ANALYSIS OF TEA LEAVES WITH DIFFERENT OXIDATION STATES BY FT-NIR SPECTROSCOPY

Eszter Benes[1, 2]*; Marietta Fodor[1]

Corresponding author: eszter.benes@gmail.com

[1] Department of Applied Chemistry, Faculty of Food Science, Szent Istvan University H-1118 Budapest, Villanyi Str. 29-31, Hungary

[2] Department of Postharvest and Sensory Evaluation, Faculty of Food Science, Szent Istvan University, H-1118 Budapest, Villanyi Str. 29-31, Hungary

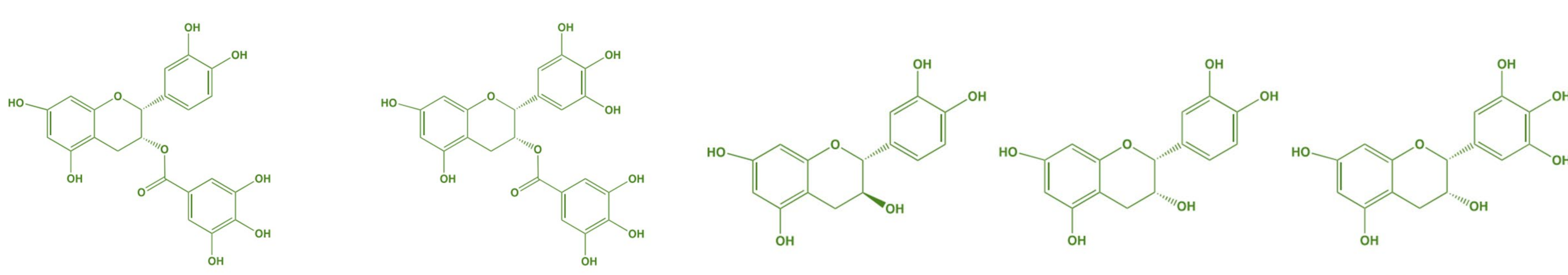
INTRODUCTION

Camellia sinensis **L**. is a perennial leafy crop. All varieties of tea are produced from the tea plants. The well-grown plants provide high-quality tea shoots, which vary with tea cultivars and the environmental conditions, such as the type of soil and altitude and climate of the growing area.

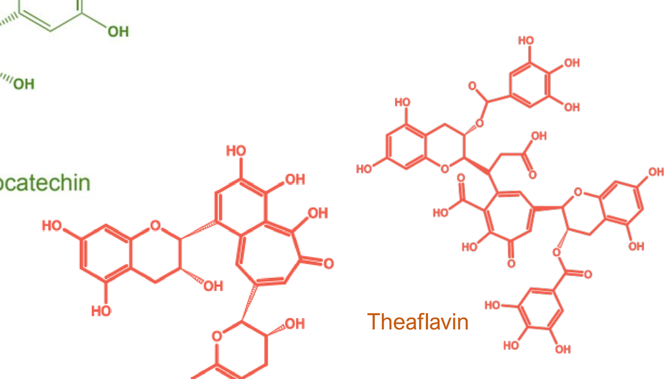
Moreover, tea quality is also determined by the processing techniques employed. For instance**, the same fresh tea leaves can be processed to black tea, oolong tea, and green tea by fermentation, semifermentation, and non-fermentation**, respectively. Those basic three types of tea have different quality characteristics, including color, aroma, taste, and appearance [1].

TANNINS are a particular class of natural products found in plants of different families and they are polyphenolics in nature. Tea contains many tannin substances in the leaves which affect the astringency and bitterness of the taste.

A large proportion of tannins in tea leaves consists of catechins, such as (-)-epicatechin, (-)-epicatechin-3-gallate, (-)-epigallocatechin, (-)-epigallocatechin-3-gallate and (+)-catechin [2].



In black teas, the major polyphenols are **theaflavin**, theaflavinic acids, proanthocyanidin polymers, and **thearubigin** or theasinensis, which are the oxidation products of polyphenols during processing usually known as catechins and gallic acid complexes [3] .



MATERIALS AND METHODS

For the study, 28 commercial tea samples were analyzed. We examined the leaves of fourteen black, four oolong, and ten green teas grown and produced in China, Nepal, Sri Lanka, India, Kenya, Taiwan, Japan and Korea.

