Title: Feasibility study to estimate the TDS of coffee from NIR spectroscopic measurement of roasted coffee beans and brew

Abstract

Coffee is the most traded plant commodity in the world. The main drivers behind this popularity are consumers who are drawn to its unique sensory attributes and cognitive effects. The strength of the coffee brew is a principal sensory quality attribute that is quantified as TDS. Traditional methods of assessing TDS require the preparation of coffee brew, which is subject to many variations and often leads to inconsistencies across different methods. Secondary analytical techniques such as NIR spectroscopy offer the advantage of minimal sample preparation and routine analysis. The method has seen extensive applications in the coffee industry.

In this study, we set out to estimate the TDS of coffee from the NIR spectra of roasted coffee beans and brew. Fifty original green coffee beans (48 Coffea arabica and 2 Coffea canephora) coffee samples were sourced from four regions: South America, Asia, Central America, and Africa. The samples were roasted with a predefined protocol (light: 40 seconds, medium: 60 seconds, and dark: 90). This generated 150 samples, which were ground (20g) to a median particle size of 600µm. The ground coffee samples were then brewed with a V60 dripper using a predefined protocol. The reference TDS of the coffee samples was then determined using oven-drying: TDS-OD and a Brix°-calibrated refractometer with a correction factor: TDS-BC.

The relation between the TDS-OD and Brix° was tested with a linear regression model which resulted in R^2 and R of 0.90 and 0.95 respectively. Furthermore, Repeated Measures Analysis of Variance (ANOVA) revealed no significant differences in the mean TDS of both reference methods within the three roasting levels; TDS-OD ($F(2;74) = 2.01$, $p = 0.14$); and TDS-BC ($F(2;74) = 3.11$, $p = 0.05$). The NIR spectra of the roasted beans as well as the brew were collected in the range of 12,500 - 3,800 cm^{-1}. The obtained spectra were averaged and pre-processed with multiple scatter correction (MSC), and Savitzky-Golay second derivative.

A visual inspection of both brew and bean spectra grouped by a row set of TDS ranges (1.3-1.5, 1.5-1.7, 1.7-1.9, 1.9-2.1) for both TDS-OD and TDS-BC did not reveal any significant patterns. Further exploratory assessment by Principal Component Analysis (PCA) of the processed

spectra resulted in an optimum of 4 PCs for the processed bean spectra (PC1: 55%, PC2: 23%, PC3: 14%, PC4: 4%) and 5 PCs for the brew spectra (PC1: 66%, PC2: 12%, PC3: 9%, PC4: 5%, PC5: 3%). The scores plot from the PCA for both bean and brew spectra did not reveal significant clusters; hence, we could not draw a meaningful conclusion from the correlations loading plot.

Further assessment with Partial Least Square Regression revealed that the calibration models for the transformed bean spectra were not significant for both reference methods; TDS-BC (R^2/ Q^2: 0.12/0.09, RMSEC :0.11%, RMSECV: 0.09%) and TDS-OD (R^2/ Q^2 : 0.16/0.06, RMSEC :0.17%, RMSECV: 0.06%). The results from calibration models for the transformed brew spectra were also not significant; TDS-BC (R^2/ Q^2 : 0.04/0, RMSEC :0.14%, RMSECV: 0.15%) and TDS-OD (R^2/ Q^2 : 0.03/0, RMSEC :0.14%, RMSECV: 0.15%). Due to the lack of robustness of the models, we did not perform a prediction with test sample set. Thus, the results show that estimating TDS from the NIR spectra with the given parameters and experimental design is not ideal.