Quality assessment of fresh tea leaves by estimating total polyphenols using near infrared spectroscopy

Ajanto Kumar Hazarika¹ · Somdeb Chanda² · Santanu Sabhapondit¹ · Sandip Sanyal¹ · Pradip Tamuly¹ · Sahnaz Tasrin¹ · Dilip Sing² · Bipan Tudu² · Rajib Bandyopadhyay^{2,3}

Revised: 11 July 2018 / Accepted: 5 September 2018 / Published online: 17 September 2018
© Association of Food Scientists & Technologists (India) 2018

Abstract This paper reports on the development of an integrated leaf quality inspecting system using near infrared reflectance (NIR) spectroscopy for quick and in situ estimation of total polyphenol (TP) content of fresh tea leaves, which is the most important quality indicator of tea. The integrated system consists of a heating system to dry the fresh tea leaves to the level of 3–4% moisture, a grinding and sieving system fitted with a 250 micron mesh sieve to make fine powder from the dried leaf. Samples thus prepared are transferred to the NIR beam and TP is measured instantaneously. The wavelength region, the number of partial least squares (PLS) component and the choice of preprocessing methods are optimized simultaneously by leave-one-sample out cross-validation during the model calibration. In order to measure polyphenol percentage in situ, the regression model is developed using PLS regression algorithm on NIR spectra of fifty-five samples. The efficacy of the model developed is evaluated by the root mean square error of cross-validation, root mean square error of prediction and correlation coefficient (R^2) which are obtained as 0.1722, 0.5162 and 0.95, respectively.

Keywords Polyphenol · Fine leaf count · Near infrared reflectance (NIR) spectroscopy · Folin Ciocalteu method · Partial least squares (PLS) · Preprocessing · Root mean square error of cross-validation (RMSECV)

Introduction

Tea, derived from young vegetative shoots of the tea plant, *Camellia sinensis* (L.), is one of the most popular non-alcoholic beverages in the world. It is of great interest to common man due to its numerous therapeutic properties (Cooper and Morre 2005; Bhuyan et al. 2013; Astill et al. 2001). Recent research suggests that antioxidants found in tea may play an important role to prevent many health related disorders (Gupta et al. 2002). Studies show that these properties are related to the antioxidant activity of tea polyphenols (Zhang et al. 2004).

Polyphenolic compounds mainly comprise of catechins (Flavan-3-ols) and their gallates and flavonol glycosides (Harbowy and Balentine 1997) [chemical structures are presented in supplementary Fig. S1(a, b)], have long been considered as one of the primary quality indicators of tea (Kumar et al. 2011; Li et al. 2015).

Tea leaves contain 180–360 mg/g of polyphenols, among which 70–80% are flavanols (Sabhapondit et al. 2011), representing a significant proportion of the tea constituents. The other important chemical constituents in tea are caffeine, amino acids, protein, polysaccharides, cellulose, lignin, lipids and ash (Pereira et al. 2014). In black tea processing polyphenols, more precisely the catechins undergo enzyme catalyzed oxidative reactions to form theaflavins (TFs), thearubigins (TRs) (Sanderson 1972) and other polymerization products. TFs and TRs, the important quality attributes of black tea, and unoxidized

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13197-018-3421-6>) contains supplementary material, which is available to authorized users.

✉ Somdeb Chanda
somdeb04@rediffmail.com

¹ Tocklai Tea Research Institute, Jorhat, Assam, India

² Department of Instrumentation and Electronics Engineering, Jadavpur University, Salt Lake Campus, Block LB, Sector III, Plot 8, Salt Lake, Kolkata 700 098, India

³ Laboratory of Artificial Sensory Systems, ITMO University, Saint Petersburg, Russia

polyphenols contribute to the distinctive sensory properties like mouth feel and astringency, brightness and colour of a tea brew (Zhang et al. 1992). Polyphenolic compounds present in tea plant decrease with leaf maturity (Obanda and Owuor 1995; Li et al. 2016). Thus, the harvesting or plucking standards of tea shoots has a direct implication on the final cup quality and market price (Forrest and Bendall 1969). The general recommendation for harvesting fresh leaves to manufacture good quality tea is two leaves and the apical bud (Eden 1976), irrespective of the type of cultivar (Mahanta 1988). To maintain this standard, fresh tea shoots are to be harvested in cycles of five to seven days round. But exclusive maintenance of such plucking standard is difficult. To ensure quality, inbound harvested leaves to a tea factory are physically examined from a random bulk. The technique is popularly known as ‘fine leaf count’, wherein, individual undamaged shoots (one leaf plus bud and two leaves plus bud) are physically counted and weighed (Sanyal 2011). This *modus operandi* is carried on a daily basis for each batch of tea leaves arriving at a tea factory. The exercise is subjective, laborious and may be influenced by external factors.

