Accurate evaluation of sugar contents in stingless bee (Heterotrigona itama) honey using a swift scheme

Kuan Wei Se¹, Raja Kamarulzaman Raja Ibrahim²,⁎, Roswanira Abdul Wahab³,⁎⁎, Sib Krishna Ghoshal⁴

¹ School of Graduate Studies, SPS, Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor, Malaysia
² Laser Centre, Department of Physics, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor, Malaysia
³ Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor, Malaysia
⁴ Department of Physics, AMORG, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor, Malaysia

⁎ Corresponding authors.
⁎⁎ E-mail addresses: rkamarulzaman@utm.my (R.K.R. Ibrahim), roswanira@kimia.fs.utm.my, roswanira711@gmail.com (R.A. Wahab).

ABSTRACT

We propose a rapid scheme for the precise determination of the sugars content in 62 stingless bee (Heterotrigona itama) honeys harvested at diverse locations in Malaysia. This scheme combined Fourier transform infrared and attenuated total reflectance (FTIR-ATR) techniques with chemometric partial least square (PLS) regression analysis, wherein high performance liquid chromatography (HPLC) is used as reference technique. Results revealed that such honeys contained (per 100 g) high amount of maltose (9.74–54.3 g) and low quantity of fructose (5.37–19.9 g) as well as glucose (3.91–27.2 g). The first derivative data transformation in the wavenumber within 1500 to 750 cm⁻¹ produced the optimum PLS prediction performance with corresponding R² value over 0.976 and very low standard error of prediction (per 100 g) for fructose (0.562 g), glucose (0.855 g) and maltose (1.55 g). The results proved that the proposed scheme is potentially useful for fast and accurate quantification of sugars content in H. itama honeys.

ARTICLE INFO

Chemical compounds studied in this article:
- Fructose (PubChem CID: 5984)
- Glucose (PubChem CID: 79025)
- Sucrose (PubChem CID: 5988)
- Maltose (PubChem CID: 25615261)
- Acetonitrile (PubChem CID: 6342)

Keywords:
- Food analysis
- Food composition
- Stingless bees (Heterotrigona itama) honey
- Sugar contents
- FTIR-ATR
- HPLC
- Chemometrics
- PLS regression analysis

1. Introduction

Honey is a type of natural sweetener produced by honey bees, collected from the nectar of flowers. Invertase, an enzyme produced in the hypopharyngeal gland of honey bees hydrolyzes the disaccharides present in the nectar into monosaccharides (Almeida-Muradian, 2013), which subsequently form the dominant sugar composition in honey. Apart from sugar and the enzyme, honey also contains minute amount of organic acids, proteins and minerals. This composition tends to vary according to the bee species, floral source and geographical region (Karabagias et al., 2014). Honey produced by stingless bees called Heterotrigona itama (H. itama) is highly preferred by consumers due to its perceived therapeutic and nutritional benefits. Such honey is believed to have curative properties for throat inflammation, gastritis, cataract as well as assisting in post-birth recovery (Vit et al., 2004). The plethora of health benefits coupled with limited production of this food commodity have inflated the economic value of H. itama honey as compared to honey produced by honey bees called Apis mellifera. It is noteworthy to mention that the market price of H. itama honey is about USD 100 per kg (Shadan et al., 2017; Razak et al., 2016). Such high price is due to elevated contents of flavonoids and polyphenols present in stingless bee honeys as compared to that produced by the Apis spp. (Biluca et al., 2016; Rodriguez-Malaver, 2013; Rodriguez-Malaver et al., 2009).

The Malaysian bee farmers prefer to foster stingless bee (H. itama) species because such populations are less vulnerable to seasonal changes and capable to survive in harsh environments (Kelly et al., 2014). These are the possible reasons for the H. itama being the most domesticated species of stingless bee in Malaysia (Razak et al., 2016). Current estimation showed that the stingless bee honey sold in the Malaysian open market is largely produced by H. itama (Kelly et al., 2014). Despite the popularity of H. itama honey among consumers, studies relating the sugar composition of this premium food remained relatively limited. Hence, we aim to develop a rapid and accurate technique to identify and quantify the individual sugar components in H. itama honey collected from different geographic locations in Malaysia.

