Evaluation and comparison of different approaches to multi-product brix calibration in near-infrared spectroscopy

Michael Kopf, Robin Gruna, Thomas Längle, Jürgen Beyerer

Fraunhofer IOSB, Visual Inspection,
Fraunhoferstr. 1, 76131 Karlsruhe

Abstract Near-infrared (NIR) spectroscopy became a widespread technology for qualitative and quantitative material analysis. New fields of application of this technology, e.g., quantitative food analysis for consumers, increase demand for multi-product calibration models. Conventional multivariate calibration methods, such as partial least squares regression (PLSR), are reported to show weakness in predictive performance [1]. Preliminary studies in multi-product calibration for quantitative analysis of food with near-infrared spectroscopy showed good results for memory-based learning (MBL) and a classification prediction hierarchy (CPH) [2]. In this study, three varieties of apples, pears and tomatoes with known °brix value are analyzed with NIR spectroscopy in the range from 900 nm to 2400 nm. Predictive performance of a linear PLSR model, two nonlinear models (CPH and MBL) and different preprocessing techniques are tested and evaluated. For error estimation, leave-one-product-out and leave-one-out cross-validation are used.

Keywords: NIR, chemometrics, nutrition, multi-product calibration.

1 Introduction

Near-infrared spectroscopy became a widespread technology for quality inspection and optical sorting issues. Due to its ability for nondestructive quantitative and qualitative analysis, it can be found in production chains and in laboratories. Unlike mid-infrared, quantitative

information can not be read directly from the spectrum. Mathematical methods and models, called chemometric methods, must be applied to gain information. For quantitative analysis in NIR spectroscopy partial least squares regression and principal component regression are gold standard methods [3]. With increasing interest in new areas of application, such as the handheld food scanners for consumers [4, 5], the requirements for chemometric methods changed. In contrast to laboratory tasks, there is no prior knowledge about the samples being scanned. Robust multi-product calibrations are necessary. Preliminary studies in memory-based learning and a classification prediction hierarchy showed promising results [2]. In this experiment, those two non-linear methods and a linear PLSR model were tested and evaluated. The task was to predict the °brix value, which is highly correlating to sugar content, for three apple varieties, tomatoes and pears. To this end, hyperspectral short wave infrared (SWIR) images were used. Five preprocessing techniques and combinations of them were tested: first and second derivative, standard normal variate transformation (SNV), multiplicative scatter correction (MSC) and absorbance transformation.

2 Material and methods

2.1 Data acquisition

Three varieties of apples (Jonagold, Gala and Elstar) were used to get product separation on different levels. For a separation on a higher product level, tomatoes and pears were added to the samples. Each fruit was cut into two halves, the ripest and the most unripe half. A SWIR line camera was used to obtain hyperspectral reflectance images in the range of 900 nm to 2400 nm. Six halogen lamps in bright field constellation were used as light source. Dark and white reference images were acquired at the beginning of each measurement and used for reflectance calculation. For the white reference image, a white teflon bar was used. A total amount of 124 pear, 454 apple (146 Elstar, 146 Gala, 162 Jonagold) and 90 tomato images were taken. After the hyperspectral image acquisition of each half, its middle third part was extracted and pressed to juice. The °brix value for each half was determined with a refractometer.

2.2 Data processing

After defective pixel elimination, reflectance calculation and segmentation, the median spectrum from each image was calculated. A median spectrum is more robust against outliers and other influences than the raw point spectra. The median spectra serve as basis for further preprocessing and analysis.

Preprocessing was used to remove scatter effects or to extract different features. First two derivatives, SNV and MSC are used as well as absorbance transformation in combination with the other preprocessing techniques.

PLSR   A conventional PLSR from the *R*-Package *PLS* [6] was used as linear calibration. The number of components was chosen via leave-one-product-out (LOPO) and leave-one-out (LOO) cross-validation. All data were pooled to one dataset for training and validation to test multi-product prediction performance.

MBL   In contrast to so called eager learning, like PLSR, there is no offline training in memory-based learning or lazy learning. To predict the response variable(s) for a sample, a distance metric is used to find nearest neighbours in the training data. A regression model is trained with those nearest neighbours on demand (see Fig. 13.1). Parameters to set are among others the distance metric to find the nearest neighbours, the number of neighbours, the regression algorithm and the use of the dissimilarity matrix. The used *R*-package *resemble* [7] offers three kinds of distance metrics and four regression methods (see [7] for more details). It is also possible to use the distance information as additional predictor variables or as weights for weighted regressions.

CPH   Another approach for multi-product calibrations is classification prediction hierarchy (CPH). For each class or subclass, a specialized regression model is trained. For prediction, the optimal model is chosen by a classifier. Then the regression model predicts the response variable. Due to the bad results for applying a specialized model to another class than it was trained on [2], the models are evaluated only with LOO cross-validation.

