Pre-processing in biochemometrics: correction for path-length and temperature effects of water in FTIR bio-spectroscopy by EMSC

The term 'biochemometrics' is proposed. Methodology is presented for simplifying high-resolution phenotyping measurements, in terms of multivariate modelling methods in order to stabilise FTIR bio-spectroscopy data. Irrelevant gas contributions from water vapour and CO2 in the instrument light path are modelled and removed, in order to avoid having to wait for N-2 purging. Variations in the infrared spectroscopy (IR) spectrum of water with temperature are described in terms of two model component spectra. These additive water variations are quantified and eliminated by Extended Multiplicative Signal Correction (EMSC), along with various physical signal variations of additive and multiplicative nature due to, for example, sample or instrument temperature. Sample temperature is predicted from the EMSC model parameters. The models developed for Attenuated Total Reflection (ATR) measurements of pure water at different temperatures are tested successfully in independent water samples as well as in the in vivo monitoring of Candida albicans growing and decaying on the ATR crystal of the same instrument.

General information
State: Published
Organisations: Department of Biotechnology, Department of Systems Biology, National Food Institute
Authors: Martens, H. (Intern), Bruun, S. W. (Intern), Adt, I. (Ekstern), Sockalingum, G. (Ekstern), Køhler, A. (Intern)
Pages: 402-417
Publication date: 2006
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of Chemometrics
Volume: 20
Issue number: 8-10
ISSN (Print): 0886-9383
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): SNIP 0.832 SJR 0.54 CiteScore 1.7
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): SJR 0.502 SNIP 0.869 CiteScore 1.71
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.556 SNIP 0.973 CiteScore 1.82
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.55 SNIP 1.28 CiteScore 1.72
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.757 SNIP 1.307 CiteScore 2.2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.673 SNIP 1.335 CiteScore 2.03
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.735 SNIP 1.044 CiteScore 1.66
Near-infrared spectra of Penicillium camemberti strains separated by extended multiplicative signal correction improved prediction of physical and chemical variations

Different methods for spectral preprocessing were compared in relation to the ability to distinguish between fungal isolates and growth stages for Penicillium camemberti grown on cheese substrate. The best classification results were obtained by temperature- and wavelength-extended multivariate signal correction (TWEMSC) preprocessing, whereby three patterns of variation in near-infrared (NIR) log(1/R) spectra of fungal colonies could be separated mathematically: (1) physical light scattering and its wavelength dependency, (2) differences in light absorption of water due to varying sample temperature, etc., and (3) differences in light absorption between different fungal isolates. With this preprocessing, discriminant partial least squares (PLS) regression yielded 100% correct classification of three isolates, both within the cross-validated calibration set and in two independent test sets of samples.
BFI (2017): BFI-level 1
Scopus rating (2017): SNIP 0.867 SJR 0.489 CiteScore 1.6
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): SJR 0.486 SNIP 0.964 CiteScore 1.76
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.64 SNIP 1.066 CiteScore 1.96
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 0.633 SNIP 1.054 CiteScore 1.96
Scopus rating (2013): SJR 0.641 SNIP 1.144 CiteScore 2.08
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 0.574 SNIP 1.088 CiteScore 1.81
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 0.609 SNIP 1.011 CiteScore 1.62
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 0.695 SNIP 0.977
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 0.635 SNIP 0.892
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 0.864 SNIP 1.076
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.782 SNIP 0.957
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.816 SNIP 1.002
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.888 SNIP 1.125
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 0.871 SNIP 1.012
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.922 SNIP 1.169
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.849 SNIP 1.109
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.903 SNIP 1.201
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 1.302 SNIP 1.189
Scopus rating (1999): SJR 1.294 SNIP 1.165

Original language: English
NIR, diffuse reflectance, preprocessing, water structure, EMSC, near infrared, P. camemberti, temperature-calibrated EMSC, extended multiplicative signal correction, temperature, light scattering
Source: orbit
Source-ID: 196413
Publication: Research - peer-review › Journal article – Annual report year: 2005
The storage protein composition from the Glu-1, Glu-3 and Gli-1 loci encoding high and low molecular weight glutenin subunits (HMW-GS and LMW-GS) and gliadins, respectively, was determined on 30 wheat (T aestivum L) genotypes from three growing seasons. The gliadins and the LMW-GS were identified as gliadin/LMW-GS pairs. All samples were analysed by two one-dimensional electrophoretic techniques, and selected samples were also subjected to two-dimensional electrophoretic separation. Different statistical/data-analytical techniques were evaluated in the study of how the presence or absence of the protein alleles, the protein content and the growing seasons are related to flour quality.

