Hungarian University of Life Sciences

NON-DESTRUCTIVE ANALYTICAL AND SENSORY EVALUATION OF CERTAIN SNACKS AND GREEN AND ROASTED COFFEE

Benes Eszter Luca

Budapest

2023

Doctoral school	Doctoral School of Food Sciences
Discipline:	Food Sciences
Head of Doctoral School:	Livia Simon-Sarkadi Professor, DSc Hungarian University of Life Sciences Institute of Food Science and Technology Department of Nutrition Science
Supervisors:	Marietta Fodor Associate Professor, PhD Hungarian University of Agriculture and Life Sciences Institute of Food Science and Technology Department of Food and Analytical Chemistry
	Attila Gere Associate Professor, PhD Hungarian University of Agriculture and Life Sciences Institute of Food Science and Technology Department of Postharvest, Supply Chain, Commerce and Sensory Science

....

Head of doctoral school Supervisors

1. INTRODUCTION AND OBJECTIVES

Today's fast-paced lifestyles increasingly lead to tiredness, lack of energy and time, giving more space to foods that can quickly alleviate these "symptoms". Snack products and coffee reach a wide range of consumers through local and global supply chains. Both the snack industry and the industry based on coffee growing and processing are fast-growing sectors, with a variety of products and actors using different processing technologies. The production and transportation of these commodities takes place in large volumes, and therefore requires methods that can effectively and quickly control product quality at different stages of the product life cycle. Near-infrared spectroscopy (NIRS), coupled with different chemometric methods (including multivariate data analysis methods), meets these requirements. It is simple and environmentally friendly to use, as food is analysed non-invasively, without the use of chemicals, through the application of prediction models, using mathematical models generated from spectral and reference data. NIRS relies heavily on computing and electronics not only to control the instrument and collect data, but also to analyse the data. Its rise is facilitated by the digital revolution, which is still ongoing today and which, with the increasing computing power of computers, also reduces the time needed to collect and process data. Since NIRS also enables the non-destructive measurement of food, data can be collected on the quality parameters and quantitative composition of a food matrix throughout the food supply chain, including the qualification of raw materials.

During my doctoral thesis, I mainly explored the potential of NIR spectroscopy in the context of various snack products, green and roasted coffee. For each of these two product groups, the determination and analysis of different characteristics (organoleptic, possibly sensory parameters) is relevant. Among these, I have focused on the analysis of parameters that have not been analysed so far, or have been analysed in a limited number of cases or from other points of view. In the case of snacks, the focus of the work was on the development of rapid methods that can be used to quantitatively estimate the average composition of products based on NIR spectra. In addition, statistical models were developed that can be used to classify snacks according to the raw material used, frying oil, country of production and production technology. For green coffee samples, I investigated compounds (caffeine, various chlorogenic

acids) that are associated with the sensorial properties of roasted coffee, and I also looked for correlations between these and spectral data. The price of green coffee is strongly influenced by its origin, and therefore the identification of origin plays a role in the qualification of the product, which can be done by various spectroscopic methods, including NIRS. To investigate this possibility, samples were obtained from four production regions, including twenty-three countries. I roasted the coffee beans and investigated the effect of different roasting times on the above-mentioned compounds, pH, colour, perceived acidity and bitterness, and the spectral properties of the beans. In addition, prediction models were developed and optimized for non-destructive quantification of the measured parameters.

I defined the goals separately for snacks, green coffee and roasted coffee.

Analysis of snack products:

- creating a spectral database;
- build a reference database by determining the carbohydrate, sugar, fat, crude protein and salt content of samples;
- development of a FT-NIR spectroscopic method to determine the average nutritional value of products;
- pattern recognition and classification according to the raw material and cooking oil used, origin and production technology.

Green and roasted coffee beans:

- creating a spectral database;
- study of pH, caffeine content, chlorogenic acid composition and spectral properties of green coffee in relation to place of origin;
- evaluation of the effect of roasting on the chemical parameters studied and on the colour of the coffee beans;
- study of the acidity and bitterness of the coffee, determined by trained sensory panel, in relation to the chemical and spectral data;
- development and optimization of FT-NIR spectroscopic method for the chemical and sensory parameters studied.

2. MATERIALS AND METHOD

2.1. Analysed samples

During the research, I examined 155 commercially available snack products, which were selected to ensure diversity, both with regard to the production technology used and the place of origin. I obtained products from twenty-five countries, mostly from Europe, but also from Japan, Canada and Africa. The samples can be grouped according to the production technology as follows: extruded (50), deep-fried potato chips (60), alkaline-cooked tortilla chips (31), puffed (4), dried (3), other (8). Most products were made using potatoes or maize as raw material, but some products were made using wheat, rice or soya.

For the experiments with coffee, I obtained fifty green coffees. Forty-eight arabica (one decaffeinated) and two robusta samples from twenty different countries were tested. In the processing of the coffee cherries, the wet process was used for most of the samples (thirty-nine), but I also tested dry, semi-dry, semi-washed and 'honey' coffee. I classified each sample into four regions according to their origin (Africa, Asia, South America, Central America). The fifty original samples were roasted on a small scale using a predefined protocol. Three levels were set according to the first crack (start of roasting), which were light, medium and dark roasting.

2.2. Determination of reference and spectral data

The fat content was determined according to the reference method of the Hungarian Standard 20501-1:2007 (MSZ) using cold extraction with petroleum ether. The carbohydrate and sugar contents of the samples were determined according to the MSZ 6369-12:1979 standard using the Schoorl method. The crude protein content was determined according to the Kjeldahl method, as described in MSZ 20501-1:2007. The salt content was determined by direct potentiometry, an electroanalytical method based on the measurement of electrode potentials.

The acidity of the green and roasted coffee samples was determined by measuring their pH, given that the perceived acidity of coffee is strongly correlated with its pH. The caffeine content was determined according to ISO 20481:2008 with minor modifications using an HPLC-UV method. Chlorogenic acids in coffee beans were quantified by HPLC-MS/MS with gradient elution. The parameters describing the colour of the samples were recorded using a tristimulus colorimeter in the CIE L*a*b* colour space. Sensory evaluation of 150 roasted coffee samples was carried out by a panel

of six panelists with hundreds of hours of experience in coffee evaluation. The sensory characteristics of the coffee were acidity and bitterness. During the tasting, the characteristics were rated on a ten-point scale (1: weak, 10: intense).

For the spectral data acquisition, a Bruker MPA^{TM} multi-purpose FT-NIR spectrometer (Bruker, Ettlingen, Germany) with diffuse reflectance measurement setup was used. I analysed 40-60 g of ground snacks and about 50 g of coffee in its original form.

2.3. Chemometric methods

Different chemometric methods were used to evaluate the reference and spectral data. Descriptive statistics were performed on the chemical and sensory data, separately for the whole data set and for the normal Arabica coffees from the four growing regions. For the latter data set (forty-seven samples), I also tested whether there were significant differences between the regions in relation to the parameters studied. To test for differences between groups, I ran the Kruskal-Wallis test (α=0.05). Principal component analysis (PCA) was used for a comprehensive examination of the dataset and to identify spectral outliers. For snacks, I used PCA for the exploratory analysis of spectral data, while for coffee I used PCA for both reference and spectral data. I ran different data pre-processing methods on the spectral data to reduce variations due to the physical properties of the samples, which were the standard normal variate (SNV), multiplicative scatter correction (MSC), derivatives (first and second derivatives, FD and SD), or the combination of FD or SD and SNV algorithms. The different data preprocessing methods were complemented by orthogonal signal correction (OSC) preprocessing in the case of coffee samples.

2.3.1. Classification methods

Supervised learning algorithms were used to analyse the data for different quality characteristics of snack products and coffee. In the case of snack products, due to the diversity of the samples, different options were available for their classification, which I summarised in Table 1. For the classification according to each criterion, I used the random forest (RF), RPropMLP algorithms and PLS-DA method, all validated by cross-validation and test validation. Accuracy (ACC), true positive rate (TPR), true negative rate (TNR) and receiver operating characteristic area under the curve (ROC AUC) were

used to determine and compare the performance of the models for each classification dataset and algorithm.

