Optik 127 (2016) 1405–1407

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Determination of active ingredients in matrine aqueous solutions by mid-infrared spectroscopy and competitive adaptive reweighted sampling

Guo Tang, Jing Hu[1], Hong Yan, Yuquan Zhao, Yanmei Xiong*, Shungeng Min*

College of Science, China Agricultural University, Beijing 100193, PR China

ARTICLE INFO

Article history:
Received 7 November 2014
Accepted 11 September 2015

Keywords:
Mid-infrared
Competitive adaptive reweighted sampling
Matrine

ABSTRACT

In this paper, we use MIR and variable selection techniques to determine the active ingredients of matrine aqueous solutions. Competitive adaptive reweighted sampling (CARS) was engaged to investigate effective variables and compared with full-spectrum partial least squares (PLS) regression. Results indicate that MIR is very efficient in aqueous solutions quantitative determination. In the meantime, CARS is proved to be an effective variable selection approach to eliminate spectral redundancy.

1. Introduction

Nowadays, mid-infrared spectroscopy (MIR) have been widely used as a rapid, non-invasion and reliable technique in many fields including agriculture [1,2], petrochemical industry [3] and pharmacy [4,5] etc. instead of chromatography.

Researches of pesticide active ingredients determination by vibrational spectroscopy have been reported by Armenta [6,7], Xiong [8,9] and Khanmohammadi et al. [10]. However, these researches just focus on emulsifiable concentrate and powders while studies on aqueous solutions have not been reported. In the recent years, aqueous solutions are gaining credibility as a more environment-friendly pesticide formulation for its solvents are usually water or alcohol. Since water has intense absorption in MIR [11], it is difficult to obtain good quality spectrum in transmission mode. However, with attenuated total reflection (ATR) by MIR, this problem can be tactfully solved.

On the other hand, modern chemometrics studies have shown that proper variable selection is necessary for multivariate spectroscopic calibration [12,13]. The competitive adaptive reweighted sampling (CARS) method is a recently proposed method, which has been proved very efficient when applied to spectral data [14].

Matrine is a widely used low toxicity botanical pesticide with contacting action, stomach poisoning and antibacterial action. It is generally used to protect crops from pest such as *Lepidoptera* and plant pathogens (e.g. anthracnose) etc. [15]. In this paper, we aim to use MIR to develop a rapid and reliable method to determine the active ingredient in aqueous solutions. MIR was employed to obtain spectra and the quantitative model was established by full-spectrum PLS and CARS–PLS, respectively. CARS here was used to investigate the effective variables and promote the full-spectrum models.

2. Experimental

Seventy-five samples were prepared in total with certain pesticide formulation, water and matrine technical for spectra collection. In order to avoid collinearity, the three components were added randomly, the concentration ranges of matrine was 0.10–1.00% (w/w). All samples were divided into calibration set (50 samples) and prediction set (25 samples) by K-Stone sampling.

Matrine aqueous solutions (0.38%, Beijing Yageer Biological Pharmacy Co., Ltd, China); Matrine technical (98.0%, FangSheng Biological Development Co., Ltd, China); carbon tetrachloride (A.R., Beijing Chemical Works, China).

MIR spectra were acquired by an FT-IR spectrometer (Nicolet 6700, Thermo Scientific, USA) over the range 650–4000 cm^{-1} at room temperature (resolution 4 cm^{-1}, 32 scans) with ATR accessory (diamond). The cuvette (for NIR) and sample cell was rinsed with carbon tetrachloride between the samples. Spectra acquisition and instrument control was performed by OMNIC (v8.2, Thermo Scientific, USA) and ATR correction was performed after recording the raw spectra.

The spectra files were imported into Matlab (v7.11, MathWorks, USA) for data analysis. The scripts used in this study are based on

* Corresponding authors. Tel.: +86 10 62733091; fax: +86 10 62733091.
E-mail addresses: yanmeix@sina.com (Y. Xiong), minsg@cau.edu.cn (S. Min).
[1] The author contributed equally as the first author to this paper.

http://dx.doi.org/10.1016/j.ijleo.2015.09.139
0030-4026/

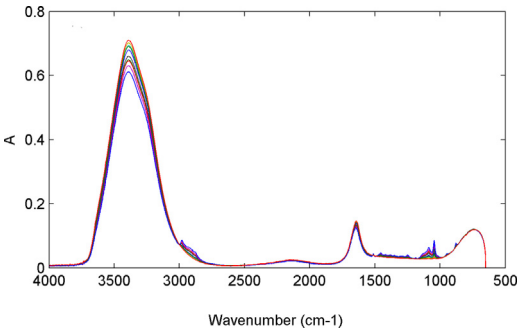

Fig. 1. MIR spectra of the prepared samples.

Table 1
Results of the MIR models.

