International Journal of Integrated Engineering, Vol. 10 No. 8 (2018) p. 112-119

DOI: https://doi.org/10.30880/ijie.2018.10.08.017

Artificial Neural Network and Savitzky Golay Derivative in Predicting Blood Hemoglobin Using Near-Infrared Spectrum

Mohd Nazrul Effendy Mohd Idrus[1,2], Kim Seng Chia[1,*], Hiew Moi Sim[3], Hasan Ali Gamal Al-kaf [2]

[1]Department of Engineering and Skills, Batu Pahat Community College, Evergreen Height, Batu Pahat, Malaysia.
[2]Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Malaysia.
[3]Faculty of Engineering and Information Technology, Southern University College, Skudai, Johor, Malaysia.

Received 1 March 2018; accepted 19 August 2018, available online 31 December 2018

Abstract: Monitoring blood hemoglobin level is essential to diagnose anaemia disease. This study aims to evaluate the capability of an artificial neural network (ANN) and Savitzky Golay (SG) pre-processing in predicting the blood hemoglobin level based on the near-infrared spectrum. The effects of the hidden neuron number and different SG pre-processing strategies were examined and discussed. ANN coupled with first order SG derivative and five hidden neurons achieved better prediction performance with root mean square error of prediction of 0.3517 g/dL and coefficient determination of prediction of 0.9849 compared to the previous studies. Results depict that ANN that coupled with first order SG derivative could improve near-infrared spectroscopic analysis in predicting blood hemoglobin level, and the proposed nonlinear model outperforms linear models without variable selections. This finding suggests that the modelling strategy is promising in establishing a better relationship between the blood hemoglobin and near-infrared spectral data.

Keywords: Blood hemoglobin prediction, Savitzky Golay pre-processing, Artificial neural network, Near-infrared

1. Introduction

Anemia disease can be diagnosed by blood hemoglobin concentration [1]. Generally, human can be defined suffering severe anemia when their hemoglobin level lower than 7.0 g/dL, 9.9 to 7.0 g/dL was moderate anemia, 10 to 11.9 g/dL was mild anemia and 12.0 g/dL or higher was defined as non-anemia [1]. Cyanmethemoglobin method is a commonly clinical practice used to measure blood hemoglobin by drawn some blood from the patient to be mixed with reagent chemicals for analysis [2]. Alternative methods, e.g. near-infrared spectroscopic technology that is promising in measuring blood hemoglobin level is worth to be investigated.

Near-infrared spectroscopy (NIRS) approach has been found to be promising as a fast response and noninvasive technique for estimating blood hemoglobin level. NIRS technique has widely used in various fields such as medical [3, 4], food [5, 6], and fuel [7]. This is because of the overtones and combinations of fundamental stretching vibrations especially the asymmetric vibrations of C–H, O–H, and N–H produces absorption bands that make near-infrared (NIR) useful for analyzing in the biological system [8]. NIRS region of the electromagnetic spectrum is from about 780 nm to 2500 nm. The advanced NIRS sensing technology produces a large volume of spectral data, which makes it an advantage in the analytical process. Multivariate calibration modelling has been used to establish a relationship between spectral data and component of interest. Preprocessing, calibration and validation are common processes rely on developing multivariate calibration modelling. However, several factors, e.g. baseline drift, weak signal to noise ratio, high dimensional data, and nonlinearity could degrade the accuracy of the NIRS approach [9]. Thus, a better understanding of preprocessing approach and predictive model is needed so that the potential of NIRS can further reveal.

The international diffuse reflectance conference (IDRC) 2010 chemometric challenge aimed to challenge the ability of participates to develop a better calibration model for blood hemoglobin level based on given spectral data [10]. Although four selected models achieved good results, only linear models, e.g. multiple linear regression (MLR) and partial least square (PLS) that coupled with wavelength selection or reduction approaches were investigated [11]. It is worth to highlight that wavelength selection or reduction approaches could accidentally remove important data, and consequently minimize the information of spectral data for modelling. On the other hand, nonlinear modelling that does not involve wavelength selection or reduction preprocessing may be able to improve performance prediction of hemoglobin. This is because one of the major concerns in developing multivariate calibration method is to deal with nonlinearity issue in spectral data [12]. Thus, nonlinear

*Corresponding author: kschia@uthm.edu.my

model e.g. artificial neural network (ANN) should be investigated to give a fair comparison among different modelling strategies. Latest research shown that ANN has high capability as a predictive model in near-infrared spectroscopy applications [13]. Furthermore, ANN which was originally inspired by neurons in the brain can be viewed as a universal model-free approximator that represent any nonlinear function with sufficient accuracy connected through complex networks. In addition, ANN as a nonlinear model may achieve a better predictive performance compared to linear models, e.g. PLS, MLR, and principal component analysis (PCA) in NIRS analysis [14].

