Analytical Methods

Accepted Manuscript

This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](http://www.rsc.org/Publishing/Journals/guidelines/AuthorGuidelines/JournalPolicy/accepted_manuscripts.asp).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](http://www.rsc.org/help/termsconditions.asp) and the Ethical quidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

www.rsc.org/methods

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx
 ARTICLE TYPE

Determination of the contents of magnesium and potassium in rapeseeds using FTIR-PAS combined with least squares support vector machine and uninformative variable elimination

Yuzhen Lu *^a* **, Changwen Du****^a* **, Changbing Yu***^b* **and Jianmin Zhou***^a*

⁵*Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX* **DOI: 10.1039/b000000x**

Fourier transform mid-infrared photoacousitc spectroscopy (FTIR-PAS) was employed to determine the contents of magnesium and potassium in rapeseeds. A total of 180 samples were collected for this purpose. A Savitzky-Golay filter was used for the spectral pretreatment. The whole sample set was ¹⁰divided into calibration and prediction sets composed of 135 and 45 samples, respectively. To build calibration models, partial least squares (PLS), least squares support vector machine (LS-SVM) and least

squares support vector machine combined with uninformative variable elimination (UVE-LS-SVM) were used. The best results for quantification of magnesium and potassium, were both achieved by UVE-LS-SVM models compared to the PLS models. The highest values of *RPD* (ratio of percentage deviation) ¹⁵were 2.5 and 2.25 for the prediction of magnesium and potassium, respectively. This work verified the good promise of FTIR-PAS combined with LS-SVM to quantify mineral nutrients of rapeseeds.

1. Introduction¹

<u>.</u>

Rapeseed (*Brassica napus* L.) is regarded among most important 20 oilseed crops^{1,2} since it supplies edible oil for human consumption or industrial applications, and feed protein for livestock. In addition, mineral components of rapeseeds are also important quality indicators deserving attention. Magnesium is an essential element for plant growth³ and also represents an $_{25}$ important nutrient in human and animal diet^{4, 5}. Potassium is a kind of indispensable nutrient for crops and potassium levels in seeds directly relate to the potassium fertilizer application. Fast and accurate determination of magnesium and potassium in rapeseeds carries significant implication for both quality ³⁰improvement in rapeseed breeding programs and fertilizer recommendation during field production.

Traditional methods to determine magnesium and potassium in rapeseeds usually include EDTA coordination titration for magnesium⁶ and flame photometry for potassium⁷. Besides, 35 inductively coupled plasma atomic emission spectroscopy (ICP-AES), is also often used for the determination of magnesium and potassium^{8,9}. However, all of these methods necessitate the sample pretreatment, that is, seed grinding, concentrated acid digestion or dry ashing. The process is time-consuming, ⁴⁰laborious, expensive and even dangerous. Also, these methods

are destructive. One measurement causes the loss of several grams of rapeseeds. Therefore, these methods fail to satisfy the need of fast characterization of rapeseed nutrient information, and are particularly unsuitable for variety selection of rapeseeds with 45 small sample size.

Analytical Methods Accepted Manuscript Analytical Methods Accepted Manuscript

Near/mid infrared spectroscopy including reflectance and transmission modes has been widely applied as a powerful alternative for qualitative and quantitative analysis in agriculture and food industry, due to its rapidity, favorable economics, 50 simplicity of sample pretreatment and absence of chemicals¹⁰⁻¹³. Recently, another unique sampling technique called Fourier transform infrared photoacoustic spectroscopy (FTIR-PAS) is gaining momentum¹⁴. FTIR-PAS requires no sample pretreatment, just tiny amounts of samples (approximately 100 mg for one ⁵⁵scanning), and can analyze nearly any type of sample in a rapid way. To date, PAS technique has found success as a potent tool to quantify property parameters of analytes including drugs¹⁵, woods¹⁶, pulps^{17, 18}, soils^{19, 20}. Thorough elaboration of the principles and applications of PAS technique is reviewed by 60 Haisch²¹. At present, however, no reports of applications of FTIR-PAS to analysis of rapeseed components are found.

