Pertanika J. Sci. & Technol. 26 (3): 1169 - 1182 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Determination of the Optimal Pre-processing Technique for Spectral Data of Oil Palm Leaves with Respect to Nutrient

Helena Anusia James Jayaselan*, Wan Ishak Wan Ismail, Nazmi Mat Nawi and Abdul Rashid Mohamed Shariff

Department of Biological and Agricultural Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

Precision agriculture with regard to crop science was introduced to apply only the required and optimal amount of fertiliser, which inspired the present study of nutrient prediction for oil palm using spectroradiometer with wavelengths ranging from 350 to 2500 nm. Partial least square (PLS) method was used to develop a statistical model to interpret spectral data for nutrient deficiency of nitrogen (N), phosphorus (P), potassium (K), magnesium (Mg), calcium (Ca) and boron (B) of oil palm. Prior to the development of the PLS model, pre-processing was conducted to ensure only the smooth and best signals were studied, which includes the multiplicative scatter correction (MSC), first and second derivatives and standard normal variate (SNV), Gaussian filter and Savitzky-Golay smoothing. The MSC technique was the optimal overall pre-treatment method for nutrients in this study, with highest prediction R^2 of 0.91 for N and lowest RMSEP value of 0.00 for P.

Keywords: Nutrients, oil palm, pre-processing, partial least square, spectroradiometer

INTRODUCTION

Oil palm is commonly known as the golden crop as it is the most globally produced and consumed oil (Phin, Stephen, & Markus, 2009). Crude palm oil (CPO) and kernel oil are not only consumed as vegetable oil that is a staple food ingredient for thousands of consumer products, but it also the source of

ARTICLE INFO

Article history:
Received: 25 July 2017
Accepted: 08 January 2018

E-mail addresses:
helena_anusia@yahoo.com (Helena Anusia James Jayaselan)
wiwi@upm.edu.my (Wan Ishak Wan Ismail)
nazmimat@upm.edu.my (Nazmi Mat Nawi)
rashidsnml@gmail.com (Abdul Rashid Mohamed Shariff)
*Corresponding Author

Current Affiliation:
Nazmi Mat Nawi
Institute of Plantation Studies, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia.

ISSN: 0128-7680

biodiesel renewable energy in fuel industry (Liaghat et al., 2014; Sheil et al., 2009; Carter, Finley, Fry, Jackson, & Willis, 2007; Graboski & McCormick, 1998). It is also widely applied in chemical industry, namely cosmetics, soap and rubber chemicals industry (Keng et al. 2009). A new study from Europe Economic Community show palm oil provides billions of dollars in gross domestic product (GDP), creating millions of sustainable jobs especially for the population in Malaysia (Basiron, 2016; Simeh & Ahmad, 2001).

The scale and reach of palm oil's positive impact creates high demand and the need to boost its productivity (Corley, 2009). Every year, more research is conducted in improving oil palm yield per hectare. Studies have shown that nutrient amendments in inorganic fertilisers lead to changes in the associated CPO constituents, while the quantitative information on the uptake of nutrients in a palm is still very sparse, including the size and growth of palm roots (Hamzah, Salleh, Wong, & Sarmani, 2016; Henson & Chai, 1997). Focusing on the application of fertilisers, study plots were conducted to evaluate the requirement of nutrients in oil palm tree to research on its nutrient intake, demands and losses as well as rate of growth (Dubos, Snoeck, & Flori, 2016; Hashim et al., 2001). Dwelling deep into precision agriculture, many methods of acquiring data were explored, namely Geographic Information System (GIS), Global Positioning System (GPS), and Remote Sensing (RS) that leads to higher potential of increasing yields (Khorramnia, Shariff, Rahim, & Mansor, 2014; Wahid, Xaviar, Tarmizi, & Ibrahim, 2002).

A new and revolutionary method is required to increase the accuracy of acquired data, leading to research on sensors such as the low altitude remote sensing (LARS) multispectral sensors remote temperature sensing and spectroradiometer (Jensen, Apan, Young, & Zeller, 2007; Sugiura, Noguchi, & Ishii, 2005; Jones, 2004). Spectroradiometer was used to obtain reflectance data of oil palm leaflets to determine the level of nutrient (nitrogen (N), phosphorus (P) and potassium (K)) (Jayaselan et al., 2017). The conventional method of detecting nutrient levels in oil palm leaflet was by using destructive methods with application of chemical analysis (Bechlin, Fortunato, Silva, Ferreira, & Neto, 2014). However, it takes a long time to determine the nutrient levels using the conventional methods; thus, immediate in-field results are desirable which would help to make quick judgement on fertiliser applications as part of the study on palm nutrient intake.

