

RESEARCH

Open Access



# The performance of consumer-grade near infrared spectrometer in traditional Chinese medicine

Zhaozhou Lin<sup>1,2†</sup>, Zhihong Zhuang<sup>1†</sup>, Chongnian Luo<sup>2</sup>, Jiayu Zhang<sup>3\*</sup> and Hao Gu<sup>4\*</sup> 

## Abstract

Technical advances in instrument manufacturing have promoted the miniaturization and cost reduction of portable NIR spectrometers. The price of a device is now affordable to ordinary consumers, which might promote the application of NIRS in real scenarios. Generally, the portable spectrometers have a lower spectral resolution and a narrower spectral region compared with the benchtop ones. Whether the consumer grade portable spectrometers is good enough for basic analysis in TCM (Traditional Chinese Medicine) remains unclear. Two real world applications were introduced in this work to evaluate the capability of consumer grade spectrometers solving complex problems. Spectra collected on bark samples were used to test the qualitative performance of the spectrometer. The result showed that the cross validation error of the hierarchical FDA (Fisher Discrimination Analysis) models was at most 0.0769. For the quantitative analysis, spectra of pharmaceutical powder were used to train a model to predict the moisture of new individual sample. The RPD (the Ratio of Performance to Deviation) value of the moisture model was 6.83. These results demonstrated the usability of the models built on NIRS data measured by consumer grade spectrometer.

**Keywords:** Consumer grade, Portable; near infrared, Qualitative analysis, Quantitative analysis, Traditional Chinese medicine

## Introduction

As a non-destructive tool providing the qualitative and quantitative predictions, NIRS (Near InfraRed Spectroscopy) has been widely used in pharmaceutical sciences [1, 2], as well as in other fields including food [3], petrochemistry [4, 5], agriculture [6], forensic chemistry [7], medicine sciences [8, 9], etc. These works can be roughly divided into four categories: at-site, in-site, on-site and off-site, depending on the location where the analysis is performed. NIRS is a type of vibrational spectroscopy probing overtones, combinations and resonances of fundamental vibrations of chemical bonds like C-H, N-H, O-H and S-H [10]. These bonds exist in almost all organic molecules. All these sound

characteristics of NIRS determine its wide use in practice. Although massive researches have demonstrated the power of NIRS technique, practical applications only account for a small fraction of them. A possible reason is the low ratio of the potential income to the investment on applying the NIRS.

Thanks to technical advances, like the micro electro-mechanical system (MEMS) and the linear variable filter (LVF) technology [11], the performance of miniaturized devices have a great increase. Multiple companies have launched NIRS spectrometers fitting into the palm of a hand. With the miniaturization of NIRS instruments, the weight have reduced considerably. The price of NIRS device has dropped from hundreds of thousands of dollars to less than a thousand dollars (consumer-grade spectrometers), which might accelerate the promotion of NIRS in real scenarios. However, the performance of these portable spectrometers still need to be tested. Because the spectral resolution of the portable spectrometers are general worse than the benchtop ones.

\* Correspondence: [zhangjiayu0615@163.com](mailto:zhangjiayu0615@163.com); [hebegu@yeah.net](mailto:hebegu@yeah.net)

<sup>†</sup>Zhaozhou Lin and Zhihong Zhuang contributed equally to this work.

<sup>3</sup>School of Pharmacy, BIN ZHOU Medical University, Yantai 260040, People's Republic of China

<sup>4</sup>Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, People's Republic of China  
Full list of author information is available at the end of the article

Recent reports on the application of portable NIRS spectrometers have proven their good performance. Eduardo Maia Paiva et al. [12] compared the performances between the portable and the benchtop instruments to determine the biodiesel content in diesel-biodiesel blends. They found that the performance of the portable instrument was comparable to the benchtop one, thus it was possible to use the portable spectrophotometer to monitor in situ the quality characteristics of diesel-biodiesel blends. Although there was a need to improve the robustness of the calibration models, SCiO™ sensor performed similarly to commercial NIR sensors when used to predict quality of four horticultural products [6]. Marco Cirilli et al. [13] demonstrated the applicability of NIR-AOTF spectroscopy as a rapid and inexpensive technique for on-site monitoring of olive drupe physical properties during ripening and at maturation. In the review of William R. de Araujo et al. [7], NIR spectroscopy was considered one of the most widely explored techniques in forensic chemistry. Their reviewed forensic studies involves drug and explosive analysis, beverages screening, blood, gemstone and mineral analysis.