Most of published work shows that infrared spectroscopy can successfully quantify organic components in agricultural products. However, the ability of this technology to determine mineral ⁶⁵contents has always been under dispute because most minerals don't produce infrared absorption. Nevertheless, many studies²²⁻²⁵ have shown that infrared spectroscopy can be used to determine some minerals including potassium and magnesium. It is explainable in that those minerals in plant materials form 70 complexes with certain organic molecules, and it is by measuring these complexes that minerals are indirectly measured by infrared spectroscopy. However, these complexes might not to be static and varied among plant species. Further study is needed to investigate the applicability of infrared spectroscopy for ⁷⁵quantification of minerals in rapeseeds.

^aNational Key Lab. of Soil and Sustainable Agriculture, Nanjing Institute of Soil Science, Chinese Academy of Sciences, Nanjing 21008, P.R.China. E-mail:chwdu@issas.ac.cn; Fax: +86-025-86881000; Tel: +86-025- 86881565

^{*b*} Key Laboratory of Biology and Genetic Improvement of Oil Crops, Oil *Crops Research Institute of the Chinese Academy of Agricultural Sciences, Ministry of Agriculture, Wuhan 430062, P.R.China*

The major objective of our study was to evaluate the potential of FTIR-PAS to determine magnesium and potassium in rapeseeds. To build calibration equations, we employed three calibration methods: partial least squares (PLS), least squares ⁵support vector machines (LS-SVM) and least squares support vector machines combined with uninformative variable elimination (UVE-LS-SVM).

2. Brief descriptions of chemometric methods

2.1.PLS

10 Partial least squares (PLS) is most widely used as a multivariate calibration method owing to its capacity of utilizing spectral information of the component of interest to ensure a robust result²⁶. PLS latent variables are extracted by taking into consideration sample spectral matrix and concentration matrix 15 simultaneously. By doing this, the variation of the spectral matrix, and the correlation between latent variables and concentration matrix, are both maximized. In PLS modeling, the number of latent variables is decided based on minimal root-mean-square error of cross validation (*RMSECV*). However, classic PLS just ²⁰handle linear regression problems. The detailed information about PLS algorithm is well elaborated by Gelahi and Kowalski²⁷.

2.2.LS-SVM

Recently, least squares support vector machine (LS-SVM) 28-29 ²⁵was introduced into the chemometric community to deal with linear and nonlinear problems. It employs a set of linear equations rather than the quadratic programming used in the original support vector machine $(SVM)^{30}$ to obtain support vectors, simplifying computational procedures of SVM. The 30 powerfulness of LS-SVM in infrared spectral analysis has been verified by many studies³¹⁻³⁴ .

In LS-SVM modeling, the selection of kernel function and its corresponding kernel parameters plays a vital role. Radical basis function (RBF) is extensively used for nonlinear problems, and ³⁵can reduce the computational complexity of the training procedures and obtain a good result under general smoothness assumptions²⁹. For RBF, two parameters, γ and σ^2 need to be a priori tuned. γ is a regulation constant which affects the generalization performance of LS-SVM models. σ^2 is the width ⁴⁰of the RBF which influences the regression errors and the number of initial eigenvectors.

2.3.UVE

- To improve and simplify calibration models, spectral variable ⁴⁵selection is often performed. Uninformative variable elimination (UVE) has been verified as a powerful tool to remove uninformative variables and improve calibration models^{35, 36}. For UVE, variables which don't contain more information than random variables are defined as uninformative variables and then ⁵⁰removed. This removal is achieved based on the comparison between reliability values of spectral variables and random variables. The main principle of UVE algorithm can be synopsized as follows.
- First, spectral variables are artificially added with the same size ⁵⁵of random variables. Then, every variable can derive a reliability value through certain calculation. The maximal absolute value of reliabilities of random variables is set as the cut-off value for variable selection. Finally, those spectral variables whose absolute reliability values are lower than the cut-off value are ⁶⁰deemed informative. UVE-α is a variant of original UVE. In

UVE- α , the cut-off value corresponds to α , e.g., 99%, the quantile of ranked reliability values of artificial variables. This modification makes original UVE more conservative and can reduce the risk of rejecting some informative variables.