28 different tea samples

determination of total tannin content with gravimetric method, according to the MSZ 20685-1980 Hungarian Standard

recording of the spectra by using Bruker MPA Multipurpose FT-NIR Analyzer (Bruker, Ettlingen, Germany)

- diffuse reflectance mode
- rotating cuvette
- resolution 16 cm^{-1}; scanning speed 10kHz

CHEMOMETRIC EVALUATION OF SPECTRAL DATA

- ✓ **Principal component analysis (PCA) – to screen spectral outliers**
- ✓ **Linear discriminant analysis (LDA) – classification by oxidation state**
- ✓ **Partial least squares regression (PLSR) – to create prediction model**

RESULTS AND DISCUSSION

Due to the diversity of samples, the **TANNIN CONTENT** varied in a wide range between **7.93 – 16.63 m/m%**. The lowest value was measured for sample 18 (oolong tea) and the highest value for sample 8, which was a black tea. Based on the results, no relationship can be found between the amount of tannins in the samples and its oxidation state.

Because **tannins are diverse class of compounds** with various chemical composition, the whole spectra can carry relevant information. Based on Chen et al. (2006), the vibration of the 2nd overtone of the carbonyl group (5352 cm^{-1}), the C–H stretch and C–H deformation vibration (7212 cm^{-1}), the $–CH_2$ (5742 cm^{-1}), and the $–CH_3$ overtone (5808 cm^{-1}) can be observed.

The absorption bands of catechins mainly involved:

- ❑ C–H second overtone (9090-8547 cm^{-1}),
- ❑ C–H first overtone and S–H first overtone (6257-5691 cm^{-1}),
- ❑ and C–H and C–H combinations paired with C–H and C–C combinations (4559-4000 cm^{-1})

