Quantitative Analysis of Routine Chemical Constituents of Tobacco Based on Thermogravimetric Analysis

Yuhan Peng

China Tobacco Zhejiang Industrial Co., Ltd

Yiming Bi

China Tobacco Zhejiang Industrial Co., Ltd

Lu Dai

China Tobacco Zhejiang Industrial Co., Ltd

Haifeng Li

China Tobacco Zhejiang Industrial Co., Ltd

Depo Cao

China Tobacco Zhejiang Industrial Co., Ltd

Qijie Qi

China Tobacco Zhejiang Industrial Co., Ltd

Fu Liao

China Tobacco Zhejiang Industrial Co., Ltd

Yudong Shen

China Tobacco Zhejiang Industrial Co., Ltd

Fangqi Du

China Tobacco Zhejiang Industrial Co., Ltd

Hui Wang (✉ wanghui@zjtobacco.com)

China Tobacco Zhejiang Industrial Co., Ltd

Research Article

Keywords: Thermogravimetric analysis, Chemometrics, Tobacco, Quantitative analysis, Routine chemical constituents

Posted Date: January 17th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1258420/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Thermogravimetric analysis (TGA) is an important method to characterize the pyrolysis reaction characteristics of tobacco, reflecting the chemical composition of tobacco from the perspective of chemical reaction. However, quantitative analysis of tobacco chemical indexes based on TG curve has not yet been reported due to the complexity to the best of our knowledge.

Results

By the combination of TG curve and partial least squares (PLS) algorithm, we established quantitative analysis models of six routine chemical constituents in tobacco, including total sugar, reducing sugar, total nitrogen, total alkaloids, chlorine and potassium. We also performed an in-depth analysis of the chemical mechanism revealed by the result of the quantitative model, namely the regression coefficient which reflected the correlation degree between the six chemicals and different stages of tobacco thermal decomposition process.

Conclusions

The quantitative analysis model of chemical constituents in tobacco based on TG curve can be used for the rapid and accurate analysis of compound content. From the perspective of pyrolysis reaction mechanisms, the quantification of total sugar and reducing sugar mainly depended on the Maillard reaction in which carbohydrate compounds participate. The correlation of total nitrogen and total alkaloids with TG curves in the whole temperature range was affected by total sugar. Although chlorine was detrimental to Maillard reaction in general, it has different effects on the different elementary steps of the Maillard reaction. Potassium can promote the thermal decomposition of hemicellulose, pectin, and water-soluble carbohydrates such as glucose and fructose.

Background

As a special lignocellulose biomass, the chemical composition of tobacco is complex [1–3]. Among these compositions, the content of routine chemical constituents, including the total sugar, reducing sugar, total nitrogen, total alkaloid, chlorine and potassium, is the most basic index to evaluate the quality of tobacco for the formulation design, quality monitoring, and classification of cigarette products [4–6]. At present, continuous-flow analysis is mainly used to detect routine chemical constituents in tobacco. Unfortunately, this method suffers from complex operation, time consumption and environmental pollution caused by the consumption of a large amount of organic reagents in the testing process [5–6]. Therefore, developing rapid and accurate quantitative analysis of routine chemical constituents in tobacco is necessary for quality assurance of cigarettes.

With the rise and development of chemometrics methods, the coupling of chemometrics with spectroscopy which can characterize the chemical information of sample has been widely used in quantitative and qualitative analysis of complex system [7–8]. In the quantitative analysis of tobacco chemical composition, previous studies mainly focused on the use of near-infrared (NIR) spectroscopy due to its high efficiency and nondestructive characteristic [9–12]. For example, Zhao *et al.* found that NIR spectroscopy combined with support vector machine (SVM) enabled rapid and accurate analysis of routine chemical compositions in tobacco [10]. Zhou *et al.* proposed an ensemble partial least squares algorithm based on variable clustering for quantitative analysis of nicotine in tobacco by NIR spectroscopy [11].

Apart from NIR, which directly characterizes the composition of tobacco through the molecular vibration of tobacco endogenous substances, we propose an indirect method, thermogravimetric analysis (TGA). TGA is an important method to characterize the pyrolysis reaction characteristics of tobacco, reflecting the chemical composition of tobacco from the perspective of chemical reaction [13–16]. The shape of thermogravimetric (TG) curve was comprehensively determined by the physical and chemical properties of tobacco, including substance composition, complex cross-linking structure between compounds, microscopic pores of tobacco, and so on [15–16]. However, quantitative analysis of tobacco chemical indexes based on TG curve has not yet been reported due to the complexity to the best of our knowledge.