⁶⁵**3. Experimental**

3.1.Rapeseed samples

A total of 180 rapeseed samples were provided by Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences. They were harvested from a field experiment in the ⁷⁰Yingtan Ecology Experimental Station. Each sample represented one variety of rapeseeds. Before spectral scanning, all samples were air-dried to make moisture content equivalent among samples, and then stored in plastic bags at room temperature.

⁷⁵3.2.**Spectral measurements**

All rapeseed samples were first subjected to spectral scanning and then to chemical determination procedures. Photoacoustic spectra were recorded for all samples using a Fourier transform infrared spectrometer (Nicolet 6700, USA) equipped with a 80 photoacoustic cell (Model 300, MTEC, USA). After placing the sample (about 100 mg) in the cell holding cup (diameter 5 mm, height 3 mm) and purging the cell with dry helium (10 ml min^{-1}) for 10 s to ensure a $CO₂$ and $H₂O$ free environment, the scans were conducted in the mid-infrared wavenumber range of 500- 854000 cm⁻¹ with a resolution of 4 cm⁻¹ and a mirror velocity of 0.32 cm/s. 32 successive scans were recorded, and the average for each sample was used in the subsequent analysis. In order to eliminate spectral features due to the infrared source, optics and PAS detector response, the obtained PAS spectra were ⁹⁰normalized by dividing the sample spectrum by the spectrum of a carbon black reference sample under the same experimental conditions.

3.3.Reference methods

⁹⁵Individual rapeseeds were weighted on a lab balance (Sartorius BS210S, Germany) with a readability of 0.0001 g, and then were digested with concentrated sulfuric acid and hydrogen peroxide in the digestion instrument (LabTech EHD36, USA). Individual sample of digested solution was properly diluted before the 100 determination with ICP-AES (Thermo element IRIS Advantage, USA). The analytical lines of magnesium and potassium were 285.2 nm and 766.4 nm, respectively.

3.4.**Data analysis and software**

- ¹⁰⁵The normalized spectra were pre-processed using a smoothing filter of Savitzky-Golay³⁷ with 21-point window and a polynomial of order 1. The second derivative of Savitzky-Golay with the same parameters was used to assist in the spectral band assignment. The smoothed spectral matrix was divided into ¹¹⁰calibration and prediction subsets of 135 and 45 samples, respectively, according to a recent sample partitioning method called SPXY³⁸. SPXY is an extended version of Kennard-Stone algorithm³⁹ by encompassing both X and Y differences in the calculation of inter-sample distances.
- 115 Calibration models were subsequently established using PLS, LS-SVM and UVE-LS-SVM algorithms. In our study, the number of latent variables of PLS model was decided based on minimal *RMSECV* through a leave-one-out cross validation procedure. In LS-LVM modeling, RBF was selected as the kernel ¹²⁰function. Its two kernel parameters were optimized using a two

dimensional grid search method. The optimal combination was decided based on the minimum *RMSECV* through a leave-fiveout cross validation. For UVE algorithm, UVE-99% was employed. Spectral preprocessing was implemented by The Unscrambler

⁵v 9.8 (CAMO Software AS, Norway). Dataset partition and modeling procedures were realized in Matlab R2011b (The Math Works, Natick, USA). Two Matlab toolboxes, PLS_Toolbox 4.2 (Eigenvector Technology, USA), and LS-SVM toolbox 1.8 (LS-SVMlab, Leuven, Belgium) were utilized for modeling.

3.5.**Model evaluation standard**

For calibration set data, calibration models were compared using the root-mean-square error of calibration (*RMSEC*) and the 15 coefficient of determination r_{cal}^2 in the process of five-fold cross validation; For validation set data, the root-mean-square error of prediction (*RMSEP*) and r_{pre}^2 were used to evaluate the performance of calibration in the prediction process⁴⁰.