However, not all applications of portable NIRS are excellent. Hui Yan [11] prepared a solid pharmaceutical formulation consisting of two excipients and three active ingredients. This formulation was used to test the quantitative ability of four different portable NIRS spectrometers in pharmaceutical scenario. Except for the caffeine, the calibration performances of the other active ingredients vary depending on the instrument. Berta Baca-Bocanegra [14] even concluded that the use of the portable device for the “in vineyard” screening of extractable polyphenols in red grape skins was hampered by several environmental and physiological factors such as the heterogeneity and the own features of grapes analyzed. Besides, the performance also depends on many technical specifications like type of energy source and the detector, the resolution, sampling accessories, the instrument and the energy intensity. Therefore, the applicability of the consumer-grade

instrument for quantification and qualification analyses in TCM requires experimental illustration.

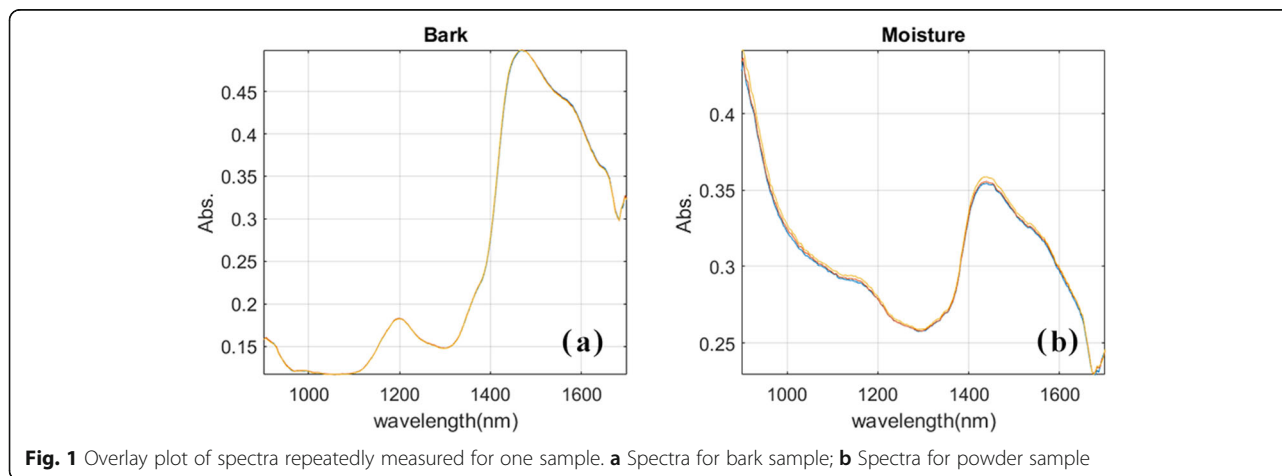
Because of the multiple practical advantages, such as providing rapid, accurate, and intact analysis, NIRS has become a popular technique in the research of TCM. Several papers have reviewed NIRS’s ability in authenticity identification, species identification, geographic origin analysis, quantitative analysis, adulteration detection, rapid detection, and on-line monitoring of TCM [15, 16]. Nevertheless, only parts of these applications was performed by portable instrument. The main aim of this work is to investigate the applicability of a consumer-grade NIRS device, NIRscan, for the species identification and the pharmaceutical preparation in TCM. The aforementioned device does not need any external probes, fiber optics or external illumination sources since all these parts are incorporated into a miniaturized module.

## Results

NIRS is generally combined with the latent variable analysis method to perform the quantitative and qualitative analysis, since the number of variables is commonly larger than the number of samples. The goodness of NIRS models is quantified by several parameters. For the qualitative models, classification accuracy, AUC curve (area under ROC (Receiver Operating Characteristic) curve), F1 score etc. are generally used. While, the metrics RMSECV (Root Mean Square Error of Cross Validation), RMSEC (RMSE of calibration), RMSEP (RMSE of Prediction), RPD,  $R^2$  between the prediction and the reference values, etc. are commonly used to score the quantitative models. In this work, part of these parameters were selected to quantify the feasibility of using consumer grade NIRS in TCM.

