Croat. J. Food Sci. Technol. (2019) 11 (1) 104-112

Croatian Journal of Food Science and Technology

journal homepage: www.ptfos.unios.hr/cjfst/

Original scientific paper

DOI: 10.17508/CJFST.2019.11.1.15

Near-infrared diffuse reflectance spectroscopy for discriminating fruit and vegetable products preserved in glass containers

MANUEL ZIMMER*, JAN SCHNEIDER

Institute for Food Technology.NRW & OWL University of Applied Sciences and Arts, Department of Life Science Technologies, Campusallee 12, 32657 Lemgo, Germany

ARTICLE INFO

Article history:
Received: April 3, 2019
Accepted: May 5, 2019

Keywords:
digital image
food preservation and processing
food safety and quality
multivariate data analysis
pasteurization

ABSTRACT

Near-infrared (NIR) diffuse reflectance spectroscopy was used in combination with multivariate analytical methods to discriminate between different fruit and vegetable products preserved in glass containers, which are commonly used as receptacles for the pasteurization of fruit and vegetable products. To investigate the samples in this way, i.e. inside the sealed glass containers, is important for this specific application in a food processing facility. In order to adapt digitalization technologies to the pasteurization process, it is necessary to investigate usually consumed products with suitable sensors and data analytics. NIR spectroscopy in combination with multivariate data analysis is a mighty tool to unravel various issues in food research and industry. Thus, this combination is in the focus of this investigation. It is shown for the first time that the discrimination between five types of preserved food in glass containers is possible by using NIR diffuse reflectance spectroscopy and multivariate data analysis (including discrimination methods). The performance parameters sensitivity, specificity, and efficiency, are determined for every product group and analyzed in a misclassification table. On average, the results show that 95% of ca. 2100 observations are correctly classified with partial least squares discriminant analysis (PLS-DA).

Introduction

Digitalization is propagating in all areas of industry and is becoming more and more important. Unsurprisingly, the already technologized food industry is strongly affected by this development. The new industrial revolution is often titled with catchphrases like Industry 4.0 (I4.0), cyber-physical systems (CPS), internet of things (IoT), cloud computing, and so on (König and Thongpull, 2015; Li et al., 2018; Pang et al., 2015; Roblek et al., 2016; Trappey et al., 2016; Verdouw et al., 2016), leaving latitude for their meaning in the respective field of research or application. An important aspect in the food industry is traceability and interconnection of production steps, which needs sensors to generate and analysis tools to handle a great amount of data. Using the possibility of sensors and analysis tools, energy, time, and waste in food processes can be reduced while the quality of the products can be increased and maintained on a high level. In order to integrate these new technologies, i.e. sensors, digitalization, automation, in a food processing facility, firstly they have to be tested and validated in laboratory studies.

One important manufacturing process in the food industry is the preservation of foodstuff. This can be done in many ways, but pasteurization (Silva and Gibbs, 2004) is still one of the most common procedures and will be also applied in Industry 4.0 food processing. The pasteurization of cabbage and fruit products is usually performed by tunnel pasteurization. The already sealed product is heated up, held for a certain time and cooled down together with packing and closures (Brody and Ryan, 1971). As there is a thermal load on the pasteurized product, degradation of nutrients and negative effects on the sensory quality, including appearance, shall be avoided by a gentle treatment. However, since the

*Corresponding author E-mail: manuel.zimmer@th-owl.de

microbiological safety may not be impaired, the quality of the product often suffers due to safety margins. The reason is that in the hitherto existing practice of pasteurization can only roughly adjust the process parameters time and temperature to the requirements of the actually treated product. In order to individually customize the two core process parameters, temperature and exposition time, detailed knowledge about the case specific microbiological inactivation kinetics is required. Today, a production line that is connected to cloud-based databases providing data of the thermal death kinetics of microorganism (D and z values) becomes possible. The LDz-Base is one example for such a free access to scientific data collection (Schwarzer et al., 2010). Those data integrated in the control system of the pasteurization plant, e g. with the help of the IoT, will be capable to determine the case-customized time-temperature settings. However, for autonomous production system the identification of the kind of incoming foodstuff is necessary first, because the relevance of different microorganism species is strongly related to the kind of product.

Therefore, for a fully automatized and autonomous pasteurization process it is required that the product is automatically identified and can be clearly discriminated from other products being processed in the same facility. To establish this step, investigating the possibilities of near infrared spectroscopy (NIRS) to become a key technology in Industry 4.0 is necessary.

NIRS is a spectroscopic method in the spectral region from 2500 to 700 nm in which combination and overtone vibrations in molecules are excited. Molecules containing C-H, O-H, N-H and S-H functional groups are the ones being in the focus by NIR spectroscopists. Therefore, it is extraordinarily well suited for investigations on foodstuff and the application in food producing facilities (Büning-Pfaue, 2003; Cozzolino, 2014; Delwiche et al., 1992; González-Caballero et al., 2010; Lee et al., 2015; Paradkar et al., 2002; Slaughter, 1995; Xie et al., 2009). Another advantage of NIRS compared to other methods is its universal application on samples in all forms and even through packaging.