+ sample to predict
+ sample used to train prediction model
+ ignored samples
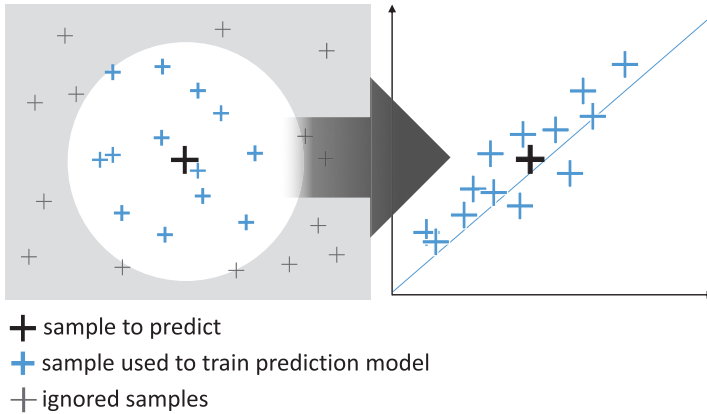
Figure 13.1: Memory-based learning: prediction model is build on demand with the nearest neighbours of the new sample.
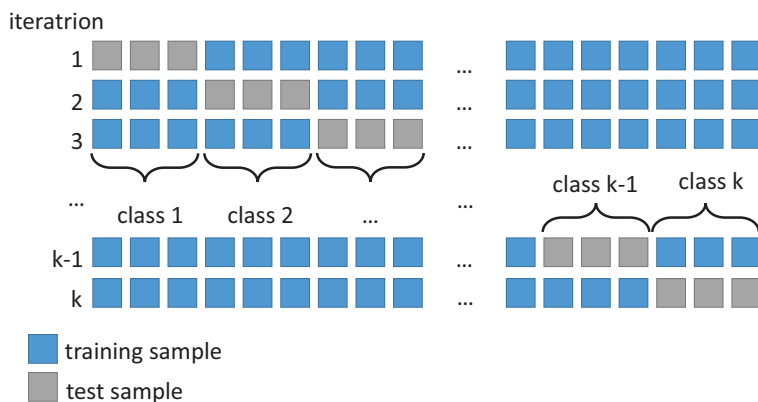
2.3 LOPO cross-validation

To validate the calibration models, a leave-one-product-out cross-vali-dation was used. To estimate the ability of prediction on materials that are not in the training set, the models were trained with all classes but one. The class left out was used for validation (Fig. 13.2). The root mean square error of cross-validation (RMSECV) was used as error function:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^{n}(p_i - r_i)^2}{n}},$$

where p_i is the i-th predicted value, r_i is the i-th observed reference value, i is in the range of $[1, n]$ and n is the number of test samples.

3 Results

PLSR   Best linear calibration model for a LOO cross-validation was trained with 16 components and absorbance transformation as pre-processing (Fig. 13.3 left). The RMSECV is 0.78 °brix. With a LOPO cross-validation best PLSR model was gained with absorbance trans-

Figure 13.2: Leave-one-product-out cross-validation with k classes.

formation and MSC as preprocessing. The RMSECV is 1.13 °brix with 6 components.

MBL Memory-based learning calibrations were tested with LOPO cross-validation to estimate the ability for predicting unknown materials. Only little improvement to the PLSR model is possible with a MBL calibration that uses euclidean distance for dissimilarity calculation, PLSR as regression algorithm with 8 components for prediction, dissimilarity matrix as additional predictor variables and 300 nearest neighbours. Samples were preprocessed by an absorbance transformation. The root mean square error of cross-validation was 1.11 °brix (see Fig. 13.4 left).

