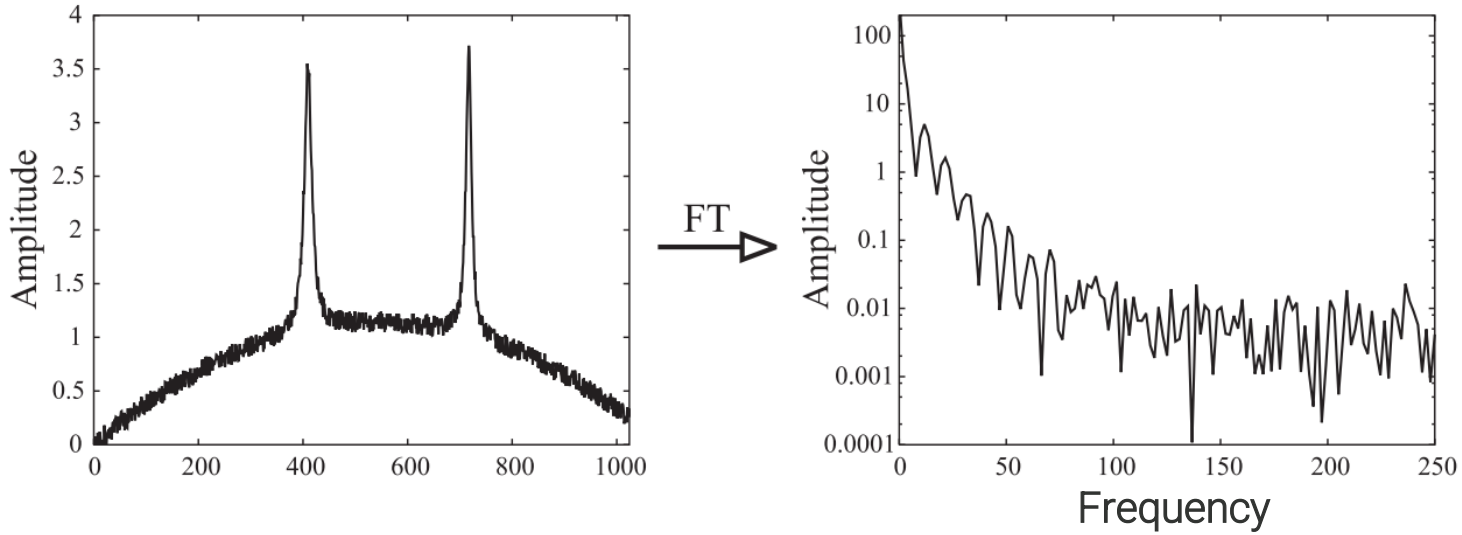
Fourier Transformation: Frequency Analysis of a Simple Signal



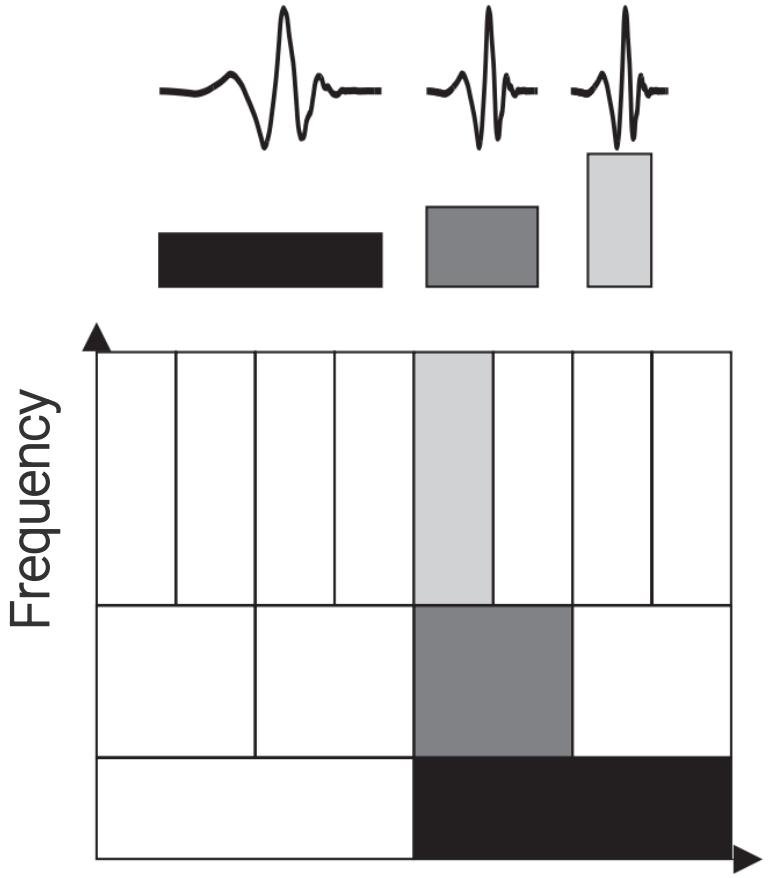


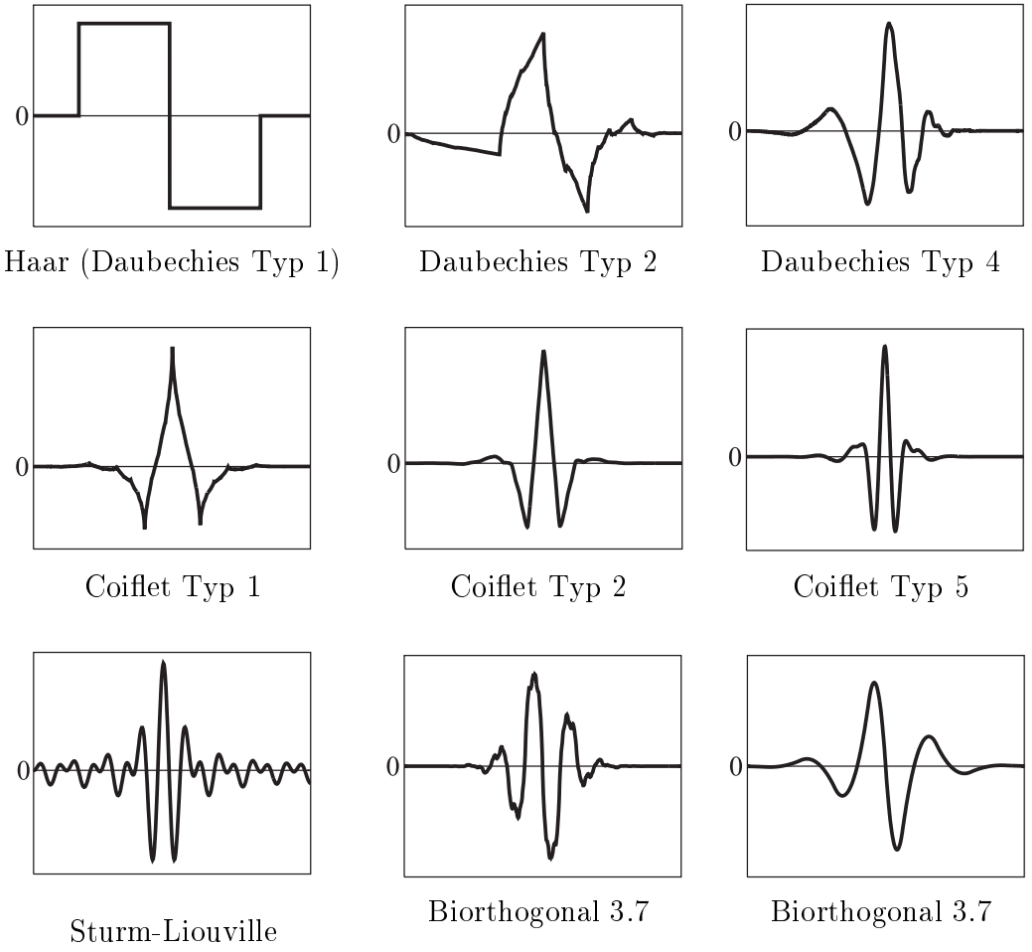
Fourier Transformation: Frequency Analysis of a Spectrum