Table 1: Summary of the snack samples according to the different classification criteria

classification criteria	groups	samples
cooking oil (dataset 1)	sunflower	56
	palm	38
raw material (dataset 2)	potato	71
	maize	55
origin (dataset 3)	Belgium	14
	Czech Republic	14
	Poland	19
	United Kingdom	12
production technology (dataset 4)	deep-frying	60
	alkaline cooking (tortilla)	31
	extrusion	50

The quality of green coffee, and therefore its price, is significantly influenced by its place of origin, so origin identification can play an important role in quality control. The samples were grouped into four data sets, which are summarised in Table 2.

Table 2: Summary of coffee samples according to the different classification criteria

classification criteria	sub-groups	samples
four regions (dataset 1)	Africa	14
	Asia	11
	South America	8
	Central America	13
three regions (dataset 2)	Africa	14
	Amerika	20
	Asia	11
two regions (dataset 3)	Africa	14
	Amerika	20
two regions (dataset 4)	Africa	14
	Asia	11

The classification of the green coffee samples was performed using the Classification Learner application of MATLAB software, which allows the development of several models, including decision trees (DT), discriminant

analysis (DA), support vector machines (SVM), nearest neighbour (k-NN), ensemble and neural network-based classification.

2.3.2. Partial least squares regression (PLSR)

For the quantitative estimation of the different parameters, I used partial least squares (PLS) regression for both product groups. To evaluate the performance of the models created, I considered different statistical indices. For the model validation, random, five-fold cross-validation and test-validation were used. In case of the latter, the data set was split 70/30 using the Kennard-Stone algorithm.

Various variable selection methods are available to improve the performance of prediction models. They can be used to identify the spectral ranges (variables) that are most likely to be related to the parameter under regression. This can reduce the complexity of the models and the computational capacity required. In my doctoral thesis, I used three different variable selection methods with different settings, which were interval partial least squares principal regression (iPLSR), genetic algorithm (GA) and projection based on the importance of variables (VIP).

For data interpretation, various statistical software were used, such as SPSS Statistics 23 (IBM, New York, USA), The Unscrambler X 10.4 (CAMO Software, Oslo, Norway), MATLAB 2019b (Mathworks Inc., Massachusetts, USA) and PLS Toolbox 9.0 (Eigenvector Research Inc., Mansos, WA, USA), KNIME analytics platform 4.0.2 (KNIME AG, Zurich, Switzerland).

3. RESULTS

3.1. Snack products

3.1.1 Overview of reference and spectral data

In a first step, the NIR spectra of all products were recorded and then the quantification of the macro components was performed. The samples were processed over a two-year period. Not all components were determined for all samples, therefore the number of available measurements varies. The most important quantitative data on measured components and calculated energy content are summarised in Table 3.

Table 3: Main descriptive statistical characteristics of the parameters measured for snacks

	average±SD	minimum	maximum	median	samples
fat [g/100 g]	26,68±9,57	2,24	56,13	27,48	156
carbohydrate [g/100 g]	56,07±10,21	20,51	84,02	55,91	90
suugar [g/100 g]	5,56±3,21	0,94	20,89	5,30	89
protein [g/100 g]	9,23±6,33	3,01	40,06	7,51	91
salt [g/100 g]	1,56±0,77	0,17	4,44	1,47	91
energy [kJ/100 g]	2029±272	1264	2520	2058	90

I compared the measured values with the data on the packaging of the products, which showed significant differences, mainly in carbohydrate and protein content, taking into account the tolerances given in Regulation (EU) No 1169/2011. This is probably due to the fact that the nutritional information is often based on calculations and databases and not on analytical tests. In addition, the inconsistent quality of the raw materials of the product may also be the cause of the discrepancy.

Prior to statistical evaluation of the data, I identified the characteristic absorption bands in the NIR spectrum that could be associated with the components of interest. Two regions of the average spectra (12800-11500 cm^{-1} and 3800-3500 cm^{-1}) were cut off as these did not carry systematic information. Snacks are usually high carbohydrate foods, the final form of which is obtained by frying in oil. It follows that the NIR spectra of the samples show substantial similarities (Figure 1). Nevertheless, there was considerable variation between the spectral data of the different samples, mainly in the wavenumbers in the 6000-5400 and 4900-4000 cm^{-1} range. This can probably be explained by

differences in the production technology and the variation in the raw and additional materials used.

The main difficulty encountered when analysing the samples by NIR spectroscopy was the differences in composition and appearance of the products. It should be mentioned that the milling process causes the particles of the higher fat content products to adhere better to each other, thus creating a more reflective surface compared to the other samples. The main differences were therefore caused by the variation in particle size, which was reflected in the shift in the spectra along the vertical axis (Figure 1).

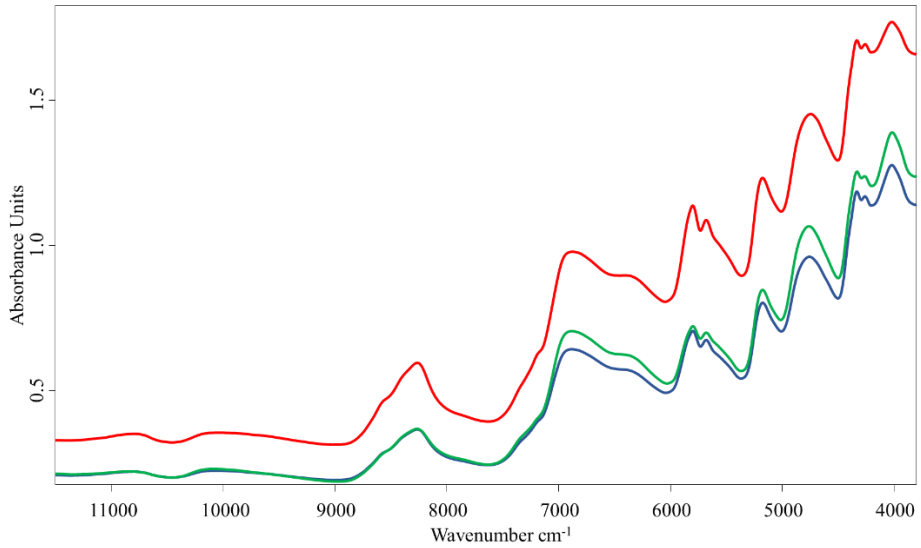


Figure 1: Average spectrum of snack products by production technology: ● deep-fried potato chips, ● tortilla, ● extruded snacks

PCA was used to detect spectral outliers at 95% confidence intervals using F-residuals and Hotelling's T^2 values. Five samples were identified as true spectral outliers, associated with matrix properties different from the others. These included salted sticks, puffed rice, beetroot and roasted pork skin based products.

3.1.2. Classification of snack products based on the NIR spectrum

Classification methods were performed after removal of spectral outliers identified during PCA. The first twenty principal components were used for

modelling. Groups with less than five samples were excluded from the evaluation because significant differences in the sample numbers of the classes would bias the model.

Classification models based on the frying oil

Sunflower and palm oil were the most commonly used frying oils for the products studied, based on which two groups were created. The PLS-DA and RPropMLP methods were found to be more effective than the RF method for cross-validation, while the PLS-DA performed worse for test validation. The latter results suggest that RPropMLP has more favourable statistical parameters than the other two algorithms. Overall, RPropMLP gave the best results in terms of ROC curves and AUC values. Thus, the RPropMLP model is the most suitable for the determination of the frying oil used in the production of snack products. Accuracy (ACC), TPR and true negative sample rate (TNR, specificity) values together with AUC values are summarized in Table 4 for both cross-validation and test validation.