Savitzky Golay (SG) preprocessing method with smoothing, first order derivatives and second order derivative were able to overcome most common issues and unwanted signal in raw spectral data [6]. Unappropriated SG preprocessing parameters selection could degrade the performance of modelling [15]. First and second order SG derivative approached can remove both baseline and slope effects and attenuate the effects of high frequency noise. Furthermore, SG preprocessing able to keep signal shape effects and guarantee better denoising. Moreover, SG preprocessing should be optimized by selecting an appropriate frame length to preserve the resolution of the derivative signal [16]. However, the optimization approach in SG preprocessing is seldom be investigated or discussed properly in previous works. Without this optimization approach, this may degrade the performance in NIR spectroscopic analysis. Therefore, the objective of this paper is to investigate the effect of ANN that coupled with different SG preprocessing strategies e.g. smoothing, first order and second order derivative in predicting blood Hemoglobin using near-infrared spectrum.

2. Material and method

There are three standard steps in developing a predictive model in NIRS, e.g. pre-processing, calibration and validation. In this study, artificial neural network (ANN) that coupled with SG pre-processing was developed and investigated as shown in Fig. 1. First, NIR spectral data were processed with smoothing, first, and second order SG derivative to remove unwanted signal. After that, treated spectral data were modelled using ANN to predict blood hemoglobin level. Results during training and testing were validated to measure the performance of the model.

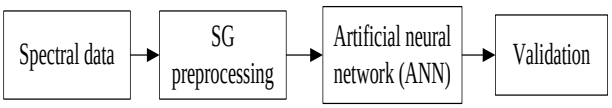


Fig. 1 Predictive model architecture.

2.1 Spectral data

The origin spectral dataset was provided by Karl Norris from IDRC shootout 2010 [10]. Blood samples were analyzed with a NIRSystems 6500 spectrometer. All spectra have 700 variables, from 1100 to 2498 nm, with a 2 nm interval as shown in Fig. 2. The blood hemoglobin reference was measured by a high-volume hematology analyzer, Coulter STKS monitor was made by the Coulter Corporation of Hialeah, FL. The dataset contains 231 sets of calibration and 194 new unseen data sets for the blind test to measure predictive accuracy of the modelling networks.

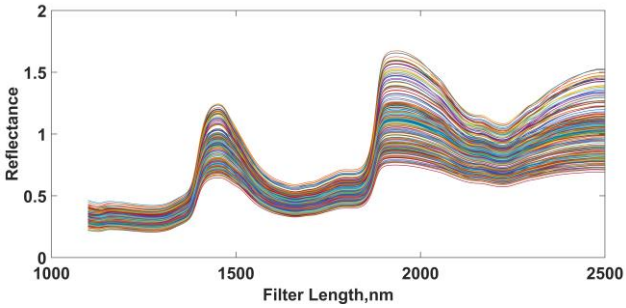


Fig. 2 Blood hemoglobin reflectance raw spectral data.

Table 1 shows the descriptive statistics of the samples and reference blood hemoglobin with a number of samples (n), minimum (Min), maximum (Max), mean and standard deviation (Std). Minimum and maximum data in calibration were 10.30 g/dL and 17.30 g/dL respectively. While minimum and maximum data in testing were 6.50 g/dL and 18.20 g/dL respectively. From these range in calibration and testing, indicated that extrapolation samples were used for modelling. This type of extrapolation samples was used in IDRC shootout 2010 [10].

Table 1 Descriptive statistics of the blood hemoglobin

Data	n	Min (g/dL)	Max (g/dL)	Mean (g/dL)	Std
Calibration	231	10.30	17.30	13.78	1.66
Testing	194	6.50	18.20	12.20	2.83
Total	425	6.50	18.20	13.06	2.40

2.2 Savitzky Golay preprocessing

The ANN is a promising modelling technique for dealing with any nonlinearities and complexities data [17]. However, the prediction performance of ANN can be degraded when various aspects e.g. preprocessing data problems, structure and learning algorithm are not considered appropriate [17]. The SG preprocessing was used to perform smoothing (SGD0), first order (SGD1) and second order (SGD2) SG derivatives to remove unwanted signals from the spectral data before calibration process. A total of 231 samples were processed with three different SG preprocessing methods e.g. SG smoothing, first order, and second order SG derivative preprocessed data. Optimal frame lengths for SG preprocessing would be measured to improve the performance of the model. The coefficients of smoothing (C_0), first order (C_1) and second order (C_2) derivatives were generated by using built in matrix routine function from MATLAB software (MATLAB® Version8.4

(R2014b)). Middle value from desired order derivative can be estimated by the dot product of each value of C_0, C_1 and C_2 represented coefficient differentiation filter with spectral data using Equation 1.

$$Y_j = (C_i \otimes y)_j = \sum_{n-\frac{m-1}{2}}^{\frac{m-1}{2}} C_i y_{i+j} \quad (1)$$

Where C_i is set of SG coefficient, y_{i+j} is a related set of data before treatment and Y_j is observed value after treatment. While *m* and *n* are the measured number and a total number of frame length, respectively.