The year of growth had a large impact on mixograph peak time. When predicting mixograph peak time from the presence or absence of significant proteins and the year of growth, 70% of the variability in mixograph peak time could be explained, whereas only 49% of the variability could be explained when the year of growth was deleted from the model. Protein had no effect on mixograph peak time as expected, and the well-known positive effect of HMW-GS 5 + 10, and the negative effects of 2 + 12 and 6 + 8 was observed. Furthermore, some of the gliadin/LMW-GS combinations influenced mixograph peak time significantly. The gliadin/LMW-GS at the combined Gli-Al, Glu-A3 loci b;f was positively related to mixograph peak time, whereas f;f and a;a was negatively related. Although the LMW-GS component f of the alleles b;f and f;f alleles appear similar on one-dimensional gels, two-dimensional separation of selected samples may suggest that the f components in these alleles are different proteins. Cross-validated partial least squares regression combined with empirical uncertainty estimates (jack-knifing) of the parameters estimated in the model, gave similar results to ANOVA in identifying quality related protein alleles. The applicability of the multivariate approach in proteomics is, however, much wider.
Sensory analysis for magnetic resonance-image analysis: Using human perception and cognition to segment and assess the interior of potatoes

The human ability to detect, interpret, assess and report visual stimuli is employed for qualitative and quantitative image analysis. Cross-sectional T2 weighted MR images were acquired from a series of intact potatoes. Descriptive sensory image: analysis was performed on the images. The sensory results were compared to conventional computer-assisted grey tone intensity histogram descriptors. A total of 60 tubers (2 varieties x 2 storage times x 15 replicate tubers) were submitted to MR-imaging. A trained panel of 9 assessors, used 16 sensory descriptors to assess the images. These descriptors were developed by the panel during preliminary training sessions, and consisted in definitions of various biological compartments inside the tubers. The results from the sensory and the computer-assisted image analyses of the shape and interior structure of the tubers were related to the experimental design and to the dry matter content of the individual tubers, by partial least-squares regression (PLSR). Finally, predictive modelling of sensory image description from computer-assisted image analysis was attempted. The results showed that both the sensory and the computer-assisted image analyses were able to detect differences between varieties as well as storage times. The sensory image analysis gave better discrimination between varieties than the computer-assisted image analysis presently employed, and was easier to interpret. Some sensory descriptors could be predicted from the computer-assisted image analysis. The present results offer new information about using sensory analysis of MR-images not only for food science but also for medical applications for analysing MR and X-ray images and for training of personnel, such as radiologists and radiographers.

(C) 2002 Elsevier Science Ltd.

General information
State: Published
Organisations: Department of Biotechnology
Authors: Martens, H. (Intern), Thybo, A. (Ekstern), Andersen, H. (Ekstern), Karlsson, A. (Ekstern), Donstrup, S. (Ekstern), Stokdilke-Jorgensen, H. (Ekstern), Martens, M. (Ekstern)
Pages: 70-79
Publication date: 2002
Main Research Area: Technical/natural sciences
Analysis of designed experiments by stabilised PLS Regression and jack-knifing

Pragmatical, visually oriented methods for assessing and optimising bi-linear regression models are described, and applied to PLS Regression (PLSR) analysis of multi-response data from controlled experiments. The paper outlines some ways to stabilise the PLSR method to extend its range of applicability to the analysis of effects in designed experiments. Two ways of passifying unreliable variables are shown. A method for estimating the reliability of the cross-validated prediction error RMSEP is demonstrated. Some recently developed jack-knifing extensions are illustrated, for estimating the reliability of the linear and bi-linear model parameter estimates. The paper illustrates how the obtained PLSR "significance" probabilities are similar to those from conventional factorial ANOVA, but the PLSR is shown to give important additional overview plots of the main relevant structures in the multi-response data. The study is part of an
ongoing effort to establish a cognitively simple and versatile approach to multivariate data analysis, with reliability assessment based on the data at hand, and with little need for abstract distribution theory [H. Martens, M. Martens, Multivariate Analysis of Quality. An Introduction, Wiley, Chichester, UK, 2001].