Another vital issue that plays a big role in the quality of tea leaves is “bought leaf”. “Bought leaf” are the leaf collected from the plantations of small tea growers (STG) in homestead garden and unutilized land along with other crops. STG contribute 33.85% of the total tea production in India (Rakshit 2016); and their contribution has been projected to increase further in the coming years. The tea industry has to adopt “bought leaf” paradigm to improve efficiency and profitability of tea.

Apart from the harvesting standard, the factors like distance of transit from field to factory and handling of the tender shoots affect the quality of the inbound tea leaves. Any mechanical stress prior to the actual processing of the tea leaves will trigger untimely and undesired enzyme catalyzed oxidative reactions involving catechins as substrates. Thus in situ scrutiny of the harvested green leaf, as they arrive at a factory is essential to maintain conformity in quality and hence the price of end product. This has never been correctly predictable in case of tea. On the retro-respect, chemical estimation of total polyphenols content in tea requires laborious and time consuming sample preparation involving enzyme deactivation, drying and grinding and skilled manpower in chemical analysis. An alternative, yet effective methodology for determination of total polyphenol (TP) content, in situ, will immensely help the tea industry in instant assessment of fresh leaf quality.

Near infrared (NIR) spectroscopy has been extensively used for its quick and non-destructive technique for qualitative and quantitative estimation of many products (Jha and Ruchi 2010; Jha and Matsuoka 2004). It has decisive

advantages compared to traditional methods, wherein samples can be rapidly analyzed. It is chemical-free (limited to the reagents required for reference analyses) and has no resultant waste (Ozaki et al. 2006; Pissard et al. 2012). The technique has host of applications, starting from agriculture, petrochemicals, polymers, biomedical, inorganic functional materials, process analysis, imaging, pharmaceutical industry, etc. to name a few (Roberts et al. 2004; Kawasaki et al. 2008); and also for on-line applications (Huang et al. 2008).

Feasibility studies on the application of NIR for qualitative and quantitative analysis of intrinsic attributes, like polyphenols, catechins, gallic acid, caffeine and theobromine in tea leaves have been reported (Luypaert et al. 2003; Chen et al. 2006, 2009). Researchers have successfully utilized NIR spectroscopy for tea variety discrimination (He et al. 2007), dry matter content in tea (Li et al. 2013), caffeine (Sinija and Mishra 2009), theaflavin and moisture content in black tea (Hall et al. 1988), simultaneous analysis of alkaloids and phenolic substances in green tea leaves (Schulz et al. 1999) and also the market price of tea (Yen 2007).

Most of the publications have focused on the application feasibility of NIR for qualitative and quantitative estimations of tea constituents. A meaningful and consistent NIR spectroscopic methodology will only be acceptable to the tea industry if it can be applied for in situ estimations of tea leaves.

Another aspect that came to light during the course of the study was that, the process of enzyme deactivation of tea leaves, followed by subsequent drying, grinding to proper size was tedious and time consuming. This is contrary to the general belief that NIR spectral analysis requires minimum or no sample preparation. This led to the development of an integrated system which could drastically reduce the sample preparation time. The absorbance spectra of the powdered samples were taken in the near infrared (NIR) wavelength range of 900–1701.5 nm. The optimum range for total polyphenol was determined to be from 1460 to 1624.50 nm (Bedini et al. 2013; Bian et al. 2010). In order to calibrate the regression model of NIR tea spectra, partial least squares (PLS) algorithm was used. The efficacy of the model developed was evaluated by the root mean square error of cross-validation (RMSECV), root mean square error of prediction (RMSEP) and correlation coefficient (R^2) (Ren et al. 2013; Chen et al. 2006).

A preliminary report based on the development of the instrument for sample preparation and measurement has been communicated to Journal of NIR News (Hazarika et al. 2018) and it will be available shortly.

