

Non-destructive determination of silymarin in *Silybum marianum* extracts with a handheld near infrared spectrometer

Hui Yan,^{a,b*} Mengxing Liang,^a Mingzhu Jiang^{a,b} and Heinz W. Siesler^b

^aSchool of Biotechnology, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212018, China. E-mail: yanh1006@163.com

^bDepartment of Physical Chemistry, University of Duisburg-Essen, D 45117 Essen, Germany. E-mail: hw.siesler@uni-due.de

The main objective of this study is to provide a potential method of detecting the silymarin in *Silybum marianum* extracts. Near infrared (NIR) spectroscopy and a handheld NIR spectrometer was used for the first time to detect the *Silybum marianum* extracts. A total of 80 samples were collected. The content of silymarin was determined in all samples by high-performance liquid chromatography. Then 56 samples were selected for the development of a calibration model, and the other 24 samples were used as test set. The diffuse reflection spectra acquired by the handheld MicroNIR 1700 spectrometer were pretreated by standard normal variate, and then a partial least squares regression model was developed. The results showed that the optimum number of factors was four. For calibration, cross validation (CV) and test set, the slopes were all larger than 0.98, the R^2 values were 0.9874, 0.9839 and 0.9888, respectively, the root mean square error of calibration, root mean square error of CV and root mean square error of prediction were 1.06%, 1.25% and 0.98%, respectively, and the residual predictive deviation was 7.56, which indicated that the calibration model was perfect and has high precision predictive ability. The overall results of this study revealed the feasibility for use of handheld NIR spectroscopy as a non-destructive method for the quantitative on-site determination of the silymarin content in *Silybum marianum* extracts.

Keywords: handheld near infrared spectroscopy, quantitative determination, silymarin, *Silybum marianum* extracts

Introduction

Medicinal and edible plants (herbs) have great effects on human health and their use has been widely accepted in recent decades.¹ Nowadays, plant extracts have received a lot of attention because of their unique role for human health. Due to the increasing recognition for life and health level, people have put forward higher requirements in the application and development of plants extracts. Strict quality control is the key factor in guaranteeing plants extracts to be safe, stable and effective.

Silybum marianum is a member of the aster family (*Asteraceae* or *Compositae*), and has been traditionally used in the treatment of liver diseases for more than 2000 years.²⁻⁴ The main active ingredient in *Silybum marianum* extract is

silymarin, which is composed of flavonolignans that include silybin, silychristin, silydianin and isosilybin.⁵ It is well-known for its antioxidant and chemoprotectant effects on the liver. In laboratory studies, silybin has been found to stabilise cell membranes to prevent toxic chemicals from entering the cell. Presently, silymarin has been widely used for toxic liver as the active ingredient in many dosage forms in clinical practice.⁶ Recently, *Silybum marianum* extract has also been marketed, and demand shows a growing trend.^{7,8}

According to International Business Standards (SWT 7-2013, China), the silymarin content that varies in dry *Silybum marianum* extract between 30% and 65% is determined by high-performance liquid chromatography (HPLC).

Correspondence

H. Yan (yanh1006@163.com)

doi: 10.1255/nir2017.135

Citation: H. Yan, M. Liang, M. Jiang and H.W. Siesler, "Non-destructive determination of silymarin in *Silybum marianum* extracts with a handheld near infrared spectrometer", in Proc. 18th Int. Conf. Near Infrared Spectrosc., Ed by S.B. Engelsen, K.M. Sørensen and F. van den Berg. IM Publications Open, Chichester, pp. 135-141 (2019). <https://doi.org/10.1255/nir2017.135>

© 2019 The Authors

This licence permits you to use, share, copy and redistribute the paper in any medium or any format provided that a full citation to the original paper is given, the use is not for commercial purposes and the paper is not changed in any way.