In order to obtain robust results using spectral reflectance data, preliminary studies involving the pre-processing method was required prior to its model development, to increase the accuracy of the models and differentiate the nutrients levels in oil palm. The pre-processing methods were divided into four categories of smoothing, derivatives, normalisations and standardisation (Nicolai et al., 2007). For this study, spectral reflectance data from oil palm leaves were acquired for nutrient level detection and classifications. However, prior to this, suitable pre-processing methods are analysed to ensure highly accurate results in this study. The content of interest for this study includes macronutrients nitrogen (N), phosphorus (P), potassium (K), magnesium (Mg) and calcium (Ca) as well as micronutrient boron (B) which are crucial for better productivity and yield (Corley & Mok, 2008).

The purpose of data pre-processing is to ensure a good correlation between spectral data and the concentration values while enhancing spectral features and removing unwanted variation of artefact in data. Hence, it is important to select the most suitable pre-treatment

technique because it can either make or break the data subsequent analysis (Engel et al., 2013). The specified objective was to determine the optimal pre-processing technique for spectral reflectance data of oil palm leaves with respect to nutrients N, P, K, Mg, Ca, and B using partial least square (PLS) model.

MATERIALS AND METHOD

Sample Acquisition

A total of 30 oil palm leaves were taken from frond number 17 (F17) of trees aged 20 years. The choice of f rond number determined was based on results and discussion of previous researchers (Hashim et al., 2001; Rajaratnam, Chan, & Goh, 1977). Plots from an oil palm plantation located at United Plantations, Jenderata, Malaysia were used for this experiment. The leaves from both right and left side of frond 17 were cut at its mid-section as shown in Figure 1. All the leaf samples were cleaned with distilled water before taking reading using a portable contact probe spectroradiometer (Jayaselan et al., 2017). This Analytical spectral device (ASD) called the Fieldspec 4 is a full range spectroradiometer (FRS) that has wavelength ranging from 350 to 2500 nm which provides uniform visible (vis), near infrared (NIR) and short-wave infrared (SWIR) manufactured by Boulder, Colorado, USA. Once the spectral reflectance data of leaf samples were acquired, the lot was sent to the nearest lab to obtain its nutrient level using conventional destructive method of chemical analysis.

Figure 1. Samples were obtained from mid-section of frond 17

The digital numbered (DN) reflectance data acquired using spectrometer was then converted into percentage of reflectance using ASD ViewSpec software. The data selected was in reflectance format rather than absorption or transmittance for better representation of nutrients level (Lelong et al., 2010; Kruse, 1994). In spectroscopy measurements, reliable calibration and prediction models are vital, thus removal of unwanted background noise and smoothing of data is necessary (Nawi, Jensen, & Chen, 2013).

Spectral Pre-Processing Techniques

It is necessary to reduce the signal contamination due to variations of sunlight scatters, other backgrounds reflectance effects, the instrumental noise and dirt on leaves (Nawi, Jensen, &

Chen, 2012; Cen & He, 2007). These disturbances usually act as baseline on the reflectance spectra, which can be a first approximation and considered as linear (Jamshidi, Minaei, Mohajerani, & Ghassemian, 2012). There is no efficient methodology yet on the selection of parameter such as smoothing window size, degree of polynomial fit and derivative order. Therefore, a few sets of combination of these parameters were tested and the optimal solution extracted. In random, six types of pre-processing techniques were explored, which consist of multiplicative scatter correction (MSC), standard normal variate (SNV), Gaussian filter (GF), Savitzky Golay smoothing (SG), gap segment (GS) and normalisation (NM). Among the six types of techniques, each were further explored among the same type, varying in terms of degree of polynomial fit and derivative order to determine the best technique.

The six techniques were further classified into four main methods of smoothing, differentiation, standardisation and normalisation. The MSC technique falls under normalisation method that is popularly used due to its ability to handle both additive and multiplicative effects of light scattering in reflectance spectroscopy. However, a number of similar effects are successfully treated with MSC, such as offset shifts and interference. It uses the mean spectrum as reference for the data set with the assumption that the calculated mean is representing all current and future spectra collected using similar field conditions without affecting the chemical signal in the data (Bi et al., 2016). The SNV technique is another type of normalisation method that removes multiplicative interference of scatter and particle size, baseline shift and curvilinearity of spectral (Rammal et al., 2014; Ding & Xu, 1999). The difference between the SNV and MSC techniques is that SNV is applied to an individual spectrum, while MSC uses a reference spectrum, such as the mean spectrum of the calibration set for correction of the spectral (Moghimi, Aghkhani, Sazgarnia, & Sarmad, 2010).

The SG and GS technique of smoothing method uses information from a localised segment of the spectrum to compute the derivative at a particular wavelength. Usually, this avoids the problem of noise enhancement from the simple difference method and applies smoothing to the data. The GS technique also requires appropriate gap size and smoothing segment size for best results. The SG and GS also fall into the category of derivation, provided the order is more than one. The SG technique was explored as it uses convolution function, and thus the number of segments in the function must be specified to one and two polynomial order to have best representation of local spectrum behaviour and avoid over smoothing (Liaghat et al., 2014). The SG technique that consists of a polynomial order of two followed by a derivative computation for both first and second order is also used to perform baseline correction especially to detect subtleness in vegetation (He, Li, & Ma, 2015; Liaghat et al., 2014). The GS derivatives are commonly applied as correction for baseline effects in spectra for the purpose of removing non-chemical effects and resolving overlapped bands, as well as aiding in emphasising small spectral variations, which are not evident in the raw data, thus creating robust calibration models. The first derivative was used since it is a measure of the slope of the spectral curve at every point, effective for removing purely additive baseline offsets (Jamshidi et al., 2012).