In this paper, we report on the development of an NIR based integrated tea leaf inspection system for onsite scrutiny of the inbound green leaf. The system carries out

enzyme deactivation, drying followed by grinding and sieving simultaneously, prior to NIR estimation of total polyphenol content. The machine, together with the NIR unit, could be easily set up at the green leaf unloading bay of a tea factory as in situ quality inspection and estimation tool. The flowchart of the whole process followed in this paper is provided in supplementary Fig. S2.

Materials and methods

Sample collection and preparation

Trials were carried out with tea leaves collected from the experimental gardens of Tocklai Tea Research Institute (TTRI), Jorhat, Assam, India and Small Tea Growers' plantations covering the full harvesting period from March to November, 2016. Green leaves were selected from cultivars from the three tea varieties, viz. Assam (*Camellia assamica*), China (*Camellia sinensis*) and Cambod (*Camellia assamica* spp. *lasiocalyx*) and their intermediates. Samples were harvested in 7 days interval maintaining 65–70% fine leaf. Initially, deactivation of enzymes in fresh leaves was carried out in a microwave oven (LG model no. MS-2342AE, Input 230 V–50 Hz, 1200 W, RF output 800 W, Frequency 2450 MHz) at radio frequency output of 800 W for 2 min, cooled and dried in the conventional endless chain pressure (ECP) tea drier at the Model Tea Factory, TTRI. The dried leaves were then grinded and sieved through 250 micron mesh and immediately packed and sealed in poly-pouches for further analysis.

Temperature and relative humidity

All experimentations were carried out at Model Tea Factory, TTRI, Jorhat (26.7465°N, 94.2026°E, Elevation: 96 m.a.s.l.) within the full harvesting period (i.e. March to November) of 2016. Average maximum and minimum temperatures for the period varied from 27 to 32 °C and 14 to 25 °C respectively; while average relative humidity during day time ranged from 55 to 70%.

Estimation of total polyphenol

Extraction: 200 mg of finely ground sample was extracted with 5 mL 70% methanol in a water bath set at 70 °C over 10 min. The extract was then cooled and centrifuged at 4000 turns for 10 min. The supernatant was decanted into a 10 mL volumetric flask. The extraction was repeated twice and volume was made up with the extraction solvent.

Total polyphenols in the samples were estimated using Folin Ciocalteu's phenol reagent by following the method

of ISO 14502-1 (ISO, 2005). This standard is an analytical tool widely used to determine the phenolic compounds for both black and green teas. The gallic acid mono hydrate is recommended and widely used as a calibration standard due to its satisfactory solubility, adequate stability and low price. The results are usually reported as Gallic acid equivalent (GAE). Folin Ciocalteu reagent analysis is presented in supplementary Fig. S3.

Development of the deactivation-drying cum grind-sieve machine

The main mechanical part consists of two co-centric stainless steel drums with plate thickness of 2 mm gauge and diameters of 11 inches and 14 inches. The outer drum is insulated with polyurethane foam and is mounted on a steel frame. The inner drum, fitted via a motor with sprocket and chain assembly system, can be rotated at a variable speed (0–50 rpm) and is heated externally by an array of LPG burners connected to a central LPG cylinder.

De-activation and drying

The enzymes polyphenol oxidase (PPO) and peroxidase (POD) in fresh tea leaves were first deactivated around 100 °C leaf temperature and simultaneously dried in the machine. During the process, the internal drum was regulated at 12 and 16 rotations per min for carrying out deactivation and drying respectively. This allowed preservation of the chemical composition of green leaf, while the moisture level of the leaves was brought down to around 3%. Depending on the 'fineness' of plucked tea shoots, the process of deactivation and drying of tea leaves were achieved in 15–18 min. A tilting arrangement of the drum drops the dried tea leaves directly into the hopper of an automatic burr mill grinder.

