Food Analytical Methods

Determination of the sugar content in commercial plant milks by near infrared spectroscopy and Luff-Schoorl total glucose titration.

--Manuscript Draft--

Manuscript Number:	FANM-D-16-00551R1
Full Title:	Determination of the sugar content in commercial plant milks by near infrared spectroscopy and Luff-Schoorl total glucose titration.
Article Type:	Original Research
Keywords:	Plant milk; Glucose determination; Luff-Schoorl method; NIR spectroscopy; Variable selection; Randomization test
Corresponding Author:	Adele Papetti, Ph.D. University of Pavia Pavia, ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Pavia
Corresponding Author's Secondary Institution:	
First Author:	Giorgio Marrubini, Ph.D.
First Author Secondary Information:	
Order of Authors:	Giorgio Marrubini, Ph.D.
	Adele Papetti, Ph.D.
	Emiliano Genorini
	Alessandro Ulrici, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	No funding was provided for this study (NOne) — Not applicable
Abstract:	Thirty-nine samples of plant milks (rice, soy, oat, barley, spelt, quinoa, almond, and a variety of wheat called kamut) were analyzed for their reducing sugars content by NIR spectroscopy, using the Luff-Schoorl official method as reference to build the calibration models. The amount of reducing sugars, expressed as grams of glucose/100 mL of beverage, ranged from 0.5 g/100 mL (soy) to 7.6 g/100 mL (rice). Both Partial Least Squares (PLS) and interval-Partial Least Squares regression (iPLS) were used to build multivariate calibration models, testing different spectra preprocessing methods. The performance in prediction of the best calibration model was evaluated on an external test set of nine randomly selected samples (RMSEP = 0.98 g/100 mL, R2PRED = 0.84), and its statistical significance was assessed using a randomization t-test based on Monte Carlo simulations. The results obtained suggest that NIR spectroscopy can be a valid alternative to the laborious reference titrimetric method for the determination of total glucose content in plant milks.

We would like to thank the reviewer, because he allowed us to improve our manuscript. We hope that the answers in the following, together with the changes made to the text and to Table 3 will cope with his comments.

In the tuning of the reference method, the step of hydrolysis of the polysaccharides described in section 2.2.1 is not sufficiently discussed with regard to the detection of the oligosaccharides that may come from the plant matrix. As the investigated matrices are rich in starch and polysaccharides that can pass to the finished product in different ways and amounts depending on the technology adopted, an investigation of the effects of hydrolysis of these fractions would be desirable.

In the processes of production of plant milks, filtration and heating are used to purify the aqueous solution from undesired oligosaccharides and carbohydrates that may have unpleasant effects for the consumers (Jiang et al. Foods 2013,2:198–212). The finished product thus should be free from indigestible oligosaccharides and carbohydrates. If residues of these compounds are present, they should not be major components of the final product.

During the reworking of the Luff-Schoorl method for the purpose presented in this study we actually investigated thoroughly the possibility that unpredictable reducing sugars, in addition to glucose or fructose may be originated during the hydrolysis step. This observation is important because if glucose and fructose were obtained during the hydrolysis step along with other hypothetical reducing sugars (e.g. because residues of fibers or starch are demolished by heating with 6M HCl) this could lead to a wrong estimation of the true total glucose content in the milk samples. We also evaluated the possibility that the hydrolysis step could cause losses of glucose/fructose, thus leading to an underestimation of the true total glucose content in the samples. However, all our results evidenced that neither of these two potential error sources were appreciable under the conditions that we applied. The accuracy study presented in Table 2 documents this conclusion. We chose finally not to expand this topic in the present manuscript, and we did not present the data regarding fructose results because we recognized that these potential sources of error were indeed very unlikely as reported also in the two new references added in the revised manuscript (Li et al. 2016; Jiang et al. 2013).

Nonetheless, we agree with the reviewer comment and thus we added in the manuscript one paragraph discussing this issue at the beginning of Section 3.5 and two more references of recent studies on the oligosaccharides in soy milk.

There is some confusion in the use of the terms "carbohydrate" and "sugars": in the text they are often used interchangeably (in particular on pagg 4,7,13). This also happens in the discussion on the values printed on the label of the analysed samples. Usually, in fact, the label of plant milks reports the amount of carbohydrates and the specification of the share represented by sugars. Carbohydrates are usually estimated for difference (dry matterproteinslipidsash). Moreover not always these two quantities coincide and often show significant differences. It would be appropriate to put these two values (carbohydrates and sugars) on Table 3 for the analyzed plant milk samples.

We revised thoroughly the text considering this criticism and found several inaccuracies which have now been amended. In addition, we revised table 3 and added the data on the carbohydrate content of the plant milk studied.

In the introduction, we added one sentence and one reference regarding the estimation of the total carbohydrate content measurement, in order to provide the interesting information on the methods used for assaying the vegetable milk during and after production.

In the light of these considerations section 3.4 should be removed or completely rewritten.

The section was renumbered (now it is section 3.5) and revised in light of the considerations made above.

In section 2.4 on page 9 lines 815 the Authors should be more specific about which routines were used.

A more detailed description of the used routines is now given at the end of section 2.4.

The independent validation set contain three samples made of different batches of the same products used in the training Set. It would be better to have in this set only samples of different origin or at least to discuss in more detail the nature of the different batches.

We would like to thank the reviewer for this comment, that helped us to amend an inaccurate definition in the legend to Table 3. Actually, all the samples included in the independent validation set are made of products that are different from those used in the training set. In fact, the three test set samples mentioned by the reviewer (CO2, SC1, OA2) come from the same producer of three samples included in the training set (CO1, SC2, OA1). However, they are not “subsequent batches” as it was erroneously reported in the original manuscript, but they are different products, since they differ for the additives claimed in the product label and, in two cases out of three, also for the amounts of carbohydrates and of sugars reported on the label. For this reason, we considered reasonable to consider CO2, SC1 and OA2 as independent samples, and to include them in the external validation set. This aspect has been clarified in the revised version of the manuscript (Sections 2.2. and 2.4, legend to Table 3).

The figures 1,2,3, and 4 appear disturbed, probably because exported by matlab as .jpg, would be more appropriate to export them as .tif

Actually, the figures have been submitted as .tif: the low quality of the figures within the pdf manuscript submitted for revision is due to the (likely .jpg) compression made while building the pdf itself. The reviewer can download the original, high quality images by clicking on the heading in the top right of the page (“Click here to download Figure XXX”).

Determination of the sugar content in commercial plant *milks* by near infrared spectroscopy and Luff-Schoorl total glucose titration.

Giorgio Marrubini[a], Adele Papetti[a*], Emiliano Genorini[b], Alessandro Ulrici[c]

[a] Department of Drug Sciences, University of Pavia, Viale Taramelli 12, 27100 Pavia, Italy.

[b] Bruker Italia S.r.l., Viale V. Lancetti 43, 20158 Milano, Italy.

[c] Department of Life Sciences, University of Modena and Reggio Emilia, padiglione Besta, via Amendola 2, 42122 Reggio Emilia, Italy.

*Corresponding author: Adele Papetti, +390382987863 +390382422975, adele.papetti@unipv.it

Abstract

Thirty-nine samples of plant *milks* (rice, soy, oat, barley, spelt, quinoa, almond, and a variety of wheat called *kamut*) were analyzed for their reducing sugars content by NIR spectroscopy, using the Luff-Schoorl official method as reference to build the calibration models. The amount of reducing sugars, expressed as grams of glucose/100 mL of beverage, ranged from 0.5 g/100 mL (soy) to 7.6 g/100 mL (rice). Both Partial Least Squares (PLS) and interval-Partial Least Squares regression (iPLS) were used to build multivariate calibration models, testing different spectra preprocessing methods. The performance in prediction of the best calibration model was evaluated on an external test set of nine randomly selected samples (RMSEP = 0.98 g/100 mL, R^2_{PRED} = 0.84), and its statistical significance was assessed using a randomization *t*-test based on Monte Carlo simulations. The results obtained suggest that NIR spectroscopy can be a valid alternative to the laborious reference titrimetric method for the determination of total glucose content in plant *milks*.

Keywords:

Plant *milk*; Glucose determination; Luff-Schoorl method; NIR spectroscopy; Variable selection; Randomization test

1. Introduction

In recent years, the food market in industrialized countries offered a number of products presented as healthier than the traditional ones consumed for many decades (Kearney 2010). This has become especially evident for milk which is often compared for its nutritional properties to non-dairy products derived from vegetables, often referred to as plant *milks*. Many products, such as cheese-like foods of vegetable origin (*e.g.* tofu) are also proposed as healthier, balanced, low-fat, and low-calories substitutes of dairy products. Plant *milks* and *dairy-like* products are free of lactose and animal proteins and are thus proposed as suitable foods for lactose-intolerant and allergic patients and for vegetarians (Bernat et al. 2014). On the Italian market, the more represented plant *milks* are those produced using almond, millet, rice, soy, oat, barley, and a variety of wheat called *kamut*.