ISBN: 978-1-906715-27-4

However, because of time requirements, high cost and constraint to the lab, the HPLC method is not necessarily the technique of choice. As an analysis method, near infrared (NIR) spectroscopy is suitable for rapid measurement. NIR spectroscopy has demonstrated its potential as a tool for rapid analysis of plants extract, such as honeysuckle extracts⁹ and bamboo leaf extract¹⁰ and so on. However, due their size, high power consumption and high price, it is difficult for benchtop spectrometers to promote their implementation in the market as on-site measurement tool.^{11,12} In recent years, portable NIR spectrometers have attracted great attention due to their portability and low energy consumption, which makes them more suitable for in-the-field detection,¹³ and has shown an extensive application prospect.

Thus, the handheld NIR spectrometer, MicroNIR 1700 (Viavi, Santa Rosa, USA) based on linear variable filter technology, is small and quite suitable for being used outside the laboratory. The portability of the MicroNIR spectrometer allows convenient non-destructive and rapid detection, providing a good basis for on-site applications. Therefore, the main objective of this work was to explore the possibility of detecting the content of silymarin in *Silybum marianum* extracts by on-site NIR spectroscopy.

Materials and methods

Samples

A total of 80 different batches of *Silybum marianum* extracts were prepared by Jiangsu Jianjia Pharmaceutical Industry Co., Ltd, in which the *Silybum marianum* extract is one of the main products.

Silymarin reference values

The reference values of silymarin were measured by HPLC based on the method provided by the Pharmacopoeia of the People's Republic of China.¹⁴

Definition of calibration and test sets

A hierarchical method was used to select the calibration and test sets.^{15,16} Samples were sorted by the silymarin concentration values from small to large, and then all the samples were marked 1, 2 and 3 in proper order, with the last sample being marked 1. The samples marked 1 and 3 were selected as the calibration set while those marked 2 were used as test set. Finally, 54 samples were

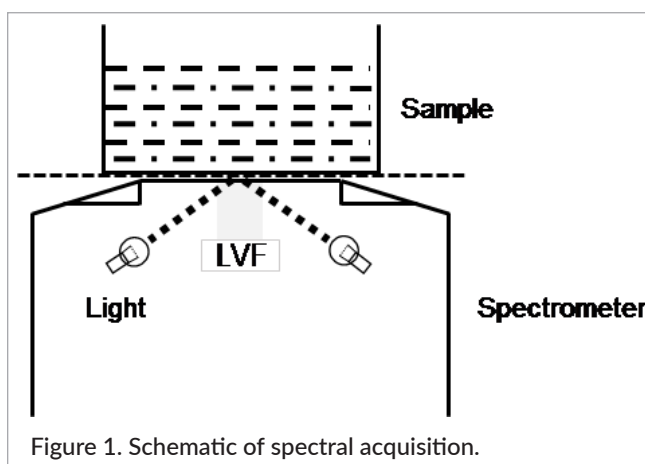


Figure 1. Schematic of spectral acquisition.

selected as calibration set and 26 samples were selected as test set.

Scanning of NIR spectra

The MicroNIR 1700 was used to collect spectra of samples as shown in Figure 1. The powder was placed in a cup with a 1 mm quartz window in the bottom, through which diffuse-reflection spectra were recorded. Each individual spectrum was the average of 50 scans with a resolution of 12.5 nm over the wavelength range of 950–1650 nm. A certified reflection standard (Labsphere, USA) was used to measure the reference spectrum. After each measurement, the sample was rotated by 120° and the average of three spectra was then used as the final raw spectrum.

Spectral pretreatment

Spectral preprocessing can significantly improve calibration model performance. NIR spectra frequently contain background information, drift and noise and in order to obtain reliable, accurate and stable calibration models, it is necessary to preprocess spectral data before modeling.¹⁷ In this work, SNV was applied as pretreatment method to eliminate the effect of light scattering.

Chemometrics and data analysis

The partial least squares (PLS) calibration was developed using The Unscrambler X [Version: 10.4 (64-bit), CAMO Software AS, Oslo, Norway], and internal cross validation was used to select the best number of factors.¹⁸ Cross validation estimated the prediction error by splitting all samples into 20 segments, and one segment was reserved for validation and the remaining 19 segments

Table 1. The population statistics of the sample reference values (in percentage).