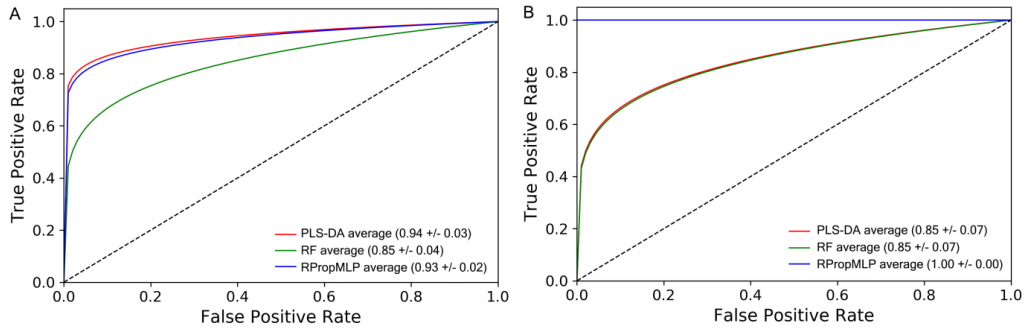


Figure 2: Summary of the average ROC curves of the three models for frying oil classification for A) cross-validation and B) external validation. PLS-DA (●), RF (●), RPropMLP (●).

Table 4: Performance parameters of the models obtained by three-fold cross-validation and test validation when classifying based on the frying oils used

method	validations	class	AUC	TPR	TNR	ACC
PLS-DA	CV	2	0,94	0,93	0,85	0,89
PLS-DA	Test	2	0,85	0,56	0,83	0,68
RF	CV	2	0,85	0,91	0,53	0,76
RF	Test	2	0,85	0,76	0,75	0,76
RPropMLP	CV	2	0,93	0,79	0,89	0,83
RPropMLP	Test	2	1,00	0,65	1,00	0,79

Classification models in terms of raw material

When examining dataset 2 (Table 1), the grouping was based on the two main raw materials: a) potatoes, b) maize. As these were the two most commonly used raw materials, I was able to work with spectral data from a total of 126 samples. Table 5 shows that all the three models produced excellent results on all performance metrics, both in cross-validation and test-validation. The accuracy of the RF model is slightly lower than the other two models. The RF had worse statistical characteristics, but the difference is not significant. In conclusion, the PLS-DA and RPropMLP models are ideal for classifying snack samples based on their raw material.

Table 5: Performance parameters of the models obtained by three-fold cross-validation and test-validation in the classification by raw material

Method	Validation	Groups	AUC	TPR	TNR	ACC
PLS-DA	CV	2	1,00	0,98	1,00	0,99
PLS-DA	Test	2	1,00	0,95	1,00	0,97
RF	CV	2	0,98	0,96	0,91	0,94
RF	Test	2	0,98	0,97	0,91	0,94
RPropMLP	CV	2	1,00	0,97	1,00	0,98
RPropMLP	Test	2	1,00	1,00	1,00	1,00

Classification models by country of production

For the classification of the dataset by country of origin, four major groups could be created, containing more than five samples: Belgium, Czech Republic, Poland and the United Kingdom. In total, fifty-nine samples were used for the classification. In this case, the neural network based RPropMLP algorithm is more favourable than the other two. The ROC curves in both the CV and test-validation cases are further away from random classification, but the differences between the models are not significant. Based on the AUC values, all three models are able to classify samples. The accuracy of PLS-DA is much lower than RPropMLP in both CV and test validation, thus concluding that RPropMLP is the recommended solution for the classification of snack products by origin (Table 6).

Table 6: Performance parameters of the models obtained by three-fold cross-validation and test-validation in the classification by country of production

Method	Validation	Groups	AUC	TPR	TNR	ACC
PLS-DA	CV	4	0,93	0,81	0,91	0,81
PLS-DA	Test	4	0,92	0,71	0,91	0,71
RF	CV	4	0,93	0,78	0,90	0,78
RF	Test	4	0,92	0,78	0,90	0,78
RpropMLP	CV	4	0,97	0,90	0,96	0,90
RpropMLP	Test	4	0,97	0,83	0,94	0,83

Classification models based on production technology

Based on production technology, three major groups could be created, into which 141 samples could be classified: frying in oil, alkaline cooking and extrusion. In terms of cross-validation results, the three models performed almost equally well. In test validation, there were larger, but still not statistically large differences. Based on the ROC curves, all three models were acceptable. The results in Table 7 show that all three models can be used to determine the production technology of snacks based on the NIR spectra.

Table 7: Performance parameters of the models obtained by five-fold cross-validation and test-validation in the classification by manufacturing technology

Method	Validation	Groups	AUC	TPR	TNR	ACC
PLS-DA	CV	3	0,96	0,92	0,96	0,92
PLS-DA	Test	3	0,97	0,88	0,95	0,88
RF	CV	3	0,97	0,89	0,93	0,89
RF	Test	3	0,95	0,91	0,94	0,91
RPropMLP	CV	3	0,96	0,91	0,96	0,91
RPropMLP	Test	3	0,91	0,84	0,92	0,84

Results of partial least squares regression (PLSR)

PLSR was used to quantitatively estimate the fat, carbohydrate, sugar, protein and salt content of snack products. In each case, random leave-five-out-cross-validation and test validation were used to determine the optimal number of factors, which were determined using RMSECV and RMSEP. After removing the spectral outliers, PLS regression was run on the raw data, which was used to study the relationships between variables and to capture the initial statistical

parameters. The scatter plots for each nutrient showed that the X-Y explained variance values were not appropriate, making the models not suitable for quantitative estimation. To overcome this, the previously listed data pre-processing methods were used.

The best performing prediction models and their most relevant statistical parameters are presented in Table 8.

Table 8: Calibration, leave-five-out cross validation and test set validation statistics of macronutrients for snack products by NIR spectroscopy

parameter	calibration				cross-validation		test validation		measurement range [g/ 100g]
	R^2	R_t^2	RMSEC [g/ 100g]	$RMSEC_t$ [g/ 100g]	Q^2	RMSECV [g/ 100g]	Q_t^2	$RMSECV_t$ [g/ 100g]	
fat	0,99	0,99	1,00	1,11	0,98	1,10	0,98	1,1	2,2-45,1
carbohydrate	0,95	0,94	1,53	1,53	0,92	1,77	0,92	1,90	45,1-69,7
sugar	0,94	0,94	0,49	0,51	0,92	0,54	0,93	0,47	1,7-8,6
protein	0,99	0,99	0,65	0,65	0,98	0,78	0,98	0,93	3,0-40,1
salt	0,96	0,98	0,11	0,11	0,93	0,14	0,91	0,16	0,7-2,5

Prediction of energy content by iPLSR

By calculating the energy content from the amount of macronutrients in the food, the whole spectrum can contain relevant information.

At first, PLS regression (full-spectrum model) was performed using all variables without data pre-processing. The X-Y explained variance values were not satisfactory in this case either, thus the model cannot be used to predict energy content. I could not significantly improve the correlation using different data treatment methods. The FD spectral transformation is not recommended for the full-spectrum model due to the scatter plot and the high distortion of the model (Table 9). The different data pre-processing methods could not improve the correlation significantly. The FD spectral transformation is not recommended for the full-spectrum model based on the scatter plot and the high bias of the model (Table 9).

To improve model performance, iPLSR was used. According to the results, models with 20 segments after MSC or SNV pre-processing showed the best performance. In both cases, the selected wavenumber ranges were 8802-8038 cm^{-1} and 4559-3795 cm^{-1}. However, as a result, absorption bands that are related to fats, oils, proteins, starch and sugars in the sample should be excluded from the evaluation based on the determined RMSECV values. The

higher mean error observed may be related to the unclear relationship between spectral changes and energy content due to the diverse composition of the products. The results clearly show that iPLSR was not suitable for reducing the mean error, but it was suitable for improving the X-Y explained variance. By including the above-mentioned absorption bands in the model, the mean error value increased, but the model bias decreased. The use of additional PLS factors would further improve the statistical parameters, however, they strongly increase the bias (overfitting). The best model after SNV data pre-treatment was obtained for 20-segment iPLSR (Table 9).