**General information**
State: Published
Organisations: Center for Microbial Biotechnology
Authors: Martens, H. (Intern), Høy, M. (Ekstern), Westad, F. (Ekstern), Folkenberg, D. (Ekstern), Martens, M. (Ekstern)
Pages: 151-170
Publication date: 2001
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Chemometrics and Intelligent Laboratory Systems
Volume: 58
Issue number: 2
ISSN (Print): 0169-7439
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 2.82 SJR 0.672 SNIP 1.222
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.6 SJR 0.652 SNIP 1.213
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.665 SNIP 1.258 CiteScore 2.68
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.88 SNIP 1.762 CiteScore 2.96
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.892 SNIP 1.43 CiteScore 2.67
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.87 SNIP 1.627 CiteScore 2.68
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.77 SNIP 1.323 CiteScore 2.27
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.761 SNIP 1.152
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 1.082 SNIP 1.314
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 0.96 SNIP 1.262
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 1.244 SNIP 1.324
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.878 SNIP 1.625
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 1.291 SNIP 1.522
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 1.513 SNIP 1.888
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 1.579 SNIP 1.785
Scopus rating (2002): SJR 0.876 SNIP 1.338
Reliable and relevant modelling of real world data: a personal account of the development of PLS Regression

Why and how the Partial Least Squares Regression (PLSR) was developed, is here described from the author's perspective. The paper outlines my frustrating experiences in the 70ies with two conflicting and equally over-ambitious and oversimplified modelling cultures - in traditional chemistry and in traditional statistics. It describes my mental progress of first learning to combine them into least squares "unmixing" of known chemical mixtures, and later extending this into the "unscrambling" of partially unknown structures as well. The bi-linear regression framework is summarised in terms of the development from Principal Component Regression into the PLSR. Finally, the versatility of the PLSR is discussed in light of the urgent need for better education in scientific data analysis.

General information
State: Published
Organisations: Department of Systems Biology
Authors: Martens, H. (Intern)
Pages: 85-95
Publication date: 2001
Main Research Area: Technical/natural sciences

Publication information
Journal: Chemometrics and Intelligent Laboratory Systems
Volume: 58
Issue number: 2
ISSN (Print): 0169-7439
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 2.82 SJR 0.672 SNIP 1.222
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.6 SJR 0.652 SNIP 1.213
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.665 SNIP 1.258 CiteScore 2.68
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.88 SNIP 1.762 CiteScore 2.96
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.892 SNIP 1.43 CiteScore 2.67
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.87 SNIP 1.627 CiteScore 2.68
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.77 SNIP 1.323 CiteScore 2.27
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.761 SNIP 1.152
BFI (2009): BFI-level 2
Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR)

General information
State: Published
Organisations: Department of Biotechnology
Authors: Martens, H. (Intern), Martens, M. (Ekstern)
Pages: 5-16
Publication date: 2000
Main Research Area: Technical/natural sciences

Publication information
Journal: Food Quality and Preference
Volume: 11
Issue number: 1-2
Original language: English
Source: orbit
Source-ID: 177388
Publication: Research - peer-review › Journal article – Annual report year: 2000

Multivariate analysis of quality

General information
State: Published
Organisations: Department of Biotechnology
Authors: Martens, H. (Intern), Martens, M. (Ekstern)
Publication date: 2000

Publication information
Place of publication: London, UK
Publisher: Wiley
Original language: English
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 177457
Publication: Research - peer-review › Book – Annual report year: 2000
Power of experimental designs, estimated by Monte Carlo simulation

General Information
State: Published
Organisations: Department of Biotechnology
Authors: Martens, H. (Intern), Dijkstra, G. B. (Ekstern), Byrne, D. V. (Ekstern)
Pages: 441-462
Publication date: 2000
Main Research Area: Technical/natural sciences

