Evaluation of Carbohydrate Concentrations in *Phalaenopsis* Using Near-infrared Spectroscopy

Yung-Kun Chuang  
*Master Program in Food Safety, Taipei Medical University, Taipei 11031, Taiwan*

I-Chang Yang  
*Taiwan Agricultural Mechanization Research and Development Center, Taipei 11051, Taiwan*

Chao-Yin Tsai  
*Department of Bio-Industrial Mechatronics Engineering, National Taiwan University, Taipei 10617, Taiwan*

Jiunn-Yan Hou  
*Department of Horticulture and Landscape Architecture, National Taiwan University, Taipei 10617, Taiwan*

Yung-Huei Chang  
*Department of Bio-Industrial Mechatronics Engineering, National Taiwan University, Taipei 10617, Taiwan*

Suming Chen  
*Department of Bio-Industrial Mechatronics Engineering, National Taiwan University, Taipei 10617, Taiwan; and Taiwan Agricultural Mechanization Research and Development Center, Taipei 11051, Taiwan*

**Additional Index Words.** Fructose, glucose, internal quality, modified partial least-squares regression, quantitative analysis, starch, stepwise multiple linear regression, sucrose

**Abstract.** Carbohydrate concentrations are important indicators of the internal quality of *Phalaenopsis*. In this study, near-infrared (NIR) spectroscopy was used for quantitative analyses of fructose, glucose, sucrose, and starch in *Phalaenopsis* plants. Both modified partial least-squares regression (MPLSR) and stepwise multiple linear regression (SMLR) methods were used for spectral analysis of 302 *Phalaenopsis* samples in the full visible NIR wavelength range (400–2498 nm). Calibration models built by MPLSR were better than those built by SMLR. For fructose, the smoothed first derivative MPLSR model provided the best results, with a correlation coefficient of calibration (Rc) of 0.96, standard error of calibration (SEC) of 0.22% dry weight (DW), standard error of validation (SEV) of 0.28% DW, and bias of -0.01% DW. For glucose, the MPLSR model based on the smoothed first derivative spectra was the best (Rc = 0.96; SEC = 0.26% DW; SEV = 0.32% DW; and bias = 0.01% DW). The best MPLSR model of sucrose was developed using the smoothed first derivative spectra (Rc = 0.96; SEC = 0.24% DW; SEV = 0.31% DW; bias = -0.03% DW). Regarding starch, the smoothed first derivative MPLSR model showed the best effects (Rc = 0.91; SEC = 0.47% DW; SEV = 0.56% DW; bias = -0.02% DW). Both the MPLSR and SMLR models showed satisfactory predictability, indicating that NIR has the potential to be adopted as an effective method of rapid and accurate inspection of the carbohydrate concentrations of *Phalaenopsis* plants. This technique could contribute substantially to quality management of *Phalaenopsis*.

The *Phalaenopsis* orchid is one of the most valued floriculture crops in the world; therefore, the production of *Phalaenopsis* is an expanding specialized international industry. Quality assurance of *Phalaenopsis* is very important because of the competition of international markets. Previous studies have indicated that carbohydrates, hormones, and minerals are related to the growth of *Phalaenopsis* plants (Chen et al., 1994; Kataoka et al., 2004). Among them, starch is one type of carbohydrate stored in plants (Theron and Jacobs, 1996) that is converted to fructose, glucose, and sucrose, which are needed for growth. Hence, these carbohydrates can be regarded as indicators of internal quality (Chen et al., 1994; Hou et al., 2011; Kataoka et al., 2004). To achieve good management of quality assurance, a rapid detection method of quantitatively analyzing fructose, glucose, sucrose, and starch in *Phalaenopsis* plants is strongly needed. The commonly used techniques of high-performance liquid chromatography (HPLC) (Di Stefano et al., 2012; Magwaza and Opara, 2015) and high-performance anion exchange chromatography (HPAEC) (Arrizon et al., 2010; Cruz-Cárdenas et al., 2015) are time-consuming and labor-intensive; therefore, these methods cannot be used for rapid comprehensive quality inspection of *Phalaenopsis* plants.

Near-infrared (NIR) spectroscopy, a rapid inspection method based on specific absorptions within a given range of wavelengths corresponding to the constituents of the sample, has been widely used to evaluate the internal quality of...
agricultural products (Chuang et al., 2014; Giovanelli et al., 2014; Maniwara et al., 2014). Until now, this technique had not been used to quantitatively monitor the carbohydrate status of *Phalaenopsis* plants. Therefore, the objective of the current study was to evaluate the carbohydrate concentrations in *Phalaenopsis* plants in terms of quantitative approaches using NIR spectroscopy. Efforts were made to build NIR spectral calibration models with high predictability to examine the internal quality of *Phalaenopsis*. To obtain plants with a wider range of carbohydrate concentrations, two different groups of plants were prepared: samples grown in the greenhouse during various durations and samples subjected to various durations of dark periods. Both samples were included in the analyses. The ability to predict fructose, glucose, sucrose, and starch using the NIR method was investigated in this study.

