Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 309 (2024) 123853

Contents lists available at ScienceDirect

Spectrochimica Acta Part A:
Molecular and Biomolecular Spectroscopy

journal homepage: www.journals.elsevier.com/spectrochimica-acta-part-a-molecular-and-biomolecular-spectroscopy

Inter seasonal validation of non-contact NIR spectroscopy for measurement of total soluble solids in high tunnel strawberries

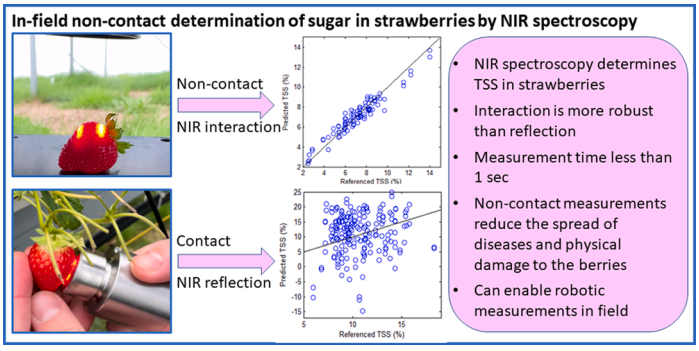
Jens Petter Wold [a,*], Petter Vejle Andersen [a], Kjersti Aaby [a], Siv Fagertun Remberg [b], Anders Hansen [c], Marion O'Farrell [c], Jon Tschudi [c]

[a] Nofima AS - Norwegian Institute of Food, Fisheries and Aquaculture Research, PB 210 NO-1431, Aas, Norway
[b] Faculty of Biosciences, Norwegian University of Life Sciences, 1430, Ås, Norway
[c] SINTEF Digital, Smart Sensor Systems, Forskningsveien 1 0373, Oslo, Norway

HIGHLIGHTS

- Interaction NIR spectroscopy is well suited for non-contact determination of total soluble solids in strawberries.
- NIR interaction measurements are less affected by irrelevant fruit surface properties than reflection.
- The method works well outside in the field with ambient daylight.
- Measurement time is less than 1 sec.
- Measurements in the 760 – 1080 nm region has a probing depth of about 6–7 mm into the strawberries.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Near-infrared spectroscopy
Non-contact interaction measurements
Field measurements
Strawberry
Total soluble solids
Sugar

ABSTRACT

Autonomous field robots are being developed for picking of fruit, where each fruit needs to be individually graded and handled. There is therefore a need for rapid and non-destructive sensing to measure critical fruit quality parameters. In this article we report how total soluble solids (TSS), a measure for total sugar content, can be measured in strawberries in the field by non-contact near-infrared (NIR) interaction spectroscopy. A specially designed prototype system working in the wavelength range 760–1080 nm was tested for this purpose. This novel instrument was compared with a commercial handheld NIR reflection instrument working in the range 900–1600 nm. The instruments were calibrated in the lab using data collected from 200 strawberries of two varieties and tested in a strawberry field on 50 berries in 2022 and 100 berries in 2023. Both systems performed well during calibration with root mean square errors of cross validation for TSS around 0.49 % and 0.57 %, for interaction and reflection, respectively. For prediction of TSS in new berries in 2023, the interaction system was superior, with a prediction error of 1.0 % versus 8.1 % for the reflection system, most likely because interaction probes deeper into the berries. The results suggest that interaction measurements of average TSS are more robust

* Corresponding author at: Norwegian Institute of Food, Fisheries and Aquaculture Research, PB 210 NO-1431, Aas, Norway.
E-mail addresses: jens.petter.wold@nofima.no (J. Petter Wold), petter.andersen@nofima.no (P. Vejle Andersen), kjersti.aaby@nofima.no (K. Aaby), siv.remberg@nmbu.no (S. Fagertun Remberg), anders.hansen@sintef.no (A. Hansen), marion.ofarrell@sintef.no (M. O'Farrell), jon.tschudi@sintef.no (J. Tschudi).

https://doi.org/10.1016/j.saa.2024.123853
Received 17 September 2023; Received in revised form 20 December 2023; Accepted 4 January 2024
Available online 10 January 2024
1386-1425/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

and would most likely require less calibration maintenance compared to reflection measurements. The non-contact feature is important since it reduces the spread of diseases and physical damage to the berries.

1. Introduction

Autonomous field robots are entering agriculture, bringing with them the potential for more automated and precise farming operations. These robots can dramatically decrease the use of pesticides and fertilizer by enabling more targeted application techniques. Automatic treatment of powdery mildew by UV-radiation in different crops has been commercially introduced [1], eliminating the need for pesticides. The development is now progressing towards more advanced automatic processing of crops, such as selective picking of fruits, where each fruit can be individually graded and handled [2]. Sensor data is needed to successfully implement automatic decision-making. Cameras are today used for detection and localization of weeds [3]. There is much research being conducted to develop robust soil sensors that can assess the need for fertilization or watering in the field [4]. RGB cameras can to some extent assess ripeness and other quality parameters on fruit based on color and shape information [5], while hyperspectral cameras have the potential to detect for instance disease related stress symptoms on plants [6]. The internal qualities of fruit and berries, such as dry matter, sugar, and acid contents, are related to ripeness and sensory properties, and are more challenging to measure in field. Reliable measurements of such qualities require probing of the interior of the fruit, not just measuring the surface. To avoid the spread of plant diseases, it is also preferred that any measurements can be done without physical contact with the fruit. Near-infrared spectroscopy (NIRS) has the potential to meet these requirements. In this article, first-time results are reported for rapid, in-the-field, non-contact determination of total soluble solids (TSS), a measure for total sugar content, in strawberries by NIR spectroscopy.

Strawberry (*Fragaria* × *ananassa* Duchesne ex Rozier) is a fruit characterized by its red color, sweetness, and delicious flavor. It is a non-climacteric fruit, meaning that further ripening occurs only to a limited degree after harvest, making the correct selection of harvest time a critical factor for post-harvest fruit quality. The decision to pick is normally made by assessing what percentage of the surface has developed its familiar red color. To also ensure acceptable sensory attributes in the fruits, additional information about chemical composition, e.g. sugars and acids, would be beneficial in the decision-making process [7]. Strawberry sweetness and acidity are important for consumer acceptance and satisfaction, and the sugar-acid ratio can be the most influential factor in how we experience sweetness [8].

Spectroscopic sensors can be used to assess chemical composition in individual strawberries without the need for destructive chemical analyses. We recently showed that Raman spectroscopy can determine TSS, the individual sugars glucose and fructose, as well as total acid and citric acid in whole berries with good accuracy [9]. It was also obtained high correlations (R^2) between the Raman spectra and the sensory properties sweetness (0.70), acidity (0.68) and sourness (0.85). The measurements were performed in a laboratory with a wide area Raman probe at a distance of about 20 cm from the berries. Depth of penetration of such measurements would be about 5–8 mm [10], thereby providing information from the interior of the berries. The methodology is promising but will have practical challenges in the field. Ambient light needs to be shielded to obtain useful Raman spectra, and measurement time with state-of-the art instrumentation would likely be at least 15 sec per berry, which is too long for practical use. In addition, Raman instrumentation is currently more expensive compared to alternative technologies.