Received 2 July 2017; Received in revised form 17 November 2017; Accepted 4 December 2017
Available online 08 December 2017
Over the years, different analytical techniques such as high performance liquid chromatography (HPLC) and high performance anion-exchange chromatography integrated with pulsed amperometric detection (HPAEC-PAD) have been used to qualitatively and quantitatively evaluate various compounds in honey including sugars, flavonoids, coumarins, and antibiotic, etc. (Kek et al., 2017; Tayeb-Cherif et al., 2016; Cherif et al., 2015; Anjos et al., 2015). Despite the proven effectiveness and accuracy of HPLC and HPAEC-PAD to analyze sugar content in honey, the overall protocol remains laborious and time-consuming, as well as requires costly reagents and cumbersome disposal practices. To overcome such limitations, a fast, portable, inexpensive, and precise method for assessing the sugars content in honey is demanded. Herein, we propose a swift scheme which combines FTIR-ATR spectrometry with chemometric PLS regression analysis to identify and quantify the individual sugar in a honey sample. Typically, each test requires less than 5 min for spectral acquisition and sugar prediction. Conceptual view of the proposed scheme is illustrated in Fig. 1. The well celebrated chemometrics regression analysis (PLS regression) is used to correlate the FTIR-ATR spectral data for precise determination of various types of sugars content in H. itama honey. The term ‘swift’ used in this study signifies the rapid and facile nature of the proposed scheme to analyze the sugar composition in a honey sample. Typically, each test requires less than 5 min for spectral acquisition and sugar prediction. Conceptual view of the proposed scheme is illustrated in Fig. 1. The well

![Conceptual view of the proposed scheme for rapid evaluation of sugar content in H. itama honey.](image)

**Fig. 1.** Conceptual view of the proposed scheme for rapid evaluation of sugar content in H. itama honey. FTIR-ATR (“Fourier transform infrared together with attenuated total reflectance spectroscopy”); HPLC-RID (“high performance liquid chromatography combined with refractive index detector”); PLS (“partial least square”); \( \mathbf{X}_{\text{cal}} \) (“X-variables in the calibration dataset”); \( \mathbf{X}_{\text{val}} \) (“X-variables in the validation dataset”); \( \mathbf{Y}_{\text{predicted}} \) (predicted Y-variables); \( \mathbf{Y}_{\text{measured}} \) (“measured Y-variables”); \( R^2 \) (“coefficient of determination”); \( \text{SEC} \) (“standard error of calibration”); \( \text{SECV} \) (“standard error of cross validation”); \( \text{SEP} \) (“standard error of prediction”).

honey, the established HPLC method is used as the reference technique to explicitly profile the main sugar components and validate the data obtained using the swift scheme. Previously, reference methods such as HPLC and HPAEC-PAD have been used to authenticate the sugars content predicted using PLS-integrated FTIR spectroscopy. Many studies used the PLS regression models constructed from raw FTIR spectra (Ruoff et al., 2006; Tewari and Irudayaraj, 2004), first derivative (Anjos et al., 2015) and second derivative (Huang et al., 2016; Wang et al., 2010) as the main data preprocessing method for sugar contents evaluation. However, a comparative study between the spectral data preprocessing (raw, first or second derivative) and the predicted accuracy of the PLS model is not yet performed. Therefore, the objective of this study is to employ PLS integrated FTIR to quantitatively assess the sugar content in H. itama honey. Performance of the corresponding PLS regression models for predicting the main sugar components is compared and evaluated in terms of the raw data, pre-treated first and second derivative data as well.

2. Material and methods

2.1. Honey samples

A total of 62 samples of pure H. itama honey are collected from reputable bee farms located in five different states of Malaysia (43 from Johor, 14 from Terengganu, 3 from Kedah, 1 from Pahang and 1 from Kelantan). School of Food Science and Technology (Universiti Malaysia Terengganu, UMT) and Johor Entrepreneur of Stingless Bee Society are the main provider of these samples. For reliable outcome with improved statistics, random sampling is performed wherein several honeys are collected from different bee hives at the same farm.
considering their color variability. Supplementary Table S1 enlists the details of collected honeys in terms of their geographical location, label, and year of harvesting. The honey samples are stored in darkness at room temperature (26 ± 2 °C) prior to the following analytical measurements.