**Materials and Methods**

**Plant materials and sample preparation.** The white-flowered clone, *Phalaenopsis* Sogo Yukidian ‘V3’ (purchased from Clone International Biotech Co., Ltd., Pingtung, Taiwan) was used in this study. Plants were mature and able to flower. They were grown in 10.5-cm-diameter, clear, soft, plastic pots (0.75 L) tightly packed with sphagnum moss. All plants were fertigated as needed with a water-soluble 20N–8.6P–16.6K fertilizer (Peters Professional 20–20–20; Everris International, Geldermalsen, the Netherlands) at 1 g L–1. Heating and cooling in the greenhouse were programmed to maintain average day/night temperatures of 28/25 °C. Automatic double-shaded cloth adjustments maintained a maximum photosynthetic photon flux (PPF) of 370 μmol m–2 s–1.

Greenhouse plants were divided into six groups and grown in the greenhouse at a constant shading level for 0, 1, 2, 3, 4, or 6 weeks to obtain plants with a wide range of carbohydrate concentrations. The first group (i.e., 0 weeks) was directly sampled for experiments. Each treatment had seven single-plant replications. Samples subjected to NIR measurements and carbohydrate analyses were taken from plants immediately after completion of each cultivation duration. The sampled parts included the first, third, and fifth leaves from the apex. In the present study, 126 leaf samples from the greenhouse plants were collected.

In addition, a total of 88 plants were subjected to various durations of dark periods for a wider range of carbohydrate concentrations for further analyses (hereafter referred to as “dark-treated plants”). The water supply was restricted for at least 7 d to dry the sphagnum moss used as the growing medium in the pots. Fertigation was also stopped for at least 7 d to reduce the risk of disease occurrence. The plant samples were then evenly divided into five groups kept at 20 °C for 0, 10, 20, 30, or 40 d. The first group (i.e., 0 d) was not subjected to the experiment and was directly sampled. The plant samples designated to undergo dark treatment were placed into corrugated fiberboard cartons with shredded newspapers and placed in a dark growth chamber at the designated temperature and duration during the experiments. The relative humidity was 40% to 50% in the cartons during storage, as measured by a data logger (H08-004-02; Onset Computer Corp., Bourne, MA). The detailed experimental designs were also mentioned in a previous study (Hou et al., 2011). After completion of each storage duration, the DW and percentage of plants with yellowed leaves were determined. Dark-treated plants were then divided into two groups: shoots including all leaves and the compressed stem and roots including all roots remaining on the stem. Each of the 88 shoot samples and 88 root samples were measured separately.

All of the 302 samples (126 leaves, 88 shoots, and 88 roots) were first washed in water and rinsed with distilled de-ionized water, frozen in liquid nitrogen, and stored at –70 °C. After freeze-drying, DWs were recorded. Then, the dried samples were ground into powder by using a high-speed grinder (RT-02A; Sun-Great Technology Co., New Taipei City, Taiwan). From these 302 samples, the NIR spectra and concentrations of fructose, glucose, sucrose, and starch were determined.

**NIR measurement and carbohydrate analysis.** Dry powder of *Phalaenopsis* was placed into a small ring cup (i.d., 35 mm; depth, 8 mm) and subjected to NIR measurements (NIRS 6500; FOSS NIRSystems, Laurel, MD) using a rapid content analyzer module. The reflectance spectra of the samples ranged from 400 to 2498 nm with 2-nm intervals, and the NIR spectrum of each sample was the average of 32 scans.

The levels of fructose, glucose, sucrose, and starch in all *Phalaenopsis* powder samples were measured using a chemical method to determine the reference values for the NIR analysis. The detailed experimental schematic of the carbohydrate analysis was introduced in a previous study (Hou et al., 2011). Each sample (50 mg) was extracted at 70 °C with 3 mL of 80% ethanol for 30 min. One milligram of raffinose was added to each sample as the internal standard. The tissue suspension was then centrifuged and re-extracted twice to collect the pooled supernatants (n = 3). After deionization, filtration, and appropriate dilution, the samples were subjected to HPAEC ( Dionex, Sunnyvale, CA) with pulsed amperometric detection (ED50; Dionex) equipped with a CarboPac PA1 column (CarboPac™ PA1 Analytical, 4 × 250 mm; Dionex) for 20 min. The separated carbohydrates (fructose, glucose, and sucrose) were quantified by comparisons with known standards (Hou et al., 2011).

Residue from the ethanol extraction was reserved for the starch analysis. Then, residue was gelatinized for 6 h, dried, suspended, and incubated. After incubation, the samples were filtered and diluted to 10 mL with HPLC-grade water. The subsequent procedure was similar to that of the soluble sugar analysis described previously. The amount of starch was determined by the glucose released after adjusting to an external standard (Hou et al., 2011). The concentrations of fructose, glucose, sucrose, and starch in organs were calculated by multiplying their concentrations by the DW of the organs.

**Data analysis.** Full visible to NIR wavelength range (400–2498 nm) was used as the input for spectral analysis. Two standard multivariate analysis methods, modified partial least-squares regression (MPLSR) and stepwise multiple linear regression (SMLR), were used to explore the relationships between the reflectance spectra and the concentrations of carbohydrates of *Phalaenopsis*. Then, the calibration models of fructose, glucose, sucrose, and starch were built.