In the food industry, a multitude of different materials are available for packaging food products, all bearing their individual pros and cons (Farmer, 2013; Robertson, 2006, 2010). Although packaging made from plastics are steadily improving, e.g. in shelf life, recyclability, and weight; glass containers are unsurpassable when aesthetic appeal and quality is important (Farmer, 2013). Glass containers, with a market share of 11% in 2010 on global consumer packaging (Farmer, 2013), are a popular receptacle to pasteurize fruit and vegetable products, as glass is inert, does not alter the flavour of the product, can easily withstand high temperatures, can be reused or recycled, and, what is probably the most important for food retailers, the product can be seen by the customers.

Measuring NIR spectra of samples in glass containers is more difficult than using an in-line probe, despite the high transmittance of glass in the NIR spectral range. For in-line measurements, transmission or transflection probes, that measure the absorption of NIR radiation by the sample, can be used directly. This is not possible for products in glass containers, especially if they are pasteurized and cannot be opened for measurement. Therefore, a diffuse reflectance (DR) NIR probe is used. DR NIRS is usually applied to products possessing a low amount of water, like powders, crystals, dried foodstuff, oils and molasses, because water reflects only a small amount of the NIR radiation (Catelani et al., 2017; Ferreira et al., 2014; Lü et al., 2013; Smeesters et al., 2016; Wedding et al., 2013). However, it is shown here that also samples with a high amount of water can be investigated. A key requirement for the successful application of NIRS in this field is the analysis of NIR-data with multivariate data analysis (MVDA). As NIR spectra usually generate more than 100 variables to be used in MVDA and many samples can be measured in a short time, it is a popular and valuable source to obtain data. Publications dealing with multivariate classification methods in regard of food-authenticity and characterization can be found abundantly. Some focus on different methods and provide an overview of the field, some address very specific problems using only a few selected methods. Oliveri et al. 2012 is showing an overview of class modelling and discriminant methods for food-authenticity. Furthermore, performance parameters for these methods are defined; as these methods are qualitative, it is necessary to use different parameters as for quantitative methods. Performance parameters for class modelling and discriminant methods are, for example, sensitivity, specificity, efficiency, precision, Matthews correlation coefficient and accuracy. Others used class modelling and discriminant methods for specific issues like adulterants in raw milk (Botelho, 2015), almond classification based on bitterness (Cortés, 2018), geographical discrimination of saffron (Liu, 2018), identification of ground meat species (Pieszczek, 2018), or adulteration of Norwegian salmon (Wu, 2018). Despite the great number of publications regarding class modelling and discriminant methods, the authors could not find a publication that is dealing with the discrimination of food preserved in glass containers.

In this study, various preserved food samples in glass containers are investigated using near-infrared diffuse reflectance spectroscopy in combination with multivariate data analysis. It is the aim of this study to investigate if these methods, applied to this type of samples, are suited to provide appropriate data to discriminate the individual samples.

Materials and methods

Samples

Commercially available fruits and vegetables preserved in glass containers (tangerine slices, pickled red cabbage, pickled cornichons, kale and applesauce) were purchased at the local supermarket, three of each kind. Specific information regarding the different samples, as nutrition facts, ingredients and best before date are summarized in Table SI 1 in the supporting information.

Spectroscopic measurements

NIR diffuse reflectance measurements were performed using a contact scan head (PSS-H-B01, equipped with a 20 W tungsten-halogen lamp) coupled to a diode array NIR spectrometer (PSS-2120) (both Polytec GmbH, Waldbronn, Germany) equipped with a 256 Pixels InGaAs detector. The samples were placed on the top of a specifically designed and 3D-printed sample holder, improving the reproducibility of the measurements (see Figure 1 for schematic depiction). Spectra were collected using PAS LABS version 1.2 (Polytec GmbH) in the wavelength range between 1100 and 2100 nm. Data were recorded in reflectance mode and are corrected by a dark spectrum and a reference spectrum, for which water was used. All spectra were recorded at room temperature. Each sample was measured 125 times (or 250 times) with 64 scans for each spectrum using an integration time of 100 ms. Due to the inhomogeneity of the samples, particularly the pickled cornichons, and the mobility of the fruit or vegetable chunks in the glass containers, especially in the case of tangerine slices and pickled cornichons, this number of measurements was required to get an appropriate number of reproduced spectra.

Data analysis and validation

Data pre-treatment and multivariate analysis was performed using the SIMCA software, version 14.1 (MKS Umetrics). Principal component analysis (PCA), projection to latent structures discriminant analysis (PLS-DA), and orthogonal projections to latent structures discriminant analysis (OPLS-DA) have been used for data analysis.
PCA is a basic tool in multivariate data analysis that can extract dominant patterns in huge datasets and can show relations between variables and observations that are otherwise hidden in the broad and mostly uncharacteristically structured NIR spectra (Cozzolino and Murray, 2004; Hämäläinen and Albano, 1992; Shumilina et al., 2016; Wold et al., 1987). PLS-DA is used to discriminate qualitatively between classes that are defined as Y-data. Similar to the PCA approach, a data reduction is performed and the observations (X-variables) are correlated with the classes, if a correlation exists (Barker and Rayens, 2003; Botelho et al., 2015; Höskuldsson, 1988; Kalivodová et al., 2015; van Ruth et al., 2010; Wold et al., 2001; Wu et al., 2018).