Plant *milks* are obtained by extraction with hot water and filtration from the starting raw material which is processed depending on the nature of the crop (*e.g.* nut or cereal). Common pretreatments include one or more steps like washing, grinding, blanching, and also peeling for nuts. After removal of insoluble and non-extractable solids through filtration, the milky liquid is collected. The pH and taste are then adjusted by adding appropriate additives (including emulsifiers, stabilizers, sweeteners), and finally the product is packed under conditions that prevent microbial and mould proliferation (Bernat et al. 2014). The final product is tested and analyzed before marketing. The quality in terms of total macro- and micronutrients content is ensured by controls made both during the production process and on the final product. Whereas carbohydrates in these products are usually estimated by subtracting water, proteins, lipids and ash from the total mass of vegetable (Jiang et al. 2013), the sugars are mostly determined by titration and colorimetry.. These methods are quite labor-intensive, but are cheap and provide results considered reliable and satisfactory for the purpose of routine controls (Guo et al. 2014). Thus, the use of Near-Infrared (NIR) spectroscopy appears appealing, because it is inexpensive, fast, and with minimal requirements of personnel supervision.

NIR spectroscopy is a non-destructive analytical technique that gained popularity in recent years due to its suitability for assaying a number of chemical products, in particular in food production (Ferrari et al. 2011; van Maarschalkerweerd and Husted 2015; Varzakas 2015; Verardo et al. 2015). NIR is mainly used as a secondary method of measurement, validated by comparison with a primary reference method. Actually, NIR bench-top and portable instruments can analyze either clear or non-transparent liquid samples and even solid samples (Blanco and Villaroya 2002). The measurement technique is based on the absorption of electromagnetic radiation at wavelengths in the 780–2500 nm range. NIR spectra of foods comprise broad bands arising from overlapping absorptions corresponding mainly to overtones and combinations of fundamental vibrational modes involving X-H chemical bonds (X = C, N, O, S, …) and many other molecular moieties. The levels of the major constituents of foods, such as proteins, fats, carbohydrates, and sugars can be measured using light absorption, if the sample is transparent. For opaque samples, solids, and suspensions, NIR is used in modes based on diffuse transmittance and reflectance, at the cost of generating spectra more complex than absorption spectra (Foca et al. 2011). NIR spectroscopy is currently used for routine controls of food ingredients, both in base research and in industrial settings where it is applied for on-line and in-line monitoring of process intermediates and final products (Blanco et al. 2000; Pan et al. 2015; Ulrici et al. 2008; Wu et al. 2012). The major advantage of NIR is that usually no sample preparation is necessary. In most applications NIR allows the direct analysis of the sample, hence the mere measurement step is fast, usually taking less than 2 minutes per sample. The analyses can be carried out also by non-trained personnel, and several constituents can be assayed simultaneously in the same sample.

The aim of the present research was to study the feasibility of NIR quantitative analyses of total sugars content in plant *milk*s expressed as total glucose, using Partial Least Squares (PLS) regression for signal processing (Wold et al. 2001). Since NIR spectra are made of highly correlated variables containing redundant information spread out over different spectral regions, variable

selection can be used to increase the performance of the NIR-based calibration models, allowing to identify and to select the spectral regions in which the information of interest is located (Ferrari et al. 2011; Foca et al. 2009; Foca et al. 2013; Ulrici et al. 2008). For this reason, a variable selection method based on interval-PLS (iPLS) (Nørgaard et al. 2000) was also considered, which led to significant improvements in the prediction results. The statistical significance of the best calibration model was assessed by randomization *t*-test based on Monte Carlo simulations (van der Voet 1994). The prediction error of the best calibration model based on NIR spectroscopy was compared with the experimental error of the Luff-Schoorl (LS) titrimetric method (AOAC 1995, method 942.15), which was selected as reference method for measuring the total glucose content in the plant *milk* samples.

2. Materials and methods

2.1. Standards and reagents

Copper (II) sulphate pentahydrate (purity >97%), citric acid monohydrate (purity > 99.5%), anhydrous sodium carbonate (purity > 99.5%), potassium hexacyanoferrate (II) trihydrate (99.5% purity), starch (purity >97%), potassium iodide (purity >99%), sulfuric acid (96%), acetic acid (96%), sodium hydroxide (purity > 97%) were purchased from Carlo Erba Reagenti (Milano, Italy). D(+)-glucose (purity > 99.8%), zinc sulphate (purity > 99.5%), and disodium hydrogenphosphate (purity > 99%) were purchased from Merck (Darmstadt, Germany). Hydrochloric acid (37%) was from Panreac Quimica (Barcelona, Spain), and ready-to-use sodium thiosulphate 0.1N in water was from Titolchimica s.p.a. (Rovigo, Italia). Deionized water HPLC-grade was produced in house using a Simpack® apparatus from Millipore (Billerica, MA, USA).

The Luff-Schoorl (LS) reactant was prepared once a week mixing 50 mL of 5M citric acid monohydrate aqueous solution with 350 mL of 4M anhydrous sodium carbonate aqueous solution. After elimination of the CO_2 produced, 500 mL of 0.2M copper (II) sulphate solution was added to

the obtained mixture and the solution was brought to the final volume of 1 L with deionized water. Subsequently, the LS reactant was left standing for one night and centrifuged before use.

The Carrez I clarification reactant was prepared dissolving 15 g of $K_4[Fe(CN)_6]\cdot 3H_2O$ in 100 mL of water, while the Carrez II reactant was made dissolving 30 g $ZnSO_4$ in 100 mL water.

The starch solution was prepared adding dropwise a dispersion of 1 g of starch in 10 mL of water to 9 mL boiling water.

2.1.1. Standard solutions of glucose and reference samples for Luff-Schoorl (LS) titration

Standard reference solutions of glucose in water were prepared at concentrations ranging from 0.5 to 10 g/L (corresponding to 0.05 to 1 g of glucose /100mL, Table 1).

For each concentration level, 5 mL of Carrez I clarification reactant were added to 20 mL of glucose solution. After vigorous shaking, the solution was left standing for 5 min and then 5 mL of Carrez II reactant were added. After further shaking, the solution was left standing for 1 h, and then 5 mL of a 10% aqueous disodium hydrogen phosphate solution (w/v) were added. Finally, the mixture was brought to 100 mL final volume and centrifuged at 800 rpm for 10 min. The supernatant was titrated.

Three trial samples were selected based on their glucose content to have a low, an intermediate and a high level of glucose reference samples. Samples #7, 20, and 22 were obtained from different producers and were assayed as such and after spiking them with glucose at two different levels of concentration corresponding to +50% and +75% of the nominal glucose level expected in the sample. Sample #7 was spiked with 0.30 and 0.45 g glucose/100ml. Sample #22 was spiked with 2.30 and 3.45 g glucose/100ml. Sample # 20 was spiked with 4.00 and 6.00 g glucose/100ml. Aqueous solutions, unspiked and spiked milk samples were assayed three times each by the LS method as reported in Table 2.

2.2. Samples

Thirty-nine different samples of plant *milk* obtained from rice, soy, oat, barley, and *kamut* were purchased in local specialized shops and supermarkets. Samples of the same product, i.e., made by the same manufacturer and labelled as containing the same plant *milk* and additives, were selected from different production batches, in order to ensure that all samples were independent. Before analysis, all samples were stored for less than one month under the conditions recommended by the manufacturers in the dark at room temperature, and were examined long before the expiry date. On the day of the analysis, each sample was divided in two aliquots and the total content of reducing sugars was assayed by the LS titrimetric method (AOAC 1995, method 942.15; Egan et al. 1981) and by NIR.

2.2.1 Plant milk sample preparation for the Luff-Schoorl titration

The volume of each plant *milk* was selected preliminarily on the basis of the amount of sugars declared on the product package. Each sample was assayed in triplicate and the results were initially evaluated in comparison with the labelled content of glucose.

The preparation of the samples was carried out as follows. The sample was clarified adding the Carrez I reactant before the sugars titration. A 50 mL aliquot of the clear supernatant was transferred into an Erlenmeyer flask, treated with 5 mL HCl 6M, and heated at 65°C for 15 min under reflux to promote the complete hydrolysis of the oligosaccharides into the corresponding monosaccharides. At the end of the reaction, the mixture was cooled under tap water and neutralized to a pH value around 7 adding 6 mL 5M sodium hydroxide. Three to five drops of glacial acetic acid were then added to the mixture under agitation until a clear solution was obtained.