It has been well established over the last three decades that NIRS can be used to assess the macro-constituents, dry matter and TSS, in thin skinned fruit with acceptable accuracy [11]. NIRS has also been evaluated for determining a range of other physicochemical properties in strawberries [11–15], such as acidity, pH, phenolic content and firmness. Many have reported good results for prediction of TSS, while results for total acids (TA) were generally not good enough for reliable grading. However, field measurement of TSS alone would be of interest for following the ripening process and determining the optimal harvesting time for each berry. In most of the studies using NIRS for strawberry measurements, the diffuse reflection geometry was chosen [12–16], meaning that mainly the surface to a depth of approximately 1–2 mm was probed. This is apparently sufficient to obtain satisfactory results for TSS but it may not be the most robust mode of measurement, since TSS in strawberries is heterogeneously distributed [17], and the surface might not be representative for the interior. By measuring in interaction mode, where the illumination spot is a distance away from the detection spot, it is possible to measure more in depth since, by design, all detected radiation must have traversed some minimum distance inside the object. Nishizawa et al. [18] did this by using the short NIR region 780–1080 nm, which enables rather deep penetration due to relatively low water absorption and obtained good results for TSS in strawberries. Interaction measurements were also proven superior to reflection for the measurement of TSS in kiwifruit [19] and dry matter in avocados [20]. NIR interaction measurements are most often done with physical contact between sample and instrument to shield against ambient light, and several commercial handheld instruments for produce grading rely on this solution [11]. However, interaction measurements can work well also without contact when the optics are well designed, and the ambient radiation can be sufficiently subtracted. Non-contact interaction is used in industrial hyperspectral NIR imaging food scanners [21,22]. Wold et al. [23] reported that the method works well for in-line quantification of dry matter in single intact potatoes. The ability to measure individual strawberries in field without contact would ease the robotic interface considerably. It would also avoid damage to the berries and reduce the risk of spreading fungal diseases.

This paper reports, for the first time, non-contact NIRS interaction measurements on single strawberries for the determination of TSS. A specially designed prototype NIRS system working in the wavelength range 760–1080 nm was tested. This novel instrument was compared with a commercial handheld contact NIR reflection instrument working in the range 900–1600 nm measuring a surface area of approximately 5 mm in diameter. The instruments were calibrated in the lab using data collected from strawberries during two periods in the same season and finally tested on two independent test set of strawberries in the field over two seasons. The motivation for the work is to design a sensor system that can be used by autonomous field robots to sense TSS in strawberries and other types of fruits.

2. Materials and methods

2.1. Berry samples

Calibration samples: The cultivars (cvs.) Favori and Murano, grown at different farms, were used in the experiment. These are two commercially important everbearing strawberry cultivars in Norway. Strawberries were grown in coconut coir in table-top systems, with automatic watering and nutrient in an open polytunnel (Haygrove Ltd., UK). Fresh strawberries were harvested in two batches during the summer of 2022 for calibration. The first batch of cv. Favori (n = 100) was harvested in week 25 and the second batch of cv. Murano (n = 100) in week 32. The berries were carefully selected to span different degrees of ripeness and visually assessed based on color development and color uniformity. The maturity range spanned from medium ripe (light red color and some parts of white on shade side) to very ripe (dark red). The berries were measured with two NIR instruments (one commercial handheld

instrument and one prototype instrument) in a laboratory after tempering the berries to room temperature (21–22 °C). The ceiling light was turned off to reduce potential disturbance from ambient light while measuring. Spectroscopic measurements were done on both the front side (the sunlit side) and the back side, with center of measurement approximately at equator (Fig. 1). The berries were then refrigerated at 4 °C overnight before reference measurements of TSS and TA.

Independent test sets measured in field: In week 38, 2022, a new set of cv. Murano strawberries (n = 50) (Test set 1) were measured in the open polytunnel during daytime, selected to span different degrees of ripeness as previously described. The berry temperature was in the range 11–13 °C. The berries were measured on the plant with the handheld NIR instrument. Two measurements were done on the front side of each berry. For the larger prototype instrument, the berries were brought to the instrument placed in the polytunnel. With this instrument, two measurements were done on the front side. The first measurement at center point as indicated in Fig. 1b, and the other slightly shifted from center position.

In week 30, 2023, a new set of cv. Favori strawberries (n = 100) (Test set 2) were measured in the polytunnel during daytime. The berries were measured on the plants with both NIR instruments, two measurements each on the front side of the berries. The berry temperature varied in the range 24–28 °C.

2.2. Spectroscopic measurements

The handheld instrument MicroNIR PAT-U (VIAVI Solutions Inc., Chandler, AZ, US) is based on a 128-pixel InGaAs photodiode array and a linear filter. It collects spectra in the wavelength region 908 – 1676 nm where spectral resolution (Full Width Half Maximum, FWHM) varies from about 12 nm to 20 nm with increasing wavelength. Two LEDs are used for illumination, and they are placed about 7 mm apart with the detector in between (Fig. 1a). A standard collar with uncoated sapphire window was mounted on the sensor for optimal working distance (3 mm) from the collection optics to the sample, and measurements were done in contact with the berries. The collar and physical contact shielded to some extent against ambient light. The illumination - detector geometry is made for reflection measurements since the illumination and detection points are overlapping at the sample surface. Some interaction effect could however be expected, but the intensity of this would be dominated by the reflected light. Each measurement integrated spectra over 1 sec. Illustration of measurement in the field is shown in Fig. 2a.

The other NIR instrument was a prototype instrument designed for

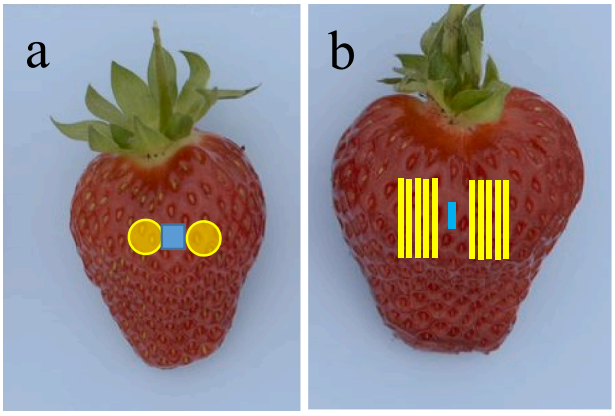


Fig. 1. (a) Strawberry with indicated illumination regions (yellow) and detection region (blue) for the NIR reflection system. (b) Strawberry with indicated illumination stripes (yellow) and detection region (blue) for the NIR interaction system. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