Data sets	Mean	Max	Min	Range	Std	Variance	Median	Q1	Q3
Total	49.00	64.62	29.28	35.34	9.29	86.32	48.49	41.97	55.73
Calibration	49.17	64.62	29.28	35.34	9.43	88.92	48.61	42.25	56.50
Test	48.65	64.02	30.81	33.21	9.17	84.09	48.36	42.09	54.84

used for calibration. This process was repeated until all segments were used for validation once.¹⁸ Calibration statistics included the slope, offset, R^2 , root mean square error of calibration ($RMSEC$), root mean square error of cross-validation ($RMSECV$), root mean square error of prediction ($RMSEP$) and bias. The residual predictive deviation (RPD) was used to evaluate how well the calibration model could predict compositional data. The higher the RPD value the better the predictive ability of the model. Generally, an RPD greater than three could be considered very good for prediction purposes.^{19,20}

Results and discussions

Silymarin reference value

The concentration of silymarin in *Silybum marianum* extracts is shown in Table 1. As for the total samples, the mean and range were 49.00% and 35.34%, respectively, and the standard deviation (std) was 9.29%. The skewness was 0.00 and kurtosis was -0.80 (close to 0), which indicated that it was a normal distribution. The statistical parameters of the calibration and test set were as similar

as that of the total samples, which suggested that they were reasonable.

Characterisation of spectral data

The raw NIR spectra of all samples are shown in Figure 2. In the spectral range used in the present study, the two characteristic O-H absorption bands at 990 nm and 1450 nm are related to the 2nd and 1st overtone of the O-H stretching vibration, respectively, whereas the band at 1190 nm is a $2 \times \nu(\text{OH}) + \delta(\text{OH})$ combination band. The absorption band around 1120 nm corresponds to the 2nd C-H overtone, and the band at 1355 nm can be assigned to a $2 \times \nu(\text{CH}_2) + \delta(\text{CH}_2)$ combination band. Finally, the peak at about 1650 nm is the 1st overtone of the C-H stretching vibration.²¹ These bands are related to the calibration of silymarin.

As shown in Figure 3, the spectra pretreated by SNV highlight the differences at these wavelength positions and enhance the spectral information.

Calibration and analysis

Outlier analysis. An outlier is an object which deviates from the other model objects and may not belong to the

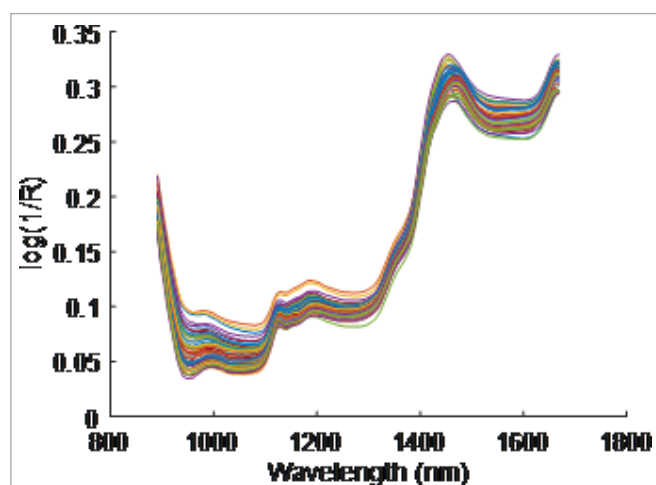


Figure 2. Original near infrared spectra of all samples.