2.2. High performance liquid chromatography (HPLC)

 Sugars content in all honey samples are quantified using HPLC equipped with refractive index detector (RID) (Agilent 1100, Santa Clara, California, United States) following the AOAC Official Method 977.20. A 5% (w/v) solution of honey dissolved in ultrapure water (Milli Q system, Merck Millipore, Burlington, Massachusetts, United States) is filtered through 0.45 µm PTFE filter and injected to the HPLC system. Zorbax Carbohydrate Column (Agilent, Santa Clara, California, United States) having 5 µm particle size, 250 mm in length and 4.6 mm inner diameter is used as the stationary phase. An isotropic elution mobile phase consisting of HPLC grade (purity ~ 99.9%) acetonitrile (Qrec, Selangor, Malaysia) and water (75:25 v/v) is used. Sample volume of 40 µL is injected with a flow rate of 1.4 mL/min with the column temperature maintained at 30 °C. Run time per sample is set at 25 min synchronized with standard analytical grade (purity > 99.0%) glucose (QRec) and sucrose (Qrec) as well as HPLC grade (purity ~ 99.9%) fructose (System, Selangor, Malaysia) and maltose (Sigma-Aldrich, St. Louis, Missouri, United States). The sugars content in the honey samples are identified by comparing them with the retention time of standard sugars (fructose, glucose, sucrose and maltose). Quantitative analyses are performed by preparing standard solutions of fructose, glucose, sucrose and maltose at different concentrations. Subsequently, the calibration curves are constructed using the respective peak area of analytes. Each measurement is repeated three times to calculate the mean and standard deviation. Supplementary Fig. S1 shows the HPLC calibration curves for fructose, glucose, sucrose and maltose.

2.3. Fourier transform infrared with attenuated total reflectance (FTIR-ATR) spectroscopy

 Room temperature FTIR spectra in the wavenumber range of 4000–650 cm⁻¹ are recorded using NIR/MIR Frontier spectrometer (Perkin Elmer, Waltham, Massachusetts, United States) which is equipped with deuterated triglycerine sulphate (DTGS) detector and interfaced to computer software (Spectrum, v10.4.0, Perkin Elmer). Attenuated total reflectance (ATR) sampling accessory used a diamond and zinc selenide (ZnSe) crystal composite for spectral acquisition. Triplicate spectra for each sample are acquired with a total accumulation of 16 scans at spectral resolution of 2 cm⁻¹. Each spectrum is standardized against the background spectrum of air and corrected for baseline prior to the further analysis. The ATR crystal is gently washed by distilled water and dried with acetone wash after every measurement. Often, the background spectrum is inspected to ensure the complete absence of any residue from the previous analysis.

2.4. Chemometric analysis of FTIR-ATR spectra

 The Unscrambler X software (v10.3, CAMO, Oslo, Norway) is used for the subsequent chemometric data analysis. Partial least square (PLS) regression is constructed based on FTIR-ATR spectral data for prediction of sugars content in the studied honeys. The PLS regression algorithm is based on the mathematical correlation between 2 independent data matrices (Li et al., 2017). PLS regression model correlates the FTIR-ATR spectral data (called X variables) to the concentration of sugars in honey acquired from the reference (HPLC) method (specified as Y-variables). The X-variables represent data matrix originated from IR spectral data with rows for samples and columns for wavenumber. The X-variables are projected into small number of latent variables named factor, and the concentration of sugars present in the honey (Y-variables) is predicted.

 Generally, PLS regression model is constructed via three steps such as building of the calibration model, optimizing the model using the cross-validation and authenticating the test dataset. The entire datasets are separated into calibration dataset and validation dataset (test dataset). The calibration dataset is used for PLS calibration and the validation dataset is utilized for the prediction which in turn checked the validity of the calibration model (Gallardo-Velazquez et al., 2009). The performance of PLS calibration is evaluated by calculating the standard error of calibration (SEC) using the following equation:

\[
SEC = \sqrt{\frac{\sum_{i=1}^{N} (A_i - A_i^*)^2}{N - f - 1}}
\]

where \(N\) is the sample size, \(f\) is the number of factor in the calibration, \(A_i\) and \(A_i^*\) are the known and calculated specimen concentrations, respectively.

