



A Rapid, Non-Destructive Method for Screening Prochloraz-Containing Water Using Near-Infrared Spectroscopy

YAN ZHANG^{1,2}, BINGREN XIANG^{2,*} and JIANPING XU³

¹Department of Organic Chemistry, China Pharmaceutical University, Nanjing, Jiangsu Province, P.R. China

²Center for Instrumental Analysis, Key Laboratory of Drug Quality Control and Pharmacovigilance, Ministry of Education, China Pharmaceutical University, Nanjing, Jiangsu Province, P.R. China

³Zhongjian Pharmaceutical Co. Ltd. Zhongshan, Guangzhou, P.R. China

*Corresponding author: E-mail: cpuxiang@yahoo.com

Received: 25 November 2013;

Accepted: 11 February 2014;

Published online: 10 May 2014;

AJC-15186

A rapid and non-destructive method, based on Near Infrared Spectroscopy (NIRS) was established for screening prochloraz in water. A changeable size moving window partial least-squares method was used for variables selection. The wavelength ranged from 1274 nm to 2300 nm. The selected variables were applied as inputs to partial least-squares discriminant analysis (PLS-DA). Without sample pre-treatment, the whole analytical process only cost about 1 min for a sample, which proved that the analytical process was simple and fast. The best model gave satisfactory classification accuracy of 100 % and a good correlation coefficient of 0.95 between the measured and reference data. The limit of detection (LOD) of this method was as low as 1 mg kg⁻¹. To test and apply the proposed method, the procedure was applied to the analysis of prochloraz in tap water and lake water samples. Good identification results (accuracy of 100 %) were also received. The results of the study showed the great potential of Near Infrared Spectroscopy as a fast, non-destructive and environmentally-acceptable method for the on-site screening of prochloraz in water.

Keywords: Near-infrared spectroscopy, Prochloraz.

INTRODUCTION

Prochloraz (N-propyl-N-[2-(2,4,6-trichlorophenoxy)-ethyl]-1H-imidazole-1-carboxamide) (Fig. 1), as an imidazole broad spectrum fungicide, suppresses the biochemical and physiological metabolism of the target phytopathogenic fungi by inhibiting sterol biosynthesis¹. Due to its relatively low toxicity, prochloraz is widely used for protecting cereal crops, fruits, field legumes and beet² against eyespot fungus. However, recent studies reported that prochloraz has been shown to act *via* multiple mechanisms of action *in vitro* as it antagonizes the androgen and the estrogen receptor, agonizes the arylhydrocarbon receptor and inhibits aromatase activity^{3,4}. Its residual metabolites 2,4,6-trichlorophenol also has toxicity of mutagenicity and carcinogenicity⁵. So a series of analytical methods have been applied to detect the prochloraz residue such as high performance liquid chromatography and gas chromatography⁶. Although these methods show high sensitivity and accuracy, they have several drawbacks which can not neglected: Firstly, large amount of organic solvents is used which is environmentally unfriendly. Secondly, the sample preparation is always tedious and time-consuming. Therefore, they are not suitable for the rapid on-site detection.

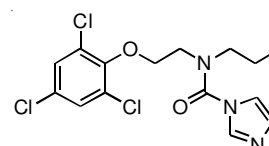


Fig. 1. Structure of prochloraz

Near Infrared Spectroscopy (NIRS) is a good alternative, because it is a fast, low cost, non-destructive method. No sample preparation is needed for NIRS, thus no organic solvents are involved, providing a safe working environment. NIRS has grown widely in use for qualitative and quantitative analysis. In recent years, quality control of pesticide products⁷, detection of pesticide residues in soil⁸ and agriculture products⁹ have been successfully performed by NIRS.