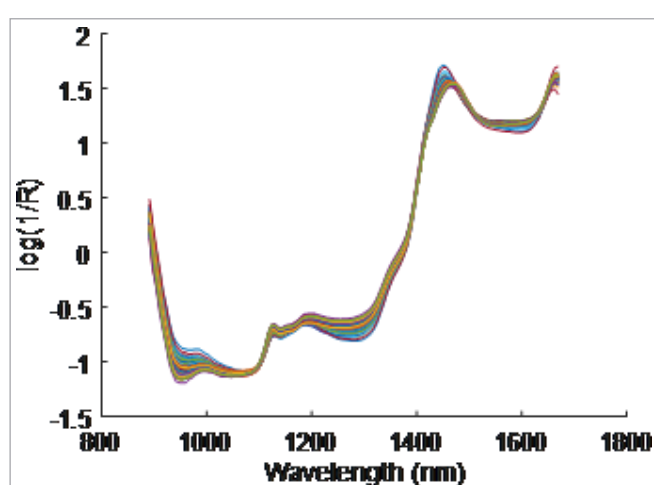
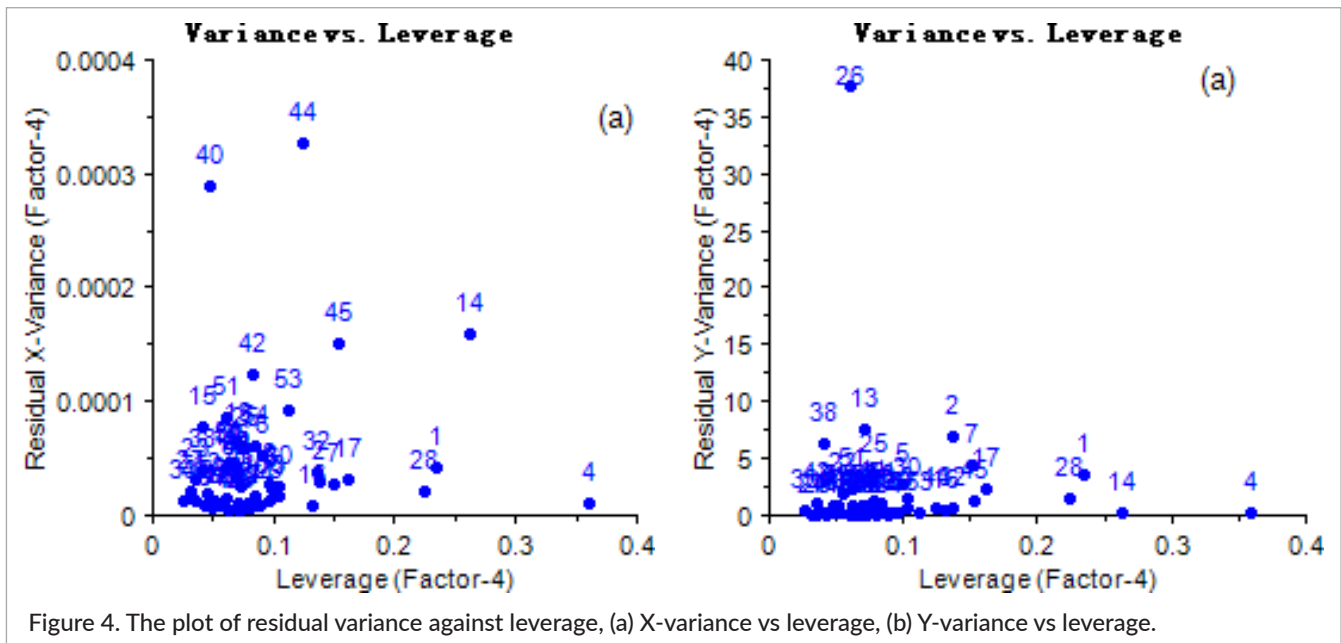