MPLSR, an extension of partial least-squares regression (PLSR) by normalization, abides by the principle of normalization of the spectra and constituent values before PLSR, which is a standard tool in chemometrics and has been widely used in the agricultural and engineering fields (Wold et al., 2001). For spectral analysis using PLSR, the spectra can be considered to be composed of several principal components (PCs) and expressed as a “factor” in the PLSR algorithm. The factor’s
sequence is determined by its influence (i.e., a more important factor is ranked earlier in the order). Because PLSR analysis uses information from spectral bands, analysis results can be improved by selecting the appropriate number of factors and specific wavelength ranges.

The SMLR analysis method selects specific wavelengths according to the F test \( (F > 3) \) of null hypothesis testing (Chang et al., 1998). To build calibration models over numerous wavelengths, the SMLR algorithm chooses the most important specific wavelength from the major molecular bonding region of the objects, and the second most important specific wavelength is usually chosen from the region located in the combination of related molecular bonding or the overtone of complementary bonding and by analogy. When adding a new wavelength for training, the algorithm is based on the previously selected specific wavelengths to continue finding the wavelength. This allows for the highest multiple coefficient of determination \( (r^2) \) and the minimum prediction error, thus determining whether such a wavelength can replace the current specific wavelength. In the case of poor competency of the newly added wavelength for training, the algorithm is trained.

The addition of light-scattering effects into the spectra is inevitable when using NIR to measure powder samples, especially when the particle size is not uniform. Therefore, multiplicative scatter correction (MSC) could be used to allow additive and multiplicative transformation of the spectra. After MSC treatment, the spectra of *Phalaenopsis* powder not only reduced the physical impact of nonuniform particles (Helland et al., 1995; Maleki et al., 2007) but also confirmed the linearity of the spectral information (Isaksson and Næs, 1988), which contributed to subsequent linear regression analyses (Thennadil et al., 2006). The spectra of *Phalaenopsis* powder after MSC was subjected to three independent treatments: 1) smoothing; 2) smoothing with the first derivative; and 3) smoothing with the second derivative. These treatments were used to choose the best pretreatment parameters, including the smoothing points and derivative gap (range, 2–20), with the gap being greater than or equal to the smoothing points. The first and second derivatives are used to remove baseline offsets and linearly sloped baselines, respectively (de Noord, 1994). Because derivatives are calculated by taking the difference between adjacent spectral points, noise in the spectrum is usually magnified. Therefore, the Savitzky-Golay algorithm (Savitzky and Golay, 1964) that smooths the data before differentiation is typically used to remove noise.

The spectral calibration models of MPLSR and SMLR were built using WinISI II chemometric software (Infrasoft International, State College, PA). The MPLSR analysis procedure included the following: 1) spectral pretreatments; 2) selecting specific wavelength regions; 3) selecting calibration and validation sets; and 4) determining the best calibration model. In steps 1 and 2, 3-fold cross-validation was used to enable objective selection of the parameters. For each type of carbohydrate (fructose, glucose, sucrose, and starch), all samples were investigated for identifying outliers. After eliminating the outliers from all of the samples, the remaining effective samples were divided into calibration sets and validation sets at a ratio of 2:1 according to the carbohydrate concentrations of the samples. Of all samples investigated, 291, 291, 290, and 283 valid samples were retained for statistical calibration and validation of the fructose, glucose, sucrose, and starch concentrations, respectively. Samples with the highest and lowest concentrations were classified into the calibration set to cover the widest range of carbohydrate concentration. All samples were ranked according to their increasing carbohydrate concentration, with these concentrations being higher in the calibration set than in the validation set; however, distribution of the carbohydrate concentration was similar in both sets. When building the calibration model, to avoid overfitting caused by use of excessive factors, the following principles were considered: 1) the maximum number of factors is one-tenth the number of calibration sets plus two to three; 2) stop if the addition of a new factor increases the standard error of validation (SEV); and 3) stop adding new factors when the SEV is lower than the standard error of calibration (SEC). The SMLR analysis procedure included the following steps: 1) selection of the calibration sets and validation sets; 2) spectral pretreatments; and 3) determination of the best calibration model and specific wavelengths. The same calibration sets and validation sets were used for both MPLSR and SMLR analyses.

### Results and Discussion

**Correlation between NIR spectra and carbohydrate concentrations.** In the greenhouse plants, the concentrations of fructose, glucose, and starch remained at higher levels of \( \approx 2\% \text{ of DW} \), and the sucrose concentration stayed at 1\% of DW (Table 1). The levels of fructose, glucose, sucrose, and starch in the shoots and roots of dark-treated plants were approximately consistent with the results reported by previous studies.

