An efficient method to reduce grain angle influence on NIR spectra for predicting extractives content from heartwood stem cores of *T. sinensis*

Yanjie Li
Research Institute of Subtropical Forestry Chinese Academy of Forestry https://orcid.org/0000-0001-7750-757X
Xin Dong
Research Institute of Subtropical Forestry Chinese Academy of Forestry
Yang Sun
Research Institute of Subtropical Forestry Chinese Academy of Forestry
Jun Liu (ucjackley@gmail.com)
Research Institute of Subtropical Forestry Chinese Academy of Forestry
Jingmin Jiang
Research Institute of Subtropical Forestry Chinese Academy of Forestry

Research

Keywords: near infrared spectroscopy, heartwood, grain angle, T. sinensis

Posted Date: December 12th, 2019

DOI: https://doi.org/10.21203/rs.2.18855/v1

Version of Record: A version of this preprint was published at Plant Methods on June 1st, 2020. See the published version at https://doi.org/10.1186/s13007-020-00623-3.

Abstract

Background

A fast and reliable non-destructive method for qualifying the content of extracts content (EC) in heartwood of *T. sinensis* cores is needed in the breeding program for studying the genetically infect on EC. However, the affecting of grain angle on near infrared (NIR) spectra prediction model for EC is unclear. In this study, NIR spectra were collected both from cross section and radial section of wood core samples in order to predict the EC in heartwood.

Results

The effect of grain angle on calibration EC model was studied. Several different spectra pre-processing methods were tested for calibration. It was found that standard normal variation (SNV) followed by 1st derivative yield the best calibration for EC. Grain angle has a significant influence on the predict model for EC when use the whole NIR spectra. However, after using the significant multivariate correlation (sMC) selection of the prior of wavenumbers for EC, the influence of grain angle have been significantly reduced.

Conclusions

It was suggested that NIR spectroscopy could be a promising methods to predict EC in solid wood without the infection of grain angle.

Background

Tree could produce varies types of wood timber for industry and society use, among these, one of the most valuable and popular type of wood, especially in China, is the natural durable wood with splendid colours. Wood which has natural durability could protect the wood against to biological decay [1, 2]. On the other hand, different colours wood, for instance, yellow or red, could potentially increase the value of the wood for end use. Therefore, natural durability wood with colour have been widely studied. In tree stem, the inner part of wood is call heartwood, and the outside of wood is recognized as sapwood, with the growth of tree, cells in the inner part of sapwood began to die and accumulated massive secondary metabolites, this process are transfer the sapwood into heartwood with natural durability and colourful [3]. However, the traditional methods to test the wood natural durability are costly and time-consuming. It has been reported that the extractives in the heartwood plays an important role for colourful and natural durability [4]. Decomposition and colourlessness will appear when the extractives are removed from durable wood [5]. Therefore, the amount of extractive content (EC) in heartwood has been studied as a proxy for natural durability [6].

The EC in heartwood has large variation and could be reduced by the genetic selection [7]. There are different ways to determine the quantity of extractives in heartwood. The traditional methods such as Soxhlet or accelerated solvent extraction (ASE) [8] which are time- and cost- consuming could not suitable for tree breeding and selection programs, which rely on the measurement of large numbers of samples. A high throughput and rapid measurement method for EC is needed.

Near infrared spectroscopy (NIR) is a non-destructive technique that is used for the analysis of agricultural products [9-11]. It has been used to determine the quantity of heartwood extractives in some tree species [12-14] and yields a promising and reliable results. The NIR spectra that collected from different sample form, wood powder or solid, could influence the performance of wood traits prediction model.

The models that build on different size of milled wood powder are different and performed higher accuracy than the model based on the solid wood when predict the chemical properties in eucalyptus wood by NIR spectroscopy [15]. However, wood sample grinding also a time consuming step and could reduce the extractive content prediction time, directly EC prediction from the solid wood samples could a suitable alterative way for NIR model calibration.

NIR spectra that take from the solid wood samples could influenced by many factors, such as moisture content [16] and grain angle [17]. It has been reported that the grain angle could influence the EC prediction of *Eucalyptus bosistoana* using NIR spectra and this influence could be reduced by using of EPO algorithm [6]. Alternatively, this influence may could reduced by using of feature selection methods. One of the most important feature selection method that has been reported that could remove confounding effects on NIR calibrations [18-20] is call significant Multivariate Correlation (sMC) [21]. Feature selection combined with PLS method could found out the most important variables in NIR spectra when predicted the target chemicals content in plant and remove the irreverent variables that influenced the accuracy of the model prediction [22]. However, The important features in the NIR spectra for the grain angles and EC are little known.

Toona *sinensis* is a native Chinese tree species that have been widely distribution in China, it has a long history of cultivation for its digestible buds in China. In addition, T. *sinensis* also hold the advantages of fast growth and bright red heartwood that has been widely used for furniture and industry use. It has been widely studied on many regions, including the cultivation, reproduction, and biological activity of *T. sinensis* and the physical characteristics and chemical of its wood. However, little is known about the nature duable and red colour timber. To establish an high quality breeding program for durable and red wood selection, it is need to find an alternative way to allow fast and efficient measure the heartwood quality of *T. sinensis*. If NIR can be successfully used to analyse the heartwood properties without any grain angle influence, the traditional methods which are time- and cost- consuming can be replaced and consequently, it will give a lot of benefit for selection.