2.3 Artificial neural network

The ANN is inspired from human brain that consists of many neurons that connected each other to form complex networks for learning and thinking. One of the famous artificial neural networks is multi-layer perceptron (MLP) that trained by Levenberg-Marquardt.

2.3.1 Multilayer perceptron

The MLP that was applied in this study has three layers e.g. an input layer connected with a hidden layer and linked to the output layer [18]. Result outputs of each neuron were established by multiplying the corresponding weights and then the weights that are first summed (combined) before passed through a transfer function to get the desired output. With 700 input of spectra variables, one to ten hidden neurons were varied to predict the value of an output layer. To minimize time consuming without degrading predictive accuracy, one hidden layer was used, which is capable of predicting problems [19-20]. Default transfer functions of tan-sigmoid for hidden layer and linear transfer function for output layer were used to receive input from the sum of the weighted inputs and the bias [21]. Nguyen-Widrow initialization was used to generate initial weight and bias values for each layer and distributed evenly over the input and has advantages with faster training process works [22]. The network has been trained 100 times with different initial weights to observe effects of different initial weights to the global performance of the model. An appropriate and optimal initial weights can reduce initial error and number of iteration required to attain specific error. After that, the best model for training will be selected based on the optimal accuracy performance. Finally, selected model was tested with new unseen data to assess the prediction accuracy of the ANN.

2.3.2 Levenberg-Marquardt

The Levenberg-Marquardt (LM) is considered as the solution to most of the problems that are nonlinear least square [23-24]. It is considered as an intermediate between the Gradient Descent and Gauss Newton algorithms. LM has been used as repetitive techniques to minimize the error of nonlinear functions. LM algorithm is extensively preferred as supervised learning technique because it could achieve minimum training time consuming in most cases. Several researches have shown that LM algorithm could achieve a slightly better accuracy compared to Scaled Conjugate Gradient (SCG) [23]. However, LM algorithm does not guarantee global optimal performance and it should be properly optimized using a proper strategy e.g. holdout validation or cross-validation [24-25]. Default training parameters for LM algorithm provided by MATLAB function has been used in this research as shown in Table 2. The training process will stop when either the maximum number of epochs is reached, the goal performance is achieved, the performance of gradient is below minimum gradient value, the momentum update is exceeded, or the failure validation is more than the maximum amount.

2.4 Performance evaluation

Root mean squared error of calibration (RMSEC) and root mean squared error of prediction (RMSEP) were used to evaluate the performance of ANN. RMSEC and RMSEP are computed using Equation 2.

$$\text{RMSEC or RMSEP} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}-y)^2}{n}} \quad (2)$$

Where *n* is the total number of samples, while $\hat{y}$ and *y* denote the predicted blood hemoglobin and reference blood hemoglobin, respectively. The difference between RMSEC and RMSEP is the data that used to compute RMSEC was from training data, while the data that used to compute RMSEP were obtained from new independent unseen data. Therefore, the best model should have the lowest value of RMSEP. Next, the coefficient of determination (R^2) was used to interpret proportion of the variance in the predicted data from reference value output of regression analysis. In other words, R^2 can be used to describe the relationship between blood hemoglobin and near-infrared spectrum as stated in Equation 3.

$$R_c^2 \text{ or } R_p^2 = 1 - \frac{\sum(y-\hat{y})^2}{\sum(y-\bar{y})^2} \quad (3)$$

Where $\bar{y}$ is mean of reference data, $\hat{y}$ and *y* denote the predicted and reference blood Hb respectively from the new unseen dataset. R_c^2 is the coefficient of determination of calibration was taken from training process, while R_p^2 is the coefficient of determination of prediction taken from the testing process. The predictive modelling can be considered as a perfect performance when the value of R^2 approaching or equal to value 1 without any measurement error.