Methods	PLS factors	Variables	RMSECV (%)	RMSEP (%)
MIR–PLS	5	6950	0.0372	0.0452
MIR–CARS–PLS	4	108	0.0158	0.0176

Ref. [16] which is available as http://code.google.com/p/carspls/. Full-range MIR data matrix: 75 rows and 6950 columns (samples and variables, respectively).

3. Result and discussion

The MIR spectra of the prepared samples, matrine technical (dissolved in tetrachloride) were shown in Fig. 1. For MIR spectra: peaks in the range of 2500–4000 cm^{-1} are the stretch vibration of X—H; 1500–1700 cm^{-1} are the stretch vibration of C═O; 1000–1200 cm^{-1} are attributed to the stretch vibration of C—H and C—N; 650–1000 cm^{-1} are the flexural vibration of C—H, O—H/H_2O and C—N [11].

3.1. Full-range models

As conventional spectral pretreatments (smoothing, normalization and derivative etc.) made no difference or even worse result, the regression was processed on the raw spectra. The regression results of full-range MIR model are shown in Table 1. MIR presented acceptable results in determination of matrine in aqueous solutions. It should be noted that, since active ingredient in this matrine formulation was around 0.4%, RMSECV/RMSEP presented by the technique may not meet the quantitative demand (predicted deviation is less than 20% when concentration is less than 1.0%) of samples with low concentration (e.g. 0.1%). Therefore, further optimization of the models should be done for better modeling results.

3.2. CARS models

Variable selection is usually an effective approach to achieve better modeling results and CARS was employed to optimize the full-range spectrum models in this study. Fig. 2 illustrates the changing trends of the number of sampled wavelengths, five-fold cross validation values with the increasing of sampling runs from one CARS running. The two-phase selection, fast selection and refined selection are clearly shown (Fig. 2a). In the fast selection, the variables were forcedly eliminated by EDF and number of sampled variables was dropped quickly. Then in the refined selection, ARS was used to select the key variables based on the regression coefficients and number of sampled variables was decreased cautiously until the optimal subset obtained. As presented in Fig. 2b, the five-fold RMSECV value were almost the same

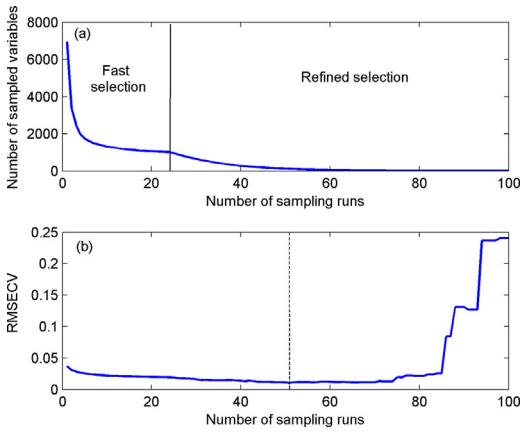

Fig. 2. The variable selection result of CARS: (a) the number of variables selected by the function of iterations; (b) the five-fold cross-validated errors at each MC sampling run.

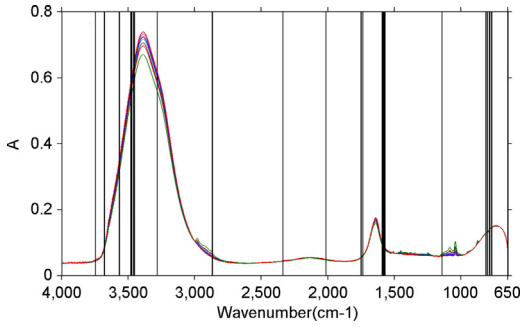

Fig. 3. The variables selected by CARS.

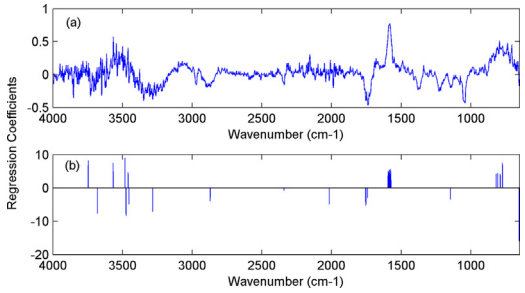

Fig. 4. Regression coefficients of the full-range and CARS models ((a) MIR–PLS model, (b) MIR–CARS–PLS model).

at first which indicates the existence of uninformative variables, then increased because of the loss of information caused by elimination of some key variables from the optimal subset (denoted by dash line).

The same as the full-range models, no spectral pretreatments were performed before modeling. The results of the CARS model are presented in Table 1 and the model was improved after CARS selection. The results indicated that CARS was an effective variable selection which can improve the PLS models effectively.