Hence, This study will focus on the effect of grain angle on NIR spectra obtained from *T. sinensis* cores, and study the possibility of applying NIR as a rapid and precise method to predict the extractives content from the solid core samples of *T. sinensis* without grain angle.

Results

PLS models with full length of NIR spectra

Table 1 shows PLS regression models for calibration and validation from the spectra with and without pre-processing constructed with EC and grain angle. Calibration model have different accuracy between EC and grain angle. Regardless of the pre-processing methods, wood core NIR spectra combined with PLS model could yield a good result to discriminate wood longitudinal growth direction (0º) and cross section(90º), PLS regression model for EC shows lower R^2 result compare to grain angle. The number of component of models appeared to need more in grain angle than in EC. Compared to other different pre-processing methods, The PLS regression models combined with SNV + 1st derivative yield best results with R^2_{cv} 0.81 and RMSECV 1.52 for calibration and R^2_V of 0.83 and $RMSE_V$ 1.43 for validation. However, different pre-processing methods did not promoted much robust model for grain angle.

NIR spectra

The result shows that SNV + first derivative performed the highest ablitity for EC prediction using PLS model. Therefore SNV + first derivative pre-processing method were used for optimal wavenumbers selected. SNV+1st derivative spectra of wood discs are shown in Fig1. Although all spectra have a very similar shape, there is still a large variability in absorbance, the significant multivariate correlation of grain angle were most observed between 9000 cm-1 and 7000 cm-1, there are some overlap in this region with sMC of EC. the effect of grain angle performed on the spectra from 9000 to 7000 cm^{-1} and 5000 to 4000 cm^{-1} are stronger than regions from 7000 to 5000 cm^{-1}, especially the regions around 8800, 7400, 7100 and 4200 cm^{-1}.

Fig. 2 shows the significant multivariate correlation (sMC) of EC (line dotted red) and grain angle (solid black) on PLS model for each wavenumber of the *T. sinensis* heartwood spectra. It can be clearly seen that the important wavelengths variables for EC and grain angle have some overlap regions. However, there are some regions can be clearly distinguished between EC and grain angle, the most important regions for grain angle are mainly located at 8200 to 7800 cm^{-1}. The regions around 6000 and 8500 cm^{-1} are highly related to EC with no grain angle influence. Therefore, the important regions that the only highly correlate to EC prediction (the both points in Fig. 2) were selected for PLS model calibration and the grain angle important regions and other irrelevant variables have been removed from the NIR spectra.

PLS models with sMC selection

The PLS calibration models for EC and grain angle with these selected wavenumbers are shown in Table 2, In total, 19 wavenumbers was selected, which corresponded to less than 1% of the total number of wavenumbers. with the prior selection of wavenumbers for EC, the R^2 of EC calibration is still high (0.82) and with a low RMSE of 1.44. Simultaneously the R^2 for grain angle has been hugely reduced from 0.90 to 0.33. Which significantly reduced the influence of grain angle on NIR spectra when calibrate for EC.

The reference values versus NIR predicted values plot of the leave-one-out cross-validation for EC in the calibration sets and validation sets are displayed in Fig. 3. It shows that the samples are reasonably well distributed both in calibration and validation. With the sMC selected variables, PLS model could yields a highly accuracy of EC prediction without the grain angle influence.

The scores plot of PLS models using both full length of NIR spectra and the sMC selection variables were shown in Fig. 4. It is clearly shown that without the sMC selection, the grain

Angle have a significantly influences on EC, the EC based on these two group have been clearly distinguished (Fig. 4 (a)). However, it has been improved with sMC selection, the EC prediction of these two angle group are mixed together, the influence has been successfully reduced (Fig. 4 (b)).

Mode check

The selected model has been tested on the validation data, the EC prediction of PLS model using the full spectra are different between 0° and 90° angle. The EC at 90° shows higher than at 0°. The results shown that there is no statistical significant difference between the measurements at 0° and 90° grain direction when use the NIR model with sMC selected wavenumbers to predict the EC (Fig. 5). Therefore it was concluded that grain orientation is not affecting the assessment of extractive content on cores by NIR.

Discussion

T. sinensis hold a long history of culture for its digestible buds in China. However, the heartwood of *T. sinensis* also valuable and have a huge potential furniture market [23]. To analysis the wood properties for genetic selection, a fast and efficient way to measure the wood properties is needed. NIR spectra of solid wood have been reported that could be a potential method for predict wood properties Pulkka, et al. [24]. Similarly to our result, we successfully use of wood cores NIR spectra to predict the EC in the heartwood of *T. sinensis*. The influence of grain angle on the NIR spectra when other chemicals need to be predicted has been studied for a long time. In our study, we found that the EC prediction of *T. sinensis* from PLS model could be influenced by the grain angles. Supported by Li and Altaner [6] who used of NIR spectroscopy to predict the EC in *Eucalyptus bosistoana* and shown that the grain angle have significantly influence on the EC prediction model, the grain angle influence have been removed after the EPO algorithm applied, which is different from our study. Gierlinger, et al. [25] evaluated the differences between axial and radial faces spectral of heartwood to classify three species of larch. They reported that spectra from axial surfaces showed less heterogeneity between larches. Schimleck, et al. [26] compared the accuracy of two different NIR model based on the spectra of radial-longitudinal and radial- transverse faces to predict wood property, the results shown that radial-longitudinal provided a stronger calibration model than radial- transverse. Fujimoto, et al. [27] found that grain angle could have an influence on the reflectance of spectra when collected spectra from the wood surface on difference angles. The differences in model accuracy for grain angle estimation from the wood cross sections and radial section can be explained by the variation of anatomical structures of the different surfaces. The exposure of parenchyma cells in radial section is higher than cross section. Thus, it made the absorption bands, which related to cellulosic features are the main factor for mechanical properties, varies from different surfaces.