Fig. 1. Schematic view of the experimental setup including contact scan head, sample holder, and sample vessel

OPLS-DA is an improvement to the PLS-DA model, leading to a better interpretation of results due to its ability to remove orthogonal (non-correlating) data from the X-variables using a build-in orthogonal signal correction (OSC) (Bylesjö et al., 2006; Bylesjö et al., 2007; Trygg, 2002; Trygg and Wold, 2002, 2003; Worley and Powers, 2016). However, the outcome from an OPLS-DA is often too optimistic as the integrated OSC filter removes systematic spectral variation that does not approve the assigned group memberships, and therefore it needs proper validation (Worley and Powers, 2016).

One way to test the resilience of a model is by performing a response permutation test. This is a procedure which fits the (O)PLS-DA model several times with unchanged X-values but with a randomly permutation of the Y-vector. By repeating these permutations, a statistical significance of the R^2- and Q^2-parameters is generated (Eriksson et al., 2008). The outcome of this test provides an indication whether the model is the product of pure chance or a systematic approach that has the ability to reliably predict new observations. To have a performance measure of qualitative discriminant models, and being able to compare different models, the following parameters are frequently used. (Oliveri and Downey, 2012) With these parameters it is also possible to define a threshold value above which a model is considered appropriate. To understand the meaning of the parameters sensitivity (sens), specificity (spec) and efficiency (eff), the terms true positive (TP), false negative (FN), false positive (FP), and true negative (TN) have to be introduced first.

TP: samples that are correctly assigned to the class they belong to. Example: Observation belongs to class 1 and is correctly assigned to class 1.

FN: samples that are not assigned to the class they belong to. Example: Observation belongs to class 1, but is assigned to any other class or no class at all.

FP: samples that are assigned to a class they do not belong to. Example: Observation belongs to any class except class 1, but is assigned to class 1.

TN: samples that do not belong to the respective class and are correctly not assigned to this class. Example: Observation belongs to any class except class 1 and is correctly assigned to a class different than class 1.

With these, the three parameters can be defined as:

$$\mathrm{sens} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (1)$$

$$\mathrm{spec} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \quad (2)$$

$$\mathrm{eff} = \sqrt{\frac{\mathrm{TP} * \mathrm{TN}}{(\mathrm{TP} + \mathrm{FN}) * (\mathrm{TN} + \mathrm{FP})}} \quad (3)$$

The sensitivity of a model can be understood as the fraction of samples that are correctly assigned by the model to the class they belong to. Whereas the specificity is the fraction of samples that are correctly rejected by the model as they do not belong to the class of interest. The efficiency is a parameter summarizing the other two parameters.

Results and discussion

In this study five groups of different fruit and vegetable products preserved in glass containers have been investigated. Each group consisted of three (slightly) different products. NIR diffuse reflectance spectroscopy and multivariate data analysis are used in combination to record and analyze the data. It is the objective of these investigations to clearly discriminate each of the five groups by the afore-mentioned method. For this purpose, models are created and split into two classes (1 and 0). Class 1 contains all the data for one specific product group. Class 0 contains the data of all the other product groups. To evaluate the model performance, the parameters sensitivity, specificity, and efficiency have been chosen. These parameters give information on the ability of a model to predict the affiliation of a product to one of the two classes.

For discrimination of different product types it is useful to apply a combination of PCA and (O-)PLS-DA as this leads to insights on general spectral trends and group-predictive spectral features (Worley and Powers, 2016). PCA is, unlike (O-)PLS-DA, an unsupervised method, which is positive as it is not biased by a response value assigned by the analyst, but also negative as it will only show differences if they have a major contribution in the variables (Worley and Powers, 2016).

The measured spectral region from 1100 to 2100 nm, which covers the complete first and second overtone region, is sampled on the 256 Pixels of the detector, leading to 256 variables for the multivariate data analysis.

In the NIR raw data (Figure 2a), it can be seen that the overall shape of the spectra of the different samples is very similar, with the strongest differences in the region around 1650 to 1750 nm. Considering the sample composition for the macronutrients: carbohydrates in the samples vary between 0.5% and 13.2%, fat between 0% and 1.7%, and proteins between 0.1% and 3.5%. It could be concluded that the differences in carbohydrate concentration are one of the possible factors for the differentiation of the samples. Furthermore, the large differences in carbohydrate content are the reason for the differences in the NIR spectra between 1650 and 1750 nm, because the first overtone of the C-H stretch vibration of carbohydrates is found in this region (Nielsen, 2017). It also becomes more evident when we look at the normalized spectra (see Figure 2b). It is found that the absorption at around 1676 nm correlates with carbohydrate (and sugar) concentration. Plotting the carbohydrate (or sugar) concentration of the different samples against the normalized log(1/R) values at 1676 nm leads to a linear relationship - if we dismiss the two samples with artificial sweeteners added, because it seems that they somehow behave differently (Figure SI 1). In Figure 2b it can also be observed that most of the samples have their peak absorption at around 1440 nm, which is most likely due to the first overtone of the free O-H vibration of water. Additionally to the differences in carbohydrates, the sample form, especially texture and particle size are also considered major factors for sample differentiation. This is due to two reasons. Firstly carbohydrate content also differs within the product groups, e.g. between 0.5 and 7.0 % for kale, and secondly, the sample forms are very similar within each sample group but differ to some extent strongly between groups, e.g. sliced tangerines and apple sauce. Sample characteristics like firmness of different fruits and beans (Lu, 2001; Mendoza et al., 2018; Munera et al., 2018; Wang et al., 2015), and particle size of cereals in flours, meal and plant milk (Ayvaz et al., 2015; Gajdoš Kljusurić et al., 2015; Zhu et al., 2017) have been the issue for many NIR investigations, showing that it is possible to use these characteristics to discriminate different samples or to predict these parameters using PLS or similar models.