Wavelet Transformation: Localization in Time and Frequency

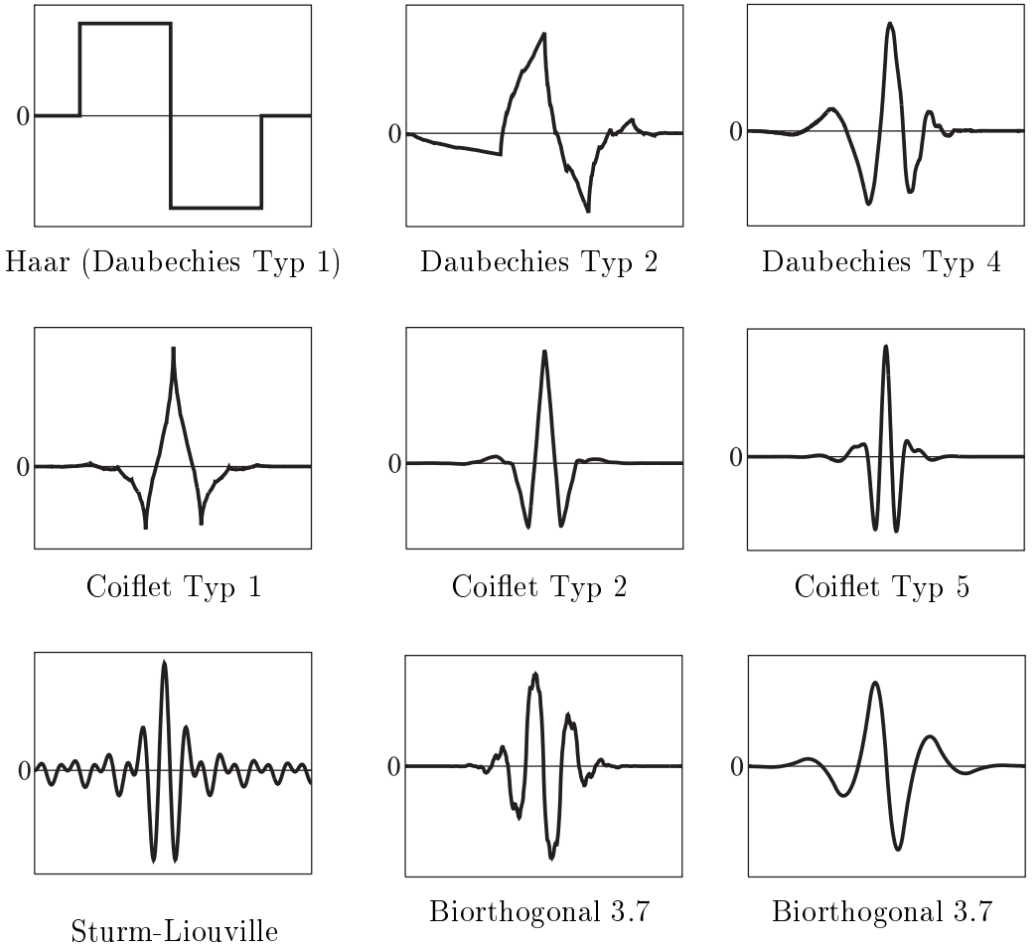




Wavelet Transformation: Lots of Wavelet Functions available

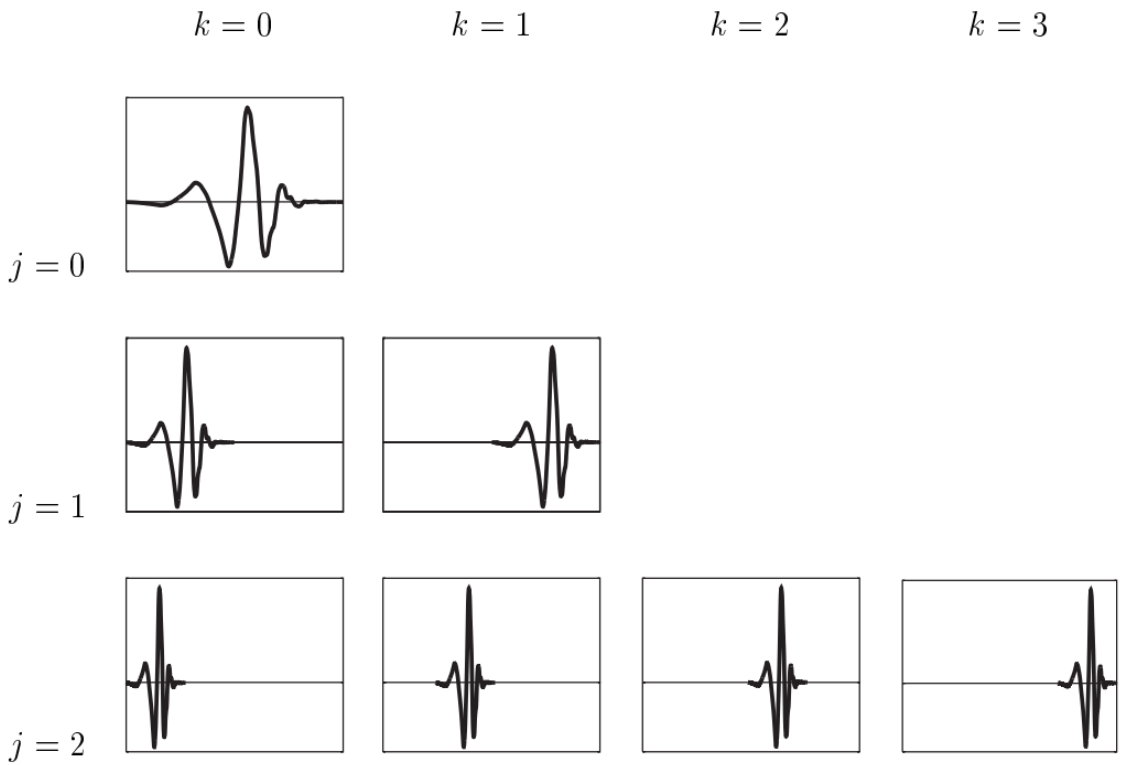
- Localization in Time and Frequency
- Compact Support
- Zero mean
- Orthogonality