Publication Information
Journal: Journal of Chemometrics
Volume: 14
Issue number: 5-6
ISSN (Print): 0886-9383
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): SNIP 0.832 SJR 0.54 CiteScore 1.7
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): SJR 0.502 SNIP 0.869 CiteScore 1.71
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.556 SNIP 0.973 CiteScore 1.82
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 0.55 SNIP 1.28 CiteScore 1.72
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 0.757 SNIP 1.307 CiteScore 2.2
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 0.673 SNIP 1.335 CiteScore 2.03
ISI indexed (2012): ISI indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 0.735 SNIP 1.044 CiteScore 1.66
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 0.6 SNIP 0.75
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 0.936 SNIP 0.993
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 0.861 SNIP 0.903
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.837 SNIP 1.075
Scopus rating (2006): SJR 1.085 SNIP 1.488
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 1.452 SNIP 1.358
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 1.425 SNIP 1.362
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 1.545 SNIP 1.467
Web of Science (2003): Indexed yes
Sensory profiling data studied by partial least squares regression

General information
State: Published
Organisations: Department of Biotechnology
Authors: Martens, M. (Ekstern), Bredie, W. L. P. (Ekstern), Martens, H. (Intern)
Pages: 147-149
Publication date: 2000
Main Research Area: Technical/natural sciences

Publication information
Journal: Food Quality and Preference
Volume: 11
Issue number: 1-2
Original language: English
Source: orbit
Source-ID: 177389
Publication: Research - peer-review › Journal article – Annual report year: 2000

Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression

General information
State: Published
Organisations: Department of Biotechnology
Authors: Westad, F. (Ekstern), Martens, H. (Intern)
Pages: 117-124
Publication date: 2000
Main Research Area: Technical/natural sciences

Publication information
Journal: Journal of Near Infrared Spectroscopy
Volume: 8
Issue number: 2
ISSN (Print): 0967-0335
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): SNIP 0.677 SJR 0.362 CiteScore 1.16
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): SJR 0.386 SNIP 0.798 CiteScore 1.23
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.415 SNIP 0.763 CiteScore 1.48
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 0.523 SNIP 0.915 CiteScore 1.57
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 0.573 SNIP 1.17 CiteScore 1.69
Multivariate data analysis for more effective R&D and better quality control in the laboratory

General information
State: Published
Organisations: Department of Biotechnology, Section for Aquatic Lipids and Oxidation, National Institute of Aquatic Resources
Authors: Martens, H. (Intern), Martens, M. (Ekstern), Jacobsen, C. (Intern)
Pages: 9-17
Publication date: 1999
Main Research Area: Technical/natural sciences

Publication information
Journal: Managing the Modern Laboratory
Volume: 4
ISSN (Print): 1082-5878
Ratings:
ISI indexed (2013): ISI indexed no
ISI indexed (2012): ISI indexed no
ISI indexed (2011): ISI indexed no
Original language: English
Source: orbit
Source-ID: 226606
Publication: Research - peer-review › Journal article – Annual report year: 1999

Projects:

Kvalitet af muskelbaserede fiskeprodukter
Department of Systems Biology  
Period: 01/10/1998 → 17/05/2004  
Number of participants: 7  
Phd Student:  
**Jensen, Kristina Nedenskov (Intern)**  
Supervisor:  
**Jørgensen, Bo Munk (Intern)**  
**Martens, Harald (Intern)**  
Main Supervisor:  
**Nielsen, Jette (Intern)**  
Examiner:  
**Jessen, Flemming (Intern)**  
**Frisvad, Jens Christian (Intern)**  
**Ofstad, Ragni (Ekstern)**  

**Financing sources**  
Source: Internal funding (public)  
Name of research programme: Forskerakademiets Samfinansier  
Project: PhD

**Applied multivariate dataanalysis and measurement techniques.**  
National Institute of Aquatic Resources  
Department of Biotechnology  
Department of Systems Biology  
University of Copenhagen  
Period: 01/01/1997 → 01/01/9999  
Number of participants: 4  
Project participant:  
**Berner, Lis (Intern)**  
**Martens, Harald (Intern)**  
**Munck, Lars (Ekstern)**  
Project Manager, organisational:  
**Jørgensen, Bo Munk (Intern)**  
Project

**Emission af kulbrinter fra lean.brun gasmotorer**  
Department of Chemical and Biochemical Engineering  
Period: 01/08/1996 → 06/09/1999  
Number of participants: 4  
Phd Student:  
**Bendtsen, Anders Broe (Intern)**  
Supervisor:  
**Glarborg, Peter (Intern)**  
Main Supervisor:  
**Dam-Johansen, Kim (Intern)**  
Examiner:  
**Martens, Harald (Intern)**  

**Financing sources**  
Source: Internal funding (public)  
Name of research programme: Kandidatstipendium ansat på DT  
Project: PhD