In this study, the feasibility of detecting prochloraz residue in water using NIRS was investigated. Since NIRS mainly contains information of vibrations of -CH, -OH, -NH and -SH bonds, which are the results of overtones or combinations of the fundamental mid-infrared bounds¹⁰, the use of chemometrics is necessary to extract relevant information from the overlapped bands of the NIRS. In this study, a changeable size moving window partial least-squares (CSMW-PLS)

method was used for variables selection. The selected variables were applied as inputs to partial least-squares discriminant analysis (PLS-DA). The proposed method was applied to the analysis of prochloraz in tap water and lake water samples.

EXPERIMENTAL

A multipurpose analyzer transform near infrared spectrometer YDZ1-1 (Nanjing Zhongdi Instruments Co., Ltd., Nanjing, China) was used in this study. Spectra treatment and data manipulation were performed using software MATLAB 7.5 (Math Works Inc. Natick, USA) with PLS Toolbox 5.8 (downloaded from http://www.eigenvector.com/software/pls_toolbox.htm).

Prochloraz (97 % purity) was purchased from Institute for the Control of Agrochemicals, Ministry of Agriculture (ICAMA). Bottled water (Guangdong Robust Co. Ltd. China) was obtained from a local supermarket. Tap water and lake water were collected from our laboratory and the Huashen Lake (Nanjing, China), respectively. All the samples were filtered through 0.2 μm filter twice prior to detection.

Samples preparation: Stock solution of prochloraz (1000 $\mu\text{g mL}^{-1}$) was prepared in double-distilled water. The working solution of prochloraz samples were prepared by diluting different amounts of the standard stock solution with water (double-distilled water, tap water and lake water, respectively) to the varying concentrations (1-200 $\mu\text{g mL}^{-1}$).

Spectral acquisition: A volume of 1.4 mL of the prepared sample was transferred to a glass vial, which was used as the measurement cell. Spectrum was recorded from the wavelength of 1100 nm (9090.909 cm^{-1}) to 2300 nm (4347.8261 cm^{-1}) at a nominal resolution of 2 cm^{-1} . The number of scans per spectrum is 2. Samples were analyzed in transflection mode, which was a hybrid of transmission and reflection. In this mode, a gold-coated reflector was placed behind the sample so that the optical path through the sample and back to the detector was doubled compared to a transmission measurement of a sample of the same thickness.

Data analysis procedure: In this study, the PLS-DA analysis method coupled with CSMW-PLS was used in the data analysis procedure. The specific steps were as follows: (1) the calibration set and prediction set was selected, (2) an appropriate method for data pre-processing was selected, (3) a correct variables region was selected, (4) the level of boundary was selected, (5) the data was imported, (6) analysis parameters was selected, (7) results was obtained.

Theory and algorithm

Partial least-squares discriminant analysis: PLS-DA is a classification method based on modelling the differences between several classes with PLS. It is essentially the inverse-least-squared approach to Linear Discriminate Analysis (LDA) and produces essentially the same result but with the noise reduction and variable selection advantages of PLS. By rotating principle components analysis (PCA), maximum separation among clusters can be achieved. In PLS-DA, PLS is used to develop a model that can predict the sample from different group. Each sample is assigned a dummy variable as a reference value, which is an arbitrary number or letter describing a particular sample group. For the present study, dummy

variables of 0 and 1 were given to the NIR spectra of prochloraz-free water and prochloraz-containing water. Thereafter, the PLS model was calculated in the usual way¹¹ and the classification was then done based on a Bayesian approach using the scores obtained from PLS. The model, of course, would not predict either a 1 or 0 accurately. A limit must be set, which is named threshold and is given by PLS-DA program. In the model used in this study, the threshold was set to 0.5 given by the PLS-DA program. Above the threshold, the sample was estimated as 1 and below the threshold, it was estimated as 0.

To validate the model, two methods were used *e.g.*, cross-validation and outer-validation. Leave-one-out is one important method of cross-validation. In the validation, all samples in calibration set except one were used to construct the calibration model and then the model was used to predict the remained one sample. The procedure was repeated, leaving out each of the samples from the calibration set in turn. The root-mean-square-error of cross validation (RMSECV) was calculated for every model. The number of PLSDA factors could also be selected according to the lowest RMSECV.