- Different properties regarding
  - Smoothness and regularity
  - Detecting signal details
  - Compression and denoising
  - Approximation power
  - Reconstructing the signal

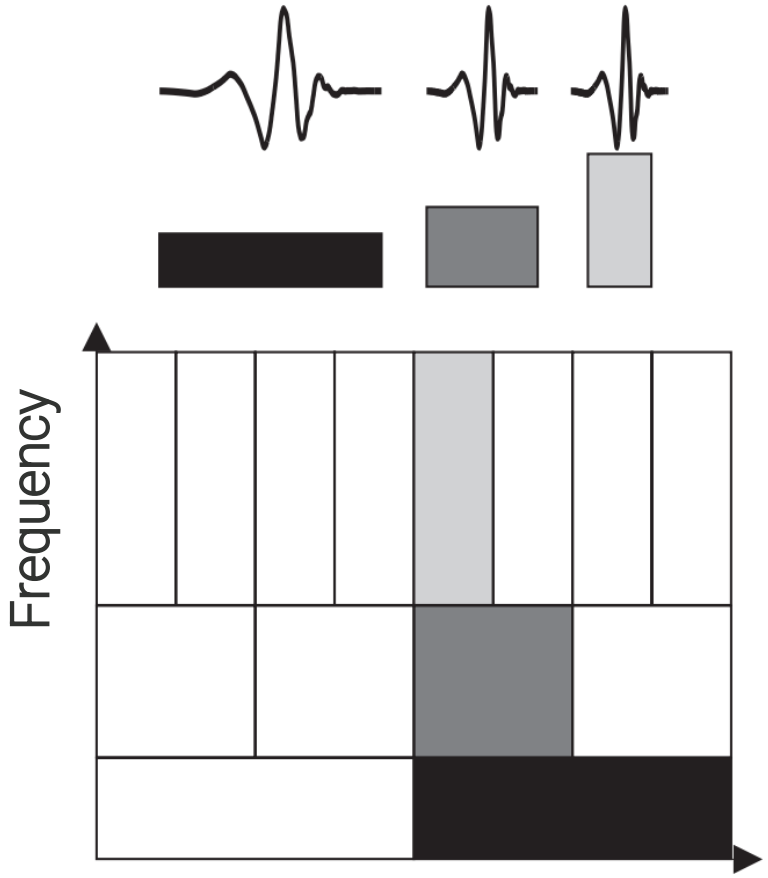


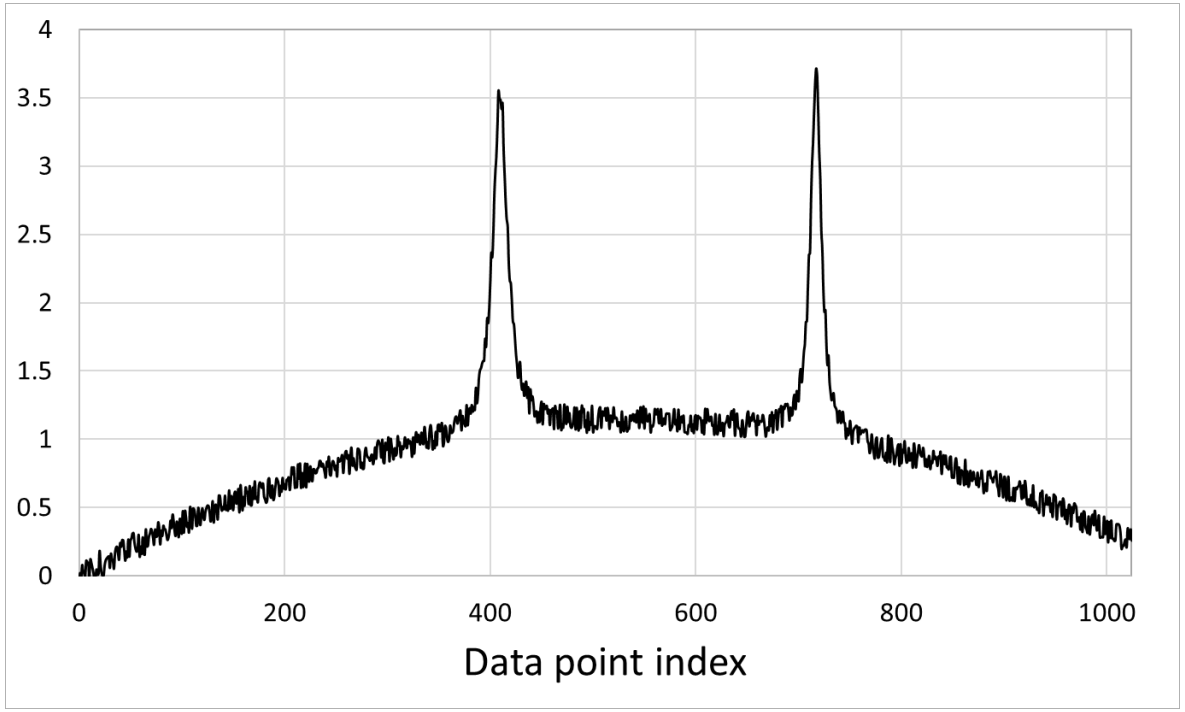
Wavelet Transformation: Dilatation and Translation

- Localization in Time and Frequency by
  - Dilation stretches or compresses the wavelet, effectively changing its frequency
  - Translation shifts the wavelet along the time axis