Besides, the ratio of percentage deviation (*RPD*) defined by Williams in 1987⁴¹ ²⁰as the ratio of standard deviation (*SD*) of

Table 1 Statistics of magnesium and potassium in rapeseeds

within the range of reference values in the calibration set. Properties Sample set Minimum Maximum Mean Standard deviation Calibration 0.2014 0.2576 0.2287 0.0106

Prediction 0.2116 0.2450 0.2270 0.0080

Calibration 0.4814 0.7418 0.6117 0.0583 Prediction 0.5089 0.6828 0.6150 0.0510

The Fig. 1a gives the average of smoothing-processed spectra of all rapeseeds, and the corresponding second derivative spectrum is shown in Fig. 1b. Well-resolved peaks could be observed from ⁴⁵the smoothed spectrum, and roughly related to the valleys in the second derivative spectrum where the baseline effect was

eliminated, and some absent peaks become evident. Major spectral features of rapeseeds could be interpreted as follows⁴³. The peaks at 3740 cm^{-1} and 3620 cm^{-1} were related with the

50 free O-H stretching. A broad peak at 3365 cm⁻¹ corresponded to O-H stretching and N-H stretching⁴⁴. The peaks at 2925 cm⁻¹ and at 2850cm-1were assigned to asymmetric and symmetric C-H stretching, respectively. A small peak at 2350 cm^{-1} corresponded

to the asymmetric stretching of C=O bonds of the residual carbon ⁵⁵ dioxide⁴⁵ in the photoacoustic cell. The peak at 1635 cm⁻¹ was associated with O-H deformation vibration overlapping with the amide I C=O stretching. A shoulder peak around 1555 cm⁻¹ corresponded to the amide II (the out-of-phase combination of the N-H deformation vibration and C-N stretching vibration)⁴⁶. The 60 peak around 1435-1380 cm⁻¹ corresponded to C-H deformation vibration. In the fingerprint range, $1200-500$ cm⁻¹, a broad band around 1060 cm^{-1} was mainly assigned to the C-O and C-C stretching. Besides, the peak around 1740 cm^{-1} corresponded to ester C=O stretching.

reference values in prediction set to the root-mean-square error of prediction (*RMSEP*), was used as an integrated criterion for evaluating the predictive ability of calibration models. According to Williams, *RPD* should be at least 3 for quality control of 25 agricultural products. As suggested by Viscarra Rossel⁴², calibration models will suffice for good quantitative prediction if *RPD* was larger than 2, and the lowest line of *RPD* for quantitative application was 1.8. In general, *RPD* should be as

The statistics including mean, range and standard deviation, of reference values of magnesium and potassium contents in sample sets are given in Table 1. A larger variability of reference values ³⁵was observed in the potassium. For the two nutrient parameters, the range of references values in the prediction set was included

high as possible for a good model.

³⁰**4. Results and discussion 4.1.Reference values**

Fig. 1 The average spectrum of smoothed photoacoustic spectra of all rapeseed samples and the corresponding second derivative

65

15

spectrum.

4.3.Calibration and prediction

⁵Table 2 shows the results of the calibration and prediction, obtained by PLS models, LS-SVM models and UVE-LS-SVM models. For all models, no pronounced differences between r_{cal}^2

and r_{pre}^2 and between *RMSEC* and *RMSEP*, were observed, which indicated the robustness of all models. According to *RPD* values, 10 the best results for the estimation of magnesium and potassium were achieved both by UVE-LS-SVM models, followed by the LS-SVM models and PLS models.

For the PLS models, the *RPD* for potassium was 1.92, which indicated an acceptable quantitative model, while the model for magnesium needed to be improved due to the *RPD* lower than 1.8. ²⁰Our result for potassium was close to the results reported by Clark et al.⁴⁷ in forage samples ($r_{pre}^2 = 0.74$ -0.81), and superior to the result report by Huang⁴⁸, in the cut straw sample $(RPD =$ 1.66). Nevertheless, for the estimation of magnesium, Clark and Huang obtained better results, with r_{pre}^2 of 0.71-0.88 in forage ²⁵samples and *RPD* of 2.11 in cut straw samples, respectively. Our poor *RPD* for magnesium was partly due to the relatively small standard deviation of reference values. For improved accuracy, more rapeseed samples should be included to increase the variation of magnesium content. The results of PLS models 30 preliminarily proved the applicability of FTIR-PAS as a fast screening tool to measure the contents of magnesium and potassium in rapeseeds.