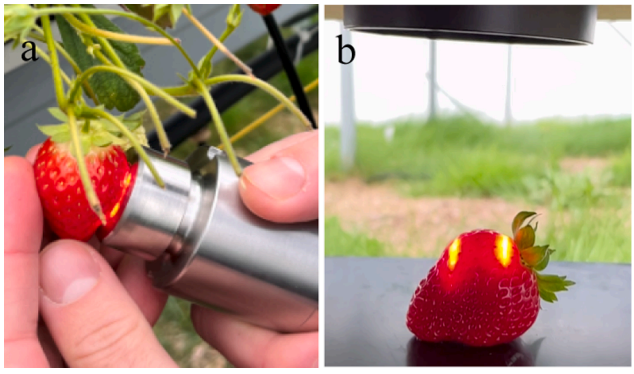


Fig. 2. (a) Illustration of contact measurement using MicroNIR in field. (b) Illustration of non-contact measurement with interaction prototype in field. This is a snapshot of interaction Distance 5 (8.5 mm), giving a distance between illumination lines of 17 mm.

non-contact interaction measurements and is previously described in an article on in-line determination of dry matter in potatoes [23]. An illumination module, with a halogen light source (50 W), was designed with a light chopper to allow for continual subtraction of the ambient light and a means for sequentially switching between five different projected illumination geometries in real time, about 50 times per second (Fig. 1b). Each geometry consisted of two illuminating lines, each of approximately 1 mm x 14 mm, and the interactance signals in the middle of the pair of lines were measured from a single 1 mm x 4 mm detection spot, or field of view (FOV) (the blue region in Fig. 1b). The light travels from the illuminated regions through the strawberry, and the radiation exiting from the FOV is detected by the custom-built spectrometer. The interaction distances (between the FOV and the pair of illumination lines) were approximately 2.5, 4, 5.5, 7 and 8.5 mm, termed Distance 1 to Distance 5, respectively. A single recording contained spectra from each of the five measurement distances taken in rapid succession, making it well-suited for determining the optimal interaction distance. Each spectrum comprised twenty evenly spaced wavelengths in the region 761–1081 nm with a spectral resolution (FWHM) of approximately 16 nm. The working distance between the instrument and strawberries was approximately 8.5 cm. Exposure time per recording was 1 sec, which means 0.2 sec per spectrum from each interaction distance. A photo of measurement in field is shown in Fig. 2b, where *one* of the interaction distances is shown and the strawberry is illuminated from above. Note that orientation of the berries relative to the illumination was as in Fig. 1b during the experimental trials. A photo of the complete illumination pattern can be seen in Wold et al. [23]. A curved white barium plate was used a reference sample.

2.3. Chemical reference analysis

The whole berry samples were homogenized, then centrifugated at 39200 *g* for 10 min (Avanti J-26 XP, Beckman Coulter, USA) and the supernatant collected. The supernatant was used for analyses of TSS and TA.

TSS was determined using a pocket °Brix-acidity meter (PAL-BX|ACID1, Atago Co., Ltd., Tokyo, Japan) and expressed as °Brix (%, g/100 g). TA was measured in supernatant diluted with purified water (1/49, w/w) by the same pocket °Brix-acidity meter. The acidity measurements were based on electroconductivity using electrical current. The concentration of TA was expressed as g citric acid equivalents per 100 g FW.

2.4. Calibration procedure

The NIR intensity spectra from both instruments were transformed to absorption spectra by taking the logarithm of the inverse of the measured spectrum (log10(1/S)), where S is the intensity of the detected

light. To minimize variation in the spectra induced by light scattering and varying distance between instrument and berry, the absorption spectra were normalized by standard normal variate (SNV): For each absorption spectrum the mean value was subtracted, and the spectrum was then divided by the standard deviation of the spectrum [24]. The spectra from MicroNIR were also tested using the second derivative as part of the pre-processing, but this did not improve calibration or prediction results. Three spectral ranges for MicroNIR were evaluated: 908–1676 nm (full), 908–1080 nm (short), and 1100–1676 nm (long). Each of the three regions was normalized by SNV.

Calibration models were made by partial least squares regression (PLSR) [25], a well-established technique for calibration of spectroscopic data. The optimal number of factors in the models was determined after cross validation by evaluation of the correlation (R^2) between measured and predicted TSS, and the root mean square error of cross validation (RMSECV). RMSE is defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y_n})^2} \quad (1)$$

where N is the total number of samples, $\hat{y}_n$ is the predicted value by cross validation (CV) or prediction (P), y_n is the measured reference value and n denotes the samples from 1 to N. Segmented cross validation was used and measurements from five berries were randomly taken out in each segment. Measurements from the same berry (front and back side) were kept in the same segment to avoid overfitted calibrations.

Separate calibrations were made for the two strawberry cultivars Murano and Favori to determine if the different calibrations could be used interchangeably. The regression models used for field testing were based on measurements on back and front sides of both cultivars, a total of 400 measurements.

Prediction results for the two test sets were evaluated by the correlation (R^2) between predicted and referenced TSS, the bias and the root mean square error of prediction (RMSEP), calculated as in Eq. (1).

To establish whether the acquired prediction errors were significantly different, a two-way analysis of variance (ANOVA) of the squared residuals was employed, as adapted from Indahl and Næs [26].

The different temperature of the berries measured in field introduced a shift in the NIR spectra relative to those measured in lab (shown and discussed below). To correct for this difference, we added the difference spectrum between a warm (22 °C) and cold (12 °C) berry to all the new spectra collected outside (Equation (2). The aim of this correction was to reduce the bias of the predicted values.

$$x_{corr,n} = x_n + k^*diff \quad (2)$$

where $x_{corr,n}$ is the temperature corrected spectrum, x_n is the original spectrum collected at lower temperature, *diff* is the difference spectrum, and k is a constant. k was used to adjust the magnitude of the difference spectrum for each data set with the aim of reducing the bias caused by the temperature difference. All spectra were SNV corrected before the correction. The strategy of using difference spectra to correct for temperature in NIR spectroscopy is well known [27]. It can be a more efficient and low-cost calibration strategy compared to measuring every or many berries at different temperatures. In the present study, it was not an aim to make a temperature invariant calibration for TSS, but to illustrate that it is important to take temperature into account and that it is possible to make such calibrations.

Multivariate regression was done with software The Unscrambler 9.8 (Camo, Norway). Other types of analysis were done with MATLAB version R2018a (MathWorks, MA, USA).

2.5. Study of NIR sampling volume

To understand the differences in measurement volume for the two different NIR instruments and due to the five interaction distances, Distance 1 – Distance 5, an experiment was conducted with a strawberry that was cut in half to a thickness of 20 mm. It was then placed, flat side down, on a 4 cm thick block of coconut fat, which was chosen because it gives an easily distinguishable spectrum from that of strawberry, with a distinct absorption fat peak at 930 nm. The curved part of the strawberry faced the NIR instrument above. For each iteration, the strawberry was then sliced from below, making it gradually thinner (a slice of approximately 1 mm was removed each time). Each time a slice was removed, the remaining berry half was placed on the coconut fat again, and an NIR spectrum was recorded. Acquisition time per measurement was 1 sec. Spectra were collected with both NIR instruments. Based on the intensity of the absorption peak for fat at 930 nm it was possible to estimate the relative contribution from fat as function of strawberry thickness. This share could be calculated as

$$\%fat(i) = 100^*(A_{930}(i) - A_{930}(15))/(A_{930}(0) - A_{930}(15)) \quad (3)$$

where i is the thickness from 1 to 15 mm, $A_{930}(15)$ is the absorption for the berry at 15 mm thickness where no fat was detected, $A_{930}(0)$ is the absorption of pure fat. This was done for each interaction distance for the prototype and for the two regions around 930 nm and 1212 nm for MicroNIR. Before the calculations, the spectra were normalized by SNV to remove variations in offset and enable meaningful results from equation (3).