Nevertheless, peaks in raw spectral usually become zero-crossing points in first derivative spectra, which cause difficulty in interpreting them. This is when the second derivative is usually desired to resolve nearby peaks and sharpen spectral features while maintaining the intensities of the original curves. The second derivative is explored as a measure of the change

in the slope of the curve to remove pure additive offset, unaffected by any linear "tilt" that may exist in the data and is therefore a very effective method for removing both the baseline offset and slope from a spectrum. The first and second derivative is effective in increasing the spectral resolution (Cen & He, 2007). The third order derivative was explored with a hope to reveal phenomena, which do not appear clearly when using lower-order derivatives and can be helpful in understanding the spectral data. Gaussian Filter is a form of smothering for spectral signals, more commonly used for pulse de-noising (Wang, Zhang, & Lu, 2016).

NM techniques falls under standardisation method that are transformations computed based on samples. There are many types of normalisation, such as area normalisation, mean normalisation, unit vector normalisation, maximum normalisation, range normalisation and peak normalisation. Area normalisation and mean normalisation (standard method) are similar methods of computing area under the curve and only differ by a constant multiplicative factor, provided the samples are all in the same units (Noviyanto, Abdulla, Yu, & Salcic, 2015). The unit vector normalisation method follows pattern normalisation, usually used for pattern recognition (Kilz & Cazes, 2001). While maximum normalisation divides each row by its maximum absolute value instead of the average like the mean normalisation method, range normalisation divides each row by its range (max value – min value) (Mandic, 2012).

The peak normalisation attempts to correct spectra for indeterminate path length by effectively removing it and normalising the spectrum to the intensity of the peak. Normalisation technique is important method of pre-processing as it performs both scaling and offset correction at once. However, only area normalisation, mean normalisation, maximum normalisation and range normalisation methods were explored in terms of its relevance to the sample features (Liaghat et al., 2014). All the pre-processing techniques were implemented using the Unscrambler 10.4 software of Camo Process As, Oslo, Norway.

Calibration and Validation Models

Effects of several pre-processing techniques on the performance of PLS models were investigated for six different nutrients. The unprocessed raw spectral reflectance data was first analysed and set as benchmark to evaluate the actual gain of pre-processing technique. The analysis of PLS regression generalises and combines features from principal component analysis (PCA) and multiple regression (MR) in order to predict a set of dependent variables from a set of independent variables or predictors. This prediction is achieved by extracting from the predictors a set of orthogonal factors also known as latent variables (LV) which provides best predictive power (Abdi, 2007).

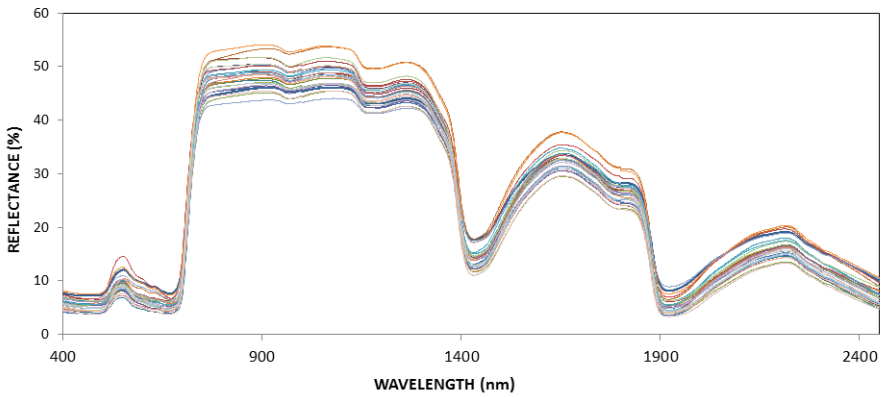
The PLS models were evaluated based on the representation of coefficient of correlation (R^2) and root mean square error (RMSE) values of both calibration and prediction models (Lazim, Nawi, Chen, Jensen, & Rasli, 2017). Before developing the calibration models, the sample data were randomly divided into calibration set (75% of the whole samples) and prediction set (25% of the whole samples) (Nawi et al., 2012). RMSE is termed the root mean square error in calibration (RMSEC) for the calibration set and the root mean square error in prediction (RMSEP) for the external prediction set. Multivariate variables such as the R^2, RMSEC, and RMSEP were computed to evaluate the quality of the PLS models. Models that

gave the lowest possible LV with the highest coefficient of determination (R^2), and lowest root mean square error of prediction (RMSEP) with respect to the lab results obtained from the destructive chemical analysis are selected for each nutrient.