Parameters that show the influence of the NIR wavelength on the scores, i.e. loadings, regression coefficients, and variable importance for the projection, do not provide evidence for especially important or unimportant regions in the NIR spectra.

As NIR data pre-treatment, which is usually performed previous to multivariate data analysis, did not provide better results according to the chosen performance parameters, raw data are used for the PLS model development.

As there are many more observations in class 0 than in class 1, the observations were chosen to be approximately the same number for class 1 and 0 in the calibration set. This was set to be half of the observations of class 1 (188 or 250 depending on the product group, respectively). Thus, there are more observations from class 0 in the prediction set than from class 1, but it turned out that this is handled well by the models. The models should be able to classify the data not used in the calibration set into the right class. Results obtained from the prediction tests have been analyzed and summarized in Table 1. For every model and performance parameter, values near 100% are achieved. The only exception is red cabbage (line C in Table 1), for which the false negative rate in class 0 is higher than for the other models, leading to a sensitivity of class 0 of only 83% and an efficiency of 91%.

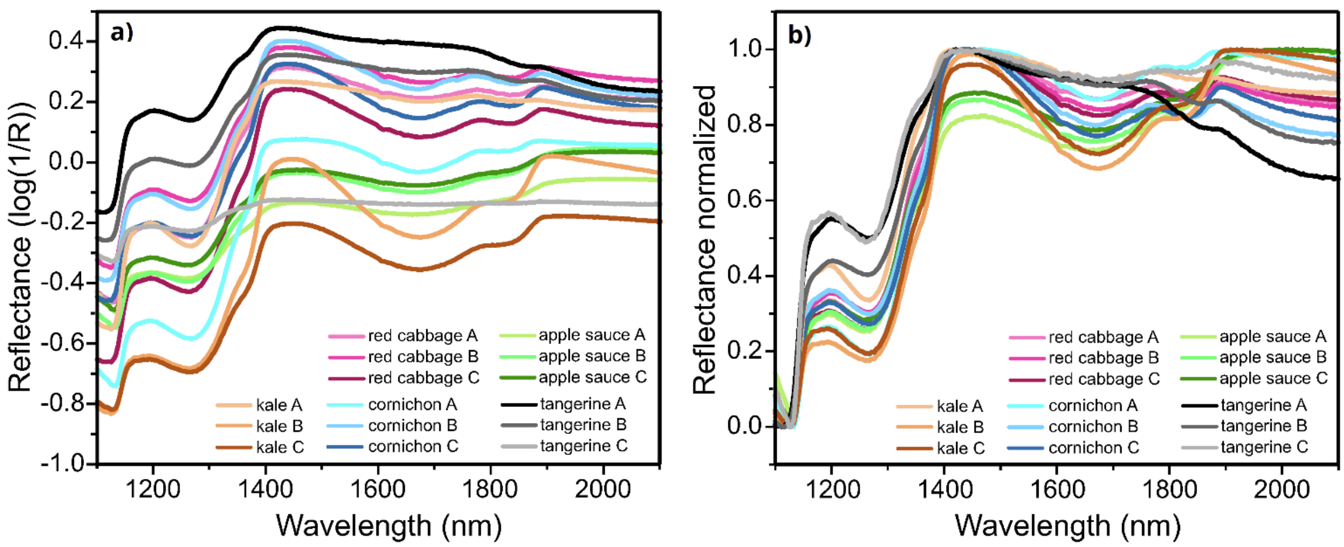

Fig. 2. a) NIR raw data of the 15 different samples, b) normalized spectra.

Table 1. Performance of the created models for each of the five product groups regarding the parameters sensitivity, specificity and efficiency.

	members prediction set class 0/1	members calibration set class 0/1	sensitivity class 0/1 (%)	specificity class 0/1 (%)	efficiency class 0/1 (%)	model type	latent variables
A	1374/250	250/250	98/98	100/100	99/99	OPLS-DA	1+13+0
B	1375/249	232/250	95/100	100/100	98/100	OPLS-DA	1+14+0
C	1566/187	183/188	83/97	100/99	91/98	OPLS-DA	1+16+0
D	1561/187	188/188	98/100	100/99	99/100	OPLS-DA	1+11+0
E	1561/188	188/187	100/100	100/100	100/100	OPLS-DA	1+10+0