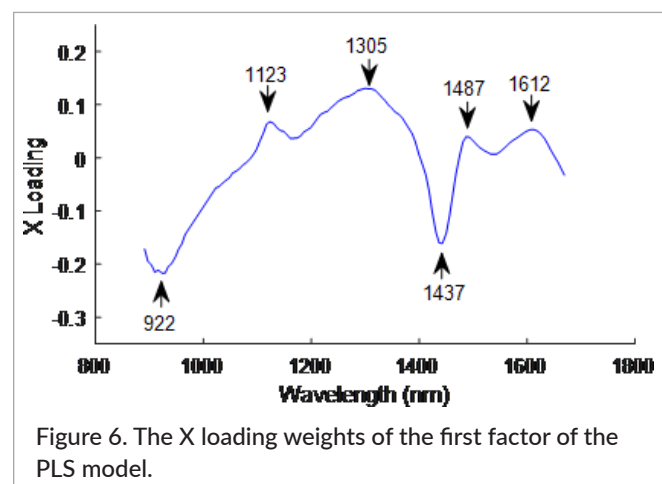
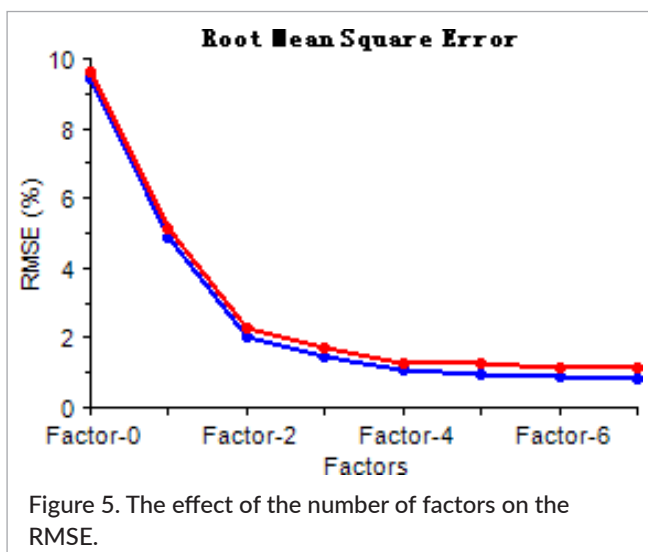
Figure 3. Original spectra after SNV pretreatment.