Upon building the model and testing it against a validation set for outer validation, the root mean square error of calibration (RMSEC) and the root mean square error of prediction (RMSEP) were calculated. The RMSEC and RMSEP measured how well the calibration data fitted the model and the expected error associated with future testing, respectively. In general, the classification accuracy was higher, when smaller RMSEC, RMSECV and RMSEP values were obtained.

Changeable size moving window partial least-squares: Studies reported NIR spectra always comprise substantial information derived from sample attributes, as well as environmental and instrumental variables¹². So, it is crucial that informative variables must be selected to construct a robust model, which can improve modelling efficiency and prediction accuracy¹³. Variable selection can also make the model simpler and obtain a better interpretation and lower measurement system costs¹⁴. So far, there are many effective methods for variable selection. Moving window wavelength selection is one of them, which is a strategy to obtain informative spectral regions for producing better prediction results¹⁵. Moving window partial least-squares (MW-PLS) was firstly introduced by Jiang *et al.*¹⁶. They demonstrated, based on a proposed theory providing the necessity of wavelength selection that MW-PLS proved to be a viable approach to eliminate the extra variability generated by non-composition related factors. In MW-PLS regression¹⁷, a spectral window commencing at certain (i)th spectral channel and terminating at the ($i + H - 1$)th spectral channel is built, in which H is the window size. PLS models with varied numbers of PLS components can then be built to relate the spectrum in the window to the analytes of interest. The window is moved through the entire spectrum. At each position, the RMSECVs are obtained from calculation with the PLS models and plotted as a function of the position of the window. A figure containing such residual lines is plotted and the plot provides information about informative regions. Based on MWPLS, changeable size moving window partial least squares (CSMWPLS)¹⁸ is a new method used to optimize an informative region, *i.e.*, to search for an optimized sub-region in a selected informative region. Its superiorities to

MWPLS are: The window size is changeable and the window is moved over the whole spectral region with fixed step. The specific process of the CSMW-PLS method is as follows:

A spectral window that starts at the $(i)^{\text{th}}$ spectral channel and ends at the $(i + w - 1)^{\text{th}}$ spectral channel is constructed, where w is the window size. The window size varies at an adjustable increment. The window is moved over the whole region with a step of 1. There are $(n - w + 1)$ windows over the whole spectra, with each window corresponding to a subset of the original spectral X . Fig. 2 explains this algorithm. PLS models can be built with varying LVs numbers from 1 to a fixed value k and then RMSECV is calculated for each subset. After calculations for all the subsets, the region with lowest RMSECV is chosen as the informative region.

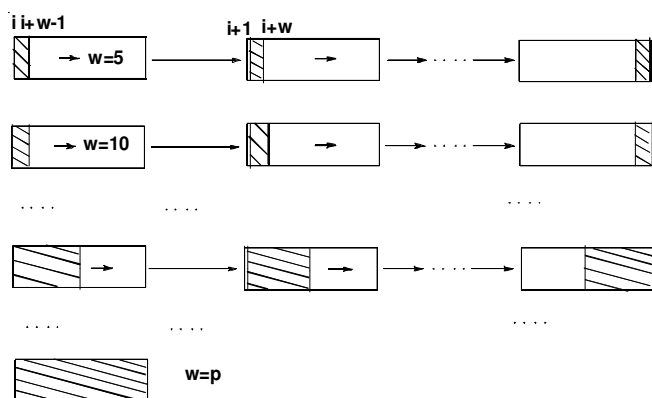


Fig. 2. Specific process of CSMW-PLS algorithm

In this study, the CSMW-PLS method was used. The window size varied from 20 to 600 with a step of 1. The max Latent Variable number was 8 and leave-one-out cross validation method was utilized. Meanwhile, CSMW-PLS procedure was coupled with PLS-DA to improve the performance of individual models and reduce the size of datasets in calibration and prediction processes.