| Sample          | Organ | Samples (no.) | Carbohydrate | Conc (% DW) |
|-----------------|-------|---------------|--------------|-------------|
| **Greenhouse**  | Leaf  | 126           | Fructose     | 0.58 4.73 2.07 0.87 0.42 |
|                 |       |               | Glucose      | 0.75 4.78 2.09 0.89 0.43 |
|                 |       |               | Sucrose      | 0.24 2.60 1.00 0.49 0.50 |
|                 |       |               | Starch       | 0.26 5.49 1.90 1.10 0.58 |
| **Dark-treated**| Shoot | 88            | Fructose     | 1.11 3.80 2.20 0.58 0.27 |
|                 |       |               | Glucose      | 0.86 3.96 2.16 0.71 0.33 |
|                 |       |               | Sucrose      | 0.36 3.46 1.22 0.76 0.63 |
|                 |       |               | Starch       | 0.17 11.89 3.12 2.35 0.75 |
| **Root**        |       | 88            | Fructose     | 0.62 4.77 2.34 0.95 0.41 |
|                 |       |               | Glucose      | 0.52 5.33 2.61 1.06 0.41 |
|                 |       |               | Sucrose      | 0.36 5.13 2.07 1.00 0.48 |
|                 |       |               | Starch       | 0.02 7.09 2.07 0.95 0.46 |

*Coefficient of variation.*
The NIR spectra of the 302 *Phalaenopsis* samples (126 leaves, 88 shoots, and 88 roots) were obtained using MSC treatment. As shown in Fig. 1A and B, absorption peaks were found in both the visible regions of blue light (448 nm) and red light (674 nm) because the chlorophyll in *Phalaenopsis* absorbs the majority of blue and red light during photosynthesis. In contrast, the root spectra (Fig. 1C) in the visible region showed a significant difference, with lower absorption occurring from the regions of blue to green light (400–528 nm) and from orange to red light (632–700 nm). This observation could have been due to less chlorophyll in the roots of *Phalaenopsis* (Motomura et al., 2008), thereby reducing the absorption of the blue and red light while reflecting the green light.

After MSC treatment, the spectra of *Phalaenopsis* were analyzed using the following pretreatments: 1) smoothing; 2) smoothing with the first derivative; and 3) smoothing with the second derivative. The results showed that the smoothed first derivative models provided the best results when analyzing fructose, glucose, sucrose, and starch. The best pretreatment parameters (smoothing points/derivative gap) of the fructose, glucose, sucrose, and starch analyses were (6/6), (6/6), (3/3), and (7/7), respectively. Both the smoothing points and derivative gap were less than 10, indicating that the NIRS 6500 spectrophotometer was stable and the spectra of *Phalaenopsis* powder exhibited minimal noise.

The correlation between the spectra of *Phalaenopsis* powder and the carbohydrates was first assessed when selecting specific wavelength regions of the spectra. The absorption bands corresponding to carbohydrates have been reported in previous studies (Chang et al., 1998; Ghosh and Roy, 1988; Giangiacomo, 2006; Giangiacomo et al., 1981; Li et al., 2015; Rambla et al., 1997). The absorption band at 2270 nm was found to be suitable for estimating fructose, glucose, and sucrose (Chang et al., 1998). Absorption bands close to 1380, 1433, 1450, 1469, 1580, 1740, 1945, 2080, 2250, and 2370 nm were correlated with fructose (Ghosh and Roy, 1988; Giangiacomo, 2006; Giangiacomo et al., 1981). Absorption bands near 1433, 1520, 1590, 1730, 1902, 1950, and 2255 nm were correlated with glucose (Ghosh and Roy, 1988; Giangiacomo, 2006). Those close to 1433, 1928, 2254, and 2286 nm were correlated with sucrose (Giangiacomo, 2006; Giangiacomo et al., 1981; Rambla et al., 1997); however,
those near 1678, 1735, and 1913 nm were correlated with starch (Li et al., 2015). Regarding the spectra with the best pretreatment parameters, the correlation coefficients of fructose, glucose, sucrose, and starch were distributed as shown in Fig. 2, and the threshold value (|r| > 0.65) was set to determine the degree of correlation. Because the influence of water absorption on the spectra of Phalaenopsis powder had been eliminated, it was not necessary to avoid the absorption bands of O-H bond near 1200, 1450, and 1950 nm (Rambla et al., 1997). The results showed that highly correlated bands were found in both the visible and NIR regions. Absorption bands of fructose, glucose, and sucrose were similar, whereas those of the first derivative spectra were located in the regions of the blue and red light, fourth overtone of the C-H bond, third overtone of the O-H bond, combination of the first overtone of the C-H bond, first overtone of the O-H bond, first overtone of the C-H bond, second overtone of the C=O bond stretch, combination of the O-H bond, and combination of two C-H bonds. Regarding starch, absorption bands of the first derivative spectra were located in the regions of the blue and red light, third overtone of the C-H bond, second overtone of the O-H bond, second overtone of the C-H bond, combination of the first overtone of the C-H bond, first overtone of the O-H bond, first overtone of the C-H bond, second overtone of the C=O bond stretch, and first overtone of the C=O bond stretch. Part of the absorption bands was correlated with the vibrations from the O-H functional groups, which are abundant in carbohydrates. The amount of O-H functional groups varies in different types of carbohydrates. Although the chemical structures of carbohydrates are similar and exhibit similar NIR absorption peaks, they can still be distinguished from the spectra based on the analysis of their wavelength positions and absorption magnitude (Rambla et al., 1997). The absorption bands of both greenhouse plant and dark-treated plant samples in the NIR region were mostly located in the overtones as well as in combinations of the C-H and O-H bonds, indicating that C-H and O-H bonds were the main absorbers of NIR light. Therefore, the wavelength ranges of 750–900, 1300–2000, and 2200–2400 nm were the major absorption bands (according to the absorption bands of C-H and O-H bonds in the spectrum) (Fig. 2), and these wavelengths can be used to provide a basis for selecting the appropriate specific wavelength regions when conducting MPLSR and SMLR analyses. It is noteworthy that the dominance of blue light and red light in the visible region could be attributed to the differences in the colors of the leaves, shoots, and roots. This observation indicated that the amount of chlorophyll contained in various organs of a Phalaenopsis plant affects the performance of the specific wavelength regions. Regarding the spectral band from 400 to 700 nm, which belonged to the absorption band of blue light to red light, color information was also reflected in the spectra.