Grinding and sieving

Dried tea leaves thus produced were grinded in a Cuisinart Supreme Grind Automatic Burr Mill fully automatic coffee grinder (Model no. DBM-8, 120 V, 60 Hz, 140 W). To suit our requirement to grind the dried tea leaves to correct particle size, the grind regulator was set at extreme fine setting. The ground powder was directly conveyed through a vibrating test sieve of 250 micron stainless steel mesh (Brand: SISCO India, B.S.S Standards-60, series: SSIC-90). The test sieve mounted on a stainless steel frame was made to vibrate by a shaft fitted to a 24 V DC motor. The reciprocating motion of the shaft imparts an eccentric movement to the sieve attached to it. A speed controller attached to the motor controls the vibration. The homogenized powdered samples were then directly conveyed

through a funnel to a small cylinder mounted on a quartz plate of the NIR detector. Schematic of fixing-drying-grinding cum sieving machine is shown in Fig. 1 and snap shot is shown in supplementary Fig. S4.

NIR setup

The NIR beam was applied to the powdered sample in the cylinder and the reflectance spectra were collected from the bottom of the quartz plate using fiber optic cables. Absorbance spectra of the powdered samples were taken in the near infrared (NIR) wavelength range of 900–1700 nm using the DWARF-Star NIR spectrometer from StellarNet Inc. USA.

The entire process, starting from enzyme deactivation of fresh leaves to NIR data acquisition and prediction was completed within 18–20 min.

2 g (2.0 ± 0.1) of finely sieved tea samples were placed in a standard quartz plate of 1 mm thickness to maintain the correct bulk density of materials. Spectral scan of each sample was collected five times and average of the five spectra collected from a sample was used for calibration. Operation of the instrument and data collection of NIR spectra were conducted by using the NIR spectrometer.

NIR spectra of tea leaves were collected in a reflectance mode. The wavelength range of NIR spectrometer is from 900 to 1701.50 nm. The data points of spectra were acquired in 1.75 nm intervals. The spectrometer produced 459 spectral points for each sample. 15 replicates were taken for each sample; thus in total $459 \times 55 \times 15$ spectral data points were obtained for 55 samples. The spectra were acquired with scan average 16 and integration time of

300 ms. This set up was made compatible with the online prediction set up shown in supplementary Fig. S4.

Pre-processing methods

In this study, several standard preprocessing techniques were utilized. Out of which, two data preprocessing techniques produce satisfactory results.

- Multiplicative scatter correction (MSC).
- Standard normal variate (SNV).

The tea powder sample may contain different size of particles and this causes scattering of light. Scattering effect may play a role for variability of the spectra from sample to sample. Multiplicative scatter correction (MSC) and standard normal variate (SNV) transformation are two different spectral preprocessing methods that eliminate the scattering effects. Figure 2a presents the raw spectral profile of tea leaves, the spectra by SNV preprocessing and the spectra by MSC preprocessing are presented in Fig. 2b, c, respectively. 1st Derivative and 2nd Derivative spectra are shown in Fig. 2d, e, respectively. On each figure the approximate absorbance wavelengths of some common functional groups between 900 and 1700 nm are marked with different colors and legends are mentioned after Fig. 2e.

Legands:

- 1143, 1170–1225 C–H second overtone.
- 1360–1395, 1415–1417, 1440–1446 C–H combination.
- 1410, 1420, 1450, 1490, 1540, O–H first overtone.
- 1471–1490, 1492–1530, 1570 N–H stretch first overtone.

Fig. 1 Schematic of fixing-drying-grinding cum sieving machine

Fig. 2 **a** Graphical representation of the near infrared wavelength versus absorbance of the original data without any pre-processing. **b** Graphical representation of the near infrared wavelength versus absorbance of the data pre-processed using standard normal variate (SNV). **c** Graphical representation of the near infrared wavelength versus absorbance of the data pre-processed using multiplicative scatter correction (MSC). **d** Graphical representation of the near infrared wavelength versus absorbance of the data after 1st Derivative. **e** Graphical representation of the near infrared wavelength versus absorbance of the data after 2nd Derivative

Fig. 2 continued

To develop a model by best preprocessing technique, results of above mentioned four preprocessing techniques were compared.

Data analysis method

Partial least squares (PLS) regression is a statistical method that bears some relation to principal components regression (Bian et al. 2010). It finds a linear regression model by projecting the predicted variables (Y) and the observable variables (X) to a new space. PLS is used to find the fundamental relations between two matrices (X and Y) using a latent variable approach to build a model of the covariance structures in these two spaces.