2.2.2 Luff-Schoorl titration

The determination of the total reducing sugars content, expressed as glucose equivalents, has been carried out following the procedure reported in AOAC method 942.15, slightly modified in order to adapt it to the measurement of sugars in plant *milks* and hence to obtain optimal accuracy and precision. Briefly, 25 mL of LS reactant were added to 25 mL of sample (or water for blank-control) in a 250 mL Erlenmeyer flask and brought to boiling temperature under reflux for 10 min. After cooling under tap water for 5 min, 10 mL 1.8M KI and 25 mL 2M H_2SO_4 were added. The excess of copper (II) in the solution was determined by iodometric titration with 0.1N $Na_2S_2O_3$ to the end point with starch. During the titration, when a pale yellow coloration of the solution appeared, one mL of starch solution was added in order to obtain a blue solution. Additional $Na_2S_2O_3$ was added drop-wise to complete the titration and reach the final change of color from blue to milky-white (X mL). The concentration of glucose-equivalents in the sample was corrected by difference between the volume of $Na_2S_2O_3$ used for the titration of the blank-control (Y mL) and the volume of $Na_2S_2O_3$ used for the titration of the sample solution. The (Y-X mL) volume represents the quantity of Copper(II) reduced by the sugars in the plant *milk* aliquot.

2.3. Near-infrared spectroscopic apparatus and operating conditions

NIR spectra were recorded using a Bruker Multi-Purpose Analyzer (MPA) near-infrared spectrophotometer (Bruker, Billerica, MA) equipped with Liquid Sampling Module (LSM) that allows measurement in transmission using a 1 mm path length quartz flow through cell. Each spectrum was obtained as the average of 24 scans in the region between 4000 and 12500 cm^{-1}, with a resolution of 7.7 cm^{-1}. The instrument was controlled by a notebook equipped with OPUS software version 7.2.

2.4 Spectral measurements and model development

NIR spectra were imported into Matlab ver. 7.11 (The MathWorks Inc., Natick, MA, USA), and the whole data were merged into a unique dataset with size {39 × 1102}, composed by the absorbance values of the NIR spectra corresponding to the 1102 wavenumber values, for each one of the 39 plant *milk* samples.

Then, for the subsequent calculation of the calibration models, the dataset was randomly split into a training set with size {30 × 1102} and a test set with size {9 × 1102}, paying attention to include in the test set products different from those included in the training set, i.e., products from different manufacturers and/or with different composition (Table 3). PLS and iPLS calibration models were elaborated and cross-validated by means of PLS-Toolbox ver. 7.8.2 (Eigenvector Research Inc., Manson,WA, USA), using the *pls.m*, *ipls.m* and *crossval.m* functions available in the PLS-Toolbox. Furthermore, some Matlab functions written *ad hoc* by some of us were used in order to fully automate both the selection of the optimal spectral regions based on the results of iPLS, as described in Section 2.4.2, and the randomization test described in Section 2.4.3. In particular, the function *massiveiPLS.m* allowed us to automatically calculate all the iPLS models corresponding to different preprocessing methods and values of the interval size; then, the function *massivePLS_selvar.m* was used to further elaborate the results of *massiveiPLS.m* as described in the second step of the variable selection described in Section 2.4.2. In order to fully automate the randomization test described in Section 2.4.3, we developed the function *massiveiPLSrandtest.m*.

2.4.1. PLS calibration models

PLS regression (Naes et al. 2002) was used to build multivariate calibration models aimed at evaluating the possibility to predict the sugar content of the plant *milk* samples using the whole NIR spectral range. To this aim, ten different spectra pre-processing methods (Zeaiter et al. 2005) were tested, namely mean centering (MNCN), standard normal variate followed by mean centering (SNV+MNCN), linear detrend followed by mean centering (DETREND+MNCN), multiplicative

scatter correction followed by mean centering (MSC+MNCN), Savitzky-Golay first-order derivative followed by mean centering (D1+MNCN), autoscaling (AUTO), standard normal variate followed by autoscaling (SNV+ AUTO), linear detrend followed by autoscaling (DETREND+ AUTO), multiplicative scatter correction followed by autoscaling (MSC+ AUTO), and Savitzky-Golay first-order derivative followed by autoscaling (D1+ AUTO).

For each preprocessing method, the corresponding PLS model was calculated. The number of latent variables (LVs) was chosen on the basis of the minimum value of the Root Mean Square Error of Cross-Validation (RMSECV), keeping the maximum possible number of LVs equal to 5. In particular, venetian blinds cross-validation method was used, considering 6 deletion groups. The performance of the PLS calibration models was expressed both in terms of Root Mean Square Error in calibration, cross-validation and prediction of the external test set (RMSEC, RMSECV and RMSEP, respectively), and using the squared value of the correlation coefficient in calibration, cross-validation and prediction of the external test set (R^2_{CAL}, R^2_{CV} and R^2_{PRED}, respectively) (Pigani et al. 2011).

2.4.2. Variable selection

In order to improve performance and robustness of the PLS calibration models (Xiaobo et al. 2010; Zeaiter et al. 2005) by keeping only those NIR spectral regions that are actually relevant for the prediction of the sugar content, a two-steps variable selection procedure was used (Foca et al. 2016).

In the first step, the interval-PLS (iPLS) algorithm was used, which essentially consists in dividing the whole spectral range in a user-defined number of intervals of equal width, and in selecting the intervals most useful for calibration by an iterative procedure (Nørgaard et al. 2000). In particular, iPLS was applied in the forward mode, where the intervals are added iteratively until the lowest value of RMSECV is reached (Xiaobo et al. 2010). Also in this case, venetian blinds cross-

validation with 6 deletion groups was used. In the present work, iPLS was run considering five different values of the interval size, i.e., 6, 11, 20, 50 and 110. For each interval size value, each one of the ten pre-processing methods described above was tested.

When using iPLS to perform variable selection, it must be recalled that the spectral regions selected by this algorithm may vary, depending both on the signal preprocessing method and on the specific value of the interval size. In fact, on the one hand, applying different preprocessing methods to the same spectral dataset can cause modifications of the position and of the relative importance of the different spectral features. On the other hand, for a given preprocessing method, the most important spectral features can be more or less efficiently selected on the basis of the interval size: for example, a useful feature corresponding to narrow absorption band can be better accounted for by a narrow interval size than by a large one (which could include also non pertinent information of the neighboring spectral regions), while another useful feature corresponding to a broader absorption band could be better accounted for by a large interval size. Therefore, since the position and width of the useful spectral features are not known in advance, it is advisable to consider different interval size values, and then to look at those regions that are the most frequently selected ones.

For these reasons, in the second step, for each preprocessing method the spectral variables were grouped on the basis of their frequency of selection in the models, obtained considering the different interval sizes, and PLS models were then calculated by adding iteratively the selected variables, in descending order according to their frequency of selection. For example, considering the five iPLS models calculated on the mean centered NIR spectra using the five different interval size values, at the beginning only the spectral regions (if any) that were selected in all the five iPLS models were considered for the calculation of a PLS regression model; then, all the spectral regions selected at least four times were included in the calculation, and so on, up to including all the spectral regions that were selected at least once. Among all these PLS models, the one leading to the minimum RMSECV value was chosen. This procedure was repeated for each spectra

preprocessing method, and the best calibration model was finally defined as the one leading to the overall minimum RMSECV value.

2.4.3. Randomization test

Since both the limited number of available samples and the feature selection procedure could lead to the risk of chance correlations, a randomization (permutation) *t*-test based on Monte Carlo simulations (van der Voet 1994; http://wiki.eigenvector.com/index.php?title=Tools:_Permutation_Test. Accessed 14 October 2016) was used to assess the statistical significance of the best calibration model. Randomization tests consist in repeatedly and randomly reordering the values of the *y* variable (i.e., of the sugar content of the plant *milk* samples), and in recalculating from the beginning the calibration model with the randomly shuffled *y* values after each reordering. In other words, for each run (i.e., for each reordering of the *y* variable), each NIR spectrum was assigned to a "wrong" *y* value (i.e., to a *y* value of another sample), and then the whole variable selection step was repeated from the beginning, in the same manner as it was done for the correct model (i.e., for the model calculated on the correctly ordered *y* vales). The same procedure was then repeated for a given number of runs, each time randomly reordering the *y* variable, and for each run the values of RMSEC, RMSECV and RMSEP were stored. In particular, in the present work 100 subsequent runs were calculated, considering the same spectra preprocessing method that led to the best calibration model. Then, a *t*-test was used to compare the RMSEC, RMSECV and RMSEP values of the best calibration model with the corresponding 100 values calculated with the randomization test.

3. Results and discussion

3.1 Validation of the Luff-Schoorl method and comparison of titration data with product label data

The titration method was initially validated using two reducing sugars, namely glucose and fructose. The slope ± standard error for the calibration curves of glucose and fructose were

identical, and therefore all data have been reported as glucose content in g/100 mL of vegetable *milk*. Linearity, specificity, and accuracy were satisfactory for the purposes of the determinations and are reported in Table 1.

The following equation was used for measuring the accuracy and overall bias of the determinations carried out on the unspiked and spiked samples at known levels of glucose (Table 2):

$$A = 100 - \left[100 \cdot \frac{(V_e - V_f)}{V_e} \right] \qquad (1)$$

V_e is the expected value of glucose in the sample, and V_f is the value found as laboratory measurement, namely in this work the LS total glucose titration value.