Wavelet Transformation of a Spectrum

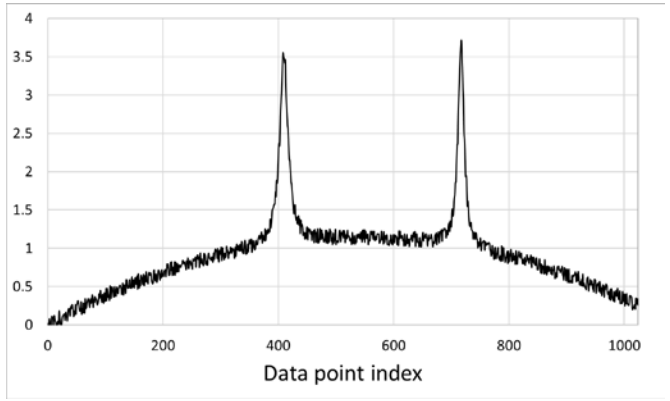


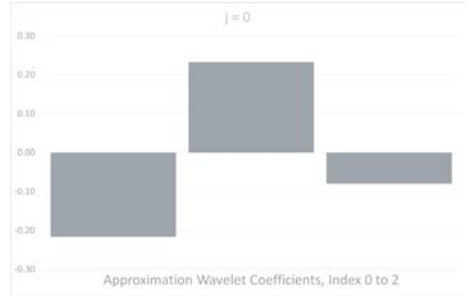


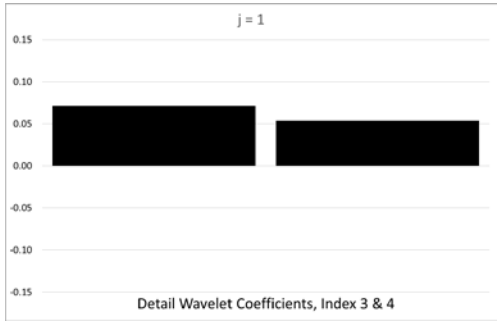
Wavelet Transformed Spectrum: Frequency Bands and Coefficients

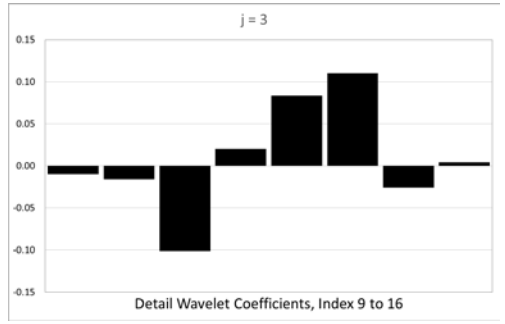
Options to work on transformed data:

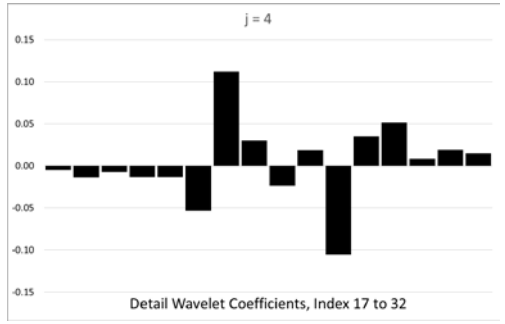
- Select & use coefficients directly
- Select & reconstruct signal again (e.g. smoothing and denoising)