2.6. Color analysis

The color of the berries was measured using a digital color measurement system (DigiEye, VeriVide Ltd., Leicester, UK). The berries were placed in a lightbox with standardized daylight (D65) and diffuse lighting and photographed with a calibrated digital camera (Nikon D7000, 35 mm lens, Nikon Corp., Tokyo, Japan). Color measurements of the pictures were performed with DigiPix software (version 2.63). The color was computed in CIELAB (Commission International de l'Eclairage) color space, expressed as the components L*a*b*, often related to lightness, redness and yellowness, respectively.

3. Results and discussion

3.1. Concentration of sugar and acids

The concentrations of TSS spanned largely the same ranges in weeks 25, 32 and 38 in 2022, but the mean values per groups were significantly different ($p < 0.01$) (Table 1). The group mean values decreased over summer, from June, via August to September. This seasonal effect was as expected [28]. Standard deviations were relatively large in the datasets because samples were selected to span from ripe to very ripe. This was intended to get a good basis for calibration. Test set 2, collected in 2023, had TSS values in the higher range, exceeding the range of the calibrations set.

For TA, there was not a significant difference in means between weeks 25 and 32. It is worth noting that the squared correlations (R^2) between TSS and TA were low: 0,03 and 0,02 in weeks 25 and 32, respectively. R^2 between TSS and the color parameter a* were 0.03 and 0.4 in week 25 and 32 respectively. Finally, R^2 between TA and a* were

Table 1
Total soluble solids (TSS; g/100 g) and total acidity (TA; g/100 g) in strawberries from weeks 25, 32 and 38 (W25, W32 and W38).

Sample set	Property	n	Mean	Min	Max	SD[a]
Week 25, 2022	TSS	100	10.1	6.2	13.4	1.30
	TA	100	0.86	0.64	1.26	0.11
Week 32, 2022	TSS	100	8.7	3.8	13.0	1.21
	TA	100	0.90	0.59	1.31	0.13
Week 38, 2022	TSS	50	7.3	2.5	14.0	2.35
Week 30, 2023	TSS	100	10.6	6.0	18.3	2.38

[a] Standard deviation

0.22 and 0.24 in week 25 and 32 respectively. Correlations with the color parameter b* were equally low for TSS and TA. The very low correlations between TSS and TA prevented indirect modelling of one parameter based on the other. The low correlations between the color parameters and TSS underline that color is not a good marker for sweetness of the berries.

3.2. Spectroscopic measurements

Typical interactance absorption spectra from the same strawberry are shown in Fig. 3a. The spectra are from the five different interaction distances. The absorption increased systematically with increasing distance since the light detected had travelled farther and probed a larger volume before it was detected. The quantitative difference in absorption between Distance 1 and Distance 5 corresponds to a signal intensity that is approximately 10 times higher for Distance 1. Note also that the contrast in the spectra (the height of the absorption peak) increased systematically with increasing distance, meaning that spectral features were more pronounced when measured at greater interaction lengths. Fig. 3b shows the corresponding reflectance spectrum collected with the MicroNIR. Note that the wavelength range here is much longer into the NIR region. In the common shaded region, the potential light penetration depth is the same. However, the different geometrical designs of the two systems facilitate different sampling depths.

Fig. 4 (a) shows spectra from a strawberry of different thicknesses measured with both systems on a block of coconut fat. Only spectra from Distance 4 for the prototype are shown. The thinner the strawberry was, the more pronounced was the C–H fat peak at 930 nm for both systems and also at 1212 nm for MicroNIR. The share of fat signal at 930 nm for different thicknesses indicates the depth of sampling. These estimated shares of fat (as percent of a spectrum of pure fat) were calculated according to equation (3) and are plotted in Fig. 4c as function of strawberry thickness. They show that the fat was detectable down to depths of 6–8 mm with the prototype interaction system. The largest share of fat at each thickness was probed with the longest interaction distance, Distance 5, and decreased with decreasing distance. With the MicroNIR, at 930 nm, little or no fat was detected at 5 mm thickness, but the system captured signals down to 2–3 mm depth. Notably, at 1212 nm, the probing depth was even shorter. This was expected, since the absorption of water is stronger at this wavelength, reducing light penetration depth. The results illustrate that: 1) Interaction enables deeper sampling than diffuse reflectance, and 2) the short NIR region (800–1000 nm) enables deeper sampling than longer wave NIR (>1100 nm) [11]. Although this is known in general, it is useful to survey such properties for the food samples being tested, to understand and interpret results in a best possible way. Since strawberries are heterogeneous with internal gradients in TSS concentration, it would most likely be favorable to probe as deep as possible provided that the measurement system has sufficient SNR. These measurements would also be less sensitive to any irrelevant variations on the surface of the strawberry.

A similar study of the penetration depth as a function of interaction distance has previously been done on potatoes [23]. In potatoes, there were more systematic and distinct differences in depth penetration between the different interaction distances than what we found in this study. The reason might be that light scattering is more pronounced in potatoes compared to strawberries, which are more translucent. Furthermore, the interaction distances in the potato study were longer (5, 8, 11, 14 and 17 mm), and thus inducing greater differences in probing depth per interaction distance.

Fig. 5 shows typical spectral features from strawberry in the 760–1080 nm region (collected with the prototype system at Distance 4). The spectra were rather featureless with the water absorption peak dominating at about 976 nm (OH-stretch) (Fig. 5a). Differences caused by varying sugar content were barely discernable, but an apparent small shift towards longer wavelengths was observed for strawberries with high TSS values. This corresponds with findings by Golic et al. [30] who documented an absorption peak for sugar around 960 nm and a shoulder at 984 nm (both OH stretching). A main absorption peak for sugar is at about 910 nm (CH-stretching) [29,30], but this was not possible to see in the spectra with the naked eye. The very small spectral differences resulting from the natural span of TSS in strawberries underline the importance of having high SNR in the spectral measurements to quantitatively model the TSS content. Since the goal is to perform measurements outside in the field, varying conditions in ambient light are likely critical and should be handled to preserve the quality of the spectra and minimize deviations in estimated values. Since the spectral features of interest are so subtle, it is difficult to validate the quantitative models by interpretation of visual spectral features, something which was also highlighted by Golic et al. [30]. However, Fig. 5b shows typical regression coefficients for TSS calibrations presented below, and they relied strongly on 910 nm, a main absorption peak for sugar. The main features were the same for all models built using data from the prototype system. The OH bands of sugar (at 960 nm and 984 nm) were less pronounced, perhaps reduced due to their temperature dependance. The regression vector carried some of the same features as reported for PLSR models for sugar in water and mango [30,31].