Since to the authors best knowledge no reference materials are available for the accuracy evaluation of the glucose determination in plant milks, accuracy was assessed by spiking three independent samples of different vegetable milks selected from three different producers. The samples were selected according with the information provided on the label of the product assuming that sample #7 would represent a reference for "low glucose level", sample #22 for "intermediate glucose level", and sample #20 for a "high glucose level".

Precision within-session (repeatability) and inter-session (intermediate precision) data were evaluated by one-way ANOVA at the levels of 0.050, 0.100, 0.200, and 1.0 g/100mL on aqueous samples, and on plant milk samples at low, intermediate and high level of glucose concentration. The titration method resulted precise at the level of probability of 95% ($\alpha = 0.05$).

The obtained amount of reducing sugars, expressed as grams of glucose per 100 mL of plant *milk*, ranged from 0.5 (in soy) to 7.6 g/100 mL (in rice *milk*) with mean value of 3.9 g/100 mL and median value of 3.6 g/100mL. The values of the standard deviations calculated on the replicate measurements of each sample (n=3) ranged from 0.007 g/100 mL to 0.8 g/100 mL.

The results of the glucose determinations in plant *milks* by the LS method presented in Table 3 show that in several cases the sugars content reported on the product label is different from the amount of reducing sugars (expressed as grams of total glucose/100mL product) measured in the sample. The anomalous error % values reported in Table 3 do not correspond to extreme (high or low) values of concentration of sugars in the plant *milk*. No dependence of the error % on the matrix was observed and no significant correlation was found between the error % values and either the sugars or the carbohydrates concentration values reported on the labels, and between the error % and the total glucose values registered for the products. As a matter of fact, the LS method is complicated and its accuracy can suffer from a number of causes, but the data collected in Tables 1-3 were thoroughly tested and the method was demonstrated to be precise and accurate on both aqueous calibration curves and trial samples. In addition, regarding the concentration values reported in the product label, the determination of carbohydrate and sugars content is not carried out systematically by the producers, but rather assayed on samples taken from randomly selected batches. Based on these reasons, it is likely that the observed differences originate from inconsistencies in the data reported on the products labels, rather than from poor accuracy of the LS method.

3.2 PLS calibration models

The LS titration is a cheap method but it has several major disadvantages. It requires burettes, glassware, and reactants prepared on purpose. The entire procedure of measurement is laborious and time-consuming requiring hours of work per sample. Indeed, a certain degree of handiness and experience is necessary to obtain consistent results. On the contrary, NIR requires a relevant investment of money for the acquisition of the instrumentation and obliges to build robust and reliable calibration curves. However, NIR provides accurate and precise determinations in very short times allowing the quantification of glucose in all the plant *milk* products. For this reason, in

order to test the feasibility of NIR quantitative analysis, the glucose concentration values measured using the LS titration method were used as dependent variable to construct PLS calibration models based on the corresponding NIR spectra.

Table 4 reports the results obtained by applying the different preprocessing methods described in Section 2.4.1. to the whole NIR spectrum. The reported results show that, in general, the performance of the calibration models calculated considering the whole spectral range is not very satisfactory, even though some pre-processing methods lead to acceptable results both in cross-validation and in prediction of the external test set samples. In particular, the Savitzky-Golay first-order derivative method followed by autoscaling (D1+AUTO) is the signal pre-processing technique that leads to the best performances in cross-validation (RMSECV = 0.63 g/100 mL, R^2_{CV} = 0.86), but the prediction of the glucose values of the test set is less satisfactory (RMSEP = 1.36 g/100 mL, R^2_{PRED} = 0.69). The reasons for the low performances of these calibration models are likely due to the fact that – as often happens – a large part of the NIR spectral range does not contain information related to the glucose content, but rather to other variability sources due to the composition of the plant *milk* samples considered in this study, which have been obtained from completely different vegetable matrices (Nørgaard et al. 2000). Therefore the inclusion in the calibration model of spectral information that is non pertinent to our specific aim (such as, e.g., spectral variability due to scattering and to the contribution of other components such as fats or proteins) leads to a general decrease of the prediction performances. Hence, this implies the need to separate the information useful to predict the glucose values present within the NIR spectrum from noise and from other variability sources, by means of proper variable selection methods.

3.3 Variable selection

The subsequent step consisted in the use of an automated variable selection method, aimed to identify only those spectral regions containing information specifically related to sugar content of

the plant milk samples. For each preprocessing method, iPLS was run five times, one for each considered interval size value (*see* Section 2.4.2.), and the convergence of the intervals selected in the five iPLS runs was evaluated through the construction of a histogram reporting the frequency of selection. As an example, Figure 1 reports the results obtained using the Savitzky-Golay first-order derivative method followed by mean centering (D1+MNCN) as spectra preprocessing method. In particular, Figure 1a reports the different spectral regions that were selected from each iPLS run. Figure 1b summarizes these selection results, highlighting that three out of the five iPLS runs converge toward the selection of a narrow spectral region located between 4406 and 4367 cm^{-1}.

In order to evaluate the convergence in the variable selection results obtained considering the different spectra preprocessing methods, the values of the frequencies obtained using the different preprocessing methods were plotted under the form of a stacked bar graph, that is reported in Figure 2 together with the average NIR spectrum for comparison purposes. This figure confirms that, in general, independently of the specific preprocessing method, iPLS tends to converge to the same spectral regions: the selected regions. The selected regions are those between 9200 and 5700 cm^{-1} and between 4800 and 4300 cm^{-1}. The general convergence in the selected regions confirms the fact that only some parts of the NIR spectrum contain useful information relevant to the evaluation of sugar content of the samples.

The next step consisted in performing, for each preprocessing method, the "selection of the selection", i.e., as it was already described in Section 2.4.2., in defining the optimal subset of variables among the previously selected ones. For example, in the case of D1+MNCN, starting from the values of the frequency of selection reported in Figure 1b, three PLS models were calculated: the first one including only the spectral variables with frequency of selection equal to three, the second one including the spectral variables with frequency of selection greater than or equal to two, and the third one including all the spectral variables that were selected at least once. Among these three models, the one showing the lowest RMSECV value was finally kept.

The overall results obtained by applying this procedure are reported in Table 5. In general, compared to the results reported in Table 4, the use of variable selection led to a significant enhancement in the model performances, both in cross-validation and in prediction, and to a drastic reduction of the number of spectral variables with respect to the 1102 of the original NIR spectrum. In particular, the best performance was obtained using D1+MNCN as pretreatment, that led to a PLS model using only 6 spectral variables in the region between 4406 and 4367 cm^{-1}. By a chemical point of view, the selection of this spectral region is perfectly consistent with the determination of sugar content, since it includes the absorption bands related to O-H stretch/C-O stretch combination and to C-H stretch/CH_2 deformation (Shenk et al. 2008).

Figure 3 reports the corresponding plot of the predicted *vs* experimental values of glucose content. It shows that, notwithstanding the tendency to underestimate the values for the samples with higher glucose content, in general the performance in prediction of the external test set samples is satisfactory, and comparable with the calibration results obtained for the training set samples. It must be highlighted that the average error in prediction of the test set samples, expressed by the RMSEP value obtained for this model (0.98), is only slightly higher than the range of standard deviations determined on replicate measurements using the reference LS method (*see* Section 3.1). Since in general the average error in prediction of a calibration model cannot be lower than the experimental error associated to the reference parameter, these results suggest that NIR can be used for reducing sugar content determination.

3.4. Randomization test

Finally, the statistical significance of this calibration model was checked by means of randomization test (Section 2.4.3.), comparing by a *t*-test the RMSEC, RMSECV and RMSEP values with the corresponding values obtained by repeating 100 times the whole feature selection procedure using randomly shuffled values of *y* (i.e., of glucose content). The results of the

randomization test are reported in Figure 4, where the RMSE values of both the correct models (full circles) and the 100 randomization test runs (empty circles) are reported as a function of the squared correlation coefficient of the randomly shuffled *y* with the original (correct) *y*. This figure clearly shows that the RMSEC, RMSECV and RMSEP values of the best calibration model are significantly lower than the corresponding values calculated on the randomly shuffled y values, as it is also demonstrated by the results of the one-tailed t-test reported in Table 6. These results confirm the statistical significance of the best calibration model. More in general, our results - although collected from a limited number of samples - suggest that NIR spectroscopy can be proposed in the future as a complement to the reference titrimetric method for the rapid screening and determination of the total glucose content in plant *milk* products.

3.5. Comparison of Label values and NIR values with Luff-Schoorl values

The label of plant milks reports the amount of carbohydrates and the specification of the share represented by sugars. The major carbohydrates occurring in soy milk for example, are low-molecular weight indigestible oligosaccharides such as stachyose, saccharose, and raffinose (Li et al. 2016; Jiang et al. 2013). By heating the raw product these oligosaccharides are generally cleaved into their non-reducing (namely D-galactose) and reducing monosaccharides constituents (D-glucose and D-fructose. Jiang et al. 2013). In addition, other high-molecular weight polysaccharides that may come from the plant matrix (such as e.g. traces of starch) can pass into the finished product in different ways and amounts depending on the technology adopted. For these reasons, the quantities of carbohydrates and sugars labelled on the vegetable milk products only occasionally may coincide, and most often show instead significant differences (Table 3).