Herein, by the combination of TG curve and chemometrics, we established quantitative analysis models of six routine chemical constituents of total sugar, reducing sugar, total nitrogen, total alkaloids, chlorine and potassium in tobacco. Due to the differences in testing principles, the response mechanism and intensity of TG curve to chemical information are different from that of NIR, which may provide a new idea for qualitative and quantitative analysis of tobacco.

Results And Discussion

Analysis of thermogravimetric (TG) curves

Figure 1 shows the thermogravimetric (TG) curves of 49 tobacco samples during the thermal decomposition process. Overall, TG curves of 49 tobacco did not appear to be significantly different. We conducted a general analysis of the correlation between the contents of six routine chemical constituents and TG curves of 49 tobacco. The contents of these constituents were detected by a continuous-flow analyzer. The correlation was analyzed based on Pearson correlation coefficients which is an index used to measure the degree of linear correlation between two variables. In general, the range in the absolute value of Pearson coefficient from 0.9 to 1.0 represents a very strong linear correlation; the range from 0.7 to 0.9 represents a strong linear correlation; the range from 0.5 to 0.7 represents a moderate linear correlation; and absolute values below 0.5 indicates a weak or negligible correlation between the two variables [18–21].

Figure 2, a and b reflect the linear correlation degree between total sugar, reducing sugar and TG curve. Total sugar and reducing sugar were mainly composed of water-soluble reducing sugars such as glucose and fructose in tobacco. The contents of these components have a highly linear correlation with TG curve at approximately 473 K. Specifically, the coefficients for total sugar and reducing sugar reached 0.95 and 0.89, respectively. The correlation may be derived from Maillard reaction. According to previous studies, the reaction occurring around 473 K mainly comprises the release of Maillard products and the volatilization of nicotine during the thermal decomposition process of tobacco [22–23]. TG curve represents the remaining mass of the sample after the release of volatile products. As an important reactant of Maillard reaction, the higher the content of water-soluble reducing sugar is, the faster the release rate of products will be, resulting in less remaining mass of the sample. Therefore, the contents of total sugar and reducing sugar were negatively correlated with the TG curve at approximately 473 K.

As shown in Figure 2c, a negative linear correlation between total nitrogen and TG curve appeared at about 423 K, reaching 0.56. This phenomenon may also be attributed to Maillard reaction. As the main components of nitrogenous compounds in tobacco, amino acids and proteins are also another reactant of Maillard reaction besides water-soluble sugars.

There was a moderate positive linear correlation (0.60–0.65) between chlorine content and TG curve in the temperature range of 450 K to 530 K (Figure 2e). As such, chlorine was detrimental to Maillard reaction. This was also consistent with previous studies that high chlorine content worsens the sensory quality of cigarettes, possibly because chlorine inhibits the release of aromatic components such as Maillard products [24–25].

The linear correlations of both total alkaloids and potassium content with TG curve were weak within the whole temperature range where the absolute values were lower than 0.5, as shown in Figure 2, d and f. In this case, the relationship between the contents of these two compounds and TG curve was relatively complex.

For clarity, we briefly summarized the correlation of chemical constituents and TG curves based on Pearson correlation coefficients. The contents of total sugar, reducing sugar, and total nitrogen were negatively correlated with the TG curve at ~473 K, ~473 K, and ~423 K respectively, which may be caused by Maillard reaction. These three substances showed different degrees of linear correlation with TG curve. Among them, total sugar and reducing sugar were strongly correlated with TG curve, while total nitrogen was moderately correlated with TG curve. A moderate positive linear correlation between chlorine content and TG curve occurred in the temperature range of 450 K to 530 K. The direct linear correlation of both total alkaloid is and potassium contents with TG curves was rather weak in the whole temperature range, indicating that the mathematical relationships between these two compounds and TG curves were relatively complex. It may be necessary to carry out principal component extraction of TG curve and find the linear combination of TG values at different temperature points as a new variable to associate with total alkaloid is and potassium content, so as to improve the correlation between TG curve and these two chemical index.