### Qualitative analysis

Figure 1a shows the overlay plot of three spectra repeatedly measured on nearly the same location of a bark



sample. It should be noticed that a whole measurement process included at least three steps: placing the spectrograph against a sample on the inner surface, activating the spectrograph and recording the data. Recording the spectral data repeatedly at a fixed location contribute little to the robustness of the resulted model. As expected, the spectral intensity were not always consistent, but varied in a small range. The variations at both ends of the spectra were fluctuant. Thus, 10 points from each end of the raw spectrum were eliminated. After that, spectra of the same sample were averaged to reduce the occasional error.

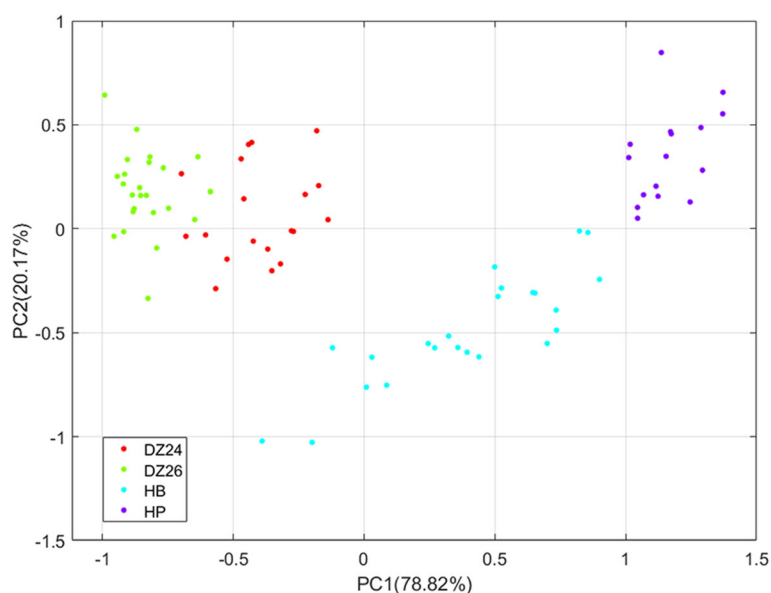
PCA was first performed on the raw spectra to illustrate the potential separation among cortex herbs. From the scatter plot of PCs in Fig. 2, it was found that all the three herbs were separated clearly. The samples of DZ (*Eucommia ulmoides*) harvested at 24 years could also be discriminated from the ones harvested at 26 years. Considering the good separation in the PC space, the FDA method was directly trained on the PC scores. Since FDA is designed for two-class classification, a dichotomous strategy was adopted in this paper. The samples were first classified into DZ and the others, and then samples in the others group were classified into HP (Officinal Magnolia Bark) and HB (Cortex Phellodendri Chinensis). To evaluate the prediction performance of FDA models, the stratified 10-fold cross validation was used and was repeated 10 times. The median of each metrics was presented in Table 1. The accuracy of the two basic FDA models with two PCs was 1 and 0.9231, the F1 score was 1 and 0.9333, respectively. The

accuracy of FDA model with two PCs for DZ sample harvested at different years was 0.9070, the F1 score was 0.8947. In general, data pre-treatment helps to filter out certain covariates, such as scattering and the variation in light path. From the above results, it can be seen that if only to classify the three herbs, no further pretreatment is required.

#### Quantitative calibration

The variation of spectra repeatedly measured is closely related to the quality of quantitative predictions. Figure 1b shows the spectra repeatedly measured for one sample. Since the variation on both ends change quickly, inner spectra from 960 nm to 1510 nm were remained to predict the moisture content. Spectra repeatedly measured are usually averaged to reduce accidental error. In this section, the original spectrum was also retained, because it is uncertain whether the effect of unwanted spectral variation can be effectively mitigated by simple averaging.

Thirty-two samples were selected from the averaged data set as calibration set by Kennard-Stone algorithm and the rest samples were incorporated into the test set. For the raw spectra data, the samples were divided based on the above indexes. The number of latent variables of the PLSR model on the training set was determined by 10-fold cross validation. Two LVs were selected to represent the averaged training data. The number of LVs was fixed the same for the following models, since it was found that the improvement of the PLS model with optimal number of LVs was limited compared with the



**Fig. 2** Scatter plot for bark spectra mapped onto the first two principal components (PCs). The percentage in the parentheses denotes the variation explained by each PCs