Table 2: Levenberg-Marquardt (LM) training parameters

Maximum number of epochs to train	1000
Performance goal	0
Maximum validation failures	6
Minimum performance gradient	1×10^{-7}
Initial momentum update	0.001
mu decrease factor	0.1
mu increase factor	10
Maximum momentum update	1×10^{10}

3. Results and Discussion

3.1 Savitzky Golay preprocessing

Spectral data has been preprocessed with different types of SG preprocessing e.g. smoothing SG, first order, and second order SG derivative with optimal number of frame length. In this research, the number of frame length for SG preprocessing were adopted from previous, i.e. smoothing SG was 77, first order SG derivatives was 27, and second order SG derivatives was 79 [26].

3.2 Hidden neurons

Fig. 3 illustrates the root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP) for training performance and predictive performance, respectively. The results indicate that most of the training performance is outperforms than predictive performance with RMSEC below than 0.3 gd/L. However, training performance cannot be benchmarking to measure the predictive performance of the model. Fig. 3 (a) shows ANN without SG preprocessing be able to achieve an optimal performance of prediction when 3 number of hidden neurons were used with RMSEP of 0.4607 gd/L. However, overfitting network occurs when more than 4 number of hidden neurons were used with increasing value of RMSEP.

Meanwhile, Fig. 3 (b) shows ANN coupled with smoothing SG tend to achieve a better performance of prediction with RMSEP of 0.4499 gd/L when 2 number of hidden neurons were used. This result indicates that preprocessed spectral data using smoothing SG tends to improve the performance of prediction and reduce the complexity of the network. Concurrently, Fig.4 (c) shows ANN coupled with first order SG derivative tends to produce an optimal performance prediction when 5 and 7 number of hidden neurons were used with RMSEP of 0.3517 gd/L and 0.4090 gd/L respectively. This result indicates that 5 hidden neurons tend to give optimal performance in the testing process. Moreover, the result was expected to be better compared to ANN without SG preprocessing and ANN coupled with smoothing SG because the baseline shift effect has been removed using first order SG derivative. Fig. 3 (d) shows an optimal prediction performance for ANN coupled with second order SG derivative was obtained with RMSEP of 0.4571 gd/L when three number of hidden neurons were used. The network going to be overfitting when more than four number of hidden neurons were used. Consequently, the results indicated that the most optimal prediction of performance achieved when less than three number of hidden neurons were used; except ANN coupled with first order SG derivative that used five number of hidden neurons. Hence, different types of SG preprocessing with different number of hidden neurons should be measured to optimize prediction of performance. Therefore, five number of hidden neurons ANN coupled with first order SG derivative are acknowledged as the best performance of prediction with its ability to perform well on unseen data.

The coefficient of determination (R^2) value represents the summary of correlation between outputs and targets as shown in Fig. 4. Based on the regression obtained, ANN with first order SG derivative shows an optimal correlation of $R_p^2 = 0.9849$ in prediction followed by ANN with smoothing SG of $R_p^2 = 0.9761$, ANN with second order SG derivative of $R_p^2 = 0.9746$, and ANN without SG derivative of $R_p^2 = 0.9734$. The result indicates that the ANN with first order SG derivative has high correlation when R_p^2 value approaching value 1 and can be considered as a good model for predictions.

3.3 Prediction performance

The performance accuracy of the artificial neural network with different SG preprocessing methods with an optimal number of hidden neurons was summarized in Table 3. Results indicate that ANN with SG preprocessing e.g. SG smoothing, first order SG derivative, and second order SG derivative achieved a better performance compared to ANN without preprocessing. This indicates that the unwanted signals in the raw spectral data were successfully reduced or removed using SG preprocessing. With five number of hidden neurons, ANN with SG1 tends to produce the optimal performance of predictions. This indicates that the complex neural networks could memorize the training data and generalize new unseen data for prediction. Results show that ANN coupled with SG1 and five number of hidden neurons achieved the best predictive accuracy with RMSEP of 0.3517 gd/L and R_p^2 of 0.9849. This result also indicates that ANN with first order SG derivative could produce better predictive accuracy compared to the previous works that used multiple linear regression (MLR), and Partial Least Square (PLS) [11][26]. This could be because the spectral data consists of nonlinear information. Besides, the combination of ANN and SG preprocessing using a proper optimization ultimately revealed the optimal predictive performance of the nonlinear model [15]. Table 3 depicts that the first order SG derivative that shows the optimal performance when nonlinear modelling approach, while second order SG derivative shows the optimal performance when linear modelling approached was applied [15]. This finding indicates that spectral data after preprocessed with first order SG derivative contains relevant information for nonlinear modelling while second SG derivative contains relevant information for linear modelling.