**Carbohydrates quantification using MPLSR.** Of all the samples investigated, 291, 291, 290, and 283 valid Phalaenopsis samples were retained for statistical calibration and validation of the fructose, glucose, sucrose, and starch concentrations, respectively. For each type of carbohydrate, the samples in the calibration sets and validation sets were contributed from the greenhouse plants, shoots, and roots. As shown in Table 2, no significant differences were found in the mean, standard deviation, or coefficient of variation (CV) of the effective samples, calibration sets, and validation sets, indicating that the distributions of carbohydrate concentrations of the two sample groups were consistent.

Quantitative analyses of the carbohydrate concentrations in Phalaenopsis plants were conducted using MPLSR with the full wavelength range spectra (400–2498 nm), as shown in Table 3. The best spectral calibration model of the fructose concentration was found with the smoothed first derivative spectra and 13 factors. The smoothing points and the derivative gap were set at 6. The wavelength ranges of 400–450, 750–900, 1300–1400, 1450–1500, 1550–2000, 2050–2100, and 2300–2400 nm were used, resulting in $R_c = 0.96$, SEC = 0.22% DW, SEV = 0.28% DW, and bias = -0.01% DW. The best spectral calibration model of the glucose concentration was identified using the

![Fig. 2. Correlation coefficient distributions between the near-infrared (NIR) spectra and the carbohydrate concentrations.](image)

Table 2. Carbohydrate concentrations of effective samples, calibration set, and validation set of Phalaenopsis plants.

| Carbohydrate | Sample group | Samples (no.) | Conc (% DW) | Minimum | Maximum | Mean | SD | CV
|--------------|--------------|---------------|-------------|---------|---------|------|----|----|
| Fructose     | Effective samples | 291          | 0.58 | 4.77 | 2.17 | 0.81 | 0.37 |
|              | Calibration set | 194          | 0.58 | 4.77 | 2.17 | 0.81 | 0.37 |
|              | Validation set | 97           | 0.62 | 4.73 | 2.17 | 0.81 | 0.37 |
| Glucose      | Effective samples | 291          | 0.52 | 4.94 | 2.25 | 0.90 | 0.40 |
|              | Calibration set | 194          | 0.52 | 4.94 | 2.25 | 0.90 | 0.40 |
|              | Validation set | 97           | 0.75 | 4.78 | 2.25 | 0.90 | 0.40 |
| Sucrose      | Effective samples | 290          | 0.24 | 5.13 | 1.34 | 0.83 | 0.62 |
|              | Calibration set | 193          | 0.24 | 5.13 | 1.35 | 0.84 | 0.62 |
|              | Validation set | 97           | 0.36 | 4.87 | 1.34 | 0.83 | 0.62 |
| Starch       | Effective samples | 283          | 0.50 | 6.85 | 2.19 | 1.14 | 0.52 |
|              | Calibration set | 189          | 0.50 | 6.85 | 2.19 | 1.14 | 0.52 |
|              | Validation set | 94           | 0.79 | 6.84 | 2.20 | 1.14 | 0.52 |

*Coefficient of variation.
Table 3. Prediction of the carbohydrate concentrations in *Phalaenopsis* plants by modified partial least-squares regression models in the wavelength 400–2498 nm.

| Carbohydrate | Spectrum | Wavelength range (nm)
|--------------|----------|-------------------------|
| Fructose     | First derivative | 400–450
|              |           | 750–900
|              |           | 1300–1400
|              |           | 1450–1500
|              |           | 1550–2000
|              |           | 2050–2100
|              |           | 2300–2400
| Glucose      | First derivative | 450–500
|              |           | 650–750
|              |           | 850–900
|              |           | 1000–1098
|              |           | 1250–1300
|              |           | 1500–1550
|              |           | 1600–1650
|              |           | 1700–1800
|              |           | 1850–2050
|              |           | 2100–2150
|              |           | 2200–2300
|              |           | 2350–2450
| Sucrose      | First derivative | 550–600
|              |           | 650–700
|              |           | 1000–1050
|              |           | 1600–1650
|              |           | 1700–1800
|              |           | 1900–1950
|              |           | 2000–2100
|              |           | 2200–2300
|              |           | 2400–2450
| Starch       | First derivative | 700–750
|              |           | 850–950
|              |           | 1550–1700
|              |           | 1750–1800
|              |           | 2200–2300