Fig. 5c shows NIR spectra from the same strawberry with core temperature at 10 °C and 20 °C. The water absorption peak at 978 nm shifts towards longer wavelengths for lower temperatures. The difference spectrum for this apparent shift shows the systematic change that occurs when temperature of the samples varies. The spectral change due to a 10 °C temperature difference was slightly bigger than the change due to

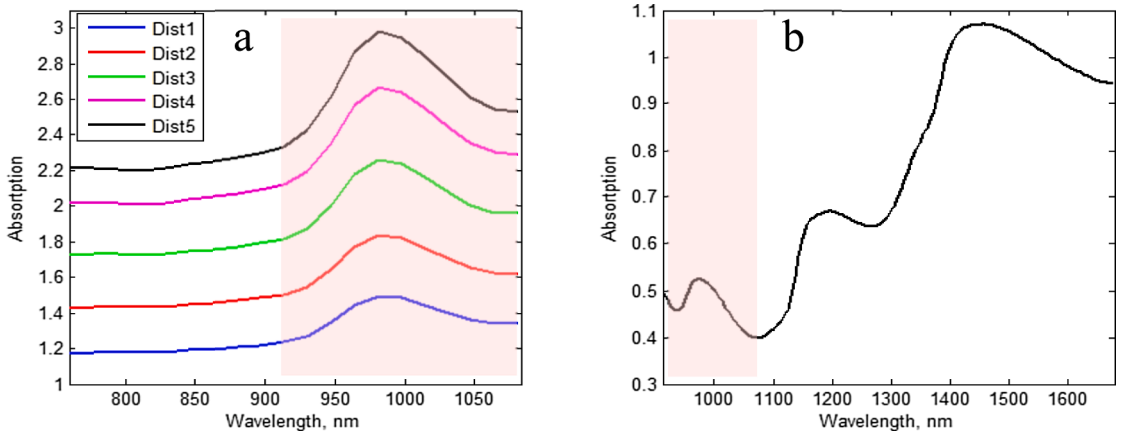


Fig. 3. a) Absorption interactance spectra for the five different distances from one strawberry, indicated with Dist 1 to Dist 5. b) Absorption spectrum measured with MicroNIR on the same strawberry. The common spectral region for the two systems is shaded.

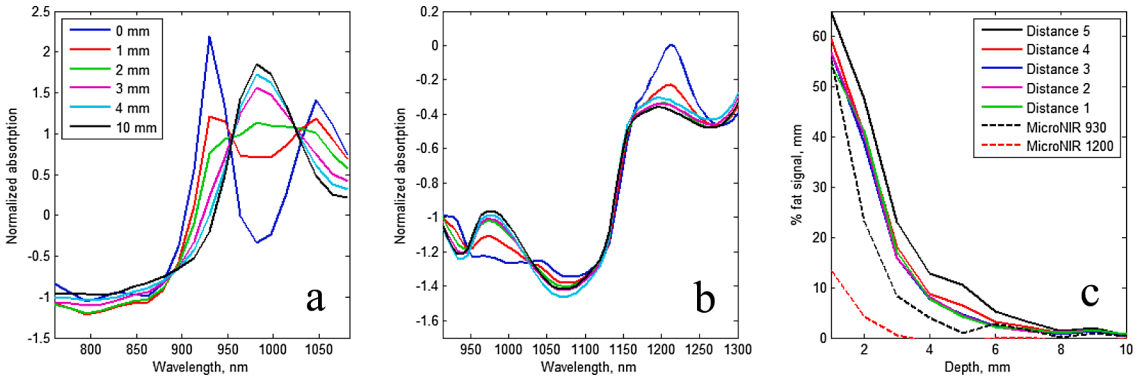


Fig. 4. Spectra from a strawberry of different thicknesses on coconut fat measured with prototype Distance 4 (a) and MicroNIR (b). (c) Approximate share of fat signal at different strawberry thicknesses for the different sensor designs.

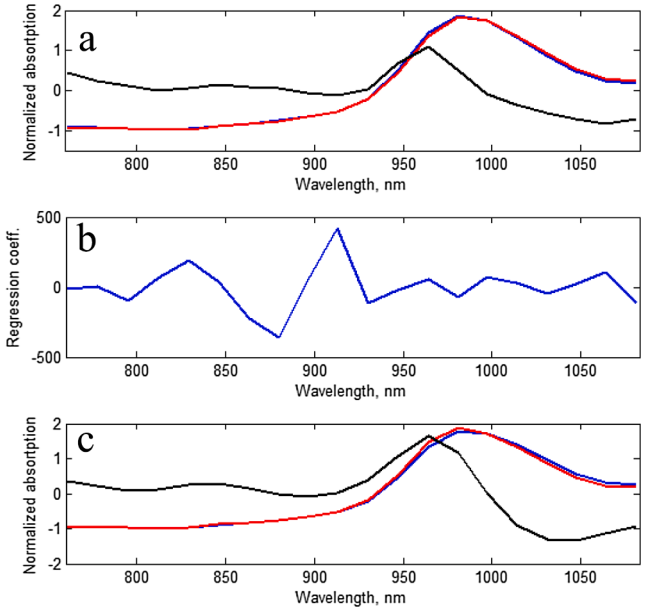


Fig. 5. (a) Normalized spectra from berries with low (blue) and high (red) TSS concentrations. Difference spectrum between the two (low TSS – high TSS) in black (multiplied by 13) (b) Typical regression vector for TSS calibration based on interaction spectra. (c) Normalized spectra from strawberry with temperature of 20 °C (red) and 12 °C (blue). Difference spectrum between the two (warm – cold) in black (multiplied by 13). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

a large variation in TSS (Fig. 5a) so it is very important to incorporate temperature in calibration models, especially when the intended measurements will be done outside in varying temperatures. This issue has been treated thoroughly by others [30,32].

3.3. Calibration results

We did not obtain meaningful calibrations for TA using data from the two NIR instruments, for any of the groups of strawberries. These results are therefore not listed in detail. Most studies on NIRS and strawberries indicate poor ability to estimate TA and citric acid [12,14,15], while a limited number of reports show very promising results [13,16], but it is not clear what chemical absorption bands the calibrations were based upon, and it is possible that the models were based on indirect correlations. For example, we have seen in our own previous work that TA sometimes can correlate closely with the amount of chlorophyll, which is measured at 675 nm.

Table 2
Calibration results for TSS (%) in Favori and Murano berries for different interaction distances and for MicroNIR.