Establishment of calibration model

From section 3.1, it was roughly proven that there was a certain correlation between TG curves and constituents content. In order to establish a quantitative analysis model, the specific mathematical relationship between them was further analyzed. The 49 tobacco samples were numbered from 1-49 and randomly divided into 39 calibration samples and 10 test samples according to the ratio (8:2) of calibration set to test set [26]. Partial least squares algorithm written by MATLAB software was used to correlate the contents of six routine chemical constituents in 39 calibration samples with the corresponding TG curves. The quantitative analysis models of six routine chemical constituents of tobacco based on TG curve were finally established *via* Root Mean Square Error for Calibration (RMSEC) and Root Mean Square Error for Cross-Validation (RMSECV). The TG curves of 10 samples in the test set were input into the calibration model to predict the contents of corresponding chemical constituents, and then compared with the measured values obtained by continuous-flow analyzer. The accuracy of the model was evaluated and verified by calculating Root Mean Square Error for Prediction (RMSEP).

Table 1 shows the number of latent variables (LV), RMSEC, RMSECV, and RMSEP of the model. The mean values (Mean (Y)) of six routine chemical constituents for 49 tobacco samples were also listed in Table 1. RMSEP of total sugar, reducing sugar, and total nitrogen were 0.46, 1.06 and 0.05, respectively, which were relatively small compared with the corresponding values of Mean(Y), indicating high accuracy of the quantitative analysis models of total sugar, reducing sugar, and total nitrogen. RMSEC and RMSEP of total alkaloids were 0.15 and 0.26, respectively, which were relatively high with respect to Mean(Y). Therefore, the prediction of the current model for total alkaloids content was less accurate than that for sugar models. The contents of chlorine and potassium in tobacco were so low that the comparison of relative deviation was not meaningful. Here, RMSEP values of 0.12 and 0.20 can guarantee the accuracy of prediction of these two indicators in an acceptable range [6, 27–28]. The predicted contents of routine chemical constituents in calibration and test samples obtained by quantitative analysis model were compared with actual values, as shown in Figure 3. Therefore, the established model can be used for accurate quantitative analysis of six conventional chemical substances including total sugar, reducing sugar, total nitrogen, total alkaloids, chlorine and potassium.

Table 1
The prediction result of the calibration model

Routine chemical constituents	LV	RMSEC	RMSECV	RMSEP	Mean (Y)
Total sugar	6	0.40	0.60	0.46	28.92
Reducing sugar	6	0.76	1.14	1.06	26.10
Total nitrogen	6	0.05	0.06	0.05	1.97
Total alkaloids	10	0.15	0.24	0.26	2.27
Chlorine	11	0.03	0.06	0.12	0.37
Potassium	10	0.05	0.11	0.20	2.02

Analysis of model results from the perspective of chemical reaction mechanism

We tried to explore the pyrolysis mechanism of tobacco through the results of the quantitative model established in section 3.2. For tobacco samples, the ordinate value (weight (%)) corresponding to each abscissa (temperature (K)) of the TG curve represents an independent variable of the sample, while the chemical content of the sample is the dependent variable. From a mathematical point of view, the establishment of the model aims to find a coefficient corresponding to each independent variable. The dot product of the coefficient and independent variable is equal to the dependent variable. This coefficient is named as the regression coefficient which can reflect the unique contribution of each independent variable [29–30].

Figure 4 shows the regression coefficients of the quantitative analysis model of six routine chemical constituents based on TG curves in the whole pyrolysis temperature range. As shown in Figure 4, a and b, the contents of total sugar and reducing sugar exhibited a strong negative correlation with TG curve at about 473 K, which may be attributed to Maillard reaction, also consistent with the conclusion in Figure 2. The correlation between total nitrogen and TG curve mainly existed before 673 K (Figure 4c). According to previous studies, the temperature range between 373 K and 473 K of TG curve mainly corresponded to the release of Maillard reaction products and the volatilization of nicotine. The temperature range between 523–623 K was mainly attributed to the pyrolysis process of glucose, pectin, hemicellulose, cellulose, and other saccharides [23, 31–32]. The results showed that there was a strong correlation between total nitrogen content and these reactions. The thermal decomposition of nitrogen-containing compounds in tobacco occurred only when the temperature was above 673 K [23, 33]. There was no obviously strong correlation between total nitrogen and this temperature range, indicating that the current model mainly applies to quantitative analysis of nitrogen-containing compounds through Maillard reaction, nicotine volatilization, and the thermal decomposition process of sugar rather than their own pyrolysis process.

The positive and negative correlation between total nitrogen and TG curve was opposite to that of total sugar in the whole temperature range. TG curve records the relative percentage content of chemical substances consumed rather than the absolute mass. The content of total sugar in tobacco (approximately 30%) was much higher than that of total nitrogen (less than 2%). The higher the relative content of sugar is, the lower the relative content of total nitrogen will be.