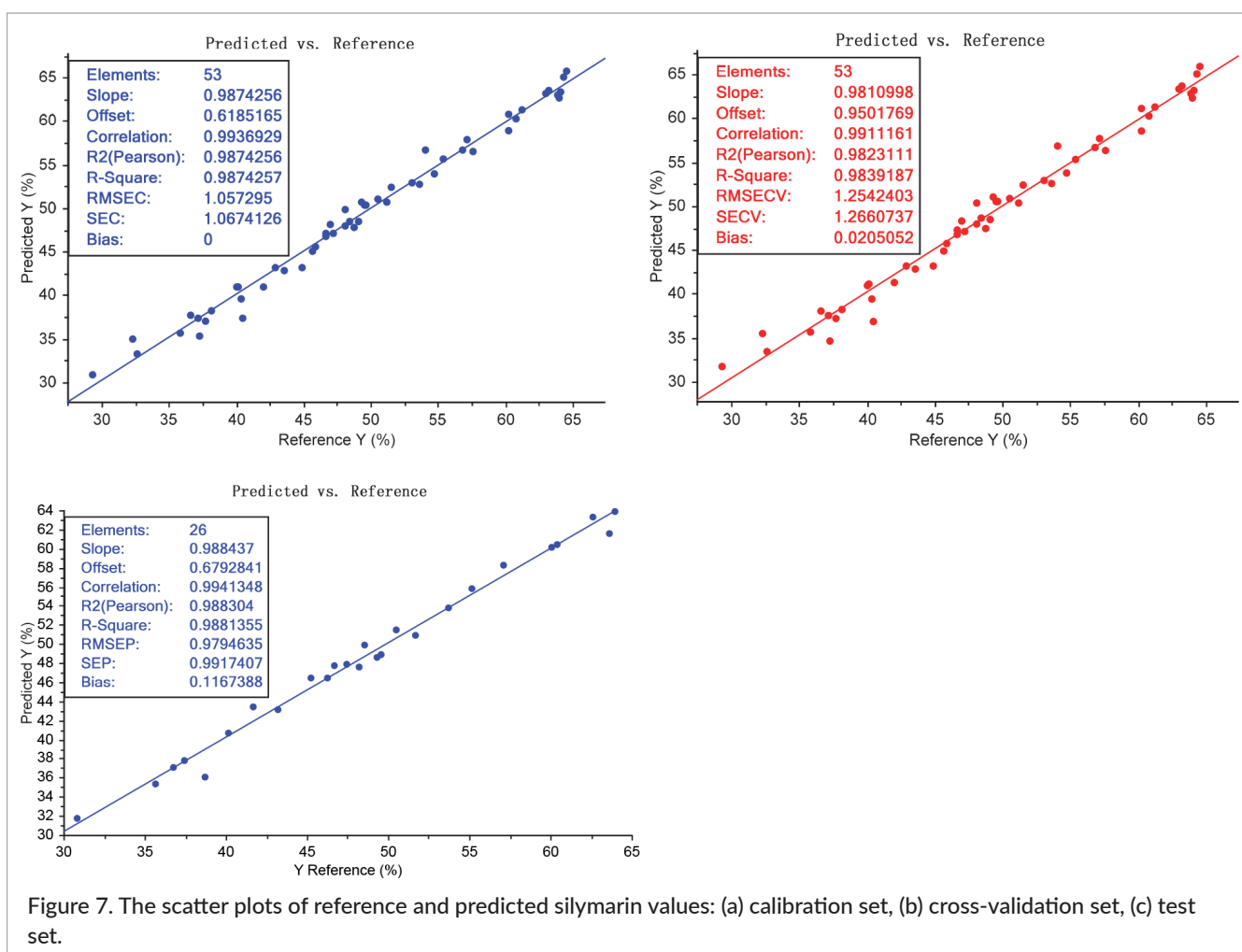


same population as the majority and therefore can disturb the model. The cause of outliers could be measurement error, noise, extreme sample and others.¹⁹ For the PLS method, outliers can be detected by using scores plots or the plots of residuals, leverage and influence. When outliers were judged by the values of leverage and residuals, the larger the leverage and the smaller the residuals, the more the data contributed to the calibration.⁹

As shown in Figure 4, the samples show a normal distribution, however, with a value of 48.32% sample 26 is far away from the others [Figure 4(b)] and was removed from the calibration set.

Number of factors. For PLS calibrations, the number of factors has a great effect on the performance. Less factors lead to lower precision, however, more factors can induce overfitting with an apparent improvement of predictive ability. Thus, an appropriate number of factors should be determined for the respective calibration. In this work, cross validation has been used and the result is shown in Figure 5. With an increase of factors, the *RMSECV* value initially decreased quickly but showed no further big changes after four components. The values of *RMSEC* and *RMSECV* became relatively low and close to each other and four was therefore chosen as the optimum number of factors.





X-loading weights analysis. Wavelengths with large loading weight values have a strong influence for the corresponding component.²² In the present application the first loading is very important because it explains 90.20% of the X variation. The X-loading weights of the 1st factor are shown in Figure 6. There are six wavelength positions with high weights, which means that they are important for the calibration:

- the wavelength position at 922 nm corresponds approximately to the 2nd OH overtone
- 1123 nm can be assigned to the 2nd CH overtone
- the 1305 nm wavelength position is in the range of the CH₂ combination band
- the 1437 nm signal is in the range of the 1st overtone of the $\nu(\text{OH})$ vibration
- at 1487 nm the 1st overtone of $\nu(\text{NH})$ vibration occurs
- the 1612 nm signal is close to the 1st overtone of C–H stretching vibrations.

Thus, the above key group vibrations play an important role in the discussed PLS calibration model.

Calibration statistics analysis. The reference/prediction scatter plots²³ of the silymarin content that make the calibration model more descriptive are shown in Figures 7a–c. Good correlation between reference and NIR-predicted values along the whole concentration range were found for the calibration, cross validation and prediction sets. Furthermore, the slopes are >0.98 and the offsets are all <1 . The R^2 values are 0.9874, 0.9839 and 0.9888, for calibration, cross validation and test set, respectively. The *RMSEC*, *RMSECV* and *RMSEP* values are 1.06%, 1.25% and 0.98%, respectively, reflecting a very good calibration model. The bias values of the three sets are 0, 0.021% and 0.1167%, respectively. Thus, based on the calibration statistics the calibration model is very good. The RPD value is 7.56 ($\gg 3$) and represents a calibration model with a good prediction performance.