The comparison between the glucose content labelled on each product, the NIR results and the LS determinations (Table 3), shows that the relative errors% are much larger between the labelled data

and the LS results than for the NIR data. This observation is evidenced by the figures reported in Table 7, which are a summary of the descriptive statistics of the data reported in Table 3.

To make a fair comparison between label and NIR data, however, it is worth emphasizing that one should consider mainly the median values of the test set data, since the median is a robust estimator for skewed/non normally distributed data as in this case. As for the training set samples, the agreement between NIR and LS data is expected to be good, since the regression model was built using the LS data themselves.

In addition it must be underlined that, in line of principle, NIR data can be collected in real-time during production. Conversely, the label values are necessarily determined from time to time by off-line titrimetric or instrumental methods (e.g. HPLC and refractive index detector or other).

Finally, it is worth emphasizing that this work is intended as a feasibility study, and for this reason it was designed to assay a wide array of different products. For the development of a calibration model to be used for real-time quality control, the measurements should be focused on a specific product and data should be collected assaying a larger number of samples. It can be reasonably assumed that the selection of a single matrix and the accumulation and evaluation of a wider number of samples should significantly reduce the prediction error.

4. Conclusions

In recent years, food controls have become increasingly more important for compositional compliance to quality standards, for public safety to protect the consumers from frauds and unsafe products, and for regulatory purposes. The NIR technology available today, providing bench-top, portable, and even microNIR instrumentation is an important tool able to produce timely results during production or before marketing, improving thus the quality and safety of food products marketed. In this context, the present work showed that, through the proper use of variable selection

techniques coupled to multivariate calibration methods, it is possible to employ NIR spectroscopy to obtain reliable estimates of the total glucose content in plant *milk* products. The promising results obtained in this study could be further improved, both by increasing significantly the number of samples used for the construction of the calibration models, and through the development of local models, i.e., of models created for specific types of plant *milk*.

Acknowledgements

The authors would like to thank Bruker Italia s.r.l. for the technical support in NIR glucose measurements on all plant *milks*.

Compliance with Ethical Standards

Funding: No funding was used for this study.

Conflict of Interest: Giorgio Marrubini declares that he has no conflict of interest. Adele Papetti declares that she has no conflict of interest. Emiliano Genorini declares that he has no conflict of interest.

Alessandro Ulrici declares that he has no conflict of interest.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: Not applicable.

References

AOAC (1995) Official methods of analysis (16th ed.). Association of the Official Analytical Chemists, Arlington, VA, USA

Bernat N, Chàfer M, Chiralt A, Gonzàlez-Martìnez C (2014) Vegetable milks and their fermented derivative products. Int J Food Studies 3:93-124

Blanco M, Coello J, Iturriaga H, Maspoch S, González Bañó R (2000) On-line monitoring of starch enzymatic hydrolysis by near infrared spectroscopy. Analyst 125:749-752

Blanco M, Villaroya I (2002) NIR spectroscopy: a rapid-response analytical tool. Trends Anal Chem 21:240-250

Egan H, Kirk R, Sawyer R (1981) The Luff Schoorl method. Sugars and preserves. In: Pearson`s chemical analysis of foods, 8th edn. Harlow, UK: Longman Scientific and Technical, pp 152-153.

Ferrari E, Foca G, Vignali M, Tassi L, Ulrici A (2011) Adulteration of the anthocyanin content of red wines: Perspectives for authentication by Fourier Transform-Near InfraRed and 1H NMR spectroscopies. Anal Chim Acta 701:139–151

Foca G, Cocchi M, Li Vigni M, Caramanico R, Corbellini M, Ulrici A (2009) Different feature selection strategies in the wavelet domain applied to NIR-based quality classification models of bread wheat flours. Chemometr Intell Lab 99:91-100

Foca G, Ferrari C, Sinelli N, Mariotti M, Lucisano M, Caramanico R, Ulrici A (2011) Minimisation of instrumental noise in the acquisition of FT-NIR spectra of bread wheat using experimental design and signal processing techniques. Anal Bioanal Chem 99:1965-1973

Foca G, Salvo D, Cino A, Ferrari C, Lo Fiego DP, Minelli G, Ulrici A (2013) Classification of pig fat samples from different subcutaneous layers by means of fast and nondestructive analytical techniques. Food Res Int 52:185-197

Foca G, Ferrari C, Ulrici A, Ielo MC, Minelli G, Lo Fiego DP (2016) Iodine Value and Fatty Acids Determination on Pig Fat Samples by FT-NIR Spectroscopy: Benefits of Variable Selection in the

Perspective of Industrial Applications. Food Anal Meth, in press, available online, DOI 10.1007/s12161-016-0478-6

Guo Q, Cui SW, Kang J (2014) Classical methods for food carbohydrate analysis. In Moreno FJ and Sanz ML (Eds.) Food oligosaccharides: production, analysis and bioactivity. Wiley, Chichester, UK, pp 284-299

Kearney J. (2010) Food consumption trends and drivers. Philos T Roy Soc B 365:2793-2807

Jiang S., Cai W., Xu B. (2013) Food Quality Improvement of Soy Milk Made from Short-Time Germinated Soybeans. Foods 2:198-212

Li Y.-R., Yun T.-T., Liu S., Qi W.-T., Zhao L.-Q., Liu J.-R., Li A.-K. (2016) Analysis of water-soluble bioactive compounds in commonly consumed soymilk in China. J Food Comp Anal 46:29–35

Naes T, Isaksson T, Fearn T, Davies T (2002) A user-friendly guide to Multivariate Calibration and Classification. NIR Publications, Chichester, UK

Nørgaard L, Saudland A, Wagner J, Nielsen JP, Munck L, Engelsen SB (2000) Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. Appl Spectrosc 54:413–419

Pan L, Zhu Q, Lu R, Mitchell McGrath J (2015) Determination of sucrose content in sugar beet by portable visible and near-infrared spectroscopy. Food Chem 167:264–271

Pigani L, Culetu A, Ulrici A, Foca G, Vignali M, Seeber R (2011) Pedot modified electrodes in amperometric sensing for analysis of red wine samples. Food Chem 129:226-233

Shenk JS, Workman JJ, Westerhaus MO (2008) Application of NIR spectroscopy to agricultural products. In: Burns DA, Ciurczak EW (Eds.) Handbook of Near-Infrared Analysis, 3rd edn., Marcel Dekker, New York, p. 357.

Ulrici A, Li Vigni M, Durante C, Foca G, Belloni P, Brettagna B, De Marco T, Cocchi M (2008) At-line monitoring of the leavening process in industrial bread making by near infrared spectroscopy. J Near Infrared Spec 16:223-231

van der Voet H. (1994) Comparing the predictive accuracy of models using a simple randomization test. Chemometr Intell Lab 25:313-323

van Maarschalkerweerd M, Husted S (2015) Recent developments in fast spectroscopy for plant mineral analysis. Front Plant Sci 6:169

Varzakas T (2015) Quality and safety aspects of cereals (wheat) and their products. Crit Rev Food Sci Nutr, in press, available online, DOI: 10.1080/10408398.2013.866070

Verardo V, Cevoli C, Pasini F, Gómez-Caravaca AM, Marconi E, Fabbri A, Caboni MF (2015) Analysis of oligomer proanthocyanidins in different barley genotype using high-performance liquid chromatography-fluorescence detection-mass spectrometry and near-infrared methodologies. J Agr Food Chem 63:4130-4137

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

Wu D, Nie P, He Y, Bao Y (2012) Determination of Calcium content in powdered milk using near and mid-infrared spectroscopy with variable selection and chemometrics. Food Bioprocess Tech 5:1402-1410

Xiaobo Z, Jiewen Z, Povey MJW, Holmes M, Hanpin M (2010) Variables selection methods in near-infrared spectroscopy. Anal Chim Acta 667:14-32

Zeaiter M, Roger JM, Bellon-Maurel V (2005) Robustness of models developed by multivariate calibration. part II: The influence of pre-processing methods. Trends Anal Chem 24:437-445.

Figure Captions

Figure 1. Regions selected by iPLS on the autoscaled spectra using the different interval size values (a) and corresponding histogram of the selection frequency (b).

Figure 2. Comparison of the regions selected by iPLS using the different spectra preprocessing methods, together with the average NIR spectrum.

Figure 3. Predicted vs. experimental values of total glucose content, as estimated by the best model obtained using variable selection.

Figure 4. Results of the randomization test performed on the RMSE values of the best calibration model.