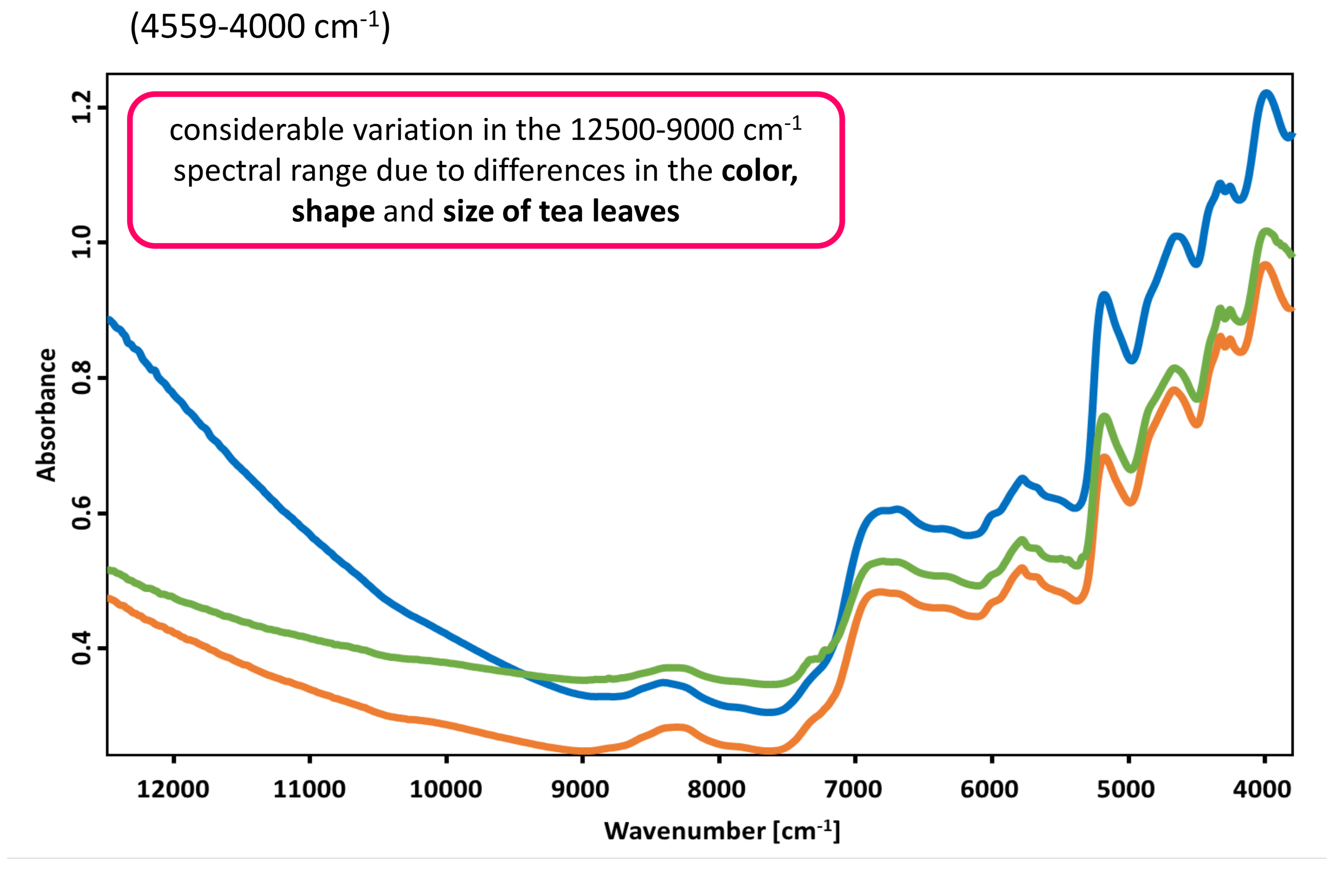


Figure 1. Spectra of tea obtained from raw data (●) black tea; (●) oolong tea; (●) green tea

PCA determines, how many PCs describe the data set and its variance. By using ten PCs, it was found that the variance of the properties is determined in 98% by the first three PCs (PC1: 79%, PC2: 19%, PC3: 1%). To detect the real spectral outliers, F-residuals and Hotelling's T^2 statistics were applied.

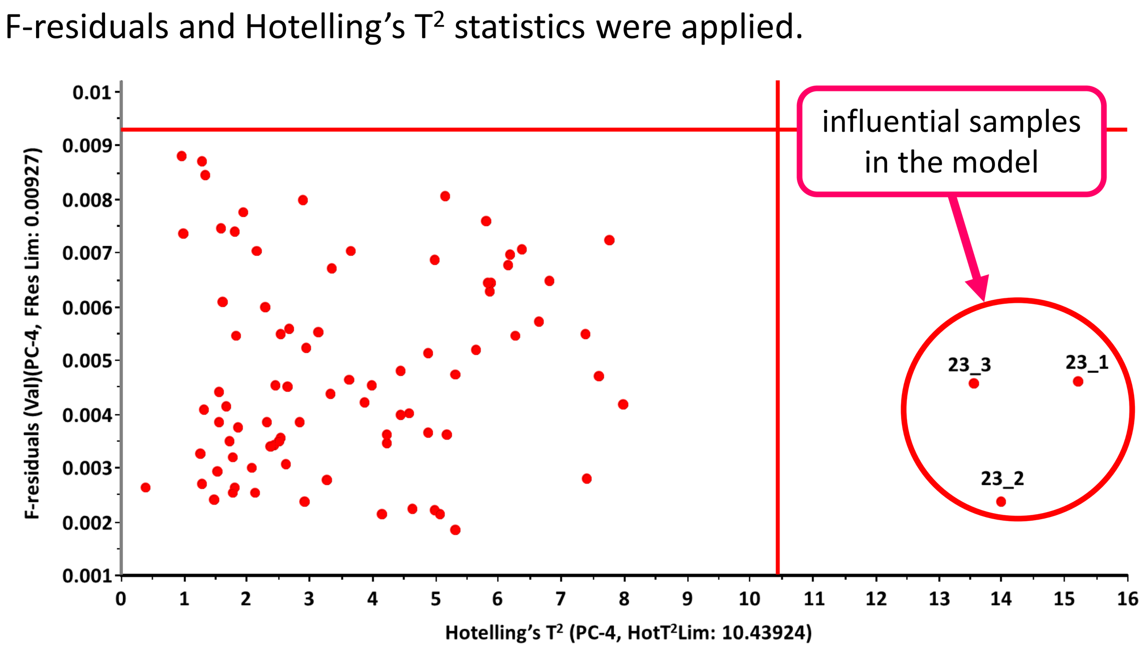


Figure 2. Influence plot of principal component analysis

For the classification of green and black tea samples, **LDA** was used in the 12500-3800 cm^{-1} spectral range. The results of the LDA classification are the predicted class for each sample. According to the confusion matrix (Table 1.), the two groups can be clearly separated. The confusion matrix carries information about the predicted and actual classifications of samples, with each row showing the instances in a predicted class, and each column representing the instances in an actual class. The accuracy of the method was 100% by using 10 components (Figure 3.).