Since PLS models failed to utilize some possible nonlinear

relation between minerals and infrared spectra of rapeseeds, LS-³⁵SVM models using the kernel function of RBF were established. The full spectrum was directly taken as the input for LS-SVM

modeling. Through the optimization of parameters, γ and σ^2 , and for magnesium and potassium were 7.12 and 13.36, 0.58 and 9.37, respectively. The resulting *RPD* values for magnesium and ⁴⁰potassium were 2.22 and 2.13, respectively. The *RPD* values were much higher than those obtained from the PLS models, and both higher than 2.0, indicating good models for quantitative prediction. The LS-SVM models had a better capacity than PLS models to quantify the contents of magnesium and potassium in ⁴⁵rapeseed samples, which implied the hypothetical relationship between spectral information and mineral components of rapeseeds was probably nonlinear, and the LS-SVM models could well exploit this nonlinear relationship for improved prediction results.

60

15

Fig. 3 Visual positions of spectral variables selected by UVE in the full spectrum for the determination of the contents of magnesium (a) and potassium (b) in rapeseeds.

Full-spectrum modeling seemed cumbersome, and even detrimental to the prediction accuracy since some noisy or uninformative variables might be incorporated into final models. Thus, UVE was adopted to refine calibration models. ¹⁰Fig. 2a and Fig. 2b show the plots of reliability value of each variable for the determination of magnesium and potassium, respectively. In the reality plot, spectral variables were at the left of the vertical line while random variables at the right. The two horizontal lines were the cut-off lines. The spectral 15 variable whose reliability was within the down and up cut-off lines was viewed as noisy or irrelevant, and to be removed. As a result, 121 and 76 variables were retained for the determination of magnesium and potassium, respectively. The visual positions of selected variables in in the full spectrum are ²⁰shown in Fig. 3a and Fig. 3b, where columns corresponded to

the selected spectral variables. Those variables might indicate the chemical complexes associated with mineral nutrients of magnesium and potassium.

Selected variables were then employed as the input of LS-25 SVM models. The optimal parameters combination γ and σ^2 were 10.26 and 13.96 for magnesium, and 1.97 and 6.39 for potassium. A further improvement of models was obtained with *RPD* increasing to 2.5 for the prediction of magnesium, and *RPD* increasing to 2.25 for the prediction of potassium. The 30 best prediction results for magnesium and potassium were both achieved with the parsimonious models of UVE-LS-SVM. This good performance of the use of UVE to improve LS-SVM models was in good agreement with the results reported by $Wu⁴⁹$ and Huang⁵⁰. Fig.4 shows the scatter plots of prediction 35 results of the optimal models for the two properties.

Fig. 4 Scatter plots of prediction results obtained by the UVE-LS-SVM models for the contents of magnesium (a) and potassium (b). The ⁴⁰solid line was the reference line corresponding to the exact prediction, and samples were distributed along the reference line.

65

5. Conclusion

Quantifying magnesium content in rapeseeds by PLS model produced unsatisfactory results, while the results obtained from ⁴⁵PLS model for the prediction of potassium content was acceptable. Further, improved prediction results were achieved by developing LS-SVM and UVE-LS-SVM models. LS-SVM could well take advantage of the nonlinearity existent between spectral variables and mineral nutrients of rapeseeds, and the ⁵⁰adoption of UVE made LS-SVM models simplified and slightly better. The highest values of *RPD* for the prediction of the contents of magnesium and potassium in rapeseeds were 2.5 and 2.25, respectively.