Figure 1

Click here to download Figure Figure1.tif

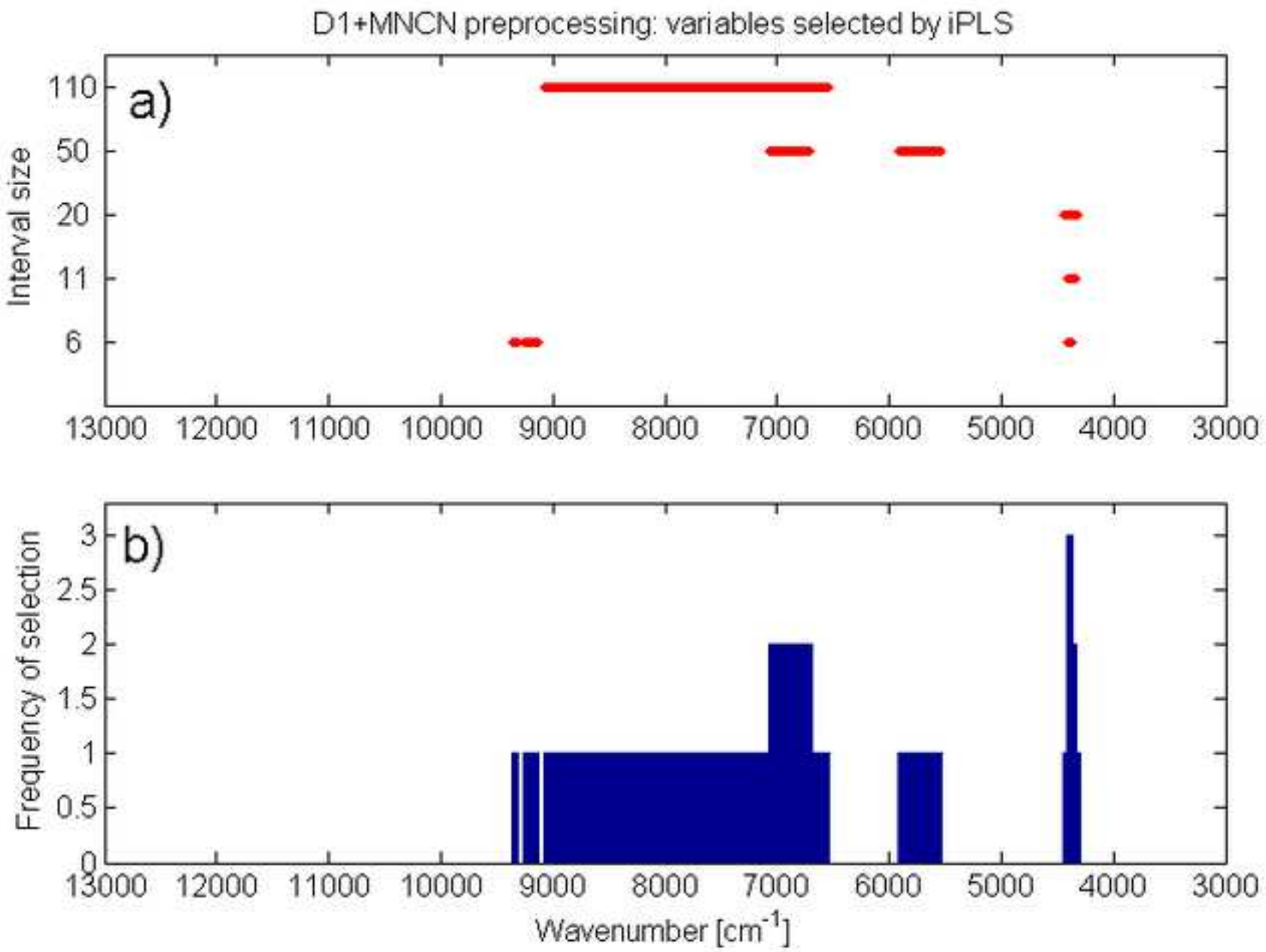

Figure 2

Click here to download Figure Figure2.tif

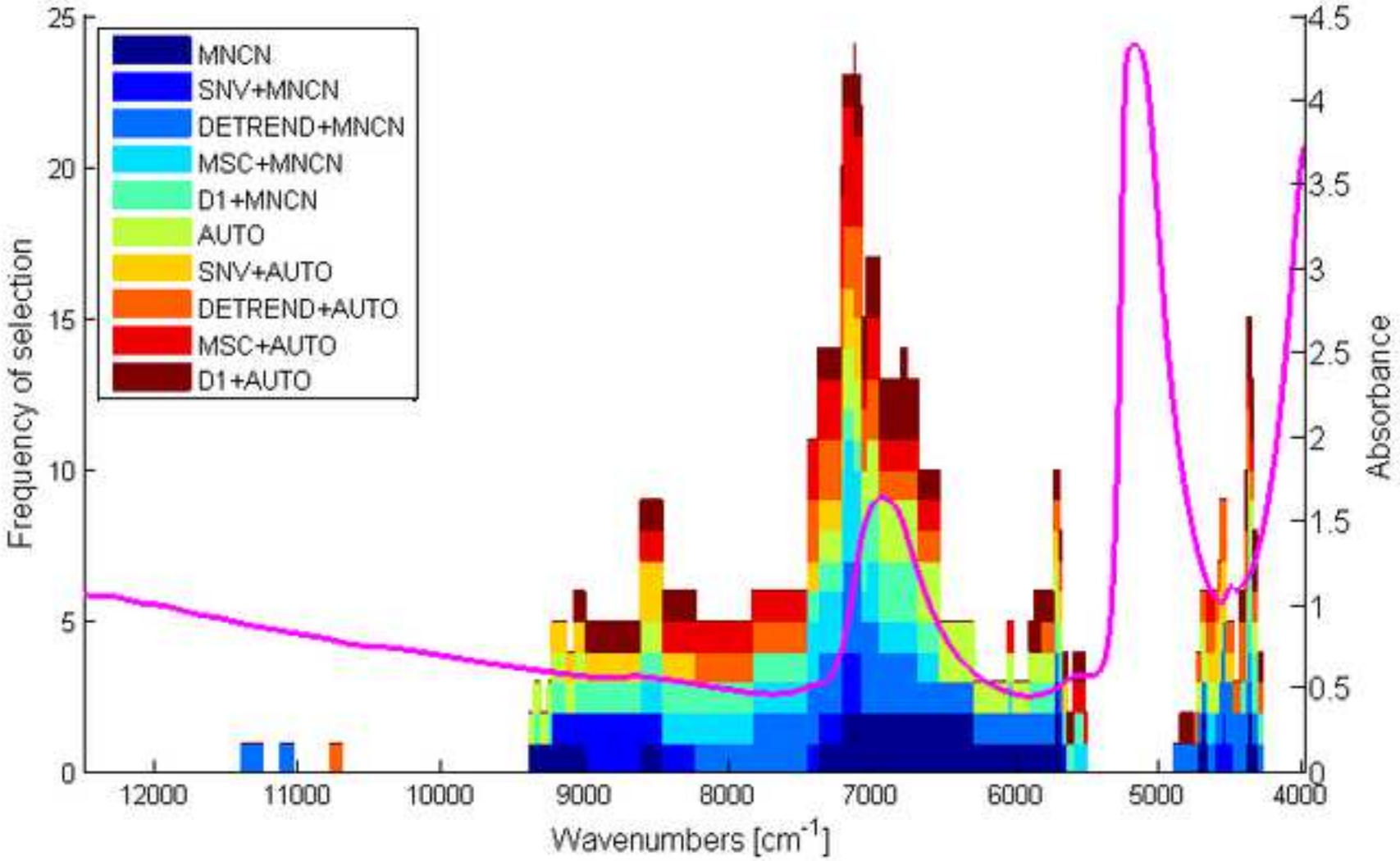

Figure 3

Click here to download Figure Figure3.tif

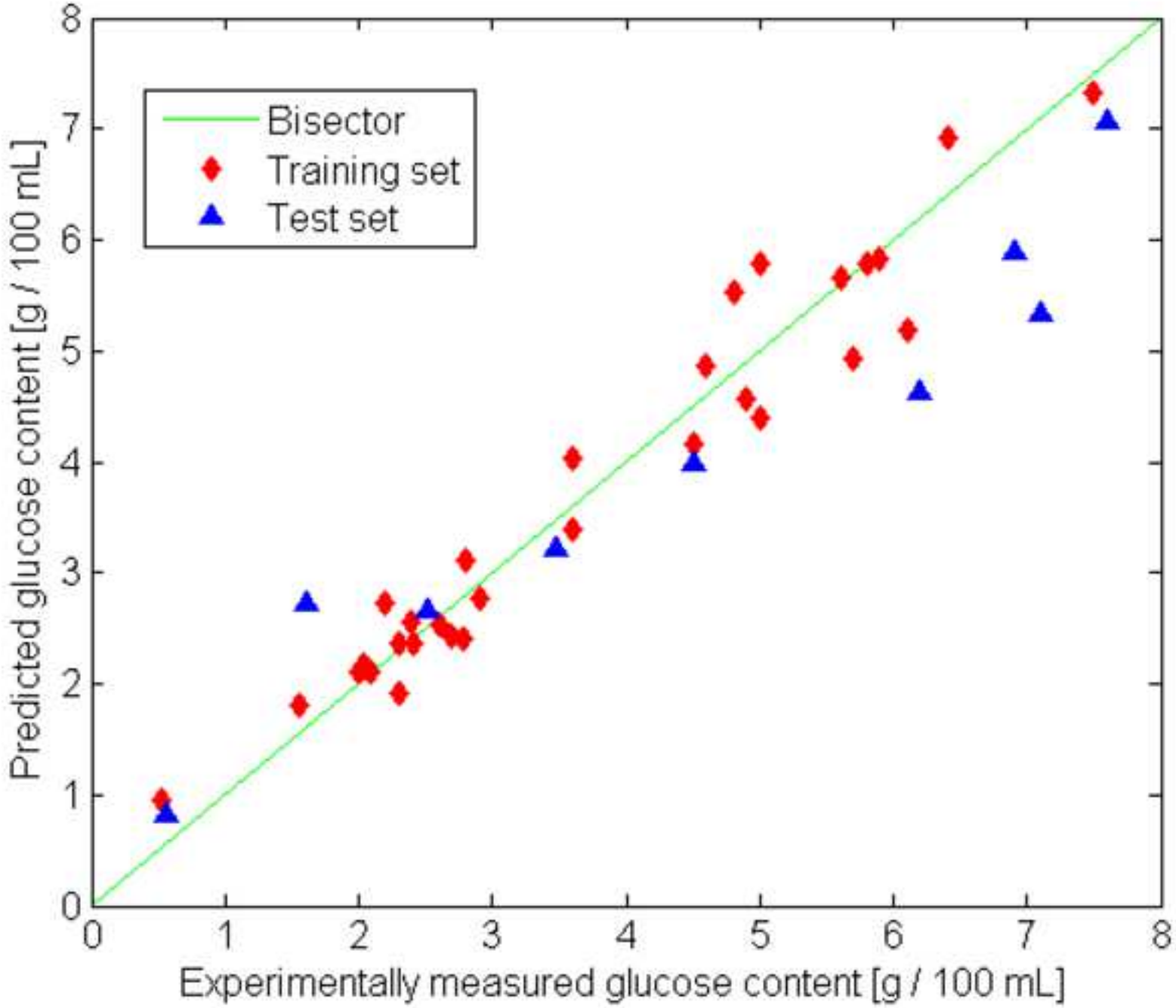

Figure 4

Click here to download Figure Figure4.tif

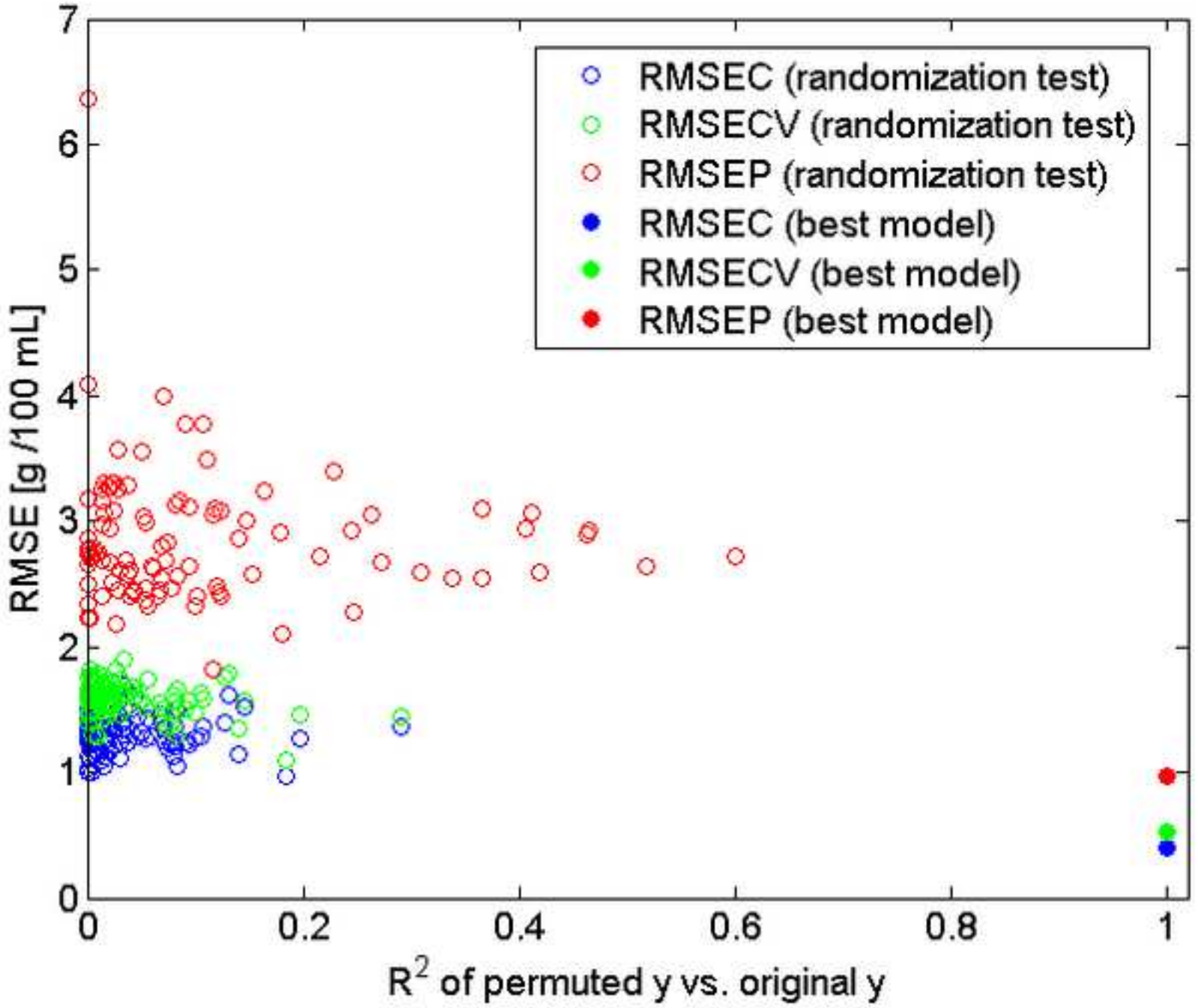

Table 1 – Study of the suitability of the Luff-Schoorl method for the assay of glucose in vegetable milks. Linearity data of glucose determination assessed in aqueous solution. Equation of the curve: Glucose level = Glucose found·(1.089±0.008) ± (0.022±0.010), correlation coefficient r = 0.9988. The analysis of the regression data showed that the slope is highly significant ($p < 0.001$) whereas the intercept is non-significant ($p = 0.07$). Every concentration level was determined in triplicate during each session of testing.

Glucose level (g /100 ml)	**Glucose found** (mean ± SD, g /100ml)	**Accuracy** (mean, g/ 100ml)	**N. Sessions**
0.050	0.053±0.002	106	3
0.088	0.082±0.001	94	1
0.100	0.094±0.004	94	3
0113	0.103±0.006	91	1
0.200	0.201±0.008	101	3
0.400	0.383±0.014	96	1
0.800	0.770±0.024	96	1