Table 9: Comparison of the relevant interval PLS (iPLS) models for the energy content of snacks (range: 1264.3-2520.2 kJ/kg).

method	variables	pre-processing	LV	R^2	Q^2	RMSEC [kJ/100g]	RMSECV [kJ/100g]
full spectrum model	1000	no	5	0,89	0,87	86,25	94,67
	1000	MSC	4	0,87	0,86	90,51	97,76
	1000	SNV	4	0,87	0,86	90,4	97,64
iPLS – 20 intervals	100	no	6	0,90	0,88	80,68	89,07
	300	***MSC***	***4***	***0,89***	***0,87***	***86,49***	***94,66***
	550	***SNV***	***4***	***0,89***	***0,87***	***86,38***	***94,66***
+ regions of 2nd and 4th absorption peaks	300	MSC	4	0,88	0,86	90,03	96,26
+ regions of 2nd and 4th absorption peaks	400	SNV	4	0,87	0,86	90,32	97,03

As the snack industry is strongly characterised by product diversification, it is recommended to analyse additional samples and to use other, even non-linear chemometric methods to estimate the energy content more accurately.

3.2. Green and roasted coffee

3.2.1. Evaluation of reference data

Green coffee

The results and the main statistical parameters of the chemical measurements on the fifty green coffee samples are summarized in Table 10. The results of the measurements were evaluated by comparing the data from samples from the four geographical regions using one-way analysis of variance (ANOVA), but no significant differences were found for any of the variables.

Table 10: Summary of results on the chemical composition of green coffee

	average ±SD	minimum	maximum	median	RSD [%]
pH	5,73±0,12	5,42	6,00	5,71	0,32
caffeine [mg/g]	11,40±2,74	0,33	22,70	11,36	1,93
5-CQA [mg/g]	47,96±7,55	18,93	61,61	47,43	4,99
4-CQA [mg/g]	11,15±2,89	5,52	19,06	10,49	5,14
3-CQA [mg/g]	7,59±2,89	3,75	18,37	6,80	4,97
3,5-diCQA [mg/g]	4,39±1,18	2,34	8,09	4,45	5,16
4,5-diCQA [mg/g]	3,08±1,08	1,85	8,00	2,89	4,72
3- FQA [mg/g]	2,35±1,07	0,98	5,37	1,93	5,62
∑CQA [mg/g]	76,07±9,76	61,18	98,07	74,19	-

RSD%: mean relative standard deviation of the reference method (n=50)

The chemical composition of samples from different regions is determined by weather, growing conditions and the processing methods used. The measurement results are in general agreement with the findings of previous scientific studies. It is worth pointing out that the analysed samples originated from different growing areas, thus the amount of the different components ranged over a wide range.

To comprehensively examine and visualise the quantitative data, PCA was run on the dataset. The decaffeinated and the two robusta samples were excluded from the evaluation as their chemical composition differs significantly from the other samples. Taking into account the results, the differences between the samples could be mainly due to variation in the concentrations of 5-CQA, 4-CQA, 3-CQA and 3-FQA, respectively, based on the factor weights obtained. To a lesser extent, the amount of 3,5-diCQA and caffeine may contribute to the dispersion of the samples through the PCs.

Roasted coffee

Roasting is a complex process during which the physico-chemical properties of green coffee are significantly altered due to the high temperatures (230-280°C). The chemical composition of roasted coffee depends largely on the roasting temperature and time set and the specific properties of the green coffee. It is common practice in the industry to roast coffee according to a specific colour scale (Agtron scale). Since the colour of coffee is influenced by its composition and the post-harvest technology, the optimum roasting level must be adjusted for each coffee to obtain the desired sensory properties.

The compositional variation of green coffee samples was analysed at three roasting levels along different physical, chemical and sensory parameters. The results obtained are summarised in Table 11.

In general, coffee acidity shows a decreasing tendency with increasing roasting level. A significant difference (α=0.05) was observed between the roast levels.

The results obtained from the caffeine determination are in agreement with the results of previous studies. The caffeine content of arabica samples ranged from 9.9 to 15.43 mg/g. As caffeine is a thermostable compound, its concentration does not change significantly during roasting, with a possible slight increase. This is confirmed by the results of the analysis of variance, which showed that there was no significant difference in caffeine concentration between the levels (Table 11).

Roasting, on the other hand, has a significant effect on the concentration and composition of chlorogenic acids in coffee. Depending on the type of processing, their concentration can vary between 0.5 and 6%. If the roasting process is more intensive, their amount can be less than 1% of the dry matter (STEFANELLO et al. 2019). The results in Table 11 show that there is a significant difference between the different roasting levels. Considering the average values, the concentrations of 5-, 4- and 3-CQA and 3,5-diCQA showed a decreasing trend during roasting.

Table 11: Average values of reference data for coffee samples by roasting level

	zöld	világos	közepes	sötét	RSD [%]
pH	5,73±0,12	4,89±0,09[a]	4,94±0,12[b]	5,07±0,16[c]	0,13
koffein [mg/g]	11,40±2,74	12,18±1,10[a]	12,27±0,87[a]	12,42±1,11[a]	1,27
5-CQA [mg/g]	47,96±7,55	20,10±13,93[a]	15,53±6,62[b]	12,25±3,67[c]	4,87
4-CQA [mg/g]	11,15±2,89	11,10±2,16[a]	9,05±1,67[b]	7,04±1,97[c]	4,97
3-CQA [mg/g]	7,59±2,89	8,70±1,91[a]	6,65±1,42[b]	5,05±1,97[c]	5,18
3,5-diCQA [mg/g]	4,39±1,18	2,36±0,70[a]	1,94±0,50[b]	0,92±0,30[c]	4,61
4,5-diCQA [mg/g]	3,08±1,08	1,96±0,76[a]	1,40±0,52[b]	1,52±0,92[b]	4,56
3-FQA [mg/g]	2,35±1,07	3,58±1,14[a]	3,65±0,95[a]	2,81±0,48[b]	5,14
sum CGA [mg/g]	76,07±9,76	47,81±8,17[a]	38,23±7,06[b]	29,60±7,64[c]	-
érzékszervi savasság	-	5,21±1,15[a]	4,26±1,31[b]	2,98±1,24[c]	-
érzékszervi keserűség	-	3,63±1,53[a]	4,33±1,45[b]	5,55±1,45[c]	-
L*	61,73±2,36	43,37±1,33[a]	41,82±1,35[b]	39,40±1,49[c]	
a*	2,08±0,95	6,91±0,39[a]	6,43±0,41[b]	5,63±0,48[c]	
b*	16,18±1,96	10,04±1,00[a]	8,56±1,17[b]	6,22±1,34[c]	

*Homogeneous subsets defined by Tukey HSD post hoc test on the basis of the variables in the rows are indicated by letters in superscript ($p<0.05$). RSD%: mean relative standard deviation of the reference method for roasted coffee (n=150)

Compared to green coffee, the total CGA content of the samples decreased on average by 35.9% for light roast, 49.2% for medium roast and 60.5% for dark roast. In general, there is indeed a downward trend in the amount of chlorogenic acids (total CGA), but there are several differences due to the different genetics, growing and processing conditions of the different green coffees. It is important to bear in mind that chlorogenic acids are degraded during roasting, but isomerisation and other various chemical transformations can also be observed. A more detailed description of these changes may highlight the difficulties in coffee quality control. The main trends in the changes of the individual chlorogenic acids studied are as follows:

- 5-CQA - a clear decrease, except for five samples; higher concentration in dark roasts than in medium roasts
- 4-CQA - no clear trend, in many cases an increase compared to green coffee; in most cases the dark roasts have the lowest concentration
- 3-CQA - two main trends - clear decrease or increase and systematic decrease
- 3,5-diCQA - two main trends - clear decrease or decrease compared to green, then increase and decrease again

- 4,5-diCQA and 3-FQA - at least six trends; 4,5-diCQA generally shows a 50% decrease at the end of roasting; 3-FQA showed an increase of about 70% at the end of roasting (35 samples)

It was also observed that, for a given coffee, the concentrations of 4-CQA and 3-CQA followed a similar trend during roasting (42 samples). This means that if the amount of one compound increased/decreased between two roasting levels, so did the other.