A: cornichons, B: tangerines, C: red cabbage, D: apple sauce, E: kale

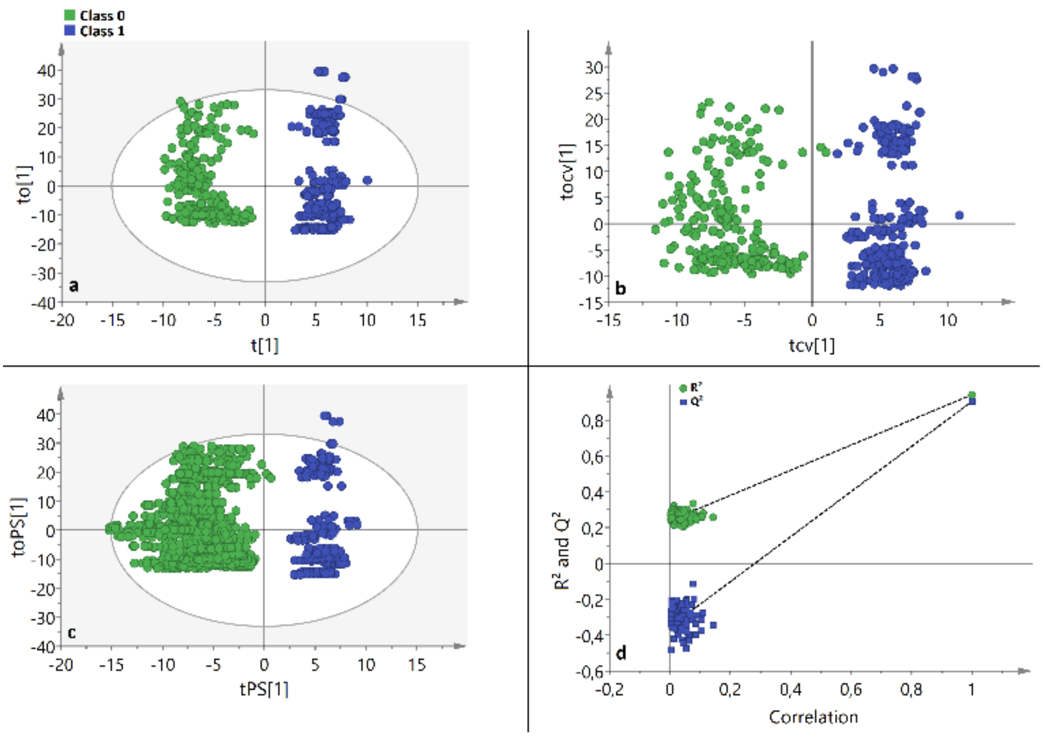

Fig. 3. Class 1 tangerines, Class 0 other samples; a OPLS-DA scores scatter plot, b CV scores scatter plot, c predicted scores scatter plot, d permutation test result (100 permutations, class 0). OPLS-DA 1+14+0 components.

One of the five models shall be discussed here in more detail. For this, the one having sliced tangerines entitled class 1 and the other samples class 0 was chosen (line B in Table 1). Figure 3 and 4 are showing the most important plots explaining the performance and validation of the model in addition to Table 1. Figure 3 shows the scatter plots created from the O-PLS-DA model, segmented in score plots (a), cross-validation score plots (b), and predicted score plots (c). Additionally, the result of the permutation test is shown in segment d. In the three O-PLS-DA scores plots (a, b, c) a clear separation between the two classes is evident, giving a descriptive representation for the very good model performance shown in Table 1. The permutation test results shown in segment d show that the model is resilient and is not over parametrized. Permutation tests have also been performed for all the other models. The results, i.e. bias and slope of the permutation tests are summarized in Table SI 2 in the supporting information. Additionally to the clear and descriptive OPLS-DA plots, also the PCA and PLS-DA scatter plots are shown (see Figure 4 a to d). In segments a and b the scores scatter plot for the PCA and PLS-DA models can be compared.