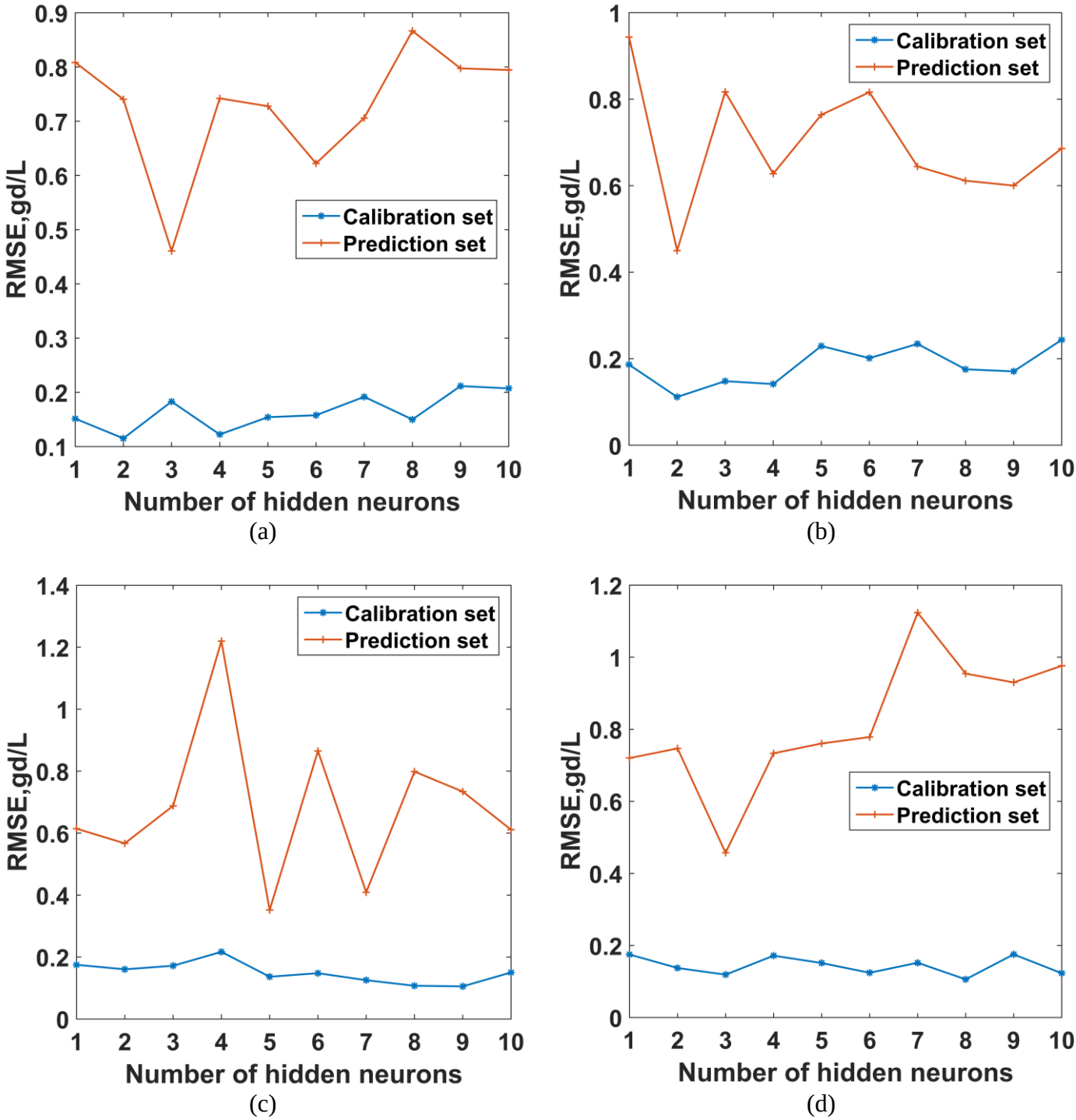
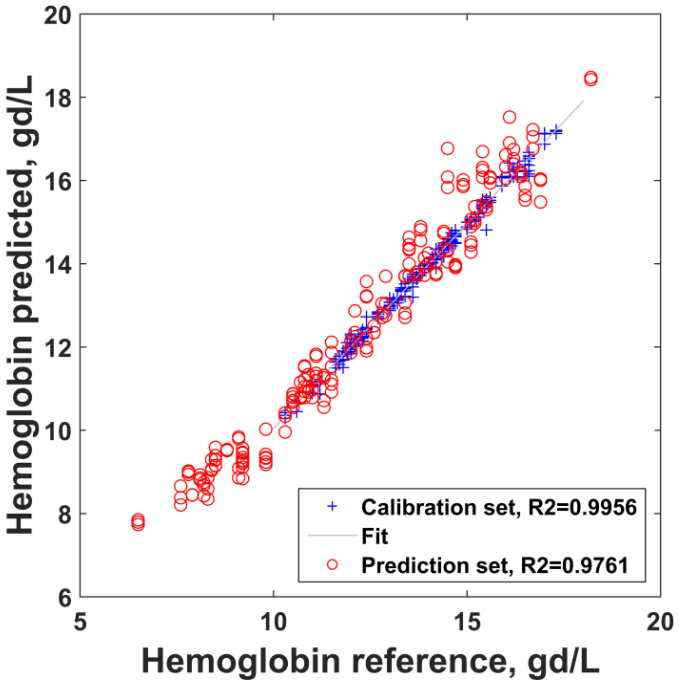
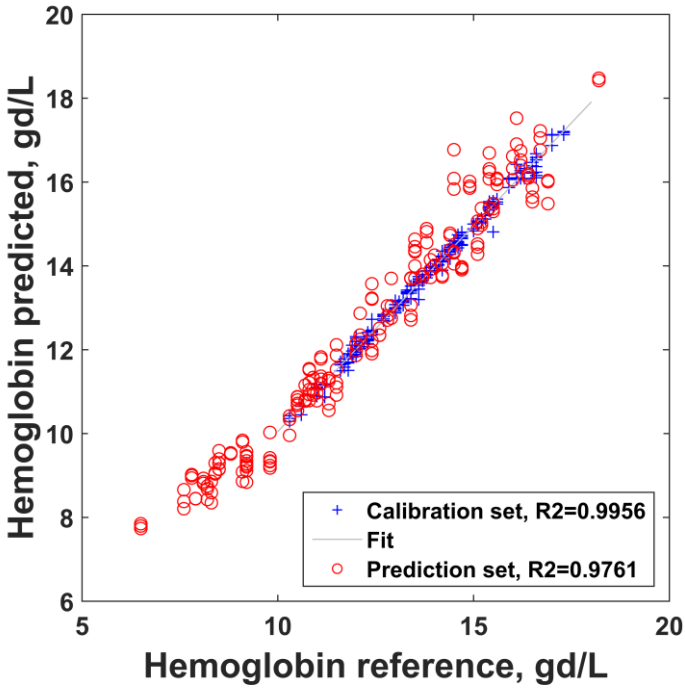