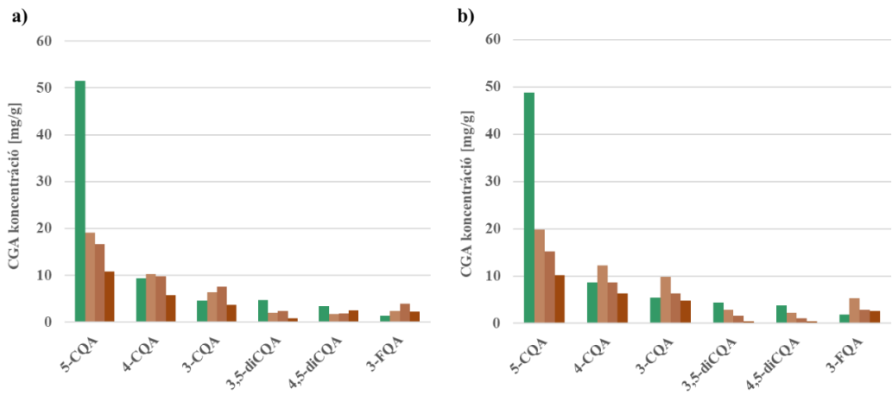
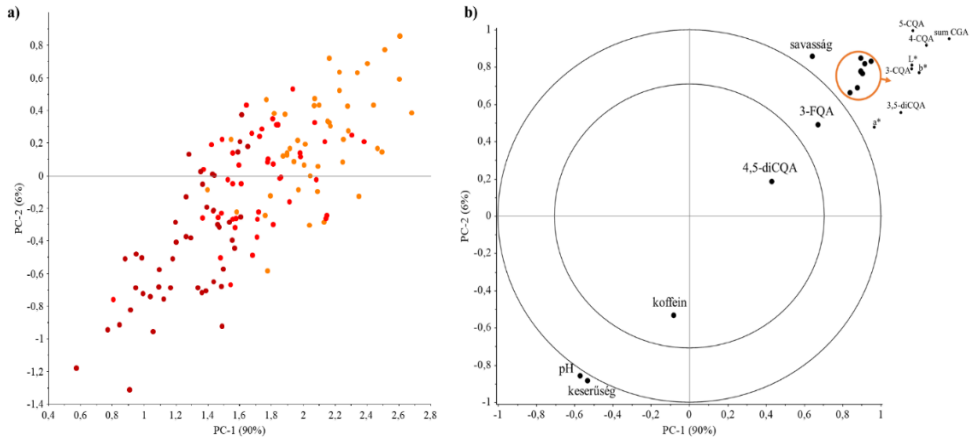


Figure 3: Variation of the concentration of each chlorogenic acid during roasting. a) Sample 4 (Ethiopia) b) Sample 46 (Uganda). ● green coffee ● light ● medium ● dark roast.

Correlation analysis and PCA were also run on the reference data, and only a few of the results are highlighted here. The PCA results clearly show (Figure 4) that the samples are scattered along the principal components according to the roasting levels. A transversal trend emerges between the variables corresponding to lower PC1 and PC2 and higher PC1 and PC2 values, from dark to light roasting. The correlation loading plot (Figure 4) shows which variables have the most significant effect on the position of the samples. Besides the colour of the coffee, the amount of chlorogenic acids present can also be an indicator of the roasting condition. The results clearly showed that there is a positive correlation between the CGA content of the samples and the defined colour parameters. The correlation exists in such a way that as the roasting progresses, the colour of the coffee becomes darker and the concentration of chlorogenic acids decreases. The relationship varies with the individual compounds, but when looking at the sum of the chlorogenic acids tested, there is a very strong correlation with L* ($r(136)=0.80$) and b*

(r(136)=0.81). In addition, there is a strong correlation with the measured a* values (r(136)=0.73). These correlations are clearly shown on the correlation loading plot of the PCA (Figure 4.b).

Figure 4: Results of principal component analysis performed on the



quantitative data of roasted coffee: a) scores plot; b) loading plot. ● green coffee ● light ● medium ● dark roast.

If we observe Figure 4.b, the location of the pH variable can be explained by the fact that a higher pH value indicates a coffee with a less acidic character. There is a strong positive correlation between bitterness and pH (r(136)=0.73). There is a similar but opposite correlation between sensory acidity and measured pH (r(136)=-0.73). The negative correlation is possible because a lower pH (more acidic) tends to be associated with a higher score in the sensory evaluation (opposite scale).

The position of the samples is also largely determined by the amount and composition of chlorogenic acids they contain, which is clearly shown by the correlation loadings. Samples containing higher amounts of CGAs were also found to have a more acidic character on the basis of the sensory evaluation. A relatively strong correlation was also observed between the CGA content of coffee and its sensory acidity (r(136)=0.64), but a moderate but negative correlation (r(136)=- 0.58) between perceived bitterness (α=0.01).

3.2.2. Evaluation of spectral data

Green coffee

Figure 5 shows the diffuse reflectance NIR spectra of green coffee samples in the range 12500- 3800 cm^{-1}. Since the spectra also contain information about the physical properties of the samples, the spectral data are influenced by the different colour and shape of the coffee beans, and by the silver skin that may be left on the surface. It can be clearly seen that the spectral characteristics are similar, but there is a visible difference for decaffeinated sample 1 (●), the two robusta (● 25, ● 33) and the 49 (●) arabica coffee. After second derivative spectral transformation, the absorption bands specific to coffee were also determined.

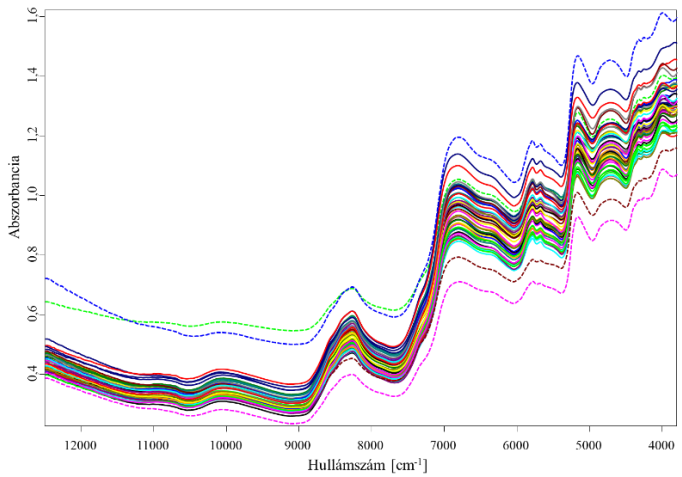


Figure 5: Diffuse reflectance NIR spectra of green coffee samples. Samples marked with dashed lines: ● decaffeinated 1G, ● 25G robusta, ● 33G robusta, ● 49G

PCA was performed using the original and SNV data for the 12500-3800 cm^{-1} range. The PCA results can also be examined in the context of the origin of the samples through the analysis of the patterns in the data set. The results show that the samples do not show clear separation with respect to the four regions, but small clusters have formed. A similar result was obtained in further publications in the analysis of American and Asian samples. In addition to chemical composition, the position of the samples is strongly influenced by the physical properties of the green coffee, such as the appearance, shape and colour of the beans. These characteristics affect the spectral range between

12500-9000 cm^{-1} most strongly. According to the defined correlation loadings, there is a strong positive correlation between the location of the samples and their physical properties.