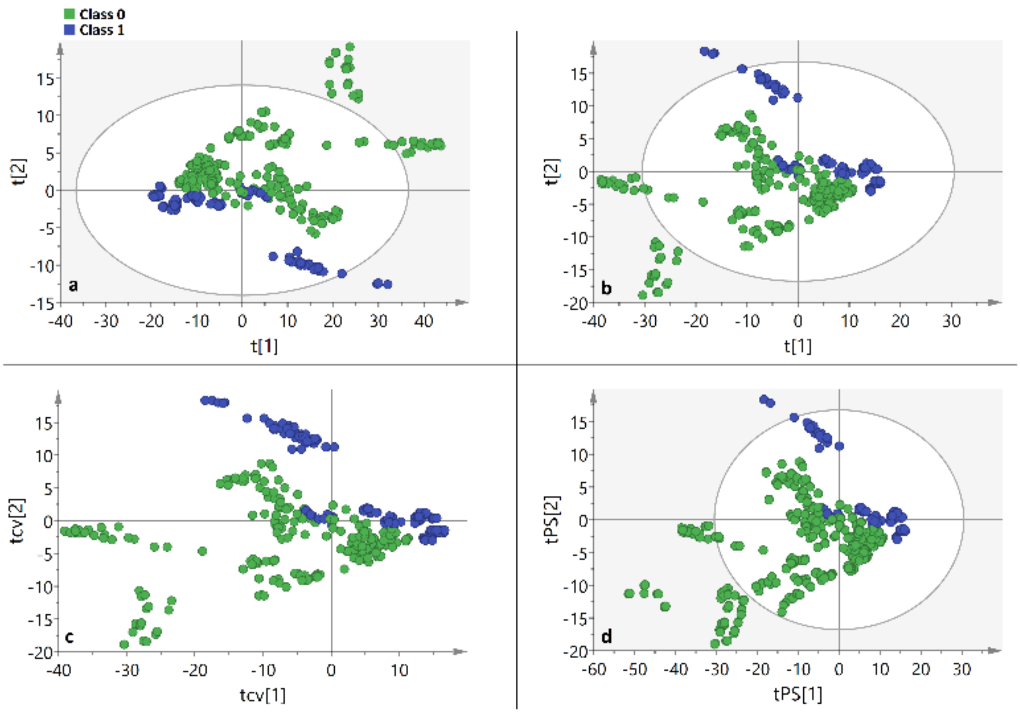

Fig. 4. Class 1 tangerines, Class 0 other samples; **a** PCA scores scatter plot, **b** PLS-DA scores scatter plot, **c** PLS-DA CV scores scatter plot, **d** PLS-DA predicted scores scatter plot. PCA 3 components, PLS-DA 15 components. t[1]: scores first principal component (PC) or latent variable (LV); t[2] scores second PC or LV; tcv: scores from cross validation; tPS: scores from prediction set

It is found that they are very alike if one of them is turned by 180°. In both score plots the classes 1 and 0 are mostly separated but they still overlap to a small degree. Nevertheless, when looking at the cross-validated and predicted PLS-DA score scatter plots, it is found that the two classes can be predicted very accurately. The predictive power of the PLS-DA model is the same as for the OPLS-DA model considering the chosen performance parameters and choosing the same amount of latent variables for the two methods

Conclusion

Near-infrared (NIR) diffuse reflectance spectroscopy in combination with multivariate data analysis (MVDA) can be used to discriminate between five groups of fruits and vegetables (sliced tangerines, pickled red cabbage, kale, apple sauce, and pickled cornichons) preserved in glass containers with an high efficiency (near 99%). Due to complex and partly heterogeneous samples, hundreds of thousands of measurements are required to obtain well working models. The results of this investigation are an important step towards a fully automated and autonomous pasteurization process which uses NIR and MVDA as its main data recording and processing unit. The OPLS-DA method provides the best results according to the score plots, compared to the PCA and PLS-DA method. Considering the performance parameters sensitivity, specificity and efficiency, OPLS-DA and PLS-DA both provide excellent results. We will carry on with similar investigations using NIRS and MVDA to distinguish between products in glass containers before and after they are pasteurized, using chemical properties that change during pasteurization. With this information it should be possible to determine an accurate thermal load that was applied to a product.

Acknowledgments

The Polytec GmbH (Waldbronn, Germany) is acknowledged for the loan of the spectrometer and scan head.

Funding

Financial support by the Federal Ministry of Education and Research (BMBF), project 13FH3E01IA, in the frame of “smart food technology OWL”, and the Ministry of Culture and Science of the German State of North Rhine-Westphalia is gratefully acknowledged.

References

Ayvaz, H., Plans, M., Towers, B.N., Auer, A., Rodriguez-Saona, L.E. (2015): The use of infrared spectrometers to predict quality parameters of cornmeal (corn grits) and differentiate between organic and conventional practices. *J. Cereal Sci.* 62, 22–30. https://doi.org/10.1016/j.jcs.2014.12.004.

Barker, M., Rayens, W. (2003): Partial least squares for discrimination. *J. Chemometr.* 17(3), 166–173. https://doi.org/10.1002/cem.785.

Botelho, B.G., Reis, N., Oliveira, L.S., Sena, M.M. (2015): Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA. *Food Chem.* 181, 31–37. https://doi.org/10.1016/j.foodchem.2015.02.077.

Brody, A.L., Ryan, D.E. (1971): Food canning in rigid and flexible packages. *C R C Critical Reviews in Food Technology* 2(2), 187–243. https://doi.org/10.1080/10408397109527122.

Büning-Pfaue, H. (2003): Analysis of water in food by near infrared spectroscopy. *Food Chem.* 82(1), 107–115. https://doi.org/10.1016/S0308-8146(02)00583-6.

Bylesjö, M., Eriksson, D., Kusano, M., Moritz, T., Trygg, J. (2007): Data integration in plant biology: The O2PLS method for combined modeling of transcript and metabolite data. *Plant J.* 52(6), 1181–1191. https://doi.org/10.1111/j.1365-313X.2007.03293.x.

Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J.K., Holmes, E., Trygg, J. (2006): OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. *J. Chemometr.* 20(8-10), 341–351. https://doi.org/10.1002/cem.1006.

Catelani, T.A., Páscoa, R.N.M.J., Santos, J.R., Pezza, L., Pezza, H.R., Lima, J.L.F.C., Lopes, J.A. (2017): A Non-invasive Real-Time Methodology for the Quantification of Antioxidant Properties in Coffee During the Roasting Process Based on Near-Infrared Spectroscopy. *Food Bioprocess Tech.* 10(4), 630–638. https://doi.org/10.1007/s11947-016-1843-6.

Cozzolino, D. (2014): An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals. *Food Res. Int.* 60, 262–265. https://doi.org/10.1016/j.foodres.2013.08.034.