**Table 1** Prediction performance estimated by performing stratified 10-fold cross-validation 10 times

Group	Data	Method	Accuracy	F1 score	AUC
DZ v.s. HB & HP	PCA scores	FDA	1	1	0.9744
HB v.s. HP	PCA scores	FDA	0.9231	0.9333	0.9125
DZ24 v.s. DZ26	PCA scores	FDA	0.9070	0.8947	0.9430

model using two LVs. Consider the variation of cross validation, the mean of 10 repeated 10-fold cross validation was calculated. As shown in Table 2, the mRMSECV on the averaged data set was 0.3601, while the mRMSECV on the raw data set was 0.5764. The prediction error of the model, built by the averaged training data, on the averaged test set was 0.3076. If the model built using the raw training set was used to predict the moisture content of spectra from the averaged test set, a RMSEP value of 0.3252 was obtained. This result consists with the consensus that the averaging can improve the accuracy of the analysis. The regression plot for the PLS model built on the averaged spectral data was shown in Fig. 3.

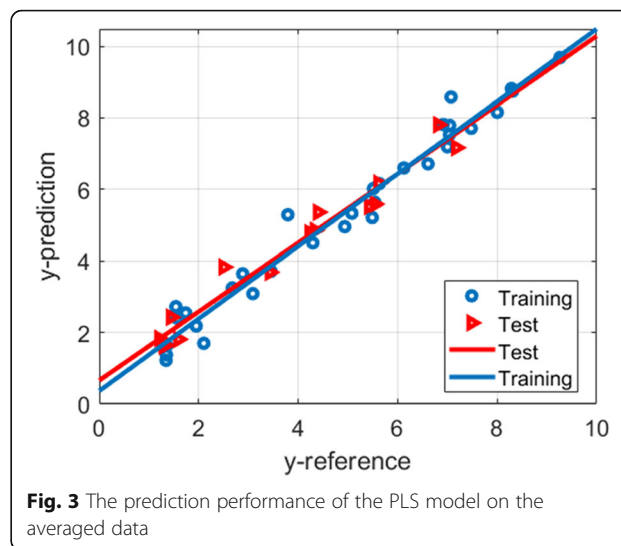
Since the value of RMSE increases with the measurement range or the mean, it is hard to draw conclusion based on RMSEs only. RPD has been widely appropriated by researchers to remove range effect. It is calculated as the ratio of the standard deviation (SD) of the reference data to RMSE of test set. The RPD value for moisture on the averaged data was 6.83, demonstrating the usability of the PLSR model built on NIRS data. The Kennard-Stone algorithm was also performed on the raw data, and the result was given in the last row of Table 2.

Scattering is one of the main factors affecting the quality of spectra. It can be mitigated by pre-treatment method, such as SNV (Standard Normal Variate transformation), ISC (Inverse Signal Correction) and MSC (Multiplicative Scatter Correction). In this work, the SNV method was adopted to reduce the negative of scattering on the averaged spectra. After pre-processing, the RMSECV of the model was 0.3688, the RMSEP on the test reduced to 0.2928. The RPD was improved to 7.17. Thus, NIRS was no doubt useful in pharmaceutical preparation process.

**Table 2** A comparison among the prediction performance of model trained on different splitting

Training data	mRMSECV	Test data	RMSEP	RPD
AVG	0.3601	AVG	0.3076	6.83
AVG	-	RAW	0.4920	4.17
RAW	0.5764	RAW	0.4171	4.92
RAW	-	AVG	0.3252	6.46
RAW <sup>1</sup>	0.5314	RAW	0.4484	4.58

<sup>1</sup>Sample splitting was performed on the raw data



**Fig. 3** The prediction performance of the PLS model on the averaged data

**Discussion**

For the two dichotomous models coupling to discriminate bark herbs, the accuracy estimated by stratified 10-fold cross validation was 0.9231 and 1, respectively. Considering the tremendous difference among the texture and components of bark herbs, these results were within expectation. Although the result of FDA on scores of DZ samples harvested at different years failed to separate the samples into two independent group, the value of AUC was 0.9430. These results demonstrated the usability of the consumer grade portable NIR instrument for qualitative analysis in TCM.

The moisture content is a critical quality control index, which appears in most solid pharmaceutical preparations and traditional Chinese decoction pieces. The moisture controlling is closely related to the heating process. According to Chinese pharmacopoeia, the drying process is generally executed under 80 or 60 °C depending on the volatileness of active ingredients [17]. Reducing the time of the pharmaceutical intermediate exposing to heat is beneficial for quality of pharmaceutical preparations. Thus, it is necessary to give a quick response to the analysis request. A possible way is to develop a method based on NIRS. The characteristic absorption of water lies in a large range of NIR spectrum around 750 nm, 930 nm, 1450 nm, 1850 nm and 2170 nm [18]. Three of these intervals lie in the spectral range of NIRscan.