It is imperative to find out which are the dominant variables from multiple sources of NIR variation related to grain angle and should be remove when calibrate PLS model for EC. The sMC method which developed by Tran, et al. [28] can be efficiently used for important variable selection, sMC could provide an optimal variables list the most correlated to the response, these variables with minimal false negative and false positive errors could improves the predictive performance of the PLS model.

Two strong water absorbance bands exist around 7070 cm-1 and 5100 cm-1, which is similar to the 1st overtone of OH- bands. The peak band approximately 6000 cm^{-1} has been reported that is due to the 1st overtone of C-H stretching vibrations of methyl, methylene and ethylene groups and mostly related to the extractives [29]. Some differences between 0° and 90° have been observed from band 9000cm^{-1} to 7000cm^{-1} and 5500cm^{-1} to 4000cm^{-1}, are related to cellulose [27, 30-32].

The significant multivariate correlation of grain angle were most observed between 8500 cm^{-1} and 7500 cm^{-1}, there are some overlap in this region with sMC of EC. However, NIR spectra have some high correlation with EC from region 6000cm^{-1} to 4000 cm^{-1} and with low correlation with grain angle, especially at ~ 6000 cm^{-1} and 5400 cm^{-1}, which shows no correlation with grain angle, Similarly results were obtained by Schwanninger, et al. [29] . It shows that it is possible to selected NIR spectra for calibration EC without grain angle influence. Hence according to the sMC results for both EC and grain angle, the region of NIR spectra with high correlation with EC but low grain angle were selected for calibrate PLS model.

The robustness of our prediction model for EC has been hugely reduce the influence of grain angle, which make it easier for measure cores, cause the grain in the cores is uncertainty, either due to the difficulty to locate the 'up and down' orientation of the tree in the core or due to grain variation inherent in the tree. The model developed in this study is a promising and reliable method for predicting EC in solid wood. Compared to grind wood samples into powder, it provided a fast and non-destructives way to save the time and cost for analysis.

Conclusions

Our study have demonstrated that it is possible to predict EC in heartwood of *T. sinensis* by established NIR PLS regression models from solid wood samples. With a prior of wavenumbers selected by sMC method which could reduce the infection of grain angle on NIR spectra, the results yield a promising and efficiency of prediction method for EC in heartwood without the impact of grain angle. Furthermore, it is not necessary to consider wherever the NIR spectra measured on the cores and to grind wood into powder to obtain powder spectra for prediction, NIR spectra of solid wood can serve as a fast and non-destructive method for forest tree breeders, and this assists the use of NIR spectroscopy to screen *T. sinensis* for heartwood quality.

Materials And Methods

Materials

Toona sinensis have been planted in 2006 at Kaihua Forest Farm, Zhejiang province, P.R. China (118°20'E，29°12'N). 52 open-pollinated families were randomly block designed as 4 tree plot and repeated 10 times. The annual average precipitation and temperature in this site is 1814 mm and 16.3°C respectively.

Sample processing

Two-hundred and twenty-three wood cores, with 14mm diameter, were taken from the bottom of tree trunk of the 12 years old *T. sinensis*. Cores were labelled directly on the surface and placed into a paper bag, all samples were air-dried for a months until a stable moisture condition obtained. After air-dried, the longitudinal growth direction (0^{o}) and cross section (90^{o}) of each core sample were marked. The surface of 0^{o} and 90^{o} were then sanded using a p100 sandpaper to obtain a consistence surface for NIR spectra collection.

NIR spectral measurements

NIR spectra from cores were taken at room temperature with a fibre optics probe (Antaris II System, Thermo Electron Company, USA) at wavelengths from 9000 to 4000 cm^{-1} at 10 cm-1 resolution. Each NIR spectra were taken at regular intervals (5 mm) from pith to bark in heartwood along the 0o and 90o surface of cores and calculate the weighted average. 64 co-averaged scans were averaged for each spectrum.

EC extraction

Each core was cut into small chips using a chisel and milled into fine powder with a 2-mm screen. The wood powder samples were oven dried at 60°C to obtain a stable moisture content. All cores samples were used for extraction after the NIR spectra collection. The accelerated solvent extraction (ASE) was processed using the thermo accelerated solvent extractor 350 (Thermo Fisher Scientific, Bellefonte, PA, USA) with ethanol. the methods were similar to Li and Altaner [6] . 4g of wood powder was placed into a stainless steel cells and the setup extraction process is as follows: static time 15 min, temperature 70°, 100% rinse volume and 2 extraction cycles. The extractive solutions of each samples were collected into a dry aluminium foil tray with known mass and placed in the fume cupboard overnight for the ethanol to evaporate. Subsequently the extracts were dried in an oven overnight at 105 °C to remove moisture. The mass of the extract was determined and the extractives content calculated on a dry mass basis. EC for the samples was ranging between 1.23 % and 16.49% with an average of 8.86%.