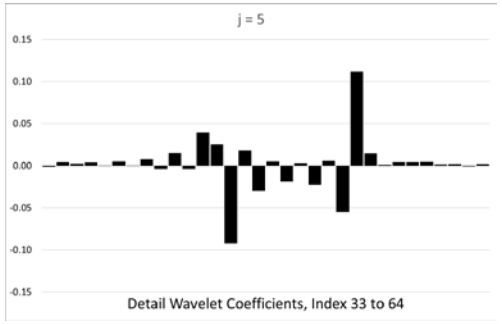


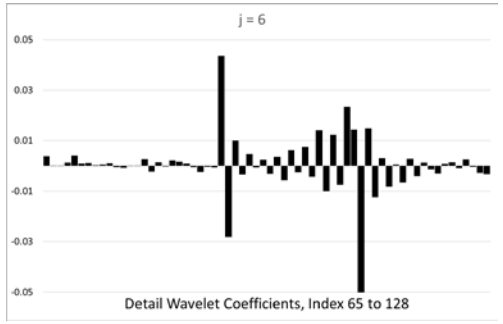


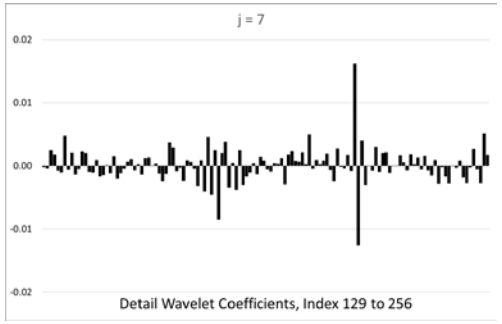


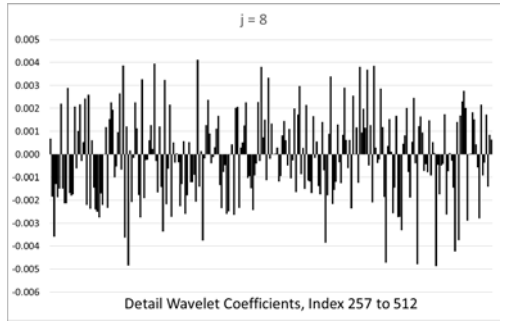












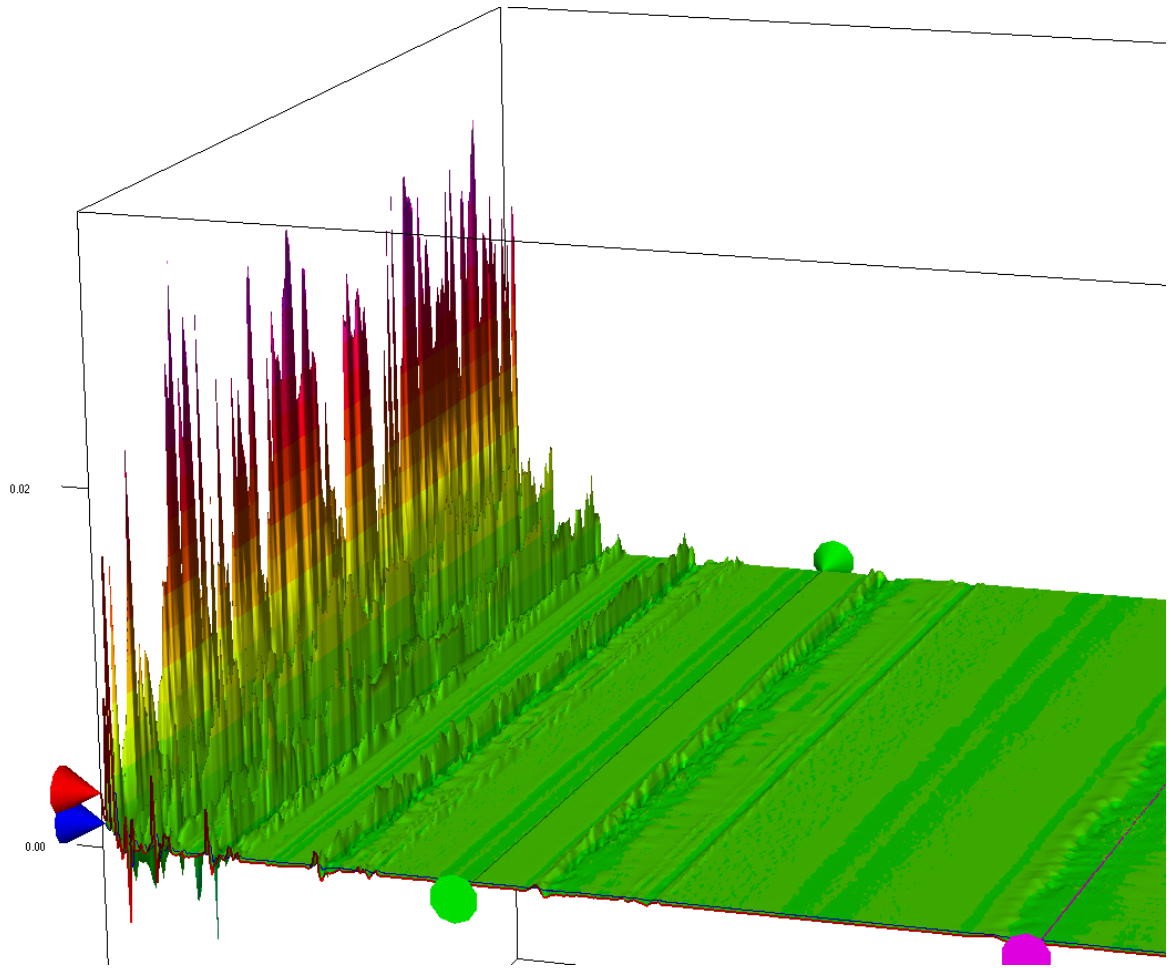
Wavelet Transformed Spectra Library

Wavelet coefficient data space

- Orthogonal latent variables
- Spectra variance reproduced
- Easy updating /expansion
- Filtering / selection possible

Local Regression

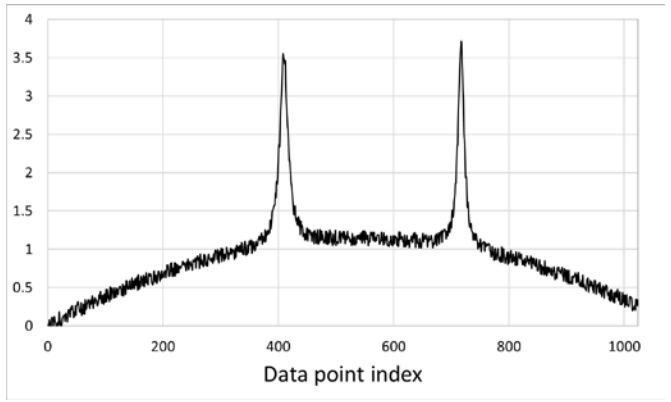
- kNN on coefficients one time
- PLS on coefficients,
  same input for all components
  PLS weighting works well on orthogonal coefficients

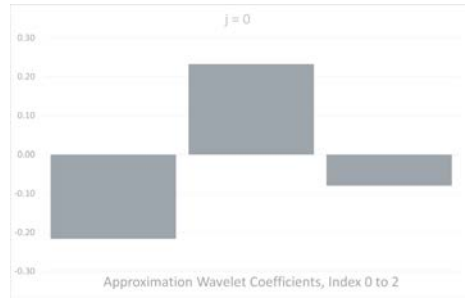


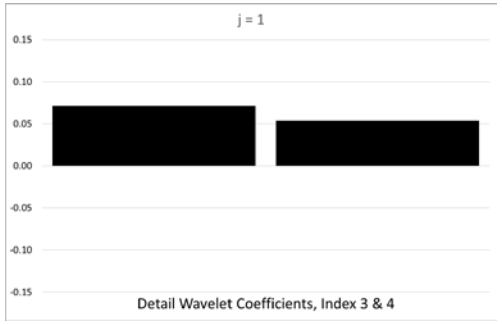
Selection by Correlation: Use only helpful Wavelet Coefficients

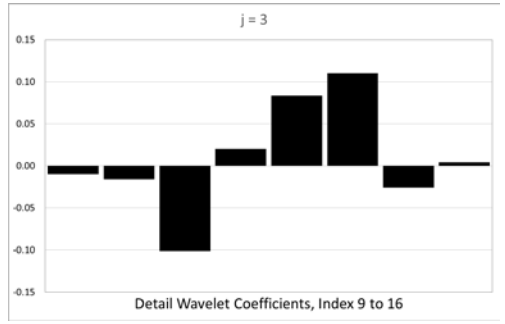
Selection of coefficients:

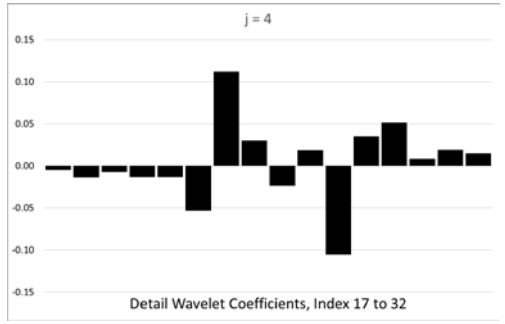
- by correlation and genetic algorithm (GA)
- by thresholding
- by frequency band