Although the *RPD* values didn't reach 3 proposed by

⁵⁵Williams for quality control, FTIR-PAS had been verified as a rapid and non-destructive tool for rapeseed scanning. Moreover, the application of FTIR-PAS to determine the contents of magnesium and potassium in rapeseeds will depend on the accuracy required by work itself. Further work will need to be ⁶⁰done by including more rapeseed samples from different agronomic conditions and plant years in order to improve the prediction accuracy of calibration models.

Acknowledgements

This work was supported by the Innovative Project for Young Scientist from Chinese Academy of Sciences (KZCX2-YW-QN411) and the National Natural Scientific Foundation of

China (40871113). We are genuinely grateful to four anonymous reviewers for valuable comments on an earlier manuscript.

References

- M. Cardone, M. Mazzoncini, S. Menini. V. Rocco, A. Senatore, M. Seggiani, S. Vitolo, *Biomass Bioenerg*., 2003, **25**, 623-636.
- Y.S. Jang, *Korean. J. Crop. Sci*., 2002, **47**, 175-185.
- 3 J. Gerendas. H. Fuhrs. *Plant Soil*, 2013. doi: 10.1007/s11104-012-1555-2.
	- E.J.M Joy, S.D. Young, C.R. Black, E.L. Ander, M.J. Watts, M.R. Broadley, *Plant Soil*, 2013, **368**, 129-137.
	- P.J White, M.R. Broadley, *New Phyto.*, 2009, **182**, 49-84.
- 6 S.D. Bao, Soil Agricultural Chemistry Analysis, China Agriculture Press, Beijing, 2000, pp. 271-275.
	- C. L. Fox, *Anal. Chem.*, 1951, **23**, 137-142.
	- N.R. McQuaker, D.F. Brown, P.D. Kluckner, *Anal. Chem*., 1979, , 1082-1084.
	- 9 M. Accto, O. Abollino, M.C. Bruzzoniti, E. Mentasti, C. Sarzanini, M. Malandrino, *Food Addit. Contam*., 2002, **19**, 126- 133.
- P. Williams, K. Norris. Near infrared technology in the agriculture and food industries, American Association of Cereal Chemists, Inc., Minnesota, 2001.
	- D. Sorak, L. Herberholz, S. Iwascek, S. Altinpinar, F. Pfeifer, H.W. Siesler, *Appl. Spectrosc. Rev*., 2012, **47**, 83-115.
	- C. Shiroma, L. Rodriguez-Saona, *J. Food Comp. Anal*., 2009, **22**, 596-605.
- 13 B.E. Madari, J.B. Reeves III, P. Machado, C. M. Guimaraes, E. Torres, G.M. McCarty, *Geoderma*, 2006, **136**, 245-259.
- J.F. McClelland, R.W. Jones, S.J. Bajic, FT-IR Photoacoustic Spectroscopy, in *Handbook of Vibrational Spectroscopy*, J.M. Chalmers, P.R. Griffiths (Eds.), John Wiley & Sons, Ltd., Chichester, 2002, pp.1231-1250.
	- R. Neubert, B. Collin, S. Wartewig, *Pharma. Res.,* 1997, **14**, 946-948.
	- S. Bjarnestad, O, Dahlman, *Anal. Chem*., 2002, **74**, 5851-5858.
	- K. Nishi, V.Q. Dang, K.L. Nguyen, *Anal. Chem*., 2006, **78**, 6818-6825.
	- V.Q. Dang, N.K. Bhardwaj, V. Hoang, K.L. Nguyen, *Carbohydr. Polym*., 2007, **68**, 489-494.
	- C.W. Du, J.M. Zhou, H.Y. Wang, *Vibr. Spectrosc*., 2009, **49**, 32- 37.
- 20 C.W. Du, J.M. Zhou, *Appl. Spectrosc. Rev*., 2011, **46**, 405-422.
	- C. Haisch, *Meas. Sci. Technol*., 2012, **23**, 012001.
	- G.D. Batten, *Aust. J. Exp. Agric*., 1998, **38**, 697-706.
	- A. Moron, D. Cozzolino, *J. Agric Sci*., 2002, **139**, 413-423.
	- D. Cozzolino, A. Moron, *Anim. Feed. Sci. Tech*., 2004, 161-173.
- 25 C. Petisco, B. Garcia-Criado, B.R. Vazquez de Aldana, I.