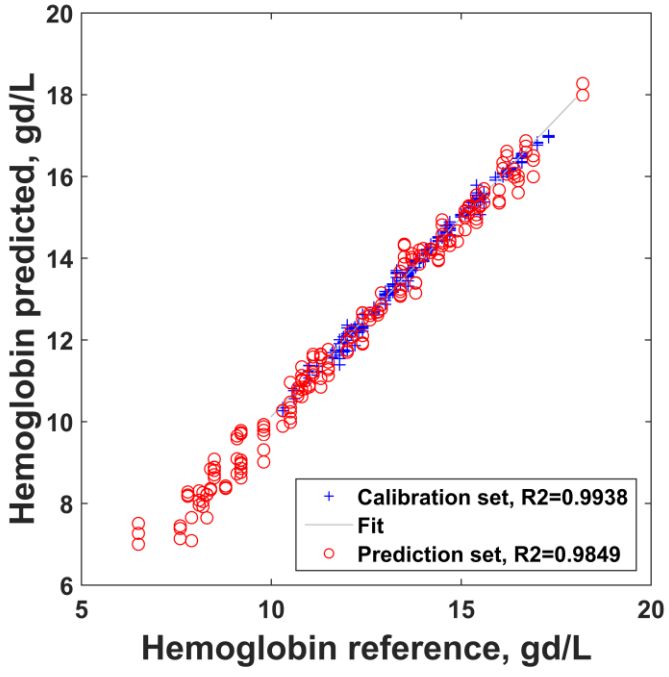
Fig. 3 The root mean square error of neural network (NN) versus the change of a number of hidden neurons with different Savitzky Golay (SG) preprocessing: (a) NN without SG preprocessing, (b) NN coupled with smoothing SG, (c) NN coupled with first order SG derivative, and (d) NN coupled second order SG derivative.