NIR spectra processing

To find out whether the grain angle have affection on the NIR spectra for EC prediction in *T. sinensis* or not. The optimal range of the spectrum for accurate calibration was investigated. PLS regression was used to perform grain angle classification analysis and to predict EC. The data from the samples were divided into two data set using Kennard-Stone sampling with Euclidean distance [33], including 200 selected samples as calibration data and the remained 37 as validation data set for PLS model calibration and validation. Different spectra pre-processing were applied on the NIR data sets before model calibration to character the best suitable pre-processing methods, which including standard normal variation (SNV), 1st and 2nd derivative using Savitzky-Golay algorithm with 2 order polynomial and 15 window sizes. A filter methods [34] which called significant multivariate correlation (sMC) was used to find out the most important variables for grain angle and EC respectively and to eliminate the unimportant variables and develop a more reliable and robust model for EC prediction. The coefficient of determination (R^2) and root-mean-square error (RMSE) derived from both the calibration (R^2_{CV} and $RMSE_{CV}$) and validation (R^2_V and $RMSE_V$) were used to track the model performance. Data analysis was conducted in R software (version 3.1.2) [35]. The pls package [36] was used for PLSR and sMC-PLSR model performing. The plsVarSel [34] for sMC variables selection. The prospectr package [37] was used for NIR spectra manipulation and Kennard-Stone sampling. The ggplot2 package [38] was used for visualization plot.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Competing interests

The authors declare that there is no conflict of interest.

Funding

This work was funded by Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding (No. 2016C02056-12) and National Natural Science Foundation of China (No. 31570658).

Authors' contributions

Yanjie Li designed the study, conducted the experiment, analysis the data and wrote the manuscript. Xin Dong, Yang Sun conducted lab experiments, Jun Liu and Jingmin Jiang supervised the experiments at all stages and reviewed the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

Not applicable.

Authors' information

Not applicable.

References

1. Taylor, A.M. et al. Heartwood formation and natural durability-a review. Wood. Fiber. Sci. 2007; 34 (4): 587-611.
2. Hillis, W.E. Heartwood and tree exudates, 1987, Springer Verlag.
3. Hawley, L.F. et al. The relation between durability and chemical composition in wood. Ind. Eng. Chem. Res. 1924; 16: 699-700.
4. Bush, D. and Walker, J. Selecting and breeding eucalypts for natural durability. Proceedings of Developing a Eucalypt Resource—Learning from Australia and Elsewhere, Blenheim, New Zealand. 2011: 3-4.
5. Scheffer, T.C. Natural resistance of wood to microbial deterioration. Annual Review of Phytopathology. 1966; 4 (1): 147-168.
6. Li, Y. and Altaner, C. Predicting extractives content of Eucalyptus bosistoana F. Muell. Heartwood from stem cores by near infrared spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2018; 198: 78-87.
7. Li, Y. et al., Genetic variation in heartwood properties and growth traits of Eucalyptus bosistoana Forest ecology and management, 2018.
8. Tappi, T., 204 cm-97, Solvent extractives of wood and pulp, TAPPI test methods, TAPPI Press, Norcross, GA, USA, 2004.
9. Roberts, C.A. et al. Near-infrared spectroscopy in agriculture, 2004, American Society of Agronomy Madison.
10. Zhao, D. et al. Identification of red and NIR spectral regions and vegetative indices for discrimination of cotton nitrogen stress and growth stage. Computers and Electronics in Agriculture. 2005; 48 (2): 155-169.
11. Baillères, H. et al. Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program. Annals of Forest Science. 2002; 59 (5-6): 479-490.
12. da Silva Allan, R. et al., Assessment of total phenols and extractives of mahogany wood by near infrared spectroscopy (NIRS), 2013, p. 1.

13. Geladi, P. et al. Data analysis of hyperspectral NIR image mosaics for the quantification of linseed oil impregnation in Scots pine wood. Wood Science and Technology. 2014; 48 (3): 467-481.
14. Maniwara, P. et al. The use of visible and near infrared spectroscopy for evaluating passion fruit postharvest quality. Journal of Food Engineering. 2014; 143: 33-43.
15. Gherardi Hein, P.R. et al. Effects of sample preparation on NIR spectroscopic estimation of chemical properties of Eucalyptus urophylla ST Blake wood. Holzforschung. 2010; 64 (1): 45-54.
16. Giordanengo, T. et al. Correction of moisture effects on near infrared calibration for the analysis of phenol content in eucalyptus wood extracts. Ann. For. Sci. 2008; 65 (8): 8.
17. Schimleck, L. et al. Estimation of the physical wood properties of green Pinus taeda radial samples by near infrared spectroscopy. Can J For Res. 2003; 33 (12): 2297-2305.
18. Wijewardane, N.K. et al. Moisture insensitive prediction of soil properties from VNIR reflectance spectra based on external parameter orthogonalization. Geoderma. 2016; 267: 92-101.
19. Barreiro, P. et al., Comparison of robust modeling techniques on NIR spectra used to estimate grape quality, IV International Symposium on Applications of Modelling as an Innovative Technology in the Agri-Food-Chain: Model-IT 802, 2008, pp. 367-372.
20. Minasny, B. et al. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. Geoderma. 2011; 167: 118-124.
21. Tran, T.N. et al. Interpretation of variable importance in partial least squares with significance multivariate correlation (sMC). Chemometr. Intell. Lab. Syst. 2014; 138: 153-160.
22. Li, Y. and Altaner, C. Effects of variable selection and processing of NIR and ATR-IR spectra on prediction of extractives content in Eucalyptus bosistoana heartwood. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2019.
23. FAN, Z.-f. et al. Comparative studies on fiber morphology and chemical composition between Toona sinensis wood from plantation and natural forest [J]. Subtropical Plant Science. 2003; 3: 34-37.
24. Pulkka, S. et al. Prediction of stilbene content from heartwood increment cores of Scots pine using near infrared spectroscopy methodology. Journal of Near Infrared Spectroscopy. 2016; 24 (6): 517-528.
25. Gierlinger, N. et al. Characteristics and classification of Fourier-transform near infrared spectra of the heartwood of different larch species (Larix sp.). Journal of Near Infrared Spectroscopy. 2004; 12: 113-120.
26. Schimleck, L.R. et al. Comparison of Pinus taeda L. wood property calibrations based on NIR spectra from the radial-longitudinal and radial-transverse faces of wooden strips. Holzforschung. 2005; 59 (2): 214-218.
27. Fujimoto, T. et al. Application of near infrared spectroscopy for estimating wood mechanical properties of small clear and full length lumber specimens. Journal of Near Infrared Spectroscopy. 2008; 16 (6): 529-537.