CPH Linear discriminant analysis was used to classify and to choose prediction models. Best models were obtained with 6 components PLSR and MSC preprocessing for pears, 12 components PLSR and absorbance transformation for Elstar apples, 13 components PLSR and absorbance transformation for Gala apples, 17 components PLSR and SNV after absorbance transformation for Jonagold apples and 7 components and SNV after absorbance transformation for tomatoes. With
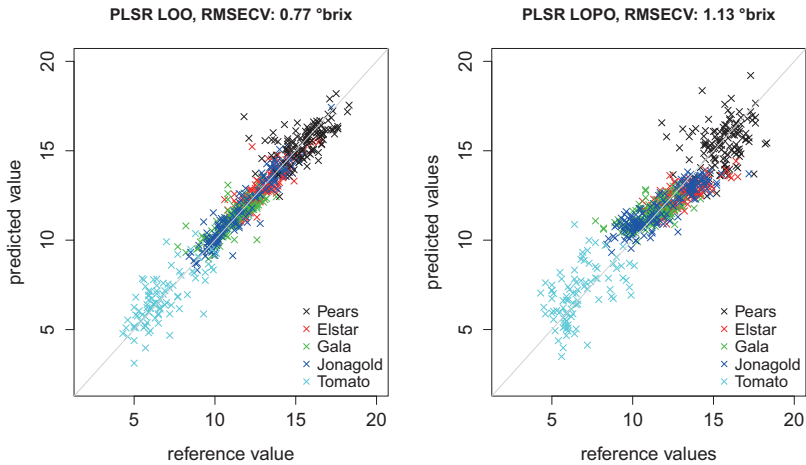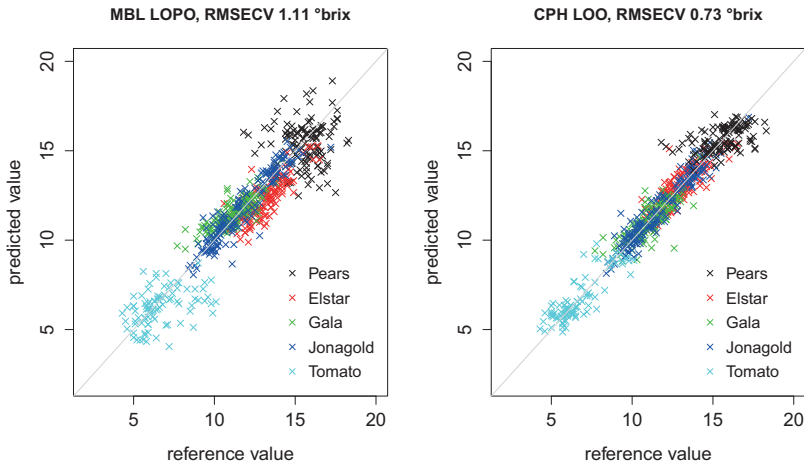
Figure 13.3: Left: best PLSR °brix calibration model, validated with LOO cross-validation, absorbance transformation as preprocessing. Right: best PLSR °brix model, validated with LOPO cross-validation, absorbance transformation and MSC as preprocessing.

those specialized models a RMSECV of 0.73 °brix was reached in LOO cross-validation (see Fig. 13.4 right).

4 Conclusion

In contrast to the study of Micklander et al. [1], where PLSR calibrations showed weakness in prediction compared to nonlinear methods as local regressions and neuronal nets, a linear multi-product calibration for °brix value on apples, pears and tomatoes shows comparable accuracy to nonlinear models. MBL and CPH can only reach little improvement in predictive performance. A hierarchical calibration, such as the classification prediction hierarchy, can increase accuracy in °brix prediction for apple varieties, pears and tomatoes from 0.78 °brix to 0.73 °brix.

Figure 13.4: Left: best MBL °brix calibration for LOPO with 300 nearest neighbours and absorbance transformation as preprocessing. Right: best CPH °brix model, validated with LOO cross-validation.

Memory-based learning as multi-product calibration approach reaches little better result for leave-one-product-out cross-validation than the pooled PLSR model. The RMSECV for best PLSR model is 1.13 °brix while MBL reaches 1.11 °brix. Especially memory-based learning has a high calculation effort by calculating big dissimilarity matrices for nearest neighbour search, which slows down prediction. As you can see in Fig. 13.3 right, the outer classes, pears and tomatoes, have a higher error and a higher deviation than the apple varieties in a LOPO cross-validation. When predicting in a LOPO cross-validation, the model has to extrapolate the response variable, which might cause the higher error, as conventional regression models are said not to be able to extrapolate without loss in accuracy [8]. Same effects are noticeable for MBL calibrations (Fig. 13.4).

References

1. E. Micklander, K. Kjeldahl, M. Egebo, and L. Norgaard, "Multi-product calibration models of near infrared spectra of foods," *Journal of Near Infrared Spectroscopy*, vol. 14, pp. 395–402, 2006.

2. M. C. Kopf and R. Gruna, "Examination of multiproduct calibration approaches for quantitative analysis of food with near infrared spectroscopy," Bachelor's thesis, 2016.

3. H. W. Siesler, Y. Ozaki, S. Kawata, and H. M. Heise, Eds., *Near-Infrared Spectroscopy: Principles, Instruments, Applications (Chemistry)*, 1st ed.   Wiley-VCH Verlag GmbH & Co. KG, 2002.

4. K. Grifantini, "Knowing what you eat: Researchers are looking for ways to help people cope with food allergies," *IEEE Pulse*, vol. 7, 2016.

5. H. Schulte, G. Brink, R. Gruna, R. Herzog, and H. Grueger, "Utilization of spectral signatures of food for daily use," in *OCM 2015 - Optical Characterization of Materials - conference proceedings*, 2015.

6. R. W. Bjorn-Helge Mevik and K. H. Liland, "pls: Partial least squares and principal component regression," CRAN repository. [Online]. Available: https://CRAN.R-project.org/package=pls

7. L. Ramirez-Lopez and A. Stevens, "resemble: Regression and similarity evaluation for memory-based learning in spectral chemometrics," CRAN repository, 2016. [Online]. Available: https://CRAN.R-project.org/package=resemble

8. E. W. Ciurczak and D. A. Burns, Eds., *Handbook of Near Infrared Analysis*, 3rd ed.   Practical Spectroscopy, 2007.