There was a strong positive correlation between total alkaloids and TG curve at about 520 K, as shown in Figure 4d. This phenomenon is out of our expectation, because more than 90% of total alkaloids in tobacco were composed of nicotine, which was mainly volatilized at 473-573 K. Accordingly, the higher the nicotine release amount, the less residual mass corresponding to TG curve. In other words, the nicotine content and TG curve should theoretically be negatively correlated. We hypothesize that this phenomenon is also affected by the correlation between sugar and TG curve, just similar to total nitrogen.

Based on regression coefficient, the numerical relationship between chlorine content and TG curve was relatively complex ((Figure 4e), although chlorine has a significant positive linear correlation with TG in the temperature range of 450 K to 530 K from the perspective of Pearson correlation coefficients as shown in Figure 2e. This may be because the Maillard reaction and the thermal decomposition of sugars are in essence a complex multi-step reaction process. The influential mechanism of chlorine on different elementary reaction step occurred at different temperature ranges varies. Therefore, the response mode and intensity of TG curve to chlorine content at different temperature points are also different.

There was a strong negative correlation between potassium content and TG curve at about 600 K, as shown in Figure 4f. This temperature range mainly corresponded to the pyrolysis process of hemicellulose, pectin and water-soluble carbohydrates such as glucose and fructose, indicating that potassium was conducive to the pyrolysis reaction of these substances. This was also consistent with previous studies that potassium salt has catalytic promoting effect on the degradation of oxygen-containing organic functional groups of carbohydrates in tobacco [34].

From the analysis of the regression coefficients, we found that the quantification of total sugar and reducing sugar mainly depended on the Maillard reaction in which carbohydrate compounds participate. The correlation of total nitrogen and total alkaloids with TG curves in the whole temperature range was affected by total sugar. Although chlorine is detrimental to Maillard reaction in general, it has different effects on the different elementary steps of the Maillard reaction. Potassium can promote the thermal decomposition of hemicellulose, pectin, and water-soluble carbohydrates such as glucose.

Conclusions

In this work, for the first time, quantitative analysis models of total sugar, reducing sugar, total nitrogen, total alkaloids, potassium and chlorine in tobacco based on TG curve were established, realizing the rapid and accurate determination of the content of these six routine chemical constituents. PLS method was adopted to realize the establishment of the quantitative analysis model. The accuracy of this model was confirmed by the value of RMSEP. We performed an in-depth analysis of the chemical mechanism

revealed by the result of the quantitative model, namely the regression coefficient which reflected the correlation degree between the six chemicals and different stages of tobacco thermal decomposition process.

We believe that this quantitative analysis model of chemical constituents in tobacco based on TG curve can not only be used for the accurate analysis of compound content, but also provide enlightenment for the study of pyrolysis reaction mechanism from the perspective of data analysis, which is of great significance for quality control, tar and damage reduction of cigarette.

Materials And Methods

Materials

49 single-grade flue-cured tobacco used in this work were obtained from the Technology Center of China Tobacco Zhejiang Industrial Co., Ltd. (Hangzhou, China). Before usage, the tobacco was ground into powder. The powder that passed through 40-mesh but was trapped on a 60-mesh sieve was collected in a sealed valve bag for subsequent measurements. The size of tobacco particles ranged from 0.250 to 0.425 mm.

Pyrolysis experiment

The weight-loss characteristics of the samples were analyzed in a thermogravimetric analyzer (Discovery, TA). In a typical run, 10 mg of sample was used. The sample was heated from room temperature to 373 K at a rate of 10 K/min, followed by incubation at 373 K for 30 min to remove free water completely. The sample weight was regarded as 100% after dehydration pretreatment. We heated the sample to 1173 K at a rate of 10 K/min and recorded the weight loss of the sample during the temperature-programmed process. During the whole tests, the flow rates of the carrier gas (high-purity N₂) and protective gas (high-purity N₂) were set at 50 mL/min and 30 mL/min, respectively. Each test was repeated three times under the same conditions.

Chemical analysis of six routine constituents of tobacco

The contents of total sugar, reducing sugar, total nitrogen, total alkaloids, chlorine and potassium in 49 tobacco powder samples were determined by a continuous flow analyzer (Alliance-Futura).

Partial least squares (PLS) modelling

Before modeling, the dataset was randomly divided into a calibration set (39 samples) and a test set (10 samples). The PLS method was used to build a calibration model. Fivefold cross validation was performed to the calibration set to calculate the RMSECV value. An F test based on the result of cross-validation was used to select the optimal number of latent variables. The significance level was set to 0.25 as previously suggested [17]. Prior to building the PLS model, all data were mean-centered.