NIR data	Calibration on Favori and Murano (n = 200)		
	LV[a]	R^2	RMSECV[b]
Interaction Dist 1	9	0.77	0.67
Interaction Dist 2	9	0.82	0.58
Interaction Dist 3	9	0.84	0.54
Interaction Dist 4	9	0.86	0.52
Interaction Dist 5	9	0.88	0.49
Reflection full[c]	10	0.82	0.57
Reflection short[c]	6	0.77	0.67
Reflection long[c]	10	0.77	0.66

[a] Latent variables.
[b] Root mean square error of cross validation.
[c] Refer to wavelength regions used from the MicroNIR system. Full: 908–1676 nm, short: 908 nm, long: 1100–1676 nm.

Table 2 summarizes calibration results for the joint sets of strawberries, cvs. Favori and Murano, harvested in weeks 25 and 32, 2022, respectively. Results are presented for all interaction distances and the three selected wavelength regions from MicroNIR. For the prototype interaction system, the regression models required 9 PLS factors. The first latent variable, which represented the shift illustrated in Fig. 5a, explained about 40 % of the variation in TSS, and then eight additional factors were needed for optimal models, based on cross validation. The models relied on subtle spectral variations, which again underlines the need for sufficient SNR in the spectra. There is a clear trend that increasing the interaction distance improves the accuracy of the calibrations. RMSECV for TSS was reduced from 0.67 % to 0.49 % when the distance between the illumination regions were increased from 2.5 to 8.5 mm. The differences in RMSECV between adjacent distances were significant ($p < 0.05$) except for between Distances 3 and 4. It indicates that the deeper we probed, the more representative the spectral measurements were of the TSS value. The advantage of probing deeper is that the model is more robust towards heterogeneity in the strawberry and any interfering features on the surface. The same increased accuracy with increasing interaction distance was also observed when measuring dry matter in potatoes [23]. Potatoes are also heterogeneous, with prominent internal gradients in dry matter [11]. The results obtained here are slightly better than a previous report on interaction NIRS on strawberries [18], even though the previous study was conducted with the instrument in full contact with the berry. A perfect match between the spectroscopically probed volume and the sugar content is difficult to obtain, and this is obviously one source of error in the calibration results.

When the full spectrum from the MicroNIR was used, calibration results were obtained that were in line with the best results using the interaction system. The regression coefficients (not shown) emphasized

principal bands for sugars, such as 1196, 1276, 1368 and 1489 nm [29]. When the short and long wavelength regions were separated, significantly poorer results were obtained, more in line with the models based on the shortest interaction distance. The *short* wavelength region from MicroNIR is not 100 % comparable with the interaction system, but they both include the major sugar absorption peak at about 910 nm. Fig. 4c indicated that the MicroNIR short region did not probe as deep as the interaction system, which could explain why it performs less well. Distance 1 and Distance 2, which are closer in measurement geometry to MicroNIR also gave lower correlations.

With increasing interactance distance, the signal to noise in the system would decrease since exposure time for each distance was the same. The instrument noise effect on the predicted TSS values was investigated by measuring the same strawberry 10 times in steady state. Standard deviation of the 10 predicted values ranged from 0.02 to 0.04 from Distance 1 to Distance 4. For the MicroNIR system, the standard deviations were 0.25 for the short region and 0.17 for the full region. Compared to the prediction errors, none of the systems were limited by instrument noise.

We do not know the detailed differences between Murano and Favori regarding sugar distribution. It is not obvious that calibration models based on *one* cultivar should work well on *the other* cultivar. We tested this (results shown in Table S1) and the results suggest that the deeper the berry is probed, the more interchangeable calibrations are obtained: This verifies the assumption by Walsh et al. [11] who emphasized that limited penetration depth would limit performance across independent populations, given variations in outer layer attributes.

3.4. Prediction of independent test sets in field

Prediction of TSS values in berries in field were based on the regression models reported in Table 2. For Test set 1, measured the same year as the calibration set, the interaction system provided good predictions of TSS. As mentioned before, the temperature differences gave a systematic shift in the spectra and without correcting this shift, the predicted values of TSS had a bias of approximately 2.3 %. With the temperature correction, the biases were greatly reduced, and those results are shown in Table 3. The prediction errors were only slightly higher than those obtained for the cross validated calibrations. A notable difference from the calibration results in Table 2, is that the lowest prediction errors were not obtained for Distance 5. Distance 3 gave the lowest prediction error, significantly lower than for Distances 2 and 4, while Distance 5 gave the highest prediction error. Increasing interaction distance results in weaker signals (due to higher absorption). It is possible that at longer distances, the weaker signal intensity is more affected by the stronger ambient light in the field. Also, at Distance 4 and 5 there is a greater chance of non-optimal measurements on small strawberries since the illuminating lines can hit the berries at the edges. Predicted versus referenced TSS concentrations for Distance 3 are shown in Fig. 6a.

Table 3
Results for common TSS calibration for Favori and Murano berries for different interaction distances and for MicroNIR. Prediction results are for Murano berries in the field.

NIR data	Test set 1 in field on Murano (n = 50)			Test set 2 in field on Favori (n = 100)		
	R^2	RMSEP[b]	bias	R^2	RMSEP[b]	bias
Interaction Dist 1	0.88	0.73	0.1	0.86	1.13	0.1
Interaction Dist 2	0.92	0.63	0.0	0.85	1.19	0.2
Interaction Dist 3	0.92	0.59	0	0.87	1.00	0.1
Interaction Dist 4	0.88	0.70	0	0.86	1.20	0.3
Interaction Dist 5	0.84	0.82	0	0.80	1.74	0.3
Reflection full[c]	0.75	1.07	0	0.02	8.1	−0.4
Reflection short[c]	0.69	1.21	0.03	0.01	9.50	−0.6
Reflection long[c]	0.72	1.78	1.35	0.01	10.4	−0.7

[a]Latent variables.
[b] Root mean square error of prediction
[c] Refer to wavelength regions used from the MicroNIR system. Full: 908–1676 nm, short: 908 nm, long: 1100–1676 nm.

Using MicroNIR, the prediction errors were significantly higher, but still promising. The temperature correction also worked well on these data, and for the *full* and *short* spectra the bias could be removed by this approach. For the *long* part of the spectra, we did not succeed in removing the bias, probably due to other effects beyond temperature. The lower correlations and higher prediction errors compared to the interaction measurements could again be explained by physiochemical differences in the outer layer of the strawberries. Fig. 6b shows predicted versus referenced TSS concentrations for the two NIR measurements done on the front side of the berries.

For Test set 2, measured the year after, the interaction system still provided good predictions of TSS, although the prediction errors were slightly higher (Table 3). Variations in berry temperature could have contributed to this higher prediction error. The reflection NIR system, however, failed completely to predict these berries satisfactory. In the spectra from both systems we could observe a shift in the absorption peaks of water compared to the 2022 berries, but this was much more pronounced in the reflection measurements. We also noticed that the visual appearance of the 2023 berries was slightly different from the 2022 berries, including a higher density of seeds on the surface of several of them. This would most likely affect the optical properties of the strawberries, especially at the surface.