Data pre-treatment methods are commonly employed to mitigate the negative effect of uncontrolled covariates, for instance, the scattering effect of power, particle size, temperature, etc. Although the SNV improved the prediction performance of PLS model on the averaged spectra in terms of RMSEP and RPD, the mRMSECV values became bigger after data pre-treatment. If the SEP of RPD in reference [19] was changed to SECV, the RPD

value would decrease after data pre-treatment. It seems that RPD is not a justice metrics for the prediction performance comparison. But if one compare the RPD values of PLS models in Table 2, it would be found that the RPD values are relatively stable than mRMSECV and RMSEP.

After pre-treated by SNV, the RPD of PLS model on the averaged spectra collected by NIRscan was improved from 6.83 to 7.17. As suggested by Chang *et. al.* [20], both PLSR models were excellent. However, there was no statistical basis for the thresholds, and other researcher gave higher thresholds [21]. Phil Williams [19] pointed out that the model with RPD larger than four could be used in any application of soils. Much the same as powder, herbal blend can at least be monitored by NIRS model with RPD larger than four.

## Conclusions

The present study demonstrated that the consumer grade portable NIRS spectrometer provide us a satisfactory tool to perform qualitative and quantitative analysis in TCM. In the space reduced by PCA, two hierarchical FDA models successfully classified the three bark herbs into three separate groups. The FDA model built with two PCs could also separate the DZ samples girdled at different years. For the quantitative analysis, the RPD value of PLSR model built on spectra pretreated by SNV was improved from 6.83 to 7.17, which means both models can be used for process control. Although the results are excellent, more experiments should be conducted to further validate the usability of the consumer grade spectrometer. Further work will be carried out in the near future.

## Methods

### Sample preparation and reference data

This study consisted of two parts: the spectral data collected for bark herbs were used to test the qualitative ability of consumer grade portable NIRS device, while the spectra of FUrong powder were used to evaluation the possible of using consumer grade device on process control in TCM.

Three kinds of cortex herbal medicines were collected from herbal market in Anguo. To make the sample as diverse as possible, four to six decoction pieces of cortex phellodendri chinensis (HB) and officinal magnolia bark (HP) were sampled from each store. Finally, 24 samples were collected for HB, 15 samples were collected for HP. The 19 samples of *Eucommia ulmoides* (DZ) girdled at their 24 years and 24 samples girdled at their 26 years were bought at a boutique herbal store.

FUrong powder is a hospital pharmaceutical preparation produced by Beijing hospital of traditional Chinese medicine. It is a blend of seven kinds of herbal powder.

As regulated in Chinese Pharmacopoeia, the moisture of powder must be lower than 9%. To monitor the moisture during drying process, five intermediate powder samples were collected for two batches. All the samples were dried for 24 h in a draught drying cabinet at 60 °C. Five gram of each sample were weighted using analytical balance (Sartorius) and then spread out on a watch glass. Each watch glasses was laid next to each other on a laboratory table and was covered with a box. Besides, there was a humidifier continuously providing aqueous vapor for the powders. Every 30 min, one sample was successively taken out from the box and weighted, and then, the NIRS of the sample was measured. The whole process was repeated five times.

### NIR spectra

The spectra of each sample were acquired by averaging 30 scans in the range of 900 to 1700 nm at 3.51 nm intervals with 228 spectral bands using a portable spectrometer (DLP NIRscan Nano Evaluation Module (EVM), Texas Instruments, Dallas, TX), in absorbance mode. The spectrometer was equipped with a reflective sampling module integrated with two tungsten lamps. The DLP NIRscan Nano Reference Software controlled the measurement process, with a diffuse reflectance standard as white reference.

For the bark samples, all samples were measured at nearly the same spots three times. Each spectrum was collected by direct contact in the central region of the inner surface of each sample. For measurement of the powder sample, every 30 min, a powder sample was transferred from the watch glass to a glass vial and the powder was pressed to make sure that there was no visible fissure at the bottom glass vial. After that, the glass vials containing powder samples were scanned by placing the vials directly on the top of the sapphire window. The whole measurement procedure was executed three times.