| Smoothing points/derivative gap | Latent variables (no.) | Calibration set | Validation set |
|---------------------------------|-------------------------|----------------|----------------|
|                                 |                         | *R*<sub>c</sub> | SEC (% DW)     | SEV (% DW) | Bias (% DW) | RPD  |
| Fructose First derivative       | 6 / 6                   | 13             | 0.96           | 0.22       | 0.28        | -0.01 | 2.86 |
| Glucose First derivative        | 6 / 6                   | 13             | 0.96           | 0.26       | 0.32        | 0.01  | 2.83 |
| Sucrose First derivative        | 3 / 3                   | 13             | 0.96           | 0.24       | 0.31        | -0.03 | 2.65 |
| Starch First derivative         | 7 / 7                   | 12             | 0.91           | 0.47       | 0.56        | -0.02 | 2.04 |

First derivative = first-order spectral differentiation.

Wavelength increment = 2 nm.

Smoothing points = the number of data points used for each smoothing; derivative gap = gap size of derivative.

Latent variables = number of factors.

*R*<sub>c</sub> = correlation coefficient of calibration; SEC = standard error of calibration.

SEV = SE of validation; bias = the average difference between the predicted values and the measured values; RPD = ratio of SEV to SD.
smoothed first derivative spectra and 13 factors. The smoothing points and the derivative gap were set at 6. The wavelength ranges of 450–500, 650–750, 850–900, 1000–1098, 1250–1300, 1500–1550, 1600–1650, 1700–1800, 1850–2050, 2100–2150, 2200–2300, and 2350–2450 nm were used, resulting in $R_c = 0.96$, SEC = 0.26% DW, SEV = 0.32% DW, and bias = 0.01% DW. The best spectral calibration model of the sucrose concentration was obtained when applying the smoothed first derivative spectra and 13 factors. The smoothing points and derivative gap were set at 3. The wavelength ranges 550–600, 650–700, 1000–1050, 1600–1650, 1700–1800, 1900–1950, 2000–2100, 2200–2300, and 2400–2450 nm were used, resulting in $R_c = 0.96$, SEC = 0.24% DW, SEV = 0.31% DW, and bias = –0.03% DW. The best spectral calibration model of the starch concentration was acquired with the smoothed first derivative spectra and 12 factors. The smoothing points and derivative gap were set at 7. The wavelength ranges 700–750, 850–950, 1550–1700, 1750–1800, and 2200–2300 nm were used, resulting in $R_c = 0.91$, SEC = 0.47% DW, SEV = 0.56% DW, and bias = –0.02% DW. Satisfactory outcomes were acquired for fructose, glucose, sucrose, and starch. The calibration models built based on the first derivative spectra were better than those based on the original and second derivative spectra (data not shown), confirming that heterogeneous particles of *Phalaenopsis* powder affect the spectral absorption. These observations indicated that spectral pretreatments enhanced the accuracy and predictability of the calibration models. The specific wavelength regions selected based on their characteristic responses to the carbohydrates were mainly distributed in 850–950 (third overtone of the C-H bond), 1400–1800 (first overtones of the O-H bond and C-H bond), 1850–2100 (second overtone of the C=O bond stretch and the combination of the O-H bond), and 2200–2450 nm (the combinations between the C-H bonds and between the C-H bond and C-C bond), which were consistent with the absorption bands shown in Fig. 2. The results indicated that the NIR region contained more information related to the carbohydrates. The differences in absorption of the leaves, shoots, and roots in the visible region also qualified both 400–500 (blue light) and 650–700 nm (red light) to be used as specific wavelength regions. Due to the spectral differences between the calibration sets and validation sets, prediction results of the validation sets were somewhat worse than those of the calibration sets when using the calibration models. Relationships between the predicted and reference concentrations of fructose, glucose, sucrose, and starch are shown in Fig. 3. Among them, the highest multiple correlation coefficient for the calibration set was 0.96. The carbohydrate concentrations predicted by the calibration models were mainly affected by bias; hence, predictability can be improved by eliminating the bias calculated from a set of representative samples.

**Carbohydrates quantification using SMLR.** The SMLR results of the carbohydrate concentrations in the *Phalaenopsis* plant using the full wavelength range of 400–2498 nm are shown in Table 4. The best spectral calibration model of the fructose concentration was found when the first derivative spectra were used. The smoothing points and derivative gap were set at 10. The specific wavelengths of 1358, 1594, 1432, 1858, 1974, and 2270 nm were used, resulting in $R_c = 0.94$, SEC = 0.28% DW, SEV = 0.35% DW, and bias = –0.05% DW. The best spectral calibration model of the glucose concentration was identified with the first derivative spectra. The smoothing points and derivative gap were set at 11, and a combination of 6 wavelengths (494, 1590, 1858, 2270, 1968, and 1442 nm) was used, resulting in $R_c = 0.92$, SEC = 0.35%; SEV = 0.42%; and bias = 0.03% DW. The best spectral calibration model of the
Table 4. Prediction of the carbohydrate concentrations in *Phalaenopsis* plants by stepwise multiple linear regression models in the wavelength 400–2498 nm.