The leave-one-out cross validation (LOOCV) is used to identify the number of optimized latent variables (factor) based on the minimum value of predicted residual error sum of squares (PRESS). This step is necessary as to avoid over fitting of a calibration model. In the process of LOOCV, the first spectrum from the calibration dataset is removed for prediction and the remaining spectra are used to construct the calibration model. Consequently, the second spectrum is omitted for prediction and the calibration is built using the remaining spectra. Similar procedure is carried out until all the sample spectra have been considered as unknown for prediction. Standard error of cross validation (SECV) is a type of mean-square error which defines the variability of the bias (accuracy) on the PLS calibration model. The calibration performance of PLS regression is also statistically evaluated through residual predictive deviation (RPD). The RPD value signified the ratio of standard deviation (SD) of population in the calibration dataset to the SECV. RPD value is calculated using the relation:

\[
RPD = \frac{SD}{SECV}
\]

RPD estimates the quality of the developed calibration model to predict the chemical data (Naes et al., 2002; Williams, 2001). A high RPD value indicates a greater probability to accurately predict the analyte of interest outside the calibration dataset (Cozzolino et al., 2014; Naes et al., 2002; Williams, 2001). A value of RPD above 3 is considered as moderate and suitable for screening purposes. Conversely, RPD value over 5 infers the adequacy of a model to be used for accurate prediction (Huang et al., 2016; Cozzolino et al., 2014; Naes et al., 2002; Williams, 2001).

Using the validation dataset (not included in the calibration dataset), the prediction capability of the calibration model is evaluated based on the standard error of prediction (SEP) via:

\[
SEP = \sqrt{\frac{\sum_{i=1}^{N} (A_i - A_i^* - Bias)^2}{M - 1}}
\]

where \(A_i^*\) is the sample concentration computed by the calibration equation and \(M\) is the sample size in the validation dataset. SEP signifies the magnitude of error when independent samples (validation dataset) are predicted using the constructed PLS calibration model (Santos et al., 2013). It is important to note that the equation for computing SECV is same as that of SEP (Mark and Workman, 2003). The only difference is the sample size that is used for cross validation and prediction. For this purpose, twenty five randomly selected samples are considered as validation dataset and the remaining 37 samples are treated as calibration dataset. Following simple data preprocessing method (Savitzky-Golay method with second polynomial order), the first derivative \((dy/dx)\) and second derivative \((d^2y/dx^2)\) of the data is utilized. In this method, five
smoothing data points are applied to the FTIR-ATR spectral data to compare their PLS prediction performances.

3. Results and discussion

3.1. HPLC analysis of sugar content

As aforementioned, HPLC is used as the reference method for qualitative and quantitative determination of sugar content in *H. itama* honey samples. Supplementary Table S2 enlists the concentrations of fructose, glucose, sucrose and maltose in the 62 studied honeys. According to the results, the *H. itama* honey showed an average fructose concentration of 13.4 ± 3.83 g/100 g, with concentration ranging from 5.37 ± 0.01 to 19.9 ± 0.35 g/100 g. Glucose concentration is slightly higher with an average of 16.4 ± 5.86 g/100 g, corresponding to concentration between 3.91 ± 0.09 to 27.2 ± 0.19 g/100 g. While there is yet a standard available for stingless bee honey, the Codex Alimentarius (2001) states that the sum of fructose and glucose content should be higher than 60 g/100 g for conventional sting bee honey. The cumulative fructose and glucose content in the studied honey are clearly well below (29.8 g/100 g) the level set forth by the above-mentioned standard (60 g/100 g). Recently, Kek et al. (2017) reported that the average value of fructose and glucose in three Malaysian stingless bees' honey (*H. itama*) is as low as 15.8 g/100 g and 9.22 g/100 g, respectively. Another study on eleven different stingless bees (*Tetragonula laeviceps-pagdeni*) honey species from Thailand also revealed the presence of low level of fructose (17.0 g/100 g) and glucose (14.0 g/100 g) (Chuttong et al., 2016). Moreover, the fructose and glucose contents in Brazilian stingless bee honey meet the Codex Standard (de Sousa et al., 2016; Biluca et al., 2016). Later, Kek et al. (2017) demonstrated that the sucrose concentrations in Malaysian *H. itama* honey are quite high (32.3 g/100 g). On the contrary, out of our 62 honeys, only five of them displayed the presence of sucrose with an average of 3.46 ± 2.49 g/100 g (ranging from 1.01 ± 0.35 to 6.50 ± 0.14 g/100 g), which is well below the permitted limit set out by the Codex Alimentarius (2001).