Here, in our calibration and validation models, PLS regression has been used to predict the reference polyphenol content values of the testing set. The entire

flow chart of data analysis method is shown in supplementary Fig. S5.

G. software

SpectraWiz, Spectrometer OS v5.3, 2013 software was used for NIR spectra acquisition. Preprocessing techniques and PLS regression algorithms were implemented in Matlab V10.0 (Mathworks Co., USA).

Results and discussion

Spectra investigation

Scrutinizing the spectra of the original data (Fig. 2a), showed absorbance mainly in the region of 1460–1620 nm (Bedini et al. 2013). The absorbances at different

wavelengths between 1410 and 1540 are corresponding to O–H first overtone. As spectral acquisitions were taken from dry powdered tea, the noticeable NIR spectral absorbance (Fig. 2b, c) are due to O–H first overtone of polyphenols in tea leaves; and therefore, selected for calibration of PLS model.

From the above figures (Fig. 2d, e), it is comprehensible that 1st and 2nd derivative do not represent any distinct peak in the region of 1460–1624.5 nm. Thus, these two derivative techniques did not reasonably provide any promising results. So, SNV and MSC were considered for further data analysis. Thus for each spectrum, we have considered only 94 spectral points in this analysis. The data points in the spectra were acquired in 1.75 nm intervals.

Calibration of model

The performance of a PLS algorithm depends significantly on PLS factors and the spectral pre-processing techniques. The data obtained from the spectrometer were processed using two pre-processing techniques as mentioned in the earlier section. Pre-processed data were subjected to PLS regression algorithm to built the calibration model. During the model calibrations following steps were followed:

- (i) Comparing the result with different pre-processing techniques.
- (ii) Performance evaluation of PLS by RMSECV results of LOSOCV.
- (iii) Finalize the best technique for processing of NIR spectral data.
- (iv) Finalize the PLS component by comparing the maximum accuracy or least root mean square error of cross validation (RMSECV).
- (v) Finally model building and prediction of Test samples.

The performance of the final PLS model was evaluated in terms of root mean square error of cross-validation (RMSECV), the root mean square error of prediction (RMSEP) and the correlation coefficient (R) (Chen et al. 2006). For RMSECV, a leave-one-sample-out cross-validation was performed: the spectrum of one sample of the training set was removed and a PLS model was built with the remaining spectra of the training set (Chen et al. 2006). The left-out sample was predicted with this model and the procedure was repeated by leaving out each of the samples of the training set. The RMSECV is calculated as follows:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

where, y_i is the chemically measurement result for the test sample i , $\hat{y}_i$ is the estimated result of the sample i and n is

the number of samples in the calibration set, when the model is constructed with the sample i removed. This procedure was repeated (for each of the preprocessed spectra). According to the lowest RMSECV, the number of PLS factors was chosen to develop the model.

In the prediction set, the root mean square error of prediction (RMSEP) (Chen et al. 2006) was calculated.

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_{pi})^2}{n}} \quad (2)$$

where, y_i is the reference measurement result for the test sample i , $\hat{y}_{pi}$ is the estimated result of the model for test sample i and n is the number of samples in the test set.

Finally, the optimal model was chosen according to the overall lowest RMSECV and prediction accuracy.

Correlation coefficients between the predicted and the measured values were calculated for both the calibration and the prediction sets, using Eq. (3), where, $\hat{y}_i$ is the predicted value for sample i , y_i is the reference measurement result for the sample i , $\bar{y}_p$ is the mean of the predicted values and $\bar{y}$ is the mean of the reference measurement results (Chen et al. 2006).

$$R = \sqrt{1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_p - \bar{y})^2}} \quad (3)$$

Results of calibration model

The number of PLS factors were optimized by determining the lowest root mean square error of cross-validation (RMSECV) by two derivative techniques (i.e. 1st Derivative and 2nd Derivative) and two different pre-processing methods, SNV and MSC.

The comparison of sample-one-out cross validation results of raw and four preprocessed spectra for 55 samples within 900–1701.5 nm and 1460–1624.5 nm are exposed in supplementary figure (Fig. S6a, b).

In comparison with raw and other four preprocessing techniques (i.e. SNV, MSC, 1st Derivative and 2nd Derivative), best cross validation result was obtained by SNV preprocessing method, within the range of 900–1701.5 nm wavelength regions. The optimized component no and lowest RMSECV is 7 and 0.2845 respectively. Let it be termed as result 1.