RESULTS AND DISCUSSION

All 155 water samples were divided into two subsets randomly. The calibration set, which was used to build the model, had 104 samples including 26 prochloraz-free samples and 78 prochloraz-containing samples. The prediction set was formed by the remaining 51 samples including 12 prochloraz-free samples and 39 prochloraz-containing samples. Table-1 showed the descriptive statistics (mean, standard deviation and range) of samples used for calibration of the PLS-DA models and their prediction.

The mean spectra of water sample with prochloraz at various concentrations are shown in Fig. 3. As can be seen, the shapes of the mean spectra of all samples are nearly the same. It is difficult to find specific band to discriminate whether prochloraz was included in the aqueous samples since NIR bands are severely overlapped due to overtones and combination modes. Moreover, the existence of the interference of irrelevant compounds, such as solvent and impurities make the data analysis more complex. So before the calibration stage, the spectral data should be pre-processed to find out how the spectra vary with the concentrations of prochloraz. The methods

TABLE-1
CHEMICAL CHARACTERIZATION OF CALIBRATION AND PREDICTION SET FOR MODELLING

	Number of samples	Mean ($\mu\text{g mL}^{-1}$)	S.D. ^a ($\mu\text{g mL}^{-1}$)
Calibration set	104	35.7	49.1
Prediction set	51	36.4	50.1

^aStandard deviation

of autoscale, orthogonal signal correction, standard normal variate and first derivative were used for data pre-processing in this study. Fig. 4 showed the change of the RMSECV with latent variable number when using different data pre-processing methods. Due to the lowest RMSECV value, autoscale was selected as the data pre-processing method. Moreover, the number of PLS-DA factors could be also selected from Fig. 4. The cross-validation suggested that up to 9 principle components (Pcs) might be best. So 9 Pcs were chosen for the final model. After data-processing, the small spectral difference of water samples with prochloraz at various concentrations could be found, as shown in Figs. 5 and 6. From Fig. 6, absorption peak near 1340 nm shows the C-H stretch second overtones.

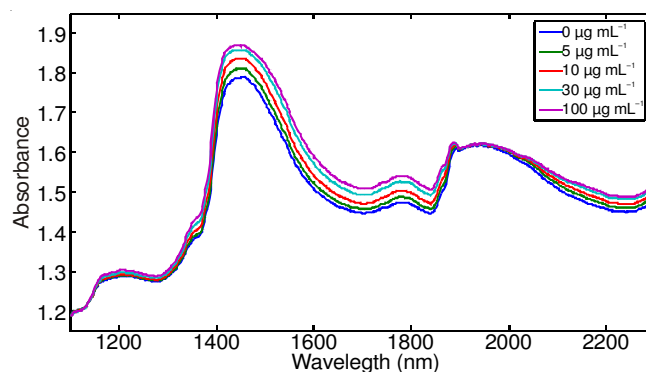


Fig. 3. Mean spectra of prochloraz at various concentrations (concentration unit: $\mu\text{g mL}^{-1}$)

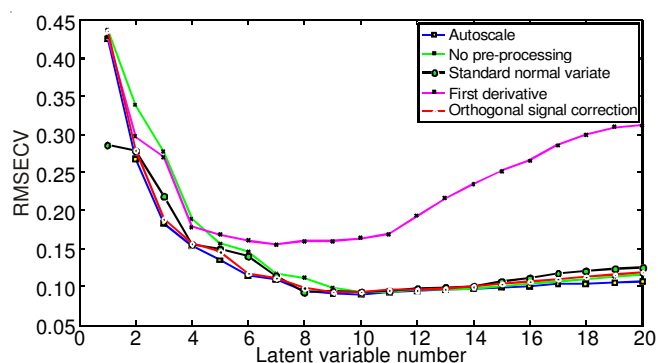


Fig. 4. Change of the RMSECV with the latent variable number when using different data pre-processing methods

CSMW-PLS algorithm was then used to find the correct variable region. The window size varied from 20 to 600 with a step of 1, the max Latent Variable number was 8 and cross validation method was leave-one-out. The wavelength region was 1274-2300 nm.