28. Tran, T.N. et al. Interpretation of variable importance in partial least squares with significance multivariate correlation (sMC). Chemometrics and Intelligent Laboratory Systems. 2014; 138: 153-160.
29. Schwanninger, M. et al. A review of band assignments in near infrared spectra of wood and wood components. Journal of Near Infrared Spectroscopy. 2011; 19 (5): 287.
30. Sinelli, N. et al. Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity. Food Research International. 2010; 43 (1): 369-375.
31. Christy, A.A. et al. The detection and quantification of adulteration in olive oil by near-infrared spectroscopy and chemometrics. Analytical Sciences. 2004; 20 (6): 935-940.
32. Cozzolino, D. et al. Relationship between wine scores and visible–near-infrared spectra of Australian red wines. Anal Bioanal Chem. 2008; 391 (3): 975-981.
33. Kennard, R.W. and Stone, L.A. Computer Aided Design of Experiments. Technometrics. 1969; 11 (1): 137-148.
34. Mehmood, T. et al. A review of variable selection methods in partial least squares regression. Chemometrics and Intelligent Laboratory Systems. 2012; 118: 62-69.
35. R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2017.
36. Mevik, B. et al., Partial Least Squares and Principal Component Regression. R package version 2.5-0, 2015.
37. Stevens, A. and Ramirez–Lopez, L. An introduction to the prospectr package. R Package Vignette, Report No.: R Package Version 0.1. 2014; 3.
38. Wickham, H. ggplot2: Elegant Graphics for Data Analysis, 2016, Springer.

Tables

Table 1 PLS regression models for calibration and validation of EC and grain angle with and without pre-treatment methods

		Calibration			validation	
	Pre-treatment	R^2_{CV}	$RMSE_{CV}$	LVs	R^2v	$RMSE_V$
EC	No (raw spectra)	0.75	1.69	10	0.79	1.74
	SNV	0.76	1.64	8	0.78	1.55
	1st derivative	0.79	1.52	6	0.81	1.42
	2nd derivative	0.80	1.61	8	0.81	1.52
	SNV+1stderivative	0.81	1.52	7	0.83	1.43
	SNV+2ndderivative	0.75	1.70	10	0.79	1.63
Grain angle	No (raw spectra)	0.87	11.32	14	0.91	9.85
	SNV	0.89	10.03	12	0.91	9.97
	1st derivative	0.89	9.98	10	0.92	9.76
	2nd derivative	0.88	10.34	10	0.90	9.99
	SNV+1stderivative	0.90	9.75	11	0.93	9.54
	SNV+2ndderivative	0.89	10.11	14	0.91	10.01

Table 2 PLS regression models for calibration and validation of EC and grain angle with selected variables and SNV+1st derivative

			Calibration			validation	
	Pre-treatment	Number of variables	R^2_{cv}	$RMSE_{CV}$	Ncomp	R^2_v	$RMSE_V$
EC	SNV+1stderivative	19	0.82	1.44	7	0.84	1.35
Grain angle	SNV+1stderivative	19	0.33	45	7	0.38	50