It was observed that the best preprocessing method within the range of 1460–1624.5 nm wavelength regions was MSC, and optimized component no and the lowest RMSECV was 8 and 0.1722 respectively. It be termed as result 2.

To develop a model by best preprocessing technique, result 1 and result 2 are compared in Table 1.

Table 1 Comparison of best preprocessing techniques

Wave length region	Best data processing technique	Optimum component no.	RMSECV
1460–1624.5 nm	MSC	8	0.1722
900–1701.5 nm	SNV	7	0.2845

In Table 1, comparison between sample-one-out cross validation results for 55 samples of two separate wavelength regions shows that the performance of MSC preprocessing technique in the wavelength region of 1460–1624.5 nm is better than SNV preprocessing technique in the wavelength region of 900–1701.5 nm wavelength region.

So, finally the selected technique for model development is MSC. The corresponding component no and wavelength region is 8 and 1460–1624.5 nm, respectively. The score plot and loading plot of finally selected model at component no 8 are shown in supplementary figures, i.e. Figs. S7 and S8, respectively.

The model was developed using 55 samples and 10 samples were for validation of the model. The calibration curve is shown in Fig. 3. Thus, this calibration model could also be used for estimation of total polyphenol values of unknown tea samples. The test results of 10 samples are shown in Table 2.

Table 2, shows the average accuracy of prediction as 91.70%. It was also observed that the root mean square error of prediction (RMSEP) of 10 samples with component number 8 was sufficiently low i.e. 0.5162 and correlation coefficient (R^2) was 0.925 which exhibits good correlation between training and testing set. The Average RPD was obtained as 8.845.

Conclusion

This paper describes an integrated green leaf quality inspecting system, based on the measurement of TP using diffuse reflectance near infrared spectroscopy. The developed integrated system drastically reduces the sample preparation time. The machine can be easily set up at the green leaf unloading bay together with the NIR unit for in situ quality estimation. The system carries out enzyme deactivation, drying followed by grinding and sieving simultaneously, prior to NIR estimation.

The nature of the data obtained from a particular sample depends on the measurement configuration and experimental set up. In this paper, the most common measurement procedure i.e. reflectance mode has been used for data collection. Since total polyphenol consists of a group of phenolic ingredients like catechin, theaflavin etc., hence the spectra carry the information about these groups of chemicals. Here, the performance of the PLS model was analyzed after preprocessing the data with well-known preprocessing techniques. It has been noted that the best model was obtained when the data has been preprocessed using MSC techniques with average accuracy of prediction of 91.70%. The RMSEP and correlation coefficient (R^2) was as 0.5162 and 0.925 respectively which demonstrates the efficacy of the model.

Thus this method can be suitably applied for in situ quality detection of inward tea leaves in tea factories,

Fig. 3 Calibration curve of model developed

Table 2 Test results of 10 samples

Reference polyphenol value (% m/m)	Predicted polyphenol value (% m/m)	Deviation on prediction	% of accuracy	Average accuracy	RMSEP	Correlation factor (R ²)
18.02	17.43	± 0.14	96.71	91.7	0.5162	0.925
15.94	18.22	± 0.27	85.69			
14.3	15.19	± 0.18	93.76			
18.29	20.71	± 0.20	86.77			
17.08	16.38	± 0.19	95.93			
23.85	23.32	± 0.28	97.78			
24.96	22.33	± 0.30	89.46			
20.13	21.82	± 0.22	91.6			
22.26	18.42	± 0.18	82.73			
23.98	24.79	± 0.09	96.6			

without undergoing the tedious chemical processes. It is expected to be of great use in price indexing of raw materials (green leaves), enabling improved product consistency and overall enhancement of end product (finished tea) quality.