4. Discussion

In this study we have seen that non-contact interaction NIRS is well suited for rapid determination of TSS in strawberries. The study does not include very large sample sets, but several calibrations have been tested on independent test sets across strawberry cultivars and time. The data is sufficient to conclude that non-contact interaction NIRS is more robust and would most likely require less calibration maintenance compared to reflection measurements. The main reason for this is the greater optical penetration depth that makes interaction measurements less sensitive to physical and chemical changes at the surface of the berries.

Making a final NIR calibration that is robust over time and temperature variations was not the focus of this article. Temperature variation in the field is a source of inaccuracy and any chosen NIRS system must handle this. The temperature correction based on difference spectra was demonstrated to work well for both the systems tested here. This approach can also be used in the calibration stage where the calibration spectra are mathematically augmented by different levels of the difference spectrum, as proposed by Segtnan et al. [27]. Or a set of strawberries can simply be measured at different temperatures and be included in the calibration set.

The subtraction of ambient daylight from the NIR measurements is critical [33] and can be a challenge, and strong daylight can reduce the performance for longer interaction distances. For a robot working in polytunnels, the ambient light will most likely be diffuse due to the tunnel cover.

The prototype designed for *non-contact* interaction measurements used in this study was rather large and bulky and not very practical for use in field. However, designs are currently being implemented to achieve a more compact optical solution that can be used by humans or robots. The non-contact feature is important and even critical since it avoids physical contact and reduces the spread of diseases and physical damage to the berries. It will also enable more flexible positioning of the probe in relation to the berry and could allow non-contact scanning of larger objects such as bunches of grapes.

5. Conclusion

This study shows that non-contact interaction NIRS is well suited for

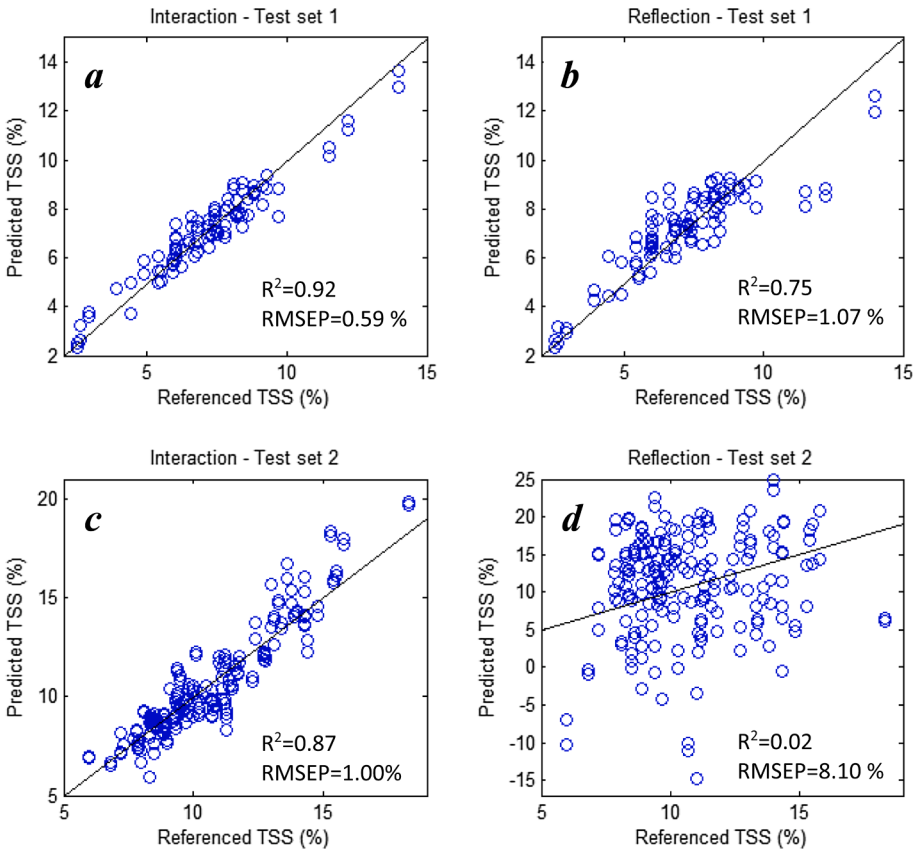


Fig. 6. Predicted versus measured TSS in test sets of strawberries measured in field with interaction system in 2022 (a) and 2023 (c), and reflection in 2022(b) and 2023 (d). Results are shown for two measurements on front side for each berry. Diagonal lines are target lines.

rapid determination of TSS in strawberries. The method was tested in field with promising results. The results suggest that interaction measurements of average TSS are more robust and would most likely require less calibration maintenance compared to diffuse reflection measurements.

Funding

This work was funded by the Research Council of Norway through the projects SFI Digital Food Quality (RCN no. 309259), Sensors for precision picking of strawberries (RCN no. 321555) and the Food Pilot Plant (RCN no. 296083); along with The Agricultural and Food Industry Research Funds through the projects Precision Food Production (RCN no. 314111) and SusHealth (grant number 314599).

CRediT authorship contribution statement

Jens Petter Wold: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. **Petter Vejle Andersen:** Data curation, Methodology, Writing – review & editing. **Kjersti Aaby:** Conceptualization, Investigation, Resources, Visualization, Writing – review & editing. **Siv Fagertun Remberg:** Conceptualization, Investigation, Methodology, Resources, Supervision, Writing – review & editing. **Anders Hansen:** Conceptualization, Investigation, Methodology, Resources, Writing – review & editing. **Marion O'Farrell:** Conceptualization, Funding acquisition, Investigation, Resources, Writing – review & editing. **Jon Tschudi:** Conceptualization, Formal analysis, Investigation, Methodology, Resources, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jens Petter Wold reports financial support was provided by Research Council of Norway. Jens Petter Wold reports financial support was provided by The Agricultural and Food Industry Research Funds (Norway).

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank Katinka Dankel, Karen Wahlstrøm Sanden and Mona Ringstad for excellent technical assistance.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.saa.2024.123853.

References

[1] L. Grimstad, P.J. From, The Thorvald II agricultural robotic system, Robotics 6 (2017) 24, https://doi.org/10.3390/robotics6040024.
[2] H. Zhou, X. Wang, W. Au, H. Kang, C. Chen, Intelligent robots for fruit harvesting: Recent developments and future challenges, Precis. Agric. 23 (2022) 1856–1907, https://doi.org/10.1007/s11119-022-09913-3.
[3] A. Wang, W. Zhang, X. Wei, A review on weed detection using ground-based machine vision and image processing techniques, Comput. Electron. Agric. 158 (2019) 226–240.