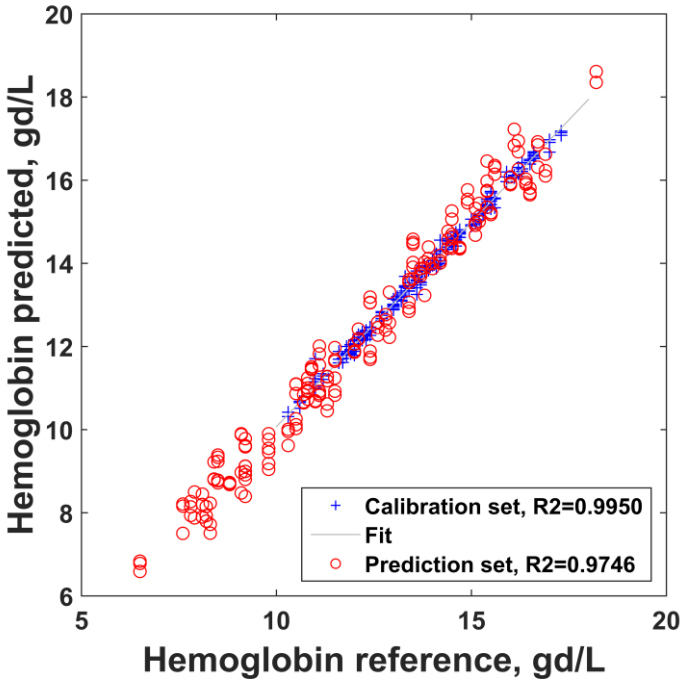
(a)



(b)



(c)



(d)

Fig. 4: Regression analysis plot of calibration and prediction set different preprocessing: (a) ANN without SG derivative, (b) ANN with smoothing SG, (c) ANN with first order SG derivative, and (d) ANN with second order SG derivative.

Table 3: The performance accuracy of the artificial neural network with different SG preprocessing methods with the optimal hidden neurons

The preprocessing method of the ANN	Hidden neurons	Frame length	Training		Prediction	
			RMSEC (gd/L)	R_c^2	RMSEP (gd/L)	R_p^2
Without preprocessing	3	N/A	0.1830	0.9890	0.4607	0.9734
SG smoothing	2	77	0.1115	0.9956	0.4499	0.9761
First order SG derivative	5	27	0.1364	0.9938	0.3517	0.9849
Second order SG derivative	3	79	0.1190	0.9950	0.4571	0.9746

4. Conclusion

This study investigated the feasibility of artificial neural network (ANN) and Savitzky Golay (SG) derivative in predicting blood hemoglobin using near-infrared spectrum. Results indicates that a combination of SG preprocessing and ANN improved predictive accuracy without wavelength selection and reduction data. In particular, the identification of an appropriate SG preprocessing and the optimal hidden neurons is important for the ANN to achieve its optimal predictive accuracy. Finding shows that ANN that coupled with the first order SG derivative with a proper optimization was able to achieve a better predictive accuracy in predicting the blood hemoglobin using near-infrared spectral data with RMSEP of 0.3517 gd/L and R_p^2 of 0.9849, compared with the previous works [11]. Nevertheless, various algorithms for multilayer perceptron and different ANN models coupled with SG preprocessing may be investigated to minimize overfitting that may occur during training process to further improve the prediction accuracy.

5. Acknowledgement

The author would like to acknowledge the Malaysia Ministry of Education for financial support, and Universiti Tun Hussein Onn Malaysia (UTHM) for providing facilities for this study.

References

[1] M. Chan, (2011). Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity: World Health Organization, Geneva, Switzerland, pp. 1–6.

[2] J. Rosenblit et al., (1999). Evaluation of three methods for hemoglobin measurement in a blood donor setting: Sao Paulo medical journal, vol. 117, no. 3, pp. 108–112.

[3] M. Goodarzi, S. Sharma, H. Ramon, and W. Saeys, (2015). Multivariate calibration of NIR spectroscopic sensors for continuous glucose monitoring: TrAC - Trends in Analytical Chemistry, vol. 67, pp. 147–158.

[4] C. Wood et al., (2016). Near infrared spectroscopy as a multivariate process analytical tool for predicting pharmaceutical co-crystal concentration: Journal of Pharmaceutical and Biomedical Analysis, vol. 129, pp. 172–181.

[5] V. Gomes, A. Fernandes, P. Martins-Lopes, L. Pereira, A. Mendes Faia, and P. Melo-Pinto, (2017). Characterization of neural network generalization in the determination of pH and anthocyanin content of wine grape in new vintages and varieties: Food Chemistry, vol. 218, pp. 40–46.

[6] N. C. T. Mariani, G. H. De Almeida Teixeira, K. M. G. De Lima, T. B. Morgenstern, V. Nardini, and L. C. C. Júnior, (2015). NIRS and iSPA-PLS for predicting total anthocyanin content in jaboticaba fruit: Food Chemistry, vol. 174, pp. 643–648.