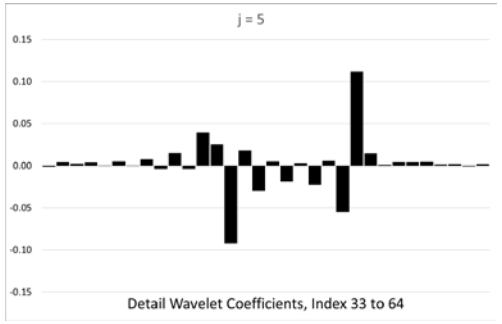


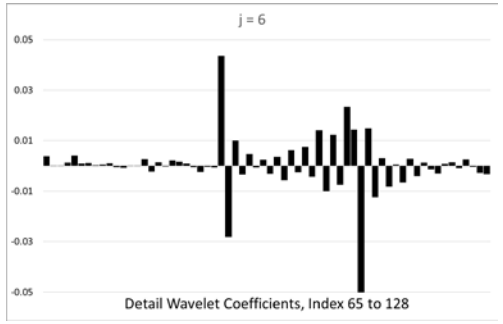


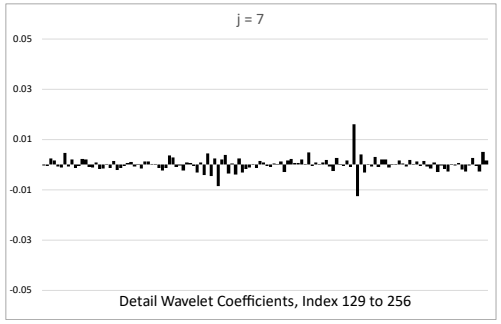


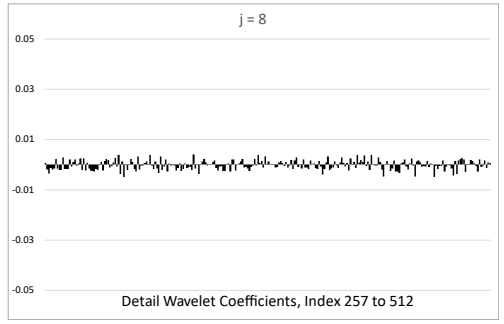








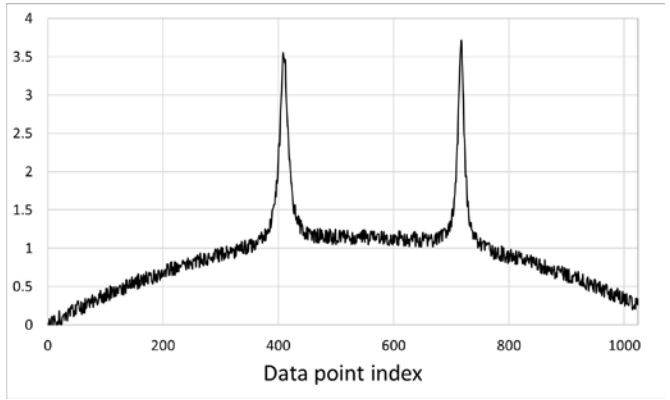


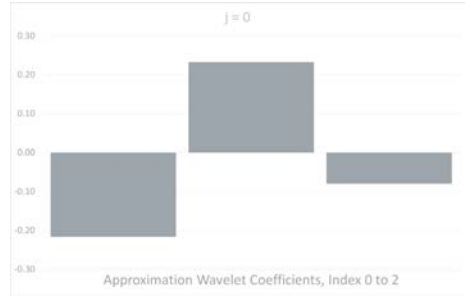


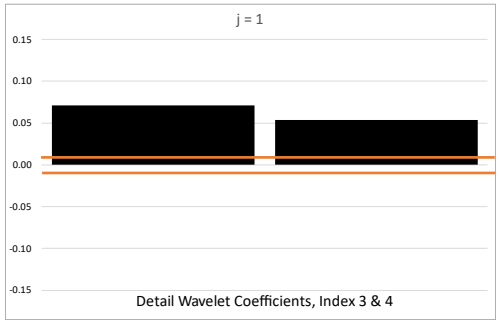
Thresholding: Discard Wavelet Coefficients which are too small

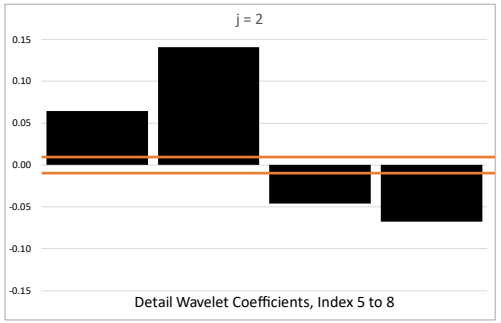
Selection of coefficients:

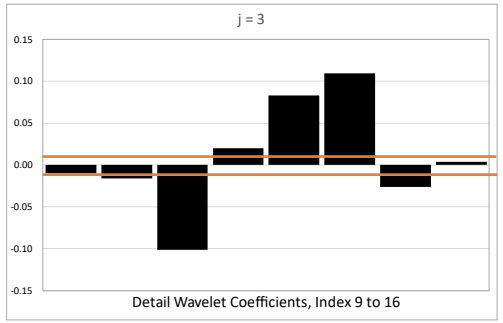
- by correlation
- by thresholding (Donoho et.al., 1994)
- by frequency band