### Data analysis

As for NIRscan, every time a sample is measured, two data file of different format will be saved in the working path. The csv format file can be accessed by most modeling software. Thus after data collection, all following computation were finished in MATLAB (2017b, MathWorks, Natick, MA). Data-pretreatments such as mean centering, SNV and Kennard-Stone splitting were carried out by functions in MATLAB. Both fisher discriminant analysis and partial least square-discriminant analysis were trained based on PLSLDA toolbox (Version 2.0) [22]. While, parameters of PLSR and PCA were optimized by functions in MATLAB.

After elimination of noisy spectral regions, PCA was adopted to visualize the samples and detect potential



outliers for both applications. Since PCA was not the last step of data analysis, the number of PCs remained was not determined but set according to the rank of spectra matrix. Based on PC scores, FDA models was developed to classify different type of bark herbs and the same type of herb harvested at different years. The number of PCs used in the FDA model was determined by stratified cross validation. The performance of the models was evaluated in terms of classification accuracy, F1 score and AUC estimated by repeated stratified 10-fold cross validation.

After observing the scatter plot of moisture data, the spectra with moisture of 0% and 12% were removed from the initial spectra matrix as potential outlier, thus the sample size was reduced to 47. Kennard-Stone splitting was then performed on spectra data with and without averaging, since it is uncertain whether the effect of unwanted spectral variation can be effectively mitigated by simple averaging. PLSR models were developed on different training set. The optimal number of latent variables for each model was determined based on RMSECV. The performances of the models were also scored based on RMSEP, RPD and repeated RMSECV. The RPD is an important index for scoring the prediction capability of a calibration model and RPD values larger than six can generally be considered sufficient for process control [19–21]. Furthermore, after the pre-treatment with SNV, the PLSR model was re-built to evaluate the potential improve of the accuracy of prediction.

#### Abbreviations

AUC: Area under the ROC curve; DZ: *Eucommia ulmoides*; FDA: Fisher discriminant analysis; HB: Cortex phellodendri chinensis; HP: Official magnolia bark; ISC: Inverse Signal Correction; MSC: Multiplicative Scatter Correction; NIR: Near-infrared; NIRS: Near-infrared spectroscopy; PCA: Principal Component Analysis; PLSR: Partial Least Squares Regression; RMSECV: Root Mean Square Error of Cross Validation; RMSEP: RMSE of Prediction; ROC: Receiver operating characteristic; RPD: The Ratio of Performance to Deviation; SNV: Standard Normal Variate transformation; TCM: Traditional Chinese Medicine

#### Acknowledgments

The authors want to thank Yuzhan Gong and Di Jiang for sample collection.

#### Authors' contributions

Jiayu Zhang and Hao Gu conceived, initiated and designed the study. Zhihong Zhuang and Chongnian Luo collected and measured the samples. Zhaozhou Lin analyzed the data and visualized the results. Zhaozhou Lin and Hao Gu wrote the manuscript. Jiayu Zhang revised the manuscript. All authors read and approved the final manuscript.

#### Funding

This research was funded by National Natural Science Foundation of China with grant number 81603396 & 81603401.

#### Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Beijing Hospital of Traditional Chinese Medicine, Capital Medical University, Beijing 100010, People's Republic of China. <sup>2</sup>Beijing Institute of Traditional Chinese Medicine, Beijing 100035, People's Republic of China. <sup>3</sup>School of Pharmacy, BIN ZHOU Medical University, Yantai 260040, People's Republic of China. <sup>4</sup>Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, People's Republic of China.