[7] R. M. Balabin, E. I. Lomakina, and R. Z. Safieva, (2011). Neural network (ANN) approach to biodiesel analysis: Analysis of biodiesel density, kinematic viscosity, methanol and water contents using near infrared (NIR) spectroscopy: Fuel, vol. 90, no. 5, pp. 2007–2015.

[8] L. Norgaard, B. Rasmus, and S. B. Engelsen, (2014).Principal Component Analysis and Near Infrared Spectroscopy: A white paper from FOSS, pp. 2–7.

[9] J. Yadav, A. Rani, V. Singh, and B. M. Murari, (2015). Prospects and limitations of non-invasive blood glucose monitoring using near-infrared spectroscopy: Biomedical Signal Processing and Control, vol. 18, pp. 214–227.

[10] B. I. Fred McClure, (2012).Chemometrics ShootOut Rules: International Diffuse Reflectance Conference 2012.

[11] B. Igne et al., (2012). The 2010 IDRC software shoot-out at a glance: NIR news, vol. 21, no. 8, p. 14.

[12] X. Shao, X. Bian, J. Liu, M. Zhang, and W. Cai, (2010).Multivariate calibration methods in near infrared spectroscopic analysis: Analytical Methods, vol. 2, no. 11, p. 1662.

[13] H. A. G. Al-kaf, K. S. Chia, and N. A. M. Alduais, (2018).A comparison between single layer and multilayer artificial neural networks in predicting diesel fuel properties using near infrared spectrum:

Petroleum Science and Technology, vol. 6466, pp. 1–8.

[14] R. M. Balabin, R. Z. Safieva, and E. I. Lomakina, (2007). Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction: Chemom. Intell. Lab. Syst., vol. 88, no. 2, pp. 183–188.

[15] Å. Rinnan, F. van den Berg, and S. B. Engelsen, (2009). Review of the most common pre-processing techniques for near-infrared spectra: TrAC - Trends in Analytical Chemistry, vol. 28, no. 10, pp. 1201–1222.

[16] K. S. Chia, H. Abdul Rahim, and R. Abdul Rahim, (2013). Evaluation of common pre-processing approaches for visible (VIS) and shortwave near infrared (SWNIR) spectroscopy in soluble solids content (SSC) assessment: Biosystems Engineering, vol. 115, no. 1, pp. 82–88.

[17] Prieto et al., (2016).Neural networks: An overview of early research, current frameworks and new challenges: Neurocomputing, vol. 214, pp. 242–268,.

[18] N. S. A. Zubir, M. A. Abas, N. Ismail, N. A. M. Ali, M. H. F. Rahiman, and N. K. Mun, (2017).Analysis of Algorithms Variation in Multilayer Perceptron Neural Network for Agarwood Oil Qualities Classification: in 2017 IEEE 8th Control and System Graduate Research Colloquium (ICSGRC).

[19] G. E. Hinton and R. R. Salakhutdinov, (2006).Reducing the dimensionality of data with neural networks: Science, vol. 313, no. 5786, pp. 504–507.

[20] T. Nakama, (2011).Comparisons of single- and multiple-hidden-layer neural networks: 8th International Symposium on Neural Networks, vol. 6675 LNCS, no. PART 1, pp. 270–279.

[21] L. Van Efferen and A. M. T. Ali-eldin, (2017).A Multi-Layer Perceptron Approach for Flow-Based Anomaly Detection: in 2017 International Symposium on Networks, Computers and Communications (ISNCC).

[22] U. Andayani, E. Nababan, A. Muhammad, H. Tigor, and I. Nasution, (2017).Optimization Backpropagation Algorithm Based on Nguyen-Widrow Adaptive Weight and: 2017 4th International Conference on Industrial Engineering and Applications (ICIEA), pp. 363–367.

[23] D. Batra, (2010).Comparison Between Levenberg-Marquardt And Scaled Conjugate Gradient Training Algorithms for Image Compression Using MLP: Devesh Batra International Journal of Image Processing, no. 86, pp. 2014–412.

[24] I. Mukherjee and S. Routroy, (2012).Comparing the performance of neural networks developed by using Levenberg-Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process: Expert Systems with Applications, vol. 39, no. 3, pp. 2397–2407.

[25] S. Yadav and S. Shukla, (2016).Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification: Proceedings - 6th International Advanced Computing Conference, IACC 2016, no. Cv, pp. 78–83.

[26] M. I. Mohd Nazrul Effendy and K. S. Chia (2018). Partial Least Square with Savitzky Golay Derivative in Predicting Blood Hemoglobin Using Near Infrared Spectrum: MATEC Web Conf., vol. 150, p. 1001.