	black	green
black	14	0
green	0	10

Table 1. Confusion matrix of LDA

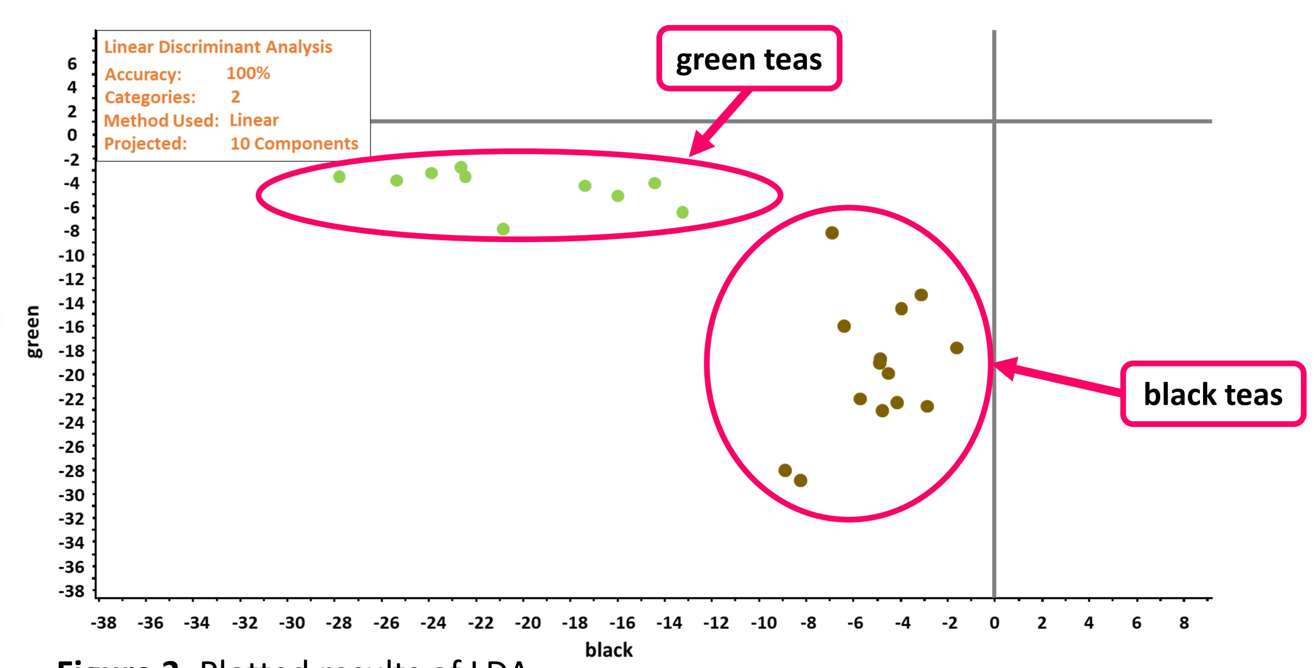


Figure 3. Plotted results of LDA

PLS regression was performed to predict the total tannin content in tea leaves. For the analysis, the three measured reference values and three spectra were used for each sample. After the first run, it was found that sample 23 really had a negative impact on the regression model. Therefore, it was threated as an outlier. Leave-one-out cross-validation was applied to validate determine the optimal number of latent variables. After using different spectrum pre-processing techniques, the best result was obtained by MSC transformation (Table 2.).

In most applications, light scattering often creates an ''absorbance shift'' in the spectrum, which can be more or less irrelevant for the chemical composition. MSC is usually used to eliminate this noise from light scattering [5]. In case of tea leaves, this technique can reduce the effect of light scattering caused by the different size and shape of tea leaves.