[4] L. Yu, W. Gao, R.R. Shamshiri, S. Tao, Y. Ren, Y. Zhang, G. Su, Review of research progress on soil moisture sensor technology, Int. J. Agric. Biol. Eng. 14 (2021) 32–42.
[5] T. Yoshida, T. Kawahara, T. Fukao, Fruit recognition method for a harvesting robot with RGB-D cameras, Robomech. J. 9 (2022) 15, https://doi.org/10.1186/s40648-022-00230-y.
[6] P. Mishra, G. Polder, N. Vilfan, Close range spectral imaging for disease detection in plants using autonomous platforms: A review on recent studies, Curr. Robot. Rep. 1 (2020) 43–48.
[7] A.A. Kader, Postharvest Technology of Horticultural Crops Oakland, Calif. University of California, Agriculture and Natural Resources, 3rd ed. 2002.
[8] K.S. Lewers, M.J. Newell, E. Park, Y. Luo, Consumer preference and physiochemical analyses of fresh strawberries from ten cultivars, Int. J. Fruit Sci. 20 (2020) 733–756, https://doi.org/10.1080/15538362.2020.1768617.
[9] P.V. Andersen, N.K. Afseth, K. Aaby, M.Ø. Gaarder, S.F. Remberg, J.P. Wold, Prediction of chemical and sensory properties in strawberries over one growing season using Raman spectroscopy, Postharvest Biol. Technol. 201 (2023) 112370, https://doi.org/10.1016/j.postharvbio.2023.112370.
[10] O. Monago-Maraña, N.K. Afseth, S.H. Knutsen, S.G. Wubshet, J.P. Wold, Quantification of soluble solids and individual sugars in apples by Raman spectroscopy: A feasibility study, Postharvest Biol. Technol. 180 (2021) 111620, https://doi.org/10.1016/j.postharvbio.2021.111620.
[11] K.B. Walsh, J. Blasco, M. Zude-Sasse, X. Sun, Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use, Postharvest Biol. Technol. 168 (2020) 111246, https://doi.org/10.1016/j.postharvbio.2020.111246.
[12] A.C. Agulheiro-Santos, S. Ricardo-Rodrigues, M. Laranjo, C. Melgão, R. Velázquez, Non-D destructive prediction of total soluble solids in strawberry using near infrared spectroscopy, J. Sci. Food Agric. 102 (2022) 4866–4872, https://doi.org/10.1002/jsfa.11849.
[13] A. Saad, M.M. Azam, B.M.A. Amer, Quality analysis prediction and discriminating strawberry maturity with a hand-held Vis–NIR spectrometer, Food Anal. Methods 15 (2022) 689–699, https://doi.org/10.1007/s12161-021-02166-2.
[14] M.T. Sánchez, M.J. De La Haba, M. Benítez-López, J. Fernández-Novales, A. Garrido-Varo, D. Pérez-Marín, Non-Destructive characterization and quality control of intact strawberries based on NIR spectral data, J. Food Eng. 11 (2012) 102–108, https://doi.org/10.1016/j.jfoodeng.2011.12.003.
[15] M.L. Amodio, F. Ceglie, M.M.A. Chaudhry, F. Piazzolla, G. Colelli, Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems, Postharvest Biol. Technol. 125 (2017) 112–121, https://doi.org/10.1016/j.postharvbio.2016.11.013.
[16] Y. Shao, Y. He, Nondestructive measurement of acidity of strawberry using VIS/NIR spectroscopy, Int. J. Food Properties 11 (2008) 102–111, https://doi.org/10.1080/10942910701257057.
[17] A. Ikegaya, T. Toyoizumi, S. Ohba, T. Nakajima, T. Kawata, S. Ito, E. Arai, Effects of distribution of sugars and organic acids on the taste of strawberries, Food Sci. Nutrit. 7 (2019) 2419–2426, https://doi.org/10.1002/fsn3.1109.
[18] T. Nishizawa, Y. Mori, S. Fukushima, M. Natsuga, Y. Maruyama, Non-destructive analysis of soluble sugar components in strawberry fruits using near-infrared spectroscopy, Nippon Shokuhin Kagaku Kogaku Kaishi 56 (2009) 229, https://doi.org/10.3136/nskkk.56.229.
[19] P.N. Schaare, D.G. Fraser, Comparison of reflectance, interactance and transmission modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit (Actinidia chinensis), Postharvest Biol. Technol. 20 (2000) (2000) 175–184.
[20] P.P. Subedi, K.B. Walsh, Assessment of avocado fruit dry matter content using portable near infrared spectroscopy: Method and instrumentation optimisation, Postharvest Biol. Technol. 161 (2020) 111078.
[21] J.P. Wold, E. Veiseth-Kent, V. Høst, A. Løvland, Rapid on-line detection and grading of wooden breast myopathy in chicken fillets by near-infrared spectroscopy, PLoS One 12 (2017) e0173384, https://doi.org/10.1371/journal.pone.0173384.
[22] J.P. Wold, M. O'Farrell, M. Høy, J. Tschudi, On-line determination and control of fat content in batches of beef trimmings by NIR imaging spectroscopy, Meat Sci. 89 (2011) 317–324.
[23] J.P. Wold, M. O'Farrell, P.V. Andersen, J. Tschudi, Optimization of instrument design for in-line monitoring of dry matter content in single potatoes by NIR interaction spectroscopy, Foods 10 (2021) 828, https://doi.org/10.3390/foods10040828.
[24] R.J. Barnes, M.J. Dhanoa, S.J. Lister, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (1989) 772–777.
[25] H. Martens, T. Næs, Multivariate calibration, Wiley, New York, 1989.
[26] U.G. Indahl, T. Næs, Evaluation of alternative spectral feature extraction methods of textural images for multivariate modelling, J. Chemom. 12 (1998) 261–278, https://doi.org/10.1002/(SICI)1099-128X(199807/08)12:4<261::AID-CEM513>3.3.CO;2-Q.
[27] V.H. Segtnan, B.H. Mevik, T. Isaksson, T. Næs, Low-cost approaches to robust temperature compensation in near-infrared calibration and prediction situations, Appl. Spectrosc. 59 (2005) 816–825.
[28] M.L. Schwieterman, T.A. Colquhoun, E.A. Jaworski, L.M. Bartoshuk, J.L. Gilbert, D.M. Tieman, A.Z. Odabasi, H.R. Moskowitz, K.M. Folta, H.J. Klee, C.A. Sims, V. M. Whitaker, D.G. Clark, Strawberry flavor: Diverse chemical compositions, a seasonal influence, and effects on sensory perception, PLoS One 9 (2014) e88446.
[29] B.G. Osborne, T. Fearn, P.H. Hindle, Practical NIR spectroscopy with applications in food and beverage analysis, Longman scientific and technical, New York, 1993.
[30] M. Golic, K.B. Walsh, P. Lawson, Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature, Appl. Spectrosc. 57 (2003) 139–145.
[31] P.P. Subedi, K.B. Walsh, Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy, Postharvest Biol. Technol. 62 (2011) 238–245, https://doi.org/10.1016/j.postharvbio.2011.06.014.
[32] X. Sun, P. Subedi, K.B. Walsh, Achieving robustness to temperature change of a NIRS-PLSR model for intact mango fruit dry matter content, Postharvest Biol. Technol. 162 (2020) 111117, https://doi.org/10.1016/j.postharvbio.2019.111117.
[33] S. Saranwong, J. Sornsrivichai, S. Kawano, On-tree evaluation of harvesting quality of mango fruit using a hand-held NIR instrument, J. Near Infrared Spectrosc. 11 (2003) 283–293.