Cozzolino, D., Murray, I. (2004): Identification of animal meat muscles by visible and near infrared reflectance spectroscopy. *LWT - Food Sci. Technol.* 37(4), 447–452. https://doi.org/10.1016/j.lwt.2003.10.013.

Delwiche, S.R., Pitt, R.E., Norris, K.H. (1992): Sensitivity of Near-Infrared Absorption to Moisture Content Versus Water Activity in Starch and Cellulose. *Cereal Chem.* 69(1), 107–109.

Eriksson, L., Trygg, J., Wold, S. (2008): CV-ANOVA for significance testing of PLS and OPLS® models. *J. Chemometr.* 22(11-12), 594–600. https://doi.org/10.1002/cem.1187.

Farmer, N. (ed) (2013): *Trends in packaging of food, beverages and other fast-moving consumer goods (FMCG): Markets, materials and technologies*. Cambridge, U.K: Woodhead Pub.

Ferreira, D.S., Galão, O.F., Pallone, J.A.L., Poppi, R.J. (2014): Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples. *Food Control* 35(1), 227–232. https://doi.org/10.1016/j.foodcont.2013.07.010.

Gajdoš Kljusurić, J., Benković, M., Bauman, I. (2015): Classification and Processing Optimization of Barley Milk Production Using NIR Spectroscopy, Particle Size, and Total Dissolved Solids Analysis. *J. Chem-NY.* 2015(1), 1–7. https://doi.org/10.1155/2015/896051.

González-Caballero, V., Sánchez, M.-T., López, M.-I., Pérez-Marín, D. (2010): First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery. *J. Food Eng.* 101(2), 158–165. https://doi.org/10.1016/j.jfoodeng.2010.06.016.

Hämäläinen, M., Albano, C. (1992): Principal variations in the chemical composition of peat: Predictive peat scales based on multivariate strategies. *J. Chemometr.* 6, 41–56.

Höskuldsson, A. (1988): PLS regression methods. *J. Chemometr.* 2(3), 211–228. https://doi.org/10.1002/cem.1180020306.

Kalivodová, A., Hron, K., Filzmoser, P., Najdekr, L., Janečková, H., Adam, T. (2015): PLS-DA for compositional data with application to metabolomics. *J. Chemometr.* 29(1), 21–28. https://doi.org/10.1002/cem.2657.

König, A., Thongpull, K. (2015): Lab-on-Spoon – a 3-D integrated hand-held multi-sensor system for low-cost food quality, safety, and processing monitoring in assisted-living systems. *J. Sens. Sens. Syst.* 4(1), 63–75. https://doi.org/10.5194/jsss-4-63-2015.

Lee, K.-M., Davis, J., Herrman, T.J., Murray, S.C., Deng, Y. (2015): An empirical evaluation of three vibrational spectroscopic methods for detection of aflatoxins in maize. *Food Chem.* 173, 629–639. https://doi.org/10.1016/j.foodchem.2014.10.099.

Li, Y., Zhang, J.-Y., Wang, Y.-Z. (2018): FT-MIR and NIR spectral data fusion: A synergetic strategy for the geographical traceability of Panax notoginseng. *Anal. Bioanal. Chem.* 410(1), 91–103. https://doi.org/10.1007/s00216-017-0692-0.

Lu, R. (2001): Predicting Firmness and Sugar Content of Sweet Cherries Using Near-Infrared Diffuse Reflectance Spectroscopy. *T. ASAE* 44(5), 1265–1271.

Lü, Y.-g., Chen, J., Li, X.-q., Wang, C., Wei, L.-l. (2013): Rapid Determination of Water and Oil Content in Instant Noodles by Fourier Transform Near-Infrared Reflectance Spectroscopy. *Food Sci. Technol. Res.* 19(3), 393–398.

Mendoza, F.A., Cichy, K.A., Sprague, C., Goffnett, A., Lu, R., Kelly, J.D. (2018): Prediction of canned black bean texture (Phaseolus vulgaris L.) from intact dry seeds using visible/near infrared spectroscopy and hyperspectral imaging data. *J. Sci. Food Agric.* 98(1), 283–290. https://doi.org/10.1002/jsfa.8469.

Munera, S., Amigo, J.M., Aleixos, N., Talens, P., Cubero, S., Blasco, J. (2018): Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine. *Food Control* 86, 1–10. https://doi.org/10.1016/j.foodcont.2017.10.037.

Nielsen, S.S. (ed) (2017): *Food Analysis*. Cham, CH: Springer.

Oliveri, P., Downey, G. (2012): Multivariate class modeling for the verification of food-authenticity claims. *TrAC-Trend. Anal. Chem.* 35, 74–86.

Pang, Z., Chen, Q., Han, W., Zheng, L. (2015): Value-centric design of the internet-of-things solution for food supply chain: Value creation, sensor portfolio and information fusion. *Inf. Syst. Front. (Information Systems Frontiers)* 17(2), 289–319. https://doi.org/10.1007/s10796-012-9374-9.

Paradkar, M.M., Sakhamuri, S., Irudayaraj, J. (2002): Comparison of FTIR, FT-Raman, and NIR Spectroscopy in a Maple Syrup Adulteration Study. *J. Food Sci.* 67(6), 2009–2015. https://doi.org/10.1111/j.1365-2621.2002.tb09493.x.