$1 \mu\text{g mL}^{-1}$ was set as the level of boundary. It means that the water samples with concentration below $1 \mu\text{g mL}^{-1}$ were assigned a dummy value 0 and others were assigned 1.

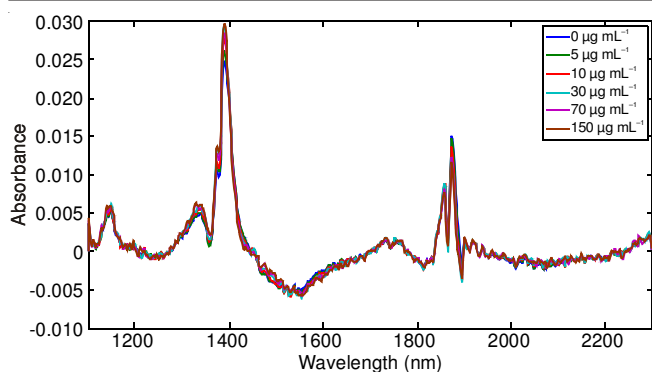


Fig. 5. First derivative spectra with prochloraz at different concentrations

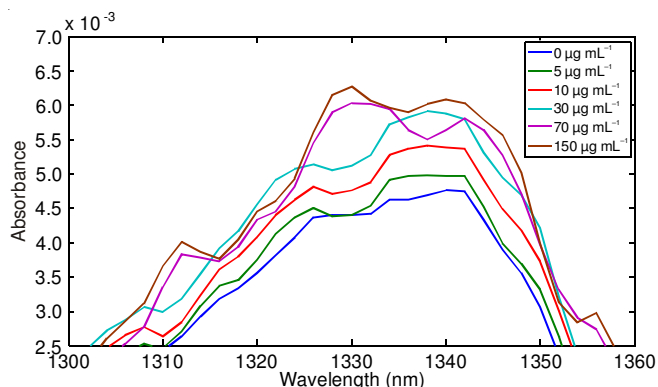


Fig. 6. Absorption peak of C-H stretch second overtones of prochloraz at different concentrations

Fig. 7 shows the classification scatter gram of PLS-DA. As can be seen intuitively: (1) the threshold given by PLS-DA was 0.5, which was represented by red dotted line, (2) black triangle points representing the prochloraz-free sample were located under the red dotted line, (3) red circle points representing the prochloraz-containing sample were located up the red dotted line, (4) the gap between the blank samples and the contaminated samples were very large. From the figure, a conclusion could be given that the classification accuracy was 100 %. The limit of detection (LOD) of this method was as low as 1 mg kg^{-1} . After further calculation, a good correlation coefficient of 0.95 between measured and reference data was obtained.

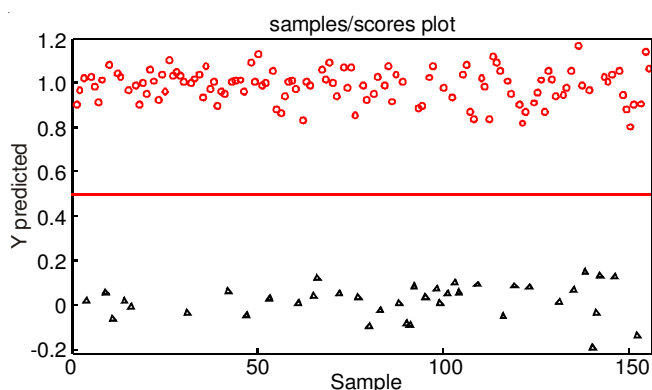


Fig. 7. Predicted results of water samples using PLS-DA, ("o", prochloraz-containing water sample; "△", prochloraz-free water sample)

Application of the technique: In order to testing and applying the proposed technique, the procedure was applied

to the analysis of prochloraz in real water samples (tap water and lake water samples). Each kind of water included 5 samples and all of them have been filtered through a $0.2 \mu\text{m}$ pore size Millipore filter prior to NIRS analysis. The prediction values of all samples were below 0.5, which indicated no prochloraz was found in the actual water. The results were verified by a HPLC method. And tap water and lake water samples containing prochloraz at a concentration of $5 \mu\text{g mL}^{-1}$ were also analysed. The prediction values of these samples exceed 0.5, which indicated prochloraz was found in these samples. It showed that the method was potential to accurately predict whether prochloraz was existed in environmental water samples.