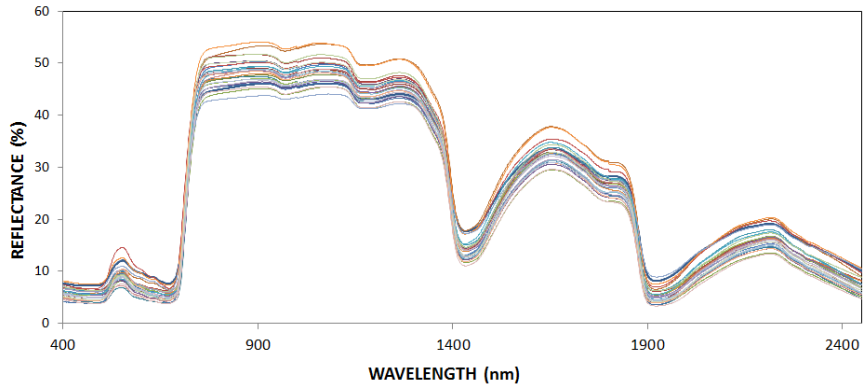
Mostly, the LV used for calibration model was 10. Full cross-validation with leave-four-out option was executed to determine the optimal number of LVs. Exploring the PLS analysis, two types of validation were compared, between the Nipals and Kernal. Generally, Nipals runs with an iterative algorithm that could impute missing values and does well with large set of samples and variables while Kernal performs non-iterative algorithms and is better with large sample group with small set of variables but does not handle missing values (Mevik, Wehrens, & Liland, 2015).

RESULTS AND DISCUSSION

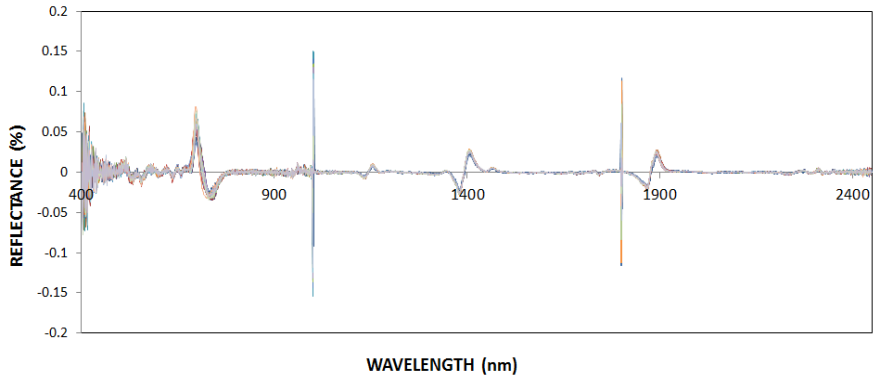
Before starting on pre-processing of spectra, the first and last 50 nm (Jamshidi et al., 2012) of wavelengths were removed because of the existence of considerable noise in these regions. Thus, only 2050 number of wavelengths are considered for pre-processing. A few numbers of pre-processing techniques were tested, alone and in combinations. The Figure 2(a) below shows the original unprocessed, raw leaf spectral reflectance data of 30 samples in percentage, from 400 nm to 2450 nm for F17 samples. Figure 2(b)-(d) below presents the result of some of the pre-processing techniques of the raw spectral. The SG (b) and raw (a) spectra looks alike to the naked eye, while the MSC(d) treated spectra with light scatter corrections has denser, slimmer and sharper spectral compared with the original raw spectral. The GS on the other hand looks very different due to the second order derivative effects. Based on research study by Noviyanto et al. (2015), there is no one best pre-processing technique for a particular field or task, but for every study, suitable spectral pre-treatment is to be determined respectively. Therefore, several pre-processing techniques were explored to find the most suitable method for oil palm leaf.



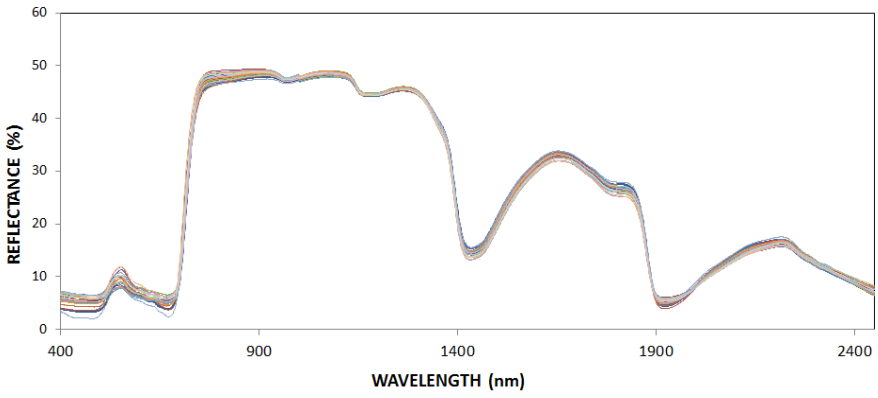
(a)