Robertson, G.L. (2006): *Food packaging: Principles and practice*. Boca Raton: CRC Press.

Robertson, G.L. (2010): *Food packaging and shelf life: A practical guide*.

Roblek, V., Meško, M., Krapež, A. (2016): A Complex View of Industry 4.0. *SAGE Open* 6(2), 215824401665398. https://doi.org/10.1177/2158244016653987.

Schwarzer, K., Schneider, J., Müller, U., Becker, B., Wilhelm, P. (2010): Lemgo Database for D- and z-values. *Brauwelt International*(V), 252–256.

Shumilina, E., Slizyte, R., Mozuraityte, R., Dykyy, A., Stein, T.A., Dikiy, A. (2016): Quality changes of salmon by-products during storage: Assessment and quantification by NMR. *Food Chem.* 211, 803–811. https://doi.org/10.1016/j.foodchem.2016.05.088.

Silva, F.V.M., Gibbs, P. (2004): Target selection in designing pasteurization processes for shelf-stable high-acid fruit products. *Crit. Rev. Food Sci.* 44(5), 353–360. https://doi.org/10.1080/10408690490489251.

Slaughter, D.C. (1995): Nondestructive Determination of Internal Quality in Peaches and Nectarines. *T. ASAE* 38(2), 617–623.

Smeesters, L., Meulebroeck, W., Raeymaekers, S., Thienpont, H. (2016): Non-destructive detection of mycotoxins in maize kernels using diffuse reflectance spectroscopy. *Food Control* 70, 48–57. https://doi.org/10.1016/j.foodcont.2016.05.039.

Trappey, A.J.C., Trappey, C.V., Govindarajan, U.H., Sun, J.J., Chuang, A.C. (2016): A Review of Technology Standards and Patent Portfolios for Enabling Cyber-Physical Systems in Advanced Manufacturing. *IEEE Access* 4, 7356–7382. https://doi.org/10.1109/ACCESS.2016.2619360.

Trygg, J. (2002): O2-PLS for qualitative and quantitative analysis in multivariate calibration. *J. Chemometr.* 16(6), 283–293. https://doi.org/10.1002/cem.724.

Trygg, J., Wold, S. (2002): Orthogonal projections to latent structures (O-PLS). *J. Chemometr.* 16(3), 119–128. https://doi.org/10.1002/cem.695.

Trygg, J., Wold, S. (2003): O2-PLS, a two-block (X-Y) latent variable regression (LVR) method with an integral OSC filter. *J. Chemometr.* 17(1), 53–64. https://doi.org/10.1002/cem.775.

van Ruth, S.M., Villegas, B., Akkermans, W., Rozijn, M., van der Kamp, H., Koot, A. (2010): Prediction of the identity of fats and oils by their fatty acid, triacylglycerol and volatile compositions using PLS-DA. *Food Chem.* 118(4), 948–955. https://doi.org/10.1016/j.foodchem.2008.10.047.

Verdouw, C.N., Wolfert, J., Beulens, A.J.M., Rialland, A. (2016): Virtualization of food supply chains with the internet of things. *J. Food Eng.* 176, 128–136. https://doi.org/10.1016/j.jfoodeng.2015.11.009.

Wang, H., Peng, J., Xie, C., Bao, Y., He, Y. (2015): Fruit quality evaluation using spectroscopy technology: A review. *Sensors* 15(5), 11889–11927. https://doi.org/10.3390/s150511889.

Wedding, B.B., Wright, C., Grauf, S., White, R.D., Tilse, B., Gadek, P. (2013): Effects of seasonal variability on FT-NIR prediction of dry matter content for whole Hass avocado fruit. *Postharvest Biol. Tec.* 75, 9–16. https://doi.org/10.1016/j.postharvbio.2012.04.016.

Wold, S., Esbensen, K., Geladi, P. (1987): Principal Component Analysis. *Chemometr. Intell. Lab.* 2, 37–52.

Wold, S., Sjöström, M., Eriksson, L. (2001): PLS-Regression: a basic tool of chemometrics. *Chemometr. Intell. Lab.* 58, 109–130.

Worley, B., Powers, R. (2016): PCA as a practical indicator of OPLS-DA model reliability. *Curr. Metabolomics* 4(2), 97–103. https://doi.org/10.2174/2213235X04666160613122429.

Wu, T., Zhong, N., Yang, L. (2018): Identification of Adulterated and Non-adulterated Norwegian Salmon Using FTIR and an Improved PLS-DA Method. *Food Anal. Methods* 11(5), 1501–1509. https://doi.org/10.1007/s12161-017-1135-4.

Xie, L., Ye, X., Liu, D., Ying, Y. (2009): Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS. *Food Chem.* 114(3), 1135–1140. https://doi.org/10.1016/j.foodchem.2008.10.076.

Zhu, Q., Xing, Y., Lu, R., Huang, M., Ng, P.K.W. (2017): Visible/shortwave near infrared spectroscopy and hyperspectral scattering for determining bulk density and particle size of wheat flour. *J. Near Infrared Spec.* 25(2), 116–126. https://doi.org/10.1177/0967033517704081.