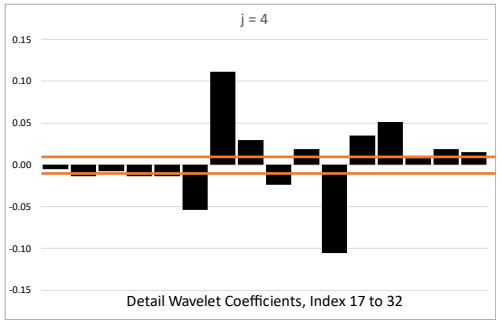


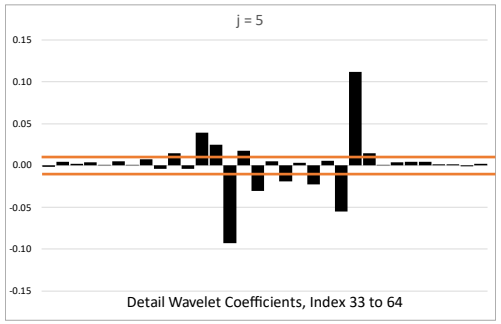


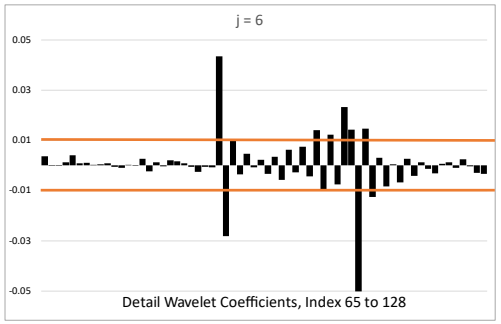


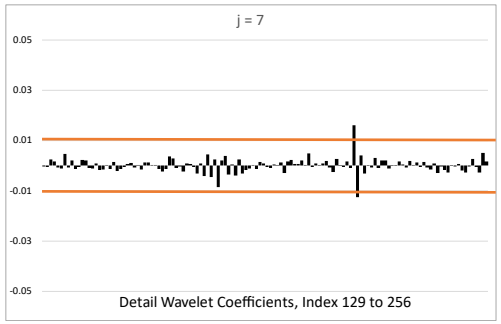


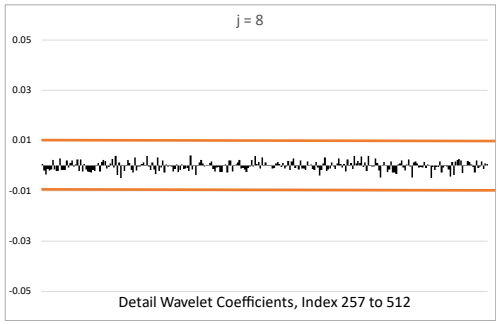








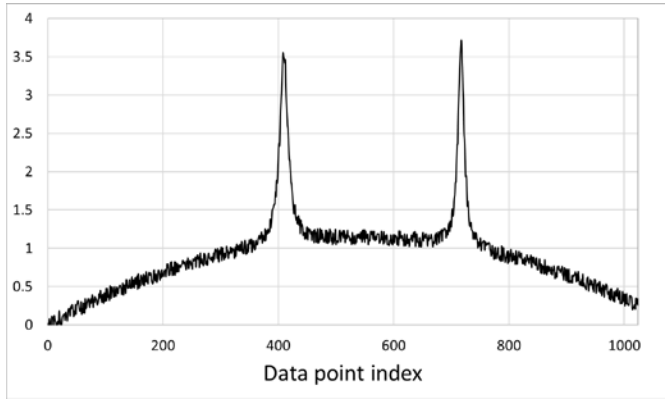


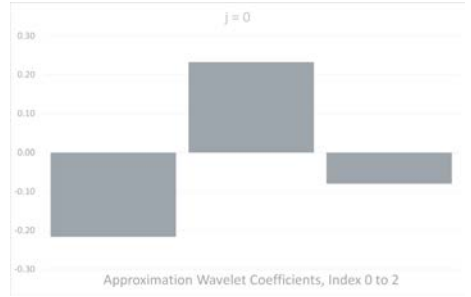


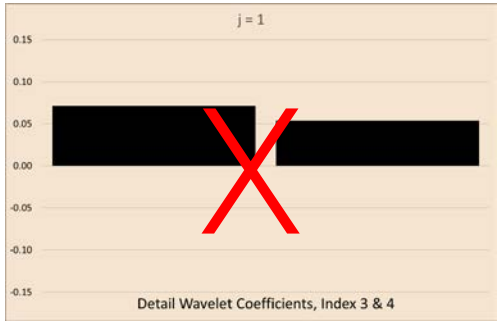
Wavelet Transformed Spectrum: Frequency Bands to be Selected

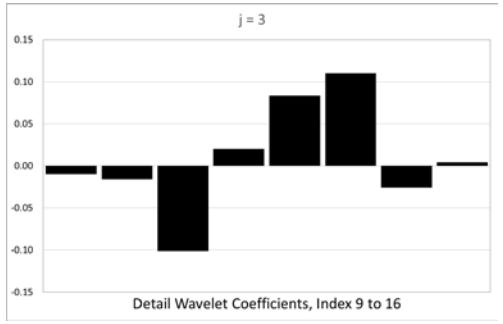
Selection of coefficients:

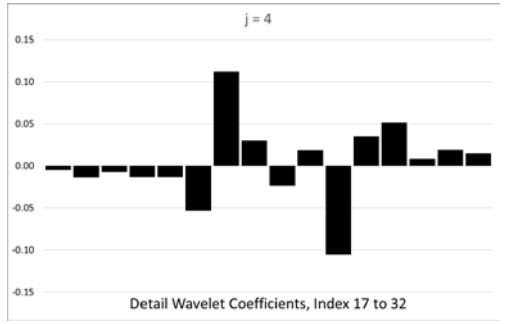
- by correlation
- by thresholding
- by frequency bands as they are
  (excluding baselines and noise)

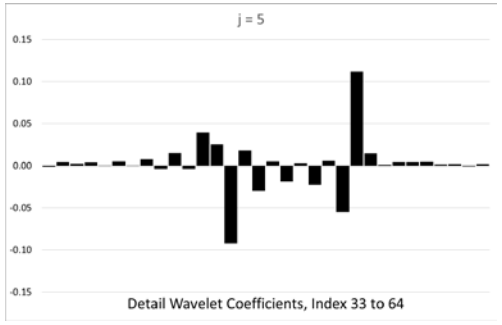


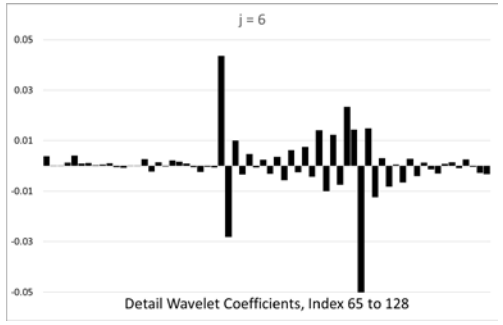












Local Regression (LR) based on Wavelet Compressed Spectra

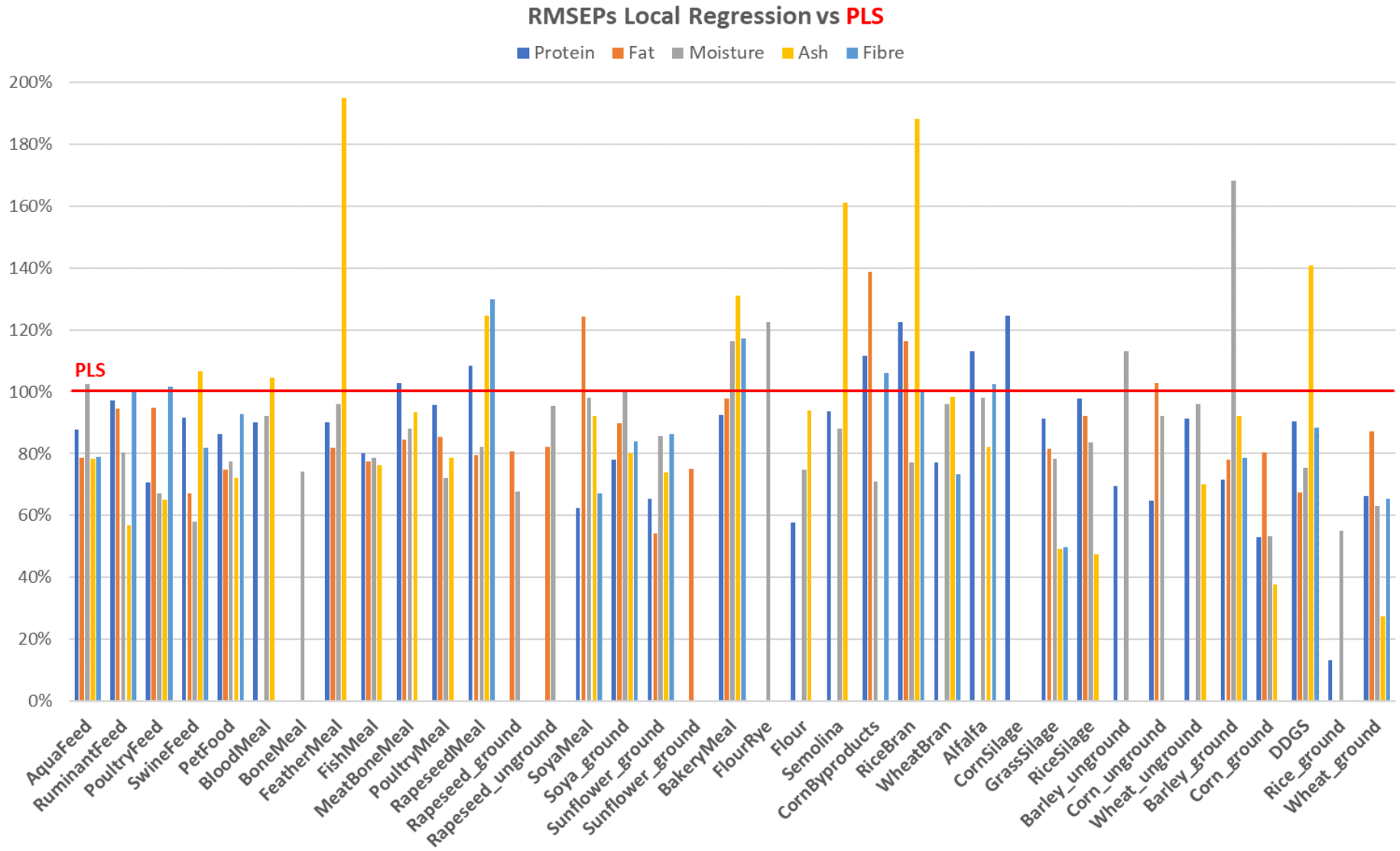
The QUANT3 approach settings:

- Broad spectral range(s) selection only (manually), no optimization required
- Due to wavelet transformation almost no pre-processing required (but possible)
- Simple selection of coefficients by compression levels (wavelet frequency bands)
- SNV on selected wavelet coefficient bands reduces noise

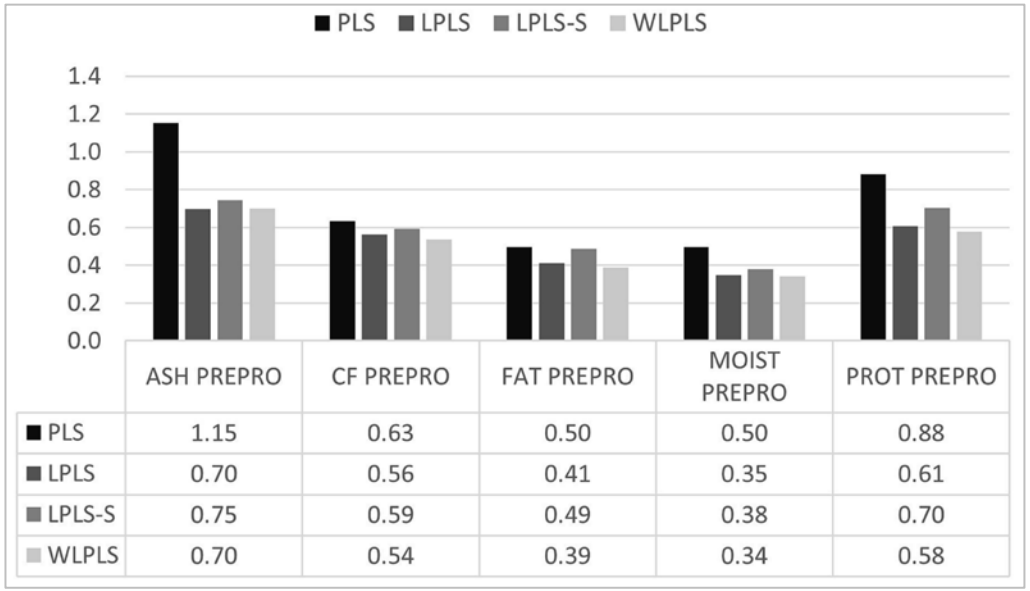
- Setting of rank
  - Global PLS rank for each component
  - Auto-Rank on averaged local suitable ranks determined by Cross Validation

New!

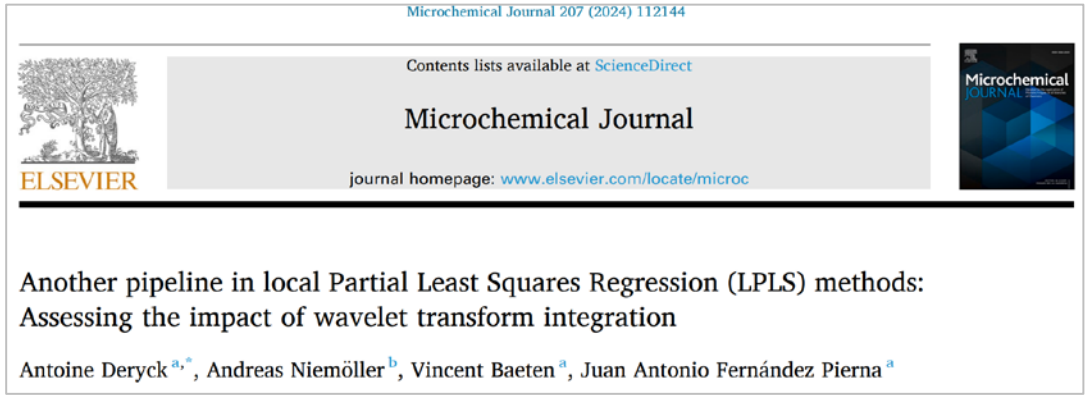
Local Regression compared to Bruker PLS Models: One Library, 37 Products and 100k Samples, 20% Validation Set



Local Regression (LR) based on Wavelet Compressed Spectra



	ASH PREPRO	CF PREPRO	FAT PREPRO	MOIST PREPRO	PROT PREPRO
PLS	1.15	0.63	0.50	0.50	0.88
LPLS	0.70	0.56	0.41	0.35	0.61
LPLS-S	0.75	0.59	0.49	0.38	0.70
WLPLS	0.70	0.54	0.39	0.34	0.58



Microchemical Journal 207 (2024) 112144

Contents lists available at ScienceDirect

Microchemical Journal

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

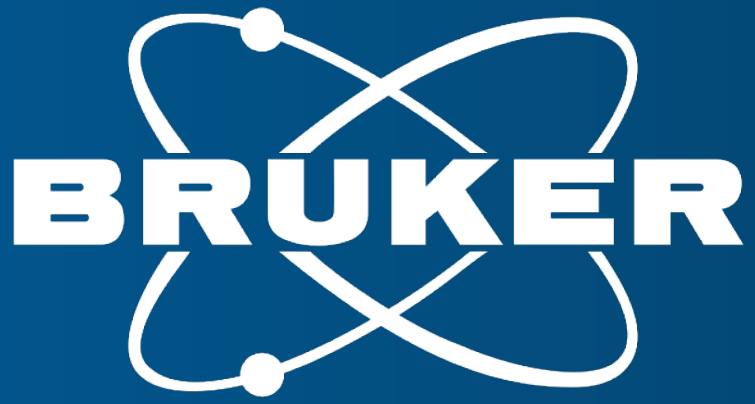
ELSEVIER

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]

Thank you, Antoine and Juan!

Wallonie recherche CRA-W



Innovation with Integrity