Figures

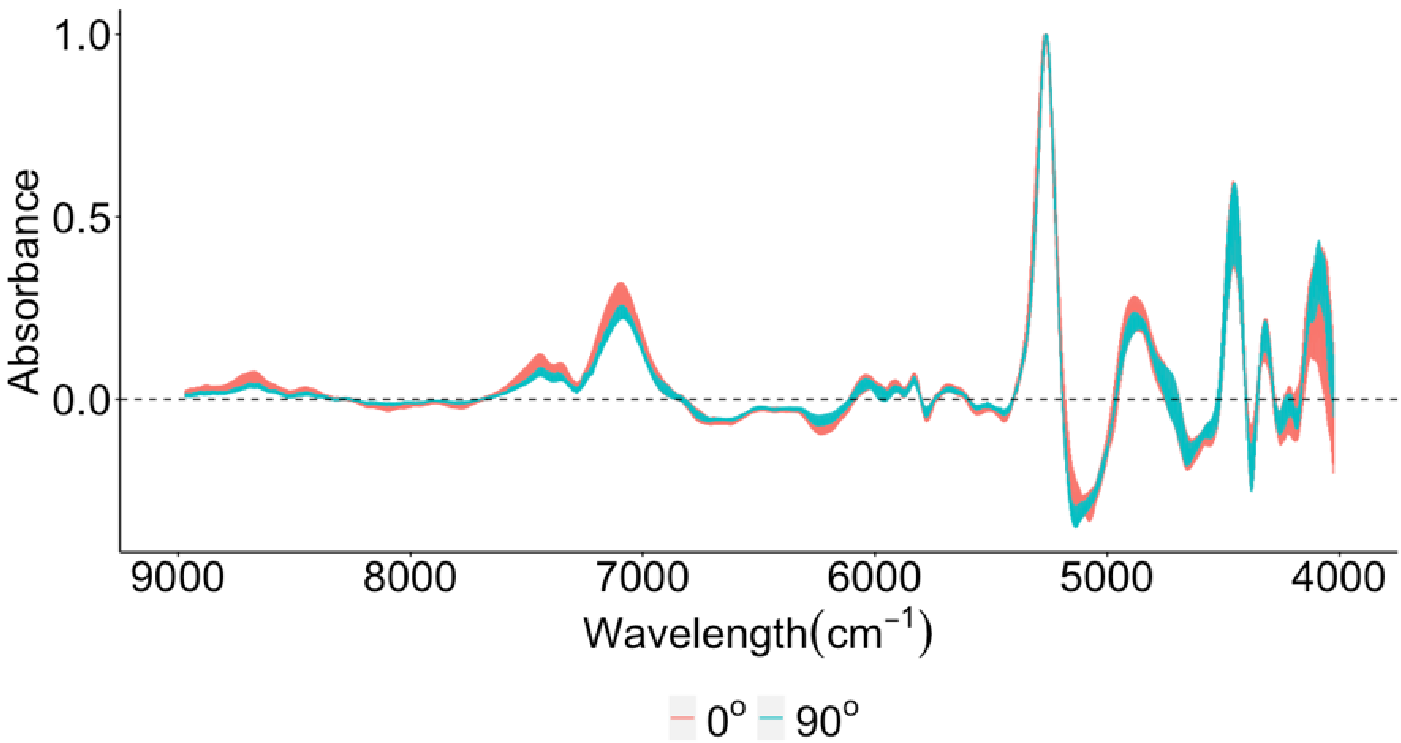

Figure 1

SNV+1st derivative absorbance spectra of all measured cores between 9000cm-1 and 4000cm-1