Conclusion

In the study, the feasibility of detecting prochloraz residue in water based on NIRS combined with PLS-DA was investigated. A CSMW-PLS method was used for variable selection. Without sample preparation, the method gave satisfactory classification accuracy of 100 % and a good correlation coefficient of 0.95 between measured and reference data. The LOD of this method was as low as 1 mg kg^{-1} . The excellent result showed that the method was a rapid and highly effective potential alternative to conventional determination of prochloraz residues in water.

ACKNOWLEDGEMENTS

This work was financially supported by the CEEUSRO combination projects of Education Ministry of Guangdong Province [No. 2007A090302100]

REFERENCES

1. Y. Yu, X. Chu, G. Pang, Y. Xiang and H. Fang, *J. Environ. Sci. (China)*, **21**, 179 (2009).
2. M. De Paoli, M.T. Barbina, V. Damiano, D. Fabbro and R. Bruno, *J. Chromatogr. A*, **765**, 127 (1997).
3. H.R. Andersen, A.M. Vinggaard, T.H. Rasmussen, I.M. Gjermansen and E.C. Bonefeld-Jørgensen, *J. Toxicol. Appl. Pharmacol.*, **179**, 1 (2002).
4. J.I. Mason, B.R. Carr and B.A. Murry, *Steroids*, **50**, 179 (1987).
5. F.W. Wiese, H.C. Chang, R.V. Lloyd, J.P. Freeman and V.M. Samokyszyn, *J. Arch. Environ. Contam. Toxicol.*, **34**, 217 (1998).
6. M.T. Lafuente and J.L. Tadeo, *J. Environ. Anal. Chem.*, **22**, 99 (1985).
7. S. Armenta, S. Garrigues and M. de la Guardia, *Vib. Spectrosc.*, **44**, 273 (2007).
8. S. Bengtsson, T. Berglöf and H. Kylin, *Bull. Environ. Contam. Toxicol.*, **78**, 295 (2007).
9. D. Fen, T. Hong, K. Zhang and Y. Hong, *Intelligent Computation Technology and Automation. International Conference*, pp. 781-783 (2010).
10. H.W. Siesler, Y. Ozaki, S. Kawata and H.M. Heise, *Near Infrared Spectroscopy: Principles, Instruments, Applications*, Wiley-VCH, Weinheim, Germany (2002).
11. M. Barker and W. Rayens, *J. Chemometr.*, **17**, 166 (2003).
12. J.A. Guthrie, K.B. Walsh and J. Australian, *Exp. Agric.*, **39**, 73 (1999).
13. W. Cai, Y. Li and X. Shao, *Chemom. Intell. Lab. Syst.*, **90**, 188 (2008).
14. C.M. Andersen and R. Bro, *J. Chemometr.*, **24**, 728 (2010).
15. N. Sorol, E. Arancibia, S.A. Bortolato and A.C. Olivieri, *Chemom. Intell. Lab. Syst.*, **102**, 100 (2010).
16. J.H. Jiang, R.J. Berry, H.W. Siesler and Y. Ozaki, *J. Anal. Chem.*, **74**, 3555 (2002).
17. Z. Xiaobo, Z. Jiewen, M.J.W. Povey, M. Holmes and M. Hanpin, *Anal. Chim. Acta*, **667**, 14 (2010).
18. Y.P. Du, Y.Z. Liang, J.H. Jiang, R.J. Berry and Y. Ozaki, *Anal. Chim. Acta*, **501**, 183 (2004).