(b)



(c)



(d)

Figure 2. Reflectance spectral from 30 leaf samples of frond 17 and results of pre-processing techniques; a) Raw spectral, b) Savitzky-Golay (SG), c) Gap-segment 2d (GS) and d) Multiplicative scatter correction (MSC)

All the techniques that were explored had calibration of R^2 value in the range from 50 to 99%, with mostly falling into the mid and upper region. Hence, the prediction model's R^2 value was the determinant factor for the optimal pre-processing technique. The differentiation method used was inclusive of first and second derivative to increase spectral resolution and remove background (Cen & He, 2007). The Savitzky-Golay (SG) technique was explored between polynomial one and two, as it was important to choose the proper width of smoothing window (Cen & He, 2007). The results obtained were mixed, where N, K, Mg and Ca performs better with SG of polynomial one while P and B were better with SG of polynomial two as shown in Table 1.

As illustrated in Table 1, SG techniques with first and second derivative were compared. Between these two, the SG with first derivative clearly shows better representation of prediction R^2 and RMSEP. Lelong et al. (2010) had 94% of overall accuracy for Ganoderma infection level classification based on spectra that was treated with second order derivative of PO 3 of oil palm canopy. Gap Segment (GS) technique of pre-processing was tested for first, second and third order derivatives settings of each with gap size of two and segment size of three. As shown in Table 1, generally the GS with second derivative shows good results for most nutrients except Mg which however had good R^2 representation for Calibration model.

However, only second derivative was found to be the most suitable pre-treatment for protein content determination in wheat flour with R^2 value of 0.98 from calibration model of modified PLS analysis (Huali et al., 2011). The smoothing method that optimises signal to noise ratio and removes random noise includes SG and moving average techniques (Cen & He, 2007; Naes, Isaksson, Fearn, & Davies, 2002). The GF technique of SS5 was considered since its prediction R^2 representation was better for nutrients of K, Mg, Ca and B with considerably low RMSEP in Table 2.

Table 1
Partial Least Square (PLS) models prediction results for N, P, K, Mg, Ca and B for raw spectral, Savitzky-Golay (SG) and Gap segment (GS) of derivative pre-processing method

Pre-treatment	Nitrogen (N)		Phosphorus (P)		Potassium (K)		Magnesium (Mg)		Calcium (Ca)		Boron (B)	
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP
Raw	0.05	0.22	0.19	0.01	0.31	0.30	0.14	0.06	-	-	0.57	7.3
S.G 1st D	0.21	0.31	0.06	0.13	0.05	2.09	0.32	0.36	0.11	0.28	0.00	10.74
S.G 2nd D	0.18	0.17	0.18	0.02	-	1.31	-	0.23	0.01	0.31	-	-
G.S 1st D	0.26	0.30	0.05	0.11	0.02	2.01	0.51	0.36	0.17	0.32	-	8.09
G.S 2nd D	0.13	0.32	0.09	0.09	-	0.08	0.20	0.21	0.4	0.23	0.26	14.98
G.S 3rd D	0.05	0.18	0.07	0.01	0.31	0.43	0.40	0.09	0.30	0.16	0.09	9.73

* - indicates no results from the Unscrambler software

Between the Gaussian Filter (GF) of segment size (SS) SS3 and SS5, it was clearly shown in Table 2 that the GF with SS5 gives better representation of R^2 for calibration as well as prediction model. Although some nutrients such as N, P, K and Ca may have higher R^2 representation for calibration, its prediction value was not satisfactory. Standardisation of spectral data was a pre-processing method used by divide the spectrum at every wavelength with its standard deviation at certain wavelength, which includes basic normalisation (NM) techniques. For the pre-processing NM technique, comparison was made among the NM-area, NM-mean, NM-range and NM-Maximum, where the results are close to each other as shown in Table 3. However, NM-range technique performed better for most nutrients of P, K, Mg, Ca and B, although this may not be the best technique available for oil palm leaf reflectance.

Table 2
Partial Least Square (PLS) models prediction results for N, P, K, Mg, Ca and B for raw spectral, Savitzky-Golay (SG) and Gaussian filter (GF) of smoothing pre-processing method

Pre-treatment	Nitrogen (N)		Phosphorus (P)		Potassium (K)		Magnesium (Mg)		Calcium (Ca)		Boron (B)	
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP
Raw	0.05	0.22	0.19	0.01	0.31	0.30	0.14	0.06	-	-	0.57	7.3
S.G(PO1)	0.40	1.83	0.09	0.15	0.63	0.08	0.59	0.08	0.46	2.00	0.46	1.98
S.G(PO2)	0.05	0.22	0.19	0.01	0.31	0.30	0.13	0.06	-	-	0.57	7.31
G.F (SS5)	0.66	1.26	0.66	0.18	0.87	0.08	0.92	0.10	0.95	0.57	0.95	1.01
G.F (SS3)	0.05	0.22	0.19	0.01	0.31	0.30	0.07	0.06	-	-	0.57	7.34