Table 2. The population statistics of the NIR predicted values (in percentage).

Predicted sets	Mean	Max	Min	Range	Std	Variance	Median	Q1	Q3
Calibration	49.19	65.73	30.84	34.89	9.46	89.47	48.54	41.00	56.62
CV	49.21	65.95	31.65	34.30	9.42	88.79	48.55	41.24	56.63
Test	48.77	63.89	31.69	32.20	9.12	83.13	48.16	43.19	55.31

The population statistical features of the sample reference and the NIR-predicted values are shown in the Tables 1 and 2. The statistical parameters of the predicted values of calibration, cross validation and test sets are similar to the reference value sets.

Conclusions

The handheld NIR spectrometer MicroNir 1700 can be used in combination with a chemometric PLS evaluation routine for the rapid quantitative determination of silymarin content in *Silybum marianum* extract. The developed calibration model shows a good performance with R^2 values of 0.9874, 0.9839 and 0.9888, for calibration, cross validation and test set, respectively, the corresponding $RMSEC$, $RMSECV$ and $RMSEP$ values were 1.06%, 1.25% and 0.98%, respectively, and an RPD value of 7.56 was determined.

Acknowledgements

This work was supported by the Chinese medicine industry under the Special Project of the Ministry of Science and Technology: Rapid detection method of Chinese herbal medicine quality (grant number 201407003).

References

1. L. Kristanc and S. Kreft, "European medicinal and edible plants associated with subacute and chronic toxicity part II: Plants with hepato-, neuro-, nephro- and immunotoxic effects", *Food Chem. Toxicol.* **2**, 38–49 (2016). <https://doi.org/10.1016/j.fct.2016.03.014>
2. M. Cavaretta, "Therapeutic review: milk thistle", *J. Exot. Pet. Med.* **24(4)**, 470–472 (2015). <https://doi.org/10.1053/j.jepm.2015.09.001>
3. B.K. Puri, "The clinical advantages of cold-pressed non-raffinated evening primrose oil over refined preparations", *Med. Hypotheses* **62(1)**, 116–118 (2004).
4. R. Saller, J. Melzer, J. Reichling, R. Brignoli and R. Meier, "An updated systematic review of the pharmacology of silymarin", *Forsch. Komplementmed.* **14(2)**, 70–80 (2007). <https://doi.org/10.1159/0000100581>
5. L. Calani, F. Brighenti, R. Bruni and D. Del Rio, "Absorption and metabolism of milk thistle flavanolignans in humans", *Phytomedicine* **20(1)**, 40–46 (2012). <https://doi.org/10.1016/j.phymed.2012.09.004>
6. M. Lahiri-Chatterjee, S.K. Katiyar, R.R. Mohan and R. Agarwal, "A flavonoid antioxidant, silymarin, affords exceptionally high protection against tumor promotion in the SENCAR mouse skin tumorigenesis model", *Cancer Res.* **59(3)**, 622–632 (1999).
7. D.Y. Lee and Y. Liu, "Molecular structure and stereochemistry of silybin A, silybin B, isosilybin A, and isosilybin B, isolated from *Silybum marianum* (milk thistle)", *J. Nat. Prod.* **66(9)**, 1171–1174 (2003). <https://doi.org/10.1021/np030163b>
8. N. Ahmad, H. Fazal, B.H. Abbasi, S. Anwar and A. Basir, "DPPH free radical scavenging activity and phenotypic difference in hepatoprotective plant (*Silybum marianum* L.), *Toxicol. Ind. Health* **29(5)**, 460–467 (2013). <https://doi.org/10.1177/0748233712436637>
9. Z. Wu, B. Xu, M. Du, C. Sui, X. Shi and Y. Qiao, "Validation of a NIR quantification method for the determination of chlorogenic acid in *Lonicera japonica* solution in ethanol precipitation process", *J. Pharm. Biomed. Anal.* **62**, 1–6 (2012). <https://doi.org/10.1016/j.jpba.2011.12.005>
10. D. Wu, J. Chen, B. Lu, L. Xiong, Y. He and Y. Zhang, "Application of near infrared spectroscopy for the rapid determination of antioxidant activity of bamboo leaf extract", *Food Chem.* **135(4)**,