Effect of roasting on NIR spectral data

The spectral characteristics of green and roasted coffee are similar with regard to the main absorption bands. The absorbance values measured in the spectral region between 8500 and 3800 cm^{-1} showed a decreasing trend with roasting (as the moisture content decreases), which is in agreement with the results of previous studies. This characteristic is reversed in the range 12500-8500 cm^{-1} (close to the visible range), with an increase in the measured absorbance values. During roasting, the coffee beans become darker, resulting in a change from the original greenish colour, through yellowish, to brown, to almost black. This change is the result of the Maillard reaction, pyrolysis and caramelisation.

PCA was also applied to the NIR spectra of roasted coffee samples after SNV data preprocessing, and these results are illustrated in Figure 6. For the evaluation data in the spectral range 9000-3800 cm^{-1} were used. In general, along PC1, light roast coffees tended to have negative values (except for samples 28L and 44L) (Figure 6.a). Medium roast coffees had both negative and positive values, 54 and 46%, respectively. This suggests that these samples are least consistent with the predefined roasting levels. The dark roast coffees typically fall in the positive range along the PC1 axis (except for sample 43D). Since the roasting conditions were the same for all coffees, the samples were actually arranged into a roasting sequence. The positioning of the samples is related to the roasting level and the variation in the concentration of chlorogenic acids during roasting.

The NIR spectra show that the boundaries between the different levels are blurred, due to differences in the quality of green coffee. Correlation loadings show that the most significant spectral ranges overlap with the results obtained for green coffee.

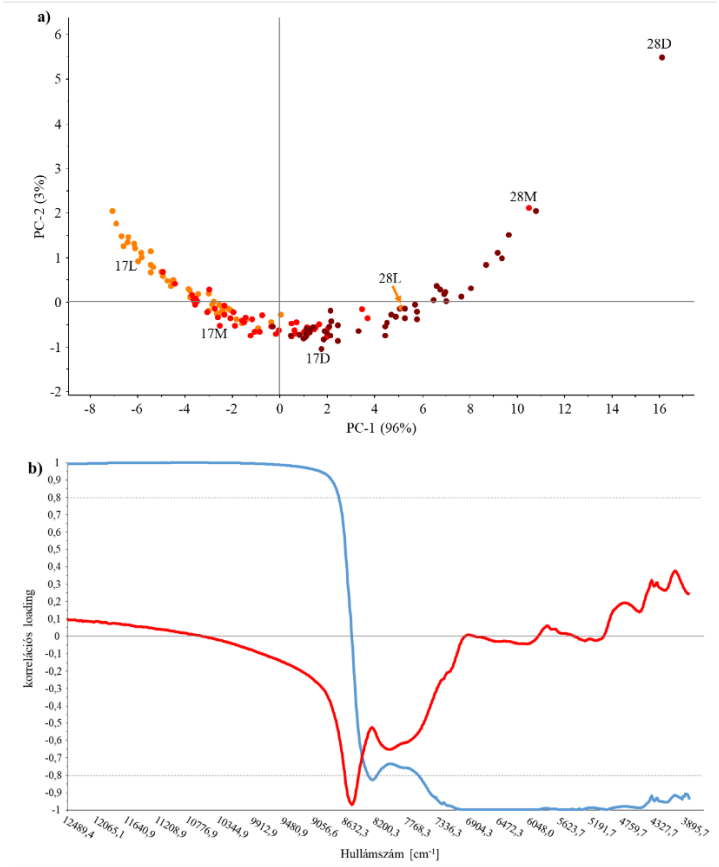


Figure 6: Results of principal component analysis performed on the NIR spectra of roasted coffee after SNV. a) scores plot: plotted according to the roasting level of the samples: ●light (L); ●medium (M); ●dark (D); b) loading plot: ● PC1, ● PC2.

Based on the measurement results presented, there is a significant difference in the concentration of CGAs at each roasting level. The most obvious changes during roasting were in the amount of 5-CQA, which in most cases decreased. On this basis, 5-CQA may be the most suitable of the compounds studied for assessing the roasting degree of coffee. When the samples are plotted on the scores plot (Figure 7) according to the concentration determined, a similar trend as for roasting levels is observed. This clearly shows that sample 28L belongs to the medium roasting samples based on spectral data and 5-CQA amount, while samples 28M and 28D belong to the dark roasting samples.

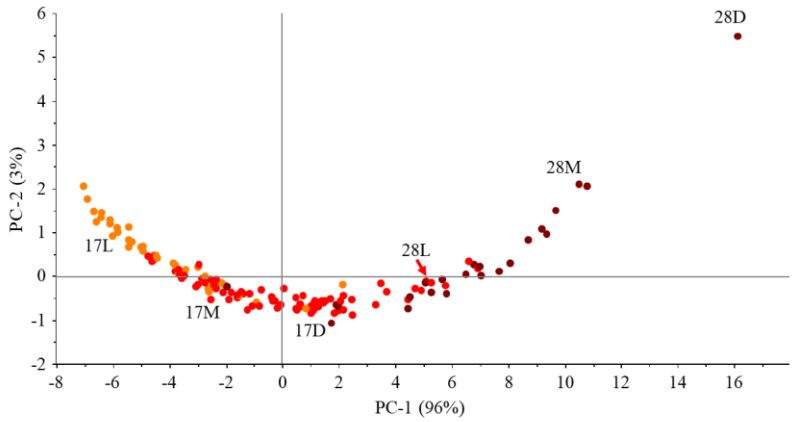


Figure 7: Scores plot of the PCA run on the NIR spectra of roasted coffee after SNV, plotted by concentration of 5-CQA: ● 19,05-26,74 mg/g; ● 11,36-19,05 mg/g; ● 3,67-11,36 mg/g

Based on the results, the roasting process can be investigated by NIR spectroscopy, but to understand the process we need a comprehensive database of the physical and chemical characteristics of the samples.

3.2.3. Identification of the origin of green coffee

Forty-six samples, covering the full spectral range (12500-3800 cm^{-1}), were used for the evaluation. The classification process was performed on four datasets, with four, three and two regions of origin as classification criteria.

The analysed green coffees originated from different parts of the world, thus showing significant variations in their physical properties and chemical composition. The samples also originate from different countries and production areas within the same region, and the available data do not indicate any parameters that show a clear correlation with the region of origin. In addition, it is important to highlight that the chemical composition is also significantly influenced by the genetics of the plant, which is not a negligible aspect in such a diverse data set.

The results for four and three growing regions show that the performance of the models is not adequate for the classification of coffee samples. The resulting clusters can probably be explained by factors unknown to me, such

as the genotype of the crop. The statistical parameters of the models were greatly improved for binary systems (Africa-America, Africa-Asia). The best results were obtained for the classification of African and Asian samples using FD+SNV data pre-processing (Figure 8), with a model accuracy of 92%.

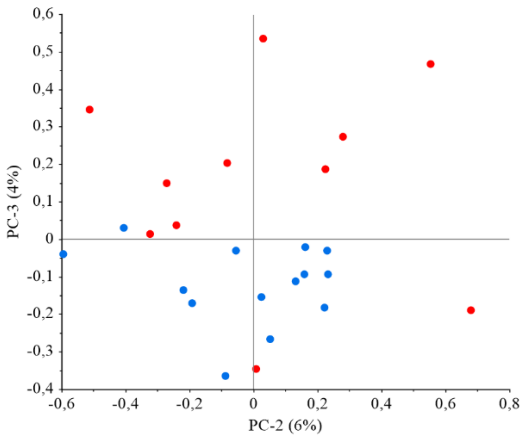


Figure 8: The obtained scores plot of the kernel Naïve Bayes method for the second and third principal component. ● Asia; ● Africa

3.2.4. Results of partial least squares regression (PLSR)

Green coffee

Due to the relatively small number of samples and the variations in their quality, three spectra were used for each sample to generate regression models (n=138). The spectral range between 10000 and 3800 cm^{-1} was included in the evaluation. Estimation models were created to quantify the pH, caffeine, 5-, 4-, 3-CQA, 3,5- and 4,5-diCQA and 3-FQA content of green coffee. The optimisation of the prediction models for each matrix and parameter under study is carried out empirically along different strategies. In this way, spectral transformation methods and different variable selection procedures, or combinations of these, are used to improve the statistical parameters of the regression relationships. The statistical parameters of the best PLS models are summarised in Table 12. For comparing the models, primarily the coefficients of determination (R^2; Q^2) and the mean squared error (RMSE) values were considered.