1.000	0.897±0.047	90	3

LEGEND Accuracy was evaluated using the equation (1).

Table 2 – Accuracy testing on samples of vegetable milks selected from different producers.

Glucose content, label value (g /100 ml)	**Glucose "spike"** (g /100 ml)	**Value expected** (mean ± SD, g /100ml)	**Value found** (mean ± SD, g /100ml)	**Accuracy** (mean ± SD, g/ 100ml)	**n. replicate measurements**
Sample #7, NB (Soy, low glucose levels)					
0.60	0	NA	0.558±0.006	93±1	3
	0.30	0.90	0.861±0.012	96±1	3
	0.45	1.05	1.01±0.01	96±1	3
Sample #22, VN2 (Kamut, median glucose level)					
4.60	0	NA	4.53±0.36	99±8	3
	2.30	6.90	6.99±0.10	101±1	3
	3.45	8.05	7.53±0.26	96±6	3
Sample #20, SC1 (Rice, high glucose level)					
8.10	0	NA	7.74±0.61	97±9	3
	4.00	12.10	11.62±0.30	96±2	3
	6.00	14.10	13.00±0.14	92±1	3

LEGEND: SD, Standard Deviation; Accuracy was evaluated using equation (1).

Table 3 – List of the 39 independent samples investigated in the study, together with the corresponding "Sugars" concentration values reported on the labels, compared with the glucose determination made with the Luff-Schoorl method and estimated by NIR. All values are reported in g/100ml vegetable *milk*.