| Carbohydrate | Spectrum *z* | Smoothing points / derivative gap *y* | Specific wavelengths (nm) | Calibration set | Validation set | Smoothing wavelengths (nm) | Calibration set | Validation set | Smoothing wavelengths (nm) | Calibration set | Validation set | Smoothing wavelengths (nm) | Calibration set | Validation set | Smoothing wavelengths (nm) | Calibration set | Validation set | Smoothing wavelengths (nm) |
|--------------|--------------|-------------------------------------|---------------------------|-------------------|----------------|---------------------------|-------------------|-----------------|---------------------------|-------------------|-----------------|---------------------------|-------------------|-----------------|---------------------------|-------------------|-----------------|---------------------------|
| Fructose     | First derivative | 10/10                               | 1358, 1594, 1432, 1858, 2270 | 0.94              | 0.28            | 0.35                       | 0.00              | 0.83           | 0.64                       | 0.61              | 0.00            | 0.00                       | 1.88              | 2.29            |
| Glucose      | First derivative | 11/11                               | 494, 1590, 1858, 2270, 1916, 1788 | 0.92              | 0.35            | 0.42                       | 0.03              | 0.03           | 0.03                       | 2.14              | 1.67            |
| Sucrose      | First derivative | 6/6                                 | 2340, 1466, 1076, 546, 2270, 1718 | 0.85              | 0.45            | 0.50                       | –0.05             | –0.03          | –0.03                      | 2.29              | 1.88            |
| Starch       | First derivative | 2/2                                 | 2490, 756, 1746, 2332, 1916, 1788 | 0.83              | 0.64            | 0.61                       | 0.00              | 0.00           | 0.00                       | 1.88              | 2.29            |

*First derivative = first-order spectral differentiation.

*Smoother points = number of data points used for each smoothing; derivative gap = gap size of derivative.

*Rc = correlation coefficient of calibration; SEC = standard error of calibration.

*SEV = SE of validation; bias = the average difference between the predicted values and the measured values; RPD = ratio of SEV to SD.

Sucrose concentration was attained when the first derivative spectra were used. The smoothing points and derivative gap were set at 6. The specific wavelengths 2340, 1466, 1076, 546, 2270, and 1718 nm were used, which yielded Rc = 0.85, SEC = 0.45% DW, SEV = 0.50% DW, and bias = –0.03% DW. The best spectral calibration model of the starch concentration was acquired when the first derivative spectra were used. The smoothing points and derivative gap were set at 2, and a combination of 6 wavelengths (2490, 756, 1746, 2332, 1916, and 1788 nm) was used, resulting in Rc = 0.83, SEC = 0.64% DW, SEV = 0.61% DW, and bias = 0% DW. The best calibration models of the carbohydrates were built using the first derivative spectra, which were consistent with the results shown in Table 3. The selected specific wavelengths provided in Table 4 were mainly distributed in 1400–1800 (first overtones of the O-H bond and C-H bond), 1850–2000 (second overtone of the C=O bond stretch), and 2250–2350 nm (the combination between the C-H bonds), consistent with the aforementioned absorption bands shown in Fig. 2. These wavelengths were also similar to those listed in Table 3, with only a small number of wavelengths beyond those selected through the MPLSR analyses. Because establishment of a SMLR model mainly based on the combination of a small number of specific wavelengths provided less spectral information than an MPLSR model, the analysis results seemed slightly inferior to those provided in Table 3. In summary, MPLSR achieved quantitative analysis results of carbohydrate concentrations with levels suitable for inspection.

**Conclusions**

In summary, MPLSR and SMLR were individually integrated with NIR spectral analysis to quantify the carbohydrate concentrations in *Phalaenopsis*, and satisfactory results were obtained. Spectral pretreatments of MSC in combination with the smoothed derivative spectra reduced the spectral noise and baseline offsets. The specific wavelength regions or specific wavelengths selected based on the absorption bands corresponding to carbohydrates can effectively improve the predictability of the calibration models. These wavelengths can also be of great use when designing instrumentation using spectral imaging techniques to achieve nondestructive determination of carbohydrate reserves in the leaves of *Phalaenopsis* plants. The results of this study showed that NIR can successfully predict the carbohydrate concentrations in *Phalaenopsis* plants and can serve as a rapid analytical screening method.

**Literature Cited**

Arrizon, J., S. Morel, A. Gschaedler, and P. Monsan. 2010. Comparison of the water-soluble carbohydrate composition and fructan structures of *Agave tequilana* plants of different ages. Food Chem. 122:123–130.

Chang, W.H., S. Chen, and C.C. Tsai. 1998. Development of a universal algorithm for use of NIR in estimation of soluble solids in fruit juices. Trans. Amer. Soc. Agr. Eng. 41:1739–1745.

Chen, W.S., H.Y. Liu, Z.H. Liu, L. Yang, and W.H. Chen. 1994. Gibberellic and temperature influence carbohydrate content and flowering in *Phalaenopsis*. Plant Physiol. 90:391–395.