Table 12: Major statistical parameters of the PLS regression models for green coffee. The units of measurement of the results for caffeine, 5-CQA, 4-CQA, 3-CQA, 3,5-diCQA, 4,5-diCQA, 3-FQA are mg/g.

parameter	pre-processing	samples	LV	cross-validation				test validation				range
				R^2	Q^2	RMSEC	RMSECV	R^2_t	Q^2_t	RMSEC	RMSEP	
pH	OSC	125	3	0,81	0,78	0,05	0,05	0,82	0,77	0,05	0,05	5,55-6
caffeine	OSC	138	2	0,94	0,89	0,23	0,32	0,95	0,81	0,22	0,30	8,32-13,6
5-CQA	OSC	107	2	0,99	0,98	0,52	0,86	0,99	0,98	0,60	0,70	33,54-61,98
4-CQA	OSC	135	3	0,97	0,90	0,48	0,91	0,98	0,86	0,43	0,91	5,52-16,92
3-CQA	OSC	130	2	0,90	0,86	0,75	0,91	0,91	0,87	0,73	0,84	3,27-13,80
3,5-diCQA	OSC	131	3	0,99	0,99	0,11	0,15	0,99	0,99	0,08	0,09	2,24-6,68
4,5-diCQA	OSC	126	4	0,99	0,99	0,06	0,08	0,99	0,99	0,06	0,05	1,71-4,34
3-FQA	OSC	133	3	0,99	0,99	0,10	0,13	0,99	0,96	0,09	0,12	0,95-4,86

All the analysed parameters were successfully modelled, with the best results in all cases after OSC data pre-processing. This technique also has a significant impact on model complexity (reducing the number of latent variables) and performance metrics.

Roasted coffee

Spectral and reference data from 138 samples were used to develop PLS models. In all cases the average values were used. The spectral range of 9000-3800 cm^{-1} was used for the evaluation, since the phase between 12500-9000 cm^{-1} is strongly influenced by the colour developed during roasting. The darker colour results in an increase of the measured absorbance, which obscures information about molecular vibrations.

In general, the results show that the statistical parameters of models were not improved by the iPLS variable selection compared to the full spectrum models. In all cases, the most efficient variable selection procedure was found to be GA, which uses all variables and their possible combinations for optimisation. The efficiency of this procedure can be explained by the complexity of the spectral data, which is due to the complexity of the molecular vibrations of coffee. Optimal results for PLS models were also obtained by using OSC.

It is important to highlight that, based on the scientific literature I have reviewed, there is no NIR model for predicting perceived acidity and bitterness that is suitable for testing coffees from different countries and with different roasting degrees.

Table 13: The main statistical parameters of the models obtained by PLS regression for roasted coffee. The units of measurement for the results for caffeine, 5-CQA, 4-CQA, 3-CQA, 3,5-diCQA, 4,5-diCQA, 3-FQA are mg/g.

parameter	pre-processing	sample	LV	cross-validation				test validation				range
				R^2	Q^2	RMSEC	RMSECV	R^2_t	Q^2_t	$RMSEC_t$	RMSEP	
pH	OSC	137	5	0,93	0,89	0,04	0,05	0,93	0,86	0,04	0,05	4,68-5,39
caffeine	OSC	135	3	0,90	0,82	0,30	0,41	0,90	0,83	0,31	0,38	9,90-15,43
5-CQA	OSC	130	2	0,90	0,89	1,46	1,52	0,90	0,90	1,51	1,29	3,67-24,96
4-CQA	OSC	127	4	0,88	0,84	0,80	0,95	0,89	0,83	0,77	0,97	3,43-14,71
3-CQA	OSC	127	4	0,87	0,81	0,72	0,89	0,87	0,86	0,70	0,83	2,40-11,58
3,5-diCQA	OSC	125	5	0,89	0,81	0,25	0,34	0,90	0,81	0,23	0,39	0,33-3,45
acidity	OSC	125	2	0,85	0,81	0,57	0,66	0,87	0,78	0,54	0,70	1,17-7,00
bitterness	OSC	130	3	0,89	0,83	0,53	0,68	0,91	0,82	0,48	0,69	1,30-9,30

Although both 4,5-diCQA and 3-FQA concentrations are greater than 0.1%, there is no clear direction of concentration change during roasting. As I mentioned earlier, the change can be described by a variety of different trends. Therefore, there is no clear relationship between the amount of compounds and the roasting level. This is also supported by the results of the analysis of variance, which shows that there is no significant difference between the concentration of 4,5-diCQA and the concentration of 3-FQA for medium and dark roasted samples, and no significant difference between light and medium roasted samples. The results suggest that for these compounds the relationship is not, or only with great difficulty, modelled by NIR spectral data. This may be supported by the fact that a PLS model with adequate statistical properties can be generated to estimate 4,5-diCQA, but this required the exclusion of several light and dark roasted samples from the evaluation. The results suggest that quantitative determination is possible for coffee samples with similar roasting degrees. In addition, it may be worth considering the use of non-linear regression methods in the modelling.

4. CONCLUSIONS AND RECOMMENDATIONS

In my PhD thesis, I performed a complex analysis of a so far little studied food matrix (snacks) by Fourier transform near-infrared spectroscopy (FT-NIRS). In addition, I investigated macromolecules and compounds that have been little or no studied before, thus extending the scientific literature.

I aimed to use chemometric methods for modelling, which have been rarely used for the analysis of spectral data, to demonstrate their potential.

Ready-to-eat foods such as snacks have become an integral part of our daily lives. The general chemical composition of snacks is a key parameter in product qualification, and is also important for the manufacturer and the authorities. The quantification of the macronutrients (carbohydrate, sugar, fat, protein, salt) is the basis for the nutrition labelling indicated on the packaging. Conventional analytical methods used in practice can take several days to complete, making it difficult to make a rapid decision. The results show that prediction models based on FT-NIR spectroscopy are suitable for the quantitative determination of macro-components, which allows a rapid and environmentally friendly determination of the average chemical composition of snacks.

Interval PLS regression was suitable for reducing model complexity in the prediction of the energy content of snacks, but it may be worthwhile to run additional variable selection methods to improve the performance parameters of the prediction model.

The snacks are characterised by a great variety of frying oils, raw materials, origins and technological processes and their combinations. However, NIRS coupled with appropriate chemometric modelling can be used to investigate the qualitative properties of snack products. The spectral data provide an opportunity to group the analysed products according to the above-mentioned characteristics using different classification methods. All three algorithms tested, PLS-DA, RF and neural network based RPropMLP, were able to classify snack samples correctly based on the four classification schemes. In addition, it is important to mention that the present research was the first to investigate the classification schemes based on the NIR spectra of snacks. However, testing further classification methods will provide a more comprehensive picture.

The quality of green coffee, and therefore its price, is greatly influenced by a number of physico-chemical and, through these, sensory characteristics, plant genetics and place of origin. In addition, information about the chemical composition of green coffee can help professional roasters to control the roasting process and thus adjust the sensory profile they want. The acidity, caffeine and chlorogenic acid content are also important parameters for the qualification of green coffee. The determination of the latter two is basically done after a relatively laborious sample preparation, using analytical methods that are both expensive and require specific knowledge. However, the developed models based on NIR spectroscopy can be easily and quickly applied directly to the non-destructive analysis of green coffee. In order to achieve optimal models, some data pre-processing is necessary and variable selection procedures are also recommended due to the complex molecular vibrations of coffee.