* - indicates no results from the Unscrambler software

Table 3
Partial Least Square (PLS) models prediction results for N, P, K, Mg, Ca and B for raw spectral, mean, area, range and maximum (nm) of standardisation pre-processing method

Pre-treatment	Nitrogen (N)		Phosphorus (P)		Potassium (K)		Magnesium (Mg)		Calcium (Ca)		Boron (B)	
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP
Raw	0.05	0.22	0.19	0.01	0.31	0.30	0.14	0.06	-	-	0.57	7.3
NM-Mean	0.07	0.23	0.03	0.01	0.46	0.28	0.01	0.09	-	-	0.64	6.90
NM-Area	0.07	0.23	0.05	0.01	0.46	0.28	0.01	0.09	-	0.18	0.64	6.90
NM-Range	0.01	0.26	0.20	0.01	0.10	0.37	0.01	0.09	-	0.16	0.80	5.14
NM-Maxi	0.19	0.25	0.29	0.01	0.50	0.28	0.08	0.09	0.30	0.18	0.64	6.79

* - indicates no results from the Unscrambler software

The normalisation method applies multiplicative effects and compensates for baseline shifts in the spectral data includes, MSC and SNV were as shown in Table 4 (Lazim et al., 2016). The MSC technique was evaluated in terms of PLS-Nipals and PLS Kernal. For most of the nutrients tested, the MSC technique applied with Kernal presents better results in terms of prediction R^2 and RMSEP, as shown in Table 4. This was seen in P and Mg in its R^2 representation for prediction and RMSEP for B. The SNV pre-processing technique performed similar to MSC and SNV, MSC-SG and MSC-SNV-SG combinations with MSC-SNV-SG being slightly better combination among this lot as seen in Table 4.

Table 4
Partial Least Square (PLS) models prediction results for N, P, K, Mg, Ca and B for raw spectral, multiplicative scatter correction (MSC) and standard normal variate (SNV) of normalisation method and mix of pre-processing techniques

Pre-treatment	Nitrogen (N)		Phosphorus (P)		Potassium (K)		Magnesium (Mg)		Calcium (Ca)		Boron (B)	
	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP	R^2	RMSEP
Raw	0.05	0.22	0.19	0.01	0.31	0.30	0.14	0.06	-	-	0.57	7.3
MSC Kernal	0.91	0.34	0.72	0.00	0.90	0.03	0.66	0.05	0.26	0.11	0.74	0.1
MSC Nipals	0.11	0.22	0.03	0.01	0.42	0.27	0.04	0.07	0.37	0.19	0.73	9.19
SNV	0.13	0.23	0.03	0.01	0.44	0.27	0.04	0.07	0.00	0.18	0.73	8.38
MSC-SNV	0.13	0.23	0.03	0.01	0.44	0.27	0.04	0.07	0.00	0.16	0.73	8.38
MSC-S.G	0.11	0.22	0.03	0.01	0.42	0.27	0.04	0.07	0.08	0.16	0.73	9.33
MSC-SNV-S.G	0.13	0.23	0.03	0.01	0.44	0.27	0.04	0.07	-	0.16	0.73	8.32

* - indicates no results from the Unscrambler software

The SG with MSC was found to have best relationship between spectral absorbance and soluble solids content (SSC Brix) with higher R^2 of 0.98 when compared to SG and MSC as well as Moving average method with MSC or SNV (Jamshidi et al., 2012). However, in an overall comparison among the best setting of each type of pre-processing technique, the MSC technique seems to fulfil the aim of high R^2 representation as well as low RMSEP. This was better than a study on grass for N using localised SNV that was found to have higher RMSEP of 0.169 (Bi et al., 2016). The PLS using spectroradiometer as sensor was proven to have good relationship for prediction of nutrients N, P, and K by Jayaselan et al. (2017).

CONCLUSION

In order to obtain highly accurate analysis and modelling results, pre-processing of spectral data is required. After evaluation of four methods and six techniques of spectral pre-processes, the MSC technique was found to be the best choice for this data form, as an overall nutrient predictor. The MSC results for N, P, K, Mg, Ca, and B in terms of R^2 representation were 0.91, 0.72, 0.90, 0.66, and 0.26, and 0.74 while in terms of RMSEP values were 0.34, 0.00, 0.03, 0.05, 0.11, and 0.1. Nevertheless, in the context of representing only one nutrient, it was found that the most suitable pre-processing technique for Mg, Ca and B was Gaussian Filter with SS5 parameter and MSC for N, K, and P nutrients.

REFERENCES

Adbi, H. (2007). Partial least square regression. PLS-regression. *Encyclopedia of Measurement and Statistics, 6(*4), 792-795.

Basiron, Y. (2016). Booster for global economies; New research shows palm oil's impact on revenues worldwide. *Palm Voices Columnist.* The Star (4 October 2016).