Acknowledgements The research work has been carried out in collaboration with Tocklai Tea Research Institute (TTRI), Jorhat, Assam, India. The work has been financially supported by National Tea Research Foundation (NTRF), India.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Astill C, Birch BR, Dacombe C, Humphrey PG, Martin PT (2001) Factors Affecting the Caffeine and Polyphenol Contents of Black and Green Tea Infusions. *J Agric Food Chem* 49(11):5340–5347
- Bedini A, Zanolli V, Zanardi S, Bersellini U, Dalcanele E, Suman M (2013) Rapid and simultaneous analysis of xanthines and polyphenols as bitter taste markers in bakery products by FT-NIR spectroscopy. *Food Anal Methods* 6(1):17–27
- Bhuyan LP, Sabhapondit S, Baruah BD, Bordoloi C, Gogoi R, Bhattacharyya P (2013) Polyphenolic compounds and antioxidant activity of CTC black tea of North-East India. *Food Chem* 141(4):3744–3751
- Bian M, Skidmore AK, Schlerf M, Fei T, Liu Y, Wang T (2010) Reflectance spectroscopy of biochemical components as indicators of tea (*Camellia Sinensis*) quality. *Photogramm Eng Remote Sens* 76(12):1385–1392
- Chen Q, Zhao J, Zhang H, Wang X (2006) Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration. *Anal Chim Acta* 572(1):77–84
- Chen Q, Zhao J, Chaitep S, Guo Z (2009) Simultaneous analysis of main catechins contents in green tea (*Camellia Sinensis* (L.)) by Fourier transform near infrared reflectance (FT-NIR) spectroscopy. *Food Chem* 113(4):1272–1277
- Cooper R, Morre J, Morre MD (2005) Medicinal benefits of green tea: part I. Review of non-cancer health benefits. *J Altern Complement Med* 11(3):521–528
- Eden T (1976) Tea. Tropical agriculture series, 3rd edn. Longman: London
- Forrest GI, Bendall DS (1969) The distribution of polyphenols in the tea plant (*Camellia sinensis* L.). *Biochem J* 113(5):741–755
- Gupta S, Saha B, Giri AK (2002) Comparative antimutagenic and anticlastogenic effects of green tea and black tea: a review. *Mutat Res* 512(1):37–65
- Hall MN, Robertson A, Scotter CNG (1988) Near-infrared reflectance prediction of quality, theaflavin content and moisture content of black tea. *Food Chem* 27(1):61–75
- Harbowy ME, Balentine DA (1997) Tea chemistry. *Crit Rev Plant Sci* 16(5):415–480
- Hazarika A, Chanda S, Sabhapondit S, Sanyal S, Tamuly P, Tasrin S, Sing D, Naskar H, Tudu B, Bandyopadhyay R (2018) On-site estimation of total polyphenol in fresh tea leaf using near-infrared spectroscopy. *NIR News* 29(1):9–14. <https://doi.org/10.1177/0960336017750326>
- He Y, Li XL, Deng XF (2007) Discrimination of varieties of tea using near-infrared spectroscopy by principal component analysis and BP model. *J Food Eng* 79(4):1238–1242
- Huang H, Yu Haiyan Xu, Huirong Ying Y (2008) Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: a review. *J Food Eng* 87(3):303–313
- Jha SN, Matsuoka T (2004) Non-destructive determination of acid-brix ratio of tomato juice using near infrared spectroscopy. *Int J Food Sci Technol* 39:425–430. <https://doi.org/10.1111/j.1365-2621.2004.00800.x>
- Jha SN, Ruchi G (2010) Non-destructive prediction of quality of intact apple using near infrared spectroscopy. *J Food Sci Technol* 47(2):207–213. <https://doi.org/10.1007/s13197-010-0033-1>
- Kawasaki M, Kawamura S, Tsukahara M, Morita S, Komiya M, Natsuga M (2008) Near-infrared spectroscopic sensing system for on-line milk quality assessment in a milking robot. *Comput Electron Agric* 63(1):22–27. <https://doi.org/10.1016/j.compag.2008.01.006>
- Kumar PVS, Basheen S, Ravi R, Thakur MS (2011) Comparative assessment of tea quality by various analytical and sensory methods with emphasis on tea polyphenols. *J Food Sci Technol* 48(4):440–446