Training set samples										
#	Code	Origin	Label		LS	Label - LS	% error Label	NIR	NIR - LS	% error NIR
			Carbohydrates	Sugars						
2	PR	Soy + Ca^{2+}	2.4	2.4	2.3	0.1	4.3%	2.4	0.1	2.9%
3	AL	Soy	2.5	2.5	2.03	0.5	23.2%	2.2	0.1	6.6%
4	VA1	Soy	3.0	2.9	2.8	0.1	3.6%	3.1	0.3	10.8%
5	CE	Soy + Ca^{2+}	2.5	1.8	1.55	0.3	16.1%	1.8	0.3	16.3%
6	VA2	Soy	3.0	2.8	2.3	0.5	21.7%	1.9	-0.4	-17.0%
8	SO	Soy + Ca^{2+}+ Vit. D	2.3	2.3	2.79	-0.5	-17.6%	2.4	-0.4	-13.6%
9	VI	Soy	2.8	2.6	2.4	0.2	8.3%	2.4	0	-1.3%
10	VA3	Soy + Ca^{2+}	3.0	2.9	2.09	0.8	38.8%	2.1	0	0.3%
11	CO1	Soy	0.8	0.7	0.53	0.2	32.1%	1.0	0.4	80.5%
12	MA1	Soy + Ca^{2+}	2.9	2.8	2.9	-0.1	-3.4%	2.8	-0.1	-4.5%
13	VA4	Soy	3	3	2.39	0.6	25.5%	2.6	0.2	7.1%
15	VA5	Soy + Ca^{2+}	3.0	2.9	2	0.9	45.0%	2.1	0.1	4.9%
16	VA6	Oat	6.6	4.6	5	-0.4	-8.0%	4.4	-0.6	-12.2%

17	VN1	Oat	7.6	5.6	5.7	-0.1	-1.8%	4.9	-0.8	-13.7%
18	BS1	Rice + Ca^{2+}	10.2	4.82	6.4	-1.6	-24.7%	6.9	0.5	8.1%
19	CO3	Rice + Vitamin D2	12.5	10.8	5.9	4.9	83.1%	5.8	-0.1	-1.4%
22	VN2	Kamut	7.5	4.6	4.5	0.1	2.2%	4.2	-0.3	-7.5%
23	NN	Rice	13.3	5.5	5.6	-0.1	-1.8%	5.7	0.1	1.0%
24	SC2	Rice + Ca^{2+}	12.5	5.4	7.5	-2.1	-28.0%	7.3	-0.2	-2.4%
25	VA7	Soy	6.7	6.1	4.9	1.2	24.5%	4.6	-0.3	-6.8%
26	CO4	Rice	12.5	10.9	5	5.9	118.0%	5.8	0.8	15.7%
27	VN3	Rice	13.3	5.5	4.8	0.7	14.6%	5.5	0.7	15.1%
28	AN1	Basmati rice	13.0	6.6	6.1	0.5	8.2%	5.2	-0.9	-15.1%
29	MA2	Spelt	10.0	5.5	5.8	-0.3	-5.2%	5.8	0	-0.2%
32	NA	Oat + Ca^{2+}	7.1	6.9	2.2	4.7	213.6%	2.7	0.5	23.6%
34	EM1	Almond	5.4	3.8	3.6	0.2	5.6%	4.0	0.4	11.9%
35	EM2	Quinoa	3.7	2.5	2.6	-0.1	-3.8%	2.5	-0.1	-2.6%
36	IB2	Kamut	7.5	4.6	3.6	1	27.8%	3.4	-0.2	-6.0%
37	OA1	Oat + Ca^{2+}	6.5	4	2.7	1.3	48.1%	2.4	-0.3	-10.3%
38	AN2	Rice	13.0	6.6	4.6	2	43.5%	4.9	0.3	5.6%
Test set samples										
#	**Code**	**Origin**	**Label**		**LS**	**Label – LS**	**% error**	**NIR**	**NIR- LS**	**% error**

			Carbohydrates	Sugars			Label			NIR
1	EN	Soy	3.6	3.3	2.52	0.8	31.0%	2.7	0.2	6.1%
7	NB	Soy	0.8	0.6	0.558	0	7.5%	0.8	0.3	48.3%
14	CO2	Soy + Ca^{2+} +Vit. B2, B12, D	3.7	3.6	3.48	0.1	3.4%	3.2	-0.3	-7.4%
20	SC1	Rice	12.5	8.1	7.6	0.5	6.6%	7.1	-0.5	-7.1%
21	BS2	Oat	7	2.9	4.5	-1.6	-35.6%	4.0	-0.5	-11.3%
30	GI	Rice + Ca^{2+}	10.5	7	6.9	0.1	1.4%	5.9	-1	-14.5%
31	PO	Rice	10.5	7	7.1	-0.1	-1.4%	5.3	-1.8	-24.8%
33	IB1	Barley	11	7	6.2	0.8	12.9%	4.6	-1.6	-25.4%
39	OA2	Oat	6.5	4	1.6	2.4	150.0%	2.7	1.1	71.0%

LEGEND
#: sample number.
Code: coding of the samples, where the first two letters indicate the producer and the number (if present) indicates different batches , when the description in column “Origin” is the same, or different products, when the description in column “Origin” is different.
Origin: indicates the starting material and additives, as stated on the label of the product.
Label: the label of plant milks reports the amount of Carbohydrates and the specification of the share represented by total content of Sugars.
LS: total glucose concentration determined by the Luff-Schoorl method.
Label-LS: difference between the values reported on the label of the product and those determined by the Luff-Schoorl method.
% error Label: relative percent error computed from the equation % error Label= 100×(Label-LS) /LS.
NIR: value of glucose concentration calculated by the NIR method for the training set samples or predicted for the test set samples, respectively.
NIR- LS: difference between the values of glucose concentrations measured by the LS method and those predicted by the NIR calibration model.
% error NIR: relative percent error computed from the equation % error NIR = 100×(NIR-LS) /LS.

Table 4 – PLS calibration results obtained considering the whole spectral range and using different preprocessing methods.

Preprocessing	LVs	RMSEC	RMSECV	RMSEP	R^2_{CAL}	R^2_{CV}	R^2_{PRED}
MNCN	5	0.53	0.74	1.70	0.90	0.82	0.52
SNV+MNCN	5	0.65	1.02	1.51	0.86	0.65	0.62
DETREND+MNCN	5	0.53	0.76	1.52	0.91	0.81	0.61
MSC+MNCN	5	0.50	0.94	1.17	0.92	0.70	0.77
D1+MNCN	4	0.65	0.85	1.49	0.86	0.75	0.63
AUTO	5	0.55	0.79	1.75	0.90	0.79	0.49
SNV+AUTO	5	0.70	1.10	1.55	0.83	0.59	0.60
DETREND+AUTO	5	0.47	0.68	1.38	0.93	0.84	0.68
MSC+AUTO	5	0.70	0.99	1.57	0.83	0.67	0.59
D1+AUTO	3	0.50	0.63	1.36	0.92	0.86	0.69

Table 5 – Calibration results obtained after variable selection by iPLS + selection frequency and using different preprocessing methods. The best calibration model is reported in italic characters.

Pre-processing	**LVs**	**N° of selected variables**	**RMSEC**	**RMSECV**	**RMSEP**	**R^2_{CAL}**	**R^2_{CV}**	**R^2_{PRED}**
MNCN	5	138	0.47	0.62	1.16	0.92	0.87	0.78
SNV+MNCN	5	75	0.51	0.66	1.29	0.91	0.85	0.72
DETREND+MNCN	4	180	0.45	0.57	1.17	0.93	0.89	0.77
MSC+MNCN	3	20	0.51	0.57	1.27	0.91	0.89	0.73
D1+MNCN	*4*	*6*	*0.40*	*0.54*	*0.98*	*0.95*	*0.90*	*0.84*
AUTO	5	138	0.47	0.62	1.16	0.92	0.87	0.78
SNV+AUTO	5	52	0.49	0.66	1.36	0.92	0.85	0.69
DETREND+AUTO	4	81	0.45	0.58	1.13	0.93	0.89	0.79
MSC+AUTO	3	20	0.51	0.57	1.27	0.91	0.89	0.73
D1+AUTO	5	73	0.45	0.60	1.27	0.93	0.88	0.73

Table 6 – Results of the 1-tailed t-test performed on the RMSEC, RMSECV and REMSEP values reported in Figure 4.

Parameter	**t_{CRIT}** *(df = 99; P = 0.05)*	**t_{CALC}**	**$P(t_{CALC})$**
RMSEC	1.66	54.82	3.45×10^{-76}
RMSECV	1.66	74.96	2.56×10^{-89}
RMSEP	1.66	34.63	1.98×10^{-57}

Table 7 – Comparison of the Label and NIR values, with respect to the Luff-Schoorl results.

TRAINING SET		
	% error Label	**% error NIR**
min	-28%	-17%
max	214%	80%
mean	24%	3%
median	11%	0%
st. dev.	47%	18%
TEST SET		
	% error Label	**% error NIR**
min	-36%	-25%
max	150%	71%
mean	20%	4%
median	7%	-7%
st. dev.	52%	33%