- 2147–2156 (2012). <https://doi.org/10.1016/j.foodchem.2012.07.011>
11. C.A.T. Dos Santos, M. Lopo, R.N. Páscoa and J.A. Lopes, “A review on the applications of portable near-infrared spectrometers in the agro-food industry”, *Appl. Spectrosc.* **67(11)**, 1215–1233 (2013). <https://doi.org/10.1366/13-07228>
 12. L. Salguero-Chaparro, B. Palagos, F. Peña-Rodríguez and J. Roger, “Calibration transfer of intact olive NIR spectra between a pre-dispersive instrument and a portable spectrometer”, *Comput. Electron. Agr.* **96**, 202–208 (2013). <https://doi.org/10.1016/j.compag.2013.05.007>
 13. K. Tsujikawa, T. Yamamuro, K. Kuwayama, T. Kanamori, Y.T. Iwata, K. Miyamoto, F. Kasuya and H. Inoue, “Application of a portable near infrared spectrometer for presumptive identification of psychoactive drugs”, *Forensic Sci. Int.* **242**, 162–171 (2014). <https://doi.org/10.1016/j.forsci-int.2014.05.020>
 14. The Pharmacopoeia Committee of People’s Republic of China, *Chinese Pharmacopoeia*. China Medical Science Press, Beijing (2015).
 15. Q. Chen, J. Ding, J. Cai and J. Zhao, “Rapid measurement of total acid content (TAC) in vinegar using near infrared spectroscopy based on efficient variables selection algorithm and nonlinear regression tools”, *Food Chem.* **135(2)**, 590–595 (2012). <https://doi.org/10.1016/j.foodchem.2012.05.011>
 16. Q. Chen, J. Zhao, H. Zhang and X. Wang, “Feasibility study on qualitative and quantitative analysis in tea by near infrared spectroscopy with multivariate calibration”, *Anal. Chim. Acta* **572(1)**, 77–84 (2006). <https://doi.org/10.1016/j.aca.2006.05.007>
 17. H. Cen and Y. He, “Theory and application of near infrared reflectance spectroscopy in determination of food quality”, *Trends Food Sci. Technol.* **18(2)**, 72–83 (2007). <https://doi.org/10.1016/j.tifs.2006.09.003>
 18. T. Næs, T. Isaksson, T. Fearn and T. Davies, *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, UK (2002).
 19. T. Fearn, “Assessing calibrations: SEP, RPD, RER and R²”, *NIR news* **13(6)**, 12–14 (2002). <https://doi.org/10.1255/nirn.689>
 20. P.C. Williams, “Application of near infrared reflectance spectroscopy to analysis of cereal grains and oilseeds”, *Cereal Chem.* **52**, 561–576 (1975).
 21. L.M. Kandpal, S. Lohumi, M.S. Kim, J.-S. Kang and B.-K. Cho, “Near-infrared hyperspectral imaging system coupled with multivariate methods to predict viability and vigor in muskmelon seeds”, *Sensor. Actuat. B* **229**, 534–544 (2016). <https://doi.org/10.1016/j.snb.2016.02.015>
 22. Y. Wang, Y. Gao, X. Yu, Y. Wang, S. Deng and J.-m. Gao, “Rapid determination of *Lycium barbarum* polysaccharide with effective wavelength selection using near-infrared diffuse reflectance spectroscopy”, *Food Anal. Method.* **9(1)**, 131–138 (2015). <https://doi.org/10.1007/s12161-015-0178-7>
 23. L. Xie, X. Ye, D. Liu and Y. Ying, “Prediction of titratable acidity, malic acid, and citric acid in bayberry fruit by near-infrared spectroscopy”, *Food Res. Int.* **44(7)**, 2198–2204 (2011). <https://doi.org/10.1016/j.foodres.2010.11.024>