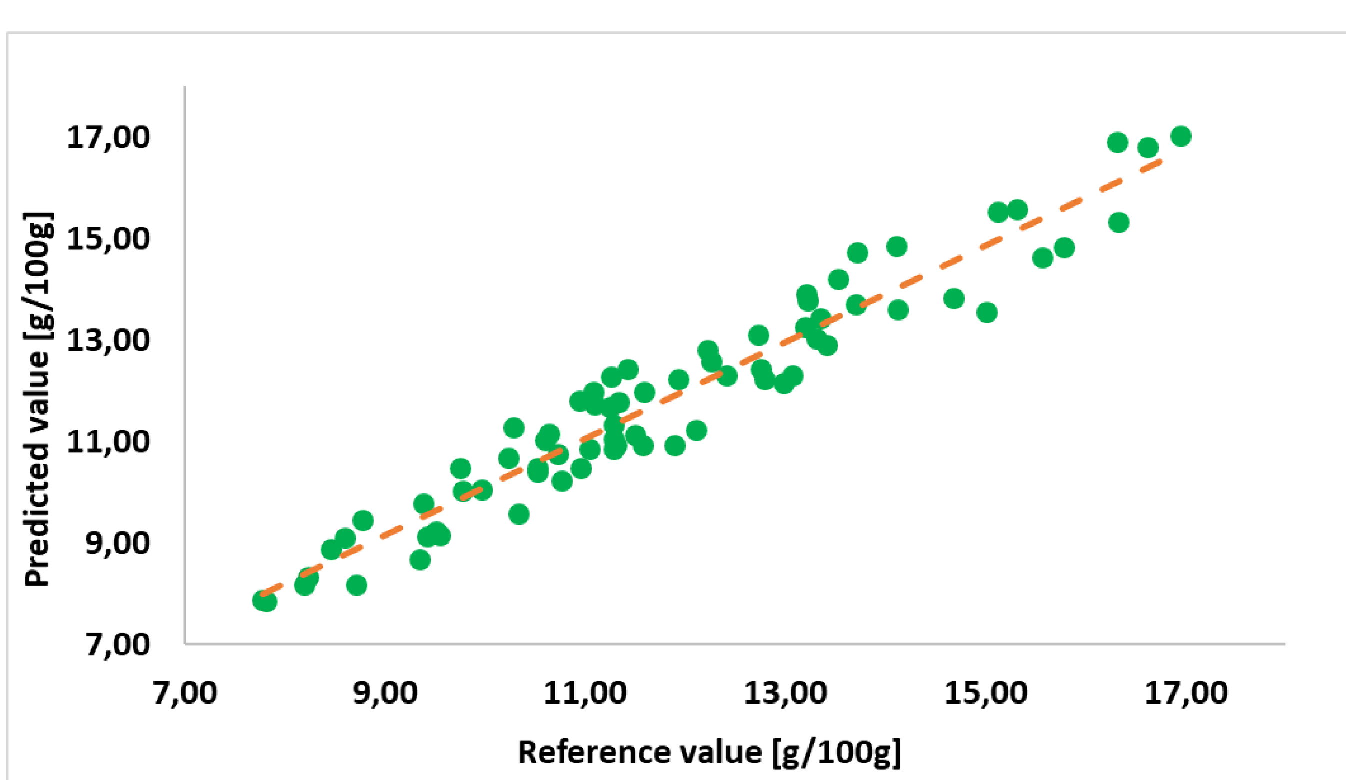


Figure 4. Validated prediction model to quantify total tannin content in tea leaves

73 spectra of 27 samples were used to the prediction model

measurement range: 7.78-16.95 m/m%

The validated model has

- high predictive ability (Q^2=93.09)
- and low root mean square error (RMSECV=0.585 m/m%)

the **evaluation ranges** of the prediction model correspond with most of the specific absorption bands of tannins

Table 2. Statistical parameters of the prediction model

Calibration			Cross-validation			PLS	Evaluation range [cm^{-1}]
R^2	RMSEE [m/m%]	RPD	Q^2	RMSECV [m/m%]	RPD		
95.58	0.504	4.76	93.09	0.585	3.81	9	7506-6094, 4605-4242

Acknowledgement The Project was supported by the ÚNKP-20-3 New National Excellence Program of the Ministry of Human Capacities and by the National Research, Development and Innovation Fund. The Project was also supported by the Doctoral School of Food Science SZIU.

References
1. *Tea and Tea Products Chemistry and Health-Promoting Properties*; Ho, C.-T., Lin, J.-K., Shahidi, F., Eds.; CRC Press, 2009; ISBN 9780849380822.
2. Takeda, Y. Differences in caffeine and tannin contents between tea cultivars, and application to tea breeding. *Japan Agricultural Research Quarterly* **1994**, *28*, 117–123.
3. Juneja, L.R.; Kapoor, M.P.; Okubo, T.; Rao, T.P. *Green Tea Polyphenols - Nutraceuticals of modern life*; CRC Press, 2013; ISBN 9781138199378.
4. Chen, Q.; Zhao, J.; Zhang, H.; Wang, X. Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration. *Analytica Chimica Acta* **2006**, *572*, 77–84, doi:10.1016/j.aca.2006.05.007.
5. Jie, D.; Xie, L.; Fu, X.; Rao, X.; Ying, Y. Variable selection for partial least squares analysis of soluble solids content in watermelon using near-infrared diffuse transmission technique.