- Li X, Luo L, He Y, Xu N (2013) Determination of dry matter content of tea by near and middle infrared spectroscopy coupled with wavelet-based data mining algorithms. *Comput Electron Agric* 98:46–53. <https://doi.org/10.1016/j.compag.2013.07.014>
- Li X, Sun C, Luo L, He Y (2015) Determination of tea polyphenols content by infrared spectroscopy coupled with IPLS and random frog techniques. *Comput Electron Agric* 112:28–35. <https://doi.org/10.1016/j.compag.2015.01.005>
- Li ZX, Yang WJ, Ahammed GJ, Shen C, Yan P, Li X, Han WY (2016) Developmental changes in carbon and nitrogen metabolism affect tea quality in different leaf position. *Plant Physiol Biochem* 106:327–335. <https://doi.org/10.1016/j.plaphy.2016.06.027>
- Luypaert J, Zhang MH, Massart DL (2003) Feasibility study for the use of near infrared spectroscopy in the qualitative and quantitative analysis of green tea. *Camellia sinensis* (L.). *Analytica Chimica Acta* 478(2):303–312
- Mahanta PK (1988) Biochemical analysis as a measure of dynamic equilibrium in genomic set up during processing of tea. *J Biosci* 13(3):343–350
- Obanda M, Owuor PO (1995) Impact of shoot maturity on chlorophyll content, composition of volatile flavour compounds and plain black tea chemical quality parameters of clonal leaf. *J Sci Food Agric* 69(4):529–534
- Ozaki Y, McClure WF, Christy AA (2006) Near infrared spectroscopy in food science and technology. Wiley, ISBN: 978-0-470-04770-5. <https://doi.org/10.1002/0470047704>
- Pereira VP, Knor FJ, Velloso JCR, Beltrame FL (2014) Determination of phenolic compounds and antioxidant activity of green, black and white teas of *Camellia sinensis* (L.) Kuntze, Theaceae. *Rev Bras Plantas Med* 16(3):490–498. https://doi.org/10.1590/1983-084x/13_061
- Pissard A, Beten V, Romnee JM, Dupont P, Mouteau A, Lateur M (2012) Classical and NIR measurements of the quality and nutritional parameters of apples: a methodological study of intra-fruit variability. *Biotechnol Agron Soc Environ* 16(3):294–306
- Rakshit A (2016) India records highest ever tea production during the financial year 2015–2016. Business standard (Kolkata, India). [cited 2016 23:52 IST Jun 16]. https://www.business-standard.com/article/markets/india-records-highest-ever-tea-production-in-2015-16-116061700687_1.html
- Ren G, Wang S, Ning J, Xu R, Wang Y, Xing Z, Wan X, Zhang Z (2013) Quantitative analysis and geographical traceability of black tea using Fourier transform near-infrared spectroscopy (FT-NIRS). *Food Res Int* 53(2):822–826
- Roberts CA, Workman J Jr., Reeves JB III (2004) Near-infrared spectroscopy in agriculture, *Agron Monogr*, vol 44. ASA, CSSA, SSSA, Madison
- Sabhapondit S, Bhuyan LP, Goswami BC, Hazarika M (2011) Seasonal and cultivar variation of catechin in North East Indian tea. *Sci Cult* 77:406–411
- Sanderson GW (1972) The chemistry of tea and tea manufacturing. *Recent Adv Phytochem* 5(C):247–316. <https://doi.org/10.1016/b978-0-12-612405-7.50015-1>
- Sanyal S (2011) Tea manufacturing manual. Tea Research Association, Tocklai Experimental Station, Jorhat. www.tocklai.net
- Schulz H, Engelhardt UH, Wegent A, Drews HH, Lapczynski S (1999) Application of near-infrared reflectance spectroscopy to the simultaneous prediction of alkaloids and phenolic substances in green tea leaves. *J Agric Food Chem* 47(12):5064–5067
- Sinija VR, Mishra HN (2009) FT-NIR spectroscopy for caffeine estimation in instant green tea powder and granules. *LWT-Food Sci Technol* 42(5):998–1002
- Yen SH (2007) NIR evaluation of the quality of tea and its market price. *Spectrosc Asia* 3(2):12–15
- Zhang D, Kuhr S, Engelhardt UH (1992) Influence of catechins and theaflavins on astringent taste of black tea brews. *Z Lebensm Unters Forsch* 195(2):108–111
- Zhang MH, Luypaert J, Pierna JAF, Xu QS, Massart DL (2004) Determination of total antioxidant capacity in green tea by near-infrared spectroscopy and multivariate calibration. *Talanta* 62(1):25–35. [https://doi.org/10.1016/S0039-9140\(03\)00397-7](https://doi.org/10.1016/S0039-9140(03)00397-7)