Spectral data are also often used for authentication purposes, which has been the subject of a number of scientific publications in the case of green coffee. Based on my experience, classification can be effective with a diverse data set, but there is no clear correlation between different growing regions and spectral properties, which makes the task difficult. The accuracy of the models is highest for binary systems (comparing two regions). In addition to the commonly used classification methods (LDA, PLS-DA, etc.), the use of several other algorithms (e.g., decision trees, ensemble methods, etc.) may be worth considering.

Based on the reference data (pH, caffeine content, 5-, 4-, 3-CQA, 3,5-diCQA, 4,5-diCQA, 3-FQA) of the analysed green coffees, I found no correlation or significant difference among the four production regions. These data suggest that a more extensive database is needed to identify differences due to the origin.

The effect of roasting on the amount of different chlorogenic acids is not fully understood even today, as it is a complex process. Of the compounds studied, 5-CQA shows a decrease in most cases as the roast develops. These changes can also be compared with the characteristic changes in the NIR spectral data. By combining the information, a clearer picture of the roasting status of coffee can be obtained. In contrast, for the other chlorogenic acids investigated, the effect of roasting is manifested through varying trends.

The quantitative assessment of caffeine in roasted coffee based on NIR spectra is complicated by the fact that while the concentration of caffeine does not

change significantly with roasting, there is a significant variation in the spectral data in the absorption ranges associated with caffeine. For successful modelling with PLS regression, derivation is an essential data pre-processing step, it is not sufficient to use conventional scatter correction operations (MSC, SNV). However, the best results are obtained with orthogonal signal correction (OSC), and therefore I suggest its use.

The results show that NIRS can be used for the quantitation of 5-, 4- and 3-CQA and 3,5-diCQA in roasted coffee. The change in concentration of 4,5-diCQA and 3-FQA during roasting has no clear direction and can be described by several different trends. Thus, there is no clear relationship between the concentration of compounds and the level of roasting. The results suggest that for these compounds the relationship is not or very difficult to model with NIR spectral data.

In general, for green and roasted coffee, the success of the regression model is largely determined by the data pre-treatment chosen and variable selection procedures play a key role in achieving optimal models for the objective. For most of the target parameters investigated, the best regression relationships are obtained using OSC, but the success of this approach is strongly influenced by the accuracy of the reference data. In addition, the use of iPLSR is not recommended for green and roasted coffee, and the VIP method is not suggested for green coffee.

5. NEW SCIENTIFIC RESULTS

1. I was the first to investigate snack products produced by different production technologies using FT-NIR spectroscopy, which included the mapping of spectral properties specific to the product group. I have found that FT-NIR spectroscopy combined with chemometric methods allows the evaluation of snacks produced with different production technologies using the same models. I have developed rapid methods for the quantitative determination of the macro components of snacks, which are suitable for the prediction of the carbohydrate, sugar, fat, protein and salt content of snacks using PLS regression. In addition, I used iPLS regression for the first time on snack matrix for the prediction of energy content.

2. Classification models with appropriate performance indicators were developed to classify snacks according to different aspects (cooking oil used, raw material, country of production, production technology), which were the first classification methods applied to the analysis of NIR spectral data of snack products.

3. It has been demonstrated that FT-NIR spectroscopy is suitable for the quantification of 5-, 4-, 3-CQA, 3,5-diCQA content in both green and roasted coffee beans, while the quantification of 4,5-diCQA and 3-FQA in green coffee was also successful. Based on the results of the analysed Arabica coffees, I found that there is no clear relationship between the amount of 4,5-diCQA and 3-FQA and the roasting level. While there was no significant difference in the concentrations of 4,5-diCQA and 3-FQA between light and medium roast samples, this was observed for the other chlorogenic acids studied.

4. I found a strong, significant positive correlation between the chlorogenic acidity of coffee and its perceived acidity, but a moderate but negative correlation between perceived bitterness.

5. Regression models were developed to predict the caffeine content and pH of green arabica bean coffees that show large variations in origin.

6. It was shown that the use of orthogonal signal correction (OSC) in PLS regression optimization greatly improves the performance of the prediction models. According to the results, I proved that the use of interval PLS regression and VIP variable selection does not significantly affect the performance of the models with respect to the parameters under study in case of coffee beans.

7. I have run previously not applied linear and non-linear chemometric methods (e.g., ensemble, Naïve Bayes algorithms, decision trees, etc.) on green coffee spectral data. I have found that the performance of the models is best in binary systems (Africa-America, Africa-Asia) using the dataset under evaluation.

PUBLICATIONS RELATED TO THE SUBJECT OF THE THESIS

IF journal articles

Eszter **Benes**; Attila Gere; Marietta Fodor (2020): Predicting macronutrients and energy content of snack products using FT-NIR analysis and chemometric techniques. JOURNAL OF FOOD ENGINEERING. SJR indicator: **D1**. IF-value: 5,354

Eszter **Benes**; Dávid Bajusz; Attila Gere; Marietta Fodor; Anita Rácz (2020): Comprehensive chemometric classification of snack products based on their near infrared spectra. LWT-FOOD SCIENCE AND TECHNOLOGY. SJR indicator: **D1**. IF-value: 4,952

Eszter **Benes**; Marietta Fodor; Sándor Kovács; Attila Gere (2020): Application of Detrended Fluctuation Analysis and Yield Stability Index to Evaluate Near Infrared Spectra of Green and Roasted Coffee Samples. PROCESSES. SJR indicator: SJR indicator: **Q2**. IF-value:2,847

John-Lewis Zinia Zaukuu; Eszter **Benes**; György Bázár; Zoltán Kovács; Marietta Fodor (2022): Agricultural Potentials of Molecular Spectroscopy and Advances for Food Authentication: An Overview. PROCESSES. SJR indicator: **Q2.** IF-value: 2,847

Brigitta Jesztl; Eszter **Benes**; Marietta Fodor (2019): FT-NIR origin identification of coffee samples. JOURNAL OF FOOD INVESTIGATION. SJR indicator: **Q4** IF-value:0,05

Zsóka Kárpáti; Eszter Benes; Marietta Fodor (2018): Nutritional analysis of coffee dregs for utilization purposes using classical, ICP-OES and FT-NIR techniques. JOURNAL OF FOOD INVESTIGATION.
SJR indicator: **Q4.** IF: 0,05

Hungarian language (full)

Benes, Eszter Luca; Fodor, Marietta. Különböző snack termékek beltartalmi értékeinek mennyiségi meghatározása FT-NIR spektroszkópia alkalmazásával. Magyar Tudományos Akadémia, Kertészeti és Élelmiszertudományi Bizottság, Élelmiszertudományi Albizottság Workshop (2018. december 6.).

International conference (full)

Eszter, Benes; Marietta, Fodor. Non-destructive measurement of carbohydrate content using NIR-spectroscopy the snack example. Third International Conference on Food Science and Technology konferenciaközlemény. 2018.

International conference (summary)

Eszter, Benes; Marietta, Fodor. Development of rapid, non-destructive method for quantitative determination of fat in different chips. II. YRICCCE: Program and Book of Abstracts p. 58 2018.

Eszter, Benes; Marietta, Fodor; Attila, Gere. Applicability of detrended fluctuation analysis (DFA) to near-infrared spectroscopic data. SZIEntific Meeting for Young Researchers. 2019

Eszter, Benes; Marietta, Fodor; Attila Gere: Prediction of coffee acidity and bitterness using sensory, near infrared spectroscopy and color measurement data. EuroSense 2020 - European Conference on Sensory and Consumer Research (Dublin, Írország)