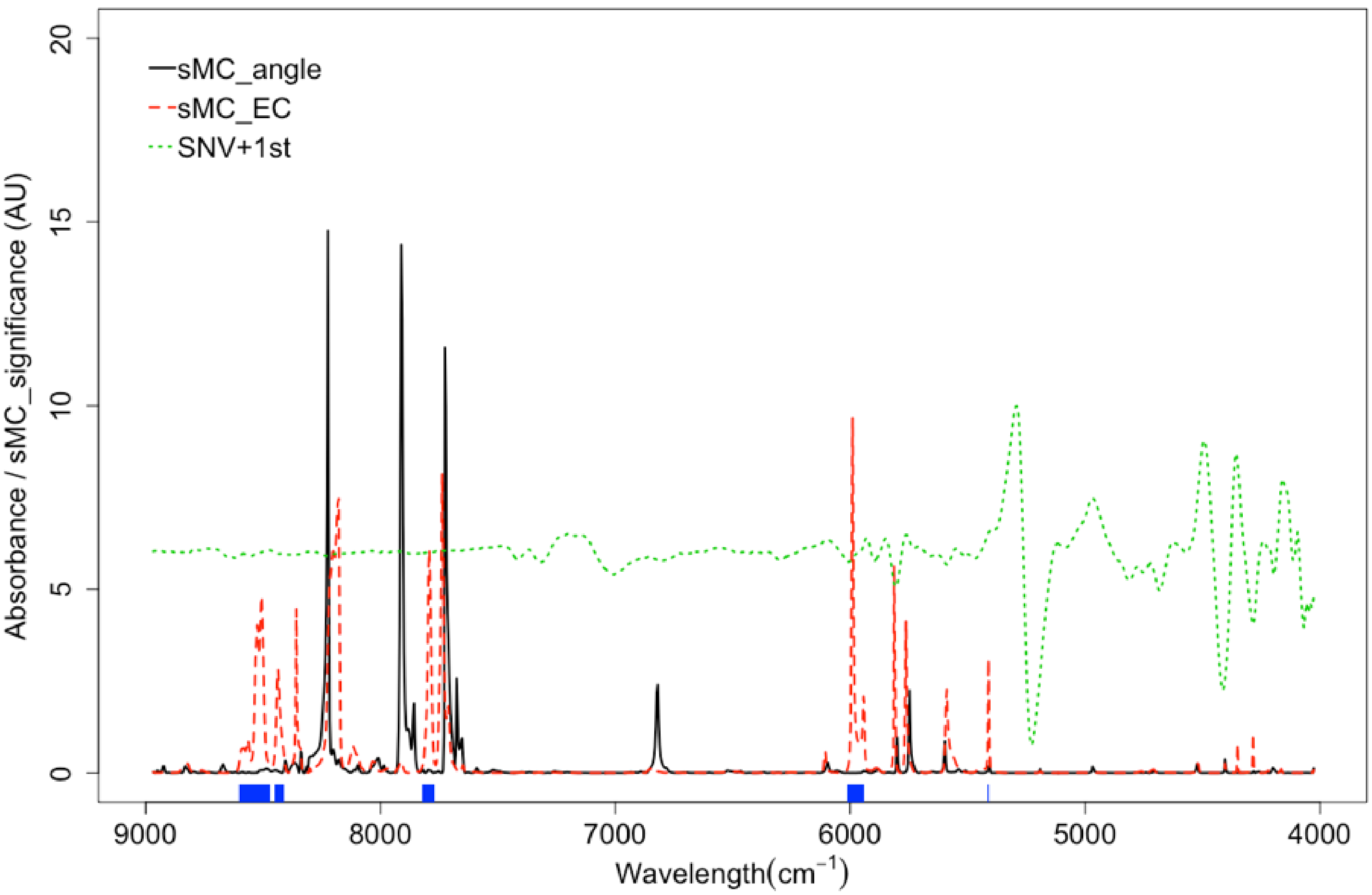

Figure 2

The influences of grain angle and EC on NIR spectra of T. sinensis. Optimum wavenumbers selected (blue), SNV+1st derivative spectra (dotted green).

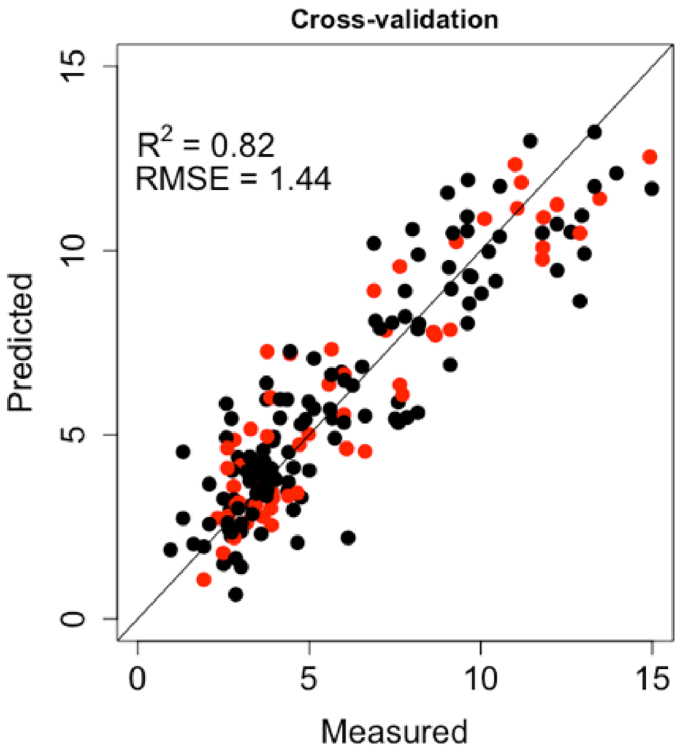

Figure 3

The best calibration model for EC with selected wavenumbers without influence of grain angle with the measured versus predicted values in the (left) calibration and (right) validation sets.