Fig. 2(a–b) shows the measured sugars content in all 62 studied stingless bee honeys. Cumulative sugar (fructose + glucose + sucrose + maltose) content in nearly all samples exceeded the limit set by Codex Alimentarius (2001). It is asserted that this variation may set a new benchmark for monitoring the quality of stingless bee honeys. This departure is because all the earlier studies on sugars profiling in...
stingless bee honey overlooked the maltose content (Kek et al., 2017; Biluca et al., 2016; Sunthiprapap et al., 2012; Guerrini et al., 2009). Interestingly, detected amount of maltose in all the studied H. itama honey is widely varied and ranged from 9.74 ± 0.46 to 54.3 ± 2.66 g/100 g with an average of 33.8 ± 10.6 g/100 g. Conversely, some studies on the sugars content in Brazilian stingless bee (Melipona subnitida Duke and Melipona scutellaris Latreille) honeys showed the complete absence of maltose (de Sousa et al., 2016). The observation on distinctively high level of maltose in the studied honey is comparable to the one found in Thai stingless bee honey species (Lepidotrigona terminata, L. flavibasia, and Tetragonula laeviceps-papdeni complex) (Chuttong et al., 2016). This wide variation among honey samples from different species of stingless bee is typically influenced by the processes carried out by complex salivary secretions, the nectar collected from flowers and regional climatic conditions, as described by Jalil et al. (2017) and Anjos et al. (2015).

3.2. FTIR-ATR spectral analysis

Fig. 3 illustrates the FTIR-ATR spectra of the studied honeys in the wavenumber region of 4000–650 cm⁻¹. Two dominant absorption bands are observed at 3270 cm⁻¹ and 1642 cm⁻¹, which are allocated to the O–H stretching and O–H deformation of water, respectively (Anjos et al., 2015). The emergence of a band around 2938 cm⁻¹ is assigned to the C–H stretching of carboxylic acid and the NH₃ stretching of free amino acids (Gallardo-Velazquez et al., 2009). The observed absorption bands originated from the presence of acid (at low concentrations) in the honey samples (Tewari and Irudayaraj, 2004; Sivakesava and Irudayaraj, 2001) is excluded from the chemometric model construction. The weak absorption band around 1740 cm⁻¹ is approved to the C=O stretching that aroused mainly from the presence of carbohydrates and tiny amount of protein (N–H bending of amide I) in the honeys (Gok et al., 2015).

The spectrum for carbohydrates typically shows multiple absorption bands within the region of 1500–750 cm⁻¹ (Wang et al., 2010; Gallardo-Velazquez et al., 2009; Tewari and Irudayaraj, 2004) in which spectroscopic variability is reportedly the highest in honey (Fig. 3). This region encapsulated the deformation of –CH₂ and angular deformation of C–C–H and H–C=O linkages at band 1500–1200 cm⁻¹ (Hineno, 1977). The significant stretching vibrational modes of carbohydrates (C–O and C–C) are appeared around 1200–950 cm⁻¹ (Pataca et al., 2007). This region also allows the assessment of the fingerprints region of anomic carbon (950–750 cm⁻¹) which is often preferred for carbohydrates analysis (Gok et al., 2015). Therefore, this carbohydrate dominated region (1500–750 cm⁻¹) is further utilized to construct the model for sugars concentration prediction using partial least square (PLS) regression.

3.3. Sugars prediction via chemometric-integrated FTIR spectroscopy

Before constructing the PLS calibration model, correlation among various samples are established using principal component analysis (PCA) based on the FTIR-ATR spectra in the region of 1500–750 cm⁻¹. This is a prerequisite because all the 62 honeys are collected from diverse hives and localities of Malaysia. Fig. 4(a–c) depicts the PCA scores plot of all studied H. itama honey from raw, first and second derivative processed FTIR-ATR data in the wavenumber within 1500–750 cm⁻¹. Principal component 1 (PC1) and component 2 (PC2) revealed an accumulation with 97%, 88% as well as 73% data variability for raw, first and second derivative data, respectively. The PCA scores plots depicted the absence of any data separation or clustering among the various samples. This confirmed the samples’ solitary class and thus chosen to build the PLS calibration model. The model is constructed based on the FTIR-ATR spectra with reference to HPLC outcome. The PLS regression model predicted the contents of fructose, glucose and maltose only. The inadequacy of detected sucrose (present only in 5 out of 62 of honey samples) made it unsuitable for chemometric data analysis.