Bechlin, M. A., Fortunato, F. M., Silva, R. M. D., Ferreira, E. C., & Neto, J. A. G. (2014). A simple and fast method for assessment of the nitrogen-phosphorus-potassium rating of the fertilisers using high-resolution continuum source atomic and molecular absorption spectrometry. *Spectrochimica Acta Part B: Atomic Spectroscopy, 101*, 240-244.

Bi, Y., Yuan, K., Xiao, W., Wu, J., Shi, C., Xia, J., … & Zhou. (2016). A local pre-processing method for near-infrared spectra combined with spectral segmentation and standard normal variate transformation. *Journal of Analytica Chinica Acta, 909*, 30-40.

Carter, C., Finley, W., Fry, J., Jackson, D., & Willis, L. (2007). Palm oil markets and future supply. *European Journal of Lipid Science and Technology, 109*(4), 307-314.

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science and Technology, 18*, 72-83.

Corley, R. H. V. (2009). How much palm oil do we need? *Journal of Environmental Sciences and Policy, 12*(2), 134-139.

Corley, R. H. V., & Mok, C. K. (2008). Effects of nitrogen, phosphorus, potassium and magnesium on growth of the oil palm. *Journal of Experimental Agriculture, 8*(4), 347-353.

Ding, H. B., & Xu, R. J. (1999). Differentiation of beef and kangaroo meat by visible/near infrared reflectance spectroscopy. *Journal of Food Science, 64*(5), 814-817.

Dubos, B., Snoeck, D., & Flori, A. (2017). Excessive use of fertiliser can increase leaching process and modify soil reserves in two Ecuadorian oil palm plantations. *Journal of Experimental Agriculture, 53*(2), 255-268.

Engel, J., Gerretzen, J., Szymanska, E., Jansen, J. J., Downey, G., Blanchet, L., & Buydens, L. M. C. (2013). Breaking with trends in pre-processing? *Journal of Trends in Analytical Chemistry, 50*, 96-106.

Graboski, M. S., & McCormick, R. L. (1998). Combustion of fat and vegetable oil derived fuels in diesel engines. *Journal of Progress in Energy and Combustion Science, 24*(2), 125-164.

Hamzah, A., Salleh, S. N. M., Wong, K. K., & Sarmani, S. (2016). Nutrient amendments of inorganic fertiliser and oil palm empty fruit bunch and their influence on bacterial species dominance and degradation of the associated crude oil constituents. *Journal of Soil and Sediment Contamination, 25*, 256-265.

Hashim, M., Ibrahim, A. L., Rasib, A. W., Shah, R., Nordin, L., & Haron, K. (2001). Detecting oil palm tree growth variability using a field spectroradiometer. *ASIAN-PACIFIC Remote Sensing and GIS Journal, 14*, 25-32.

He, Z., Li, M., & Ma, Z. (2015). Prediction of dry matter, protein and acidity in corn steep liquor using near infrared spectroscopy. In *7th International Conference on Awareness Science and Technology (iCAST)* (pp. 212 – 215). IEEE Conference Publications.

Henson, I. E., & Chai, S. H. (1997). Palm productivity: Biomass distribution, productivity and turnover of the root system. *Elaeis, 9*(2), 78-92.

Huali, J., Jinshui, W., Lihui, Y., Rui, G., & Li, L. (2011). Establishment of non-destructive testing model of the protein content in wheat flour by NIR spectroscopy. In *New Technology of Agricultural Engineering (ICAE), 2011 International Conferenc*e (pp. 1087-1091). IEEE.

Jamshidi, B., Minaei, S., Mohajerani, E., & Ghassemian, H. (2012). Reflectance Vis/NIR spectroscopy for non-destructive taste characterization of Valencia oranges. *Journal of Computers and Electronics in Agriculture, 85*, 64-69.

Jayaselan, H. A. J., Nawi, N. M., Ismail, W. I. W., Shariff, A. R. M., Rajah, V. J., & Arulandoo, X. (2017). Application of spectroscopic method to predict nutrient content of oil palm from leaf. *Journal of Experimental Agriculture, 15*(3), 1-9.

Jensen, T., Apan, A., Young, F., & Zeller, L. (2007). Detecting the attributes of a wheat crop using digital imagery acquired from a low-altitude platform. *Journal of Computers and Electronics in Agriculture, 59*(1), 66-77.

Jones, G. H. (2004). Application of thermal imaging and infrared sensing in plant physiology and ecophysiology. *Journal of Advances in Botanical Research, 41*, 107-163.

Keng, P. S., Baasri, M., Zakaria, M. R. S., Rahman, M. B. A., Ariff, A. B., Rahman, R. N. Z. A., & Salleh, A. B. (2009). Newly synthesized palm esters for cosmetics industry. *Journal of Industrial Crops and Products, 29*(1), 37-44.