Evaluation

The root mean square error (RMSE) is used as a measure of model performance. RMSE is defined as follows,

$$\text{RMSE} = \sqrt{\frac{\sum (y_{pre} - y_{ref})^2}{N}} \quad (1)$$

where y_{pre} is the predicted value, y_{ref} is the actual value, and N is the number of samples. RMSEC is RMSE calculated from the calibration samples, that is, a measure of fit. RMSECV is calculated from the cross-validated samples. The root mean square error of prediction (RMSEP) is calculated from test set.

Computation

Computations were performed in MATLAB 7.14 (Mathworks, Inc., Natick, MA, USA). All of the programs were written in-house and run on a personal computer with a 3.20 GHz Intel Core i5 processor, 8 GB RAM, and Windows 7 operating system.

Declarations

Acknowledgements

This research was supported by the Science Foundation of China Tobacco Zhejiang Industrial (Grants no. ZJZY2021A009 and ZJZY2021A013).

Authors' contributions

Yuhan Peng performed the research and data analysis, and prepared the manuscript. Yiming Bi establish the model by using PLS algorithm. Lu Dai and Haifeng Li supervised the study and prepared figures. Depo Cao, Qijie Qi, Fu Liao and Yudong Shen collected the tobacco samples, ground the sample into powder and performed TGA experiments. Fangqi Du and Hui Wang developed the idea for the study, set up the methodology, and edited the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Rodgman A, Perfetti TA. The Chemical Components of Tobacco and Tobacco Smoke. Boca Raton FL: CRC Press. 2012.
2. Margham J, McAdam K, Forster M, Liu C, Wright C, Mariner D, Proctor C. Chemical Composition of Aerosol from an E-Cigarette: A Quantitative Comparison with Cigarette Smoke. *Chemical Research in Toxicology*. 2016; 29:1662-1678.
3. Banožić M, Babić J, Jokić S. Recent advances in extraction of bioactive compounds from tobacco industrial waste-a review. *Industrial Crops and Products*. 2020;144:112009.
4. Shao Y, He Y, Wang Y. A new approach to discriminate varieties of tobacco using vis/near infrared spectra. *European Food Research and Technology*. 2006;224:591-596.
5. Zimmer GF, Santos RO, Teixeira ID, Schneider RDD, Helfer GA, Costa AB. Rapid quantification of constituents in tobacco by NIR fiber-optic probe. *Journal of Chemometrics*. 2020;34.
6. Duan J, Huang Y, Li ZH, Zheng B, Li QQ, Xiong YM, Wu LG, Min SG. Determination of 27 chemical constituents in Chinese southwest tobacco by FT-NIR spectroscopy. *Industrial Crops and Products*. 2012;40:21-26.
7. Casale M, Casolino C, Oliveri P, Forina M. The potential of coupling information using three analytical techniques for identifying the geographical origin of Liguria extra virgin olive oil. *Food Chem*. 2009;118:163-170.
8. Woodcock T, Downey G, O'Donnell CP. Confirmation of declared provenance of European extra virgin olive oil samples by NIR spectroscopy. *J. Agric. Food Chem*. 2008;56:11520-11525.
9. Tan C, Wang J, Wu T, Qin T, Li M. Determination of nicotine in tobacco samples by near-infrared spectroscopy and boosting partial least squares. *Vibrational Spectroscopy*. 2010;54:35-41.
10. Zhang Y, Cong Q, Xie Y, Yang J, Zhao B. Quantitative analysis of routine chemical constituents in tobacco by near-infrared spectroscopy and support vector machine. *Spectrochim. Acta A*. 2008;71:1408-1413.
11. Bi YM, Chu GH, Wu JZ, Yuan KL, Wu J, Liao F, Xia J, Zhang GX, Zhou GJ. Ensemble Partial Least Squares Algorithm Based on Variable Clustering for Quantitative Infrared Spectrometric Analysis. *Chinese Journal of Analytical Chemistry*. 2015;43:1086-1091.
12. Liu X, Chen HC, Liu TA, Li YL, Lu ZR, Lu WC. Application of PCA-SVR to NIR prediction model for tobacco chemical composition. *Spectrosc. Spectral Anal*. 2007;27:2460-2463.
13. Zhou L, Luo T, Huang Q. Copyrolysis characteristics and kinetics of coal and plastic blends. *Energy Convers Manag*. 2009;50:705-710.
14. Várhegyi G, Czégény Z, Jakab E, McAdam K, Liu C. Tobacco pyrolysis. Kinetic evaluation of thermogravimetric-mass spectrometric experiments. *J Anal Appl Pyrolysis*. 2009;86:310-322.
15. Saldarriaga JF, Aguado R, Pablos A, Amutio M, Olazar M, Bilbao J. Fast characterization of biomass fuels by thermogravimetric analysis (TGA). *Fuel*. 2015;140:44-51.
16. Biagini E, Tognotti L. A generalized procedure for the devolatilization of biomass fuels based on the chemical components. *Energy Fuels* 2014;28:614-623.