Fig. 5(a–c) displays the plot of PLS regression coefficient against the spectral region of interest (1500–750 cm⁻¹) for fructose, glucose and maltose. The high positive and negative values of PLS regression coefficient signify a strong correlation with the actual concentration of the investigated sugar. Conversely, PLS regression coefficient equivalent to zero indicates the lowest correlation. Fig. 5 reveals the difference in the contribution levels of wavenumbers in the range 1500–750 cm⁻¹ correspond to actual concentration of fructose, glucose and maltose present in the studied samples. This finding again verified that the wavenumber within 1500–750 cm⁻¹ is comprised of the most important vibrational mode of sugars and thereby affirmed the usability of this spectral region for sugars prediction.

Table 1 summarizes the PLS calibration parameters such as optimal number of factor, coefficient of determination (R²), SEC, SECV, RPD, and prediction parameters including SEP based on FTIR-ATR data (raw as well as processed first derivative and second derivative) in the region of 1500–750 cm⁻¹. The R² value denotes the correlation between the
actual and predicted values with $R^2 \approx 1$ implies perfect correlation.

The PLS calibration model produced high $R^2$ (Table 1) for fructose (0.969–0.979), glucose (0.923–0.976), and maltose (0.958–0.980) corresponding to raw, first derivative and second derivative data, respectively. The computed SEC of all PLS calibration model are in the range of 0.557–2.19 which is with the acceptable level of 3 (Gallardo-Velazquez et al., 2009). Before predicting the sugars content using validation dataset, the constructed PLS calibration model is first optimized using LOOCV. Furthermore, the calibration performance is assessed using the magnitude of SECV and RPD. The calculated RPD values are discerned to be all above 3, which indicated the good quality of the constructed calibration model for predicting the sugars content in stingless bee honeys. Meanwhile, the first derivative of FTIR-ATR data provided the highest RPD values for fructose (5.48), glucose (4.73) and
maltoose (4.74). Yet again, the emergence of high RPD values (close to 5) strongly indicated the good predictive capability (very accurate) of the calibration models for quantifying the sugars concentration in *H. itama* honeys.

Validation dataset is projected to the pre-constructed PLS calibration model to further validate the model's quality. The similar magnitude of calculated SEPs to that of SECV clearly indicated the reliability and accurate prediction performance of the constructed PLS regression model. High $R^2$ values are achieved for the predicted concentrations of fructose (0.942–0.981), glucose (0.934–0.981) and maltoose (0.898–0.977). These high $R^2$ values corresponded to the low SEP values for fructose (0.562 g/100 g), glucose (0.855 g/100 g) and maltoose (1.55 g/100 g) obtained from the first derivative data. This observation reaffirmed that the first derivative transformation of FTIR-ATR spectral data could yield the best PLS prediction performance. Conversely, it is revealed that the second derivative transformation of FTIR-ATR data is inappropriate towards predicting the maltoose content (corresponding SEP value above 3) in terms of calibration and assessments.

Table 2 depicts a comparative evaluation of the present scheme with existing art of the techniques integrated with PLS regression analysis. The prediction accuracy of PLS is compared in terms of the calibration parameter, RPD values and prediction parameters including $R^2$ and SEP (or root mean square error of prediction, RMSEP) which are extracted from the previous findings. The RPD values for fructose (5.48), glucose (4.73) and maltoose (4.74) those are acquired from our calibration models are found to be higher than the one reported by Anjos et al. (2015). Following similar scheme (Ruoff et al., 2006) the values of $R^2$ (SEP) for fructose, glucose and maltoose are obtained to be 0.841 (1.20 g/100 g), 0.943 (0.900 g/100 g) and 0.250 (0.600 g/100 g), respectively. Furthermore, utilization of near infrared (NIR) spectroscopy with PLS regression (Ruoff et al., 2007) yielded lower $R^2$ and higher SEP values compared to our prediction. Based on these findings, Ruoff concluded that the NIR and MIR spectra for maltoose ($R^2 \approx 0.25$) is uncorrelated, which indicated the incapability of such model towards predicting the maltoose content in honey.