The variables selected by CARS are shown in Fig. 3. The selected variables scatted in the main absorption peaks of the spectrum: 2800–3700 cm^{-1} (stretch vibration of X—H), 1500–1700 cm^{-1} (stretch vibration of C═O and C═C) and 650–1000 cm^{-1} (flexural vibration of C—H, O—H and C—N). Fig. 4 shows the regression coefficients of the corresponding MIR model. As shown in the figure that the regression coefficients of the variables selected by CARS were all the ones with relatively large coefficients (e.g. informative)

in the regression. The above results demonstrated that the variables selected by CARS were effective and chemically meaningful.

4. Conclusions

In this study, we successfully apply MIR to determine the active ingredient in matrine aqueous solutions by conventional PLS and CARS–PLS. Results indicated that MIR is efficient in aqueous system. Since with ATR technique, the signal saturation caused by intense water absorption can effectively be settled, loss of informative variables in the range can be avoided. Thus, more information can be employed and better results were achieved. On the other hand, CARS is also proved to be an effective variable selection technique that makes the model simpler and more efficient.

Acknowledgements

This research is financially supported by National Natural Science Foundation of China (Grant No. 31301685) and Project supported by National Science Foundation for Fostering Talents in Basic Research of China (No. J1210064).

References

[1] Y. Zhu, C. Xu, J. Huang, G. Li, X. Liu, S. Sun, J.J. Wang, Rapid discrimination of cultivated *Codonopsis lanceolata* in different ages by FT-IR and 2DCOS-IR, J. Mol. Struct. 1069 (2014) 272–279.
[2] J. Duan, Y. Huang, Z. Li, B. Zheng, Q. Li, Y. Xiong, L. Wu, S. Min, Determination of 27 chemical constituents in Chinese southwest tobacco by FT-NIR spectroscopy, Ind. Crops Prod. 40 (2012) 21–26.
[3] R.M. Balabin, R.Z. Safieva, E.I. Lomakina, Gasoline classification using near infrared (NIR) spectroscopy data: comparison of multivariate techniques, Anal. Chim. Acta 671 (2010) 27–35.
[4] S.T.H. Sherazi, M. Ali, S.A. Mahesar, Application of Fourier-transform infrared (FT-IR) transmission spectroscopy for the estimation of roxithromycin in pharmaceutical formulations, Vib. Spectrosc. 55 (2011) 115–118.
[5] C.H. Howland, S.W. Hoag, Analysis of curing of a sustained release coating formulation by application of NIR spectroscopy to monitor changes physical–mechanical properties, Int. J. Pharm. 452 (2013) 82–91.
[6] J. Moros, S. Armenta, S. Garrigues, M. Guardia, Univariate near infrared methods for determination of pesticides in agrochemicals, Anal. Chim. Acta 579 (2006) 17–24.
[7] S. Armenta, S. Garrigues, M. Guardia, Partial least squares-near infrared determination of pesticides in commercial formulations, Vib. Spectrosc. 44 (2007) 273–278.
[8] Y. Xiong, G. Tang, J. Duan, C. Li, D. Wang, S. Min, Quantitative analysis of the content of deltamethrin in agrochemicals by near-infrared attenuated total reflectance infrared and Raman spectroscopy, Spectrosc. Spectral Anal. 30 (2010) 2936–2940.
[9] Y. Xiong, X. Song, C. Chen, H. Lin, G. Tang, Y. Huang, J. Duan, S. Min, The establishment and evaluation of near infrared universal model to determinate the effective ingredient content in pesticide rapidly, Chin. Chem. Lett. 10 (2012) 162–167.
[10] M. Khanmohammadi, M.A. Karimi, K. Ghasemi, M. Jabbari, A. Garmarudi, Quantitative determination of Malathion in pesticide by modified attenuated total reflectance-Fourier transform infrared spectrometry applying genetic algorithm wavelength selection method, Talanta 72 (2007) 620–625.
[11] M. Silverstein, G. Clayton Basseler, C. Morill, Spectrometric Identification of Organic Compounds, Wiley, New York, NY, 1981.
[12] H. Xu, Z. Liu, W. Cai, X. Shao, A wavelength selection method based on randomization test for near-infrared spectral analysis, Chemom. Intell. Lab. Syst. 97 (2009) 189–193.
[13] S. Ye, D. Wang, S. Min, Successive projections algorithm combined with uninformative variable elimination for spectral variable selection, Chemom. Intell. Lab. Syst. 91 (2008) 194–199.
[14] H. Li, Y. Liang, Q. Xu, D.S. Cao, Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration, Anal. Chim. Acta 648 (2009) 77–84.
[15] J. Yuan, Z. Zhang, B. Cong, Bioactivity and research progress of matrine, Pesticide 42 (2003) 1–4.
[16] J.J. Kelly, J.B. Gallis, Nondestructive analytical procedure for simultaneous estimation of the major classes of hydrocarbon constituents of finished gasolines, Anal. Chem. 62 (1990) 1444–1451.