Zabalgogeazcoa, S. Mediavilla, A. Garcia-Ciudad, *Anal. Bioanal. Chem*., 2005, **382**, 458-465.

- J.J.B. Nevado, J.R. Flores, G.C. Penalvo, *Anal. Chim. Acta*, 1997, , 257-265.
- 27 P. Geladi, B.R. Kowalski, *Anal. Chim. Acta*, 1986, **185**, 1-17.
	- J.A.K. Sukens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least squares support vector machines. Singapore: World Scientific Publishing, 2002.
- R.P. Cogdill, P. Dardenne, *J. Near Infrared Spectrosc.*, 2004, **12**, 60 93-100.
	- C. Cortes, V. Vapnik, Machine Learning, 1995, **20**, 273-297.
	- F. Chauchard, R. Cogdill, S. Roussel, J.M. Roger, V. Bellon-Maurel, *Chemom. Intell. Lab. Syst*., 2004, **71**, 141-150.
	- U. Thissen, B. Ustun, W.J. Melssen, L.M.C. Buydens, *Anal. Chem*. 2004, **76**, 3099-3105.
	- A. Borin, M.F. Ferrao, C. Mello, D.A. Maretto, R. J. Poppi, *Anal. Chim. Acta*, 2006, **579**, 25-32.
	- J.Y. Shi, X.B. Zou, X.W. Huang, J.W. Zhao, Y.X. Li, L.M. Hao, *Food Chem*., 2013, **138**, 192-199.
- 35 V. Center, D.L. Massart, O.E. de Noord, S. de Jong, B.M. Vandeginste, C. Sterna, *Anal. Chem*. 1996, **68**, 3851-3858.
	- W.S. Cai, Y.K. Li, X.G. Shao, *Chemom. Intell. Lab. Syst*., 2008, , 188-194.
	- A. Savitzky, M.J.E. Golay, *Anal. Chem*., 1964, **36**, 1627-1632.
- 38 R.K.H. Galvao, M.C.U. Araujo, G.E. Jose, M.J.C. Pontes, E.C. Silva, T.C.B. Saldanha, *Talanta*, 2005, **67**, 736-740.
	- R.W. Kennard, L.A. Stone, *Technometrics*, 1969, **11**, 137-148.
	- D. Wu, Y. He, P.C. Nie, F. Cao, Y.D. Bao, *Anal. Chim. Acta*, 2010, **659**, 229-237.
- 41 P. Williams, Variables affecting near-infrared reflectance spectroscopic analysis, in *Near-Infrared Technology in the Agriculture and Food Industries*, P. Williams, K. Norris (Eds), American Association of Cereal Chemists, St. Paul, MN, 1987, pp. 143-166.
- 42 R.A. Viscarra Rossel, R.N. McGlynn, A.B. McBratney, *Geoderma*, 2006, **137**, 70-82.
	- K.N. Fan, Introduction to spectroscopy, High Education Press, Beijing, 2001, pp. 63-73.
	- H. Yang, J. Irudayaraj, *J. Am. Oil Chem. Soc*., 2000, **77**, 291-295
- 45 M.P. Bernstein, *Icarus*, 2005, **179**, 527-534.
	- A. Barth, *Biochim. Biophys. Acta*., 2007, **1767**, 1073-1101
	- D.H. Clark, H.F. Mayland, R.C. Lamb, *Agron. J*., 1987, **79**, 485- 490.
- C.J. Huang, L.J. Han, Z.L. Yang, X. Liu, *Fuel*, 2009, **88**, 163-
	- D. Wu, Y. He, J.H. Shi, S.J. Feng, *J. Agric. Food. Chem*., 2009, , 1697-1704.
	- Z.G. Huang, S. Sha, Z.Q. Rong, J.H. Chen, Q.L. He, D.M. Khan, *Ind. Crop. Prod*., 2013, **43**, 654-660.