Received: 3 October 2019 Accepted: 27 January 2020

Published online: 10 February 2020

#### References

1. Kallakunta, V.R., Sarabu, S., Bandari, S., Tiwari, R., Patil, H., Repka, M.A.: An update on the contribution of hot-melt extrusion technology to novel drug delivery in the twenty-first century: part I. *Expert Opin Drug Deliv.* **16**, 539–550 (2019)
2. Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., Jent, N.: A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J Pharm Biomed Anal.* **44**, 683–700 (2007)
3. Prieto, N., Roehe, R., Lavín, P., Batten, G., Andrés, S.: Application of near infrared reflectance spectroscopy to predict meat and meat products quality: a review. *Meat Sci.* **83**, 175–186 (2009)
4. Lovatti, B.P.O., Silva, S.R.C., Portela, N.D.A., Sad, C.M.S., Rainha, K.P., Rocha, J.T.C., Romão, W., Castro, E.V.R., Filgueiras, P.R.: Identification of petroleum profiles by infrared spectroscopy and chemometrics. *Fuel.* **254**, 115670 (2019)
5. Pasquini, C., Bueno, A.F.: Characterization of petroleum using near-infrared spectroscopy: quantitative modeling for the true boiling point curve and specific gravity. *Fuel.* **86**, 1927–1934 (2007)
6. Li, M., Qian, Z., Shi, B., Medlicott, J., East, A.: Evaluating the performance of a consumer scale SCIO™ molecular sensor to predict quality of horticultural products. *Postharvest Biol. Technol.* **145**, 183–192 (2018)
7. de Araujo, W.R., Cardoso, T.M.G., da Rocha, R.G., Santana, M.H.P., Muñoz, R.A., Richter, E.M., Paixão, T.R.L.C., Coltro, W.K.T.: Portable analytical platforms for forensic chemistry: a review. *Anal. Chim. Acta.* **1034**, 1–21 (2018)
8. Quaresima, V., Ferrari, M.: A mini-review on functional near-infrared spectroscopy (fNIRS): where do we stand, and where should we go? *Photonics.* (2019). <https://doi.org/10.3390/photonics6030087>
9. Eiken, F.L., Pedersen, B.L., Bækgaard, N., Eiberg, J.P.: Diagnostic methods for measurement of peripheral blood flow during exercise in patients with type 2 diabetes and peripheral artery disease: a systematic review. *Int. Angiol.* **38**, 62–69 (2019)
10. Pasquini, C.: Near infrared spectroscopy: a mature analytical technique with new perspectives – a review. *Anal. Chim. Acta.* **1026**, 8–36 (2018)
11. Yan, H., Siesler, H.W.: Quantitative analysis of a pharmaceutical formulation: performance comparison of different handheld near-infrared spectrometers. *J Pharm Biomed Anal.* **160**, 179–186 (2018)
12. Paiva, E.M., Rohwedder, J.J.R., Pasquini, C., Pimentel, M.F., Pereira, C.F.: Quantification of biodiesel and adulteration with vegetable oils in diesel/biodiesel blends using portable near-infrared spectrometer. *Fuel.* **160**, 57–63 (2015)
13. Cirilli, M., Bellincontro, A., Urbani, S., Servili, M., Esposto, S., Mencarelli, F., Muleo, R.: On-field monitoring of fruit ripening evolution and quality parameters in olive mutants using a portable NIR-AOTF device. *Food Chem.* **199**, 96–104 (2016)
14. Baca-Bocanegra, B., Hernandez-Hierro, J.M., Nogales-Bueno, J., Heredia, F.J.: Feasibility study on the use of a portable micro near infrared spectroscopy device for the "in vineyard" screening of extractable polyphenols in red grape skins. *Talanta.* **192**, 353–359 (2019)
15. Wang, P., Yu, Z.: Species authentication and geographical origin discrimination of herbal medicines by near infrared spectroscopy: a review. *J. Pharm Anal.* **5**, 277–284 (2015)
16. Yin, L., Zhou, J., Chen, D., Han, T., Zheng, B., Younis, a., Shao, Q.: A review of the application of near-infrared spectroscopy to rare traditional Chinese

medicine. *Spectrochim Acta Part A*. **221**, 1–9 (2019). <https://doi.org/10.1016/j.saa.2019.117208>

17. Chinese Pharmacopoeia 2015. IV, 0116
18. Xiaobo, Z., Jiewen, Z., Povey, M.J.W., Holmes, M., Hanpin, M.: Variables selection methods in near-infrared spectroscopy. *Anal Chim Acta*. **667**, 14–32 (2010)
19. Williams, P.: The RPD statistic: a tutorial note. *NIR News*. **25**, 22–26 (2010)
20. Chang, C.W., Laird, D. A., Mausbach, J., Hurburgh, C, M.: Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Sci Soc Am J*, **65**. 480–490 (2001)
21. Williams, P.C., Sobering, D.C.: Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spectrosc.* **1**, 25–32 (1993)
22. Hong-Dong, L., Yi-Zeng, L. [http://freesourcecode.net/matlabprojects/71983/a-library-of-pls-and-pls-lda-in-matlab-\(2019\)](http://freesourcecode.net/matlabprojects/71983/a-library-of-pls-and-pls-lda-in-matlab-(2019)). Accessed 02 Oct. 2019

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---