Chuang, Y.K., Y.P. Hu, I.C. Yang, S.R. Delwiche, Y.M. Lo, C.Y. Tsai, and S. Chen. 2014. Integration of independent component analysis with near infrared spectroscopy for evaluation of rice freshness. J. Cereal Sci. 60:238–242.

Cruz-Cárdenas, C.I., M.L. Miranda-Ham, L.A. Castro-Concha, J.R. Ku-Cauich, R. Vergauwen, T. Reijnders, W. Van den Ende, and R.M. Escobedo-Gracia Medrano. 2015. Fructans and other water soluble carbohydrates in vegetative organs and fruits of different *Musa* spp. accessions. Front. Plant Sci. 6:395–404.

de Noord, O.E. 1994. The influence of data preprocessing on the robustness and parsimony of multivariate calibration models. Chemom. Intell. Lab. Syst. 23:65–70.

Di Stefano, V., G. Avellone, D. Bongiorno, V. Cunsolo, V. Muccilli, S. Sforza, A. Dossena, L. Drahos, and K. Vékey. 2012. Applications of liquid chromatography-mass spectrometry for food analysis. J. Chromatography 1259:74–85.

Ghosh, S. and R.B. Roy. 1988. Quantitative near-infra-red analysis of reducing sugar from the surface of cotton. J. Textil. Inst. 79:504–510.

Giangiacomo, R. 2006. Study of water-sugar interactions at increasing sugar concentration by NIR spectroscopy. Food Chem. 96:371–379.

Giangiacomo, R., J.B. Magee, G.S. Birth, and G.G. Dull. 1981. Near-infrared reflectance spectroscopy. J. Food Sci. 46:531–534.

Giovanni, G., N. Sinelli, R. Beghi, R. Guidetti, and E. Casiraghi. 2014. NIR spectroscopy for the optimization of postharvest apple management. Postharvest Biol. Technol. 87:13–20.

Helland, I.S., T. Næs, and T. Isaksson. 1995. Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data. Chemom. Intell. Lab. Syst. 29:233–241.

Hou, J.Y., W.B. Miller, and Y.C.A. Chang. 2011. Effects of simulated temperature on carbohydrate concentrations in *Phalaenopsis* plants and can serve as a rapid analytical screening method.

*J. Amer. Soc. Hort. Sci.* 143(6):494–502. 2018.
Isaksson, T. and T. Næs. 1988. The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. Appl. Spectrosc. 42:1273–1284.
Kataoka, K., K. Sumitomo, T. Fudano, and K. Kawase. 2004. Changes in sugar content of Phalaenopsis leaves before floral transition. Scientia Hort. 102:121–132.
Li, J., M.-G.C. Danao, S.F. Chen, S. Li, V. Singh, and P.J. Brown. 2015. Prediction of starch content and ethanol yields of sorghum grain using near infrared spectroscopy. J. Near Infrared Spectrosc. 23:85–92.
Liu, Y.C., C.H. Liu, Y.C. Lin, C.H. Lu, W.H. Chen, and H.L. Wang. 2016. Effect of low irradiance on the photosynthetic performance and spiking of Phalaenopsis. Photosynthetica 54:259–266.
Magwaza, L.S. and U.L. Opara. 2015. Analytical methods for determination of sugars and sweetness of horticultural products - A review. Scientia Hort. 184:179–192.
Maleki, M.R., A.M. Mouazen, H. Ramon, and J. De Baerdemaecker. 2007. Multiplicative scatter correction during on-line measurement with near infrared spectroscopy. Biosyst. Eng. 96:427–433.
Maniwara, P., K. Nakano, D. Boonyakiat, S. Ohashi, M. Hiroi, and T. Tohyama. 2014. The use of visible and near infrared spectroscopy for evaluating passion fruit postharvest quality. J. Food Eng. 143:33–43.
Motomura, H., O. Ueno, A. Kagawa, and T. Yukawa. 2008. Carbon isotope ratios and the variation in the diurnal pattern of malate accumulation in aerial roots of CAM species of Phalaenopsis (Orchidaceae). Photosynthetica 46:531–536.
Rambla, F.J., S. Garrigues, and M. de la Guardia. 1997. PLS-NIR determination of total sugar, glucose, fructose and sucrose in aqueous solutions of fruit juices. Anal. Chim. Acta 344:41–53.
Savitzky, A. and M.J.E. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36:1627–1639.
Thennadil, S.N., H. Martens, and A. Kohler. 2006. Physics-based multiplicative scatter correction approaches for improving the performance of calibration models. Appl. Spectrosc. 60:315–321.
Theron, K.I. and G. Jacobs. 1996. Changes in carbohydrate composition of the different bulb components of Nerine bowdemi W. Watson (Amaryllidaceae). J. Amer. Soc. Hort. Sci. 121:343–346.
Williams, P.C. and D.C. Sobering. 1993. Comparison of commercial near-infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J. Near Infrared Spectrosc. 1:25–32.
Wold, S., M. Sjöström, and L. Eriksson. 2001. PLS-regression: A basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58:109–130.