Khorramnia, K., Shariff, A. R. M., Rahim, A. A., & Mansor, S. (2014). Toward malaysian sustainable agriculture in 21st century. In *IOP Conference Series: Earth and Environmental Science* (Vol. 18, No. 1, p. 012142). IOP Publishing.

Kilz, P., & Cazes, J. (2001). *Encyclopaedia of chromatography* (pp. 195-199). Marcel Dekkar Inc.

Kruse, F. A. (1994). Imaging spectrometer analysis: A Tutorial. In *Proceeding of International Symposium on Spectral Sensing Research (ISSSR)* (pp: 10-15). San Diego, CA.

Lazim, S. S. R. M., Nawi, N. M., Chen, G., Jensen, T., & Rasli, A. M. M. (2016). Influence of different pre-processing methods in predicting sugarcane quality from near infrared (NIR) spectral data. *International Food Research Journal, 23*, 231-236.

Lelong, C. C. D., Roger, J. M., Bregand, S., Dubertret, F., Lanore, M., Sitorus, … & Caliman, J. P. (2010). Evaluation of oil palm fungal disease infestation with canopy hyperspectral reflectance data. *Journal of Sensors, 10*, 734-747.

Liaghat, S., Mansor, S., Ehsani, R., Shafri, H. Z. M., Meon. S., & Sankaran, S. (2014). Mid-infrared spectroscopies for early detection of basal stem rot disease in oil palm. *Journal of Computers and Electronics in Agriculture, 101*, 48-54.

Mandic, Z. (2012). *Physio chemical methods in drug discovery and development.* Zagreb, Croatia: IAPC Publishing.

Mevik, B. H., Wehrens, R., & Liland, K. H. (2015). *Package 'PLS'.* Retrieved from http://mevik.net/work/software/pls.html

Moghimi, A., Aghkhani, M. H., Sazgarnia, A., & Sarmad, M. (2010). Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Journal of Biosystem Engineering, 106*(3), 295-302.

Nawi, N. M., Jensen, T., & Chen, G. (2012). The application of spectroscopic methods to predict sugarcane quality based on stalk cross-sectional scanning. *Journal of American Society of Sugar Cane Technologists, 32*, 16-27.

Nawi, N. M., Jensen, T., & Chen, G. (2013). Application of spectroscopic method to predict sugar content of sugarcane internodes. *Journal of Tropical Agriculture and Food Science, 41*(2), 211-220.

Nicolai, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*, 99-118.

Noviyanto, A., Abdulla, W., Yu, W., & Salcic, Z. (2015). Research trends on optical spectrum for honey analysis. In *Proceedings of Asia-Pacific Signal and Information Processing Association (APSIPA) Annual Summit and Conference* (pp. 416-425). IEEE Conference Publications.

Næs, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A user-friendly guide to multivariate calibration and classification.* Charlton, Chichester, UK: NIR Publications.

Phin, C. K., Stephen, R., & Markus, A. (2009). In vitro antimicrobial activity and fungi toxicity of syringic acid, caffeic acid and 4-hydroxybenzoic acid against Ganoderma Boninense. *Journal of Agriculture Science, 1*(2), 15-20.

Rajaratnam, J. A., Chan, K. W., & Goh, K. H. (1977). The foundation for selecting leaf 17 for nutrient requirements of mature oil palms. In *Proceeding Conference on Classification and Management of Tropical Soils (Joseph, K.T., Ed.)* (pp: 340-348). Malaysian Society of Soil Science, Kuala Lumpur, Malaysia.

Rammal, A., Perrin, E., Vrabie, V., Bertrand, I., Habrant, A., & Chabbert, B. (2014). Optimal pre-processing and FCM clustering of MIR, NIR and combined MIR-NIR spectra for classification of maize roots. In *The Third International Conference on e-Technologies and Networks for Development* (pp: 110–115). IEEE Conference Publications.

Sheil, D., Casson, A., Meijaard, E., Nordwijk, V. M., Gaskell, J., Groves, S. J., … & Kanninen, M. (2009). *The impacts and opportunities of oil palm in Southeast Asia: What do we know and what do we need to know?* (No. 51). CIFOR, Bogor, Indonesia.

Simeh, A., & Ahmad, T. M. A. T. (2001). The case study on Malaysian palm oil. In *Regional Workshop on Commodity Export Diversification and Poverty Reduction in South and South-East Asia* (pp 3-5). Bangkok.

Sugiura, R., Noguchi, N., & Ishii, K. (2005). Remote sensing technology for vegetation monitoring using an unmanned helicopter. *Biosystem Engineering, 90*(4), 369-379.

Wahid, O., Xaviar, A., Tarmizi, A. M., & Ibrahim, S. (2002). Precision agriculture: Fertiliser management MAP: Spatial and correlation analysis of yield and leaf nutrient. *MPOB Information Series*.

Wang, D., Zhang, D., & Lu, G. (2016). A robust signal pre-processing framework for wrist pulse analysis. *Journal of Biomedical Signal Processing and Control, 23*, 62-75.