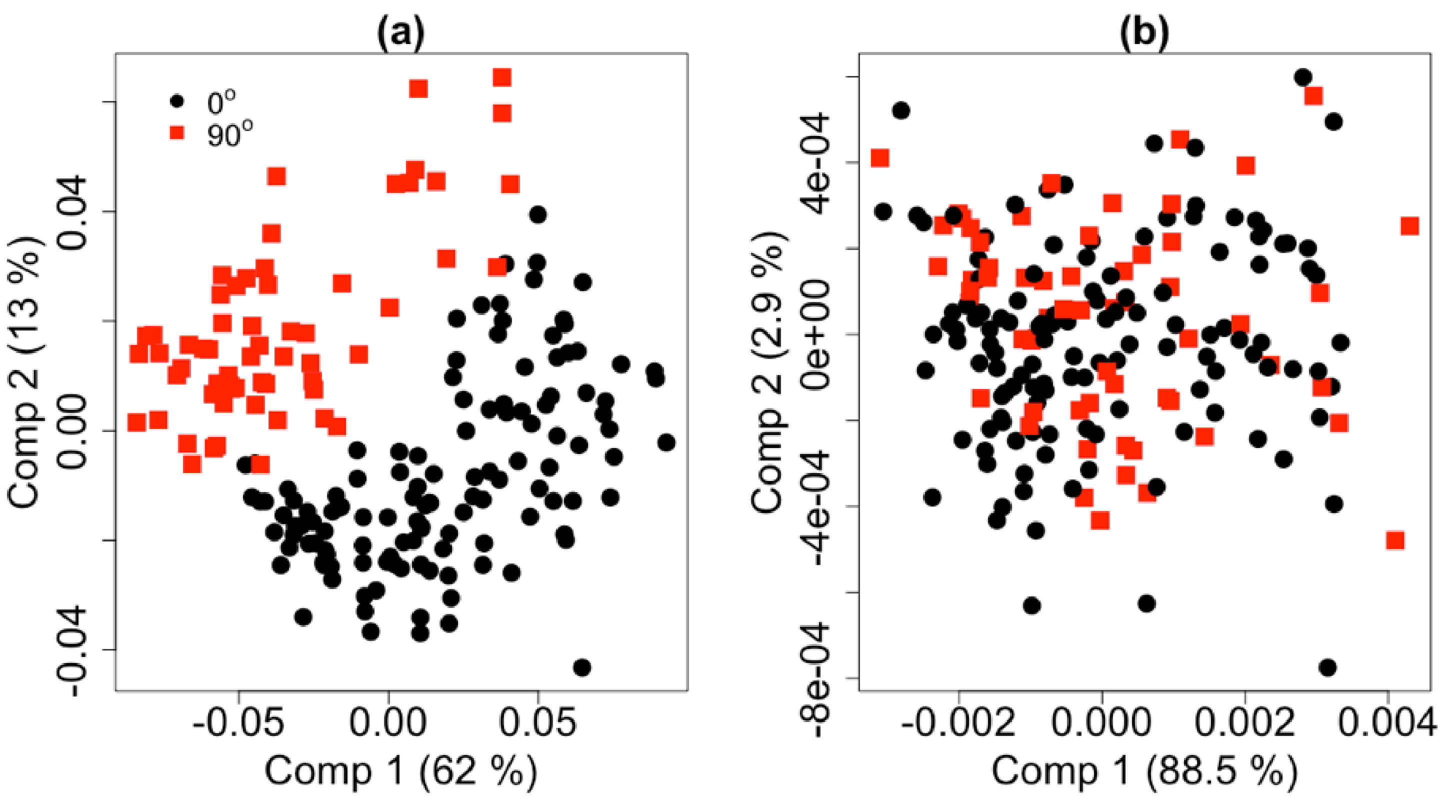

Figure 4

the scores plot of PLS model for EC prediction based on the full length of NIR spectra and the sMC optimal selected spectra variables.

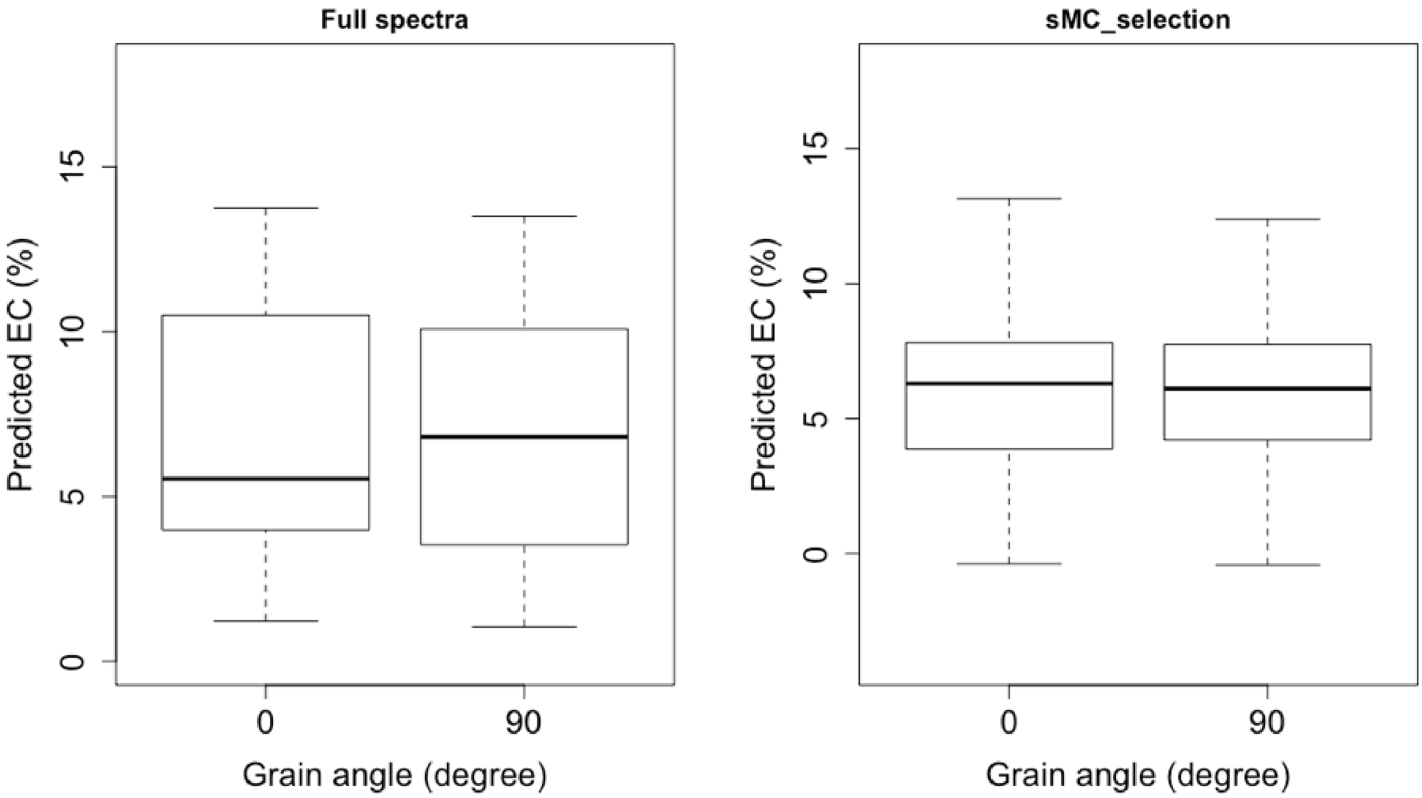

Figure 5

predicted EC on T. sinensis validation cores samples dependent on grain angle with and without sMC variables selection.