17. Haaland DM, Thomas EV. Partial least-squares methods for spectral analyses.1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Anal. Chem.* 1988;60:1193-1202.
18. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. 2018;126:1763-1768.
19. Mukaka MM. a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24:69-71.
20. Akoglu H. User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine.* 2018;18:91-93.
21. Overholser BR, Sowinski KM. Biostatistics primer: part 2. *Nutr Clin Pract.* 2008;23:76-84.
22. Sung YJ, Seo YB. Thermogravimetric study on stem biomass of *Nicotiana tabacum*. *Thermochimica Acta.* 2009;486:1-4.
23. Peng YH, Tang XD, Xuan RQ, Wang L, Dai L, Zhang LL, Liao F, Li HF, Li X, Shen YD, Su Y, Wang H. Analysis of pyrolysis behaviors of biomass extractives via nonlinear stepwise heating program based on Gaussian multipeak fitting of differential thermogravimetric curve. *Thermochimica Acta.* 2021;702:178976.
24. Xu ZC, Li DD, Bi QW, Yin Y, Xing XJ, Zhang XB. Relationship between chlorine content in flue-cured tobacco leaf and main volatile aroma components as well as sensory evaluation. *Acta Tabacaria Sinica.* 2008;14:27-32.
25. Darvishzadeh R, Alavi R. Genetic analysis of chloride concentration in oriental tobacco genotypes. *Journal of Plant Nutrition* 2011;34:1070-1078.
26. Xu Y, Goodacre R. On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of Analysis and Testing.* 2018;2:249-262.
27. Jiang JF, Li L, Zhao MY. Rapid determination of main chemical components in tobacco leaf by NIR diffuse reflectance spectroscopy. *Acta Tabacaria Sinica.* 2006;12:8-12.
28. Tan C, Li M, Qin X. Random Subspace Regression Ensemble for Near-Infrared Spectroscopic Calibration of Tobacco Samples. *Analytical Sciences.* 2008;24:647-653.
29. Yun YH, Cao DS, Tan ML, Yan J, Ren DB, Xu QS, Yu L, Liang YZ. A simple idea on applying large regression coefficient to improve the genetic algorithm-PLS for variable selection in multivariate calibration. *Chemometrics and Intelligent Laboratory Systems.* 2014;130:76-83.
30. Forina M, Casolino C, Millan CP. Iterative predictor weighting (IPW) PLS: a technique for the elimination of useless predictors in regression problems. *J. Chemometrics.* 1999;13:165-184.
31. Tuzzin G, Godinho M, Dettmer A, Zattera AJ. Nanofibrillated cellulose from tobacco industry wastes. *Carbohydrate Polymers.* 2016;148:69-77.
32. Yang HP, Yan R, Chen HP, Lee DH, Zheng CG. Characteristics of hemicellulose, cellulose and lignin pyrolysis. *Fuel.* 86 2007;86:1781-1788.

33. Giudicianni P, Cardone G, Sorrentino G, Ragucci R. Hemicellulose, cellulose and lignin interactions on *Arundo donax* steam assisted pyrolysis. *J. Anal. Appl. Pyrolysis*. 2014;110:138-146.
34. Zhou S, Wang XF, Ning M, Xu YB, He Q, Zhang YP, Yang Y, Tian Z, Bao S, Zhu DL, She SK, Wang H, Chen G. Functional mechanism of potassium salts on combustion and pyrolysis of tobacco. *Tobacco Science & Technology*. 2016;49:33-41.

Figures

Figure 1

Thermogravimetric (TG) curves of 49 tobacco samples.

Figure 2

Pearson correlation coefficients between thermogravimetric curves and the content of six routine chemical constituents detected by continuous flow analyzer.

Figure 3

Scatter plot of models of six chemical constituents, where o represents calibration set data, * represents test set data.

Figure 4

The regression coefficients of the quantitative analysis model of six routine chemical constituents based on TG curves.