Apart from combining infrared spectroscopy with PLS, some studies mingled Raman spectroscopy with PLS regression analysis (Mignani et al., 2016; Ozbalci et al., 2013). Moreover, the values of $R^2$ achieved by us are superior than the earlier report based on Raman spectroscopy (Table 2), which signified better prediction accuracy of the proposed models. In short, present study demonstrated that first derivative of FTIR-ATR data permits the construction of a reliable and accurate PLS model to predict the concentration of fructose, glucose, and maltoose present in *H. itama* honeys. This rapid, facile and yet accurate scheme can be adopted in practice to monitor the quality of Malaysian stingless bee honey.

The proposed swift scheme integrated the FTIR-ATR spectroscopy with PLS regression. It is more convenient due to easy handling and possibly incurring lower operating costs without requiring any extensive sample preparation steps. The present scheme typically takes around 5 min to complete sugars prediction in *H. itama* honey. Meanwhile, current standard method to determine sugar concentration in honey (AOAC Official Method 977.20) employs HPLC which is not suitable for large scale sample analysis because it not only requires toxic reagents and skilful operator but also involves tedious sample preparation procedure. As a conclusion, the proposed swift scheme offers facile and rapid evaluation of sugars content in honey, which is suitable in real field application and large scale investigation.

4. Conclusion

For the first time, this paper proposed an efficient scheme to profile accurately the sugars content in diverse Malaysian stingless bee (*Heterotrigona itama*) honeys. It is demonstrated that *H. itama* honeys contained an elevated level of maltoose (33.8 g/100 g), low level of fructose (13.4 g/100 g) and glucose (16.4 g/100 g) as well as insignificant amount of sucrose (3.46 g/100 g). It is established that FTIR-ATR spectrometry integrated with chemometric analysis may constitute a basis for routine analysis of sugars content in *H. itama* honey. The constructed PLS regression models are optimized by cross validation and validated using independent test dataset. The developed models revealed best predictive capability to estimate the fructose, glucose and maltoose contents wherein the first derivative treatment of the FTIR-ATR spectral data is exploited. The excellent attributes of the results suggested that the simplistic combination of FTIR-ATR spectrometry with PLS regression analysis is advantageous for the fast and precise prediction of sugars content in honey. This study is an effort to develop a standard for stingless bee honey’s sugar content.
Comparative evaluation of the proposed scheme with existing art-of-the techniques integrated with PLS regression analysis.

Table 1

<table>
<thead>
<tr>
<th>Pre-processed FTIR-ATR data</th>
<th>Property of interest</th>
<th>Factors</th>
<th>Calibration</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fructose</td>
<td>6</td>
<td>0.969 0.685</td>
<td>3.87 0.888</td>
</tr>
<tr>
<td></td>
<td>Glucose</td>
<td>7</td>
<td>0.976 0.870</td>
<td>5.63 1.31</td>
</tr>
<tr>
<td></td>
<td>Maltose</td>
<td>8</td>
<td>0.978 1.61</td>
<td>10.8 2.49</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Technique</th>
<th>Reference</th>
<th>Calibration parameter</th>
<th>Prediction parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R²a</td>
<td>SECb (g/100 g)</td>
</tr>
<tr>
<td>Pre-processed FTIR-ATR data</td>
<td>Anjos et al. (2015)</td>
<td>2.55</td>
<td>0.842 0.880</td>
</tr>
<tr>
<td></td>
<td>Ruoif et al. (2006)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>NIRb spectroscopy</td>
<td>Ruoif et al. (2007)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Raman spectroscopy</td>
<td>Mignani et al. (2016)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Ozbalci et al. (2013)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Present scheme</td>
<td></td>
<td>5.48</td>
<td>4.73</td>
</tr>
</tbody>
</table>

References


Conflict of interest

All authors have read and understood the Journal of Food Composition and Analysis’s policy and declare that we have no conflict of interests.

Acknowledgements

Authors are thankful to the Malaysian Ministry of Higher Education and UTM for financial support through grants (FRGS: R.J130000.7826.4F869) and (RU/GUP: Q.J130000.2526.13H67 and 17H19). We are also grateful to the President of Johor Entrepreneur of Stingless Bee Society and School of Food Science and Technology, UMT for supplying the honeys.

Appendix A. Supplementary data

Methods 6, 7710–7715.


